PROCEEDINGS OF FSTP3 CONGRESS - A SUSTAINABLE DURUM WHEAT CHAIN FOR FOOD SECURITY AND HEALTHY LIVES

EDITED BY: Brian L. Beres, Aldo Ceriotti, Agata Gadaleta, Curtis J. Pozniak, Roberto Tuberosa and Luigi Cattivelli

PUBLISHED IN: Frontiers in Plant Science and Frontiers in Sustainable Food Systems

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88966-854-0
DOI 10.3389/978-2-88966-854-0

About Frontiers

Frontiers is more than just an open-access publisher of scholarly articles: it is a pioneering approach to the world of academia, radically improving the way scholarly research is managed. The grand vision of Frontiers is a world where all people have an equal opportunity to seek, share and generate knowledge. Frontiers provides immediate and permanent online open access to all its publications, but this alone is not enough to realize our grand goals.

Frontiers Journal Series

The Frontiers Journal Series is a multi-tier and interdisciplinary set of open-access, online journals, promising a paradigm shift from the current review, selection and dissemination processes in academic publishing. All Frontiers journals are driven by researchers for researchers; therefore, they constitute a service to the scholarly community. At the same time, the Frontiers Journal Series operates on a revolutionary invention, the tiered publishing system, initially addressing specific communities of scholars, and gradually climbing up to broader public understanding, thus serving the interests of the lay society, too.

Dedication to Quality

Each Frontiers article is a landmark of the highest quality, thanks to genuinely collaborative interactions between authors and review editors, who include some of the world's best academicians. Research must be certified by peers before entering a stream of knowledge that may eventually reach the public - and shape society; therefore, Frontiers only applies the most rigorous and unbiased reviews.
Frontiers revolutionizes research publishing by freely delivering the most outstanding research, evaluated with no bias from both the academic and social point of view.
By applying the most advanced information technologies, Frontiers is catapulting scholarly publishing into a new generation.

What are Frontiers Research Topics?

Frontiers Research Topics are very popular trademarks of the Frontiers Journals Series: they are collections of at least ten articles, all centered on a particular subject. With their unique mix of varied contributions from Original Research to Review Articles, Frontiers Research Topics unify the most influential researchers, the latest key findings and historical advances in a hot research area! Find out more on how to host your own Frontiers Research Topic or contribute to one as an author by contacting the Frontiers Editorial Office: frontiersin.org/about/contact

PROCEEDINGS OF FSTP3 CONGRESS - A SUSTAINABLE DURUM WHEAT CHAIN FOR FOOD SECURITY AND HEALTHY LIVES

Topic Editors:
Brian L. Beres, Agriculture and Agri-Food Canada, Canada
Aldo Ceriotti, National Research Council (CNR), Italy
Agata Gadaleta, University of Bari Aldo Moro, Italy
Curtis J. Pozniak, University of Saskatchewan, Canada
Roberto Tuberosa, University of Bologna, Italy
Luigi Cattivelli, Research Centre for Genomics and Bioinformatics, Council for Agricultural and Economics Research, Italy

Citation: Beres, B. L., Ceriotti, A., Gadaleta, A., Pozniak, C. J., Tuberosa, R., Cattivelli, L., eds. (2021). Proceedings of FSTP3 Congress - A Sustainable Durum Wheat Chain for Food Security and Healthy Lives. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-88966-854-0

Table of Contents

frontiers in Plant Science

EDITORIAL
published: 09 April 2021
doi: 10.3389/fpls.2021.675510

Editorial: Proceedings of FSTP3 Congress–A Sustainable Durum Wheat Chain for Food Security and Healthy Lives

Roberto Tuberosa[1]*, *Luigi Cattivelli*[2], *Aldo Ceriotti*[3], *Agata Gadaleta*[4], *Brian L. Beres*[5] and *Curtis J. Pozniak*[6]

[1] Department of Agricultural and Food Sciences (DISTAL), University of Bologna, Bologna, Italy, [2] Council for Agricultural Research and Economics, Research Centre for Genomics and Bioinformatics, Fiorenzuola d'Arda, Italy, [3] Institute of Agricultural Biology and Biotechnology, National Research Council (CNR), Rome, Italy, [4] Department of Agricultural and Environmental Science (DISAAT), University of Bari "Aldo Moro", Bari, Italy, [5] Lethbridge Research and Development Centre, Lethbridge, AB, Canada, [6] Crop Development Centre, University of Saskatchewan, Saskatoon, SK, Canada

Keywords: durum wheat, germplasm collections, drought resistance, disease resistance, genome-wide association study (GWAS), quantitative trait locus (QTL), yield sustainability, grain quality

Editorial on the Research Topic

Proceedings of FSTP3 Congress–A Sustainable Durum Wheat Chain for Food Security and Healthy Lives

This Special Issue of Frontiers in Plant Science collects 14 manuscripts presented at the Congress 'From Seed to Pasta 3' (www.fromseedtopasta.com/). The papers highlight some of the most recent achievements in durum wheat (*Triticum turgidum* ssp. *durum*) research, toward a sustainable durum wheat chain able to enhance food security and healthier grain production. A timely update that gauges the impressive progress achieved since the publication of the first Special Issue on durum wheat genomics (Tuberosa and Pozniak, 2014).

The studies address issues relevant for climate-resilient production of durum wheat, particularly in North African countries where durum wheat is a staple to more than 100 million people. Beres et al. describe how meeting this challenge requires matching new cultivars with the best modern management practices. Further, the authors argue that durum production faces stresses that curtail yields and quality specifications desired by export market end-users. Thus, successful biotic and abiotic threat mitigation are ideal case studies in Genotype (G) × Environment (E) × Management (M) interactions where superior cultivars (G) are grown in at-risk regions (E) and require unique approaches to management (M) for sustainable durum production that suit specific agro-ecozones and close the gap between genetic potential and on-farm achieved yield.

From a breeding standpoint, a pivotal issue remains having access to and leveraging beneficial allelic diversity to support genetic studies (Pozniak et al., 2012) and to enhance the performance of elite cultivars (Salvi and Tuberosa, 2015; Khalid et al., 2019; Brinton et al., 2020). A stepping stone to address this issue is the assembly of the Global Durum wheat Panel (GDP), a collection of 1,018 lines designed to identify beneficial alleles for durum wheat improvement (Mazzucotelli et al.). The GDP captures the majority of the diversity available in modern durum wheat germplasm and landraces along with a selection of Emmer and primitive tetraploid wheats to maximize biodiversity. Public seed availability and complete genetic characterization of the GDP provides a unique resource to identify and exchange beneficial genes and alleles to enhance durum wheat breeding world-wide. The GDP is now complemented and extended by the Tetraploid wheat Global Panel (TGC), suitable for a more accurate haplotyping owing to its lower decay of linkage disequilibrium (Maccaferri et al., 2019) and for rare allele identification (**Figure 1**).

Edited and reviewed by:
Hanwei Mei,
Shanghai Agrobiological Gene Center, China

***Correspondence:**
Roberto Tuberosa
roberto.tuberosa@unibo.it

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *03 March 2021*
Accepted: *09 March 2021*
Published: *09 April 2021*

Citation:
Tuberosa R, Cattivelli L, Ceriotti A, Gadaleta A, Beres BL and Pozniak CJ (2021) Editorial: Proceedings of FSTP3 Congress–A Sustainable Durum Wheat Chain for Food Security and Healthy Lives. Front. Plant Sci. 12:675510. doi: 10.3389/fpls.2021.675510

A multi-trait, multi-environment genomic prediction model based on genomic best linear unbiased predictor (GBLUP) coupled with a deep-learning method (DLM) was adopted by Montesinos-López et al. to predict grain yield, days to heading and plant height of 270 durum wheat lines evaluated in 43 environments. The results of the multi-trait DLM were compared with univariate predictions of the GBLUP method and the univariate counterpart of the multi-trait DLM. The best predictions were observed in the absence of genotype × environment interaction term in the univariate and multivariate DLM. Overall, the deep DLM proved to be a practical approach for predicting traits in the context of genomic selection.

The remaining manuscripts addressed the genetic dissection of key traits for durum wheat yield and grain quality. Root system architecture features are increasingly investigated as proxy traits to optimize yield under drought conditions (Maccaferri et al., 2016; Bektas et al., 2020; Alemu et al., 2021). A GWA study allowed Alahmad et al. to identify a major quantitative trait locus (QTL) for seminal root angle on chromosome 6A harboring candidate genes related to gravitropism, polar growth, and hormonal signaling. The authors discussed the potential to deploy root architectural traits to enhance yield stability in environments with limited rainfall.

A more profitable and sustainable durum production scheme relies on novel cultivars that express enhanced resistance to Fusarium Head Blight (FHB) and the three rust diseases. Using a cytogenetic strategy, Kuzmanović et al. equipped durum wheat-*Thinopyron ponticum* recombinant lines with a *Thinopyrum elongatum* major QTL for FHB resistance. Then, a *Th. ponticum* 7el1L arm segment containing the exceptionally effective FHB resistance QTL from *Th. elongatum* together with *Lr19* (leaf rust resistance) and *Yp* (yellow pigment content), was also inserted onto 7DL of bread wheat lines. Chromosome engineering was also deployed by Othmeni et al. to introgress *Amblyopyrum muticum* and D-genome chromosome segments into durum wheat using pentaploid crosses. Results highlighted the importance of the parental genotype when attempting to transfer/develop introgressions into durum wheat from pentaploid crosses. Novel leaf rust QTLs were identified by Kthiri et al. based on mapping populations derived from crosses of resistant cultivars to Mexican races of *P. triticina*, with a susceptible line. Genetic analyses of host reactions suggested oligogenic control of resistance in all populations and identified seven QTLs physically anchored to the durum wheat reference sequence (cv. Svevo; Maccaferri et al., 2019). These QTLs and closely linked markers are useful resources for gene pyramiding and breeding for durable leaf rust resistance in durum wheat.

The search of novel alleles is also pursued through mutagenesis. Harrington et al. screened the cv. Kronos TILLinG population to identify a locus controlling a dominant, environmentally-dependent chlorosis phenotype. A segregating population was classified into discrete phenotypic groups and subjected to bulked-segregant analysis using exome capture followed by next-generation sequencing which highlighted the association on chromosome 3A of *Yellow Early Senescence 1* with the mutant phenotype. Coupling next-generation sequencing with phenotyping of large TILLinG collections allows high-throughput mutation discovery and selection by genotyping. Fruzangohar et al. mutagenized an advanced durum breeding line and performed short-read Illumina sequencing of the exome of 100 lines. An exome reference generated from Svevo and Kronos facilitated reads alignment from the mutants which produced a 484.4 Mbp exome reference. The authors also developed a user-friendly, searchable database and bioinformatic analysis pipeline that predicts zygosity of the mutations discovered and extracts flanking sequences for rapid marker development.

Among yield components, kernel size and shape are important parameters for wheat profitability. Based on data from three environments and 118 RILs from a cross of landrace Iran 249 with cultivar Zardak, Desiderio et al. identified 31 QTLs and 9 QTL interactions for kernel size, and 21 and 5 QTL interactions for kernel shape. Landrace Iran 249 contributed the beneficial allele for most QTLs for kernel shape suggesting its considerable potential for further yield improvement.

End-use quality traits are critical to the success of durum world in the production of pasta, couscous, and other related products. The carotenoid pigment content confers a bright yellow color to pasta and provides some antioxidant capacity. Colasuonno et al. reviewed the genetics of pigment accumulation in durum wheat grain. The major QTLs, accounting for up to 60% variation, were mapped on 7L homoeologous chromosome arms and were explained by allelic variations of the Phytoene Synthase (PSY) genes.

Seed storage proteins are crucial in determining end-use properties of wheat and its nutritional value. Giancaspro et al. reviewed the genetic variability for quantity and composition of grain protein content (GPC), together with grain yield/spike (GYS) and thousand-kernel weight (TKW) in a durum wheat population obtained from an interspecific cross between a common wheat accession and the durum cv. Saragolla. Three major QTLs were detected for both GPC and GYS while eight QTLs influenced TKW. QTLs for GYS, TKW, and GPC overlapped only marginally, with beneficial alleles contributed by both parents. Gluten strength is also key to determine pasta quality of durum wheat grain. Using 162 DH lines from Strongfield × Pelissier, Ruan et al. reported two major QTLs on chromosomes 1B and 1A explaining 25.4–40.1% and 13.7–18.7% of the gluten strength variability, respectively. These QTLs expressed consistently across environments are of great importance to maintain gluten strength of Canadian durum wheat to current market standards while selecting for other traits.

As compared to bread wheat, the comparatively more limited processing and food functionality of durum wheat relates to kernel texture (hardness) and gluten strength. Morris et al. addressed both traits using *ph1b*-mediated translocations from bread wheat. For kernel texture, ca. 28 Mbp of chromosome 5DS replaced about 20 Mbp of 5BS. Single Kernel Characterization System (SKCS) hardness was decreased from ca. 80 to 20 by the expressed puroindolines that softened the endosperm. Crosses with CIMMYT durum lines all produced soft kernel progeny, showing that soft durum can be considered a "tetraploid soft white spring wheat"; notably, excellent bread making potential

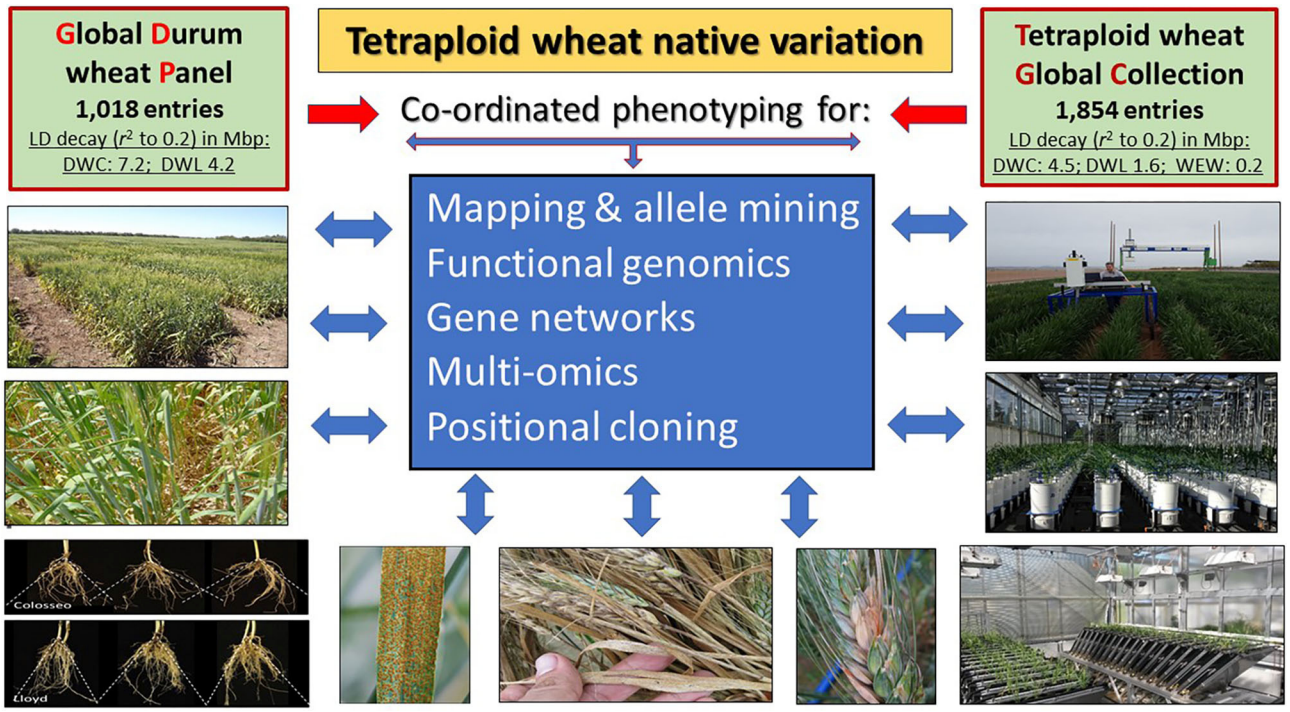


FIGURE 1 | The Global Durum wheat Panel (GDP; Mazzucotelli et al.) and the Tetraploid wheat Global Collection (TGC; Maccaferri et al., 2019) are instrumental to mine the vast biodiversity present in the A and B genomes of *Triticum* species. The lower linkage disequilibrium (LD) decay of the TGC suggests its suitability for cloning QTLs while the GDP is more suitable for breeding purposes. DWC, Durum Wheat Cultivars; DWL, Durum Wheat Landraces; WEW, Wild Emmer Wheat.

was achieved by introgressing Dx2+Dy12 *Glu-D1* alleles in the Soft Svevo background.

In conclusion, facing and tackling the formidable challenges posed by the ongoing climate crisis will increasingly rely on faster genetic gains resulting from a systems-based approach integrating Triticeae multi-omics databases (Li et al., 2021), sequencing (Avni et al., 2017; Appels et al., 2018; Maccaferri et al., 2019; Brinton et al., 2020), breeder friendly phenotyping (Reynolds et al., 2020), deep learning (Wang et al., 2020) and modeling (Condon, 2020). The studies presented here underline the potential of leveraging durum wheat genomic resources (http://plants.ensembl.org/Triticum_turgidum/Info/Index; https://wheat.pw.usda.gov/GG3/node/759) and the tetraploid wheat diversity organized in the GDP and TGC panels (**Figure 1**; https://wheat.pw.usda.gov/GG3/content/november-2020-global-durum-genomic-resources-graingenes-0) as a genomic bridge with bread wheat (Maccaferri et al., 2015). The innovations offered by more sustainable wheat cultivars, once melded into resilient GxExM systems, will no doubt flourish (Beres et al.) for the benefit of farmers, consumers and the environment.

AUTHOR CONTRIBUTIONS

All authors listed have made a substantial, direct and intellectual contribution to the work, and approved it for publication.

ACKNOWLEDGMENTS

This contribution was made possible by the networking offered through the activities of the Expert Working Group (EWG) on Durum Wheat Genomics and Breeding, and the Agronomy EWG. For this, we thank the Wheat Initiative for supporting such networking activities. We greatly appreciate the efforts of all authors, journal editors and peer reviewers.

REFERENCES

Alemu, A., Feyissa, T., Maccaferri, M., Sciara, G., Tuberosa, R., Ammar, K., et al. (2021). Genome-wide association analysis unveils novel QTLs for seminal root system architecture traits in Ethiopian durum wheat. *BMC Genomics* 22:20. doi: 10.1186/s12864-020-07320-4

Appels, R., Eversole, K., Stein, N., Feuillet, C., Keller, B., Rogers, J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., et al. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* 357, 93–97. doi: 10.1126/science.aan0032

Bektas, H., Hohn, C. E., and Waines, J. G. (2020). Dissection of quantitative trait loci for root characters and day length sensitivity in SynOpDH wheat (*Triticum aestivum* L.) bi-parental mapping population. *Plant Genet. Resourc.* 18, 130–142. doi: 10.1017/S1479262120000192

Brinton, J., Ramirez-Gonzalez, R. H., Simmonds, J., Wingen, L., Orford, S., Griffiths, S., et al. (2020). A haplotype-led approach to increase the precision of wheat breeding. *Commun. Biol.* 3:712. doi: 10.1038/s42003-020-01413-2

Condon, A. G. (2020). Drying times: plant traits to improve crop water use efficiency and yield. *J. Exp. Bot.* 71, 2239–2252. doi: 10.1093/jxb/eraa002

Khalid, M., Afzal, F., Gul, A., Amir, R., Subhani, A., Ahmed, Z., et al. (2019). Molecular characterization of 87 functional genes in wheat diversity panel and their association with phenotypes under well-watered and water-limited conditions. Front. Plant Sci. 10:717. doi: 10.3389/fpls.2019.00717

Li, J., Yu, Z., Zhang, A., Yin, Y., Tang, L., Li, G., et al. (2021). Physical mapping of chromosome 7J and a purple coleoptile gene from *Thinopyrum intermedium* in the common wheat background. *Planta* 253:22. doi: 10.1007/s00425-020-03552-6

Maccaferri, M., El-Feki, W., Nazemi, G., Salvi, S., Canè, M. A., Colalongo, M. C., et al. (2016). Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. *J. Exp. Bot.* 67, 1161–1178. doi: 10.1093/jxb/erw039

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Gen.* 51, 885–895. doi: 10.1038/s41588-019-0381-3

Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., et al. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 13, 648–663. doi: 10.1111/pbi.12288

Pozniak, C. J., Clarke, J. M., and Clarke, F. R. (2012). Potential for detection of marker–trait associations in durum wheat using unbalanced, historical phenotypic datasets. *Mol. Breed.* 30, 1537–1550. doi: 10.1007/s11032-012-9737-4

Reynolds, M., Chapman, S., Crespo-Herrera, L., Molero, G., Mondal, S., Pequeno, D. N., et al. (2020). Breeder friendly phenotyping. *Plant Sci.* 295:110396. doi: 10.1016/j.plantsci.2019.110396

Salvi, S., and Tuberosa, R. (2015). The crop QTLome comes of age. *Curr. Opin. Biotec.* 32, 179–185. doi: 10.1016/j.copbio.2015.01.001

Tuberosa, R., and Pozniak, C. (2014). Durum wheat genomics comes of age. *Mol. Breed.* 4, 1527–1530. doi: 10.1007/s11032-014-0188-y

Wang, H., Cimen, E., Singh, N., and Buckler, E. (2020). Deep learning for plant genomics and crop improvement. *Curr. Opin. Plant Biol.* 54, 34–41. doi: 10.1016/j.pbi.2019.12.010

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 April 2019
doi: 10.3389/fpls.2019.00436

A Major Root Architecture QTL Responding to Water Limitation in Durum Wheat

Samir Alahmad[1], Khaoula El Hassouni[2,3], Filippo M. Bassi[3], Eric Dinglasan[1], Chvan Youssef[1], Georgia Quarry[1], Alpaslan Aksoy[1], Elisabetta Mazzucotelli[4], Angéla Juhász[5], Jason A. Able[6], Jack Christopher[7], Kai P. Voss-Fels[1] and Lee T. Hickey[1]**

[1] Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, QLD, Australia, [2] Laboratory of Microbiology and Molecular Biology, Faculty of Sciences, Mohammed V University, Rabat, Morocco, [3] International Center for Agricultural Research in the Dry Areas, Rabat, Morocco, [4] CREA - Genomics Research Centre, Fiorenzuola d'Arda, Italy, [5] School of Science, Edith Cowan University, Joondalup, WA, Australia, [6] School of Agriculture, Food & Wine, Waite Research Institute, The University of Adelaide, Urrbrae, SA, Australia, [7] Leslie Research Facility, Queensland Alliance for Agriculture and Food Innovation, The University of Queensland, Brisbane, QLD, Australia

OPEN ACCESS

Edited by:
Roberto Tuberosa, University of Bologna, Italy

Reviewed by:
Ana María Fita, Universitat Politècnica de València, Spain
Pasquale De Vita, Council for Agricultural Research and Economics, Italy

***Correspondence:**
Kai P. Voss-Fels
k.vossfels@uq.edu.au
Lee T. Hickey
l.hickey@uq.edu.au

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 19 December 2018
Accepted: 22 March 2019
Published: 10 April 2019

Citation:
Alahmad S, El Hassouni K, Bassi FM, Dinglasan E, Youssef C, Quarry G, Aksoy A, Mazzucotelli E, Juhász A, Able JA, Christopher J, Voss-Fels KP and Hickey LT (2019) A Major Root Architecture QTL Responding to Water Limitation in Durum Wheat. Front. Plant Sci. 10:436. doi: 10.3389/fpls.2019.00436

The optimal root system architecture (RSA) of a crop is context dependent and critical for efficient resource capture in the soil. Narrow root growth angle promoting deeper root growth is often associated with improved access to water and nutrients in deep soils during terminal drought. RSA, therefore is a drought-adaptive trait that could minimize yield losses in regions with limited rainfall. Here, GWAS for seminal root angle (SRA) identified seven marker-trait associations clustered on chromosome 6A, representing a major quantitative trait locus (*qSRA-6A*) which also displayed high levels of pairwise LD (r^2 = 0.67). Subsequent haplotype analysis revealed significant differences between major groups. Candidate gene analysis revealed loci related to gravitropism, polar growth and hormonal signaling. No differences were observed for root biomass between lines carrying hap1 and hap2 for *qSRA-6A*, highlighting the opportunity to perform marker-assisted selection for the *qSRA-6A* locus and directly select for wide or narrow RSA, without influencing root biomass. Our study revealed that the genetic predisposition for deep rooting was best expressed under water-limitation, yet the root system displayed plasticity producing root growth in response to water availability in upper soil layers. We discuss the potential to deploy root architectural traits in cultivars to enhance yield stability in environments that experience limited rainfall.

Keywords: root angle, seminal roots, root architecture, GWAS, QTL, haplotype, drought adaptation

INTRODUCTION

Durum wheat (*Triticum durum* Desf.) is a major staple crop in the Mediterranean region (Shewry and Hey, 2015) and other semi-arid regions of the world (Araus et al., 2002). The crop is typically grown under rain-fed conditions where water scarcity is a major limiting factor for productivity, particularly when drought occurs during the flowering or grain filling period (Loss and Siddique, 1994; Belaid, 2000; Mohammadi et al., 2011; Bassi and Sanchez-Garcia, 2017).

Due to climate change, rainfall patterns are predicted to change in most durum production regions worldwide, particularly in the Mediterranean region (Christensen et al., 2007; Carvalho et al., 2014). Therefore, breeding durum for water-limiting environments is a priority (Cattivelli et al., 2008; Boutraa, 2010).

Until recently, breeding programs have focused on above ground traits and direct selection for yield *per se*, while the crop's "hidden-half," i.e., the roots have been largely overlooked. Plant roots are important organs in determining grain productivity driven by water uptake and nutrient acquisition (Sharma et al., 2009; Ehdaie et al., 2012; Shen et al., 2013; Palta and Yang, 2014). Hence, improving RSA in breeding programs is a promising strategy to increase the resilience of durum wheat genotypes in drought-prone environments (Sanguineti et al., 2007; Manschadi et al., 2008). RSA has been recognized as one of the foundations for crop adaptation under water stress conditions (Manschadi et al., 2006; Christopher et al., 2008; Gregory et al., 2009; Asif and Kamran, 2011). Root length, density and root depth are the main components of RSA influencing water extraction in deep soils (King et al., 2003; Asif and Kamran, 2011; Carvalho et al., 2014). These adaptive features determining the root distribution in the soil profile have been associated with root growth angle (Nakamoto et al., 1991; Oyanagi et al., 1993; Oyanagi, 1994; Borrell et al., 2014). In durum wheat, seminal root angle (SRA) is representative of the mature RSA and provides a useful proxy because the trait can be easily phenotyped at the seedling stage (Tuberosa et al., 2002a,b, 2007; de Dorlodot et al., 2007; Fang et al., 2017; El Hassouni et al., 2018). For instance, a narrow SRA is associated with a higher proportion of roots at depth at the mature stage in wheat (Nakamoto and Oyanagi, 1994; Bengough et al., 2004; Manschadi et al., 2008), similar to root growth angle reported in other major crops like sorghum and rice (Omori and Mano, 2007; Uga et al., 2011; Mace et al., 2012). A narrow SRA can improve access to residual moisture in deep soils, particularly under terminal drought conditions (Manschadi et al., 2006; Reynolds et al., 2007; Christopher et al., 2008; Acuña and Wade, 2012; Hamada et al., 2012) and can prolong the grain filling period to improve yield (Blum et al., 1983; Lynch, 1995; Kashiwagi et al., 2005). On the other hand, wide SRA is associated with a shallow root system that may be beneficial for exploring the superficial soil layers and capturing in-season rainfall. Therefore, identifying the optimal RSA in each target environment is critical to guide breeding efforts (El Hassouni et al., 2018). Minor differences in the distribution of roots in the soil space can lead to major impacts on yield. For instance, results from modeling studies suggest that wheat yield would increase by 55 $kg.ha^{-1}$ for each additional millimeter of water extracted from the soil during the critical grain filling stage (Manschadi et al., 2006; Kirkegaard et al., 2007; Christopher et al., 2013). Furthermore, a recent study examining RSA in durum wheat suggests that genotypes with deep root systems could increase grain yield up to 35% and thousand kernel weight by 9% in environments with limited moisture, compared to genotypes with shallow root systems (El Hassouni et al., 2018). The availability of large genetic variability in terms of rooting patterns and the high heritability of SRA (Manschadi et al., 2006, 2008; Maccaferri et al., 2016; Alahmad et al., 2018; El Hassouni et al., 2018) are two key factors suggesting that optimization of the roots could potentially deliver high yielding durum cultivars in water-limiting environments.

In comparison to aboveground traits, studying root traits have been a challenge for plant breeders (Zhang et al., 2009), largely due to lack of efficient and reliable root phenotyping methods and limited knowledge of the genetic control of root development (Tuberosa et al., 2002b; Zhang et al., 2009; Richards et al., 2010; Mace et al., 2012; Shen et al., 2013; Carvalho et al., 2014; Wang et al., 2014). Recently, a high-throughput, affordable and scalable phenotyping method for screening seminal root angle under controlled conditions has been developed, known as the 'clear pot' method (Richard et al., 2015), and has been successfully applied to durum wheat, barley, and bread wheat. The technique has facilitated direct phenotypic selection of SRA (Alahmad et al., 2018; Richard et al., 2018), phenotyping cultivars and breeding lines to investigate yield trends (El Hassouni et al., 2018; Robinson et al., 2018), and phenotyping of mapping populations required for QTL discovery (Robinson et al., 2016). While evaluation of mature RSA in the field is challenging, moderately efficient techniques have been developed, such as 'shovelomics' (Trachsel et al., 2011), soil coring (Wasson et al., 2014) and the 'pasta strainer' method (El Hassouni et al., 2018). Despite the challenges, good progress has been made to identify some of the genomic regions influencing RSA in durum wheat, with several bi-parental and association mapping studies published to date (Sanguineti et al., 2007; Cane et al., 2014; Maccaferri et al., 2016). A recent prioritization analysis of QTL detected in bi-parental and association mapping studies identified nine main QTL clusters on chromosomes 2A, 2B, 4B, 6A, 7A, and 7B, which appear to be most valuable for breeding applications (Maccaferri et al., 2016). However, further research is required to dissect the genetics of RSA in durum wheat that is relevant to breeders, along with the discovery of large effect QTL that are most desirable for marker-assisted breeding.

This study applied the 'clear pot' method to phenotype elite durum populations derived from crosses between Australian and ICARDA germplasm pools and performed a genome-wide association study (GWAS) using DArT-seq markers. A major QTL was identified on chromosome 6A that modulates growth angle, but not root biomass. This major QTL could be exploited and combined with root biomass, thus facilitating the development of new varieties with designer root systems that optimize resource capture in the soil profile targeting different environments.

MATERIALS AND METHODS

Plant Material

A panel of 14 genotypes (**Table 1**) was evaluated for SRA under controlled conditions and nodal root angle in the field to investigate correlation between these traits. This included eight genotypes imported into Australia in 2015 from ICARDA's durum wheat breeding program in Morocco (Fastoz2, Fastoz3, Fastoz6, Fastoz7, Fastoz8, Fastoz10, Outrob4, and Fadda98). The lines were preselected for drought adaptation and used as parents in breeding programs targeting marginal rainfall regions of West

TABLE 1 | Details for the panel of 14 durum wheat and bread wheat standards examined in this study.

Genotype ID	Ploidy	Origin	Pedigree
DBA-Aurora	Tetraploid	Australia	Tamaroi*2/Kalka//RH920318/Kalka///Kalka*2/Tamaroi
Jandaroi	Tetraploid	Australia	110780/111587
Yawa	Tetraploid	Australia	Westonia/Kalka//Kalka/Tamaroi///RAC875/Kalka//Tamaroi
Outrob4	Tetraploid	ICARDA	Ouassel1/4/GdoVZ512/Cit//Ruff/Fg/3/Pin/Gre//Trob
Fadda98	Tetraploid	ICARDA	Awl2/Bit
Fastoz2	Tetraploid	ICARDA	T.polonicumTurkeyIG45272/6/ICAMORTA0463/5/Mra1/4/Aus1/3/Scar/GdoVZ579//Bit
Fastoz3	Tetraploid	ICARDA	Msbl1//Awl2/Bit/3/T.dicoccoidesSYRIG117887
Fastoz6	Tetraploid	ICARDA	Azeghar1/6/Zna1/5/Awl1/4/Ruff//Jo/Cr/3/F9.3/7/Azeghar1//Msbl1/Quarmal
Fastoz7	Tetraploid	ICARDA	CandocrossH25/Ysf1//CM829/CandocrossH25
Fastoz8	Tetraploid	ICARDA	MorlF38//Bcrch1/Kund1149/3/Bicrederaa1/Miki
Fastoz10	Tetraploid	ICARDA	Younes/TdicoAlpCol//Korifla
Mace	Hexaploid	Australia	Wyalkatchem/Stylet//Wyalkatchem[3798]
Scout	Hexaploid	Australia	Sunstate/QH-71-6//Yitpi[4113][4174][4177]
Wylie	Hexaploid	Australia	QT-2327/Cook//QT-2804[3596][3784]

Asia and North Africa. Three Australian durum commercial varieties were also included (DBA Aurora, Jandaroi, Yawa) which are preferred by growers and the pasta industry due to high yield potential and protein content. In addition, bread wheat varieties Mace, Wylie and Scout were included with Mace and Scout used as standards of known root angle phenotype (**Table 1**).

A subset of 393 durum recombinant inbred lines from a nested association mapping (NAM) population were evaluated for SRA and used for GWAS. The NAM population was generated by crossing the eight ICARDA lines listed above as 'founders' to the 'reference' Australian durum varieties Jandaroi and DBA Aurora. The speed breeding facility at The University of Queensland was used to rapidly progress through six generations of spring durum wheat in a year (Ghosh et al., 2018; Watson et al., 2018). The NAM resource comprises 10 donor × reference sub-populations of 92 F_6 lines each (**Figure 1**). The subset of 393 lines evaluated for SRA was selected from the ten families based on agronomic appearance in the field.

Phenotyping Seminal Root Angle Under Controlled Conditions

The panel of 14 genotypes including NAM parents and bread wheat standards (**Table 1**) were phenotyped for SRA, using the 'clear pot' method which is suitable for screening small grain crops (Richard et al., 2015; Robinson et al., 2016; Alahmad et al., 2018). In this experiment, clear pots were filled with composted fine, black-colored pine bark, consisting of 70% particles 0–5 mm in size, pre-mixed with 30% coco peat to increase the water-holding capacity. A randomized complete block design (RCBD) was adopted using *R V3.4.3* (R Core Team, 2017), with 15 replicates per genotype and 24 positions per 4 L pot. Pots were placed on the bench in a distinct column/row grid according to the RCBD design. Seeds were planted in the pots carefully positioning the embryo facing the wall of the pot and vertically with the radical pointed downward. This allows enhanced visibility of the seminal roots following germination. Plants were grown in the glasshouse under diurnal natural light conditions and constant temperature (17 ± 2°C) as recommended by Richard et al. (2015). Images were captured 5 days after sowing (seminal roots 3–5 cm in length) using a Canon PowerShot SX600 HS 16MP Ultra–Zoom Digital camera. The angle between the first pair of seminal roots was measured from the images using *ImageJ* software[1]. Two bread wheat genotypes were tested as standards, including Mace for wide and Scout for narrow SRA (Alahmad et al., 2018). The subset of 393 NAM lines, parents and standards were subsequently phenotyped for SRA using the same procedure as described above.

Phenotyping Nodal Root Angle in the Field

The panel of 14 genotypes was evaluated for nodal root angle in the field using a 'shovelomics' approach (Trachsel et al., 2011). The field experiment was conducted at The University of Queensland Gatton Research Station (27°32′45″ S; 152°19′44″ E) known for summer dominant rainfall and clay soils, Queensland, Australia, in 2017. RCBD was adopted using *R V3.4.3* (R Core Team, 2017) where each genotype was replicated 3 times in 7.5 m^2 yield plots (5 rows spaced 0.3 m × 5 m long). Once all genotypes had reached anthesis, 10 plants per genotype were randomly selected and excavated from the internal two rows. Plants were manually removed to a depth of 20 cm. Excavated plants were then vigorously shaken to remove the loose dry soil before images of the crown roots were captured using a smart-phone camera in the field. These images were then analyzed and the outer angle of the nodal roots was measured using *Image J* software.

Analysis of Phenotype Data

All phenotypic data analyses were performed in *R V3.4.3* (R Core Team, 2017). The root growth angle phenotypes for parental lines (n = 14) in the glasshouse and the field, as well as the subset of NAM lines (n = 393 F_6 lines) in the

[1]http://imagej.nih.gov/ij/

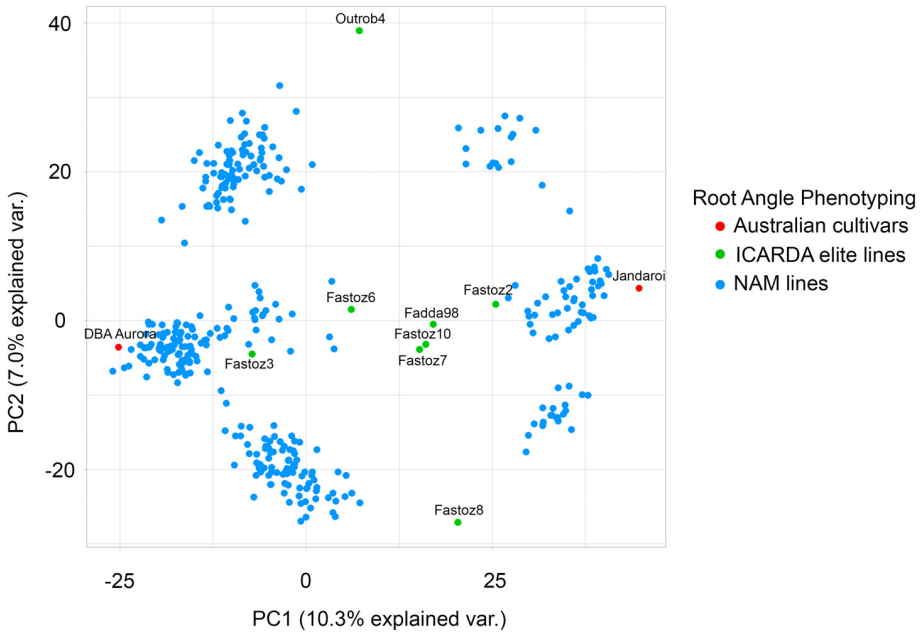


FIGURE 1 | Population structure for the durum NAM lines evaluated for seminal root angle using the clear pot method. A principal component analysis based on pairwise modified Roger's distances calculated from 2,541 polymorphic DArTseq markers was performed for the 393 NAM lines. The ten families NAM lines are derived from two Australian reference varieties (red) and ICARDA elite lines (green).

glasshouse, were measured using ImageJ (Schneider et al., 2012). Statistical analyses of the root growth angle measurements were performed using *ASReml-R* (Butler et al., 2009). Best linear unbiased estimates (BLUEs) were calculated for each individual, including the parental lines and the NAM lines, based on root angle data generated in the glasshouse experiment using the 'clear pot' method. To account for spatial variation a mixed linear model was fitted. In this model, the genotypes were fitted as fixed effect, while the replicate, pot and position were fitted as random terms. The field experiment of the parental lines was conducted to investigate the correlation of mature root growth angle under field conditions with measurements of roots from plants grown under glasshouse conditions at early growth stage. BLUEs for the field data were obtained by fitting a linear mixed model with genotype as a fixed effect and the plot coordinates (row and column) as random effects. The BLUEs for the subset of NAM lines were used as phenotypes in the GWAS analyses. Significance of differences in root biomass between genotypes and between SRA haplotypes were tested using Tukey's test for general linear hypothesis testing based on the linear models described above. Data derived from image analysis of rhizoboxes, and the anatomical traits from the cross sections were also analyzed and differences between the means of genotypes were tested for significance based on a Fisher's Least Significant Difference (LSD) test for multiple comparison with a family-wise error rate 5%.

Genotyping and Curation of Marker Data

All 10 families of the durum NAM population were genotyped using the Diversity Arrays Technology (DArT) genotyping-by-sequencing platform (DArTseq; **Figure 1**). Leaf tissues were sampled from F_6 plants and genomic DNA was extracted according to the protocol provided by DArT. Genotyping resulted in a total of 13,395 DArTseq markers which were ordered according to their genetic positions in the consensus map (version 4.0), provided by Diversity Arrays Technology Pty Ltd., Canberra, Australia. Markers with a frequency of heterozygotes of $\geq$0.1 and missing calls of $\geq$20% were omitted. Markers with $\geq$30% missing data and a minor allele frequency of <3% were omitted and only genotypes with $\leq$20% missing marker information were considered, resulting in a selection of 2,541 high-quality, polymorphic DArTseq markers in 393 durum wheat lines which were used for the subsequent genetic analyses.

Genome-Wide Association Mapping

The 2,541 high-quality genome-wide markers were used to investigate marker-trait-associations (MTA) for SRA. Significances for MTAs were calculated in a two-step mixed linear model approach that increases detection power without increasing the empirical type I error (Stich, 2009). We used a mixed model implemented in the R package *GenABEL* (Aulchenko et al., 2007) which adjusted for population stratification by including identity-by-state estimates for genotype pairs (as a kinship matrix) and a principal component adjustment that uses the first four principal components as fixed covariates. To reduce the type I error rate, we applied a stringent Bonferroni cut-off threshold of $-\log_{10}(p\text{-value}) = 4.67$ ($\alpha = 0.05$) for SRA (Bland and Altman, 1995). The major SRA QTL exceeding this threshold was then compared with

previously identified drought-related and yield component QTL in an alignment approach.

Local LD of the significant markers on chromosome 6A for SRA was calculated and used to group markers into one QTL. Markers with pairwise r^2 values > 0.60 were assigned to an LD block and included in the haplotype analysis, resulting in eight haplotype variants which were observed in the population. Haplotype networks, showing TCS genealogies between haplotype variants (Clement et al., 2000), were calculated using *PopART*[2] (Leigh and Bryant, 2015). The network nodes were colored according to the average SRA in the respective haplotype groups. To investigate the effect of the SRA QTL on the growth angle measurements while correcting for variability due to genetic background, we selected three sub-NAM populations, segregating for the SRA allele combinations associated with narrow and wide SRA including DBA Aurora × Outrob4, DBA Aurora × Fastoz8, DBA Aurora × Fastoz3. We compared the mean SRA of lines that carried the two most frequent haplotypes hap1 and hap2 within the families separately. A Tukey's test was performed to test phenotypic differences in SRA between the haplotype groups within each family. The haplotype effects on root angle phenotypes, were visualized using *GraphPad Prism V6* (Graphpad Software Inc.).

Evaluating Root and Shoot Biomass Effects of Root Angle QTL

To investigate whether the identified major SRA QTL is also associated with pleiotropic differences in root or shoot biomass, a glasshouse experiment was conducted under controlled conditions. A total of 40 closely related genotypes segregating for SRA QTL were evaluated, including 20 lines carrying hap1 and 20 lines carrying hap2. The panel was phenotyped for root biomass using the method reported by Voss-Fels et al., 2017 with some modifications. Here, ANOVA pots (ANOVApot®, 137 mm diameter, 140 mm height) were filled with 1,650 g of sand (with particle size ranging from 0.075–4.75 mm) to facilitate efficient root washing. An RCBD was used for the experimental design, with four plants per genotype in each 1.4 L pot, in three replicates. Fifteen pots were placed in a container fitted with capillary mats to provide sufficient water and nutrient supply. A hydroponic solution was added to each container (1.50 mL of Cultiplex per L of deionized water) and was maintained at the same level over the course of the experiment. The concentration of the nutrient solution was gradually increased as the plants developed and required additional nutrient supply (days 1–10: 1.50 mL/L, days 11–17: 2 mL/L, days 18–22: 2.50 mL/L, days 23–26: 3 mL/L).

Seeds were germinated using a cold treatment (4°C) for 3 days to promote synchronous germination. The germinated seeds were transplanted to the sand-filled plastic pots and grown under diurnal (12 h) photoperiod in a temperature-controlled glasshouse (22/17°C; day/night). At 26 days after sowing (early tillering stage) plants were extracted with minimum disruption to the roots by placing the pot in a water-filled container and carefully washing off the remaining sand in clean water. The roots and shoots from each pot were separated and placed in a dehydrator at 65°C for 72 h before dry weight was measured.

Evaluation of Root Ideotypes Under Well-Watered and Drought Conditions

To investigate the potential for breeding cultivars with different root ideotypes, durum NAM lines representative of four distinct root ideotypes (root angle-root biomass; wide-low, wide-high, narrow-high, and narrow-low) were evaluated using rhizoboxes, similar to those described by Singh et al. (2010). Representative lines were selected based on extreme root angle and biomass phenotypes, as well as haplotype information for the major SRA QTL. Briefly, germinated seeds were sown in rhizoboxes (4 cm × 26 cm × 60 cm) at a depth of 3 cm and maintained under diurnal photoperiod (12 h) and a temperature of 22/17°C (day/night). An RCBD design was adopted in three replicates as blocks in two treatments (well-watered and drought). Four plants per ideotype were planted in each rhizobox. Four rhizoboxes were placed in a container filled with 300 mL water to supply plants with water from the bottom of the rhizoboxes in both treatments. Following sowing, all chambers were watered daily until 1 week after sowing. The well-watered (control) treatment received daily watering from the top of the rhizobox while the drought treatment received no additional water and was subjected to severe water-limitation in the upper layer of the soil. The percentage of soil moisture was measured weekly over the course of the experiment using a soil moisture meter (PMS-714; Lutron Electronic; probe length 22 cm and probe diameter 1 cm) at a depth of 50 cm. Images of the rhizoboxes were captured after 5 weeks and analyzed using *GIA Roots* software (Galkovskyi et al., 2012). The images were cropped into three equal sections at 0–20, 20–40, and 40–60 cm to evaluate root distribution at various soil depths.

To investigate differences in root anatomy associated with the root angle QTL or root ideotype, the stele diameter (SD) and metaxylem area (MXA) was measured for root tissue sampled 10 cm from the seminal root apex in both well-watered and drought treatments. Roots were hand sectioned with a razor blade using a dissecting microscope. The sections were stained with Toluidine Blue O. Images of the root sections were processed using a Zeiss Axio Microscope (Scope.A1 with 100× magnification). All image analyses were processed using *ZEN lite 2012 software* (blue edition, Jena, Germany).

Alignment of Previously Reported QTL for Root and Yield Component Traits

The QTL reported in this study was positioned on the Svevo durum physical map (Maccaferri et al., 2019). The previously reported QTL associated with RSA, distribution and growth angle (Maccaferri et al., 2016) were also projected onto the map using *MapChart V2.3* (Voorrips, 2002). In addition, the previously reported QTL associated with the yield components (TKW, grain yield per spike) and the quality parameter yellow pigment concentration were aligned on the chromosomal region of interest (Golabadi et al., 2011; Roncallo et al., 2012; Maccaferri et al., 2016; Mengistu et al., 2016).

[2]http://popart.otago.ac.nz

Candidate Gene Analysis

Mapping of Marker Genes in the Bread Wheat Reference Genome IWGSC RefSeq v.1.0

Identified peak markers were mapped onto the homologous bread wheat pseudochromosome 6A using the recently published RefSeq v1.0 annotations (Appels et al., 2018). High confidence (HC) and low confidence (LC) RefSeq v1.0 gene models were extracted from the identified region and used in further analyses. Similarity searches were carried out using BLASTn with high stringency settings (with an *e*-value cut-off of 1e-100). Collinearity analysis of Chromosome 6AL between *T. durum* and *Triticum aestivum* regions were performed using Pretzel[3]. Mapped markers and genes expressed in root tissues in seedling stage were used for the analysis.

Gene Expression Analysis and Functional Predictions

Gene expression patterns of the selected bread wheat gene homologs on chromosome 6A were analyzed using the developmental gene expression atlas of polyploid wheat (Ramírez-González et al., 2018); Wheat eFP Browser at http://bar.utoronto.ca/efp_wheat/cgi-bin/efpWeb.cgi and visualized in R using the *Morpheus* package[4].

Translated sequences of selected durum gene models were subjected to functional KEGG pathway analysis using blastKOALA (Kanehisa et al., 2016). Potential interacting proteins were analyzed in STRING (Szklarczyk et al., 2015) using the reference genomes of *Brachypodium distachyon*, *Hordeum vulgare*, *Oryza sativa*, *Zea mays*, and *Arabidopsis thaliana* as data background.

[3]https://github.com/plantinformatics/pretzel

[4]https://software.broadinstitute.org/morpheus

RESULTS

Variation for Root Angle: From Glasshouse to Field

In this study, a panel consisting of the parents of the NAM population and standard lines with previously analyzed root characteristics was evaluated for SRA under controlled conditions in the glasshouse (**Figure 2A**) and nodal root angle under field conditions (**Figure 2B**). Phenotypes displayed by standards were as expected under glasshouse and field conditions, however, less variation under field conditions was observed. For example, the SRA for standard lines under glasshouse conditions were 110.1° (Mace) and 62.6° (Scout) compared to 76.8° (Mace) and 69.9° (Scout) for nodal roots under field conditions (**Figure 2C**). Although the absolute values varied between glasshouse and the field, Mace consistently displayed a wider root angle than Scout across both experiments.

In both experiments, the ICARDA founder lines generally displayed a narrow root growth angle in comparison to the Australian durum cultivars. For example, the SRA for ICARDA founder lines ranged from 50.2–60.7° under glasshouse conditions and 51.8–61.6° in the field. Australian cultivars ranged

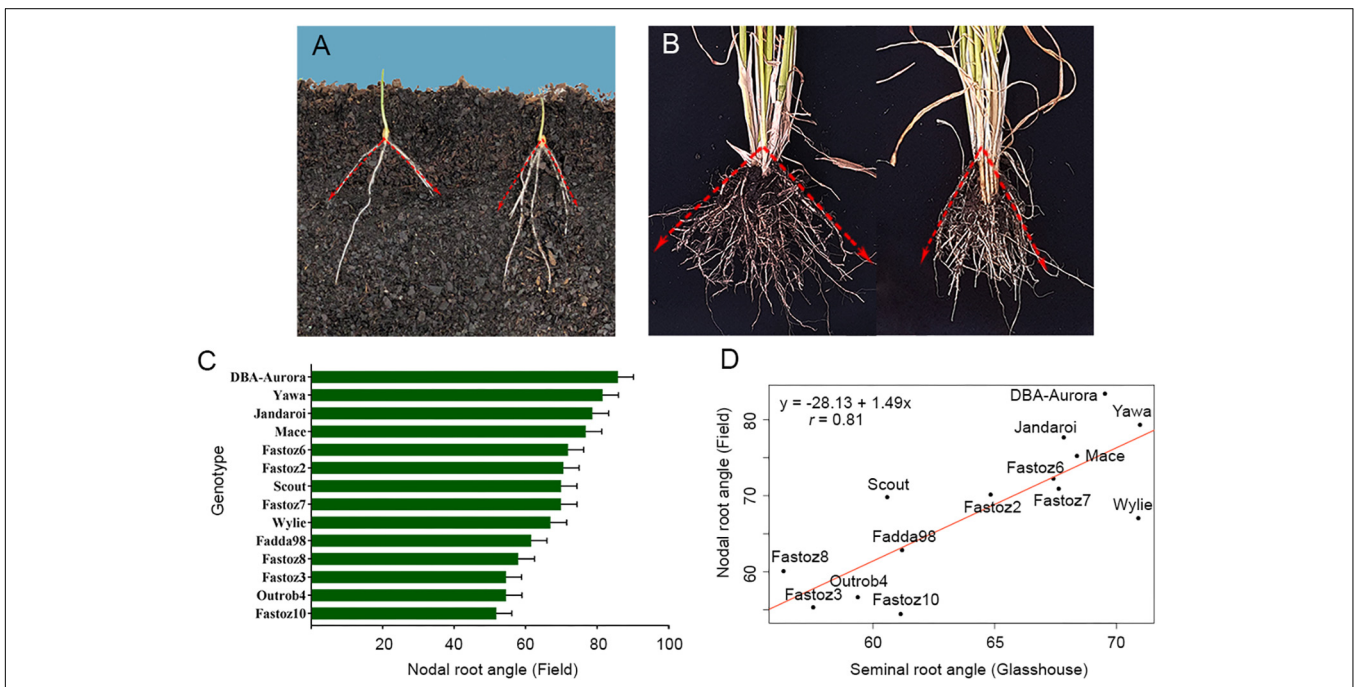


FIGURE 2 | Root growth angle phenotypes measured in important durum wheat cultivars from Australia and ICARDA: **(A)** seminal root angle for Australian variety DBA Aurora (left, wide root angle) and an ICARDA elite founder line Outrob4 (right, narrow root angle) screened using the clear pot method, and **(B)** in the field using shovelomics method. **(C)** Nodal Root growth angle field measurements of 14 parental lines used for NAM population development. Correlation between seminal root angle in the glasshouse and mature roots in the field, $r = 0.81$, $P = 0.00038$ **(D)**.

from 83.0–97.8° under glasshouse conditions and 78.7–85.8° in the field. A strong correlation between seminal root angle in the glasshouse and mature root angle in the field was observed as shown in the regression analysis in **Figure 2** (r = 0.81, P = 0.00038), wherein the panel of 14 lines showed consistent root growth angle phenotypes (**Figure 2D**).

Segregation for Root Angle in the NAM Lines

A high degree of variation for SRA was observed among the 393 NAM lines, with adjusted means ranging from 36.6–91.1° (**Figure 3**). In families derived from DBA Aurora (SRA = 81.1°), the SRA ranged from 38. 5–91.1° and in the families derived from Jandaroi (SRA = 75.5°), the SRA ranged from 36.6–85.4°. In particular, three families (Family 2, Family 3 and Family 5) derived from crosses between DBA Aurora (widest root angle) and three ICARDA founder lines with the narrowest root angle (Outrob4 = 48.7°, Fastoz8 = 39.7° and Fastoz3° = 41.8°, respectively) displayed little transgressive segregation, with a number of lines showing slightly narrower or wider SRA phenotypes than the respective parents. For example, SRA of the individuals ranged from 40.7–91.1°, 38.5–86.0°, and 43.9–86.6° for Families 2, 3, and 5, respectively. Family 1 displayed a higher degree of transgressive segregation. In addition, two families (Jandaroi × Fastoz8 and Jandaroi × Outrob4, i.e., families 6 and 10, respectively) also displayed transgressive segregation, ranging from 41.5–85.1° (Family 6) and 36.6–85.4° (Family 10).

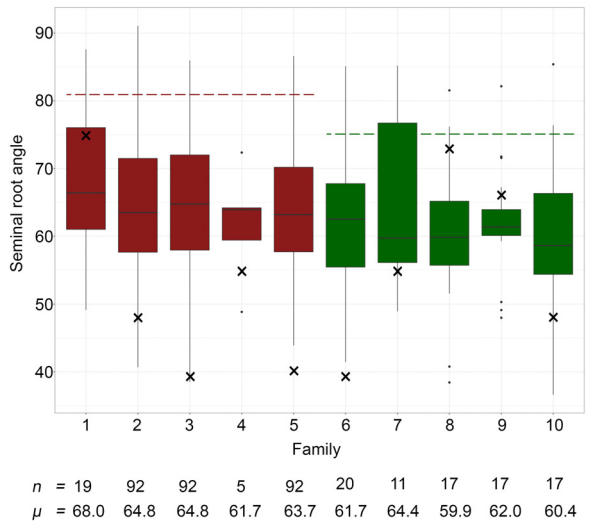


FIGURE 3 | Seminal root growth angle measurements of the 10 NAM families. Families 1 to 5 (red) share DBA Aurora as the common parent, and families 6–10 (green) share Jandaroi as the common parent. Family 1 = DBA Aurora × Fastoz7; Family 2 = DBA Aurora × Outrob4; Family 3 = DBA Aurora × Fastoz8; Family 4 = DBA Aurora × Fadda98; Family 5 = DBA Aurora × Fastoz3, Family 6 = Jandaroi × Fastoz8; Family 7 = Jandaroi × Fastoz10; Family 8 = Jandaroi × Fastoz6; Family 9 = Jandaroi × Fastoz2; Family 10 = Jandaroi × Outrob4. Boxplots display the quartile range and median SRA (horizontal line) of individuals within each of the 10 sub-NAM populations. The broken red line displays the mean SRA value of DBA Aurora and the broken green line displays the mean SRA value of Jandaroi; × represents the mean SRA value of ICARDA founder lines; n represents the number of individuals in each family; μ represents the mean SRA value of each family.

A Major QTL for Root Growth Angle Is Located on Chromosome 6A

A total of seven highly significant markers for SRA were detected on chromosome 6A [at Bonferroni threshold = $-\log_{10}(P)$ 4.67; **Figure 4A**]. A single major QTL region was defined based on high LD (r^2 > 0.60) between pairwise markers, resulting in a QTL interval defined by the outer flanking markers 2256226 (86.46 cM DArTseq V4 consensus map) and1127634 (94.68 cM DArTseq V4 consensus map) (**Figure 4B**). For this QTL, eight main haplotypes were detected (**Figure 4C**). Hap1 and hap2 were the most frequent allelic variants in the subset of NAM lines (frequency = 36.1 and 30.3%, respectively) (**Figure 4D**). The mean SRA for genotypes in the eight defined haplotype groups ranged from 57.8–71.0° (**Figure 4E**). Comparison of SRA between the most frequent haplotypes hap1 and hap2 revealed a highly significant difference of 7.7° (SE = 1.2, P = <0.001) across all families segregating for the QTL in both genetic reference backgrounds DBA Aurora and Jandaroi.

The QTL detected in this study and previously reported QTL in the same chromosomal region (Golabadi et al., 2011; Roncallo et al., 2012; Mengistu et al., 2016; Maccaferri et al., 2016) were positioned onto the durum reference genome (Svevo physical map, Maccaferri et al., 2019; **Figure 5**). The major QTL found in our study (*qSRA-6A*) was found to be co-located with previously reported durum QTL for root growth angle, total root length and root biomass, as well as QTL for yield components and quality traits (**Figure 5**).

qSRA-6A Influences Root Angle but Not Root Biomass in Different Genetic Backgrounds

To evaluate haplotype effects for *qSRA-6A*, we compared the most common haplotypes in three families that were segregating for the QTL. The three families derived from crossing Outrob4, Fastoz8 and Fastoz3 with the common reference parent DBA Aurora were tested as these families segregated for hap1 and hap2 of the major QTL (**Figure 6**). The phenotypic differences in SRA between individuals carrying hap1 and hap2 were significant for family 2 and 3 and followed a similar trend in family 5. Amongst the families, the difference in SRA for lines carrying hap1 versus hap2 ranged between 4.4–9.3°. The largest effect was evident in the DBA Aurora × Fastoz8 family (60.3 and 69.6° for hap1 and hap2, respectively).

To investigate if the contrasting main haplotypes for *qSRA-6A* were only associated with root architectural differences or with overall plant development we conducted a subsequent experiment in which we assayed root and shoot biomass for a subset of 40 genotypes that represented hap1 (n = 20) and hap2 (n = 20). Comparing dried total root biomass, total shoot biomass and the root/shoot ratio of this subset showed no significant differences between the two main haplotype groups. Average values for hap1 and hap2 were 0.641 g/line and 0.638 g/line for

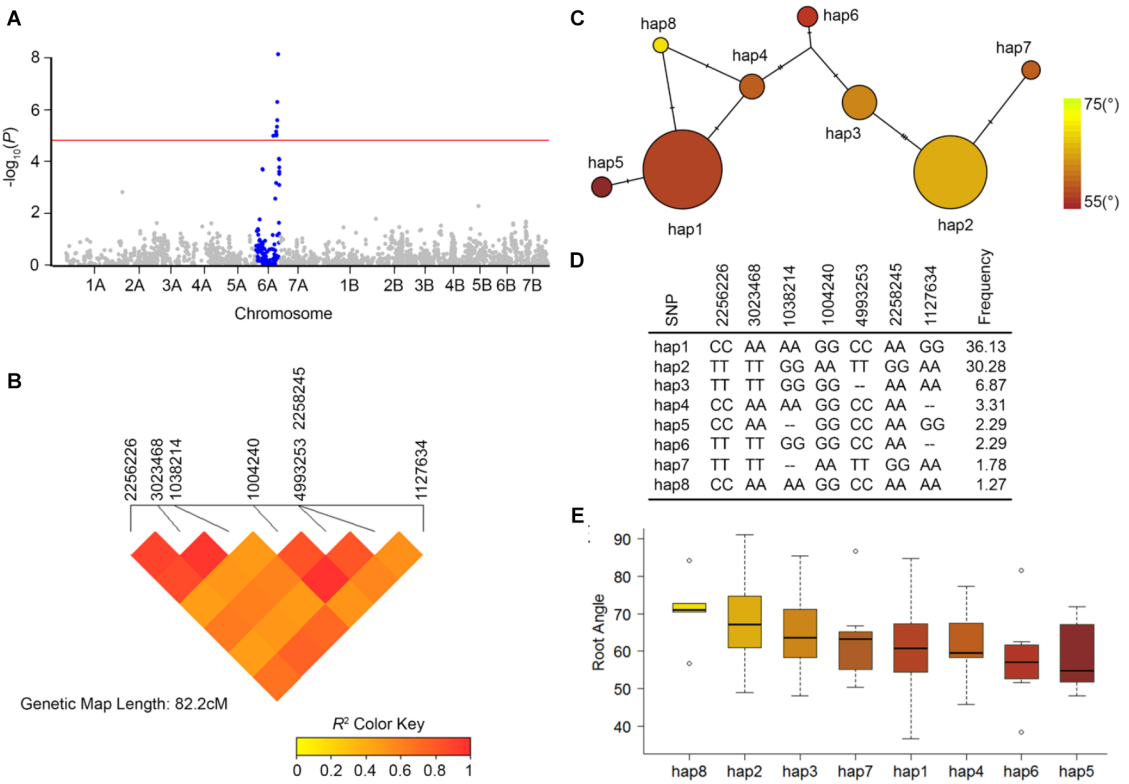


FIGURE 4 | Genome-wide association mapping for seminal root angle in 393 durum lines using 2,541 high quality DArTseq markers (minor allele frequency > 5%). **(A)** Manhattan plot showing chromosome 6A (blue) with significant marker-trait association at Bonferroni significant threshold 4.67 (red horizontal line). The *x*-axis displays the DArTseq markers on 14 chromosomes; y-axis is the $-\log_{10}(P)$. **(B)** Heat map showing pairwise linkage disequilibrium (LD) between 7 significant markers representing major seminal root angle QTL on chromosome 6A (*qSRA-6A*). Color gradient represents LD as r^2. **(C)** Haplotype network of 8 haplotype variants of the *qSRA-6A* that were found in the 393 NAM lines. Size of the circles represents the frequency of each haplotype in the population. Node color indicates mean seminal root angle for lines carrying the haplotype. **(D)** Allelic marker-combination of the 8 haplotypes for the 7 DArTseq markers and the frequency value of each haplotype. **(E)** Seminal root angle variation in each haplogroup.

total root biomass, 1.007 g/line and 1.023 g/line for total shoot biomass and 0.643 and 0.631 for root/shoot ratio.

Root Distribution and Anatomy of Four Root System Ideotypes

Root distribution under well-watered and drought conditions was investigated for four root system ideotypes in rhizoboxes (**Supplementary Figure S1**). Soil moisture of the rhizoboxes decreased dramatically with significant differences between treatments from 3 weeks after sowing (**Supplementary Figure S2**). Overall, plants in the drought treatment had less total root area (area of the roots in the images) (**Figure 7A**) and significantly reduced crown root growth (**Figures 7B,C**). Unexpectedly, the lines carrying the narrow allele were responsive to localized water availability in the upper strata (**Figure 7B**), while in the drought treatment root proliferation shifted deeper into the strata in response to soil moisture at depth (**Figure 7C**). For example, root ideotype 'narrow-high' produced significantly higher root area (23.54 cm^2) in comparison to wide ideotypes ($P < 0.05$, wide-high = 21.15 cm^2; wide-low = 15.11 cm^2) in the upper soil layer (0–20 cm) of the rhizobox under well-watered conditions. In addition, the 'narrow-high' ideotype produced a significantly higher root area distribution under drought conditions in the middle and deepest soil layers (20–40 cm = 18.15 cm^2; 40–60 cm = 12.13 cm^2) when compared to both wide ideotypes (20–40 cm; 10.95–11.60 cm^2, $P < 0.05$–0.1 and 40–60 cm; 2.60–9.35 cm^2, $P < 0.01$–0.4; **Figure 7**).

To investigate associations between root architecture and root anatomical features, stele diameter (SD) and metaxylem area (MXA) were measured for the four ideotypes in the rhizobox experiment. Results suggested a strong link between root angle QTL *qSRA-6A* and SD (**Figures 8A,B**), as well as MXA at depth under well-watered conditions. Under well-watered conditions, the mean MXA for wide-low and wide-high root ideotypes were 11,930 ± 2,100 μm^2 and 11,781 ± 3,287 μm^2, respectively, in comparison to 3,137 ± 353 μm^2 and 3,821 ± 794 μm^2 for the narrow-high and narrow-low, respectively. However, the link was not evident under drought conditions. Under drought conditions, the mean MXA for wide-low and wide-high root ideotypes were 6,627 ± 1,203 μm^2 and 4,543 ± 1,076 μm^2, respectively, in comparison to 3,986 ± 256 μm^2 and 5,577 ± 765 μm^2 for the

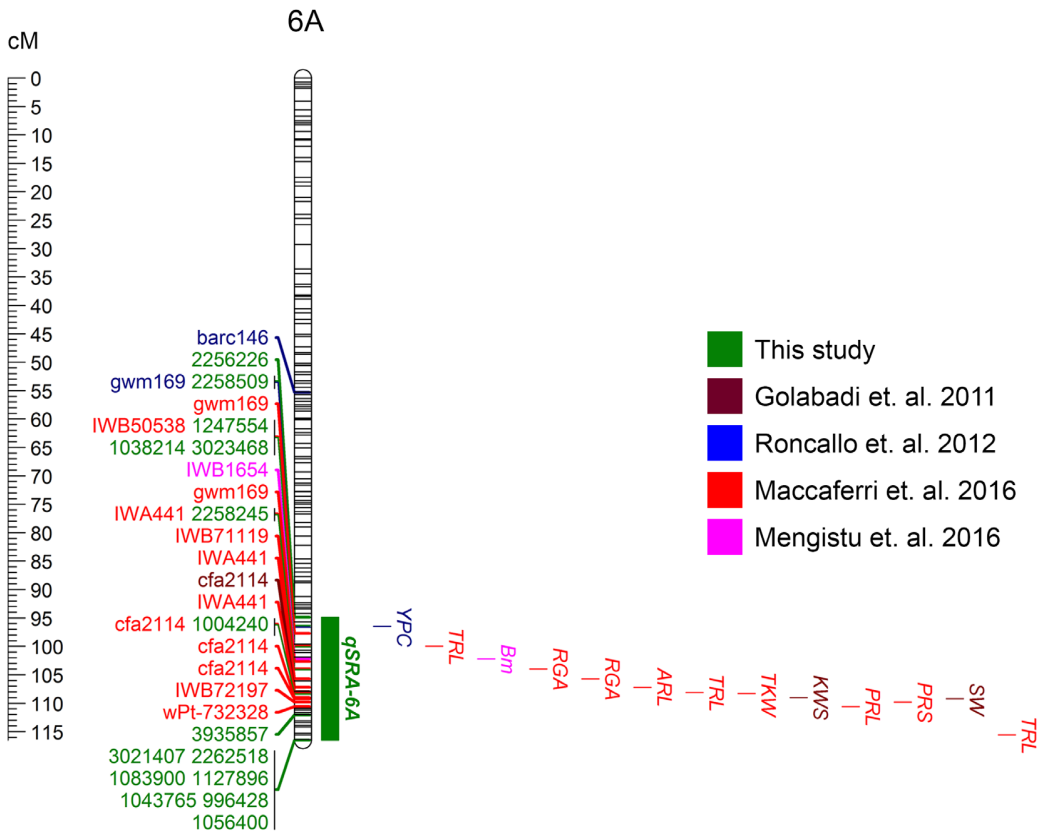


FIGURE 5 | Major QTL for seminal root angle (*qSRA-6A*) positioned on the Svevo durum physical map (Mbp), along with QTL reported in previous mapping studies including root system architecture traits (TRL, total root length; RGA, root growth angle; ARL, average root length; TRL, total root length; PRL, primary root length; PRS, primary root surface), yield component traits (Bm, biomass; TKW, thousand kernel weight; KWS, grain yield per spike; SW, spike dry matter) and a quality trait (YPC, yellow pigment concentration).

narrow-high and narrow-low, respectively (**Figure 8C**). Overall, wide root angle genotypes showed significantly reduced SD and MXA under drought conditions ($P < 0.05$). In addition, the 'narrow-high' ideotype which displayed the highest proportion of roots at depth, also showed smaller MXA under drought conditions at depth.

Candidate Genes Underpinning the 6A QTL

Markers that were found to be significantly associated with SRA mapped to the distal end of chromosome 6A in durum and bread wheat. The length of the marked region was 22.82 Mbp in durum and 22.81 Mbp in bread wheat. These regions contain 393 gene models in durum, while 515 HC and 34 LC gene models were identified in bread wheat. The homologous genes had a high level of collinearity between the terminal regions of chromosome 6A in the *T. durum* reference cultivar Svevo and the bread wheat reference genome RefSeq v1.0 (**Supplementary Figure S3**).

Gene expression patterns were analyzed using the high-resolution tissue and stage-specific RNAseq data of Azurhnaya spring wheat (Winter et al., 2007; Ramírez-González et al., 2018). Altogether 206 genes show root specific expression during the plant life cycle, from which 76 genes show significant expression during early root development stages (radicle and roots at the seedling stage, one leaf and three leaf stage roots and root apical meristem tissues; **Supplementary Figure S3**).

Transcript expression patterns from various tissues both at seedling stage, vegetative and reproductive stages are represented in **Figure 9**. Of these, 15 genes were primarily enriched in the root tissues during early root development (**Figure 9** and **Table 2**).

In the bread wheat genome (RefSeq v1.0 chr 6A region) the mapped markers overlap with gene models representing a NAC transcription factor (TraesCS6A1G386700), a fatty acid hydroxylase family protein (TraesCS6A1G384600), a PRONE protein (TraesCS6A1G405000) and a SAWADEE homeodomain protein 2 (TraesCS6A1G420700). The position of the peak marker from the SRA QTL *qSRA-6A* mapped to the exon region both in the NAC domain-containing protein (3023468) and the PRONE protein (3935857). The observed SNP caused 8:T > A and 7:C > G nucleotide changes, respectively, which resulted in an amino acid change to the translated protein.

Translated durum protein sequences mapped to the QTL region were subjected to KEGG pathway analysis. From the 387 *T. durum* sequences, 114 proteins had significant blast hits in the KEGG database. The following KEGG pathways were

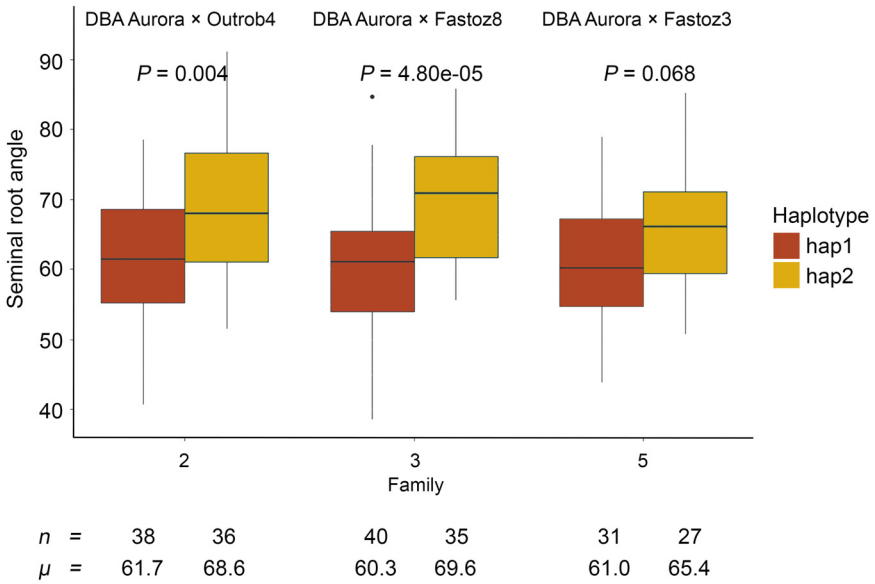


FIGURE 6 | Seminal root angle measurements in three NAM families segregating for the most common haplotype of the QTL *qSRA-6A*. Families were derived from crosses between DBA Aurora to three ICARDA lines (Outrob4, Fastoz3, and Fastoz8). In each family, mean SRA value of individuals carrying hap1 and hap2 was compared. The colors represent the two haplotype groups, *n* represents the number of individuals carrying different haplotype groups, μ represents the mean SRA value of each haplotype group, and *P* represents significance from an Honestly Significant Difference (HSD) test for the difference between the two haplotype groups within each family.

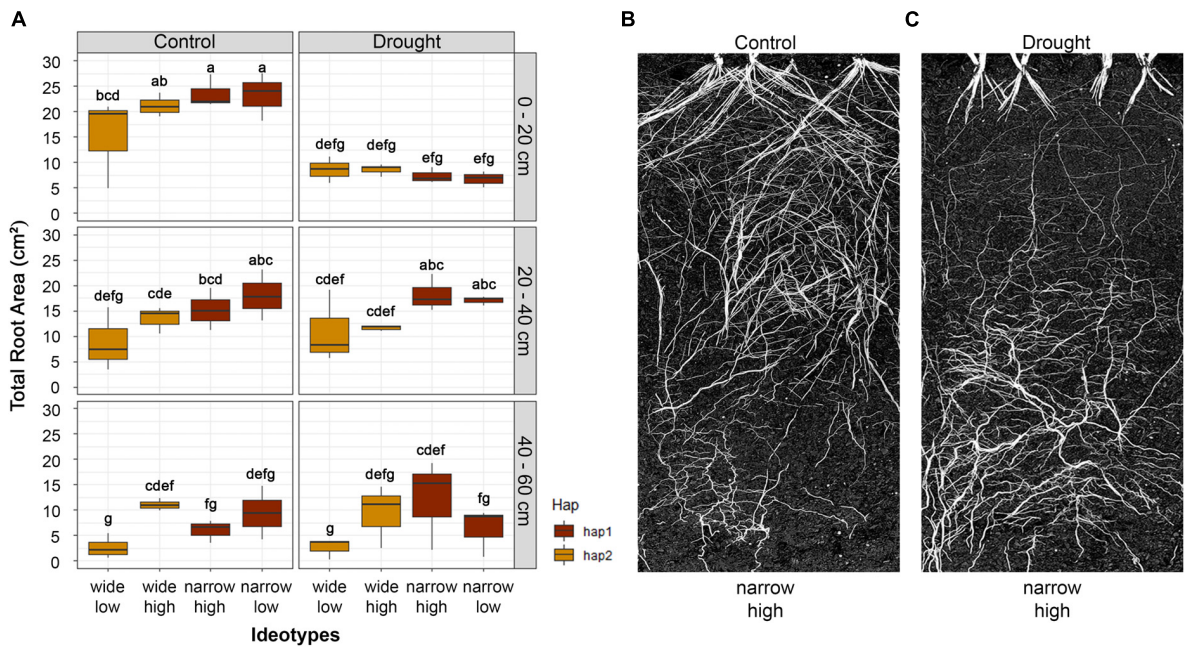


FIGURE 7 | Root area distribution of the four root ideotypes wide-low, wide-high, narrow-high and narrow-low at different depths of the growth chamber. **(A)** Boxplots display root distribution of the four ideotypes under control (well-watered) and drought conditions. The colors represent the two haplotype groups of the root angle *qSRA-6A*. Letters above boxplots indicate significance difference between the four root ideotypes using least significant difference (LSD) test at $\alpha = 0.05$. Visualization of a narrow-high root ideotype under **(B)** controlled (well-watered) and **(C)** drought conditions is shown.

enriched: 13 proteins involved in pathogen defense mechanisms, seven proteins in secondary metabolite biosynthesis (monolignol biosynthesis) and five proteins in fatty acid metabolism [fatty acid biosynthesis, jasmonic acid (JA) biosynthesis and beta-oxidation]. Genes significantly expressed in the radicle and roots at the seedling stage as well as roots and root

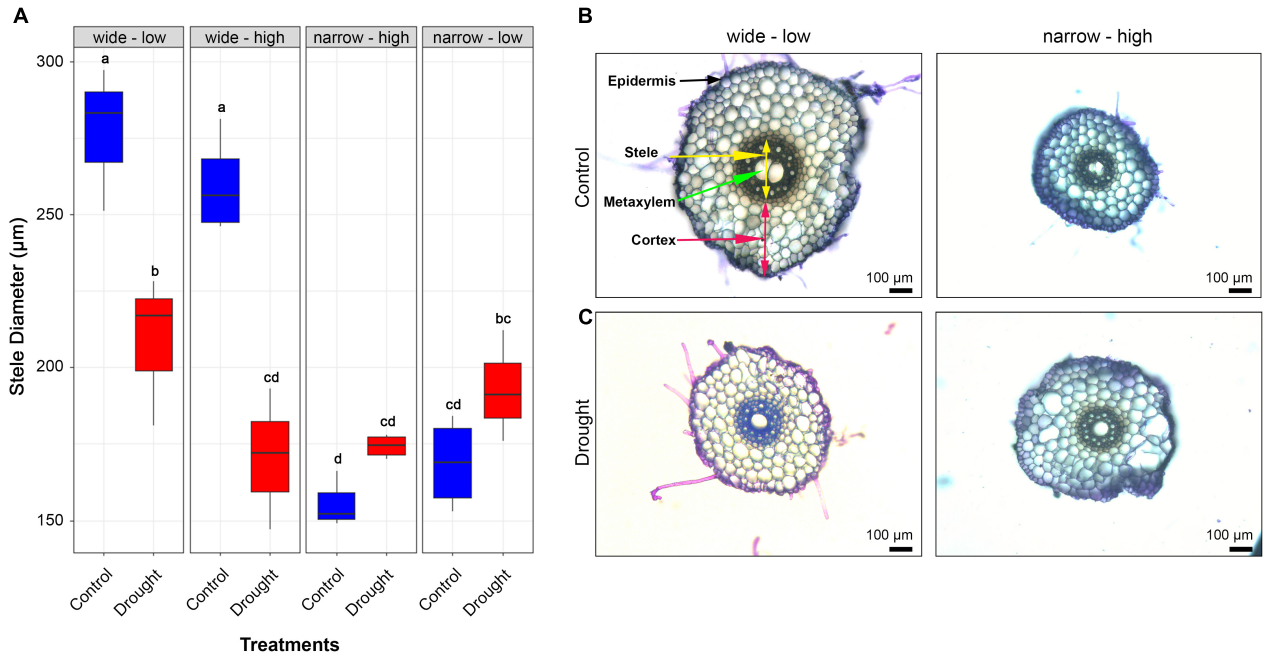


FIGURE 8 | Anatomical features of roots sampled from durum wheat genotypes representative of distinct root system ideotypes. **(A)** Stele diameter of the four root ideotypes of samples collected 10 cm from root apex, under well-watered (control; boxplot colored in blue) and drought conditions (boxplot colored in red). Mean stele diameter with different letters above the boxplot are significantly different. Radial root cross sections on seminal root at 10 cm from root apex displaying anatomical variation in the root ideotype wide root angle with low biomass **(B)** and narrow root angle with high root biomass **(C)** under well-watered and drought conditions, scale bars in the cross sections = 100 µm.

apical meristem at the three leaf stage were analyzed in more detail, using the STRING database to predict potential interacting protein networks. Studies using related monocot species (*B. distachyon*, *H. vulgare*, *O. sativa*, and *Z. mays*) as well as the model dicot species *A. thaliana* indicated the conserved patterns of interacting proteins enriched in functions involved in monolignol biosynthesis, fatty acid metabolism, jasmonic acid metabolism and beta-oxidation. Proteins belonging to plant–pathogen interaction pathways were also detected using both the moncot and dicot data backgrounds. However, homologous proteins of *Arabidopsis* were also related to fatty acid metabolism pathways. Extended interaction networks in all backgrounds also highlighted proteins that are identified in auxin metabolism.

DISCUSSION

A Major QTL on 6A Determines Seminal Root Growth Angle

Here, we report a QTL on chromosome 6A (*qSRA-6A*) that has a significant effect on root growth angle in a subset of 393 durum NAM lines. The co-location of *qSRA-6A* with previously mapped QTL in durum wheat for various root traits, including root length, root surface and root biomass (Maccaferri et al., 2016; Mengistu et al., 2016) suggests that this chromosomal region has a major impact on root development. In addition, *qSRA-6A* also aligned with genomic regions influencing yield components and quality parameters, such as thousand kernel weight and yellow pigment content (Golabadi et al., 2011; Roncallo et al., 2012; Mengistu et al., 2016). Therefore, this region appears important for root system development, which may impact other agronomically important traits including grain yield and end-use quality parameters. Analysis of local LD around the main QTL peak showed high levels of pairwise LD between seven SRA-associated markers. Similar to reported observations for root traits in bread wheat (Voss-Fels et al., 2017) this suggests strong directional selection for this chromosomal block, resulting in a block-wise co-inheritance of markers in tight LD due to strong allelic fixation in important durum wheat germplasm (Tuberosa et al., 2002a; Hayes et al., 2007). Since root architecture directly affects the Source-sink relationship, it is likely that the underlying genetic mechanisms for drought-adaptive traits, such as root growth characteristics, also influence above ground traits like spike grain weight, TKW, spike dry weight and grain quality, which facilitates detection of similar QTL in segregating populations.

A recent study on SRA in bread wheat suggests that root angle is under complex genetic control with multiple small effect QTL involved (Richard et al., 2018). In barley, similar to our study, seminal root traits were reported as being affected by a major QTL on chromosome 5H (Robinson et al., 2016). In maize, a major QTL was reported as constitutive and was associated with root growth angle, root branching and root thickness. This QTL exhibited consistent strong effects under glasshouse and field conditions with different water treatments (Giuliani et al., 2005).

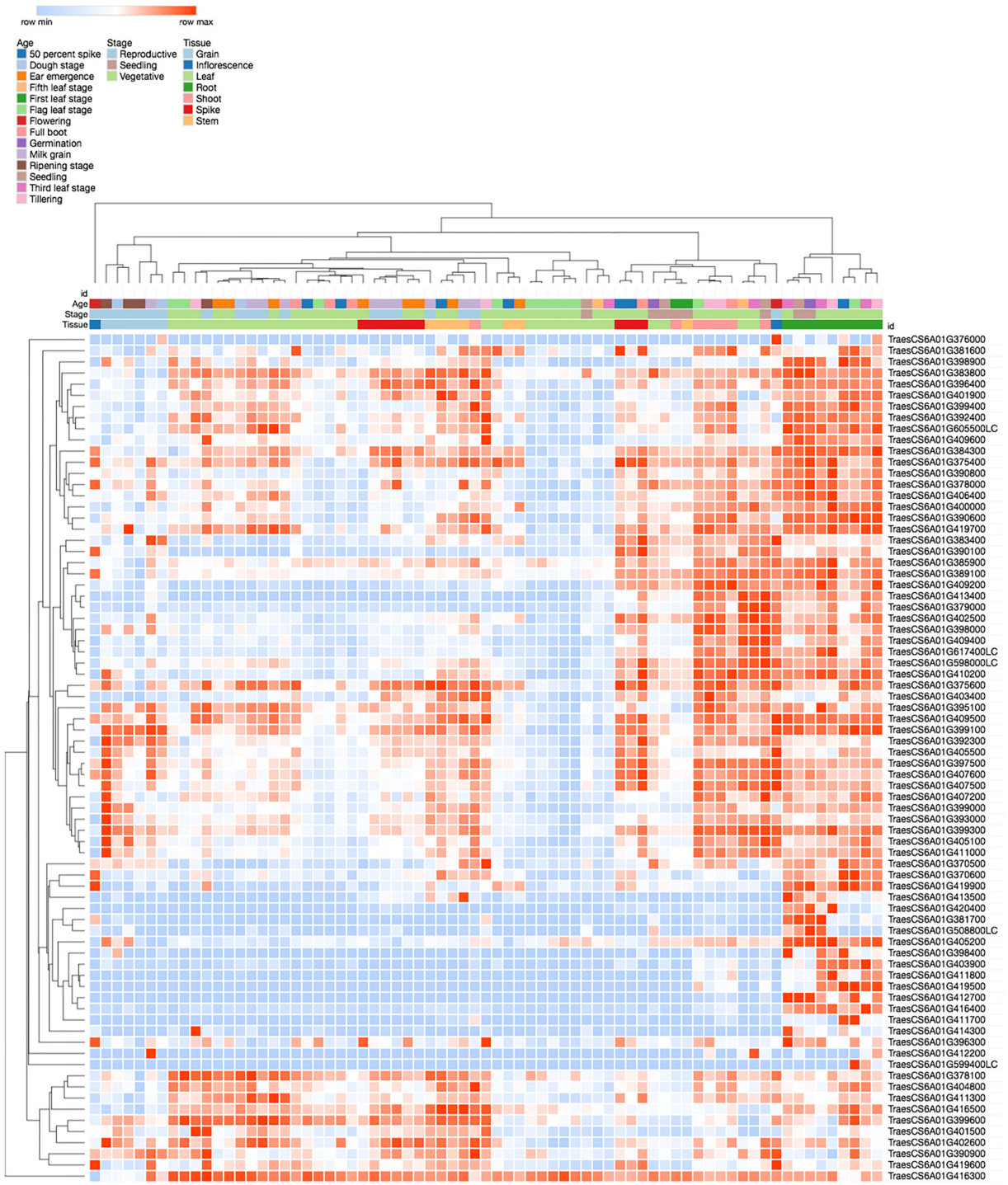


FIGURE 9 | Expression patterns of bread wheat homologs at different age, stage, and tissue specific RNAseq libraries. Heatmap (blue = low; red = high) displayed high expression in root tissues (green) during seedling (pink) and vegetative stage (light green). Genes primarily expressed in the root tissues at seedling and vegetative stages are listed in **Table 2**.

In sorghum, four QTL for nodal root angle were mapped, two of which had a major effect and appeared to co-locate with previously identified QTL for stay-green expressed under low moisture conditions (Mace et al., 2012).

On the other hand, *VERNALIZATION1* (*VRN1*), which controls flowering time in cereals like wheat and barley, was found to modulate RSA in bread wheat and barley (Voss-Fels et al., 2018a). The QTL identified in this study provides

TABLE 2 | List of 15 candidate genes identified using the homologous chromosome 6A genomic region of bread wheat through functional analysis of *qSRA-6A*.

Gene ID	Homologous RefSeq v1.0 Gene ID	Molecular function	KEGG pathway
TRITD6Av1G217760	TraesCS6A01G381700	Cinnamoyl CoA reductase	Monolignol biosynthesis
TRITD6Av1G218270	TraesCS6A01G383800	Glutathione cytosolic	Glutathione metabolism
TRITD6Av1G218390	TraesCS6A01G384300	desumoylating isopeptidase 1-like	
TRITD6Av1G220070	TraesCS6A01G392400	3-ketoacyl- thiolase peroxisomal	Fatty acid metabolism, Jasmonic acid biosynthesis, Beta-oxidation
TRITD6Av1G221500	TraesCS6A01G396400	Serine/threonine-protein phosphatase	Translation/mRNA surveillance pathway
TRITD6Av1G222970	TraesCS6A01G605500LC	Membrane protein	
TRITD6Av1G223490	TraesCS6A01G407600	F-box family protein	
TRITD6Av1G223520	TraesCS6A01G407600	F-box family protein	
TRITD6Av1G223760	TraesCS6A01G409500	Transmembrane protein	
TRITD6Av1G223780	TraesCS6A01G409600	Electron transfer flavoprotein beta-subunit	leucine catabolism and in phytol degradation
TRITD6Av1G224460	TraesCS6A01G412700	Protein kinase, Wall-associated receptor kinase 2	MAPK signaling
TRITD6Av1G224970	TraesCS6A01G413500	Ripening-related protein, RIPER1	
TRITD6Av1G225080	TraesCS6A01G508800LC	12-oxophytodienoate reductase 2	Fatty acid metabolism, Jasmonic acid biosynthesis
TRITD6Av1G225140	TraesCS6A01G414300	Disease resistance protein RPM1	Plant/pathogen interaction
TRITD6Av1G225520	TraesCS6A01G414300	Disease resistance protein RPM1	Plant/pathogen interaction

the opportunity to introgress novel diversity into durum and bread wheat to modulate RSA, potentially leading to improved performance under specific environmental conditions.

Promising Candidate Genes for *qSRA-6A* Have Root Growth-Related Functions

Analyses revealed that the position of the mapped markers of *qSRA-6A* overlaps with a genomic region enriched with genes related to gravitropism, polar growth and hormonal signaling. Notably, genes that are expressed only in root tissues at early stages of plant development (e.g., seedling and one-leaf stage) are involved in pathways such as fatty acid metabolism, jasmonic acid biosynthesis, and monolignol biosynthesis, that might be related to root angle variations in the analyzed phenotypes. Fatty acid metabolism and beta-oxidation play a significant role in the early germination steps when reserve lipids are mobilized to serve as respiratory substrates and to sustain the growth of the seedling. Among the identified early root development-related genes, genes encoding a 3-ketoacyl-CoA thiolase-like protein, Electron transfer flavoprotein beta-subunit and 12-oxophytodienoate reductase 2 and glutathione reductase are related to fatty acid metabolism and beta-oxidation in barley, rice, maize.

During triacylglycerol degradation, fatty acids are released and channeled into gluconeogenesis. Beta-oxidation is also essential to produce secondary metabolites in oxylipin signaling such as the 12−oxophytodienoic acid (OPR2) and jasmonic acid, which serve as signaling compounds in plant growth and pathogen defense mechanisms (Christine and Clifford, 2009; Wasternack and Hause, 2013). An inhibiting role of OPR2 on seed germination has been described (Dave et al., 2011; Dave and Graham, 2012), showing the interaction between 12−oxophytodienoic acid and abscisic acid that leads to increased ABA isensitive5 (ABI5) gene expression and suppressed germination. The identified durum OPR2 shows close homology to OPR5 in maize and OPR2 in barley; both of which are known to be involved in Jasmonic-acid biosynthesis pathways (Helmut et al., 2000). In roots, the growth inhibition by JA and OPR2 occurs via cross-talk with auxin and possibly other hormones, such as gibberelic acid (GA) and brassinoteroids (BR), mostly as an indirect effect via auxin. Jasmonic acid also regulates root gravitropism through affecting the biosynthesis of auxin. It directly influences the gradient formation by modulating its polar distribution of auxin (Singh et al., 2017). JA induced gene expression is most characteristic in the outer layers of the roots (Gasperini et al., 2015).

The monolignol biosynthesis primarily regulated by Cinnamoyl CoA reductases plays an essential role in cell wall lignification in the casparian strips of the root. In the analyzed genomic region of the durum chromosome 6A, there are five cinnamoyl reductase genes encoded, and additional genes (e.g., ABC transporter, laccase) related to monolignol biosynthesis were also found in the region. Drought conditions were reported to enhance the monolignol biosynthesis in the root elongation zone of the seedlings by the inhibition of the cell wall extensibility and root growth (Ma, 2007). Similarly, lignin-related phenolics biosynthesis was also reported during biotic stress (Silva et al., 2010). It is plausible to therefore suggest that these candidate genes may be having a role to play in the constitiution of the durum root ideotype.

The concentration of genes functioning in fatty acid metabolism, monolignol biosynthesis and jasmonic acid biosynthesis pathways in the identified QTL region highlight the importance of jasmonic acid-auxin crosstalk in gravitropism perception and primary root angle formation. The potential target genes include enzymes involved in cell wall expansion (monolignol biosynthesis) and JA biosynthetic pathways.

JA signaling in young root tissues acts contrary to auxin signaling effects, is also related to gravitropism and therefore might be directly related to root angle variations. Genetic variations observed in the encoding genes or their regulating *cis*-promoter elements can help to identify phenotypes where the coordinated negative impact of increasing JA levels and lignification is controlled. Next to the growth-related functions, these genes are also involved in abiotic and biotic stress responses, including drought stress.

The Context-Dependency of Root System Ideotypes in Different Environments

The architecture of roots has great importance for sourcing underground water and nutrients which is essential for plant growth, particularly in marginal environments characterized by water limitation (Manschadi et al., 2006; Asif and Kamran, 2011). In barley, it has been hypothesized that shallow root growth as characterized by a wide root growth angle may be advantageous for accessing nutrients in the upper soil surface under environments where plants experience sporadic rainfall throughout the growing season (Robinson et al., 2016). However, studies showed that this may not be always the case. For example, in the Mediterranean climate of South Australia, which experiences high in-season precipitation, narrow root angle seems advantageous and tends to be associated with higher grain yield (McDonald, 2010). On the other hand, 'steep, deep, and cheap' ideotypes with longer roots or more root branching at depth are most desirable for enhanced access to nutrients and water stored in deep layers of the soil under environments experiencing terminal drought (Manschadi et al., 2008; Christopher et al., 2013; Lynch, 2013). Moreover, deep roots could be ideal to reduce between-plants competition for resources under high-density planting in high-input conditions (Manske and Vlek, 2002). Plants that express drought-adaptive traits under water-limited environments have been shown to sustain increased yield (Manschadi et al., 2010). This is likely due to increased water access post-anthesis which can be through deeper and more efficient root systems (Mace et al., 2012). Manschadi et al. (2006) demonstrated in their modeling study a yield increase of an extra 55 kg/ha for each millimeter of water extracted from the soil after anthesis and during the grain filling stage. The key reason for that was an increase in marginal water use efficiency to almost three times after anthesis, due to enhanced access to water available in deep soils (Kirkegaard et al., 2007; Christopher et al., 2013). The *qSRA-6A* QTL identified in this study is highly associated with root growth angle. This suggests deployment of the narrow (hap1) allele for *qSRA-6A* could be beneficial in breeding programs targeting production environments with deep soils that often experience water stress. The root plasticity under drought suggests that durum genotypes carrying the narrow allele may not have a yield penalty in high rainfall seasons because root growth appears to respond and take advantage of resource availability in the upper soil layers. The G × E for root development and utility of this feature should be further explored.

Interestingly, no association between *qSRA-6A* QTL with root and shoot biomass was found when comparing contrasting haplotype groups with similar genetic backgrounds. This implies that root growth angle and root biomass are under separate genetic control, opening up the possibility to create customized root systems, e.g., by using marker-assisted introgression approaches. Results of our study also highlighted that the 'narrow-high' ideotype produced the highest root proliferation at the deepest soil level with the smallest MXA under drought. This suggests the mechanism for accumulation of root biomass may not only be related to root branching at depth but also associated with an adaptive mechanism involving reduced water use uptake during early stages of crop development. If the loci controlling root biomass are deployed with loci influencing the direction of root growth, root proliferation could be directed and concentrated at desired soil depths. Such allelic combinations assembled through plant breeding could give rise to improved commercial varieties with designer roots tailored for specific target environments (Voss-Fels et al., 2018b). A similar observation was made in a rice study where a major RSA gene called *DEEPER ROOTING 1* (*DRO1*) was cloned (Uga et al., 2013). They showed that *DRO1* is involved in gravitropic response of root cells, thereby influencing root growth direction, but without a significant effect on root biomass. The *qSRA-6A* QTL is unlikely to be *DRO1* because the ortholog is located on the group 5 chromosomes of wheat and wild emmer. Additional QTL related to root growth angle (*DRO2*, *DRO3*, *DRO4*, and *DRO5*) have also been reported (Kitomi et al., 2015). The QTL region *DRO4* is located on the long arm of chromosome 2 in rice and *Aux/IAA8*, *OsPIN1*, and *SAUR* were reported as major contributors to the measured phenotypes (Kitomi et al., 2015; Lou et al., 2015). Using these rice orthologs we found that the corresponding genes are not overlapping with the genomic region determined in our study and are located closer to the centromere of chromosome 6A both in hexaploid and tetraploid wheat. Recently, another QTL has been reported by Wang et al. (2018), showing a significant role of *OsPIN2* gene in root growth angle in rice. *OsPIN2* encodes an Auxin-efflux carrier (Os06g0660200) that showed the highest sequence homology to the homologous gene group on chromosome 7 in wheat (TraesCS7A02G492400, TraesCS7B01G398100, and TraesCS7D01G478800). Overall, the lack of alignment between known genes in rice and the QTL on 6A mapped in the present study implies the region likely contains novel or currently uncharacterized gene(s).

The genetically stable effects of the *qSRA-6A* haplotypes across three different families implies that marker-assisted backcrossing strategies using the marker sequences identified in this study could effectively modulate RSA in future breeding attempts. Numerous studies have reported that root growth angle at the seedling stage was predictive for root growth angle in the field (Tuberosa et al., 2002a; Landi et al., 2010; Li et al., 2015; Richard et al., 2015; Uga et al., 2015; Kanehisa et al., 2016). Furthermore, recent studies have shown that RSA can be manipulated through recurrent phenotypic selection at the seedling stage under glasshouse conditions in which root traits could be measured at high broad-sense heritabilities (ranging from 0.62

to 0.79), leading to significant shifts in population distributions after a few cycles of selection (Alahmad et al., 2018; Richard et al., 2018). This offers plant breeders different options to directly manipulate RSA.

One major limitation for the direct consideration of root traits in defined breeding goals is the high context-dependency of varying RSA in different environments and the interplay of roots with other key phenology traits like flowering time in the expression of the end-point trait such as grain yield (Voss-Fels et al., 2018b). It was recently shown in a comprehensive study involving multi-environment trials in barley that the genetic correlation of root growth angle and yield was highly context dependent, ranging from situations in which shallow roots were associated with increased yield performance and vice versa (Robinson et al., 2018). Multi-environment field trials are required to thoroughly evaluate the value of *qSRA-6A* and different root ideotypes to improve or stabilize durum grain yield in a range of environmental circumstances.

AUTHOR CONTRIBUTIONS

SA and LH conceived and designed the experiments. SA, CY, AA, and GQ performed the experiments. SA, ED, CY, KV-F, EM, and AJ analyzed the data. SA, JA, FB, and LH provided germplasm. SA, KEH, ED, CY, EM, AJ, KV-F, JA, FB, JC, and LH wrote or reviewed the manuscript. All authors read and approved the final manuscript.

FUNDING

SA was supported by Monsanto's Beachell-Borlaug International Scholars Program (MBBISP) Iowa, the United States and a Ph.D. scholarship from The University of Queensland, Australia (UQRS). LH was supported through an Early Career Discovery Research Award (DE170101296) from the Australian Research Council. Genotyping was supported by the Food and Agriculture Organization (FAO; Grant LoA/TF/W3B-PR-02/JORDAN/2016/AGDT).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00436/full#supplementary-material

REFERENCES

Acuña, T. B., and Wade, L. (2012). Genotype × environment interactions for root depth of wheat. *Field Crops Res.* 137, 117–125. doi: 10.1093/aob/mcs251

Alahmad, S., Dinglasan, E., Leung, K. M., Riaz, A., Derbal, N., Voss-Fels, K. P., et al. (2018). Speed breeding for multiple quantitative traits in durum wheat. *Plant Methods* 14:36. doi: 10.1186/s13007-018-0302-y

Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., Stein, N., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191

Araus, J., Slafer, G., Reynolds, M., and Royo, C. (2002). Plant breeding and drought in C3 cereals: What should we breed for? *Ann. Bot.* 89, 925–940. doi: 10.1093/aob/mcf049

Asif, M., and Kamran, A. (2011). Plant breeding for water-limited environments. *Crop Sci.* 51, 2911–2912. doi: 10.2135/cropsci2011.12.0004br

Aulchenko, Y. S., Ripke, S., Isaacs, A., and Van Duijn, C. M. (2007). GenABEL: an R library for genome-wide association analysis. *Bioinformatics* 23, 1294–1296. doi: 10.1093/bioinformatics/btm108

Bassi, F., and Sanchez-Garcia, M. (2017). Adaptation and Stability Analysis of ICARDA Durum Wheat Elites across 18 Countries. *Crop Sci.* 57, 2419–2430. doi: 10.2135/cropsci2016.11.0916

Belaid, A. (2000). "Durum wheat in WANA: production, trade and gains from technological change," in *Durum Wheat Improvement in the Mediterranean Region: New challenges. Options Méditerranéennes*, eds C. M. Royo, M. Nachit, N. Di Fonzo and J.L. Araus (Zaragoza: CIHEAM), 63–70.

Bengough, A., Gordon, D., Al-Menaie, H., Ellis, R., Allan, D., Keith, R., et al. (2004). Gel observation chamber for rapid screening of root traits in cereal seedlings. *Plant Soil* 262, 63–70. doi: 10.1023/B:PLSO.0000037029.82618.27

Bland, J. M., and Altman, D. G. (1995). Multiple significance tests: the Bonferroni method. *BMJ* 310:170. doi: 10.1136/bmj.310.6973.170

Blum, A., Mayer, J., and Gozlan, G. (1983). Associations between plant production and some physiological components of drought resistance in wheat. *Plant Cell Environ.* 6, 219–225.

Borrell, A. K., Mullet, J. E., George-Jaeggli, B., Van Oosterom, E. J., Hammer, G. L., Klein, P. E., et al. (2014). Drought adaptation of stay-green sorghum is associated with canopy development, leaf anatomy, root growth, and water uptake. *J. Exp. Bot.* 65, 6251–6263. doi: 10.1093/jxb/eru232

Boutraa, T. (2010). Improvement of water use efficiency in irrigated agriculture: a review. *J. Agron.* 9, 1–8. doi: 10.3389/fchem.2018.00092

Butler, D., Cullis, B. R., Gilmour, A., and Gogel, B. (2009). *ASReml-R Reference Manual*. Brisbane: Department of Primary Industries and Fisheries.

Cane, M. A., Maccaferri, M., Nazemi, G., Salvi, S., Francia, R., Colalongo, C., et al. (2014). Association mapping for root architectural traits in durum wheat seedlings as related to agronomic performance. *Mol. Breed.* 34, 1629–1645. doi: 10.1007/s11032-014-0177-1

Carvalho, P., Azam-Ali, S., and Foulkes, M. J. (2014). Quantifying relationships between rooting traits and water uptake under drought in Mediterranean barley and durum wheat. *J. Integr. Plant Biol.* 56, 455–469. doi: 10.1111/jipb.12109

Cattivelli, L., Rizza, F., Badeck, F.-W., Mazzucotelli, E., Mastrangelo, A. M., Francia, E., et al. (2008). Drought tolerance improvement in crop plants: an integrated view from breeding to genomics. *Field Crops Res.* 105, 1–14. doi: 10.3389/fpls.2015.00563

Christensen, J. H., Hewitson, B., Busuioc, A., Chen, A., Gao, X., Held, R., et al. (2007). "Regional Climate Projections," in *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*, eds S. Solomon, D. Qin, M. Manning, Z. Chen, M. Marquis, K. B. Averyt, et al. (Cambridge: Cambridge University Press).

Christine, M., and Clifford, W. (2009). Influence of mitochondrial β-oxidation on early pea seedling development. *New Phytol.* 181, 832–842. doi: 10.1111/j.1469-8137.2008.02717.x

Christopher, J., Christopher, M., Jennings, R., Jones, S., Fletcher, S., Borrell, A., et al. (2013). QTL for root angle and number in a population developed from bread wheats (*Triticum aestivum*) with contrasting adaptation to water-limited environments. *Theor. Appl. Genet.* 126, 1563–1574. doi: 10.1007/s00122-013-2074-0

Christopher, J., Manschadi, A., Hammer, G., and Borrell, A. (2008). Developmental and physiological traits associated with high yield and stay-green phenotype in wheat. *Aust. J. Agric. Res.* 59, 354–364. doi: 10.1071/AR07193

Clement, M., Posada, D., and Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9, 1657–1659. doi: 10.1046/j.1365-294x.2000.01020.x

Dave, A., and Graham, I. (2012). Oxylipin Signaling: a Distinct Role for the Jasmonic Acid Precursor cis-(+)-12-Oxo-Phytodienoic Acid (cis-OPDA). *Front. Plant Sci.* 3:42. doi: 10.3389/fpls.2012.00042

Dave, A., Hernández, M. L., He, Z., Andriotis, V. M. E., Vaistij, F. E., Larson, T. R., et al. (2011). 12-Oxo-Phytodienoic Acid Accumulation during Seed Development Represses Seed Germination in *Arabidopsis*. *Plant Cell* 23, 583–599. doi: 10.1105/tpc.110.081489

de Dorlodot, S., Forster, B., Pagès, L., Price, A., Tuberosa, R., and Draye, X. (2007). Root system architecture: opportunities and constraints for genetic improvement of crops. *Trends Plant Sci.* 12, 474–481. doi: 10.1016/j.tplants.2007.08.012

Ehdaie, B., Layne, A. P., and Waines, J. G. (2012). Root system plasticity to drought influences grain yield in bread wheat. *Euphytica* 186, 219–232. doi: 10.1007/s10681-011-0585-9

El Hassouni, K., Alahmad, S., Belkadi, B., Filali-Maltouf, A., Hickey, L., and Bassi, F. (2018). Root system architecture and its association with yield under different water regimes in durum wheat. *Crop Sci.* 58, 1–16. doi: 10.2135/cropsci2018.01.0076

Fang, Y., Du, Y., Wang, J., Wu, A., Qiao, S., Xu, B., et al. (2017). Moderate drought stress affected root growth and grain yield in old, modern and newly released cultivars of winter wheat. *Front. Plant Sci.* 8:672. doi: 10.3389/fpls.2017.00672

Galkovskyi, T., Mileyko, Y., Bucksch, A., Moore, B., Symonova, O., Price, C. A., et al. (2012). GiA Roots: software for the high throughput analysis of plant root system architecture. *BMC Plant Biol.* 12:116. doi: 10.1186/1471-2229-12-116

Gasperini, D., Chételat, A., Acosta, I. F., Goossens, J., Pauwels, L., Goossens, A., et al. (2015). Multilayered organization of jasmonate signalling in the regulation of root growth. *PLoS Genet.* 11:e1005300. doi: 10.1371/journal.pgen.1005300

Ghosh, S., Watson, A., Gonzalez-Navarro, O. E., Ramirez-Gonzalez, R. H., Yanes, L., Mendoza-Suárez, M., et al. (2018). Speed breeding in growth chambers and glasshouses for crop breeding and model plant research. *Nat. Protoc.* 13, 2944–2963. doi: 10.1038/s41596-018-0072-z

Giuliani, S., Sanguineti, M. C., Tuberosa, R., Bellotti, M., Salvi, S., and Landi, P. (2005). *Root-ABA1*, a major constitutive QTL, affects maize root architecture and leaf ABA concentration at different water regimes. *J. Exp. Bot.* 56, 3061–3070. doi: 10.1093/jxb/eri303

Golabadi, M., Arzani, A., Maibody, S. M., Tabatabaei, B. S., and Mohammadi, S. (2011). Identification of microsatellite markers linked with yield components under drought stress at terminal growth stages in durum wheat. *Euphytica* 177, 207–221. doi: 10.1007/s10681-010-0242-8

Gregory, P. J., Bengough, A. G., Grinev, D., Schmidt, S., Thomas, W. B. T., Wojciechowski, T., et al. (2009). Root phenomics of crops: opportunities and challenges. *Funct. Plant Biol.* 36, 922–929. doi: 10.1071/FP09150

Hamada, A., Nitta, M., Nasuda, S., Kato, K., Fujita, M., Matsunaka, H., et al. (2012). Novel QTLs for growth angle of seminal roots in wheat (*Triticum aestivum* L.). *Plant Soil* 354, 395–405. doi: 10.1007/s11104-011-1075-5

Hayes, B., Chamberlain, A., Mcpartlan, H., Macleod, I., Sethuraman, L., and Goddard, M. (2007). Accuracy of marker-assisted selection with single markers and marker haplotypes in cattle. *Genet. Res.* 89, 215–220. doi: 10.1017/S0016672307008865

Helmut, M., Bettina, H., Ivo, F., Jörg, Z., and Claus, W. (2000). Allene oxide synthases of barley (*Hordeum vulgare* cv. Salome): tissue specific regulation in seedling development. *Plant J.* 21, 199–213. doi: 10.1046/j.1365-313x.2000.00669.x

Kanehisa, M., Sato, Y., and Morishima, K. (2016). BlastKOALA and GhostKOALA: KEGG tools for functional characterization of genome and metagenome sequences. *J. Mol. Biol.* 428, 726–731. doi: 10.1016/j.jmb.2015.11.006

Kashiwagi, J., Krishnamurthy, L., Upadhyaya, H. D., Krishna, H., Chandra, S., Vadez, V., et al. (2005). Genetic variability of drought-avoidance root traits in the mini-core germplasm collection of chickpea (*Cicer arietinum* L.). *Euphytica* 146, 213–222. doi: 10.1007/s10681-005-9007-1

King, J., Gay, A., Sylvester-Bradley, R., Bingham, I., Foulkes, J., Gregory, P., et al. (2003). Modelling cereal root systems for water and nitrogen capture: towards an economic optimum. *Ann. Bot.* 91, 383–390. doi: 10.1093/aob/mcg033

Kirkegaard, J., Lilley, J., Howe, G., and Graham, J. M. (2007). Impact of subsoil water use on wheat yield. *Aust. J. Agric. Res.* 58, 303–315. doi: 10.1071/AR06285

Kitomi, Y., Kanno, N., Kawai, S., Mizubayashi, T., Fukuoka, S., and Uga, Y. J. R. (2015). QTLs underlying natural variation of root growth angle among rice cultivars with the same functional allele of DEEPER ROOTING 1. *Rice* 8:16. doi: 10.1186/s12284-015-0049-2

Landi, P., Giuliani, S., Salvi, S., Ferri, M., Tuberosa, R., and Sanguineti, M. C. (2010). Characterization of root-yield-1.06, a major constitutive QTL for root and agronomic traits in maize across water regimes. *J. Exp. Bot.* 61, 3553–3562. doi: 10.1093/jxb/erq192

Leigh, J. W., and Bryant, D. (2015). Popart: full-feature software for haplotype network construction. *Methods Ecol. Evol.* 6, 1110–1116. doi: 10.1111/2041-210X.12410

Li, P., Chen, F., Cai, H., Liu, J., Pan, Q., Liu, Z., et al. (2015). A genetic relationship between nitrogen use efficiency and seedling root traits in maize as revealed by QTL analysis. *J. Exp. Bot.* 66, 3175–3188. doi: 10.1093/jxb/erv127

Loss, S. P., and Siddique, K. (1994). Morphological and physiological traits associated with wheat yield increases in Mediterranean environments. *Adv. Agron.* 52, 229–276. doi: 10.1016/S0065-2113(08)60625-2

Lou, Q., Chen, L., Mei, H., Wei, H., Feng, F., Wang, P., et al. (2015). Quantitative trait locus mapping of deep rooting by linkage and association analysis in rice. *J. Exp. Bot.* 66, 4749–4757. doi: 10.1093/jxb/erv246

Lynch, J. (1995). Root architecture and plant productivity. *Plant Physiol.* 109, 7–13. doi: 10.1104/pp.109.1.7

Lynch, J. P. (2013). Steep, cheap and deep: an ideotype to optimize water and N acquisition by maize root systems. *Ann. Bot.* 112, 347–357. doi: 10.1093/aob/mcs293

Ma, Q.-H. (2007). Characterization of a cinnamoyl-CoA reductase that is associated with stem development in wheat. *J. Exp. Bot.* 58, 2011–2021. doi: 10.1093/jxb/erm064

Maccaferri, M., El-Feki, W., Nazemi, G., Salvi, S., Canè, M. A., Colalongo, M. C., et al. (2016). Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. *J. Exp. Bot.* 67, 1161–1178. doi: 10.1093/jxb/erw039

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome reveals past domestication signatures and future improvement targets. *Nat. Genet.* doi: 10.1038/s41588-019-0381-3

Mace, E., Singh, V., Van Oosterom, E., Hammer, G., Hunt, C., and Jordan, D. (2012). QTL for nodal root angle in sorghum (*Sorghum bicolor* L. Moench) co-locate with QTL for traits associated with drought adaptation. *Theor. Appl. Genet.* 124, 97–109. doi: 10.1007/s00122-011-1690-9

Manschadi, A., Christopher, J., Hammer, G., and Devoil, P. (2010). Experimental and modelling studies of drought-adaptive root architectural traits in wheat (*Triticum aestivum* L.). *Plant Biosystems* 144, 458–462. doi: 10.1080/11263501003731805

Manschadi, A. M., Christopher, J., and Hammer, G. L. (2006). The role of root architectural traits in adaptation of wheat to water-limited environments. *Funct. Plant Biol.* 33, 823–837. doi: 10.1071/FP06055

Manschadi, A. M., Hammer, G. L., Christopher, J. T., and Devoil, P. (2008). Genotypic variation in seedling root architectural traits and implications for drought adaptation in wheat (*Triticum aestivum* L.). *Plant Soil* 303, 115–129. doi: 10.1007/s11104-007-9492-1

Manske, G. G., and Vlek, P. L. (2002). Root architecture–wheat as a model plant. *Plant Roots* 3, 249–259. doi: 10.1201/9780203909423.ch15

McDonald, G. (2010). "The effects of root angle on root growth and yield of wheat in the Australian cereal belt," in *Proceedings of 15th Agronomy Conference: Food Security from Sustainable Agriculture*, Lincoln.

Mengistu, D. K., Kidane, Y. G., Catellani, M., Frascaroli, E., Fadda, C., Pè, M. E., et al. (2016). High-density molecular characterization and association mapping in Ethiopian durum wheat landraces reveals high diversity and potential for wheat breeding. *Plant Biotechnol. J.* 14, 1800–1812. doi: 10.1111/pbi.12538

Mohammadi, R., Sadeghzadeh, D., Armion, M., and Amri, A. (2011). Evaluation of durum wheat experimental lines under different climate and water

regime conditions of Iran. *Crop Pasture Sci.* 62, 137–151. doi: 10.1071/CP10284

Nakamoto, T., and Oyanagi, A. (1994). The direction of growth of seminal roots of *Triticum aestivum* L. and experimental modification thereof. *Ann. Bot.* 73, 363–367. doi: 10.1006/anbo.1994.1045

Nakamoto, T., Shimoda, K., and Matsuzaki, A. (1991). Elongation angle of nodal roots and its possible relation to spatial root distribution in maize and foxtail millet. *Jpn. J. Crop Sci.* 60, 543–549. doi: 10.1626/jcs.60.543

Omori, F., and Mano, Y. (2007). QTL mapping of root angle in F2 populations from maize 'B73' × teosinte 'Zea luxurians'. *Plant Root* 1, 57–65. doi: 10.3117/plantroot.1.57

Oyanagi, A. (1994). Gravitropic response growth angle and vertical distribution of roots of wheat (*Triticum aestivum* L.). *Plant Soil* 165, 323–326. doi: 10.1007/BF00008076

Oyanagi, A., Nakamoto, T., and Wada, M. (1993). Relationship between root growth angle of seedlings and vertical distribution of roots in the field in wheat cultivars. *Jpn. J. Crop Sci.* 62, 565–570. doi: 10.1626/jcs.62.565

Palta, J. A., and Yang, J. (2014). Crop root system behaviour and yield. *Field Crops Res.* 165, 1–4. doi: 10.1016/j.fcr.2014.06.024

R Core Team (2017). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ramírez-González, R., Borrill, P., Lang, D., Harrington, S., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* 361:eaar6089. doi: 10.1126/science.aar6089

Reynolds, M. P., Pierre, C. S., Saad, A. S., Vargas, M., and Condon, A. G. (2007). Evaluating potential genetic gains in wheat associated with stress-adaptive trait expression in elite genetic resources under drought and heat stress. *Crop Sci.* 47, S-172–S-189. doi: 10.2135/cropsci2007.10.002 2IPBS

Richard, C., Christopher, J., Chenu, K., Borrell, A., Christopher, M., and Hickey, L. (2018). Selection in early generations to shift allele frequency for seminal root angle in wheat. *Plant Genome* 11:170071. doi: 10.3835/plantgenome2017.08.0071

Richard, C. A., Hickey, L. T., Fletcher, S., Jennings, R., Chenu, K., and Christopher, J. T. (2015). High-throughput phenotyping of seminal root traits in wheat. *Plant Methods* 11:13. doi: 10.1186/s13007-015-0055-9

Richards, R. A., Rebetzke, G. J., Watt, M., Condon, A. G., Spielmeyer, W., and Dolferus, R. (2010). Breeding for improved water productivity in temperate cereals: phenotyping, quantitative trait loci, markers and the selection environment. *Funct. Plant Biol.* 37, 85–97. doi: 10.1071/FP09219

Robinson, H., Hickey, L., Richard, C., Mace, E., Kelly, A., Borrell, A., et al. (2016). Genomic regions influencing seminal root traits in barley. *Plant Genome* 9, 1–13. doi: 10.3835/plantgenome2015.03.0012

Robinson, H., Kelly, A., Fox, G., Franckowiak, J., Borrell, A., and Hickey, L. (2018). Root architectural traits and yield: exploring the relationship in barley breeding trials. *Euphytica* 214:151. doi: 10.1007/s10681-018-2219-y

Roncallo, P. F., Cervigni, G. L., Jensen, C., Miranda, R., Carrera, A. D., Helguera, M., et al. (2012). QTL analysis of main and epistatic effects for flour color traits in durum wheat. *Euphytica* 185, 77–92. doi: 10.1007/s10681-012-0628-x

Sanguineti, M., Li, S., Maccaferri, M., Corneti, S., Rotondo, F., Chiari, T., et al. (2007). Genetic dissection of seminal root architecture in elite durum wheat germplasm. *Ann. Appl. Biol.* 151, 291–305. doi: 10.1111/j.1744-7348.2007.00198.x

Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012). NIH Image to ImageJ: 25 years of image analysis. *Nat. Methods* 9, 671–675. doi: 10.1038/nmeth.2089

Sharma, S., Bhat, P. R., Ehdaie, B., Close, T. J., Lukaszewski, A. J., and Waines, J. G. (2009). Integrated genetic map and genetic analysis of a region associated with root traits on the short arm of rye chromosome 1 in bread wheat. *Theor. Appl. Genet.* 119, 783–793. doi: 10.1007/s00122-009-1088-0

Shen, Y., Li, S., and Shao, M. (2013). Effects of spatial coupling of water and fertilizer applications on root growth characteristics and water use of winter wheat. *J. Plant Nutr.* 36, 515–528. doi: 10.1080/01904167.2012.717160

Shewry, P. R., and Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food Energy Secur.* 4, 178–202. doi: 10.1002/fes3.64

Silva, M. J. C. M., Valencise, B. C. A., Juliana, D. O. F. V., Carnier, D. M., and Paulo, M. (2010). Abiotic and biotic stresses and changes in the lignin content and composition in plants. *J. Integr. Plant Biol.* 52, 360–376. doi: 10.1111/j.1744-7909.2010.00892.x

Singh, M., Gupta, A., and Laxmi, A. (2017). Striking the right chord: signaling enigma during root gravitropism. *Front. Plant Sci.* 8:1304. doi: 10.3389/fpls.2017.01304

Singh, V., Van Oosterom, E. J., Jordan, D. R., Messina, C. D., Cooper, M., and Hammer, G. L. (2010). Morphological and architectural development of root systems in sorghum and maize. *Plant Soil* 333, 287–299. doi: 10.1007/s11104-010-0343-0

Stich, B. (2009). Comparison of mating designs for establishing nested association mapping populations in maize and *Arabidopsis thaliana*. *Genetics* 183, 1525–1534. doi: 10.1534/genetics.109.108449

Szklarczyk, D., Franceschini, A., Wyder, S., Forslund, K., Heller, D., Huerta-Cepas, J., Simonovic, M., et al. (2015). STRING v10: protein-protein interaction networks, integrated over the tree of life. *Nucleic Acids Res.* 43, D447–D452. doi: 10.1093/nar/gku1003

Trachsel, S., Kaeppler, S. M., Brown, K. M., and Lynch, J. P. (2011). Shovelomics: high throughput phenotyping of maize (*Zea mays* L.) root architecture in the field. *Plant Soil* 341, 75–87. doi: 10.1007/s11104-010-0623-8

Tuberosa, R., Giuliani, S., Parry, M., and Araus, J. (2007). Improving water use efficiency in Mediterranean agriculture: What limits the adoption of new technologies? *Ann. Appl. Biol.* 150, 157–162. doi: 10.1111/j.1744-7348.2007.00127.x

Tuberosa, R., Salvi, S., Sanguineti, M. C., Landi, P., Maccaferri, M., and Conti, S. (2002a). Mapping QTLs regulating morpho-physiological traits and yield: case studies, shortcomings and perspectives in drought-stressed maize. *Ann. Bot.* 89, 941–963.

Tuberosa, R., Sanguineti, M. C., Landi, P., Giuliani, M. M., Salvi, S., and Conti, S. (2002b). Identification of QTLs for root characteristics in maize grown in hydroponics and analysis of their overlap with QTLs for grain yield in the field at two water regimes. *Plant Mol. Biol.* 48, 697–712.

Uga, Y., Kitomi, Y., Ishikawa, S., and Yano, M. (2015). Genetic improvement for root growth angle to enhance crop production. *Breed. Sci.* 65, 111–119. doi: 10.1270/jsbbs.65.111

Uga, Y., Okuno, K., and Yano, M. (2011). Dro1, a major QTL involved in deep rooting of rice under upland field conditions. *J. Exp. Bot.* 62, 2485–2494. doi: 10.1093/jxb/erq429

Uga, Y., Sugimoto, K., Ogawa, S., Rane, J., Ishitani, M., Hara, N., et al. (2013). Control of root system architecture by DEEPER ROOTING 1 increases rice yield under drought conditions. *Nat. Genet.* 45, 1097–1102. doi: 10.1038/ng.2725

Voorrips, R. (2002). MapChart: software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78. doi: 10.1093/jhered/93.1.77

Voss-Fels, K. P., Qian, L., Parra-Londono, S., Uptmoor, R., Frisch, M., Keeble-Gagnère, G., Appels, R., and Snowdon, R. J. (2017). Linkage drag constrains the roots of modern wheat. *Plant Cell Environ.* 40, 717—725. doi: 10.1111/pce.12888

Voss-Fels, K. P., Robinson, H., Mudge, S. R., Richard, C., Newman, S., Wittkop, B., et al. (2018a). VERNALIZATION1 modulates root system architecture in wheat and barley. *Mol. Plant* 11, 226–229. doi: 10.1016/j.molp.2017.10.005

Voss-Fels, K. P., Snowdon, R. J., and Hickey, L. T. (2018b). Designer roots for future crops. *Trends Plant Sci.* 23, 957–960. doi: 10.1016/j.tplants.2018.08.004

Wang, C., Liu, W., Li, Q., Ma, D., Lu, H., Feng, W., et al. (2014). Effects of different irrigation and nitrogen regimes on root growth and its correlation with above-ground plant parts in high-yielding wheat under field conditions. *Field Crops Res.* 165, 138–149. doi: 10.1016/j.fcr.2014.04.011

Wang, L., Guo, M., Li, Y., Ruan, W., Mo, X., Wu, Z., et al. (2018). LARGE ROOT ANGLE1, encoding OsPIN2, is involved in root system architecture in rice. *J. Exp. Bot.* 69, 385–397. doi: 10.1093/jxb/erx427

Wasson, A., Rebetzke, G., Kirkegaard, J., Christopher, J., Richards, R., and Watt, M. (2014). Soil coring at multiple field environments can directly quantify variation in deep root traits to select wheat genotypes for breeding. *J. Exp. Bot.* 65, 6231–6249. doi: 10.1093/jxb/eru250

Wasternack, C., and Hause, B. (2013). Jasmonates: biosynthesis, perception, signal transduction and action in plant stress response, growth and development. An update to the 2007 review in Annals of Botany. *Ann. Bot.* 111, 1021–1058. doi: 10.1093/aob/mct067

Watson, A., Ghosh, S., Williams, M. J., Cuddy, W. S., Simmonds, J., Rey, M.-D., et al. (2018). Speed breeding is a powerful tool to accelerate crop research and breeding. *Nat. Plants* 4, 23–29. doi: 10.1038/s41477-017-0083-8

Winter, D., Vinegar, B., Nahal, H., Ammar, R., Wilson, G. V., and Provart, N. J. (2007). An "Electronic Fluorescent Pictograph" browser for exploring and analyzing large-scale biological data sets. *PLoS One* 2:e718. doi: 10.1371/journal.pone.0000718

Zhang, X., Chen, S., Sun, H., Wang, Y., and Shao, L. (2009). Root size, distribution and soil water depletion as affected by cultivars and environmental factors. *Field Crops Res.* 114, 75–83. doi: 10.1016/j.fcr.2009.07.006

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 April 2019
doi: 10.3389/fpls.2019.00448

Genomic Regions From an Iranian Landrace Increase Kernel Size in Durum Wheat

Francesca Desiderio[1], Leila Zarei[2], Stefania Licciardello[3], Kianoosh Cheghamirza[2], Ezatollah Farshadfar[2], Nino Virzi[3], Fabiola Sciacca[3], Paolo Bagnaresi[1], Raffaella Battaglia[1], Davide Guerra[1], Massimo Palumbo[3], Luigi Cattivelli[1] and Elisabetta Mazzucotelli[1]**

[1] Council for Agricultural Research and Economics, Research Centre for Genomics and Bioinformatics, Fiorenzuola d'Arda, Italy, [2] Department of Agronomy and Plant Breeding, Razi University, Kermanshah, Iran, [3] Council for Agricultural Research and Economics, Research Centre for Cereal and Industrial Crops, Acireale, Italy

Kernel size and shape are important parameters determining the wheat profitability, being main determinants of yield and its technological quality. In this study, a segregating population of 118 recombinant inbred lines, derived from a cross between the Iranian durum landrace accession "Iran_249" and the Iranian durum cultivar "Zardak", was used to investigate durum wheat kernel morphology factors and their relationships with kernel weight, and to map the corresponding QTLs. A high density genetic map, based on wheat 90k iSelect Infinium SNP assay, comprising 6,195 markers, was developed and used to perform the QTL analysis for kernel length and width, traits related to kernel shape and weight, and heading date, using phenotypic data from three environments. Overall, a total of 31 different QTLs and 9 QTL interactions for kernel size, and 21 different QTLs and 5 QTL interactions for kernel shape were identified. The landrace Iran_249 contributed the allele with positive effect for most of the QTLs related to kernel length and kernel weight suggesting that the landrace might have considerable potential toward enhancing the existing gene pool for grain shape and size traits and for further yield improvement in wheat. The correlation among traits and co-localization of corresponding QTLs permitted to define 11 clusters suggesting causal relationships between simplest kernel size trait, like kernel length and width, and more complex secondary trait, like kernel shape and weight related traits. Lastly, the recent release of the *T. durum* reference genome sequence allowed to define the physical interval of our QTL/clusters and to hypothesize novel candidate genes inspecting the gene content of the genomic regions associated to target traits.

Keywords: *T. durum*, landrace, QTL, kernel size, kernel weight

OPEN ACCESS

Edited by:
Thomas Miedaner, University of Hohenheim, Germany

Reviewed by:
Conxita Royo, Institute of Agrifood Research and Technology (IRTA), Spain
Ildikó Karsai, Centre for Agricultural Research (MTA), Hungary

****Correspondence:***
Francesca Desiderio
francesca.desiderio@crea.gov.it
Elisabetta Mazzucotelli
elisabetta.mazzucotelli@crea.gov.it

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *15 January 2019*
Accepted: *25 March 2019*
Published: *18 April 2019*

Citation:
Desiderio F, Zarei L, Licciardello S, Cheghamirza K, Farshadfar E, Virzi N, Sciacca F, Bagnaresi P, Battaglia R, Guerra D, Palumbo M, Cattivelli L and Mazzucotelli E (2019) Genomic Regions From an Iranian Landrace Increase Kernel Size in Durum Wheat. Front. Plant Sci. 10:448. doi: 10.3389/fpls.2019.00448

INTRODUCTION

Durum wheat (*T. turgidum* L. var. *durum*) is a major crop in Mediterranean regions with a total of about 14 million hectares cultivated worldwide. Commercial wheat cultivars have a rather narrow genetic base (Van de Wouw et al., 2010) therefore investigation and exploitation of new genetic diversity is a fundamental requirement for modern breeding programs. Landraces, the locally adapted germplasm as result of the natural and farmers' selection,

represent interesting genetic materials, they usually exhibit a high genetic diversity with relevant allele variations including rare variants and/or potentially new alleles (Lopes et al., 2015).

The development of high yielding wheat cultivars is a major objective of modern breeding programs. Since grain yield is a complex trait, it is often dissected in two main components that are kernel weight, expressed as 1,000 kernel weight (TKW), and number of seeds per square meter resulting from the number of spikes per unit area and number of kernels per spike. Kernel dimensions, as kernel length (KL) and kernel width (KW), greatly influence the TKW. Moreover, especially for durum wheat, kernel size and shape also influence the test weight (TW), which, in turn, has an effect on semolina yield (Gegas et al., 2010). For these reasons, increasing TKW and TW are main targets in wheat breeding, in addition to total yield. Larger kernels not only impact on grain yield but also have favorable effects on seedling vigor and early growth (Peng et al., 2003). These traits are quantitative and complex, highly influenced by the environment (E) and display high Genotype × Environment interactions (GxE). Modern durum wheat varieties exhibit large kernels and rather uniform seed size, because of domestication and breeding for increased yield and TW. On the contrary, durum wheat landraces show a much greater variability for kernel size and shape (Moore, 2015; Liu et al., 2017).

The understanding of the genetic and molecular determinants of grain size and grain shape might provide valuable information on genetic diversity and corresponding markers to be used for improving grain yield. The most advanced genetic knowledge on the genetic factors controlling grain size and shape is available in rice where many genes have been functionally characterized. An update about genetic pathways controlling kernel size and weight in rice and Arabidopsis has been recently reported in Li and Yang (2017). Some genes (for instance: *D1*, *D61*, and *SRS5*) have pleiotropic effects on organ growth, including a reduction in seed size in the corresponding mutants, due to alteration of phytohormones signaling (Yamamuro et al., 2000; Ashikari et al., 2005; Segami et al., 2012). Others (for instance: *GW2*, *GS2*, *GS5*, *GLW7, GIF1*) appear to specifically affect grain morphology (Song et al., 2007; Wang et al., 2008; Li et al., 2011; Hu et al., 2016; Si et al., 2016). At the cellular level, increase of the grain size could be a consequence of an increase in cell number, such as for the activity of *D1* and *GS5*, or of cell size expansion, as for the role of *D61* and *GLW7*, or of both as observed for *GS2*.

The direct translation of genetic knowledge gained from rice to wheat allowed the identification of several orthologs. As in rice, *TaGW2*, encoding an E3 RING ligase (Su et al., 2011; Simmonds et al., 2016), is a negative regulator of grain size and weight (Hong et al., 2014), and showed natural allelic variation in extensive studies in both tetraploid and hexaploid wheat (Su et al., 2011; Qin et al., 2014; Jaiswal et al., 2015; Simmonds et al., 2016). Similarly, allelic variation at *TaGS-D1,* the wheat homolog of the rice *GS3* (Wang et al., 2012), and at *TaTGW6*, an enzyme related to the auxin metabolism (Hanif et al., 2015; Hu et al., 2016), showed main effects on TKW and kernel size. *TaGS5* is a positive regulator of grain size (Ma et al., 2016) and *TaCwi*, homolog of *GIF1*, encodes a cell wall invertase with effects on TKW (Jiang et al., 2015). Other genes have been found in wheat as related to kernel weight, they include *TaSAP-A1*, *TaGS1a*, *6-SFT-A2*, *TaSus1*, and *TaSus2* (Jiang et al., 2011; Chang et al., 2013; Guo et al., 2013; Hou et al., 2014; Yue et al., 2015). All these genes except *TaSAP-A1* have specific roles during the grain filling.

Many studies have been conducted to identify quantitative trait loci (QTL) associated to kernel traits, TKW above all, but also parameters related to kernel size in common wheat (Sun et al., 2008; Gegas et al., 2010; Ramya et al., 2010; Tsilo et al., 2010; Prashant et al., 2012; Maphosa et al., 2014; Rasheed et al., 2014; Williams and Sorrells, 2014; Wu et al., 2015; Kumar et al., 2016; Cheng et al., 2017; Su et al., 2018; Würschum et al., 2018). In these studies, some QTLs for TKW co-localized with QTL of kernel size, thus confirming also at genetic level the positive correlation between grain size and grain weight. Furthermore, a co-location of yield related traits was also found with QTL for flowering time and plant height suggesting pleiotropic effects on fundamental agronomic traits (Gegas et al., 2010; Bogard et al., 2011). In tetraploid wheat, only two studies unravel the genetic bases of kernel size (Russo et al., 2014; Golan et al., 2015), but much more identified regions related to kernel weight (Maccaferri et al., 2016; Kidane et al., 2017; Roncallo et al., 2017; Soriano et al., 2017; Mangini et al., 2018). All findings have been collected by a global metaQTL analysis which summarized and projected all known QTLs on the durum wheat reference genome providing a tool for comparison between QTLs and candidate genes (cv Svevo; Maccaferri et al., 2019).

The current study was designed to identify novel regions of the durum wheat genome controlling kernel related traits in a RIL population derived from a cross between the Iranian cultivar Zardak and the Iranian landrace Iran_249. For this purpose, we developed a high-density genetic map, and conducted a QTL mapping whose results were physically mapped on the recently published durum wheat reference genome (Maccaferri et al., 2019). The results provide the physical position of QTLs directly on the durum wheat pseudomolecules and a list of candidate genes laying within the QTL confidential regions.

MATERIALS AND METHODS

Genetic Materials

A population of 118 F_{7-8} RILs, derived from a cross between the landrace accession "Iran_249" originated from Western Iran, and the old cultivar "Zardak" from the Iranian Kermanshah province, was used in the current study. A leaf of each line was ground using the Retsch_MM300 Mixer Mill instrument, then the DNA was isolated and purified with the Wizard_Magnetic 96 DNA Plant System (Promega) following the manufacturer's instructions.

Field Experiments and Phenotypic Evaluation

Seed increase was done in the experimental farm of the CREA-Research Centre for Genomics and Bioinformatics in Fiorenzuola d'Arda (Italy). The RIL population and the two parents were evaluated in Libertinia (Sicily island, southern Italy) during the 2013–2014 (L14) and 2014–2015 (L15) seasons, and in

TABLE 1 | Kernel morphological traits evaluated through image analysis by WinSeedle software.

Traits	Description
Length (L, mm)	Line connecting the two farthest points on the projected image perimeter
Width (W, mm)	The maximum width perpendicular to length
Perimeter (P, mm)	The length of an object's projected area boundary
Area (A, mm^2)	The two-dimensional projected area of a three-dimensional object
Curvature (C)	Defined as (a/b), where (a) is a perpendicular distance from the center of the object at the point of maximum straight width to the straight length and (b) is the straight length
WL ratio (WL)	Width to length ratio
Form coefficient (FC)	Is 4*A/P2 where A = cell area and P = cell perimeter. It can take values between 0 and 1, 1 being a perfect circle and 0 a filiform object (perfect line)

Fiorenzuola d'Arda (northern Italy) in 2014–2015 (F15), thus providing phenotypic data for three environments.

In each environment, a randomized complete block design with three replications was used; the experimental units consisted of 1.8 m^2 in Libertinia and 3 m^2 plot in Fiorenzuola d'Arda. Trials were fertilized following the standard agronomic practices for each location, weeds were chemically controlled. **Supplementary Table 1** reports the details about field experimental conditions and relevant environmental parameters. Heading date (HD) was recorded as number of days from the April 1st to the time when 50% of tillers within a plot have spike emerged from the flag leaf sheet. Test Weight (TW) was recorded for each plot/environment. After several months of storage at constant temperature and humidity, three samples of 100 kernels were randomly chosen from the seed bulk of each plot/experiment and weighted to calculate the corresponding Thousand Kernel Weight (TKW). One 100 kernel punch for each plot/experiment was randomly sampled out and used for batch scanner imaging. Then through image analysis by the software Winseedle pro[1] (2011 Regents Instruments Inc., Canada) kernels were measured for several descriptors of seed morphology as reported in **Table 1**.

Statistical Analysis of Phenotypic Data

For each environment and trait, the frequency distribution of the RIL phenotypic data was evaluated and analysis of variance (ANOVA) was performed. Overall data were analyzed by fitting a model by the REstricted Maximum Likelihood (REML) method to assess significance of Genotype (G), environment (E), and Genotype × Environment interaction (GxE). Broad sense heritability (H) was calculated according to Nyquist (1991): $H = \delta^2_G/[\delta^2_G + \delta^2_{GE}/E) + \delta^2_e/rE)]$, where δ^2_G is the genetic variance, δ^2_{GE} is the GxE interaction variance, δ^2_e is the residual variance, E is the number of environments, and r the number of replicates. δ^2_G was calculated as $(MS_G - MS_{GxE})/n$ where MS_G is the genotype mean square and MS_{GxE} is the mean square of GxE. All these statistical analyses were conducted by using JMP version 7 software (SAS Institute Inc., 2007).

Pearson's correlation coefficients were calculated for all trait combinations based on data recorded for each year/environment, and using overall data across the three environments using the standard cor.test function in R. The significance of correlations was assessed with the *t*-test implemented in the cor.test function.

For each trait, QTL analysis was performed based on mean values of the three replicates for each single environment.

Molecular Marker Analysis

Both simple sequence repeat (SSR) and single-nucleotide polymorphism (SNP) molecular markers were used to analyze the parental lines and the RILs.

The parental lines were screened with a total of 360 SSR markers selected from the published wheat map (Röder et al., 1998; Eujayl et al., 2002; Gupta et al., 2002; Guyomarc'h et al., 2002; Sourdille et al., 2003; Peng and Lapitan, 2005; Song et al., 2005; Xue et al., 2008). The PCR and fragment analysis were carried out as described in Desiderio et al. (2014).

Genotyping for SNPs was performed at the Trait Genetics Laboratory (Gatersleben, Germany) with the Infinium iSelect 90K wheat SNP BeadChip array (Illumina Inc., San Diego, USA), which contains 81,587 functional markers (Wang et al., 2014).

Linkage Analysis

Linkage analysis was performed using CarthaGene software (de Givry et al., 2005) with a logarithm of odds (LOD) score threshold of 9.0, maximum distance of 20 cM and the Kosambi mapping function to calculate map distances (Kosambi, 1944). The linkage groups obtained were assigned to chromosomes by comparing markers of the generated maps to the high-density consensus durum map (Maccaferri et al., 2015). Within each linkage group, the best order of markers and the genetic distances were established using different CarthaGene functions: "build," "greedy," "flips," and "polish." All mapped markers were tested for the expected 1:1 segregation ratio using a Chi squared ($\chi 2$) goodness-of-fit test.

QTL Analysis

QTL mapping was conducted with the R/qtl module of the R statistical computing package (Broman et al., 2003). For each trait, an initial QTL scan was performed using simple interval mapping with a 1-cM step (Lander and Botstein, 1989) and the position of the highest LOD was recorded. A genome-wide significance level of 5% was calculated after 1,000 permutations (Churchill and Doerge, 1994). The position and the effect of the QTL were then estimated using the multiple imputation method (Sen and Churchill, 2001) by executing the "sim.geno" command, followed by the "fitqtl" command. To search additional QTLs, the "addqtl" command was used. If a second QTL was detected, "fitqtl" was used to test a model containing both QTLs and their interaction effect. If both QTL remained significant, the "refineqtl" command was used to re-estimate the QTL positions based on the full model including

[1] Winseedle pro (2011 Regents Instruments Inc., Canada).

both QTLs. QTL interactions were analyzed and the significant locus combinations are reported based on F value. The additive effects of QTLs were estimated as half the difference between the phenotypic values of the respective homozygotes.

The confidence interval (CI) of each QTL was determined as proposed by Darvasi and Soller (1997). The QTLs were named according to the rule "trait.gb + chromosome.locus number."

Analysis of Physical Regions Carrying QTLs Related to Kernel Traits

The most significant QTLs identified in the present study were projected on the *T. durum* reference genome sequence (cv. Svevo) (Maccaferri et al., 2019). Peak markers and flanking markers corresponding to the CIs were located on the reference genome based on the Blast matches of the corresponding SNP's nucleotide sequences. Whenever the marker was a singleton and/or found similarity hit within the unassembled fraction of the Svevo genome (chromosome 0), the marker was searched on the consensus durum map (Maccaferri et al., 2015), the cosegregant markers from the consensus were identified and the corresponding sequence localized by Blast on the Svevo genome. This approach was used to roughly locate the QTLs on the reference genome for three different comparison analyses. Firstly, the likely position of the identified QTLs was compared with that of durum wheat homologs of common wheat and rice cloned genes whose function is known to be associated to kernel related traits. Secondly, the physical position of the identified QTLs was compared with QTLs previously genetically mapped and published in tetraploid wheat for the same traits and recently anchored to the durum reference genome by Maccaferri et al. (2019). Finally, the physical region underlined by the most significant QTLs was inspected to identify candidate genes, their functional annotation, and the expression data available for the homologous genes in bread wheat.

Toward this end, durum genes were annotated via blast2GOPRO (Götz et al., 2008) using as queries proteins run against *viridiplantae* database (NCBI non-redundant protein dataset; available at FIGSHARE (https://figshare.com/s/2629b4b8166217890971).

Next, best reciprocal hit (BRH) blasts of durum wheat (cv. Svevo) CDS queries (longest representative isoforms for each gene in the physical region of interest) were conducted against a database consisting of bread wheat (Chinese Spring) CDS (longest representative isoforms for each gene; only genes located in the chromosome homologous to Svevo's query genes chromosome). The bread wheat best hits (filtered for a percent identity threshold of at least 90%) were subsequently used as queries for blasts (blast2; version 2.2.26) against the WheatExp database at https://wheat.pw.usda.gov/WheatExp/. Blasts were fine-tuned by testing several parameters (gapped vs. ungapped blasts and various score penalties for gap opening and gap extension). Finally, blast results were again filtered for a minimum identity of 90%. The expected chromosome location of hits as well as the consistency of their annotation with respect to original Svevo queries annotated with blast2GOPRO was evaluated. To be able to compare the expression profile of all the genes mapping under a specific QTL and to represent these data into a heat map, the z-scores of the FPKM log mean values have been calculated.

FIGURE 1 | Kernel morphology of Iran_249 **(a)** and Zardak **(b)**.

RESULTS

Phenotypic Evaluation

The two parents, the cultivar Zardak and the landrace Iran_249, and the RILs were evaluated for traits related to kernel morphology and size (**Table 1**), grain weight, and for heading time in 3 years × environment combinations (Libertinia 2014 -L14- and 2015 -L15- Fiorenzuola d'Arda 2015 -F15).

Mean values of Zardak, Iran_249, and RILs across the three environments are reported in **Table 2**, single environment data are in **Supplementary Table 2**. The two parents showed significant differences for kernel length, perimeter, area and shape related traits (WL, FC) in all environments, while for TKW only in L15. In detail, kernels of Iran_249 were longer and not significantly but generally narrower and heavier compared to those of Zardak *cv* (**Figure 1**; **Table 2**), while kernels of Zardak had a higher degree of roundness (FC, WL) and a higher test weight, as a consequence. Data about the RILs population showed a continuous variation and a normal distribution for most of the traits, suggesting a polygenic inheritance (**Figure 2**). For kernel width, transgressive segregation was observed in both directions, while for kernel length, perimeter and area only RILs with kernels shorter/smaller than the worse parent were present within the population. Consequently, RILs producing kernels with a roundness degree higher than the better parent were reported. Interestingly, some RILs showed values higher than the better parent for TKW, and TW. Overall, this evidence suggests the presence of superior QTL alleles for TKW and TW in both parents, likely supported by larger kernels and higher grain roundness.

Variation for the phenotypic measures was assessed by ANOVA for each single environment and for the overall dataset, evaluating the effects of G, E and GxE (**Table 2**; **Supplementary Table 3**). In each single environment the variability for replications was significant for almost all traits, but much more of the variation was attributed to the genotype effect. Considering overall data across the three environments, all effects were significant for all traits. However, although the strong environment effect, the genotype variability was higher than GxE component for all traits, with the exception of the

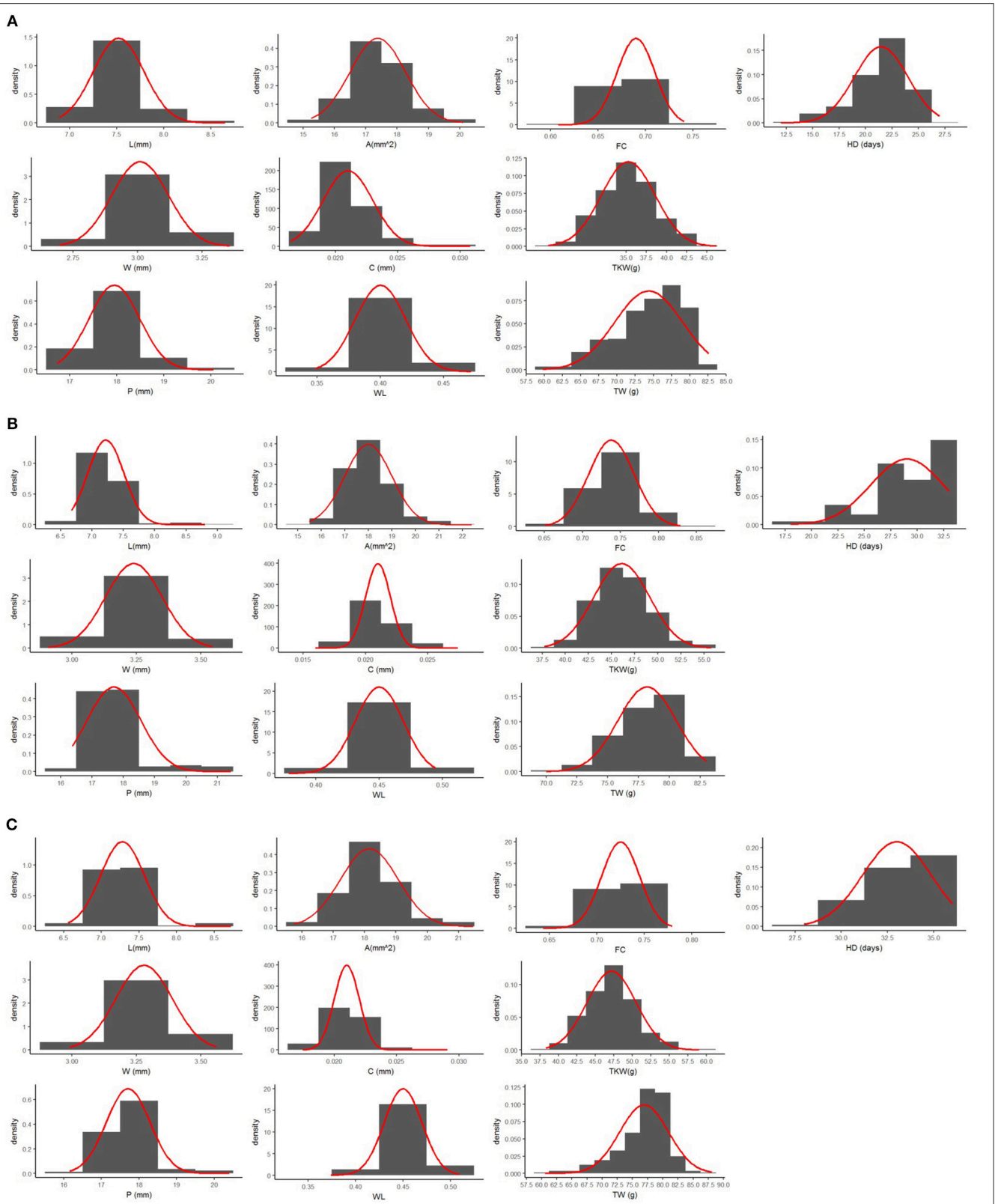


FIGURE 2 | Frequency distribution for all the traits analyzed in this study. The normal distribution was represented as solid red line. Mean data for each environment **(A)** L14, **(B)** L15, and **(C)** F15 have been reported.

TABLE 2 | Summary of phenotypic data and variation parameters for parental lines and RILs for kernel shape (C, curvature; WL, width length ratio; FC, Form coefficient), size (L, length; W, width; P, perimeter; A, area), and weight related traits (TKW, Thousand kernel weight; TW, Test weight), and heading date (HD).

Trait	Iran_249	Zardak	*P*-value*	RIL mean	RIL min	RIL max	CV%	MS_G	MS_{GxE}	H
L	10.02	7.581	*	7.339	6.706	8.726	1.6	0.577	0.030	0.94
W	3.09	3.185	ns	3.179	2.865	3.489	2.2	0.040	0.013	0.65
P	22.74	18.169	*	17.790	16.420	20.631	2.2	2.349	0.358	0.82
A	23.08	18.505	*	17.844	15.502	21.319	3.1	4.099	0.758	0.79
C	0.022	0.021	ns	0.021	0.017	0.029	1.8	1.09 E-05	3.2 E-06	0.66
WL	0.309	0.420	*	0.434	0.367	0.492	1.8	2.64 E-03	1.72 E-04	0.92
FC	0.57	0.702	*	0.713	0.634	0.783	2	2.77 E-03	3.2 E-04	0.86
TKW	48.93	43.136	ns	42.843	33.816	53.640	5	34.36	11.77	0.62
TW	64.39	73.167	*	76.453	63.467	84.567	4.7	23.31	13.84	0.34
HD	29.7	28.3	ns	27.8	19.33	32	4.8	37.18	3.38	0.88

Mean data across the three environments have been reported.

**Significant difference at 0.05% among parents based on t-test; CV, coefficient of variation; MS_G, genotype mean square; MS_{GxE}, mean square GxE; H, broad sense heritability.*

TW. As a consequence, high values of broad sense heritability were obtained for kernel size and shape related traits, ranging from 0.65 to 0.94, with generally lowest values for kernel width and curvature, and highest values for kernel length (**Table 2**). Moderate to low heritability values were calculated for TKW and TW (0.62 and 0.34, respectively). Finally, a high heritability (0.88) was found for heading date. Based on the highly significant GxE interaction showed by some traits, QTLs were determined using the mean values of the three replicates for each environment.

Pearson's correlation coefficients between all possible couple of target traits were calculated based on both single-environment data (**Supplementary Table 4**; **Supplementary Figure 1**), and overall dataset of the three environments (**Supplementary Table 4**; **Figure 3**). As expected, some traits were inherently correlated, like perimeter vs. length ($r = 0.97$), and WL vs. FC ($r = 0.99$). TKW and TW showed a high positive correlation with kernel width ($r = 0.98$ and $r = 0.7$, respectively), and kernel roundness as showed by WL ($r = 0.88$ and 0.82) and FC traits ($r = 0.88$ and $r = 0.83$), while they had negative correlations with seed length ($r = -0.66$ and -0.8, respectively). However, considering single environment data, a significant positive correlation was found between L and TKW in L15. Finally, HD was negatively correlated to L and P (r-values ranging from -0.6 to -0.7) and positively associated to traits about kernel width (W, WL, $r = 0.88$ and 0.85, respectively), and kernel weight related traits (TKW: $r = 0.93$, TW: $r = 0.68$).

Molecular Analysis and Map Construction

The Zardak × Iran_249 genetic linkage map integrated both SSR and SNP markers. Out of 360 SSRs used to screen the parental lines, 87 (24%) were polymorphic between parents and were tested on the whole segregating population. Within the 81,587 markers of the 90k iSelect Infinium, 5,591 SNPs failed the hybridization and were discarded, while 8,220 (10.8%) were polymorphic between the two parents. Within the polymorphic marker set, we further removed markers showing more than 10% missing values and markers with a minor allele frequency (MAF) significantly deviating from the expected 1:1 ratio (MAF < 0.3). After these checks, 6,452 high-quality SNP markers represented the valuable SNP data set. On the overall, 6,539 polymorphic loci (comprising 87 SSR and 6,452 SNP markers) were therefore identified and used for the construction of the molecular marker map. After elimination of the unlinked loci, the genotype data relating to 6,195 informative marker loci were assembled into 18 linkage groups corresponding to the 14 durum wheat chromosomes, for a total of 977 unique loci (**Table 3** and **Supplementary Table 5**). Two linkage groups were identified for chromosomes 1B, 2B, 6A and 6B.

The overall length of the map was 2,884.5 cM with individual chromosome genetic length ranging from 314.7 cM (chromosome 3A) to 117.4 cM (chromosome 6B) and average chromosome length of 206.04 cM. The total number of mapped loci per chromosome ranged from 196 (chromosome 4A) to 794 (chromosome 1B) with an average of 442.5 loci. The genome-wide marker density was 0.47 cM, varying from 0.21 cM (chromosome 6B) to 0.85 cM (chromosome 4A).

Considering the two sub-genomes (A and B), genome B showed a higher number of loci (3,643) and a higher marker density (mean of 0.36 cM/marker), while genome A the longer map length (1,567.4 cM, **Table 3**).

QTL Mapping Analysis for Kernel Size

QTL analysis was performed for traits related to kernel size, shape and weight, and HD, using phenotypic data from single environments (L14, L15 and F15). Overall, 94 QTLs distributed on all chromosomes were identified, in addition to 16 epistatic interactions (**Tables 4A,B**). Chromosomes 6B and 2B reported the highest QTL frequency (24 and 19, respectively). The kernel length identified the highest number of QTLs (17), followed by WL ratio (14), and perimeter (11). QTLs for the same trait, identified in different environment and with overlapping CIs or QTL peak at < 20 cM were considered the same (Maccaferri et al., 2008). Upon this merging, we identified a total of 31 different QTLs and 9 QTL interactions for kernel size (L, W) and the correlated measures (P, A), and 21 different QTLs and 5 interactions for kernel shape (C, WL, and FC; **Table 4**).

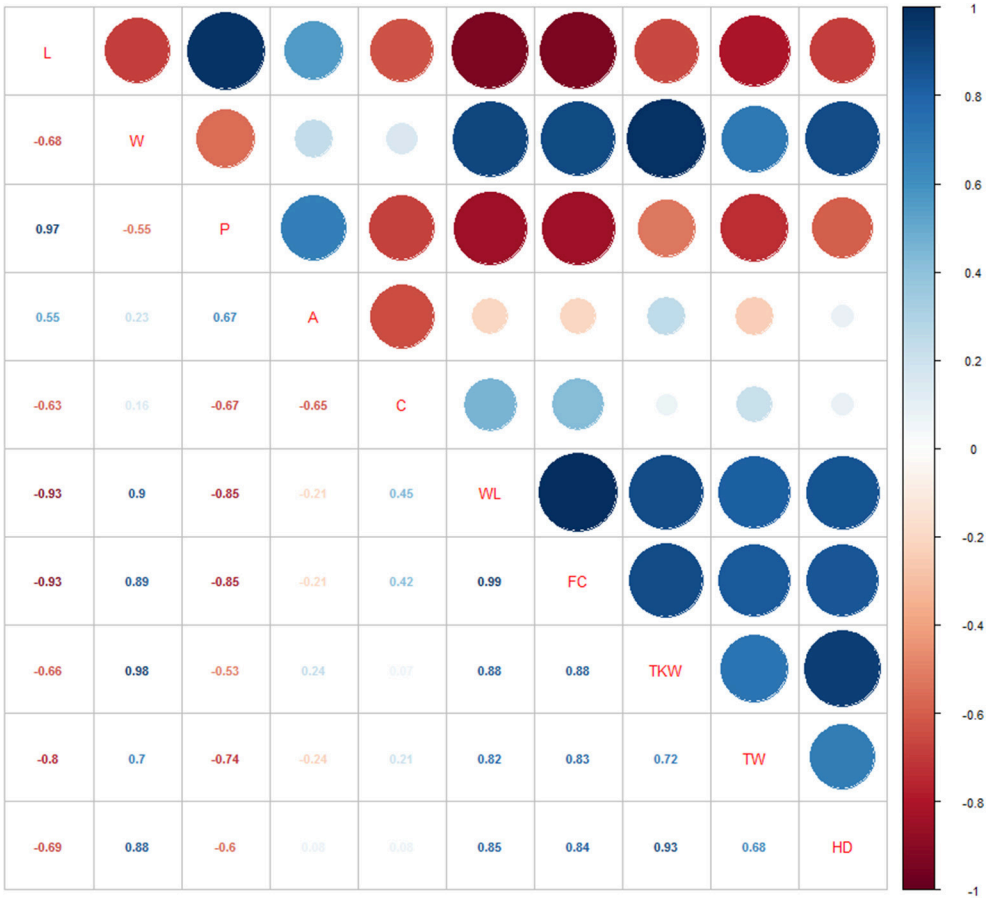


FIGURE 3 | Pearson correlations among the phenotypic traits analyzed using overall data.

TABLE 3 | Distribution of molecular markers in the chromosomes and in the homeologous groups of the Zardak × Iran249 map.

Chromosome	N° of linkage group	Total marker	Map lenght (cM)	Marker density (cM/marker)
1A	1	359 (76)	231.7	0.65 (3.05)
1B	2	794 (109)	246.2	0.31 (2.26)
2A	1	390 (59)	189.4	0.49 (3.21)
2B	2	529 (78)	236	0.45 (3.03)
3A	1	456 (72)	314.7	0.69 (4.37)
3B	1	471 (62)	155.7	0.33 (2.51)
4A	1	196 (41)	166.9	0.85 (4.07)
4B	1	296 (52)	165.1	0.56 (3.18)
5A	1	366 (63)	228	0.62 (3.62)
5B	1	519 (81)	229.9	0.44 (2.84)
6A	2	469 (66)	186.8	0.40 (2.83)
6B	2	563 (67)	117.4	0.21 (1.75)
7A	1	316 (73)	249.9	0.79 (3.42)
7B	1	471 (78)	166.8	0.35 (2.14)
Total	18	6,195 (977)	2884.5	0.47 (2.95)
Genome A	8	2,552 (450)	1567.4	0.61 (3.48)
Genome B	9	3,643 (527)	1317.1	0.36 (2.5)

Within brackets, we reported the number of non-cosegregant markers and the marker density calculated accordingly.

TABLE 4 | QTLs **(A)** and their interactions **(B)** detected in Zardak × Iran249 segregating population for traits related to kernel morphology (L, W, P, A, C, WL, FC), kernel weight (TKW, TW), and heading date (HD).

QTL name	Traits	Environment	Chr.	LG	cM	LOD	R^2 (%)	CI start (cM)	CI end (cM)	Additive effect
(A)										
QA.gb-2A	A	F15	2A	1	177.6	5.29	13.52	172.50	182.70	−0.35
QA.gb-2B	A	F15	2B	1	47.3	11.78	34.44	45.30	49.30	−0.56
QA.gb-2B	A	L15	2B	1	57.5	5.40	11.10	51.30	63.70	−0.28
QA.gb-5B.1	A	L15	5B	1	85.2	8.30	18.10	81.40	89.00	−0.36
QA.gb-5B.2	A	F15	5B	1	179.5	4.11	10.27	172.80	186.20	−0.29
QA.gb-6B.1	A	L15	6B	1	0.5	4.35	15.26	0.00	4.99	−0.37
QA.gb-6B.2	A	L14	6B	1	53.8	7.63	24.15	50.96	56.64	−0.56
QA.gb-6B.2	A	F15	6B	1	54.2	4.20	14.46	49.46	58.94	−0.10
QA.gb-6B.3	A	L14	6B	2	8.6	3.94	11.58	2.69	14.51	0.17
QA.gb-6B.4	A	F15	6B	2	35.1	3.52	11.95	29.37	40.83	−0.28
QC.gb-1A	C	L15	1A	1	93.8	3.57	10.37	87.10	100.50	0.00
QC.gb-1B	C	F15	1B	1	27.9	7.35	18.88	24.20	31.60	0.00
QC.gb-2B	C	L14	2B	1	37.5	6.52	18.97	33.90	41.10	0.00
QC.gb-2B	C	F15	2B	2	30.4	4.57	11.09	24.20	36.60	0.00
QC.gb-6B	C	F15	6B	2	5.5	3.50	12.78	0.14	10.86	0.00
QC.gb-6B	C	L15	6B	2	6.0	5.81	20.30	2.63	9.37	0.00
QC.gb-6B	C	L14	6B	2	11.8	4.52	16.17	7.56	16.04	0.00
QFC.gb-1B.1	FC	F15	1B	1	2.7	7.09	14.07	0.00	7.60	0.01
QFC.gb-1B.2	FC	L14	1B	1	144.9	17.34	30.27	142.60	147.20	0.00
QFC.gb-2B	FC	F15	2B	1	61.9	6.48	12.70	56.50	67.30	0.01
QFC.gb-2B	FC	L14	2B	1	76.2	3.15	4.13	59.50	92.90	0.00
QFC.gb-6B.1	FC	L14	6B	1	0.0	5.49	19.29	0.00	3.55	0.01
QFC.gb-6B.1	FC	F15	6B	1	0.0	4.27	12.91	0.00	5.30	0.01
QFC.gb-6B.2	FC	F15	6B	2	2.0	3.37	10.00	0.00	8.85	0.01
QFC.gb-6B.3	FC	L15	6B	2	34.7	2.37	8.83	26.94	42.46	0.01
QFC.gb-7A	FC	L14	7A	1	20.1	12.23	19.31	16.52	23.68	0.07
QFC.gb-7B	FC	L15	7B	1	60.3	4.36	13.52	55.20	65.40	0.01
QFC.gb-7B	FC	F15	7B	1	77.5	9.34	20.15	74.10	80.90	0.01
QHD.gb-2A	HD	L14	2A	1	61.8	10.91	14.64	57.10	66.50	−1.23
QHD.gb-2A	HD	F15	2A	1	61.8	9.22	15.65	57.40	66.20	−0.80
QHD.gb-2B	HD	L15	2B	1	19.8	15.80	40.65	18.10	21.50	−1.80
QHD.gb-2B	HD	L14	2B	1	24.1	7.85	9.89	17.10	31.10	−0.62
QHD.gb-2B	HD	F15	2B	1	24.1	9.66	16.54	19.90	28.30	−0.83
QHD.gb-3B	HD	L15	3B	1	140.0	7.79	16.95	135.90	144.10	−1.09
QHD.gb-4B	HD	F15	4B	1	19.3	6.13	9.77	12.20	26.40	−0.57
QHD.gb-5A	HD	F15	5A	1	208.7	5.57	8.77	200.80	216.60	0.55
QHD.gb-5B	HD	L14	5B	1	86.5	10.08	13.30	81.30	91.70	−1.05
QHD.gb-5B	HD	L15	5B	1	86.5	3.67	7.34	77.10	95.90	−0.67
QHD.gb-7A	HD	L14	7A	1	20.1	8.67	11.10	13.88	26.32	−0.91
QL.gb-2B.1	L	L15	2B	1	57.5	3.60	7.14	47.80	67.20	−0.07
QL.gb-2B.2	L	L14	2B	1	78.5	6.86	8.47	70.30	86.70	−0.07
QL.gb-2B.2	L	F15	2B	1	78.5	13.45	9.40	71.20	85.80	−0.09
QL.gb-4A.1	L	L14	4A	1	23.2	7.85	9.90	16.20	30.20	−0.11
QL.gb-4A.1	L	F15	4A	1	23.2	12.64	8.68	15.20	31.20	−0.10
QL.gb-4A.2	L	L14	4A	1	154.5	9.91	13.05	149.20	159.80	0.13
QL.gb-4A.2	L	F15	4A	1	154.5	30.03	30.33	152.20	156.80	0.16
QL.gb-4B	L	L14	4B	1	110.7	7.00	8.69	102.80	118.60	−0.07

(Continued)

TABLE 4 | Continued

QTL name	Traits	Environment	Chr.	LG	cM	LOD	R^2 (%)	CI start (cM)	CI end (cM)	Additive effect
QL.gb-4B	L	F15	4B	1	110.7	15.47	11.28	104.60	116.80	−0.13
QL.gb-5B	L	L15	5B	1	205.8	5.80	12.10	200.10	211.50	0.09
QL.gb-6B.1	L	L15	6B	1	0.0	5.89	18.01	0.00	3.80	−0.13
QL.gb-6B.1	L	F15	6B	1	0.5	8.47	27.42	0.00	3.00	−0.18
QL.gb-6B.1	L	L14	6B	1	1.0	5.68	14.77	0.00	5.64	−0.11
QL.gb-6B.2	L	L14	6B	2	2.0	6.83	18.18	0.00	5.77	−0.12
QL.gb-6B.3	L	F15	6B	2	8.2	3.68	10.82	1.87	14.53	0.04
QL.gb-7B	L	L14	7B	1	63.8	13.46	19.08	60.20	67.40	−0.17
QL.gb-7B	L	F15	7B	1	63.8	20.28	16.41	59.60	68.00	−0.16
QP.gb-2A	P	F15	2A	1	32.4	9.27	20.12	29.00	35.80	−0.35
QP.gb-2B.1	P	F15	2B	1	57.5	6.33	12.93	52.20	62.80	−0.22
QP.gb-2B.1	P	L14	2B	1	61.9	4.62	9.49	54.60	69.20	−0.16
QP.gb-2B.2	P	L15	2B	1	174.2	12.92	32.30	172.10	176.30	−0.43
QP.gb-3A	P	L15	3A	1	134.2	3.45	7.10	124.50	143.90	−0.20
QP.gb-6B.1	P	L15	6B	1	0.0	6.24	21.51	0.00	3.18	−0.32
QP.gb-6B.1	P	L14	6B	1	1.0	5.41	14.68	0.00	5.67	−0.22
QP.gb-6B.1	P	F15	6B	1	1.0	4.94	17.54	0.00	4.90	−0.26
QP.gb-6B.2	P	L14	6B	2	2.0	5.84	16.00	0.00	6.28	−0.21
QP.gb-7A	P	L15	7A	1	108.7	6.86	15.13	113.26	113.26	−0.21
QP.gb-7B	P	L14	7B	1	55.9	5.78	11.92	50.10	61.70	−0.18
QTKW.gb-1B	TKW	L14	1B	1	66.5	7.63	22.46	63.40	69.60	1.08
QTKW.gb-2B	TKW	L15	2B	1	46.8	3.90	10.00	39.90	53.70	−0.87
QTKW.gb-3B	TKW	F15	3B	1	130.1	5.82	17.60	126.20	134.00	−1.32
QTKW.gb-5B	TKW	L14	5B	1	75.2	5.41	15.21	70.70	79.70	−0.86
QTKW.gb-5B	TKW	F15	5B	1	75.2	4.19	12.27	69.60	80.80	−1.00
QTKW.gb-5B	TKW	L15	5B	1	85.2	8.00	22.36	82.10	88.30	−1.26
QTW.gb-6A.1	TW	L15	6A	2	54.1	5.30	18.70	50.40	57.80	−0.75
QTW.gb-6A.2	TW	F15	6A	2	97.4	5.39	18.97	93.80	101.00	1.11
QW.gb-1B	W	L14	1B	1	4.5	9.64	27.66	2.00	7.00	0.04
QW.gb-3A	W	L15	3A	1	175.1	3.10	9.14	167.50	182.70	0.02
QW.gb-5A	W	L14	5A	1	53.6	5.75	15.22	49.10	58.10	0.03
QW.gb-5B	W	L15	5B	1	75.2	5.90	18.45	71.50	78.90	−0.04
QW.gb-6A	W	F15	6A	2	97.4	5.60	19.63	93.90	100.90	0.04
QWL.gb-1B	WL	F15	1B	1	2.7	8.17	13.24	0.00	7.90	0.01
QWL.gb-1B	WL	L14	1B	1	4.5	8.84	12.24	0.00	10.10	0.01
QWL.gb-1B	WL	L15	1B	1	4.5	9.10	17.70	0.60	8.40	0.01
QWL.gb-2B	WL	F15	2B	1	61.9	7.84	12.70	56.50	67.30	0.01
QWL.gb-2B	WL	L14	2B	1	71.3	7.90	10.73	64.90	77.70	0.01
QWL.gb-2B	WL	L15	2B	1	85.6	4.75	8.45	77.40	93.80	0.01
QWL.gb-4A	WL	L14	4A	1	23.2	6.81	9.04	15.60	30.80	0.01
QWL.gb-6A	WL	F15	6A	2	97.0	6.40	10.00	90.10	103.90	0.01
QWL.gb-6B.1	WL	L14	6B	1	0.0	3.37	10.76	0.00	6.36	0.06
QWL.gb-6B.2	WL	L15	6B	2	2.0	5.70	19.94	0.00	5.43	0.01
QWL.gb-6B.2	WL	F15	6B	2	14.9	5.65	19.79	11.44	18.36	0.01
QWL.gb-7B	WL	L14	7B	1	55.9	21.91	40.15	54.20	57.60	0.01
QWL.gb-7B	WL	L15	7B	1	55.9	11.25	22.88	52.90	58.90	0.01
QWL.gb-7B	WL	F15	7B	1	55.9	6.37	9.95	49.00	62.80	0.01

(Continued)

TABLE 4 | Continued

QTL interaction	Traits	Environment	LOD	R^2 (%)	Additive effect
(B)					
QL.gb-4A.2*QL.gb-6B.1	L	L14	8.11	10.29	−0.12
QL.gb-4A.2*QL.gb-6B.1	L	F15	25.9	23.8	−0.18
QL.gb-4B*QL.gb-6B.1	L	L14	6.21	7.57	0.09
QL.gb-5B*QL.gb-6B.1	L	L15	5.2	10.77	−0.1
QP.gb-2B*QP.gb-7A	P	L15	4.63	9.76	0.24
QP.gb-2A*QP.gb-6B.1	P	F15	3.91	7.62	0.23
QC.gb-2B*QC.gb-6B.2	C	L14	2.62	7.04	0.00049
QC.gb-1A*QC.gb-6B.2	C	L15	2.79	7.96	0.00049
QFC.gb-1B.2*QFC.gb-6B.1	FC	L14	11.6	18.06	0.01
QFC.gb-1B.1*QFC.gb-6B.2	FC	F15	3	5.45	−0.005
QFC.gb-1B*QFC.gb-7A	FC	L14	5.43	7.45	0.0053
QHD.gb-2A*QHD.gb-7A	HD	L14	6.17	7.51	−0.82
QHD.gb-2B*QHD.gb-3B	HD	L15	3.81	7.65	−0.9
QA.gb-6B.2*QA.gb-6B.4	A	F15	3.15	10.64	0.343
QL.gb-6B.1*QL.gb-6B.3	L	F15	3.46	10.11	−0.114
QA.gb-6B.2*QA.gb-6B.3	A	L14	3.53	10.28	−0.366

Kernel Length

Ten QTLs were found to be significantly associated with kernel length (L, **Table 4**). Among them, QL.gb-6B.1 was reported in all environments, while other five QTLs were reported in the two environments L14 and F15, on chromosomes 2B, 4B, and 7B, and two QTLs on chromosome 4A (QL.gb-4A.1, QL.gb-4A.2). For all these QTLs, excepted for QL.gb-4A.2, the landrace Iran_249 contributed the allele for longer kernels. Major QTLs were QL.gb-4A.2 and QL.gb-6B.1, which showed up to 30.3 and 27.4% of phenotypic explained variance (PEV), respectively, thus defining confidence intervals narrower than 5 cM. Out of four different epistatic effects among L-related QTLs, QL.gb-6B.1 had environmentally stable relationships with QL.gb-4A.2, which explained further phenotypic variation of 8.1–25.9%.

Kernel Width

Five QTLs were associated with kernel width (W) on chromosomes 1B, 3A, 5A, 5B, and 6A, but none of them were conserved among environments. The region that explained the highest value for LOD and phenotypic variance (9.6 and 27.7%, respectively) was detected on chromosome 1B, QW.gb-1B (**Table 4**), with a confidence interval of 5cM. Zardak contributed the allele for larger kernel for all QTLs, except for the region identified on chromosome 5B based on data from L15 (QW.gb-5B, with $R^2 = 18.45\%$).

Kernel Perimeter

Eight QTLs were identified for kernel perimeter (P) on chromosomes 2A, 2B, 3A, 6B, 7A, and 7B. QP.gb-6B.1 was identified based on phenotypic data recorded from all environments, while QP.gb-2B.1 was reported in two environments (L14 and F15). QP.gb-6B.1 had from 14.7 to 21.5% of PEV and showed epistatic interaction with the region QP.gb-2A. QP.gb-2B.1 phenotypic variation ranged from 9.5 and 12.9% for L14 and F15 data analysis, respectively. However, the major QTL QP.gb-2B.2, with a confidence interval of 4.2 cM and around 32% PEV, was obtained based on L15 phenotypic data only and showed the highest additive effect value (0.43). In addition, this QTL showed epistatic interaction with QP.gb-7A, thus explaining a further 10% quote of PEV. For all these QTLs, the alleles for increased perimeter were contributed by Iran_249.

Kernel Area

Eight QTLs were detected for kernel area (A) on four different chromosomes, 2A, 2B, 5B, 6B, but only the QA.gb-6B.2 was reported in two environments (L14, F15). This QTL explained up to 24.1% of PEV and an additional quote of 10% resulted from the interaction with other two different regions identified on the same chromosome (QA.gb-6B.3 and QA.gb-6B.4). The QTL with the largest effect was identified on chromosome 2B based on F15 data. It was named QA.gb-2B.1 and explained 34.4% of phenotypic variation. The parent landrace Iran_249 contributed the positive allele for all QTLs associated to kernel area, except for QA.gb-6B.3 (**Table 4**).

QTL Mapping Analysis for Kernel Shape

The analysis of three kernel shape parameters, the curvature (C), the WL ratio and the form coefficient (FC), discovered a total of 21 different QTLs and 5 QTL interactions (**Table 4**).

Curvature

Five QTLs were associated with kernel curvature (C), but only QC.gb-6B was stable across the three environments and showed the highest PEV (up to 20.3%). For all QTLs identified, with the only exception of QC.gb-2B.2 detected using data from F15, Zardak positively contributed for increased curvature, as a combination of greater width and/or shorter length.

WL Ratio

Seven QTLs associated with width/length phenotypic variability were found where Zardak contributed the allele with the

positive effect on the target trait (**Table 4**). Notably, QWL.gb-1B, QWL.gb-2B, and QWL.gb-7B were stable across the three environments. In detail, QWL.gb-7B registered the highest LOD and R^2 values based on data from L14 (21.91 and 40.15%, respectively), thus defining a confidence interval of 3.4 cM. QWL.gb-1B had LOD values comprised between 8.17 and 9.1 and explained a phenotypic variation ranging from 12.2 to 17.7%. About the region on chromosome 2B, the data from F15 identified the highest LOD and R^2 values (LOD = 7.9, R^2 = 12.7%). The region QWL.gb-6B.2 conserved across environments L15 and F15 showed till 20% of PEV.

Form Coefficient

Out of nine QTLs associated with FC, three were identified across two environments and located on chromosomes 2B, 6B and 7B (named as QFC.gb-2B, QFC.gb-6B.1, QFC.gb-7B, respectively). The major conserved QTLs were QFC.gb-7B and QFC.gb-6B.1, which explained around 20% PEV each. Overall, the cultivar Zardak contributed the positive allele at all loci, with the only exception of another major QTL, detected on chromosome 1B using data from L14 and explaining 30.3% of phenotypic variance (QFC.gb-1B.2). Additional quote of explained variance was retrieved by the interactions among QTLs detected, particularly for QFC.gb-1B.2 and QFC.gb-6B.1 (R^2 = 18.1%).

QTL Mapping Analysis for Kernel Weight Related Traits and Heading Date

Thousand Kernel Weight

Four QTLs associated with TKW were detected on chromosomes 1B, 2B, 3B and 5B explaining 10–22.5% of PEV (**Table 4**). Notably, the QTL detected on 5B (QTKW.gb-5B) was stable across three environments, explaining 12.3–22.4% of the phenotypic variation. The allele of Iran_249 positively contributed to most of the QTLs.

Test Weight

Only two significant QTLs were found both on chromosome 6A and explaining around 18% of phenotypic variance with positive allele contributed by both parents.

Heading Date

Seven QTLs for heading date (HD) were detected on chromosomes 2A, 2B, 3B, 4B, 5A, 5B, 7A, and three of these were environmentally stable. In detail, the QTL located on 2B (named as QHD.gb-2B) was conserved among the three sites and explained from 9.9 to 40.6% of variation. Other two QTLs on chromosomes 2A (QHD.gb-2A) and 5B (QHD.gb-5B) were stable in two environments and explained up to 15.6% of PEV. For all these regions, the additive effect responsible for late flowering was contributed by Iran_249, with the only exception of QHD.gb-5A.

Cluster of QTLs

Since the parameters used to characterize the kernel are all geometrically or biologically related, we expected to identify coincident loci for different traits. Indeed, the co-localization of QTLs for different traits, proven the coherence about parent providing the QTL additive effect, allowed to define 11 QTL clusters (**Table 5** and **Figure 4**). Clusters included up to 14 QTLs, and the largest clusters were found on chromosomes 2B (cluster 2) and 7B (cluster 11). Regarding cluster 2, the overlapping covered a region that spanned for about 60 cM. We can suppose that this cluster include at least two different associated regions located on chromosome 2B, but based on the resolution of our data they were indistinguishable. Cluster 11 spanned for 30 cM based on two associations for FC whose peaks were located < 20 cM and thus considered the same QTL. Clusters 7 and 8, located on chromosome 6B, were considered as different based on the opposite additive effect values shown by QTL identified for length trait.

Notably, the overlapping of QTL, also supported by correlation between traits, can suggest causative relationships among the different kernel parameters. Coincidences might derive from different parameters describing the same kernel trait, like WL, FC and C, and thus indicate simple relationships. For instance, the chromosome 6B (clusters 6 and 9) hosted coincident QTLs for all three shape related traits, while three regions on chromosomes 1B, 2B, and 7B were identified using both WL and FC data, and defined clusters 1, 2, and 11, respectively. More intriguingly, QTL co-localizations might depend on geometric relationships between primary characters, like L and W, and the secondary traits like WL, P, A and FC, which directly derive from L and W based on geometric formulas (**Table 1**). As an example, cluster 1 on chromosome 1B grouped QTL related to traits W, WL and FC, suggesting that phenotypic variation for kernel shape associated to the cluster 1 might depend on width variation. Contrarily, in clusters 2 (chromosome 2B), 3 (4A) and 11 (7B), the kernel length was the primary trait associated to a QTL together with a WL locus, indicating a main effect of kernel length on the grain shape. Finally, other clusters included kernel size/shape QTLs and regions associated to TKW and TW. These co-localizations, together with significant correlation between traits, can suggest causal relationships between the simplest kernel trait, kernel length, width and shape, and the more complex relevant agronomic traits, TKW and TW, which indirectly depend on kernel size. This kind of coincidences was indeed revealed by cluster 2, 4 and 5. In detail, the cluster 2 included QTLs for both TKW and kernel size and morphology (A, P, L, FC and C). Notably, cluster 4 contains the QTKW.gb-5B stable in three environments and QTLs for W and A, but also for HD, highlighting a possible effect of phenology on kernel weight, through an impact on specific kernel dimension. An interesting coincidence was also found in cluster 5 (6A) between TW, WL and W, as expected based on the known positive impact of kernel roundness on TW. Notably, Iran_249 contributes the positive allele at cluster 2 and 4, respectively through an allele with positive effect for kernel length and width, respectively. This finding suggested that increase of kernel size from the landrace might improve important agronomic traits like TKW.

Analysis of Physical Regions Carrying QTLs Related to Kernel Traits

The recent durum wheat reference genome was used as common framework to compare our results with QTLs related to

TABLE 5 | Clusters of QTLs.

QTL name	Environments	Peak cM	R^2 (%)	CI_start	CI_end	Cluster
QFC.gb-1B.1	F15	2.7	14.07	0	7.6	1
QWL.gb-1B	L14, L15, F15	2.7–4.5	12.24–17.7	0	10.1	
QW.gb-1B	L14	4.5	27.66	2	7	
QC.gb-2B.1	L14	37.5	18.97	33.9	41.1	2
QTKW.gb-2B	L15	46.8	10	39.9	53.7	
QA.gb-2B	L15, F15	47.3–57.5	11.1–34.44	45.3	63.7	
QP.gb-2B.1	L14, F15	57.5–61.9	9.49–12.93	52.2	69.2	
QL.gb-2B.1	L15	57.5	7.14	47.8	67.2	
QL.gb-2B.2	L14, F15	78.5	8.47–9.4	70.3	86.7	
QFC.gb-2B	L14, F15	61.9–76.2	4.13–12.7	56.5	92.9	
QWL.gb-2B	L14, L15, F15	61.9–85.6	8.45–12.7	56.5	93.8	
QL.gb-4A.1	L14, F15	23.2	8.68–9.9	15.2	31.2	3
QWL.gb-4A	L14	23.2	9.04	15.6	30.8	
QW.gb-5B	L15	75.2	18.45	71.5	78.9	4
QTKW.gb-5B	L14, L15, F15	75.2–85.2	12.27–22.36	69.6	88.3	
QA.gb-5B	L15	85.2	18.1	81.4	89	
QHD.gb-5B	L14, L15	86.5	7.34–13.3	77.1	95.9	
QWL.gb-6A	F15	97.0	10	90.1	103.9	5
QW.gb-6A	F15	97.4	19.63	93.9	100.9	
QTW.gb-6A.2	F15	97.4	18.97	93.8	101	
QP.gb-6B.2	L14	2.0	16	0.0	6.3	6
QWL.gb-6B.2	L15, F15	2–14.9	19.85	0.0	18.4	
QFC.gb-6B.2	F15	2.0	10	0.0	8.8	
QL.gb-6B.2	L14	2.0	18.18	0.0	5.8	
QC.gb-6B	L14, L15, F15	5.5–11.8	12.78–20.3	0.1	16.0	
QL.gb-6B.3	F15	8.2	10.82	1.9	14.5	7
QA.gb-6B.3	L14	8.6	11.58	2.7	14.5	
QFC.gb-6B.3	L15	34.7	8.83	26.9	42.5	8
QA.gb-6B.4	F15	35.1	11.95	29.4	40.8	
QWL.gb-6B.1	L14	0.0	10.76	0.0	6.4	9
QP.gb-6B.1	L14, L15, F15	0–1.0	14.68–21.51	0.0	5.7	
QFC.gb-6B.1	L14, F15	0.0	12.91–19.29	0.0	5.3	
QL.gb-6B.1	L14, L15, F15	0–1.0	14.77–27.42	0.0	5.6	
QA.gb-6B.1	L15	0.5	15.26	0.0	5.0	
QFC.gb-7A	L14	20.1	19.31	16.5	23.7	10
QHD.gb-7A	L14	20.1	11.1	13.9	26.3	
QP.gb-7B	L14	55.9	11.92	50.1	61.7	11
QWL.gb-7B	L14, L15, F15	55.9	9.95–40.15	49	62.8	
QFC.gb-7B	L15	60.3–77.5	13.52–20.15	55.2	80.9	
QL.gb-7B	L14, F15	63.8	16.41–19.08	59.6	68	

Each cluster grouped QTLs with overlapping CIs but related to different traits.

kernel traits already published. To this aim, the CI of the most consistent QTLs as well as the extreme positions of the QTL clusters were anchored on the Svevo genome assembly through the projection of the associated markers. Analogously, the nucleotide sequences of all known (bread) wheat genes or rice genes related to kernel morphology/weight were used as Blast queries to identify the durum wheat orthologs and define their physical position on the Svevo pseudomolecules (**Supplementary Table 6**).

The comparison of physical position of QTLs and of these orthologs revealed some interesting overlapping which might suggest worth candidate genes (**Figure 4**; **Table 5** and **Supplementary Table 6**). When anchored to the Svevo genome sequence, the large cluster 2 on chromosome 2B, including a total of 8 QTLs, encompassed several wheat genes or wheat homologs cloned for their effect on kernel size and weight, namely *TaSus2, SRS1, GW7, GLW7* and *D11*. Interesting coincidences were also found on chromosome 4A and 6B where the *TaTGW6* and

TaGS1a felt in cluster 3 and 9, respectively. In addition, QL-gb.4A.2 maps at around 3Mb from the candidate gene *6-SFT-A2* (Yue et al., 2015).

To further support these genes as candidates of mapped QTLs, we checked if possible sequence variations at these genes are represented by SNPs of the Illumina 90K wheat SNP BeadChip array which also proved to be polymorphic and mapped within our population. Although most of the candidates were covered by Illumina 90K SNP markers, a polymorphic marker was found only for *TaGS1*, in detail IWB13090 mapped at 8.2 cM on chromosome 6B (linkage group 2) in the Zardak × Iran_249 genetic map. This finding allowed us to genetically map the gene *TaGS1b* under the QTL cluster 9.

To assess the novelty of our results, we compared the clusters identified in this work with known QTLs for related traits reported in tetraploid wheat species. Firstly, we checked the physical position of our clusters on the durum reference genome together with those of the QTL previously reported for kernel shape and size by Russo et al. (2014) and Golan et al. (2015) and recently physically anchored on the durum wheat reference genome by the whole metaQTL analysis conducted by Maccaferri et al. (2019). This analysis did not reveal any overlapping. Analogously, we checked the coincidence of our clusters with the physical positions of QTLs genetically mapped for weight related traits and HD in previous studies in tetraploid wheat and physically defined by Maccaferri et al. (2019) on the Svevo genome. In this case, coincidences were found for all 11 clusters (**Table 6**). In detail, clusters 2, 4, 5, and 10 included QTLs for TKW, TW and HD located in genome regions where QTLs for the same traits have been already detected. Other clusters (3, 6, 7, 8, 9, 11), which grouped QTLs for kernel morphology and size, co-localized with regions known to be associated with yield related traits, thus remarking the functional/biological relationship between grain size and weight. Finally, for all

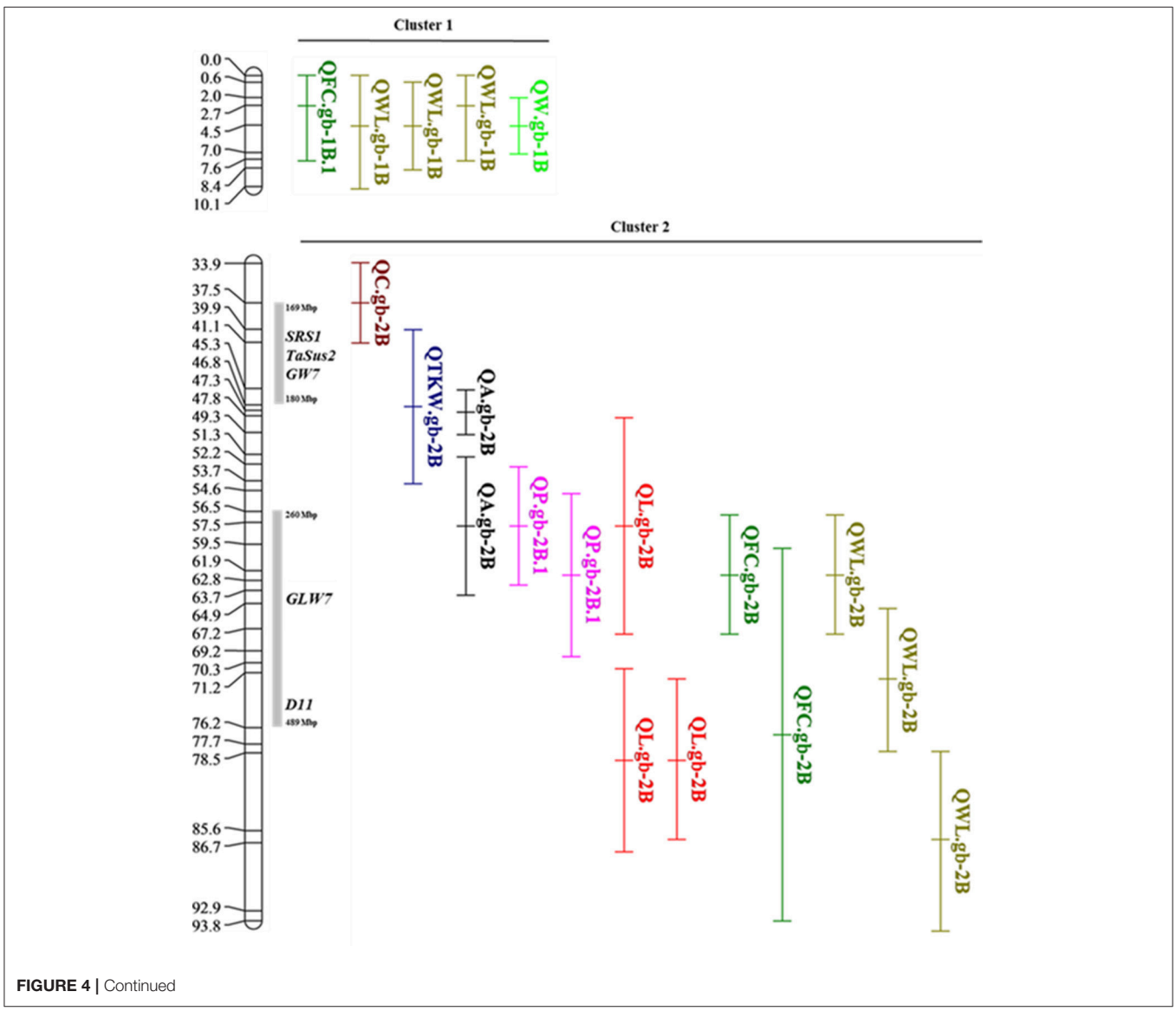


FIGURE 4 | Continued

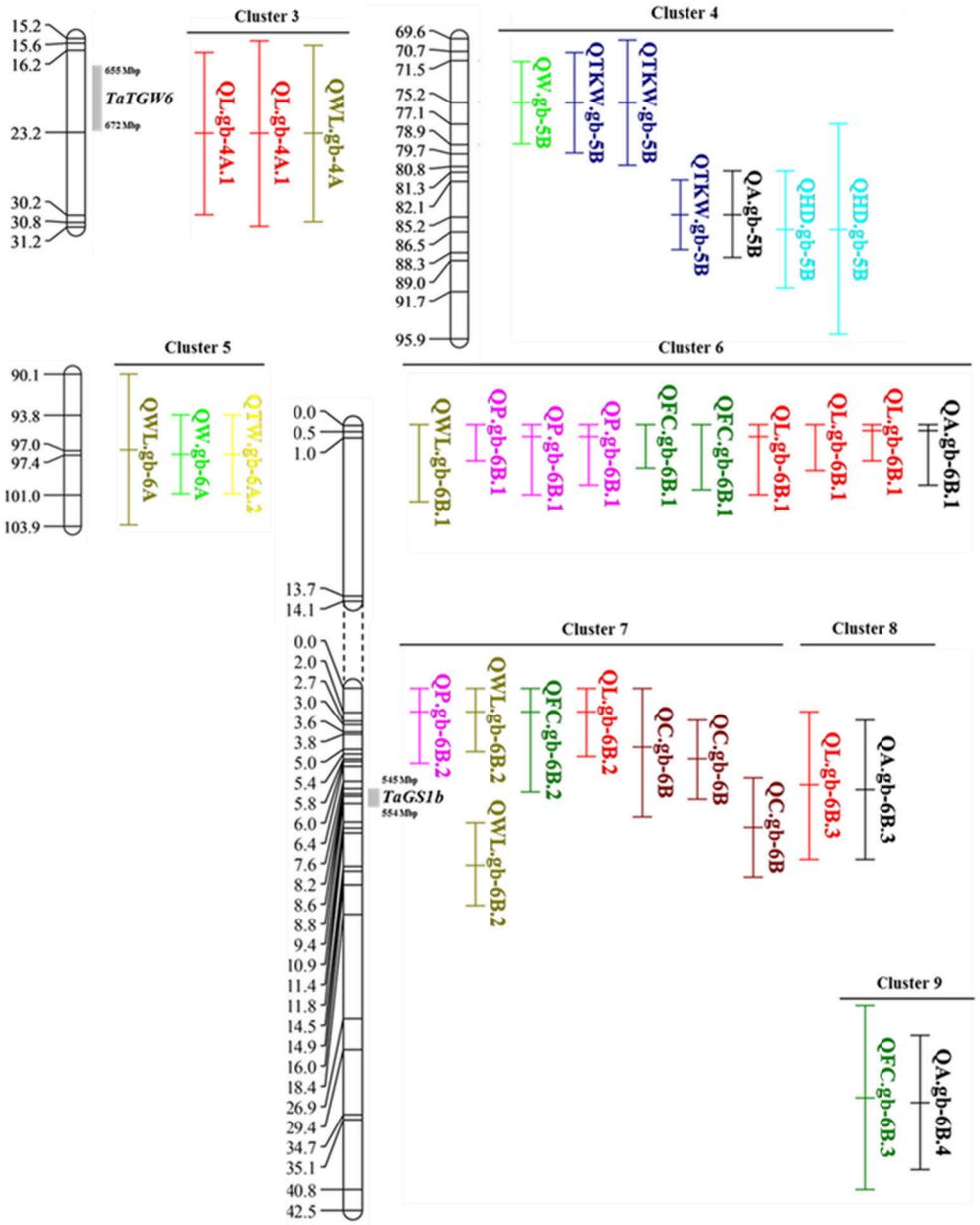


FIGURE 4 | Continued

clusters, except 4 and 9, the correspondence with QTLs for HD was reported suggesting the probable pleiotropic effect of phenology on traits about grain size and weight.

For each cluster, the physical region underlined by the CI of the most consistent QTLs was inspected for candidate genes. To this purpose, we took advantage of the durum wheat reference genome (Maccaferri et al., 2019) together with the expression data available for the orthologous bread wheat genes. All predicted genes on the *T. durum* reference genome were functionally annotated through Blast2Go available at FIGSHARE (https://figshare.com/s/2629b4b8166217890971), while for *T. durum* genes lying under the anchored QTLs the *T. aestivum* ortholog was identified. Expression data of these bread wheat genes were retrieved and reported as a heat map in **Supplementary Table 7**. This approach was expected to support the identification of candidate genes based both on the functional annotation and expression profile in the closely related species *T. aestivum*. Focusing the attention on expression data, grain and spike specific genes have been identified in the genomic regions controlling the following traits: kernel width (chromosome 1B and 6A), kernel length (4B, 6B), kernel area (6B), kernel shape (7B and 7A), and TKW trait (5B).

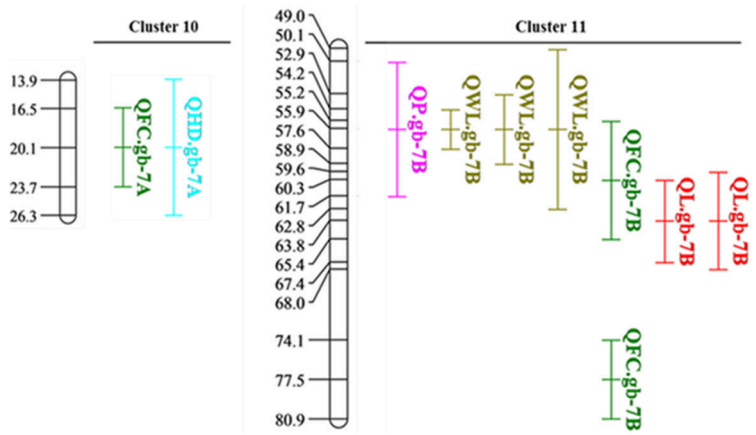


FIGURE 4 | Distribution of co-localized QTLs in clusters detected in this work. Known genes co-localized with our regions have been reported.

TABLE 6 | Co-localization of QTL clusters with known wheat and/or rice genes and known QTLs in tetraploid germplasm related to seed size and shape.

Cluster					Best QTL		Known genes	Tetraploid QTL/MTA
N°	Chr	Start (Mbp)	End (Mbp)	Traits	Start (Mbp)	End (Mbp)		
1	1B	8.2	14.6	FC, WL, **W**	6.8	12.5		HD[1]
2	2B	106.7	536.81	C, TKW, **A**, P, L, FC, WL	188.1	222.8	*TaSUS2, SRS1, GW7, GLW7, D11*	TKW[2], TW[3], HD[4]
3	4A	628.9	686.7	L, WL	\	\	*TaTGW6*	TW[3], HD[3], TKW[5]
4	5B	420.8	489.7	W, **TKW,** A, HD	435.7	452.4		TKW[5,6,7]
5	6A	539.4	576.8	**W**, WL, TW	549.5	590.4		TKW[4,8], TW[3], HD [9,10]
6	6B	2.0	31.3	P, **WL**, FC, L, C	24.4	31.4		TKW[11,12], TW[13], HD[3,9]
7	6B	15.9	26.9	L, **A**	15.9	26.8		TKW[11,12], TW[13], HD[14,9,3]
8	6B	41.8	149.6	FC, **A**	54.9	124.1		TKW[5,6,7,15], TW[3,6], HD[16,3,10]
9	6B	552.9	622.6	WL, P, FC, **L**, A	610.7	622.7	*TaGS1b*	TKW[5,7,9,17]
10	7A	673	684.9	**FC**, HD	673.1	683.5		HD[14]
11	7B	606.3	687.9	P, **WL**, FC, L	609.8	633.6		TKW[9], HD[9,14,15]

The best QTLs are reported in bold.

[1](Maccaferri et al., 2008); [2](Faris et al., 2014); [3](Maccaferri et al., 2011); [4](Kidane et al., 2017); [5](Mangini et al., 2018); [6](Graziani et al., 2014); [7](Peleg et al., 2011); [8](Golabadi et al., 2011); [9](Roncallo et al., 2017); [10](Milner et al., 2016); [11](Soriano et al., 2017); [12](Peng et al., 2003); [13](Canè et al., 2014); [14](Maccaferri et al., 2014); [15](Giraldo et al., 2016); [16](Elouafi and Nachit, 2004); [17](Blanco et al., 2012).

Among these candidate genes, seed and spike specific chromatin remodeling factors (TRITD4Bv1G205360, TRITD5Bv1G146200, TRITD6Av1G202880, TRITD6Bv1G197750, and TRITD7Bv1 G204890), ubiquitin ligases (TRITD5Bv1G144430, TRITD 5Bv1G144440, TRITD6Av1G195410, TRITD6Av1G212580, and TRITD6Av1G212590), and cell wall modeling factors (TRITD6Av1G205500 and TRITD6Av1G205580) might play a role in controlling seed morphology.

DISCUSSION

Kernel weight and shape are important parameters determining the wheat profitability, being the main determinants of yield and its technological quality. Indeed durum wheat breeding has constantly pursued the improvement of TKW and TW. In parallel, a plethora of studies dissected the genetic bases of TKW and TW in wheat. However, while some works investigated the genetic bases of grain shape and size traits and their relationship with TKW and TW in bread wheat (Sun et al., 2008; Gegas et al., 2010; Ramya et al., 2010; Tsilo et al., 2010; Prashant et al., 2012; Maphosa et al., 2014; Rasheed et al., 2014; Williams and Sorrells, 2014; Wu et al., 2015; Kumar et al., 2016; Cheng et al., 2017; Würschum et al., 2018), very few were dedicated to durum wheat (Russo et al., 2014; Golan et al., 2015). Moreover, only a few studies based on the linkage mapping approach (Russo et al., 2014; Wu et al., 2015; Kumar et al., 2016; Cheng et al., 2017; Su et al., 2018) used a high-density genetic map to analyze kernel size related traits. Therefore, an understanding of the genetic basis of kernel size/shape traits is an important objective whose results could be deployed in future (durum) wheat breeding. Furthermore, this study was

also conceived to inspect the relevant genetic diversity present in less cultivated materials, such as landraces. Therefore, a RIL population derived from a cross among two Iranian durum wheat genotypes, a landrace and a local old cultivar (Iran_249 and Zardak, respectively), was used to investigate durum wheat kernel morphology factors and their relationships with kernel weight, and to map the corresponding QTLs. The two genotypes derive from different regions of Iran and show significant differences for morphology of kernel and spike, with Iran_249 being similar to *T. turanicum*. This wheat, currently cultivated in Iran, is a tetraploid subspecies also called Khorasan wheat, but it is genetically not dissimilar from durum landrace as shown in Maccaferri et al. (2019). For our analysis, we considered the most common parameters used to describe kernel size and shape, in the above mentioned genetic studies, and we applied high-throughput phenotyping based on digital image analysis to get accurate scoring data from a higher amount of seeds per samples from two experimental sites, thus addressing the variability present in seed sample as well as the environment effect. In addition, a high density genetic map, comprising 6,195 markers, was developed and used to perform the QTL analysis. Lastly, we anchored the mapped QTLs on the recently released *T. durum* reference genome.

The experimental field provided phenotypic data that highlight significant variability for the genotype effect for all traits considered, thus allowing to conduct QTL analysis on each single environment data. As possible for large field trial that may likely encompass non-uniform soil parameters, the replicate effect was also significant for almost all traits, however most of the variation was accounted by the genotype component.

About overall data across the three environments, all effects (G, E, and GxE interaction) were significant for all traits, with E accounting for most of the variability. The two sites used for field trail represent two durum wheat growing areas in Italy characterized by strong differences in soil fertility and climatic conditions (**Supplementary Table 1**). Consequently, for some traits known to be influenced by environment (like HD, TW and TKW), a large environmental effect, even larger than the genotype component, was observed, that is large differences among environmental means causing most of the variation in genotype performances. This confirms that the experimental sites were enough different to highlight the environment and possible GxE effect on the target traits. We can suppose that major differences in these trait phenotypes were associated to rainfall levels and temperature values. This was already reported in studies about the performances of durum genotypes conducted in the same two experimental sites (De Vita et al., 2010), and in general for the target traits across different environments (Graziani et al., 2014; Kumar et al., 2016). The observation that environment affects also kernel width, and consequently WL ratio and FC, may suggest that the environment impacts on TKW and TW through effects on width of kernel, and in a minor extent through length of kernel. More interesting is the impact of GxE interaction on total variation. We found significant GxE variability for almost all traits, but interactions contributed significantly less to the phenotypic variations, compared with the genotypic effects. Indeed, we were able to identify QTLs stable across the three environments. The only exception is TW that, with a GxE variance higher than that due to genotype, revealed its low heritability level. Accordingly, lowest level of heritability was observed for kernel width which is the morphology trait more correlated with TW. Previous studies have already reported lower level of heritability for width of kernel in comparison to length of kernel (Russo et al., 2014; Kumar et al., 2016; Su et al., 2018), thus length promises to be an effective target for breeding.

Correlations among size and shape related traits, as well as with kernel weight have been addressed in all the mentioned studies in wheat, aiming to highlight distinct genetic controls and to disentangle complex traits in their simplest but likely causative primary traits. In our case, we observed positive correlation between size related traits and grain weight, a higher correlation of width with weight of kernels as opposed to length of kernels, and a negative correlation among kernel width and kernel length. These observations, in agreement with insights from previous studies (Russo et al., 2014; Kumar et al., 2016; Cheng et al., 2017; Su et al., 2018), suggest that kernel width should be the main contributor to the increased grain weight and that kernel length and width are probably under different genetic control. Analogous results have been obtained in bread wheat, through a detailed analysis which has dissected the phenotypic and genetic structure of kernel size and shape (Gegas et al., 2010). The authors developed a phenotypic model integrating grain size and shape parameters, thus demonstrating that the kernel length and width traits are probably under the control of distinct genetic components.

The present study identified a total of 94 QTLs along all chromosomes; in detail, 43 QTLs for traits related to kernel size (L, W, P, A), 32 QTLs associated with kernel shape (C, WL, and FC), 8 QTLs for kernel weight (TKW and TW) and 11 regions associated with heading date. The phenotypic variation explained by each QTL ranged from 4.1% (QFC.gb-2B) to 40.1% (QWL.gb-7B), with an average of 15.7%. Thus, both few major and several minor QTLs for all the grain characteristics were identified, confirming the polygenic control of these traits suggested by the distribution of phenotypic data as already reported. Many of the QTLs identified were environment specific as expected according to the significant GxE effect observed for all traits. However, we were able to identify robust QTLs stable across two or three environments. Indeed, three regions associated to WL (QWL.gb-1B, QWL.gb-2B, and QWL.gb-7B), two regions identified by length data (QL.gb-2B and QL.gb-6B.1) and one region for P (QP.gb-6B.1), C (QC.gb-6B), HD (QHD.gb-2B) and TKW (QTKW.gb-5B) were effective in all evaluation trials. Other 14 QTLs detected for traits A, C, P, FC, HD, L, and WL, and spanning on different chromosomes (2A, 2B, 4A, 4B, 5B, 6B, and 7B) were expressed in two environments.

Focusing on the parent contribution, Iran_249 contributed the allele with increasing effect for most of the QTLs related to kernel length, vice versa for QTLs related to kernel width is Zardak the parent conferring the allele with positive effect. Moreover, Iran_249 conferred positive allele at 4 out of 6 loci related to kernel weight (TKW and TW), although kernel width showed consistently higher positive correlation with kernel weight than kernel length. Landraces are considered

valuable resource to enlarge the genetic diversity of modern cultivated genetic pools (Moore, 2015), however, to the better of our knowledge, the variance available for kernel length has been rarely addressed so far for the landraces, neither in durum wheat nor in bread wheat (Abdipour et al., 2016). For common wheat, a detailed analysis of the kernel size and shape trait assessed the genetic variation available among/within wheat subspecies, including primitives. In contrast to modern wheat varieties, these primitives exhibited broader variation in grain size and shape with grain width being the least variable trait, meaning that the modern breeding germplasm has lost grain morphology variation, probably due to selection for more uniform grain shape in the élite varieties (Gegas et al., 2010). In this context, our finding suggest that landraces, as exemplified by Iran_249, might have considerable potential toward enhancing the existing gene pool for grain shape and size traits and for further yield improvement in wheat, without the issue of linkage drag related to using primitive wheat's.

QTLs identified in the present study were grouped according to their genetic positions and the parent responsible of positive additive effect, thus identifying 11 cluster regions which include both loci for the primary traits L and W, as well as their corresponding derivative traits (WL, P, A and FC), and relevant agronomic traits (TKW, TW and HD). QTL clustering or coincidence is common in wheat for a number of traits and has been already reported for kernel morphology and weight (Gegas et al., 2010; Russo et al., 2014; Wu et al., 2015). It suggests that associated loci either have pleiotropic effect or are closely linked, both resulting in phenotypic correlations among corresponding traits (Kumar et al., 2016; Cheng et al., 2017). The cloning of several genes for grain shape and size genes in rice and wheat also confirmed the pleiotropic effects of those genes (Fan et al., 2006; Song et al., 2007). Following this rationale, because of the geometric and/or physiologic relationship among the traits, clusters are expected to suggest which primary kernel trait, between kernel length and width, might more strongly impact on a co-located and more complex secondary trait, like kernel shape and, more intriguingly, weight related traits. For instance, based on this assumption, we could hypothesize that for the cluster 1 (chromosome 1B) phenotypic variation for WL might depend on the co-located QTL for kernel width. Analogously, in clusters 2 (2B), 3 (4A), and 11 (7B) variation for length was likely responsible for the identification of the WL and FC loci. Notably, the comparison of the regions associated to TKW with QTLs for kernel size might identify relevant relationships between kernel size and yield related traits. This coincidence was revealed by two clusters, evidencing effect of both kernel length (cluster 2 on 2B) and width (cluster 4 on 5B) on kernel weight. Interestingly, cluster 4 contains the robust QTKW.gb-5B, repeatedly identified in three environments, QTLs for W and A, but also for HD, highlighting a possible effect of phenology on kernel weight, putatively through an impact of regulation of HD on specific kernel dimension (W). Another interesting coincidence might indicate the positive impact of kernel roundness on TW. This is the case of cluster 5 (6A) which included QTLs for TW, WL and W. Notably, Iran_249 contributes the positive allele at cluster 2 and 4, respectively through an allele with positive effect for kernel length and width, respectively. As we already pointed out, this finding suggested that increase of kernel size from the landrace might improve important agronomic traits like TKW.

The projection of the clusters on the *T. durum* reference genome sequence allowed to enlarge this approach considering QTLs for the target traits so far mapped in tetraploid wheat germplasm. Within all clusters some potentially coincidences emerged with QTL previously identified through both linkage and association mapping and recently physically mapped on the reference genome (Maccaferri et al., 2019). Interestingly, most of our clusters of QTLs for kernel morphology and size co-located with known QTLs for TKW and TW (Peng et al., 2003; Elouafi and Nachit, 2004; Maccaferri et al., 2008; Peleg et al., 2011; Canè et al., 2014; Graziani et al., 2014; Roncallo et al., 2017; Soriano et al., 2017; Mangini et al., 2018). This result further indicates that kernel size/shape genetic determinants are responsible for variability in kernel weight, suggesting that selection for these traits can indirectly improve grain weight. In other cases, coincidence was found between QTLs for the same traits, thus validating the results shown in the present study, also for those QTLs that are expressed only in one environment. For example, Faris et al. (2014) and Mangini et al. (2018) reported QTL for TKW on chromosome 2B that may correspond to QTKW.gb-2B, while QTKW.gb-5B and QTW.gb-6A could correspond to the QTLs previously reported on chromosomes 5B and 6A for the same traits (Maccaferri et al., 2011; Peleg et al., 2011; Graziani et al., 2014; Mangini et al., 2018).

The recent release of the *T. durum* reference genome (cv. Svevo) allowed us to identify durum wheat homologs to rice genes known to be involved in the regulation of kernel size and weight [as summarized by Huang et al. (2013); Li and Yang (2017)] as well as new candidate genes. The base assumption supporting this approach is that most of the gene content is conserved among the cv. Svevo and parental lines selected for this study. Consequently, the diversity observed in the current work is supposed to be mainly related to allelic variation at conserved loci, at to a lesser extent to different gene content. The large cluster 2 on chromosome 2B, including 8 QTLs detected for kernel size/shape (C, A, P, L, FC) and weight traits (TKW), encompassed several wheat genes or wheat homologs cloned for they effect on kernel size and weight (*TaSus2*, *SRS1*, *GW7*, *GLW7* and *D11*). In details, *TaSus2*, involved in the starch synthesis pathway had a direct association with grain yield in wheat representing one of the major target of indirect selection in wheat breeding for higher yield. Other cloned rice genes (*SRS1*, *GW7*, *GLW7* and *D11*) appear to be involved directly in the seed morphology, by determining the spatial control of cell division, and indirectly in the regulation of yield. However, the extension of cluster 2 impaired us to hypothesized which gene represents the best candidate gene involved in the trait determination. Interesting coincidences were also found on chromosome 4A and 6B where the *TaTGW6* and *TaGS1b* felt in cluster 3 and 9, respectively. We were also able to genetically

map the *TaGS1b* gene in the Zardak × Iran_249 population under the QTL cluster 9. Both these genes were associated to high grain weight but for the homeolog form *TaGS1a* functions for grain size and shape functions have been also hypothesized (Bernard et al., 2008; Guo et al., 2013). Therefore, the QTL present in cluster 9 as associated to kernel size could correspond to *TaGS1b*.

Besides known genes, novel candidates emerged by inspecting the gene content of the genomic regions underlined by the most consistent QTL for each cluster. Both functional annotation and expression data, as predicted based on the tissue specific RNA-seq data available for *T. aestivum*, were considered. Among the tens of genes located under the target QTLs, we were able to find a subset of genes specifically expressed in grain and spikes and having functional annotation already reported for genes related to grain size and shape in rice and wheat. However, fine mapping approaches together with detailed expression profile analysis in the parental lines are required to increase the mapping resolution and thus identify best candidate genes.

In the most recent breeding above all, the market and industry requirements for almost spherical grains led to selection for larger grains. However, yield was unaffected due to reduced kernel number (Wiersma et al., 2001), as consequence of the physiological trade-offs between individual components of yield (kernel number, kernel weight, kernel shape, etc.). These complex physiological relationships hinder improvement of grain yield when trying to manipulate single yield component using only phenotypic data. The knowledge of the genetic bases of such complex quantitative traits, together with relevant new alleles from less cultivated germplasm can contribute to model the interactions among components, to find effective combinations of traits and candidate genes, toward the improvement of wheat kernel size and yield.

AUTHOR CONTRIBUTIONS

FD and EM carried out the data analyses and wrote the manuscript. FD, EM, DG, SL, NV, and FS participated in field and post-harvest evaluations of phenotypic traits. PB and RB performed the bioinformatic analysis. LZ, KC, and EF developed RIL population. FD, MP, LC, and EM designed the study. All authors have read and approved the final manuscript.

ACKNOWLEDGMENTS

This work was supported by PON01_01145 project ISCOCEM, H2020-MSCA-RISE-2015 project ExpoSEED (grant no. 691109) and project BIOTECH-WHEADIT (Approcci di genome editing per ottimizzare la performance dei cereali tramite il controllo dei pathway ormonali).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00448/full#supplementary-material

Supplementary Figure 1 | Pearson correlations among the phenotypic traits analyzed using single environment data: L14 **(A)**, L15 **(B)**, and F15 **(C)**.

Supplementary Table 1 | Field experimental conditions and environmental parameters.

Supplementary Table 2 | Summary of phenotypic data and variation parameters for parental lines and RILs for kernel shape, size and weight related traits, and heading date. Mean data for each environment (L14, L15, and F15) has been reported.

Supplementary Table 3 | Variance components of RIL population for each trait using both mean data across the three environments and data from single environments (L14, L15, and F15).

Supplementary Table 4 | Pearson correlations and corresponding p-value, among the phenotypic traits analyzed using both overall data and data from single environments (L14, L15, and F15).

Supplementary Table 5 | Zardak × Iran_249 linkage map. Linkage groups, chromosomes, marker names and their position (cM, obtained using the Kosambi mapping function) are reported.

Supplementary Table 6 | Genes controlling kernel related phenotypes (TKW, L, W) identified in wheat and/or rice. Gene name, gene ID, wheat chromosome and the physical position on the *T. durum* reference genome are reported.

Supplementary Table 7 | Gene content under the selected QTL regions. Functional annotation and expression data are also reported. Expression data related to the ortholog bread wheat genes were downloaded from the WheatExp database (https://wheat.pw.usda.gov/WheatExp/). Tissues and stages are the following: Grains collected at the Zadoks scale 71, 75, and 85, whole endosperm collected at 10 and 20 days After Pollination (DAP), endosperm tissues as starchy endosperm, transfer cells, the aleurone, maternal tissues as the inner and outer pericarp, spikes collected at the Zadoks scale 32, 39, and 65, leaves collected at the Zadoks scale 10, 23, and 71, root collected at the Zadoks scale 10, 13, and 39, and stem collect at the Zadoks scale 30, 32, 65.

REFERENCES

Abdipour, M., Ebrahimi, M., Izadi-Darbandi, A., Mastrangelo, A. M., Najafian, G., Arshad, Y., et al. (2016). Association between grain size and shape and quality traits, and path analysis of thousand grain weight in Iranian bread wheat landraces from different geographic regions. *Not. Bot. Hort. Agrobo.* 44, 228–236. doi: 10.15835/nbha44110256

Abe, Y., Mieda, K., Ando, T., Kono, I., Yano, M., Kitano, H., et al. (2010). The SMALL AND ROUND SEED1 (SRS1/DEP2) gene is involved in the regulation of seed size in rice. *Genes Genet. Syst.* 85, 327–339. doi: 10.1266/ggs.85.327

Ashikari, M., Sakakibara, H., Lin, S., Yamamoto, T., Takashi, T., Nishimura, A., et al. (2005). Cytokinin oxidase regulates rice grain production. *Science.* 309, 741–745. doi: 10.1126/science.1113373

Ashikari, M., Wu, J., Yano, M., Sasaki, T., and Yoshimura, A. (1999). Rice gibberellin-insensitive dwarf mutant gene Dwarf 1 encodes the alpha-subunit of GTP-binding protein. *Proc. Natl. Acad. Sci. U.S.A.* 96, 10284–10289. doi: 10.1073/pnas.96.18.10284

Bernard, S. M., Molle, R. A. L. B., Dionisio, G., Kichey, T., Jahn, T. P., Dubois, F., et al. (2008). Gene expression, cellular localisation and function of glutamine synthetase isozymes in wheat (*Triticum aestivum* L.). *Plant Mol. Biol.* 67, 89–105. doi: 10.1007/s11103-008-9303-y

Blanco, A., Mangini, G., Giancaspro, A., Giove, S., Colasuonno, P., Simeone, R., et al. (2012). Relationships between grain protein content and grain yield components through quantitative trait locus analyses in a recombinant inbred line population derived from two elite durum wheat cultivars. *Mol. Breed.* 30, 79–92. doi: 10.1007/s11032-011-9600-z

Bogard, M., Jourdan, M., Allard, V., Martre, P., Perretant, M. R., Ravel, C., et al. (2011). Anthesis date mainly explained correlations between post-anthesis leaf

senescence, grain yield, and grain protein concentration in a winter wheat population segregating for flowering time QTLs. *J. Exp. Bot.* 62, 3621–3636. doi: 10.1093/jxb/err061

Broman, K. W., Wu, H., Sen, S., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics.* 19, 889–890. doi: 10.1093/bioinformatics/btg112

Canè, M. A., Maccaferri, M., Nazemi, G., Salvi, S., Francia, R., Colalongo, C., et al. (2014). Association mapping for root architectural traits in durum wheat seedlings as related to agronomic performance. *Mol. Breed.* 34, 1629–1645. doi: 10.1007/s11032-014-0177-1

Chang, J., Zhang, J., Mao, X., Li, A., Jia, J., and Jing, R. (2013). Polymorphism of TaSAP1-A1 and its association with agronomic traits in wheat. *Planta.* 237, 1495–1508. doi: 10.1007/s00425-013-1860-x

Chen, J., Gao, H., Zheng, X.-M., Jin, M., Weng, J.-F., Ma, J., et al. (2015). An evolutionarily conserved gene, FUWA, plays a role in determining panicle architecture, grain shape and grain weight in rice. *Plant J.* 83, 427–438. doi: 10.1111/tpj.12895

Cheng, R., Kong, Z., Zhang, L., Xie, Q., Jia, H., Yu, D., et al. (2017). Mapping QTLs controlling kernel dimensions in a wheat inter-varietal RIL mapping population. *Theor. Appl. Genet.* 130, 1405–1414. doi: 10.1007/s00122-017-2896-2

Churchill, G. A., and Doerge, R. W. (1994). Empirical threshold values for quantitative trait mapping. *Genetics* 138, 963–971.

Darvasi, A., and Soller, M. (1997). A simple method to calculate resolving power and confidence interval of QTL map location. *Behav. Genet.* 27, 125–132. doi: 10.1023/A:1025685324830

de Givry, S., Bouchez, M., Chabrier, P., Milan, D., and Schiex, T. (2005). CARHTA GENE: multipopulation integrated genetic and radiation hybrid mapping. *Bioinformatics.* 21, 1703–1704. doi: 10.1093/bioinformatics/bti222

De Vita, P., Mastrangelo, A. M., Matteu, L., Mazzucotelli, E., Virzi, N., Palumbo, M., et al. (2010). Genetic improvement effects of yield stability in durum wheat genotypes grown in Italy. *Field Crops Res.* 119, 68–77. doi: 10.1016/j.fcr.2010.06.016

Desiderio, F., Guerra, D., Rubiales, D., Piarulli, L., Pasquini, M., Mastrangelo, A. M., et al. (2014). Identification and mapping of quantitative trait loci for leaf rust resistance derived from a tetraploid wheat Triticum dicoccum accession. *Mol. Breed.* 34, 1659–1675. doi: 10.1007/s11032-014-0186-0

Elouafi, I., and Nachit, M. M. (2004). A genetic linkage map of the durum × *Triticum dicoccoides* backcross population based on SSRs and AFLP markers, and QTL analysis for milling traits. *Theor. Appl. Genet.* 108,401–413. doi: 10.1007/s00122-003-1440-8

Eujayl, I., Sorrells, M. E., Baum, M., Wolters, P., and Powell, W. (2002). Isolation of EST-derived microsatellite markers for genotyping the A and B genomes of wheat. *Theor. Appl. Genet.* 104, 399–407. doi: 10.1007/s001220100738

Fan, C., Xing, Y., Mao, H., Lu, T., Han, B., Xu, C., et al. (2006). GS3, a major QTL for grain length and weight and minor QTL for grain width and thickness in rice, encodes a putative transmembrane protein. *Theor. Appl. Genet.* 112, 1164–1171. doi: 10.1007/s00122-006-0218-1

Faris, J. D., Zhang, Q., Chao, S., Zhang, Z., and Xu, S. S. (2014). Analysis of agronomic and domestication traits in a durum × cultivated emmer wheat population using a high-density single nucleotide polymorphism-based linkage map. *Theor. Appl. Genet.* 127, 2333–2348. doi: 10.1007/s00122-014-2380-1

Gegas, V. C., Nazari, A., Griffiths, S., Simmonds, J., Fish, L., Orford, S., et al. (2010). A genetic framework for grain size and shape variation in wheat. *Plant Cell.* 22, 1046–1056. doi: 10.1105/tpc.110.074153

Giraldo, P., Royo, C., González, M., Carrillo, J. M., and Ruiz, M. (2016). Genetic diversity and association mapping for agromorphological and grain quality traits of a structured collection of durum wheat landraces including subsp. durum, turgidum and diccocon. *PLoS ONE* 11:e0166577. doi: 10.1371/journal.pone.0166577

Golabadi, M., Arzani, A., Mirmohammadi Maibody, S. A. M., Tabatabaei, B. E. S., and Mohammadi, S. A. (2011). Identification of microsatellite markers linked with yield components under drought stress at terminal growth stages in durum wheat. *Euphytica* 177, 207–221. doi: 10.1007/s10681-010-0242-8

Golan, G., Oksenberg, A., and Peleg, Z. (2015). Genetic evidence for differential selection of grain and embryo weight during wheat evolution under domestication. *J. Exp. Bot.* 66, 5703–5711. doi: 10.1093/jxb/erv249

Götz, S., García-Gómez, J. M., Terol, J., Williams, T. D., Nagaraj, S. H., Nueda, M. J., et al. (2008). High-throughput functional annotation and data mining with the Blast2GO suite. *Nucleic Acids Res.* 36, 3420–3435. doi: 10.1093/nar/gkn176

Graziani, M., Maccaferri, M., Royo, C., Salvatorelli, F., and Tuberosa, R. (2014). QTL dissection of yield components and morpho-physiological traits in a durum wheat elite population tested in contrasting thermo-pluviometric conditions. *Crop Pasture Sci.* 65, 80–95. doi: 10.1071/CP13349

Guo, Y., Sun, J., Zhang, G., Wang, Y., Kong, F., Zhao, Y., et al. (2013). Haplotype, molecular marker and phenotype effects associated with mineral nutrient and grain size traits of TaGS1a in wheat. *Field Crops Res.* 154, 119–125. doi: 10.1016/j.fcr.2013.07.012

Gupta, K., Balyan, S., Edwards, J., Isaac, P., Korzun, V., Röder, M., et al. (2002). Genetic mapping of 66 new microsatellite (SSR) loci in bread wheat. *Theor. Appl. Genet.* 105, 413–422. doi: 10.1007/s00122-002-0865-9

Guyomarc'h, H., Sourdille, P., Edwards, J., and Bernard, M. (2002). Studies of the transferability of microsatellites derived from *Triticum tauschii* to hexaploid wheat and to diploid related species using amplification, hybridization and sequence comparisons. *Theor. Appl. Genet.* 105, 736–744. doi: 10.1007/s00122-002-0963-8

Hanif, M., Gao, F., Liu, J., Wen, W., Zhang, Y., Rasheed, A., et al. (2015). TaTGW6-A1, an ortholog of rice TGW6, is associated with grain weight and yield in bread wheat. *Mol Breeding* 36, 1. doi: 10.1007/s11032-015-0425-z

Hong, Y., Chen, L., Du, L., Su, Z., Wang, J., Ye, X., et al. (2014). Transcript suppression of TaGW2 increased grain width and weight in bread wheat. *Funct. Integr. Genomics.* 14, 341–349. doi: 10.1007/s10142-014-0380-5

Hong, Z., Ueguchi-Tanaka, M., Umemura, K., Uozu, S., Fujioka, S., Takatsuto, S., et al. (2003). A rice brassinosteroid-deficient mutant, ebisu dwarf (d2), is caused by a loss of function of a new member of cytochrome P450. *Plant Cell.* 15, 2900–2910. doi: 10.1105/tpc.014712

Hou, J., Jiang, Q., Hao, C., Wang, Y., Zhang, H., and Zhang, X. (2014). Global selection on sucrose synthase haplotypes during a century of wheat breeding. *Plant Physiol.* 164, 1918–1929. doi: 10.1104/pp.113.232454

Hu, M.-J., Zhang, H.-P., Cao, J.-J., Zhu, X.-F., Wang, S.-X., Jiang, H., et al. (2016). Characterization of an IAA-glucose hydrolase gene TaTGW6 associated with grain weight in common wheat (*Triticum aestivum* L.). *Mol Breeding* 36, 25. doi: 10.1007/s11032-016-0449-z

Huang, R., Jiang, L., Zheng, J., Wang, T., Wang, H., Huang, Y., et al. (2013). Genetic bases of rice grain shape: so many genes, so little known. *Trends Plant Sci.* 18, 218–226. doi: 10.1016/j.tplants.2012.11.001

Ishimaru, K., Hirotsu, N., Madoka, Y., Murakami, N., Hara, N., Onodera, H., et al. (2013). Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. *Nat. Genet.* 45, 707–711. doi: 10.1038/ng.2612

Jaiswal, V., Gahlaut, V., Mathur, S., Agarwal, P., Khandelwal, M. K., Khurana, J. P., et al. (2015). Identification of novel SNP in promoter sequence of TaGW2-6A associated with grain weight and other agronomic traits in wheat (*Triticum aestivum* L.). *PLoS ONE.* 10:e0129400. doi: 10.1371/journal.pone.0129400

Jiang, Q., Hou, J., Hao, C., Wang, L., Ge, H., Dong, Y., et al. (2011). The wheat (*T. aestivum*) sucrose synthase 2 gene (TaSus2) active in endosperm development is associated with yield traits. *Funct. Integr. Genomics.* 11, 49–61. doi: 10.1007/s10142-010-0188-x

Jiang, Y., Jiang, Q., Hao, C., Hou, J., Wang, L., Zhang, H., et al. (2015). A yield-associated gene TaCWI, in wheat: its function, selection and evolution in global breeding revealed by haplotype analysis. *Theor. Appl. Genet.* 128, 131–143. doi: 10.1007/s00122-014-2417-5

Kidane, Y. G., Hailemariam, B. N., Mengistu, D. K., Fadda, C., Pè, M. E., and Dell'Acqua, M. (2017). Genome-wide association study of Septoria tritici blotch resistance in Ethiopian durum wheat landraces. *Front. Plant Sci.* 8:1586. doi: 10.3389/fpls.2017.01586

Kitagawa, K., Kurinami, S., Oki, K., Abe, Y., Ando, T., Kono, I., et al. (2010). A novel kinesin 13 protein regulating rice seed length. *Plant Cell Physiol.* 51, 1315–1329. doi: 10.1093/pcp/pcq092

Kosambi, D. D. (1944). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Kumar, A., Mantovani, E. E., Seetan, R., Soltani, A., Echeverry-Solarte, M., Jain, S., et al. (2016). Dissection of genetic factors underlying wheat kernel shape

and size in an elite × nonadapted cross using a high density SNP linkage map. *Plant Genome*. 9:81. doi: 10.3835/plantgenome2015.09.0081

Lander, E. S., and Botstein, D. (1989). Mapping mendelian factors underlying quantitative traits using RFLP linkage maps. *Genetics*. 121, 185–199.

Li, W., and Yang, B. (2017). Translational genomics of grain size regulation in wheat. *Theor. Appl. Genet.* 130, 1765–1771. doi: 10.1007/s00122-017-2953-x

Li, Y., Fan, C., Xing, Y., Jiang, Y., Luo, L., Sun, L., et al. (2011). Natural variation in GS5 plays an important role in regulating grain size and yield in rice. *Nat. Genet.* 43, 1266–1269. doi: 10.1038/ng.977

Liu, Y., Lin, Y., Gao, S., Li, Z., Ma, J., Deng, M., et al. (2017). A genome-wide association study of 23 agronomic traits in Chinese wheat landraces. *Plant J.* 91, 861–873. doi: 10.1111/tpj.13614

Lopes, S. M., El-Basyoni, I., Baenziger, P. S., Singh, S., Royo, C., Ozbek, K., et al. (2015). Exploiting genetic diversity from landraces in wheat breeding for adaptation to climate change. *J. Exp. Bot*. 66, 3625–3638. doi: 10.1093/jxb/erv122

Ma, D., Yan, J., He, Z., Wu, L., and Xia, X. (2012). Characterization of a cell wall invertase gene TaCwi-A1 on common wheat chromosome 2A and development of functional markers. *Mol. Breed.* 29, 43–52. doi: 10.1007/s11032-010-9524-z

Ma, L., Li, T., Hao, C., Wang, Y., Chen, X., and Zhang, X. (2016). TaGS5-3A, a grain size gene selected during wheat improvement for larger kernel and yield. *Plant Biotechnol. J.* 14, 1269–1280. doi: 10.1111/pbi.12492

Maccaferri, M., Cane, M. A., Sanguineti, M. C., Salvi, S., Colalongo, M. C., Massi, A., et al. (2014). A consensus framework map of durum wheat (Triticum durum Desf.) suitable for linkage disequilibrium analysis and genome-wide association mapping. *BMC Genom.* 15:873. doi: 10.1186/1471-2164-15-873

Maccaferri, M., El-Feki, W., Nazemi, G., Salvi, S., Canè, M. A., Colalongo, M. C., et al. (2016). Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. *J. Exp. Bot.* 67, 1161–1178. doi: 10.1093/jxb/erw039

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome reveals past domestication signatures and future improvement targets. *Nature Genetics*. doi: 10.1038/s41588-019-0381-3

Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., et al. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 13, 648–663. doi: 10.1111/pbi.12288

Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L. A., Salem, M. B., Bort, J., et al. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics*. 178, 489–511. doi: 10.1534/genetics.107.077297

Maccaferri, M., Sanguineti, M. C., Demontis, A., El-Ahmed, A., Garcia del Moral, L., Maalouf, F., et al. (2011). Association mapping in durum wheat grown across a broad range of water regimes. *J. Exp. Bot.* 62, 409–438. doi: 10.1093/jxb/erq287

Mangini, G., Gadaleta, A., Colasuonno, P., Marcotuli, I., Signorile, A. M., Simeone, R., et al. (2018). Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats. *PLoS ONE.* 13:e0190162. doi: 10.1371/journal.pone.0190162

Mao, H., Sun, S., Yao, J., Wang, C., Yu, S., Xu, C., et al. (2010). Linking differential domain functions of the GS3 protein to natural variation of grain size in rice. *Proc. Natl. Acad. Sci. U.S.A.* 107, 19579–19584. doi: 10.1073/pnas.1014419107

Maphosa, L., Langridge, P., Taylor, H., Parent, B., Emebiri, L. C., Kuchel, H., et al. (2014). Genetic control of grain yield and grain physical characteristics in a bread wheat population grown under a range of environmental conditions. *Theor. Appl. Genet.* 127, 1607–1624. doi: 10.1007/s00122-014-2322-y

Milner, S. G., Maccaferri, M., Huang, B. E., Mantovani, P., Massi, A., Frascaroli, E., et al. (2016). A multiparental cross population for mapping QTL for agronomic traits in durum wheat (*Triticum turgidum* ssp. durum). *Plant Biotechnol. J.* 14, 735–748. doi: 10.1111/pbi.12424

Mohler, V., Albrecht, T., Castell, A., Diethelm, M., Schweizer, G., and Hartl, L. (2016). Considering causal genes in the genetic dissection of kernel traits in common wheat. *J. Appl. Genet.* 57, 467–476. doi: 10.1007/s13353-016-0349-2

Moore, G. (2015). Strategic pre-breeding for wheat improvement. *Nat. Plants.* 1:15018.doi: 10.1038/nplants.2015.18

Nyquist, W. E. (1991). Estimation of heritability and prediction of selection response in plant populations. *Crit. Rev. Plant Sci.* 10, 235–322. doi: 10.1080/07352689109382313

Peleg, Z., Fahima, T., Korol, A. B., Abbo, S., and Saranga, Y. (2011). Genetic analysis of wheat domestication and evolution under domestication. *J. Exp. Bot.* 62, 5051–5061. doi: 10.1093/jxb/err206

Peng, J., Ronin, Y., Fahima, T., Röder, M. S., Li, Y., Nevo, E., et al. (2003). Domestication quantitative trait loci in *Triticum dicoccoides*, the progenitor of wheat. *Proc. Natl. Acad. Sci. U.S.A.* 100, 2489–2494. doi: 10.1073/pnas.252763199

Peng, J. H., and Lapitan, N. L. V. (2005). Characterization of EST-derived microsatellites in the wheat genome and development of eSSR markers. *Funct. Integr. Genomics.* 5, 80–96. doi: 10.1007/s10142-004-0128-8

Prashant, R., Kadoo, N., Desale, C., Kore, P., Dhaliwal, H. S., Chhuneja, P., et al. (2012). Kernel morphometric traits in hexaploid wheat (*Triticum aestivum* L.) are modulated by intricate QTL × QTL and genotype × environment interactions. *J. Cereal Sci.* 56, 432–439. doi: 10.1016/j.jcs.2012.05.010

Qin, L., Hao, C., Hou, J., Wang, Y., Li, T., Wang, L., et al. (2014). Homologous haplotypes, expression, genetic effects and geographic distribution of the wheat yield gene TaGW2. *BMC Plant Biol.* 14:107. doi: 10.1186/1471-2229-14-107

Ramya, P., Chaubal, A., Kulkarni, K., Gupta, L., Kadoo, N., Dhaliwal, H. S., et al. (2010). QTL mapping of 1000-kernel weight, kernel length, and kernel width in bread wheat (*Triticum aestivum* L.). *J. Appl. Genet.* 51, 421–429. doi: 10.1007/BF03208872

Rasheed, A., Xia, X., Ogbonnaya, F., Mahmood, T., Zhang, Z., Mujeeb-Kazi, A., et al. (2014). Genome-wide association for grain morphology in synthetic hexaploid wheats using digital imaging analysis. *BMC Plant Biol.* 14:128. doi: 10.1186/1471-2229-14-128

Röder, M. S., Korzun, V., Wendehake, K., Plaschke, J., Tixier, M. H., Leroy, P., et al. (1998). A microsatellite map of wheat. *Genetics.* 149, 2007–2023.

Roncallo, P. F., Akkiraju, P. C., Cervigni, G. L., and Echenique, V. C. (2017). QTL mapping and analysis of epistatic interactions for grain yield and yield-related traits in *Triticum turgidum* L. *var. durum. Euphytica.* 213:277. doi: 10.1007/s10681-017-2058-2

Russo, M. A., Ficco, D. B. M., Laidò, G., Marone, D., Papa, R., Blanco, A., et al. (2014). A dense durum wheat × *T. dicoccum* linkage map based on SNP markers for the study of seed morphology. *Mol. Breed.* 34, 1579–1597. doi: 10.1007/s11032-014-0181-5

Sajjad, M., Ma, X., Habibullah Khan, S., Shoaib, M., Song, Y., Yang, W., et al. (2017). TaFlo2-A1, an ortholog of rice Flo2, is associated with thousand grain weight in bread wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 17:164. doi: 10.1186/s12870-017-1114-3

SAS Institute Inc. (2007). *JMP v. 7*. Cary, NC: SAS Institute Inc.

Segami, S., Kono, I., Ando, T., Yano, M., Kitano, H., Miura, K., et al. (2012). Small and round seed 5 gene encodes alpha-tubulin regulating seed cell elongation in rice. *Rice.* 5:4. doi: 10.1186/1939-8433-5-4

Sen, S., and Churchill, G. A. (2001). A statistical framework for quantitative trait mapping. *Genetics.* 159, 371–387.

She, K.-C., Kusano, H., Koizumi, K., Yamakawa, H., Hakata, M., Imamura, T., et al. (2010). A novel factor FLOURY ENDOSPERM2 is involved in regulation of rice grain size and starch quality. *Plant Cell.* 22, 3280–3294. doi: 10.1105/tpc.109.070821

Shomura, A., Izawa, T., Ebana, K., Ebitani, T., Kanegae, H., Konishi, S., et al. (2008). Deletion in a gene associated with grain size increased yields during rice domestication. *Nat. Genet.* 40, 1023–1028. doi: 10.1038/ng.169

Si, L., Chen, J., Huang, X., Gong, H., Luo, J., Hou, Q., et al. (2016). OsSPL13 controls grain size in cultivated rice. *Nat. Genet.* 48, 447–456. doi: 10.1038/ng.3518

Simmonds, J., Scott, P., Brinton, J., Mestre, T. C., Bush, M., Del Blanco, A., et al. (2016). A splice acceptor site mutation in TaGW2-A1 increases thousand grain weight in tetraploid and hexaploid wheat through wider and longer grains. *Theor. Appl. Genet.* 129, 1099–1112. doi: 10.1007/s00122-016-2686-2

Song, Q. J., Shi, J. R., Singh, S., Fickus, E. W., Costa, J. M., Lewis, J., et al. (2005). Development and mapping of microsatellite (SSR) markers in wheat. *Theor. Appl. Genet.* 110, 550–560. doi: 10.1007/s00122-004-1871-x

Song, X.-J., Huang, W., Shi, M., Zhu, M.-Z., and Lin, H.-X. (2007). A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* 39, 623–630. doi: 10.1038/ng2014

Soriano, J. M., Malosetti, M., Roselló, M., Sorrells, M. E., and Royo, C. (2017). Dissecting the old Mediterranean durum wheat genetic architecture for phenology, biomass and yield formation by association mapping and QTL meta-analysis. *PLoS ONE.* 12:e0178290. doi: 10.1371/journal.pone.0178290

Sourdille, P., Cadalen, T., Guyomarc'h, H., Snape, J. W., Perretant, M. R., Charmet, G., et al. (2003). An update of the Courtot × Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. *Theor. Appl. Genet.* 106, 530–538. doi: 10.1007/s00122-002-1044-8

Su, Q., Zhang, X., Zhang, W., Zhang, N., Song, L., Liu, L., et al. (2018). QTL detection for kernel size and weight in bread wheat (*Triticum aestivum* L.) using a high-density SNP and SSR-based linkage map. *Front. Plant Sci.* 9:1484. doi: 10.3389/fpls.2018.01484

Su, Z., Hao, C., Wang, L., Dong, Y., and Zhang, X. (2011). Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 122, 211–223. doi: 10.1007/s00122-010-1437-z

Sun, C., Zhang, F., Yan, X., Zhang, X., Dong, Z., Cui, D., et al. (2017). Genome-wide association study for 13 agronomic traits reveals distribution of superior alleles in bread wheat from the Yellow and Huai Valley of China. *Plant Biotechnol. J.* 15, 953–969. doi: 10.1111/pbi.12690

Sun, X.-Y., Wu, K., Zhao, Y., Kong, F.-M., Han, G.-Z., Jiang, H.-M., et al. (2008). QTL analysis of kernel shape and weight using recombinant inbred lines in wheat. *Euphytica* 165:615. doi: 10.1007/s10681-008-9794-2

Takano-Kai, N., Jiang, H., Kubo, T., Sweeney, M., Matsumoto, T., Kanamori, H., et al. (2009). Evolutionary history of GS3, a gene conferring grain length in rice. *Genetics.* 182, 1323–1334. doi: 10.1534/genetics.109.103002

Tanabe, S., Ashikari, M., Fujioka, S., Takatsuto, S., Yoshida, S., Yano, M., et al. (2005). A novel cytochrome P450 is implicated in brassinosteroid biosynthesis via the characterization of a rice dwarf mutant, dwarf11, with reduced seed length. *Plant Cell.* 17, 776–790. doi: 10.1105/tpc.104.024950

Tsilo, T. J., Hareland, G. A., Simsek, S., Chao, S., and Anderson, J. A. (2010). Genome mapping of kernel characteristics in hard red spring wheat breeding lines. *Theor. Appl. Genet.* 121, 717–730. doi: 10.1007/s00122-010-1343-4

Van de Wouw, M., Kik, C., Van Hintum, T., Van Treuren, R., and Visser, B. (2010). Genetic erosion in crops: Concept, research results and challenges. *Plant Genet. Resour.* 8, 1–15. doi: 10.1017/S1479262109990062

Wang, E., Wang, J., Zhu, X., Hao, W., Wang, L., Li, Q., et al. (2008). Control of rice grain-filling and yield by a gene with a potential signature of domestication. *Nat. Genet.* 40, 1370–1374. doi: 10.1038/ng.220

Wang, S., Li, S., Liu, Q., Wu, K., Zhang, J., Wang, S., et al. (2015b). The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat. Genet.* 47, 949–954. doi: 10.1038/ng.3352

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Wang, S., Wu, K., Yuan, Q., Liu, X., Liu, Z., Lin, X., et al. (2012). Control of grain size, shape and quality by OsSPL16 in rice. *Nat. Genet.* 44, 950–954. doi: 10.1038/ng.2327

Wang, S., Zhang, X., Chen, F., and Cui, D. (2015a). A single-nucleotide polymorphism of TaGS5 gene revealed its association with kernel weight in Chinese bread wheat. *Front. Plant Sci.* 6:1166.doi: 10.3389/fpls.2015.01166

Wang, Y., Xiong, G., Hu, J., Jiang, L., Yu, H., Xu, J., et al. (2015c). Copy number variation at the GL7 locus contributes to grain size diversity in rice. *Nat. Genet.* 47, 944–948. doi: 10.1038/ng.3346

Wei, X., Jiao, G., Lin, H., Sheng, Z., Shao, G., Xie, L., et al. (2017). GRAIN INCOMPLETE FILLING 2 regulates grain filling and starch synthesis during rice caryopsis development. *J. Integr. Plant Biol.* 59, 134–153. doi: 10.1111/jipb.12510

Weng, J., Gu, S., Wan, X., Gao, H., Guo, T., Su, N., et al. (2008). Isolation and initial characterization of GW5, a major QTL associated with rice grain width and weight. *Cell Res.* 18, 1199–1209. doi: 10.1038/cr.2008.307

Wiersma, J. J., Busch, R. H., Fulcher, G. G., and Hareland, G. A. (2001). Recurrent selection for kernel weight in spring wheat. *Crop Sci.* 41, 999–1005. doi: 10.2135/cropsci2001.414999x

Williams, K., and Sorrells, M. E. (2014). Three-dimensional seed size and shape QTL in hexaploid wheat (*Triticum aestivum* L.) populations. *Crop Sci.* 54, 98–110. doi: 10.2135/cropsci2012.10.0609

Wu, Q.-H., Chen, Y.-X., Zhou, S.-H., Fu, L., Chen, J.-J., Xiao, Y., et al. (2015). High-density genetic linkage map construction and QTL mapping of grain shape and size in the wheat population Yanda1817 × Beinong6. *PLoS ONE.* 10:e0118144. doi: 10.1371/journal.pone.0118144

Würschum, T., Leiser, W. L., Langer, S. M., Tucker, M. R., and Longin, C. F. H. (2018). Phenotypic and genetic analysis of spike and kernel characteristics in wheat reveals long-term genetic trends of grain yield components. *Theor. Appl. Genet.* 131, 2071–2084. doi: 10.1007/s00122-018-3133-3

Xue, S., Zhang, Z., Lin, F., Kong, Z., Cao, Y., Li, C., et al. (2008). A high-density intervarietal map of the wheat genome enriched with markers derived from expressed sequence tags. *Theor. Appl. Genet.* 117, 181–189. doi: 10.1007/s00122-008-0764-9

Yamamuro, C., Ihara, Y., Wu, X., Noguchi, T., Fujioka, S., Takatsuto, S., et al. (2000). Loss of function of a rice brassinosteroid insensitive1 homolog prevents internode elongation and bending of the lamina joint. *Plant Cell.* 12, 1591–1606. doi: 10.1105/tpc.12.9.1591

Yu, J., Xiong, H., Zhu, X., Zhang, H., Li, H., Miao, J., et al. (2017). OsLG3 contributing to rice grain length and yield was mined by Ho-LAMap. *BMC Biol.* 15:28. doi: 10.1186/s12915-017-0365-7

Yue, A., Li, A., Mao, X., Chang, X., Li, R., and Jing, R. (2015). Identification and development of a functional marker from 6-SFT-A2 associated with grain weight in wheat. *Mol. Breed.* 35:63. doi: 10.1007/s11032-015-0266-9

Zhai, H., Feng, Z., Du, X., Song, Y., Liu, X., Qi, Z., et al. (2018). A novel allele of TaGW2-A1 is located in a finely mapped QTL that increases grain weight but decreases grain number in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 131, 539–553. doi: 10.1007/s00122-017-3017-y

Zhang, L., Zhao, Y.-L., Gao, L.-F., Zhao, G.-Y., Zhou, R.-H., Zhang, B.-S., et al. (2012a). TaCKX6-D1, the ortholog of rice OsCKX2, is associated with grain weight in hexaploid wheat. *New Phytol.* 195, 574–584. doi: 10.1111/j.1469-8137.2012.04194.x

Zhang, X., Wang, J., Huang, J., Lan, H., Wang, C., Yin, C., et al. (2012b). Rare allele of OsPPKL1 associated with grain length causes extra-large grain and a significant yield increase in rice. *Proc. Natl. Acad. Sci. U.S.A.* 109, 21534–21539. doi: 10.1073/pnas.1219776110

Zhang, Y., Liu, J., Xia, X., and He, Z. (2014). TaGS-D1, an ortholog of rice OsGS3, is associated with grain weight and grain length in common wheat. *Mol. Breed.* 34, 1097–1107. doi: 10.1007/s11032-014-0102-7

Zuo, J., and Li, J. (2014). Molecular genetic dissection of quantitative trait loci regulating rice grain size. *Annu. Rev. Genet.* 48, 99–118. doi: 10.1146/annurev-genet-120213-092138

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 24 July 2019
doi: 10.3389/fpls.2019.00963

Identification of a Dominant Chlorosis Phenotype Through a Forward Screen of the *Triticum turgidum* cv. Kronos TILLING Population

*Sophie A. Harrington[1], Nicolas Cobo[2], Miroslava Karafiátová[3], Jaroslav Doležel[3], Philippa Borrill[1,4] and Cristobal Uauy[1]**

[1] John Innes Centre, Norwich, United Kingdom, [2] Department of Plant Sciences, University of California, Davis, Davis, CA, United States, [3] Institute of Experimental Botany, Centre of the Region Haná for Biotechnological and Agricultural Research, Olomouc, Czechia, [4] School of Biosciences, University of Birmingham, Birmingham, United Kingdom

OPEN ACCESS

Edited by:
Agata Gadaleta, University of Bari Aldo Moro, Italy

Reviewed by:
Filippo Maria Bassi, International Center for Agricultural Research in the Dry Areas (ICARDA), Morocco
Shichen Wang, Texas A&M University, United States

***Correspondence:**
Cristobal Uauy
cristobal.uauy@jic.ac.uk

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *09 May 2019*
Accepted: *10 July 2019*
Published: *24 July 2019*

Citation:
Harrington SA, Cobo N, Karafiátová M, Doležel J, Borrill P and Uauy C (2019) Identification of a Dominant Chlorosis Phenotype Through a Forward Screen of the Triticum turgidum cv. Kronos TILLING Population. Front. Plant Sci. 10:963. doi: 10.3389/fpls.2019.00963

Durum wheat (*Triticum turgidum*) derives from a hybridization event approximately 400,000 years ago which led to the creation of an allotetraploid genome. The evolutionary recent origin of durum wheat means that its genome has not yet been fully diploidised. As a result, many of the genes present in the durum genome act in a redundant fashion, where loss-of-function mutations must be present in both gene copies to observe a phenotypic effect. Here, we use a novel set of induced variation within the cv. Kronos TILLING population to identify a locus controlling a dominant, environmentally dependent chlorosis phenotype. We carried out a forward screen of the sequenced cv. Kronos TILLING lines for senescence phenotypes and identified a line with a dominant early senescence and chlorosis phenotype. Mutant plants contained less chlorophyll throughout their development and displayed premature flag leaf senescence. A segregating population was classified into discrete phenotypic groups and subjected to bulked-segregant analysis using exome capture followed by next-generation sequencing. This allowed the identification of a single region on chromosome 3A, *Yellow Early Senescence 1* (*YES-1*), which was associated with the mutant phenotype. While this phenotype was consistent across 4 years of field trials in the United Kingdom, the mutant phenotype was not observed when grown in Davis, CA (United States). To obtain further SNPs for fine-mapping, we isolated chromosome 3A using flow sorting and sequenced the entire chromosome. By mapping these reads against both the cv. Chinese Spring reference sequence and the cv. Kronos assembly, we could identify high-quality, novel EMS-induced SNPs in non-coding regions within *YES-1* that were previously missed in the exome capture data. This allowed us to fine-map *YES-1* to 4.3 Mb, containing 59 genes. Our study shows that populations containing induced variation can be sources of novel dominant variation in polyploid crop species, highlighting their importance in future genetic screens. We also demonstrate

the value of using cultivar-specific genome assemblies alongside the gold-standard reference genomes particularly when working with non-coding regions of the genome. Further fine-mapping of the *YES-1* locus will be pursued to identify the causal SNP underpinning this dominant, environmentally dependent phenotype.

Keywords: durum wheat, genomics, senescence, chlorosis, bulked-segregant analysis, TILLING, mapping-by-sequencing

INTRODUCTION

Polyploidisation events underpin plant evolution and have been suggested to be key drivers of innovation, particularly within the angiosperms (Soltis and Soltis, 2016). All angiosperm species, including important crops such as wheat, rice, and maize, carry signatures within their genomes of ancient whole genome duplication (WGD) events that occurred within their lineage, such as the monocot-specific duplication, τ (Paterson et al., 2012). These polyploidisation events lead to the presence of multiple copies of genes which previously carried out the same function. It has been proposed that, following WGD, the resulting diploidisation of the genome leads to neo-functionalization or sub-functionalisation of gene copies derived from the original WGD (Dodsworth et al., 2016; Clark and Donoghue, 2018). The diploidisation process reduces the redundancy present within the genome by minimizing the number of genes with duplicate functions.

However, unlike rice and maize, wheat has also undergone two more recent allopolyploidisation events, where inter-species hybridizations bring together the chromosomes of each parent, creating a hybrid species with higher ploidy. The first event, approximately 400,000 years ago, occurred when two wild grasses hybridized to produce a tetraploid grass (wild emmer) which would go on to be domesticated as durum wheat (*Triticum turgidum*) (Dubcovsky and Dvorak, 2007; Borrill et al., 2019). The second polyploidisation event occurred more recently, only 10,000 years ago, when the tetraploid emmer hybridized with another diploid wild grass, leading to a hexaploid species which was then domesticated as bread wheat (*Triticum aestivum*). Unlike in ancient WGDs, these polyploidisation events have occurred relatively recently, such that most wheat genes are present as homoeologous duos or triads in durum and bread wheat, respectively, and may often have redundant functions (Ramírez-González et al., 2018).

A direct result of this homoeolog redundancy is that the inheritance of many traits in polyploid wheat tend to be quantitative, with multiple homoeologous loci contributing partly to the phenotype (Borrill et al., 2019; Brinton and Uauy, 2019). The phenotypic consequences of mutations in single homoeologs in wheat can be broadly classified into three categories — dominant (e.g., *VRN1*), whereby the mutant allele leads to a complete change in phenotype akin to mutations in diploids (Yan et al., 2003); additive (e.g., *NAM*, *GW2*), whereby mutants in each homoeolog lead to a partial change in phenotype which becomes additive as mutations are combined (Avni et al., 2014; Pearce et al., 2014; Borrill et al., 2018; Wang et al., 2018); and full redundancy (e.g., *MLO*), whereby the single and double mutants are similar to wildtype individuals, and only the full triple mutant (in hexaploid wheat) leads to significant phenotypic variation (Acevedo-Garcia et al., 2017). The presence of homoeolog redundancy, therefore, can hinder the use of forward genetic screens in polyploid wheat.

Therefore, beyond its status as an important crop, tetraploid durum wheat can provide a useful system to reduce the redundancy inherent in polyploid wheat. New advances in wheat genomics resources are increasing the speed and resolution with which we can now map loci corresponding to quantitative traits (Uauy, 2017). Recently gold-standard reference genomes for wheat were released, based on the hexaploid landrace Chinese Spring (International Wheat Genome Sequencing Consortium, 2018) and the tetraploid cultivar Svevo (Maccaferri et al., 2019). Additional wheat cultivars from across the globe are being sequenced as part of the wheat 10+ pan-genome project (10+Wheat Genomes Project, 2016). Crucially, this also includes durum wheat cultivar Kronos, which was used in the development of an *in silico* TILLING population (Krasileva et al., 2017). This mutant resource contains over 4 million chemically induced point mutation variation that can be rapidly accessed for a gene of interest through Ensembl Plants (Vullo et al., 2017).

An additional challenge when working in wheat is the sheer size of the genome, approximately 16 Gb in hexaploid and 11 Gb in tetraploid wheat. This is particularly important when designing sequencing strategies of mutant populations or individuals for mapping-by-sequencing. Various reduced representation methods exist for subsampling the wheat genome. These include gene-based methods through exome capture (Mamanova et al., 2010; Krasileva et al., 2017) or sequencing a specific gene family, as in R-gene enrichment sequencing (RenSeq) (Jupe et al., 2013; Steuernagel et al., 2016). However, these methods are less successful in obtaining variant information from non-coding regions due to their focus on genic regions. This is particularly important in the case of dominant phenotypes, which are often due to variations in regulatory regions that are not within the gene body (Yan et al., 2004; Fu et al., 2005; Borrill et al., 2015), although not exclusively (Simons et al., 2006; Greenwood et al., 2017). Methods do exist, however, to facilitate subsampling of the wheat genome while still retaining information from non-coding regions. In particular, chromosome flow sorting reduces the size of the genome by isolating an entire chromosome which can then be sequenced (Doležel et al., 2012). Other techniques (implemented in rice) include skim sequencing, which uses low coverage to obtain information about deletions or duplications, as well as SNPs, across the genome (Huang et al., 2009).

The new genomic resources and techniques for use in wheat now allows the more in-depth study of traits at the genetic level. One of these traits is chlorosis, characterized by the degradation of chlorophyll pigments in the leaves of the wheat plant. Chlorosis can be a symptom of disease, such as for yellow (stripe) rust and barley yellow dwarf virus (BYDV) (Kimura et al., 2016), as well as of many different nutrient deficiencies, such as nitrogen, magnesium, and potassium (Snowball and Robson, 1991). The presence of chlorosis, therefore, is often dependent on environmental conditions, such as the level of nutrients present in the soil. As a result, while chlorosis is not a desirable phenotype for breeding purposes, understanding the genetics underpinning chlorosis in different circumstances can increase our understanding of complex processes involved in nutrient homoeostasis and disease resistance.

Here, we used the Kronos TILLING population as a case study to identify and fine-map a novel locus in a tetraploid background (Krasileva et al., 2017). We performed a forward screen of the Kronos TILLING population for lines that exhibited late or early senescence phenotypes. From this set, we identified a line that segregated for a dominant chlorosis phenotype and was consistent across multiple years of field trials in the United Kingdom. We used mapping-by-sequencing to define the dominant phenotype as a single Mendelian locus on chromosome 3A, which we called *Yellow Early Senescence-1*. Using exome capture and chromosome flow-sorting to subsample the large wheat genome, we utilized the new RefSeqv1.0 hexaploid reference genome (International Wheat Genome Sequencing Consortium, 2018) alongside an assembly of the durum cultivar Kronos to identify SNPs across the region of interest. Following this, we mapped the *Yellow Early Senescence-1* locus to 4.3 Mb, containing 59 high-confidence genes.

MATERIALS AND METHODS

Field Trials

TILLING Population Screen

The initial screen of the sequenced Kronos TILLING population (N = 951 M_4 lines) was carried out on un-replicated single 1 m rows (**Supplementary Figure 1A**), sown in November 2015 at Church Farm, Bawburgh (52°38′N 1°10′E). Note that all John Innes Centre (JIC) trials were sown at Church Farm, but in different fields at the farm in each year. Lines were sown in numerical order (i.e., line Kronos0423 was followed by Kronos0427). For simplicity, TILLING lines will be referred to as KXXXX throughout the manuscript (i.e., Kronos0423 as K0423). Wild-type controls (cvs. Kronos, Paragon, and Soissons) were sown randomly throughout the population. Rows were phenotyped for senescence as detailed below. Following scoring, independent M_4 individuals from 10 mutant lines with early flag leaf and/or peduncle senescence and 11 mutant lines with late flag leaf and/or peduncle senescence were crossed in the glasshouse to wild-type Kronos (**Supplementary Table 1**). The F_1 plants were then self-pollinated to obtain F_2 seed (**Figure 1E**). For three mutant lines (K0331, K3085, and K3117) we recovered insufficient F_2 seeds and hence these populations were not pursued further. All original mutant lines described are available through the JIC Germplasm Resources Unit[1].

Recombinant Scoring

F_2 populations of the selected TILLING lines (backcrossed to cv. Kronos) were sown at Church Farm in March 2016 and grown as described previously (Harrington et al., 2019b). Briefly, individual F_2 seeds were hand-sown in 6 × 6 1 m^2 grids, leaving approximately 17 cm between each plant (**Supplementary Figure 1B**). In total, we sowed 31 F_2 populations representing 18 distinct TILLING mutant lines. For K2282, two F_2 populations were sown, K2282-A and K2282-B, and phenotyped. Seeds from both K2282 populations were taken forward for further field trials. 119 kg/ha of Nitrogen was applied to the trials during 2016.

In 2017 and 2018, the F_3 seed from the K2282 F_2 plants that were either heterozygous across the identified region on chromosome 3A or contained recombination within the mapped interval were grown. In 2017, we selected 30 lines from the K2282-A population and 8 lines from the K2282-B population. F_3 seed from these 38 lines were sown in a randomized block design, replicated between 1 to 4 times depending on seed availability. Each experimental unit consisted of a 1 m^2 plot that contained three 1 m rows of a single lines, separated from each other by ∼17 cm (**Supplementary Figure 1A**). The primary tillers of 12 individual plants from each row were tagged before heading. In 2018, 374 individual seeds derived from 16 F_2 plants completely heterozygous across the SH467/SH969 region were hand-sown into a 1 m^2 grid (**Supplementary Figure 1B**) and scored as in 2016. In each year, tissue for genotyping was sampled from the tagged plants (2017) or each individual plant (2018). Senescence phenotyping was carried out as detailed below. Precipitation data for the JIC field trials was obtained from a weather station at 52°37′ 52.29′′ N, 1°10′ 23.57′′ E. Nitrogen was applied to the field as fertilizer in both years, 147 kg/ha in 2017 and 124.5 kg/ha in 2018.

Phenotypic Characterization

Based on the 2017 genotypic information, nine individual F_3 lines genotyped as fully homozygous mutant (N = 4) or homozygous wild-type (N = 5) across the initial mapped region (from marker SH467 to SH969) were selected. Plants from these nine genotypes were sown in 2018 in 1 m^2 plots (two double 1 m rows separated by approximately 33 cm; **Supplementary Figure 1C**) and replicated 3 times in a complete randomized design. Wild-type Kronos and M_5 seed from the K2282 line were also sown as controls (N = 3). Two tillers in each row were tagged at heading and used for SPAD readings and genotyping. Senescence phenotyping was carried out as detailed below.

Putative Candidate Gene Phenotyping

In 2018, TILLING lines containing mutations in TraesCS3A02G414000 were grown in the field at JIC. Twenty-four seeds were hand sown per line in single, unreplicated 1 m rows (**Supplementary Figure 1A**). The TILLING lines were selected using the *in silico* TILLING browser implemented

[1]www.seedstor.ac.uk

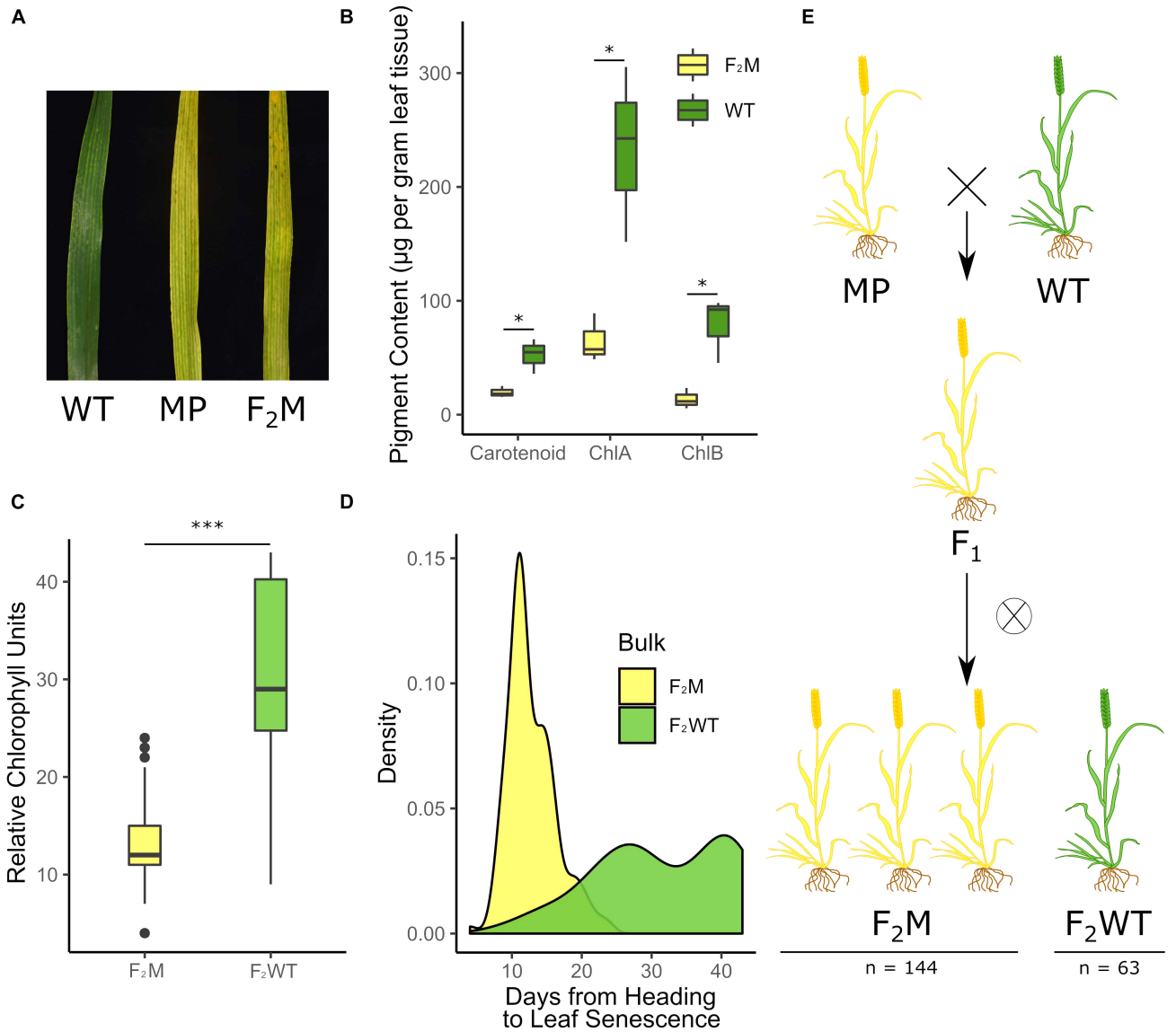


FIGURE 1 | A premature yellowing phenotype from the Kronos TILLING population segregates as a single dominant locus. F_2 populations of the K2282 Kronos TILLING line grown at the JIC in 2016 showed an early yellowing phenotype **(A)** Pigment content was measured in the yellow mutant plants (F_2M) compared to the wild-type plants (F_2WT) (**B**; n = 3 per genotype) and was also quantified using SPAD (**C**; n = 153 F_2M, n = 61 F_2WT). The yellow group (F_2M) senesced significantly earlier than the late bulk (F_2WT) (**D**; n = 148 F_2M, n = 56 F_2WT). Scoring of the plants demonstrated that the F_2 population was segregating 3:1 for the yellow trait, indicative of a dominant single locus where heterozygous plants, such as the F_1 generation, also present the mutant phenotype (**E**; numbers are combined for both populations). F_2M and F_2WT refer to plants which are yellow and green, respectively, and which derive from the F_2 population (see bottom of **E**), while WT and MP refer to Kronos WT plants or M_4 K2282 plants, respectively (see top of **E**).

through Ensembl Plants (based on reanalysis of data from Krasileva et al., 2017). Lines containing frameshift and deleterious missense mutations (SIFT = 0) were identified, and three were selected for testing in the field based on mutation location and germination success of the TILLING line.

Davis, California Trial

91 F_3 lines from population K2282-A and 55 F_3 lines from population K2282-B were sown at the University of California Field Station near Davis, California (38° 31′ N, 121° 46′ W) in November 2016. Lines were selected if the F_2 parent contained recombination within the SH467/SH969 region or was fully heterozygous across the region. In addition, seed from F_2 parents completely mutant or wild-type across the region (N = 12 each) were also selected. Lines were sown in a complete randomized design, as double 1 m rows each separated by an empty row (as in the JIC 2018 trial; **Supplementary Figure 1C**). Eight individual plants were tagged per row at heading for plants derived from heterozygous parents, to allow genotyping and scoring of individual plants. At least two plants per double

row were also tagged and sampled to verify the genotype of the completely wild-type or mutant lines. Heading and visual senescence was scored as in 2018 at the JIC, detailed below. 225 kg N/hectare was applied to the trial (as ammonium sulfate), half before planting and half on March 31st (Z30 stage). The trial was treated with appropriate fungicides to prevent stripe rust (*Puccinia striiformis* f. sp. *tritici*). Precipitation data for the Davis trial was obtained from the Davis, California weather station (38° 32′ 07″ N, 121° 46′ 30″ W).

Glasshouse Trial

F_3 plants derived from mutant or wild-type F_2 parents, genotyped across the SH179-SH969 region, were pre-germinated on damp filter paper for 48 h at 4°C in the dark. The seedlings were sown into P96 trays with 85% fine peat and 15% horticultural grit. Plants were transplanted to 1 L pots at the 3-leaf stage. The pots contained either (a) Petersfield Cereal Mix (Petersfield, Leicester, United Kingdom), (b) Horticultural Sand (J. Arthur Bower's, Westland Horticulture), or (c) Soil taken from the Church Farm site used for JIC field trials (Bawburgh, United Kingdom). Plants sown into sand were also supplied with 100 mL of Hoagland solution every 3 days (Hoagland and Arnon, 1950). K2282 mutant and wild-type F_3 plants were also tested under low water conditions in each of the three soil conditions listed above. Under the low water conditions, the plants were watered once weekly, and additionally to maintain a soil volumetric water content of approximately 20%, as measured with the Decagon GS3 sensor (ICT International, Armidale, Australia). Three plants of each genotype were treated in each condition. Plants were visually phenotyped for chlorosis onset, determined as a visual yellowing of the main flag leaf (see **Figure 1A** for a visual example).

Plant Phenotyping

Visual Color Scoring

Plants were scored as either yellow or green based on their visual appearance in the field. Those which phenocopied the wild-type Kronos plants were scored as green, while those which phenocopied the TILLING mutant parent (MP) were scored as yellow; the two categories are illustrated in **Figure 1A**. These initial visual scores were then corroborated with SPAD and chlorophyll scoring; see methods below.

Senescence Phenotyping

Plants were scored for senescence across the different field trials as detailed previously (Harrington et al., 2019b). Briefly, when scoring individual plants, all phenotyping was carried out on the main tiller, tagged upon heading. Heading was scored at Zadoks growth stage 57, when the spike was 25% emerged (Zadoks et al., 1974). Flag leaf senescence was scored for the main tiller when 25% of the flag leaf showed visual yellowing and tissue death (necrosis) from the tip. Senescence of the main peduncle was scored when the top 3 cm were fully yellow. When scoring rows of the same genotype, all stages were scored across the entirety of the row. Rows were considered to have reached heading when 75% of the main spikes reached Zadoks growth stage 57. Leaf senescence was similarly scored when 75% of the flag leaves were yellowing and necrotic across 25% of the leaf, from the tip. Peduncle senescence was scored when the top inch of 75% of the peduncles were completely yellow.

Alongside visual scoring, we utilized the SPAD-502 meter (Konica Minolta, Osaka, Japan) to obtain non-destructive chlorophyll content readings. For measurements of individual plants (2016, 2017, 2018) eight readings were taken along the flag leaf on each side of the midrib and averaged to obtain a final reading which was considered the SPAD score for that biological replicate. For measurements of rows (2018), the two tagged tillers were both measured in the same way, and the average of their measurements was taken as the SPAD reading for that biological replicate.

Chlorophyll Quantification

Chlorophyll content was measured directly from sampled leaf tissue in 2016 and 2018 at JIC. In 2016, flag leaf tissue was sampled at heading ($N = 3$ per genotype); in 2018 flag leaf tissue was sampled at anthesis and the third leaf was sampled at the third leaf stage (Zadoks 13–14), approximately 24 days before anthesis (Mutant, $N = 8$; Wild-type, $N = 10$). In 2016, one leaf was sampled per individual plant and was treated as an independent biological replicate. Similarly, in 2018, one leaf was sampled per row, and treated as an independent biological replicate. Three 1 cm^2 discs were extracted from each leaf, one at the base of the leaf, one in the middle, and one from near the leaf tip. Chlorophyll was extracted as described previously (Wellburn, 1994); briefly, the discs of tissue were soaked in N, *N*-Dimethylformamide (analytical grade, Sigma-Aldrich, United Kingdom) for 48–64 h until all pigment was completely removed from the leaf tissue. Pigment content was then quantified as previously described (Wellburn, 1994).

Leaf Mineral Content

Mineral content was taken from leaf tissue samples (2018). Leaf samples of approximately 0.2 *g* were dried and ground to a fine powder before digestion with 2 mL nitric acid (67–69%, low-metal) and 0.5 mL hydrogen peroxide (30–32%, low-metal) for 12 h at 95°C. Samples were then diluted 1:11 in ultrapure water before analysis with ICP-OES (Vista- PRO CCD Simultaneous ICP-OES; Agilent). Calibration was carried out using standards of Zn, Fe, and Mg at 0.2, 0.4, 0.6, 0.8, and 1 mg L^{-1} and Mn and P at 1, 2, 3, 4, and 5 mg L^{-1}.

Light Microscopy

Thin sections of flag leaves were cut using a razor from mutant and wild-type plants in 2018 were imaged using a Leica MZ16 light microscope (Meyer Instruments, Houston, United States; $N = 3$ per genotype).

Bulked Segregant Analysis

Individual plants with green and yellow phenotypes from the K2282 F_2 populations sown at the JIC in 2016 were selected for bulked segregant analysis. DNA from plant tissues, sampled at seedling stage, was extracted using the QIAGEN DNeasy Plant Mini Kit. The quality and quantity of the DNA was checked using

a DeNovix DS-11 Spectrophotometer, Qubit (High Sensitivity dsDNA assay, Q32854, Thermo Fisher), and by running a sample of the DNA on an agarose gel (1%) to visualize the high molecular weight DNA. Four bulks were assembled by pooling DNA from plants which had been scored as either "yellow" or "green" (K2282-A, $N = 75$ for yellow, $N = 16$ for green; K2282-B, $N = 33$ for yellow, $N = 22$ for green). Equal quantities of DNA from the individual plants were pooled into each bulk to minimize bias.

Library preparation and sequencing was carried out at the Earlham Institute (Norwich, United Kingdom) as follows. DNA quality control was carried out using the High Sensitivity Qubit assay, before library preparation was carried out with a KAPA HTP Library Prep Kit. Size selection was carried out using Beckman Coulter XP beads, and DNA was sheared to approximately 350 bp using the Covaris S2 sonicator. Four libraries were produced, one for each bulk detailed above, which were barcoded and pooled. Five cycles of PCR were carried out on the libraries before carrying out exome capture.

Hybridization to the wheat NimbleGen target capture, previously described in Krasileva et al. (2017), was carried out using the SeqCapEZ protocol v5.0, with the following changes: 2.8 μL of Universal Blocking Oligos was used, and the Cot-1 DNA was replaced with 14 μL of Developer Reagent. Hybridisation was carried out at 47°C for 72 h in a PCR machine with a lid heated to 57°C.

The library pool was diluted to 2 nM with NaOH and 10 μL transferred into 990 μL HT1 (Illumina) to give a final concentration of 20 pM. This was diluted further to an appropriate loading concentration in a volume of 120 μL and spiked with 1% PhiX Control v3 before loading onto the Illumina cBot. The flow cell was clustered using HiSeq PE Cluster Kit v4, utilizing the Illumina PE_HiSeq_Cluster_Kit_V4_cBot_recipe_V9.0 method on the Illumina cBot. After clustering, the flow cell was loaded onto the Illumina HiSeq2500 instrument following the manufacturer's instructions. The sequencing chemistry used was HiSeq SBS Kit v4. The library pool was run on two lanes with 125 bp paired end reads. Reads in bcl format were demultiplexed using the 6 bp Illumina index by CASAVA 1.8, allowing for a one base-pair mismatch per library, and converted to FASTQ format by bcl2fastq.

Chromosome Flow-Sorting and Sequencing

Seeds from the original K2282 M_5 mutant line were used for the chromosome sorting and sequencing to ensure all parental SNPs were included. Suspensions of intact mitotic chromosomes were prepared from synchronized root tip meristems according to Vrána et al. (2000). To achieve better discrimination of individual chromosomes by flow cytometry, GAA microsatellite loci were fluorescently labeled by FISHIS (Giorgi et al., 2013) using FITC-labeled $(GAA)_7$ oligonucleotides as described (Vrána et al., 2016). Chromosomal DNA was then stained by 4′,6-diamidine-2′-phenylindole (DAPI) at final concentration 2 μg/ml and the chromosome suspensions were analyzed by FACSAria SORP II flow sorter (BD Biosciences, San Jose, United States) at rates of 1000–2000 particles/s. Bivariate flow karyotypes DAPI vs. GAA-FITC were obtained and individual populations were flow sorted to identify the population representing chromosome 3A and to estimate the extent of contamination by other chromosomes (**Supplementary Figure 2**). Briefly, 2000 chromosomes were sorted onto a microscopic slide and evaluated by fluorescence microscopy after FISH with probes for GAA microsatellite and Afa-family repeat (Kubaláková et al., 2002). Three batches of 30,000 copies of chromosome 3A corresponding to ∼50 ng of DNA each were then sorted into PCR tubes containing 40 μl sterile deionized water. Chromosomal DNA was purified and amplified by Illustra GenomiPhi V2 DNA amplification Kit (GE Healthcare, Piscataway, United States) according to Šimková et al. (2008).

Library preparation and sequencing were carried out at Novogene. DNA integrity was confirmed on 1% agarose gels. A PCR-free library preparation was carried out, using the NEBNext Ultra II DNA Library Prep Kit for Illumina, following manufacturer's instructions. Libraries were sequenced using a HiseqX platform, generating 150 bp paired end reads.

Sequencing Alignments and SNP Calling

For the bulked segregant analysis, the raw Illumina reads were aligned to the Chinese Spring reference genome, RefSeqv1.0 (International Wheat Genome Sequencing Consortium, 2018), using bwa-mem (v 0.7.5) with the default settings (-k 20, -d 100) (Li, 2013). Alignments were sorted, indexed, and PCR duplicates removed using SAMtools (v 1.3.1) (Wysoker et al., 2009), and SNPs were called using freebayes (v 1.1.0, default settings) (Garrison and Marth, 2012). Depth of coverage was calculated using the exome capture size detailed previously (Krasileva et al., 2017; **Supplementary Table 2**). Following SNP calling, we then filtered the original output to obtain only SNPs that were previously called in the K2282 parent line (Krasileva et al., 2017) using an original script available online to convert SNP coordinates to the RefSeq v1.0 genome[2]. The relative enrichment of each SNP in the yellow and green bulks was visualized across the wheat genome using the Circos package (Krzywinski et al., 2009). To do this, the AO/DP ratio was calculated for each of the green and yellow bulks, where AO is the number of reads with the mutant, or "alternate," allele at the position in question, and DP is the depth of reads at that position. A Δ value was calculated by subtracting the AO/DP ratio for the green bulk from the yellow bulk. A schematic of the pipeline is provided in **Supplementary Figure 3**.

Following flow-sorting of chromosome 3A, reads were aligned to both RefSeq v1.0 and the Kronos assembly. We obtained access to the draft Kronos assembly produced at the Earlham Institute, which was assembled using the methods previously described (Clavijo et al., 2017a,b). The Kronos assembly is available in advance of publication from Grassroots Genomics[3]. In both cases, the alignment was carried out with bwa-mem (v 0.7.5; default

[2]https://github.com/Uauy-Lab/K2282_scripts; SNPs from K2282 were obtained from www.wheat-tilling.com

[3]https://opendata.earlham.ac.uk/opendata/data/Triticum_turgidum/

settings -k 20, -d 100) (Li, 2013). Illumina reads from the wild-type Kronos assembly were aligned to RefSeq v1.0 using hisat (v 2.0.4, default settings with -p 8) (Kim et al., 2015). In all cases, files were sorted, indexed, and PCR duplicates removed with SAMtools (v 1.3.1) (Wysoker et al., 2009). For alignments to RefSeq v1.0, depth of coverage across part 2 of chromosome 3A was calculated using genomic windows of 1 Mb (**Supplementary Table 2**). Depth of coverage was not calculated for the complete Kronos alignment, as the scaffolds are not associated with a chromosome. SNPs were called on the respective alignments using freebayes (v 1.1.0) at default settings in all cases. BCFtools (Wysoker et al., 2009) was used to filter the SNPs based on quality (QUAL $\geq$ 20), depth ($DP > 10$), zygosity (only homozygous), and EMS-like status (G/A or C/T SNPs).

For the alignment against the Kronos genome, SNPs were manually filtered to remove those in regions of very high SNP density. We then identified scaffolds from the Kronos genome which fall within the *YES-1* locus in the Chinese Spring RefSeq v1.0 genome using BLAST (v2.2.30; default parameters with -max_hsps 1 and -outfmt 6) (Altschul et al., 1990) against the gene sequences annotated within that region, using the v1.1 gene annotation. All further analysis of the SNP data for mapping and marker design focused solely on the 32.9 Mb *YES-1* region. This retained 18 scaffolds which contained high quality SNPs within the *YES-1* region.

Varietal SNPs between Kronos and Chinese Spring were identified by aligning the raw wild-type Kronos reads against RefSeqv1.0 (as detailed above). These varietal SNPs were then removed from the SNPs called from the alignment of the chromosome 3A reads to RefSeqv1.0. These SNPs were also manually curated to remove any which fell in regions of unreasonably high SNP density. A schematic of this workflow is provided in **Supplementary Figure 4**.

KASP Marker Genotyping

Markers were designed for the identified SNPs predominantly using the PolyMarker pipeline (Ramirez-Gonzalez et al., 2015b). Those not successful in PolyMarker were designed manually to be homoeolog specific. Markers were run on the recombinant populations using KASP genotyping, as previously described (Ramirez-Gonzalez et al., 2015a). Markers specific to K2282 are listed in **Supplementary Table 3**. Markers used for *NAM-A1* genotyping were previously published (Harrington et al., 2019b).

Data Analysis

Appropriate statistical tests for all data analyses were carried out and are detailed explicitly in the section "Results". When needed, adjustments for false discovery rate were carried out using the Benjamini–Hochberg adjustment. This is referred to in the results as "adjusted for FDR." All statistics were carried out in R (v3.5.1) (R Core Team, 2018), and data was manipulated using packages tidyr (Wickham and Henry, 2018) and dplyr (Wickham et al., 2019). Graphs of phenotyping and expression data were produced using ggplot2 (Wickham, 2016) and gplots (Warnes et al., 2019), respectively.

RESULTS

A Forward Screen of the Kronos TILLING Population Identifies a Line Segregating for a Dominant Chlorosis Phenotype

951 M_4 lines of the Kronos TILLING population (Krasileva et al., 2017) were grown at the JIC in 2015 and scored for flag leaf and peduncle senescence timing. Ten lines showed early senescence phenotypes, while 11 showed late senescence phenotypes relative to Kronos wild-type (**Supplementary Figure 5** and **Supplementary Table 1**). We developed F_2 populations for these 21 lines crossed to wild-type Kronos. In 2016 the F_2 mapping populations for 18 of these 21 lines were grown at JIC, and again scored for the senescence. From these populations, two showed significantly delayed peduncle senescence; K1107, with delayed peduncle senescence present in two independent F_2 populations, and K2711, with delayed peduncle senescence in one of two F_2 populations (**Supplementary Figure 6**). These two lines both contained mutations in the *NAM-A1* gene, known to be a positive regulator of senescence (Uauy et al., 2006). The presence of the *NAM-A1* mutation was sufficient to account for the variation in peduncle senescence timing found within the F_2 populations for both K1107 and K2711 (Tukey's HSD, $p < 0.01$, **Supplementary Figure 7**), indicating that the *NAM-A1* SNPs were causal. The effect of the *NAM-A1* mutations was followed up separately (Harrington et al., 2019b).

Based on the data from the 2016 field trials, we identified a single line, K2282, which showed a significant deviation in the timing of flag leaf senescence onset between the F_2 population and the wild-type controls ($p < 0.001$, Kolmogorov-Smirnov test, adjusted for FDR; **Supplementary Figure 6**). Two F_2 populations derived from K2282, K2282-A, and K2282-B, both showed earlier senescence compared to the wild-type controls. This phenotype, however, did not appear to be typical of a leaf senescence mutant. Although leaf senescence (scored based on leaf-tip necrosis) was indeed earlier in the K2282 populations, by anthesis the leaf tissue of individual plants was already highly chlorotic (**Figure 1A**). Quantification of chlorophyll levels confirmed that the yellow F_2 individuals from both populations contained significantly less pigment than green F_2 individuals ($p < 0.05$, Student's *t*-test, **Figure 1B**). We also observed that the chlorosis phenotype predominated in the interveinal regions in the yellow plants, leading to a characteristic striated phenotype (**Supplementary Figure 8**).

We scored the K2282 F_2 populations for chlorosis as a binary trait; i.e., plants were scored as yellow or green [see **Figure 1A** for an image of yellow (MP/F_2M) and green (WT) flag leaves]. We confirmed that our visual scoring of the plants corresponded to the true chlorotic phenotype using non-destructive measurements of relative chlorophyll units. This identified a significant reduction in chlorophyll in the yellow (F_2M) plants compared to the green (F_2WT) plants, as expected ($p < 0.001$, Student's *t*-test; **Figure 1C**). After classifying the F_2 population into the green (F_2WT) and yellow (F_2M) groups, we found that the yellow group had significantly earlier leaf senescence (when scored to include necrotic symptoms)

than the green group ($p < 0.001$ Kolmogorov–Smirnov test, **Figure 1D**). The yellow group showed a slight but significant delay of approximately 2 days in heading ($p < 0.001$, Wilcox test; **Supplementary Figure 9**) compared to the green group. This suggests that the early senescence observed is not due to the yellow group flowering earlier than the green group. The segregation of the chlorotic phenotype within the two populations was not significantly different from a 3:1 yellow to green ratio (X^2, $p = 0.07$; **Figure 1E**), consistent with the trait being underpinned by a single dominant locus, hereafter referred to as *Yellow Early Senescence 1* (*YES-1*).

The *YES-1* Locus Maps to the Long Arm of Chromosome 3A

To map the trait, we carried out bulked segregant analysis on the two independent populations, K2282-A and K2282-B. A diagram of the analysis pipeline used is provided in **Supplementary Figure 3**. Following library preparation and exome capture, reads were aligned against the RefSeqv1.0 genome (International Wheat Genome Sequencing Consortium, 2018) and SNPs were called (**Supplementary Table 2**). To reduce the number of false SNP calls, we initially filtered the SNPs to only include those previously identified in the original M_2 TILLING line (Krasileva et al., 2017). We recovered 1,548 SNPs out of the 3,060 SNPs present in the original K2282 M_2 line which was sequenced. We expected to recover fewer SNPs than those identified in the original TILLING line as SNPs that were initially heterozygous in the M_2 generation, may have been lost in the following two generations. Similarly, ∼50% of heterozygous mutations present in the M_4 line crossed to wild-type Kronos to produce the F_2 population would also have been lost.

We initially focused our analysis on the K2282-A population and calculated the ratio of the mutant (alternate) allele over total depth of coverage (AO/DP) at each SNP location in the yellow and green bulks (**Figure 2A**, inner track). From this, we then calculated the Δ value representing the enrichment of the mutant allele in the yellow bulk compared to the green bulk (**Figure 2A**, outer track). The segregation ratio seen in the field suggested this was a dominant single locus trait. Hence, we assumed that the yellow bulk would contain individuals homozygous or heterozygous for the causal mutant allele, while the green bulk should only contain homozygous wild-type plants. As a result, the AO/DP value should approach 0 in the green bulk, and 0.66 in the mutant bulk, and thus have a Δ value of 0.66. Using a conservative limit of 0.5 for the Δ value (gray line, outer track of **Figure 2A**), we identified only one region, on chromosome 3A, that was enriched for the mutant allele (**Figure 2B**). This result was consistent with that obtained from mapping carried out on the second population, K2282-B (**Supplementary Figure 10**).

To validate this mapping, we developed KASP markers for the SNPs within and surrounding the region of interest (**Figure 2C** and **Supplementary Table 3**). Mapping of the individual F_2 plants which were used to perform the exome capture confirmed the location of the region of interest on the long arm of chromosome 3A. Using the recombination events within this region and requiring at least two independent F_2 plants to define the mapping interval, we narrowed the *YES-1* region to between markers SH179 and SH969, a region of 32.9 Mb in the RefSeq v1.0 genome containing 345 genes (RefSeq v1.1 gene annotation) (**Figure 2C**).

Leaf Chlorosis Precedes Anthesis but Is Inconsistent Across Environments

To further characterize the phenotype, individual lines which were genotyped as completely mutant or wild-type across the *YES-1* region were grown at the JIC in 2018. The mutant lines contained less chlorophyll A, B, and carotenoid pigment as early as the 3rd leaf stage (Zadoks 13–14) (Student's *t*-test, $p < 0.01$; **Figure 3A**). This difference was increased at anthesis (Student's *t*-test, $p < 0.005$), at which stage there was a larger spread in pigment content within the mutant lines than the wild-type lines. Chlorophyll content, measured with SPAD units, was also monitored across the development of the plants, from 14 days before anthesis to 39 days post-anthesis. SPAD readings were consistently lower in the mutant lines up to 24 days post-anthesis ($p < 0.01$, Pairwise Wilcoxon Rank Sum adjusted for FDR). The chlorotic phenotype remained highly visible on the leaves of the mutant plants, compared to wild-type (shown at 20 DPA, **Figure 3C**). In both wild-type and mutant lines, the level of chlorophyll in the flag leaf peaked at approximately 6 DPA (**Figure 3B**). No significant decline in SPAD units was observed in the wild-type plants until 24 DPA ($p < 0.01$, Pairwise Wilcoxon Rank Sum adjusted for FDR). In contrast, the mutant plants contained significantly less chlorophyll at 18 DPA compared to the peak at 6 DPA ($p < 0.01$, Pairwise Wilcoxon Rank Sum adjusted for FDR). Despite this earlier onset of senescence, the mutant lines continued to lose chlorophyll until the final stage of the time course (39 DPA), in line with the wild-type plants. We also found that the chlorosis phenotype is associated with significant decreases in leaf mineral content, with chlorotic leaves containing less magnesium at the 3rd leaf stage, and less of all four measured minerals at anthesis (Mg, Fe, Zn, $p < 0.05$; Mn, $p = 0.05$; **Supplementary Figure 11**).

Mutant and wild-type lines were also grown at UC Davis during the summer of 2016. Unlike in the United Kingdom, no chlorosis or senescence phenotype was observed either through visual or SPAD scoring (**Supplementary Figure 12**). This suggested that the causal locus underpinning *YES-1* was environmentally dependent. Given the similarity between the interveinal chlorosis phenotype observed in the *YES-1* mutant plants to that seen in plants with varying forms of nutrient deficiency (Snowball and Robson, 1991) and the decrease in leaf mineral content seen in the mutant plants (**Supplementary Figure 11**), we hypothesized that the environmental variation in phenotype may be due to nutrient content in the soil. To test this, F_3 plants fully mutant across the *YES-1* region were grown under glasshouse conditions in three soil types: standard glasshouse cereal mix, soil taken from the JIC field site in 2017, and horticultural sand supplemented with nutrient-replete Hoagland solution. However, none of the three conditions tested recapitulated the yellowing phenotype observed in the United Kingdom field (**Supplementary Figure 13**). This was surprising given the consistency of the phenotype at the JIC

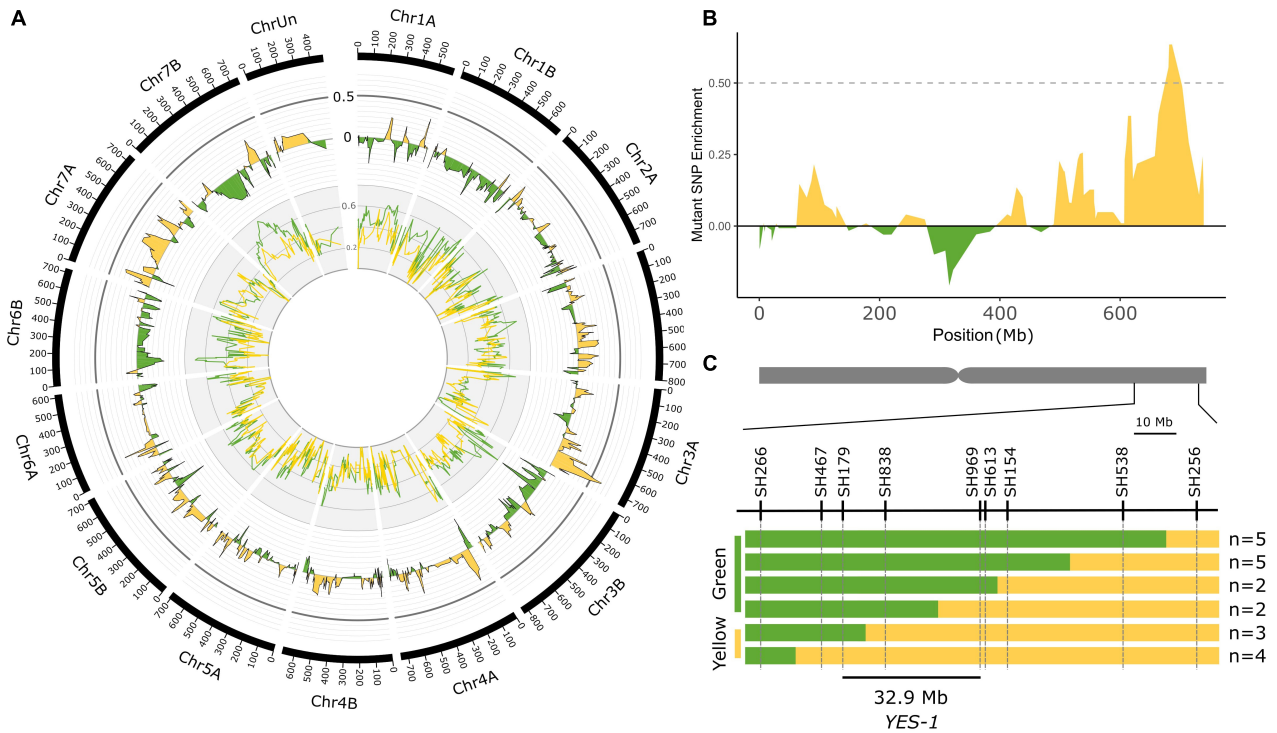


FIGURE 2 | Bulked segregant analysis identifies the *YES-1* locus on chromosome 3A. Exome capture was carried out on yellow and green bulks from K2282 × Kronos F_2 populations grown at the JIC in 2016. The K2282-A yellow bulk (yellow line, inner track; smoothed to a moving average of 4) and green bulk (green line) were scored at each SNP locus identified for enrichment of the mutant allele. The level of enrichment in the green bulk was subtracted from that of the yellow bulk to obtain the Δ value (outer track; smoothed to a moving average of 4). A high Δ value, indicative of a region enriched for mutant alleles within the yellow bulks, was identified on the long arm of chromosome 3A (**B**; smoothed to a moving average of 4). Markers designed on known TILLING SNPs within this region mapped the *YES-1* locus to a 32.9 Mb interval within the F_2 population **(C)**. Green bars indicate wild-type calls, while yellow bars indicate mutant or heterozygous calls. The relevant phenotype of each recombinant group is shown to the left of the panel. The numbers of individual plants that fell into each recombination interval are shown to the right. The chromosome scale in **(A)** is given in Mb.

field site across four different fields during four successive field seasons (2015–2018).

We then investigated weather-related environmental variation across the two field sites and across years. We obtained rainfall and temperature data from Davis, CA, for the 2016–2017 growing season, and from the JIC field site for the 2016, 2017, and 2018 growing seasons. The trials carried out in California in 2017 received substantially more rainfall between sowing and heading than in any of the JIC trials (**Supplementary Table 4** and **Supplementary Figure 14**). This suggested that perhaps reduced rain levels were correlated with the appearance of the mutant yellow phenotype. However, attempts to recapitulate the yellowing phenotype in the glasshouse through reduced watering of plants was also unsuccessful, as no early chlorosis or senescence was observed under different watering conditions (**Supplementary Figure 13**).

Fine-Mapping Reduces the *YES-1* Locus to 4 Mb on Chromosome 3A

To identify further SNPs within the *YES-1* locus, we purified chromosome 3A from the K2282 mutant by flow cytometry sorting. However, as the population of 3A chromosomes partially overlapped with the population of 7A chromosomes on a bivariate flow karyotype DAPI vs. GAA-FITC (**Supplementary Figure 2**), flow-sorted fractions comprised 80% of chromosome 3A and 20% of chromosome 7A as determined by microscopic observation. For sequencing, three batches of 30,000 chromosomes (∼50 ng) were flow-sorted and subsequent DNA amplification of three independent samples resulted in a total of 4.51 μg DNA.

Following sequencing, reads were mapped against the A and B genomes of the wheat RefSeq v1.0 genome (**Supplementary Figure 4**). 60.38% of reads aligned to chromosome 3A while 25.37% aligned to chromosome 7A, consistent with the expected contamination. The remaining reads (14.25%) mapped against the rest of the genome. We obtained on average 82X coverage across chromosome 3A, using genomic windows of 1 Mb.

In order to maximize our ability to discover novel SNPs in the *YES-1* region, we carried out a simultaneous approach to SNP discovery utilizing both the Chinese Spring reference genome as well as the draft Kronos assembly, as depicted in **Supplementary Figure 4**. In brief, paired end sequencing of the K2282 mutant chromosome 3A was used to obtain high-quality SNPs outside of the previously captured exome. We used the Kronos assembly to identify SNPs in non-coding regions that are less conserved

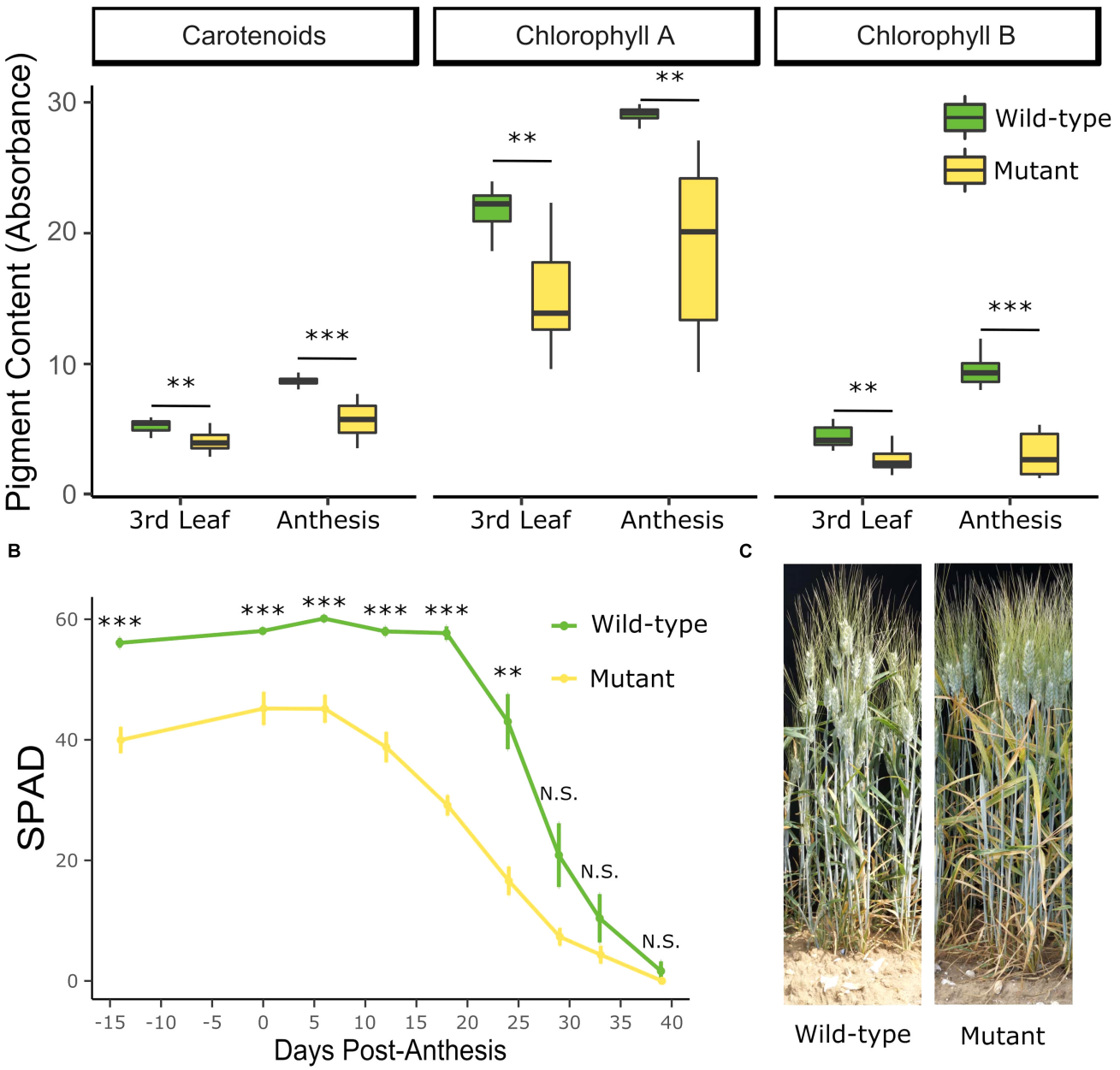


FIGURE 3 | The *YES-1* locus causes lower chlorophyll levels before anthesis and earlier onset of senescence. The early chlorosis phenotype was recapitulated in the JIC 2018 field trials. Pigment content in the mutant lines is significantly lower at the third leaf stage (Zadoks 13–14, 24 days before anthesis) and becomes more extreme by anthesis (**A**; $^{**}p < 0.01$; $^{***}p < 0.001$ Student's *t*-test). Relatively chlorophyll content, as measured with a SPAD meter, is significantly decreased in the mutant lines before anthesis, and remains significantly lower until 29 days post-anthesis (**B**; $^{**}p < 0.01$; $^{***}p < 0.001$ Pairwise Wilcoxon Rank Sum, adjusted for FDR). The yellowing phenotype in the leaves were clear in the field at 20 DPA **(C)**.

between the Kronos and Chinese Spring cultivars. In tandem, we took advantage of the contiguity of the RefSeq v1.0 genome which facilitated the identification of high-quality SNPs in and around all genes within the *YES-1* locus.

Reads from the mutant K2282 chromosome 3A were mapped against the draft Kronos assembly and were filtered for homozygous, EMS-like SNPs, passing minimum quality and depth thresholds. To obtain only SNPs that fell within the physical region encompassed by the *YES-1* locus, we carried out a BLAST between the Kronos scaffolds which contained SNPs and the Chinese Spring gene sequences within part of the *YES-1* region. Conducting a BLAST against gene sequences within the *YES-1* region, rather than the entirety of the region, reduced the number of scaffolds that mapped to the *YES-1* region due

to shared repetitive sequences rather than true synteny. Based on recombination seen in individual plants, we focussed on a region spanning 3 Mb upstream of marker SH179 and 3 Mb downstream of marker SH838, approximately 16 Mb in size. Within this region, we identified 18 unique Kronos scaffolds which both contained SNPs and at least one gene found in the RefSeqv1.0 *YES-1* physical interval (**Supplementary Table 5**). 26 of the genes within the *YES-1* region in Chinese Spring were identified (out of 345 total) within these 18 Kronos scaffolds. Genes that were not identified in the Kronos scaffolds may fall in scaffolds that contained no high-quality SNPs, may be split across multiple scaffolds, or may be absent from the Kronos genome. The SNPs within these scaffolds were manually curated, to exclude any regions that contained an unexpectedly high density of SNPs, leaving a final list of 16 scaffolds containing high-quality SNPs (**Supplementary Table 5**). These scaffolds covered an approximately 2.2 Mb of genome sequence, containing SNPs which fell mostly in non-coding regions (**Supplementary Table 5**). The SNPs underlying markers SH838 and SH179, initially identified in the exome capture data, were also recovered in the Kronos genome, validating the use of this method. KASP primers were designed for a subset of the SNPs and were used, together with the previous phenotypic data, to map *YES-1* to a 6.6 Mb region between markers SH123480 and SH59985 (**Figure 4A**).

To obtain more markers across the region, we also called SNPs against the Chinese Spring reference. To account for varietal SNPs between Kronos and Chinese Spring, we aligned raw reads from wild-type Kronos to the RefSeq v1.0 genome (**Supplementary Figure 4**). Using a subset of reads, we obtained a coverage of approximately 30X across chromosome 3A. SNP calling was then carried out against RefSeq v1.0 to obtain a list of varietal SNPs between Chinese Spring and Kronos. A total of 968,482 homozygous SNPs with quality greater than 20 and depth of coverage greater than 10 were identified across the second part of chromosome 3A, encompassing *YES-1*.

SNPs were then called between the K2282 mutant chromosome 3A reads and the RefSeq v1.0 genome. The set of SNPs was filtered for quality and depth and to exclude the varietal SNPs identified above. Following this filtering, a total of 7,153 SNPs were identified between markers SH123480 and SH969, a region of approximately 30 Mb. This was substantially more SNPs than would be expected from the known mutation density of 23 mutations/Mb for the Kronos TILLING lines (Krasileva et al., 2017). However, SNP density across the region was highly irregular which we hypothesized was due to mismapping and spurious SNP calling in repetitive regions.

To reduce the impact of repetitive regions on SNP calling, we extracted SNPs only from regions encompassing 1 Kb up and downstream of annotated genes within the *YES-1* region. Following manual curation of SNPs, we identified a set of 15 SNPs that were located near genes within the annotated region (**Supplementary Table 6**). Of these SNPs, three were located in gene bodies (including the known TILLING SNP SH838), while the remainder were intronic (5) or fell in the promoter (5) or 3′ UTR (2). Note that some SNPs are in sufficient proximity to two gene models to be counted twice. Of the mutations in the coding region, SH838 and SH858 are both missense variants with low SIFT scores (0.00 and 0.03, respectively), while SH567 is a synonymous mutation. We designed markers based on these new SNPs and based on the JIC 2017 and 2018 phenotypic data we mapped *YES-1* to a 4.3 Mb interval, between markers SH044 and SH59985 (**Figure 4A**).

Genes Within the Region

Within this region we identified 59 high-confidence genes based on the RefSeq v1.1 gene annotation (**Supplementary Table 7**). Using developmental time course data from two wheat varieties (Chinese Spring and Azhurnaya) (Borrill et al., 2016; Ramírez-González et al., 2018), we found that 25 genes within the region are expressed above 0.5 transcripts per million (TPM) in at least one stage of leaf or shoot tissue during development, consistent with our observation of a leaf-based phenotype (**Figure 4**). Of these genes, 18 were expressed above 0.5 TPM in leaf and shoot tissue during both vegetative and reproductive stages (**Supplementary Table 7**). This set of genes includes a putative magnesium transporter, TraesCS3A02G414000 (Gebert et al., 2009), which contains a missense mutation in the first exon of the gene which is predicted to be highly deleterious (SIFT = 0). This is the only gene within the 4 Mb region that contains a coding-region SNP, however, no chlorosis phenotype was observed for three additional lines with mutations in this gene (**Supplementary Table 8**). Within the 59 high-confidence genes, five genes were found to have senescence-related functions in their closest rice orthologues. A set of four tandem duplicated genes, TraesCS3A02G412900 to TraesCS3A02G413200, are orthologues to the rice gene *OsSAG12-1*, a negative regulator of senescence (Singh et al., 2013). A fifth gene, TraesCS3A02G410800, is orthologous to *Tryptophan Decarboxylase 2*, a rice gene that causes higher serotonin levels and delayed leaf senescence when over-expressed (Kang et al., 2009). All five genes with senescence-related phenotypes are lowly expressed or non-expressed across the set of tissues and developmental stages considered. However, the majority of the genes within the region are un-annotated and lack orthologous copies in either rice or Arabidopsis.

DISCUSSION

Here we have fine-mapped a region causing a dominant, environmentally dependent early-chlorosis phenotype. We have taken advantage of the recently released genetic and genomic resources for wheat to increase our ability to identify SNPs *de novo* in a Kronos TILLING mutant line. We have shown how the use of cultivar specific genome assemblies can be used to increase the ability to identify high-quality SNPs in non-genic regions which are often relatively less conserved between varieties than coding sequences.

Induced SNP Variation Can Lead to Novel Dominant Phenotypes

Many of the critical domestication alleles in polyploid wheat are derived from dominant mutations (Borrill et al., 2015;

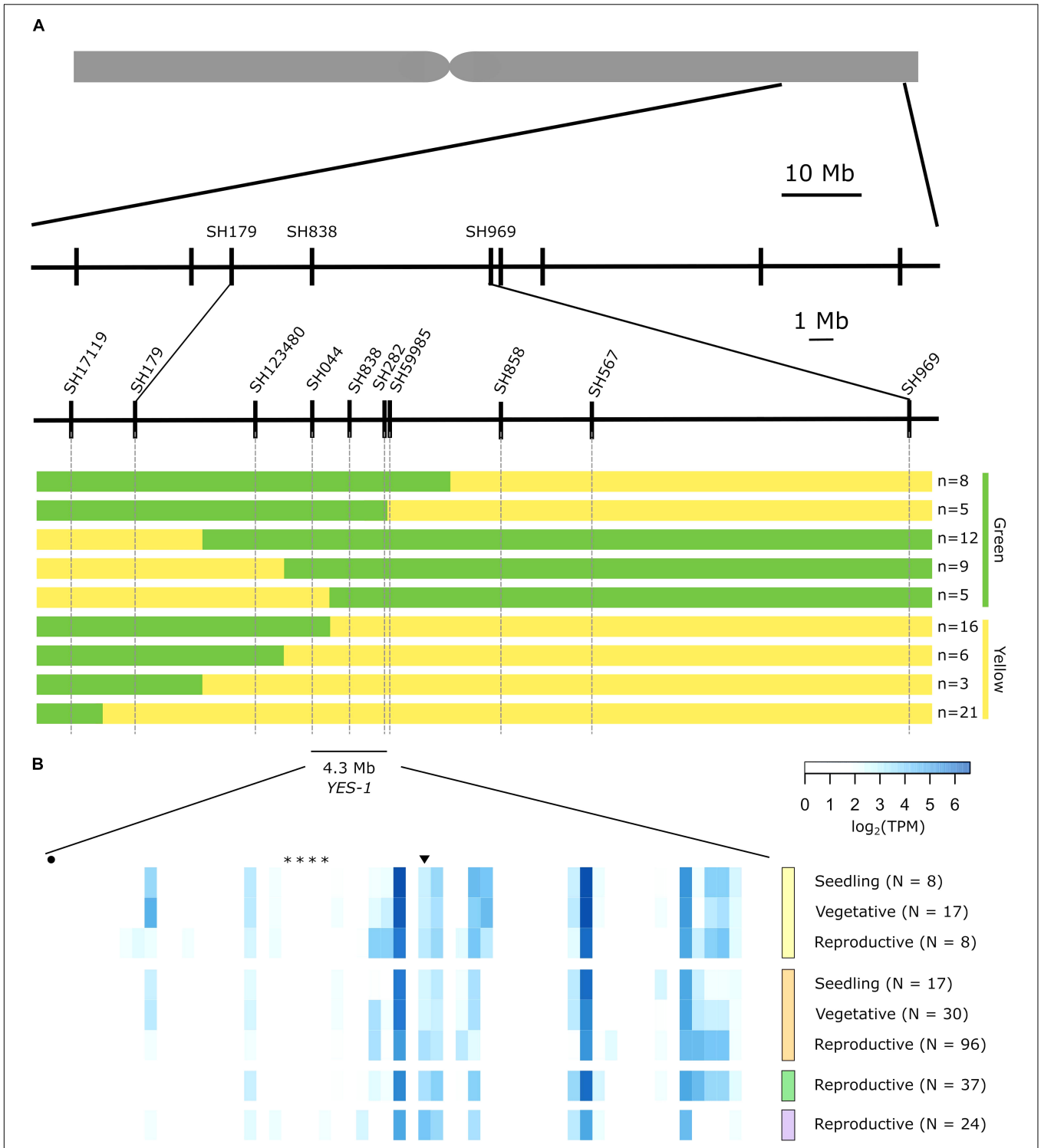


FIGURE 4 | The *YES-1* locus fine-maps to a 4.3 Mb region containing 59 genes. Markers were developed for novel SNPs identified in non-coding regions. We used phenotypic data from JIC 2017 and 2018 field trials to classify recombinant lines as green or yellow, indicated to the right of the panel. **(A)**. These markers mapped *YES-1* to a 4.3 Mb interval between markers SH044 and SH59985. Expression data for the 59 high-confidence genes in the region **(B)** from developmental time course data (Ramírez-González et al., 2018) highlights gene expression in root (yellow, top), leaf/shoot (orange, second from top), spike (green, second from bottom) and grain (purple, bottom) tissues across developmental stages. Genes mentioned in the text are highlighted by an asterisk (TraesCS3A02G412900 to TraesCS3A02G413200; *OsSAG12* orthologs), a circle (TraesCS3A02G410800; *Tryptophan Decarboxylase 2*) or an inverted triangle (TraesCS3A02G414000; putative magnesium transporter). The full list of genes is provided in physical order in **Supplementary Table 7**.

Uauy et al., 2017). This includes genes with critical variation in flowering time and free-threshing alleles resulting from dominant mutations (Yan et al., 2004; Fu et al., 2005; Simons et al., 2006; Greenwood et al., 2017). In wheat, the high level of redundancy between homoeologous genes adds to the importance of identifying dominant alleles to develop novel traits. Dominant alleles have retained their importance in modern breeding programs, underpinning the Green Revolution via the dominant dwarfing *Rht-1* allele (Peng et al., 1999; Borrill et al., 2015). Most traits selected for in modern breeding programs, however, lack standing variation of dominant alleles in both the modern breeding pool and in older wheat landraces and progenitors. Instead, forward screens for phenotypes of interest typically identify multiple quantitative trait loci (QTL) that each contribute toward a small portion of the desired phenotype. These more complex effects, often caused by loss-of-function mutations, are inherently more difficult to identify due to the need to acquire/combine mutations in both or all homoeologous copies of a gene to attain a clear phenotypic effect (Borrill et al., 2015, 2019).

Here we have demonstrated that novel dominant alleles can be identified in chemically mutagenized TILLING populations (Krasileva et al., 2017). Forward screens of the TILLING population are most likely to identify novel traits caused by dominant mutations, given the low likelihood of obtaining simultaneous mutations in multiple homoeologous copies of the same gene. Indeed, the fact that mutations in *NAM-A1* underpinned the only other senescence phenotype identified during this forward screen underscores this. The B-genome homoeolog of *NAM-A1* is non-functional in Kronos; as a result, a single mutation in the A-homoeolog equates to a complete null and was sufficient to show a strong and consistent phenotype (Pearce et al., 2014; Harrington et al., 2019b).

The dominant phenotype identified in the K2282 line was particularly clear in that individual plants could unambiguously be scored for a binary green/yellow trait. However, we suggest that the TILLING population is equally well suited for forward genetic screens to identify novel dominant alleles governing other phenotypes. Recently, the Kronos TILLING population was used to identify a line which contained a deletion of *Rht-B1*, the partially dominant dwarfing allele (Mo et al., 2018). Here we have identified a novel dominant allele with no previously characterized genes located within the candidate region. This highlights the potential for novel dominant alleles to be identified in populations with induced variation, such as the Kronos and Cadenza TILLING populations (Krasileva et al., 2017).

The Use of Cultivar-Specific Assemblies Facilitates the Identification of Non-genic SNPs

A complication of working with dominant induced variation, however, is that dominant mutations may often act through changes to regulatory elements. Variation in the promoter and intron sequence of the flowering time gene *VRN1* underpins the transition from winter to spring growth habit in wheat and barley (Yan et al., 2004; Fu et al., 2005). More recently, CRISPR editing has been used in tomato to edit the promoter region of various yield-related genes, leading to a high level of variation in trait morphology (Rodríguez-Leal et al., 2017). These results, amongst many others, highlight the importance of non-coding regions in regulating agronomically relevant traits. However, many reduced representation methods focus on enrichment of coding regions (Borrill et al., 2019). Such methods of genome complexity reduction, therefore, are less likely to contain the information needed to identify a dominant causal SNP in a regulatory region. Compounding this difficulty is the fact that non-coding regions of the genome are typically less conserved between cultivars. As a result, SNP identification against the reference variety may fail to identify critical SNPs or, conversely, identify a large number of spurious SNPs.

We have shown here that the draft Kronos assembly can instead be used, alongside non-biased methods of genome size reduction (e.g., chromosome flow sorting), to identify cultivar-specific SNPs in non-coding regions of the genome. We started by calling SNPs against scaffolds of the Kronos assembly, obtaining a large amount of SNP variation between the wild-type and mutant lines. Once we had this data, we then positioned the scaffolds which contained SNPs against the reference genome, identifying the SNPs which were located within our region of interest (here *YES-1*). This approach overcame two of the main drawbacks to using the reference genome and the Kronos assembly. On one hand, the reference genome would be expected to have different sequence content to another variety, such as Kronos, limiting its utility for SNP identification. On the other hand, unlike the gold-standard reference genome, the Kronos assembly does not have long-range assemblies needed to obtain positional information for SNPs. Long-range contiguous assemblies of additional cultivars, such as the recently published Svevo genome (Maccaferri et al., 2019), will greatly improve this current limitation. Until then, using variety-specific genomes, such as those being produced by the 10+Wheat Genomes project, alongside the highly contiguous reference genomes will facilitate the identification of non-genic SNPs.

Variability in Phenotype Points to an Environmentally Dependent Causal Locus

The early chlorosis and senescence phenotype caused by the *YES-1* locus was consistent across 4 years in field trials at the JIC. However, mutant lines showed no evidence of a chlorotic phenotype when grown in Davis, CA. Comparison of rainfall and temperature patterns between the years and locations highlighted the fact that the plants received a high level of rainfall in Davis before flowering, substantially more than that received in any of the years at the JIC (**Supplementary Table 4**). This was due to the highly unusual wet winter that occurred in California in 2016/2017, with an average rainfall of 781 mm across the state from October 2016 to March 2017 (NOAA National Centers For Environmental Information, 2017). This suggested initially that the chlorosis response may be a response to higher water stress, yet we were unable to recapitulate the phenotype when grown in the glasshouse under different watering conditions. It is also

possible that differences in temperature and photoperiod may influence the presence of the chlorosis phenotype. In particular, given the longer growing season in Davis, CA compared to at the JIC, as well as the difference in latitude, the photoperiod experienced by the Davis field trial would vary substantially from that experienced in the United Kingdom. However, as we also failed to recapitulate the yellow chlorosis phenotype in the glasshouse in the United Kingdom, under natural photoperiod in the spring and summer season of 2017, this does not seem to be the most likely source of environmental variation affecting the phenotype.

We also considered whether the phenotype was due to variation in soil nutrient content. The presence of a missense mutation within the coding region of a putative Mg^{2+} transporter (Gebert et al., 2009) highlighted this as a promising candidate gene. Similarly, the observed interveinal chlorosis phenotype (**Supplementary Figure 8**) is reminiscent of that characteristic of a magnesium deficiency (Snowball and Robson, 1991). However, we failed to recapitulate the phenotype when grown in the glasshouse using soil taken from the field at JIC, where the three trials were grown that showed a clear chlorotic phenotype (**Supplementary Figure 13**). Compounding this, we found that three Kronos TILLING lines which contained other SNPs within the transporter gene sequence did not show the same chlorotic phenotype (**Supplementary Table 8**). This included lines with both missense and premature stop codon mutations which lacked the exon containing the identified SNP in K2282. This implies that, if the magnesium transporter were the cause of the *YES-1* phenotype, the specific missense mutation present in K2282 has a unique ability to cause a dominant change in function. As the transporter is predicted to function in a pentamer (Payandeh et al., 2013), it is possible that the mutation could be sufficient to prevent the pentamer to function effectively once formed, but not sufficient to prevent the mutant monomer from being incorporated into the pentamer. In this way it may be possible that plants heterozygous for the mutation show an equally strong phenotype as homozygous mutants as incorporation of a mutant monomer disrupts completely the function of the hexameric complex. This hypothesis could be tested in the future using Cas9-driven base editing in wheat to recapitulate the exact mutation in an independent background (Zong et al., 2017).

An alternative possibility is that a separate SNP located in a regulatory region may be acting either on the identified magnesium transporter, or on a separate, currently uncharacterised gene. Few dominant chlorosis phenotypes have previously been reported in the literature. A dominant chlorosis phenotype was previously reported in *Brassica napus*, however, this phenotype disappeared after budding (equivalent to heading in cereals) unlike here, where the yellowing phenotype became increasingly strong post-heading (Wang et al., 2016). In wheat, a *Ygm* (yellow-green leaf color) mutant has been identified with a semi-dominant phenotype where the heterozygous plants are an intermediate yellow-green color between the wild-type and homozygous mutant plants (Wu et al., 2018). This phenotype is underpinned by abnormal chloroplast development and is associated with differential expression of genes involved in chlorophyll biosynthesis and carbon fixation, amongst other traits. Further work to fine-map the *YES-1* locus will hopefully shed light on the specific causal SNP underpinning the environmentally dependent chlorosis phenotype observed here, as well as on mechanisms governing dominant traits in polyploid wheat.

DATA AVAILABILITY

The raw reads from the exome capture and the flow-sorting experiments have been deposited on the SRA (PRJNA540141). The Kronos assembly is available from http://opendata.earlham.ac.uk/Triticum_turgidum/. Codes used for coordinate conversion and creation of the Circos plots (**Figure 2** and **Supplementary Figure 9**) are stored on Github (https://github.com/Uauy-Lab/K2282_scripts). The original phenotypic data from the screen of all TILLING lines has been deposited on Dryad (doi: 10.5061/dryad.g3r3hp7).

AUTHOR CONTRIBUTIONS

CU and PB conceived the study. SH, PB, and NC carried out the field trials and phenotyping. SH carried out the mapping. MK and JD flow-sorted chromosome 3A, determined the purity in flow-sorted fractions and amplified chromosomal DNA for sequencing. SH and CU wrote the manuscript. All authors have read and approved the final manuscript.

FUNDING

This work was supported by the United Kingdom Biotechnology and Biological Sciences Research Council (BBSRC) through the Designing Future Wheat (BB/P016855/1) and GEN (BB/P013511/1) ISPs and an Anniversary Future Leader Fellowship to PB (BB/M014045/1). SH was supported by the John Innes Foundation. MK and JD were supported from ERDF project "Plants as a tool for sustainable global development" (No. CZ.02.1.01/0.0/0.0/16_019/0000827). NC acknowledges the funding received from Fulbright and Comisión Nacional de Investigación Científica y Tecnológica (CONICYT) Becas-Chile 72111195. This research was also supported in part by the NBI Computing infrastructure for Science (CiS) group through the HPC resources.

ACKNOWLEDGMENTS

We would like to acknowledge A. Smith and M. Banfield (John Innes Centre) for helpful input during this project. We thank J. Vrána, M. Kubaláková and R. Šperková (Institute of Experimental Botany, Olomouc) for assistance with chromosome flow sorting and chromosome DNA amplification. We also thank J. Dubcovsky (University of California, Davis) for hosting the field trials at the UC Davis in 2016–2017. We also acknowledge the help of the field and glasshouse teams at the JIC and the UC Davis, as well as the help of G. Chilvers at the University of East Anglia ICP-OES

platform. Finally, we thank B. Clavijo and the wheat pan-genome team at the Earlham Institute for allowing the use of the Kronos assembly before final publication. This manuscript has been released as a Pre-Print at https://www.biorxiv.org/content/10.1101/622076v1 (Harrington et al., 2019a).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.00963/full#supplementary-material

REFERENCES

10+Wheat Genomes Project (2016). *The Wheat 'Pan Genome'*. Available at: http://www.10wheatgenomes.com/ (accessed January 25, 2019).

Acevedo-Garcia, J., Spencer, D., Thieron, H., Reinstädler, A., Hammond-Kosack, K., Phillips, A. L., et al. (2017). *mlo*-based powdery mildew resistance in hexaploid bread wheat generated by a non-transgenic Tilling approach. *Plant Biotechnol. J.* 15, 367–378. doi: 10.1111/pbi.12631

Altschul, S. F., Gish, W., Miller, W., Myers, E. W., and Lipman, D. J. (1990). Basic local alignment search tool. *J. Mol. Biol.* 215, 403–410. doi: 10.1006/jmbi.1990.9999

Avni, R., Zhao, R., Pearce, S., Jun, Y., Uauy, C., Tabbita, F., et al. (2014). Functional characterization of *Gpc-1* genes in hexaploid wheat. *Planta* 239, 313–324. doi: 10.1007/s00425-013-1977-y

Borrill, P., Adamski, N., and Uauy, C. (2015). Genomics as the key to unlocking the polyploid potential of wheat. *New Phytol.* 208, 1008–1022. doi: 10.1111/nph.13533

Borrill, P., Harrington, S. A., Simmonds, J., and Uauy, C. (2018). Identification of transcription factors regulating senescence in wheat through gene regulatory network modelling. *bioRxiv* 456749. doi: 10.1104/pp.19.00380

Borrill, P., Harrington, S. A., and Uauy, C. (2019). Applying the latest advances in genomics and phenomics for trait discovery in polyploid wheat. *Plant J.* 97, 56–72. doi: 10.1111/tpj.14150

Borrill, P., Ramirez-Gonzalez, R., and Uauy, C. (2016). EXPVIP: a customizable RNA-seq data analysis and visualization platform. *Plant Physiol.* 170:2172. doi: 10.1104/pp.15.01667

Brinton, J., and Uauy, C. (2019). A reductionist approach to dissecting grain weight and yield in wheat. *J. Integr. Plant Biol.* 61, 337–358. doi: 10.1111/jipb.12741

Clark, J. W., and Donoghue, P. C. J. (2018). Whole-genome duplication and plant macroevolution. *Trends Plant Sci.* 23, 933–945. doi: 10.1016/j.tplants.2018.07.006

Clavijo, B. J., Garcia Accinelli, G., Wright, J., Heavens, D., Barr, K., Yanes, L., et al. (2017a). W2rap: a pipeline for high quality, robust assemblies of large complex genomes from short read data. *bioRxiv* 110999.

Clavijo, B. J., Venturini, L., Schudoma, C., Accinelli, G. G., Kaithakottil, G., Wright, J., et al. (2017b). An improved assembly and annotation of the allohexaploid wheat genome identifies complete families of agronomic genes and provides genomic evidence for chromosomal translocations. *Genome Res.* 27, 885–896. doi: 10.1101/gr.217117.116

Dodsworth, S., Chase, M. W., and Leitch, A. R. (2016). Is post-polyploidization diploidization the key to the evolutionary success of angiosperms? *Bot. J. Linn. Soc.* 180, 1–5. doi: 10.1111/boj.12357

Doležel, J., Vrána, J., Šafář, J., Bartoš, J., Kubaláková, M., and Šimková, H. (2012). Chromosomes in the flow to simplify genome analysis. *Funct. Integr. Genomics* 12, 397–416. doi: 10.1007/s10142-012-0293-0

Dubcovsky, J., and Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316:1862. doi: 10.1126/science.1143986

Fu, D., Szücs, P., Yan, L., Helguera, M., Skinner, J. S., Von Zitzewitz, J., et al. (2005). Large deletions within the first intron in *Vrn-1* are associated with spring growth habit in barley and wheat. *Mol. Genet. Genomics* 273, 54–65. doi: 10.1007/s00438-004-1095-4

Garrison, E., and Marth, G. (2012). *Haplotype-based Variant Detection from Short-read Sequencing*. Available: https://ui.adsabs.harvard.edu/abs/2012arXiv1207.3907G (accessed July 01, 2012).

Gebert, M., Meschenmoser, K., Svidová, S., Weghuber, J., Schweyen, R., Eifler, K., et al. (2009). A root-expressed magnesium transporter of the *Mrs2/Mgt* gene family in *Arabidopsis thaliana* allows for growth in Low-Mg2+ environments. *Plant Cell* 21:4018. doi: 10.1105/tpc.109.070557

Giorgi, D., Farina, A., Grosso, V., Gennaro, A., Ceoloni, C., and Lucretti, S. (2013). FISHIS: fluorescence in situ hybridization in suspension and chromosome flow sorting made easy. *PLoS One* 8:e57994. doi: 10.1371/journal.pone.0057994

Greenwood, J. R., Finnegan, E. J., Watanabe, N., Trevaskis, B., and Swain, S. M. (2017). New alleles of the wheat domestication gene *Q* reveal multiple roles in growth and reproductive development. *Development* 144:1959. doi: 10.1242/dev.146407

Harrington, S. A., Cobo, N., Karafiátová, M., DoleŽel, J., Borrill, P., and Uauy, C. (2019a). Identification of a dominant chlorosis phenotype through a forward screen of the *Triticum turgidum* cv. Kronos TILLING population. *bioRxiv* 622076.

Harrington, S. A., Overend, L. E., Cobo, N., Borrill, P., and Uauy, C. (2019b). Conserved residues in the wheat (*Triticum aestivum*) Nam-A1 Nac domain are required for protein binding and when mutated lead to delayed peduncle and flag leaf senescence. *bioRxiv* 573881.

Hoagland, D. R., and Arnon, D. I. (1950). *The Water-Culture Method for Growing Plants Without Soil*, Vol. 347. Berkeley, CA: College of Agriculture, University of California, 32.

Huang, X., Feng, Q., Qian, Q., Zhao, Q., Wang, L., Wang, A., et al. (2009). High-throughput genotyping by whole-genome resequencing. *Genome Res.* 19, 1068–1076. doi: 10.1101/gr.089516.108

International Wheat Genome Sequencing Consortium (2018). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361:eaar7191. doi: 10.1126/science.aar7191

Jupe, F., Witek, K., Verweij, W., Sliwka, J., Pritchard, L., Etherington, G. J., et al. (2013). Resistance gene enrichment sequencing (RenSeq) enables reannotation of THE NB-LRR gene family from sequenced plant genomes and rapid mapping of resistance loci in segregating populations. *Plant J.* 76, 530–544. doi: 10.1111/tpj.12307

Kang, K., Kim, Y.-S., Park, S., and Back, K. (2009). Senescence-induced serotonin biosynthesis and its role in delaying senescence in rice leaves. *Plant Physiol.* 150, 1380–1393. doi: 10.1104/pp.109.138552

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12, 357–360. doi: 10.1038/nmeth.3317

Kimura, E., Bell, J., Trostle, C., Neely, C., and Drake, D. (2016). *Potential Causes of Yellowing During the Tillering Stage of Wheat in Texas. Texas A&M AgriLife Extension*. Available: https://agrilifecdn.tamu.edu/texaslocalproduce-2/files/2018/07/Potential-Causes-of-Yellowing-During-the-Tillering-Stage-of-Wheat-in-Texas.pdf (accessed June 06, 2019).

Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., et al. (2017). Uncovering hidden variation in polyploid wheat. *Proc. Natl. Acad. Sci. U.S.A.* 114:E913. doi: 10.1073/pnas.1619268114

Krzywinski, M., Schein, J., Birol, İ, Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19, 1639–1645. doi: 10.1101/gr.092759.109

Kubaláková, M., Vrána, J., Číhalíková, J., Šimková, H., and Doležel, J. (2002). Flow karyotyping and chromosome sorting in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 104, 1362–1372. doi: 10.1007/s00122-002-0888-2

Li, H. (2013). *Aligning Sequence Reads, Clone Sequences and Assembly Contigs with BWA-MEM*. Available: https://ui.adsabs.harvard.edu/abs/2013arXiv1303.3997L (accessed March 01, 2013).

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51, 885–895. doi: 10.1038/s41588-019-0381-3

Mamanova, L., Coffey, A. J., Scott, C. E., Kozarewa, I., Turner, E. H., Kumar, A., et al. (2010). Target-enrichment strategies for next-generation sequencing. *Nat. Methods* 7, 111–118. doi: 10.1038/nmeth.1419

Mo, Y., Howell, T., Vasquez-Gross, H., De Haro, L. A., Dubcovsky, J., and Pearce, S. (2018). Mapping causal mutations by exome sequencing in a wheat Tilling population: a tall mutant case study. *Mol. Genet. Genomics* 293, 463–477. doi: 10.1007/s00438-017-1401-6

NOAA National Centers For Environmental Information (2017). *State of the Climate: National Climate Report for Annual 2017*. Asheville, NC: National Centers for Environmental Information.

Paterson, A. H., Wang, X., Li, J., and Tang, H. (2012). "Ancient and recent polyploidy in monocots," in *Polyploidy and Genome Evolution*, eds P. Soltis and D. E. Soltis (Heidelberg: Springer).

Payandeh, J., Pfoh, R., and Pai, E. F. (2013). The structure and regulation of magnesium selective ion channels. *Biochim. Biophys. Acta* 1828, 2778–2792. doi: 10.1016/j.bbamem.2013.08.002

Pearce, S., Tabbita, F., Cantu, D., Buffalo, V., Avni, R., Vazquez-Gross, H., et al. (2014). Regulation of Zn and Fe transporters by the *Gpc1* gene during early wheat monocarpic senescence. *BMC Plant Biol.* 14:368. doi: 10.1186/s12870-014-0368-2

Peng, J., Richards, D. E., Hartley, N. M., Murphy, G. P., Devos, K. M., Flintham, J. E., et al. (1999). 'Green revolution' genes encode mutant gibberellin response modulators. *Nature* 400, 256–261. doi: 10.1038/22307

R Core Team (2018). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing.

Ramírez-González, R. H., Borrill, P., Lang, D., Harrington, S. A., Brinton, J., Venturini, L., et al. (2018). The transcriptional landscape of polyploid wheat. *Science* 361:eaar6089. doi: 10.1126/science.aar6089

Ramirez-Gonzalez, R. H., Segovia, V., Bird, N., Fenwick, P., Holdgate, S., Berry, S., et al. (2015a). Rna-Seq bulked segregant analysis enables the identification of high-resolution genetic markers for breeding in hexaploid wheat. *Plant Biotechnol. J.* 13, 613–624. doi: 10.1111/pbi.12281

Ramirez-Gonzalez, R. H., Uauy, C., and Caccamo, M. (2015b). PolyMarker: a fast polyploid primer design pipeline. *Bioinformatics* 31, 2038–2039. doi: 10.1093/bioinformatics/btv069

Rodríguez-Leal, D., Lemmon, Z. H., Man, J., Bartlett, M. E., and Lippman, Z. B. (2017). Engineering quantitative trait variation for crop improvement by genome editing. *Cell* 171, 470–480.e8. doi: 10.1016/j.cell.2017.08.030

Šimková, H., Svensson, J. T., Condamine, P., Hřibová, E., Suchánková, P., Bhat, P. R., et al. (2008). Coupling amplified Dna from flow-sorted chromosomes to high-density Snp mapping in barley. *BMC Genomics* 9:294. doi: 10.1186/1471-2164-9-294

Simons, K. J., Fellers, J. P., Trick, H. N., Zhang, Z., Tai, Y.-S., Gill, B. S., et al. (2006). Molecular Characterization of the major wheat domestication gene *Q*. *Genetics* 172, 547–555. doi: 10.1534/genetics.105.044727

Singh, S., Giri, M. K., Singh, P. K., Siddiqui, A., and Nandi, A. K. (2013). Down-regulation of *OSSAG12-1* results in enhanced senescence and pathogen-induced cell death in transgenic rice plants. *J. Biosci.* 38, 583–592. doi: 10.1007/s12038-013-9334-7

Snowball, K., and Robson, A. D. (1991). *Nutrient Deficiencies and Toxicities in Wheat: A Guide for Field Identification*. Mexico City, MX: CIMMYT.

Soltis, P. S., and Soltis, D. E. (2016). Ancient WGD events as drivers of key innovations in angiosperms. *Curr. Opin. Plant Biol.* 30, 159–165. doi: 10.1016/j.pbi.2016.03.015

Steuernagel, B., Periyannan, S. K., Hernández-Pinzón, I., Witek, K., Rouse, M. N., Yu, G., et al. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* 34, 652–655. doi: 10.1038/nbt.3543

Uauy, C. (2017). Wheat genomics comes of age. *Curr. Opin. Plant Biol.* 36, 142–148. doi: 10.1016/j.pbi.2017.01.007

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., and Dubcovsky, J. (2006). A NAC GENE regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314, 1298–1301. doi: 10.1126/science.1133649

Uauy, C., Wulff, B. B. H., and Dubcovsky, J. (2017). Combining traditional mutagenesis with new high-throughput sequencing and genome editing to reveal hidden variation in polyploid wheat. *Annu. Rev. Genet.* 51, 435–454. doi: 10.1146/annurev-genet-120116-024533

Vrána, J., Cápal, P., Šimková, H., Karafiátová, M., Čížková, J., and Doležel, J. (2016). Flow analysis and sorting of plant chromosomes. *Curr. Protoc. Cytom.* 78, 5.3.1–5.3.43. doi: 10.1002/cpcy.9

Vrána, J., Kubaláková, M., Šimková, H., Číhalíková, J., Lysák, M. A., and Doležel, J. (2000). Flow sorting of mitotic chromosomes in common wheat (*Triticum aestivum* L.). *Genetics* 156, 2033–2041.

Vullo, A., Allot, A., Zadissia, A., Yates, A., Luciani, A., Moore, B., et al. (2017). Ensembl Genomes 2018: an integrated omics infrastructure for non-vertebrate species. *Nucleic Acids Res.* 46, D802–D808. doi: 10.1093/nar/gkx1011

Wang, W., Simmonds, J., Pan, Q., Davidson, D., He, F., Battal, A., et al. (2018). Gene editing and mutagenesis reveal inter-cultivar differences and additivity in the contribution of *TAGW2* homoeologues to grain size and weight in wheat. *Theor. Appl. Genet.* 131, 2463–2475. doi: 10.1007/s00122-018-3166-7

Wang, Y., He, Y., Yang, M., He, J., Xu, P., Shao, M., et al. (2016). Fine mapping of a dominant gene conferring chlorophyll-deficiency in *Brassica napus*. *Sci. Rep.* 6:31419. doi: 10.1038/srep31419

Warnes, G. R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A., et al. (2019). *gplots: Various R Programming Tools for Plotting Data. R package version 3.0.1.1*.

Wellburn, A. R. (1994). The spectral determination of chlorophylls a and b, as well as total carotenoids, using various solvents with spectrophotometers of different resolution. *J. Plant Physiol.* 144, 307–313. doi: 10.1007/s11120-013-9834-1

Wickham, H. (2016). *ggplot2: Elegant Graphics for Data Analysis*. New York, NY: Springer-Verlag.

Wickham, H., François, R., Henry, L., and Müller, K. (2019). *dplyr: A Grammar of Data Manipulation. R package version 0.8.0.1*.

Wickham, H., and Henry, L. (2018). *tidyr: Easily Tidy Data with 'spread()' and 'gather()' Functions. R package version 0.8.2*.

Wu, H., Shi, N., An, X., Liu, C., Fu, H., Cao, L., et al. (2018). Candidate genes for yellow leaf color in common wheat (*Triticum aestivum* L.) and major related metabolic pathways according to transcriptome profiling. *Int. J. Mol. Sci.* 19:1594. doi: 10.3390/ijms19061594

Wysoker, A., Handsaker, B., Marth, G., Abecasis, G., Li, H., Ruan, J., et al. (2009). The sequence alignment/map format and samtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Yan, L., Helguera, M., Kato, K., Fukuyama, S., Sherman, J., and Dubcovsky, J. (2004). Allelic variation at THE *VRN-1* PROMOTER region in polyploid wheat. *Theor. Appl. Genet.* 109, 1677–1686. doi: 10.1007/s00122-004-1796-4

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene *Vrn1*. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6263–6268. doi: 10.1073/pnas.0937399100

Zadoks, J. C., Chang, T. T., and Konzak, C. F. (1974). A decimal code for the growth stages of cereals. *Weed Res.* 14, 415–421. doi: 10.1111/j.1365-3180.1974.tb01084.x

Zong, Y., Wang, Y., Li, C., Zhang, R., Chen, K., Ran, Y., et al. (2017). Precise base editing in rice, wheat and maize with a Cas9-cytidine deaminase fusion. *Nat. Biotechnol.* 35, 438–440. doi: 10.1038/nbt.3811

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 18 September 2019
doi: 10.3389/fpls.2019.01110

The Use of Pentaploid Crosses for the Introgression of *Amblyopyrum muticum* and D-Genome Chromosome Segments Into Durum Wheat

*Manel Othmeni[1], Surbhi Grewal[1], Stella Hubbart-Edwards[1], Caiyun Yang[1], Duncan Scholefield[1], Stephen Ashling[1], Amor Yahyaoui[2], Perry Gustafson[3], Pawan K. Singh[2], Ian P. King[1] and Julie King[1]**

[1] Nottingham BBSRC Wheat Research Centre, Division of Plant and Cop Sciences, School of Biosciences, The University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire, United Kingdom, [2] International Maize and Wheat Improvement Center (CIMMYT) Mexico, Mexico City, Mexico, [3] Division of Plant Sciences, University of Missouri, Columbia, MO, United States

OPEN ACCESS

Edited by:
Agata Gadaleta,
University of Bari Aldo Moro, Italy

Reviewed by:
Daryl LaVerne Klindworth,
Edward T. Schafer Agricultural Research Center, United States
Assaf Distelfeld,
Tel Aviv University, Israel
Peng Zhang,
University of Sydney, Australia

Correspondence:
Julie King
Julie.King@nottingham.ac.uk

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 14 June 2019
Accepted: 13 August 2019
Published: 18 September 2019

Citation:
Othmeni M, Grewal S, Hubbart-Edwards S, Yang C, Scholefield D, Ashling S, Yahyaoui A, Gustafson P, Singh PK, King IP and King J (2019) The Use of Pentaploid Crosses for the Introgression of Amblyopyrum muticum and D-Genome Chromosome Segments Into Durum Wheat. Front. Plant Sci. 10:1110. doi: 10.3389/fpls.2019.01110

The wild relatives of wheat provide an important source of genetic variation for wheat improvement. Much of the work in the past aimed at transferring genetic variation from wild relatives into wheat has relied on the exploitation of the *ph1b* mutant, located on the long arm of chromosome 5B. This mutation allows homologous recombination to occur between chromosomes from related but different genomes, e.g. between the chromosomes of wheat and related chromosomes from a wild relative resulting in the generation of interspecific recombinant chromosomes. However, the *ph1b* mutant also enables recombination to occur between the homologous genomes of wheat, e.g. A/B, A/D, B/D, resulting in the generation of wheat intergenomic recombinant chromosomes. In this work we report on the presence of wheat intergenomic recombinants in the genomic background of hexaploid wheat/*Amblyopyrum muticum* introgression lines. The transfer of genomic rearrangements involving the D-genome through pentaploid crosses provides a strategy by which the D-genome of wheat can be introgressed into durum wheat. Hence, a pentaploid crossing strategy was used to transfer D-genome segments, introgressed with either the A- and/or the B-genome, into the tetraploid background of two durum wheat genotypes Karim and Om Rabi 5 in either the presence or absence of different *Am. muticum* (2n = 2x = 14, TT) introgressions. Introgressions were monitored in backcross generations to the durum wheat parents *via* multi-color genomic *in situ* hybridization (mc-GISH). Tetraploid lines carrying homozygous D-genome introgressions, as well as simultaneous homozygous D- and T-genome introgressions, were developed. Introgression lines were characterized *via* Kompetitive Allele-Specific PCR (KASP) markers and multi-color fluorescence *in situ* hybridization (FISH). Results showed that new wheat sub-genomic translocations were generated at each generation in progeny that carried any *Am. muticum* chromosome introgression irrespective of the linkage group that the segment was derived from. The highest frequencies of homologous recombination were observed

between the A- and the D-genomes. Results indicated that the genotype Karim had a higher tolerance to genomic rearrangements and T-genome introgressions compared to Om Rabi 5. This indicates the importance of the selection of the parental genotype when attempting to transfer/develop introgressions into durum wheat from pentaploid crosses.

Keywords: durum wheat, pentaploid crosses, *Amblyopyrum muticum*, introgression, *in situ* hybridization, Kompetitive Allele-Specific PCR markers

INTRODUCTION

The most important cultivated *Triticum* species are hexaploid bread wheat (2n = 2x = 42; AABBDD, *Triticum aestivum* L. ssp. *aestivum*) and tetraploid durum wheat (2n = 2x = 28; AABB, *Triticum turgidum* L. ssp. *durum*). Tetraploid wheat arose 500,000 years ago from a cross between the wild ancestors of the A-genome, *Triticum urartu* Thum ex. Gandil (2n = 2x = 14; A^uA^u) (Feldman and Levy, 2005), and the B-genome from an *Aegilops speltoides*-like progenitor (Haider, 2013). After domestication, a spontaneous cross of tetraploid wheat as the female parent with the goat grass *Ae. tauschii* Coss. (2n = 2x = 14; DD) approximately 8,000 years ago gave rise to hexaploid bread wheat (Kihara, 1944; McFadden and Sears, 1944, Matsuoka and Nasuda, 2004). The addition of the D-genome to hexaploid wheat conferred baking characteristics and a wide climatic adaptation compared to durum wheat (Zohary et al., 1969) resulting in bread wheat becoming one of the most widely grown crops due to its high yields and nutritional and processing qualities (Shewry and Hey, 2015).

Despite the relatively small growing area (8%) and lower annual production compared to bread wheat, durum wheat remains a major crop in the Mediterranean basin where about 75% of the world's durum wheat is produced (Li et al., 2013; Kabbaj et al., 2017) although Europe and North Africa are also the largest importers of durum wheat (Bonjean et al., 2016). According to data from the International Grain Council, durum wheat production has shown annual fluctuations, largely attributable to abiotic and biotic stresses, e.g., in the Mediterranean area, crops are often exposed to environmental stresses such as high temperature and drought during grain filling (Nazco et al., 2012). Breeding programs have greatly improved durum wheat yield and quality (Magallanes-López et al., 2017). However, the incorporation of new alleles into wheat germplasm is considered essential for the continued improvement of durum wheat productivity.

Wheat is related to a large number of other species, many of which are wild and uncultivated. These wild relatives provide a vast and largely untapped reservoir of genetic variation for agronomically important traits (Friebe et al., 1996; Jauhar and Chibbar, 1999; Qi et al., 2007; Schneider et al., 2008). The incorporation of these traits into wheat has the potential to increase the yield potential. For example, *Ae. speltoides* has been shown to be insect and disease resistant (Elek et al., 2014) and *Thinopyrum bessarabicum* salt tolerant (King et al., 1997).

Abbreviations: WMI, wheat/*Am. muticum* introgression; SA, short arm; LA, long arm.

Among the wild relatives of wheat, *Am. muticum* (2n = 2x = 14; TT) is an annual, native species of Turkey and Armenia (Kilian et al., 2013). This species has been reported to be resistant to environmental stresses (Iefimenko et al., 2015), powdery mildew (Eser, 1998), and leaf rust (Dundas et al., 2015). The introgression of *Am. muticum* into bread wheat is an ongoing project at the Wheat Research Centre (WRC) at the University of Nottingham (King et al., 2013; King et al., 2017) where 218 genome- wide bread wheat/*Am. muticum* introgressions have been developed covering the seven linkage groups of *Am. muticum* (King et al., 2017). Genomic *in situ* hybridization (GISH) analysis revealed that some of the introgression lines also contained intergenomic rearrangements between the A, B, and D sub-genomes of wheat. These intergenomic recombinants, and particularly those that involve the D-genome, can be transferred into durum wheat. Hybridization between bread and durum wheat leads to the production of a pentaploid hybrid (AABBD) with a chromosomal constitution of 2n = 5x = 35 (Kihara, 1924). Depending on the direction of the backcrosses, pentaploid hybrids have the potential to improve both bread wheat and durum (Eberhard et al., 2010; Martin et al., 2013; Kalous et al., 2015).

This paper describes the introgression of both wheat inter-genome rearrangements involving the D-genome and T-genome segments of *Am. muticum* present in hexaploid wheat/*Am. muticum* introgression (WMI) lines into two durum wheat genotypes using pentaploid crosses. The effect of the presence of the T-genome in the WMI lines, the efficiency of the crossing strategy as well as the choice of the durum wheat are discussed.

MATERIALS AND METHODS

Plant Material

The self-fertilized or back-crossed seed of eight hexaploid wheat/*Am. muticum* introgression lines, designated as WMI (wheat/*Am. muticum* introgression) lines, were obtained from the Nottingham/BBSRC Wheat Research Centre (WRC) (King et al., 2017). The WMI lines were characterized by multi-color genomic *in situ* hybridization (mc-GISH) in the BC_3 generation and shown to carry wheat inter-genomic rearrangements involving the D-genome. The genome rearrangements were designated by the letter of the genome involved (A, B, or D). An upper case letter designated the larger segment, a lower case letter the smaller segment. In the case of non centromeric translocations, the two letters were separated by a dash (e.g. A-d), whereas for centromeric translocations, a dot was used

TABLE 1 | Type and number of the D- genome and T-genome introgressions present in the parental introgression lines and the reference of the WMI lines used in the crosses.

Group	Parental lines	Genome translocation*No.	Number of T-genome introgressions	T-genome introgression linkage group	WMI lines used to cross
G-1	BC_3-F_1-157-C	A-d*1	0	–	BC_4-F_1-129
	BC_3-F_1-157-D	D-a*1	0	–	BC_4-F_1-130
	BC_3-F_1-157-E	A-d*1	0	–	BC_3-F_2-130
	BC_3-F_1-172-C	D-a*1	0	–	BC_3-F_2-132
G-2	BC_3-F_1-172-E	D-a-b*1 + A-d*1	3	1T, 3TL, 5T	BC_3-F_2-133
	BC_3-F_1-177-E	D.a-b*1 + A.D*1	2	2T, 4T	BC_3-F_2-134
	BC_3-F_1-244-A	d-A-d*1 + D-a*1	1	6TS.7TL	BC_3-F_2-135
	BC_3-F_1-244-B	A-d*1	2	1TS.3TL, 6TS.7TL	BC_3-F_2-136

*NB: *No.indicates the number of copies, G1, WMI parental lines without a T-genome segment; G2, WMI parental lines carrying T-genome segments.*

(e.g., A.D). Four of the WMI lines also carried one to three different large T-genome segments characterized using the Axiom® Wheat-Relative Genotyping Array (King et al., 2017) (**Table1**). Hence, the WMI lines were categorized into two groups, the G-1 lines without a T introgression/chromosome and the G-2 lines carrying a T introgression/chromosome. Four seeds of each line were germinated and screened for the presence of the D-genome introgression using mc-GISH. Lines that retained the introgressions were then used as the female parent in a pentaploid crossing strategy involving two durum wheat genotypes Karim and Om Rabi 5 (**Figure 1**).

Genomic *In Situ* Hybridization (GISH)

Slides of chromosome spreads were obtained as described in Kato et al. (2004) and King et al. (2017). Mc-GISH of the slides was conducted using the labeled total genomic DNA of the three putative progenitor species of wheat; *T. urartu* (A-genome), *Ae. speltoides* (B-genome) and *Ae. tauschii* (D-genome), as well as *Am. muticum* (T-genome). DNA was extracted from the young leaves using a CTAB method (Zhang et al., 2013) and labeled using the nick translation procedure (Luchniak et al., 2002). Slides were probed with *T. urartu* labeled with Chroma Tide Alexa Fluor 488-5-dUTP (Invitrogen; C11397; green), *Ae. tauschii* with Alexa Fluor 594-5-dUTP (Invitrogen; C11400; red), *Am. muticum* with Alexa Fluor 546 (Invitrogen; C11401; yellow) and the genomic DNA of *Ae. speltoides* fragmented to 300–500bp (using a heat block for 15 min at 110°C) used as a blocking DNA in a ratio of 1:1:2:30.

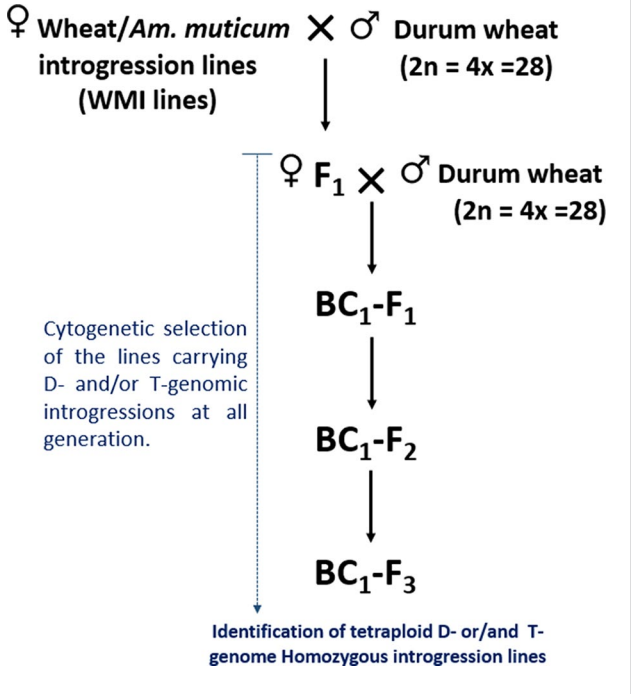


FIGURE 1 | Crossing diagram for the introgression of the D- and T-genome segments identified in WMI lines into durum wheat.

For the detection of T-genome introgression alone some of the lines were probed by single color GISH using the labeled *Am. muticum* genomic DNA with Chroma Tide Alexa Fluor 488-5-dUTP (Invitrogen; C11397; green) and the fragmented genomic DNA of wheat cv. Chinese Spring (300–500bp) as blocking DNA in a ratio of 1:50 per slide.

Slides were counterstained with 4'-6-diamidino-2- phenylindole (DAPI) and analyzed using a high throughput, fully automated Zeiss Axio Imager.Z2 upright epifluorescence microscope (Carl Zeiss Ltd, Oberkochen, Germany). Photographs were taken using a MetaSystems Coolcube 1 m CCD camera. Further slide analysis was carried out using an automated metaphase image capture software, Metafer4, and the ISIS software for image processing (Metasystems GmbH, Altlussheim, Germany).

Fluorescence *in Situ* Hybridization (FISH)

For multi-color fluorescence *in situ* hybridization (mc-FISH), two repetitive DNA sequences, pSc119.2 (McIntyre et al., 1990) and pAs1 (Rayburn and Gill, 1986), were labeled by nick translation with Alexa Fluor 488-5-dUTP (green) and Alexa Fluor 594-5-dUTP (red), respectively, and hybridized to the slides. Subsequent counterstaining and image capture were performed as described for GISH.

Genotyping With KASP™ Markers

Genomic DNA was isolated from leaf tissue of 10-day old seedlings in a 96-well plate as described by Thomson and Henry (1995). All lines showing T- and/or D-genome introgressions were genotyped alongside the two durum wheat genotypes, one Ae. tauschii accession P95-81.1.1-1 obtained from USDA

TABLE 2 | Number of crosses, percentage of crosses setting seed, number of seeds produced and percentage of seed germination at every generation in the two cross-combination of the WMI lines and their subsequent backcross generations to both the Om Rabi 5 and Karim durum wheat genotypes.

Cross-combination	Generation	Number of crosses	Percentage of crosses set seed	Number of crossed seeds produced	Percentage of germination
WMI line/Om Rabi 5	WMI x Om Rabi 5	28	100%	246	64%
	BC_1	31	67%	149	87%
	BC_1-F_2	**	**	**	93.45%
	BC_1-F_3	**	**	**	100%
WMI line/Karim	WMI x Karim	35	82%	242	76%
	BC_1	37	54%	105	85%
	BC_1-F_2	**	**	**	94%
	BC_1-F_3	**	**	**	76.90%

and two bread wheat genotypes, Chinese Spring and Paragon, used as controls. The full set of Langdon disomic D-genome substitution lines (obtained from the USDA), were also used as control lines to verify the specificity of the Kompetitive Allele-Specific PCR (KASP) markers to the D-genome. In these lines a pair of D-genome chromosomes substitute a pair of either the A- or the B-genome chromosomes of the same linkage group (Joppa and Williams, 1988).

A total of 80 D-genome specific KASP™ markers (Grewal et al., 2019) of which 29 markers were polymorphic between wheat and *Am. muticum*, were used for simultaneous detection of the D- and T-genome introgression. For each KASP™ marker, two allele-specific forward primers and one common reverse primer were used (**Supplementary Material**). Genotyping reactions were performed in a ProFlex PCR system (Applied Biosystems by Life Technology) in a final volume of 5 µl with 1 ng genomic DNA, 2.5 µl KASP reaction mix (ROX), 0.068 µl primer mix and 2.43 µl nuclease free water. PCR conditions were set as 15 min at 94°C; 10 touchdown cycles of 10 s at 94°C, 1 min at 65–57°C (dropping 0.8°C per cycle); and 35 cycles of 10 s at 94°C, 1 min at 57°C.

Fluorescence detection of the reactions was performed using a QuantStudio 5 (Applied Biosystems) and the data analyzed using the QuantStudio™ Design and Analysis Software V1.5.0 (Applied Biosystems).

RESULTS

The Development of Durum Wheat D- and/or T-Genome Introgression Lines

Only 23 seed, out of the 32 randomly selected from the eight WMI parental lines, germinated and reached maturity. Cytogenetic screening *via* mc-GISH showed that 15 of these lines had retained at least one copy of the D-genome introgression. Sixty-three crosses were made between these lines and the two durum genotypes to produce the F_1 plants. A further 68 back-crosses were made between the F_1 plants and durum wheat. The total number of crosses, percentage of crosses setting seed, number of seed produced and percentage germination are shown in **Table 2**.

Chromosome counts showed that 70% of the BC_1-F_2 generation and 88.4% of the BC_1-F_3 lines carrying D-genome introgressions had 28 chromosomes. Hence, backcrossing to the durum wheat parent had gradually decreased the average chromosome number through the loss of D-genome univalents.

Mc-GISH Analysis of the D-Genome Introgression Lines

Only lines carrying D-genome introgressions were selected using mc-GISH at every generation. The percentage of F_1, BC_1-F_1, BC_1-F_2, and BC_1-F_3 carrying D-genome introgressions was 33.3, 35.5, 36.5, and 85, respectively, with the percentage retention slightly higher in the lines produced with Karim in all generations except the BC_1-F_3 (**Table 3**). The number of D-genome introgressions per line varied between one to three introgressions, with most lines carrying a single D-genome introgression. A higher number of lines carrying inter-genome introgressions were identified in the lines belonging to the G-2 plants, i.e. those in which the initial parental line had one to three introgressions of *Am. muticum*. New D-genome introgressions that were not present in the WMI parental lines were identified in all generations, with the highest number occurring in the BC_1-F_1 and BC_1-F_2 generations in Karim and Om Rabi 5 cross-combination, respectively (**Table 3**). However, these new introgressions occurred only in the G-2 group and mainly in the progeny from the crosses to Karim.

The most frequent introgressions identified initially in the WMI parental lines were D-introgressions into the A-genome, with recombination in the telomeric region (A-d or D-a introgressions—**Table 4**). The newly formed introgressions mainly involved either the D-genome with the A-genome or with both the A- and the B-genomes. Overall, a higher number of different AD (e.g., A-d, d-A-d, D-a, and A.D) and ABD (e.g., D.a-b, D-a-b, D.a-d, and B.a-d) recombinants were identified compared to BD (e.g., B.D and B-d) or AB (e.g., A.B and B-a) recombinants (**Table 4**). A-d recombinants, consisting of a small D-genome segment introgressed into either the long arm (LA) or the short arm (SA) of an A-chromosome, were retained the most between consecutive selfed generations.

The D-genome introgressions identified in the BC_1-F_2 progeny from Om Rabi 5 originated from only two BC_1-F_1 plants—a G-1 plant and a G-2 plant. Five progenies from the G-1 plant were found to contain a homozygous D-genome introgression of the A-$d_{(SA)}$ type. This introgression was initially identified in

TABLE 3 | Summary result table on the percentage of the retention and occurrence of new D-genome introgressions at the F_1 and subsequent backcross generations of the WMI lines to Om Rabi 5 and Karim.

Cross-combination	Generation	Number of lines screened	Percentage of lines with a D-genome translocation	Average total chromosome number	Percentage of lines that retained D-genome translocation	Percentage of lines with new D-genome translocation
WMI line/Om Rabi 5	F_1	50	32%	34	62%	38%
	BC_1-F_1	28	35%	31	100%	0%
	BC_1-F_2	43	30%	31	46%	54%
	BC_1-F_3	9	100%	28	100%	0%
WMI line/Karim	F_1	54	35%	34	79%	21%
	BC_1-F_1	36	36%	30	46%	54%
	BC_1-F_2	47	43%	28	85%	15%
	BC_1-F_3	41	70%	28	96%	4%

TABLE 4 | Summary table of the introgressions identified and retained at the F_1 and subsequent backcross generations of the WMI lines crossed to Om Rabi 5 and Karim genotypes and occurrence of new D-genome introgressions in the G-2 group.

Cross-combination	Generation	Type of D-genomic introgressions retained from previous generation (G-1 and G-2)	Type of the newly formed recombinant chromosomes (G-2)
WMI line/Om Rabi 5	F_1	A-$d_{(SA)}$, D-a, D-a-b,	D-a-b, A.D, d-A-d, D-a, A.B
	BC_1-F_1	A-$d_{(SA)}$, D-a, D-a-b	0
	BC_1-F_2	A-$d_{(SA)}$	D-a, D-a-b, A.D, d-A-d, B.D, B-A-d, B-d, B-a-d
	BC_1-F_3	A-$d_{(SA)}$	0
WMI line/Karim	F_1	A-$d_{(SA)}$, D-a, D.a-b, A.D	D-a-b, A.D, B.D, d-A-d
	BC_1-F_1	A-$d_{(SA)}$, D-a-b, A.D, d-A-d	A-$d_{(LA)}$, B-$d_{(SA)}$, B-$d_{(LA)}$, D-a, A.B
	BC_1-F_2	A-$d_{(SA)}$, A-$d_{(LA)}$, D.a-b, A.D, d-A-d	D-a, B-a
	BC_1-F_3	A-$d_{(SA)}$, A-$d_{(LA)}$, D.a-b, A.D, D-a	B-A-d

NB: SA and LA stands for the introgression of the small segment (lowercase letter) in the short or long arm of the chromosome (uppercase letter), respectively. G-1, WMI parental lines without T-genome segment; G-2, WMI parental lines carrying T-genome segments.

the parental hexaploid WMI line and hence, was successfully transferred into Om Rabi 5. Five of the seven progeny from the G-2 BC_1-F_1 line showed the presence of either a single copy (3 lines) or two copies (2 lines) of a large T-genome introgression. One to three new D-genome introgressions were also identified in all seven lines screened.

The BC_1-F_2 Karim lines containing D-genome introgressions were progeny of the same G-2 WMI parental line (BC_3-F_2-134). The D-genome introgressions were all telomeric, recombined with an A-chromosome (A-$d_{(SA)}$ or A-$d_{(LA)}$). Ninety-three percent of the lines carrying D-genome introgressions also contained at least one T-genome introgression or chromosome. Hence, simultaneous introgression of the D- and the T-genomes was identified in the tetraploid lines. Mc-GISH showed that the T-genome introgression was recombined with a B-genome chromosome near the telomere (T-b) and had substituted a B-genome chromosome. Homozygous tetraploid introgression lines for both the D- and the T-genomes were identified in the BC_1-F_3 generation in some of the selfed progeny of the tetraploid BC_1-F_2 lines.

Genotyping of the Introgression Lines

Five D-genome introgression lines (one in the Om Rabi 5 background and four in the Karim background) were isolated in the BC_1-F_3 generation. While the D-genome introgression in Om Rabi 5 background was present in the parental WMI line, the four D-genome introgressions into Karim were identified in four BC_1-F_1 lines which were derived from the same WMI line. A total of 80 KASP markers distributed across the seven linkage groups of the D- and T-genomes (except for the 1DS and 1TS arms) were used to characterise both the D-genome and T-genome introgressions in the progenies of the G2-WMI lines (**Figure 2**). A total of 16 D-genome introgression lines were genotyped (including at least two sister lines carrying each of the five introgressions described above).

Genotyping identified the D-genome introgression into Om Rabi 5 as the telomeric region of the 5DS chromosome arm *via* the amplification of the closely linked KASP markers WRC0669 and WRC0670 located at 24,574,003 and 24,971,617 bp (base pair) on wheat chromosome 5D (International Wheat Genome Sequercincy Consortium et al., 2018). However, the absence of

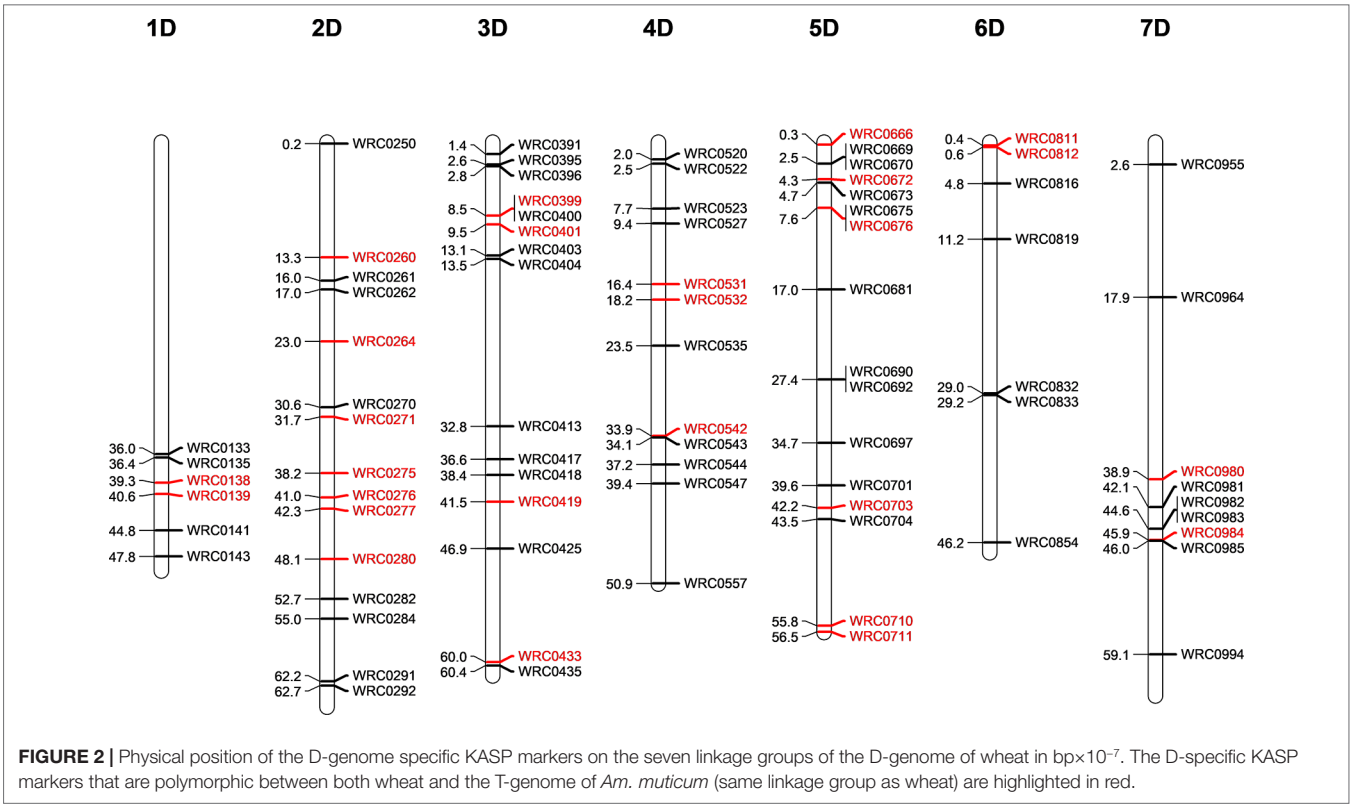


FIGURE 2 | Physical position of the D-genome specific KASP markers on the seven linkage groups of the D-genome of wheat in bp×10^{-7}. The D-specific KASP markers that are polymorphic between both wheat and the T-genome of *Am. muticum* (same linkage group as wheat) are highlighted in red.

amplification of marker WRC0666 (located at 3,031,923 bp) on chromosome 5D indicates that a deletion might have occurred in this region of the 5DS introgressed segment. The introgression was confirmed as homozygous by mc-GISH.

Three of the four D-introgressions into Karim were A-$d_{(LA)}$ introgressions. Two of these segments were characterized as the telomeric region of 1DL and one as the telomeric region of 6DL. KASP markers detected a small difference in segment size between the two 1DL introgressions. Only one marker, WRC0143, located at the telomeric region of 1DL amplified in the two BC_1-F_3-202-A and -B sister lines. However, the introgression in the BC_1-F_3-214, 215, 312, and 315 sister lines were shown to be larger due to the amplification of WRC0143 and WRC0141. The 6DL introgression was identified as a telomeric segment through amplification of marker WRC0854. The fourth introgression into Karim (A-$d_{(SA)}$) was characterized as the very telomeric region of 2DS *via* the amplification of marker WRC0250.

The polymorphic wheat/*Am. muticum* KASP markers (highlighted in red in **Figure 2**) were able to detect the presence of T-genome introgressions in all the introgression lines from Karim. The two 6DL introgression lines had retained both the 2T and 4T *Am. muticum* introgressions, originally present in the WMI parental line. The remaining lines, however, had retained only the 4T introgression (mc-GISH showed that the large 4T introgression in all these lines had recombined with a B-genome chromosome). Combined analysis with genotyping and mc-GISH identified lines containing simultaneous homozygous introgressions of 4T and either 1DL or 6DL. The introgression remained heterozygous in the BC_1-F_3 line analyzed, although the 4T introgression was again homozygous. GISH analysis of the two tetraploid BC_1-F_3-324 sister lines showed that the 2T and 4T *Am. muticum* introgressions, identified *via* KASP, were both homozygous substituting two A- and two B-chromosomes. However, these two lines were both sterile and failed to produce seed (BC_1-F_4 seed was produced from the rest of the introgression lines).

Mc-FISH Characterization of the Introgression Lines

Mc-FISH based karyotyping of the introgression lines was used to identify the wheat chromosomes involved in the introgressions by comparison with the mc-FISH karyotype of Chinese Spring (Tang et al., 2014). Mc-FISH of the homozygous Om Rabi 5 5DS introgression identified it as being recombined with the short arm of chromosome 5A (**Figures 3 A–C**). Only two of the D-genome introgressions into Karim could be characterized as the 1DL introgression identified in the BC_3-F_2-202 sister lines and the 6DL introgression were too small to detect. The 1DL introgression identified in the BC_1-F_3-214, 215, 312, and 315 sister lines, however, was recombined with the long arm of chromosome 1A (**Figures 3 D–F**) and the 2DS introgression with the short arm of chromosome 2A (**Figures 3 G–I**). The B-genome introgression, recombined with the large 4T introgression, was also too small to detect. However, this single or homozygous 4T-b recombinant chromosome was found to have substituted either a single or a pair of 4B chromosomes (**Figures 3 E–J**).

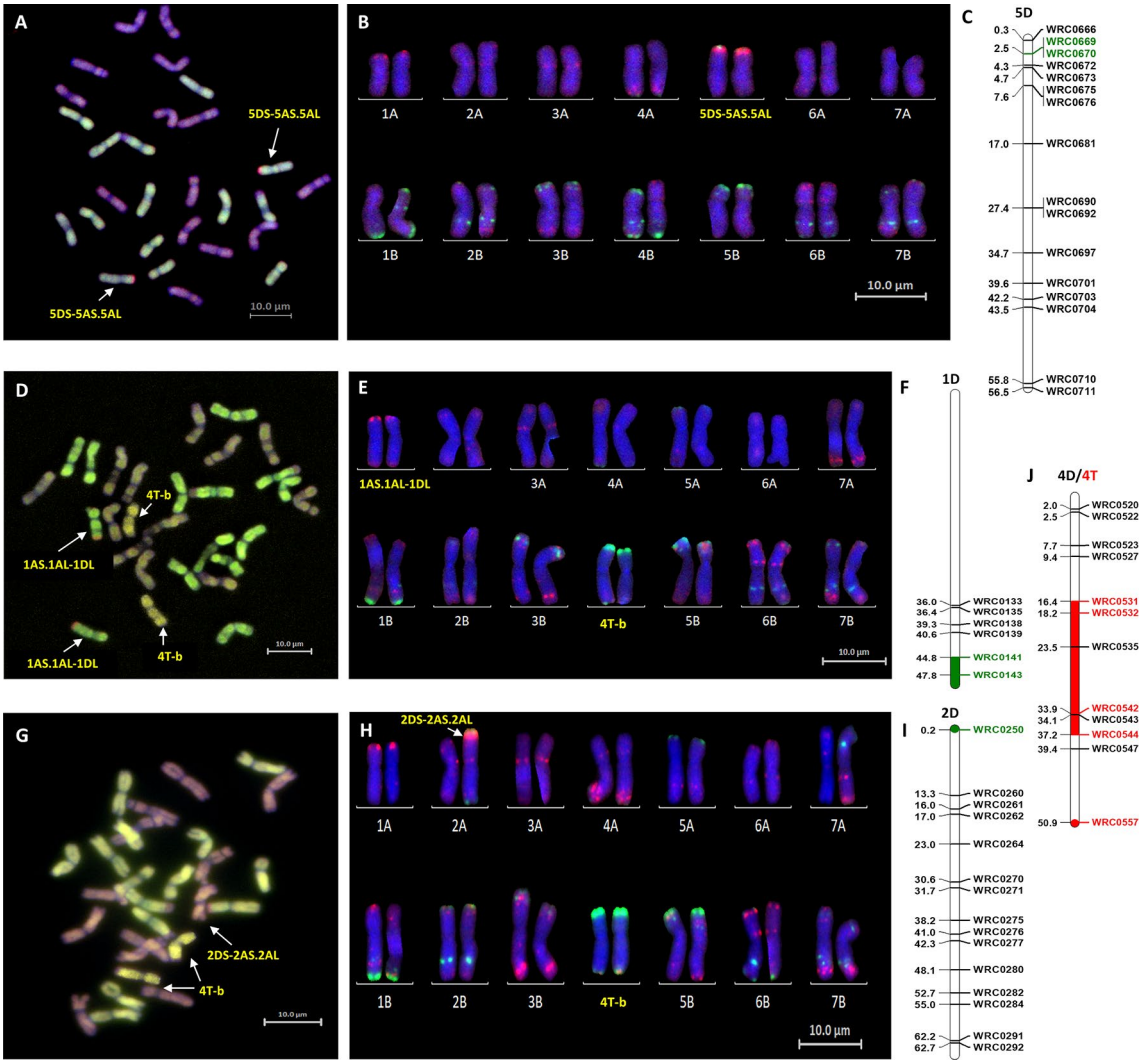


FIGURE 3 | Molecular and cytogenetic characterisation of D-genome and T-genome introgression lines. **(A**, **D**, **G)** Mc-GISH showing the D-genome and T-genome introgression (A-genome in green, B-genome in purple, D-genome in red and T-genome in yellow), **(B**, **E**, **H)** mc-FISH based karyotype using the Oligo-pAs.1 (red) and Oligo-pSc119.2 (green) probes counterstained with DAPI (blue) **(C**,**F**, **I)** physical position (in bp×10^{-7}) of the 5DS, 1DL and 2DS introgressions (green markers and region) using D-genome specific KASP markers showing the D-genome introgressions as **(A**, **B**, **C)** 5DS-5AS.5AL in the genomic backgroung of Om Rabi 5, and as **(D**, **E**, **F)** 1AS.1AL-1DL and **(G**, **H**, **I)** 2DS-2AS.2AL in the BC_1-F_3-315-E and BC_1-F_3-141-A lines, respectively, in the genomic background of Karim. **(J)** characterization of the T-genome introgression as a 4T chromosome recombined in its telomeric long arm with a small B-genome segment noted as 4T-b substituting the pair 4B chromosomes using wheat/T-genome polymorphic KASP markers (red marker and region) in both the BC_1-F_3-214-B and BC_1-F_3-141-A lines, respectively.

DISCUSSION

Pentaploid crosses between bread and durum wheat have previously been shown to generate viable F_1 seed that can be used in a backcrossing programme to either of the parents (Eberhard et al., 2010; Martin et al., 2013; Kalous et al., 2015). The presence of inter-genomic rearrangements in the hexaploid background of wheat/wild relative introgression lines [such as those identified in wheat/*Am. muticum* introgression lines by King et al. (2017)] can thus be used for the introgression of the D-genome of bread wheat into durum wheat. While the overall aim of the current programme was to introgress the D-genome, the crosses also had the potential to increase the genetic variability of the durum A- and B-genomes through recombination with their homologues of bread wheat.

In crosses involving parents of different ploidy levels, it has been shown that using the higher ploidy level genotype as the maternal parent is generally more successful in producing viable F_1 progeny (Ramsey and Schemske, 1998; Kalous et al., 2015). In pentaploid wheat crosses, the hexaploid parent is usually used as the female parent (Padmanaban et al., 2017a; Padmanaban et al., 2017b) and thus, the hexaploid WMI lines were used here as the female parent. Viable F_1 seeds were obtained with both durum wheat parents, Om Rabi 5 and Karim. However, a higher seed set was obtained in the F_1 and the BC_1 generations using Om Rabi 5 as compared to Karim suggesting a higher crossing compatibility of the Om Rabi

5 genotype in the crosses with the bread wheat used here. Other studies have highlighted the importance of the parental choice, as well as the direction of the cross, in the production of viable pentaploid hybrids (reviewed in Padmanaban et al., 2017b).

Mc-GISH was used to visualize the three genomes of wheat and thus, the presence of inter-genomic recombinant chromosomes within the wheat genome of the original introgression lines. This technique has been widely used for studying genome rearrangements, alien introgressions and the discrimination between different genomes in polyploid cereals (Schwarzacher et al., 1989; Schwarzacher et al., 1992; Schubert et al., 2001; Silva and Souza, 2013).

A higher number of lines carrying D-genome introgressions were distinguished in the progeny of Karim in all generations as compared to Om Rabi 5. This could indicate a higher tolerance of Karim to the presence of the D-genome introgressions emphasizing the importance of the durum parent selected for the work. However, a higher seed set with Om Rabi 5 does indicate that the choice of durum parent might not be straight forward. In addition to the D-genome introgressions present in the parental WMI lines, new wheat sub-genome introgressions involving the D- with either the A- and/or the B-genome and the A- with the B-genome were identified at all generations from the F_1 to the BC_1-F_2. A mc-FISH karyotype [based on the karyotype for Chinese Spring developed by Tang et al. (2014)] was used to identify the wheat chromosomes present in the tetraploid BC_1-F_3 plants containing single or homozygous D-genome introgressions from both cross combinations. This showed the presence of a pair of 5B chromosomes in all lines. Since wild-type wheat (Paragon) was used to develop the WMI parental lines, instead of a *ph1b* mutant wheat (King et al., 2017), the inter-genomic rearrangements that occurred in the later generations were not due to the absence of the *Ph1* gene. However, the presence of more than two genomes (A, B, D, and T) and unequal chromosome numbers in one cell could have promoted abnormal meiotic behavior leading to homologous paring. Wheat chromosomes in the selfed progeny of wheat/rye monosomic addition lines, such as 1R, 4R and 6R, show abnormal behavior at meiosis resulting in the elimination or the addition of some of the wheat chromosomes e.g., three 4A-chromosomes were observed in one of the progeny from a 7R monosomic addition line and chromosomes 5A and 4B were eliminated from some of the progeny of the 6R monosomic addition line in addition to alterations of the wheat chromosomes (Fu et al., 2012; Fu et al., 2013).

In the present study, new recombinant events occurred only in lines belonging to the G-2 group with at least one large T-genome introgression/chromosome present in the parental lines. *Am. muticum* is known to contain genes that promote pairing between homologous chromosomes/suppress the effect of the *Ph1* gene in hybrids with allopolyploid wheat (Dvorak, 1972; Dover and Riley, 1972). Similarly, two major *Ph1* suppressor loci, *Su1-Ph1* and *Su2-Ph1* were mapped on the distal end of the long arm of chromosomes 3S and 7S, respectively, in *Ae. speltoides* (Dvorak et al., 2006; Li et al., 2017). It may be possible that some of the introgressed segments from *Am. muticum* also carry a *Ph1* suppressor gene. However, new introgressions were distinguished in the progeny of G-2 WMI lines carrying different introgressions of *Am. muticum* such as 2T, 4T, 6TS.7TL and 1TS.3TL. Hence, it is possible that the stress caused by the presence of *Am. muticum* introgression(s) might be one factor inducing recombination.

The new inter-genomic rearrangements were found to be made up of 80% D-genome with either the A- and/or the B-genome. The univalent state of the D-chromosomes in these lines may also have promoted the rearrangements. The A- and B-genomes have previously been shown to be more similar to the D-genome than they are to each other (Marcussen et al., 2014). Pairing is frequently observed between the A- and the D-genomes in wheat-rye hybrids denoting a much lower differentiation between these two genomes than between the A- and B- or B- and D-genomes, at least in the regions of high recombination in the distal chromosome regions (Naranjo et al., 1987; Marcussen et al., 2014). This is consistent with the high level of A-D recombinant chromosomes observed in the present study, especially in the telomeric regions of the chromosomes. For example, for the introgressions that could be identified with mc-FISH, analysis showed that the slightly larger 1DL introgression had recombined with the short arm of 1A and the 2DS introgression with the short arm of 2A.

Only the small telomeric D-genome introgressions were successfully transferred into the tetraploid background of both durum wheat varieties indicating that introgressions of a smaller size have a higher chance of being transmitted compared to larger D-genome introgressions. If the large D-genome segments do not have the ability for genetic compensation for the homologous A- or B- genome chromosome segments, it less likely they will be retained. The inter-genomic recombinant chromosomes that were present as additions were generally lost due to a lack of pairing at meiosis. For instance, the A-d translocation when present as a monosomic addition, in the tetraploid background with 29 chromosomes, was not retained after self-fertilisation. Whereas, the recombinant chromosomes that had substituted one of the wheat chromosomes had a higher rate of retention and transmission.

KASP marker analysis showed that the *Am. muticum* introgression in all the Karim D-genome introgression lines was a large 4T introgression previously confirmed as present in the WMI parental line, together with a 2T introgression, using the Axiom® Wheat-Relative Genotyping Array (King et al., 2017). The 4T introgression was highly retained in the progeny of Karim. Lines homozygous for both 4T and 1DL were identified in the BC_1-F_3 where FISH analysis showed that the pair of 4T recombinant chromosomes were substituting the pair of 4B-chromosomes. Under glasshouse conditions, these introgression lines were fertile with a normal spring wheat growth cycle and a durum wheat head type. Thus, the disomic 4T-b(4B) substitution did not affect fertility in these lines. Among the full set of Chinese Spring nullisomic–tetrasomic lines, only the 4B nullisomic tetrasomic line (N4BT4D) was completely male sterile (and had to be maintained as a monosomic tetrasomic line, M4BT4D) suggesting the presence an essential gene for male fertility on this chromosome (Sears, 1966). In addition, Endo and Gill (1996),

failed to establish a homozygous deletion line for the short arm of chromosome 4B in a hexaploidy background, because plants were male sterile. However, Langdon durum 4D(4B) disomic substitution line is also fertile and can be selfed in the absence of the 4B chromosomes (Joppa and Williams, 1988). Hence, the 4D disomic substitution compensates for the absence of both copies of chromosome 4B at the tetraploid level but does not compensate when present as tetrasomic in the 4B nullisomic tetrasomic line at the hexaploid level. This can possibly be due to the interaction of several genes. Similar to the Langdon durum 4D(4B) disomic substitution line, the 4T-b introgression fully compensate for the absence of the male fertility gene, *Ms1* (Driscoll, 1975), on chromosome 4B in durum wheat.

To our knowledge, this is the first study to transfer D-genome introgressions into either the A- or B-genomes, present in hexaploid wheat/wild relative introgression lines, into durum wheat. Advances in cytology and mc-GISH have made it possible to identify, characterize and track these genome rearrangements, together with wild relative introgressions, enabling their transfer *via* pentaploid crosses. Mc-GISH, however, is labour intensive and relatively low throughput. KASP markers, able to detect the presence of *Am. muticum* introgressions in wheat, have been developed at the WRC. Many of the KASP markers are wheat genome specific and those that are specific to the D-genome were used for the detection of the D-genome introgressions in the later generations. However, for future work, these markers will be used in the earlier generations such as the F_1 and BC_1-F_1 with the mc-GISH analysis used for validation and chromosome counting in the later generations. The developed introgression lines can be of use in durum wheat breeding through marker assisted selection, to screen for several traits of interest such as disease resistance or agronomic traits. Once multiplied, D- and T-genome introgression lines as well as the KASP markers associated with the introgressed segments will be made freely available upon request from the GRU.

DATA AVAILABILITY

The raw genotyping data supporting the conclusions of this manuscript will be made available by the authors, without undue reservation, to any qualified researcher.

AUTHOR CONTRIBUTIONS

MO conducted the crossing program. MO and CY performed the *in situ* hybridization experiments. MO, SH-E, DS, and SA performed the genotyping analysis. MO, JK, IK, and SG wrote the manuscript. MO, AY, PG, PS, IK, and JK conceived the experimental design. All authors read and approved the final manuscript.

FUNDING

This study was supported by the Monsanto Beachell Borlaug International Scholarship program (2015), The Crop Trust and the Biotechnology and Biological Sciences Research Council [grant number BB/P016855/1] as part of the Designing Future Wheat programme (DFW). The funding body played no role in the design of the study and collection, analysis and interpretation of data and in writing the manuscript.

ACKNOWLEDGMENTS

Langdon D-genome disomic substitution lines were kindly provided by Justin Faris (USDA-ARS, Fargo).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01110/full#supplementary-material

TABLE S1 | Primer sequences for all KASP markers used in this study.

REFERENCES

Bonjean, A. P., Angus, W. J., and van Ginkel, M. (2016). *The World Wheat Book: A History of Wheat Breeding* Vol. 3. Paris: Lavoisier.

Dover, G. A., and Riley, R. (1972). The prevention of pairing of homoeologous meiotic chromosomes of wheat by a genetic activity of supernumerary chromosomes of *Aegilops*. *Nature* 240, 159–161. doi: 10.1038/240159a0

Driscoll, C. J. (1975). Cytogenetic analysis of two chromosomal male-sterility mutants in hexaploidy wheat. *Aus. J. Biol. Sci.* 28, 413–416. doi: 10.1071/BI9750413

Dundas, I., Verlin, D., and Islam, R. (2015). Chromosomal locations of stem and leaf rust resistance genes from *Ae. caudata, Ae. searsii* and *Ae. mutica*. Abstract retrieved from *BGRI Workshop, September 17–20, Sydney.*

Dvorak, J. (1972). Genetic variability in *Aegilops speltoides* affecting homoeologous pairing in wheat. *Can. 5. Genet. CytoL.* 14, 371–380. doi: 10.1139/g72-046

Dvorak, J., Deal, K. R., and Luo, M. C. (2006). Discovery and mapping of wheat *Ph1* suppressors. *Genetics* 174, 17–27. doi: 10.1534/genetics.106.058115

Eberhard, F. S., Zhang, P., Lehmensiek, A., Hare, R. A., Simpfendorfer, S., and Sutherland, W. M. (2010). Chromosome composition of an F_2 *Triticum aestivum* ¾ *T. turgidum* spp. durum cross analysed by DArT markers and MCFISH. *Crop Pasture Sci.* 61, 619–624. doi: 10.1071/CP10131

Elek, H., Smart, L., Ahmad, S., Anda, A., Werner, C. P., and Pickett, J. A. (2014). A comparison of the levels of hydroxamic acids in *Aegilops speltoides* and a hexaploid wheat and effects on *Rhopalosiphum* padi behaviour and fecundity. *Acta Biol. Hung.* 65, 38–46. doi: 10.1556/ABiol.65.2014.1.4

Endo, T. R. and Gill, B. S. (1996). The deletion stock of common wheat. *Heredity*. 87, 295–307. doi: 10.1093/oxfordjournals.jhered.a023003

Eser, V. (1998). Characterisation of powdery mildew resistant lines derived from crosses between *Triticum aestivum* and *Aegilops speltoides* and *Ae. mutica*. *Euphytica* 100, 269–272. doi: 10.1023/A:1018372726968

Feldman, M., and Levy, A. A. (2005). Allopolyploidy: a shaping force in the evolution of wheat genomes. *Cytogenet. Genome Res.* 109, 250–258. doi: 10.1159/000082407

Friebe, B., Jiang, J., Raupp, W. J., McIntosh, R. A., and Gill, B. S. (1996). Characterization of wheat alien translocations conferring resistance to

diseases and pests: current status. *Euphytica* 71, 59–87. doi: 10.1007/BF00035277

Fu, S., Lv, Z., Qi, B., Guo, X., Li, J., Liu, B., et al. (2012). Molecular cytogenetic characterization of wheat-*Thinopyrum elongatum* addition, substitution and translocation lines with a novel source of resistance to wheat Fusarium head blight. *J. Genet. Genom.* 39, 103–110. doi: 10.1016/j.jgg.2011.11.008

Fu, S., Yang, M., Fei, Y., Tan, F., Ren, Z., Yan, B., et al. (2013). Alterations and abnormal mitosis of wheat chromosomes induced by wheat-rye monosomic addition lines. *PLoS One* 8, 70483. doi: 10.1371/journal.pone.0070483

Grewal, S., Hubbart-Edwards, S., Yang, C., Devi, U., Baker, L., Heath, J., et al. (2019). Rapid identification of homozygosity and site of wild relative introgressions in wheat through chromosome-specific KASP genotyping assays. *Plant Biotechnology Journal. In Press.*

Haider, N. (2013). The origin of the B-genome of bread wheat (*Triticum aestivum* L.). *Russ. J. Genet.* 49, 263–274. doi: 10.1134/S1022795413030071

Iefimenko, T. S., Fedak, , Yu. G., Antonyuk, M. Z., and Ternovska, T. K. (2015). Microsatellite analysis of chromosomes from the fifth homoeologous group in the introgressive *Triticum aestivum/Amblyopyrum muticum* wheat lines. *Cytol. Genet.* 49, 183–191. doi: 10.3103/S0095452715030056

International Wheat Genome Sequencing Consortium,, Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., et al. (2018). Shifting the limits in wheat research and breeding using a fully annotated reference sequence. *Science* 361, 7191. doi: 10.1126/science.aar7191

Jauhar, P. P., and Chibbar, R. N. (1999). Chromosome-mediated and direct gene transfers in wheat. *Genome* 42, 570–583. doi: 10.1139/g99-045

Joppa, L. R., and Williams, N. D. (1988). Langdon durum disomic substitution lines and aneuploid analysis in tetraploid wheat. *Genome.* 30, 222–228. doi: 10.1139/g88-038

Kabbaj, H., Sall, A. T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., et al. (2017). Genetic diversity within a global panel of durum wheat (*Triticum durum*) landraces and modern germplasm reveals the history of allele's exchange. *Front. Plant Sci.* 8, 1277. doi: 10.3389/fpls.2017.01277

Kalous, J. R., Martin, J. M., Sherman, J. D., Heo, H. Y., Blake, N. K., Lanning, S. P., et al. (2015). Impact of the D genome and quantitative trait loci on quantitative traits in a spring durum by spring bread wheat cross. *Theor. Appl. Genet.* 128, 1799–1811. doi: 10.1007/s00122-015-2548-3

Kato, A., Lamb, J. C., and Birchler, J. A. (2004). Chromosome painting using repetitive DNA sequences as probes for somatic chromosome identification in maize. *Proc. Natl Acad. Sci. U.S.A.* 101, 13554–13559. doi: 10.1073/pnas.0403659101

Kihara, H. (1924). Cytologische und genetische Studien bei wichtigen Getreidearten mit besondere Rüchsicht auf das Verhalten der Chromosomen und die Sterilität in den Bastarden. *Mem. Coll. Sci. Kyoto Imp. Univ.* 1, 1–200.

Kihara, H. (1944). Discovery of the DD-analyser, one of the ancestors of *Triticum* vulgare (abstr) (in Japanese). *Agric. Hortic.* 19, 889–890. doi: 10.1016/j.jgg.2011.07.002

Kilian, B., Mammen, K., Millet, E., Sharma, R., Graner, A., Salamini, F. et al., (2013). "*Aegilops*," in *Wild crop relatives: genomic and breeding resources*. Ed. Ch Kole (Berlin, Heidelberg: Cereals. Springer), 1–76. doi: 10.1007/978-3-642-14228-4

King, I. P., Forster, B. P., Law, C. C., Cant, K. A., Orford, S. E., Gorham, J., et al. (1997). Introgression of salt-tolerance genes from *Thinopyrum bessarabicum* into wheat. *New Phytol.* 137, 75–81. doi: 10.1046/j.1469-8137.1997.00828.x

King, J., Armstead, I., Harper, J., Ramsey, L., Snape, J., Waugh, R., et al. (2013). Exploitation of interspecific diversity for monocot crop improvement. *Heredity* 110, 475–483. doi: 10.1038/hdy.2012.116

King, J., Grewal, S., Yang, C., Hubbart, S., Scholefield, D., Ashling, S., et al. (2017). A step change in the transfer of interspecific variation into wheat from *Amblyopyrum muticum*.. *Plant Biotech. J.* 15, 217–226. doi: 10.1111/pbi.12606

Li, H., Deal, K. R., Luo, M.-C., Ji, W., Distelfeld, A., and Dvorak, J. (2017). Introgression of the *Aegilops speltoides Su1-Ph1* Suppressor into Wheat. *Front. Plant Sci.* 8, 2163. doi: 10.3389/fpls.2017.02163

Li, Y. F., Wu, Y., Hernandez-Espinosa, N., and Peña, R. J. (2013). Heat and drought stress on durum wheat: responses of genotypes, yield, and quality parameters. *J. Cereal Sci.* 57, 398–404. doi: 10.1016/j.jcs.2013.01.005

Luchniak, P., Maluszynska, J., and Olszewska, M. J. (2002). *In situ* fluorescent nick translation procedure for plant chromosomes. *Biotech. Histochemistry* 77, 15–19. doi: 10.1080/bih.77.1.15.19

Magallanes-López, A. M., Ammar, K., Morales-Dorantes, A., González-Santoyo, H., Crossa, J. and Guzman, C. (2017). Grain quality traits of commercial durum wheat varieties and their relationships with drought stress and glutenins composition. *J. Cereal Sci.* 75, 1–9. doi: 10.1016/j.jcs.2017.03.005

Marcussen, T., Sandve, S. R., Heier, L., Spannagl, M., Pfeifer, M., The International Wheat Genome Sequencing Consortium, et al. (2014). Ancient hybridizations among the ancestral genomes of bread wheat. *Science* 345, 6194. doi: 10.1126/science.1250092

Martin, A., Simpfendorfer, S., Hare, R., and Sutherland, M. (2013). Introgression of hexaploid sources of crown rot resistance into durum wheat. *Euphytica* 192, 463–470. doi: 10.1007/s10681-013-0890-6

Matsuoka, Y., and Nasuda, S. (2004). Durum wheat as a candidate for the unknown female progenitor of bread wheat: An empirical study with a highly fertile F1 hybrid with *Aegilops tauschii* Coss. *Theor. Appl. Genet.* 109, 1710–1717. doi: 10.1007/s00122-004-1806-6

McFadden, E. S., and Sears, E. R. (1944). The artificial synthesis of *Triticum spelta*. *Rec. Genet. Soc. Am.* 13, 26–27. doi: 10.4236/as.2014.514154

McIntyre, C. L., Pereira, S., Moran, L. B., and Apples, R. (1990). New *Secale cereale* (rye) DNA derivatives for the detection of rye chromosome segments in wheat. *Genome* 33, 317–323. doi: 10.1139/g90-094

Naranjo, T., Roca, A., Goicoechea, P. G., and Giraldez, R. (1987). Arm homoeology of wheat and rye chromosomes. *Genome* 29, 873–882. doi: 10.1139/g87-149

Nazco, R., Villegas, D., Ammar, K., Peña, R. J., Moragues, M., and Royo, C. (2012). Can Mediterranean durum wheat landraces contribute to improved grain quality attributes in modern cultivars? *Euphytica* 185, 1–17. doi: 10.1007/s10681-011-0588-6

Padmanaban, S., Sutherland, M. W., Knight, N. L., and Martin, A. (2017b). Genome inheritance in populations derived from hexaploid/tetraploid and tetraploid/hexaploid wheat crosses. *Mol. Breed.* 37, 48. doi: 10.1007/s11032-017-0647-3

Padmanaban, S., Zhang, P., Hare, R. A., Sutherland, M. W., and Martin, A. (2017a). Pentaploid Wheat Hybrids: Applications, Characterisation, and Challenges. *Front. Plant Sci.* 8, 358. doi: 10.3389/fpls.2017.00358

Qi, L. L., Friebe, B., Zhang, P., and Gill, B. S. (2007). Homoeologous recombination, chromosome engineering and crop improvement. *Chromosome Res.* 15, 3–19. doi: 10.1007/s10577-006-1108-8

Ramsey, J., and Schemske, D. W. (1998). Pathways, mechanisms, and rates of polyploid formation in flowering plants. *Annu. Rev. Ecol. Syst.* 29, 467–501. doi: 10.1146/annurev.ecolsys.29.1.467

Rayburn, A. L., and Gill, B. S. (1986). Isolation of a D-genome specific repeated DNA sequence from *Aegilops squarrosa*. *Plant. Mol. Biol. Rep.* 4, 102–109. doi: 10.1007/BF02732107

Schneider, A., Molnár, I., and Molnár-Láng, M. (2008). Utilisation of *Aegilops* (Goatgrass) species to widen the genetic diversity of cultivated wheat. *Euphytica* 163, 1–19. doi: 10.1007/s10681-007-9624-y

Schubert, I., Fransz, P. F., Fuchs, J., and De Jong, J. H. (2001). Chromosome painting in plants. In Chromosome Painting. *Methods Cell Sci.* 23, 57. doi: 10.1023/A:1013137415093

Schwarzacher, T., Anamthawat-Jónsson, K., Harrison, G. E., Islam, A. K. M. R., Jia, J. Z., King, I. P., et al. (1992). Genomic *in situ* hybridization to identify alien chromosomes and chromosome segments in wheat. *Theor. Appl. Genet.* 84, 778–786. doi: 10.1007/BF00227384

Schwarzacher, T., Leitch, A., Bennett, M., and Heslop-Harrison, J. (1989). *In situ* localization of parental genomes in a wide hybrid. *Ann. Bot.* 64, 315–324. doi: 10.1093/oxfordjournals.aob.a087847

Sears, E. (1966). Chromosome mapping with the aid of telocentrics. *Hereditas* 2, 370-381

Shewry, P. R., and Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food Energy Secur.* 4, 178–202. doi: 10.1002/fes3.64

Silva, G., and Souza, M. (2013). Genomic *in situ* hybridization in plants. *Genet. Mol. Res.* 12, 2953–2965. doi: 10.4238/2013.August.12.11

Tang, Z. X., Yang, Z. J., and Fu, S. L. (2014). Oligonucleotides replacing the roles of repetitive sequences *pAs1, pSc119.2, pTa-535, pTa71, CCS1*, and *pAWRC*.1 for FISH analysis. *J. Appl. Genet.* 55, 313–318. doi: 10.1007/s13353-014-0215-z

Thomson, D., and Henry, R. (1995). Single-step protocol for preparation of plant tissue for analysis by PCR. *BioTechniques* 19, 394–397, 400.

Zhang, H., Bian, Y., Gou, X., Zhu, B., Xu, C., Qi, B., et al. (2013). Persistent whole-chromosome aneuploidy is generally associated with nascent allohexaploid wheat. *Proc. Natl Acad. Sci. U.S.A.* 110, 3447–3452. doi: 10.1073/pnas.1300153110

Zohary, D., Harlan, J. R., and Vardi, A. (1969). The wild diploid progenitors of wheat and their breeding value. *Euphytica* 18, 58–65. doi: 10.1007/BF00021982

Conflict of Interest Statement: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 October 2019
doi: 10.3389/fpls.2019.01247

Mapping of Genetic Loci Conferring Resistance to Leaf Rust From Three Globally Resistant Durum Wheat Sources

*Dhouha Kthiri[1], Alexander Loladze[2], Amidou N'Diaye[1], Kirby T. Nilsen[1], Sean Walkowiak[1], Susanne Dreisigacker[2], Karim Ammar[2] and Curtis J. Pozniak[1]**

[1] Department of Plant Sciences, Crop Development Centre, University of Saskatchewan, Saskatoon, SK, Canada,
[2] International Maize and Wheat Improvement Center (CIMMYT), Mexico City, Mexico

Genetic resistance in the host plant is the most economical and environmentally friendly strategy for controlling wheat leaf rust, caused by *Puccinia triticina* Eriks. The durum wheat lines Gaza (Middle East), Arnacoris (France) and Saragolla (Italy) express high levels of resistance to the Mexican races of *P. triticina*. Three recombinant inbred line (RIL) populations, derived from crosses of each of these resistance sources to the susceptible line ATRED #2, were evaluated for leaf rust reactions at CIMMYT's leaf rust nurseries in Mexico. Genetic analyses of host reactions suggested oligogenic control of resistance in all populations. The F_8 RILs from each cross were genotyped using the Illumina iSelect 90K array, and high-density genetic maps were constructed for each population. Using composite interval mapping, a total of seven quantitative trait loci (QTL) that provide resistance to leaf rust were identified. Two QTL designated as *QLr.usw-6BS* and *QLr.usw-6BL* were identified on chromosome 6B in Gaza, which explained up to 78.5% and 21.3% of the observed leaf rust severity variance, respectively. A major QTL designated as *QLr.usw-7BL* was detected on the long arm of chromosome 7B in Arnacoris, which accounted for up to 65.9% of the disease severity variance. Arnacoris also carried a minor QTL on chromosome 1BL, designated as *QLr.usw-1BL.1* that explained up to 17.7% of the phenotypic variance. Three QTL conferred leaf rust resistance in Saragolla, namely *QLr.usw-2BS*, *QLr.usw-3B,* and *QLr.usw-1BL.2*, which accounted for up to 42.3, 9.4, and 7.1% of the phenotypic variance, respectively. Markers flanking each QTL were physically mapped against the durum wheat reference sequence and candidate genes involved in disease resistance were identified within the QTL intervals. The QTL identified in this study and their closely linked markers are useful resources for gene pyramiding and breeding for durable leaf rust resistance in durum wheat.

Keywords: durum wheat, leaf rust, *Puccinia triticina*, resistance, quantitative trait loci, single nucleotide polymorphism (SNP)

OPEN ACCESS

Edited by:
Jacqueline Batley, University of Western Australia, Australia

Reviewed by:
James Kolmer, Agricultural Research Service, United States Department of Agriculture, United States
Filippo Maria Bassi, International Center for Agricultural Research in the Dry Areas (ICARDA), Morocco

****Correspondence:***
Curtis J. Pozniak
curtis.pozniak@usask.ca

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *14 May 2019*
Accepted: *06 September 2019*
Published: *08 October 2019*

Citation:
Kthiri D, Loladze A, N'Diaye A, Nilsen KT, Walkowiak S, Dreisigacker S, Ammar K and Pozniak CJ (2019) Mapping of Genetic Loci Conferring Resistance to Leaf Rust From Three Globally Resistant Durum Wheat Sources. Front. Plant Sci. 10:1247. doi: 10.3389/fpls.2019.01247

INTRODUCTION

The importance of leaf rust as a threat to global durum wheat (*Triticum turgidum* L. ssp *durum*) production has increased dramatically during the last decade, due to the occurrence of highly virulent races of *Puccinia triticina* and the breakdown of the resistance genes that were widely deployed (Singh et al., 2004; Huerta-Espino et al., 2011; Herrera-Foessel et al., 2014a; Kolmer, 2015b; Soleiman et al., 2016; Kolmer and Hughes, 2017). Genetic control offers

a cost-effective and environmentally friendly alternative to avoid yield losses due to this pathogen. Resistance to wheat leaf rust is commonly categorized into two classes based on their genetic control and phenotypic effect: race-specific, all-stage resistance, which is usually expressed as a hypersensitive response leading to host cell death, and adult plant resistance (APR), which is usually expressed as a slow-rusting phenotype (Knott, 1989; Lagudah, 2011; Kolmer, 2013; Singh et al., 2016). To date, over 77 leaf rust resistance (*Lr*) genes have been characterized and catalogued in wheat (McIntosh et al., 2017). Most *Lr* genes have major effects and confer race-specific, all-stage near-immunity. However, this class of resistance is prone to rapid breakdown as the pathogen population evolves, and new virulent races emerge (Suenaga et al., 2003; Huerta-Espino et al., 2011; Lowe et al., 2011; Ellis et al., 2014; Herrera-Foessel et al., 2014b). The over-reliance on a single race-specific resistance gene has led to leaf rust epidemics and considerable losses in the state of Sonora in Mexico, when the new fungal race BBG/BN overcame *Lr72* in 2001 (Singh et al., 2004; Herrera-Foessel et al., 2014a). Subsequently, loss of the resistance conferred by the complementary genes *Lr27+Lr31* occurred in 2008, with the emergence of race BBG/BP (Singh et al., 2004; Huerta-Espino et al., 2011). In contrast to race-specific resistance genes, APR genes express at the adult stage and only provide partial resistance which results in longer latent periods and smaller and fewer fungal spores or uredinia (Knott, 1989; Herrera-Foessel et al., 2008c; Lagudah, 2011; Lowe et al., 2011). Currently, at least eight genes that confer APR to leaf rust have been characterized in hexaploid wheat, including *Lr34* (Suenaga et al., 2003), *Lr46* (Singh et al., 1998), *Lr67* (Hiebert et al., 2010), *Lr68* (Herrera-Foessel et al., 2012), *Lr74* (McIntosh et al., 2015), *Lr75* (Singla et al., 2017), *Lr77* (Kolmer et al., 2018c), and *Lr78* (Kolmer et al., 2018a). Only *Lr46* has been reported in durum wheat (*T. turgidum* L. ssp. *durum*) (Herrera-Foessel et al., 2011).

Mapping quantitative trait loci (QTL) enables the detection of genes with both major and minor effects and the identification of linked molecular markers that could be used for gene stacking and breeding for durable rust resistance (Soriano and Royo, 2015). The advent of next generation sequencing technologies and high-throughput SNP genotyping platforms facilitated the development of high density genetic maps in wheat (Wang et al., 2014; Maccaferri et al., 2015; Winfield et al., 2016), enhancing our ability to dissect economically important traits such as disease resistance. Several studies have used QTL mapping to identify and tag genomic regions involved in leaf rust resistance in hexaploid wheat (Schnurbusch et al., 2004; Rosewarne et al., 2008; Rosewarne et al., 2012; Buerstmayr et al., 2014; Kolmer, 2015a; Soriano and Royo, 2015), however, only a few studies have been conducted to map QTL for leaf rust resistance in durum wheat (Marone et al., 2009; Singh et al., 2013a; Lan et al., 2017). The objective of this study was to use the iSelect 90K SNP array to characterize and map genetic loci conferring leaf rust resistance in three globally resistant durum wheat genotypes and to identify linked SNP markers useful for gene pyramiding and marker-assisted breeding.

MATERIALS AND METHODS

Plant Materials and Field Phenotyping

Three sources of resistance to leaf rust (*P. triticina*), including the Middle Eastern landrace Gaza (unknown pedigree, CIMMYT genotype ID 233), the French cultivar Arnacoris (unpublished pedigree, CIMMYT genotype ID 6048080), and the Italian cultivar Saragolla (pedigree: Iride/Linea PSB 0114, CIMMYT genotype ID 255301), were identified by CIMMYT's durum wheat breeding program, through extensive multi-race, multisite testing (data not shown). Each source was crossed to the susceptible line ATRED #2 (pedigree: Atil*2/LocalRed, CIMMYT genotype ID 5460557), and RIL populations of over 200 RILs each were developed by advancing generations through single-seed-descent, at the two locations used by CIMMYT's breeding program: CENEB experimental station near Ciudad Obregon in Sonora (latitude 27.33, longitude -109.93, altitude 35 meters above sea level (masl)), with a wheat crop season from mid-November to late April, and El Batán experimental station, northeast of Mexico City (latitude 19.53, longitude -98.84, altitude 2250 masl), with a wheat crop season from mid-May to mid-October. Reactions to the widely virulent race of *P. triticina* BBG/BP (Huerta-Espino et al., 2011; Loladze et al., 2014) were scored on the F_2 progenies during the spring of 2011 (CENEB). During summer 2011, the F_2-derived F_3 families ($F_{2:3}$) were space planted in double 1.2-meter-long rows at El Batán. In 2013, F_6 families were grown in replicated 1.2 m rows at the CENEB station, while paired 1.2 m rows of the F_8 RILs were phenotyped at El Batán, in summer 2014. In all experiments, parental lines and progenies were inoculated at the tillering stage of plant development, using a mineral oil suspension of urediniospores of race BBG/BP of *P. triticina*, at a concentration of 5 to 10 mg of urediniospores per 5 ml of oil (Soltrol 170). Susceptible spreader rows surrounding plots and consisting of a mixture of the cultivars Banamichi C2004 and Jupare C2001 (susceptible only to race BBG/BP in Mexico) were also inoculated. The race BBG/BP of *P. triticina* is the predominant durum-specific race in Mexico, with the following avirulence/virulence formula: *Lr1*, *2a*, *2b*, *2c*, *3*, *3ka*, *3bg*, *9*, *13*, *14a*, *15*, *16*, *17*, *18*, *19*, *21*, *22a*, *24*, *25*, *26*, *28*, *29*, *30*, *32*, *35*, *37/Lr10*, *11*, *12*, *14b*, *20*, *23*, *27 + 31*, *33*, *72* (Huerta-Espino et al., 2008; Huerta-Espino et al., 2011; Loladze et al., 2014).

For the $F_{2:3}$ families and the F_6 RIL populations from the crosses Gaza/ATRED #2, Arnacoris/ATRED #2, and Saragolla/ATRED #2, as well as the F_8 RILs from the Gaza/ATRED #2 cross, the modified Cobb scale was used to visually estimate the percentage of pustule-infected leaf area or leaf rust severity (LRS) on the parental lines and their progenies (Peterson et al., 1948). Three LRS scores were recorded at weekly intervals, starting at 14 days post inoculation, and the area under the disease progress curve (AUDPC) was calculated as before (Maccaferri et al., 2010). Host reactions were also recorded using four categories: "R" to indicate resistance or uredinia traces, "MR" to indicate moderate resistance with small uredinia surrounded by necrosis, "MS" to indicate moderate susceptibility expressed as chlorosis surrounding moderate sized uredinia, and "S" to indicate full susceptibility with large uredinia lacking necrosis

or chlorosis (Roelfs et al., 1992). Based on the host reactions of plants within each family, the $F_{2:3}$ families were categorized as homozygous resistant (R), homozygous susceptible (S), and heterozygous (Het). The F_8 RILs from the crosses Arnacoris/ATRED #2 and Saragolla/ATRED #2 were scored as R or S. The chi-square (χ^2) test was used to estimate the number of genes involved in the inheritance of leaf rust resistance in these populations. For analyses with a single degree of freedom, the chi-square values were adjusted with the Yates's correction for continuity (Yates, 1934).

Seedling Stage Evaluations of the $F_{2:3}$ Families

Seedlings of $F_{2:3}$ families from the crosses Gaza/ATRED #2, Arnacoris/ATRED #2, and Saragolla/ATRED #2 were evaluated for resistance to race BBG/BP of *P. triticina*, under controlled conditions in CIMMYT's greenhouse, at El Batán experimental station. Approximately 25 to 35 seedlings from each $F_{2:3}$ family were grown in 7-by-7-by-10 cm pots and inoculated with a suspension of urediniospores of race BBG/BP in light mineral oil (Loladze et al., 2014). Infection types (IT) were recorded using the 0–4 scale where "0" = no visible leaf rust symptoms; ";" = hypersensitive flecks without any uredinia; "1" = small uredinia surrounded by necrosis; "2" = small to medium uredinia surrounded by chlorosis or necrosis; "3" = medium-sized uredinia with or without chlorosis; "4" = large uredinia without chlorosis or necrosis; "X" = random distribution of variable-sized uredinia (mesothetic reaction), and "+" and "-" were used when uredinia were somewhat larger or smaller than the average for the IT class (McIntosh et al., 1995). ITs of 3, 3+, and 4 indicate susceptible host reactions, whereas all of the other ITs were considered resistant. Based on their ITs at 10 to 12 days after inoculation, the $F_{2:3}$ families were classified as homozygous resistant "R," homozygous susceptible "S," or heterozygous "Het." The χ^2 test was used to estimate the number of genes involved in the inheritance of leaf rust resistance at the seedling stage.

Allelism Tests Between Gaza and Carriers of Known *Lr* Genes

Allelism between the resistance genes in Gaza and the known *Lr* genes *Lr61* and Lr_{Camayo} was investigated using 177 F_2 plants from the cross Gaza/Sooty_9/Rascon_37//Guayacan INIA, and 273 F_2 plants from the cross Gaza/Cirno C2008, respectively. Allelism to the *Lr* genes in the durum lines Geromtel_3 and Tunsyr_2 was also studied in populations of 326 and 181 F_2 plants, respectively. The F_2 progenies were tested for their field reaction to race BBG/BP of *P. triticina*, at El Batán. The resulting resistant/susceptible ratios were tested for goodness of fit to various gene models, using chi-square analysis. When no susceptible recombinants were detected in the F_2 progenies, it was assumed that the two resistant parents carried allelic or closely linked leaf rust resistance genes.

DNA Extraction and Illumina iSekect 90K SNP Array Genotyping

Genomic DNA was extracted from the parental lines and the F_8 RILs from the three crosses Gaza/ATRED #2, Arnacoris/ATRED #2, and Saragolla/ATRED #2, using a BIOMEK FXp liquid handling station and the Sbeadex mini plant kit (LGC, Teddington, Middlesex, UK) (Dreisigacker et al., 2016). The Illumina iSelect 90K SNP array (Illumina, San Diego, CA, USA) was used for genotyping of the RILs and the parental lines (Wang et al., 2014). Genotype calling was performed using GenomeStudio software (Illumina, San Diego, CA, USA). The genotyping data for all three biparental populations are provided as **Supplementary Materials**. Prior to mapping, data filtering was carried out and markers that showed significant segregation distortion or more than 10% missing values were excluded.

Construction of Linkage Maps and QTL Mapping

Curated SNP data were used to build linkage maps for each of the three populations. Initial linkage groups (LG) were obtained using the MSTMap software (Wu et al., 2008) with a stringent cutoff *p*-value of $1E^{-10}$ and a maximum distance between markers of 15.0 centimorgan (cM). The LGs were assigned to individual wheat chromosomes based on existing high-density SNP maps (Cavanagh et al., 2013; Wang et al., 2014; Maccaferri et al., 2015) and the tetraploid wheat reference sequences of wild emmer wheat (WEW) (Avni et al., 2017) and the modern durum wheat cultivar Svevo (Maccaferri et al., 2019). Once LGs were attributed to chromosomes, their genotypic data were pooled on a chromosome-by-chromosome basis and final genetic maps were constructed using MapDisto 1.7.7 software (Lorieux, 2012) at threshold values of recombination frequency = 0.3 and logarithm of the odds (LOD) = 3. Markers sharing the same segregation pattern (co-segregating) were identified in each LG, and the marker with the lowest percentage of missing data was chosen to represent each cluster or bin. Double recombinants were corrected using the functions "Show double recombinants," "Show error candidates," and "Replace error candidates by flanking genotype" as implemented in the MapDisto software (Lorieux, 2012). The order of the markers was determined using the Order, Ripple, and Check inversions commands. The Kosambi function was used to convert the recombination fractions to cM (Kosambi, 1943).

QTL detection was performed using the composite interval mapping (CIM) method implemented in QGene 4.4.0 software (Joehanes and Nelson, 2008). QTL were identified at a scan interval of 1 cM. The stepwise regression method was used to select cofactors and the LOD thresholds for determining statistically significant QTL were calculated by 1000 permutations with $P < 0.05$. The additive effect and percentage of phenotypic variance explained by each QTL were obtained from the final CIM results. Phenotypic traits analyzed included LRS, AUDPC and the host reaction (HR) recorded for the F_8 RILs from the crosses Arnacoris/ATRED #2 and Saragolla/ATRED #2.

QTL Interaction Tests in Each Population

A single marker with the highest LOD score was selected from each QTL to estimate its phenotypic effect. The mean phenotypic data for all the RILs from each population were grouped based on their genotype combinations at these selected loci

and mean separation was performed using Fisher's protected least significant difference (LSD) test in SAS 9.4 software (SAS Institute Inc., Cary, NC, USA).

Genotyping of Molecular Markers Linked to Known *Lr* Genes

To test for presence of the APR gene *Lr46* (http://maswheat.ucdavis.edu/protocols/Lr46), the two simple sequence repeat (SSR) markers *Xwmc44* and *Xgwm259* (William et al., 2006) and a kompetitive allele specific polymerase chain reaction (KASP) marker were used to genotype the parental lines Arnacoris, Saragolla, and ATRED #2, as well as subsets of resistant and susceptible RILs from each population. The durum cultivars Kofa, CDC Verona, and Strongfield were used as checks. Three SSR markers (i.e. *Xwmc764*, *Xgwm210,* and *Xwmc661*) and two KASP markers (i.e. *kwm677* and *kwm744*) linked to the resistance gene *Lr16* (McCartney et al., 2005; Kassa et al., 2017) were used to genotype both parents and selected lines from the Saragolla/ATRED #2 population, as well as the check cultivar AC Domain. The parental lines Gaza and ATRED #2 and check cultivars including the *Lr61*-carrying Guyacan INIA were genotyped using the SSR marker *Xwmc487* linked to *Lr61* (Herrera-Foessel et al., 2008a). The four parental lines Gaza, Arnacoris, Saragolla, and ATRED #2 were also screened using the nucleotide-binding site leucine-rich repeat (NBS-LRR)-specific primers *4406F/4840R*, previously determined to be linked to *Lr14a* (Kthiri et al., 2018). The durum cultivars Sachem and Somayoa were used as positive controls that carry *Lr14a*. The amplification reactions were performed according to published protocols (McCartney et al., 2005; William et al., 2006; Herrera-Foessel et al., 2008a; Kthiri et al., 2018). Polymorphisms for the SSR markers and the NBS-LRR-specific primers were scored on 2% agarose gels.

Physical Mapping to the Durum Wheat Reference Genome

The program GMAP (Wu and Watanabe, 2005) was used to align the sequences of the SNP markers that localized within each QTL LOD plot area to the genome sequence of the durum wheat cultivar "Svevo" (Maccaferri et al., 2019). Putative physical intervals for each QTL were identified using a cut-off value of 98% for sequence identity and sequence coverage. Genes that mapped within these physical intervals were identified using the available annotations for the durum wheat reference genome (Maccaferri et al., 2019).

RESULTS

The Inheritance of Leaf Rust Resistance in Gaza, Arnacoris, and Saragolla

Leaf rust symptoms developed adequately for all the field evaluation trials at the CENEB and El Batán experimental stations. The RILs from the three mapping populations expressed a wide range in disease severity, with the resistant parents Gaza, Arnacoris, and Saragolla showing the lowest scores for leaf rust severity (0–5%), and the highest scores (90–100%) being observed on the susceptible parent ATRED #2. No transgressive segregation was observed among the RILs from the three crosses, which confirmed that the susceptible parent ATRED #2 does not contribute any genes for leaf rust resistance that could be detected under the present phenotyping conditions.

The $F_{2:3}$ seedlings from the cross Gaza/ATRED #2 segregated at 63R:94Het:50S, which is a good fit to the 1:2:1 ratio expected for a single dominant seedling resistance gene, based on *p*-value ($P_{(< 0.05)}$ = 0.18) of the χ^2 test at a 95% level of confidence (**Table 1**). However, field-based segregation ratio of 7R:8Het:1S ($P_{(< 0.05)}$ = 0.12) of the F_3 adult plants suggested the presence of two resistance genes in Gaza (**Table 1**). This discrepancy between seedlings and adult plants evaluation results led to the conclusion that Gaza could carry one APR and one seedling resistance gene. The involvement of an APR gene was further investigated by selecting 10 $F_{2:3}$ families that were uniformly resistant in the field at the adult stage and uniformly susceptible at the seedling stage and testing them again at the adult plant stage, under greenhouse conditions. The results from these tests confirmed that the 10 selected families were indeed resistant

TABLE 1 | Segregation ratios of the $F_{2:3}$ and F_8 progenies from three crosses of durum wheat lines evaluated for leaf rust resistance to race BBG/BP of *P. triticina*, at the seedling and adult plant stages.

Cross	Seedlings					Adult plants								
	$F_{2:3}$ Families					$F_{2:3}$ Families					F_8 RILs			
	R	Het	S	Ratio	P	R	Het	S	Ratio	P	R	S	Ratio	P^a
Gaza/ATRED #2	63	94	50	1:2:1	0.18	91	135	18	7:8:1	0.12	123	115	1:1	0.65
Arnacoris/ATRED #2	10	126	88	1:8:7	0.14	11	133	91	1:8:7	0.12	120	99	1:1	0.18
Saragolla/ATRED #2	59	102	46	1:2:1	0.43	9	95	104	1:8:7	0.15	82	122	1:1	0.01*
Saragolla/ATRED #2											82	122	3:1	0*

*$F_{2:3}$ families were categorized as resistant (R), heterozygous (Het), or susceptible (S). Mendelian ratios and their corresponding p-value (P) from chi-square (χ^2) analysis are shown. The null hypothesis for the χ^2 test was rejected at P < 0.05. * P < 0.05 indicating that the observed segregation ratio is significantly different from the expected segregation ratio at a 95% level of confidence. [a] The chi-square values were adjusted with the Yates's correction for continuity.*

at the adult stage. The F_8 RILs from the cross Gaza/ATRED #2 showed a segregation of 1R:1S consistent with the segregation pattern of a single resistance gene ($P_{(< 0.05)} = 0.65$) (**Table 1**). This may be explained by the fact that APR genes are less effective in El Batán compared to the CENEB station (K. Ammar and A. Loladze, *unpublished*).

Similar results were observed for the Saragolla/ATRED #2 population since segregation of the $F_{2:3}$ seedlings fit a 1R:2Het:1S ratio ($P_{(< 0.05)} = 0.43$) expected for a single resistance gene, while the $F_{2:3}$ adult plants segregation ratio fit a 1R:8Het:7S ($P_{(< 0.05)} = 0.15$) ratio, expected for the segregation of two resistance genes (**Table 1**). The screening of the F_8 RILs, however, resulted in a distorted segregation of 82R:122S, which did not fit the expected ratios for Mendelian inheritance of one ($P_{(< 0.05)} = 0.01$) or two ($P_{(< 0.05)} = 0$) resistance genes (**Table 1**).

The 1R:8Het:7S segregation ratio observed for the $F_{2:3}$ progenies from the cross Arnacoris/ATRED #2 at both the seedling ($P_{(< 0.05)} = 0.14$) and adult plant ($P_{(< 0.05)} = 0.12$) stages suggested the presence of two resistance genes in Arnacoris. However, the segregation ratio of 1R:1S ($P_{(< 0.05)} = 0.18$) observed in the F_8 RILs was more consistent with that expected for a single resistance gene (**Table 1**).

The frequency distributions of the LRS and the AUDPC were determined for the three RIL populations (**Figure 1**). Although both disease severity and AUDPC data of the three populations showed continuous distributions, there was an obvious tendency of skewness towards resistance, which suggested that, while leaf rust resistance in Gaza, Arnacoris, and Saragolla was not monogenically inherited, there may be major gene effects in these populations.

QTL Analyses Detect Several Leaf Rust Resistance Loci in Gaza, Arnacoris, and Saragolla

Polymorphic SNP markers with call frequencies ≥ 90% that fit the expected segregation ratio of 1:1 were considered reliable for mapping and were subsequently used to construct linkage maps for each of the three F_8 RIL populations from the following crosses: Gaza/ATRED #2 (6248 SNP), Arnacoris/ATRED #2 (7315 SNP), and Saragolla/ATRED #2 (5345 SNP) (**Table 2**). The number of LGs ranged from 29 to 35, with all 14 durum chromosomes represented by at least one LG (**Table 2**). Comparison across genomes indicated that the maximum number of markers mapped to the B genome, except for Gaza where 51.2% of the SNPs mapped to the A genome. The high proportion of co-segregating SNP markers significantly reduced the final number of bins (unique marker loci) in the three maps. Marker density was greatest for the Arnacoris/ATRED #2 population, with an average inter-bin interval of 2.4 cM. However, the Gaza/ATRED #2 population produced the longest map with a total length of 4476.2 cM and the lowest marker density. The Saragolla/ATRED #2 population produced the shortest map with a total length of 3647.2 cM (**Table 2**). Final genetic maps for the three RIL populations are available as supplementary material (**Supplementary File S1**).

QTL in the Gaza/ATRED #2 Population

Two QTL on chromosome 6B were associated with leaf rust resistance in the Gaza/ATRED #2 population. Both were derived from the resistant parent Gaza. The QTL were detected at a LOD significance threshold of 3.5, based on 1,000 permutation tests at a type I error rate of $\alpha < 0.05$ (**Table 3**). The first QTL, *QLr.usw-6BS*, peaked at the locus *CAP7_c10772_156* and was detected in both F_6 and F_8 RILs, which were evaluated for leaf rust at CENEB in 2013 and at El Batán in 2014, respectively. This QTL was detected for both traits analyzed (LRS and AUDPC) at both locations, and explained up to 34.5% of the total phenotypic variance for AUDPC at CENEB (LOD 21.1) and up to 78.5% at El Batán (LOD 79.5) (**Table 3**). The second QTL detected in the Gaza/ATRED #2 progeny was designated as *QLr.usw-6BL* and peaked at the SNP marker *GENE-3689_293*. *QLr.usw-6BL* was also detected at both locations for all the traits analyzed, and accounted for 20.5% (LOD 11.4) and 18.7% (LOD 10.7) of the final LRS variance at CENEB and at El Batán, respectively (**Table 3**).

The RILs from the cross Gaza/ATRED #2 were grouped into four categories based on the allelic state of SNP markers *CAP7_c10772_156* and *GENE-3689_293* and mean separation was performed using Fisher's LSD test. As expected, RILs that carried no resistance alleles scored the highest LRS and AUDPC of the allelic combinations (**Figures 2A**, **B**). *QLr.usw-6BS* had the strongest effect, since the mean LRS and AUDPC expressed by carriers of this QTL did not exceed 6% and 50.2, respectively (**Figures 2A**, **B**). *QLr.usw-6BL* was also able to reduce the disease symptoms, and though the reduction was not as strong as with *QLr.usw-6BS*, it was statistically significant. In general, the presence of *QLr.usw-6BS* reduced leaf rust symptoms to its lowest significant level, thereby masking the potential expression of *QLr.usw-6BL*.

The F_2 progenies from the crosses Gaza/Sooty_9/Rascon_37// Guayacan INIA and Gaza/Cirno C2008 segregated at 55R:9S ($P_{(< 0.05)} = 0.93$) and 61R:3S ($P_{(< 0.05)} = 0.44$), respectively (**Supplementary Table S1**), indicating that the leaf rust resistance genes in Gaza are neither allelic nor linked to *Lr61* or Lr_{Camayo}. The absence of susceptible recombinants in the F_2 progeny from the cross Gaza/Geromtel_3 suggests that these two cultivars carry either allelic or closely linked *Lr* genes. However, the 61R:3S ($P_{(< 0.05)} = 0.16$) segregation ratio observed in the F_2 population from the cross Gaza/Tunsyr_2 indicates that the *Lr* genes in these two cultivars are unrelated (**Supplementary Table S1**). The SSR marker *Xwmc487* linked to *Lr61* showed a clear difference in PCR product size between Gaza and the *Lr61*-carrying Guayacan INIA (**Supplementary Figure S1**).

QTL in the Arnacoris/ATRED #2 Population

Composite interval mapping revealed two QTL associated with resistance to leaf rust in the Arnacoris/ATRED #2 population (**Table 4**). A major QTL on the distal region of chromosome 7BL, designated as *QLr.usw-7BL*, with a peak LOD value at marker *BS00010355_51*, explained 65.9% of LRS variance (LOD 49.3) for the 2013 field trials at CENEB, and was highly significant for the AUDPC (LOD 46.1). *QLr.usw-7BL* was the only QTL detected for the host reactions of the F_8 RILs evaluated at El Batán, and accounted for 99.6% of the phenotypic variance. A less significant

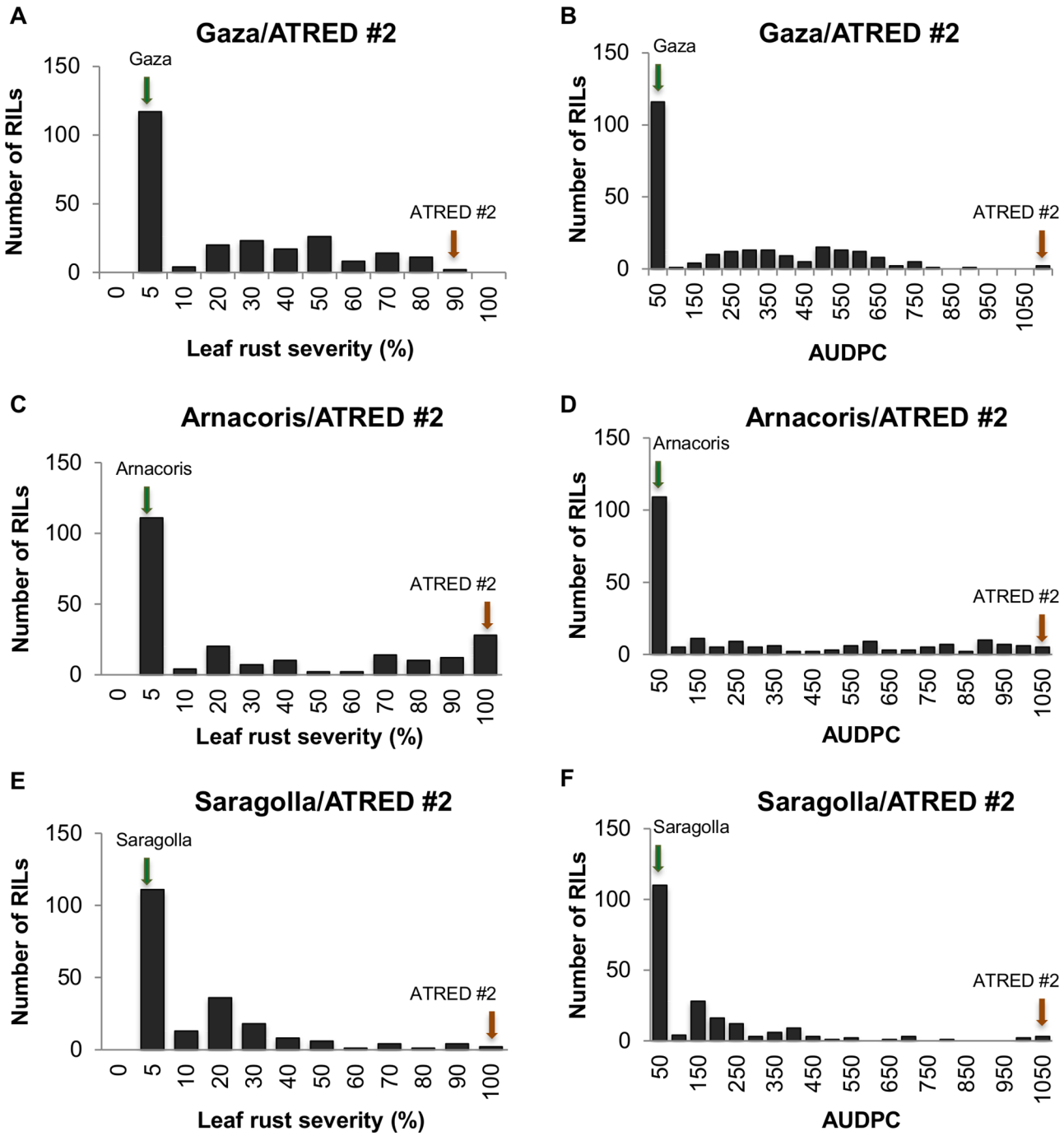


FIGURE 1 | Frequency distributions of leaf rust severity and AUDPC in adult plants of the three RIL populations, evaluated for leaf rust reaction to the Mexican race BBG/BP in field plot tests. Panels represent **(A)** LRS for Gaza/ATRED #2, **(B)** AUDPC for Gaza/ATRED #2, **(C)** LRS for Arnacoris/ATRED #2, **(D)** AUDPC for Arnacoris/ATRED #2, **(E)** LRS for Saragolla/ATRED #2, **(F)** AUDPC for Saragolla/ATRED #2. Green arrows indicate positions of the resistant parents, and red arrows indicate positions of the susceptible parent ATRED #2.

QTL, designated as *QLr.usw-1BL.1*, was detected on the long arm of chromosome 1B for the 2013 field trials at CENEB (**Table 4**), explaining up to 17.7% of the phenotypic variance for the AUDPC (LOD 8.9) and *BS00060686_51* was the most significant marker within the interval.

Based on the allelic state of the SNP markers *BS00010355_51* and *BS00060686_51*, the RILs from the cross Arnacoris/ATRED #2 were classified into carriers and non-carriers of *QLr.usw-7BL* and *QLr.usw-1BL.1*, and mean separation for LRS and AUDPC was performed using Fisher's LSD test. Clearly, *QLr.*

TABLE 2 | Statistics of the three linkage maps from the three RIL populations Gaza/ATRED #2, Arnacoris/ATRED #2 and Saragolla/ATRED #2.

Resistant parent	Mapped SNPs	# of LGs	% A Genome	% B Genome	Map length (cM)	# of Bins	Average inter-bin distance (cM)
Gaza	6248	35	51.2	48.8	4476.2	1431	3.1
Arnacoris	7315	29	43	57	3745	1549	2.4
Saragolla	5345	33	35.8	64.2	3647.2	1293	2.8

TABLE 3 | Leaf rust resistance QTL detected by CIM in the Gaza/ATRED #2 population.

QTL	Flanking markers	Peak position (cM)	Trait	LOD	R^2 (%)	Additive effect
QLr.usw-6BS	*RAC875_c82406_177- CAP7_c10772_156*	1.3	LRS_F_6	15.5	26.7	-7.1
			AUDPC_F_6	21.1	34.5	-94.3
			LRS_F_8	70.7	74.5	-20.1
			AUDPC_F_8	79.5	78.5	-189.6
QLr.usw-6BL	*wsnp_Ex_c45713_51429315-GENE-3689_293*	91	LRS_F_6	11.4	20.5	-6.5
			AUDPC_F_6	10.9	19.6	-69.3
			LRS_F_8	10.7	18.7	-5.9
			AUDPC_F_8	12.3	21.3	-53.1

usw-7BL had the strongest effect on all the traits and its presence alone conferred the highest level of resistance (**Figures 3A**, **B**). The results also showed that *QLr.usw-1BL.1* was significant in reducing leaf rust symptoms only in the absence of *QLr.usw-7BL* (**Figures 3A**, **B**).

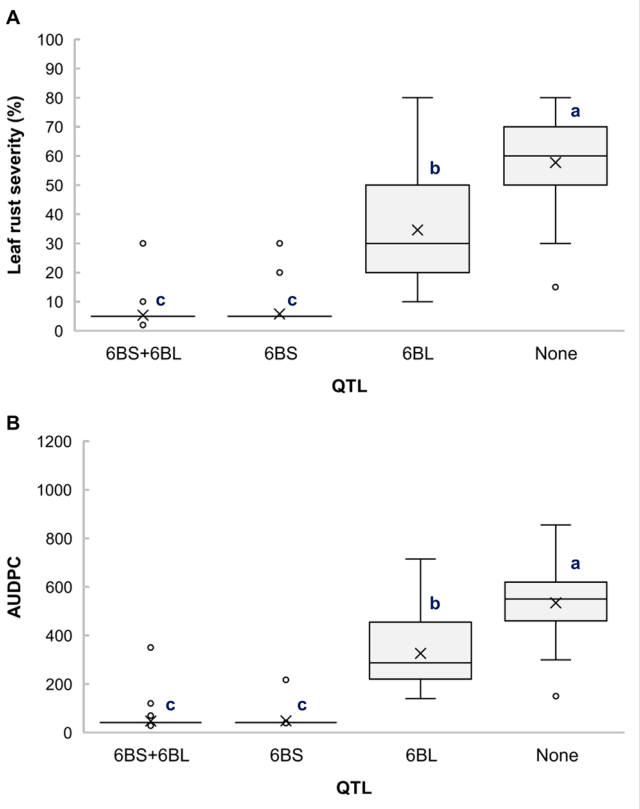


FIGURE 2 | Boxplots illustrating the effects of *QLr.usw-6BS* and *QLr.usw-6BL* on **(A)** LRS and **(B)** AUDPC in the Gaza/ATRED #2 population. The QTL effects are shown in isolation and in combination. Means with the same letter are not significantly different from one another at $P < 0.05$.

Molecular marker testing using the SSR markers *Xwmc44* and *Xgwm259* linked to the APR gene *Lr46* on chromosome 1BL showed inconclusive results, since both positive (Kofa) and negative (CDC Verona) controls had PCR products of similar size for *Xwmc44* (**Supplementary Figure S2A**), whereas marker *Xwmg259* had no PCR product in Arnacoris (**Supplementary Figure S2B**). However, genotyping with the KASP marker linked to *Lr46* showed that Arnacoris, as well as the RILs with and without *QLr.usw-1BL.1*, clustered with the susceptible parent ATRED #2 and the negative controls CDC Verona and Strongfield, suggesting that Arnacoris may not be carrying *Lr46* (**Supplementary Figure S3**). PCR amplification using the NBS-LRR primers *4406F/4840R*, which are linked to *Lr14a*, showed that the positive controls Somayoa and Sachem carry this marker, while Arnacoris had the null allele, which indicates that *QLr.usw-7BL* is different from *Lr14a* (**Supplementary Figure S4**).

QTL in the Saragolla/ATRED #2 Population

Three QTL controlled leaf rust resistance in the Saragolla/ATRED #2 population, which were detected by CIM on chromosomes 1BL, 2BS, and 3B (**Table 5**). *QLr.usw-2BS* peaked at 82 cM on chromosome 2B with *wsnp_Ex_c18354_27181086* being the most significant marker linked to this major QTL. *QLr.usw-2BS* explained up to 42.3% of the final leaf rust severity variance, and was the only QTL detected for the host reactions of the F_8 RILs with a LOD score of 25.18. *QLr.usw-3B* peaked at marker *RAC875_rep_c82061_78* on chromosome 3B and accounted for 9.4% of the AUDPC variance. The minor QTL on the long arm of chromosome 1B, designated as *QLr.usw-1BL.2*, was flanked by markers *wsnp_Ex_c4436_7981188* and *BS00000010_51*, and explained up to 7.1% of the observed variance for AUDPC (**Table 5**).

Analysis of different combinations of QTL based on the allelic state of SNP markers *BS00000010_51*, *wsnp_Ex_c18354_27181086,* and *RAC875_rep_c82061_78* showed that the RILs carrying all three QTL had the lowest averages for LRS (5.8%) (**Figure 4A**) and AUDPC (59) (**Figure 4B**), which indicates the additive effects of these QTL. It is also noticeable

TABLE 4 | Leaf rust resistance QTL detected by CIM in the Arnacoris/ATRED #2 population.

QTL	Flanking markers	Peak position (cM)	Trait	LOD	R^2 (%)	Additive effect
QLr.usw-1BL.1	*BS00060686_51 - Kukri_c46030_412*	0	LRS_F_6	8.7	17.3	-8.9
			$AUDPC\_F_6$	8.9	17.7	-84.5
QLr.usw-7BL	*Tdurum_contig30545_715 – Bobwhite_c42202_158*	127	LRS_F_6	49.3	65.9	-27.3
			$AUDPC\_F_6$	46.1	63.4	-239.3
			HR_F_8	263.3	99.6	0.5

TABLE 5 | Leaf rust resistance QTL detected by CIM in the Saragolla/ATRED #2 population.

QTL	Flanking markers	Peak position (cM)	Trait	LOD	R^2 (%)	Additive effect
QLr.usw-1BL.2	wsnp_Ex_c4436_7981188 - BS00000010_51	27	LRS_F_6	2.02	4.6	-9.1
			$AUDPC\_F_6$	3.16	7.1	-109.8
QLr.usw-2BS	Tdurum_contig76118_145 - wsnp_Ex_c18354_27181086	82	LRS_F_6	23.67	42.3	-12
			$AUDPC\_F_6$	20.95	38.6	-108.7
			HR_F_8	25.18	44.3	0.3
QLr.usw-3B	Tdurum_contig33168_461 - RAC875_rep_c82061_78	13	LRS_F_6	3.96	8.8	-4.1
			$AUDPC\_F_6$	4.24	9.4	-41.8

that *QLr.usw-2BS* had the major effect on leaf rust symptoms compared to both *QLr.usw-3B* and *QLr.usw-1BL.2*. However, both *QLr.usw-3B* and *QLr.usw-1BL.2* significantly reduced the leaf rust symptoms in the absence of *QLr.usw-2BS* (**Figure 4**).

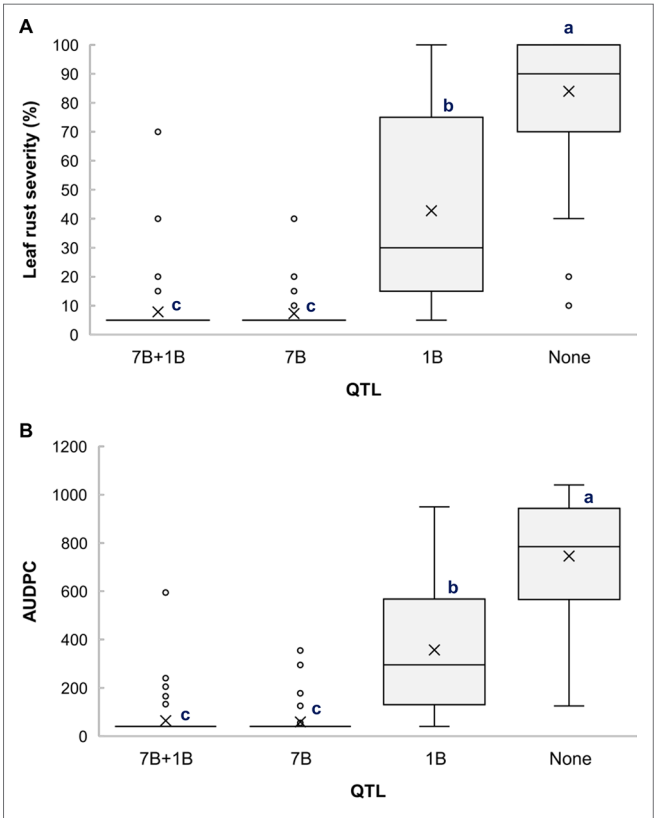


FIGURE 3 | Boxplots illustrating the effects of *QLr.usw-7BL* and *QLr.usw-1BL.1* on **(A)** LRS and **(B)** AUDPC in the Arnacoris/ATRED #2 population. The QTL effects are shown in isolation and in combination. Means with the same letter are not significantly different from one another at $P < 0.05$.

Both parents and selected lines from the Saragolla/ATRED #2 population were genotyped with three SSR markers (*Xwmc764*, *Xgwm210,* and *Xwmc661*) linked to *Lr16* on chromosome 2BS (**Supplementary Figure S5**). The results showed polymorphism between the amplicons from Saragolla and AC Domain for the three markers, which suggests that *QLr.usw-2BS* is different from *Lr16*. The KASP markers *kwm677* (**Supplementary Figure S6A**) and *kwm744* (**Supplementary Figure S6B**) also showed polymorphism between Saragolla and the *Lr16*-carrying AC Domain.

Physical Mapping to the Durum Wheat Reference Genome

The DNA sequences associated with the SNP markers mapping within each QTL LOD plot area in Gaza, Arnacoris, and Saragolla were physically mapped to the durum wheat reference genome of Svevo (Maccaferri et al., 2019) to identify candidate genes for leaf rust resistance (**Supplementary File S2**). Several genes identified within these QTL intervals encode proteins with motifs known to be associated with disease resistance such as NBS-LRR receptor proteins, calcium-dependant lipid-binding (CaLB domain) proteins, ATP-binding cassette (ABC) transporter proteins, as well as several receptor-like protein kinases (RLKs) (**Supplementary File S2**).

DISCUSSION

Diversification of the genetic basis of leaf rust resistance and breeding for durable resistance are both priorities in many durum wheat breeding programs worldwide, especially after the emergence of new virulent races of *P. triticina*. The goal of the present study was to characterize and map putatively uncharacterized genes for leaf rust resistance in the three durum genotypes: Gaza, Arnacoris, and Saragolla. Inheritance studies indicated the involvement of several loci that controlled leaf rust resistance in these lines, including at least one APR gene

in each of Gaza and Saragolla. The segregation ratios of the $F_{2:3}$ progenies from the Gaza/ATRED #2 population suggested one resistance gene in the seedlings evaluation and two in the adult plants evaluation. However, segregation of the F_8 RILs adult plants suggested the presence of a single resistance gene. This discrepancy may be caused by the ineffectiveness of the Gaza APR gene in El Batán, since certain APR genes are known to be environment or temperature sensitive (Kaul and Shaner, 1989; McIntosh et al., 1995; Risk et al., 2012). For the Arnacoris/ATRED #2 population, the seedling and adult plant stage evaluations of the $F_{2:3}$ progenies suggested the presence of two resistance genes, while the segregation ratio of the F_8 RILs supported a single gene theory. Nonetheless, the $F_{2:3}$ results were indeed supported by the QTL mapping results, since two QTL were detected in this population. The major phenotypic effect of *QLr.usw-7BL* suggests that this is an all-stage resistance gene, while *QLr.usw-1BL.1* is likely an APR gene, with a minor phenotypic effect. The discrepancy between the segregation ratios could be due to the differences between the climatic conditions during the years of evaluations. The $F_{2:3}$ progenies were evaluated in El Batán in 2011, while the F_8 RILs evaluation was conducted in 2014, and it is possible that the APR gene was not effective in 2014. The environmental effects on this gene warrant for follow-up field studies under various conditions. Although the material was inoculated with the race BBG/BP of *P. triticina* in both years, it is also possible that the racial constitution of the natural field population of the pathogen has been different between 2011 and 2014, thus affecting the results. The Saragolla/ATRED #2 population showed similar results for the $F_{2:3}$ progenies, with the seedling data suggesting the presence of a single resistance gene, while the adult plant data suggested the presence of at least one additional APR gene. However, the segregation ratio from the F_8 generation could not fit neither one nor two-gene models and we could not find a reasonable explanation for this aberration, which also calls for further evaluation of this cultivar under various environments.

Composite interval mapping identified genomic regions on chromosomes 1BL, 2BS, 3B, 6BS, 6BL, and 7BL, which were associated with leaf rust resistance in these three durum cultivars. These may be particularly valuable in breeding since they can be strategically combined to produce more durable resistance. The SNP markers linked to all the leaf rust resistance QTL identified in the present study were converted into KASP markers (**Supplementary Table S2**) and are currently being validated at CIMMYT.

QTL on Chromosome 1B

The distal region of chromosome 1BL is known to carry the APR gene *Lr46*, identified in the wheat cultivar Pavon76 (Singh et al., 1998). The flanking SNP markers for *QLr.usw-1BL.1* in Arnacoris, *BS00060686_51,* and *Kukri_c46030_412*, mapped at 152.5 cM and 162.3 cM, respectively, on chromosome 1BL in the SNP-based consensus map of tetraploid wheat (Maccaferri et al., 2015). Likewise, the SNP marker *wsnp_Ex_c4436_7981188* that flanks *QLr.usw-1BL.2* in Saragolla, mapped at 145.4 cM on 1BL. The two SSR markers linked to *Lr46*, *Xwmc44* and *Xgwm259* (William et al., 2006), mapped at 140.6 and 150.6 cM,

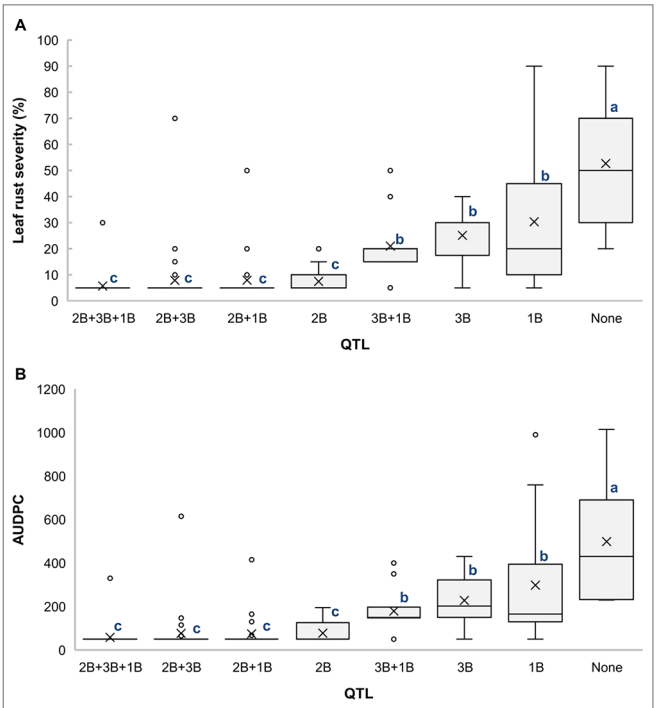


FIGURE 4 | Boxplots illustrating the effects of *QLr.usw-2BS, QLr.usw-3B*, and *QLr.usw-1BL.2* on **(A)** LRS and **(B)** AUDPC in the Saragolla/ATRED #2 population. The QTL effects are shown in isolation and in combination. Means with the same letter are not significantly different from one another at $P < 0.05$.

respectively, in the same durum consensus map (Maccaferri et al., 2015). Despite the proximity between these marker intervals, further molecular marker analyses showed that Arnacoris may not be carrying *Lr46* (**Supplementary Figure S3**). This is supported by the observation that *QLr.usw-1BL.1* alone has a much stronger phenotypic effect on disease symptoms reduction (**Figure 3**) compared to *Lr46*, when present alone in durum wheat (K. Ammar, *unpublished*). However, *Qlr.usw-1BL.2* present in Saragolla is likely *Lr46*, which is consistent with the observation that *QLr.usw-1BL.2* has a very moderate effect on disease symptoms reduction (**Figure 4**).

Other leaf rust resistance genes that map to chromosome 1B include *Lr33* (Dyck et al., 1987), *Lr44* (Dyck and Sykes, 1994), *Lr71* (Singh et al., 2013b), and *Lr26* (Mago et al., 2002; Mago et al., 2005). However, virulence in durum specific races is common for *Lr33*, such as the Mexican race BBG/BP that is used in the present study (Herrera-Foessel et al., 2008b; Huerta-Espino et al., 2011; Lan et al., 2017). Genes *Lr44* and *Lr71* were both identified in spelt wheat (*T. spelta*) and no reports indicate their presence in durum wheat. Furthermore, the SSR markers *Xgwm18* and *Xbarc187*, which are closely linked to *Lr71* (Singh et al., 2013b), mapped at 35.6 and 35.7 cM in the tetraploid consensus map, respectively (Maccaferri et al., 2015). Based on the position of the SNP markers flanking *QLr.usw-1BL.1* in the same consensus map at 152.5 and 162.3 cM, it is possible to conclude that Arnacoris does not carry *Lr71*. The short arm of chromosome 1R of rye (*Secale cereale*) carries the leaf rust

resistance gene *Lr26* and has been widely used in wheat breeding programs through the 1BL.1RS (wheat-rye) translocation (Mago et al., 2005). However, there is no indication of the presence of the 1BL.1RS translocation in Arnacoris. While it is highly unlikely that *QLr.usw-1BL.1* harbor any of the three designated genes on chromosome 1B, definitive proof could not be obtained in the present study.

QTL on Chromosome 2B

QLr.usw-2BS mapped to a chromosome region known to carry at least six designated leaf rust resistance genes: *Lr13* (Seyfarth et al., 2000), *Lr16* (McCartney et al., 2005), *Lr23* (McIntosh and Dyck, 1975), *Lr35* (Seyfarth et al., 1999), *Lr48* (Saini et al., 2002), and *Lr73* (Park et al., 2014). *Lr13*, *Lr35* and *Lr48* are reportedly APR genes, whereas Saragolla likely carries a major seedling resistance gene, based on the seedling evaluations conducted at CIMMYT. In addition, the molecular markers *Xbarc55* and *IWB35283*, previously reported to be linked to *Lr13* (Zhang et al., 2016), mapped at 72.1 cM and 74.2 cM, respectively, in the tetraploid wheat consensus map (Maccaferri et al., 2015), while the SNP markers *Tdurum_contig76118_145* and *wsnp_Ex_c18354_27181086* flanking *QLr.usw-2BS*, mapped at 8.4 cM and 12.3 cM, respectively, in the same consensus map. Also, *Lr35* is unlikely to be present in Saragolla since it was introgressed into hexaploid wheat from *Aegilops speltoides* and we did not detect molecular evidence that Saragolla is carrying this introgression. The SNP markers *IWB31002*, *IWB39832*, *IWB34324*, *IWB72894*, *IWB36920,* and *IWB70147* are reported to be co-segregating with the leaf rust resistance gene *Lr48* (Nsabiyera et al., 2016). However, none of these SNP markers are located within the genetic interval of *QLr.usw-2BS* in both the Saragolla/ATRED #2 genetic map (**Supplementary File S1**) and the tetraploid consensus map (Maccaferri et al., 2015). Therefore, it can be assumed that *QLr.usw-2BS* is different from *Lr48*. *Lr23* is also an unlikely candidate for *QLr.usw-2BS*, since the race BBG/BP used in this study is virulent to *Lr23* (Huerta-Espino et al., 2008; Huerta-Espino et al., 2011; Loladze et al., 2014). *Lr73* was identified in Australia in the common wheat genotype Morocco, which is widely susceptible to isolates of *P. triticina* (Park et al., 2014), including race BBBQJ with a virulence phenotype and SSR genotype similar to BBG/BP (Aoun et al., 2017). Hence, it is unlikely that *Lr73* is present in Sragolla. The all-stage resistance gene *Lr16* maps to the distal end of chromosome 2BS, and is closely linked to the SSR markers *Xwmc764*, *Xgwm210,* and *Xwmc661* (McCartney et al., 2005) and to KASP markers *kwm677* and *Kwm744* (Kassa et al., 2017). However, genotyping with these markers showed polymorphism between Saragolla and the *Lr16*-carrying AC Domain (**Supplementary Figures S5**, **S6**), which suggests that *QLr.usw-2BS* is different from *Lr16*. Aoun et al. (2017) identified *LrPI244061* that confers resistance to race BBBQJ of *P. triticina*, on chromosome 2BS of the durum landrace PI 244061. Based on the tetraploid wheat consensus map of chromosome 2B (Maccaferri et al., 2015), markers *IWB6117* and *IWB72183* linked to *LrPI244061* (Aoun et al., 2017) mapped at 42.8 cM and 45.8 cM, respectively, whereas markers *Tdurum_contig76118_145* and *wsnp_Ex_c18354_27181086*, flanking *QLr.usw-2BS*, mapped at 8.4 cM and 12.3 cM, respectively. In addition, marker *IWB72183* mapped at 133.8 cM on chromosome 2B_1, in the Saragolla/ATRED #2 genetic map (**Supplementary File S1**), while *QLr.usw-2BS* peaked at 82 cM. In summary, *QLr.usw-2BS* does not seem to include any of the six previously designated *Lr* genes and therefore is likely to harbor a previously uncharacterized gene.

QTL on Chromosome 3B

Chromosome 3B is known to carry the race-specific resistance gene *Lr27* that requires the presence of the complementary gene *Lr31* on chromosome 4B, to confer leaf rust resistance in wheat (Mago et al., 2011). However, race BBG/BP emerged in 2008 in the state of Sonora, in northwestern Mexico, after acquiring virulence to the adult plant race-specific resistance gene *Lr12* and the seedling complementary resistance genes *Lr27+Lr31* (Huerta-Espino et al., 2008; Huerta-Espino et al., 2011). Therefore, *Lr27* is unlikely to be involved in the resistance to BBG/BP in Saragolla. The APR gene *Lr74* was identified in the hexaploid wheat population Ning7840 × Clark, and mapped to the short arm of chromosome 3B, closely linked to the SNP markers *IWA6651*, *IWA3724*, *IWA4654*, *IWA1702*, *IWA5203*, *IWA5202,* and *IWA5201* (Li et al., 2017). Kolmer et al. (2018b) identified a QTL for adult plant leaf rust resistance in the soft red winter wheat cultivar Caldwell, and mapped it very close to the *Lr74* locus. Based on the tetraploid wheat consensus map (Maccaferri et al., 2015), all of the SNP markers linked to *Lr74* mapped between positions 6 and 7.1 cM on chromosome 3BS, while markers *Tdurum_contig33168_461* and *RAC875_rep_c82061_78*, which flank *QLr.usw-3B* in Saragolla, mapped at 87 and 88 cM, respectively, in the same consensus map. In addition, marker *IWA3724* mapped at 15.9 cM on LG 3B_1 in the Saragolla/ATRED #2 map (**Supplementary File S1**), while *QLr.usw-3B* spanned the interval 11.5 to 13.5 cM on LG 3B_2. Therefore, we can assume that *QLr.usw-3B* is distinct from the APR gene *Lr74*. Recently, the new APR gene *Lr77* was mapped to chromosome 3BL in the hard red winter wheat cultivar "Santa Fe" (Kolmer et al., 2018c). The SNP markers *IWB32805* and *IWB10344* co-segregating with *Lr77*, mapped respectively at 148.4 cM and 151.6 cM in the tetraploid wheat consensus map, approximately 62 cM distal to the markers flanking *QLr.usw-3B* (Maccaferri et al., 2015). Furthermore, marker *IWB32805* mapped at 150.8 cM on LG 3B_2 in the Saragolla/ATRED #2 map, while *QLr.usw-3B* peaked at 13 cM on the same LG (**Supplementary File S1**). These large distances between the two intervals harboring *Lr77* and *QLr.usw-3B* indicate that they are two distinct *Lr* genes. *Lr79* is another newly mapped all-stage leaf rust resistance gene on chromosome 3BL, from the durum wheat landrace Aus26582 (Qureshi et al., 2018). Based on comparative analysis using the consensus 90K SNP genetic map (Wang et al., 2014) and the physical map of Chinese Spring (RefSeq v1.0), the authors estimated the distance between *Lr77* and *Lr79* at approximately 12 cM or 9.2 Mbp (Qureshi et al., 2018). Hence, it is possible to argue that *Lr79* does also map at a large genetic distance from *QLr.usw-3B*, and that the latter may be an uncharacterized leaf rust resistance gene.

QTL on Chromosome 6B

Previous studies have reported three designated leaf rust resistance genes that map to the short arm of chromosome 6B: *Lr36* (Dvorak and Knott, 1990), *Lr53* (Marais et al., 2005), and *Lr61* (Herrera-Foessel et al., 2008a). Genes *Lr36* and *Lr53* were introgressed into chromosome 6BS of hexaploid wheat from *T. speltoides* and *T. dicoccoides*, respectively (Dvorak and Knott, 1990; Marais et al., 2005), but no reports of the transfer of either of these genes into durum wheat are available. Hence, these two genes are unlikely candidates for the major QTL *QLr.usw-6BS*, identified on the short arm of chromosome 6B in Gaza. However, since *T. dicoccoides* (genome AABB) is the wild progenitor of durum wheat, further genetic analysis and allelism tests would be required to fully rule out *Lr53* as a candidate. *Lr61* was identified in the CIMMYT-derived Chilean durum wheat cultivar Guayacan INIA, with linkage to marker *Xwmc487*, and is effective against the *P. triticina* race BBG/BP, predominant in northwestern Mexico (Herrera-Foessel et al., 2008a; Loladze et al., 2014). The presence of susceptible plants in the F_2 progeny from the cross between Gaza and the *Lr61*-carrying Sooty_9/Rascon_37//Guayacan INIA indicated that *QLr.usw-6BS* was neither allelic nor linked to *Lr61* (**Supplementary Table S1**). Furthermore, Gaza showed polymorphism compared to the *Lr61* carrier Guayacan INIA, for the SSR marker *Xwmc487* (**Supplementary Figure S1**). Kthiri et al. (2018) identified two leaf rust resistance genes in the two durum cultivars Geromtel_3 and Tunsyr_2, conferring resistance to race BBG/BP of *P. triticina*, and mapping to the short arm of chromosome 6B. The authors showed that *Lr_Geromtel_3* and *Lr_Tunsyr_2* were either allelic or closely linked to each other and to *Lr61* (Kthiri et al., 2018; Kthiri, 2017). Allelism tests showed that the resistance gene in Gaza is either allelic or closely linked to the gene in Geromtel_3 but not to the gene in Tunsyr_2 (**Supplementary Table S1**). The Australian breeding line PI 209274 carries resistance to races BBBQJ and BBB/BN_*Lr61*vir of *P. triticina*, on chromosome 6BS (Aoun et al., 2017). Molecular markers *IWB39456* and *IWB416*, both linked to *LrPI209274* (Aoun et al., 2017), mapped at 5.6 cM in the Gaza/ATRED #2 linkage map (**Supplementary File S1**), close to the peak position of *QLr.usw-6BS* at 1.3 cM. Additional fine mapping studies and allelism tests are required to determine the relationship between *QLr.usw-6BS* and *LrPI209274.*

Several *Lr* genes have also been previously identified on the long arm of chromosome 6B. These include *Lr9*, which was translocated into hexaploid wheat from *Ae. umbellulata* (Schachermayr et al., 1994). However, there are no reports of this gene being transferred into durum wheat, thus, *Lr9* is unlikely a candidate gene for *QLr.usw-6BL*. The all-stage resistance gene *Lr3* is a known locus on the long arm of chromosome 6B with four reported alleles: *Lr3a*, *Lr3bg*, *Lr3ka,* and *Lr3d* (Kolmer, 2015c). Herrera-Foessel et al. (2007) identified and mapped *Lr3* and the closely linked gene Lr_{Camayo} in the two durum wheat lines Storlom and Camayo, respectively. Although these two closely linked genes on chromosome 6BL confer resistance telo *P. triticina* races prevalent on durum wheat in Northwestern Mexico, allelism tests between Gaza and Cirno C2008, a carrier of Lr_{Camayo}, suggested that the resistance to leaf rust in Gaza is unrelated to Lr_{Camayo} (**Supplementary Table S1**). Furthermore, analysis of seedling and adult plant evaluation results confirmed the involvement of an APR gene for leaf rust resistance in Gaza, while both *Lr3* and Lr_{Camayo} are all-stage resistance genes. Recently, Lan et al. (2017) identified *QLr.cim-6BL* that confers adult plant resistance to race BBG/BP of *P. triticina*, in the CIMMYT durum wheat line Bairds. Likewise, Aoun et al. (2017) mapped *LrPI387263* to the long arm of chromosome 6B, in the Ethiopian durum landrace PI 387263. Based on comparative mapping analysis, molecular markers linked to these genes mapped either very close to or overlapping the *QLr.usw-6BL* interval in the tetraploid wheat consensus map (Maccaferri et al., 2015) or the Gaza/ATRED #2 map (**Supplementary File S1**). Therefore, *QLr.usw-6BL*, *QLr.cim-6BL,* and *LrPI387263* could possibly harbor the same gene.

QTL on Chromosome 7B

QLr.usw-7BL maps in a gene-dense region with several genes/QTL for resistance to rusts and other fungal diseases, including *Lr14a*, one of the most widely exploited *Lr* genes in wheat (Herrera-Foessel et al., 2008b; Terracciano et al., 2013), and the closely linked gene *Lr14b* (Dyck and Samborski, 1970), the slow-rusting APR gene *Lr68* (Herrera-Foessel et al., 2012), *LrBi16* which is allelic to *Lr14a* (Zhang et al., 2015), and *LrFun*, which is closely linked to *Lr14a* (Xing et al., 2014). The prevalent Mexican races of *P. triticina*, including race BBG/BP that is used in the present study, are virulent for *Lr14b* (Ordoñez and Kolmer, 2007; Herrera-Foessel et al., 2008b); therefore, it is unlikely that this gene is conferring leaf rust resistance in Arnacoris. *Lr68* is an adult-plant resistance gene that confers a slow-rusting phenotype, however, *QLr.usw-7BL* had a major effect on leaf rust resistance in Arnacoris, at both the seedling and adult plant stages (**Figure 3**), making *Lr68* an unlikely candidate for *QLr.usw-7BL*. When tested with the NBS-LRR primers *4406F/4840R*, both parental lines Arnacoris and ATRED #2 had the null allele (**Supplementary Figure S4**), indicating that *Lr14a* is not segregating in this population. Further investigation will be required to identify the specific gene responsible for the leaf rust resistance conferred by *QLr.usw-7BL*.

Physical Mapping and Candidate Gene Identification

Anchoring of the SNP markers associated with the various QTL detected in the present study to the durum wheat reference genome identified several genes that encode for proteins known to be involved in plant pathogen interactions and disease resistance. So far, three race specific leaf rust resistance genes (*Lr1*, *Lr10,* and *Lr21*) have been cloned in wheat, and all three proteins contained NBS-LRR motifs (Feuillet et al., 2003; Huang et al., 2003; Cloutier et al., 2007). The APR gene *Lr34* protein is a full-size ABC transporter (Krattinger et al., 2009) while *Lr67* was shown to encode a recently evolved hexose transporter (Moore et al., 2015). Map-based cloning of *Yr36*, a gene that confers non-race-specific adult plant resistance to stripe rust in wild emmer wheat (*T. turgidum* ssp. *dicoccoides*), showed that it encodes a protein with a predicted kinase domain (Fu et al., 2009). The genes identified within the leaf rust resistance QTL intervals in Gaza, Arnacoris,

and Saragolla have structures typical of disease resistance proteins and are excellent gene candidates for further studies.

CONCLUSION

The present study identified six genomic regions involved in leaf rust resistance in durum wheat. The resistance to race BBG/BP of *P. triticina* in the Middle Eastern landrace Gaza was controlled by two QTL on chromosome 6B. *Qlr.usw-6BS* accounted for most of the phenotypic variance and is neither allelic nor linked to *Lr61*. The second QTL, *QLr.usw-6BL*, may be a new APR gene for leaf rust resistance in wheat. Likewise, the French cultivar Arnacoris carried two QTL for leaf rust resistance. *QLr.usw-1BL.1* mapped to the *Lr46* region; however, Arnacoris did not carry any of the *Lr46* molecular markers. The major QTL on chromosome 7B in Arnacoris, *QLr.usw-7BL*, explained most of the leaf rust phenotypic variance and was shown to be different from the widely deployed gene *Lr14a*. The Italian durum variety Saragolla carried a major QTL on chromosome 2B, designated as *QLr.usw-2BS*, which accounted for most of the phenotypic variance, as well as two minor QTL on chromosomes 3B and 1BL. Molecular marker analysis suggested that *QLr.usw-2BS* is distinct from *Lr16* while *QLr.usw-1BL.2* is likely the APR leaf rust resistance gene *Lr46*, and *QLr.usw-3B* is a potentially uncharacterized leaf rust resistance gene. Physical mapping of the SNP markers associated with these QTL to the durum wheat reference sequence enabled the identification of candidate genes for leaf rust resistance in these cultivars. With the availability of SNP markers tightly linked to all these QTL, some with major and other with minor effects, the durum wheat lines used in the present study can be used as donors to strategically combine genes with different modes of action to produce a more durable leaf rust resistance in durum wheat.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the manuscript/**Supplementary Files**.

AUTHOR CONTRIBUTIONS

CP and KA conceived and designed the study. KA developed the populations. AL produced the purified inoculum and conducted phenotyping and data collection. AL and DK analyzed the phenotypic data. SD performed DNA extraction. DK performed the genotyping and QTL analysis and drafted the manuscript. AN'D built the genetic maps. KN conducted the physical mapping. SW revised the manuscript. All authors contributed to manuscript revision, read, and approved the submitted version.

FUNDING

We are grateful for the financial support of Genome Canada, Genome Prairie, Western Grains Research Foundation, Saskatchewan Wheat Development Commission, Alberta Wheat Development Commission, the Saskatchewan Ministry of Agriculture, Viterra, the Saskatchewan Innovation and Opportunity Scholarship, and the Rene Vandeveld Postgraduate Scholarship. CIMMYT's contribution to the present work has been made possible thanks to funding from "Patronato para la Investigación y Experimentación Agrícola del Estado de Sonora (PIEAES)," "Fundacíón Produce Sonora, México," and the Durable Rust Resistance Wheat (DRRW) project as part of the Borlaug Global Rust Initiative (BGRI).

ACKNOWLEDGMENTS

This manuscript is part of DK's Ph.D. thesis at the University of Saskatchewan. This research was conducted as part of the Canadian Triticum Applied Genomics (CTAG2) project.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01247/full#supplementary-material

REFERENCES

Aoun, M., Kolmer, J. A., Rouse, M. N., Chao, S., Denbel Bulbula, W., Elias, E. M., et al. (2017). Inheritance and bulked segregant analysis of leaf rust and stem rust resistance in durum wheat genotypes. *Phytopathology* 107, 1496–1506. doi: 10.1094/PHYTO-12-16-0444-R

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., et al. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* 357, 93–97. doi: 10.1126/science.aan0032

Buerstmayr, M., Matiasch, L., Mascher, F., Vida, G., Ittu, M., Robert, O., et al. (2014). Mapping of quantitative adult plant field resistance to leaf rust and stripe rust in two European winter wheat populations reveals co-location of three QTL conferring resistance to both rust pathogens. *Theor. Appl. Genet.* 127, 2011–2028. doi: 10.1007/s00122-014-2357-0

Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., et al. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci.* 110, 8057–8062. doi: 10.1073/pnas.1217133110

Cloutier, S., McCallum, B. D., Loutre, C., Banks, T. W., Wicker, T., Feuillet, C., et al. (2007). Leaf rust resistance gene Lr1, isolated from bread wheat (Triticum aestivum L.) is a member of the large psr567 gene family. *Plant Mol. Biol.* 65, 93–106. doi: 10.1007/s11103-007-9201-8

Dreisigacker, S., Sehgal, D., Reyes-Jaimez, A. E., Luna-Garrido, B., Muñoz-Zavala, S., Núñez-Ríos, C. et al., (2016). *CIMMYT wheat molecular genetics: laboratory protocols and applications to wheat breeding.* Mexico, D.F: CIMMYT.

Dvorak, J., and Knott, D. R. (1990). Location of a Triticum speltoides chromosome segment conferring resistance to leaf rust in Triticum aestivum. *Genome* 33, 892–897. doi: 10.1139/g90-134

Dyck, P. L., Kerber, E. R., and Lukow, O. M. (1987). Chromosome location and linkage of a new gene (Lr33) for reaction to Puccinia recondita in common wheat. *Genome* 29, 463–466. doi: 10.1139/g87-080

Dyck, P. L., and Samborski, D. J. (1970). The genetics of two alleles for leaf rust resistance at the Lr14 locus in wheat. *Can. J. Genet. Cytol.* 12, 689–694. doi: 10.1139/g70-091

Dyck, P. L., and Sykes, E. E. (1994). Genetics of leaf-rust resistance in three spelt wheats. *Can. J. Plant Sci.* 74, 231–233. doi: 10.4141/cjps94-047

Ellis, J. G., Lagudah, E. S., Spielmeyer, W., and Dodds, P. N. (2014). The past, present and future of breeding rust resistant wheat. *Front. Plant Sci.* 5, 641. doi: 10.3389/fpls.2014.00641

Feuillet, C., Travella, S., Stein, N., Albar, L., Nublat, A. L., and Keller, B. (2003). Map-based isolation of the leaf rust disease resistance gene Lr10 from the hexaploid wheat (Triticum aestivum L.) genome. *Proc. Natl. Acad. Sci. U. S. A.* 100, 15253–15258. doi: 10.1073/pnas.2435133100

Fu, D., Uauy, C., Distelfeld, A., Blechl, A., Epstein, L., Chen, X., et al. (2009). A kinase-START gene confers temperature-dependent resistance to wheat stripe rust. *Science* 323, 1357–1360. doi: 10.1126/science.1166289

Herrera-Foessel, S., Huerta-Espino, J., Calvo-Salazar, V., Lan, C. X., and Singh, R. P. (2014a). Lr72 confers resistance to leaf rust in durum wheat cultivar Atil C2000. *Plant Dis.* 98, 631–635. doi: 10.1094/PDIS-07-13-0741-RE

Herrera-Foessel, S., Lagudah, E. S., McIntosh, R., Salazar, V. C., Huerta-Espino, J., and Singh, R. P., (2011). First report of slow rusting gene *Lr46* in durum wheat. *Proceedings of the Borlaug global rust initiative technical workshop*, St. Paul, MN.

Herrera-Foessel, S., Singh, R. P., Huerta-Espino, J., Rosewarne, G. M., Periyannan, S. K., Viccars, L., et al. (2012). Lr68: a new gene conferring slow rusting resistance to leaf rust in wheat. *Theor. Appl. Genet.* 124, 1475–1486. doi: 10.1007/s00122-012-1802-1

Herrera-Foessel, S., Singh, R. P., Huerta-Espino, J., William, H. M., Djurle, A., and Yuen, J. (2008a). Molecular mapping of a leaf rust resistance gene on the short arm of chromosome 6B of durum wheat. *Plant Dis.* 92, 1650–1654. doi: 10.1094/PDIS-92-12-1650

Herrera-Foessel, S., Singh, R. P., Huerta-Espino, J., William, H. M., Garcia, V., Djurle, A., et al. (2008b). Identification and molecular characterization of leaf rust resistance gene Lr14a in durum wheat. *Plant Dis.* 92, 469–473. doi: 10.1094/PDIS-92-3-0469

Herrera-Foessel, S., Singh, R. P., Lillemo, M., Huerta-Espino, J., Bhavani, S., Singh, S., et al. (2014b). Lr67/Yr46 confers adult plant resistance to stem rust and powdery mildew in wheat. *Theor. Appl. Genet.* 127, 781–789. doi: 10.1007/s00122-013-2256-9

Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., Crossa, J., Djurle, A., and Yuen, J. (2008c). Genetic analysis of slow-rusting resistance to leaf rust in durum wheat. *Crop Sci.* 48, 2132–2140. doi: 10.2135/cropsci2007.11.0606

Herrera-Foessel, S. A., Singh, R. P., Huerta-Espino, J., William, M., Rosewarne, G., Djurle, A., et al. (2007). Identification and mapping of Lr3 and a linked leaf rust resistance gene in durum wheat. *Crop Sci.* 47, 1459–1466. doi: 10.2135/cropsci2006.10.0663

Hiebert, C. W., Thomas, J. B., McCallum, B. D., Humphreys, D. G., DePauw, R. M., Hayden, M. J., et al. (2010). An introgression on wheat chromosome 4DL in RL6077 (Thatcher*6/PI 250413) confers adult plant resistance to stripe rust and leaf rust (Lr67). *Theor. Appl. Genet.* 121, 1083–1091. doi: 10.1007/s00122-010-1373-y

Huang, L., Brooks, S. A., Li, W., Fellers, J. P., Trick, H. N., and Gill, B. S. (2003). Map-based cloning of leaf rust resistance gene Lr21 from the large and polyploid genome of bread wheat. *Genetics* 164, 655.

Huerta-Espino, J., Singh, R. P., German, S., McCallum, B. D., Park, R. F., Chen, W. Q., et al. (2011). Global status of wheat leaf rust caused by Puccinia triticina. *Euphytica* 179, 143–160. doi: 10.1007/s10681-011-0361-x

Huerta-Espino, J., Singh, R. P., Herrera-Foessel, S. A., Pérez-López, J. B., and Figueroa-López, P. (2008). First detection of virulence in Puccinia triticina to resistance genes Lr27 + Lr31 present in durum wheat in Mexico. *Plant Dis.* 93, 110–110. doi: 10.1094/PDIS-93-1-0110C

Joehanes, R., and Nelson, J. C. (2008). QGene 4.0, an extensible Java QTL-analysis platform. *Bioinformatics* 24, 2788–2789. doi: 10.1093/bioinformatics/btn523

Kassa, M. T., You, F. M., Hiebert, C. W., Pozniak, C. J., Fobert, P. R., Sharpe, A. G., et al. (2017). Highly predictive SNP markers for efficient selection of the wheat leaf rust resistance gene Lr16. *BMC Plant Biol.* 17, 45. doi: 10.1186/s12870-017-0993-7

Kaul, K., and Shaner, G. (1989). Effect of temperature on adult-plant resistance to leaf rust in wheat. *Phytopathology* 79, 391–394. doi: 10.1094/Phyto-79-391

Knott, D. R. (1989). *The wheat rusts - Breeding for resistance*. Berlin, Heidelberg: Springer-Verlag. doi: 10.1007/978-3-642-83641-1

Kolmer, J. A. (2013). Leaf rust of wheat: pathogen biology, variation and host resistance. *Forests* 4 (1), 70–84. doi: 10.3390/f4010070

Kolmer, J. A. (2015a). A QTL on chromosome 5BL in wheat enhances leaf rust resistance of Lr46. *Mol. Breed.* 35, 74. doi: 10.1007/s11032-015-0274-9

Kolmer, J. A. (2015b). First report of a wheat leaf rust (Puccinia triticina) phenotype with high virulence to durum wheat in the Great Plains region of the United States. *Plant Dis.* 99, 156–156. doi: 10.1094/PDIS-06-14-0667-PDN

Kolmer, J. A. (2015c). Leaf rust resistance in wheat line RL6062 is an allele at the Lr3 locus. *Crop Sci.* 55, 2186–2190. doi: 10.2135/cropsci2015.01.0031

Kolmer, J. A., Bernardo, A., Bai, G., Hayden, M. J., and Chao, S. (2018a). Adult plant leaf rust resistance derived from toropi wheat is conditioned by Lr78 and three minor QTL. *Phytopathology* 108, 246–253. doi: 10.1094/PHYTO-07-17-0254-R

Kolmer, J. A., Chao, S., Brown-Guedira, G., Bansal, U., and Bariana, H. (2018b). Adult plant leaf rust resistance derived from the soft red winter wheat cultivar 'caldwell' maps to chromosome 3BS. *Crop Sci.* 58, 152–158. doi: 10.2135/cropsci2017.05.0272

Kolmer, J. A., and Hughes, M. E. (2017). Physiologic specialization of Puccinia triticina on wheat in the United States in 2015. *Plant Dis.* 101, 1968–1973. doi: 10.1094/PDIS-02-17-0200-SR

Kolmer, J. A., Su, Z., Bernardo, A., Bai, G., and Chao, S. (2018c). Mapping and characterization of the new adult plant leaf rust resistance gene Lr77 derived from Santa Fe winter wheat. *Theor. Appl. Genet.* 131, 1553–1560. doi: 10.1007/s00122-018-3097-3

Kosambi, D. D. (1943). The estimation of map distances from recombination values. *Ann. Eugen.* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Krattinger, S. G., Lagudah, E. S., Spielmeyer, W., Singh, R. P., Huerta-Espino, J., McFadden, H., et al. (2009). A putative ABC transporter confers durable resistance to multiple fungal pathogens in wheat. *Science* 323, 1360–1363. doi: 10.1126/science.1166453

Kthiri, D. (2017). Genetic and molecular characterization of leaf rust resistance from uncharacterized sources of durum wheat (Triticum turgidum L. ssp. durum). [dissertation/PhD. thesis]. Saskatoon (SK): University of Saskatchewan. http://hdl.handle.net/10388/7928.

Kthiri, D., Loladze, A., MacLachlan, P. R., N'Diaye, A., Walkowiak, S., Nilsen, K., et al. (2018). Characterization and mapping of leaf rust resistance in four durum wheat cultivars. *PLoS ONE* 13, e0197317. doi: 10.1371/journal.pone.0197317

Lagudah, E. S. (2011). Molecular genetics of race non-specific rust resistance in wheat. *Euphytica* 179, 81–91. doi: 10.1007/s10681-010-0336-3

Lan, C., Basnet, B. R., Singh, R. P., Huerta-Espino, J., Herrera-Foessel, S., Ren, Y., et al. (2017). Genetic analysis and mapping of adult plant resistance loci to leaf rust in durum wheat cultivar Bairds. *Theor. Appl. Genet.* 130, 609–619. doi: 10.1007/s00122-016-2839-3

Li, C., Wang, Z., Li, C., Bowden, R., Bai, G., Su, Z., et al. (2017). Mapping of quantitative trait loci for leaf rust resistance in the wheat population Ning7840 × Clark. *Plant Dis.* 101, 1974–1979. doi: 10.1094/PDIS-12-16-1743-RE

Loladze, A., Kthiri, D., Pozniak, C., and Ammar, K. (2014). Genetic analysis of leaf rust resistance in six durum wheat genotypes. *Phytopathology* 104, 1322–1328. doi: 10.1094/PHYTO-03-14-0065-R

Lorieux, M. (2012). MapDisto: fast and efficient computation of genetic linkage maps. *Mol. Breed.* 30, 1231–1235. doi: 10.1007/s11032-012-9706-y

Lowe, I., Cantu, D., and Dubcovsky, J. (2011). Durable resistance to the wheat rusts: integrating systems biology and traditional phenotype-based research methods to guide the deployment of resistance genes. *Euphytica* 179, 69–79. doi: 10.1007/s10681-010-0311-z

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51, 885–895. doi: 10.1038/s41588-019-0381-3

Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., et al. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 13, 648–663. doi: 10.1111/pbi.12288

Maccaferri, M., Sanguineti, M. C., Mantovani, P., Demontis, A., Massi, A., Ammar, K., et al. (2010). Association mapping of leaf rust response in durum wheat. *Mol. Breed.* 26, 189–228. doi: 10.1007/s11032-009-9353-0

Mago, R., Miah, H., Lawrence, G. J., Wellings, C. R., Spielmeyer, W., Bariana, H. S., et al. (2005). High-resolution mapping and mutation analysis separate the rust resistance genes Sr31, Lr26 and Yr9 on the short arm of rye chromosome 1. *Theor. Appl. Genet.* 112, 41–50. doi: 10.1007/s00122-005-0098-9

Mago, R., Spielmeyer, W., Lawrence, G., Lagudah, E., Ellis, J., and Pryor, A. (2002). Identification and mapping of molecular markers linked to rust resistance

genes located on chromosome 1RS of rye using wheat-rye translocation lines. *Theor. Appl. Genet.* 104, 1317–1324. doi: 10.1007/s00122-002-0879-3

Mago, R., Tabe, L., McIntosh, R. A., Pretorius, Z., Kota, R., Paux, E., et al. (2011). A multiple resistance locus on chromosome arm 3BS in wheat confers resistance to stem rust (Sr2), leaf rust (Lr27) and powdery mildew. *Theor. Appl. Genet.* 123, 615–623. doi: 10.1007/s00122-011-1611-y

Marais, G. F., Pretorius, Z. A., Wellings, C. R., McCallum, B., and Marais, A. S. (2005). Leaf rust and stripe rust resistance genes transferred to common wheat from Triticum dicoccoides. *Euphytica* 143, 115–123. doi: 10.1007/s10681-005-2911-6

Marone, D., Del Olmo, A. I., Laidò, G., Sillero, J. C., Emeran, A. A., Russo, M. A., et al. (2009). Genetic analysis of durable resistance against leaf rust in durum wheat. *Mol. Breed.* 24, 25–39. doi: 10.1007/s11032-009-9268-9

McCartney, C. A., Somers, D. J., McCallum, B. D., Thomas, J., Humphreys, D. G., Menzies, J. G., et al. (2005). Microsatellite tagging of the leaf rust resistance gene Lr16 on wheat chromosome 2BSc. *Mol. Breed.* 15, 329–337. doi: 10.1007/s11032-004-5948-7

McIntosh, R., and Dyck, P. L. (1975). Cytogenetical studies in wheat. VII Gene Lr23 for reaction to Puccinia recondita in gabo and related cultivars. *Aust. J. Biol. Sci.* 28, 201–211. doi: 10.1071/BI9750201

McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., Appels, R., and Xia, X. C., *Catalogue of gene symbols for wheat: 2015-2016 supplement*, 2015, Available at: https://shigen.nig.ac.jp/wheat/komugi/genes/macgene/supplement2015.pdf.

McIntosh, R. A., Dubcovsky, J., Rogers, W. J., Morris, C., and Xia, X. C., *Catalogue of gene symbols for wheat: 2017 supplement*, 2017, Available at: https://shigen.nig.ac.jp/wheat/komugi/genes/macgene/supplement2017.pdf.

McIntosh, R. A., Wellings, C. R., and Park, R. F., (1995). *Wheat rusts: an atlas of resistance genes*. Netherlands: Springer. doi: 10.1071/9780643101463

Moore, J. W., Herrera-Foessel, S., Lan, C., Schnippenkoetter, W., Ayliffe, M., Huerta-Espino, J., et al. (2015). A recently evolved hexose transporter variant confers resistance to multiple pathogens in wheat. *Nat. Genet.* 47, 1494–1498. doi: 10.1038/ng.3439

Nsabiyera, V., Qureshi, N., Bariana, H. S., Wong, D., Forrest, K. L., Hayden, M. J., et al. (2016). Molecular markers for adult plant leaf rust resistance gene Lr48 in wheat. *Mol. Breed.* 36, 65. doi: 10.1007/s11032-016-0488-5

Ordoñez, M. E., and Kolmer, J. A. (2007). Virulence phenotypes of a worldwide collection of Puccinia triticina from durum wheat. *Phytopathology* 97, 344–351. doi: 10.1094/PHYTO-97-3-0344

Park, R. F., Mohler, V., Nazari, K., and Singh, D. (2014). Characterisation and mapping of gene Lr73 conferring seedling resistance to Puccinia triticina in common wheat. *Theor. Appl. Genet.* 127, 2041–2049. doi: 10.1007/s00122-014-2359-y

Peterson, R. F., Campbell, A. B., and Hannah, A. E. (1948). A diagrammatic scale for estimating rust intensity on leaves and stems of cereals. *Can. J. Res.* 26c, 496–500. doi: 10.1139/cjr48c-033

Qureshi, N., Bariana, H., Kumran, V. V., Muruga, S., Forrest, K. L., Hayden, M. J., et al. (2018). A new leaf rust resistance gene Lr79 mapped in chromosome 3BL from the durum wheat landrace Aus26582. *Theor. Appl. Genet.* 131, 1091–1098. doi: 10.1007/s00122-018-3060-3

Risk, J. M., Selter, L. L., Krattinger, S. G., Viccars, L. A., Richardson, T. M., Buesing, G., et al. (2012). Functional variability of the Lr34 durable resistance gene in transgenic wheat. *Plant Biotechnol. J.* 10, 477–487. doi: 10.1111/j.1467-7652.2012.00683.x

Roelfs, A. P., Singh, R. P., and Saari, E. E., (1992). *Rust diseases of wheat: concepts and methods of disease management*. Mexico: CIMMYT.

Rosewarne, G. M., Singh, R. P., Huerta-Espino, J., Herrera-Foessel, S., Forrest, K. L., Hayden, M. J., et al. (2012). Analysis of leaf and stripe rust severities reveals pathotype changes and multiple minor QTLs associated with resistance in an Avocet x Pastor wheat population. *Theor. Appl. Genet.* 124, 1283–1294. doi: 10.1007/s00122-012-1786-x

Rosewarne, G. M., Singh, R. P., Huerta-Espino, J., and Rebetzke, G. J. (2008). Quantitative trait loci for slow-rusting resistance in wheat to leaf rust and stripe rust identified with multi-environment analysis. *Theor. Appl. Genet.* 116, 1027–1034. doi: 10.1007/s00122-008-0736-0

Saini, R. G., Kaur, M., Singh, B., Sharma, S., Nanda, G. S., Nayar, S. K., et al. (2002). Genes Lr48 and Lr49 for hypersensitive adult plant leaf rust resistance in wheat (Triticum aestivum L.). *Euphytica* 124, 365–370. doi: 10.1023/A:1015762812907

Schachermayr, G., Siedler, H., Gale, M. D., Winzeler, H., Winzeler, M., and Keller, B. (1994). Identification and localization of molecular markers linked to the Lr9 leaf rust resistance gene of wheat. *Theor. Appl. Genet.* 88, 110–115. doi: 10.1007/BF00222402

Schnurbusch, T., Paillard, S., Schori, A., Messmer, M., Schachermayr, G., Winzeler, M., et al. (2004). Dissection of quantitative and durable leaf rust resistance in Swiss winter wheat reveals a major resistance QTL in the Lr34 chromosomal region. *Theor. Appl. Genet.* 108, 477–484. doi: 10.1007/s00122-003-1444-4

Seyfarth, R., Feuillet, C., Schachermayr, G., Messmer, M., Winzeler, M., and Keller, B. (2000). Molecular mapping of the adult-plant leaf rust resistance gene Lr13 in wheat (Triticum aestivum L.). *J. Genet. Breed.* 54, 193–198.

Seyfarth, R., Feuillet, C., Schachermayr, G., Winzeler, M., and Keller, B. (1999). Development of a molecular marker for the adult plant leaf rust resistance gene Lr35 in wheat. *Theor. Appl. Genet.* 99, 554–560. doi: 10.1007/s001220051268

Singh, A., Pandey, M. P., Singh, A. K., Knox, R. E., Ammar, K., Clarke, J. M., et al. (2013a). Identification and mapping of leaf, stem and stripe rust resistance quantitative trait loci and their interactions in durum wheat. *Mol. Breed.* 31, 405–418. doi: 10.1007/s11032-012-9798-4

Singh, D., Mohler, V., and Park, R. F. (2013b). Discovery, characterisation and mapping of wheat leaf rust resistance gene Lr71. *Euphytica* 190, 131–136. doi: 10.1007/s10681-012-0786-x

Singh, R. P., Huerta-Espino, J., Pfeiffer, W., and Figueroa-Lopez, P. (2004). Occurrence and impact of a new leaf rust race on durum wheat in northwestern Mexico from 2001 to 2003. *Plant Dis.* 88, 703–708. doi: 10.1094/PDIS.2004.88.7.703

Singh, R. P., Mujeeb-Kazi, A., and Huerta-Espino, J. (1998). Lr46: a gene conferring slow-rusting resistance to leaf rust in wheat. *Phytopathology* 88, 890–894. doi: 10.1094/PHYTO.1998.88.9.890

Singh, R. P., Singh, P. K., Rutkoski, J., Hodson, D. P., He, X., Jørgensen, L. N., et al. (2016). Disease impact on wheat yield potential and prospects of genetic control. *Annu. Rev. Phytopathol.* 54, 303–322. doi: 10.1146/annurev-phyto-080615-095835

Singla, J., Lüthi, L., Wicker, T., Bansal, U., Krattinger, S. G., and Keller, B. (2017). Characterization of Lr75: a partial, broad-spectrum leaf rust resistance gene in wheat. *Theor. Appl. Genet.* 130, 1–12. doi: 10.1007/s00122-016-2784-1

Soleiman, N. H., Solis, I., Soliman, M. H., Sillero, J. C., Villegas, D., Alvaro, F., et al. (2016). Short communication: emergence of a new race of leaf rust with combined virulence to Lr14a and Lr72 genes on durum wheat. *Span. J. Agric. Res.* 14, e10SC02. doi: 10.5424/sjar/2016143-9184

Soriano, J. M., and Royo, C. (2015). Dissecting the genetic architecture of leaf rust resistance in wheat by QTL meta-analysis. *Phytopathology* 105, 1585–1593. doi: 10.1094/PHYTO-05-15-0130-R

Suenaga, K., Singh, R. P., Huerta-Espino, J., and William, H. M. (2003). Microsatellite markers for genes Lr34/Yr18 and other quantitative trait loci for leaf rust and stripe rust resistance in bread wheat. *Phytopathology* 93, 881–890. doi: 10.1094/PHYTO.2003.93.7.881

Terracciano, I., Maccaferri, M., Bassi, F., Mantovani, P., Sanguineti, M. C., Salvi, S., et al. (2013). Development of COS-SNP and HRM markers for high-throughput and reliable haplotype-based detection of Lr14a in durum wheat (Triticum durum Desf.). *Theor. Appl. Genet.* 126, 1077–1101. doi: 10.1007/s00122-012-2038-9

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

William, H. M., Singh, R. P., Huerta-Espino, J., Palacios, G., and Suenaga, K. (2006). Characterization of genetic loci conferring adult plant resistance to leaf rust and stripe rust in spring wheat. *Genome* 49, 977–990. doi: 10.1139/g06-052

Winfield, M. O., Allen, A. M., Burridge, A. J., Barker, G. L. A., Benbow, H. R., Wilkinson, P. A., et al. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnol. J.* 14, 1195–1206. doi: 10.1111/pbi.12485

Wu, T. D., and Watanabe, C. K. (2005). GMAP: a genomic mapping and alignment program for mRNA and EST sequences. *Bioinformatics* 21, 1859–1875. doi: 10.1093/bioinformatics/bti310

Wu, Y., Bhat, P. R., Close, T. J., and Lonardi, S. (2008). Efficient and accurate construction of genetic linkage maps from the minimum spanning tree of a graph. *PLoS Genet.* 4, e1000212. doi: 10.1371/journal.pgen.1000212

Xing, L., Wang, C., Xia, X., He, Z., Chen, W., Liu, T., et al. (2014). Molecular mapping of leaf rust resistance gene LrFun in Romanian wheat line Fundulea 900. *Mol. Breed.* 33, 931–937. doi: 10.1007/s11032-013-0007-x

Yates, F. (1934). Contingency tables involving small numbers and the $\chi 2$ test. Supplement *J. R. Stat. Soc.* 1, 217–235. doi: 10.2307/2983604

Zhang, P., Hiebert, C. W., McIntosh, R. A., McCallum, B. D., Thomas, J. B., Hoxha, S., et al. (2016). The relationship of leaf rust resistance gene Lr13 and hybrid necrosis gene Ne2m on wheat chromosome 2BS. *Theor. Appl. Genet.* 129, 485–493. doi: 10.1007/s00122-015-2642-6

Zhang, P., Zhou, H., Lan, C., Li, Z., and Liu, D. (2015). An AFLP marker linked to the leaf rust resistance gene LrBi16 and test of allelism with Lr14a on chromosome arm 7BL. *Crop J.* 3, 152–156. doi: 10.1016/j.cj.2014.11.004

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 October 2019
doi: 10.3389/fpls.2019.01244

Novel Informatic Tools to Support Functional Annotation of the Durum Wheat Genome

*Mario Fruzangohar[1], Elena Kalashyan[1], Priyanka Kalambettu[1], Jennifer Ens[2], Krysta Wiebe[2], Curtis J. Pozniak[2], Penny J. Tricker[1] and Ute Baumann[1]**

[1] School of Agriculture, Food and Wine, The University of Adelaide, Adelaide, SA, Australia, [2] Department of Plant Sciences and Crop Development Centre, University of Saskatchewan, Saskatoon, SK, Canada

Seed mutagenesis is one strategy to create a population with thousands of useful mutations for the direct selection of desirable traits, to introduce diversity into varietal improvement programs, or to generate a mutant collection to support gene functional analysis. However, phenotyping such large collections, where each individual may carry many mutations, is a bottleneck for downstream analysis. Targeting Induced Local Lesions in Genomes (TILLinG), when coupled with next-generation sequencing allows high-throughput mutation discovery and selection by genotyping. We mutagenized an advanced durum breeding line, UAD0951096_F2:5 and performed short-read (2x125 bp) Illumina sequencing of the exome of 100 lines using an available exome capture platform. To improve variant calling, we generated a consolidated exome reference using the recently available genome sequences of the cultivars Svevo and Kronos to facilitate the alignment of reads from the UAD0951096_F2:5 derived mutants. The resulting exome reference was 484.4 Mbp. We also developed a user-friendly, searchable database and bioinformatic analysis pipeline that allowed us to predict zygosity of the mutations discovered and extracts flanking sequences for rapid marker development. Here, we present these tools with the aim of allowing researchers fast and accurate downstream selection of mutations discovered by TILLinG by sequencing to support functional annotation of the durum wheat genome.

Keywords: exome capture, mutagenesis, reverse genetics, durum wheat, polyploidy, TILLinG

OPEN ACCESS

Edited by:
Roberto Papa,
Marche Polytechnic University, Italy

Reviewed by:
Francesca Taranto,
Council for Agricultural and Economics Research, Italy
Matteo Dell'Acqua,
Sant'Anna School of Advanced Studies, Italy

****Correspondence:***
Ute Baumann
ute.baumann@adelaide.edu.au

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *08 June 2019*
Accepted: *06 September 2019*
Published: *10 October 2019*

Citation:
Fruzangohar M, Kalashyan E, Kalambettu P, Ens J, Wiebe K, Pozniak CJ, Tricker PJ and Baumann U (2019) Novel Informatic Tools to Support Functional Annotation of the Durum Wheat Genome. Front. Plant Sci. 10:1244. doi: 10.3389/fpls.2019.01244

INTRODUCTION

Mutants are valuable tools for the identification and functional analysis of genes. Mutations can arise spontaneously or can be induced physically (e.g., radiation), chemically (e.g., alkylating agents), and by transposon insertions or through gene editing, such as with the CRISPR/Cas9 system for specific modifications of target genes (Adli, 2018).

The use of chemical mutagenesis has had a renaissance with the development of Targeting Induced Local Lesions in Genomes (TILLinG) method in *Arabidopsis* (McCallum et al., 2000). TILLinG is a high-throughput method of inducing and identifying genetic variations in target genes. Its main advantage is that it can be employed as a functional genomics platform for virtually any species, independent of genome size and ploidy. It is hence not surprising that TILLinG populations have been generated for various animal and plant species as described in (Kurowska et al., 2011).

While a range of methods have been developed for mutation detection in a TILLinG population (Yang et al., 2000; Colbert et al., 2001; Caldwell et al., 2004; Till et al., 2006; Raghavan et al., 2007; Suzuki et al., 2008; Dong et al., 2009), most were designed for the identification of mutations in a relatively small set of genes and become costly and labor-intensive when scaled to hundreds of genes. While pooling strategies (Tsai et al., 2011; Chi et al., 2014) combined with next-generation sequencing (NGS) have increased the number of genes (amplicons) that can be interrogated simultaneously, the background error rate can be high due to the numerous PCR steps in the protocol. In polyploid species, the presence of homeologs can additionally lead to false negatives and the interpretation of the sequence data may require customized bioinformatics pipelines (Tsai et al., 2011).

An alternative approach is to integrate NGS with capture methodologies. Saintenac et al. (2011) demonstrated that sequencing of DNA targeting non-repetitive genic regions can be highly reproducible and region-/locus-specific which can allow large-scale variant discovery in tetraploid wheat. Since sequence-capture methodologies offer the possibility of restricting sequencing to the coding portion of the genome, i.e., the exome (Winfield et al., 2012; Allen et al., 2013), they are especially suited to species with large or highly repetitive genomes, like wheat, where whole-genome sequencing would be excessively expensive (Tucker et al., 2011; Henry et al., 2014).

Exome capture probe design requires knowledge of the gene sequences preferably from full-genome assemblies. However, with a total of ~16 Gbp for bread wheat and ~11 Gbp for durum wheat, the wheat genome is one of the largest in the grass family, and full-genome assemblies of hexaploid and tetraploid wheat have only recently been released (IWGSC, 2018; Maccaferri et al., 2019). Therefore, all commercially available exome capture platforms were developed from wheat gene sequences in public databases such as NCBI and TriFL-DB (RIKEN) and EST and transcriptome assemblies. This carries the risk of underrepresenting low abundance genes and tightly regulated gene family members. Since exome capture is a hybridization process, not only will (near) identical sequences be captured but also non-target sequences (also called off-target reads) depending on the probes' lengths and GC contents (Asan et al., 2011; Chilamakuri et al., 2014). Off-targets can include adjacent intronic regions, closely related genes, or homeologous sequences. Without an annotated reference sequence or knowledge of the complete gene set of an organism, these off-target sequences may be mistaken for allelic variants of a target gene. Thus, the potential for off-target alignments must be considered during the analysis and interpretation of mutant read alignments to mitigate false-positive mutation calls.

We developed a TILLinG population suitable for southern Australian environmental conditions. We chose an advanced spring-habit breeding line semi-dwarf tetraploid durum wheat (vernalization- and photoperiod-insensitive) which yields well in southern Australia and has given rise to the commercially grown cultivar DBA-Aurora. We used a subset of the population, 99 M_2 plants, for an exome capture experiment using the Roche NimbleGen Wheat Exome Design. To overcome the complications caused by the Roche NimbleGen incomplete reference sequence for read alignment, such as missing homeologs, gaps, undefined nucleotides (i.e., "N"), and presence of homopolymer artifacts, we devised a novel method to construct a suitable reference sequence for mutation calling. We developed a bioinformatics pipeline for mutation calling and a web client application for querying and retrieval of mutation information.

MATERIALS AND METHODS

Plant Material

Approximately 2,000 seeds from three individual plants of an advanced *Triticum turgidum* durum F2:5 breeding line (ex:UAD0951096 with the pedigree Tamaroi*2/Kalka//RH920318/Kalka///Kalka*2/Tamaroi) were mutagenized with 0.7% ethyl methanesulfonate (EMS) by gentle agitation in the solution on an orbital shaker overnight (18 hr) as described by Dong et al. (2009). Following rinsing, four seeds per pot were sown in 12-cm pots filled with coco peat with additional slow release fertilizer. After 20 days, when 76% of seeds had germinated, plants were thinned to one plant per pot in order to obtain a population of 500 mutant plants. Main spikes were isolated in bags pre-anthesis to ensure self-pollination. At full maturity, seeds were harvested separately from each mutant plant.

DNA Isolation and Exome Capture

DNA Isolation and Exome Capture was extracted from a single 2-week old seedling of 99 randomly chosen M2 mutant plants and the unmutagenized control as described by Pallotta et al. (2000). Library preparation and hybridization followed Jordan et al. (2015) with modifications. Briefly, 1mg gDNA was fragmented by sonification to an average fragments length of 300bp. Illumina TruSeq libraries were prepared with fragmented DNA, indexed, size-selected, and pooled (n = 6) for exome capture. Pooled libraries were hybridized using the Roche's NimbleGen wheat exome capture design (120426_Wheat_WEC_D02) (Roche) and protocol as described in Jordan et al. (2015).

Building a Durum Exome Reference Sequence

Available genome sequences for the two tetraploid durum wheat cultivars Kronos ('Kronos EI v1') and Svevo (Maccaferri et al., 2019) were used for read alignment. In total, 245M paired-end reads of the unmutagenized control line UAD0951096_F2:5 were processed including adapter and quality trimming by Trimmomatic 0.36 (Bolger et al., 2014) using the following parameters: ILLUMINACLIP : TruSeq3-PE.fa:2:30:10 and LEADING:22 TRAILING:22 SLIDINGWINDOW:4:15 MINLEN:50.

The resulting reads were aligned using BioKanga version 4.3.6 (Stephen, 2012) (align –pemode 1 -s 2) to the Svevo pseudomolecules allowing for a 2% mismatch rate and no gaps. This resulted in 60% mapped reads. Since the genomic annotation for Svevo was not available when we conducted the project, we developed an in-house Java application (https://github.com/CroBiAd/TILLinG-mutants) for the retrieval of coding regions by making use of the coverage depth of aligned

reads as an indicator. Start and end positions of genomic regions with a coverage of 17 reads or more were firstly marked and subsequently retrieved together with 500-bp flanking sequences on either side. The reason to add these tails was the observation that coverage never drops abruptly at the intron–exon boundaries of exome captured aligned reads. If two regions with high coverage were in close proximity, i.e., less than 301 bp, they were merged (**Supplementary Figure 1**). The resulting 191,892 contigs covered a total length of 443 Mbp.

In the second step, reads that did not map to the Svevo genome (97 M) were aligned to Kronos by BioKanga as above resulting in 10.6% mapped reads. Regions were retrieved as described above. In the third step, the remaining 86.7 million unaligned reads were assembled using ABySS version 2.0.2 (Jackman et al., 2017) with k-mer size = 96. We selected contigs with a minimum length of 500 bp resulting in 552 contigs with a total length of 420 Kbp. We performed BlastX searches to explore which proteins might potentially be encoded by the 552 assembled contigs against rice (MSU Rice Genome Annotation Project Release 7) (Kawahara et al., 2013) and *Arabidopsis* (TAIR10) (Berardini et al., 2015) protein sequence databases (e-value cutoff 10^{-5}).

Combining the three sets of contigs (i.e., from Svevo, Kronos, and the ABySS assembly) gave us our 484.4 Mbp reference sequence for read alignment and mutation detection, hereafter called DECaR (DurumExomeCaptureReference). DECaR can be downloaded from doi: 10.25909/5d258fa699358

Read Alignment to Decar and Mutation Calling Pipeline

Following quality and adapter trimming, exome captured reads (minimum 50 bp) from unmutagenized control line, and the M2 lines were aligned to DECaR using Bowtie 2 version 2.3.0 (Langmead and Salzberg, 2012) allowing a 2% mismatch rate with the following parameters: –end-to-end –very-sensitive –n-ceil L,0,0.1 –rdg 3,3 –rfg 3,3 –no-unal –mp 6,6 –np 4 –no-mixed -score-min L,0,-0.12

After alignment PCR duplicates were detected and removed from BAM files using our in-house Java application.

One pileup file was generated from the bam files using SAMtools version 1.6 (Li et al., 2009) with a minimum mapping quality (MAPQ) of 2 to mitigate multi-mapping and mapping errors.

We used three criteria to identify mutations in the TILLinG population. First, any variation from DECaR was considered a potential mutation if it was present in only one mutant sample and non-polymorphic between the control line and DECaR. Secondly, we demanded a mutation to be covered by at least three reads to be confident that the mutation was not derived from sequencing error. Finally, because coverage at a reference position was variable from sample to sample, a mutation was only called in a mutant sample if we had sufficient coverage for the control allele in at least 50 other mutant samples. An in-house developed Java application was used to implement this logic

Initially, 9.5M mutations were called across all the 81 mutant samples that had sufficient coverage. In order to reduce false-positive calls, we applied two conditions that had been used in previous studies (Henry et al., 2014; King et al., 2015). Firstly, EMS preferentially changes C- > T and G- > A. It has been shown that the higher percentage of CG- > TA transitions in EMS-induced mutant populations was associated with better mutation calling (Henry et al., 2014; King et al., 2015). Secondly, we expected a ratio of 2:1 heterozygous to homozygous mutations in M_2 populations (Henry et al., 2014).

Database

To make the results easily accessible, we created a Web application for querying the mutations.

First, details of all detected mutations (i.e., position, zygosity, flanking sequence) were deposited into an SQLite database. Then, the stand-alone version of BLAST© Command Line Application (ncbi-blast-2.7.1+) (Camacho et al., 2009) was installed locally, and a nucleotide BLAST database was generated from the DECaR. Next, an ASP.Net 4.6 Web client, published on Microsoft Internet Information Services (IIS), was developed to allow BLAST searches of the DECaR and querying of the mutations. Finally, the complete application, named Durum Wheat TILLinG (DuWTill) was hosted on Microsoft Windows Server© 2016 Standard edition and is publicly accessible at http://duwtill.acpfg.com.au/.

In addition, DuWTill is available for download through GitHub (https://github.com/CroBiAd/DuWTill) where steps to build it locally are described. After installation DuWTill can be run locally either with our data or on data sets prepared by researchers from their own populations. The distribution is provided for the Windows operating system, which requires Windows IIS to be turned on and Visual Studio 2015 or later (Microsoft) installed on the development computer.

RESULTS

The Novel Durum Exome Reference Sequence DECaR

We mapped reads from the unmutagenized control line to the two publically available durum wheat genome assemblies [Svevo (Maccaferri et al., 2019) and Kronos ('Kronos EI v1')]. By combining these aligned read data with contigs assembled from unmapped reads, we constructed a new durum exome reference, DECaR, which consists of 220,114 contigs with a total length of 484,479,862 bp covering ca. 4% of the estimated 11-Gbp durum wheat genome. A comparison of alignment rates of the control sample to the NimbleGen reference and DECaR showed 20% alignment *versus* 51%, respectively. We also observed an increase in average alignment quality (MAPping Quality, MAPQ, SAMtools (Li et al., 2009)) from 28.75 to 29.74.

JBrowse was used to visually compare the alignment of the unmutagenized control sample reads to the original Roche NimbleGen exome reference and to DECaR, respectively. **Figure 1A** shows reads aligned to contig05736 of the original Roche exome reference. In this example, it is clear that, within the 1.37-kbp region, there are several putatively mutated/polymorphic positions (depicted as colored bars in the coverage track and indicated by black triangles). **Figures 1B**, **C** show alignments of the same reads to the corresponding regions located on DECaR

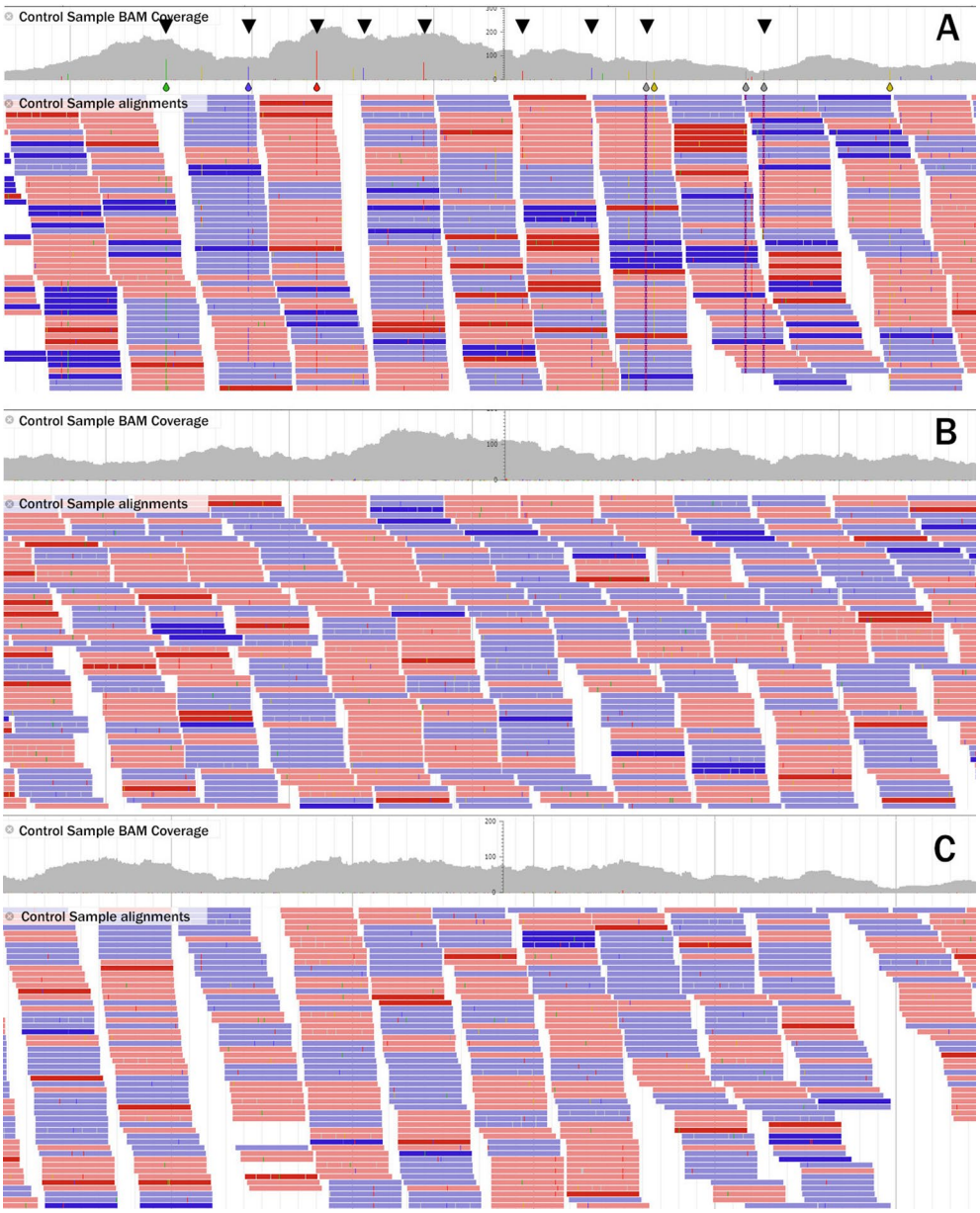


FIGURE 1 | Example of read alignment of the control line to the a 1.37-Kbp region of **(A)** the Roche NimbleGen exome reference contig05736, **(B)** to the corresponding regions of contig05736 in DECaR originating from Svevo chromosome 3A, and **(C)** Svevo chromosome 3B. Reads in red align to the (+) strand, those in blue to the (-) strand. Location of potential mutations/polymorphisms are indicated by blue (cytosine), green (adenine), red (thymine), and yellow (guanine) bars for the DNA base called and highlighted by black triangles in the coverage track in A; in B and C, no mutated bases were called.

contigs derived from chromosomes 3A and 3B, respectively. No mutated positions are visible. This example demonstrated the advantage of DECaR to position reads properly to the A and B chromosomes, whereas alignments of the same reads to the NimbleGen exome capture reference created false positives.

Mutations' Discovery

The average read coverage per base position in all samples was estimated, and results are given in **Supplementary Table 1**. Coverage within mutant samples ranged from 1.2 to 11.3 with an average of 6 (reads/base position). Alignment rates ranged from 13 to 76% with an average of 57%. In order to find the reason for a low alignment rate for some samples, we selected the three samples with the lowest alignment percentage (673, 677, 661) and mapped their reads to the entire Svevo reference genome. Surprisingly, alignment rates increased to 77, 81, and 80% respectively. Closer inspection indicated that these samples contained a significant amount of non-exonic DNA; however, they nevertheless showed sufficient coverage in the coding regions to be included. On the other hand, we excluded

samples with low coverage either due to lower exome capture efficiency or high PCR duplication rates. In summary, for 18 of the 99 mutant samples, the data failed to be of sufficient quality to proceed; therefore, these were excluded from further analysis (**Supplementary Table 1**).

To reduce false-positive calling of mutations, we gradually increased the minimum number of reads confirming a mutated base as shown in **Figure 2**. **Table 1** shows the results for a minimum coverage of 10, for which 83,573 mutations were called (49,652 heterozygous, 33,921 homozygous) of which 94% were of CG- > TA type.

Since coverage varied from sample to sample, the number of detected mutations per sample ranged from 12 (in mutant sample 653) to 2,603 (in mutant sample 417) (see **Supplementary Table 1**). The mutation rate among the 81 samples varied from 2.4 to 20.7 mutations/Mbp, with an average of 10.3 (derived by dividing the total number of mutations by number of positions that are covered by 10 reads or more).

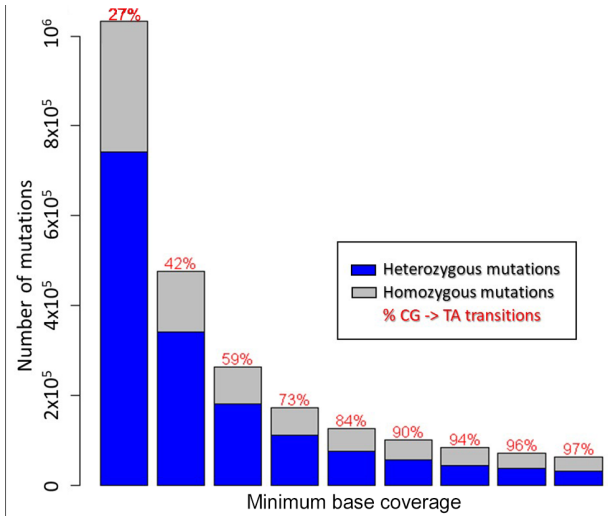


FIGURE 2 | Number of mutations and their zygosity depending on mutant allele coverage.

TABLE 1 | Types of mutations and their frequencies detected in the 81 mutant lines for a minimum base coverage of 10.

Mutation	Base	Number
Deletion	A	19
Deletion	C	15
Deletion	G	23
Deletion	T	20
Substitution	A- > C	53
Substitution	A- > G	285
Substitution	A- > T	277
Substitution	C- > A	1,705
Substitution	C- > G	51
Substitution	C- > T	39,294
Substitution	G- > A	39,472
Substitution	G- > C	34
Substitution	G- > T	1,610
Substitution	T- > A	385
Substitution	T- > C	240
Substitution	T- > G	90
SUM		83,573

The unmutagenized control sample was sequenced to a higher depth, and average base coverage (29.5 reads/base position) was deeper (>3 times) than that of mutant samples (**Supplementary Table 1**). The higher coverage of the control sample helped us to distinguish SNPs specific to the line from true EMS-derived mutations.

Svevo is an Italian durum wheat cultivar derived from crossing CIMMYT selection with Zenit in the 1990s. Kronos, on the other hand, was released by Arizona Plant Breeders in 1992 and is derived from a male-sterile-facilitated recurrent selection population (APB MSFRS Pop, selection D03–21) (Jackson, 2011; Berg, 2014). Whereas the advanced breeding line used in our study has Australian cultivars Kalka adapted to the boron-toxic soil of Southern Australia and Tamoroi in its pedigree. A recent study into genetic diversity across durum wheat by Kabbaj et al., 2017 shows that the Australian cultivars are distinct from Kronos and Svevo. It is therefore not surprising that we not only saw varietal SNPs but also differences in gene content between the accessions. One example of gene families that rapidly evolved is the NBS-LRR disease resistance genes (Steuernagel et al., 2016). These tend to vary significantly between elite cultivars due to selective breeding. Indeed, we found that 41 of the assembled contigs showed homology to disease resistance genes, but there were also members of the cytochrome P450 and oxidoreductase families (**Supplementary Table 2**).

The DuWTill Database

We deposited all identified mutations into a database and developed the online tool DuWTill for access to the collection. DuWTill is publicly available at http://duwtill.acpfg.com.au/.

DuWTill's intuitive interface has principally one main 'Search' page (**Figure 3**) where the mutations table is displayed. The database can be searched by two types of identifiers (with restriction on region, if preferred):

1. Contig ID (using DECaR nomenclature) to obtain mutations for all mutant lines occurring within a specific contig.
2. Mutant ID (individual mutant plants) to get mutations on all contigs for a specific mutant line.

The output table contains one row for each mutation found and the mutation position (in bases) relative to the start of the respective DECaR contig, induced mutation type, base call in the non-mutagenized control and in the EMS-mutagenized individual. It also includes predicted zygosity, chromosome location, and mutant allele coverage as a measure of confidence that the mutation has been called correctly. Clicking on the flanking sequence link will expand the sequence fragment with minimum 50 to 200 bp on either side of the putative mutation.

Alternatively, the DECaR can be queried with a FASTA-formatted sequence of interest using the internal BLAST portal on a separate utility "BLAST" page (**Figure 4**). The top hits to the available reference will be displayed. Selecting a hit will redirect to the "Search" page showing all putative mutations in

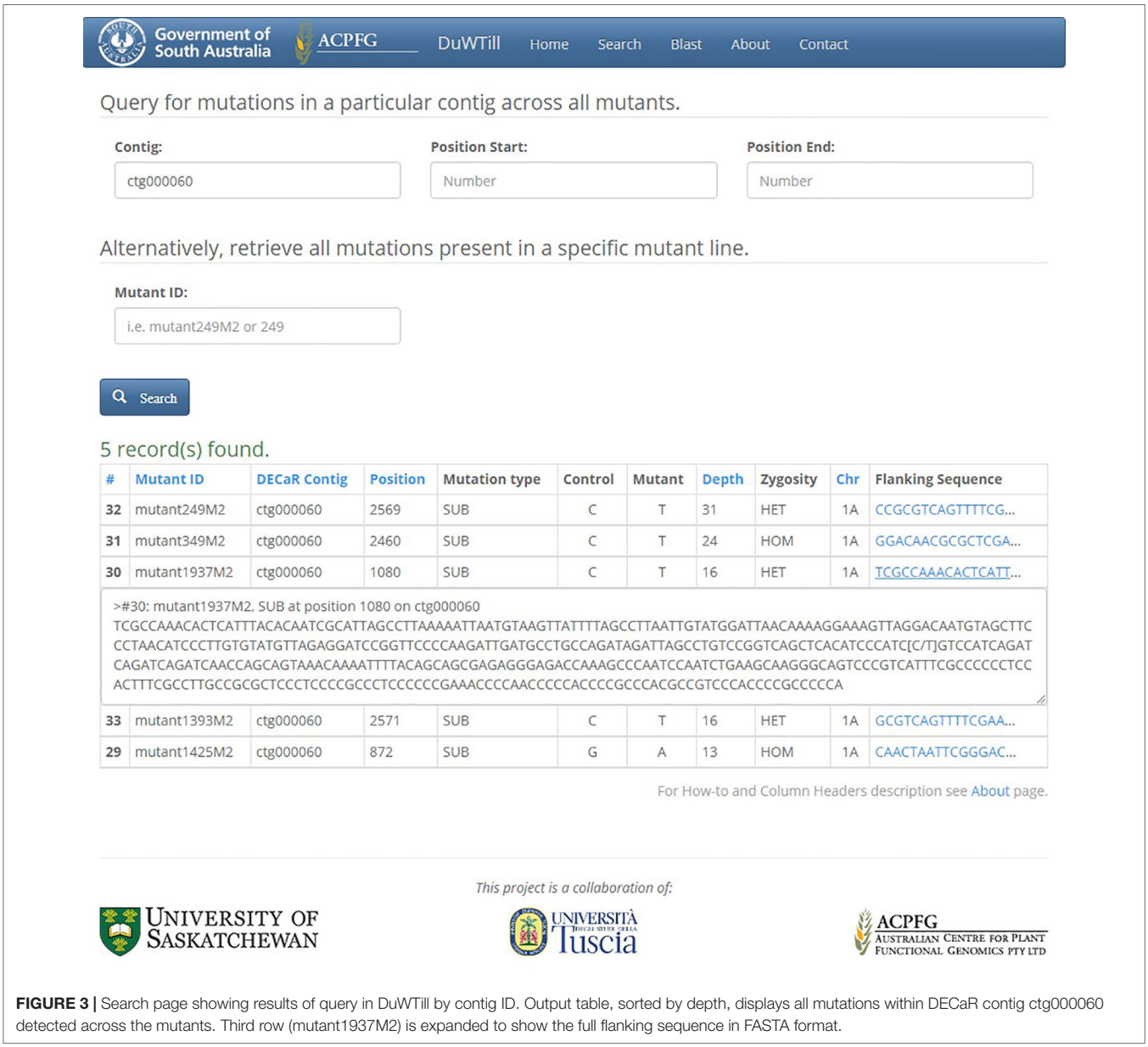


FIGURE 3 | Search page showing results of query in DuWTill by contig ID. Output table, sorted by depth, displays all mutations within DECaR contig ctg000060 detected across the mutants. Third row (mutant1937M2) is expanded to show the full flanking sequence in FASTA format.

the sequence of interest called within a contig for all mutants of the population.

In addition, short background information, how-to instructions and a description of the table headers are presented on the "About" page.

DuWTill source code and mutation call table were deposited in a public repository on GitHub (https://github.com/CroBiAd/DuWTill). It also includes a console application, YourDB, which has been written to help with formatting the mutation calls into an SQLite database.

The DuWTill application can be used for TILLinG data from any organism and is also independent of the mutation calling pipelines. YourDB application will format any comma separated value (.csv) file which contains the appropriate fields into an SQLite database. Obviously, if the TILLinG data does not make use of DECaR, the exome reference BLAST database needs to be substituted with an appropriate one. The GitHub repository contains all necessary instructions.

DISCUSSION

We tested the suitability of the Roche NimbleGen wheat exome capture platform for mutation discovery in a subset of M2 plants generated with EMS of an Australian adapted breeding line. As researchers before us had, we faced the challenge of developing a mutation calling pipeline that would maximize true positive calls and minimize false positives but not result in missing many potentially interesting mutations. For example, Henry et al. (2014) applied their bioinformatics pipeline (MAPS) which was

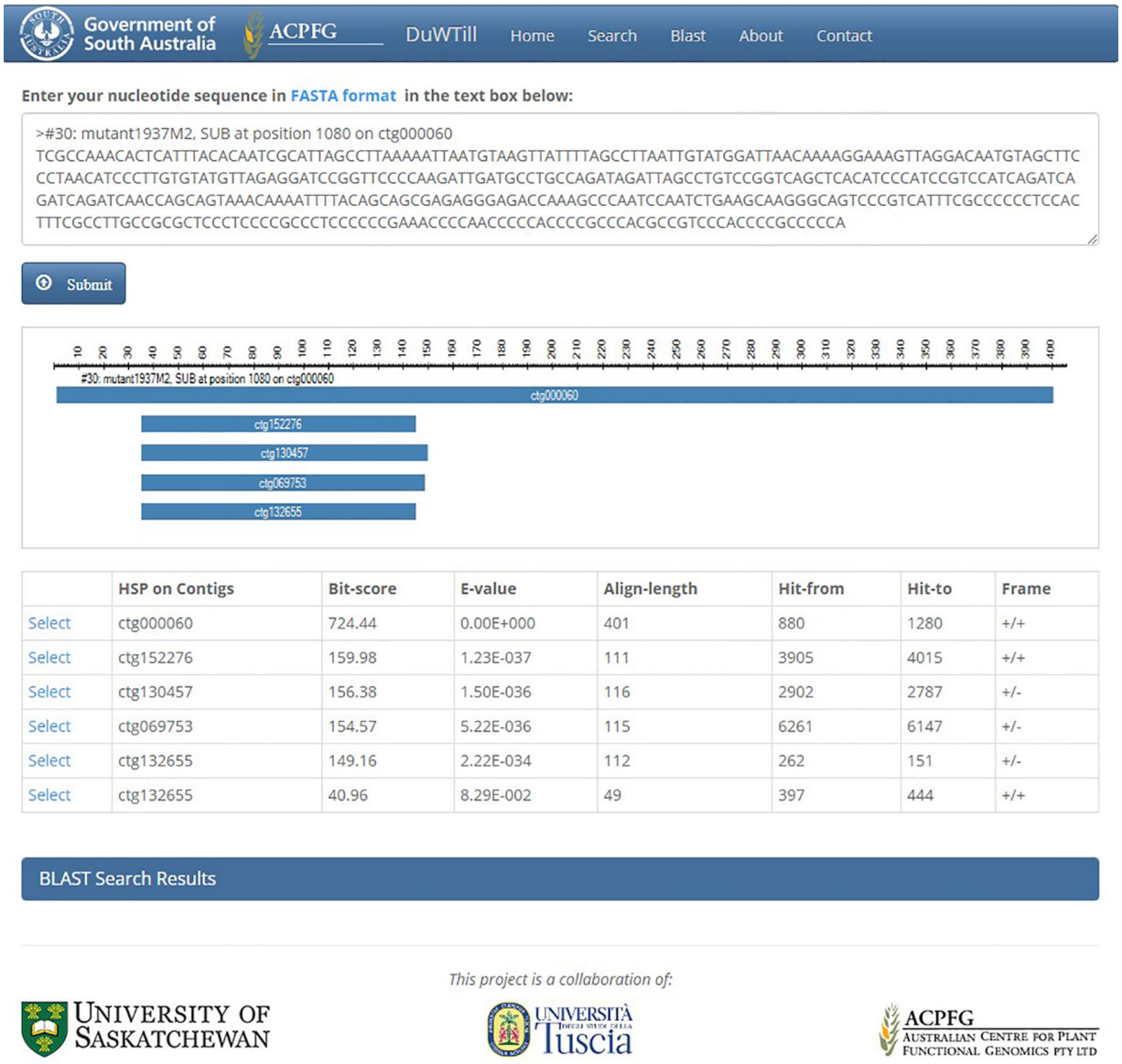


FIGURE 4 | BLAST page showing the results of a BLASTn search query in DuWTill. BLASTn search was executed using the same sequence that was presented in the table of **Figure 3**. Top hits to DECaR are shown in two views: alignment overview and hit table. Detailed HSP alignments are reported in the additional expandable panel "BLAST Search Results."

initially designed for detection of mutations in EMS-mutagenized rice to exome capture data of the Kronos and six tetraploid M2 wheat plants derived from Kronos. The researchers used the Roche NimbleGen exome capture reference contigs for alignment of the captured reads using the short-read aligner BWA-SW. A variant was called a mutation if it was present in only one of the samples and absent in Kronos, based on the assumption that the probability of the same mutation appearing in two independent plants is low. The applied minimum read coverage was set to seven for heterozygous and five for homozygous mutations and resulted in more than 90% of CG- > TA mutation rate. Setting the minimum coverage threshold to a higher value will decrease the false-positive rate but can simultaneously also reduce the number of true mutations called. In other words, the mutations with lower coverage are not all false. For our dataset and using the DECaR reference, we chose a minimum coverage of at least 10 reads per base position based on 94% of mutations called being CG- > TA transitions. By settling on a minimum coverage of 10 reads per base position for a mutation call, we erred on the side of caution in order to keep false positives low. This may explain the comparatively lower average mutation rate of 10.3/Mbp and a heterozygous to homozygous ratio of 1.47. The stringency of mutation calling can be adjusted by for example choosing a lower read depth/base position and a different minor allele ratio when running our Java script. Thus, researchers have the flexibility to analyze the data to what is most appropriate for their needs.

Wheat exome capture designs have developed along with improved sequence and genome assembly knowledge. King et al. (2015) used a custom designed capture array with 1,846 full-length cDNAs (approximately 2 Mbp capture space) for targeted sequencing to examine *TaGA20ox1* homeologs across three Cadenza bread wheat EMS-induced M5 mutant lines. The sequence reads were aligned to the IWGSC Chinese Spring–derived Chromosome Survey Sequences CSS (IWGSC, 2014). They obtained a relatively low alignment rate of 26% to target genes compared to similar targeted capture experiments. Despite low alignment rates, by filtering based on CG- > TA mutation rates and expected hetero-homozygous ratios, King et al. (2015) were able to validate 75–80% mutations called.

Another significant wheat EMS-induced exome capture experiment was performed by Krasileva et al. (2017). In that study, first, a new and improved whole-exome capture design was developed that targeted 84 Mbp sequence space. One thousand and three hundred thirty-five EMS-induced M_2 mutants of the tetraploid cultivar Kronos and 1,200 mutants of the hexaploid, bread wheat cultivar Cadenza were exome captured and sequenced. One hundred base pairs of paired-end reads were aligned to A and B genome contigs of the CSS (IWGSC, 2014). Similarly to our study, Krasileva et al. (2017) then improved their reference in order to improve alignment. They did this by assembling the remaining unmapped reads from their samples to expand their sequence-capture space by an additional 33.4 Mbp for the durum, i.e., 117.4 Mbp in total.

With the availability of two full durum wheat reference sequences and the recently released assembly (Maccaferri et al., 2019), we were able to create a new reference specifically for this durum wheat exome capture experiment. By combining regions from the Svevo and Kronos genomes (where reads from the control line aligned) with contigs assembled from unaligned reads, we created a 484.4 Mbp new reference which was more than four times larger than the original Roche reference (106.9 Mbp) (Roche). The process of first aligning reads from our deeper-sequenced unmutagenized control sample also allowed us to distinguish varietal SNPs confidently. This approach could be used for any new TILLinG population where knowledge of the complete gene set is not known.

King et al. (2015) also demonstrated that absence of one or two copies of a gene in the wheat reference could cause homozygous mutations to be erroneously called heterozygous, because reads containing the mutated position were diluted by wild-type reads. A true mutation located in a homeolog that is not represented in the reference sequences can lead to its assignment to the wrong homeolog (off-target homeolog). Roche's NimbleGen Exome capture reference is mostly homeolog-insensitive and was designed by including sequences from multiple hexaploid wheat varieties generated by different sequencing technologies. The advantage of using DECaR over the NimbleGen reference is: having a durum-based reference, including absent homeologs, an improved reference sequence quality (e.g., removal of homopolymer errors, and inclusion of intronic regions for better mapping); to be as inclusive as possible (i.e., include genes that were captured but not represented in the original reference); and finally to keep the alignment space small for ease of computation. Following this adjustment, the mutation rate was estimated to be 20.1 mutations/Mbp, consistent with a previously reported mutation rate (Uauy et al., 2009).

The DuWTill database was developed as a tool to mine for mutants of interest following exome capture. We required an intuitive interface for collaborators to query the data and obtain sufficient information for follow-up work such as primer design to test for the presence of the mutation of interest. Until very recently, no such tools existed and especially not for durum wheat.

DuWTill application was designed to accommodate information on a large number of individuals and their mutations and is easily adaptable to other organisms than durum wheat. It is a small and simple tool which can be easily installed locally on any windows platform even a laptop, or can be run as an open web service application.

For wheat, DuWTill is comparable with the established and widely used database at wheat-tilling.com which houses mutants' and mutations' information from Krasileva et al. (2017). The wheat-tilling database additionally incorporates useful mutation effects and oligo primer designs where these have been predicted or tested, whereas DuWTill does not. The DuWtill interface has been designed to be simple, portable, and user-friendly and displays flanking sequence with the mutation in position for primer design on the same page as all other information. However, we believe its main advantage is the ability to readily update the reference which should continue to make it an effective tool for mining variant information for the future.

CONCLUSIONS

We have optimized a reference sequence for tetraploid wheat to use with the Roche Wheat Exome Capture Design for diversity and mutation studies. Furthermore, we have developed a bioinformatics pipeline for the analysis of TILLinG mutants in conjunction with the new reference and have called mutations for a subset of an Australian durum TILLinG population. A software application has been written that allows online or local interrogation of the TILLinG collection and can also be used to host propriety data.

All resources are publically available to interested researchers and can be adapted to their needs.

DATA AVAILABILITY STATEMENT

The sequence data generated for this study can be found in SRA under https://dataview.ncbi.nlm.nih.gov/object/PRJNA574238?reviewer=lbstjt0r17312aeo6n7nqjdtv6

AUTHOR CONTRIBUTIONS

MF analysed the exome capture data and wrote the in house Java application. EK designed and programmed the DuWTill. PK isolated the DNA; JE and KW made the libraries and performed exome capture experiments. UB, MF, EK, PT, and CP wrote the manuscript; CP, PT, and UB designed the study.

FUNDING

This project was supported by a Premier's Research and Industry Fund grant (no. IRGP15) provided by the Government of South Australia Department of State Development.

ACKNOWLEDGMENTS

We thank A/Prof. Jason Able (University of Adelaide and Durum Breeding Australia - Southern node) for providing seeds of the breeding line UAD0951096 used to create the TILLinG population and for his early input. We also thank the Durum Wheat Sequencing Consortium for pre-publication access to the reference whole-genome sequence of Svevo.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01244/full#supplementary-material

SUPPLEMENTARY TABLE 1 | Detailed information on alignment rates and mutations for each Tilling line.

SUPPLEMENTARY TABLE 2 | BLAST search results for the assembled contigs.

REFERENCES

Adli, M. (2018). The CRISPR tool kit for genome editing and beyond. *Nat. Commun.* 9, 1911. doi: 10.1038/s41467-018-04252-2

Allen, A. M., Barker, G. L., Wilkinson, P., Burridge, A., Winfield, M., Coghill, J., et al. (2013). Discovery and development of exome-based, co-dominant single nucleotide polymorphism markers in hexaploid wheat (Triticum aestivum L.). *Plant Biotechnol. J.* 11, 279–295. doi: 10.1111/pbi.12009

Asan, Xu, Y., Jiang, H., Tyler-Smith, C., Xue, Y., Jiang, T., et al. (2011). Comprehensive comparison of three commercial human whole-exome capture platforms. *Genome Biol.* 12, R95. doi: 10.1186/gb-2011-12-9-r95

Berardini, T. Z., Reiser, L., Li, D., Mezheritsky, Y., Muller, R., Strait, E., et al. (2015). The Arabidopsis information resource: making and mining the "gold standard" annotated reference plant genome. *Genesis* 53, 474–485. doi: 10.1002/dvg.22877

Berg, Jim. (2014). *North American wheat varieties released since 1990*. Montana State University. http://plantsciences.montana.edu/foundationseed/quickfacts/Wheat%20Varieties%201990-.pdf

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Caldwell, D. G., McCallum, N., Shaw, P., Muehlbauer, G. J., Marshall, D. F., and Waugh, R. (2004). A structured mutant population for forward and reverse genetics in Barley (Hordeum vulgare L.). *Plant J.* 40, 143–150. doi: 10.1111/j.1365-313X.2004.02190.x

Camacho, C., Coulouris, G., Avagyan, V., Ma, N., Papadopoulos, J., Bealer, K., et al. (2009). BLAST+: architecture and applications. *BMC Bioinform.* 10, 421. doi: 10.1186/1471-2105-10-421

Chi, X., Zhang, Y., Xue, Z., Feng, L., Liu, H., Wang, F., et al. (2014). Discovery of rare mutations in extensively pooled DNA samples using multiple target enrichment. *Plant Biotechnol. J.* 12, 709–717. doi: 10.1111/pbi.12174

Chilamakuri, C. S., Lorenz, S., Madoui, M. A., Vodak, D., Sun, J., Hovig, E., et al. (2014). Performance comparison of four exome capture systems for deep sequencing. *BMC Genomics* 15, 449. doi: 10.1186/1471-2164-15-449

Colbert, T., Till, B. J., Tompa, R., Reynolds, S., Steine, M. N., Yeung, A. T., et al. (2001). High-throughput screening for induced point mutations. *Plant Physiol.* 126, 480–484. doi: 10.1104/pp.126.2.480

Dong, C., Vincent, K., and Sharp, P. (2009). Simultaneous mutation detection of three homoeologous genes in wheat by high resolution melting analysis and mutation surveyor. *BMC Plant Biol.* 9, 143. doi: 10.1186/1471-2229-9-143

Henry, I. M., Nagalakshmi, U., Lieberman, M. C., Ngo, K. J., Krasileva, K. V., Vasquez-Gross, H., et al. (2014). Efficient genome-wide detection and cataloging of EMS-induced mutations using exome capture and next-generation sequencing. *Plant Cell* 26, 1382–1397. doi: 10.1105/tpc.113.121590

IWGSC, International Wheat Genome Sequencing Consortium (2014). A chromosome-based draft sequence of the hexaploid bread wheat (Triticum aestivum) genome. *Science* 345, 1251788. doi: 10.1126/science.1251788

IWGSC, International Wheat Genome Sequencing Consortium (2018). IWGSC RefSeq principal investigators, Appels, R., Eversole, K., Feuillet, C., Keller, B., Rogers, J., et al. Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361. doi: 10.1126/science.aar7191

Jackman, S. D., Vandervalk, B. P., Mohamadi, H., Chu, J., Yeo, S., Hammond, S. A., et al. (2017). ABySS 2.0: resource-efficient assembly of large genomes using a Bloom filter. *Genome Res.* 27, 768–777. doi: 10.1101/gr.214346.116

Jackson, Lee. (2011). *Wheat cultivars for California*. Davis: University of California. https://smallgrains.ucdavis.edu/cereal_files/WhtCVDescLJ11.pdf

Jordan, K. W., Wang, S., Lun, Y., Gardiner, L. J., MacLachlan, R., Hucl, P., et al. (2015). A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. *Genome Biol.* 16, 48. doi: 10.1186/s13059-015-0606-4

Kabbaj, H., Sall, A. T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., et al. (2017). Genetic diversity within a global panel of durum wheat (Triticum durum) landraces and modern germplasm reveals the history of alleles exchange. *Front. Plant Sci.* 8, 1277. doi: 10.3389/fpls.2017.01277

Kawahara, Y., de la Bastide, M., Hamilton, J. P., Kanamori, H., McCombie, W. R., Ouyang, S., et al. (2013). Improvement of the Oryza sativa Nipponbare reference genome using next generation sequence and optical map data. *Rice (N. Y.)* 6, 4. doi: 10.1186/1939-8433-6-4

King, R., Bird, N., Ramirez-Gonzalez, R., Coghill, J. A., Patil, A., Hassani-Pak, K., et al. (2015). Mutation scanning in wheat by exon capture and next-generation sequencing. *PLoS One* 10, e0137549. doi: 10.1371/journal.pone.0137549

Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., et al. (2017). Uncovering hidden variation in polyploid wheat. *Proc. Natl. Acad. Sci. U. S. A.* 114, E913–EE21. doi: 10.1073/pnas.1619268114

'Kronos EI v1' Grassroots Data Repository, Earlham Institute. https://opendata.earlham.ac.uk/opendata/data/Triticum_turgidum/EI/v1/

Kurowska, M., Daszkowska-Golec, A., Gruszka, D., Marzec, M., Szurman, M., Szarejko, I., et al. (2011). TILLING: a shortcut in functional genomics. *J. Appl. Genet.* 52, 371–390. doi: 10.1007/s13353-011-0061-1

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence alignment/map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51 (5), 885–895. doi: 10.1038/s41588-019-0381-3

McCallum, C. M., Comai, L., Greene, E. A., and Henikoff, S. (2000). Targeted screening for induced mutations. *Nat. Biotechnol.* 18, 455–457. doi: 10.1038/74542

Microsoft. 'Visual Studio 2015'. https://visualstudio.microsoft.com/

Pallotta, M. A., Graham, R. D., Langridge, P., Sparrow, D. H. B., and Barker, S. J. (2000). RFLP mapping of manganese efficiency in barley. *Theor. Appl. Genet.* 101, 1100–1108. doi: 10.1007/s001220051585

Raghavan, C., Naredo, M. E. B., Wang, H. H., Atienza, G., Liu, B., Qiu, F. L., et al. (2007). Rapid method for detecting SNPs on agarose gels and its application in candidate gene mapping. *Mol. Breed.* 19, 87–101. doi: 10.1007/s11032-006-9046-x

Roche. 'Wheat exome'. https://sequencing.roche.com/en/products-solutions/by-category/target-enrichment/shareddesigns.html

Saintenac, C., Jiang, D., and Akhunov, E. D. (2011). Targeted analysis of nucleotide and copy number variation by exon capture in allotetraploid wheat genome. *Genome Biol.* 12, R88. doi: 10.1186/gb-2011-12-9-r88

Stephen, S. (2012). 'BioKanga', CSIRO. https://github.com/csiro-crop-informatics/biokanga/tree/v4.3.6.

Steuernagel, B., Periyannan, S. K., Hernandez-Pinzon, I., Witek, K., Rouse, M. N., Yu, G., et al. (2016). Rapid cloning of disease-resistance genes in plants using mutagenesis and sequence capture. *Nat. Biotechnol.* 34, 652–655. doi: 10.1038/nbt.3543

Suzuki, T., Eiguchi, M., Kumamaru, T., Satoh, H., Matsusaka, H., Moriguchi, K., et al. (2008). MNU-induced mutant pools and high performance TILLING enable finding of any gene mutation in rice. *Mol. Genet. Genomics* 279, 213–223. doi: 10.1007/s00438-007-0293-2

Till, B. J., Zerr, T., Comai, L., and Henikoff, S. (2006). A protocol for TILLING and Ecotilling in plants and animals. *Nat. Protoc.* 1, 2465–2477. doi: 10.1038/nprot.2006.329

Tsai, H., Howell, T., Nitcher, R., Missirian, V., Watson, B., Ngo, K. J., et al. (2011). Discovery of rare mutations in populations: TILLING by sequencing. *Plant Physiol.* 156, 1257–1268. doi: 10.1104/pp.110.169748

Tucker, B. A., Scheetz, T. E., Mullins, R. F., DeLuca, A. P., Hoffmann, J. M., Johnston, R. M., et al. (2011). Exome sequencing and analysis of induced pluripotent stem cells identify the cilia-related gene male germ cell-associated kinase (MAK) as a cause of retinitis pigmentosa. *Proc. Natl. Acad. Sci. U. S. A.* 108, E569–E576. doi: 10.1073/pnas.1108918108

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R. K., Tsai, H., Berardi, S., et al. (2009). A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC Plant Biol.* 9, 115. doi: 10.1186/1471-2229-9-115

Winfield, M. O., Wilkinson, P. A., Allen, A. M., Barker, G. L., Coghill, J. A., Burridge, A., et al. (2012). Targeted re-sequencing of the allohexaploid wheat exome. *Plant Biotechnol. J.* 10, 733–742. doi: 10.1111/j.1467-7652.2012.00713.x

Yang, B., Wen, X., Kodali, N. S., Oleykowski, C. A., Miller, C. G., Kulinski, J., et al. (2000). Purification, cloning, and characterization of the CEL I nuclease. *Biochemistry* 39, 3533–3541. doi: 10.1021/bi992376z

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 22 October 2019
doi: 10.3389/fpls.2019.01324

Equipping Durum Wheat—*Thinopyrum ponticum* Recombinant Lines With a *Thinopyrum elongatum* Major QTL for Resistance to Fusarium Diseases Through a Cytogenetic Strategy

*Ljiljana Kuzmanović[1], Giulia Mandalà[1], Silvio Tundo[1†], Roberto Ciorba[1], Matteo Frangella[1], Roberto Ruggeri[1], Francesco Rossini[1], Federica Gevi[1], Sara Rinalducci[2] and Carla Ceoloni[1]**

[1] Department of Agricultural and Forest Sciences (DAFNE), University of Tuscia, Viterbo, Italy, [2] Department of Ecological and Biological Sciences (DEB), University of Tuscia, Viterbo, Italy

OPEN ACCESS

Edited by:
Brian L. Beres,
Agriculture and Agri-Food Canada, Canada

Reviewed by:
Filippo Maria Bassi,
International Center for Agricultural Research in the Dry Areas (ICARDA), Morocco
Ken Chalmers,
University of Adelaide, Australia

***Correspondence:**
Carla Ceoloni
ceoloni@unitus.it

†Present Address:
Silvio Tundo
Department of Land, Environment, Agriculture and Forestry (TeSAF), Research Group in Plant Pathology, University of Padova, Legnaro, Italy

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 14 June 2019
Accepted: 24 September 2019
Published: 22 October 2019

Citation:
Kuzmanović L, Mandalà G, Tundo S, Ciorba R, Frangella M, Ruggeri R, Rossini F, Gevi F, Rinalducci S and Ceoloni C (2019) Equipping Durum Wheat—Thinopyrum ponticum Recombinant Lines With a Thinopyrum elongatum Major QTL for Resistance to Fusarium Diseases Through a Cytogenetic Strategy. Front. Plant Sci. 10:1324. doi: 10.3389/fpls.2019.01324

Prompted by recent changes in climate trends, cropping areas, and management practices, *Fusarium* head blight (FHB), a threatening disease of cereals worldwide, is also spreading in unusual environments, where bread wheat (BW) and durum wheat (DW) are largely cultivated. The scarcity of efficient resistance sources within adapted germplasm is particularly alarming for DW, mainly utilized for human consumption, which is therefore at high risk of kernel contamination by health-dangerous mycotoxins (e.g., deoxynivalenol = DON). To cope with this scenario, we looked outside the wheat primary gene pool and recently transferred an exceptionally effective FHB resistance QTL (*Fhb-7EL*) from *Thinopyrum elongatum* 7EL chromosome arm onto a *Thinopyrum ponticum* $7el_1L$ arm segment, containing additional valuable genes (including *Lr19* for leaf rust resistance and *Yp* for yellow pigment content), distally inserted onto 7DL of BW lines. Two such lines were crossed with two previously developed DW-*Th. ponticum* recombinants, having $7el_1L$ distal portions on 7AL arms. Genomic *in situ* hybridization (GISH) analysis showed homologous pairing, which is enabled by $7el_1L$ segments common to the BW and DW recombinant chromosomes, to occur with 42-78% frequency, depending on the shared $7el_1L$ amount. Aided by 7EL/$7el_1L$-linked markers, 7EL+$7el_1L$ tetraploid recombinant types were isolated in BC_1 progenies to DW of all cross combinations. Homozygous 7EL+$7el_1L$ recombinant plants and null segregates selected in BC_2F_2 progenies were challenged by *Fusarium graminearum* spike inoculation to verify the *Fhb-7EL* efficacy in DW. Infection outcomes confirmed previous observations in BW, with >90% reduction of disease severity associated with *Fhb-7EL* presence *vs*. its absence. The same differential effect was detected on seed set and weight of inoculated spikes, with genotypes lacking *Fhb-7EL* having ~80% reduction compared with unaffected values of *Fhb-7EL* carriers. In parallel, DON content in flour extracts of resistant recombinants averaged 0.67 ppm, a value >800 times lower than that of susceptible controls. Furthermore, as observed

in BW, the same *Fhb-7EL* also provided the novel DW recombinants with resistance to *Fusarium* crown rot (~60% symptom reduction) as from seedling infection with *Fusarium culmorum*. Through alien segment stacking, we succeeded in equipping DW with a very effective barrier against different *Fusarium* diseases and other positive attributes for crop security and safety.

Keywords: alien gene transfer, chromosome engineering, chromosome pairing, GISH, marker-assisted selection, *Triticum*, wild wheat relatives, sustainability

INTRODUCTION

With about 8% coverage of the world's wheat area, durum wheat (*Triticum durum* Desf., 2n = 4x = 28, genome AABB) is the 10th most important crop in the world (Bassi and Sanchez-Garcia, 2017). Not only does it represent a strategic commodity for the three world areas where it is mainly cropped (the Mediterranean basin, the North America's Great Plains, and the desert areas of South-Western United States and Northern Mexico; Ranieri, 2015), but durum wheat cultivation is also expanding in Canada, India, and even the Senegal River basin in sub-Saharan Africa (Sall et al., 2018). As with all other crops, it is experiencing the effects of climate changes, hence requiring dedicated breeding efforts to cope with them and concurrent challenges to the present and projected demand for higher food supply (e.g., Ray et al., 2019).

As a result of climate extremes, particularly rising temperatures, not only do the conventional distribution areas of crops tend to be modified (e.g., Ceoloni et al., 2014a), but the ecology, epidemiology, and virulence/aggressiveness of their pathogens are also subject to considerable variation (Fones and Gurr, 2017). Plants suffering abiotic stresses such as heat and drought are more susceptible to unspecialized necrotrophic pathogens, the same stress conditions also accelerating pathogen evolution (Chakraborty, 2013; Vaughan et al., 2016). Typical necrotrophs are fungal pathogens belonging to the *Fusarium* genus, responsible for some of the most threatening diseases of wheat and other cereals, namely, Fusarium head blight (FHB) and Fusarium crown rot (FCR). Environments where humid and warm conditions occur around the flowering stage are typically prone to FHB, while FCR is prevalent under drier conditions. On a world scale, FHB is predominantly caused by *F. graminearum*, while *F. culmorum* and *F. pseudograminearum* are the main agents of FCR (Gilbert and Haber, 2013; Scherm et al., 2013; Matny, 2015). They are all toxigenic fungi, secreting secondary metabolites that play a significant role in pathogen virulence *in planta*, likely due to their ability to inhibit eukaryotic protein synthesis (reviewed in Bakker et al., 2018). In wheat, the most frequently detected of such mycotoxins is deoxynivalenol (DON), belonging to the trichothecenes, whose role as virulence factor in FHB and FCR was consistently demonstrated in bread and durum wheat subjected to inoculation with the *Fusarium* species mentioned earlier (Mudge et al., 2006; Scherm et al., 2011; Sella et al., 2014; Mandalà et al., 2019). Alongside its role in pathogenesis, DON is a highly hazardous compound for human and animal health (Maresca, 2013), and strict rules and legislative limits for maximum levels in food and feed have been defined worldwide (Romer Labs Division Holding GmbH, 2016). The economic value of contaminated crops is affected not only by safety problems but also by grain yield and quality penalties, due to failed development or shrivelling, discoloration, and low test weight of infected kernels (e.g., McMullen et al., 2012; Matny, 2015; Salgado et al., 2015).

Impacts on safety, security, and processing issues are particularly alarming for durum wheat, used almost exclusively for transformation into human food products. In a sustainable agricultural perspective, and also considering that agronomic practices and fungicides can only partially reduce the infection risks, the use of resistant cultivars is widely recognized as the most effective tool for controlling Fusarium diseases (e.g., Steiner et al., 2017). However, the needed genetic variation for successful breeding actions addressing such diseases appears to be quite scarce within the cultivated and closely related tetraploid gene pools, being limited to quantitative trait loci (QTL) of minor individual effect (Prat et al., 2014).

In bread wheat (*Triticum aestivum* L., 2n = 6x = 42, genome AABBDD), breeding for FHB resistance has so far been centered mostly on a large-effect QTL, namely *Fhb1*, located on the 3BS chromosome arm of the bread wheat Chinese cultivar Sumai 3 and its derivatives (Gilbert and Haber, 2013; Steiner et al., 2017). Similarly, a single major QTL, identified on 3BL of hexaploid germplasm (*T. spelta*), is being exploited for FCR resistance breeding (Liu and Ogbonnaya, 2015). Being located on a shared chromosome, i.e., 3B, transfer of both QTL from bread wheat into durum wheat represented a relatively amenable option. However, results indicated dependency on the cultivar background for the expression of *Fhb1*-linked resistance (Prat et al., 2017), with lack of any FCR improvement associated with presence of the 3BL locus (Ma et al., 2012). Whether the higher susceptibility of durum wheat than bread wheat toward Fusarium diseases, and hence the partial and unpredictable effect of interspecific transfers, might be due to durum wheat-specific susceptibility factors, or to the so far minor exposure of the crop to relevant disease pressure, remains to be elucidated (Giancaspro et al., 2016). No doubt, the current lack of highly resistant genotypes among cultivated durum wheat worldwide is also the result of limited breeding efforts to date targeting *Fusarium* spp. resistance in durum wheat compared with bread wheat (Giancaspro et al., 2016; Prat et al., 2017).

A wide array of beneficial traits, rarely or not represented in cultivated wheat or closely allied gene pools, such as resistance to Fusarium diseases, resides in more distant relatives, including perennial Triticeae of the *Thinopyrum* genus (Forte et al., 2014; Ceoloni et al., 2015 and references therein). Belonging to the wheat

tertiary gene pool, they still retain considerable cytogenetic affinity with wheat chromosomes, albeit often characterized by segmental homoeology (Ceoloni et al., 2015). A meaningful example of positive impact of a *Thinopyrum* source on enhancement of *Fusarium* spp. resistance in cultivated wheat germplasm is that of a major QTL, named *Fhb-7el*$_2$ (Forte et al., 2014) and later *Fhb7* (Guo et al., 2015b), originating from chromosome 7el$_2$ of decaploid *Th. ponticum*. This strong FHB resistance QTL was pyramided into bread wheat (Shen and Ohm, 2007; Zhang et al., 2011; Forte et al., 2014) and also durum wheat (Forte et al., 2014) by exploiting the close relatedness of 7el$_2$ with 7el$_1$ (Dvorak, 1975; Guo et al., 2015a), deriving from a different accession of the same species. The latter bear genes for effective rust resistances (*Lr19* and *Sr25*) and for yield-contributing traits (see Ceoloni et al., 2015 for a review). In both wheat species, the effect of the 7el$_2$ QTL was highly significant when compared with susceptible sibs, reducing FHB severity on infected spikes by 70–85% (Forte et al., 2014).

Previously obtained wheat-alien translocation and recombinant lines with portions of the respective *Thinopyrum* donor chromosome containing the target genes were instrumental to the successful pyramiding of the 7el$_2$+7el$_1$ genes/QTL *via* 7el$_1$–7el$_2$ pairing and recombination. These wheat-alien primary transfer lines were, in turn, the result of chromosome engineering, i.e., a suite of cytogenetic methodologies which enable alien segmental introgressions into wheat homoeologous chromosomes, mostly *via* pairing mediated by wheat *ph1* mutations (reviewed in Ceoloni and Jauhar, 2006; Qi et al., 2007). However, in wheat-alien combinations, *ph1* mutations promote autosyndetic (wheat–wheat), as well as allosyndetic (wheat–alien) homoeologous chromosome pairing and recombination. This represents a drawback, which limits recovery and affects background stability of target allosyndetic recombinants (see, e.g., Ceoloni and Jauhar, 2006; Zhang et al., 2017).

To circumvent these problems, an alternative strategy was followed in the recent transfer of a major QTL for resistance to Fusarium diseases from chromosome 7E of *diploid Th. elongatum* into bread wheat (Ceoloni et al., 2017a). This did not rely on the *ph1* promotion but took advantage of the close homoeology relating *Th. elongatum* chromosome 7E and *Th. ponticum* 7el$_1$ (Dvorak, 1975). As a result, spontaneous pairing and recombination occurred between the 7E long arm (7EL) and a sizable 7el$_1$L segment (70% of the arm length), present in a 7E(7D) substitution line and in the 7DS·7DL-7el$_1$L of the T4 translocation line, respectively. Pyramiding of the positive traits controlled by 7el$_1$L genes/QTL (see above) with the 7EL-linked Fusarium resistance QTL (named *Fhb-7EL*) was thus achieved, with *Fhb-7EL* being shown to map more distally than the 7el$_1$L genes (Ceoloni et al., 2017a). The presence of small 7EL terminal segments containing the *Fhb-7EL* QTL was shown to determine an exceptionally effective FHB resistance in bread wheat recombinant lines inoculated with *F. graminearum*, of the same degree as that previously associated with the entire 7E or 7EL (Shen et al., 2004; Miller et al., 2011). The "type II" resistance, i.e., resistance to fungal spread within host tissues (Mesterházy et al., 1999), was expressed at its maximum level, with spread of the pathogen limited to the immediate vicinity of the inoculated floret (Ceoloni et al., 2017a), and an average 95% reduction of disease severity in inoculated spikes of *Fhb-7EL* carrier *vs.* non-carrier lines. In the same work, the *Fhb-7EL* QTL was for the first time also demonstrated to substantially reduce FCR, incited by seedling inoculation with *F. culmorum* and *F. pseudograminearum*. Marker- and phenotype-based assessments showed some of the recombinants bearing *Fhb-7EL* to possess additional desirable 7el$_1$L genes. In particular, *Fusarium* spp. resistant recombinant lines R69-9 and R74-10 were shown to include in their proximal 7el$_1$L segments the leaf rust resistance gene *Lr19* and different alleles at the *Psy1* (*Phytoene synthase 1*) locus (*Psy1-7el$_1$L* in R69-9, and *Psy1-7EL* in R74-10; see **Figure 1**), consistently associated with increases of yellow pigment (*Yp*) content (Zhang and Dubcosky, 2008; Ceoloni et al., 2017a).

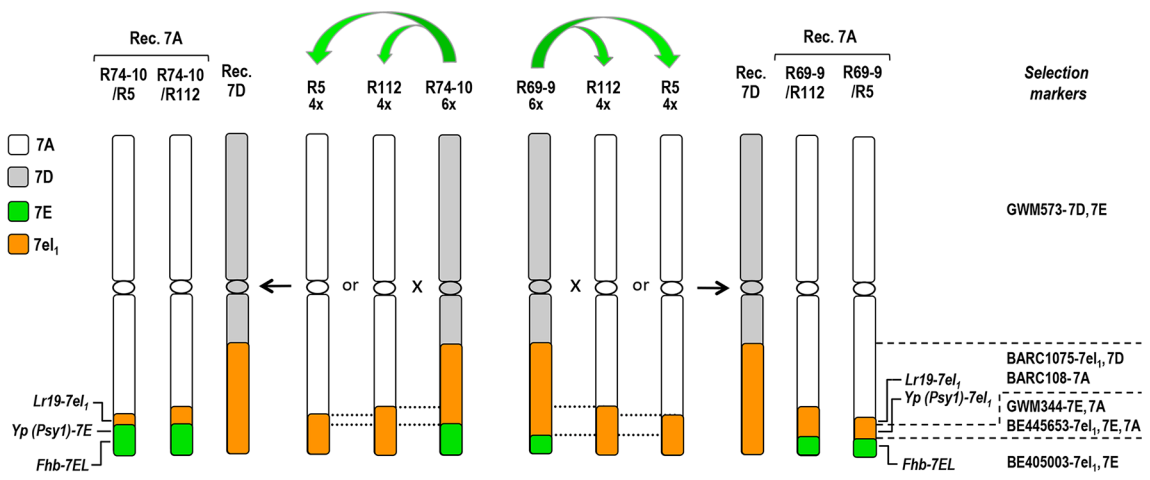


FIGURE 1 | Cytogenetic maps of group-7 wheat-*Thinopyrum* spp. chromosomes involved in the pyramiding scheme of distal 7EL segments carrying the *Fhb-7EL* QTL (from 6x lines R69-9 or R74-10) into 7el$_1$L-7AL arms of 4x lines R112 or R5. Dotted lines delimit the regions where homologous pairing and crossing-over in the shared 7el$_1$L regions can occur and give rise to the desired pyramiding. Dashed lines on the right indicate the chromosomal regions where marker loci and target genes are located. For detailed genetic and physical mapping data of *Thinopyrum* spp. segments into wheat chromosomes, see Ceoloni et al., 2017a.

For their chromosomal and genetic makeup, as well as the good agronomic performance in preliminary tests (Ceoloni et al., 2017a), bread wheat recombinant lines such as R69-9 and R74-10 appeared as attractive candidates for the incorporation of the 7EL+7el$_1$L gene/QTL package into durum wheat as well. The envisaged strategy was to rely on potential homologous pairing of such donor chromosomes with recipient ones sharing some 7el$_1$L chromatin. These were present in previously obtained 7AL-7el$_1$L durum wheat introgression lines (Ceoloni et al., 2005). Among them, lines R5 and R112 not only possess the *Lr19+Sr25+Yp* genes from 7el$_1$L but also showed 7el$_1$L-linked positive effects on various yield components in various environments (Kuzmanović et al., 2014, Kuzmanović et al., 2016, Kuzmanović et al., 2018). The 7el$_1$L segment spans 23% of the recombinant 7AL in R5 and 28% in R112 (Ceoloni et al., 2005). In principle, the same scheme adopted for the bread wheat transfer, i.e., use of the bread wheat 7E(7D) substitution line as donor of the *Fhb-7EL* resistance QTL, might have been attempted for its introduction into 7el$_1$L segments of R5 and R112. However, previous experience suggested this route to be quite impractical for durum wheat. In fact, mainly due to the different chromosomal contexts (pentaploid *vs*. hexaploid hybrids), spontaneous pairing between the 7EL arm from the 7E(7D) substitution line and the 7el$_1$L segments of R5 or R112 was expected to be sharply reduced compared with that using the T4 translocation line, as observed in the aforementioned 7el$_1$L+7el$_2$L pyramiding (Forte et al., 2014).

The objectives of the work described here were (i) to engineer the R5 and R112 7el$_1$L segments with telomeric 7EL portions, bearing the *Fhb-7EL* QTL, by exploiting the homologous pairing potential of 7el$_1$L segments shared by recipient and donor chromosomes; (ii) to verify the ability and extent of the *Fhb-7EL* QTL in conferring FHB and FCR resistances once stably introgressed into the new genomic context of durum wheat; and (iii) to provide a preliminary assessment of stability and performance of novel recombinant types, in relation to their exploitation in breeding.

MATERIALS AND METHODS

Plant Materials and Transfer Scheme

Donors used for the transfer of the *Th. elongatum Fhb-7EL* QTL into durum wheat were two bread wheat recombinant lines, named R74-10 and R69-9 (7DS·7DL-7el$_1$L/7EL; see **Figure 1**), with a terminal 7EL portion, including *Fhb-7EL*, embedded into a 7el$_1$L *Th. ponticum* segment extending to 0.7 fractional length of the 7DL arm (T4 translocation line; see Ceoloni et al., 2017a). To combine the *Fhb-7EL* locus with 7el$_1$L-linked positive genes/QTL for durum wheat performance (see, e.g., Gennaro et al., 2003, Gennaro et al., 2007; Kuzmanović et al., 2014, Kuzmanović et al., 2016), R74-10 and R69-9 were each crossed with two *Th. ponticum*–durum wheat recombinant lines, named R5 and R112, having 23 and 28%, respectively, of their distal 7AL arms replaced by homoeologous 7el$_1$L portions (7AS·7AL-7el$_1$L; see **Figure 1**) in a background near-isogenic to that of the Italian durum wheat cv. Simeto (Ceoloni et al., 2005). Pentaploid (5x) hybrid progeny of each of the four cross combinations was then subjected to further cross with cv. Simeto, hence consisting of backcrosses (BCs) to the same recurrent background. To identify recombinant types, BC$_1$ plants (e.g., R74-10/R112//Simeto = R74-10/2**T. durum*) were analyzed by suitable polymorphic markers (**Figure 1**), and the chromosome number of recombinant individuals determined (see below). BC$_2$ progenies were then obtained from plants whose marker profile was indicative of the location on 7AL of the targeted 7EL+7el$_1$L assembly, and these, with the majority having reached a euploid condition (2n = 28), were self-pollinated. BC$_2$F$_2$ offsprings were then genotyped, and the resulting homozygous carriers (HOM+) and non-carriers (HOM–) of the specific 7EL+7el$_1$L combination, as well as their self-fertilized progeny, were used in various comparative tests. In these, depending on the type of experiment, the Chinese Spring (CS) 7E(7D) substitution line (2n = 42), original donor to R74-10 and R69-9 of the *Fhb-7EL* QTL (see Ceoloni et al., 2017a), the R112 and R5 recombinants, as well as durum wheat cv. Simeto, were included as control lines. Data from R112 and R5 plants (FHB infection and subsequent assays on inoculated plants, see below) were pooled (hereafter indicated as R112+R5), as the two genotypes did not show appreciable differences for such traits (see, e.g., Forte et al., 2014).

Cytogenetic Analyses

Standard Feulgen or aceto-carmine staining techniques were applied both to assess the somatic chromosome number in root tip cells of selected genotypes and for quick anther screening from freshly collected young spikes, in view of meiotic metaphase I preparations. To this aim, selected anthers with pollen mother cells (PMCs) at the target phase, kept at −20°C in 3:1 fixative (absolute alcohol:acetic acid) for up to several weeks, were rinsed in 45% acetic acid, transferred for about 1 h to 2% aceto-carmine in 60% acetic acid at 37°C and squashed in 45% acetic acid, before freezing the slide in liquid nitrogen. For pairing analyses, metaphase I spreads were subjected to GISH (genomic *in situ* hybridization), using total DNAs of *T. aestivum* and *Th. ponticum* as genomic probes. Due to the close relatedness between *Th. elongatum* and *Th. ponticum* genomes, the latter equally highlights any *Thinopyrum* spp. introgression into wheat. Total DNAs were extracted from leaves following Tai and Tanksley (1990), mechanically sheared to 8–10 kb fragments and labeled by nick translation, including biotin-11-dUTP (Fermentas) or digoxigenin 11-dUTP (Roche Diagnostics) in the deoxyribonucleotide (dNTP) mix. The hybridization protocol followed Anamthawat-Jónsson and Reader (1995) with some modifications. In particular, to enrich the hybridization mixture in genome-specific sequences, equal quantities (100 ng) of denatured and differently labeled wheat and *Thinopyrum* probes were allowed to preanneal for 30 min at 58°C. Prior to hybridization with the pre-annealed probes, a blocking mixture, containing 1 mg of autoclaved and unlabeled DNA of *Aegilops speltoides* (2n = 14, genome SS, closely related to the B genome of polyploid wheats), was applied for 1.5 h at 63°C onto denatured chromosome preparations. This additional step led to a preferential block of B-genome chromosomes (not involved in wheat-*Thinopyrum* rearrangements), which enhanced the overall

differentiation among chromosomes/segments of different genomic origin. Hybridization was then carried out for 2 h at 63°C, after which digoxigenin- and biotin-labeled probes were correspondingly detected using anti-digoxigenin conjugated with FITC (Roche; green fluorescence) and streptavidin conjugated with Cy3 (Amersham; red fluorescence).

All chromosome preparations were analyzed using a Leica DM5000B epifluorescence microscope, equipped with a SPOT-RT3 (Diagnostic Instruments, Inc.) color digital camera and the SPOT™ Advanced Plus imaging software.

Marker-Assisted Selection (MAS)

The choice of suitable markers enabling discrimination between parental and recombinant types in BC_1 progeny (pentaploid F_1s × Simeto) and further genotyping in subsequent generations was facilitated by previously established inter-genomic polymorphism and genetic/physical mapping of several wheat and *Thinopyrum* spp. group 7 markers in the chromosomal regions of interest (see Ceoloni et al., 2017a, Ceoloni et al., 2017b and additional references therein). Therefore, only a limited number of PCR-based, mostly codominant markers were employed (**Figure 1**, **Table 1**). Most of these markers were used to isolate and confirm identity of recombinant types in the BC_1 progeny to Simeto, while only one (e.g., BE405003) was sufficient to select for heterozygous recombinants in all BC_2 progenies (presence of the 7E allele), and to discriminate heterozygotes (HET) from HOM+ and HOM– for the 7EL+7el$_1$L segment assembly in BC_2F_2 progenies of 7A recombinants (e.g., BE445653 or GWM344). The STS*Lr19*$_{130}$ marker, closely linked to *Lr19* (Prins et al., 2001), was employed to confirm presence of the *Th. ponticum* leaf rust resistance gene. A previously developed STS-CAPS assay (Ceoloni et al., 2017a), enabling discrimination of 7EL *vs.* 7el$_1$L alleles, was applied to tag the *Psy1* gene, associated to the *Yp* phenotype (see Introduction).

For PCR reactions, DNA was extracted from young leaves or half-kernels according to Dellaporta et al. (1983). Primer sequences were retrieved from the public GrainGenes databases (http://wheat.pw.usda.gov/GG3/). For each 10 µl PCR reaction, 1× GoTaq® G2 Master Mix (Promega, #M7822) and 25 ng of DNA were used for all primer pairs, while primer concentration, annealing temperature, and use of additional reagents varied, as reported in **Table 1**. Except for BARC1075 and BARC108 markers, for which a multiplex assay was developed, all other markers were amplified in a simple PCR. Amplified products were separated on 1.5–3% agarose gel, visualized by ethidium bromide staining and images captured with Kodak EDAS 290 digital system.

Fusarium spp. Inoculation and DON Assays

FHB: Spike Inoculations With *F. graminearum*

Tetraploid homozygous carriers (HOM+) and non-carriers (HOM–) of the *Fhb-7EL* locus (based on marker analyses), isolated in BC_2F_2 progenies after crossing with Simeto (see above) of R74-10/R112, R74-10/R5 and R69-9/R112 F_1s, together with the recurrent parent cv. Simeto, as well as R112 and R5 recipient lines and the CS7E(7D) substitution line (as FHB resistant control), were employed for a single-floret *F. graminearum* inoculation experiment. The R69-9/R5 corresponding progeny was not available at the time of inoculations, hence was not included in the assay. The infection assay was conducted under controlled conditions (16 h light/8 h dark photoperiod and 22–24°C/20°C corresponding temperature regimes) when plants were at mid-anthesis stage. The inoculum consisted of 1,000 macroconidia of *F. graminearum* strain 3824 (Tundo et al., 2016), freshly cultured on synthetic nutrient agar (SNA) medium (Urban et al., 2002), suspended in 20 µl of sterile distilled water (5×10^4 ml^{-1} concentration), and supplemented with 0.05% Tween-20. The conidia suspension was pipetted through the glumes onto the basal floret of one central spikelet from the tip of the first spike of each plant. Inoculated spikes were covered with a plastic bag for 48 h to maintain high relative humidity. Disease symptoms were assessed at 7, 14, and 21 days post-inoculation (dpi), by calculating the percentage number of visually diseased florets (NDF) out of the total number of florets per spike. Differences in disease severity among genotypes were estimated by means of NDF ± SE (standard error) of 8–10 plants/genotype (4 each for R5 and R112, pooled) and by one-way analysis of variance (ANOVA). Seed number and weight (thousand grain weight, TGW) were assessed for inoculated and non-inoculated spikes of the same infected plants, and the differences among genotypes assessed as described above for disease severity. The same seeds were also used to extract wholemeal flour and determine the DON content (see below).

TABLE 1 | Group 7 molecular markers used to identify wheat—*Thinopyrum* spp. genotypes in the course of the work.

Marker	Type	Primer concentration (nM)	Other reagents	Annealing temperature (°C)	Alleles amplified (bp)					
					7el$_1$L	7ES	7EL	7AL	7DS	7DL
BE405003	EST	400	–	55	700	–	600	–	–	–
BE445653	EST	250	5% DMSO	52	930	–	1200	750	–	–
GWM344	SSR	200	–	55	–	–	100	130–150	–	–
GWM573	SSR	200	–	50	–	200	–	–	180	–
BARC1075	SSR	200	–	53	250	–	–	–	–	200
BARC108	SSR	250	–	53	–	–	–	160	–	–
STS*Lr19*$_{130}$	STS	200	–	58	130	–	–	–	–	–
STSPsy1	STS-CAPS	200	–	60	730	–	705	450+270	–	–

BARC1075 and BARC108 markers were used in a multiplex assay (see Materials and Methods); for details of the CAPS assay applied for STSPsy1, see Ceoloni et al., 2017a.

FCR: Seedling Inoculations With *F. culmorum*

Homozygous BC_2F_3 plants from the progeny of one of the tetraploid 7EL-7el$_1$L HOM+, FHB resistant recombinants (R69-9/R112 cross derivatives), as well as of sib HOM− plants and of cv. Simeto as controls, were used in two independent infection experiments with *F. culmorum*. Twenty plants per genotype were included in each experiment. Seeds were surface sterilized with sodium hypochlorite (0.5% vol/vol) for 20 min and then rinsed thoroughly in sterile water. Seedlings were individually grown in 5 × 5 × 5-cm pots and arranged in plastic trays and maintained at the same light and temperatures regimes as described for the FHB assay throughout the experiments. *F. culmorum* strain UK99 macroconidia were produced by fungal culture on SNA medium and harvested by washing the culture surface with 2 ml sterile water (Urban et al., 2002). The inoculum solution contained 2×10^6 ml^{-1} conidia (Beccari et al., 2011) and 0.05% Tween 20. As described in Ceoloni et al. (2017a), the inoculation procedure consisted of evenly spreading (with a small paintbrush) 20 µl of conidia suspension on the stem base leaf sheaths of plantlets at the first-leaf stage. Trays with inoculated plants were covered with a plastic film for 48 h to maintain high humidity conditions. Disease symptoms were assessed at 7, 11, 14, 18, and 21 days post-inoculation (dpi) measuring two parameters on the infected tissue: symptom extension (SE; cm) and browning index (BI, visual rating of the degree of extension of necrosis, as indicated by brown discoloration, based on a five-point scale: 0, symptomless; 1, slightly necrotic; 2, moderately necrotic; 3, severely necrotic; 4, completely necrotic). The final score, indicated as disease index (DI), was determined as SE × BI (Beccari et al., 2011). DI values, expressed as means ± SE of 20 plants per genotype and per experiment at each time-point, were subjected to two-way ANOVA.

Quantification of DON

DON content was determined in wholemeal flour of kernels produced by plants subjected to *F. graminearum* infection. Extraction and analytical procedures were performed as described in Mandalà et al. (2019). Briefly, the metabolite was extracted from 100 mg wholemeal flour dissolved in 400 µl of 86:14 acetonitrile:water (v/v) solution by prolonged shaking (24 h, 180 rpm, 4°C). After centrifugation, supernatants were injected into a UHPLC system (Ultimate 3000, Thermo) and run in positive ion mode. A Reprosil C18 column (2.0 mm × 150 mm, 2.5 µm—Dr. Maisch, Germany) was used for metabolite separation. The UHPLC system was coupled online with a mass spectrometer Q Exactive (Thermo) scanning in full MS mode (2 µ scans) at 70,000 resolution in the 60 to 1,000 m/z range. Data files were processed by MAVEN.52 (http://genomics-pubs.princeton.edu/mzroll/) upon conversion of raw files into mzXML format through MassMatrix (Cleveland, OH). Standard curves were obtained with six calibration points (2mg–0.00002 mg) of DON analytical standard (Romer Labs). To assess the effect of presence *vs*. absence of the *Fhb-7EL* QTL on DON content, each of the three *Fhb-7EL* carriers (R74-10/R112, R74-10/R5 and R69-9/R112 HOM+ derivatives) and non-carrier (bulked HOM− segregates, bulked R112+R5, and Simeto) genotypes was considered as a biological replicate. Bulks were a necessary option, due to limited amount of flour extracted from shriveled seeds of heavily diseased genotypes. For each replicate, seeds from all infected spikes were used to produce a single flour sample, from which three technical replicates were obtained. Values of all biological × technical replicates were analyzed by analysis of co-variance (ANCOVA), which, better than ANOVA, could eliminate the undesirable variable represented by genetic background heterogeneity across genotypes.

Evaluation of Yield-Related and Quality Traits

Homozygous durum wheat BC_2F_{3-4} recombinant plants (HOM+) from the R74-10/R112/2*Simeto and R69-9/R112/2*Simeto cross combinations, carrying different amounts of 7EL chromatin including the *Fhb-7EL* QTL, stacked into the same 7el$_1$L segment of R112-7AL arm (see **Figure 1**), were field grown for 2 years (2017–18 and 2018–19 seasons) and in one locality (Viterbo, Central Italy, University of Tuscia Experimental Station), alongside sib plants of null segregates (HOM−) from the same progeny, as well as Simeto plants. In both seasons, plants were grown under common cultural practices and no fungicide application. In the 1st season, BC_2F_3 plants were organized in randomized, triplicate rows (1 m long), at a 25-cm distance between rows and 10-cm distance along the row. In the 2nd experimental year, the trial consisted of spike rows of BC_2F_4 selections of each HOM+ and HOM− genotype and of cv. Simeto. On separate plants (1st year), data were collected for spike number/plant (SNP), grain number/plant (GNP), TGW, grain yield/plant (GYP), plant height (PH), days to heading (HD), and spike traits, including grain number/spike (GNS), grain yield/spike (GYS), spikelet number/spike (SPN), grain number/spikelet (GNSP), and spike fertility index (SFI). SFI, indicating the ability of the plant to set seeds in relation to the spike biomass, was calculated as the ratio between GNS and weight of spike chaff (g) of mature and oven-dried (48 h at 65°C) spikes. For each parameter, values from 20 to 30 plants per genotype, expressed as means ± SE, were subjected to one-way ANOVA. In the 2nd year trial, besides PH and HD average values/row, the same spike traits mentioned above were analyzed on 25 spikes/genotype (five from each of five rows).

Harvested seed from BC_2F_4 plant rows was milled into semolina to measure the yellow index (YI) of contrasting genotypes for *Psy1* alleles. Using the reflectance colorimeter CR-400 Chroma Meter (Minolta), absolute measurements for L* (lightness), a* (red-green chromaticity), and b* (yellow-blue chromaticity) coordinates in the Munsell color system were taken using D65 lightning (reviewed in Ficco et al., 2014). The b* parameter, representing the variation in semolina YI, is known to be highly correlated with yellow pigment content (YPC) of whole-meal flour extracts (Ravel et al., 2013; Ficco et al., 2014). Semolina samples, each analyzed in triplicate (technical replicates), derived from seeds of three plants of HOM+ and HOM− sister lines/genotype and of Simeto.

As to leaf rust evaluation, aimed at confirming the efficacy of *Lr19*-based resistance, accurate scoring of disease severity was carried out in the 2018–19 season. A commonly used double-digit scale was adopted, in which the first digit indicates the

rise of the disease, from the 1st leaf (1) to spike (9; typically 8 = flag-leaf for leaf rust), with 0 = no disease, and the second digit corresponds to a one-value percentage of the average infection intensity on the leaf area (e.g., 3 = 30%), based on the modified Cobb scale (Peterson et al., 1948).

Statistical Analyses

ANOVA and ANCOVA were performed using SYSTAT12 Software (Systat Software Incorporated, San Jose, CA, USA). The variable parameter (i.e., percentage of diseased florets for FHB, DI for FCR, and quantification of DON content in flour, each of various agronomic parameters) was considered as a dependent factor against the independent factor "genotype" (G). Additionally, "replica" (R) was included as independent factor in the two-way ANOVA performed for FCR assessment, or as a covariate in the ANCOVA used for DON and YI assays. Three levels of significance ($P < 0.05$, $P < 0.01$, and $P < 0.001$) were considered for F values. When significant values were observed, a pairwise analysis was carried out by the Tukey Honestly Significant Difference test (Tukey test) at 0.95 confidence level.

RESULTS

Meiotic Pairing Analysis

The ability to undergo meiotic metaphase I pairing by 7DS·7DL-7el$_1$L/7EL chromosomes of R74-10 and R69-9 6x lines (bearing the *Fhb-7EL* QTL) and 7AS·7AL-7el$_1$L chromosomes of R112 and R5 4x lines within their shared (homologous) 7el$_1$L regions was assessed in PMCs of their 5x F_1 plants processed by GISH. Because of the presence of a normal 7A from the 6x parent, homologous to the 7A of the durum parents at the short (S) arm level, a trivalent configuration was expected to occur if, in addition to 7AS-7AS pairing, that between the 7el$_1$L homologous regions of R74-10/R69-9 and R112/R5 would have also taken place. This was in fact the type of association in which a GISH site at the level of a chiasmate region was observed in the largest majority of PMCs in all F_1 types (**Figures 2A**, **B**; **Table 2**). Only a small percentage (ranging from 1.6 to 6.7) of PMCs showed the R74-10 or R69-9 chromosome paired with R112 or R5 in the form of a rod bivalent (**Figure 2C**), as a consequence of failure of 7AS-7AS pairing. Concerning the trivalent associations, these were prevailingly of the open type (**Figure 2A**), but a

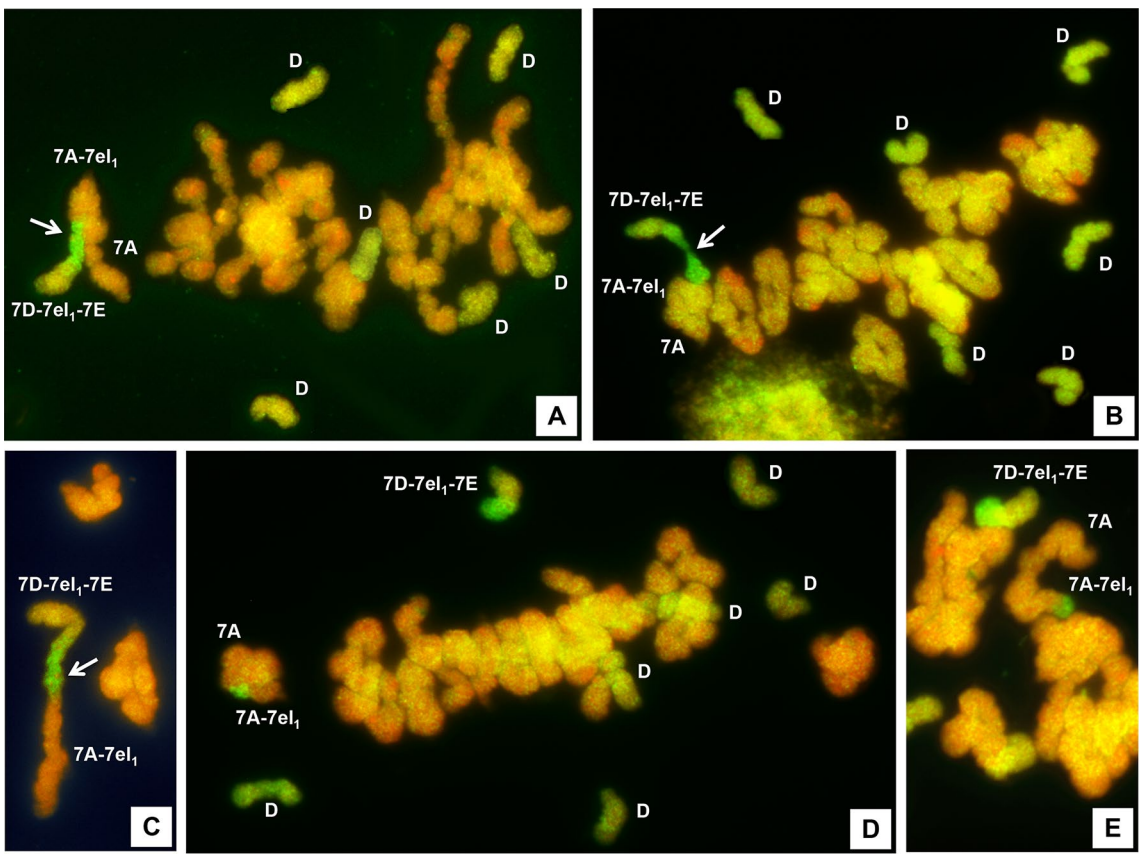


FIGURE 2 | GISH of pollen mother cells (PMCs) at meiotic metaphase I stage of 5x F_1 plants from the cross between R69-9 or R74-10 6x 7D-7el$_1$-7E recombinant lines and R112 or R5 4x 7A-7el$_1$ recombinant lines. Pairing in the 7el$_1$L segments shared by the respective recombinant chromosomes (arrowed) is highlighted by the GISH site (bright green fluorescence) in the open **(A)** or closed **(B)** trivalent, and in the rod bivalent **(C)**. In **(D)** and **(E)**, the R69-9/R74-10 recombinant chromosome is unpaired (univalent), while a ring **(D)** or rod **(E)** bivalent is established between a complete 7A (from the 6x parent) and the 7A-7el$_1$ chromosome from R112/R5. The greenish, univalent chromosomes from the D genome of the 6x parent are indicated **(D)** in plates **(A)**, **(B)** and **(D)**.

TABLE 2 | Meiotic metaphase I pairing behavior of 7D-7el$_1$-7E and 7A-7el$_1$ chromosomes in pollen mother cells (PMCs) of pentaploid hybrids from crosses between 6x recombinants (R74-10 or R69-9) and 4x recombinants (R112 or R5).

Cross combination (6x/4x)	No. PMCs	% 7el$_1$-7el$_1$ pairing	Mode of 7E-7el$_1$ pairing (%)		
			Open trival.	Closed trival.	Rod bival.
R74-10/R112	68	72.5 ± 1.5	55.3 ± 4.2	38.5 ± 0.5	6.2 ± 3.1
R74-10/R5	91	42.2 ± 1.0	63.4 ± 1.9	30.0 ± 1.9	2.2 ± 2.2
R69-9/R112	96	78.3 ± 3.3	72.9 ± 0.4	25.5 ± 1.3	1.6 ± 1.0
R69-9/R5	64	60.7 ± 3.2	60.5 ± 7.2	32.8 ± 4.3	6.7 ± 3.8

Pairing figures are expressed as means ± standard errors; values concerning the percentage of 7el$_1$-7el$_1$ pairing derive from PMCs extracted from 2–3 plants/cross combination.

considerable proportion was of the closed type (**Figure 2D**), evidently resulting from formation of a second chiasma, proximal to that between 7el$_1$L portions, between the 7AL arms. In PMCs where the critical chromosomes, carrying *Thinopyrum* spp. chromatin, were unpaired, the 7D-7el$_1$-7E chromosome of either R69-9 or R74-10 was invariably observed as a univalent (**Figures 2D**, **E**). On the contrary, because of the presence in the same cells of a normal 7A (see above), the 7A-7el$_1$ chromosome of R5 or R112 paired with the latter in over 95% of PMCs, mostly as ring bivalent (60% of cases) rather than as a rod bivalent, both associations clearly marked by the GISH hybridization site of the 7el$_1$L segment of R5 or R112 (**Figures 2D**, **E**, respectively).

Overall, the total amount of 7el$_1$L–7el$_1$L pairing varied in proportion to the length of homologous 7el$_1$L portion shared by the two parental chromosomes in each F_1 type. Pairing frequency (*pf*) was higher in cross combinations involving the longer 7el$_1$L segment of R112 than in those involving R5, and, concomitantly, in combinations where the shorter 7EL segment of R69-9 was involved compared to those including the R74-10 chromosome (**Figure 1**). As a result, higher *pf* were observed between R69-9 or R74-10 and R112 chromosomes (78.3 and 72.5%, respectively; **Table 2**), as compared with those involving R5 (60.7 and 42.2%, respectively).

Isolation of 7EL+7el$_1$L Durum Wheat Recombinants

To isolate recombinant types within progeny of the crosses of cv. Simeto with the various 5x types (considered equivalent to BC_1 to durum wheat of 6x parents; see Materials and Methods), marker-based genotyping was carried out (**Table 1**). In particular, BE405003 was useful at revealing presence of the associated *Fhb-7EL* QTL, and BE445653 or GWM344 confirmed the origin of the proximally adjacent region (7el$_1$L for R69-9, 7EL for R74-10). Afterwards, a PCR assay for the further proximal segment enabled discrimination between parental and recombinant chromosomes, as well as between recombinant types bearing the 7EL+7el$_1$L assembly on 7AL rather than on 7DL (**Figure 1**). Presence/absence of the 7DS marker GWM573 provided a further validation of both recombinant and parental genotypes. In 7A recombinants, the presence of 7el$_1$L target genes, such as *Lr19* and *Psy1* (for the *Yp* trait; see Introduction), was confirmed by the respective markers (**Table 1**). As expected (see Introduction), none of the isolated recombinants showed dissociation of most distal loci with respect to the parental allelic makeup; this indicated, at least at the resolution level allowed by the markers used, that no homoeologous 7EL-7el$_1$L recombination occurred.

A total of 38.3% of recombinant types were isolated in the BC_1 to durum wheat (**Table 3**). The remaining genotypes were prevailingly of the parental type, either R69-9/R74-10 (P1, 24.8%) or R5/R112 (P2, 18.4%). A minor percentage was that of non-recombinant genotypes in which the P1 and P2 chromosomes, due to pairing failure (hence behaving as univalents at meiosis), underwent abnormal segregation, being eventually either both incorporated (P1+P2 types) or excluded ("7A only" types) from gametes. The relative percentage of these abnormal types was expectedly higher in progenies of cross combinations exhibiting the lowest pairing values, i.e., R74-10/R5 and R69-9/R5 (**Table 3**).

TABLE 3 | Recombination frequency and genotypes isolated in the cross progeny to durum wheat cv. Simeto of pentaploid F_1s (6x recombinants, R74-10 or R69-9 × 4x recombinants, R112 or R5).

F_1 hybrid	Progeny types (29 < 2n < 32)								
	No. plants	Recombinants		Rec. frequency (%)		6x parental chromosome	4x parental chromosome	Co-presence	7A only
		7A	7D	TOT.	Gametic (7A)	R74-10 or R69-9 (P1)	R112 or R5 (P2)	P1 + P2	
R74-10/R112	21	4	7	52.4	36.4	3	5	2	–
R74-10/R5	39	2	9	28.2	18.2	6	9	6	7
R69-9/R112	36	4	13	47.2	23.5	16	2	–	1
R69-9/R5	45	4	11	33.3	26.7	10	10	6	4
Total	141	14	40			35	26	14	12
%	100	9.9	28.4			24.8	18.4	9.9	8.5

Recombination frequency (*rf*) resulting from exchanges within the $7el_1L$ shared chromatin between R74-10/R69-9 and R112/R5 showed the expected trend from the *pf* values. Except for the R69-9/R5 *rf*, the other values exceeded the expected 50% of the respective *pf*, probably due to metaphase I observations providing an underestimate of actual (early prophase I) pairing events (see also Gennaro et al., 2012). The *rf* data confirmed the main contribution to $7el_1L$-$7el_1L$ pairing and crossover occurrence of the 5%-long segment differentiating R112 from R5, the highest values corresponding to F_1s containing the R112 chromosome *vs*. those including R5 (**Table 3**).

In the progeny of all cross combinations, recombinant 7D chromosomes (and, to a lesser extent, parental 7D types) prevailed over recombinant 7As. Nonetheless, sufficient representatives of all 7A novel (7EL+$7el_1L$) recombinant types were isolated, with chromosome numbers ranging from 2n = 29 to 2n = 32 (all other genotypes of the BC_1 progeny to durum wheat of 5x F_1s fell within the same range). A further backcross of selected plants (2n = 29, 30 or 31) to the same cv. Simeto brought most of them (over 80%) to the euploid (2n = 28) condition, hence to stabilization of 7A-$7el_1L$-7EL recombinant genotypes in view of further analyses.

A first check of their stability, also aimed at obtaining homozygous recombinant (HOM+) and non-recombinant (HOM–) individuals, was carried out by screening the BC_2F_2 progeny of euploid recombinant plants by a single codominant marker, namely, BE445653 in the case of R69-9 derivatives ($7el_1$ and 7A alleles; **Figure 1**) and GWM344 for R74-10 derivatives (7E and 7A alleles; **Figure 1**). Such markers allowed discrimination of HOM+, HOM–, and heterozygous (HET) segregates, whose ratio was compared with the expected 1:2:1 for normal segregation (**Table 4**). The χ^2 test was in all cases associated with probability (*P*) levels indicative of normal gametic transmission (> 5%), although *P* values were higher for progenies involving the R69-9 chromosome than for those involving R74-10.

Effects of *Fusarium* spp. Infections in the Presence vs. Absence of the *Fhb-7EL* QTL

Reaction to Spike Inoculation With *F. graminearum*

Progression of infection through the three time-points (7, 14, and 21 dpi) following single-floret inoculation with *F. graminearum* unequivocally discriminated carriers [HOM+ segregates of BC_2F_2 progenies from the cross to durum wheat of 5x F_1s, see above, and the CS7E(7D) substitution line] from non-carriers (HOM– segregates, R112+R5, and Simeto) of the *Fhb-7EL* QTL (**Figure 3A**). Regarding the former group, number of diseased florets (NDF) in the HOM+ plants of three tetraploid recombinant genotypes did not exceed 8% even at 21 dpi, with values at this time-point being not significantly different among the three cross combinations. This NDF value was only slightly superior than that recorded at 14 dpi, which in turn was only somewhat higher than that at 7 dpi of the corresponding genotypes. This trend is indicative of a very minor progression of the FHB disease from the inoculation time and site (**Figures 4A**, **B**). Considering altogether the reaction of 4x resistant (*Fhb-7EL*+) *vs*. susceptible (*Fhb-7EL*–) genotypes, the reduction in FHB severity in the former amounted to nearly 93%.

As indicated by the Tukey test ranking (**Figure 3A**), values of NDF expressed by 4x HOM+ segregates at all time-points were not significantly different from those of the 6x CS7E(7D) FHB resistant control. By contrast, in all genotypes known to lack the *Fhb-7EL* QTL infection progress was much faster, reaching the majority and even 100% of florets and spikelets of the inoculated spike between 14 and 21 dpi (**Figures 3A** and **4C–F**).

The conspicuous difference in FHB severity among 4x genotypes sharing a similar background, i.e., that of cv. Simeto, revealed by the NDF parameter was likewise obvious when the seed set and the grain weight of inoculated and non-inoculated spikes of the same plants were measured (**Table 5**). The Tukey test ranking showed that seed set of inoculated spikes of all genotypes carrying the *Fhb-7EL* QTL was significantly greater than that of genotypes lacking the QTL (average 23.4 seeds/spike *vs*. 6.2, respectively), corresponding to 73.5% reduction in seed number/spike in the susceptible plants. In parallel, average TGW calculated for inoculated spikes of FHB resistant 4x genotypes was 33.8 g, in sharp contrast with the 6.7 g average TGW of susceptible genotypes (over 80% reduction). In fact, conspicuous shrivelling (**Figures 4H**, **I**) was consistently observed in the few seeds occasionally produced by severely diseased spikes, whereas no significant alteration of plumpness and weight was detected in seeds of infected spikes of FHB resistant genotypes (**Figure 4G**).

That the defects in grain number and weight were ascribable to the fungal attack is demonstrated by the GNS and TWG values of the remaining (non-inoculated) spikes of the same plants (**Table 5**). Both R112+R5 and Simeto showed a normal seed set in such spikes, which was not significantly different from that of non-inoculated (or even inoculated) spikes of FHB resistant genotypes (average for all genotypes around 23 seeds/spike). Concomitantly, TGW was very similar among genotypes when non-infected spikes were compared (**Table 5**).

UHPLC-MS analyses were performed to quantify the content of DON mycotoxin in flour extracted from mature grains of the 4x FHB resistant recombinants (*Fhb-7EL* HOM+) and susceptible controls (*Fhb-7EL* HOM–). The 3 resistant recombinants, taken as biological replicates of the *Fhb-7EL*+ condition, showed an average value of 0.67 ppm, more than 800 times lower than the 547.4 ppm mean figure of the 3 genotypes representing the *Fhb-7EL*– condition (**Table 6**). No appreciable difference was detected among

TABLE 4 | Segregation ratios of novel 7A recombinant chromosomes in tetraploid BC_2F_2 progenies from crosses to durum wheat (R112 or R5/2*cv. Simeto) of 6x recombinants (R74-10 or R69-9).

Cross combination (6x/4x)	No. BC_2F_2 plants	Segregation			χ^2 (1:2:1)	*P* value (%)
		HOM+	HET	HOM–		
R74-10/R112	61	9	31	21	4.74	9.4
R74-10/R5	54	6	31	17	5.67	5.9
R69-9/R112	60	14	32	14	0.27	87.4
R69-9/R5	88	23	39	26	1.34	51.0

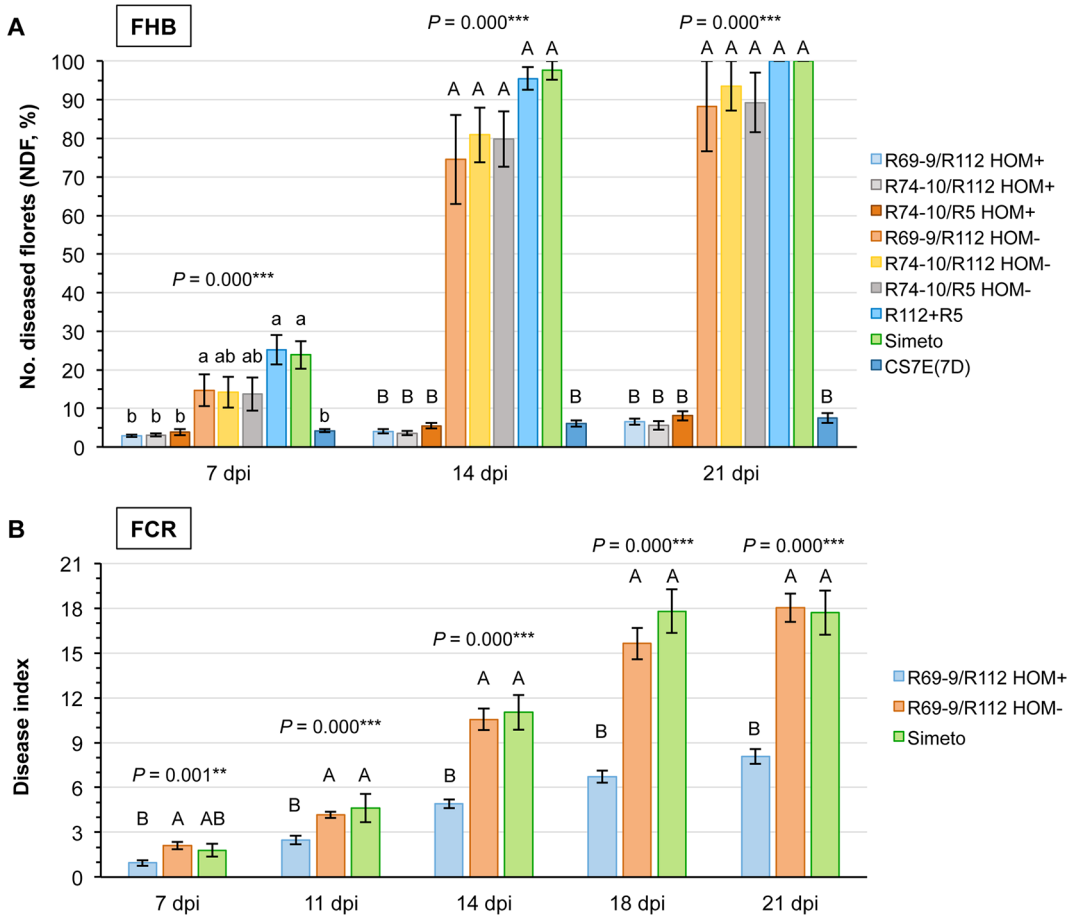


FIGURE 3 | Evaluation of FHB **(A)** and FCR **(B)** symptom development at different time-points following inoculation (dpi = days post-inoculation) in durum wheat homozygous carriers (HOM+) and non-carriers (HOM– segregates and cv. Simeto) of the *Fhb-7EL* QTL. The hexaploid CS7E(7D) original donor line of *Fhb-7EL* is included as FHB resistant control in **(A)**. Data at all time points were subjected to ANOVA analysis, and significant F values indicated by $^{**}P < 0.01$ and $^{***}P < 0.001$, respectively.

HOM+ lines, whereas DON content of different genotypes lacking *Fhb-7EL* varied. The lower DON content exhibited by HOM– segregates relative to R112+R5 and Simeto (with a similar, though not significant trend present in FHB infection data; see **Figure 3A**) is probably due to minor FHB resistance QTL in their background (including CS, from the original donor line of the *Fhb-7EL* QTL; see Materials and Methods and Ceoloni et al., 2017a).

Reaction to Spike Inoculation With *F. culmorum*

To monitor visible disease progress accurately, the time-course of the present FCR infection assay, conducted on seedlings of tetraploid *Fhb-7EL* HOM+, FHB resistant recombinant plants (R69-9/R112 cross derivatives), as well as on FHB susceptible control plants (HOM– sibs and cv. Simeto), was extended of 1 week with respect to a previous experiment (Ceoloni et al., 2017a), with five time-points between inoculation and 21 dpi (**Figure 3B**). The DI values of HOM– (*Fhb-7EL*–) and Simeto plants largely overlapped throughout the time-points, collectively showing a highly significant difference *vs.* *Fhb-7EL*-bearing plants, especially from 11 dpi onward (**Figure 3B**). From this time-point, FCR symptoms increased rapidly in seedlings of genotypes lacking *Fhb-7EL*, reaching average DI values of about 11 at 14 dpi and 17-18 at 18-21 dpi, with peaks of up to 28 recorded (**Figure 5**). By contrast, a much slower progression was observed in *Fhb-7EL* HOM+ plants, exhibiting a maximum DI of around 8 (21 dpi), characterized by limited SE and brown discoloration of the infected tissue (**Figures 3B** and **5**). As a whole, during 14 to 21 dpi, FCR symptom severity was consistently reduced by 55–60% in *Fhb-7EL*+ compared with *Fhb-7EL*– plants. No major disease intensification was observed on the former plants beyond the 21 dpi assessment, while several *Fhb-7EL*– seedlings withered completely (not shown).

Agronomic and Quality Features of Novel Recombinant Genotypes

Sufficient seed was available to run preliminary tests of performance under field conditions of two of the newly obtained durum wheat

FIGURE 4 | Phenotypes of *F. graminearum* inoculated spikes (21 dpi) and of corresponding harvested seeds of durum wheat-*Thinopyrum* spp. 7A-7el$_1$L-7EL recombinant lines and control lines. **(A-B)** FHB resistant (*Fhb-7EL*+) recombinant (HOM+), with arrows pointing at the diseased floret(s)/spikelet(s); fully diseased spikes of HOM– segregates in the same BC_2F_2 progenies from the cross to durum wheat of 5x F_1s **(C-D)** and of durum wheat cv. Simeto **(E-F)**. Mature seeds of HOM+ **(G)**, HOM– **(H)**, and Simeto **(I)** genotypes show a sharp difference in plumpness.

TABLE 5 | Effects of *F. graminearum* infection on fertility traits of mature spikes of durum wheat homozygous carriers (HOM+) and non-carriers (HOM–) of the *Fhb-7EL* QTL.

Genotype	Fhb-7EL	Inoculated spike		Remaining spikes	
		No. seeds	TGW	No. seeds	TGW
R69-9/R112	HOM+	24.0 ± 2.8 A	33.8 ± 2.1 A	23.8 ± 1.4	33.2 ± 2.0
	HOM–	4.6 ± 2.1 B	3.5 ± 1.9 B	21.5 ± 2.4	34.8 ± 3.0
R74-10/R112	HOM+	25.3 ± 1.2 A	32.9 ± 1.3 A	23.5 ± 2.2	34.3 ± 1.9
	HOM–	8.2 ± 2.3 B	6.8 ± 1.0 B	22.1 ± 1.5	32.4 ± 2.6
R74-10/R5	HOM+	21.0 ± 2.1 A	34.6 ± 2.6 A	21.7 ± 1.0	38.3 ± 2.6
	HOM–	2.5 ± 1.2 B	12.2 ± 1.3 B	21.2 ± 2.7	36.8 ± 2.6
R112+R5	HOM–	10.1 ± 1.8 B	5.7 ± 1.1 B	23.0 ± 1.3	38.9 ± 2.1
Simeto	HOM–	5.6 ± 1.1 B	5.1 ± 0.9 B	28.4 ± 1.1	35.9 ± 2.9
ANOVA *P*-value		0.000***	0.000***	0.062	0.439

*Values are expressed as means ± standard errors; letters indicate ranking of the Tukey test at P < 0.01; *** indicates significant F values at P < 0.001. The first three sets of genotypes (HOM+/HOM–) are BC_2F_2 segregates from the cross of R69-9 or R74-10 with T. durum cv. Simeto background (see text).*

recombinant lines, deriving from the R69-9/R112 and R74-10/R112 cross combinations. The small-scale trials (see *Materials and Methods*) included F_{3-4} HOM+ and HOM– sib plants from R74-10 and R69-9/R112/2*Simeto crosses, as well as the recurrent cv. Simeto. In both experimental years, no major penalty on yield-related traits was found to be associated with presence of alien segments. In the 1st season (2017–18), yield-related traits such as spike number, grain number, and grain yield per plant (SNP, GNP, and GYP; **Table 7A**) gave higher values in both HOM+ lines *vs*. their respective HOM– sibs and Simeto, although ANOVA showed differences to be significant only for SNP. As to spike traits, SPN differed significantly among genotypes, though not in a clear-cut relation with presence/absence of any alien introgression. GNS and SFI also showed some variation among genotypes (**Table 7A**). However, for most spike traits, differences became highly significant and genotype-dependent in the 2018–19 season (**Table 7B**). In particular, both recombinant (HOM+) genotypes outperformed their HOM– sibs for GNS and GNSP. A positive background effect was evident in HOM+ and HOM– R69-9/R112 selections, resulting in significantly higher SPN and GYS values compared with all other genotypes (**Table 7B**).

A mild leaf rust attack occurred during the 2017–18 season. Despite this, the pathogen produced visible pustules on HOM– and Simeto plants, while leaving HOM+ sibs, carriers of the

TABLE 6 | Deoxynivalenol (DON) content in wholemeal flour from seeds of infected spikes of carrier (HOM+) and non-carrier (HOM–) genotypes of the *Fhb-7EL* QTL.

Genotype	Fhb-7EL	DON (ppm)	
R69-9/R112	HOM+	0.47 ± 0.1 c	
R74-10/R112	HOM+	0.66 ± 0.0 c	0.673 ± 0.1
R74-10/R5	HOM+	0.89 ± 0.1 c	
Null segregates	HOM–	176.64 ± 19.3 b	
R112+R5	HOM–	685.97 ± 44.9 a	547.4 ± 96.8
Simeto	HOM–	779.73 ± 69.5 a	
ANCOVA *P*-value		0.000***	0.000***

*DON values are expressed as means ± standard errors; those regarding individual genotypes are followed by letters corresponding to ranking of the Tukey test at P < 0.05; values reported in the last column refer to the 3 HOM+ and the 3 HOM– genotypes, taken as biological replicates (see Materials and Methods). *** indicates significant F values at P < 0.001. Genotypes are the same as described in* **Table 5**.

Lr19 gene within their $7el_1L$ segments (**Figure 1**), totally rust-free. In the 2018–19 season, a stronger natural infection took place, which allowed clearer discrimination among genotypes. As values were rather consistent within each genotype, a single double-digit record has been reported/genotype (**Table 7B**), corresponding to the peak-time of the disease progress. No disease symptom was recorded on the novel recombinant types, whereas in their HOM– sibs and Simeto the infection reached the flag leaf (score "8") or the penultimate leaf (score "7"), with pustules covering 50–60% of the leaf area ("5" and "6" second digit in **Table 7B**). In the test environment, there was no evidence of FHB presence in the 1st season, and only a sporadic appearance in the 2nd one, which, however, did not involve any of the materials under assay. In the absence of stem rust epidemics, the presence of the $7el_1L$-linked *Sr25*, to be excluded in R74-10 derivatives on the basis of mapping data (Ceoloni et al., 2017a), remains to be ascertained in R69-9 derivatives.

A highly significant difference was revealed by the colorimetric test for the semolina YI of genotypes, alternatively carrying the *Psy1-7el$_1$L* (R69-9/R112 HOM+) or the *Psy1-7EL* (R74-10/R112 HOM+) allele in place of a *Psy1-7AL* allele (HOM– lines and Simeto). Presence of *Psy1-7el$_1$L* from *Th. ponticum* determined a 37–42% increase compared with all other genotypes, while the *Th. elongatum Psy1-7EL* allele had no incremental effect *vs*. the *T. durum* 7AL resident allele (**Table 7B**).

DISCUSSION

Effectiveness of the Transfer Strategy

In the present study, we successfully exploited meiotic recombination, confined to a homologous *Th. ponticum*

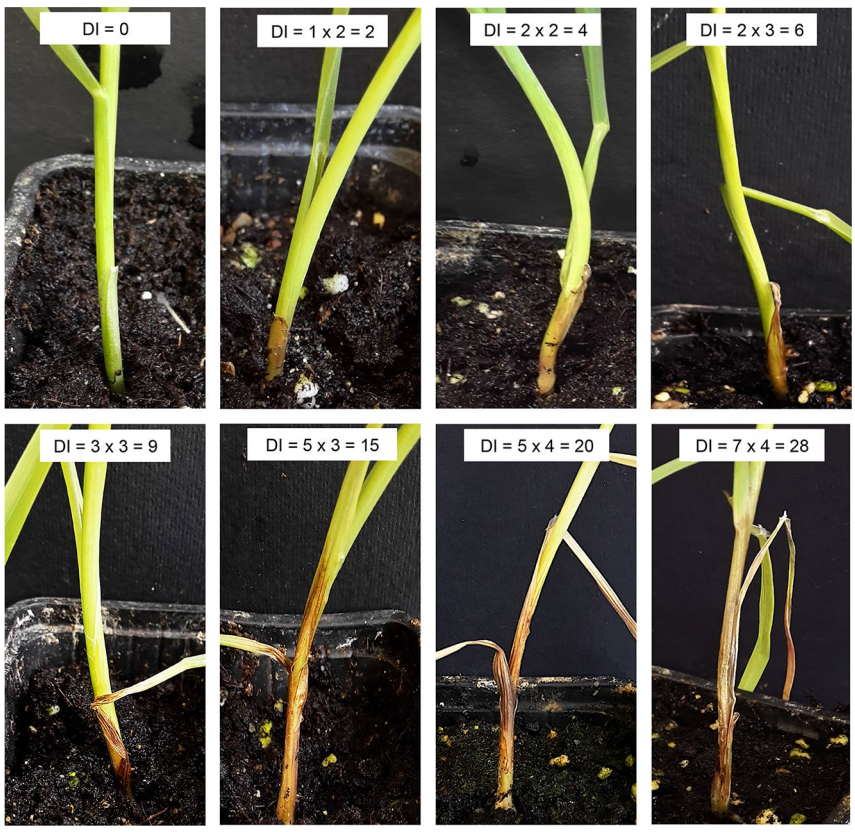


FIGURE 5 | Examples of FCR disease symptoms recorded on seedling stem base leaf sheath of R69-9/R112/2*Simeto derivatives. DI, disease index = SE (symptom extension, cm) x BI (browning index; see Materials and Methods for further details.

TABLE 7 | Spike and plant traits of field-grown homozygous (HOM+) recombinants (R74-10 or R69-9/R112/2*Simeto F_{3-4} derivatives) compared to corresponding HOM– segregates and to the recurrent cv. Simeto.

Trait	R74-10/R112 HOM+	R74-10/R112 HOM–	R69-9/R112 HOM+	R69-9/R112 HOM–	Simeto	ANOVA *P* value
A. 2017–18						
SNP	7.6 ± 1.0	5.1 ± 0.5	6.2 ± 0.4	5.3 ± 0.6	4.9 ± 0.6	0.038*
GNP	231.4 ± 32.7	176.0 ± 15.6	184.0 ± 13.2	162.7 ± 14.0	148.0 ± 18.7	0.087
GYP	11.1 ± 1.5	8.8 ± 0.8	8.7 ± 0.7	7.6 ± 0.7	7.6 ± 1.1	0.128
GNS	47.6 ± 2.2	46.1 ± 2.3	52.1 ± 1.6	53.5 ± 2.3	45.5 ± 1.8	0.018*
SPN	17.7 ± 0.4 bc	16.8 ± 0.4 c	18.3 ± 0.3 b	19.7 ± 0.3 a	16.9 ± 0.3 c	0.000***
GNSP	2.7 ± 0.1	2.7 ± 0.1	2.8 ± 0.1	2.7 ± 0.1	2.7 ± 0.1	0.601
SFI	56.4 ± 2.0	53.1 ± 2.2	57.1 ± 1.2	61.1 ± 1.5	54.2 ± 1.7	0.038*
TGW	48.2 ± 0.7	50.2 ± 1.2	47.3 ± 0.8	46.9 ± 1.3	50.5 ± 1.4	0.101
GYS	2.4 ± 0.1	2.6 ± 0.1	2.6 ± 0.1	2.7 ± 0.1	2.5 ± 0.1	0.401
HD	117.3 ± 0.3	117.6 ± 0.4	118.7 ± 0.5	118.4 ± 0.3	119.0 ± 0.6	0.055
PH	70.1 ± 1.2 bc	68.3 ± 1.3 c	72.5 ± 0.8 ab	68.2 ± 1.4 c	75.1 ± 1.1 a	0.000***
B. 2018–19						
GNS	49.5 ± 1.2 c	44.2 ± 1.5 d	61.2 ± 1.0 a	54.9 ± 1.3 b	49.7 ± 0.9 c	0.000***
SPN	17.9 ± 0.2 b	17.3 ± 0.3 bc	19.2 ± 0.2 a	20.0 ± 0.2 a	16.9 ± 0.2 c	0.000***
GNSP	2.8 ± 0.1 b	2.5 ± 0.0 c	3.2 ± 0.1 a	2.7 ± 0.1 bc	2.9 ± 0.0 b	0.000***
SFI	46.5 ± 1.2 b	50.7 ± 1.5 ab	55.3 ± 1.3 a	52.7 ± 1.7 a	53.0 ± 1.4 a	0.001**
TGW	53.9 ± 0.8	52.2 ± 1.5	49.6 ± 1.0	51.8 ± 0.2	52.6 ± 0.7	0.063
GYS	2.7 ± 0.1 bc	2.3 ± 0.1 b	3.0 ± 0.1 a	2.8 ± 0.1 ab	2.6 ± 0.1 c	0.000***
HD	112.4 ± 0.5 b	115.3 ± 0.5 a	115.4 ± 0.2 a	114.5 ± 0.6 a	114.5 ± 0.3 a	0.001**
PH	88.9 ± 1.3 ab	81.7 ± 1.8 c	89.9 ± 1.2 a	83.7 ± 1.6 bc ab	87.5 ± 1.6 abc	0.001**
YI	22.1 ± 0.5 B	22.2 ± 0.4 B	31.4 ± 0.5 A	23.0 ± 0.3 B	22.9 ± 0.2 B	0.000***
LR	0	8-5	0	7-6	8-5	–

*SNP, spike number/plant; GNP, grain number/plant; GYP, grain yield/plant; GNS, grain number/spike; GYS, grain yield/spike; SPN, spikelet number/spike; GNSP, grain number/spikelet; SFI, spike fertility index; TGW, thousand grain weight; HD, days to heading (from January 1st); PH, plant height; YI, yellow index; LR, leaf rust. Except for LR, trait values are given as means ± standard errors; in case of significant differences among genotypes, these are followed by letters corresponding to ranking of the Tukey test at P < 0.01 (capital) and P < 0.05 (lower case) levels. *, **, *** indicate significant F values at P < 0.05, P < 0.01, and P < 0.001, respectively.*

$7el_1$ chromosome segment shared by two selected pairing partners, to create a new pyramid of positive alien genes/QTL, including a potent *Fusarium* spp. resistance locus, into durum wheat chromosome 7A. Previous studies (e.g., Ceoloni and Jauhar, 2006 for review) have widely demonstrated that homoeologous pairing-based wheat-alien chromosome engineering carried out at the tetraploid level, i.e., with durum wheat as the primary recipient crop species, leads to much less success than when hexaploid bread wheat is targeted. Besides the overall reduced tolerance to chromosome manipulations associated to the lower ploidy level, a further limiting factor is represented by closer affinity between certain alien genomes, such as those of some widely exploited *Thinopyrum* species, to the wheat D genome, compared with A and B genomes (see, e.g., Forte et al., 2014; Ceoloni et al., 2017a, and references therein). This results in excellent performance of corresponding recombinant products even in the presence of sizable introgressions. The latter case is well exemplified by the bread wheat T4 translocation line, widely used in breeding (reviewed in Ceoloni et al., 2015).

In the present work, the recently obtained T4 derivatives R74-10 and R69-9 (Ceoloni et al., 2017a), containing the *Fhb-7EL* resistance QTL at the distal end of their 7DL-$7el_1$L arm (i.e., 7DL-$7el_1$L-7EL), were selected to transfer *Fhb-7EL* into the 7AL-$7el_1$L durum wheat recombinant chromosomes of R5 and R112 lines (Ceoloni et al., 2005). The choice of parental recombinant types was inherent to their structure; in fact, they could provide the physical basis for spontaneous pairing and recombination to occur in the common $7el_1$L portion to their otherwise homoeologous target chromosomes, both in the most distal end of the same arm (7EL *vs.* $7el_1$L) and in the remaining portions (7D *vs.* 7A; see **Figure 1**). Even the limited extension of the shared $7el_1$L segment in all R74-10/R69-9 with R5/R112 chromosome combinations turned out to be sufficient to recover the novel 7EL+$7el_1$L recombinant types at a relatively high rate (**Table 3**). Although never tested previously in the same chromosomal and genomic context as presented here, this result was not totally unexpected. As a matter of fact, in wheat and in many other species, the distribution of pairing and crossover (CO) events follows a telomere-to-centromere gradient, with concentration of such events in the distal half or even less of the physical arm length, both between homologous and homoeologous chromosomes (Lukaszewski and Curtis, 1993; Lukaszewski, 1995; Saintenac et al., 2009, Saintenac et al., 2011; Higgins et al., 2012; Darrier et al., 2017; Jordan et al., 2018). In general terms, the location of the shared $7el_1$L segment in the different cross combinations could be considered to fall within the high recombinogenic chromosomal space. Moreover, the virtual absence of pairing in the homoeologous most terminal 7EL-$7el_1$L regions in all 5x F_1s, accompanied by further interruption of homology in the more proximal arm portions of the same pairing partners, evidently favored pairing and recombination in the only $7el_1$L homologous interval available to the respective parental recombinant chromosomes. In this respect, several examples have demonstrated dramatic effects on pairing and CO frequency and

distribution as a result of regional differences in the structure of potential pairing partners, mainly at the telomeric ends (see, e.g., Lukaszewski et al., 2004 and references therein). Furthermore, the current results indicate a particularly high propensity for pairing and CO of the roughly 5% 7el$_1$L chromatin differentiating R112 from R5 (see crosses with R112 of either R69-9 or R74-10 in **Table 3**), even irrespective of the somewhat wider space in the more distal vicinity, as in corresponding crosses with R5. Interestingly, this is the interval within which 7AL-7el$_1$L *ph1*-induced homoeologous pairing gave rise to three recombination products (one being R112), compared with two recovered in the same progeny in the more distal region, spanning the remaining 23% telomeric end of the arm (Ceoloni et al., 2005). For its consistent behavior in homologous and homoeologous contexts, the 5% 7el$_1$L stretch included in R112 appears as a recombination hotspot, similar to several others frequently mapped to subterminal regions in wheat and related Triticeae chromosomes (e.g., Saintenac et al., 2011; Zhang et al., 2018).

Efficacy of the *Fhb-7EL* QTL and Value of Novel Recombinant Types

In view of its exploitation in durum wheat breeding, verification of the full expression of the *Fhb-7EL* resistance QTL into the target species background was an essential step. All evaluation parameters, from assessment of FHB severity following controlled inoculations, to measurement of seed setting and development and, importantly, quantitation of DON content, provided consistent evidence of remarkably high reduction of all symptoms and effects of *F. graminearum* infection in genotypes carrying *Fhb-7EL* as compared with non-carrier lines. As confirmed by inclusion of the CS7E(7D) bread wheat substitution line in the infection assay (**Figure 3A**), the over 90% reduction of the FHB severity in inoculated spikes of novel 4x recombinant lines was of the same extent than that observed in the bread wheat background (**Figure 3A** and Ceoloni et al., 2017a), and even higher than that provided to both 6x and 4x wheat by the *Th. ponticum* 7el$_2$ QTL (see also Introduction), averaging 80% (Forte et al., 2014). Transfer of the *Fhb1* major resistance QTL from Sumai 3 into *T. durum* cultivars led to a reduction of FHB severity from 6 to 36%, depending on the background (Prat et al., 2017). Moreover, undesirable effects on agronomic traits were reported when using Sumai 3 in breeding efforts (reviewed in Gilbert and Haber, 2013). On the other hand, particularly accurate and complex selection strategies are needed to effectively exploit multiple small-effect QTL (Tuberosa and Pozniak, 2014; Miedaner et al., 2017; Sari et al., 2018; Steiner et al., 2019).

By contrast, the completely dominant expression of a single major QTL, for whose selection a single PCR assay is sufficient, undoubtedly represents the ideal, breeder-friendly situation, and this is offered by the *Fhb-7EL* QTL. This locus has been shown to confer the same protection against FHB to different bread wheat lines, from the standard CS to the Italian elite cultivar Blasco (Ceoloni et al., 2017a), and to unrelated durum wheat genotypes, such as the Italian cv. Simeto (this work) and Langdon, an old North Dakota variety and laboratory line to which the complete chromosome 7E was recently added (Liu et al., 2017). Moreover, comparing the resistance demonstrated by the latter work with the results presented here confirms that the exceptional FHB resistance associated with *Th. elongatum* 7E chromosome is completely determined by the *Fhb-7EL* locus previously mapped to the distal end of 7EL (Ceoloni et al., 2017a).

Visual assessments of head blight were fully consistent with the prominently reduced accumulation of the DON toxin (**Table 6**), hence considerably reducing the health risk from FHB infection of the novel materials. Whether the *Fhb-7EL*-linked low DON content is due to its more efficient *in planta* conversion into the less active DON-3-glucoside derivative, which is identified as the main detoxification strategy in wheat and correlated to the *Fhb1* resistant response (e.g., Kluger et al., 2015; Lemmens et al., 2016; Mandalà et al., 2019 and references therein), or to alternative mechanisms (e.g., Miller et al., 2011), remains to be elucidated. Near-isogenic lines of durum wheat recombinants with and without *Fhb-7EL*, currently under development, will be ideal tools for comparative analyses aimed at elucidating the mechanism(s) of action underlying this unique resistance gene.

Among the intriguing characteristics of the *Fhb-7EL* QTL is its efficacy also toward another important Fusarium disease, i.e., crown rot (FCR), both in bread wheat (Ceoloni et al., 2017a) and in durum wheat (this work). Particularly prevalent in semi-arid regions amenable to the latter crop, FCR is increasingly showing an upsurge in incidence and severity in durum wheat, thereby causing even higher yield losses than in other susceptible cereals (GRDC Grains Research and Development Corporator, 2009; Fernandez and Conner, 2011; Scherm et al., 2013; Chekali et al., 2016). Preliminary evidence on durum wheat recombinant lines carrying the FHB resistance QTL from *Th. ponticum* 7el$_2$ (Forte et al., 2014) similarly showed that QTL to confer resistance also to FCR (Ceoloni et al., unpublished results). Effectiveness toward both diseases, incited by different *Fusarium* species (*F. graminearum*, *F. culmorum* and *F. pseudograminearum*; see also Ceoloni et al., 2017a), is an exceptional attribute of the *Thinopyrum* spp. QTL, not paralleled by the situation in wheat germplasm, within which such a genetic and phenotypic coincidence finds no clear-cut example (Li et al., 2010). The largely comparable phenotype, combined with the corresponding location at the most distal end on the respective arms, 7el$_2$L and 7EL (see Forte et al., 2014 and Ceoloni et al., 2017a), suggests the *Th. ponticum* and *Th. elongatum* Fusarium resistance QTL to be orthologous. Whereas high resolution maps of the respective chromosomal regions will be a necessary tool to verify this hypothesis, comparison of the gene content of the distal portions of *Th. ponticum* 7elL and *Th. elongatum* 7EL reveals additional similarities, including a *Psy1* gene, a likely candidate for the "yellow pigment" phenotype, common to 7el$_1$L, 7el$_2$L and 7EL (Forte et al., 2014, Ceoloni et al., 2017a; see also Introduction), and *Sd* (segregation distortion) genes spread along the arms, particularly in their proximal halves (Ceoloni et al., 2014b, Ceoloni et al., 2017b).

Regarding the effect of *Thinopyrum Psy1* alleles, this work has offered for the first time the possibility to assess the relative strength of *Psy1-7el$_1$L* from *Th. ponticum* and *Psy1-7EL* from *Th. elongatum* once inserted into durum wheat. In contrast to what observed at the bread wheat level, where both contributed to a YPC increase, with the former providing a more conspicuous effect than the latter

(Ceoloni et al., 2017a), only *Psy1-7el$_1$L* was found to determine a significant increment of semolina YI in durum wheat. A likely explanation for what resulted in the different species contexts could be that the effect of the weaker *Psy1-7EL* allele can be detected when it replaces the non-contributing *Psy-D1a* allele, foremost widespread in bread wheat worldwide collections (Ravel et al., 2013), but not when it substitutes for alleles at the *Psy-A1* locus, as in the present durum wheat recombinant lines. Both in bread wheat (Ravel et al., 2013) and in durum wheat (Pozniak et al., 2007) major QTL for YPC have been mapped on 7AL and 7BL arms, which co-locate with *Psy1* alleles. The evidently stronger effect on semolina yellowness of *Psy1-7el$_1$L* over the resident *Psy1-7AL* allele (see also Gennaro et al., 2007) confers to the R69-9 derivatives a particularly desirable attribute for transformation into pasta products.

The preliminary field trials showed no penalty on yield-related traits at the plant and spike levels associated with presence of either 7el$_1$L+7EL segment. Instead, positive effects on spike fertility traits of both R74-10/R112 and R69-9/R112 recombinant lines (notably grain number per spike and per spikelet) were mostly evident in the 2nd year trial (**Table 7B**). In both experimental seasons, the presence of the 7el$_1$L leaf rust resistance gene *Lr19*, initially tracked by the STS*Lr19*$_{130}$ closely linked marker, was validated in field grown plants. Its remarkable and durable efficacy is an additional, important asset in sustainable breeding.

While larger-scale and multi-location field trials are planned to better evaluate yield-related characteristics of all novel recombinant types, their highly valuable package of genes/QTL has already prompted marker-assisted crossing programs to incorporate the composite *Thinopyrum* segment into elite durum wheat varieties of different geographical origin. Further, the R5- or R112-type segments, involving a 7A chromosome, are being transferred into bread wheat as well, to evaluate their relative performance as compared with that of 7D recombinants previously engineered with the same 7EL portions but in a much longer 7el$_1$L segment (Ceoloni et al., 2017a).

In conclusion, the chromosome engineering work described here marks a significant step forward in equipping durum wheat with highly desirable attributes, primarily the largely missing resistance to Fusarium diseases, which can sustainably enhance security and safety, as well as market and trade values of this important crop.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/supplementary material.

AUTHOR CONTRIBUTIONS

LK carried out most of the research, contributing to germplasm development, molecular and phenotypic selection and cytogenetic characterization; GM and ST performed the *Fusarium* spp. infection assays and analysed the data; RC and MF contributed to germplasm development and screening; RR and FR carried out the field test, the agronomic evaluations and statistical analyses; FG and SR performed the DON analysis. CC conceived the project, coordinated the research and prepared the manuscript. All authors read and approved the final manuscript.

FUNDING

The research was partially supported by MIUR (Italian Ministry for education, University and Research) in the context of the initiative "Departments of excellence" (law 232/216), and by Lazio region, FILAS project "MIGLIORA".

ACKNOWLEDGMENTS

The excellent technical assistance of Mrs. Alessandra Bitti is gratefully acknowledged. Special thanks are due to Dr. Alessandro Cammerata, CREA - Centro di ricerca Ingegneria e Trasformazioni agroalimentari, Roma, Italy, for having performed the semolina colorimetric assay, and to Prof. S. A. Quarrie (Belgrade and Newcaste Universities) for critically reviewing the manuscript.

REFERENCES

Anamthawat-Jónsson, K., and Reader, S. M. (1995). Pre-annealing of total genomic DNA probes for simultaneous genomic in situ hybridization. *Genome* 38, 814–816. doi: 10.1139/g95-104

Bakker, M. G., Brown, D. W., Kelly, A. C., Kim, H.-S., Kurtzman, C. P., Mccormick, S. P., et al. (2018). Fusarium mycotoxins: a trans-disciplinary overview. *Can. J. Plant Pathol.* 40, 161–171. doi: 10.1080/07060661.2018.1433720

Bassi, F. M., and Sanchez-Garcia, M. (2017). Adaptation and stability analysis of ICARDA durum wheat elites across 18 countries. *Crop Sci.* 57, 2419–2430. doi: 10.2135/cropsci2016.11.0916

Beccari, G., Covarelli, L., and Nicholson, P. (2011). Infection processes and soft wheat response to root rot and crown rot caused by *Fusarium culmorum*. *Plant Pathol.* 60, 671–684. doi: 10.1111/j.1365-3059.2011.02425.x

Ceoloni, C., and Jauhar, P. P. (2006). "Chromosome engineering of the durum wheat genome: Strategies and applications of potential breeding value," in *Genetic Resources, Chromosome Engineering, and Crop Improvement: Cereals*. Eds. R. J. Singh and P. P. Jauhar (Boca Raton, FL, USA: CRC Press), 27–59, ISBN: ISBN-10: 0-8493-1432-1. doi: 10.1201/9780203489260.ch2

Ceoloni, C., Forte, P., Gennaro, A., Micali, S., Carozza, R., and Bitti, A. (2005). Recent developments in durum wheat chromosome engineering. *Cytogenet. Genome Res.* 109, 328–344. doi: 10.1159/000082416

Ceoloni, C., Kuzmanović, L., Forte, P., Gennaro, A., and Bitti, A. (2014a). Targeted exploitation of gene pools of alien Triticeae species for sustainable and multi-faceted improvement of the durum wheat crop. *Crop Pasture Sci.* 65, 96–111. doi: 10.1071/CP13335

Ceoloni, C., Kuzmanović, L., Gennaro, A., Forte, P., Giorgi, D., Grossi, M. R., et al. (2014b). "Genomes, chromosomes and genes of perennial Triticeae of the genus *Thinopyrum*: the value of their transfer into wheat for gains in cytogenomic knowledge and 'precision' breeding," in *Advances in Genomics of Plant Genetic Resources*. Eds. R. Tuberosa, A. Graner, and E. Frison (Dordrecht, The Netherlands: Springer), 333–358. doi: 10.1007/978-94-007-7575-6_14

Ceoloni, C., Kuzmanović, L., Forte, P., Virili, M. E., and Bitti, A. (2015). "Wheat-perennial Triticeae introgressions: Major achievements and prospects," in *Alien Introgression in Wheat -Cytogenetics, Molecular Biology, and Genomics*. Eds. M.

Molnár-Láng, C. Ceoloni, and J. Doležel (Switzerland: Springer International Publishing), 273–313. doi: 10.1007/978-3-319-23494-6_11

Ceoloni, C., Forte, P., Kuzmanović, L., Tundo, S., Moscetti, I., De Vita, P., et al. (2017a). Cytogenetic mapping of a major locus for resistance to Fusarium head blight and crown rot of wheat on *Thinopyrum elongatum* 7EL and its pyramiding with valuable genes from a *Th. ponticum* homoeologous arm onto bread wheat 7DL. *Theor. Appl. Genet.* 130, 2005–2024. doi: 10.1007/s00122-017-2939-8

Ceoloni, C., Kuzmanović, L., Ruggeri, R., Rossini, F., Forte, P., Cuccurullo, A., et al. (2017b). Harnessing genetic diversity of wild gene pools to enhance wheat crop production and sustainability: challenges and opportunities. *Diversity* 9, 55. doi: 10.3390/d9040055

Chakraborty, S. (2013). Migrate or evolve: options for plant pathogens under climate change. *Global Change Biol.* 19, 1985–2000. doi: 10.1111/gcb.12205

Chekali, S., Gargouri, S., Rezgui, M., Paulitz, T., and Nasraoui, B. (2016). Impacts of previous crops on Fusarium foot and root rot, and on yields of durum wheat in North West Tunisia. *Phytopathol. Mediterr.* 55, 253–261. doi: 10.14601//Phytopathol_Mediterr-17933

Darrier, B., Rimbert, H., Balfourier, F., Pingault, L., Josselin, A. A., Servin, B., et al. (2017). High-resolution mapping of crossover events in the hexaploid wheat genome suggests a universal recombination mechanism. *Genetics* 206, 1373–1388. doi: 10.1534/genetics.116.196014

Dellaporta, S. L., Wood, J., and Hicks, J. B. (1983). A plant DNA minipreparation: version II. *Plant Mol. Biol. Rep.* 1, 19–21. doi: 10.1007/BF02712670

Dvorak, J. (1975). Meiotic pairing between single chromosomes of diploid *Agropyron elongatum* and decaploid *A. elongatum* in *Triticum aestivum*. *Can. J. Genet. Cytol.* 17, 329–336. doi: 10.1139/g75-044

Fernandez, M. R., and Conner, R. L. (2011). Root and Crown Rot of Wheat. *Prairie Soils Crops J.* 4, 151–157.

Ficco, D. B. M., Mastrangelo, A. M., Trono, D., Borrelli, G. M., De Vita, P., Fares, C., et al. (2014). The colours of durum wheat: a review. *Crop Pasture Sci.* 65, 1–15. doi: 10.1071/CP13293

Fones, H. N., and Gurr, S. J. (2017). NO_xious gases and the unpredictability of emerging plant pathogens under climate change. *BMC Biol.* 15, 36. doi: 10.1186/s12915-017-0376-4

Forte, P., Virili, M. E., Kuzmanović, L., Moscetti, I., Gennaro, A., D'Ovidio, R., et al. (2014). A novel assembly of *Thinopyrum ponticum* genes into the durum wheat genome: pyramiding Fusarium head blight resistance onto recombinant lines previously engineered for other beneficial traits from the same alien species. *Mol. Breed.* 34, 1701–1716. doi: 10.1007/s11032-014-0175-3

Gennaro, A., Borrelli, G. M., D'Egidio, M. G., De Vita, P., Ravaglia, S., and Ceoloni, C. (2003). "A chromosomally engineered durum wheat-*Thinopyrum ponticum* recombinant line with novel and promising attributes for varietal development," in *Proc. 10th Int. Wheat Genetics Symp*. Eds. N. E. Pogna, M. Romanò, E. A. Pogna, and G. Galterio (Italy, S.I.M.I., Rome: Paestum), 881–883.

Gennaro, A., Forte, P., Carozza, R., Savo Sardaro, M. L., Ferri, D., Bitti, A., et al. (2007). Pyramiding different alien chromosome segments in durum wheat: Feasibility and breeding potential. *Isr. J. Plant Sci.* 55, 267–276. doi: 10.1560/IJPS.55.3-4.267

Gennaro, A., Forte, P., Panichi, D., Lafiandra, D., Pagnotta, M. A., D'Egidio, M. G., et al. (2012). Stacking small segments of the 1D chromosome of bread wheat containing major gluten quality genes into durum wheat: transfer strategy and breeding prospects. *Mol. Breed.* 30, 149–167. doi: 10.1007/s11032-011-9606-6

Giancaspro, A., Giove, S. L., Zito, D., Blanco, A., and Gadaleta, A. (2016). Mapping QTLs for Fusarium head blight resistance in an interspecific wheat population. *Front. Plant Sci.* 7, 1381. doi: 10.3389/fpls.2016.01381

Gilbert, J., and Haber, S. (2013). Overview of some recent developments in fusarium head blight of wheat. *Can. J. Plant Pathol.* 35, 149–174. doi: 10.1080/07060661.2013.772921

GRDC Grains Research & Development Corporation, (2009). Crown rot in cereals – fact sheet. https://grdc.com.au/__data/assets/pdf_file/0021/210918/grdc-crown-rot-fact-sheet-southern-and-western-regions.pdf.pdf.

Guo, J., He, F., Cai, J.-J., Wang, H.-W., Li, A.-F., Wang, H.-G., et al. (2015a). Molecular and cytological comparisons of chromosomes $7el_1$, $7el_2$, $7E^e$, and $7E^i$ derived from *Thinopyrum*. *Cytogenet. Genome Res.* 145, 68–74. doi: 10.1159/000381838

Guo, J., Zhang, X., Hou, Y., Cai, J., Shen, X., Zhou, T., et al. (2015b). High-density mapping of the major FHB resistance gene *Fhb7* derived from *Thinopyrum ponticum* and its pyramiding with *Fhb1* by marker-assisted selection. *Theor. Appl. Genet.* 128, 2301–2316. doi: 10.1007/s00122-015-2586-x

Higgins, J. D., Perry, R. M., Barakate, A., Ramsay, L., Waugh, R., Halpin, C., et al. (2012). Spatiotemporal asymmetry of the meiotic program underlies the predominantly distal distribution of meiotic crossovers in barley. *Plant Cell* 24, 4096–4109. doi: 10.1105/tpc.112.102483

Jordan, K. W., Wang, S., He, F., Chao, S., Lun, Y., Paux, E., et al. (2018). The genetic architecture of genome-wide recombination rate variation in allopolyploid wheat revealed by nested association mapping. *Plant J.* 95, 1039–1054. doi: 10.1111/tpj.14009

Kluger, B., Bueschl, C., Lemmens, M., Michlmayr, H., Malachova, A., Koutnik, A., et al. (2015). Biotransformation of the mycotoxin deoxynivalenol in fusarium resistant and susceptible near isogenic wheat lines. *PLoS One* 10, e0119656. doi: 10.1371/journal.pone.0119656

Kuzmanović, L., Gennaro, A., Benedettelli, S., Dodd, I. C., Quarrie, S. A., and Ceoloni, C. (2014). Structural-functional dissection and characterization of yield-contributing traits originating from a group 7 chromosome of the wheatgrass species *Thinopyrum ponticum* after transfer into durum wheat. *J. Exp. Bot.* 65, 509–525. doi: 10.1093/jxb/ert393

Kuzmanović, L., Ruggeri, R., Virili, M. E., Rossini, F., and Ceoloni, C. (2016). Effects of *Thinopyrum ponticum* chromosome segments transferred into durum wheat on yield components and related morpho-physiological traits in Mediterranean rain-fed conditions. *Field Crops Res.* 186, 86–98. doi: 10.1016/j.fcr.2015.11.007

Kuzmanović, L., Ruggeri, R., Able, J. A., Bassi, F. M., Maccaferri, M., Tuberosa, R., et al. (2018). Yield of chromosomally engineered durum wheat-*Thinopyrum ponticum* recombinant lines in a range of contrasting rain-fed environments. *Field Crops Res.* 228, 147–157. doi: 10.1016/j.fcr.2018.08.014

Lemmens, M., Steiner, B., Sulyok, M., Nicholson, P., Mesterhazy, A., and Buerstmayr, H. (2016). Masked mycotoxins: does breeding for enhanced Fusarium head blight resistance result in more deoxynivalenol-3-glucoside in new wheat varieties? *World Mycotoxin J.* 9, 741–754. doi: 10.3920/WMJ2015.2029

Li, H. B., Xie, G. Q., Ma, J., Liu, G. R., Wen, S. M., Ban, T., et al. (2010). Genetic relationships between resistances to Fusarium head blight and crown rot in bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 121, 941–950. doi: 10.1007/s00122-010-1363-0

Liu, C. J., and Ogbonnaya, F. C. (2015). Resistance to Fusarium crown rot in wheat and barley: a review. *Plant Breed.* 134, 365–372. doi: 10.1111/pbr.12274

Liu, H., Dai, Y., Chi, D., Huang, S., Li, H., Duan, Y., et al. (2017). Production and molecular cytogenetic characterization of a durum wheat-*Thinopyrum elongatum* 7E disomic addition line with resistance to Fusarium head blight. *Cytogenet. Genome Res.* 153, 165–173. doi: 10.1159/000486382

Lukaszewski, A. J. (1995). Physical distribution of translocation breakpoints in homoeologous recombinants induced by the absence of the *Ph1* gene in wheat and triticale. *Theor. Appl. Genet.* 90, 714–719. doi: 10.1007/BF00222138

Lukaszewski, A. J., and Curtis, C. A. (1993). Physical distribution of recombination in B-genome chromosomes of tetraploid wheat. *Theor. Appl. Genet.* 86, 121–127. doi: 10.1007/BF00223816

Lukaszewski, A. J., Rybka, K., Korzun, V., Malyshev, S. V., Lapinski, B., and Whitkus, R. (2004). Genetic and physical mapping of homoeologous recombination points involving wheat chromosome 2B and rye chromosome 2R. *Genome* 47, 36–45. doi: 10.1139/g03-089

Ma, J., Zhang, C. Y., Liu, Y. X., Yan, G. J., and Liu, C. J. (2012). Enhancing Fusarium crown rot resistance of durum wheat by introgressing chromosome segments from hexaploid wheat. *Euphytica* 186, 67–73. doi: 10.1007/s10681-011-0492-0

Mandalà, G., Tundo, S., Francesconi, S., Gevi, F., Zolla, L., Ceoloni, C., et al. (2019). Deoxynivalenol detoxification in transgenic wheat confers resistance to Fusarium head blight and crown rot diseases. *Mol. Plant Microbe Interact.* 32, 583–592. doi: 10.1094/MPMI-06-18-0155-R

Maresca, M. (2013). From the gut to the brain: journey and pathophysiological effects of the food-associated trichothecene mycotoxin deoxynivalenol. *Toxins (Basel)* 5, 784–820. doi: 10.3390/toxins5040784

Matny, O. N. (2015). Fusarium head blight and crown rot on wheat & barley: losses and health risks. *Adv. Plants Agric. Res.* 2, 38–43. doi: 10.15406/apar.2015.02.00039

McMullen, M., Bergstrom, G., De Wolf, E., Dill-Macky, R., Hershman, D., Shaner, G., et al. (2012). A unified effort to fight an enemy of wheat and barley: Fusarium head blight. *Plant Dis.* 96, 1712–1728. doi: 10.1094/PDIS-03-12-0291-FE

Mesterházy, Á., Bartok, T., Mirocha, C. G., and Komoroczy, R. (1999). Nature of wheat resistance to Fusarium head blight and the role of deoxynivalenol for breeding. *Plant Breed.* 118, 97–110. doi: 10.1046/j.1439-0523.1999.118002097.x

Miedaner, T., Sieber, A. N., Desaint, H., Buerstmayr, H., Longin, C. F. H., and Wurschum, T. (2017). The potential of genomic-assisted breeding to improve Fusarium head blight resistance in winter durum wheat. *Plant Breed.* 136, 610–619. doi: 10.1111/pbr.12515

Miller, S. S., Watson, E. M., Lazebnik, J., Gulden, S., Balcerzak, M., Fedak, G., et al. (2011). Characterization of an alien source of resistance to Fusarium head blight transferred to Chinese Spring wheat. *Botany* 89, 301–311. doi: 10.1139/b11-017

Mudge, A. M., Dill-Macky, R., Dong, Y., Gardiner, D. M., White, R. G., and Manners, J. M. (2006). A role for the mycotoxin deoxynivalenol in stem colonisation during crown rot disease of wheat caused by *Fusarium graminearum* and *Fusarium pseudograminearum*. *Physiol. Mol. Plant Pathol.* 69, 73–85. doi: 10.1016/j.pmpp.2007.01.003

Peterson, R. F., Campbell, A. B., and Hannah, A. E. (1948). A diagrammatic scale for estimating rust intensity of leaves and stem of cereals. *Can. J. Res.* 26, 496–500. doi: 10.1139/cjr48c-033

Pozniak, C. J., Knox, R. E., Clarke, F. R., and Clarke, J. M. (2007). Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. *Theor. Appl. Genet.* 114, 525–537. doi: 10.1007/s00122-006-0453-5

Prat, N., Buerstmayr, M., Steiner, B., Robert, O., and Buerstmayr, H. (2014). Current knowledge on resistance to Fusarium head blight in tetraploid wheat. *Mol. Breed.* 34, 1689–1699. doi: 10.1007/s11032-014-0184-2

Prat, N., Guilbert, C., Prah, U., Wachter, E., Steiner, B., Langin, T., et al. (2017). QTL mapping of Fusarium head blight resistance in three related durum wheat populations. *Theor. Appl. Genet.* 130, 13–27. doi: 10.1007/s00122-016-2785-0

Qi, L., Friebe, B., Zhang, P., and Gill, B. S. (2007). Homoeologous recombination, chromosome engineering and crop improvement. *Chromos. Res.* 15, 3–19. doi: 10.1007/s10577-006-1108-8

Ravel, C., Dardevet, M., Leenhardt, F., Bordes, J., Joseph, J. L., Perretant, M. R., et al. (2013). Improving the yellow pigment content of bread wheat flour by selecting the three homoeologous copies of *Psy1*. *Mol Breeding* 31, 87–99. doi: 10.1007/s11032-012-9772-1

Ranieri, R. (2015). Geography of the durum wheat crop. *Pastaria Int.* 6, 24–36.

Ray, D. K., West, P. C., Clark, M., Gerber, J. S., Prishchepov, A. V., and Chatterjee, S. (2019). Climate change has likely already affected global food production. *PLoS One* 14, e0217148. doi: 10.1371/journal.pone.0217148

Romer Labs Division Holding GmbH. (2016), *Worldwide Mycotoxin Regulations*, https://www.romerlabs.com/knowledge-center/knowledge-library/articles/news/worldwide-mycotoxin-regulations/.

Saintenac, C., Falque, M., Martin, O. C., Paux, E., Feuillet, C., and Sourdille, P. (2009). Detailed recombination studies along chromosome 3B provide new insights on crossover distribution in wheat (*Triticum aestivum* L.). *Genetics* 181, 393–403. doi: 10.1534/genetics.108.097469

Saintenac, C., Faure, S., Remay, A., Choulet, F., Ravel, C., Paux, E., et al. (2011). Variation in crossover rates across a 3-Mb contig of bread wheat (*Triticum aestivum*) reveals the presence of a meiotic recombination hotspot. *Chromosoma* 120, 185–198. doi: 10.1007/s00412-010-0302-9

Salgado, J. D., Madden, L. V., and Paul, P. A. (2015). Quantifying the effects of fusarium head blight on grain yield and test weight in soft red winter wheat. *Phytopathology* 105, 295–306. doi: 10.1094/PHYTO-08-14-0215-R

Sall, A. T., Cisse, M., Gueye, H., Kabbaj, H., Ndoye, I., Filali-Maltouf, A., et al. (2018). Heat tolerance of durum wheat (*Triticum durum* Desf.) elite germplasm tested along the Senegal River. *J. Agr. Sci.* 10, 217–233. doi: 10.5539/jas.v10n2p217

Sari, E., Berraies, S., Knox, R. E., Singh, A. K., Ruan, Y., Cuthbert, R. D., et al. (2018). High density genetic mapping of Fusarium head blight resistance QTL in tetraploid wheat. *PLoS One* 13, e0204362. doi: 10.1371/journal.pone.0204362

Scherm, B., Balmas, V., Spanu, F., Pani, G., Delogu, G., Pasquali, M., et al. (2013). *Fusarium culmorum*: causal agent of foot and root rot and head blight on wheat. *Mol. Plant Pathol.* 14, 323–341. doi: 10.1111/mpp.12011

Scherm, B., Orrù, M., Balmas, V., Spanu, F., Azara, E., Delogu, G., et al. (2011). Altered trichothecene biosynthesis in *TRI6*-silenced transformants of *Fusarium culmorum* influences the severity of crown and foot rot on durum wheat seedlings. *Mol. Plant Pathol.* 12, 759–771. doi: 10.1111/j.1364-3703.2011.00709.x

Sella, L., Gazzetti, K., Castiglioni, C., Schäfer, W., and Favaron, F. (2014). *Fusarium graminearum* possesses virulence factors common to Fusarium head blight of wheat and seedling rot of soybean but differing in their impact on disease severity. *Phytopathology* 104, 1201–1207. doi: 10.1094/PHYTO-12-13-0355-R

Shen, X., Kong, L., and Ohm, H. (2004). Fusarium head blight resistance in hexaploid wheat (*Triticum aestivum*)-*Lophopyrum* genetic lines and tagging of the alien chromatin by PCR markers. *Theor. Appl. Genet.* 108, 808–813. doi: 10.1007/s00122-003-1492-9

Shen, X., and Ohm, H. (2007). Molecular mapping of *Thinopyrum*-derived Fusarium head blight resistance in common wheat. *Mol. Breed.* 20, 131–140. doi: 10.1007/s11032-007-9079-9

Steiner, B., Buerstmayr, M., Michel, S., Schweiger, W., Lemmens, M., and Buerstmayr, H. (2017). Breeding strategies and advances in line selection for Fusarium head blight resistance in wheat. *Trop. Plant pathol.* 42, 165–174. doi: 10.1007/s40858-017-0127-7

Steiner, B., Michel, S., Maccaferri, M., Lemmens, M., Tuberosa, R., and Buerstmayr, H. (2019). Exploring and exploiting the genetic variation of Fusarium head blight resistance for genomic-assisted breeding in the elite durum wheat gene pool. *Theor. Appl. Genet.* 132, 969–988. doi: 10.1007/s00122-018-3253-9

Tai, T. H., and Tanksley, S. D. (1990). A rapid and inexpensive method for isolation of total DNA from dehydrated plant tissue. *Plant Mol. Biol. Rep.* 8, 297–303. doi: 10.1007/BF02668766

Tuberosa, R., and Pozniak, C. (2014). Durum wheat genomics comes of age. *Mol. Breed.* 34, 1527–1530. doi: 10.1007/s11032-014-0188-y

Tundo, S., Janni, M., Moscetti, I., Mandalá, G., Savatin, D., Blechl, A., et al. (2016). PvPGIP2 accumulation in specific floral tissues but not in the endosperm limits *Fusarium graminearum* infection in wheat. *Mol. Plant-Microbe Interact.* 29, 815–821. doi: 10.1094/MPMI-07-16-0148-R

Urban, M., Daniels, S., Mott, E., and Hammond-Kosack, K. (2002). *Arabidopsis* is susceptible to the cereal ear blight fungal pathogens *Fusarium graminearum* and *Fusarium culmorum*. *Plant J.* 32, 961–973. doi: 10.1046/j.1365-313X.2002.01480.x

Vaughan, M. M., Backhouse, D., and Ponte, E. M. D. (2016). Climate change impacts on the ecology of *Fusarium graminearum* species complex and susceptibility of wheat to *Fusarium* head blight: a review. *World Mycotoxin J.* 9, 685–700. doi: 10.3920/WMJ2016.2053

Zhang, W., and Dubcosky, J. (2008). Association between allelic variation at the *Phytoene synthase 1* gene and yellow pigment content in the wheat grain. *Theor. Appl. Genet.* 116, 635–645. doi: 10.1007/s00122-011-1596-6

Zhang, W., Cao, Y., Zhang, M., Zhu, X., Ren, S., Long, Y., et al. (2017). Meiotic homoeologous recombination-based alien gene introgression in the genomics era of wheat. *Crop Sci.* 57, 1189–1198. doi: 10.2135/cropsci2016.09.0819

Zhang, X. L., Shen, X. R., Hao, Y. F., Cai, J. J., Ohm, H. W., and Kong, L. (2011). A genetic map of *Lophopyrum ponticum* chromosome 7E, harboring resistance genes to Fusarium head blight and leaf rust. *Theor. Appl. Genet.* 122, 263–270. doi: 10.1007/s00122-010-1441-3

Zhang, W., Zhu, X., Zhang, M., Chao, S., Xu, S., and Cai, X. (2018). Meiotic homoeologous recombination-based mapping of wheat chromosome 2B and its homoeologues in *Aegilops speltoides* and *Thinopyrum elongatum*. *Theor. Appl. Genet.* 131, 2381–2395. doi: 10.1007/s00122-018-3160-0

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

REVIEW
published: 07 November 2019
doi: 10.3389/fpls.2019.01347

Carotenoid Pigment Content in Durum Wheat (*Triticum turgidum* L. var *durum*): An Overview of Quantitative Trait Loci and Candidate Genes

Pasqualina Colasuonno[1†], Ilaria Marcotuli[1†], Antonio Blanco[1], Marco Maccaferri[2], Giuseppe Emanuele Condorelli[2], Roberto Tuberosa[2], Roberto Parada[3], Adriano Costa de Camargo[3], Andrés R. Schwember[3] and Agata Gadaleta[1]*

[1] Department of Agricultural and Environmental Science (DISAAT), University of Bari "Aldo Moro", Bari, Italy, [2] Department of Agricultural and Food Sciences (DISTAL), University of Bologna, Bologna, Italy, [3] Facultad de Agronomía e Ingeniería Forestal, Pontificia Universidad Católica de Chile, Santiago, Chile

Carotenoid pigment content is an important quality trait as it confers a natural bright yellow color to pasta preferred by consumers (whiteness vs. yellowness) and nutrients, such as provitamin A and antioxidants, essential for human diet. The main goal of the present review is to summarize the knowledge about the genetic regulation of the accumulation of pigment content in durum wheat grain and describe the genetic improvements obtained by using breeding approaches in the last two decades. Although carotenoid pigment content is a quantitative character regulated by various genes with additive effects, its high heritability has facilitated the durum breeding progress for this quality trait. Mapping research for yellow index and yellow pigment content has identified quantitative trait loci (QTL) on all wheat chromosomes. The major QTL, accounting for up to 60%, were mapped on 7L homoeologous chromosome arms, and they are explained by allelic variations of the phytoene synthase (*PSY*) genes. Minor QTL were detected on all chromosomes and associated to significant molecular markers, indicating the complexity of the trait. Despite there being currently a better knowledge of the mechanisms controlling carotenoid content and composition, there are gaps that require further investigation and bridging to better understand the genetic architecture of this important trait. The development and the utilization of molecular markers in marker-assisted selection (MAS) programs for improving grain quality have been reviewed and discussed.

Keywords: durum wheat, grain yellow pigment content, carotenoids, yellow index, marker-assisted selection

OPEN ACCESS

Edited by:
Jose Luis Gonzalez Hernandez, South Dakota State University, United States

Reviewed by:
Amidou N'Diaye, University of Saskatchewan, Canada
Sergio G. Atienza, Spanish National Research Council (CSIC), Spain
Ajay Kumar, North Dakota State University, United States

***Correspondence:**
Andrés R. Schwember
aschwember@uc.cl

†These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *12 June 2019*
Accepted: *27 September 2019*
Published: *07 November 2019*

Citation:
Colasuonno P, Marcotuli I, Blanco A, Maccaferri M, Condorelli GE, Tuberosa R, Parada R, de Camargo AC, Schwember AR and Gadaleta A (2019) Carotenoid Pigment Content in Durum Wheat (Triticum turgidum L. var durum): An Overview of Quantitative Trait Loci and Candidate Genes. Front. Plant Sci. 10:1347. doi: 10.3389/fpls.2019.01347

INTRODUCTION

Durum wheat (*Triticum turgidum* L. ssp. *durum*) is a cereal crop grown around the world on about 17 million hectares and with about 37 million tons produced annually during the last decade, with wide variation from 32 to 42 million tons caused mainly by drought and heat stresses (data FAO, 2017). Globally durum wheat represents only 8% of the whole area cultivated with wheat and about 5% of world wheat production. The principal durum-producing countries are the European Union, Canada, Turkey, USA, Algeria, Kazakhstan, and Mexico, whereas minor production countries encompass Syria, Morocco, Tunisia, India, Australia, and Argentina and Chile, among others.

Major producers in the EU are Italy, France, Greece, and Spain. Although durum wheat is a relatively minor crop worldwide, it is the main crop for many areas of the Mediterranean basin and makes up the raw material for various finished products such as pasta and couscous consumed all over the world (Kabbaj et al., 2017).

Grain protein content and conformation together with yellow color are the most valued wheat quality traits, which are important in the commercial, nutritional, and technological values of grain and end products of both common and durum wheat (Sissons, 2008; Nazco et al., 2012; Subira et al., 2014; Mazzeo et al., 2017).

Semolina and pasta color are the consequence of two distinct constituents: yellow (desirable) and brown (undesirable) pigments. The yellow color, principally explained by carotenoid accumulation in kernels, has been considered a source of significant nutrients/antioxidant compounds and a factor for the commercial value since consumers prefer a bright yellow color of semolina and the pasta products.

It is a typical quantitative trait controlled by a complex genetic system (quantitative trait loci, QTL) and influenced by environmental factors. As confirmed by the high value of heritability, the genetic component is predominant, and this has facilitated the success of breeding programs (Elouafi et al., 2001). A consequence of this intense breeding activity has been proved by the higher carotenoid concentration in durum wheat cultivated varieties compared to the wild ones (Digesù et al., 2009).

Genetic analyses based on molecular markers have mapped major QTL for carotenoid content on homoeologous group 7. Minor QTL, associated to significant markers, were detected on almost all chromosomes of the durum wheat genome. Significant marker-trait associations for carotenoid content have been detected on all of the chromosomes by linkage mapping. Association mapping has been used as a new strategy for the dissection of this trait and highlighted its complexity.

Due to the importance of this quality trait, the aim of the present review is to summarize the information available on the detection of QTL for carotenoid content and individual components and the identification of candidate genes in durum wheat.

CAROTENOIDS AND NUTRITIONAL ASPECTS

In the first years of the 21st century, major breeding programs were focused on improving the durum productivity traits of wheat, such as grain yield and biotic and abiotic stress resistance. Recently, the attention on food quality over quantity has switched the research aims at increasing wheat nutritional value estimated through different parameters, like protein content, water absorption, and flour color. The latter one is due to the carotenoid pigments, whose nutritional benefits in human health is worldwide recognized (Sommer and Davidson, 2002).

Over 600 carotenoids have been identified in plants and microorganisms. They are one of the most studied groups of natural pigments, because of their broad distribution, structural variety, and multiple functions. All fruit and plant color, ranging from yellow to red, are good sources of carotenoids (Britton, 1998).

Bendich and Olson were the first scientists that characterize more than 750 carotenoid compounds (Bendich and Olson, 1989; Olson and Krinsky, 1995; DellaPenna and Pogson, 2006). *Carotenoids* is the generic term indicating the majority of red, orange, and yellow pigments naturally encountered in photosynthetic organisms and in certain fungi and bacteria (Britton 1995; Khoo et al., 2011).

Most carotenoids are tetraterpenoids (C40 compounds), which are composed of eight isoprenic units linked in a linear and symmetrical structure. The basic cyclic structure can be changed by dehydrogenation, hydrogenation, cyclization, and oxidation reactions, while the high chemical reactivity has been induced by a complex system of double bonds (Oliver and Palou, 2000). Two classes of carotenoids are found in nature: (a) the carotenes, linear tetraterpenoid hydrocarbons (i.e., β-carotene) that can be cyclized at one or both ends of the molecule, and (b) the xanthophylls, composed by one or more oxygen groups (i.e., lutein, violaxanthin, neoxanthin, and zeaxanthin) (Van den Berg et al., 2000). Overall, carotenoids possess general properties common to all carotenoids (i.e., antioxidant nature) from specific ones (i.e., provitamin A, only particular ones) (Young and Lowe, 2018).

Food, animal feed, and pharmaceutical industries are using carotenoids for their color properties (in fruit juices, pasta, candies, cheese, chicken skin, and egg yolk) or for their contribution in the flavor and fragrances of some foods (Landrum, 2009). In addition, in a context of increasing interest in improving health through the consumption of natural products, they are considered also in food fortification. It has been demonstrated that they are precursors of vitamin A, generating health advantages, such as antioxidant properties, reinforcing the immune system, decreasing the risk of degenerative and cardiovascular ailments, anti-obesity/hypolipidemic properties, and defense of the macula region of the retina (Mezzomo et al., 2015).

Among the carotenoid compounds, following the presence/absence of the provitamin A in the molecule structure, fewer than 10% show a significant biological activity and act as vitamin A precursors. The precursors of vitamin A have a minimum of one β-ionone ring and a polienic chain with 11 carbons at least (Kelly et al., 2018).

Bioactive compounds and vitamin A are categorized as antioxidants, playing a crucial function in humans' health such as in growth, in the development and maintenance of epithelial tissues, in the immune system strength, and in the first protection mechanism against oxidative stress. Antioxidants reduce the singlet oxygen in the human body and scavenge free radicals, i.e., reactive nitrogen species (RNS) and reactive oxygen species (ROS) (Choe and Min, 2005; Leong et al., 2018). The oxidation of carotenoids by ROS causes the loss of their characteristic color inducing cell protection and the prevention of degenerative diseases (Halliwell, 2011; Guaadaoui et al., 2014).

Carotenoids are transformed to vitamin A to satisfy the body requirements, changing the levels of conversion efficiency (Mezzomo and Ferreira, 2016). A prolonged deficiency of

vitamin A can produce skin modifications, corneal ulcers, and night blindness, while a surplus is toxic and may be associated to congenital malformation in pregnancy, bone diseases in patients with chronic renal malfunction, blindness, xerophthalmia, and death (Mezzomo et al., 2015).

According to the high unsaturation degree of carotenoids, light, heat, acids, and enzymatic oxidation can change their structure from the trans-isomers (the most stable type in nature) to the cis-structure, producing a minor decrease of color and provitamin activity (Schroeder and Johnson, 1995).

THE CAROTENOIDS' VALUE IN YELLOW COLOR OF WHEAT PRODUCTS

The yellow-amber color of semolina is caused by the carotenoid (yellow) pigment content (YPC) in the entire grain, which is known as the yellow index (YI) of semolina at a commercial level (CIE, 1986). The average carotenoid concentration in durum wheat is 6.2 ± 0.13 mg/kg in dry weight, determining the pasta color (Beleggia et al., 2011; Brandolini et al., 2015).

In wheat kernel, a wide range of carotenoids have been detected such as lutein, β-carotene, zeaxanthin, β-cryptoxanthin, β-apocarotenal, antheraxanthin, taraxanthin (lutein-5,6-epoxide), flavoxanthin, and triticoxanthin (Lachman et al., 2017).

The pigments are variable: α- and β-carotene (7.7%) are mainly located in the germ, while lutein, the most abundant pigment (86–94%) (Konopka et al., 2004; Digesù et al., 2009), is equally distributed across the layers (Borrelli et al., 2008). Specifically, aleurone layer, starchy endosperm, and germ contain 0.425, 0.557, and 2.157 mg/kg of lutein, respectively. In parallel, aleurone and germ contain 0.776 and 3.094 mg/kg zeaxanthin.

During the milling process, a large amount of these components are gradually reduced, depending on the extraction rate (Paznocht et al., 2019). Lutein, and a small amount of zeaxanthin, has higher cooking stability compared to other carotenoids commonly present in foods, for example, β-cryptoxanthin and β-carotene (Britton and Khachik, 2009; Kean et al., 2011).

In the wheat end products, steady-state level of carotenoids is dependent on the equilibrium between biosynthesis and degradation in the processing phase. This last process has been principally attributed to oxidative enzymes, such as the lipoxygenases (LOX), polyphenol oxidase (PPO), and peroxidases (PER), that can cause the loss of yellow color of flour and pasta (Hessler et al., 2002; Leenhardt et al., 2006; Garbus et al., 2013; Morris, 2018).

CAROTENOID BIOSYNTHESIS AND DEGRADATION PATHWAYS

The carotenoid metabolic biosynthetic pathway has been thoroughly investigated in some plants, including *Arabidopsis*, rice, maize, pepper, tomato, and orange (see studies by Colasuonno et al., 2017a; Rodriguez-Concepcion et al., 2018; Sun et al., 2018). This biosynthetic route has been examined in depth (**Figure 1**), and all genes and enzymes involved have been isolated and well characterized. It starts with the condensation of two molecules of geranylgeranyl diphosphate (GGPP) by phytoene synthase (*PSY*) to generate phytoene, which normally is not accumulated in tissues. This step is a key rate-limiting stage of carotenoid biosynthesis, since it might affect the carotenoid pool content (Cazzonelli and Pogson, 2010; Ke et al., 2019). Following a sequence of desaturation and isomerization

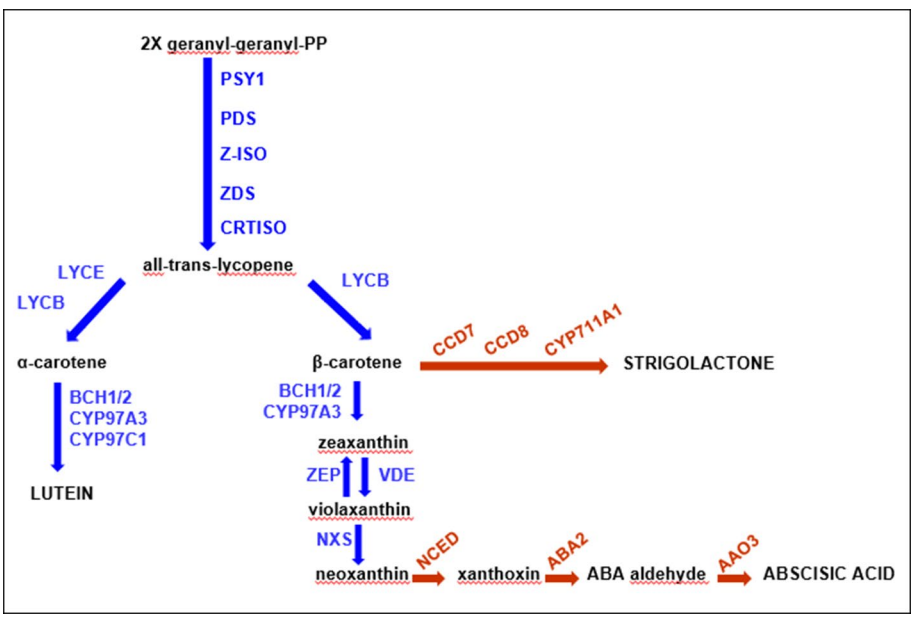


FIGURE 1 | Schematic carotenoid pathway. The main components of the biosynthetic chain have been reported in the figure in black font, while all the enzymes involved are blue. The proteins from the related pathways are indicated in red.

reactions catalyzed by phytoene desaturase (*PDS*), zeta-carotene isomerase (*Z-ISO*), zeta-carotene desaturase (*ZDS*), and carotenoid isomerase (*CRTISO*), the lycopene biosynthesis, a red-colored carotenoid, takes place. Double lycopene cyclization produces orange β-carotene (branch β-β) or α-carotene (branch β-ε). Further hydroxylation of α-carotene generates yellow zeinoxanthin and lutein, while the modification of β-carotene produces β-cryptoxanthin, zeaxanthin, antheraxanthin, violaxanthin, and neoxanthin (**Figure 1**). These steps are catalyzed by two non-heme β-carotene hydroxylases (*BCH1* and *BCH2*) and two heme hydroxylases (*CYP97A* and *CYP97C*), respectively (Sun et al., 2018). The consequent epoxidation and de-epoxidation of zeaxanthin by zeaxanthin epoxidase (*ZEP*) and violaxanthin de-epoxidase (*VDE*) induce the production of xanthophylls, molecules involved in plant protection's mechanisms (Jahns and Holzwarth, 2012). The last phase of carotenoid biosynthesis consists in the transformation of violaxanthin into neoxanthin by neoxanthinsynthase (*NXS*) (Sun et al., 2018). Further oxidative cleavage reactions of violaxanthin and neoxanthin produce xanthoxin, transformed to the abscisic acid (ABA) plant hormone by ABA-aldehyde enzymes (Seo and Koshiba, 2002).

Strigolactones are carotenoid derivatives and originated from the functioning of carotenoid cleavage dioxygenases (CCDs), contributing to the regular accumulation and balanced levels of pigments (Al-Babili and Bouwmeester, 2015; Nisar et al., 2015).

Numerous studies focused on all the carotenoid enzymes have shown that the inheritance of this trait in wheat is of quantitative nature, and it is highly heritable (Digesù et al., 2009; Blanco et al., 2011; Roncallo et al., 2012; Schulthess and Schwember, 2013; Schulthess et al., 2013). Consequently, the correct approach to study carotenoid pigments is to conduct a QTL mapping strategy.

DETERMINATION OF YELLOW PIGMENTS CONTENT AND THEIR SINGLE COMPONENTS

One of the principal problems of carotenoid analyses is testing for their composition and concentration through reproducible and accurate methods, often using only small amounts of seeds. The detection of carotenoid levels is technically restricted by several limitations such as the interference with some regulation/degradation processes and the product storage (Sandmann, 2001). Carotenoid content is a complex trait, and there are several procedures for measuring total pigment concentration and individual pigment compounds.

The reference methods for the total YPC are Standard Method 152 (ICC, 1990) of the International Association for Cereal Science and Technology (ICC) and AACC International Official Method 14-50.01 (AACC International, 2013). These two procedures rely on the extraction of total pigments in water-saturated n-butanol followed by a spectrophotometric quantification of the absorbance of the alcoholic extract at 435.8 nm (the wavelength of maximum absorption of lutein, the dominant carotenoid in durum wheat), using USA/Canada standards (Fu et al., 2017).

Alternatively, pigment content can be measured by the YI determination based on the quantification of flour light reflectance. Analysis with Minolta CR-300 Chroma Meter (Konica Minolta Pty Ltd, Macquarie Park, NSW) geared up with a pulsed xenon arc lamp is the most used instrument for YI analysis. It provides the brown and yellow indexes, after calibration with standard flour control samples. Specially, the measurements consist of the L* (lightness, ranging from 0 for darkness to 100 for total light), a* (red–green chromaticity), and b* (yellow–blue chromaticity) coordinates of the Munsell color system, employing D65 lightning (CIE, 1986). The b* value is directly linked to the lutein and carotene contents, determining the variation in the yellow intensity (Rodriguez-Amaya and Kimura, 2004). Near-infrared (NIR) spectroscopy is a fast, non-destructive, and economic technique, useful for prediction of individual carotenoid pigments in maize and wheat semolina (McCaig et al., 1992).

Accurate measurements of carotenoids and their components can be exclusively obtained by high-performance liquid chromatography (HPLC) analysis (Brandolini et al., 2008; Fu et al., 2017). HPLC, with grade solvents of methyl tert-butyl ether and methanol, and photodiode array detector, allows the identification and separation of carotenoid compounds (Panfili et al., 2004). Detailed HPLC protocols to identify carotenoids in cereals and end products have been designed by several research groups (Burkhardt and Böhm, 2007; Digesù et al., 2009; Irakli et al., 2011; Shen et al., 2015).

QTL DETECTION IN DURUM WHEAT

YPC is controlled by various genes with additive effects that are affected by different environmental conditions (Schulthess et al., 2013). In classic quantitative genetics research, the creation of linkage maps in biparental populations allowed studying the number of loci regulating the trait, the phenotypic and pleiotropic effects, as well as the epistatic interactions with other QTL, enabling the identification and characterization of candidate genes. Mapping studies for YPC and YI in various biparental populations have led to QTL detection across all wheat chromosomes (**Table 1**).

A suitable population for carotenoid analysis are the recombinant inbred lines (RILs) in advanced selfing generations, doubled haploid (DH) populations, or populations derived from backcrosses (Elouafi et al., 2001; Pozniak et al., 2007; Singh et al., 2009; Tsilo et al., 2011; Colasuonno et al., 2014).

When wheat germplasm, including cultivars, advanced breeding lines, or germplasm collections, are considered, mapping methods of genome-wide association study (GWAS) have been applied to link some marker haplotypes with trait expression (Vargas et al., 2016; Colasuonno et al., 2017a; Fiedler et al., 2017). The principle behind the method is to estimate correlations among the genotypes and the phenotypes in panels of lines, based on the linkage disequilibrium between the allelic variants of molecular markers and causal genes (Gupta et al., 2005; Bush and Moore, 2012). This approach has been the official method for many years. In the last decade, it has been

TABLE 1 | Summary of quantitative trait loci (QTL) clusters for grain yellow pigment content (YPC) and yellow index (YI) reported in tetraploid wheat from the literature and from this study.

Chr.	Marker, marker interval *	Map position (cM)**	Carotenoid trait	QTL type***	R2 (%)****	Plant material	References
1AS	hap_1A_1	1.3–1.7	YPC	GWAS	8.5	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
1AS	gwm136-1A	0–4.6	YPC	BP	12.3	Colosseo × Lloyd (176 RILs)	This study
1AS	hap_1A_3	58.4–61.2	YPC	GWAS	7	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
1AS	barc83-gwm135ª	47.8–52.8–57.8	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
1AS	IWB72019	51–59.7–68.4	YPC	BP	5.5	Meridiano × Claudio (181 RILs)	This study
1AL	hap_1A_6	149.5–149.5	YI	GWAS	3.8	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
1AL	barc158-barc17ª	146.4–151.4–156.4	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
1BS	wPt-2694	22.2–27.2–32.5	YI	GWAS	6	Landraces (72 lines), modern cultivars (20 lines)	Rosello et al., 2018
1BS	barc137-wmc626ª	29.6–36.7–41.7	YI	GWAS	3.8	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
1BS	hap_1B_3	35.7–38.8	YPC	GWAS	12.7	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
1BL	BE443797_436–barc302	38–60	YPC, YI	BP	9	UC1113 × Kofa (93 RILs)	Zhang et al., 2008
1BL	BE443797_436–barc302	38–60	YI	BP	10.8	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
1BL	IWB73028-IWB27784	103.5–(106–110)–113.5	YPC	BP	19.3	Svevo × Ciccio (120 RILs)	Colasuonno et al., 2014
1BL	hap_1B_6	109–109.8	YPC	GWAS	5.2	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
1BL	hap_1B_7	115.7–119.1	YPC	GWAS	13.8	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
1BL	Cap3_137325-1B, IWB6947, IWB73429	115.2–117.4–118.6	YPC	BP	2.8	Colosseo × Lloyd	This study
1BL	gwm268-1B	106.8–118.9–131.0	YPC	BP	2.2	Kofa × Svevo (249 RILs)	This study
1BL	wmc44	135.6–140.6–145.6	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
1BL	IWB435ª	145.3–156.3–161.3	YPC	BP	4	Colosseo × Lloyd	This study
2AS	gwm1115-2A	53.2–69.9–86.6	YPC	BP	2.3	Kofa × Svevo	This study
2AS	IWB77592-IWB79691	74.6–78	YPC, YI	GWAS	15.1	Canadian durum wheat collection (169 lines)	N'Diaye et al., 2017
2AS	gwm425	99	YPC, YI	BP	NA	W9262-260D3 × Kofa (155 RILs)	Pozniak et al., 2007
2AS×	wmc522-wmc296-gwm425-gwm95	63.6–105.6	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
2AS	gwm425-gwm372ª	99–107.7	YPC, YI	BP	11.8–24.5	Latino Primadur (121 RILs)	Blanco et al., 2011
2AS	IWB72639	102.7–107.7–112.7	YPC	BP	3.4–13.5	Svevo × Ciccio (120 RILs)	Colasuonno et al., 2014
2BL	IWB73809ª	90–(94–113)–117	YPC	BP	16.3–16.4	Svevo × Ciccio (120 RILs)	Colasuonno et al., 2014
2BL	IWB73809	108.3–112.3–118.3	YPC	BP	16.3–16.4	Svevo × Ciccio (120 RILs)	Colasuonno et al., 2014
3BS	gwm389	0-5–6	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
3BS	CS-ssr7-3B	5.4–12.1–18.8	YPC	BP	4.8	Kofa × Svevo	This study
3BS	wPt-8140	42.8–47.8–52.8	YI	GWAS	7	Landraces (72 lines), modern cultivars (20 lines)	Rosello et al., 2018
3BS	wmc505-3Bª	48.6–63.3–78.0	YPC	BP	1.9	Kofa × Svevo	This study
3BL	wPt-4401-3B, IWB71566	170.9–183.8–196.6	YPC	BP	2	Colosseo × Lloyd	This study
3BL	barc77-gwm299	170.7–175.7–197.1	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
3BL	gwm299-barc84	192–198	YPC, YI	BP	11.5–16.2 (YPC), 9.4–16.6 (YI)	Latino × Primadur (121 RILs)	Blanco et al., 2011
3BL	IWB45539-IWB58810	200.5–205.5	YPC, YI	GWAS	13.4	Canadian durum wheat collection (169 lines)	N'Diaye et al., 2017
3BL	wmc632-gwm340	206.9–211.9	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008

(*Continued*)

TABLE 1 | Continued

Chr.	Marker, marker interval *	Map position (cM)**	Carotenoid trait	QTL type***	R2 (%)****	Plant material	References
3BL	IWB8780-IWB72417	208–209.6	YPC, YI	GWAS	14.6	Canadian durum wheat collection (169 lines)	N'Diaye et al., 2017
3BL	wPt-2416	211.3–216.3–221.3	YI	GWAS	5	Landraces (72 lines), modern cultivars (20 lines)	Rosello et al., 2018
4AL	Lpx-A3-gwm192-wmc617[α]	63.6–81	YPC, YI	BP	10.6–12	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
4AL	hap_4A_2	64.1–64.1	YPC	GWAS	5.5	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
4AL	wPt-0162	69.7	YI	GWAS	5	Landraces (72 lines), modern cultivars (20 lines)	Rosello et al., 2018
4AL	dupw4-barc170	88.9–90.4	YI	BP	8.4–9.1	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
4AL	barc170	85.4–90.4–95.4	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
4AL	hap_4A_3	90.6–90.6	YPC	GWAS	3.6	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
4AL	barc327-gwm160-barc52[α]	160–172	YPC	BP	5	UC1113 × Kofa (93 RILs)	Zhang et al., 2008
4AL	wmc219-psr573.2	165.2–167.8	YPC	BP	6.7–12.0	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
4AL	wmc219-4A	158.3–165.2–172.1	YPC	BP	4.1	Kofa × Svevo	This study
4AL	hap_4A_6	173.3–175.8	YPC	GWAS	8	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
4AL	hap_4A_7	173.3–175.9	YI	GWAS	6.8	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
4BS	gwm368-barc20	34.4–39.4–41.3–46.3	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
4BS	IWB72011[α]	38.9–43.9–48.9	YPC, YI	GWAS	11.1	Durum collection (124 lines)	Colasuonno et al., 2017a
4BS	hap_4B_2	59.8–60.4	YI	GWAS	13.3	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
4BS	hap_4B_2	59.8–60.4	YPC	GWAS	6.8	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
4BS	IWB70599	58.7–60.4	YPC, YI	GWAS	14.1	Canadian durum wheat collection (169 lines)	N'Diaye et al., 2017
4BS	gwm495	60–65	YPC, YI	BP	NA	W9262-260D3 × Kofa (155 RILs)	Pozniak et al., 2007
4BL	IWB73624	88.7–95.1–101.5	YPC	BP	7.1	Colosseo × Lloyd	This study
4BL	hap_4B_4	105.5–110.2	YPC	GWAS	9.1	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
5AS	hap_5A_1	43.8–43.8	YPC	GWAS	4.8	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
5AS	gwm293-gwm304	41.9–46.9–51.9	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
5AS	gwm304-IWB73092[α]	45–60	YPC, YI	BP	9.4–17.9 (YPC), 7.4–16.4 (YI)	Latino × Primadur (121 RILs)	Blanco et al., 2011
5AS	hap_5A_2	50.5–54.9	YPC	GWAS	5.7	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
5AS	IWB73092	54.2–59.2–64.2	YPC	BP	12.6	Svevo × Ciccio (120 RILs)	Colasuonno et al., 2014
5AS	IWB12396	53.9–68.2–82.5	YPC	BP	3.1	Meridiano × Claudio	This study
5AS	hap_5A_3	64.2–73.6	YPC	GWAS	15.3	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
5BL	wmc790-cfa2019	90.4–100.4	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
5BL	gwm499-BE495277_399	90–94	YPC, YI	BP	9–12.2	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
5BL	gwm408-barc232[α]	140.3–146.3	YPC	BP	7.3	PDW 233 × Bhalegaon 4 (140 RILs)	Patil et al., 2008
5BL	tPt-1253	142–145–148	YPC	BP	21.8	Svevo × Ciccio (120 RILs)	Colasuonno et al., 2014
5BL	cfd86-wmc508	152.5–157.5–162.5	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
5BL	311891-5B, IWB70782, IWB63422	139.7–152.5–165.3	YPC	BP	3	Colosseo × Lloyd	This study
5BL	wPt-8125-5B	145.5–154.5–163.5	YPC	BP	6.2	Meridiano × Claudio	This study

(*Continued*)

TABLE 1 | Continued

Chr.	Marker, marker interval *	Map position (cM)**	Carotenoid trait	QTL type***	R2 (%)****	Plant material	References
6AS	gwm334	0–4.3–9.3	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
6AS	wPt-8443	0–10	YI	GWAS	5	Landraces (72 lines), modern cultivars (20 lines)	Rosello et al., 2018
6AS	hap_6A_3	39.5–41	YPC	GWAS	9.7	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
6AS	hap_6A_4	39.5–41	YI	GWAS	5.4	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
6AL	Cap3_141157-6A, GWM132, IWB71882	50.8–52.5–54.2	YPC	BP	18.8	Colosseo × Lloyd	This study
6AL	barc146	47.552.5–57.5	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
6AL	barc146-gwm132	53–63	YPC, YI	BP	16.1–21.4	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
6AL	barc1077-6A	54.7–57.7–60.7	YPC	BP	9.4	Kofa × Svevo	This study
6AL	barc113-gwm570-wmc553	73–95	YPC, YI	BP	14	UC1113 × Kofa (93 RILs)	Zhang et al., 2008
6AL	barc113-wmc553ª	73.6–95	YI	BP	29 (YPC), 17.9 (YI)	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
6AL	IWB14365	88.4–93.4–98.4	YI	GWAS	NA	Durum collection (124 lines)	Colasuonno et al., 2017a
6AL	barc353-gwm169	97.6–114	YPC, YI	BP	9.8–12.4	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
6AL	gwm169-BE483091_472	114–124	YI	BP	10.4	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
6BL	gwm193	69.9–74.9–79.9	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
6BL	gwm193-wmc539ª	74–90	YPC, YI	BP	NA	W9262-260D3 × Kofa (155 RILs)	Pozniak et al., 2007
6BL	hap_6B_5	92.3–96	YPC	GWAS	6	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7AS	gwm471	0–0.3–5.3	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
7AS	wmc168-barc219ª	1–5	YI	BP	12.6	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
7AS	hap_7A_1	21.9–21.9	YPC	GWAS	4.6	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7AS	hap_7A_3	55.9–62.5	YPC	GWAS	21.3	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7AS	barc127-cfa2028	50.4–55.4–60.4	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
7AS	IWB8374	56.6–61.6–66.6	YI	GWAS	12.6	Durum collection (124 lines)	Colasuonno et al., 2017a
7AS	hap_7A_4	82.4–84.4	YPC	GWAS	8.8	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7AS	BQ170462_176-barc174	85–90–95	YPC, YI	BP	11.7	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
7AS	hap_7A_5	90.9–90.9	YPC	GWAS	3.4	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7AL	gwm1065-7Aª	88.1–90.9–93.7	YPC	BP	10	Kofa × Svevo	This study
7AL	IWB72567	97.3–102.3–107.3	YPC, YI	GWAS	18.4	Durum collection (124 lines)	Colasuonno et al., 2017a
7AL	IWB11003	98–102.3–106.6	YPC	BP	11.2	Meridiano × Claudio	This study
7AL	barc108-wmc283-wmc603	107.8–113.4–118.4	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
7AL	hap_7A_7	112.2–118	YPC	GWAS	9.4	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7A	gwm276-cfd6	144.8–145.5	YPC	BP	22	UC1113 × Kofa (93 RILs)	Zhang and Dubcovsky, 2008
7AL	gwm276-wmc116-cfd6	144.8–145.5	YPC	BP	6.3 (YPC), 9.8–22.5 (YI)	UC1113 × Kofa (93 RILs)	Zhang et al., 2008
7AL	wmc116-cfd6	144.8–145.5	YI	BP	9.8–22.5	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
7AL	gwn282-IWB59875ª	170–183	YPC, YI	BP	19.8–30.4 (YPC), 13–15.7 (YI)	Latino × Primadur (121 RILs)	Blanco et al., 2011
7A	IWB59875	177.3–180.3–183.3	YPC	BP	51.6	Svevo × Ciccio (120 RILs)	Colasuonno et al., 2014

(*Continued*)

TABLE 1 | Continued

Chr.	Marker, marker interval *	Map position (cM)**	Carotenoid trait	QTL type***	R2 (%)****	Plant material	References
7AL	IWB72397	180.2–181.8	YPC, YI	GWAS	35.6	Canadian durum wheat collection (169 lines)	N'Diaye et al., 2017
7AL	IWB59875	175.3–180.3–185.3	YI	GWAS	12.2	Durum collection (124 lines)	Colasuonno et al., 2017a
7AL	IWB28063	179.5–181.8–184.1	YPC	BP	19.5	Meridiano × Claudio	This study
7AL	wgwm63-gwm282º	192–206	YPC	BP	NA	Omrabi5 × *Triticum dicoccoides* 600545 114 RILs	Elouafi et al., 2001
7AL	Xscar3362º	192–206	YPC	BP	22.6–55.2	PDW 233 (YAV'S'/TEN'S') × Bhalegaon 4 140 RILs	Patil et al., 2008
7AL	*Psy1-A1*º	192–206	YPC, YI	BP	NA	Commander × DT733 (110 RILs); Strongfield × Blackbird (89 DHs); Strongfield × Commander (106 RILs)	Singh et al., 2009
7AL	*Psy1-A1*º	192–206	YI	BP	NA	Advanced breeding lines (100 lines, F7–F10 generations)	He et al., 2009
7AL	D_304196-*PsyA1*	192–206	YPC, YI	BP	42–53.2 (YPC), 26.1–32.4 (YI)	Latino × Primadur (121 RILs)	Blanco et al., 2011
7AL	*Psy1-A1*	192–206	YPC,YI	BP	NA	Breeding lines (65), landraces (155 lines)	Campos et al., 2016
7AL	hap_7A_11	193.9–194.6	YPC	GWAS	5.7	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7AL	hap_7A_11	193.9–194.6	YI	GWAS	3.9	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7AL	IWB49295	198.4–203.4–208.4	YPC	GWAS	10.4	Durum collection (124 lines)	Colasuonno et al., 2017a
7BS	wmc546-wmc335	55–73	YPC	BP	8.75	PDW 233 (YAV'S'/TEN'S') × Bhalegaon 4 140 RILs	Patil et al., 2008
7BS	wmc182-7B	50–51.6–53.2	YPC	BP	19.3	Kofa × Svevo	This study
7BS	Cap3_173782-7B, IWB72147, IWB12844	53.5–58.3–63.1	YPC	BP	11.7	Colosseo × Lloyd	This study
7BS	barc23-barc72-gwm297º	66.3–67.9–72.8	YPC, YI	BP	8.5–12.8	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
7BL	wmc758-wmc475-gwm333-wmc396	81.2–103.2	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
7BL	Cap3_127025-7B, IWB8805, IWB11767, IWB12371	88.9–104.4–119.8	YPC	BP	4.5	Colosseo × Lloyd	This study
7BL	hap_7B_3	120.4–127.4	YPC	GWAS	11.6	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7BL	hap_7B_3	120.4–127.4	YI	GWAS	10.3	Canadian durum breeding lines (192 lines)	N'Diaye et al., 2018
7BL	wmc311-cfa2257	181.2–185	YPC, YI	BP	7	UC1113 × Kofa (93 RILs)	Zhang et al., 2008
7BL	wmc311-wmc276	181.2–185	YPC	BP	16.9	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
7BL	wPt-5138	180–189	YI	GWAS	6	Landraces (72 lines), modern cultivars (20 lines)	Rosello et al., 2018
7BL	cfa2040-Psy-B1-barc1073-cfa2257º	181.2–207	YI, YPC	BP	6.6–16.9	UC1113 × Kofa (93 RILs)	Roncallo et al., 2012
7BL	wPt-744987-7B	194.8–202.9–211.0	YPC	BP	4.3	Meridiano × Claudio	This study
7BL	Psy1-B1º	204	YI	BP	NA	100 advanced breeding lines (F7–F10 generations)	He et al., 2009
7BL	Psy1-B1º	204	YPC, YI	BP	NA	W9262-260D3 × Kofa (155 RILs)	Pozniak et al., 2007
7BL	Psy1-B1-wmc10-gwm146	204	YPC	GWAS	NA	Worldwide elite durum wheat collection (93 lines)	Reimer et al., 2008
7BL	ubw18b-7B	196.3–204.9–213.5	YPC	BP	4.5	Kofa × Svevo	This study
7BL	IWB34193-IWB12638	202.9–206.3	YPC, YI	GWAS	9.3	Canadian durum wheat collection (169 lines)	N'Diaye et al., 2017

(*Continued*)

TABLE 1 | Continued

Chr.	Marker, marker interval *	Map position (cM)**	Carotenoid trait	QTL type***	R2 (%)****	Plant material	References
7BL	gwm344	203	YPC	BP	52	Omrabi5 × *T. dicoccoides* 600545 114 RILs	Elouafi et al., 2001
7BL	barc340-cfa2257	195.9–207	YPC	BP	7	UC1113 × Kofa (93 RILs)	Zhang and Dubcovsky, 2008

**QTL markers are SSR, DART, single nucleotide polymorphism (SNP) markers mapped in the durum wheat consensus map (Maccaferri et al., 2014).*
***QTL marker, marker interval: reported by projecting the QTL interval and QTL peak markers onto the durum wheat consensus map (Maccaferri et al., 2014).*
****BP = for QTL obtained by using biparental mapping populations; GW = for QTL revealed in a genome-wide association (GWAS) study; RIL = recombinant inbred line.*
*****R2 (%) single value indicates the average QTL R2 across environments. A two-values range reports the QTL R2 range across environments.*
[a]Stable QTL detected in more than one population.

extensively used for different traits owing to the availability of high numbers of DNA-based markers uniformly distributed in the genome (such as the high-density maps obtained from single nucleotide polymorphisms, SNPs, Wang et al., 2014 and Maccaferri et al., 2014) and the improvement of statistical tools. These included improved mixed models that effectively take into account the interfering panel population structure effects (Breseghello and Sorrells, 2006; Maccaferri et al., 2006; Zhang et al., 2010; Maccaferri et al., 2011; Lipka et al., 2012). Although several reports of SNP markers, associated to QTL for carotenoid pigments, have been published to date (**Table 1**), they mostly relied on linkage mapping and biparental RIL populations. Only four studies were deeply focused on SSR and SNP-based association mapping in durum wheat (Reimer et al., 2008; Colasuonno et al., 2017a; N'Diaye et al., 2017; Rosello et al., 2018), and only one considered haplotypes instead of single bi-allelic markers (N'Diaye et al., 2018).

CAROTENOID QTL REGIONS

The results of recent peer-reviewed studies and additional studies provided by the authors reporting QTL for carotenoid pigments in durum wheat are compiled in **Table 1** (QTL clusters) and **Table S1** (all QTL), encompassing information on mapping population, examined phenotypic traits, and markers associated to carotenoid QTL. Map position of the major QTL clusters listed in **Tables 1** and **S1** are reported in **Figure 2**. The schematic representation is based upon the durum wheat consensus map published by Maccaferri et al. (2014), and revised for the carotenoid composition by Colasuonno et al. (2017a). Considering that each QTL/MTA identified by different studies was located in the same position of the durum wheat consensus map, we used the cM reported by Maccaferri et al. (2014) to identify the genome location.

Every chromosome is depicted with the first and the last SNP marker, and one marker every 20 cM. SSR markers have also been incorporated to anchor the actual consensus map with the previous ones. For most of the detected QTL, the connection between the different maps has been quite simple to be achieved.

The QTL mapping study highlights the differences in number and map position of the QTL identified in several analyses. This could be due to the specific plant material and/or the analytical and statistical procedures adopted in each study (e.g., linkage mapping vs. GWAS). In fact, the presence of numerous genes with an additive effect on the trait, the parental influence on the genotypes of the mapping populations, the genotype × environment interaction, and the number of markers used may affect the results. Moreover, a different carotenoid measurement approach and statistical procedures adopted might influence the reproducibility of each QTL analysis.

Mapping studies for YPC and YI identified 81 QTL (including singletons and QTL clusters) distributed on all wheat chromosomes (**Table S1**). Some of these QTL have been detected in more than one map/population, indicating the presence of stable alleles valuable for enhancing color and nutritional value of wheat grain (QTL clusters, listed in **Table 1**). Twenty stable QTL (highlighted in **Table 1**) were detected on chromosomes 1A (two), 1B (two), 2A, 2B, 3B, 4A (two), 4B, 5A, 5B (two), 6A, 6B, 7A (four), and 7B (two). Therefore, these QTL can be considered useful for MAS in breeding programs. Major, moderate, and minor QTL were defined basing on the percentage of the phenotypic variation (> 40%, 10–40%, and <10%, respectively) as reported in the different studies. The major QTL have been mapped on chromosomes 7AL (Elouafi et al., 2001; Patil et al., 2008; Zhang and Dubcovsky, 2008) and 7BL distal regions (Elouafi et al., 2001; Pozniak et al., 2007; Zhang and Dubcovsky, 2008). In particular, two different QTL have been identified on 7AL (Zhang and Dubcovsky, 2008; Blanco et al., 2011; Crawford and Francki, 2013): the first one (QTL-72) associated with allelic variations of the *AO* gene (Colasuonno et al., 2014; Colasuonno et al., 2017b) with a negative effect (R^2 22.1% for YPC and 18.4% for YI) on carotenoid content, and the second one (QTL-73) in the region of the *PSY1* gene (He et al., 2008; Zhang and Dubcovsky, 2008; He et al., 2009) with a positive and consistent effect up to 60% on YPC. The same QTL–gene associations are reported on chromosome 7BL with R^2 up to 29.1% for the first QTL (QTL-80, at 180–189 cM) and 52% for the one at 181–211 cM (QTL-81). All the loci detected on chromosome group 7 resulted as negatively correlated with grain yield per spike (GYS) and thousand grain weights (TKWs) but positively correlated with protein content. The same association highlighted on chromosomes 7A and 7B between QTL and PSY2 gene has been reported on chromosome 5AS with moderate effect (R2 9.4–17.9% for YPC and 7.4–16.4 for YI) and negative correlation with GYS and TKW.

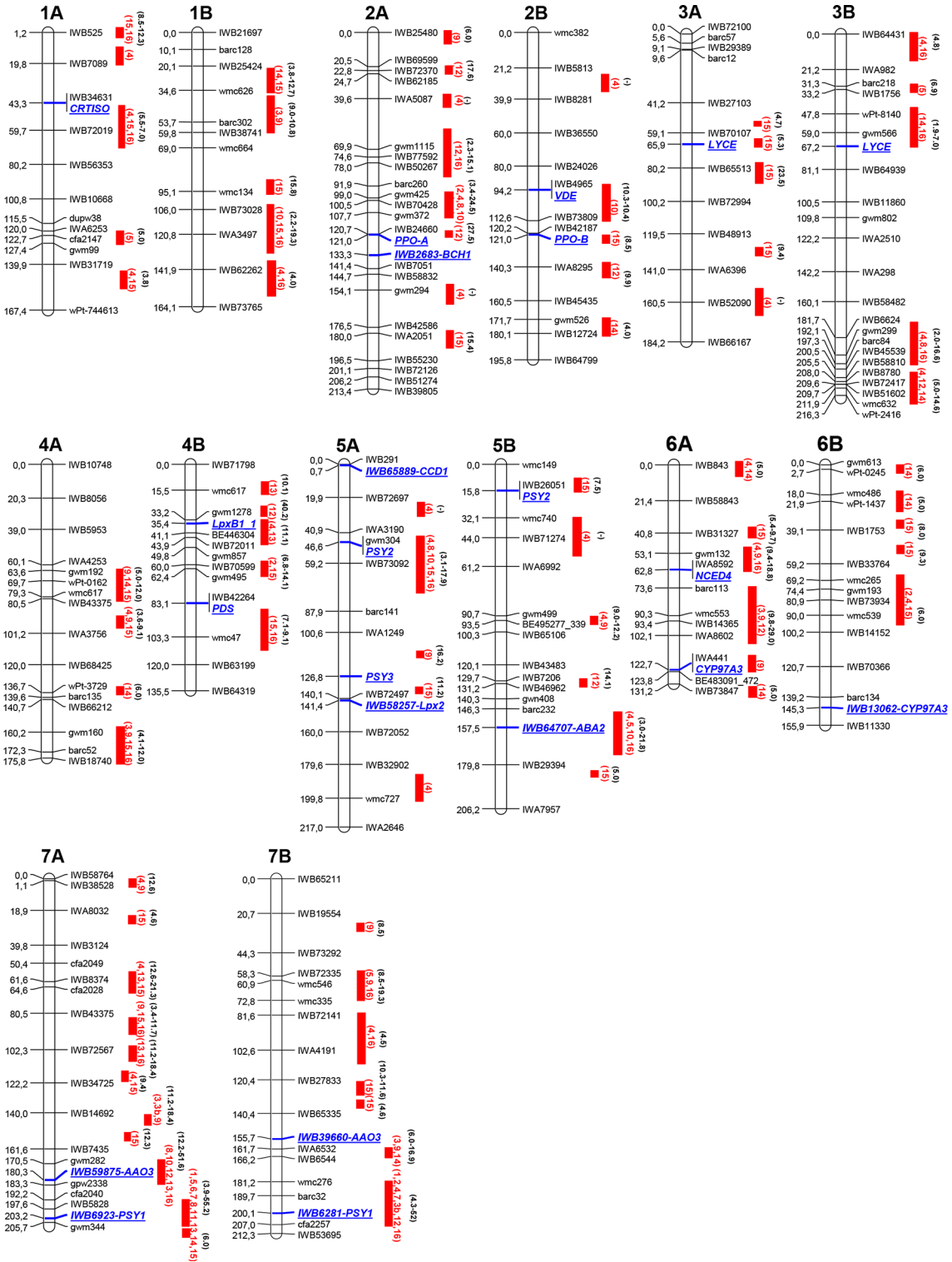


FIGURE 2 | Schematic representation of durum wheat chromosomes [based on the Maccaferri et al. (2014) map] and quantitative trait loci (QTL) summary for yellow pigment content (YPC) or yellow index (YI) trait detected in references in **Table 1**. Markers, on the right chromosome side, are reported every 20 cM approximately. cM distances are indicated on the left side of the bar. Red solid bars indicate the QTL confidence interval regions. The red bars report the reference publication numbers in red (as in **Table 1** and **Table S1**) and the QTL R2 range value in black. 1, Elouafi et al. (2001); 2, Pozniak et al. (2007); 3, Zhang et al. (2008); 3b, Zhang and Dubcovsky (2008); 4, Reimer et al. (2008); 5, Patil et al. (2008); 6, Singh et al. (2009); 7, He et al. (2009); 8, Blanco et al. (2011); 9, Roncallo et al. (2012); 10, Colasuonno et al. (2014); 11, Campos et al. (2016); 12, N'Diaye et al. (2017); 13, Colasuonno et al. (2017a); 14, Rosello et al. (2018); 15, N'Diaye et al. (2018); 16, this study.

Various studies have confirmed the role of the violaxanthin de-epoxidase (VDE) on 2B (Tsilo et al., 2011; Roncallo et al., 2012) in YPC and YI content. The association between VDE and the QTL on 2BL at 94–113 cM has been validated by many results gained thanks to biparental mapping populations. Colasuonno et al. (2014) identified a QTL on 2BL at 94–113 cM with moderate and negative effect from the "Ciccio" alleles (16.4%).

Minor QTL have been detected on chromosomes 1A, 1B, 2A, 3A, 3B, 4A, 4B, 5A, 5B, and 6B (Pozniak et al., 2007; Zhang et al., 2008). Among them, 2A, 4B, and 6A showed variable effect with R^2 of 3.5–15% for YPC and 12.3–21% for YI, 4–11.9% for YPC and 26.2 for YI, and 14–17% for YPC and 28.3% for YI, respectively (**Table 1** and **Table S1**). Lipoxygenase (*LpxB1.1*) gene is located on chromosome 4B, in the same site of the QTL described in literature by several authors (QTL-40; Carrera et al., 2007; Reimer et al., 2008; Verlotta et al., 2010; Colasuonno et al., 2017a).

QTL for individual carotenoid constituents in wheat were outlined by Howitt et al. (2009) and Blanco et al. (2011). Considering wheat rice synteny, two genes, phytoene synthase (*Psy-A1*) and ε-cyclase (*ε-LCY*), were identified by Howitt et al. (2009) as candidate genes for two of the QTL involved lutein content in wheat endosperm. A segregant population, achieved from crossing the durum wheat cultivars Primadur and Latino, was used by Blanco et al. (2011) to detect QTL for individual carotenoid compounds (lutein, zeaxanthin, β-cryptoxanthin, α-carotene, and β-carotene), YI, and yellow pigment concentration. Total carotenoid concentration accounted for 30–50% of the yellow pigment quantities in durum wheat (Hentschel et al., 2002), reflecting unknown color-producing substances in the durum extracts. Lutein was the most prominent carotenoid compound identified, followed by zeaxanthin, α-carotene, and β-carotene, whereas β-cryptoxanthin was a secondary component. QTL mapping identified clusters of QTL for total and/or one or more carotenoid compounds (α-carotene and β-carotene) in the same chromosomal zones (2A, 3B, 5A, and 7A) where QTL for yellow pigment concentration and YI were detected. The existence of molecular markers related to the main QTL previously indicated is a valuable tool for marker-assisted selection (MAS) programs to increase high carotenoid concentration and the nutritional value of wheat grains.

CANDIDATE GENES FOR ENDOSPERM YELLOWNESS

Candidate Genes for Carotenoid Biosynthesis

The "candidate gene approach" has been used in QTL or association mapping to test SNPs within a candidate gene for a significant association with the yellow color character. Although most studies were focused in increasing the carotenoid content or altering the relative components through conventional breeding (Schulthess and Schwember, 2013), some carotenoid genes have been characterized and/or linked to QTL for carotenoids. In wheat, significant attention has been given to the carotenoid biosynthesis genes (*PSY* (Pozniak et al., 2007; He et al., 2008; Dibari et al., 2012; Campos et al., 2016; Vargas et al., 2016), lycopene ε-cyclase (*LCYE*) (Howitt et al., 2009; Crawford and Francki, 2013), lycopene β-cyclase (*LYCB*) (Zeng et al., 2015), carotenoid β-hydroxylase (*HYD*) (Qin et al., 2012), carotene desaturase (*PDS*), and *ZDS* (Cong et al., 2010), while the degradation of carotenoids has been studied by some catabolic genes, such as aldehyde oxidase (*AO*) (Colasuonno et al., 2017b), polyphenol oxidase (*PPO*) (Si et al., 2012), lipoxygenase (*LOX* or *LPX*) (Verlotta et al., 2010), and peroxidase (*PER*) (Asins and Perez de la Vega, 1985; Fraignier et al., 2000; Ficco et al., 2014).

As previously noted, the gene with major effect on YPC and YI traits is the *PSY*, which is essential for the carotenoid accumulation in the kernels (Gallagher et al., 2004). Phylogenetic approaches have identified three different *PSY* isoforms: *PSY1*, *PSY2*, and *PSY3* mapped on homoeologous chromosome groups 7, 5, and 3, respectively (Dibari et al., 2012). Atienza et al. (2007) was the first study to show that *PSY1* was located on chromosomes 7A and 7B of durum wheat. The gene coding for *PSY1* on chromosome 7B was found to be co-segregant with a carotenoid QTL, while the *PSY1* on chromosome 7A behaved as a co-dominant marker explaining 20–28% of the phenotypic variability (Pozniak et al., 2007). The effect of the *PSY1* loci on the endosperm yellowness of wheat is variable depending on the genetic background. Therefore, the percentage of explained variability of endosperm yellowness for *PSY1-A1* has been found ranging from medium (10–30%), to high (30–50%) and very high (> 50%) (Blanco et al., 2011; Colasuonno et al., 2014), while the reported values for *PSY1-B1* are from low (< 10%) to medium (10–30%) (Zhang and Dubcovsky, 2008; Roncallo et al., 2012) in durum wheat. Overall, the effect of alternative alleles of *PSY-A1* appears to be the most important in the variation of semolina yellowness (SY) using different durum wheat populations (Campos et al., 2016; Vargas et al., 2016). In these studies, *PSY1-Al* was associated to higher SY due to the allelic variant *l* of this gene. The use of molecular markers linked to *PSY1-A1* (i.e., Psy1-A1_STS and YP7A-2 studied jointly) in MAS was suitable to enhance grain yellow pigmentation (Campos et al., 2016). In addition, at 35 days after anthesis, *PSY1-A1* was 21-fold higher expressed in the high-yellowness compared to the low-yellowness genotypes evaluated, corroborating the major role of *PSY1-A1* in the genotypes associated to high SY (Vargas et al., 2016).

Additional evidence highlights how other genes are involved in the control of grain amber color. The full-length DNA sequence of a *ZDS* gene on wheat chromosome 2A, designated *Zds-A1*, was cloned, and a co-dominant functional marker, YP2A-1, was designed based on the polymorphisms of two alleles (Dong et al., 2012). The functional marker, explaining 11.3% of the phenotypic variance for YP content, was co-segregating with a QTL for YP content detected on chromosome 2A in a DH population.

The lycopene ε-cyclase gene (*LCYE*) associated to a QTL on chromosome 3A, playing an important role in the regulation of lutein content in wheat grain (Howitt et al., 2009). An SNP marker in *LCYE* was detected between two Australian wheat genotypes, and a highly significant ($P < 0.01$) association with a QTL on chromosome 3A in two mapping populations showed that *LCYE* is involved in color differences at a functional level (Crawford and Francki, 2013).

Zeng et al. (2015) were able to clone the lycopene β-cyclase gene (*LCYB)* and describe its function and connection with β-carotene biosynthesis in wheat grains. Their results suggest that *LCYB* has a key function in β-carotene biosynthesis in wheat and that *LCYB* may be useful to breed new wheat cultivars with high provitamin A content using RNA interference (RNAi) to hinder specific carotenogenic genes in the wheat endosperm.

Qin et al. (2012) characterized two genes, *HYD1* and *HYD2*, encoding β-hydroxylases in wheat. They observed different expression patterns for different *HYD* genes and homeologs in vegetative tissues and developing grains of tetraploid and hexaploid wheats, indicating their distinct regulatory control in tissue, grain development, and ploidy-specific fashions. The expression of *HYD-B1* achieved highest levels at the last stage of grain filling, showing how carotenoids were still synthesized in mature grains, raising the nutritional value of kernels.

In a recent investigation by Colasuonno et al. (2017), 24 candidate genes encompassed in the biosynthesis and catabolism of carotenoid compounds have been reported using wheat comparative genomics. SNPs identified in the coding sequences of 19 candidate genes enabled their chromosomal location and precise map positions on the two bread and durum reference consensus maps studied (Maccaferri et al., 2014; Wang et al., 2014). Six candidate genes (*PSY1, PSY2, CYP97A3, VDE, ABA2*, and *AAO3*) showing between one to five SNPs were significantly associated to YI by genome-wide association mapping in a collection of 233 accessions of tetraploid wheat, suggesting their involvement in the yellow pigment biosynthesis or catabolism. *PSY1, BCH1, CYP97A3, VDE*, and *ABA2* were also associated to YPC. The phenotypic variation (R^2) explained by each of these markers ranged between 5.9% and 16.3% for YI and from 7.4% to 14.8% for YPC.

Candidate Genes for Carotenoid Degradation

Lipoxygenase (*LPX*) genes, involved in the catabolic pathway, are the most characterized. LPX enzymes are non-heme iron-containing and dioxygenases are found in all kingdoms (Zhang et al., 2015), catalyzing the addition of oxygen in polyunsaturated fatty acids that possesses a cis, cis-1,4 pentadiene system (Garbus et al., 2009). In plants, lipoxygenases are found in leaves, seedlings, and seeds. LPX activity produces ROS that can induce oxidation and degradation of carotenoids (Borrelli et al., 2003). In durum wheat, there are different lipoxygenase genes and alleles contributing to the variation of pasta yellowness (Verlotta et al., 2010; Borrelli and Trono, 2016).

Holtman et al. (1996) reported that lipoxygenase-1 is responsible for LPX activity in barley seeds. Hessler et al. (2002) sequenced several wheat fragments, which were assigned to the *Lpx-1* locus based on their similarity to barley genes. De Simone et al. (2010) reported different levels of *Lpx-1* and *Lpx-3* transcripts at maturity between cultivars with contrasting LPX activities, whereas *Lpx-2* transcripts were absent at this stage. *Lpx-B1* locus was located on the short arm of chromosome 4B (Hessler et al., 2002; Garbus et al., 2009; Verlotta et al., 2010), and five related genes and allele sequences have been reported: *Lpx-B1.1a* (Genbank HM126466), *Lpx-B1.1b* (Genbank HM126468), and *Lpx-B1.1c* (Genbank HM126470) for the *Lpx-B1.1* locus, and *Lpx-B1.2* (Genbank HM126467) and *Lpx-B1.3* (Genbank HM126469) (Hessler et al., 2002; Carrera et al., 2007; Verlotta et al., 2010).

QTL analyses in durum wheat showed that 36–54% of the variation in LPX activity is explained by *Lpx-B1* (Carrera et al., 2007; Verlotta et al., 2010). Carrera et al. (2007) reported a deletion in the *Lpx-B1.1* gene, later named *Lpx-B1.1c* allele, which possesses a deletion covering from the second intron up to the last exon. This allele correlates with higher levels of pasta yellowness, due to the large deletion on its sequence, but it is not correlated with semolina or flour color, suggesting that the role of lipoxygenase on carotenoid degradation occurs in the pasta-making process rather than in the development of the grains (Carrera et al., 2007; Verlotta et al., 2010).

Verlotta et al. (2010) genotyped the presence of the *Lpx-B1.1* alleles in combination with either *Lpx-B1.2* or *Lpx-B1.3* in a diverse modern/old durum wheat population, and found three haplotypes: haplotype I (*Lpx-B1.3* and *Lpx-B1.1b*), haplotype II (*Lpx-B1.2* and *Lpx-B1.1a*), and haplotype III (*Lpx-B1.2* and *Lpx-B1.1c*), exhibiting high, intermediate, and low levels of functional *Lpx-B1* transcripts and enzymatic activity in mature grains, respectively, which is also correlated with β-carotene bleaching. Carrera et al. (2007) reported sequences corresponding to the *Lpx-2* and *Lpx-3* on chromosome group 5 and 4, respectively, and *Lpx-A3* showed significant effects on semolina color, but not on LPX mature grain activity, proving that the LPX activity given by Lpx-A3 acts earlier on grain development.

PPO (EC 1.14.18.1) catalyzes the oxidation of phenolic acids, producing short-chain polymers related to undesirable browning or darkening of pasta products, reducing its apparent quality (Watanabe et al., 2006). There are two non-linked genes controlling PPO activity on durum wheat that have been identified on the homoeologous chromosome 2A and 2B (Jimenez and Dubcovsky, 1999; Nair and Tomar, 2001; Simeone et al., 2002; Watanabe et al., 2004). Watanabe et al. (2006) reported that the locus on chromosome 2A was linked to PPO activity explaining 49.1% of the trait variation in an RIL population made from the crosses between Jennah Khetifa and Cham 1. Two *Ppo* paralogous families were mapped on the homoeologous group 2, named *Ppo-1* (*Ppo-A1* and *Ppo-B1*) and *Ppo-2* (*Ppo-A2* and *Ppo-B2*) (Beecher et al., 2012). *Ppo-A1* was found to have a major role in PPO activity (Simeone et al., 2002; Sun et al., 2005; Anderson et al., 2006; He et al., 2007). Using a marker developed for common wheat (PPO18) (Sun et al., 2005), four *Ppo-A1* alleles, named *Ppo-A1b*, *Ppo-A1f*, *Ppo-A1e*, and *Ppo-A1g*, were detected on durum wheat (He et al., 2009). Taranto et al. (2012) linked the different *Ppo-A1* alleles to distinct levels of PPO activity in 113 accessions of tetraploid wheat. *Ppo-A1f* was associated to high PPO activity, whereas *Ppo-A1b* and *Ppo-A1g* were related to low PPO activity, although they argued there was a consistent variability on the PPO activities associated with each allele. Using the reverse primer of marker PPO18 and a new forward primer specific for the *Ppo-A1* allele, Taranto et al. (2012) developed a new marker (MG18) able to detect the same alleles than PPO18 in a collection of 228 accessions of old, intermediate, and modern

tetraploid wheats, but in a more efficient manner, reducing the variability on PPO activity of each allele-related group. Further decrease on PPO activity can be achieved by selecting also alleles for low PPO activity of the *Ppo-B1* and *Ppo-B2* paralogous genes, using markers MG08 and MG33, respectively, developed by Taranto et al. (2015). *Ppo-B1* and *Ppo-B2* were located 11 cM apart on chromosome arm 2BL. Marker MG08 identified four *Ppo-B1* alleles, related to high (*Ppo-B1c*) and low (*Ppo-B1a, Ppo-B1b*, and *Ppo-B1d*) PPO activity, whereas marker MG33 recognized two *Ppo-B2* alleles, associated to high (*Ppo-B2d*) and low (*Ppo-B2a*) PPO activity levels. The use of these markers in MAS breeding programs has the potential to further improve the color of pasta and durum wheat derived end products.

The aldehyde oxidase 3 (*AO-A3*) gene, located on chromosome 7AL, has been significantly associated to YI and to a QTL linked for YPC (Colasuonno et al., 2017b). Aldehyde oxidase enzymes (AO; EC 1.2.3.1) play roles in the final catalytic steps from carotenoids to ABA (Seo and Koshiba, 2002). qRT-PCR experiments revealed higher levels of *AO-A3* expression in the low YPC cultivar Ciccio in comparison to the high YPC cultivar Svevo. This gene also showed higher expression levels in the later stages of seed formation than *AO-A1* and *AO-A2*, suggesting a major role in the final stages of seed development (Colasuonno et al., 2017b). Colasuonno et al. (2017b) developed a marker for *AO-A3* for DHPLC, which could be useful for MAS programs.

Peroxidase (*PER*) genes have received less attention than the other carotenoid degradation genes on durum wheat, with most of the studies being conducted on common wheat, and no durum wheat specific markers for low peroxidase activity are available to date. Peroxidases (EC 1.11.1.7.) are oxidoreductases that oxidize a vast array of compounds present in plants, using hydrogen peroxide as substrate. They are related to pasta brownness due to the oxidation of phenolic substances. Studies showed that peroxidase genes are located on the homoeologous chromosome groups 1, 2, 3, 4, and 7 (Kobrehel and Feillet, 1975; Benito and de la Vega, 1979; Bosch et al., 1987; Liu et al., 1990; Wei et al., 2015). Up to 12 peroxidase isoforms are present in the durum grain, varying in quantity during kernel development, maturation, and germination, with some isoforms having specific locations between milling fractions (bran, semolina, and embryo), in which isoform P-5 is of importance because of its endosperm specific location, contributing to the darkening of pasta products (Feillet et al., 2000; Fraignier et al., 2000). Fortunately, PER do not show activity during pasta processing, likely due to the unavailability of hydrogen peroxide, its main substrate, whereas it is abundant in semolina (Feillet et al., 2000; Ficco et al., 2014).

NOVEL MUTATIONS IN CAROTENOID GENES

New advances in wheat genomics resources and in molecular technologies contributed to increase the knowledge of carotenoid genes. This included the screening of mutant resource containing chemically induced point mutation variation in candidate genes through TILLING strategy (Targeting Induced Local Lesions in Genomes) (Uauy et al., 2009; Colasuonno et al., 2016; Richaud et al., 2018).

Among all genes involved in the carotenoid pathway, LCYE and LCYB were the only genes extensively studied with this approach, highlighting the complexity of the trait and the difficulty of its modification.

Colasuonno et al. (2016) screened 1,140 mutant lines (0.70–0.85% ethyl methane sulfonate, EMS) focusing on these two target genes. The denaturing high-performance liquid chromatography (DHPLC) allowed to identify a total of 38 and 21 mutations for LCYE and LCYB genes, respectively, with a mutation density of 1/77 kb. Similarly, the analysis of 1,370 DNA individuals from the durum wheat Kronos TILLING mutant population (Uauy et al., 2009) allowed to identify 39 mutants for the LCYE homologues (Richaud et al., 2018) through CelI/agarose method.

Although in both studies, premature stop codons or deleterious missense mutations had been identified, no significant differences in the increment of β-carotene and total carotenoid content among the lines and the relative control were detected. This could be attributed to a high number of effective genes underlying this complex trait and their influence on the final phenotype.

Additional availability of an *in silico* TILLING population (Krasileva et al., 2017) allowed to identify knockout alleles in these target genes and give information about their mutation rate. For instance, of over 1,500 EMS mutagenized lines from the Kronos cultivar, 76 and 128 mutations had been detected in the LCYE and LCYB protein coding regions, respectively. Among all these mutations, only for LCYE gene 7 premature stop codons or deleterious missense mutations resulted to have significant effect on the change of protein composition. The low rate of deleterious SNPs for the target genes marked their main role into the biosynthesis and how some unknown mechanism prevented mutations in these key carotenoid enzymes.

Extensive studies specific for other carotenoid genes are needed to understand the available mutations and their potential effects if combined in double mutants on the final phenotype.

TRANSFER OF QTL OR GENES LINKED TO HIGH YPC

Backcross breeding has been used to transfer gene(s) or QTL of interest from a certain genetic background into an elite cultivar lacking for carotenoids. Subsequently, MAS technology validated the additive effect of the locus/candidate gene and assessed its impact on the new genetic background (Hospital, 2005).

Even though there is a great number of works of QTL linked to high YPC identification (**Table 1**), direct validation on durum wheat and use through introgression is limited. Patil et al. (2018) developed a marker, PSY-1SSR, based on the microsatellite variations in the promoter region of Psy-1, allowing the identification of eight alleles of Psy-A1 and seven alleles of Psy-B1 simultaneously, linked to Qyp.macs-7A, a major QTL for YPC on the long arm of chromosome 7A identified in a PDW 233/Bhalegaon 4 RIL population (Patil et al., 2018). They used this marker to improve YPC through MAS of two different low YPC

Indian cultivars, MACS 3125 and HI 8498, and they were able to follow the introgression of the allele Psy-A1SSRe (linked to high yellowness) from PDW 23, using backcross breeding. MACS 3125 backcrossed lines showed a significant increase in YPC (6.16–7.7 ppm) over the recurrent parent MACS 3125 (3.57 ppm). HI 8498 introgressed lines also showed a significant YPC increase (5.0–7.46 ppm) in comparison to their recurrent parent (3.26 ppm).

MAS is currently being used by CIMMYT and by the Canadian durum wheat breeding programs (Randhawa et al., 2013; N'Diaye et al., 2018) by selecting materials with low LOX activity, with the implementation of the LOXA marker (Carrera et al., 2007), targeting the Lpx-B1.1c allele (Verlotta et al., 2010) for the generation of breeding lines (Randhawa et al., 2013; Dreisigacker et al., 2016).

Even though there is a great number of works of QTL linked to high YPC identification (**Table 1**), direct validation on durum wheat and use through introgression is limited. Patil et al. (2018) developed a marker, PSY-1SSR, based on the microsatellite variations in the promoter region of *Psy-1*, allowing the identification of eight alleles of *Psy-A1* and seven alleles of *Psy-B1* simultaneously, linked to *Qyp.macs-7A*, a major QTL for YPC on the long arm of chromosome 7A identified in a PDW 233/Bhalegaon 4 RIL population (Patil et al., 2018). They used this marker to improve YPC through MAS of two different low YPC Indian cultivars, MACS 3125 and HI 8498, and they were able to follow the introgression of the allele *Psy-A1SSRe* (linked to high yellowness) from PDW 23, using backcross breeding. MACS 3125 backcrossed lines showed a significant increase in YPC (6.16–7.7 ppm) over the recurrent parent MACS 3125 (3.57 ppm). HI 8498 introgressed lines also showed a significant YPC increase (5.0–7.46 ppm) in comparison to their recurrent parent (3.26 ppm).

MAS is currently being used by CIMMYT and by the Canadian durum wheat breeding programs (Randhawa et al., 2013; N'Diaye et al., 2018) by selecting materials with low LOX activity, with the implementation of the LOXA marker (Carrera et al., 2007), targeting the *Lpx-B1.1c* allele (Verlotta et al., 2010) for the generation of breeding lines (Randhawa et al., 2013; Dreisigacker et al., 2016).

FUTURE PERSPECTIVES AND CONCLUSIONS

Understanding the biosynthetic pathway for carotenoid pigment accumulation requires many efforts due to the durum wheat polyploidy and its quantitative nature. The information examined in this article explains the significant goals that have been reached in the last two decades in understanding the genetic and the molecular mechanisms underlying the metabolism and regulation of wheat carotenoids. Furthermore, the characterization of specific plant materials and the release of the durum wheat genome sequences, together with the development of more accurate classes of DNA-based markers and consensus maps, have allowed the identification of important genes involved in the control of carotenoid biosynthesis and catabolism.

Clearly, the most studied and repeatable QTL are those located on chromosomes 3AS (linked to the *LCYE* gene), 7AL, and 7BL (both tightly linked to the *PSY1* genes). Diagnostic markers are available in all these regions for MAS application. Hopefully, other carotenoid QTL regions will likely be further characterized in the future, taking advantage of the recent results and tools for identifying the candidate genes involved in the accumulation/degradation of the carotenoid compounds. This will certainly increase the speed of the genetic gains of carotenoid accumulation, which will benefit the breeding programs and the pasta industry. According to these new resources, we can anticipate an implementation in genotypic selection flanking the traditional phenotypic selection in the durum wheat breeding programs. At the same time, the additive effects of the genes involved in yellowness will generate improved plants through several breeding cycles able to incorporate the beneficial alleles introgressed. Future developments on MAS breeding will focus on selecting many genes alleles at once in order to reach such purpose. Despite all the research in this subject, efforts should be taken on the transfer of knowledge between the bench and the field, because of the current use of the markers described in this review that could potentially benefit the durum wheat breeding programs globally.

Finally, further emphasis of future activities will encompass the analysis of the genetic variability present in the durum wheat germplasm collections (i.e., pre-breeding work), and the TILLING populations, to better understand the functions of the genes involved in the structural and the regulatory system responsible for the YPC trait. Advanced techniques (i.e., CRISPR-Cas9–based genome editing method), will be useful if combined and used to understand the homoeologous silenced gene acting additively and imposing effects on both the total gene expression and the resulting phenotype. Taking these strategies together, the characterization of each gene could provide opportunities for diversifying the genetic architecture of carotenoid pigments and expand the existing allelic variation available for wheat improvement.

AUTHOR CONTRIBUTIONS

PC, AG, RP, and AS designed the review. MM, GC, RT, AB, IM, JA-F, and AC prepared the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

The article processing charge was funded by Conicyt (Fondecyt Regular n° 1161298). This work was partially supported by Conicyt [grants Fondecyt Regular n° 1161298 (AS) and Fondecyt Postdoctorado n° 3180432 (AC)]. This research was also partially funded by a research grant from ISEA AGROSERVICE SpA (S. Severino Marche, MC, Italy).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01347/full#supplementary-material

REFERENCES

AACC International. (2013). "AACC International Official Method 14-50.01", In: *Approved Methods of the American Association of Cereal Chemists, tenth ed.* (MN, U.S.A: St. Paul).

Al-Babili, S., and Bouwmeester, H. J. (2015). Strigolactones, a Novel Carotenoid-Derived Plant Hormone. *Annu. Rev. Plant Biol.* 66 (1), 161–186. doi: 10.1146/annurev-arplant-043014-114759

Anderson, J. V., Patrick Fuerst, E., Hurkman, W. J., Vensel, W. H., and Morris, C. F. (2006). Biochemical and genetic characterization of wheat (*Triticum* spp.) kernel polyphenol oxidases. *J. Cereal Sci.* 44, 353–367. doi: 10.1016/j.jcs.2006.06.008

Asins, M. J., and Perez de la Vega, M. (1985). The inheritance of tetraploid wheat seed peroxidases. *Theor. Appl. Genet.* 71, 61–67. doi: 10.1007/BF00278255

Atienza, S. G., Ballesteros, J., Martin, A., and Hornero-Mendes, D. (2007). Genetic variability of carotenoid concentration and degree of esterification among Tritordeum (xTritordeum Ascherson et Graebner) and durum wheat accessions. *Agric Food Chem.* doi: 10.1021/jf070342p

Beecher, B. S., Carter, A. H., and See, D. R. (2012). Genetic mapping of new seed-expressed polyphenol oxidase genes in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 124, 1463–1473. doi: 10.1007/s00122-012-1801-2

Beleggia, R., Platani, C., Nigro, F., and Papa, R. (2011). Yellow pigment determination for single kernels of durum wheat (*Triticum durum* Desf.). *Cereal Chem.* 88, 504–508. doi: 10.1094/cchem-02-11-0013

Bendich, A., and Olson, J. A. (1989). Biological actions of carotenoids. *FASEB J.* 3, 1927–1932. doi: 10.1096/fasebj.3.8.2656356

Benito, C., and de la Vega, M. P. (1979). The chromosomal location of peroxidase isozymes of the wheat kernel. *Theor. Appl. Genet.* 55, 73–76. doi: 10.1007/bf00285193

Blanco, A., Colasuonno, P., Gadaleta, A., Mangini, G., Schiavulli, A., Simeone, R., et al. (2011). Quantitative trait loci for yellow pigment concentration and individual carotenoid compounds in durum wheat. *J. Cereal Sci.* 54, 255–264. doi: 10.1016/j.jcs.2011.07.002

Borrelli, G. M., De Leonardis, A. M., Fares, C., Platani, C., and Di Fonzo, N. (2003). Effects of modified processing conditions on oxidative properties of semolina dough and pasta. *Cereal Chem.* 80, 225–231. doi: 10.1094/cchem.2003.80.2.225

Borrelli, G. M., De Leonardis, A. M., Platani, C., and Troccoli, A. (2008). Distribution along durum wheat kernel of the components involved in semolina colour. *J. Cereal Sci.* 48, 494–502. doi: 10.1016/j.jcs.2007.11.007

Borrelli, G. M., and Trono, D. (2016). Molecular approaches to genetically improve the accumulation of health-promoting secondary metabolites in staple crops-A case study: the Lipoxygenase-B1 genes and regulation of the carotenoid content in pasta products. *Int. J. Mol. Sci.* 17, 1177. doi: 10.3390/ijms17071177

Bosch, A., Vega, C., and Benito, C. (1987). The peroxidase isozymes of the wheat kernel: tissue and substrate specificity and their chromosomal location. *Theor. Appl. Genet.* 73, 701–706. doi: 10.1007/bf00260779

Brandolini, A., Hidalgo, A., and Moscaritolo, S. (2008). Chemical composition and pasting properties of einkorn (*Triticum monococcum* L. subsp. monococcum) whole meal flour. *J. Cereal Sci.* 47, 599–609. doi: 10.1016/j.jcs.2007.07.005

Brandolini, A., Hidalgo, A., Gabriele, S., and Heun, M. (2015). Chemical composition of wild and feral diploid wheats and their bearing on domesticated wheats. *J. Cereal Sci.* 63, 122–127. doi: 10.1016/j.jcs.2015.03.005

Breseghello, F., and Sorrells, M. E. (2006). Association mapping of kernel size and milling quality in wheat (Triticum aestivum L.) cultivars. *Genetics.* 172(2), 1165–1177.

Britton, G. (1995). Structure and properties of carotenoids in relation to function. *FASEB J.* 9, 1551–1558. doi: 10.1096/fasebj.9.15.8529834

Britton, G. (1998). "Overview of carotenoid biosynthesis," in *Carotenoids: Biosynthesis and Metabolism*. Eds. Britton, G., Liaaen-Jenen, S., and Pfander., H. (Basel, Switzerland: Birkhauser Verlag), 13–147.

Britton, G., and Khachik, F., (2009). "Carotenoids in Food," in *Carotenoids: Nutrition and Health*, vol. 5 . Eds. Britton, G., Pfander, H., and Liaaen-Jensen, S. (Basel: Birkhäuser Basel), 45–66.

Burkhardt, S., and Böhm, V. (2007). Development of a new method for the complete extraction of carotenoids from cereals with special reference to durum wheat (*Triticum durum* Desf.). *J. Agric. Food Chem.* 55, 8295–8301. doi: 10.1021/jf0712853

Bush, W. S., and Moore, J. H. (2012). Chapter 11: genome-wide association studies. *PloS Comput. Biol.* 8, e1002822. doi: 10.1371/journal.pcbi.1002822

Campos, K. M., Royo, C., Schulthess, A., Villegas, D., Matus, I., Ammar, K., et al (2016). Association of phytoene synthase Psy1-A1 and Psy1-B1 allelic variants with semolina yellowness in durum wheat (*Triticum turgidum* L. var. durum). *Euphytica* 207, 109–117. doi: 10.1007/s10681-015-1541-x

Carrera, A., Echenique, V., Zhang, W., Helguera, M., Manthey, F., Schrager, A., et al. (2007). A deletion at the Lpx-B1 locus is associated with low lipoxygenase activity and improved pasta color in durum wheat (*Triticum* turgidum ssp. *durum). J. Cereal Sci.* 45, 67–77. doi: 10.1016/j.jcs.2006.07.001

Cazzonelli, C. I., and Pogson, B. J. (2010). Source to sink: regulation of carotenoid biosynthesis in plants. *Trends Plant Sci.* 15, 266–274. doi: 10.1016/j.tplants.2010.02.003

CIE. (1986). *Publication 15.2 Colorimetry*. Vienna: Commission Internationale de l'Eclairage (CIE).

Colasuonno, P., Gadaleta, A., Giancaspro, A., Nigro, D., Giove, S., Incerti, O., et al. (2014). Development of a high-density SNP-based linkage map and detection of yellow pigment content QTLs in durum wheat. *Mol. Breed.* 34, 1563–1578. doi: 10.1007/s11032-014-0183-3

Colasuonno, P., Incerti, O., Lozito, M. L., Simeone, R., Gadaleta, A., and Blanco, A. (2016). DHPLC technology for high-throughput detection of mutations in a durum wheat TILLING population. *BMC Genet.* 17, 17:43. doi: 10.1186/s12863-016-0350-0

Colasuonno, P., Lozito, M. L., Marcotuli, I., Nigro, D., Giancaspro, A., Mangini, G., et al. (2017a). The carotenoid biosynthetic and catabolic genes in wheat and their association with yellow pigments. *BMC Genomics* 18, 122. doi: 10.1186/s12864-016-3395-6

Colasuonno, P., Marcotuli, I., Lozito, M. L., Simeone, R., Blanco, A., and Gadaleta, A. (2017b). Characterization of aldehyde oxidase (AO) genes involved in the accumulation of carotenoid pigments in wheat grain. *Front. Plant Sci.* 8, 863. doi: 10.3389/fpls.2017.00863

Cong, L., Wang, C., Li, Z., Chen, L., Yang, G., Wang, Y., et al. (2010). cDNA cloning and expression analysis of wheat (*Triticum aestivum* L.) phytoene and ζ-carotene desaturase genes. *Mol. Biol. Rep.* 37, 3351–3361. doi: 10.1007/s11033-009-9922-7

Crawford, A. C., and Francki, M. G. (2013). Chromosomal location of wheat genes of the carotenoid biosynthetic pathway and evidence for a catalase gene on chromosome 7A functionally associated with flour b* colour variation. *Mol. Genet. Genomics* 288, 483–493. doi: 10.1007/s00438-013-0767-3

Choe, E., and Min, D. B. (2005). Chemistry and reactions of reactive oxygen species in foods. *J. Food Sci.* 70, R142–R159. doi: 10.1111/j.1365-2621.2005.tb08329.x

De Simone, V., Menzo, V., De Leonardis, A. M., Ficco, D. B. M., Trono, D., Cattivelli, L., et al. (2010). Different mechanisms control lipoxygenase activity in durum wheat kernels. *J. Cereal Sci.* 52, 121–128. doi: 10.1016/j.jcs.2010.04.003

DellaPenna, D., and Pogson, B. J. (2006). Vitamin synthesis in plants: tocopherols and carotenoids. *Annu. Rev. Plant Biol.* 57, 711–738. doi: 10.1146/annurev.arplant.56.032604.144301

Dibari, B., Murat, F., Chosson, A., Gautier, V., Poncet, C., Lecomte, P., et al. (2012). Deciphering the genomic structure, function and evolution of carotenogenesis related phytoene synthases in grasses. *BMC Genomics* 13, 221. doi: 10.1186/1471-2164-13-221

Digesù, A. M., Platani, C., Cattivelli, L., Mangini, G., and Blanco, A. (2009). Genetic variability in yellow pigment components in cultivated and wild tetraploid wheats. *J. Cereal Sci.* 50, 210–218. doi: 10.1016/j.jcs.2009.05.002

Dong, C. H., Ma, Z. Y., Xia, X. C., Zhang, L. P., and He, Z. H. (2012). Allelic variation at the TaZds-A1 locus on wheat chromosome 2A and development of a functional marker in common wheat. *J. Integr. Agric.* 11, 1067–1074. doi: 10.1016/S2095-3119(12)60099-9

Dreisigacker, S., Sehgal, D., E. Reyes Jaimez, A., Luna Garrido, B., Muñoz Zavala, S., Núñez Ríos, C., et al., (2016). "CIMMYT Wheat Molecular Genetics," in *Laboratory Protocols and Applications to Wheat Breeding* (Mexico: CIMMYT).

Elouafi, I., Nachit, M. M., and Martin, L. M. (2001). Identification of a microsatellite on chromosome 7B showing a strong linkage with yellow pigment in durum wheat (*Triticum turgidum* L. var. durum). *Hereditas* 135, 255–261. doi: 10.1111/j.1601-5223.2001.t01-1-00255.x

Feillet, P., Autran, J.-C., and Icard-Vernière, C. (2000). MINI REVIEW Pasta brownness: an Assessment. *J. Cereal Sci.* 32, 215–233. doi: 10.1006/jcrs.2000.0326

Ficco, D. B. M., Mastrangelo, A. M., Trono, D., Borrelli, G. M., De Vita, P., Fares, C., et al. (2014). The colours of durum wheat: a review. *Crop Pasture Sci.* 65, 1–15. doi: 10.1071/cp13293

Fiedler, J. D., Salsman, E., Liu, Y., Michalak de Jiménez, M., Hegstad, J. B., Chen, B., et al. (2017). Genome-wide association and prediction of grain and semolina quality traits in durum wheat breeding populations. *Plant Genome* 10 (3), 1–12. doi: 10.3835/plantgenome2017.05.0038.

Fraignier, M.-P., Michaux-Ferrière, N., and Kobrehel, K. (2000). Distribution of peroxidases in durum wheat (*Triticum durum*). *Cereal Chem.* 77, 11–17. doi: 10.1094/cchem.2000.77.1.11

Fu, B. X., Chiremba, C., Pozniak, C. J., Wang, K., and Nam, S. (2017). Total phenolic and yellow pigment contents and antioxidant activities of durum wheat milling fractions. *Antioxidants* 6, 78. doi: 10.3390/antiox6040078

Gallagher, C. E., Matthews, P. D., Li, F., and Wurtzel, E. T. (2004). Gene duplication in the carotenoid biosynthetic pathway preceded evolution of the grasses. *Plant Physiol.* 135, 1776–1783. doi: 10.1104/pp.104.039818

Garbus, I., Carrera, A. D., Dubcovsky, J., and Echenique, V. (2009). Physical mapping of durum wheat lipoxygenase genes. *J. Cereal Sci.* 50, 67–73. doi: 10.1016/j.jcs.2009.02.008

Garbus, I., Soresi, D., Romero, J., and Echenique, V. (2013). Identification, mapping and evolutionary course of wheat lipoxygenase-1 genes located on the A genome. *J. Cereal Sci.* 58, 298–304. doi: 10.1016/j.jcs.2013.05.012

Guaadaoui, A., Benaicha, S., Elmajdoub, N., Bellaoui, M., and Hamal, A. (2014). What is a bioactive compound? A combined definition for a preliminary consensus. *Int. J. Nutr. Food Sci.* 3, 174–179. doi: 10.11648/j.ijnfs.20140303.16

Gupta, P. K., Rustgi, S., and Kulwal, P. L. (2005). Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Mol. Biol.* 57, 461–485. doi: 10.1007/s11103-005-0257-z

Halliwell, B. (2011). Free radicals and antioxidants - quo vadis? *Trends Pharmacol. Sci.* 32, 125–130. doi: 10.1016/j.tips.2010.12.002

He, X. Y., He, Z. H., Morris, C. F., and Xia, X. C. (2009). Cloning and phylogenetic analysis of polyphenol oxidase genes in common wheat and related species. *Genet. Resour. Crop Evol.* 56, 311. doi: 10.1007/s10722-008-9365-3

He, X. Y., He, Z. H., Zhang, L. P., Sun, D. J., Morris, C. F., Fuerst, E. P., et al. (2007). Allelic variation of polyphenol oxidase (PPO) genes located on chromosomes 2A and 2D and development of functional markers for the PPO genes in common wheat. *Theor. Appl. Genet.* 115, 47–58. doi: 10.1007/s00122-007-0539-8

He, X. Y., Zhang, Y. L., He, Z. H., Wu, Y. P., Xiao, Y. G., Ma, C. X., et al. (2008). Characterization of phytoene synthase 1 gene (Psy1) located on common wheat chromosome 7A and development of a functional marker. *Theor. Appl. Genet.* 116, 213–221. doi: 10.1007/s00122-007-0660-8

Hentschel, V., Kranl, K., Hollmann, J., Lindhaeur, M. G., Bohm, V., and Bitsch, R. (2002). Spectrophotometric determination of yellow pigment content and evaluation of carotenoids by high-performance liquid chromatography in durum wheat grain. *J. Agric. Food Chem.* 50, 6663–6668. doi: 10.1021/jf025701p

Hessler, T. G., Thomson, M. J., Benscher, D., Nachit, M., and Sorrells, M., (2002). Association of a lipoxygenase locus, Lpx-B1, with variation in lipoxygenase activity in durum wheat seeds. *Crop Sci.* 42, 1695–1700. doi: 10.2135/cropsci2002.1695.

Holtman, W., Gert van, D., Norbert, J. A. S., and Anneke, C. D. (1996). Differential expression of lipoxygenase isoenzymes in embryos of germinating barley. *Plant Physiol.* 111, 569–576. doi: 10.1104/pp.111.2.569

Hospital, F. (2005). Selection in backcross programmes. *Philos. Trans. R. Soc Lond. B. Biol. Sci.* 360, 1503–1511. doi: 10.1098/rstb.2005.1670

Howitt, C. A., Cavanagh, C. R., Bowerman, A. F., Cazzonelli, C., Rampling, L., Mimica, J. L., et al. (2009). Alternative splicing, activation of cryptic exons and amino acid substitutions in carotenoid biosynthetic genes are associated with lutein accumulation in wheat endosperm. *Funct. Integr. Genomics* 9, 363–376. doi: 10.1007/s10142-009-0121-3

ICC. (1990). "ICC Method 152," in *Standard Methodsof the International Association for Cereal Science and Technology* (Detmold, Germany: (Verlag Moritz SchäferICC)).

Irakli, M. N., Samanidou, V. F., and Papadoyannis, I. N. (2011). Development and validation of an HPLC method for the simultaneous determination of tocopherols, tocotrienols and carotenoids in cereals after solid-phase extraction. *J. Sep. Sci.* 34, 1375–1382. doi: 10.1002/jssc.201100077

Jahns, P., and Holzwarth, A. R. (2012). The role of the xanthophyll cycle and of lutein in photoprotection of photosystem II. *Biochim. Biophys. Acta* 1817, 182–193. doi: 10.1016/j.bbabio.2011.04.012

Jimenez, M., and Dubcovsky, J. (1999). Chromosome location of genes affecting polyphenol oxidase activity in seeds of common and durum wheat. *Plant Breed.* 118, 395–398. doi: 10.1046/j.1439-0523.1999.00393.x

Kabbaj, H., Sall, A. T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., et al. (2017). Genetic diversity within a Global panel of durum wheat (*Triticum durum*) landraces and modern germplasm reveals the history of alleles exchange. *Front. Plant Sci.* 8, 1277. doi: 10.3389/fpls.2017.01277

Ke, Q., Kang, L., Kim, H. S., Xie, T., Liu, C., Ji, C. Y., et al. (2019). Down-regulation of lycopene ε-cyclase expression in transgenic sweetpotato plants increases the carotenoid content and tolerance to abiotic stress. *Plant Sci.* 281, 52–60. doi: 10.1016/j.plantsci.2019.01.002

Kean, E. G., Bordenave, N., Ejeta, G., Hamaker, B. R., and Ferruzzi, M. G. (2011). Carotenoid bioaccessibility from whole grain and decorticated yellow endosperm sorghum porridge. *J. Cereal Sci.* 54, 450–459. doi: 10.1016/j.jcs.2011.08.010

Kelly, M. E., Ramkumar, S., Sun, W., Colon Ortiz, C., Kiser, P. D., Golczak, M., et al. (2018). The biochemical basis of vitamin A production from the asymmetric carotenoid β-cryptoxanthin. *ACS Chem. Biol.* 13, 2121–2129. doi: 10.1021/acschembio.8b00290

Khoo, H. E., Prasad, K. N., Kong, K. W., Jiang, Y., and Ismail, A. (2011). Carotenoids and their isomers: color pigments in fruits and vegetables. *Molecules* 16, 1710–1738. doi: 10.3390/molecules16021710

Kobrehel, K., and Feillet, P. (1975). Identification of genomes and chromosomes involved in peroxidase synthesis of wheat seeds. *Can. J. Bot.* 53, 2336–2344. doi: 10.1139/b75-259

Konopka, I., Kozirok, W., and Rotkiewicz, D. (2004). Lipids and carotenoids of wheat grain and flour and attempt of correlating them with digital image analysis of kernel surface and cross-sections. *Food Res. Int.* 37, 429–438. doi: 10.1016/j.foodres.2003.12.009

Krasileva, K. V., Vasquez-Gross, H. A., Howell, T., Bailey, P., Paraiso, F., Clissold, L., et al. (2017). Uncovering hidden variation in polyploidy wheat. *Proceed Nat. Acad. Sci.* 114, E913. doi: 10.1073/pnas.1619268114

Lachman, J., Martinek, P., Kotikova, Z., Orsák, M., and Šulc, M. (2017). Genetics and chemistry of pigments in wheat grain -A review. *J. Cereral Sci.* 74, 145–154. doi: 10.1016/j.jcs.2017.02.007

Landrum, J. T. (2009). "Carotenoids: physical, chemical, and biological functions and properties,". Ed. Landrum., J. T. (Boca Raton, FL: CRC Press). doi: 10.1201/9781420052312

Leenhardt, F., Lyan, B., Rock, E., Boussard, A., Potus, J., Chanliaud, E., et al. (2006). Genetic variability of carotenoid concentration, and lipoxygenase and peroxidase activities among cultivated wheat species and bread wheat varieties. *Eur. J. Agron.* 25, 170–176. doi: 10.1016/j.eja.2006.04.010

Leong, H. Y., Show, P. L., Lim, M. H., Ooi, C. W., and Ling, T. C. (2018). Natural red pigments from plants and their health benefits: a review. *Food Rev. Int.* 34, 463–482. doi: 10.1080/87559129.2017.1326935

Lipka, A. E., Buckler, E. S., Tian, F., Peiffer, J., Li, M., Gore, M. A., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444

Liu, C. J., Chao, S., and Gale, M. D. (1990). The genetical control of tissue-specific peroxidases, Per-1, Per-2, Per-3, Per-4, and Per-5 in wheat. *Theor. Appl. Genet.* 79 (3), 305–313. doi: 10.1007/bf01186072

Maccaferri, M., Sanguineti, M. C., Natoli, V., Ortega, J. L. A., Salem, M. B., Bort, J., et al. (2006). A panel of elite accessions of durum wheat (Triticum durum Desf.) suitable for association mapping studies. *Plant Gen. Res.* 4 (1), 79–85.

Maccaferri, M., Cane', M. A., Sanguineti, M. C., Salvi, S., Colalongo, M. C., Massi, A., et al. (2014). A consensus framework map of durum wheat (*Triticum durum* Desf.) suitable for linkage disequilibrium analysis and genome-wide association mapping. *BMC Genomics* 15, 873. doi: 10.1186/1471-2164-15-873

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51(5), 885–895. doi: 10.1038/s41588-019-0381-3

Maccaferri, M., Sanguineti, M. C., Demontis, A., El-Ahmed, A., Garcia del Moral, L., Maalouf, F., et al. (2011). Association mapping in durum wheat grown across a broad range of water regimes. *J. Exp. Bot.* 62(2), 409–438.

Mazzeo, M. F., Di Stasio, L., D'Ambrosio, C., Arena, S., Scaloni, A., Corneti, S., et al. (2017). Identification of early represented gluten proteins during durum wheat grain development. *J. Agric. Food Chem.* 65, 3242–3250. doi: 10.1021/acs.jafc.7b00571

McCaig, T. N., McLeod, J. G., Clarke, J. M., and DePauw, R. M. (1992). Measurement of durum pigment with an NIR instrument operating in the visible range. *Cereal Chem.* 69, 671–672.

Mezzomo, N., and Ferreira, S. R. S. (2016). Carotenoids functionality, sources, and processing by supercritical technology: a review. *J. Chem.* 1–16. doi: 10.1155/2016/3164312

Mezzomo, N., Tenfen, L., Farias, M. S., Friedrich, M. T., Pedrosa, R. C., and Ferreira, S. R. S. (2015). Evidence of anti-obesity and mixed hypolipidemic effects of extracts from pink shrimp (Penaeus brasiliensis and Penaeus paulensis) processing residue. *J. Supercrit. Fluids* 96, 252–261. doi: 10.1016/j.supflu.2014.09.021

Morris, C. F. (2018). Determinants of wheat noodle color. *J. Sci. Food Agric.* 98, 5171–5180. doi: 10.1002/jsfa.9134

N'Diaye, A., Haile, J. K., Cory, A. T., Clarke, F. R., Clarke, J. M., Knox, R. E., et al. (2017). Single marker and haplotype-based association analysis of semolina and pasta colour in elite durum wheat breeding lines using a high-density consensus map. *PloS One* 12, e0170941. doi: 10.1371/journal.pone.0170941

N'Diaye, A., Haile, J. K., Nilsen, K. T., Walkowiak, S., Ruan, Y., Singh, A. K., et al. (2018). Haplotype loci under selection in Canadian durum wheat germplasm over 60 years of breeding: Association with grain yield, quality traits, protein loss, and plant height. *Front. Plant Sci.* 9 (1589), 1–19. doi: 10.3389/fpls.2018.01589

Nair, S. K., and Tomar, S. M. S. (2001). Genetics of phenol colour reaction of grains and glumes in tetraploid and hexaploid wheats. *J. Genet. Breed.* 55, 369–373.

Nazco, R., Villegas, D., Ammar, K., Peña, R. J., Moragues, M., and Royo, C. (2012). Can Mediterranean durum wheat landraces contribute to improved grain quality attributes in modern cultivars? *Euphytica* 185, 1–17. doi: 10.1007/s10681-011-0588-6

Nisar, N., Li, L., Lu, S., Khin, , Nay C., Pogson, and Barry, J. (2015). Carotenoid metabolism in plants. *Mol. Plant* 8, 68–82. doi: 10.1016/j.molp.2014.12.007

Oliver, J., and Palou, A. (2000). Chromatographic determination of carotenoids in foods. *J. Chromatogr. A.* 881, 543–555. doi: 10.1016/S0021-9673(00)00329-0

Olson, J. A., and Krinsky, N. I. (1995). Introduction: the colorful, fascinating world of the carotenoids: important physiologic modulators. *FASEB J.* 9, 1547–1550. doi: 10.1096/fasebj.9.15.8529833

Panfili, G., Fratianni, A., and Irano, M. (2004). Improved normal-phase high-performance liquid chromatography procedure for the determination of carotenoids in cereals. *J. Agric. Food Chem.* 52, 6373–6377. doi: 10.1021/jf0402025

Patil, R., Oak, M., Deshpande, A., and Tamhankar, S. (2018). Development of a robust marker for Psy-1 homoeologs and its application in improvement of yellow pigment content in durum wheat. *Mol. Breed.* 38, 136. doi: 10.1007/s11032-018-0895-x

Patil, R. M., Oak, M. D., Tamhankar, S. A., Sourdille, P., and Rao, V. S. (2008). Mapping and validation of a major QTL for yellow pigment content on 7AL in durum wheat (*Triticum turgidum* L. ssp durum). *Mol. Breed.* 21, 485–496. doi: 10.1007/s11032-007-9147-1

Paznocht, L., Kotikova, Z., Orsak, M., Lachman, J., and Martinek, P. (2019). Carotenoid changes of colored-grain wheat flours during bun-making. *Food Chem.* 277, 725–734. doi: 10.1016/j.foodchem.2018.11.019

Pozniak, C. J., Knox, R. E., Clarke, F. R., and Clarke, J. M. (2007). Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. *Theor. Appl. Genet.* 114, 525–537. doi: 10.1007/s00122-006-0453-5

Qin, X., Zhang, W., Dubcovsky, J., and Tian, L. (2012). Cloning and comparative analysis of carotenoid beta-hydroxylase genes provides new insights into carotenoid metabolism in tetraploid (*Triticum turgidum* ssp. durum) and hexaploid (*Triticum aestivum*) wheat grains. *Plant Mol. Biol.* 80, 631–646. doi: 10.1007/s11103-012-9972-4

Randhawa, H. S., Asif, M., Pozniak, C., Clarke, J. M., Graf, R. J., Fox, S. L., et al. (2013). Application of molecular markers to wheat breeding in Canada. *Plant Breed.* 132, 458–471. doi: 10.1111/pbr.12057

Reimer, S., Pozniak, C. J., Clarke, F. R., Clarke, J. M., Somers, D. J., Knox, R. E., et al. (2008). Association mapping of yellow pigment in an elite collection of durum wheat cultivars and breeding lines. *Genome* 51, 1016–1025. doi: 10.1139/g08-083

Richaud, D., Stange, C., Gadaleta, A., Colasuonno, P., Parada, R., and Schwember, A. R. (2018). Identification of Lycopene epsilon cyclase (LCYE) gene mutants to potentially increase β-carotene content in durum wheat (*Triticum* turgidum L. ssp. durum) through TILLING. *PloS One* 13, e0208948. doi: 10.1371/journal.pone.0208948

Rodriguez-Amaya, D., and Kimura, M., (2004). "HarvestPlus Handbook for Carotenoids Analysis," in *HarvestPlus Technical Monographs Series*, vol. 2. (Washington DC, USA and Cali, Colombia). IFPRI CIAT,.

Rodriguez-Concepcion, M., Avalos, J., Bonet, M. L., Boronat, A., Gomez-Gomez, L., Hornero-Mendez, D., et al. (2018). A global perspective on carotenoids: Metabolism, biotechnology, and benefits for nutrition and health. *Prog. Lipid Res.* 70, 62–93. doi: 10.1016/j.plipres.2018.04.004

Roncallo, P. F., Cervigni, G. L., Jensen, C., Miranda, R., Carrera, A. D., Helguera, M., et al. (2012). QTL analysis of main and epistatic effects for flour color traits in durum wheat. *Euphytica* 185, 77–92. doi: 10.1007/s10681-012-0628-x

Rosello, M., Royo, C., Alvaro, F., Villegas, D., Nazco, R., and Soriano, J. M. (2018). Pasta-making quality QTLome from Mediterranean durum wheat landraces. *Front. Plant Sci.* 9, 1512. doi: 10.3389/fpls.2018.01512

Sandmann, G. (2001). Genetic manipulation of carotenoid biosynthesis: strategies, problems and achievements. *Trends Plant Sci.* 6, 14–17. doi: 10.1016/S1360-1385(00)01817-3

Schroeder, W. A., and Johnson, E. A. (1995). Singlet oxygen and peroxyl radicals regulate carotenoid biosynthesis in Phaffia rhodozyma. *J. Biol. Chem.* 270, 18374–18379. doi: 10.1074/jbc.270.31.18374

Schulthess, A., Matus, I., and Schwember, A. (2013). Genotypic and environmental factors and their interactions determine semolina color of elite genotypes of durum wheat (*Triticum turgidum* L. var. durum) grown in different environments of Chile. *Field Crops Res.* 149, 234–244. doi: 10.1016/j.fcr.2013.05.001

Schulthess, A., and Schwember, A. R. (2013). Improving durum wheat (*Triticum turgidum* L. *var durum*) grain yellow pigment content through plant breeding. *Cienc. Invest. Agr.* 40, 475–490. doi: http://dx.doi.org/10.7764/rcia.v40i3.1157

Seo, M., and Koshiba, T. (2002). Complex regulation of ABA biosynthesis in plants. *Trends Plant Sci.* 7, 41–48.

Shen, R., Yang, S., Zhao, G., Shen, Q., and Diao, X. (2015). Identification of carotenoids in foxtail millet (Setaria italica) and the effects of cooking methods on carotenoid content. *J. Cereal Sci.* 61, 86–93. doi: 10.1016/j.jcs.2014.10.009

Si, H., Ma, C., Wang, X., and He, X. (2012). Variability of polyphenol oxidase (PPO) alleles located on chromosomes 2A and 2D can change the wheat kernel PPO activity. *Aust. J. Crop Sci.* 6, 444–449.

Simeone, R., Pasqualone, A., Clodoveo, M., and Blanco, A. (2002). Genetic mapping of polyphenol oxidase in tetraploid wheat. *Cell. Mol. Biol. Lett.* 7, 763–769.

Singh, A., Reimer, S., Pozniak, C. J., Clarke, F. R., Clarke, J. M., Knox, R. E., et al. (2009). Allelic variation at Psy1-A1 and association with yellow pigment in durum wheat grain. *Theor. Appl. Genet.* 118, 1539–1548. doi: 10.1007/s00122-009-1001-x

Sissons, M. (2008). Role of durum wheat composition on the quality of pasta and bread. *Food* 2, 75–90.

Sommer, A., and Davidson, F. R. (2002). Assessment and control of vitamin A deficiency: The annecy accords. *J. Nutr.* 132, 2845S–2850S. doi: 10.1093/jn/132.9.2845S

Subira, J., Peña, R. J., Álvaro, F., Ammar, K., Ramdani, A., and Royo, C. (2014). Breeding progress in the pasta-making quality of durum wheat cultivars released in Italy and Spain during the 20th Century. *Crop Pasture Sci.* 65, 16–26. doi: 10.1071/CP13238

Sun, D. J., He, Z. H., Xia, X. C., Zhang, L. P., Morris, C. F., Appels, R., et al. (2005). A novel STS marker for polyphenol oxidase activity in bread wheat. *Mol. Breed.* 16, 209–218. doi: 10.1007/s11032-005-6618-0

Sun, T., Yuan, H., Cao, H., Yazdani, M., Tadmor, Y., and Li, L. (2018). Carotenoid metabolism in plants: The role of plastids. *Mol. Plant* 11, 58–74. doi: 10.1016/j.molp.2017.09.010

Taranto, F., Delvecchio, L. N., Mangini, G., Del Faro, L., Blanco, A., and Pasqualone, A. (2012). Molecular and physico-chemical evaluation of enzymatic browning of whole meal and dough in a collection of tetraploid wheats. *J. Cereal Sci.* 55, 405–414. doi: 10.1016/j.jcs.2012.02.005

Taranto, F., Mangini, G., Pasqualone, A., Gadaleta, A., and Blanco, A. (2015). Mapping and allelic variations of Ppo-B1 and Ppo-B2 gene-related polyphenol oxidase activity in durum wheat. *Mol. Breed.* 35, 80. doi: 10.1007/s11032-015-0272-y

Tsilo, T., Hareland, A., Chao, S., and Anderson, J. A. (2011). Genetic mapping and QTL analysis of flour color and milling yield related traits using recombinant inbred lines in hard red spring wheat. *Crop Sci.* 51, 237–246. doi: 10.2135/cropsci2009.12.0711

Uauy, C., Paraiso, F., Colasuonno, P., Tran, R. K., Tsai, H., Berardi, S., et al. (2009). A modified TILLING approach to detect induced mutations in tetraploid and hexaploid wheat. *BMC Plant Biol.* 9 (115), 1–14. doi: 10.1186/1471-2229-9-115

van den Berg, H., Faulks, R., Granado, H. F., Hirschberg, J., Olmedilla, B., Sandmann, G., et al. (2000). The potential for the improvement of carotenoid levels in foods and the likely systemic effects. *J. Sci. Food Agric.* 80, 880–912.

Vargas, V. H., Schulthess, A., Royo, C., Matus, I., and Schwember, A. R. (2016). Transcripts levels of Phytoene synthase 1 (Psy-1) are associated to semolina yellowness variation in *durum* wheat (*Triticum turgidum* L. *ssp. J. Cereal Sci.* 68, 155–163. doi: 10.1016/j.jcs.2016.01.011

Verlotta, A., De Simone, V., Mastrangelo, A. M., Cattivelli, L., Papa, R., and Trono, D. (2010). Insight into durum wheat Lpx-B1: a small gene family coding for the lipoxygenase responsible for carotenoid bleaching in mature grains. *BMC Plant Biol.* 10, 263. doi: 10.1186/1471-2229-10-263

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Watanabe, N., Akond, A. S. M. G. M., and Nachit, M. M. (2006). Genetic mapping of the gene affecting polyphenol oxidase activity in tetraploid durum wheat. *J. Appl. Genet.* 47, 201–205. doi: 10.1007/BF03194624

Watanabe, N., Takeuchi, A., and Nakayama, A. (2004). Inheritance and chromosomal location of the homoeologous genes affecting phenol colour reaction of kernels in durum wheat. *Euphytica* 139, 87–93. doi: 10.1007/s10681-004-2255-7

Wei, J., Geng, H., Zhang, Y., Liu, J., Wen, W., Zhang, Y., et al. (2015). Mapping quantitative trait loci for peroxidase activity and developing gene-specific markers for TaPod-A1 on wheat chromosome 3AL. *Theor. Appl. Genet.* 128, 2067–2076. doi: 10.1007/s00122-015-2567-0

Young, A., and Lowe, G. (2018). Carotenoids-Antioxidant properties. *Antioxidants* 7, 28. doi: doi 10.3390/antiox7020028

Zeng, J., Wang, X., Miao, Y., Wang, C., Zang, M., Chen, X., et al. (2015). Metabolic engineering of wheat provitamin A by simultaneously overexpressing CrtB and silencing carotenoid hydroxylase (TaHYD). *J. Agric. Food Chem.* 63, 9083–9092. doi: 10.1021/acs.jafc.5b04279

Zhang, B., Liu, J., Cheng, L., Zhang, Y., Hou, S., Sun, Z., et al. (2019). Carotenoid composition and expression of biosynthetic genes in yellow and white foxtail millet [Setaria italica (L.) Beauv]. *J. Cereal Sci.* 85, 84–90. doi: 10.1016/j.jcs.2018.11.005

Zhang, F., Chen, F., Wu, P., Zhang, N., and Cui, D. (2015). Molecular characterization of lipoxygenase genes on chromosome 4BS in Chinese bread wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 128, 1467–1479. doi: 10.1007/s00122-015-2518-9

Zhang, W., Chao, S., Manthey, F., Chicaiza, O., Brevis, J. C., Echenique, V., et al. (2008). QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. *Theor. Appl. Genet.* 117, 1361–1377. doi: 10.1007/s00122-008-0869-1

Zhang, W., and Dubcovsky, J. (2008). Association between allelic variation at the Phytoene synthase 1 gene and yellow pigment content in the wheat grain. *Theor. Appl. Genet.* 116, 635–645. doi: 10.1007/s00122-007-0697-8

Zhang, Z., Ersoz, E., Lai, C.-Q., Todhunter, R. J., Tiwari, H. K., Gore, M. A., et al. (2010). Mixed linear model approach adapted for genome-wide association studies. *Nat. Genet.* 42, 355. doi: 10.1038/ng.546

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 November 2019
doi: 10.3389/fpls.2019.01311

Multi-Trait, Multi-Environment Genomic Prediction of Durum Wheat With Genomic Best Linear Unbiased Predictor and Deep Learning Methods

Osval A. Montesinos-López[1], Abelardo Montesinos-López[2], Roberto Tuberosa[3], Marco Maccaferri[3], Giuseppe Sciara[3], Karim Ammar[4] and José Crossa[4]**

[1] Facultad de Telemática, Universidad de Colima, Colima, Mexico, [2] Departamento de Matemáticas, Centro Universitario de Ciencias Exactas e Ingenierías (CUCEI), Universidad de Guadalajara, Guadalajara, Mexico, [3] Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy, [4] Global Wheat Breeding Program, International Maize and Wheat Improvement Center (CIMMYT), Mexico City, Mexico

OPEN ACCESS

Edited by:
Roberto Papa, Marche Polytechnic University, Italy

Reviewed by:
Tania Gioia, University of Basilicata, Italy
Ester Murube Torcida, Servicio Regional de Investigación y Desarrollo Agroalimentario (SERIDA), Spain

***Correspondence:**
Roberto Tuberosa
roberto.tuberosa@unibo.it
Jose Crossa
j.crossa@cgiar.org

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *18 June 2019*
Accepted: *20 September 2019*
Published: *08 November 2019*

Citation:
Montesinos-López OA, Montesinos-López A, Tuberosa R, Maccaferri M, Sciara G, Ammar K and Crossa J (2019) Multi-Trait, Multi-Environment Genomic Prediction of Durum Wheat With Genomic Best Linear Unbiased Predictor and Deep Learning Methods. Front. Plant Sci. 10:1311. doi: 10.3389/fpls.2019.01311

Although durum wheat (*Triticum turgidum* var. *durum* Desf.) is a minor cereal crop representing just 5–7% of the world's total wheat crop, it is a staple food in Mediterranean countries, where it is used to produce pasta, couscous, bulgur and bread. In this paper, we cover multi-trait prediction of grain yield (GY), days to heading (DH) and plant height (PH) of 270 durum wheat lines that were evaluated in 43 environments (country–location–year combinations) across a broad range of water regimes in the Mediterranean Basin and other locations. Multi-trait prediction analyses were performed by implementing a multi-trait deep learning model (MTDL) with a feed-forward network topology and a rectified linear unit activation function with a grid search approach for the selection of hyper-parameters. The results of the multi-trait deep learning method were also compared with univariate predictions of the genomic best linear unbiased predictor (GBLUP) method and the univariate counterpart of the multi-trait deep learning method (UDL). All models were implemented with and without the genotype × environment interaction term. We found that the best predictions were observed without the genotype × environment interaction term in the UDL and MTDL methods. However, under the GBLUP method, the best predictions were observed when the genotype × environment interaction term was taken into account. We also found that in general the best predictions were observed under the GBLUP model; however, the predictions of the MTDL were very similar to those of the GBLUP model. This result provides more evidence that the GBLUP model is a powerful approach for genomic prediction, but also that the deep learning method is a practical approach for predicting univariate and multivariate traits in the context of genomic selection.

Keywords: durum wheat, deep learning, multi-trait, univariate trait, GBLUP, genomic selection

INTRODUCTION

Nowadays, wheat is the most widespread crop around the world, for it is cultivated on approximately 219 million hectares. It is a basic staple food of mankind, since it provides 18% of the daily intake of calories and 20% of protein (Royo et al., 2017). In the Mediterranean region, wheat covers 27% of arable land, 60% of which is cultivated with durum wheat (*Triticum turgidum* var. *durum* Desf.) used for pasta, couscous, bulgur and bread production. Although durum represents just 5–7% of the world's total wheat crop, it is a staple food of the Mediterranean diet, widely recognized for its health benefits and the prevention of cardiovascular disease. One reason that durum wheat is chosen for manufacturing premium pasta is that it is the hardest of all wheats due to its density, high protein content and gluten strength, which are essential features for producing firm pasta with consistent cooking quality (Royo et al., 2017).

Durum breeding based on genomics-assisted approaches will play an increasingly important role in delivering cultivars more resilient to climate change and with more nutritious and high-quality semolina (Tuberosa and Pozniak, 2014). Accordingly, some durum wheat breeding programs are adopting genomic selection (GS) to accelerate early identification and selection of superior genotypes (Michel et al., 2019; Steiner et al., 2019). GS was first proposed by Meuwissen et al. (2001) as a method for selecting candidate individuals with the help of a regression model that uses all the dense molecular markers simultaneously as independent variables; with this information, a model is trained using a reference population composed of genomic information and phenotypic information. The trained model is then used with a testing population to produce predictions of the individuals in the testing population that were not phenotyped but only genotyped. Applications of GS are common in many areas of animal and plant breeding since there is empirical evidence that this technology has the power to: (a) significantly reduce the time needed to develop new varieties or animals, (b) increase genetic gain in a shorter period of time, and (c) revolutionize the traditional way of developing plant and animals.

For durum wheat breeding, grain yield and semolina quality traits are important selection criteria usually applied and tested in late generations in relatively few lines due to high screening cost, which lowers selection efficiency due to the advancement of undesirable lines into large and expensive yield trials for grain yield and quality trait testing (Fiedler et al., 2017). For this reason, the potential application of genomic selection (GS) in a durum wheat breeding program using 1,184 durum wheat breeding lines was investigated by Fiedler et al. (2017), with prediction accuracies ranging from 0.27 to 0.66 for five quality traits, which pointed out the importance of GS for further enhancing breeding efficiency in durum cultivar development. On the other hand, Crossa et al. (2016) used a genomic marker × environment interaction model for (i) making genome-based predictions of untested individuals, as well as (ii) identifying genomic regions whose effects are stable across environments and other genomic regions that show environmental specificity. The same authors used a multi-parental durum wheat population that was evaluated for grain yield, grain volume weight, thousand-kernel weight and heading date in four environments in Italy. The marker × environment interaction model had better genomic-enabled prediction accuracy than the single-environment or across-environment models. The marker × environment model found that genes controlling heading date, *Ppd* and *FT* on chromosomes 2A, 2B and 7A, showed stable effects across environments as well as environment-specific effects. For grain yield, regions in chromosomes 2B and 7A had large marker effects.

Additionally, since more than one trait was measured in the durum wheat experiments of this study, we performed multi-trait analyses that outperformed univariate analyses. Empirical evidence indicates that multi-trait analyses outperform univariate analyses in terms of prediction performance when the correlation between traits is moderate or large (Jia and Jannink, 2012; Jiang et al., 2015; Montesinos-Lopez et al., 2016; Schulthess et al., 2018; Covarrubias-Pazaran et al., 2018; Montesinos-López et al., 2018b; Montesinos-López et al., 2019a). However, multi-trait analysis is more complex than univariate analysis and more prone to overfitting; for this reason, in many cases the prediction performance of the former is worse than that of the latter. The problem of overfitting is very challenging, not only in conventional multi-trait analysis, but also in conventional statistical and machine learning algorithms like deep learning models.

Deep learning is a generalization of artificial neural networks where the number of layers used is more than one, which helps to capture complex patterns in the data at the cost of increasing the computational resources required due to the fact that more neurons are used. It should be pointed out that deep learning methodology is being implemented in different areas of knowledge such as astrophysics (for classifying exoplanets), geology (for predicting earthquakes), information technology (for classifying emails), botany (for classifying species using photos), engineering (for developing self-driving cars) and meteorology (for predicting time series), among others. In biological sciences there are also many successful applications of deep learning (Alipanahi et al., 2015; Cole et al., 2017; Pan and Shen, 2017; Tavanaei et al., 2017), while in the context of GS, this methodology is applied to continuous (univariate and multivariate), binary and ordinal (univariate) and mixed traits (continuous, binary and ordinal) (Montesinos-López et al., 2018a; Montesinos-López et al., 2018b; Montesinos-López et al., 2019a; Montesinos-López et al., 2019b).

In order to study the feasibility of using GS methodology to select durum wheat genotypes early in time, we evaluated the prediction performance of three statistical learning methods, namely (i) the univariate best linear unbiased predictor (GBLUP), (ii) the multi-trait deep learning model (MTDL), and (iii) the univariate deep learning (UDL) model. Each type of model was evaluated taking into account genotype × environment interaction (I) and ignoring it (WI) in order to evaluate the impact of the interaction term on the prediction performance of out-of-sample genotypes not used for training the model. All these models were implemented using real data sets from multi-environment trials conducted mainly in the Mediterranean Basin.

MATERIALS AND METHODS

Experimental Data

Phenotypic Data

This data set originated in durum wheat experiments conducted with (i) 189 elite cultivars collectively named the Durum Wheat Panel assembled at the University of Bologna and first described by Maccaferri et al. (2006), and (ii) an additional set of 81 elite cultivars contributed by collaborators worldwide. Field experiments were conducted mainly in Mediterranean countries, Hungary and Mexico under both rainfed and irrigated conditions. **Table 1** reports the acronyms and further details of the environments under study.

The analyzed traits were grain yield (GY), days to heading (DH) and plant height (PH) evaluated in 270 lines. Field trials were conducted over 11 years, and in each year only some locations were evaluated. A total of 43 environments (country-site-year combinations) were included in this study. The years under the study were 2003–2009 and 2012–2015. Other traits of interest were also measured in addition to GY, DH and PH, but due to the scarcity of the data, this study only considered these three traits and information from 43 environments. It is important to mention that the number of lines in each environment ranged from a minimum of 57 lines to a maximum of 193 lines, with a mean and median of 180.9 and 186 lines, respectively.

Genotypic Data

The genotypes of all 270 lines were obtained using 24,576 single nucleotide polymorphisms (SNPs) from the consensus map of tetraploid wheat assembled by Maccaferri et al. (2015) as a bridge to integrate durum and bread wheat genomics and breeding mapped in tetraploid wheat. The tetraploid consensus map incorporates SSR, DArT® and SNP markers from 13 mapping populations. The SSR and DArT® profiles (Mantovani et al., 2008; Maccaferri et al., 2011) were integrated with the high-density Infinium® iSelect® Illumina 90K SNP array (Wang et al., 2014; Maccaferri et al., 2015). DNA was extracted from a bulk of 25 one-week-old seedlings per accession using the DNeasy 96 Plant Kit (Qiagen GmbH, Hilden, Germany). The final array included 81,587 transcript-associated SNPs, 8,000 of which are durum-specific SNPs (Wang et al., 2014). Those SNPs with >10% missing values or <0.05 minor allele frequency were excluded. After line-specific quality control, 14,163 SNPs were retained.

Statistical Models

Multiple-Environment Genomic Best Linear Unbiased Predictor (GBLUP) Method

The univariate GBLUP model including the random interaction term between the genomic effect of the *j*th line and the *i*th environment method is represented by the model

$$y_{ij} = E_i + g_j + gE_{ij} + e_{ij} \quad (1)$$

where (i) y_{ij} represents the response of the *j*th line in the *i*th environment (i=1,2,…,I, j=1,2,…J); (ii) E_i denotes the fixed effect of the *i*th environment; (iii) g_j represents the random genomic effect of the *j*th line, with $\boldsymbol{g} = \boldsymbol{g} = \left(g_1,\ldots,g_J\right)^T \sim N\left(\boldsymbol{0},\sigma_1^2\,\boldsymbol{G}_g\right)$, σ_1^2 is a genomic variance and $\boldsymbol{G}_g$ is of order $J \times J$ and represents the genomic relationship matrix (GRM); $\boldsymbol{G}_g$ is calculated (Van Raden, 2008) as $G_g = \frac{\boldsymbol{W}\boldsymbol{W}^T}{p}$, where p denotes the number of markers and $\boldsymbol{W}$ is the matrix of markers of order $J \times p$; the $\boldsymbol{G}_g$ matrix is constructed using the observed similarity at the genomic level between lines, rather than the expected similarity based on pedigree (iv) gE_{ij} is the random interaction term between the genomic effect of the *j*th line and the *i*th environment with $\boldsymbol{gE} = \left(gE_{11},\ldots,gE_{IJ}\right)^T \sim N\left(\boldsymbol{0},\sigma_2^2\ \mathrm{I}_I \otimes \boldsymbol{G}\right)$, where σ_2^2 is an interaction variance; and (v) e_{ij} is a random residual associated with the *j*th line in the *i*th environment distributed as $N\left(0,\sigma^2\right)$, where σ^2 is the residual variance.

Multi-Trait Deep Learning Method

There are many network topologies in deep learning; however, in this application we used densely connected networks, also known as feed-forward networks (see **Figure 1**). This network topology is

TABLE 1 | Acronyms of the 43 environments and the notations used to represent them. Field trials were conducted in nine countries under rainfed (r) and/or irrigated (i) conditions in 11 years (2003-04-05-06-07-08-09-12-13-14-15).

Environment*	Notation	Environment	Notation	Environment	Notation
Hng-i12	E1	Itl5-r15	E16	Mxc-r14	E31
Hng-i13	E2	Lbn-i04	E17	Spn1-r04	E32
Hng-r12	E3	Lbn-i05	E18	Spn2-r05	E33
Hng-r13	E4	Lbn-r04	E19	Syr-i05	E34
Itl-r06	E5	Lbn-r05	E20	Syr-i06	E35
Itl1-r03	E6	Mrc-i04	E21	Syr-i07	E36
Itl1-r04	E7	Mrc-r04	E22	Syr-r05	E37
Itl1-r12	E8	Mxc-i14	E23	Syr-r07	E38
Itl1-r13	E9	Mxc-i06	E24	Syr2-r06	E39
Itl2-r04	E10	Mxc-i07	E25	Tns-i05	E40
Itl2-r05	E11	Mxc-i14	E26	Tns-r05	E41
Itl3-r08	E12	Mxc-i06	E27	Trk-i12	E42
Itl4-r07	E13	Mxc-n07	E28	Trk-r12	E43
Itl5-n15	E14	Mxc-r06	E29	–	–
Itl5-r09	E15	Mxc-r07	E30	–	–

**Hung, Hungary; Itl, Italy; Lbn, Lebanon; Mrc, Morocco; Mxc, Mexico; Spn, Spain; Syr, Syria; Tns, Tunisia; Trk, Turkey.*

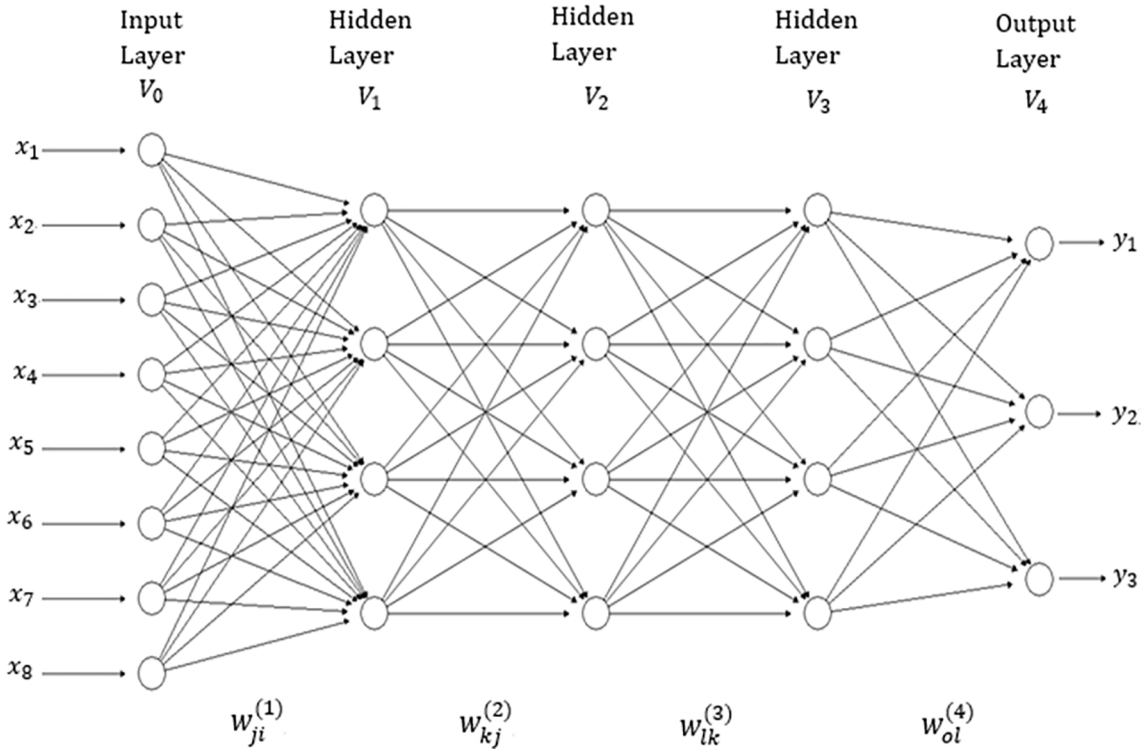


FIGURE 1 | A feedforward deep neural network with one input layer, three hidden layers and one output layer. There are eight neurons in the input layer that corresponds to the input information, four neurons in each of three hidden layers, while there are three neurons in the output layer that correspond to the number of traits to be predicted.

a typical feedforward neural network (where there are no feedback interconnections) and a specific structure is not assumed in the input features (Goodfellow et al., 2016). This network consists of an input layer, an output layer and multiple hidden layers between the input and output layers. The number of features correspond to the input layer neurons (units). Hidden layer neurons perform non-linear transformation on the original input attributes (Lewis, 2016). The number of output neurons depends on the number of response variables to be predicted for continuous response variables (traits in plant breeding) which receive as input the output of hidden neurons, and produce as output the prediction values of interest (Goodfellow et al., 2016). The connected neurons form the network and the strength of the weights of each connection controls the contribution of each neuron. We implemented a type of regularization called "dropout", which consists of temporarily removing (setting to zero) a random subset (%) of neurons and their connections during training.

The network in **Figure 1** has four layers (V_0,V_1,V_2,V_3 and V_4); V_0 represents the input layer (which is not counted), V_1 to V_3 are the hidden layers and V_4 denotes the output layer. This means that the "depth" of the network is four. The size of the network given in **Figure 1** is $|V| = \left|\bigcup_{t=0}^{T}\right| = |9+5+5+5+3| = 27$. In each layer, a +1 was added to the observed unit to represent the lacking node of the bias (or intercept). The width of the network given in **Figure 1** is $max_t|V_t|$=9. It is important to point out that the networks implemented in this research are similar to the network in **Figure 1**, but with many more input and hidden neurons. In addition to implementing networks with three hidden layers, we also implemented networks with one and two hidden layers.

The analytical forms of the model given in **Figure 1** for o output, with d inputs, M_1 hidden neurons (units) in hidden layer 1, M_2 hidden units in hidden layer 2, M_3 hidden units in hidden layer 3, and O output neurons, are given by the following Equations (1)–(4):

$$V_{1j} = g_1\left(\sum_{i=1}^{d} w_{ji}^{(1)} x_i\right) \text{for } j = 1,\ldots, M_1 \quad (1)$$

$$V_{2k} = g_2\left(\sum_{j=1}^{M_1} w_{kj}^{(2)} V_{1j}\right) \text{for } k = 1,\ldots, M_2 \quad (2)$$

$$V_{3l} = g_3\left(\sum_{k=1}^{M_2} w_{lk}^{(3)} V_{2k}\right) \text{for } l = 1,\ldots, M_3 \quad (3)$$

$$y_o = g_4\left(\sum_{l=1}^{M_3} w_{ol}^{(4)} V_{3l}\right) \text{for } o = 1,\ldots, O \quad (4)$$

where Equation (1) produces the output of each of the neurons in the first hidden layer, Equation (2) produces the output of each of the neurons in the second hidden layer, Equation (3) produces the output of each of the neurons in the third hidden layer and, finally, Equation (4) produces the output of each response variable of interest. The learning process is carried out with the weights ($w_{ji}^{(1)}, w_{kj}^{(2)}, w_{lk}^{(3)}$ and $w_{ol}^{(4)}$) that correspond to the input, to the first hidden layer, second hidden layer, third hidden layer and the output hidden layer, respectively. All the activation functions used (g_1, g_2, g_3 and g_4) in this paper were the rectified linear activation unit (RELU) function, since the response variables we wish to predict are all continuous. In Equations (1) to (4), the intercepts or bias terms were ignored.

The input variables of the multi-trait deep learning model (MTDL) were the result of concatenating the information of (i) environments, (ii) markers through the Cholesky decomposition of the genomic relationship matrix, and (iii) genotype × environment interaction ($G\times E$). This meant that first we built the design matrices of environments ($\boldsymbol{Z}_E$), genotypes ($\boldsymbol{Z}_G$) and $G\times E$ ($\boldsymbol{Z}_{GE}$), followed by the Cholesky decomposition of the genomic relationship matrix ($\boldsymbol{G}$). After that, the design matrix of genotypes was post-multiplied by the transpose of the upper triangular factor of the Cholesky decomposition ($\boldsymbol{Q}^T$), $\boldsymbol{Z}_G^* = \boldsymbol{Z}_G\boldsymbol{Q}^T$, followed by the calculation of the $G\times E$ term as the product of the design matrix of the $G\times E$ term post-multiplied by the Kronecker product of the identity matrix of order equal to the number of environments and $\boldsymbol{Q}^T$, that is, $\boldsymbol{Z}_{GE}^* = \boldsymbol{Z}_{GE}(\boldsymbol{I}_I \otimes \boldsymbol{Q}^T)$. Finally, the matrix with input covariates used for implementing both deep learning models was equal to $\boldsymbol{X} = \left[\boldsymbol{Z}_E, \boldsymbol{Z}_G^*, \boldsymbol{Z}_{GE}^*\right]$.

As discussed by Montesinos-López et al. (2018a), appropriate selection of hyper-parameters is fundamental for successfully implementing deep learning models. For this reason, we used the grid search method for tuning the required number of neurons and epochs. That is, we discretized these hyper-parameters into a desired set of values of interest, and the models were trained and evaluated for all combinations of these values (that is, a "grid"); from there we selected the best combination of each hyper-parameter. The values of units used for the grid search were 20 to 200 with increments of 20, for the number of epochs we used 200 with increments of 1, while for layers we used 1, 2 and 3. This meant that the grid search consisted of 6,000 combinations of units, epochs and layers. For the remaining hyper-parameters, we chose their values according to a literature review. The percentage of dropouts used in our application was 30% (Srivastava et al., 2014; Chollet and Allaire, 2017).

Univariate Deep Learning

It is important to point out that a univariate deep learning (UDL) version of the multi-trait deep learning model described above was implemented, but with a feedforward neural network with only one neuron in the output layer, which meant that three independent UDL models were implemented for the three traits under study. For comparison purposes, the three models (GBLUP, MTDL and UDL) were implemented with (I) and without (WI), the interaction term. The predictor was composed of the main effects due to environments, lines and the genotype × environment (G×E) interaction term, while in the last scenario the G×E was ignored.

Evaluating Prediction Performance With Cross-Validation

For evaluating the prediction accuracy of the Bologna data set under the three models (GBLUP, MTDL and UDL), we implemented cross-validation. The implemented random cross-validation is denoted as CV1 and consists of dividing the whole data set into a training (TRN) and a testing (TST) set. The percentages of the whole data set assigned to the TRN and TST sets were 80 and 20%, respectively. In this cross-validation, some individuals can never be part of the training set. Our random CV1 used sampling with replacement, which means that one observation can appear in more than one partition. The design we implemented mimics a prediction problem faced by breeders in incomplete field trials where lines are evaluated in some, but not all, target environments. More explicitly, TRN–TST partitions were obtained as follows: since the total number of records per trait available for the data set with multi-environments is $N=J\times I$, to select lines in the TST data set, we fixed the percentage of data to be used for TST (PTesting = 20%). Then we chose 0.20×N (lines) at random, and subsequently we randomly picked one environment per line from I environments. The resulting cells (*ij)* were assigned to the TST data set, while cells not selected through this algorithm were allocated to the TRN data set. Lines were sampled without replacement if $J \geq 0.20\times$N, and with replacement otherwise (López-Cruz et al., 2015). For each data set under CV1, five random partitions were implemented, and with the observed and predicted values of each testing data set, we calculated the mean arctangent absolute percentage error (MAAPE) as a measure of prediction accuracy (Kim and Kim, 2016) (the smallest MAAPE indicates the best genome-based prediction model). The MAAPE is defined as the arctan of the absolute value of the difference between the observed value minus the predicted value divided by the observed value. Its main advantage is that it is defined in radians and therefore scale-free with the acceptance of missing values, and that it approaches Pi over 2 when dividing by zero.

All the analyses done were implemented in the R statistical software (R Core Team, 2019) using the keras library for implementing the DL method (Chollet and Allaire, 2017) and the BGLR library for implementing the GBLUP model (de los Campos and Pérez-Rodríguez, 2014).

RESULTS

The results are given in four sections: (1) prediction performance for DH, (2) prediction performance for GY, (3) prediction performance for PH and (4) prediction performance across environments for the three traits. Data on the genome-based predictive values using the MAAPE criterion are displayed in **Figures 2**–**5**. The same results used to construct these figures are included in the Supplemental Material at the following link: http://hdl.handle.net/11529/10548262.

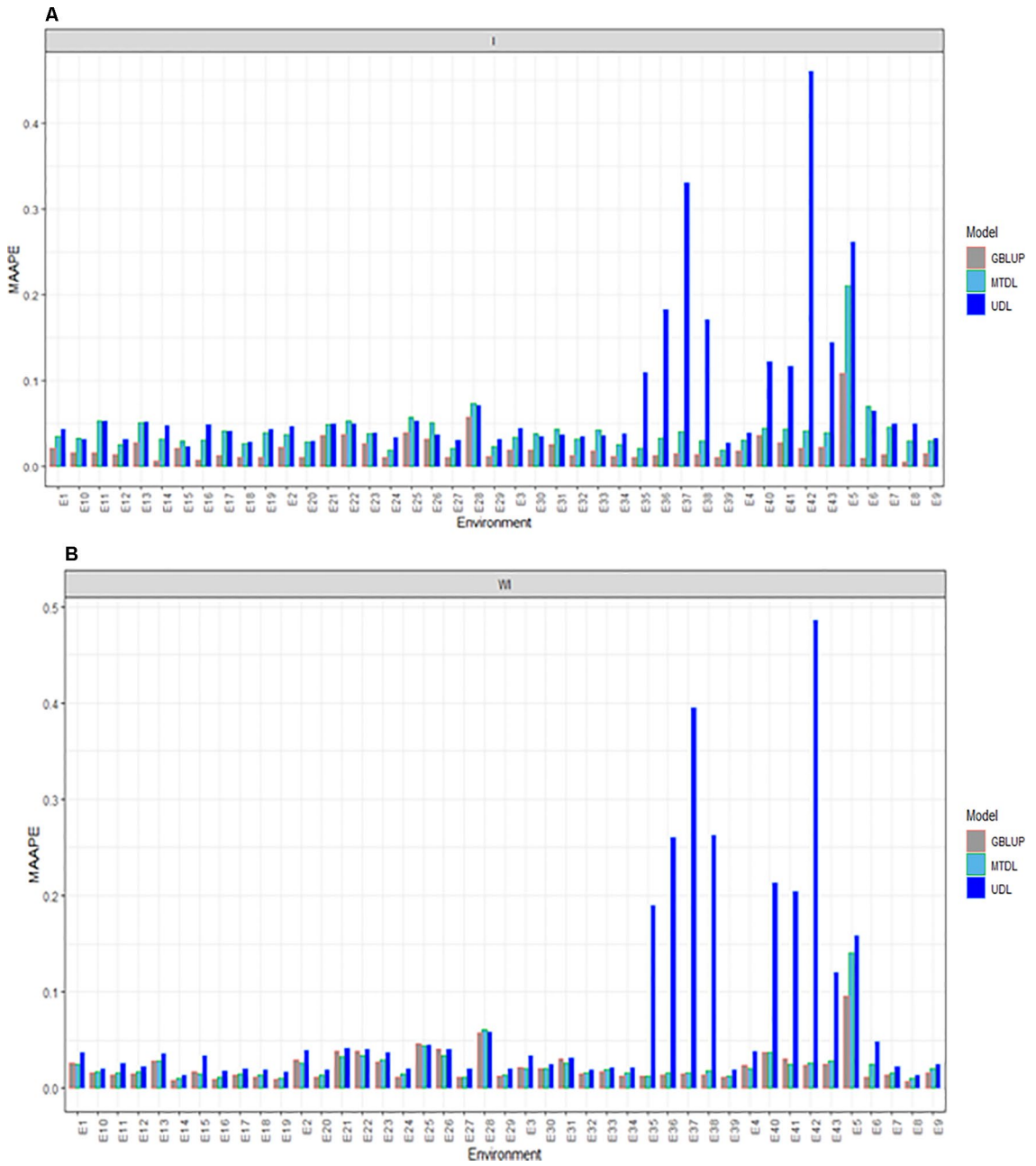


FIGURE 2 | Prediction accuracy of GBLUP, MTDL and UDL in terms of MAAPE for DH in 43 environments (E1–E43) **(A)** including genotype × environment interaction (I), and **(B)** without genotype × environment (WI).

Prediction Performance for DH

Here we report the prediction accuracy in terms of MAAPE obtained by implementing the random cross-validation with 80% TRN and 20% TST for each of the three methods, GBLUP, MTDL and UDL. Each method was implemented with (I) and without (WI) interaction. With interaction (I) under MAAPE, the predictions of the three methods ranged between 0.00550 and 0.4602, with a mean and median equal to 0.0469 and 0.0334, respectively. **Figure 2A** shows that in all cases the GBLUP model was better than the other two deep learning models; the MTDL model was the second best model, while the worst predictions were those of the UDL model.

Without (WI) the interaction term, we found that the range of MAAPE was between 0.0066 and 0.4856, with a mean and median equal to 0.0409 and 0.0211, respectively. **Figure 2B** indicates that the GBLUP model was also better than the deep learning methods (MTDL and UDL), but in 14 out of 43 environments, the MTDL outperformed the GBLUP model. It is important to point out that with and without genotype ×environment interaction, the worst

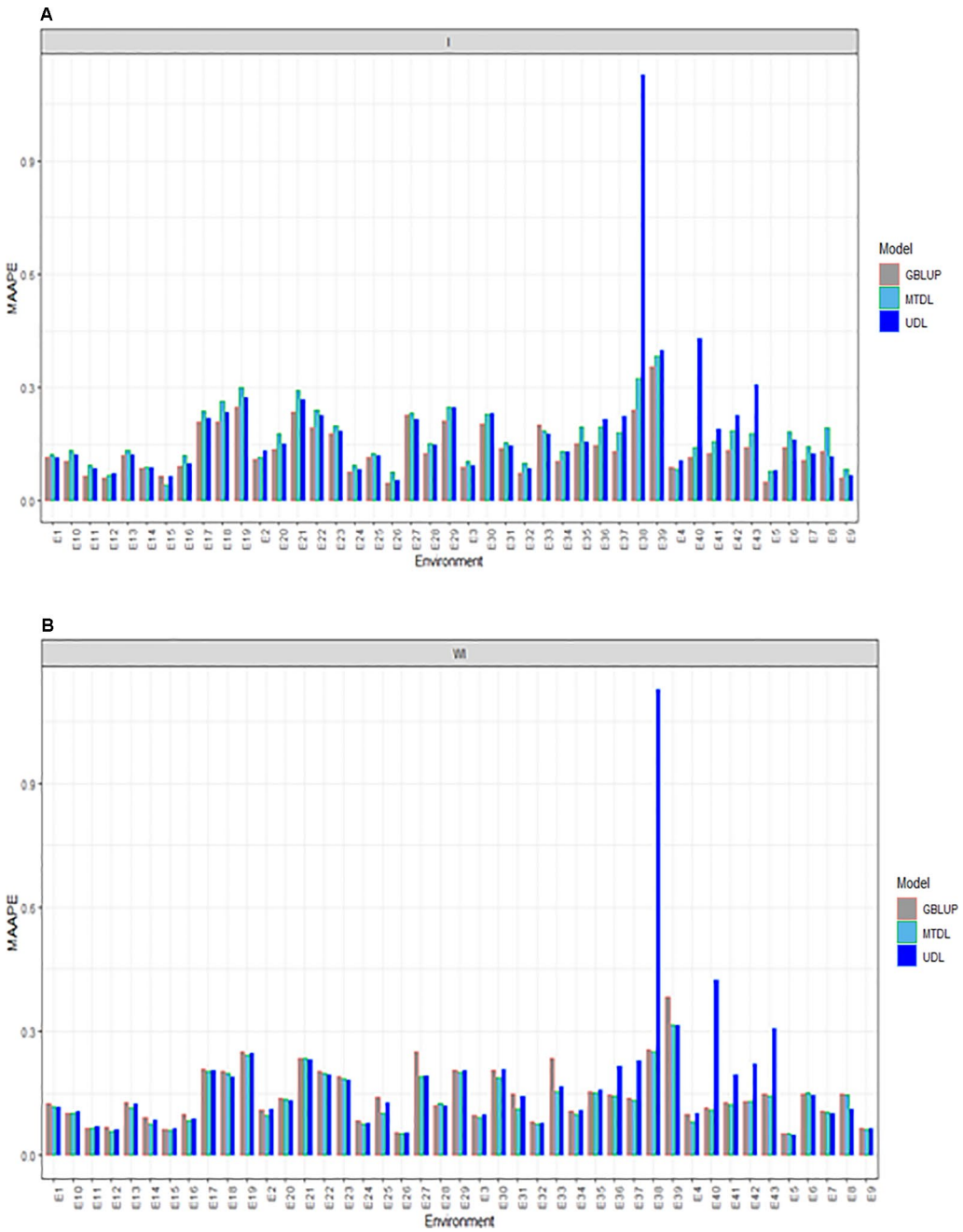


FIGURE 3 | Prediction accuracy of GBLUP, MTDL and UDL in terms of MAAPE for GY in 43 environments (E1–E43) **(A)** including genotype × environment interaction (I), and **(B)** without genotype × environment (WI).

predictions of the three methods were found under the UDL model in environments E35–38, E5–E6 and E40–E43.

Prediction Performance for GY

For GY, we provide the prediction performance under three methods (GBLUP, MTDL and UDL) taking into account the interaction term (I) and ignoring it (WI). Under I (interaction term) we found that the predictions in terms of MAAPE ranged between 0.0418 and 1.127, with a mean and median of 0.1652 and 0.1397, respectively. Also for GY, the best predictions were observed under the GBLUP model and the worst under the UDL model, while the predictions of the MTDL model were quite similar to those of the GBLUP model. However, in all environments, the best predictions were observed under the GBLUP model (**Figure 3A**). When the interaction term was ignored (WI), GY predictions

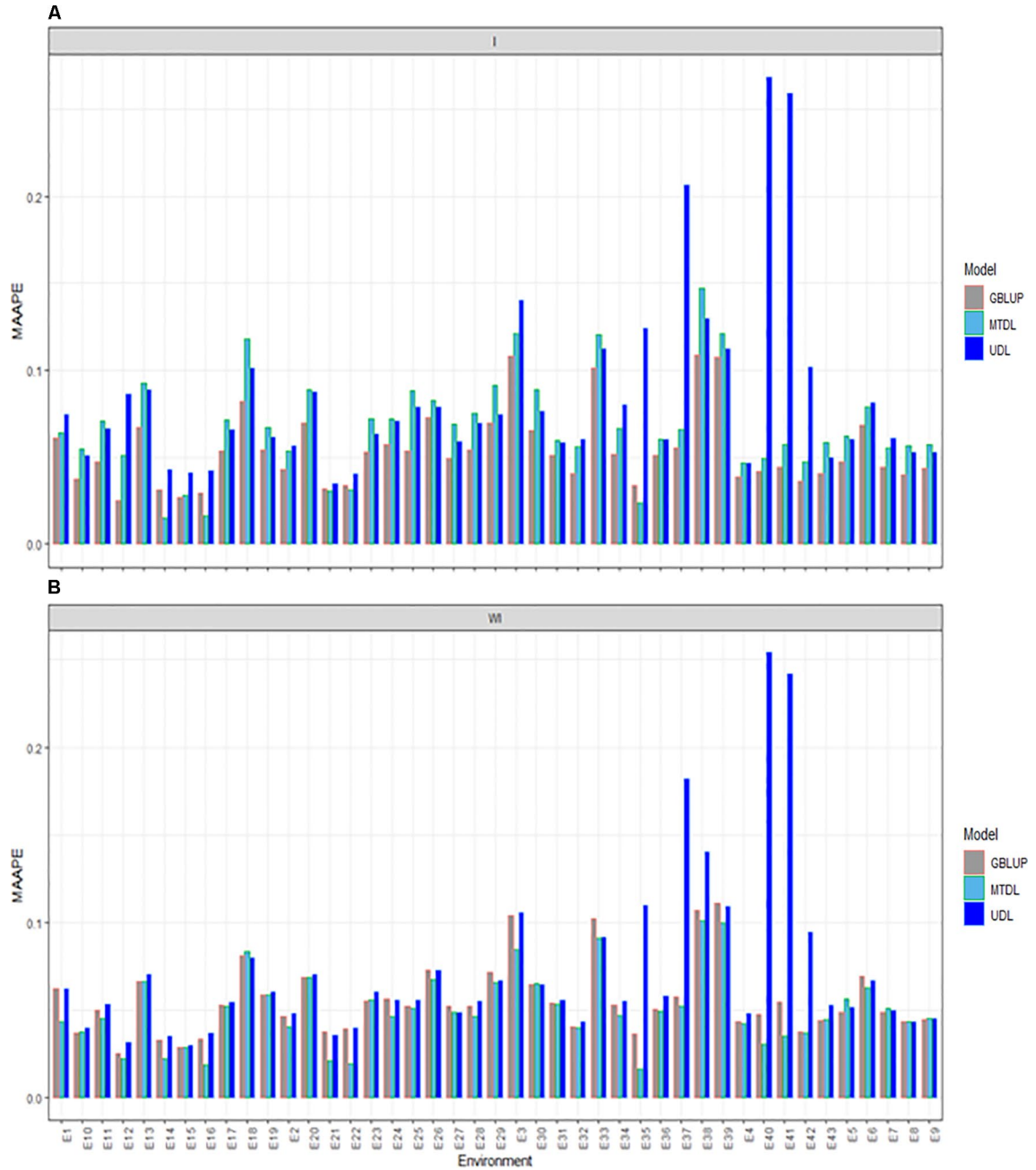


FIGURE 4 | Prediction accuracy of GBLUP, MTDL and UDL in terms of MAAPE for PH in 43 environments (E1–E43) **(A)** including genotype × environment interaction (I), and **(B)** without genotype × environment (WI).

under the MAAPE metric ranged between 0.0512 and 1.1269, with a mean and median equal to 0.1522 and 0.1287, respectively (**Figure 3B**). However, now the best predictions were observed under the MTDL model, since in 40 out of 43 environments, it outperformed the GBLUP model, and in 32 out of 43 environments, it outperformed the UDL model. It is important to point out that the worst predictions of all methods with and without the interaction term were found under the UDL model in environments E38 and E40–E43.

Prediction Performance for PH

The three methods under study (GBLUP, MTDL and UDL) were compared for PH with (I) and without (WI) the interaction term. With the interaction term, we found that the predictions under the MAAPE metric ranged between 0.0148 and 0.2685, with a mean and median equal to 0.0687 and 0.0604, respectively. The best predictions were observed under the GBLUP model, since it was superior to the MTDL model in 38 out of the 43 environments, and in all 43 environments was better than

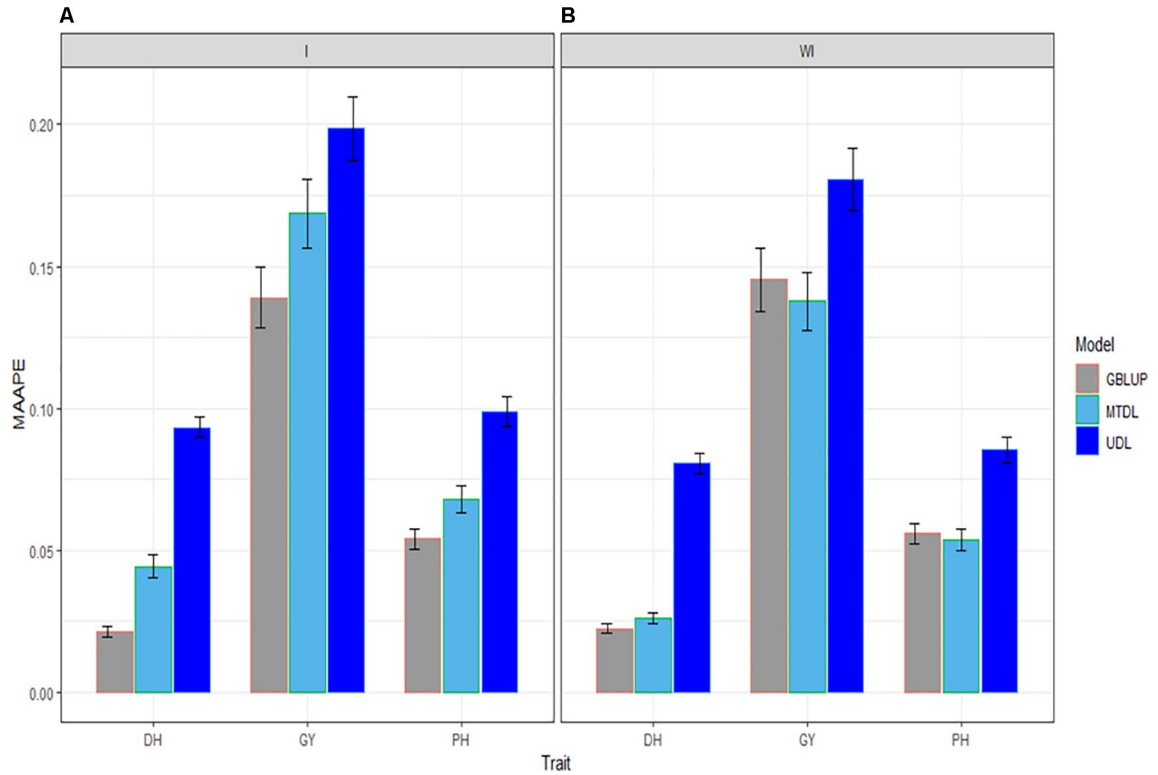


FIGURE 5 | Prediction accuracy of GBLUP, MTDL and UDL in terms of average MAAPE for traits DH, GY and PH across 43 environments (E1–E43) **(A)** including genotype × environment interaction (I), and **(B)** without genotype × environment (WI).

the UDL model (**Figure 4A**). When the interaction term was ignored (WI), we found that the MAAPE metric ranged between 0.0165 and 0.2541, with the mean and median equal to 0.0598 and 0.0526, respectively (**Figure 4B**). However, now the best predictions were observed under the MTDL model in 33 out of 43 environments compared to the GBLUP model, and in 38 out 43 environments compared to the UDL model. Finally, it is important to point out that in general the worst predictions were under the UDL model, since under this method the worst predictions without genotype by environment interaction were observed in environments E35, E37–E38 and E40–42, while with the interaction term, these environments were the worst, except environment E38.

Prediction Performance Across Environments for the Three Traits

In this subsection we provide a summary of the prediction performance of the three traits across environments using GBLUP, MTDL and UDL. When the interaction term (I) was taken into account, in **Figure 5** it is clear that the best predictions were observed under the GBLUP model and the worst under the UDL model, and the second best under the MTDL method. However, when the interaction term was ignored, the best predictions were observed under the GBLUP method and MTDL model and the worst under the UDL model; non-relevant differences can be observed in the predictions between the GBLUP and MTDL. Finally, the best predictions for DH were slightly better under the GBLUP method with the interaction term compared to the MTDL method (without the interaction term), while for traits GY and PH, there were no differences between the predictions of the GBLUP (with the interaction term) and the MTDL (without the interaction term) models.

DISCUSSION

Genomic selection is not new in the context of durum wheat breeding, and interest in applying this metodology for developing new varieties of durum wheat continues (Michel et al., 2019; Steiner et al., 2019). For this reason, in this study we applied GS to durum wheat with the purpose of studying the usefulness of this methodology for choosing candidate genotypes for early selection. The prediction performance of GS depends on many factors such as: (1) the genetic architecture of the target traits, (2) the number of traits under study, (3) the statistical or machine learning models used, (4) the quality of the marker information, (5) the strength of the relationship between individuals and (6) the relationship between the reference population and the validation population, among others. For this reason, in this study we evaluated the prediction performance of 270 durum wheat lines in 43 environments for traits DH, GY and PH.

Among the GWAS of durum wheat, only a few have investigated genomic predictions for grain yield and quality traits

(Asbati et al., 2000; Patil et al., 2012; Crossa et al., 2017; Fiedler et al., 2017; Sukumaran et al., 2018a; Sukumaran et al., 2018b; Johnson et al., 2019; Merida-Garcıa et al., 2019). Maccaferri et al. (2011) used SSR markers and the data of a panel of 189 elite cultivars evaluated in 15 field trials in different Mediterranean countries to dissect the genetic basis of durum wheat plasticity across a broad range of water availability during the crop cycle (from 146 to 711 mm). More recently, the GWAS carried out by Sukumaran et al. (2018a; 2018b) used DArTseq markers to evaluate CIMMYT durum wheats grown under three different conditions (yield potential, drought and heat stresses). Among these studies, those of Fiedler et al. (2017), Crossa et al. (2017) and Sukumaran et al. (2018b) have shown promising genome-enabled prediction accuracy for grain yield and other traits. Notably, Crossa et al. (2017) concluded that genomic selection models incorporating marker × environment interaction are useful in durum wheat breeding for increasing genetic gains in rapid cycle selection. The present study further supports the effectiveness of GS in durum wheat.

The results obtained with GBLUP, MTDL and UDL show that the best predictions were found under the GBLUP model, followed by the MTDL model, while the worst predictions were observed under the UDL model. These results are in agreement with those reported by Montesinos-López et al. (2018a), Montesinos-López et al. (2018b) and Montesinos-López et al. (2019a), Montesinos-López et al. (2019b) who compared GBLUP models against deep learning models. Our results are also in agreement with those of Bellot et al. (2018), who found that conventional genomic prediction models outperformed deep learning methods.

Collectively, our results support the feasibility of applying genomic selection as a cost-effective means to enhance genetic gain for complex and expensive-to-measure traits in durum wheat, in agreement with Fiedler et al. (2017) and Haile et al. (2018). However, unlike the results of Haile et al. (2018), the best predictions in our study were found under a univariate model (GBLUP), while they found that multi-trait models were more accurate than the univariate trait models only for grain yield. Although the low number of lines in this study might limit the performance of deep learning models (UDL and MTDL), results show that the performance of DL models is competitive and can be applied for GS in durum wheat to develop faster and more efficient genetic gains in breeding programs.

The low prediction accuracies of deep learning methods can be attributed to the fact that deep learning methods outperform other techniques in the context of large or very large data sets, but with smaller data sets, traditional statistical and machine learning algorithms are preferable. However, although the durum wheat data set used here is not large, the prediction performance of the MTDL models was very competitive compared to the GBLUP model, and most of the time outperformed the UDL model. This result can be attributed to the fact that the training process under the MTDL model is more efficient because it simultaneously uses the information of the three traits under study, which is equivalent to increasing the sample size (Montesinos-López et al., 2018b; Montesinos-López et al., 2019b). There is some evidence that deep learning methods are better than conventional statistical and machine learning methods when tackling very complex problems such as natural language processing, image classification and speech recognition. However, there is also enough evidence that DL methods are worse in linearly separable problems and in the context of small data sets.

Although the universal approximation theorem states that "a feedforward network with a single hidden layer containing a finite number of neurons can approximate any continuous function to any degree of precision" (Cybenko, 1989), in many applications it has been shown that artificial neural networks and deep learning models (with more than one hidden layer) are not very efficient in terms of prediction performance due to the fact that the required number of hidden neurons is so incredibly large that it is not possible to implement them; for this reason, these models fail to learn and generalize correctly.

In this study, we evaluated the prediction performance of durum wheat and the results obtained show that even when the number of lines under study is low, it is possible to implement the GS methodology for selecting candidate genotypes early in time. However, the results obtained depend on the type of statistical model used, since we observed that the best predictions were under the GBLUP model, followed by the MTDL model, while the worst predictions were under the UDL model. Notably, the predictions of the MTDL model were very similar to those of the GBLUP model. From all the above, it is clear that better prediction statistical machine learning models are still required for improving the efficacy of GS methodology; for this reason, other exercises for benchmarking the existing models with real data are needed and of course new statistical or machine learning models are welcome.

CONCLUSIONS

In this study, univariate and multivariate deep learning methods were applied to evaluate the prediction performance of durum wheat. These predictions were also compared to the predictions of the GBLUP method. In general, the best predictions were observed under the GBLUP model, although the predictions of the multi-trait deep learning model were very close to those of the GBLUP model, while the univariate deep learning model provided the worst predictions. These results can be attributed in part to the fact that the durum wheat data set used is small, which did not help the performance of deep learning methods, an approach better suited to huge amounts of training data. This notwithstanding, the multi-trait deep learning model produced predictions that were close to those of the GBLUP model. However, due to the number of hyper-parameters that need to be tuned in deep learning models, their implementation remains very challenging, because the DL process is still a combination of art and science. Even with these restrictions, we are confident that deep learning has a lot of potential to be successfully applied in the context of genomic selection.

DATA AVAILABILITY STATEMENT

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

OAML and AML performed the statistical analyses of the data and wrote the first version of the article; RT, MM, GS and KA produced the data analyzed in this study; JC proposed the initial idea and provided and participated in writing the manuscript.

FUNDING

We are thankful for the financial support provided by the Bill & Melinda Gates Foundation (for maize and wheat breeding programs), the CIMMYT CGIAR CRP (maize and wheat), as well the USAID projects. We acknowledge the financial support provided by the Foundation for Research Levy on Agricultural Products (FFL) and the Agricultural Agreement Research Fund (JA) in Norway through NFR grant 267806. We are equally thankful to the European project FP7-244374 (DROPS) and the Ager Agroalimentare e Ricerca—Project 'From Seed to Pasta: Multidisciplinary approaches for a more sustainable and high quality durum wheat production' for the financial support to implement the field trials used in this study.

REFERENCES

Alipanahi, B., Delong, A., Weirauch, M. T., and Frey, B. J. (2015). Predicting the sequence specificities of DNA- and RNA-binding proteins by deep learning. *Nat. Biotechnol.* 33, 831–838. doi: 10.1038/nbt.3300

Asbati, A., Elouafi, I., Elsaleh, A., Mather, D. E., and Nachit, M. (2000). "QTL-mapping of genomic regions controlling gluten strength in durum (*Triticum turgidum* L. var. *durum*)," in *Durum wheat improvement in the Mediterranean region: new challenges*, vol. 40. Eds. C. Royo, M. Nachit, N. Di Fonzo, and J. L. Araus (Zaragoza: CIHEAM (Options Méditerranéennes: Série A Séminaires Méditerranéens), 505–509.

Bellot, P., de los Campos, G., and Pérez-Enciso, M. (2018). Can deep learning improve genomic prediction of complex human traits? *Genetics* 210 (3), 809–819. doi: 10.1534/genetics.118.301298

Chollet, F., and Allaire, J. J. (2017). *Deep learning with R*. 1st ed. (New Delhi, India: Manning Publications, Manning Early Access Program (MEA)).

Cole, J. H., Rudra, P. K., Poudel, D. T., Matthan, W. A., Caan, C. S., Tim, D., et al. (2017). Predicting brain age with deep learning from raw imaging data results in a reliable and heritable biomarker. *NeuroImage* 163 (1), 115–124. doi: 10.1016/j.neuroimage.2017.07.059

Covarrubias-Pazaran, G., Schlautman, B., Diaz-Garcia, L., Grygleski, E., Polashock, J., Johnson-Cicalese, J., et al. (2018). Multivariate GBLUP improves accuracy of genomic selection for yield and fruit weight in biparental populations of Vaccinium macrocarpon Ait. *Front. Plant Sci.* 9, 1310. doi: 10.3389/fpls.2018.01310

Crossa, J., de los Campos, G., Maccaferri, M., Tuberosa, R., Burgueño, J., and Pérez-Rodríguez, P. (2016). Extending the marker × environment interaction model for genomic-enabled prediction and genome-wide association analysis in durum wheat. *Crop Sci.* 5 (56), 2193.2209. doi: 10.2135/cropsci2015.04.0260

Crossa, J., Pérez-Rodríguez, P., Cuevas, J., Montesinos-López, O. A., Jarquín, D., de Los Campos, G., et al. (2017). Genomic selection in plant breeding: methods, models, and perspectives. *Trends Plant Sci.* 22 (11), 961–975. doi: 10.1016/j.tplants.2017.08.011

Cybenko, G. (1989). Approximations by superpositions of sigmoidal functions. *Math. Control Signals Syst.* 2 (4), 303–314. doi: 10.1007/BF02551274

de los Campos, G., and Pérez-Rodríguez, P. (2014). "Bayesian generalized linear regression," in *R package version 1.0.4*. http://CRAN.R-project.org/package=BGLR.

Fiedler, J. D., Salsman, E., Liu, Y., Michalak de Jiménez, M., Hegstad, J. B., Chen, B., et al. (2017). Genome-wide association and prediction of grain and semolina quality traits in durum wheat breeding populations. *Plant Genome* 10 (3), 1–12. doi: 10.3835/plantgenome2017.05.0038

Goodfellow, I., Bengio, Y., and Courville, A. (2016). *Deep learning*. Cambridge, Massachusetts, USA: MIT Press.

Haile, J. K., N'Diaye, A., Clarke, F., Clarke, J., Knox, R., Rutkoski, J., et al. (2018). Genomic selection for grain yield and quality traits in durum wheat. *Mol. Breeding* 38, 75. doi: 10.1007/s11032-018-0818-x

Jia, Y., and Jannink, J.-L. (2012). Multiple-trait genomic selection methods increase genetic value prediction accuracy. *Genetics* 192 (4), 1513–1522. doi: 10.1534/genetics.112.144246

Jiang, J., Zhang, Q., Ma, L., Li, J., Wang, Z., and Liu, J. F. (2015). Joint prediction of multiple quantitative traits using a Bayesian multivariate antedependence model. *Heredity* 115 (1), 29–36. doi: 10.1038/hdy.2015.9

Johnson, M., Kumar, A., Oladzad-Abbasabadi, A., Salsman, E., Aoun, M., Manthey, F. A., et al. (2019). Association mapping for 24 traits related to protein content, gluten strength, color, cooking, and milling quality using balanced and unbalanced data in durum wheat [*Triticum turgidum* L. var. *durum* (Desf).]. *Front. Genet.* 10, 717. doi: 10.3389/fgene.2019.00717

Kim, H., and Kim, H. (2016). A new metric of absolute percentage error for intermittent demand forecasts. *Int. J. Forecast.* 32 (3), 669–679. doi: 10.1016/j.ijforecast.2015.12.003

Lewis, N. D. (2016). "Deep learning made easy with R," in *A gentle introduction for data science* (CreateSpace Independent Publishing Platform).

López-Cruz, M., Crossa, J., Bonnett, D., Dreisigacker, S., Poland, J., Jannink, J.-L., et al. (2015). Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection method. *G3: Genes Genomes Genet.* 5 (4), 569–582. doi: 10.1534/g3.114.016097

Maccaferri, M., Sanguineti, M. C., Natoli, V., Ortega, J. L. A., Salem, M. B., Bort, J., et al. (2006). A panel of elite accessions of durum wheat (*Triticum durum* Desf.) suitable for association mapping studies. *Plant Genetic Resour.* 4 (1), 79–85. doi: 10.1079/PGR2006117

Maccaferri, M., Sanguineti, M. C., Demontis, A., El-Ahmed, A., Garcia del Moral, L., Maalouf, F., et al. (2011). Association mapping in durum wheat grown across a broad range of water regimes. *J. Exp. Bot.* 62 (2), 409–438. doi: 10.1093/jxb/erq287

Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., et al. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 13 (5), 648–663. doi: 10.1111/pbi.12288

Mantovani, P., Maccaferri, M., Sanguineti, M. C., Tuberosa, R., Catizone, I., Wenzl, P., et al. (2008). Integrated DArT-SSR linkage map of durum wheat. *Mol. Breeding* 22 (4), 629–648. doi: 10.1007/s11032-008-9205-3

Merida-Garcıa, R., Liu, G., He, S., Gonzalez-Dugo, V., Dorado, G., Galvez, S., et al. (2019). Genetic dissection of agronomic and quality traits based on association mapping and genomic selection approaches in durum wheat grown in Southern Spain. *PLoS One* 14 (2), e0211718. doi: 10.1371/journal.pone.0211718

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829.

Michel, S., Löschenberger, F., Ametz, C., Pachler, B., Sparry, E., and Bürstmayr, H. (2019). Combining grain yield, protein content and protein quality by multi-trait genomic selection in bread wheat. *Theor. Appl. Genet.* 132 (6), 1745–1760. doi: 10.1007/s00122-019-03312-5

Montesinos-Lopez, O. A., Montesinos-Lopez, A., Crossa, J., Toledo, F. H., Perez- Hernandez, O., Eskridge, K. M., et al. (2016). A genomic Bayesian multi-trait and multi-environment model. *G3: Genes Genomes Genet.* 6 (9), 2725–2744. doi: 10.1534/g3.116.032359

Montesinos-López, A., Montesinos-López, O. A., Gianola, D., Crossa, J., and Hernández-Suárez, C. M. (2018a). Multi-environment genomic prediction of plant traits using deep learners with a dense architecture. *G3: Genes Genomes Genet.* 8 (12), 3813–3828. doi: 10.1534/g3.118.200740

Montesinos-López, O. A., Montesinos-López, A., Gianola, D., Crossa, J., Hernández-Suárez, C. M., and Javier Martín-Vallejo (2018b). Multi-trait, multi-environment deep learning modeling for genomic-enabled prediction of plant. *G3: Genes, Genomes, Genetics* 8 (12), 3829–3840. doi: 10.1534/g3.118.200728

Montesinos-López, A., Montesinos-López, O. A., Gianola, D., Crossa, J., and Hernández-Suárez, C. M. (2019a). Multivariate Bayesian analysis of on-farm

trials with multiple-trait and multiple-environment data. *Agron. J.* 3 (1), 1–12. doi: 10.2134/agronj2018.06.0362

Montesinos-López, O. A., Martín-Vallejo, J., Crossa, J., Gianola, D., Hernández-Suárez, C. M., Montesinos-López, A., et al. (2019b). New deep learning genomic prediction model for multi-traits with mixed binary, ordinal, and continuous phenotypes. *G3: Genes Genomes Genet.* 9 (5), 1545–1556. doi: 10.1534/g3.119.300585

Pan, X., and Shen, H.-B. (2017). RNA-protein binding motifs mining with a new hybrid deep learning based cross-domain knowledge integration approach. *BMC Bioinform.* 18 (136), 1–14. doi: 10.1186/s12859-017-1561-8

Patil, R. M., Tamhankar, S. A., Oak, M. D., Raut, A. L., Honrao, B. K., Rao, V. S., et al. (2012). Mapping of QTL for agronomic traits and kernel characters in durum wheat (*Triticum durum* Desf.). *Euphytica* 190 (1), 117–129. doi: 10.1007/s10681-012-0785-y

R Core Team. (2019). *R: A language and environment for statistical computing*. Vienna. Austria. ISBN 3-900051-07-0: R Foundation for Statistical Computing. URL http://www.R-project.org/.

Royo, C., Soriano, J. M., and Alfaro, F. (2017). "Wheat: a crop in the bottom of the Mediterranean diet pyramid," in *Chapter 16 of the Book of Mediterranean identities—environment, society, culture*, 381–399. doi: 10.5772/intechopen.69184

Schulthess, A. W., Zhao, Y., Longin, C. F. H., Reif, J. C. (2018). Advantages and limitations of multiple-trait genomic prediction for Fusarium head blight severity in hybrid wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 131 (3), 685–701. doi: 10.1007/s00122-017-3029-7

Steiner, B., Michel, S., Maccaferri, M., Lemmens, M., Tuberosa, R., and Buerstmayr, H. (2019). Exploring and exploiting the genetic variation of Fusarium head blight resistance for genomic-assisted breeding in the elite durum wheat gene pool. *Theor. Appl. Genet*. 132 (4), 969–988. doi: 10.1007/s00122-018-3253-9

Sukumaran., S., Reynolds, M. P., and Sansaloni, C. (2018a). Genome-wide association analyses identify QTL hotspots for yield and component traits in durum wheat grown under yield potential, drought, and heat stress environments. *Front. Plant Sci.* 9, 81. doi: 10.3389/fpls.2018.00081

Sukumaran, S., Jarquin, D., Crossa, J., and Reynolds, M. (2018b). Genomic-enabled prediction accuracies increased by modeling genotype × environment interaction in durum wheat. *Plant Genome* 11 (2), 1–11. doi: 10.3835/plantgenome2017.12.0112

Tavanaei, A., Anandanadarajah, N., Maida, A. S., and Loganantharaj, R. (2017). A deep learning model for predicting tumor suppressor genes and oncogenes from PDB structure. *bioRiv*, 177378. doi: 10.1109/BIBM.2017.8217722

Tuberosa, R., and Pozniak, C. (2014). Durum wheat genomics comes of age. *Mol. Breeding* 4 (34), 1527–1530. doi: 10.1007/s11032-014-0188-y

Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *J. Mach. Learn. Res.* 15 (6), 1929–1958.

Van Raden, P. M. (2008). Efficient method to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90 000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12 (6), 787–796. doi: 10.1111/pbi.12183

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Sustainable Food Systems

MINI REVIEW
published: 15 November 2019
doi: 10.3389/fsufs.2019.00103

Re-evolution of Durum Wheat by Introducing the *Hardness* and *Glu-D1* Loci

Craig F. Morris[1]*, *Alecia M. Kiszonas*[1], *Jessica Murray*[1], *Jeff Boehm Jr.*[1], *Maria Itria Ibba*[1], *Mingyi Zhang*[2] *and Xiwen Cai*[2]

[1] Western Wheat Quality Laboratory, Agricultural Research Service, United States Department of Agriculture, Pullman, WA, United States, [2] Department of Plant Sciences, North Dakota State University, Fargo, ND, United States

Durum wheat is an important crop worldwide. In many areas, durum wheat appears to have competitive yield, and biotic and abiotic advantages over bread wheat. What limits durum production? In one respect, the comparatively more limited processing and food functionality. Two traits directly relate to these limitations: kernel texture (hardness) and gluten strength. We have addressed both using *ph1b*-mediated translocations from bread wheat. For kernel texture, ca. 28 Mbp of chromosome 5D short arm replaced about 20 Mbp of 5B short arm. Single Kernel Characterization System (SKCS) hardness was reduced from ca. 80 to 20 as the puroindolines were expressed and softened the endosperm. Break flour yields increased from 17 to >40%. Straight-grade flour had low starch damage (2%), and a mean particle size of 75 μm. Crosses with CIMMYT durum lines all produced soft kernel progeny and a high degree of genetic variance for milling and baking quality. Solvent Retention Capacities (SRC) and cookie diameters were similar to soft white hexaploid wheat, showing that soft durum can be considered a "tetraploid soft white spring wheat." Regarding gluten strength, CIMMYT durums contributed a high genetic variance, with the "best" progeny exhibiting Na-dodecylsulfate (SDS) sedimentation volume, SRC Lactic Acid and Mixograph characteristics that were similar to medium-gluten-strength U.S. hard red winter. The best loaf volume among these progeny was 846 cm^3 at ca. 12.8% flour protein. To further address the issue of gluten strength, Soft Svevo was crossed with durum lines possessing Dx2+Dy12 and Dx5+Dy10. Bread baking showed that Dx5+Dy10 was overly strong, whereas Dx2+Dy12 significantly improved bread loaf volume. The best progeny produced a loaf volume of 1,010 cm^3 at 12.1% protein. As a comparison, the long-term in-house regression for loaf volume-flour protein for hard "bread" wheats is 926 cm^3 at 12.1% protein. Obviously, from these results, excellent bread making potential has been achieved.

Keywords: durum wheat, kernel texture, cookie quality, bread baking, gluten strength

OPEN ACCESS

Edited by:
Aldo Ceriotti, Italian National Research Council (CNR), Italy

Reviewed by:
Somnath Mandal, Uttar Banga Krishi Viswavidyalaya, India
Didier Marion, INRA Centre Angers-Nantes Pays de la Loire, France

***Correspondence:**
Craig F. Morris
craig.morris@usda.gov

Specialty section:
This article was submitted to Nutrition and Sustainable Diets, a section of the journal Frontiers in Sustainable Food Systems

Received: *12 April 2019*
Accepted: *23 October 2019*
Published: *15 November 2019*

Citation:
Morris CF, Kiszonas AM, Murray J, Boehm J Jr, Ibba MI, Zhang M and Cai X (2019) Re-evolution of Durum Wheat by Introducing the Hardness and Glu-D1 Loci. Front. Sustain. Food Syst. 3:103. doi: 10.3389/fsufs.2019.00103

INTRODUCTION

High kernel hardness (*texture*) is a defining trait of durum wheat (*Triticum turgidum* subsp. *durum*) grain. Kernel texture dictates many aspects of durum milling and utilization, and in some ways, limits its culinary uses. Durum wheat also lacks the D genome, and thus it also lacks the *Glu-D1* locus for the high molecular weight (HMW) glutenins Dx2+Dy12 and Dx5+Dy10.

Consequently, the elasticity and extensibility of durum doughs are often viewed as inferior to bread wheat (*Triticum aestivum*) (Ammar et al., 2000). The research reviewed here shows how both kernel texture and dough rheology can be manipulated via *ph1b*-mediated homoeologous recombination and the transfer of genetic material from bread wheat to durum wheat.

Endosperm softness in wheat is controlled by the *Puroinoline* genes/proteins, *Pina* and *Pinb*, which reside at the *Hardness* (*Ha*) locus on the distal end of chromosome 5D short arm (5DS) (Morris, 2002; Bhave and Morris, 2008). When both genes are in a functional state, the endosperm is soft, but when either gene is absent or its sequence altered, harder endosperm is observed (Giroux and Morris, 1997, 1998; Morris and Beecher, 2012). When durum wheat formed, both genes from both diploid progenitors (A and S=B sub-genomes) were lost, and thus durum has the hardest kernels of all wheats. Because of the high kernel hardness of durum grain, roller milling does not aim to produce flour, but rather coarse semolina. Attempts to further reduce particle size of semolina result in unacceptably high starch damage and excessive dough water absorption.

Dough strength is a complex interplay between the HMW glutenin subunits, low molecular weight (LMW) glutenins, gliadins, and non-protein endosperm constituents. In bread wheat, the most prominent locus that contributes to dough elasticity and extensibility is *Glu-D1*, with two allelic variants, Dx2+Dy12 and Dx5+Dy10. In general, the Dx5+Dy10 allele is considered the "stronger" allele and is more desirable for bread quality.

HOMOEOLOGOUS RECOMBINATION

Homoeologous recombination in polyploid wheat can be achieved by eliminating the restrictive control of the *Pairing homoeologous-1* (*Ph1*) locus, which restricts pairing to homologous chromosomes and prevents homoeologs from pairing. A line carrying an induced mutation in *Ph1* (*ph1b*) was used in crossing a Langdon durum disomic substitution line carrying the pair of 5D chromosomes from Chinese Spring (Morris et al., 2011, 2015). Subsequently, several recombinant lines were isolated. Interestingly, all carry an identical 28 Mbp of 5DS which replaced 20 Mbp of 5BS (Boehm et al., 2017c; Ibba et al., in press). The specific cross-over occurred in a 39-bp region in the middle of a putative gene. The translocated 5DS fragment carries an entire and intact *Ha* locus with normal expression and endosperm softening.

MILLING AND BAKING PERFORMANCE

Soft kernel durum wheat was found to mill similar to soft white hexaploid wheats. Break flour yields increased from ∼17% (normal durum) to >40% (Murray et al., 2016). Straight-grade flour had low starch damage (2%), and a mean particle size of 75 μm. Ash contents of flours from soft durums were lower than those obtained from hard durum. All crosses with a number of CIMMYT durum lines produced soft kernel progeny and a high degree of genetic variance for milling and baking quality (Boehm et al., 2017a,b). Family mean Single Kernel Characterization System (SKCS) hardness ranged from 5.8 to 23.0. Family mean break flour yields ranged from 38.2 to 42.8%. Ash and starch damage of the straight-grade flours were ∼0.41 and ∼1.5%, respectively. Solvent Retention Capacity (SRC) Water, Na-carbonate, and Sucrose were low and typical of soft white hexaploid wheats. Cookie diameters ranged from 9.16 to 9.48 cm, and were similar to soft white hexaploid wheat. Thus, soft durum can be considered a "tetraploid soft white spring wheat." On a per unit weight of flour produced, soft durum required only from one-fifth to one-third the energy as hard durum (Heinze et al., 2016).

Durum wheat has variable but limited baking quality (Ammar et al., 2000). The very hard kernel texture affects milling, particle size, starch damage, and dough water absorption. Consequently, it is difficult to make direct comparisons between durum and

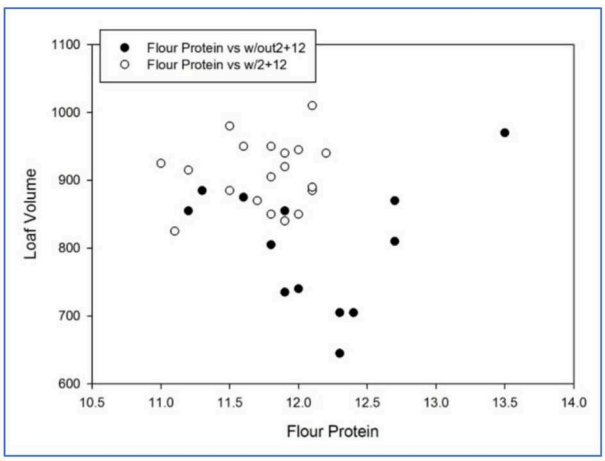


FIGURE 1 | Bread loaf volume of soft durum lines with or without Dx2+Dy12 (flour protein in percent, loaf volume in cm^3).

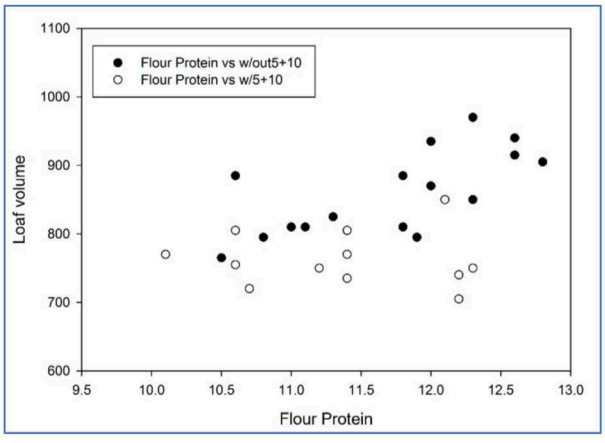


FIGURE 2 | Bread loaf volume of soft durum lines with or without Dx5+Dy10 (flour protein in percent, loaf volume in cm^3).

soft durum. Murray et al. (2017) used near-isogenic lines of Svevo durum to show that the softer kernel was associated with about a 3% decrease in Na-dodecylsulfate (SDS) sedimentation volume, 17% lower SRC Water, 9% lower SRC Lactic acid, and about 10% lower SRC Sucrose. Dough water absorption on the 10-g Mixograph was 5% lower for soft durum flour, and about 10% lower on the 65-g Farinograph. Alveograph parameters were dramatically affected since analyses were performed at constant dough water absorption. Consequently, soft durum flours exhibited lower *W* and *P,* but similar *L*. Since the near-isogenic lines were at similar protein levels, differences were interpreted as being a direct result of "over hydrating" the soft durum doughs. Similarly, with AACCI 100-g "pup" bread loaf testing, optimum water absorption for durum flour was 66.5% and 58% for soft durum flour. Among the CIMMYT progeny, a wide range of SDS sedimentation volume and SRC Lactic Acid was observed. Similarly, bread loaf volumes varied significantly both within, but more so among families. Overall the loaf volume range for individual lines ranged from a very poor 629 cm^3 to a moderate 864 cm^3 at about 12% flour protein.

INTROGRESSION OF *GLU-D1*

More recently, the *Glu-D1* alleles Dx2+Dy12 and Dx5+Dy10 of (Lukaszewski, 2003) were introgressed into the soft kernel durum variety Soft Svevo (**Figures 1**, **2**). Multiple full sibs possessing each glutenin allele were evaluated for milling and pan bread baking.

Those lines with allele Dx2+Dy12 exhibited superior loaf volumes (**Figure 1**). The best progeny line produced a loaf volume of 1,010 cm^3 at 12.1% protein. As a comparison, the long-term in-house regression for loaf volume-flour protein for hard red 'bread" wheats is 926 cm^3 at 12.1% protein. **Figure 2** shows the sibs with or without Dx5+Dy10. Across flour protein contents, those lines without the *Glu-D1* translocation were superior. The lines with Dx5+Dy10 generally lacked extensibility, were termed "bucky" and could not reach full volume potential. This allele actually decreased bread quality.

And as described above, the best line derived from the CIMMYT crosses had a loaf volume of 864 cm^3 at ∼12% protein.

Obviously, from these results, excellent bread making potential can be achieved using Dx2+Dy12 in the Soft Svevo background. Sissons et al. (2019) however found no improvement in bread quality by adding Dx2+Dy12 or Dx5+Dy10 to hard Svevo.

CONCLUSIONS

The soft kernel trait in durum affects nearly every aspect of milling and baking quality. SKCS hardness, break flour yield and flour yield were similar to commercial soft white wheat cultivars. With the exception of dough water absorption, dough strength was essentially unchanged and reflected the inherent gluten properties of the durum background. That said, the introgression of *Glu-D1* alleles dramatically changed dough strength and bread volume, with Dx2+Dy12 showing superiority over Dx5+Dy10. With the caveat of dough water absorption, soft kernel texture and bread quality are not in opposition to one another. The soft kernel trait itself appears to exert no negative affect on yield, agronomic performance or pest resistance (Kiszonas et al., 2019).

AUTHOR CONTRIBUTIONS

CM wrote the manuscript. CM, AK, JM, JB, MI, and XC conceived the research. AK, JM, JB, MI, MZ, and XC performed the research and analyses, and contributed to the writing.

FUNDING

Funding was provided in part by USDA NIFA 2013-67013-21226 and 2019-67013-29164, and USDA ARS CRIS Proj. 2090 43440-007-00-D.

ACKNOWLEDGMENTS

Carlos Guzman, Karim Ammar, Claudia Carter, Teng Vang, Domenico Lafiandra, Marco Simeone, Leonard Joppa, Stacey Sykes, and the staff of the Western Wheat Quality Lab.

REFERENCES

Ammar, K., Kronstad, W. E., and Morris, C. F. (2000). Breadmaking quality of selected durum wheat genotypes and its relationship with high molecular weight glutenin subunits, allelic variation and gluten protein polymeric composition. *Cereal Chem*. 77, 230-236. doi: 10.1094/CCHEM.2000.77.2.230

Bhave, M., and Morris, C. F. (2008). Molecular genetics of puroindolines and related genes: allelic diversity in wheat and other grasses. *Plant Mol. Biol*. 66, 205–219. doi: 10.1007/s11103-007-9263-7

Boehm, J. D. Jr., Ibba, M. I., Kiszonas, A. M., and Morris, C. F. (2017a). End-use quality of CIMMYT-derived soft-kernel durum wheat germplasm: I. Grain, milling and soft wheat quality. *Crop. Sci*. 57, 1475–1484. doi: 10.2135/cropsci2016.09.0774

Boehm, J. D. Jr., Ibba, M. I., Kiszonas, A. M., and Morris, C. F. (2017b). End-use quality of CIMMYT-derived soft-kernel durum wheat germplasm: II. Dough strength and pan bread quality. *Crop Sci*. 57, 1485–1494. doi: 10.2135/cropsci2016.09.0775

Boehm, J. D. Jr., Zhang, M., Cai, X., and Morris, C.F. (2017c). Molecular and cytogenetic characterization of the 5DS-5BS chromosome translocation conditioning soft kernel texture in durum wheat. *Plant Genome* 10, 1–11. doi: 10.3835/plantgenome2017.04.0031

Giroux, M. J., and Morris, C. F. (1997). A glycine to serine change in puroindoline b is associated with wheat grain hardness and low levels of starch-surface friabilin. *Theor. Appl. Genet*. 95, 857–864. doi: 10.1007/s001220050636

Giroux, M. J., and Morris, C. F. (1998). Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. *Proc. Natl. Acad. Sci. U.S.A*. 95, 6262-6266. doi: 10.1073/pnas.95.11.6262

Heinze, K., Kiszonas, A. M., Murray, J. C., Morris, C. F., and Lullien-Pellerin, V. (2016). Puroindoline genes introduced into durum wheat reduce milling energy and change milling behavior similar to soft common wheats. *J. Cereal Sci*. 71, 183-189. doi: 10.1016/j.jcs.2016.08.016

Ibba, M. I., Zhang, M., Cai, X., and Morris, C. F. (in press). Identification of a conserved *ph1b*-mediated 5DS-5BS crossing over site in soft-kernel

durum wheat (*Triticum turgidum* subsp. *durum*) lines. *Euphytica*. doi: 10.1007/s10681-019-2518-y

Kiszonas, A. M., Higginbotham, R., Chen, X. M., Garland-Campbell, K., Bosque-Perez, N. A., Pumphrey, M., et al. (2019). Agronomic traits in durum wheat germplasm possessing puroindoline genes. *Agron. J.* 111, 1254–1265 doi: 10.2134/agronj2018.08.0534

Lukaszewski, A. (2003). Registration of six germplasms of durum wheat with introgressions of the Glu-D1 locus. *Crop Sci.* 43, 1138–1139. doi: 10.2135/cropsci2003.1138

Morris, C. F. (2002). Puroindolines: the molecular genetic basis of wheat grain hardness. *Plant Mol. Biol.* 48, 633–647. doi: 10.1023/A:1014837431178

Morris, C. F., and Beecher, B. S. (2012). The distal portion of the short arm of wheat (*Triticum aestivum* L.) chromosome 5D controls endosperm vitreosity and grain hardness. *Theor. Appl. Genet.* 125, 247–254. doi: 10.1007/s00122-012-1830-x

Morris, C. F., Casper, J., Kiszonas, A. M., Fuerst, E. P., Murray, J., Simeone, M. C., et al. (2015). Soft kernel durum wheat–a new bakery ingredient? *Cereal Foods World* 60, 76–83. doi: 10.1094/CFW-60-2-0076

Morris, C. F., Simeone, M. C., King, G. E., and Lafiandra, D. (2011). Transfer of soft kernel texture from *Triticum aestivum* to durum wheat, *Triticum turgidum* ssp. durum. *Crop Sci.* 51, 114–122. doi: 10.2135/cropsci2010.05.0306

Murray, J. C., Kiszonas, A. M., and Morris, C. F. (2017). Influence of soft kernel texture on the flour, water absorption, rheology, and baking quality of durum wheat. *Cereal Chem.* 94, 215-222. doi: 10.1094/CCHEM-06-16-0163-R

Murray, J. C., Kiszonas, A. M., Wilson, J. D., and Morris, C. F. (2016). Effect of soft kernel texture on the milling properties of soft durum wheat. *Cereal Chem.* 93, 513–517. doi: 10.1094/CCHEM-06-15-0136-R

Sissons, M., Pleming, D., Sestili, F., and Lafiandra, D. (2019). Effect of Glu-D1 gene introgression and amylose content on breadmaking potential of blends of durum and hexaploid wheat. *Cereal Chem.* 96, 193–206. doi: 10.1002/cche.10136

Conflict of Interest: CM is a co-inventor of soft durum wheat whose rights are assigned to the USDA.

The remaining authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 22 November 2019
doi: 10.3389/fpls.2019.01509

Genetic Variation for Protein Content and Yield-Related Traits in a Durum Population Derived From an Inter-Specific Cross Between Hexaploid and Tetraploid Wheat Cultivars

*Angelica Giancaspro, Stefania L. Giove, Silvana A. Zacheo, Antonio Blanco and Agata Gadaleta**

Department of Agricultural and Environmental Sciences (DiSAAT), Research Unit of "Genetics and Plant Biotechnology", University of Bari Aldo Moro, Bari, Italy

OPEN ACCESS

Edited by:
Jose Luis Gonzalez Hernandez, South Dakota State University, United States

Reviewed by:
Maria Itria Ibba, International Maize and Wheat Improvement Center (Mexico), Mexico
Qijun Zhang, North Dakota State University, United States
Karl D. Glover, South Dakota State University, United States

***Correspondence:**
Agata Gadaleta
agata.gadaleta@uniba.it

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *11 June 2019*
Accepted: *30 October 2019*
Published: *22 November 2019*

Citation:
Giancaspro A, Giove SL, Zacheo SA, Blanco A and Gadaleta A (2019) Genetic Variation for Protein Content and Yield-Related Traits in a Durum Population Derived From an Inter-Specific Cross Between Hexaploid and Tetraploid Wheat Cultivars. Front. Plant Sci. 10:1509. doi: 10.3389/fpls.2019.01509

Wheat grain protein content (GPC) and yield components are complex quantitative traits influenced by a multi-factorial system consisting of both genetic and environmental factors. Although seed storage proteins represent less than 15% of mature kernels, they are crucial in determining end-use properties of wheat, as well as the nutritional value of derived products. Yield and GPC are negatively correlated, and this hampers breeding programs of commercially valuable wheat varieties. The goal of this work was the evaluation of genetic variability for quantity and composition of seed storage proteins, together with yield components [grain yield per spike (GYS) and thousand-kernel weight (TKW)] in a durum wheat population obtained by an inter-specific cross between a common wheat accession and the durum cv. Saragolla. Quantitative trait loci (QTL) analysis was conducted and closely associated markers identified on a genetic map composed of 4,366 SNP markers previously obtained in the same durum population genotyped with the 90K iSelect SNP assay. A total of 22 QTL were detected for traits related to durum wheat quality. Six genomic regions responsible for GPC control were mapped on chromosomes 2B, 3A, 4A, 4B, 5B, and 7B, with major QTL on chromosomes 2B, 4A, and 5B. Nine loci were detected for GYS: two on chromosome 5B and 7A and one on chromosomes 2A, 2B, 4A, 4B, 7B, with the strongest QTL on 2B. Eight QTL were identified for TKW, three of which located on chromosome 3A, two on 1B and one on 4B, 5A, and 5B. Only small overlapping was found among QTL for GYS, TKW, and GPC, and increasing alleles coming from both parents on different chromosomes. Good candidate genes were identified in the QTL confidence intervals for GYS and TKW.

Keywords: grain protein content (GPC), protein quality, quantitative trait loci (QTL), wheat, genetic map

INTRODUCTION

Mediterranean countries rely heavily on cereal production as main commodity crop of economic importance. Amongst cereals, durum wheat is the leading commercial crop as its consumption is the highest amongst Mediterranean countries. Given the always growing consumers' attention to health, objectives of wheat breeding programs are recently focusing not only on increasing productivity,

but also on endowing derived products with a higher nutritional value (Battenfield et al., 2016). Wheat quality is a very complex issue including several components (Marcotuli et al., 2015). Usually, kernel protein quantity and quality are considered key factors of "wheat quality." Wheat quality, that is its ability to be processed into derived foods, is largely determined both by protein quantity—measured as grain protein content (GPC), and by protein quality—given by protein composition. Wheat proteins show high complexity and different interaction with each other, thus making them difficult to characterize. Wheat grain storage proteins are a complex mix of different polypeptide chains traditionally classified according to their solubility (following the sequential Osborne extraction procedure) or composition and structure (Zilic et al., 2011). Grain storage proteins include gliadin, glutenin, albumin, and globulin. In particular, gluten proteins that is gliadins and glutenins, give wheat flour unique extensibility and processing properties and consequently a good quality to derived products. Gluten is made up by proteins able to form a cohesive and visco-elastic dough by mixing flour with water. Differences in wheat processing properties are due to gluten quantity and composition. According to their solubility in aqueous alcoholic solutions, gluten proteins can be divided into soluble gliadins and insoluble glutenins. Gliadins are monomeric proteins responsible for gluten viscosity. Glutenins are polipeptidic aggregates responsible for dough strength and elasticity. After reduction of inter-molecular disulfide bonds, glutenin monomers are classified into high molecular weight (HMW) and low molecular weight (LMW) subunits (Payne et al., 1982; Jackson et al., 1983). Glutenin chains have high molecular weights up to tens of millions (Wrigley, 1996). Differences in glutenin subunits strongly influence end-use quality (Kasarda, 1999), in fact when managing the protein quality, one of the main aim of breeding programs is to select wheat varieties or lines that possess glutenin alleles associated with superior processing characteristics.

Wheat storage proteins are encoded by gene families with highly variable alleles. Gliadins are encoded by complex loci located on the short arms of homoeologous group 1 chromosomes (*Gli-A1*, *Gli-B1*, *Gli-D1* for omega and gamma) and 6 (*Gli-A2*, *Gli-B2*, *Gli-D2* for alpha- and beta-) (Shewry et al., 2003). For what concerns gutenins, LMW-GS derive from loci (*Glu*-3) on the short arms of group 1 chromosome (linked to *Gli-1* loci), whereas HMW-GS are coded by complex loci on the group 1 long arms (*Glu-1*). Each *Glu*-1 locus contains two tightly linked paralogous genes encoding two different types of HMW-GS, namely the "x-" and a "y-" type, based on their electrophoretic mobility, with a wide combination of different alleles at each locus responsible for high polymorphism.

Hexaploid wheat contains six HMW-GS genes (*Glu-1*) located on 1A, 1B, and 1D chromosomes, whereas tetraploid wheat accounts for four HMW-GS loci on 1A and 1B. In hexaploid wheat there are five classes of HMW-GS proteins: 1Ax (encoded by *Glu-A1*),1Bx and 1By (encoded by *Glu-B1*), 1Dx and 1Dy (encoded by *Glu-D1*). The y-type genes at the *Glu-A1* locus are missing in hexaploid wheat (Galdi and Feldman, 1983). HMW glutenin-subunits from 1D are lacking in durum wheat. In hexaploid wheat, the 1Bx, 1Dx, and 1Dy HMW-GS are constitutively present in all cultivars, whereas 1Ay genes are always silent.

Gliadin and glutenins contribute together to determine wheat dough quality. In particular, the presence of some HMW-GS is crucial for the determination of wheat quality and its attitude to make pasta or bread (Jang et al., 2017). Noteworthy, 80% of the tetraploid genotypes miss the HMW-GS encoded by *Glu-A1* locus, with the consequence to be unsuitable for bread making quality (Branlard et al., 1989), in fact durum wheat is usually used to making pasta.

By influencing both the nutritional value and the processing properties of flour/semolina, GPC deeply contributes to determine the economic value of the crop. GPC is a typical quantitative trait controlled by several genes located on almost all chromosomes and influenced by both environment and management practices. As mature wheat grains usually contain 8 to 20% of proteins (Zilic et al., 2011), one of the main objectives of breeders is to identify stable QTL and relative favorable alleles, which can be successfully transferred from high-GPC donor lines to low-GPC varieties.(Kumar et al., 2018). However, reaching of this goal is hampered by the quantitative nature of the trait, which is controlled by multiple genes and strongly influenced by environmental factors and agronomical practices (Blanco et al., 2012). In fact, heritability for GPC has been reported to range from 0.41 (Kramer, 1979) to 0.70 (Suprayogi et al., 2009) depending on genotypes, location of field trials, and computational analysis.

Protein quality and quantity have always been among the crucial topics of wheat breeding. However, the increase of grain protein content is also complicated by its negative association with productivity, the strong environmental influence and the very narrow genetic variation found within the cultivated gene pool (Barneix, 2007; Iqbal et al., 2016). Modern wheat varieties are relatively low in GPC compared to their wild relatives, which could be employed as donors of useful alleles (Avivi, 1978; Nevo et al., 2002; Cakmak et al., 2004; Blanco et al., 2006; Chatzav et al., 2010). For example, wild emmer (*Triticum turgidum* ssp. *dicoccoides*) has a wide genotypic variation in several agronomic traits, including grain constituents (Uauy et al., 2006) and abiotic/biotic stress tolerance (Peleg et al., 2005; Ben-David et al., 2010). For this reason, the high genetic variability in the gene pool of wild emmer can be exploited, by crossing with durum wheat, for the improvement of several agronomic and economical important traits. However, its very low yield complicates wheat breeding for yield components.

Improvement of GPC by means of classical breeding methods have always given poor results. Recently, the release of genome sequences and comparative studies with model species (Colasuonno et al., 2013), and the consequent availability of advanced genetic tools such as linkage maps and molecular markers has fasten the improvement of GPC especially thanks to the application of marker-assisted selection (MAS) programs. Genetic factors and candidate genes influencing protein concentration in cultivated and wild wheats are located on all chromosomes. A recent review by Kumar et al. (2018) reported 49 QTL mapping publications with a total of 325 main effect QTL and 42 epistatic QTL for GPC.

Given the complexity of GPC genetic control and the existence of a very narrow genetic variation within cultivated gene pool, the improvement of this trait has always been a challenge for genetic breeders. Biodiversity is becoming an increasingly important issue in agriculture. Significant research of new valuable germplasm, breeding efforts, and subsequent plant multiplication are needed to improve the performance of cereal sector through better-suited varieties. It is fundamental continuing to achieve successful genotypes and to guarantee creation of new genetic diversity to allow a continued genetic gain in a dynamic agricultural environment.

Together with grain protein content, wheat yield is another complex agronomic trait resulting from the interaction of several components which are deeply influenced by environment. Understanding the molecular mechanisms underlying yield-related genes is one of the main objective of breeders. Several QTL have been mapped for some components such as thousand kernel weight, grain weight per spike, tiller number etc. (Huang et al., 2004; Narasimhamoorthy et al., 2006; Zhang et al., 2012), and some functional markers have been developed for thousand-kernel weight (TKW) (Su et al., 2011). However, gene isolation by map-based cloning is hampered by the hugeness and complexity of wheat genome, thus only few genes have been isolated through comparative genetics based on gene synteny (Su et al., 2011; Yang et al., 2012).

In the present work, an inter-specific cross between hexaploid and tetraploid wheat was used to map QTL for GPC and yield-related traits [grain yield per spike (GYS) and TKW], and identify putative candidate genes for the most associated markers. Moreover, the use of an accession of bread wheat as donor of useful genes for durum wheat, allowed to survey new genetic variability for grain quantity and composition, and identify new high quality lines useful in durum breeding programs.

MATERIALS AND METHODS

Genetic Materials and Field Experiments

A total of 135 recombinant inbred lines (RILs) were obtained by the interspecific cross between the elite durum wheat cultivar Saragolla and the bread wheat accession 02-5B-318 (derived from the Chinese cv. Sumai-3) by advancing single F_2 plants to the F_7 by single seed descent (SSD) (Giancaspro et al., 2016). The cross produced two RIL populations: a sub-set of RILs were classified as hexaploid as carrying all the D chromosomes, and a sub-set of 135 durum RILs consisted of lines with no D chromosomes. In this work, only the durum RIL was taken into consideration for the study of grain protein content and yield-related traits The tetraploid lines were harvested, and seeds used for replicated field trials and DNA extraction. The parental lines for this interspecific cross were selected because different for several traits including yield components and protein content. The parents and the 135 RILs were evaluated for grain protein content and yield in replicated field trials at the location of Valenzano Bari-Italy for three years (2015–2017). The RILs were evaluated using a randomized complete block design with four replications. Each plots consisted of 1-m rows, 30 cm apart, with 80 seeds sown in each plot and supplemented with nitrogen (10 g/m^2). Grain protein content was determined on a 2 gsample of whole meal flour by near-infrared reflectance (NIR) spectroscopy using a SpectraAlyzer device (Basic model, Zeuton). The instrument was calibrated by using a set of 25 whole-meal flour samples belonging to *T. turgidum* ssp. *durum*, ssp. *dicoccum,* and ssp. *dicoccoides* with known protein concentration and moisture, previously calculated according to official standard methods. Final GPC was expressed as a percentage of proteins on a dry weight basis. On each wheat line, GPC measurement was performed twice then the final value was averaged between the two technical replicates. GYS was measured as the total grain yield per row on the number of spikes per row (about 70–80 spikes). TKW was assessed from a 15 g seed sample per each plot (line). TKW of each wheat line was determined as the average of two technical replicates.

HMW-GS and gliadins analyses were performed as described by the method of Payne et al. (1980) and by Bushuk and Zillman (1978), respectively. The cv Chinese Spring of bread wheat and cv Svevo of durum wheat, containing known HMW-GS, were used as standards. HMW-glutenins composition was scored according to Payne's catalogue (Payne and Lawrence, 1983) that named HMW glutenins gene loci as *Glu-A1*, *Glu-1B,* and *Glu-1D* and proteins subunits as 0, 1, 2*, 2 + 12, 5 + 10, 6 + 8, 7 + 9, and 17 + 18.

Genetic Map and QTL Analysis

A durum wheat genetic linkage map obtained by Giancaspro et al. (2016) was used for QTL analysis. The map, covering a total length of 4,227.37 cM, consists of 4,366 SNPs surveyed from the 81,587 sequences of the 90K iSelect array by Illumina CSProR (SanDiego, CA, USA) described by Wang et al. (2014).

ANOVA was conducted for each trait with standard procedures using X-Stat software and genetic variance (σ^2_G) and broad-sense heritability (h^2_B) calculated using the variance component estimates. Pearson phenotypic correlation coefficients were calculated by using M-STAT-C software between GPC, GYS, and TKW. The Inclusive Composite Interval Mapping (ICIM) method (Li et al., 2007) was used for QTL mapping with QGene 4.0 software (Joehanes and Nelson, 2008). Scanning interval of 2 cM between markers, and putative QTL with a window size of 10 cM was used for QTL detection. The number of marker cofactors for background control was set by forward regression with a maximum of five controlling markers. Putative QTL were defined as two or more linked markers associated with a trait at LOD ≥3. Positive and negative signs of the estimates QTL effect indicate the contribution of cv Saragolla and the 02-5B-318 accession, respectively. The phenotypic variance explained by each single QTL was quantified by the square of the partial correlation coefficient (R^2). Graphical representation of linkage groups and QTL was obtained with MapChart 2.2 Software (Voorrips, 2002).

HMW-GS Analysis

Wheat grains were grinded with porcelain mortar and pestle and boiled for 5' in an extraction buffer - ratio 10:1 (µl/g)—consisting of 0.4 ml β-mercaptoethanol, 4 ml pure water, and 1.7 ml γ-piromin dye (15 mg γ piromin; 2.0 g 100% SDS; 6.25 ml Tris-Hcl, pH 6.8; 10.5 ml water). β-mercaptoethanol served for the reduction of

intermolecular disulphide bonds between glutenin subunits, SDS for protein denaturation, and γ-piromin for the visualization of protein bands on electrophoretic gel. Samples were incubated 2 h at room temperature, then centrifuged to recover supernatant. Electrophoresis was performed in SDS-PAGE with 4% stacking gel (pH 6.8) and 10% separating gel (pH8.8) for 24 h at 500 Volts, in a1X running buffer (Tris-SDS-Glycin, pH 8.3). Gels were left 24 h in a dying solution of Comassie Blue R250 composed by 12% tricloroacetic acid and 1%Comassie Brilliant Blue (19:1, v/v), then rinsed overnight with distilled water and visualized on UV. Common wheat cv. Chinese Spring and durum wheat cv. Svevo, with known HMW-GS, were used as standards.

Gliadins Analysis

Gliadin extraction was performed on 30 µg of single-seed flour by following the protocol described in Bushuk and Zillman (1978). 100 µl of 70% ethanol were added to flour and incubated at 24°C for 1 h with brief vortexing at 10 min intervals. Supernatant was collected by centrifuging tubes at 12,000 rpm for 10 min at room temperature. Then, 26 µl of gel buffer was added to 20 µl of supernatant and loaded onto the gel for PAGE separation. Details on gel preparation and run are described in Bushuk and Zillman (1978). The electrophoregrams were evaluated based on gliadin relative mobility.

Candidate Gene Analysis

In order to assign genes to SNPs, sequences corresponding to all the SNP markers mapping in the confidence intervals of QTL for GPC, GYS and TKW were used in a BLAST search using TBLASTX algorithm (http://ncbi.nlm.nih.gov/BLAST), against the wheat draft genome sequence of the tetraploid *dicoccoides* accession Zavitan (Avni et al., 2017), the durum cv. Svevo (Maccaferri et al., 2019), and the hexaploid Chinese Spring (Appels et al., 2018) in order to assign a putative function to SNPs, and identify candidate genes (CGs) for the traits of interest. Alignments with at least 80% identity and E-value >10^{-7} were considered, and the corresponding putative genes evaluated for their involvement in metabolic pathways related to protein content and yield components.

Molecular Structure (Exons/Introns) and Function (Corresponding Coding Protein) Was Determined for the Most Associated CG to TWK QTL: Elongation Factor (EF) Gene. Comparison Was Carried on Among Diploid (*Triticum Urartu*), Tetraploid (*T. Durum*, Cv. Svevo, *T. Dicoccoides*), and Hexaploid (*T. Aestivum*, Cv. Chinese Spring) Wheats. Molecular Characterization Was Obtained by Browsing the Corresponding SNP Sequence (IWB30162, Excalibur_Rep_C106003_475) in the Svevo Genome Browser at: https://Interomics.Eu/Durum-Wheat-Genome and in Plant Ensemble Databases (https://Plants.Ensembl.Org).

RESULTS

Field Traits Analysis

GYS and TKW were evaluated at Valenzano (Bari) in southern Italy for three years (2015, 2016, 2017). **Table 1** reports the mean value and the range of GPC and yield components for parents Saragolla and 02-5B-318 and for RI lines, together with variance components and broad-sense heritability estimates, in each

TABLE 1 | Means, ranges, coefficients of variation (CV), genetic variance (σ^2_G), and heritability (h^2_B) of grain protein content and grain yield components in the Saragolla x 02-5B-318 RIL population and parental lines, evaluated in three environments.

Trait	Environments			
	Valenzano 2015	Valenzano 2016	Valenzano 2017	Mean
Grain Protein Content, GPC (%)				
02-5B-318	13.88	14.45	12.37	13.56
Saragolla	12.33	15.12	13.25	13.56
Mean RIL	16.37	16.66	15.06	16.03
Range	12.8-22.10	13.44-21.80	12.41-18.12	12.88-20.67
CV(%)	6.66	6.50	5.90	6.35
σ^2_G	1.04	1.00	1.10	1.04
h^2_B	0.48	0.47	0.47	0.47
Grain Yield per Spike, GYS (g)				
02-5B-318	1.30	1.47	0.99	1.25
Saragolla	2.26	2.78	1.68	2.24
Mean RIL	1.13	1.26	1.05	1.15
Range	0.22-2.40	0.38-2.45	0.47-1.88	0.36-2.24
CV(%)	20.88	19.90	20.00	20.26
σ^2_G	0.08	0.07	0.08	0.08
h^2_B	0.57	0.50	0.52	0.53
1000 Kernel Weight, TKW (g)				
02-5B-318	32.68	43.56	30.58	35.61
Saragolla	41.76	51.52	38.16	43.81
Mean RIL	35.71	36.32	36.55	36.19
Range	22.51-67.84	15.50-60.41	9.93-53.00	15.98-60.41
CV(%)	11.01	12.00	11.92	11.64
σ^2_G	35.18	34.20	36.00	35.13
h^2_B	0.69	0.68	0.69	0.62

of the three environments. GPC values were different between the two parents both in Valenzano 2015, 2016, and 2017, with a mean value of 14.23 for cv Saragolla and 13.56 for 02-5B-318. As expected for a typical quantitative trait, GPC values showed a broad variability across the RI lines in each of the three environments, in fact the phenotypic means of RILs were distributed across a normal curve (**Figure 1**). The RIL population means (16.4, 16.7, and 15.0 for the three years, respectively), were significantly higher of both mid-parental values and with a minimum value of 12.8 and a maximum value of 22.10 recorded at Valenzano 2015. Differences in mean values and variances of parental lines and RI population observed in the different locations, were very likely due to the different environmental factors. Broad sense heritability (genotype mean basis) of GPC ranged from 0.47 to 0.48 in the three environments.

Saragolla and 02-5B-318 showed a significant difference also for the yield components GYS and TKW; consequently, a large variation was observed for these two traits in the RIL population in all the three environments. The durum cv. Saragolla showed higher values in each environment for both traits: in particular, the mean values for GYS and TKW were 2.24 and 43.81 g, respect to 1.14 and 35.60 g of the same traits in 02-5B-318 parent. Interestingly, for both GYS and TKW, the trait variation was much larger in the RILs population than between the parental lines, with lowest values of 0.22 and 9.93 g and highest values of 2.45 and 67.84, for GYS and TKW, respectively, which suggested the presence of favorable alleles increasing the trait in both parents. For example, in all the three environments, the difference between the highest and the lowest value of GYS in the RI lines was more than twice the difference between the two parents. Heritability across environments was high for both GYS (mean value of 0.53) and TKW (mean value of 0.69).

As expected, in the present work we found a negative correlation between GPC and yield components (**Table 2**). In particular, GPC resulted negatively correlated with GYS (r values ranging from -0.09 to -0.16) in all environments, and with 1,000 kernel weight only at Valenzano 2017 (r = -1.00).

QTL Analysis

QTL analysis was performed according to the Inclusive Composite Interval Mapping method (ICIM, Li et al., 2007). Putative QTL (LOD ≥ 3.0) for GPC, GYS, and TKW in individual environments are listed in **Tables 3–5** respectively. QTL positions on the durum wheat linkage map are reported in **Figure 2**. Suggestive QTL, with a LOD value comprised between 2.0 and 3.0, were also reported.

In the present work we identified six major QTL for GPC, each on chromosome 2B, 3A, 4A (two loci), 4B, 5B, and 7B, responsible for the highest percentage of the overall phenotypic variation of the trait in the three environments. Interestingly, we found trait-increasing alleles coming from both the high-GPC parent Saragolla and from the low-GPC parent 02-5B-318. Specifically, alleles from Saragolla were those mapped on chromosomes 4A, 5B, and 7B, whereas alleles from 02-5B-318 were those on 2B, 3A, and 4B. The additive effect values ranged from 0.27 to 0.71 of protein content unit. The percentage of

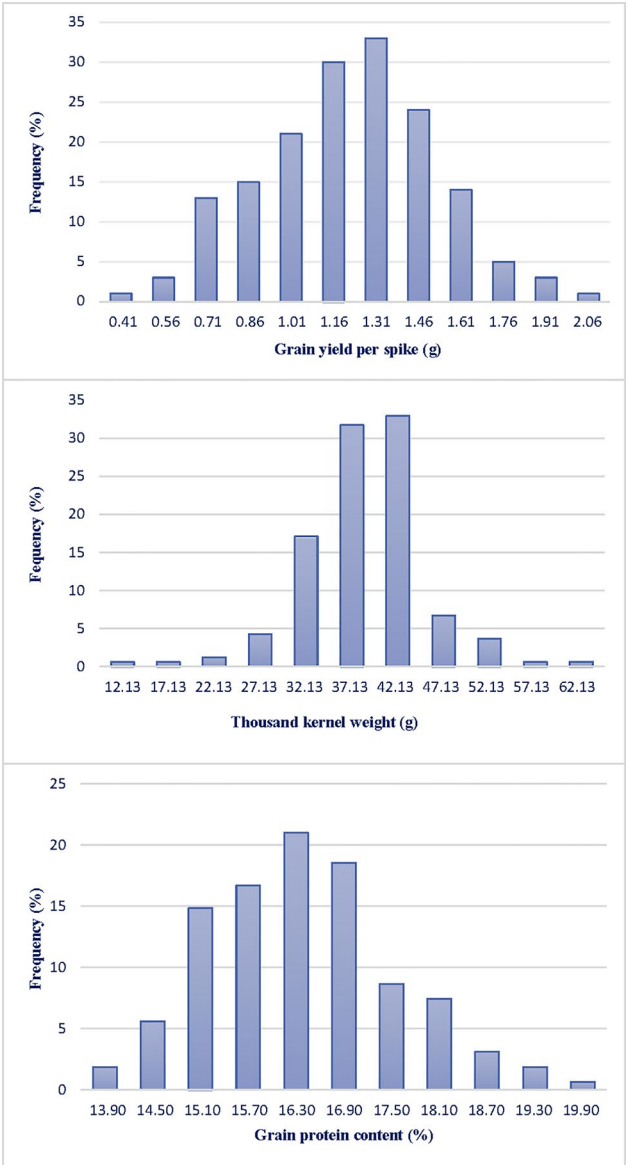


FIGURE 1 | Frequency distribution of grain yield per spike (GYS), thousand kernel weight (TKW) and grain protein content (GPC) in a RIL population derived by the cross between the durum cv Saragolla and the bread wheat accession 02-5B-318. Traits were evaluated at Valenzano (Bari, southern Italy) for three years (2015, 2016, 2017); the mean value of the three years are presented.

TABLE 2 | Correlation coefficients between grain protein content and grain yield components in the Saragolla x 02-5B-318 RIL mapping population, evaluated in three environments. (Valenzano 2015, Valenzano 2016, Valenzano 2017).

Trait	Grain protein content		
	Valenzano 2015	Valenzano 2016	Valenzano 2017
Grain yield per spike	–0.11***	–0.09***	–0.16***
1000 kernel weight	–0.07	0.43	–1.00*

, * Significant differences at 0,05 P, 0,01 P e 0,001 P respectively.*

TABLE 3 | QTL for grain protein content (GPC) mapped in the durum wheat RIL population derived from the cross between the bread wheat accession 02-5B-318 and the durum wheat cv. Saragolla, evaluated in three different environments (Valenzano 2015, Valenzano 2016, Valenzano 2017). Analyses were performed by ICIM (Inclusive Composite Interval Mapping).

Chrom.	Linkage group	Marker interval	Most associated marker	Map position (cM)	Valenzano 2015			Valenzano 2016			Valenzano 2017			Environment mean		
					Add [a]	LOD	R^2	Add	LOD	R^2	Add	LOD	R^2	Add	LOD	R^2
2BS	2B-2	IWB31001-IWA897	IWB72906	67.6	–0.7	5.6	20	–0.4	2.2°	08	–0.3	2.4°	09	–0.4	4.9	17
3AS	3A-2	IWB64668-IWB72529	IWB72484	108.5	–0.4	2.4°	09	–0.4	3.3	12	–	–	–	–0.3	2.2°	08
4AL	4A-2	IWB7798-IWB39495	IWB54916	76.3	–.6	4.1	15	0.6	6.4	22	–	–	–	–.4	5.6	19
4BL	4B-2	IWB36086-IWB34895	IWB55835	58.6	–0.6	3.8	14	–	–	–	–	–	–	–0.4	4.3	15
5BL	5B-4	IWB54873-IWB11747	IWB11571	31.6	0.7	6.2	22	0.5	3.7	13	0.4	5.5	19	0.5	7.0	23
7BL	7B-3	IWB10498-IWB69574	IWB69002	39.0	0. 6	3.9	15	0.4	3.2	12	–	–	–	0.4	4.4	15

° Suggestive QTLs at the sub-threshold 2.0 < LOD < 3.0.

- Not significant QTL.

[a]Additive Effect: positive additive effects are associated with an increased effect from Saragolla alleles and negative additive effects with an increased effect from 02-5B-318 alleles

R^2: Percentage of phenotypic variance explained by the additive effects of the mapped QTL.

TABLE 4 | QTL for grain yield per spike (GYS) mapped in the durum wheat RIL population derived from the cross between the bread wheat accession 02-5B-318 and the durum wheat cv. Saragolla, evaluated in three different environments (Valenzano 2015, Valenzano 2016, Valenzano 2017). Analyses were performed by ICIM (Inclusive Composite Interval Mapping).

Chrom.	Linkage group	Marker interval	Most associated marker	Map position (cM)	Valenzano 2015			Valenzano 2016			Valenzano 2017			Environment mean		
					Add [a]	LOD	R^2	Add	LOD	R^2	Add	LOD	R^2	Add	LOD	R^2
2AL	2A-3	IWA32-IWB71282	IWB7315	6.9	0.1	6.2	21	0.2	7.7	27	0.1	3.6	13	0.1	10.4	32
2BS	2B-2	IWA1665-IWA897	IWB320054	67.0	0.3	23.6	29	0.3	3.3	12	0.3	23.0	29	0.3	23.6	08
4AL	4A-2	IWB87-IWB12722	IWB59450	170.7	–0.1	5.1	18	–	–	–	–	–	–	–0.1	3.7	13
4BL	4B-2	IWB71836-IWB70999	IWB64615	13.3	0.1	6.7	23	–	–	–	–	–	–	–0.1	5.2	18
5BL	5B-2	IWB72334-IWB20927	IWB72334	45.3	–0.1	4.8	17	–	–	–	–	–	–	–0.1	4.2	15
5BL	5B-3	IWA8097-IWB149734	IWB34530	66.7	–0.1	5.1	18	–	–	–	–	–	–	–	–	–
7AL	7A-6	IWB7367-IWA8312	IWA6576	72.8	0.1	3.9	14	–	–	–	–	–	–	0. 8	3.2	11
7BL	7B-3	IWB1711-IWB9018	IWA8570	16.7	–0.1	6.1	21	–	–	–	–0.1	5.4	19	–0.1	8.0	26

- Not significant QTL.

[a]Additive Effect: positive additive effects are associated with an increased effect from Saragolla alleles and negative additive effects with an increased effect from 02-5B-318 alleles

R^2: Percentage of phenotypic variance explained by the additive effects of the mapped QTL.

TABLE 5 | QTL for thousand kernel weight (TKW) mapped in the durum wheat RIL population derived from the cross between the bread wheat accession 02-5B-318 and the durum wheat cv. Saragolla, evaluated in three different environments (Valenzano 2015, Valenzano 2016, Valenzano 2017). Analyses were performed by ICIM (Inclusive Composite Interval Mapping).

Chrom.	Linkage group	Marker interval	Most associated marker	Map position (cM)	Valenzano 2015			Valenzano 2016			Valenzano 2017			Environment mean		
					Add[a]	LOD	R^2	Add	LOD	R^2	Add	LOD	R^2	Add	LOD	R^2
1BS	1B-1	IWB62561-IWB8572	IWB10407	72.0				–2.4	4.7	17						
1BL	1B-3	IWB69144-IWB14436	IWB69144	0.00	–2.1	3.6	13	–3.6	9.6	32	–2.0	4.2	15	–2.3	5.9	20
3AS	3A-2	IWA950-IWB73247	IWB73711	37.2	3. 6	8.2	27	2.8	6.0	21	2.9	7.4	25	3.3	10.3	32
3AS	3A-2	IWB71453-IWB27100	IWB8477	91.6	-2.8	4.9	18	–3.0	6.1	22	–3.6	10.1	32	–3.1	8.5	27
3AL	3A-3	IWB58806-IWB70483	IWB37509	49.7	–	–	–	–2.2	3.8	14	–	–	–	–1.6	2.9°	11
4BL	4B-3	IWB7164-IWB24289	IWA892	33.0	–	–	–	–3.0	6.6	23	–	–	–	–	–	–
5AL	5A-4	IWA1258-IWB72888	IWA1258	0.00	–	–	–	–2.1	3.3	13	–	–	–	–	–	–
5BL	5B-4	IWB42947-IWB764	IWB7719	15.9	-4. 8	5.1	28	-5.7	5.8	28	-4.5	46.2	83	-4.8	5.5	28

° Suggestive QTLs at the sub-threshold 2.0 < LOD < 3.0.

- Not significant QTL.

[a]Additive Effect: positive additive effects are associated with an increased effect from Saragolla alleles and negative additive effects with an increased effect from 02-5B-318 alleles

R^2: Percentage of phenotypic variance explained by the additive effects of the mapped QTL.

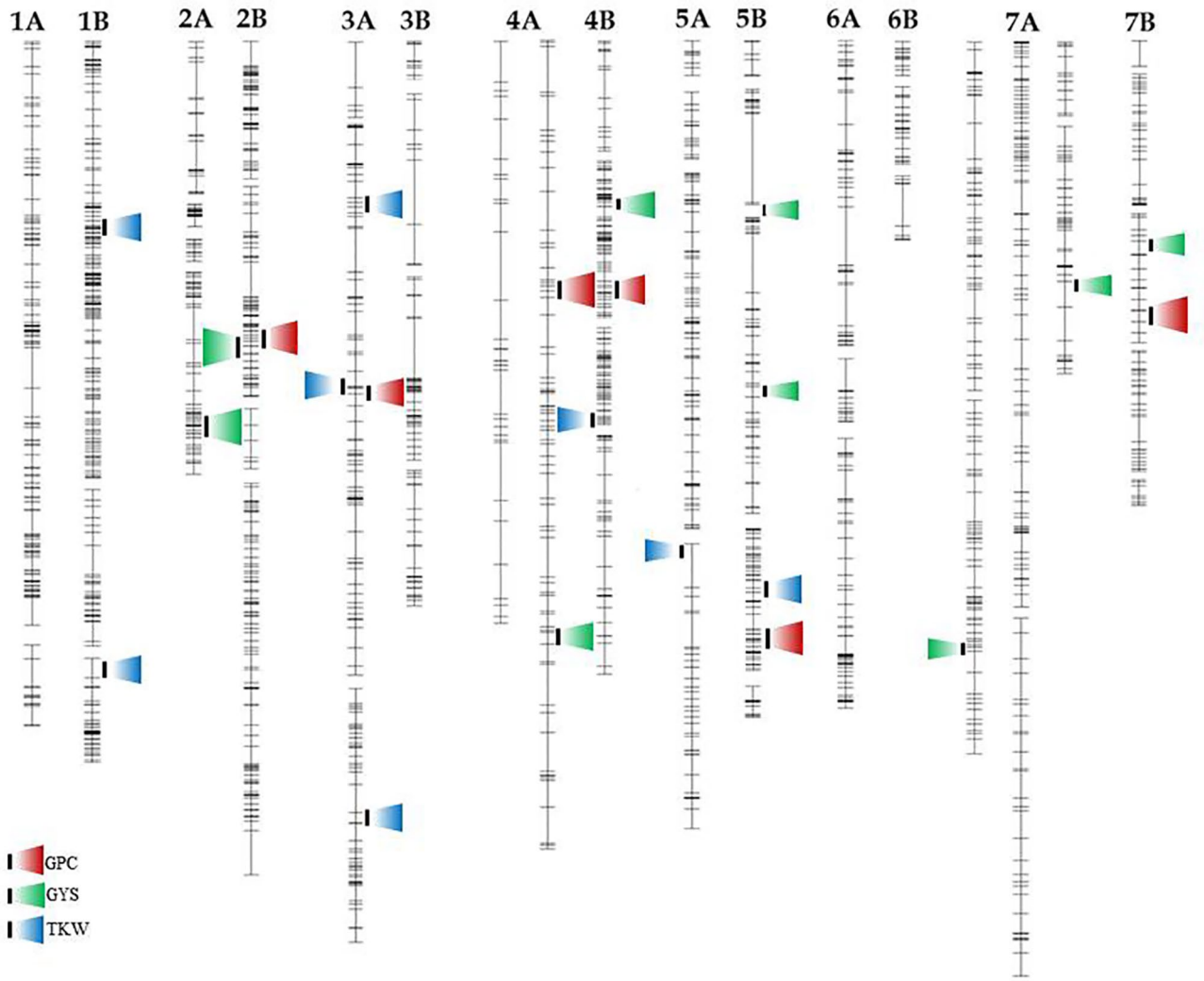


FIGURE 2 | Genetic location of QTL for GYP, TKW, and GPC mapped on the durum wheat linkage map obtained in the RIL population derived by the cross between the durum cv. Saragolla and the bread wheat accession 02-5B-318. Loci for the different traits are reported in different colors (GPC, red; GYS, green; TKW, blue).

phenotypic variation (R^2) contributed by each single QTL in each environment was comprised between 8 and 23%. QTL on 2B and 5B were found significant in all the environments, whereas three were significant in two environments, and one was significant in one single location (**Table 3**).

Eight chromosomal regions were detected for GYS (**Table 4**), two on chromosome 5B and 7A and one on chromosomes 2A, 2B, 4A, 4B, 7B. Five QTL came from cv Saragolla and four from the 02-5B-318 accession line, three of which were found in all the environments. The strongest QTL was the one detected on 2B chromosome with a LOD score of 23; this stable QTL was responsible of up to 29% of GYS variation, and the positive allele came from Saragolla, the durum parental line with the highest GYS. The GYS QTL on 7B was significant in two environments, and the positive allele came from 02-5B-318. The remaining five QTL mapped for GYS were significant only in one single environment, with a quote of phenotypic variation comprised between 14 and 23%.

In the present work we mapped eight QTL for TKW, three of which were located on chromosome 3A, two on 1B and one on 4B, 5A, and 5B (**Table 5**). Four QTL were stable in the three environments and significant in the mean of environments, while the others QTL were significant only in one single environment. Each QTL had an additive effect ranging from of 1.67 g (QTL on 3A-3) to 5.67 g (QTL on 5B), and explained 11 to 32% of trait variation. Out of the eight QTL detected, six came from the bread parental line accession 02-5B-318, whereas only two derived from the durum wheat cv Saragolla.

Interestingly, only few overlapping occurred between QTL for grain yield components and QTL for GPC on 2B and 3A chromosomes (**Figure 2**). In most cases, the QTL was responsible for the genetic control of only one of the traits.

Putative Function and Candidate Gene Detection for GPC and Yield Components

One of the major steps of a gene-based analysis is the assignment of SNPs to genes. In order to identify a putative function for SNP markers and search for CGs for the traits of interest, SNP sequences were BLAST searched against GeneBank non redundant (nr) database. A total of 378 SNP sequences included in QTL regions were analysed, specifically 83 for GPC, 129 for TKW and 166 for GYP (for each QTL, SNPs from co-migrating loci mapped in a confidence interval of 15 cM were considered). **Tables 6–8** list SNP loci for which sequence homology was retrieved for GPC, GYS and TKW, respectively. Overall, 22% of the total query sequences matched with known function genes (E value > 10^{-7}), but only 10% corresponded to protein products involved in metabolic pathways related to the traits of interest. Sequences similarity searches revealed storage proteins loci, and genes involved in different cellular processes such as: DNA transcription, chromatin structure determination, response to biotic and abiotic stress, carbon metabolisms, modification of cell wall structure, and intracellular signalling.

As reported in **Table 6**, putative functions retrieved for SNPs included in the GPC-QTL were seed storage proteins (ω/γ/δ gliadins), or enzymes involved in RNA processing (RNA ligase), amino acid modification (serine carboxypeptidase, tryptophan aminotransferase), determination of protein tertiary structure (PDI-like protein) and modification of proteins destined to degradation (Ubiquitin_protein ligase, GW2-B gene promoter).

Table 7 shows the putative function detected for the SNP markers mapped in the GYS-QTL. Some sequences matched with genes for gliadin storage protein; other SNPs were found to fall into genes related to plant response to stress (Secologanin synthase - Cyt P450- and Pyrroline 5-carboxylate synthase) or involved in sugar metabolism (Alcohol dehydrogenase, ADH1A). In this work, a particular attention was paid to the candidate gene matching with the marker IAAV2296 (IWB34530). Bioinformatics research showed that SNP sequence has a high similarity percentage (95%) with the *Aegilops tauschii* gene for *Trasparent-testa-glabra1* (*TTG-1*). Molecular structure of *TTG-1* was compared in diploid (*T. urartu*, *Aegilops tauschii*), tetraploid (*T. durum*, cv. Svevo, *T. dicoccoides*), and hexaploid (*T. aestivum*, cv. Chinese Spring) wheats, where genes were located on A, B, and D genome. Details on gene structure of *TTG* in wheat, together with *Arabidopsis*, tobacco and rice are reported in **Supplementary Table S2**.

For what concerns TKW, the identified candidate genes in the QTL region were storage ω/γ/δ gliadins, RNA helicase, proteins involved in photosynthesis (**Table 8**). Attention was focused on candidate genes identified in the confidence interval for the QTL on 5B chromosome, where major QTL for TKW have been cited

TABLE 6 | Candidate genes identified in the QTL for grain protein content (GPC) mapped in the interspecific RIL population derived by the cross between the hexaploid wheat accession 02-5B-318 and the durum wheat cv. Saragolla (putative function is reported for the SNP markers in the QTL confidence interval showing an identity percentage > 80%).

QTL confidence interval	SNP name	SNP id	Chromosome arm	Linkage Group	Map position (cM)	Putative candidate gene	Species	E-value	Identity percentage (%)
IWB31001-IWA897	GENE-0644_370	IWB32005	2BS	2B-2	67.1	E3_Ubiquitin_protein ligase (SP1)	*B. distachyon*	$3e^{-30}$	89.8
	Excalibur_c1305_662	IWB22202	2BS	2B-2	67.1	E3_Ubiquitin_protein ligase (SP1)	*B. distachyon*	$3e^{-30}$	89.8
	GENE-0644_421	IWB32007	2BS	2B-2	67.1	E3_Ubiquitin_protein ligase (SP1)	*B. distachyon*	$3e^{-30}$	89.8
	Ra_c106376_879	IWB50813	2BS	2B-2	67.1	RNA ligase isoform 2	*T. aestivum*	$8e^{-44}$	99.0
IWB64668-IWB72529	RFL_Contig4403_1034	IWB64668	3AS	3A-2	105.4	Serine carboxypeptidase-like 33 (GS5)	*T. aestivum*	$8e^{-45}$	99.0
	wsnp_Ex_c10272_16842803	IWA1308	3AS	3A-2	105.4	Serine carboxypeptidase-like 33 (GS5)	*T. aestivum*	$8e^{-45}$	99.0
	Excalibur_c8930_548	IWB29267	3AL	3A-2	119.6	Tryptophan aminotransferase related 3 (TAR3.1-3B)	*T. aestivum*	$6e^{-15}$	100.0
IWB36086-IWB34895	IAAV8654	IWB35513	4BL	4B-2	55.7	GW2-B gene, promoter region	*T. turgidum* ssp. *durum*	$1e^{-28}$	87.6
IWB10498-IWB69574	BS00070791_51	IWB10498	7BL	7B-3	34.7	Omega/gamma/delta gliadin (LMW-D1/D2/D3/D6/D7)	*T. aestivum*	$9e^{-10}$	78.2
	Tdurum_contig23504_196	IWB69002	7BL	7B-3	39.0	Omega/gamma/delta gliadin (LMW-D1/D2/D3/D6/D7)	*T. aestivum*	$9e^{-10}$	78.2
	Excalibur_c57808_355	IWB27761	7BL	7B-3	39.1	Omega/gamma/delta gliadin (LMW-D1/D2/D3/D6/D7)	*T. aestivum*	$9e^{-10}$	78.2
	Tdurum_contig28630_245	IWB69574	7BL	7B-3	43.6	PDI-like protein (pdil8-1)	*T. aestivum*	$3e^{-06}$	88.1

TABLE 7 | Candidate genes identified in the QTL for grain yield per spike (GYS) mapped in the interspecific RIL population derived from the cross between the hexaploid wheat accession 02-5B-318 and the durum wheat cv. Saragolla (putative function is reported for the SNP markers in the QTL confidence interval showing an identity percentage > 80%).

QTL confidence interval	SNP name	SNP id	Chromosome arm	Linkage Group	Map position (cM)	Putative candidate gene	Species	E-value	Identity percentage
IWA1665-IWA897	BS00083078_51	IWB11285	2BS	2B-2	62.2	Gamma gliadin-A1/A3/A4 and LMW-A2 genes	*T. aestivum*	$3e^{-12}$	85.7
	JD_c39990_130	IWB37419	2BS	2B-2	64.2	Gamma gliadin-A1/A3/A4 and LMW-A2 genes	*T. aestivum*	$3e^{-12}$	85.7
	Tdurum_contig62595_466	IWB72906	2BS	2B-2	67.7	Secologanin synthase (Cyt P450)	*Ae. tauschii*	$5e^{-37}$	94.1
IWB71836-IWB70999	RFL_Contig4212_597	IWB64614	4BS	4B-2	13.6	Pyrroline 5-carboxylate synthetase (P5CS1)	*T. turgidum* ssp. *durum*	$8e^{-15}$	80.2
	Tdurum_contig98478_400	IWB74042	4BS	4B-2	13.8	Alcohol dehydrogenase (ADH1A)	*T. turgidum* ssp. *dicoccoides*	$1e^{-06}$	100.0
	Tdurum_contig98478_494	IWB74043	4BS	4B-2	13.8	Alcohol dehydrogenase (ADH1A)	*T. turgidum* ssp. *dicoccoides*	$1e^{-06}$	100.0
IWB72334-IWB20927	wsnp_Ku_c14202_22436656	IWA6516	5BL	5B-2	46.2	Gamma,delta,omega gliadin	*T. aestivum*	$1e^{-13}$	80.9
	Tdurum_contig11060_433	IWB66909	5BL	5B-2	46.2	Gamma,delta,omega gliadin	*T. aestivum*	$1e^{-13}$	80.9
IWA8097-IWB14973	IAAV2296	IWB34530	5BL	5B-3	66.7	Transparent testa glabra1 *(TTG-1)*	*Ae. tauschii*	$3e^{-30}$	95.0

TABLE 8 | Candidate genes identified in the QTL for thousand kernel weight (TKW) mapped in the interspecific RIL population derived by the cross between the hexaploid wheat accession 02-5B-318 and the durum wheat cv. Saragolla (putative function is reported for the SNP markers in the QTL confidence interval showing an identity percentage > 80%).

QTL confidence interval	SNP name	SNP id	Chromosome arm	Linkage Group	Map position (cM)	Putative candidate gene	Species	E-value	Identity percentage
IWB62561-IWB8572	Excalibur_c20610_149	IWB23524	1BS	1B-1	66.9	Annexin 6-2	*T. turgidum* ssp. *Durum*	$2e^{-21}$	98.4
	wsnp_BF291549B_Ta_1_1	IWA435	1BS	1B-1	66.9	GSK-like kinase 1B (GSK1B)	*T. aestivum*	$2e^{-53}$	83.2
	Excalibur_c20610_251	IWB23525	1BS	1B-1	66.9	Annexin 6-2	*T. turgidum* ssp. *Durum*	$2e^{-21}$	98.4
	IAAV2366	IWB34541	1BS	1B-1	75.1	Gamma-glutamylcysteine synthetase (GSH1)	*T. aestivum*	$9e^{-56}$	99.0
IWB69144-IWB14436	CAP7_rep_c6352_375	IWB14436	1BL	1B-3	11.8	Chlorophyll a-b binding protein of LHCII	*Ae. Tauschii*	$7e^{-41}$	97.1
IWA950-IWB73247	CAP8_c2839_118	IWB14646	3AS	3A-2	20.3	GID2-A1 protein (gid2-A1)	*T. aestivum*	$5e^{-35}$	98.9
	RAC875_c64107_404	IWB59845	3AS	3A-2	42.0	Gamma gliadin-A	*T. aestivum*	$7e^{-08}$	75.5
IWB58806-IWB70483	Tdurum_contig34075_98	IWB70483	3AL	3A-3	54.2	Gamma/delta/omega gliadin-B	*T. aestivum*	7e-08	75.5
IWB7164-IWB24289	CAP8_rep_c3658_272	IWB15007	4BL	4B-3	35.6	Catalase (CAT)	*T. turgidum* ssp. *durum*	$1e^{-42}$	99.0
IWA1258-IWB72888	Tdurum_contig9074_2085	IWB73761	5AL	5A-4	17.6	RNA helicase (DEAD1-B)	*T. aestivum*	$1e^{-18}$	87.6
IWB42947-IWB764	Excalibur_rep_c106003_475	IWB30162	5BL	5B-4	11.0	Protein elongation factor	*T. aestivum*	$7e^{-43}$	99.0
	Excalibur_c9370_944	IWB29437	5BL	5B-4	11.9	Squamosa promoter-binding-like protein 6 (SPL6)	*T. aestivum*	$9e^{-04}$	73.3
	BobWhite_c15241_604	IWB764	5BL	5B-4	20.9	Ammonium transporter (AMT2.1)	*T. aestivum*	$6e^{-40}$	97.0

in literature (Wang et al., 2012; Assanga et al., 2017; Su et al., 2018). In this region, we found very good candidates as some of the SNP markers fell into genes in some way involved in the determination of kernel weight, such as a protein elongation factor (Excalibur_rep_c106003_475), a squamosa promoter-binding-like protein (Excalibur_c9370_944), and an ammonium transporter (BobWhite_c15241_604). Among these, we decided to better characterize the EF gene, whose homologous on 7A chromosome has recently (Zheng et al., 2014) been assessed to be significantly associated with grain number per spike and to potentially increase

TABLE 9 | HMW-GS and γ-42/γ-45 gliadin segregation in the durum wheat RIL population derived from the cross between the bread wheat accession 02-5B-318 and the durum cv. Saragolla.

	Glu-A1		Gli-B1		Gli-B1	
	2*	null	7+9	6+8	γ-42	γ-45
02-5B-318	+	−	+	−	+	−
Saragolla	−	+	−	+	−	+
N. of RILs	3	132	72	63	55	80

wheat grain yield and yield-related traits. The molecular structure was compared in diploid (*Triticum urartu*), tetraploid (*T. durum*, cv. Svevo, *T. dicoccoides*), and hexaploid (*T. aestivum*, cv. Chinese Spring) wheats (**Supplementary Table 1**). The gene could be identified only on 5B and 5D homoeologous chromosomes, whereas no genes were found on A genome, in both tetraploid and hexaploid genotypes. *EF* gene is characterized from having several alternative splicing forms, with a variable number of exons/introns.

Grain Protein Composition

In the present work both parents and the complete RIL population were characterized for HMW-GS and for the gliadin *Gli-B1* locus (**Table 9**). For HMW-GS the bread parent has for *Glu-A1* gene the *1Ax2** allele encoding for 2* HMW-GS, for *Glu-B1* the *Bx7* and *By9* alleles encoding 7 + 9 HMW-GS and γ42 gliadin subunit for *Gli-B1* gene, while the durum wheat cv Saragolla presented the following allele combination: 6 + 8 for HMW-GS and the γ-45 for *Gli-B1* gene. A segregation for all genes, reported in **Table 9**, was observed in the RIL population, and in particular on 135 analyzed RI lines, 72 showed the HMW-GS profile of Saragolla parent (6 + 8) and 63 RI lines the HMW-GS of the bread wheat parent (7 + 9). Moreover, 55 RI lines showed the γ-45 gliadin allele coming from cv Saragolla and 80 the γ-42 allele of the bread wheat parent.

DISCUSSION

In the present work, an inter-specific RIL population of 135 tetraploid lines, originally developed by Giancaspro et al. (2016) for investigating the genetic basis of Fusarium head blight (FHB) resistance in durum wheat, was further characterized for GPC, protein composition, and yield components (TKW and GYS). In all the environments, the traits showed a wide difference between the two parents and a big variability in the RIL population. In particular, the range of phenotypic values in the RIL population was larger than between the two parents, clearly indicating the establishment of new genetic variability in the RI lines due to the combination of favorable alleles originating from both parents. This has been confirmed by the fact that the QTL identified for GYS, TKW, and GPC came from both parents (**Tables 1**–**5**).

Although wheat is an important source for protein, most of the modern varieties are relatively low in grain protein (10–14%) especially in cultivars containing the *Rht1* dwarfing gene (McClung et al., 1986). Breeding efforts to increase GPC have been influenced by the strong impact of environmental conditions and by the negative effect on yield components, especially when the alleles come from wild emmers (Simmonds, 1995; Feil, 1997; Oury et al., 2003; Oury and Godin, 2007).

In this study the evaluation of grain yield components and GPC in three field trials lead us to the identification of 22 QTL distributed among all chromosomes excluding 1A, 3B, and 6A. This confirmed previous studies where QTL for GPC and grain yield were identified on almost all wheat chromosomes (Börner et al., 2002; Marza et al., 2005; McCartney et al., 2005; Quarrie et al., 2005; Huang et al., 2006; Gadaleta et al., 2007; Kuchel et al., 2007; Kumar et al., 2007; Maccaferri et al., 2008; McIntyre et al., 2010; Marcotuli et al., 2017; Mangini et al., 2018). Several QTL were found significant only in a one environment, and several authors reported that when a QTL was detected in more than one environment, a variation in its effects occurred (Huang et al., 2003; Huang et al., 2004; Kumar et al., 2007; Kuchel et al., 2007; Maccaferri et al., 2008; McIntyre et al., 2010).

By comparing the GPC-QTL found in the present study with those reported in the most recent literature, we found that loci on 2BS, 4BL, and 5BL were also reported by Nigro et al. (2019) and they occupy almost the same genomic region. Moreover, QTL on 2B, 4A, and 7B were also found in the work by Marcotuli et al. (2017). On chromosomes 2B and 4A we found QTL on a different chromosome arm (2BS instead of 2BL and 4AL rather than 4AS). For what concerns 7B, different stable QTL for GPC have been reported in different studies (reviewed by Kumar et al., 2018), most of which were located on the short arm of chromosome 7B. However, accordingly to what reported by Blanco et al. (2006) and Marcotuli et al. (2017), we found the QTL on 7B located on the long arm of the chromosome. Also in the work by Roselló et al. (2018) we found GPC loci on 4A, 4B, 5B, and 7B. Differences observed in map position could be due to the different mapping populations used for QTL analyses, the marker coverage of linkage maps, and the high number of genes controlling the trait.

Our work confirmed the negative relationship between GPC and yield component traits (**Table 2**), although only the QTL for GYS and GPC on 2B chromosome were co-localized. The other QTL on 3A, 4A, 4B, 5B, and 7B showing significant effects on GPC values were interesting because not liked to yield potential, probably because contain genes that influence GPC independently from variation in the grain yield components, and could be used to improve the GPC. The high-density consensus linkage map for durum wheat described by Maccaferri et al. (2019) was used as reference for chromosome localization and SNP markers position to compare our results with those reported in literature.

A total of 16 QTL for GYS and TKW were also detected on almost all chromosomes of durum wheat, of which none coincident. These data are consistent with several works that reported significant QTL for yield components almost on all wheat chromosomes (Araus et al., 2008; Cui et al., 2014; Gupta et al., 2017). QTL for TKW were found on 1B, 3A, 4B, 5A, and 5B chromosomes, while QTL for GYS were localized on homoeologous chromosome group 2, 4, 5, and 7 as also reported by Mangini et al. (2018) and Marcotuli et al. (2017), respectively for loci on 2BS and 4AL, and 2BS. Most of the yield loci mapped in the present work confirm the mapping data reported in recent literature: a very good correspondence was observed for the TKW-QTL detected on 1B (1A), 3A, 4B, 5A, and 5B which were also reported by Wang et al. (2012); Assanga et al.

(2017) and Su et al. (2018). In particular, the TKW-QTL identified in the present work on chromosomes 1B, 4B, and 5B localized on the same map confidence interval of those mapped by Wang et al. (2012) and Su et al. (2018).

In the present work, beside surveying new genetic variability for GPC and yield components, some alleles related to grain storage protein composition were transferred from bread wheat to durum wheat. Glutenin subunits or their alleles can be used as indicators of wheat quality and can be used as "markers" for breeding purposes. In particular, we obtained durum lines with 1Ax2* allele encoding for 2* HMW-GS, and lines with alleles encoding 7 + 9 HMW-GS. Branlard and Dardevet (1985) observed that Zeleny sedimentation value had positive correlation with subunits 7 + 9 and 5 + 10, and negative correlation with 2 + 12; moreover, extensibility had relationship with subunits 2* and 17 + 18. γ-gliadins 45 and 42 are valuable markers for good and poor pasta quality, respectively, and this is because the genetic linkage with low molecular weight glutenin subunits (Sapirstein et al., 2007). In the present work, durum lines with γ-42 gliadin subunit have been obtained and the identification of such lines suggests the possibility to improve durum wheat for bread-making processing, which is a common practice in the bakery tradition of certain Italian regions.

In conclusion, in the present work an enlargement of genetic variability has been achieved in a RIL population obtained by the cross of a bread and a durum wheat line with different quality characteristics. Lines with higher protein content and/or composition could be usefully employed directly or as donor lines in future breeding programs. Moreover, for the two yield components, confident CGs were identified associated to the QTL on 5BL: *TTG* (Transparent testa glabra) for GYS, and *EF* (Elongation factor) for TKW. In both cases we can assume these genes as good candidates as they have been reported in literature to be correlated to seed storage proteins accumulation or determination of grain number per spike. In particular, TTG proteins have been well characterized in *Arabidopsis* and *Nicotiana*: they are involved in the regulation of several processes of plant development and immunity (Wang et al., 2009; Li et al, 2012). In this work, *TTG* was chosen as putative candidate gene for GYS because other than taking part to anthocyanin biosynthesis (Pang et al., 2009; Zhang and Schrader, 2017), plant thricome formation (Morohashi et al., 2007; Wang et al., 2009), and pathogen resistance (Wang et al., 2009; Truman et al., 2010; Li et al., 2012), it was recently reported to be involved in seed storage proteins accumulation in *Arabidopsis* (Chen et al., 2015) and *Nicotiana* (Zhu et al., 2013; Ge et al., 2016).

For what concerns *EF*, in the work by Zheng et al. (2014) *Elongation Factor* expression has been reported to increase gradually during common wheat grain filling, expecially in young spikes and developing seedlings. Moreover, *Arabidopsis* transgenic plants harbouring a *TaTEF-7A* gene construct, showed a better vegetative growth and an increased silique length, silique number, and grain length. *TaTEF-7A* was also mapped in wheat using a DH population, resulting in a co-localization with several reported QTL for yield-related traits (spikelet number per spike, flour yield, test weight, and grain yield).

In conclusion, both candidate genes identified in this work well represent a good starting point for dissecting the molecular basis of some yield-related traits, and understanding the genetic mechanism underlying this complex quantitative trait. Moreover, they may also serve as useful tools for developing genetic markers suitable for MAS breeding programs.

DATA AVAILABILITY STATEMENT

All datasets for this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

AgG designed the experiments. SG performed phenotypic characterization of wheat lines for protein content and yield components. SZ performed protein extraction and electrophoretic evaluation. AnG performed QTL analyses and candidate gene identification. AgG and AnG collaborated to data interpretation and writing the manuscript. All authors read and approved the final manuscript.

ACKNOWLEDGMENTS

This research was funded by Research Grant from ISEA AGROSERVICE SpA (S. Severino Marche, MC, Italy); Italian Ministry of University and Research (MIUR); PRIN Grant n. 2010T7247Z.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2019.01509/full#supplementary-material

REFERENCES

Appels, R., Eversole, K., Feuille, C, Keller, B., Rogers, J., Stein, N., et al. (2018). The International Wheat Genome Sequencing Consortium (IWGSC). Shifting the limits in wheat research and breeding using a fully annotated reference genome. *Science* 361, 6403. doi: 10.1126/science.aar7191

Araus, J. L., Slafer, G. A., Royo, C., and Serret, M. D. (2008). Breeding for yield potential and stress adaptation in cereals. *Crit. Rev. Plant Sci.* 27, 377–412. doi: 10.1080/07352680802467736

Assanga, S., Zhang, G., Tan, C.-T., Rudd, J. C., Ibrahim, A., and Xue, Q. (2017). Saturated genetic mapping of wheat streak mosaic virus resistance gene Wsm2 in wheat. *Crop Sci.* 57 (1), 332–339. doi: 10.2135/cropsci2016.04.0233

Avivi, L. (1978). High grain protein content in wild tetraploid wheat Triticum dicoccoides Korn, in *5th Int. Wheat Genet. Symp*. Ed.Ramanujan, S. (India:New Delhi), 372–380.

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., and Gundlach, H. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* 357 (346), 93–97. doi: 10.1126/science.aan0032

Börner, A., Schumann, E., Fürste, A., Cöster, H., Leithold, B., and Röder, M. (2002). Mapping of quantitative trait loci determining agronomic important characters in hexaploid wheat (Triticum aestivum L.). *Theor. Appl. Genet.* 105 (6-7), 921–936. doi: 10.1007/s00122-002-0994-1

Barneix, A. J. (2007). Physiology and biochemistry of source-regulated protein accumulation in the wheat grain. *J. Plant Physiol.* 164 (5), 581–590. doi: 10.1016/j.jplph.2006.03.009

Battenfield, S. D., Guzmán, C., Gaynor, R. C., Singh, R. P., Dreisigacker, S., et al. (2016). Genomic selection for processing and end-use quality traits in the CIMMYT spring bread wheat breeding program. *Plant Genome* 9, 2. doi: 10.3835/plantgenome2016.01.0005

Ben-David, R., Xie, W. L., Peleg, Z., Saranga, Y., Dinoor, A., and Fahima, T. (2010). Identification and mapping of PmG16, a powdery mildew resistance gene derived from wild emmer wheat. *Theor. Appl. Genet.* 121, 499–510. doi: 10.1007/s00122-010-1326-5

Blanco, A., Simeone, R., and Gadaleta, A. (2006). Detection of QTLs for grain protein content in durum wheat. *Theor. Appl. Genet.* 112, 1195–1204. doi: 10.1007/s00122-006-0221-6

Blanco, A., Mangini, G., Giancaspro, A., Giove, S., Colasuonno, P., Simeone, R., et al. (2012) Relationships between grain protein content and grain yield components through QTL analyses in a RIL population derived from two elite durum wheat cultivars. *Mol. Breeding* 30 (1), 79–92.

Branlard, G., and Dardevet., M. (1985). Diversity of grain protein and bread quality. II. Correlation between high molecular weight subunits of glutenin and flour characteristics. *J. Cereal Sci.* 3, 345–354. doi: 10.1016/S0733-5210(85)80007-2

Branlard, G., Autran, J. C., and Monneveux, P. (1989). High molecular weight glutenin subunits in durum wheat (T. durum). *Theor. Appl. Genet.* 78, 353–358. doi: 10.1007/BF00265296

Bushuk, W., and Zillman, R. R. (1978). Wheat cultivar identification by gliadin electrophoregrams. i. apparatus, method and nomenclature. *Can. J. Plant Sci.* 58 (2), 505– 515. doi: 10.4141/cjps78-076

Cakmak, I., Torun, A., Millet, E., Feldman, M., Fahima, T., and Korol, A. B. (2004). Triticum dicoccoides: an important genetic resource for increasing zinc and iron concentration in modern cultivated wheat. *Soil. Sci. Plant Nutr.* 50, 1047–1054. doi: 10.1080/00380768.2004.10408573

Chatzav, M., Peleg, Z., Ozturk, L., Yazici, A., and Fahima, T. I. (2010). Genetic diversity for grain nutrients in wild emmer wheat: potential for wheat improvement. *Ann. Bot.* 105 (7), 1211–1220. doi: 10.1093/aob/mcq024

Chen, M., Zhang, B., Li, C., Kulaveerasingam, H., Chew, F. T., and Yu, H. (2015). TRANSPARENT TESTA GLABRA1 Regulates the Accumulation of Seed Storage Reserves in Arabidopsis. *Plant Physiol.* 169, 391–402. doi: 10.1104/pp.15.00943

Colasuonno, P., Maria, M. A., Blanco, A., and Gadaleta, A. (2013). Description of durum wheat linkage map and comparative sequence analysis of wheat mapped DArT markers with rice and Brachypodium genomes. *BMC Genet.* 14, 114. doi: 10.1186/1471-2156-14-114

Cui, F., Zhao, C., Ding, A., Li, J., Wang, L., and Li, X. (2014). Construction of an integrative linkare map and QTL mapping of grain yield-related traits using three related wheat RIL populations. *Theor. Appl. Genet.* 127, 659–675. doi: 10.1007/s00122-013-2249-8

Feil, B. (1997). The inverse yield-protein relationship in cereals: possibilities and limitations for genetically improving the grain protein yield. *Trends In Agron.* 1, 103–119.

Gadaleta, A., Mangini, G., Mulè, G., and Blanco, A. (2007) Characterization of dinucleotide and trinucleotide EST-derived microsatellites in the wheat genome. *Euphytica* 153 (1-2), 73–85.

Galdi, G., and Feldman, M. (1983). Diploidization of endosperm protein genes in polyploid wheats. *Proc. 6th Int. Wheat Genetics Symp.* (Kyoto, Japan), 1119–1123.

Ge, J., Li, B., Shen, D., Xie, J., Long, J., and Dong, H. (2016).Tobacco TTG2 regulates vegetative growth and seed production via the predominant role of ARF8 in cooperation with ARF17 and ARF19. *BMC Plant Biol.* 16, 126. doi: 10.1186/s12870-016-0815-3

Giancaspro, A., Giove, S. L., Zito, D., and Blanco, A. (2016). Mapping QTLs for Fusarium head blight resistance in an interspecific wheat population. *Front. Plant Sci.* 7, 1381. doi: 10.3389/fpls.2016.0138124

Gupta, P. K., Balyan, H. S., and Gahlaut, V. (2017). QTL analysis for drought tolerance in wheat: present status and future possibilities. *Agronomy* 7, 5. doi: 10.3390/agronomy7010005

Huang, X. Q., Cöster, H., Ganal, M. W., and Röder, M. S. (2003). Advanced backcross QTL analysis for the identification of quantitative trait loci alleles from wild relatives of wheat (Triticum aestivumL.). *Theor. Appl. Genet.* 106, 1379–1389. doi: 10.1007/s00122-002-1179-7

Huang, X. Q., Kempf, H., Ganal, M. W., and Röder, M. S. (2004). Advanced backcross QTL analysis in progenies derived from a cross between a German elite winter wheat variety and a synthetic wheat (Triticum aestivum L.). *Theor. Appl. Genet.* 109, 933–943. doi: 10.1007/s00122-004-1708-7

Huang, X. Q., Cloutier, S., Lycar, L., Radovanovic, N., Humphreys, D. G., Noll, J. S., et al. (2006). Molecular dissection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (Triticum aestivum L.). *Theor. Appl. Genet.* 113, 753–766.29. doi: 10.1007/s00122-006-0346-7

Iqbal, M., Moakhar, N. P., Strenzke, K., Haile, T., Pozniak, C., Hucl, P., et al. (2016). Genetic improvement in grain yield and other traits of wheat grown in Western Canada. *Crop Sci.* 56, 613–624. doi: 10.2135/cropsci2015.06.0348

Jackson, E. A., Holt, L. M., and Payne, P. I. (1983). Characterization of high-molecular weight gliadin and low-molecular-weight glutenin subunits of wheat endosperm by two dimensional electrophoresis and the chromosomal localization of their controlling genes. *Theor. Appl. Genet.* 66, 29–37. doi: 10.1007/BF00281844

Jang, Y. R., Beom, H. R., Altenbach, S. B., Lee, M. K., Lim, S. H., and Lee, J. Y. (2017). Improved method for reliable HMW-GS identification by RP-HPLC and SDS-PAGE in common wheat cultivars. *Molecules* 22 (7), 1055. doi: 10.3390/molecules22071055

Joehanes, R., and Nelson, J. C. (2008).QGene 4.0, an extensible Java QTL-analysis platform. *Bioinformatics* 24, 2788–2789. doi: 10.1093/bioinformatics/btn523

Kasarda, D. D. (1999). Glutenin polymers: The *in vitro* to *in vivo* transition. *Cereal Foods World* 44, 566–571.

Kramer, T. (1979). Environmental and genetic variation for protein content in winter wheat (Triticum aestivum L.). *Euphytica* 28, 209–218. doi: 10.1007/BF00056577

Kuchel, H., Williams, K. J., Langridge, P., Eagles, H. A., and Jefferies, S. P. (2007). Genetic dissection of grain yield in bread wheat: I. *QTL Anal. Theor. Appl. Genet.* 115, 1029–1041. doi: 10.1007/s00122-007-0629-7

Kumar, N., Kulwal, P. L., Balyan, H. S., and Gupta, P. K. (2007). QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. *Mol. Breed.* 19, 163–177. doi: 10.1007/s11032-006-9056-8

Kumar, A., Jain, S., Elias, E. M., Ibrahim, M., and Sharma, L. K. (2018). an overview of QTL identification and marker-assisted selection for grain protein content in wheat, in *Eco-friendly Agro-biological Techniques for Enhancing Crop Productivity*. Eds.Sengar, A., and Singh, R. (Singapore: Springer). doi: 10.1007/978-981-10-6934-5_11

Li, H., Ye, G., and Wang, J. (2007). A modified algorithm for the improvement of composite interval mapping. *Genetics* 175, 361–374. doi: 10.1534/genetics.106.066811.40

Li, B., Gao, R., Cui, R., Lü, B., Li, X., Zhao, Y., et al. (2012). Tobacco TTG2 suppresses resistance to pathogens by sequestering NPR1 from the nucleus. *J. Cell Sci.* 125, 4913–4922. doi: 10.1242/jcs.111922

Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L., Salem, M. B., and Bort, J. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat (Triticum durum Desf.) across a wide range of water availability. *Genetics* 178, 489–511. doi: 10.1534/genetics.107.077297

Maccaferri, M. Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51, 885–895. doi: 10.1038/s41588-019-0381-3

Mangini, G., Gadaleta, A., Colasuonno, P., Marcotuli, I., Signorile, A. M., and Simeone, R. (2018). Genetic dissection of the relationships between grain yield components by genome-wide association mapping in a collection of tetraploid wheats. *PloS One* 13 (1), e0190162. doi: 10.1371/journal.pone.0190162

Marcotuli, I., Houston, K., Waugh, R., Fincher, G. B., Burton,R. A., Blanco, A., Gadaleta, A. (2015) Genome wide association mapping for arabinoxylan content in a collection of tetraploid wheats. *PloS One*: 10 (7): e0132787

Marcotuli, I., Gadaleta, A., Mangini, G., Signorile, A. M., Zacheo, S. A., Blanco, A., et al. (2017). Development of a high-density snp-based linkage map and detection of QTL for ß-glucans, protein content, grain yield per

spike and heading time in durum wheat. *Int. J. Mol. Sci.* 18, 1329. doi: 10.3390/ijms18061329

Marza, F., Bai, G. H., Carver, B. F., and Zhou, W. C. (2005). Quantitative trait loci for yield and related traits in the wheat population Ning7840 Clark. *Theor. Appl. Genet.* 112, 688–698. doi: 10.1007/s00122-005-0172-3

McCartney, C. A., Somers, D. J., Humphreys, D. G., Lukow, O., Ames, N., and Noll, J. (2005). Mapping quantitative trait loci controlling agronomic traits in spring wheat cross RL4452× 'AC Domain'. *Genome* 48 (5), 870–883. doi: 10.1139/g05-055

McClung, A. M., Cantrell, R. G., Quick, J. S., and Gregory, R. S. (1986). Influence of the Rht1 semidwarf gene on yield, yield components, and grain protein in durum wheat1. *Crop Sci.* 26, 1095–1099. doi: 10.2135/cropsci1986.0011183X002600060001x

McIntyre, C. L., Mathews, K. L., Rattey, A., Chapman, S. C., Drenth, J., and Ghaderi, M. (2010). Molecular detection of genomic regions associated with grain yield and yield-related components in an elite bread wheat cross evaluated under irrigated and rainfed conditions. *Theor. Appl. Genet.* 120, 527–541. doi: 10.1007/s00122-009-1173-4

Morohashi, K., Zhao, M., Yang, M., Read, B., Lloyd, A., Lamb, R., et al. (2007). Participation of the Arabidopsis bHLH factor GL3 in trichome initiation regulatory events. *Plant Physiol.* 145, 736–746. doi: 10.1104/pp.107.104521

Narasimhamoorthy, B., Gill, B. S., Fritz, A. K., Nelson, J. C., and Brown-Guedira, G. L. (2006). Advanced backcross QTL analysis of a hard winter wheat×synthetic wheat population. *Theor. Appl. Genet.* 112, 787–796. doi: 10.1007/s00122-005-0159-0

Nevo, E., Korol, A. B., Beiles, A., and Fahima, T. (2002). *Evolution of wild emmer and wheat improvement: population genetics, genetic resources, and genome organization of wheat's progenitor, Triticum dicoccoides*. New York:Springer, Berlin Heidelberg. doi: 10.1007/978-3-662-07140-3

Nigro, D., Gadaleta, A., Mangini, G., Colasuonno, P., Marcotuli, I., Giancaspro, A., et al. (2019). Candidate genes and genome-wide association study of grain protein content and protein deviation in durum wheat. *Planta*. doi: 10.1007/s00425-018-03075-1

Oury, F. X., and Godin, C. (2007). Yield and grain protein concentration in bread wheat: how to use the negative relationship between the two characters to identify favourable genotypes?. *Euphytica* 157 (1-2), 45–57. doi: 10.1007/s10681-007-9395-5

Oury, F. X., Berard, P., Brancourt-Hulmel, M., Depatureaux, C., Doussignault, G., and Galic, N. (2003). Yield and grain protein concentration in bread wheat: a review and a study of multi-annual data from a French breeding program. *J. Genet. Breed.* 57, 59–68.

Pang, Y., Wenger, J. P., Saathoff, K., Peel, G. J., Wen, J., Huhman, D., et al. (2009). A WD40 repeat protein from Medicago truncatula is necessary for tissue specific anthocyanin and proanthocyanidin biosynthesis but not for trichome development. *Plant Physiol.* 151, 1114–1129. doi: 10.1104/pp.109.144022

Payne, P. I., and Lawrence, G. J. (1983). Catalogue of alleles for the complex gene loci, Glu-A1, Glu-B1, and Glu-D1which code for HMW-GS of glutenin in hexaploid wheat. *Cereal Res. Commun.* 11, 29–35.

Payne, P. I., Lawn, C. N., and Mudd, E. E. (1980). Control by homoeologous group 1 chromosomes of the high-molecular-weight subunits of glutenin, a major protein of wheat endosperm. *Theor. Appl. Genet.* 58 (3-4), 113–120.59. doi: 10.1007/BF00263101

Payne, P. I., Holt, L. M., Worland, A. J., and Law, C. N. (1982). Structural and genetical studies on the high-molecular-weight subunits of wheat. Part 3. Telocentric mapping of the subunit genes on the long arms of homoeologous group 1 chromosomes. *Theor. Appl. Genet.* 63, 129–128. doi: 10.1007/BF00303695

Peleg, Z., Fahima, T., Abbo, S., Krugman, T., Nevo, E., and Yakir, D. (2005). Genetic diversity for drought resistance in wild emmer wheat and its ecogeographical associations. *Plant Cell Environ.* 28, 176–191. doi: 10.1111/j.1365-3040.2005.01259.x

Quarrie, S. A., Steed, A., Calestani, C., Semikbodskii, A., Lebreton, C., and Chinoy, -VC. (2005). A high-density genetic map of hexaploid wheat (Triticum aestivum L.) from the cross Chinese Spring x SQ1 and its use to compare QTLs for grain yield across a range of environments. *Theor. Appl. Genet.* 110, 865–880. doi: 10.1007/s00122-004-1902-7

Roselló, M., Royo, C., Álvaro, F., Villegas, D., Nazco, R., and Soriano, J. M. (2018). Pasta-making quality qtlome from mediterranean durum wheat landraces. *Front. Plant Sci.* 9, 1512. doi: 10.3389/fpls.2018.01512

Sapirstein, H. D., David, P., Preston, K. R., and Dexter, J. E. (2007). Durum wheat breadmaking quality: Effects of gluten strength, protein composition, semolina particle size and fermentation time. *J. Cereal Sci.* 45, 150–161. doi: 10.1016/j.jcs.2006.08.006

Shewry, P. R., Halford, N. G., and Lafiandra, D. (2003). Genetics of wheat gluten proteins. *Adv. Genet.* 49, 111–184. doi: 10.1016/S0065-2660(03)01003-4

Simmonds, N. W. (1995). The relation between yield and protein in cereal grain. *J. Sci. Food Agric.* 67, 309–315. doi: 10.1002/jsfa.2740670306

Su, Z. Q., Hao, C. Y., Wang, L. F., Dong, Y. C., and Zhang, X. Y. (2011). Identification and development of a functional marker of TaGW2 associated with grain weight in bread wheat (Triticum aestivum L.). *Theor. Appl. Genet.* 122, 211–223. doi: 10.1007/s00122-010-1437-z

Su, Q., Zhang, X., Zhang, W., Zhang, N., Song, L., and Liu, L. (2018). QTL detection for Kernel size and weight in bread wheat (Triticum aestivum L.) using a high-density SNP and SSR-based linkage map. *Front. Plant Sci.* 9, 1484. doi: 10.3389/fpls.2018.0148467

Suprayogi, Y., Pozniak, C. J., Clarke, F. R., Clarke, J. M., and Knox, R. E. (2009). Identification and validation of quantitative trait loci for grain protein concentration in adapted Canadiam durum wheat populations. *Theor. Appl. Genet.* 119, 437–448. doi: 10.1007/s00122-009-1050-1

Truman, W. M., Bennett, M. H., Turnbull, C. G., and Grant, M. R. (2010). Arabidopsis auxin mutants are compromised in systemic acquired resistance and exhibit aberrant accumulation of various indolic compounds. *Plant Physiol.* 152, 1562–1573. doi: 10.1104/pp.109.152173

Uauy, C., Brevis, J. C., and Dubcovsky, J. (2006). The high grain protein content gene Gpc-B1 accelerates senescence and has pleiotropic effects on protein content in wheat. *J. Exp. Bot.* 57, 2785–2794. doi: 10.1093/jxb/erl047

Voorrips, R. E. (2002).MapChart: Software for the graphical presentation of linkage maps and QTLs. J. *Hered.* 93, 77–78. doi: 10.1093/jhered/93.1.77

Wang, Y., Liu, R., Chen, L., Liang, Y., Wu, X., Li, B., et al. (2009). Nicotiana tabacum TTG1 contributes to ParA1-induced signalling and cell death in leaf trichomes. *J. Cell Sci.* 122, 2673–2685. doi: 10.1242/jcs.049023

Wang, L., Ge, H., Hao, C., Dong, Y., and Zhang, X. (2012). Identifying loci influencing 1,000-Kernel weight in wheat by microsatellite screening for evidence of selection during breeding. *PloS One* 7, e29432. doi: 10.1371/journal.pone.0029432

Wang S. W., Forrest D., Allen K., Chao A., Huang S., Huang B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90,000 single nucleotide polymorphismarray. *Plant Biotechnol. J.* 12, 787–796.

Wrigley, C. W. (1996). Giant proteins with flour power. *Nature* 381, 738–739. doi: 10.1038/381738a0

Yang, Z. B., Bai, Z. Y., Li, X. L., Wan, P., Wu, Q. X., Yang, L., et al. (2012). SNP identification and allelic-specific PCR markers development for TaGW2, a gene linked to wheat kernel weight. *Theor. Appl. Genet.* 125, 1057–1068. doi: 10.1007/s00122-012-1895-6

Zhang, B., and Schrader, A. (2017). Transparent testa glabra 1-dependent regulation of flavonoid biosynthesis. *Plants* 6, 65. doi: 10.3390/plants6040065

Zhang, D. L., Hao, C. Y., Wang, L. F., and Zhang, X. Y. (2012). Identifying loci influencing grain number by microsatellite screening in bread wheat (Triticum aestivum L.). *Planta* 236, 1507–1517. doi: 10.1007/s00425-012-1708-9

Zheng, J., Liu, H., Wang, Y., Wang, L., Chang, X., and Jing, R. (2014). TEF-7A, a transcript elongation factor gene, influence yeld-related traits in bread wheat (Triticum aestivum L.). *J. Exp. Bot.* 65 (18), 5351–5365. doi: 10.1093/jxb/eru306

Zhu, Q., Li, B. Y., Mu, S. Y., Han, B., Cui, R. Z., Xu, M. Y., et al. (2013). TTG2- regulated development is related to expression of putative AUXIN RESPONSE FACTOR genes in tobacco. *BMC Genomics* 14, 806. doi: 10.1186/1471-2164-14-806

Zilic, S., Barac, M., Pešic, M., Dodig, D., and Ignjatovic-Micic, D. (2011). Characterization of proteins from grain of different bread and durum wheat genotypes. *Int. J. Mol. Sci.* 12 (9), 5878–5894. doi: 10.3390/ijms12095878

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 04 March 2020
doi: 10.3389/fpls.2020.00170

High Density Mapping of Quantitative Trait Loci Conferring Gluten Strength in Canadian Durum Wheat

Yuefeng Ruan[1], Bianyun Yu[2], Ron E. Knox[1], Asheesh K. Singh[1†], Ron DePauw[1†], Richard Cuthbert[1], Wentao Zhang[2], Isabelle Piche[1], Peng Gao[2], Andrew Sharpe[2†] and Pierre Fobert[3]*

[1] Swift Current Research and Development Centre, Agriculture and Agri-Food Canada, Swift Current, SK, Canada, [2] Aquatic and Crop Resource Development, National Research Council Canada, Saskatoon, SK, Canada, [3] Aquatic and Crop Resource Development, National Research Council Canada, Ottawa, ON, Canada

OPEN ACCESS

Edited by:
Roberto Tuberosa,
University of Bologna, Italy

Reviewed by:
Angelica Giancaspro,
University of Bari Aldo Moro, Italy
Filippo Maria Bassi,
International Center for Agricultural Research in the Dry Areas (ICARDA), Morocco

***Correspondence:**
Bianyun Yu
Bianyun.Yu@nrc-cnrc.gc.ca

†Present address:
Asheesh K. Singh,
Department of Agronomy, Iowa State University, Ames, IA, United States
Ron DePauw,
Advancing Wheat Technologies, Swift Current, SK, Canada
Andrew Sharpe,
The Global Institute for Food Security, Saskatoon, SK, Canada

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 15 September 2019
Accepted: 04 February 2020
Published: 04 March 2020

Citation:
Ruan Y, Yu B, Knox RE, Singh AK, DePauw R, Cuthbert R, Zhang W, Piche I, Gao P, Sharpe A and Fobert P (2020) High Density Mapping of Quantitative Trait Loci Conferring Gluten Strength in Canadian Durum Wheat. Front. Plant Sci. 11:170. doi: 10.3389/fpls.2020.00170

Gluten strength is one of the factors that determine the end-use quality of durum wheat and is an important breeding target for this crop. To characterize the quantitative trait loci (QTL) controlling gluten strength in Canadian durum wheat cultivars, a population of 162 doubled haploid (DH) lines segregating for gluten strength and derived from cv. Pelissier × cv. Strongfield was used in this study. The DH lines, parents, and controls were grown in 3 years and two seeding dates in each year and gluten strength of grain samples was measured by sodium dodecyl sulfate (SDS)-sedimentation volume (SV). With a genetic map created by genotyping the DH lines using the Illumina Infinium iSelect Wheat 90K SNP (single nucleotide polymorphism) chip, QTL contributing to gluten strength were detected on chromosome 1A, 1B, 2B, and 3A. Two major and stable QTL detected on chromosome 1A (*QGlu.spa-1A*) and 1B (*QGlu.spa-1B.1*) explaining 13.7–18.7% and 25.4–40.1% of the gluten strength variability respectively were consistently detected over 3 years, with the trait increasing alleles derived from Strongfield. Putative candidate genes underlying the major QTL were identified. Two novel minor QTL (*QGlu.spa-3A.1* and *QGlu.spa-3A.2*) with the trait increasing allele derived from Pelissier were mapped on chromosome 3A explaining up to 8.9% of the phenotypic variance; another three minor QTL (*QGlu.spa-2B.1*, *QGlu.spa-2B.2*, and *QGlu.spa-2B.3*) located on chromosome 2B explained up to 8.7% of the phenotypic variance with the trait increasing allele derived from Pelissier. *QGlu.spa-2B.1* is a new QTL and has not been reported in the literature. Multi-environment analysis revealed genetic (QTL) × environment interaction due to the difference of effect in magnitude rather than the direction of the QTL. Eleven pairs of digenic epistatic QTL were identified, with an epistatic effect between the two major QTL of *QGlu.spa-1A* and *QGlu.spa-1B.1* detected in four out of six environments. The peak SNPs and SNPs flanking the QTL interval of *QGlu.spa-1A* and *QGlu.spa-1B.1* were converted to Kompetitive Allele Specific PCR (KASP) markers, which can be deployed in marker-assisted breeding to increase the efficiency and accuracy of phenotypic selection for gluten strength in durum wheat. The QTL that were expressed consistently across

environments are of great importance to maintain the gluten strength of Canadian durum wheat to current market standards during the genetic improvement.

Keywords: durum wheat, gluten strength, sodium dodecyl sulphate-sedimentation volume, quantitative trait loci, single nucleotide polymorphism

INTRODUCTION

Durum wheat (*Triticum turgidum* L. ssp. durum), a tetraploid with A and B genomes (AABB), is an economically important crop and the source of semolina for the production of pasta, couscous, and various types of baked products particularly in Mediterranean countries (Sapirstein et al., 2007). Global durum production reached 40.2 million metric tons in 2016 (http://agfax.com/2017/03/23/wheat-market-global-durum-production-expected-to-fall-in-201718/) with 7.8 million tons produced in Canada (http://www.world-grain.com/articles/news_home/World_Grain_News/2016/12/Canada_wheat_production_up_15.aspx?ID=%7BD9C6D337-5F18-480D-B635-E996394D6E6C%7D&cck=1). Gluten strength, the ability of the gluten proteins to form a satisfactory protein/starch network that promotes good cooking quality, is a key determinant of the end-use quality in durum wheat (Dexter et al., 1980). Strong gluten is a prerequisite for the production of dough with excellent rheological characteristics and hence desired quality in the finished pasta products with greater textual characteristics and increased stability to overcooking (Irvine, 1971). Gluten strength relates to the balance between viscosity and elasticity (Sissons, 2008). A positive relationship between gluten strength and low temperature dried pasta viscoelasticity has been reported (Ames et al., 2003). Strong gluten with high elastic recovery gives better cooking stability and higher cooked firmness scores (Liu et al., 1996). Rheological properties of semolina, determined by the mixograph, farinograph, extensigraph, and alveograph, are generally used to predict the cooked pasta quality (Kovacs et al., 1997). It is widely accepted that semolina from extra strong durum wheat produces firmer pasta, although the optimal level of gluten strength required for firm pasta is not clear (Sissons, 2008). Pasta quality factors of commercial importance have been the primary focus of cultivar improvement and tested for the acceptability of any new durum cultivar in Canada resulting in substantial improvement over time (Clarke et al., 2010). As such, gluten strength is an important target for genetic improvement of Canadian durum varieties.

Gluten strength variation among genotypes is mainly affected by quality and quantity of gluten proteins which are composed of polymeric glutenins and monomeric gliadins categorized by their solubility in aqueous alcohol (Autran and Feillet, 1985; Du Cros, 1987; Feillet et al., 1989; Kovacs et al., 1991; Kovacs et al., 1993). Glutenins and gliadins, together accounting for about 75–80% of total flour protein, contribute to the rheological properties of the dough (Kumar et al., 2013). Gliadins are classified as α/β, γ, and ω gliadins according to their different mobility in an acid-polyacrylamide gel electrophoresis system [reviewed by (Barak et al., 2015)]. Glutenins can be further classified into two groups based on high and low molecular weight subunits (HMW-GS and LMW-GS) reflected by their mobility during sodium dodecyl sulfate polyacrylamide gel electrophoresis (SDS-PAGE). The HMW-GS comprise about 20–30% of the glutenin (Shewry et al., 1992; Henkrar et al., 2017). LMW-GS, the major class of glutenin subunits, accounts for 70–80% of the glutenin and a strong positive correlation of LMW-GS with durum wheat quality has been reported [reviewed by (Sissons, 2008)]. The ratio of glutenin to gliadin and the ratio of HMW-GS to LMW-GS are directly related to the functional properties of the dough (Wrigley et al., 2006; Sissons et al., 2007).

Various tests were used for the prediction of gluten strength, such as SDS-sedimentation test, gluten index, alveograph, and mixograph. The SDS-sedimentation test has positive correlation with gluten strength and has been widely used for the evaluation of quality of gluten protein and for fast screening in durum wheat breeding programs due to a few advantages such as the small sample size required, simplicity, and rapidness (Dexter et al., 1980; Quick and Donnelly, 1980; Clarke et al., 1998). SDS-sedimentation volume (SV) was reported to be a good predictor of cooked pasta disk viscoelasticity (Kovacs et al., 1995a) and has been widely used for evaluation of gluten strength in durum wheat breeding programs (Clarke et al., 1998). The efficacy of SV as the predictor for gluten strength might be confounded by the low to moderate positive correlation between SV and grain protein concentration (GPC) (Kovacs et al., 1995b; Clarke et al., 2010). However, no correlation between GPC and SV was reported as well (Brites and Carrillo, 2001).

Genetic studies have proposed the quantitative nature of the gluten strength trait with multiple genes coding glutenins and gliadins. Gliadins are encoded by loci *Gli-1* and *Gli-2* located on the short arm of the homoeologous group of chromosome 1 and 6 (Payne, 1987; Anderson et al., 2009). The *Gli-B1* locus on the short arm of chromosome 1B encoding γ-gliadins bands (γ-45/γ-42) was reported to be associated with gluten strength (Joppa et al., 1983; Pasqualone et al., 2015). Selection for the favorable γ-45 gliadin allele using a monoclonal antibody was implemented in very early generation of durum breeding (Clarke et al., 1998). However, later studies indicated that it was the linked LMW-2 rather than the γ-45 gliadin that was directly associated with gluten strength (Pogna et al., 1990). LMW-GS are encoded by gene clusters at *Glu-A3* and *Glu-B3* loci tightly linked with *Gli-1* on the short arms of chromosome 1 (D'Ovidio and Masci, 2004).

The HMW-GS displayed a high level of polymorphism and are encoded by *Glu-1* loci (*Glu-A1*, *Glu-B1*) on the long arms of chromosomes 1A and 1B (Payne and Lawrence, 1983). Each *Glu-*

1 locus contains two closely linked genes encoding two different types of HMW-GS, higher molecular weight x-type subunit and lower molecular weight y-type subunits (Shewry et al., 1992). Not all of these *Glu-1* genes are expressed in certain cultivars, resulting in variation in HMW-GS subunit number between genotypes (Xu et al., 2009). The *Glu-B1* locus presented higher polymorphism compared with *Glu-A1*. There are considerable allelic variations at *Glu-A1* and *Glu-B1* loci and a total of 40 alleles (6 for *Glu-A1* and 34 for *Glu-B1*) and 62 subunit combinations, were detected among 205 accessions of cultivated emmer wheat (*T. turgidum* ssp. *dicoccum* Schrank) collected from different regions of Europe and China (Li et al., 2006). Similarly, a total of 43 alleles, including 5 at *Glu-A1* and 38 at *Glu-B1*, resulting in 60 different allele combinations were identified in 232 accessions of durum wheat (*T. turgidum* L. ssp. *durum*) originated from various countries (Elfatih et al., 2013).

Moreover, quantitative trait loci (QTL) associated with gluten strength of durum wheat have been reported on a number of chromosomes, including chromosomes 1 and 6. Along with the major QTL on chromosome 1B and 1AL, Blanco et al. (1998) identified six additional loci on chromosomes 3AS, 3BL, 5AL, 6AL, and 7BS associated with gluten strength. As for the most quantitative traits, it has been reported that interaction among minor QTL, and between minor QTL and environment in addition to major effect QTL determine the expression of gluten strength. Patil et al. (2009) reported three major effect QTL located on chromosome 1B in proximity to glutenin coding loci *Glu-B1*, *Glu-B2*, and *Glu-B3* along with seven epistatic QTL distributed on six chromosomes (1A, 1B, 4A, 5B, 6A, and 7A) involved in four digenic epistatic interactions (Q × Q). QTL × environment (Q × E) interactions also contributed to the variation in gluten strength (Patil et al., 2009). However, a recent study (Kumar et al., 2013) identified only one QTL consistently expressed across three environments on chromosome 1BS explaining up to 90% of the phenotypic variation and no Q x Q or Q x E interactions were observed. The differences in these studies, at least in part, could result from the different genetic background of the mapping populations. Haplotype-trait association analysis detected five loci associated with gluten index on chromosomes 1A, 1B, 2B, 4B, and 7A with the locus on 4B explaining the highest amount of phenotypic variation in 192 Canadian durum wheat breeding lines (N'Diaye et al., 2018).

Reconstituting gluten strength to current market standards during genetic improvement for other traits is difficult due to the complex quantitative nature and the environmental effect on the expression of the trait. Therefore, molecular markers closely associated with QTL underlying gluten strength are of great value for developing marker-assisted selection in the durum breeding programs. In this study, we aimed to characterize genetic components controlling gluten strength in Canadian durum wheat. Along with identification of QTL, the epistatic interaction among QTL and the interactions between QTL and environmental factors, and putative candidate genes are also reported. The findings here will facilitate the marker assisted breeding for gluten strength in durum wheat.

MATERIALS AND METHODS

Population and Field Trials

A durum wheat population of 162 doubled haploid (DH) lines developed with the maize pollen method (Humphreys and Knox, 2015) and derived from Pelissier × Strongfield segregating for gluten strength was used in this study. Strongfield, selected from the cross AC Avonlea/DT665, is a registered Canada Western Amber Durum variety with strong gluten and low cadmium, developed at the Agriculture and Agri-Food Canada-Swift Current Research and Development Centre, Swift Current, SK (Clarke et al., 2005). Pelissier, a founder influencing the Canadian durum wheat gene pool, is a variety introduced from North Africa in 1929 (Dexter, 2008). It has high cadmium and lipoxygenase. The DH lines, along with their parents and controls were grown in field trials during year 2014, 2015, and 2016. Experiment was conducted as a randomized complete block design with two replicates at each of two seeding dates (early, E; late, L) and 1 week interval between two seeding dates per year. The field trial of each seeding date was grown at the different locations near Swift Current, SK, Canada. For phenotypic data analysis and QTL mapping, each different seeding date in each year was considered as one environment providing a total of six environments labeled as E14, L14, E15, L15, E16, and L16. Pre-plant soil testing was conducted each year to determine the rate of fertilizer application. The fertilizers were applied to target 112 kg ha^{-1} nitrogen, 67 kg ha^{-1} phosphorus, and 22 kg ha^{-1} sulfur. The soil is naturally high in potassium and did not require additional application.

Gluten Strength Measurement

The seeds harvested from each replicate of each seeding date were subjected to gluten strength measurement. Therefore, a total of four replicates of samples from each year/location over 3 years were analyzed. The durum whole grain samples were ground on an Udy mill with 1-mm screen at 13% moisture basis. The gluten strength was determined on 2.5 g samples of whole grain flour samples using the SDS-sedimentation volume (SV) method of Dick and Quick (1983) as modified by Agriculture and Agri-Food Canada (AAFC) with the addition of 25 ml of distilled water and 25 ml SDS solution to each sample. The higher the SV, the stronger the gluten.

Statistical Analysis and Biplot Analysis of Genotype-by-Environment Interaction

Pairwise phenotypic correlations were calculated using the Pearson's correlation coefficient in the R package Hmisc (version 4.2-0, http://cran.r-project.org/web/packages/Hmisc/index.html).

Analysis of variance (ANOVA) was performed using the PROC MIXED procedure of SAS 9.3 (SAS Institute, Cary, NC, USA). In the mixed model, lines were considered as fixed effects, and years, seeding dates, line × year interactions, line × seeding date interactions, line × year × seeding date interactions, seeding dates nested in years, and replications nested in years and

seeding dates were considered as random effects. The heritability of SV was calculated as the ratio of the genetic variance and the phenotypic variance across years using $\sigma_g^2/(\sigma_g^2 + \sigma_{gy}^2/y + \sigma_{gs}^2/s + \sigma_{gys}^2/ys + \sigma_\epsilon^2/rys)$, where σ_g^2, σ_{gy}^2, σ_{gs}^2, σ_{gys}^2, and σ_ϵ^2 were estimates of line, line × year interaction, line × seeding date interaction, line × year × seeding date interaction, and residual variance, respectively, and y, s, and r represented the numbers of year, seeding date, and replication, respectively, The repeatability of SV was calculated as the ratio of the genetic variance and the phenotypic variance of individual year using $\sigma_g^2/(\sigma_g^2 + \sigma_{gs}^2/s + \sigma_\epsilon^2/rs)$, where σ_g^2, σ_{gs}^2, and σ_ϵ^2 were estimates of line, line × seeding date interaction and residual variance, respectively, and s and r represented the numbers of seeding date and replication. For the estimations of the heritability and repeatability, all effects were considered random.

Biplot analysis of genotype-by-environment interaction was performed with the GGEBiplotGUI R (R version 3.0.3) package (Frutos et al., 2014). The analysis was based on a "tester-centered (G + GE)" table and row metric preserving, without any scaling.

Genotyping and Genetic Map Construction

DNA was extracted from leaves of 2-week-old seedlings of DH lines and parents using the AutoGenprep 965 (AutoGen Inc, Holliston, MA). The Infinium iSelect Wheat 90K SNP chip was used for genotyping according to the manufacturer's protocols (Illumina). Single nucleotide polymorphism (SNP) allele clustering and genotype calling was performed with GenomeStudio v2011.1 as described by Cavanagh et al. (2013). The default clustering algorithm implemented in GenomeStudio was first used to identify assays that produced three distinct clusters expected for bi-allelic SNPs. Manual curation was performed for assays that produced compressed SNP allele clusters that could not be discriminated by the default algorithm. The accuracy and robustness of SNP clustering was visually validated. SNPs with poor clustering quality, more than 30% missing data, or segregation distortion of more than 0.35 were removed. Redundant SNPs were also removed in R/qtl (Broman et al., 2003).

A total of 1,212 polymorphic SNP markers were used for genetic map construction in MapDisto version 2.0 software (Heffelfinger et al., 2017). Markers were classified into linkage groups based on a logarithm of odds (LOD) score threshold of 7.0 and recombination of 0.3. Genetic distances in cM were estimated using Kosambi's mapping function. Markers within each group were ordered using the AutoOrder command with the Seriation II method. The marker order was refined using CheckInversion and Ripple command with the sum of adjacent recombination frequencies (SARF) option. Markers showing double recombination events were re-scored. Markers detected with genotyping errors were replaced by missing values. All calculations were repeated for new linkage groups. The markers were distributed over 25 linkage groups (LGs). LGs were assigned to chromosomes based on comparison with an existing high-density SNP-based consensus map of durum wheat (Maccaferri et al., 2015). Parents were genotyped with the published molecular markers that discriminate glutenin and gliadin to test if they are polymorphic at these loci and to facilitate the comparative mapping.

Quantitative Trait Locus Mapping

Each different seeding date in each year was considered as one environment. Mean values for the trait from two replicates in each environment were used for the detection of QTL. Outliers of trait values were detected and removed using a Z-score transformation with a threshold of 3. QTL detection was performed using composite interval mapping (CIM) in WinQTL Cartographer v.2.5 software (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm; Wang et al., 2012). A walking speed of 1 cM was used. Forward regression was used for the selection of the markers to control the genetic background (control markers or cofactors) with up to five control markers. A window size of 10 cM was used to exclude closely linked control markers at the testing site. The LOD threshold at a significance level of 0.05 for declaring statistically significant QTL was calculated by 1,000 permutations. The additive effect (a) and phenotypic variance explained by each QTL (R^2) were estimated by CIM. The identified QTL (LOD > threshold) were automatically localized with the following parameters: minimal space between peaks = 30 cM; and minimum LOD from top to valley = 1.4. QTL detected in different environments were considered to be the same if the confidence intervals overlapped and the trait enhancing allele was contributed by the same parent.

The digenic epistatic interactions among all pairwise combinations of QTL were analyzed with multiple interval mapping (MIM) in the WinQTL Cartographer v.2.5 software. The initial QTL model was set using the CIM results obtained in each environment. The QTL model was progressively refined by searching and testing QTL or epistasis, and re-estimating. Both main additive effects of QTL and their epistatic interactions were tested for significance using the Bayesian information criterion (BIC). Not only main QTL (QTL with statistically significant main effect) and interactions among main QTL, but epistatic QTL (QTL that has no or small main effect but statistically significant interaction effect with another QTL) interacting with the main QTL were searched.

Multiple-trait composite interval mapping (Mt-CIM) implemented in WinQTL Cartographer v.2.5 was used to test for the presence of Q × E interaction at the main chromosome regions affecting the target trait (Maccaferri et al., 2008; Wang et al., 2012). The value of the trait in each environment was treated as a separate trait for the common genotypes. The G × E (H4) hypothesis was tested. All reported QTL were designated according to the Recommended Rules for Gene Symbolization in Wheat (http://wheat.pw.usda.gov).

Comparative Mapping and Projection of Quantitative Trait Locus Markers Onto the Durum Wheat Consensus Genetic Map and Onto the Reference Genomes of Durum and Wild Emmer Wheat

QTL reported in the literature and identified in this study were projected onto the durum high-density consensus genetic map

developed by Maccaferri et al. (2015) which includes SNP, simple sequence repeat (SSR), and diversity array technology (DArT) markers by projecting either a single marker near the QTL peak position or a pair of flanking markers within the QTL interval. The genetic linkage map and the QTLs were drawn using MapChart (version 2.3) software (Voorrips, 2002). Pairwise Spearman's rank correlation was performed in R version 3.3.2 to compare the collinearity of the marker order on the chromosomes of the durum consensus map and the genetic map generated in this study.

The sequences of the 90K SNPs were downloaded from the Kansas State University SNP marker database (http://wheatgenomics.plantpath.ksu.edu/snp/). Sequences of SSR markers were retrieved from the GrainGenes database (https://wheat.pw.usda.gov/GG3/). Sequences of DArT markers were downloaded from https://www.diversityarrays.com/technology-and-resources/sequences. Physical map positions of SNP, SSR, and DArT markers on genomes of durum wheat cv. Svevo (Maccaferri et al., 2019) and wild emmer wheat accession Zavitan (Avni et al., 2017) were aligned using the BLAST from Durum Wheat Genome Database (http://d-data.interomics.eu) and GrainGenes database (https://wheat.pw.usda.gov/GG3/wildemmer_blast). QTL markers on the physical map of durum wheat cv. Svevo and wild emmer wheat accession Zavitan were drawn with PhenoGram software (http://visualization.ritchielab.org/phenograms/plot).

Development of Kompetitive Allele Specific PCR Markers

Firstly, several SNPs in the interval of each of the target QTL were tested for 22 DH lines plus parents using the Kompetitive Allele Specific PCR (KASP) primers available for the Infinium iSelect Wheat 90K SNP chip (http://www.polymarker.info/designed_primers). Then, two closest KASP markers to each target QTL were used to genotype the population. The KASP assays were performed as described by Rasheed et al. (2016).

RESULTS

Phenotypic Variation Among Doubled Haploid Population

The gluten strength of the DH population was measured using SV. The summary statistics including mean SV values, range [minimum and maximum values, and standard deviation (SD)] are shown in **Figure 1**. Strongfield had significant higher SV value than Pelissier in all environments except E15 and E16 (**Figure 1** and **Table S1**). The population had the highest mean value in environment L15 (mean = 34.0) and the lowest in L16 (mean = 23.6) which reflects the environmental effect on gluten strength. Nevertheless, except in year 2015, no significant difference was observed for the mean SV of the population between two seeding dates. Although seeding date had no significant effect, the interaction of line by year and by seeding date was significant (**Table S2**). SV showed high Pearson's correlations among DH lines across environments ranging from r = 0.85 to 0.92 (**Figure S1**). The population had the largest phenotypic variation in environment L14, as indicated by the highest standard deviation (SD) and coefficient of variation

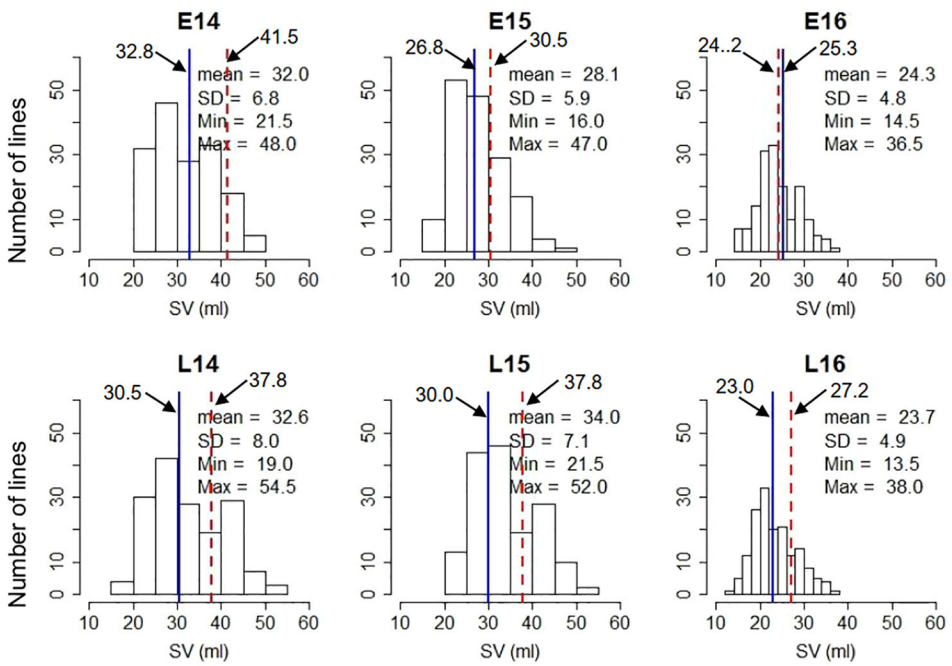


FIGURE 1 | Frequency distribution of SDS-sedimentation volume (SV) in the Pelissier × Strongfield population from 2014 to 2016 field trials with two seeding dates in each year and two replicates at each seeding date. Top panel, early seeding date; bottom panel, late seeding date. The blue solid line represents Pelissier; the red dashed line represents Strongfield; the mean SV of parents in each seeding date were shown; SD, standard deviation; Max, maximum; Min, minimum.

(CV), and the least variation in environment E16. Individual DH lines displayed bi-directional transgressive segregation, as shown by the maximum and minimum values relative to the parents (**Figure 1**). The transgressive segregation might result from the recombination of favorable or deleterious additive alleles from the parents, epistatic interactions of two genes, or any combinations of these mechanisms. The lines carrying favorable alleles from both parents showed higher SV than parent Strongfield, while the lines with the trait decreasing alleles from both parents had lower SV than Pelissier.

The percentages of GGE (Genotypic main effect plus Genotype-by-Environment interaction) explained by the first principal component was 90.4% and second principal component was 8.8% (**Figures 2A, B**). The DH lines were ranked based on both mean performance and stability across environments. The single arrowed line in **Figure 2A** points to higher mean SV value across environments. Therefore, line 162 had the highest mean SV value, while line 15 had the lowest mean SV value. The AEC (average-environment coordination) ordinate (dashed line) points to the greater variability (poor stability) in either direction. Thus, lines 90 and 93 were the most stable lines across environments (**Figure 2A**). The position of the ideal genotype which has the highest performance in all environments, is indicated by the arrow in **Figure 2B**. The DH lines located closer to the ideal genotype are more desirable than others. Taken into account both mean SV and stability, line 93 was the most desirable genotype.

Significant weak to moderate positive correlation between SV and GPC was observed in three out of six environments ($r = 0.3$–0.36) in this population (**Figure S2A**). In year 2014, significant correlation was shown in both seeding dates when the population had lower GPC compared to other field years. No significant correlation existed for any seeding date in year 2015. Similarly, significant weak to moderate negative correlation between SV and grain yield (GY) was also observed in the same three out of six environments [$r = -0.334$–(-0.216)] (**Figure S2B**). Significant negative correlation was displayed in both seeding dates in year 2014 when higher GY was obtained. While no significant correlation was observed in the year 2015 with lower GY.

Quantitative Trait Locus Mapping by Composite Interval Mapping in Single Environments

ANOVA (**Table S2**) indicates that genotype by year and by seeding date interaction had significant effect on the SV. Therefore, QTL analysis was first performed for SV in each environment. Variable numbers of significant QTL, from two to five, were detected in each environment. Globally, the largest number of QTL (5) was detected in environment L14 and L15. A total of nine different QTL were detected across environments, four of which were specific for a single environment (**Table 1**). Both parental lines contributed the favorable alleles depending on the QTL (2 by Strongfield and 7 by Pelissier). A major QTL on chromosome 1B (*QGlu.spa-1B.1*) explaining up to 40.1% of the phenotypic variance (R^2) and a second major QTL on chromosome 1A (*QGlu.spa-1A*) explaining up to 18.7% of the phenotypic variance, were detected across all environments with high SV allele derived from Strongfield. Two minor QTL, *QGlu.spa-1B.2* and *QGlu.spa-1B.3* with R^2 values of 4.1% and 6.3%, were also detected on chromosome 1B but only in a single environment. Additionally, three QTL on chromosome 2B and two QTL on chromosome 3A were detected with R^2 values ranging from 3.4 to 8.9%. The QTL *QGlu.spa-3A.1* and *QGlu.spa-3A.2* were repeatedly detected in at least two environments. Except *QGlu.spa-1A* and *QGlu.spa-1B.1*, no other minor QTL was detected in L16. Pelissier contributed trait-enhancing alleles to all minor QTL.

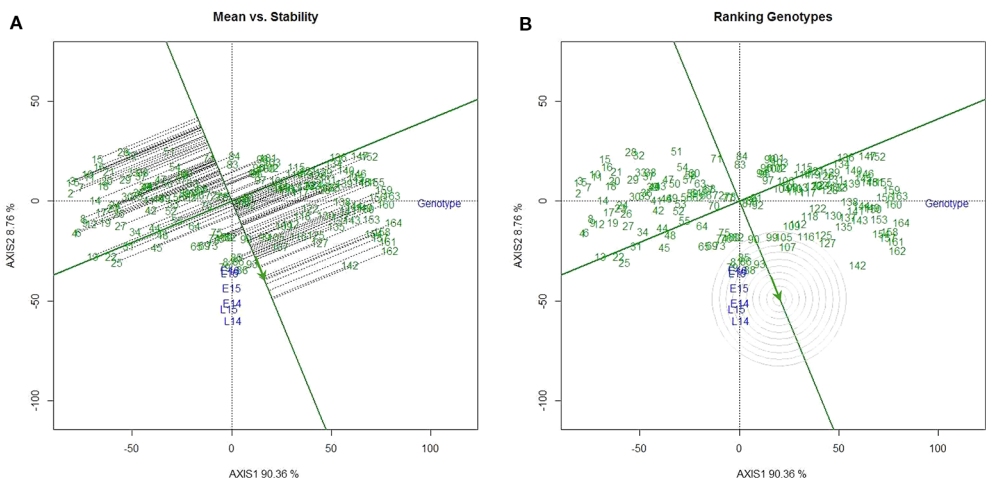


FIGURE 2 | GGE (Genotypic main effect plus Genotype-by-Environment interaction) analysis of SDS-sedimentation volume (SV) in DH lines of Pelissier × Strongfield. **(A)** Average-environment coordination (AEC) view of the GGE biplot; the single-arrowed line is the AEC abscissa (or AEA) and points to the higher mean SV value across environments. **(B)** Ranking doubled haploid (DH) lines relative to the ideal genotype (a genotype that GGE predicted has the best performance across environments for SV) on SV performance. The arrow is where an ideal genotype should be. The DH lines located closer to the ideal genotype are more desirable than others.

TABLE 1 | Overview of quantitative trait loci (QTL) identified for SDS-sedimentation volume (SV) across six environments.

Chr[a]	QTL	Env[b]	Peak marker	LOD	Additive[c]	R^2(%)[d]	Interval (two LOD drop)
1A	*QGlu.spa-1A*	E14	wsnp_Ex_c13186_20822127	16.1	3.00	18.7	BS00088136_51 - Kukri_c10405_1277
		E15	wsnp_Ex_c13186_20822127	15.1	2.49	17.6	IAAV1142 - RAC875_c31031_387
		E16	wsnp_Ex_c13186_20822127	10.7	1.83	13.7	IAAV1142 - RAC875_c31031_387
		L14	wsnp_Ex_c13186_20822127	14.3	3.09	14.6	BS00088136_51 - RAC875_c31031_387
		L15	wsnp_Ex_c13186_20822127	18.1	3.14	18.2	IAAV1142 - RAC875_c31031_387
		L16	wsnp_Ex_c13186_20822127	10.8	2.03	16.9	IAAV1142 - RAC875_c31031_387
1B	*QGlu.spa-1B.1*	E14	Kukri_c38353_67	26.5	4.11	35.6	BS00085235_51 - RCA875_rep_c74067_541
		E15	Kukri_c38353_67	25.6	3.48	34.6	BS00085235_51 - RCA875_rep_c74067_541
		E16	Kukri_c38353_67	27.2	3.01	35.1	BS00085235_51 - RCA875_rep_c74067_541
		L14	Kukri_c38353_67	22.2	4.71	25.4	BS00085235_51 - RCA875_rep_c74067_541
		L15	Kukri_c38353_67	32.7	4.60	40.1	BS00085235_51 - RCA875_rep_c74067_541
		L16	Kukri_c38353_67	20.4	2.98	36.5	BS00085235_51 - RCA875_rep_c74067_541
	QGlu.spa-1B.2	L14	Excalibur_c50079_420	5.4	-1.80	4.1	Ku_c241_460 - BS00078029_51
	QGlu.spa-1B.3	L15	BS00067436_51	7.4	-1.91	6.3	Tdurum_contig7449_800 - RAC875_c47427_235
2B	*QGlu.spa-2B.1*	L15	RAC875_c38003_164	4.2	-1.37	3.4	Excalibur_c19499_948 - D_F5XZDLF01CFO7W_135
	QGlu.spa-2B.2	E16	Kukri_c25868_56	7.3	-1.44	8.7	Kukri_c25868_56 - Ex_c55735_1012
	QGlu.spa-2B.3	L14	Excalibur_c91034_141	5.5	-1.79	5.6	Excalibur_c33221_681 - CAP7_6910_523
3A	*QGlu.spa-3A.1*	E14	RAC875_c64107_404	4.1	-1.41	4.0	RAC875_c64107_404 - BS00021981_51
		L15	RAC875_c64107_404	4.4	-1.40	3.6	RAC875_c64107_404 - BS00021981_51
	QGlu.spa-3A.2	E15	Excalibur_c14216_692	5.6	-1.46	5.9	Tdurum_contig98188_239 - RAC875_c775_1264
		E16	wsnp_Ex_rep_c69864_68824236	6.7	-1.42	8.3	CAP_c3367_68 - Tdurum_contig98188_239
		L14	Excalibur_c14216_692	9.7	-2.45	8.9	Tdurum_contig98188_239 - RAC875_c775_1264

[a]*Chromosome;* [b]*environment.*
[c]*Additive effect, the positive values indicate that the alleles from Strongfield have the effect of increasing the trait value.*
[d]R^2 *is the percentage of phenotypic variation explained by each QTL.*

Multi-Environment Quantitative Trait Loci Analysis

Multi-environment QTL analysis was performed to detect the significant QTL across environments and Q × E effect (**Figure 3**). Two significant QTL *QGlu.spa-1A* and *QGlu.spa-1B.1* were detected by multi-environment analysis which agreed with single-environment analysis. In these analyses, environment L15 influenced joint analysis the most for QTL *QGlu.spa-1A* while environment L14 influenced joint analysis the most for QTL *QGlu.spa-1B.1*. However, environment L14 influenced SV the least for QTL *QGlu.spa-1A*. Both *QGlu.spa-1A* and *QGlu.spa-1B.1* were statistically significant for multi-environment joint analysis. Although the Q × E effect was significant, these two QTL were stably expressed across all environments. The QTL mapped on other chromosomes in only one or two environments did not reach the significance threshold value in the multi-environment QTL analysis. Notably, the Q × E effect observed was due to the difference of the effect in magnitude and not the direction of QTL. This was also evidenced by the consistent sign of the effects of QTL detected across environments (**Table 1**).

Implication of Quantitative Trait Loci on 1A and 1B and Development of Kompetitive Allele Specific PCR Markers

Figure 4A is a graphical illustration of the chromosome 1A region harboring QTL *QGlu.spa-1A* of a selection of 20 DH

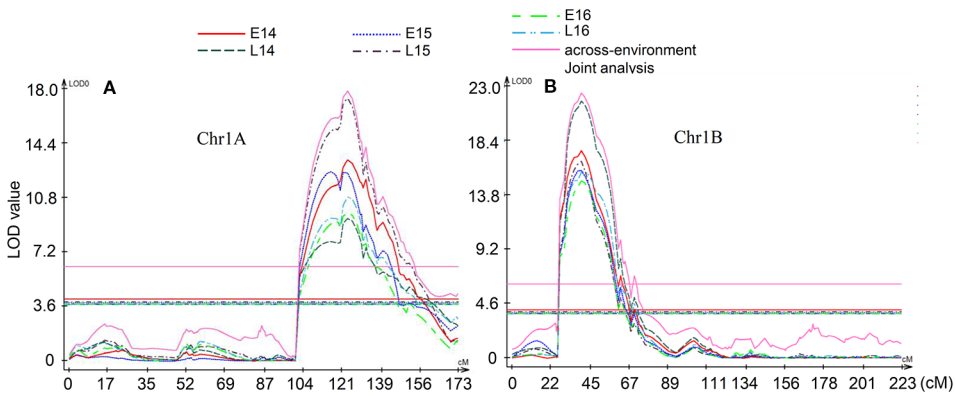


FIGURE 3 | Multi-environment quantitative trait loci (QTL) analysis. **(A)** QTL *QGlu.spa-1A* on chromosome 1A, **(B)** QTL *QGlu.spa-1B.1* on chromosome 1B. The horizontal lines indicate logarithm of odds (LOD) significance threshold determined by 1,000 permutations at a significance level of 0.05.

genotypes and **Figure 4B** of the chromosome 1B region containing *QGlu.spa-1B.1* for the same genotypes with high and low SV values. On both chromosome 1A and 1B, the Strongfield alleles occurred in high SV genotypes, whereas the Pelissier alleles contributed to the low SV lines. This agreed with QTL analysis results (**Table 1**). Colored fragments along the chromosome region outside the green line of the peak marker, refer to the loci belonging to other traits may be not related to SV. Based on the genotypes of two flanking markers in the QTL region of *QGlu.spa-1A* (IAAV1142 and RAC875_c31031_387) and *QGlu.spa-1B.1* (Kukri c38553_67 and RCA875_rep_c74067_541), the DH lines were separated into two groups with significantly different means (*t* test, $p < 10^{-4}$) between the two groups (**Figures 5A, B**). More distinct separation was shown for the *QGlu.spa-1B.1* than *QGlu.spa-1A*, which is in agreement with a larger portion of the phenotypic variation explained by *QGlu.spa-1B.1* (**Table 1**). Based on the genotypes of the flanking markers of both aforementioned QTL combined, two main groups with clearer separation was observed within the population: one group of DH lines with high SV value (strong gluten) having flanking marker alleles from Strongfield and the other group of lines with low SV value (weak gluten) carrying flanking marker alleles from Pelissier (**Figure 5C**).

Two KASP assays were developed for each of QTL, *QGlu.spa-1A* (wsnp_Ex_c13186_20822127 and IAAV1142), and *QGlu.spa-1B.1* (RAC875_rep_c74067_541 and Kukri_c38553_67). All KASP assays were validated against the SV values of the population in each environment (**Figure 6**). In all cases, the genotypes carrying Strongfield allele had significantly higher SV than those with Pelissier allele.

Combined Haplotype Analysis Across Multiple Quantitative Trait Loci

To investigate the accumulated effects of the favorable alleles on SV across multiple QTL, the combined haplotype analysis was performed on QTL detected in two or more environments, *QGlu.spa-1A*, *QGlu.spa-1B.1*, *QGlu.spa-3A.1*, and *QGlu.spa-3A.2*. The SNPs in the two LOD interval of each QTL were used for haplotype analysis. A total of 11 different haplotypes (Hap1–Hap11) were identified at different frequencies, with each haplotype containing three or more DH lines (**Figure 7**). The DH lines with Hap2 has the best combination of all favorable alleles at each QTL, as evidenced by the highest mean SV across all environments. The most desirable genotype, line 93, has this haplotype. While the lines with Hap10 has the least favorable combination of the alleles from each QTL. Significant difference was observed for SV in these two haplotype groups across all environments. Significant difference in SV between Hap1 and Hap8 across all environments agreed with the effect of the *QGlu.spa-1A*. Likewise, the significant difference in SV between Hap1 and Hap4, Hap2 and Hap7, Hap8 and Hap10, confirmed the effect of major QTL *QGlu.spa-1B.1*. Except in E16, no significant difference was observed between Hap1 and Hap2. This is not surprising given the environment specific expression and minor effect of *QGlu.spa-3A.1* and *QGlu.spa-3A.2*.

Identification of Epistatic Interaction of Quantitative Trait Loci

Multiple interval mapping (MIM) has been used for mapping multiple QTL with epistasis (Laurie et al., 2014). In this study MIM was used for identification of digenic epistatic interactions

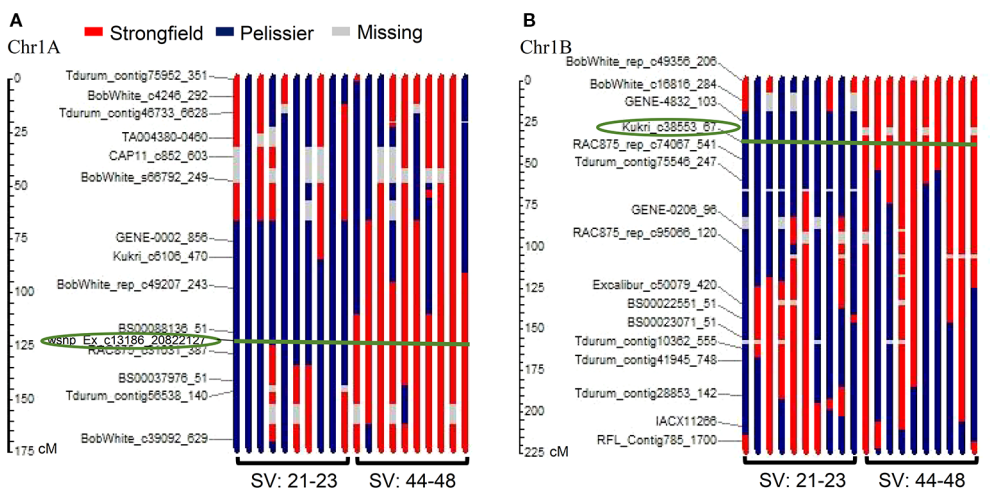


FIGURE 4 | Graphical illustration of genotypes of 20 doubled haploid (DH) lines with recombination pattern of major quantitative trait loci (QTL) **(A)** *QGlu.spa-1A* on chromosome 1A and **(B)** *QGlu.spa-1B.1* on chromosome 1B. Circled marker is the peak marker in the QTL region and the green line indicates the peak marker position on each genotype. The blue bar represents the fragment derived from weak gluten strength parent Pelissier and the red bar represents the fragment derived from strong gluten strength parent Strongfield. SV, SDS-sedimentation volume.

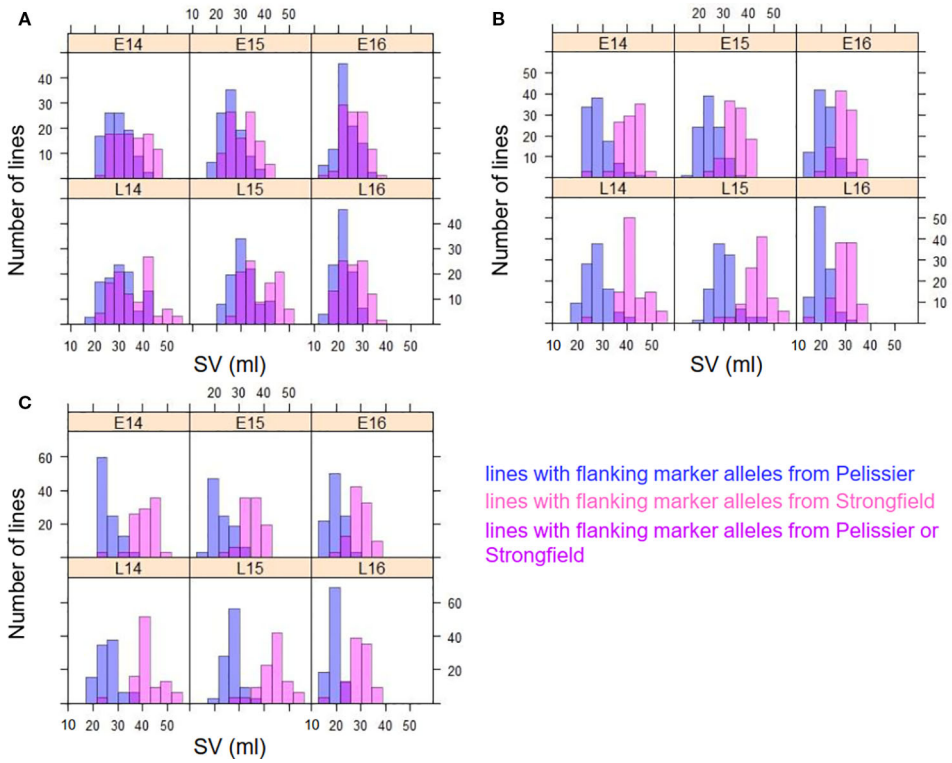


FIGURE 5 | Frequency distribution of SDS-sedimentation volume (SV) in two groups of DH lines separated on the genotype of two flanking markers of **(A)** *QGlu.spa-1A* (IAAV1142 and RAC875_c31031_387) on chromosome 1A, **(B)** *QGlu.spa-1B.1* (Kukri c38553_67 and RCA875_rep_c74067_541) on chromosome 1B, and **(C)** both *QGlu.spa-1A* on chromosome 1A and *QGlu.spa-1B.1* on chromosome 1B across six environments (field year 2014–2016 with two seeding dates in each year. early, E; late, L). The blue bar represents the lines with alleles from weak gluten strength parent Pelissier, the pink bar represents the lines with alleles from strong gluten strength parent Strongfield, and the magenta bar represents the lines with alleles either from Strongfield or Pelissier.

among all pairwise combinations of QTL. Compared with the results obtained by CIM analysis, two additional significant QTL, *QGlu.spa-1B.4* in L14, and *QGlu.spa-5A* in L15, were detected with R^2 values of 9.4% and 2.3% respectively (**Table 2**). The epistatic interactions detected together with their average effects and R^2 values are reported in **Table 2**. A total of 11 pairwise QTL interactions (additive × additive) were detected in different environments at a significance level of 0.05. Of note, the epistatic effect between the two major QTL *QGlu.spa-1A* and *QGlu.spa-1B.1* was detected in four out of six environments with R^2 values of 1.4–1.9%. The other 10 pairs of environment-specific interactions with R^2 values of 0.8–3.7% were detected only in a single environment. Not only interactions among main QTL (QTL with statistically significant main effect) but also epistatic QTL, QTL that has no or small main additive effect but statistically significant interaction effects with another QTL, interacting with main QTL were identified. It is interesting to note that the additive effect of *QGlu.spa-5B was* too small to reach the genome-wide significance level in CIM scans but it had significant interaction with identified QTL *QGlu.spa-1B.4* and *QGlu.spa-2B.3* in environment L14, as well as with QTL *QGlu.spa-2B.4* in L16. Likewise, epistatic QTL *QGlu.spa-6B* had significant interaction with other QTL in two environments, E14 and E16, but no main effect.

Comparison With Previously Reported Quantitative Trait Loci

The marker order in the genetic map generated in this study was highly collinear with the durum consensus map developed by Maccaferri et al. (2015), as indicated by the Pairwise Spearman's rank correlations (r = 0.992–0.999) (**Figure S4**). QTL reported for SV in the literature and identified in this study were projected onto the durum wheat consensus genetic map by projecting either a single marker near the QTL peak position or a pair of flanking markers within the QTL interval (**Table S3** and **Figure 8**). The QTL *QGlu.spa-1A* (peak marker: *wsnp_Ex_c13186_20822127*) detected in this study was projected at position 89.5 cM on chromosome 1A of the durum consensus genetic map. It is approximately 6 cM away from SSR marker *wmc312* reported to be associated with SV in durum wheat by Conti et al. (2011). The QTL *QGlu.spa-1B.1* (peak SNP: *Kukri_c38353_67*) was projected on the short arm of chromosome 1B on the durum wheat consensus

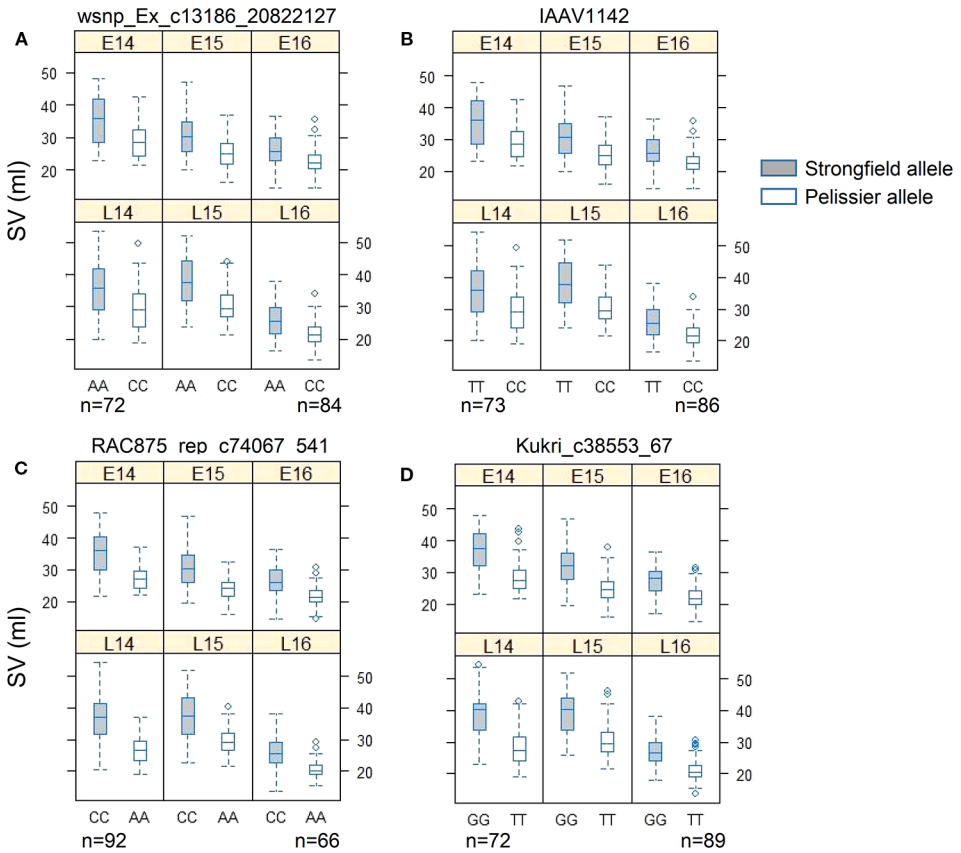


FIGURE 6 | Phenotypic validation of Kompetitive Allele Specific PCR (KASP) assays for **(A)** wsnp_Ex_c13186_20822127 and **(B)** IAAV1142 in the interval of *QGlu.spa-1A* on chromosome 1A, **(C)** RAC875_rep_c74067_541, and **(D)** Kukri_c38553_67 in the interval of *QGlu.spa-1B.1* on chromosome 1B across six environments (field year 2014–2016 with two seeding dates in each year. early, E; late, L). The *p* value of *t* test of two genotype groups for each marker in each environment is smaller than 0.001. SV, SDS-sedimentation volume.

map, approximately 4.4 cM away from SSR marker *gwm550* reported by Patil et al. (2009). Likewise, QTL *QGlu.spa-1B.2* (peak SNP: *Excalibur_c50079_420*) on the long arm of chromosome 1B was 6 cM away from the QTL interval (*barc181-psr162*) identified by Zhang et al. (2008) and Conti et al. (2011) in durum wheat, and 0.5 cM apart from the QTL (peak SNP: *CAP8_c818_370*) identified by Jernigan et al. (2018) in bread wheat. Of the two QTL reported by Roselló et al. (2018), QTL associated with marker *wPt-1140* was located within *QGlu.spa-2B.2* (peak SNP: *Kukri_c25868_56*) and another QTL associated with marker *wPt-6894* within *QGlu.spa-2B.3* (peak SNP: *Excalibur_c91034_141*). There have been no reports for QTL on the short arm of chromosome 2B close to *QGlu.spa-2B.1* (peak SNP: *RAC875_c38003_164*). A QTL on the short arm of chromosome 3A reported by Roselló et al. (2018) is about 32 cM away from *QGlu.spa-3A.1* (peak SNP: *RAC875_c64107_404*) identified in this study, indicating that these two QTL might be different and *QGlu.spa-3A.1* was a novel QTL. In addition, another QTL *QGlu.spa-3A.2* (peak SNP: *Excalibur_c14216_692* and *wsnp_Ex_rep_c69864_68824236*) on chromosome 3A was likely a novel QTL for SV.

Identification of Putative Candidate Genes for Major Quantitative Trait Loci

To predict the putative candidate genes at major QTL on chromosome 1A and 1B and to facilitate comparative mapping analysis, the sequences of the peak and flanking markers associated with QTL for SV were anchored to their physical location on the genome by aligning the marker sequence to the wild emmer wheat accession Zavitan (*Triticum dicoccoides*, WEWSeq_v.1.0; Avni et al., 2017) (**Table S4** and **Figure 9A**) and the durum wheat cv. Svevo assemblies (Maccaferri et al., 2019) (**Table S4** and **Figure 9B**). The gene content in the two LOD drop of QTL region, corresponding to a region of 10.75 Mb on 1A and 11.78 Mb on 1B, was searched.

Among the annotated high confidence genes, the gene *TRIDC1AG047310* in proximity to SNP *wsnp_Ex_c13186_20822127* (M6) on chromosome 1A of Zavitan, has five transcript splice variants. Three out of five splice variants encode a protein with three domains of HMW glutenin (**Figure S5**). *TRIDC1AG047310.3* encodes a protein with sequence similarity with *Glu-A1* (GenBank accession: ANJ03342) from wild emmer wheat accession TD-256 (*T.*

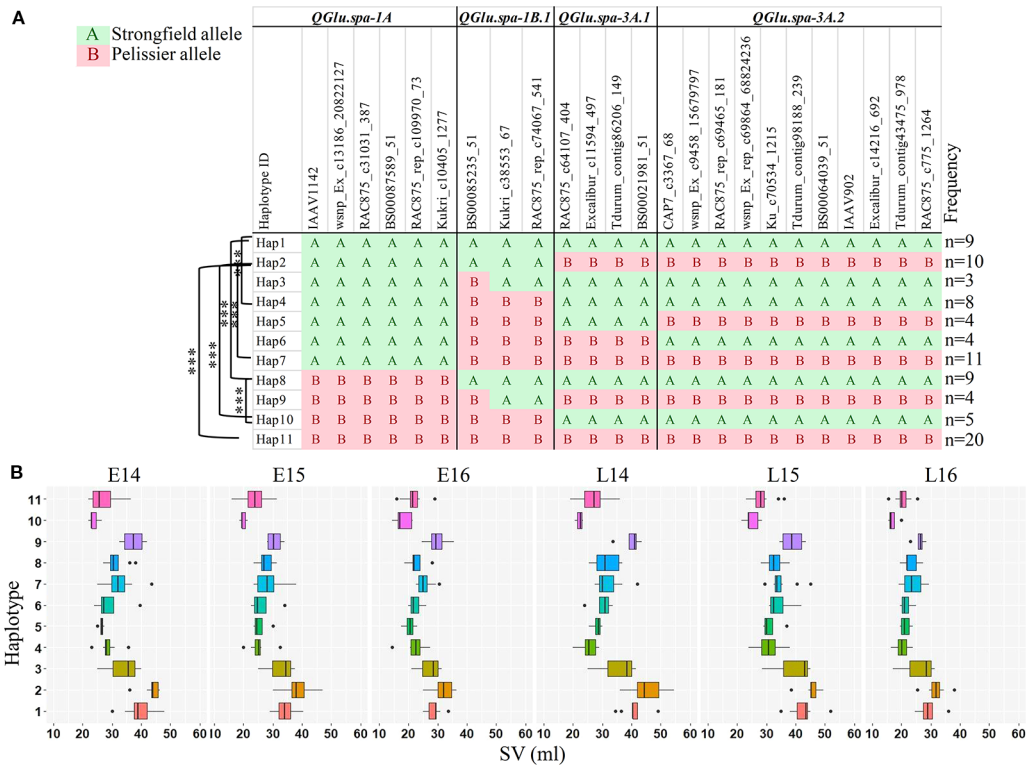


FIGURE 7 | Haplotype analysis across four quantitative trait loci (QTL) [two logarithm of odds (LOD) interval] which were identified in at least two environments. **(A)** Haplotype block based on single nucleotide polymorphism (SNP) markers in each QTL region. **(B)** Boxplots of the phenotype values corresponding to 11 different haplotype groups in each environment. Haplotypes containing less than three doubled haploid (DH) lines were omitted from the table. The DH lines with undetermined haplotype were not shown. Haplotypes were assigned using R package Haplotyper. ***, significant at $p < 0.001$ (*t* test). SV, SDS-sedimentation volume.

TABLE 2 | Epistatic interaction between quantitative trait loci (QTL).

QTL1[a]	QTL2	Environment	LOD	Effect[b]	R^2 (%)[c]	Empirical *p*-Value
QGlu.spa-1A	*QGlu.spa-1B.1*	E14	3.0	2.3	1.5	0.002
		E15	2.1	1.8	1.8	0.015
		L14	3.6	2.3	1.4	<0.0001
		L16	2.2	1.9	1.9	0.028
QGlu.spa-1A	*QGlu.spa-1B.2*	L16	3.2	3.5	3.0	<0.0001
QGlu.spa-1A	*QGlu.spa-1B.3*	L15	1.6	1.8	0.8	0.022
QGlu.spa-5A		E16	4.2	4.6	2.2	<0.0001
		L15	2.9	2.7	2.3	0.001
QGlu.spa-1B.1	*QGlu.spa-5A*	L15	1.8	1.8	1.0	0.02
QGlu.spa-1B.1	*QGlu.spa-6B**	E14	1.7	-1.7	1.2	0.036
QGlu.spa-1B.4		L14	7.1	2.0	9.4	<0.0001
QGlu.spa-1B.4	*QGlu.spa-5B**	L14	2.8	-2.0	1.5	<0.0001
*QGlu.spa-6B**		E14	0.3	0.4	0.5	0.777
		E16	0.5	0.6	0.5	0.477
QGlu.spa-2B.3	*QGlu.spa-5B**	L14	5.0	6.5	1.8	<0.0001
QGlu.spa-2B.3	*QGlu.spa-6B**	E16	3.6	4.7	2.2	<0.0001
*QGlu.spa-2B.3**	*QGlu.spa-2B.4**	L16	3.4	3.2	1.4	<0.0001
*QGlu.spa-2B.4**	*QGlu.spa-3A.2*	L16	2.8	3.4	3.7	<0.0001
*QGlu.spa-2B.4**	*QGlu.spa-5B**	L16	3.0	3.6	2.0	<0.0001

[a]*QTL1 and QTL2 are a pair of interacting QTL.*

[b]*Epistatic effect of QTL1 and QTL2.*

[c]R^2 *is the percentage of phenotypic variation explained by QTL or QTL epistasis.*

**Epistatic QTL, QTL that has no or small main additive effect but statistically significant interaction effect with another QTL.*

dicoccoides) and *Glu-1Ax1* (GenBank accession: CAA43331) from bread wheat (**Figure S6**). The HMW glutenin locus *Glu-A1* reported by Li et al. (2009) was projected on chromosome 1A where *TRIDC1AG047310* is located (**Figure 9A**). Similarly, SNP *wsnp_Ex_c13186_20822127* (M6) is about 8.66 Mb away from *Glul-A1* on chromosome 1A of durum wheat cv. Svevo (**Figure 9B**). This suggested that *TRIDC1AG047310* might be a candidate gene underlying QTL *QGlu.spa-1A*. In addition, three annotated high confidence genes (*TRITD1Av1G002310*, *TRITD1Av1G002360* and *TRITD1Av1G002790*) encode LMW-GS on the short arm of chromosome 1A of Svevo. However, no QTL was detected in this region in the present study.

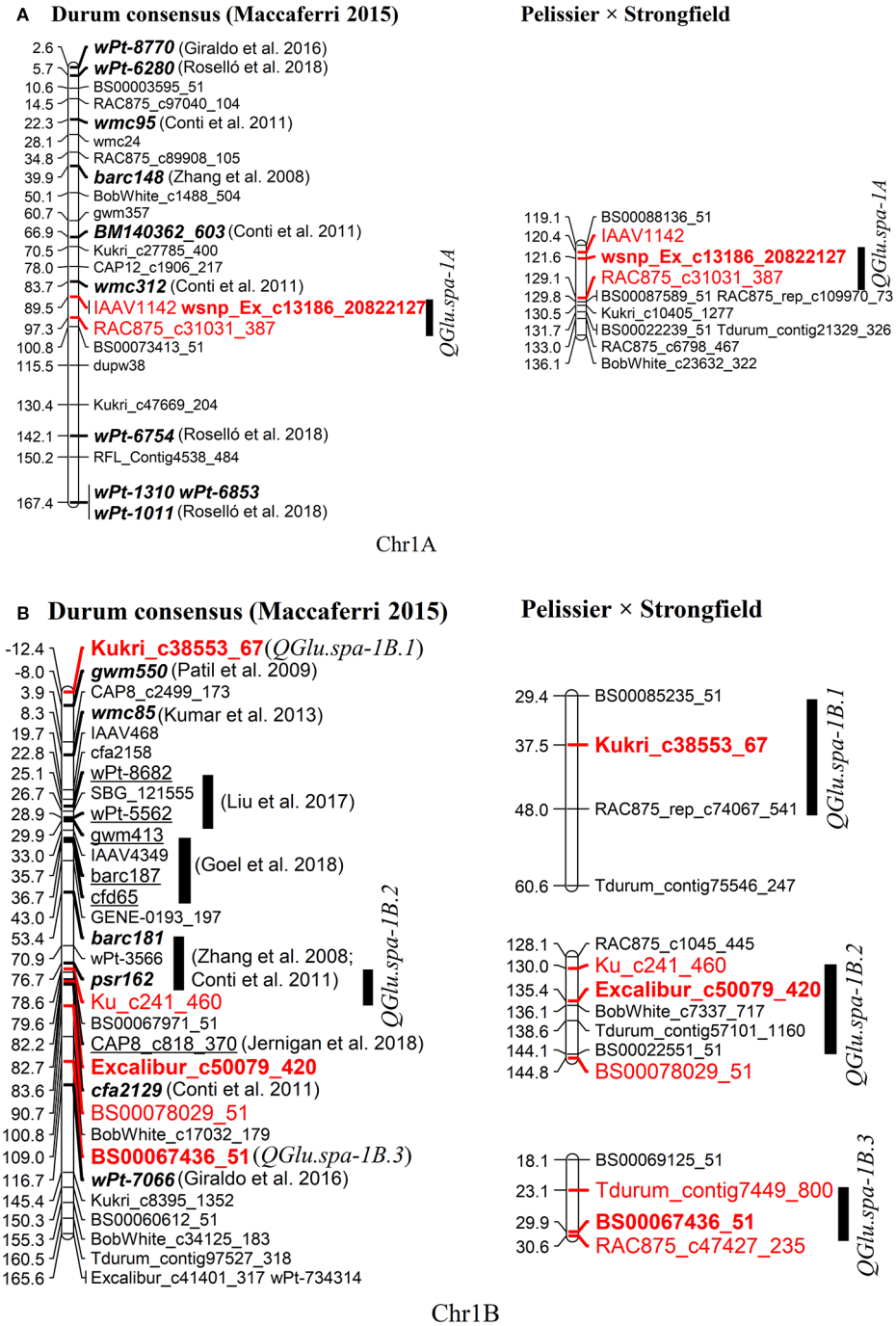


FIGURE 8 | Continued

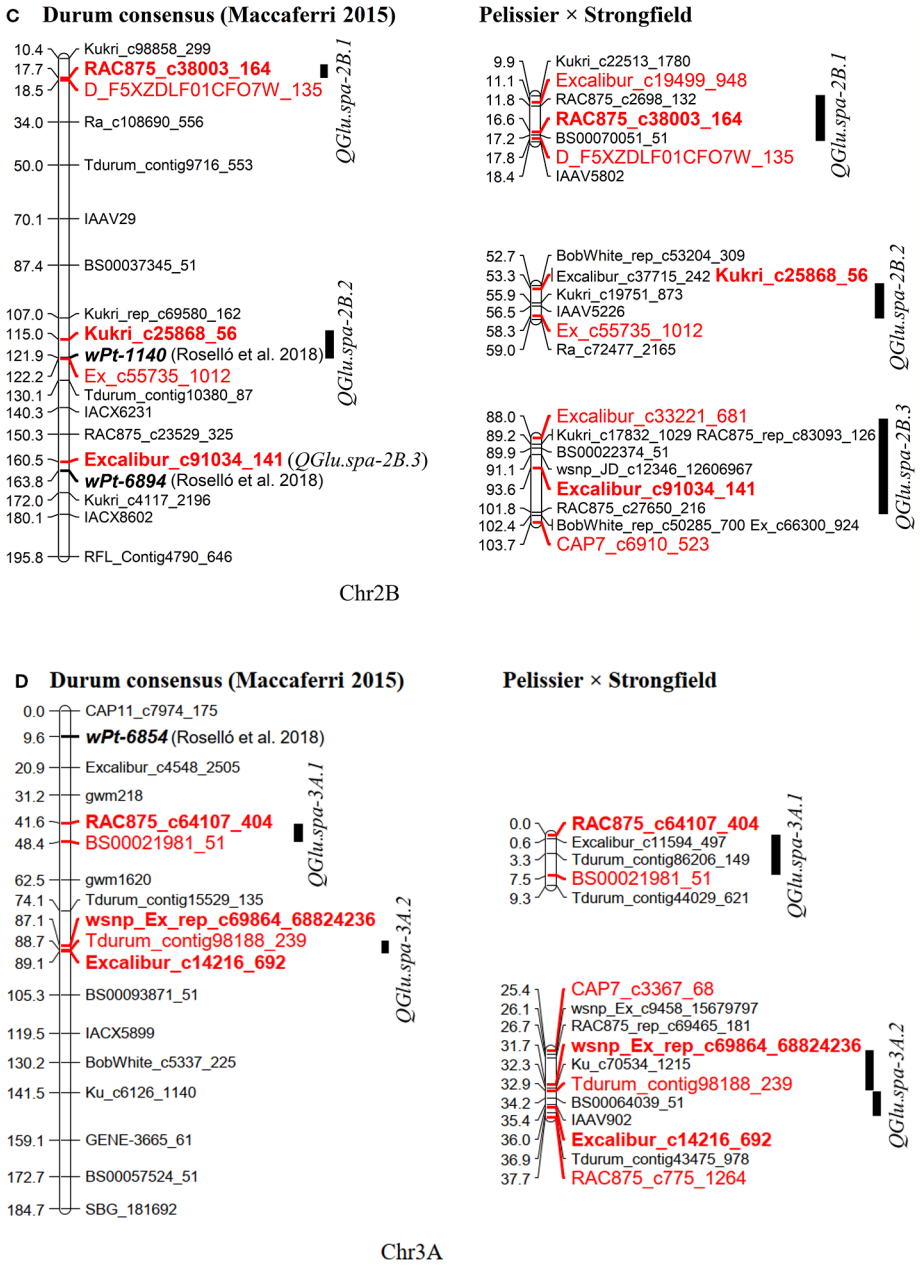


FIGURE 8 | Projection of quantitative trait loci (QTL) for SDS-sedimentation volume (SV) reported in the literature (in both bread wheat and durum wheat) and QTL identified in this study onto the durum wheat consensus genetic map developed by Maccaferri et al. (2015) (left). The QTL on the genetic map from this present study is shown on the right. **(A)** Chromosome 1A, **(B)** 1B, **(C)** 2B, and **(D)** 3A. The markers highlighted in red and bold are peak markers of QTL identified in this study and those highlighted in red are flanking markers in two LOD drop of interval; the markers in bold and italics were reported in durum wheat; the markers underlined were reported in bread wheat.

The gene *TRIDC1BG001970* on chromosome 1B of Zavitan, with a distance of 108 Kb from SNP *Kukri_c38553_67* (M12), has three domains of gliadin/LMW glutenin. Three transcript splice variants were identified for gene T*RIDC1BG001970* with *TRIDC1BG001970.2* encoding a protein of 298 aa and *TRIDC1BG001970.3* for a protein of 182 aa. The transcript *TRIDC1BG001970.1* encodes a protein with 139 aa without gliadin/LMW glutenin domain (**Figure S7**). Three paralogous genes of *TRIDC1BG001970* on chromosome 1B of Zavitan were identified as *TRIDC1BG001560* (1B: 7,719,618 -7,720,229 bp), *TRIDC1BG001740* (1B: 8,606,652–8,608,041 bp) and *TRIDC1BG004610* (1B: 20,578,493–20,579,630 bp) (**Figure 9A**). However, the protein structure of *TRIDC1BG001970* is more similar to that of *TRIDC1BG001740* (**Figure S8**).

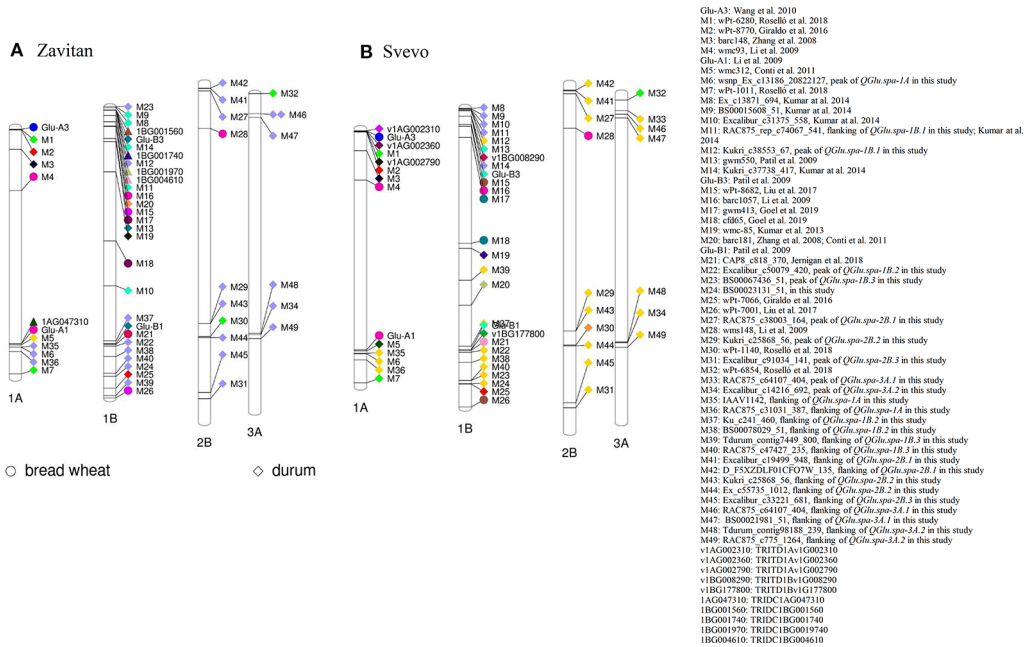


FIGURE 9 | Projection of quantitative trait loci (QTL) for SDS-sedimentation volume (SV) reported in the literature (in both bread wheat and durum wheat) and the QTL identified in this study onto the reference genomes of **(A)** wild emmer wheat accession Zavitan and **(B)** durum wheat cv. Svevo.

TRIDC1BG001970.2 shares 85% identity at the protein level with *Glu-B3* (GenBank # AVI69508.1) in durum cv. Langdon (**Figure S9**). On chromosome 1B of Svevo, gene *TRITD1Bv1G008290* (**Figure S9**) encoding a portion of LMW-GS is about 105 Kb away from the *Glu-B3* marker and 13.1 Mb from SNP *Kukri_c38553_67* (M12, *QGlu.spa-1B.1*). Similarly, gene *TRITD1Bv1G177800* encoding a HMW-GS is 186 Kb away from the *Glu-B1* marker and 5.88 Mb from SNP *Excalibur_c50079_420* (M22, *QGlu.spa-1B.2*) on the long arm of chromosome 1B of Svevo. This comparative mapping indicated that *TRITD1Bv1G008290* and *TRITD1Bv1G177800* are likely the candidate genes for *QGlu.spa-1B.1* and *QGlu.spa-1B.2* respectively, although the possible existence of other paralogs in these regions cannot be excluded.

DISCUSSION

Gluten strength is one of the most important quality criteria in durum breeding. Previous studies have shown that gluten strength of durum wheat is quantitatively controlled by a few major QTL and some minor QTL whose expression is affected by environmental conditions. In this study, a total of nine QTL were detected for gluten strength measured by SV. Two major QTL positioned on chromosome 1A and 1B were detected across all environments. These two QTL together accounted for up to 59% of the phenotypic variance. The present work also allowed the identification of a few minor QTL on chromosome 1B, 2B, and 3A, with inconsistent expression over different environments. Favorable alleles were identified from both parents at different loci.

Quantitative Trait Loci Associated With Low Molecular Weight Glutenin Subunit

The major QTL *QGlu.spa-1B.1* explaining up to 40.1% of the phenotypic variance in the present study was identified on the short arm of chromosome 1B close to the LMW-GS locus *Glu-B3*. This finding supports the possibility that allelic variation for LMW-GS encoded by the *Glu-B3* locus on chromosome 1BS is the major contributor for the difference of gluten strength in durum wheat (Pogna et al., 1990). A major QTL on 1BS near the *Glu-B3* locus has been previously reported in a variety of durum wheat germplasm (Blanco et al., 1998; Elouafi et al., 2000; Patil et al., 2009; Kumar et al., 2013; Kumar et al., 2014), which is indicative of the importance of the *Glu-B3* region for gluten strength of durum wheat although with various levels of expression in different genetic backgrounds and environments. Apart from a strong positive correlation between the LMW-GS *Glu-B3* locus and gluten strength, a strong association of LMW-GS with pasta-cooking quality has been well documented (Pogna et al., 1990; Kovacs et al., 1995b; Ruiz and Carrillo, 1995). Similarly, nine protein alleles of the *Glu-B3* (a, b, c, d, e, f, g, h, i) with various effect on dough quality have been reported in bread wheat (Metakovsky, 1990; Gupta et al., 1991).

Strongfield and Pelissier display different profiles of LMW-GS and HMW-GS (**Figure S3**). The analysis of allele-specific PCR markers showed that there is no polymorphism for gliadin alleles *GliB1.1* and *GliB1.2* between two parents, while polymorphism exists for LMW-glutenin alleles *gluB3c* and *gluB3i* (**Table S5**). The two characteristic subunits (39,623 and 42,930 Da in Strongfield; 39,627 and 42,906 Da in Pelissier) of Glu-B3c (Wang et al., 2015) showed differential ratio in the two parents

(**Figure S3**), which might result from the polymorphism of *gluB3c*. Combining with the variation of LMW-GS composition between parents, this indicated that *Glu-B3* might be associated with *QGlu.spa-1B.1* in this population. Furthermore, candidate gene analysis suggested that both *TRIDC1BG001970* and *TRIDC1BG001740* could be associated with QTL *QGlu.spa-1B.1*. The physical location of *TRIDC1BG001970* and *TRIDC1BG001740* is next to a major QTL (SNP: *Kukri_c37738_417*) contributing up to 90% of the phenotypic variation for gluten strength measured using SV in durum (Kumar et al., 2013; Kumar et al., 2014), followed by the *Glu-B3* locus on the physical sequence map of wild emmer accession Zavitan. High similarity of protein structure between the identified genes and *Glu-B3* suggested there could be a gene cluster in the *Glu-B3* region responsible for gluten strength, which is in agreement with the report that LMW-GS are encoded by multi-gene families at the *Glu-A3* and *Glu-B3* loci (D'Ovidio and Masci, 2004). Further studies are needed to clone and differentiate the functions of identified genes. In addition, a new LMW-GS allele of 43,351 Da was identified in Strongfield but not in Pelissier. Further studies are required to characterize this new allele.

Quantitative Trait Loci Associated With High Molecular Weight Glutenin Subunit

In the present study, a major QTL (*QGlu.spa-1A*) in proximity to the HMW-GS locus *Glu-A1* on chromosome 1A was detected across all environments and explained 13.7–18.3% of variation in gluten strength. *Glu-A1* allele 1Ay was present in Strongfield but not in Pelissier (**Table S5**). The gene encoding 1Ay subunit is always silent in hexaploid wheat, while expressed in some diploid and tetraploid wheats (Jiang et al., 2009). However, no extra peak corresponding to 1Ay was detected in Strongfield by MALDI-TOF-MS, which might be due to the inactivation of the 1Ay allele. Previous studies identified similar inactivation of 1Ay allele in tetraploid wheat (Jiang et al., 2009). The variation of HMW-GS subunits 1Ax2* (**Figure S3**) is likely associated with *QGlu.spa-1A*, although no polymorphism was detected for allele 1Ax2* using PCR based analysis (**Table S5**). The discrepancy was most likely caused by the differential expression of the gene. In addition, the putative candidate gene for *QGlu.spa-1A* encodes HMW glutenin with high protein similarity to *Glu-A1*, suggesting that the genes associated with *QGlu.spa-1A* and *Glu-A1* could be the same. However, the cloning of the putative candidate gene and conversion into KASP markers are required to confirm this assumption. Likewise, a QTL for gluten strength in durum wheat was detected on chromosome 1AL, but only in one environment in a recombinant inbred line (RIL) population derived from line UC1113 and cv. Kofa (Conti et al., 2011). Another QTL associated with the HMW-GS loci detected in this study is *QGlu.spa-1B.2* on the long arm of chromosome 1B in a position close to *Glu-B1*. This result agrees with the findings reported by Conti et al. (2011) that a stable QTL associated with SV was identified on chromosome 1BL (*Glu-B1*) across multiple environments. Similarly, a QTL linked to the *Glu-B1* locus was also found to be associated with gluten strength in durum wheat (Patil et al., 2009). The subunit 1Bx7 (82,441 Da) detected only in Pelissier (**Figure S3**) is likely associated with *QGlu.spa-1B.2*. All these results confirmed the significant positive association between HMW-GS loci and gluten strength in durum wheat. QTL for SV associated with *Glu-A1* (Li et al., 2009) and *Glu-B1* (Jernigan et al., 2018) were also reported in bread wheat. However, a weaker association was reported between HMW-GS loci and gluten strength in modern durum wheat cultivars likely as a result of limited genetic variation at *Glu-1* (Sissons, 2008). A weak but significant relationship between the HMW-GS and spaghetti quality was previously reported, while some studies showed no clear relationship between these two (reviewed by Liu et al., 1996). A direct measurement of rheological properties of the dough might be needed to determine the gluten strength associated with allelic variation at *Glu-A1* and *Glu-B1* in durum wheat.

The *Glu-A1* locus presented less polymorphism compared to *Glu-B1* in both durum landraces and modern cultivars. In addition, the HMW-GS genes on chromosome lA were reported to have a negligible relationship with durum quality parameters when compared to the genes on chromosome 1B, although active *Glu-A1* alleles were found to have a favorable influence on baking properties of some Italian durum [(Liu et al., 1996) and references therein]. Conti et al. (2011) identified that the most important and stable QTL for gluten strength is associated with *Glu-B1* on chromosome 1BL. In contrast, in the present study, QTL associated with *Glu-A1* had stronger effect than the QTL at *Glu-B1* as evidenced by the higher percentage of phenotypic variance explained. The difference of the findings could be related to the genetic background of the parental lines used for the population development. A null allele at the *Glu-A1* locus was found in Mediterranean durum wheat cultivars while non-null alleles exist in about 40% of the landraces studied (Nazco et al., 2013). Likewise, over 83% of a collection of 502 durum wheat varieties from 23 countries were found to have the *GluA1c* (null) allele (Branlard et al., 1989). The presence of some alleles at the *Glu-B3* locus can offset the effect of the *Glu-B1* alleles. Removal of the *Glu-B3* effect resulted in the detection of the greatest influence of *Glu-B1* (Martínez et al., 2005). In our study, the largest effect QTL (*QGlu.spa-1B.1*) on 1BS in the *Glu-B3* region might mask the effect of *Glu-B1* alleles in some environments, although no significant interaction was observed between these two loci. Further studies are necessary to confirm the assumption and elucidate the underlying mechanism.

Stability of Quantitative Trait Loci

High broad sense heritability of 0.96 was observed for SV in this study (**Table S2**), indicating the phenotypic variation was attributable mainly to the genetic variation. Similar high heritability value of gluten strength measured by SV has been reported in other studies carried out in durum wheat (Clarke et al., 2010; Conti et al., 2011; Kumar et al., 2013). Two stable QTL located on chromosome 1A and 1B near *Glu-A1* and *Glu-B3*, respectively, were detected across all environments tested, with the trait increasing alleles derived from Strongfield. These two QTL are highly desirable for MAS as the selected favorable alleles confer high SV in all years tested and therefore are easy to

be incorporated in breeding programs. QTL × E interaction is an important contributor to the variation in the expression of complex traits. Although the genotype was the main source of variation for SV, significant QTL × E were detected from the multi-trait CIM analysis of two major QTL, *QGlu.spa-1A* and *QGlu.spa-1B.1*, which were significant across all environments and displaying fluctuations in the magnitude of the effects. Another seven QTL detected on chromosome 1B, 2B, and 3A in one or two environments had favorable alleles from Pelissier. This indicates that the expression of the alleles from Pelissier is more prone to be affected by the environment and may be favored in one environment but neutral in others. As demonstrated by previous studies, gluten strength was influenced by genotype and environment, and to some extent by the interaction of genotype × environment, suggesting trials in multiple environments are required for the selection of this trait (Patil et al., 2009; Conti et al., 2011).

Furthermore, gluten strength measured by SV could be positively correlated with GPC, which depends on the genotypes and environments (Clarke et al., 2010). The moderate positive correlation between SV and GPC and weak to moderate negative correlation between SV and GY were observed in three out of six environments in this population (**Figure S2**). However, our studies showed that the stable QTL on 1A and 1B identified in this population do not contribute to GPC and GY (data unpublished).

Epistatic Quantitative Trait Loci Interaction

The identification of epistatic interactions for the QTL whose effects mostly dependent on the genotypes of other loci, can provide a more comprehensive understanding of genetic components controlling the expression of complex traits and a more accurate prediction for the phenotypic traits (Bocianowski, 2013). Of note, in the present study the epistatic interaction between *QGlu.spa-1A* and *QGlu.spa-1B.1* was repeatedly detected in 4 out of 6 environments indicating a positive interaction between alleles of HMW-GS and LMW-GS. Therefore, it is important to take into account such epistatic effects for marker assisted selection (MAS). Significant interactions between *Glu-B3* and other glutenin loci were observed in a previous study (Martínez et al., 2005). Likewise, QTL for gluten strength on 1BL was reported to have an epistatic effect with other loci having no main effect (Conti et al., 2011).

Taking together, *QGlu.spa-1A* and *QGlu.spa-1B.1* contributed the most desirable alleles derived from parental line Strongfield and were consistently expressed over multiple environments. Two flanking markers, Kukri c38553_67 and RCA875_rep_c74067_541, in the QTL region of *QGlu.spa-1B.1* can be used to effectively separate the DH lines into two groups with significantly different mean SV values. More distinct separation was obtained using flanking markers of both QTL *QGlu.spa-1A* (IAAV1142 and RAC875_c31031_387) and *QGlu.spa-1B.1*. The KASP assays for *QGlu.spa-1A* and *QGlu.spa-1B.1* showed the good clusters and reliable results, demonstrating the effectiveness of using these KASP markers for selecting the lines with higher SV/gluten strength in durum wheat although the validation in a diverse panel is required. As such, these KASP markers have the potential to be applied for MAS in durum breeding programs. These two QTL should be subjected to map-based cloning. Although the candidate genes were predicted for these two QTL based on the QTL position on the reference genomes of durum wheat cv. Svevo and wild emmer accession Zavitan along with the comparison with previously published studies, further studies are needed to confirm our assumption. Haplotype analysis of these two QTL along with another two QTL on 3A indicated the DH lines with the combination of all favorable alleles at each QTL had the highest mean SV across all environments. These DH lines are the good candidates as parental lines for developing new varieties with strong gluten strength. Similar haplotype analysis of QTL*QGlu.spa-1A* and *QGlu.spa-1B.1* in a diverse durum panel will enhance our understanding of the allelic variants of *Glu-A1* and *Glu-B3* and may facilitate more effective use of favorable alleles in further improving gluten strength of durum wheat.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/ **Supplementary Material**.

AUTHOR CONTRIBUTIONS

YR and RK conceptualized this study. YR, RK, AS, RD and RC generated the population and implemented the field trials and phenotyping of the population. IP and PG conducted the marker validation and the KASP assays. WZ, AS and PF provided the genotyping platform. BY conducted data analysis. BY and YR contributed to data management and visualization. BY wrote the original manuscript and YR, RK, RD, and PF contributed to the review and editing of the manuscript. YR was the principal investigator and acquired the fund for this study.

FUNDING

Financial support was received from the National Wheat Improvement Program and the Canadian Agricultural Partnership with support from Agriculture and Agri-Food Canada, Western Grains Research Foundation, Alberta Wheat Commission, Saskatchewan Wheat Development Commission, and Manitoba Wheat and Barley Growers Association. The work was also supported by Canadian Wheat Improvement Flagship funded by National Research Council Canada.

ACKNOWLEDGMENTS

We gratefully acknowledge the support of Agriculture and Agri-food Canada (AAFC), Western Grain Research Foundation (WGRF), National Research Council Canada (NRC), Alberta Wheat Commission, Saskatchewan Wheat Development Commission, and Manitoba Wheat and Barley Growers Association in carrying out this study. We thank Christine Sidebottom and Janet Condie at NRC for technical assistance

with DNA extraction and genotyping using the Infinium iSelect Wheat 90K SNP chip. The technical support of Wheat Breeding Group at the Swift Current Research and Development Centre and Cereal Quality Lab in Winnipeg from AAFC is sincerely appreciated.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.00170/full#supplementary-material

REFERENCES

Ames, N. P., Clarke, J. M., Marchylo, B. A., Dexter, J. E., Schlichting, L. M., and Woods, S. M. (2003). The effect of extra-strong gluten on quality parameters in durum wheat. *Can. J. Plant Sci.* 83, 525–532. doi: 10.4141/P02-072

Anderson, O. D., Gu, Y. Q., Kong, X., Lazo, G. R., and Wu, J. (2009). The wheat ω-gliadin genes: structure and EST analysis. *Funct. Integr. Genomics* 9, 397–410. doi: 10.1007/s10142-009-0122-2

Autran, J. C., and Feillet, P. (1985). "Genetic and technological basis of protein quality for durum wheat in pasta," in *Protein Evaluation in Cereals and Legumes. Report EUR 10404*. Ed. A. Pattakou (Luxemburg: Commission of the European Communities), 59–71.

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., et al. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* 97, 93–97. doi: 10.1126/science.aan0032

Barak, S., Mudgil, D., and Khatkar, B. S. (2015). Biochemical and functional properties of wheat gliadins: a review. *Crit. Rev. Food Sci.* 55, 357–368. doi: 10.1080/10408398.2012.654863

Blanco, A., Bellomo, M. P., Lotti, C., Maniglio, T., Pasqualone, A., Simeone, R., et al. (1998). Genetic mapping of sedimentation volume across environments using recombinant inbred lines of durum wheat. *Plant Breed.* 117, 413–417. doi: 10.1111/j.1439-0523.1998.tb01965.x

Bocianowski, J. (2013). Epistasis interaction of QTL effects as a genetic parameter influencing estimation of the genetic additive effect. *Genet. Mol. Biol.* 36, 93–100. doi: 10.1590/S1415-47572013000100013

Branlard, G., Autran, J. C., and Monneveaux, P. (1989). High molecular weight glutenin subunits of durum wheat (*T. durum*). *Theor. Appl. Genet.* 78, 353–358. doi: 10.1007/BF00265296

Brites, C., and Carrillo, J. M. (2001). Influence of high molecular weight (HMW) and low molecular weight (LMW) glutenin subunits controlled by *Glu-1* and *Glu-3* loci on durum wheat quality. *Cereal Chem.* 78, 59–63. doi: 10.1094/CCHEM.2001.78.1.59

Broman, K. W., Wu, H., Sen, Ś., and Churchill, G. A. (2003). R/qtl: QTL mapping in experimental crosses. *Bioinformatics* 19, 889–890. doi: 10.1093/bioinformatics/btg112

Cavanagh, C. R., Chao, S. M., Wang, S. C., Huang, B. E., Stephen, S., Kiani, S., et al. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8057–8062. doi: 10.1073/pnas.1217133110

Clarke, J. M., Marchylo, B., Kovac, M. I. P., Noll, J. S., McCaig, T. N., and Howes, N. K. (1998). Breeding durum wheat for pasta quality in Canada. *Euphytica* 100, 163–170. doi: 10.1023/A:1018313603344

Clarke, J. M., McCaig, T. N., DePauw., R. M., Knox, R. E., Clarke, F. R., Fernandez, M. R., et al. (2005). Strongfield durum wheat. *Can. J. Plant Sci.* 85, 651–654. doi: 10.4141/P04-119

Clarke, F. R., Clarke, J. M., Ames, N. A., Knox, R. E., and Ross, J. R. (2010). Gluten index compared with SDS-sedimentation volume for early generation selection for gluten strength in durum wheat. *Can. J. Plant Sci.* 90, 1–11. doi: 10.4141/CJPS09035

Conti, V., Roncallo, P. F., Beaufort, V., Cervigni, G. L., Miranda, R., Jensen, C. A., et al. (2011). Mapping of main and epistatic effect QTLs associated to grain protein and gluten strength using a RIL population of durum wheat. *J. Appl. Genet.* 52, 287–298. doi: 10.1007/s13353-011-0045-1

D'Ovidio, R., and Masci, S. (2004). The low-molecular-weight glutenin subunits of wheat gluten. *J. Cereal Sci.* 39, 321–339. doi: 10.1016/j.jcs.2003.12.002

Dexter, J. E., Matsuo, R. R., Kosmolak, F. G., Leisle, D., and Marchylo, B. A. (1980). The suitability of the SDS-sedimentation test for assessing gluten strength in durum wheat. *Can. J. Plant Sci.* 60, 25–29. doi: 10.4141/cjps80-004

Dexter, J. E. (2008). *The history of durum wheat breeding in Canada and summaries of recent research at the Canadian Grain Commission on factors associated with durum wheat processing, in Bosphorus 2008 ICC* (Istanbul: International Cereal Congress).

Dick, J. W., and Quick, J. S. (1983). A modified screening test for rapid estimation of gluten strength in early generation durum wheat breeding. *Cereal Chem.* 60, 315–318.

Du Cros, D. L. (1987). Gluten proteins and gluten strength in durum wheat. *J. Cereal Sci.* 5, 3–12. doi: 10.1016/S0733-5210(87)80003-6

Elfatih, S., Peng, Y., Ma, J., Peng, J., Sun, D., and Ma, W. (2013). High frequency of unusual high molecular weight glutenin alleles in 232 tetraploid durum wheat accessions (*Triticum turgidum*L. ssp. *durum* Desf). *Cereal Res. Commun.* 41, 583–592. doi: 10.1556/CRC.2013.0031

Elouafi, I., Nachit, M., Elsaleh, A., Asbati, A., and Mather, D. E. (2000). "QTL-mapping of genomic regions controlling gluten strength in durum (*Triticum turgidum* L. var. durum)," in *Durum Wheat Improvement in The Mediterranean Region: New Challenges*. Eds. C. Royo, M. Nachit, N. Di Fonzo and J. L. Araus (Zaragoza: CIHEAM), 505–509.

Feillet, P., Ait-Mouh, O., Kobrehel, K., and Autran, J. C. (1989). The role of low molecular weight glutenin proteins in the determination of cooking quality of pasta products: an overview. *Cereal Chem.* 66, 26–30.

Frutos, E., Galindo, M. P., and Leiva, V. (2014). An interactive biplot implementation in R for modeling genotype-by-environment interaction. *Stoch Environ. Res. Risk Assess.* 28, 1629–1641. doi: 10.1007/s00477-013-0821-z

Giraldo, P., Royo, C., González, M., Carrillo, J. M., and Ruiz, M. (2016). Genetic diversity and association mapping for agromorphological and grain quality traits of a structured collection of durum wheat landraces including subsp. *durum*, *turgidum* and *diccocon*. *PloS One* 11, e0166577. doi: 10.1371/journal.pone.0166577

Goel, S., Singh, K., Singh, B., Grewa, S., Dwivedi, N., Alqarawi, A. A., et al. (2019). Analysis of genetic control and QTL mapping of essential wheat grain quality traits in a recombinant inbred population. *PloS One*. 14, e0200669. doi: 10.1371/journal.pone.0200669

Gupta, R. B., Bekes, F., and Wrigley, C. W. (1991). Prediction of physical dough properties from glutenin subunit composition in bread wheats. *Cereal Chem.* 68, 328–333.

Heffelfinger, C., Fragoso, C. A., and Lorieux, M. (2017). Constructing linkage maps in the genomics era with MapDisto 2.0. *Bioinformatics*. 33, 2224–2225. doi: 10.1093/bioinformatics/btx177

Henkrar, F., El-Haddoury, J., Iraqi, D., Bendaou, N., and Udupa, S. M. (2017). Allelic variation at high-molecular weight and low-molecular weight glutenin subunit genes in Moroccan bread wheat and durum wheat cultivars. *3 Biotech.* 7, 287. doi: 10.1007/s13205-017-0908-1

Humphreys, G. D., and Knox, R. E. (2015). "Doubled-Haploid Breeding in Cereals," in *Advances in Plant Breeding Strategies: Breeding, Biotechnology and Molecular Tools*. Eds. J. M. Al-Khayri, S. M. Jain and D. V. Johnson (Switzerland: Springer International Publishing). doi: 10.1007/978-3-319-22521-0_9

Irvine, G. N. (1971). "Durum wheat and paste products," in *Wheat: Chemistry and Technology*, vol. III . Ed. Y. Promeranz (St. Paul, Minn: American Association of Cereal Chemists), 777–798.

Jernigan, K. L., Godoy, J. V., Huang, M., Zhou, Y., Morris, C. F., Garland-Campbell, K. A., et al. (2018). Genetic dissection of end-use quality traits in adapted soft white winter wheat. *Front. Plant Sci.* 9, 271. doi: 10.3389/fpls.2018.00271

Jiang, Q., Wei, Y., Wang, F., Wang, J.-R., Yan, Z.-H., and Zheng, Y.-L. (2009). Characterization and comparative analysis of HMW glutenin 1Ay alleles with differential expressions. *BMC Plant Biol.* 9, 16. doi: 10.1186/1471-2229-9-16

Joppa, L. R., Khan, K., and Williams, N. D. (1983). Chromosomal location of genes for gliadin polypeptides in durum wheat *Triticum turgidum*. *Theor. Appl. Genet.* 64, 289–293. doi: 10.1007/BF00274164

Kovacs, M. I. P., Howes, N. K., Leisle, D., and Skerritt, J. H. (1991). "The Effect of High Molecular Weight Glutenin Subunit Composition on Tests Used to Predict Durum Wheat Quality," in *Cereals Int. Proc. Victoria*. Eds. D. F. Martin and C. W. Wrigley (Melbourne, Australia: Royal Australian Chemical Institute), 257–262.

Kovacs, M. I. P., Howes, N. K., Leisle, D., and Skerritt, J. H. (1993). The effect of high Mr, glutenin subunit composition on the results from tests used to predict durum wheat quality. *J. Cereal Sci.* 18, 43–51. doi: 10.1006/jcrs.1993.1032

Kovacs, M. I. P., Noll, J. S., Dahlke, G., and Leisle, D. (1995a). Pasta viscoelasticity: its usefulness in the Canadian durum wheat breeding program. *J. Cereal Sci.* 22, 115–121. doi: 10.1016/0733-5210(95)90040-3

Kovacs, M. I. P., Howes, N. K., Leisle, D., and Zawistowski, J. (1995b). Effect of two different low molecular weight glutenin subunits on durum wheat pasta quality parameters. *Cereal Chem.* 72, 85–87.

Kovacs, M. I. P., Poste, L. M., Butler, G., Woods, S. M., Leisle, D., Noll, J. S., et al. (1997). Durum wheat quality: comparison of chemical and rheological screening tests with sensory analysis. *J. Cereal Sci.* 25, 65–75. doi: 10.1006/jcrs.1996.0069

Kumar, A., Jain, S., Iqbal, M. J., Elias, E. M., Kianian, S., and Mergoum, M.Fine Mapping and Validation of a Major QTL for Gluten Strength, in: *ASA, CSSA & SSSA International Annual Meeting*, Long Beach, CA., Nov. 2-5, 2014.

Kumar, A., Elias, E. M., Ghavami, F., Xu, X., Jain, S., Manthey, F. A., et al. (2013). A major QTL for gluten strength in durum wheat (*Triticum turgidum* L. var. *durum*). *J. Cereal Sci.* 57, 21–29. doi: 10.1016/j.jcs.2012.09.006

Laurie, C., Wang, S., Carlini-Garcia, L. A., and Zeng, Z. B. (2014). Mapping epistatic quantitative trait loci. *BMC Genet.* 15, 112. doi: 10.1186/s12863-014-0112-9

Li, Q. Y., Yan, Y. M., Wang, A. L., An, X. L., Zhang, Y. Z., Hsam, S. L. K., et al. (2006). Detection of HMW glutenin subunit variations among 205 cultivated emmer accessions (*Triticum turgidum* spp. *dicoccum*). *Plant Breed.* 125, 120–124. doi: 10.1111/j.1439-0523.2006.01173.x

Li, Y., Song, Y., Zhou, R., Branlard, G., and Jia, J. (2009). Detection of QTLs for bread-making quality in wheat using a recombinant inbred line population. *Plant Breed.* 128, 235–243. doi: 10.1111/j.1439-0523.2008.01578.x

Liu, C. Y., Shepherd, K. W., and Rathjen, A. J. (1996). Improvement of durum wheat pasta making and bread making qualities. *Cereal Chem.* 73, 155–166.

Liu, T.-T., Liu, K., Wang, F.-F., Zhang, Y., Li, Q.-F., Zhang, K.-R., et al. (2017). Conditional and unconditional QTLs mapping of gluten strength in common wheat (*Triticum aestivum* L.). *J. Integr. Agr.* 16, 2145–2155. doi: 10.1016/S2095-3119(16)61564-2

Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L., Salem, M. B., Bort, J., et al. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat (*Triticum durum* Desf.) across a wide range of water availability. *Genetics* 178, 489–511. doi: 10.1534/genetics.107.077297

Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., et al. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 13, 648–663. doi: 10.1111/pbi.12288

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51, 885–895. doi: 10.1038/s41588-019-0381-3

Martínez, M. C., Ruiz, M., and Carrillo, J. M. (2005). Effects of different prolamin alleles on durum wheat quality properties. *J. Cereal Sci.* 41, 123–131. doi: 10.1016/j.jcs.2004.10.005

Metakovsky, E. V., Wrigley, C. W., Bekes, F., and Gupta, R. B. (1990). Gluten polypeptides as useful genetic markers of dough quality in Australian wheats. *Aust. J. Agric. Res.* 41, 289–306. doi: 10.1071/AR9900289

N'Diaye, A., Haile, J. K., Nilsen, K. T., Walkowiak, S., Ruan, Y., Singh, A. K., et al. (2018). Haplotype loci under selection in Canadian durum wheat germplasm over 60 years of breeding: association with grain yield, quality traits, protein loss, and plant height. *Front. Plant Sci.* 9, 1589. doi: 10.3389/fpls.2018.01589

Nazco, R., Peña, R. J., Ammar, K., Villegas, D., Crossa, J., Moragues, M., et al. (2013). Variability in glutenin subunit composition of Mediterranean durum wheat germplasm and its relationship with gluten strength. *J. Agric. Sci.* 152, 379–393. doi: 10.1017/S0021859613000117

Pasqualone, A., Piarulli, L., Mangini, G., Gadaleta, A., Blanco, A., and Simeone, R. (2015). Quality characteristics of parental lines of wheat mapping populations. *Agric. Food Sc.* 24, 118–127. doi: 10.23986/afsci.49570

Patil, R. M., Oak, M. D., Tamhankar, S. A., and Rao, V. S. (2009). Molecular mapping of QTLs for gluten strength as measured by sedimentation volume and mixograph in durum wheat (*Triticum turgidum* L. ssp *durum*). *J. Cereal Sci.* 49, 378–386. doi: 10.1016/j.jcs.2009.01.001

Payne, P. I., and Lawrence, G. J. (1983). Catalogue of alleles for the complex loci, Glu-A1, Glu-B1 and Glu-D1 which code for high-molecular-weight subunits of glutenin in hexaploid wheat. *Cereal Res. Commun.* 11, 29–35.

Payne, P. I. (1987). Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Ann. Rev. Plant Physiol.* 38, 141–153. doi: 10.1146/annurev.pp.38.060187.001041

Pogna, N. E., Autran, J. C., Mellini, F., Lafiandra, D., and Feillet, P. (1990). Chromosome 1B-encoded gliadins and glutenin subunits in durum wheat: genetics and relationship to gluten strength. *J. Cereal Sci.* 11, 15–34. doi: 10.1016/S0733-5210(09)80178-1

Quick, J. S., and Donnelly, B. J. (1980). A rapid test for estimating durum wheat gluten quality. *Crop Sci.* 20, 816–818. doi: 10.2135/cropsci1980.0011183X002000060037x

Rasheed, A., Wen, W., Gao, F., Zhai, S., Jin, H., Liu, J., et al. (2016). Development and validation of KASP assays for genes underpinning key economic traits in bread wheat. *Theor. Appl. Genet.* 129, 1843–1860. doi: 10.1007/s00122-016-2743-x

Roselló, M., Royo, C., Álvaro, F., Villegas, D., Nazco, R., and Soriano, J. M. (2018). Pasta-making quality QTLome from mediterranean durum wheat landraces. *Front. Plant Sci.* 9, 1512. doi: 10.3389/fpls.2018.01512

Ruiz, M., and Carrillo, J. M. (1995). Separate effects on gluten strength of *Gli-1* and *Glu-3* prolamin genes on chromosomes 1A and 1B in durum wheat. *J. Cereal Sci.* 21, 137–144. doi: 10.1016/0733-5210(95)90029-2

Sapirstein, H. D., David, P., Preston, K. R., and Dexter, J. E. (2007). Durum wheat breadmaking quality: effects of gluten strength, protein composition, semolina particle size and fermentation time. *J. Cereal Sci.* 45, 150–161. doi: 10.1016/j.jcs.2006.08.006

Shewry, P. R., Halford, N. G., and Tatham, A. S. (1992). The high molecular weight subunits of wheat glutenin. *J. Cereal Sci.* 15, 105–120. doi: 10.1016/S0733-5210(09)80062-3

Sissons, M. J., Soh, H. N., and Turner, M. A. (2007). Role of gluten and its components in influencing durum wheat dough properties and spaghetti cooking quality. *J. Sci. Food Agric.* 87, 1874–1885. doi: 10.1002/jsfa.2915

Sissons, M. (2008). Role of durum wheat composition on the quality of pasta and bread. *Food* 2, 75–90.

Voorrips, R. E. (2002). MapChart: Software for the graphical presentation of linkage maps and QTLs. *J. Hered.* 93, 77–78. doi: 10.1093/jhered/93.1.77

Wang, L. H., Li, G. Y., Peña, R. J., Xia, X. C., and He, Z. H. (2010). Development of STS markers and establishment of multiplex PCR for *Glu-A3* alleles in common wheat (*Triticum aestivum* L.). *J. Cereal Sci.* 51, 305–312. doi: 10.1016/j.jcs.2010.01.005

Wang, S., Basten, C. J., and Zeng, Z.-B. (2012). *Windows QTL Cartographer 2.5. Department of Statistics* (Raleigh, NC: North Carolina State University). (http://statgen.ncsu.edu/qtlcart/WQTLCart.htm).

Wang, A., Liu, L., Peng, Y., Islam, S., Applebee, M., Appels, R., et al. (2015). Identification of low molecular weight glutenin alleles by matrix-assisted laser desorption/ionization time-of-flight mass spectrometry (MALDI-TOF-MS) in common wheat (*Triticum aestivum* L.). *PloS One* 10 (9), e0138981. doi: 10.1371/journal.pone.0138981

Wrigley, C., Bekes, F., and Bushuk, W. (2006). "Gliadin and Glutenin," in *The Unique Balance of Wheat Quality, 1st Ed* (MN: AACC International), 3–32.

Xu, L. L., Li, W., Wei, Y. M., and Zheng, Y. L. (2009). Genetic diversity of HMW glutenin subunits in diploid, tetraploid and hexaploid *Triticum* species. *Genet. Resour. Crop Evol.* 56, 377–391. doi: 10.1007/s10722-008-9373-3

Zhang, W., Chao, S., Manthey, F., Chicaiza, O., Brevis, J. C., Echenique, V., et al. (2008). QTL analysis of pasta quality using a composite microsatellite and SNP map of durum wheat. *Theor. Appl. Genet.* 117, 1361–1377. doi: 10.1007/s00122-008-0869-1

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

frontiers in Plant Science

REVIEW
published: 29 October 2020
doi: 10.3389/fpls.2020.568657

A Systematic Review of Durum Wheat: Enhancing Production Systems by Exploring Genotype, Environment, and Management (G × E × M) Synergies

Brian L. Beres[1], Elham Rahmani[1], John M. Clarke[2†], Patricio Grassini[3], Curtis J. Pozniak[2], Charles M. Geddes[1], Kenton D. Porker[4], William E. May[5] and Joel K. Ransom[6]*

[1] Agriculture and Agri-Food Canada, Lethbridge Research and Development Centre, Lethbridge, AB, Canada, [2] Crop Development Centre and Department of Plant Sciences, University of Saskatchewan, Saskatoon, SK, Canada, [3] Department of Agronomy and Horticulture, University of Nebraska, Lincoln, NE, United States, [4] South Australia Research and Development Institute, Adelaide, SA, Australia, [5] Agriculture and Agri-Food Canada, Indian Head Research Station, Saskatchewan, SK, Canada, [6] Department of Plant Sciences, North Dakota State University, Fargo, ND, United States

OPEN ACCESS

Edited by:
Jose Luis Gonzalez Hernandez, South Dakota State University, United States

Reviewed by:
Ivan A. Matus, Institute of Agricultural Research, Chile
Hamid Khazaei, University of Saskatchewan, Canada

***Correspondence:**
Brian L. Beres
brian.beres@canada.ca

[†] *Deceased*

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: *01 June 2020*
Accepted: *06 October 2020*
Published: *29 October 2020*

Citation:
Beres BL, Rahmani E, Clarke JM, Grassini P, Pozniak CJ, Geddes CM, Porker KD, May WE and Ransom JK (2020) A Systematic Review of Durum Wheat: Enhancing Production Systems by Exploring Genotype, Environment, and Management (G × E × M) Synergies. Front. Plant Sci. 11:568657. doi: 10.3389/fpls.2020.568657

According to the UN-FAO, agricultural production must increase by 50% by 2050 to meet global demand for food. This goal can be accomplished, in part, by the development of improved cultivars coupled with modern best management practices. Overall, wheat production on farms will have to increase significantly to meet future demand, and in the face of a changing climate that poses risk to even current rates of production. Durum wheat [*Triticum turgidum* L. ssp. durum (Desf.)] is used largely for pasta, couscous and bulgur production. Durum producers face a range of factors spanning abiotic (frost damage, drought, and sprouting) and biotic (weed, disease, and insect pests) stresses that impact yields and quality specifications desired by export market end-users. Serious biotic threats include Fusarium head blight (FHB) and weed pest pressures, which have increased as a result of herbicide resistance. While genetic progress for yield and quality is on pace with common wheat (*Triticum aestivum* L.), development of resistant durum cultivars to FHB is still lagging. Thus, successful biotic and abiotic threat mitigation are ideal case studies in Genotype (G) × Environment (E) × Management (M) interactions where superior cultivars (G) are grown in at-risk regions (E) and require unique approaches to management (M) for sustainable durum production. Transformational approaches to research are needed in order for agronomists, breeders and durum producers to overcome production constraints. Designing robust agronomic systems for durum demands scientific creativity and foresight based on a deep understanding of constitutive components and their innumerable interactions with each other and the environment. This encompasses development of durum production systems that suit specific agro-ecozones and close the yield gap between genetic potential and on-farm achieved yield. Advances in individual technologies (e.g., genetic improvements, new pesticides,

seeding technologies) are of little benefit until they are melded into resilient G × E × M systems that will flourish in the field under unpredictable conditions of prairie farmlands. We explore how recent genetic progress and selected management innovations can lead to a resilient and transformative durum production system.

Keywords: durum wheat, genotype, environment, management, G × E × M, agronomy

INTRODUCTION

An estimate by the UN-FAO indicates that, by 2050, the global demand for agricultural products will have risen by 50%. Meeting this demand will require traditional development of improved cultivars coupled with modern best management practices as well as innovations that are transformational. Achievement of this goal on existing cropland will require a significant increase in rates of genetic gain in grain yield for crops such as wheat (*Triticum aestivum* L.), increasing the current rate of gain (ca. 1% p.a.) by 30–40% (FAO, 2017; Cassman and Grassini, 2020). Climate change will be an added challenge to productivity improvement (Reynolds et al., 2016). Durum wheat (*Triticum turgidum* L. var durum Desf.) is the 10th most important and commonly cultivated cereal worldwide with a yearly production average of 40 million tonnes (MT) (2016/17). Typically, durum wheat production represents 5% of total wheat production with a planting area of 16 M hectares globally (International Grains Council [IGC], 2020). Wheat and wheat products could account for 20% of protein and calories consumption per capita for a global population of 9.7 billion in 2050 (CRP-WHEAT, 2016). Durum is produced primarily for making pasta, but is also an important ingredient for couscous and bulgur, particularly in North Africa and the Middle East. These products use durum semolina resulting from milling of the hard-textured durum wheat kernel. In some countries such as Italy, regulatory standards specify that pasta must be made with 100% durum semolina (Sopiwnyk, 2018). The production of pasta requires grain with high protein content, gluten strength, and high yellow pigment content (resulting largely from lutein), which provides the characteristic yellowness that is expected from the pasta. Only a few regions in the world are capable of producing durum that meets the high standards for end-use suitability. In this review article, we will discuss the current state of durum production and explore how recent genetic progress and selected management innovations can lead to a resilient and transformative durum production system.

Statistics and Regional Specific Summaries (Canada, Australia, and the United States)

Durum wheat is grown in many of the same countries that produce common wheat, with Italy as an important producer (4.95 MT) within Europa, along with the Commonwealth of Independent States (CIS), North America, South America, Asia, Africa, Oceania, and Turkey (3.62 MT) (International Grains Council [IGC], 2020). A majority of the world's durum wheat is planted in North America, with Canadian production typically around 7.8 MT, which is almost three times the production of the United States (US) and Mexico (International Grains Council [IGC], 2020). In Canada, southern Saskatchewan is the largest wheat durum producer, supplying 81 percent of the total Canadian durum wheat produced from 1990 to 2017. Canada's durum export volume has increased from 2.7 to 4.5 MT from 1990 to 2017, which underscores the significant contribution of durum production to the Canadian export market (Statistics Canada, 2019). Canada now leads the world in durum wheat exports, with about half of all durum wheat available for export grown in Canada. Mexico and the EU contribute 17 and 16%, respectively. The remaining 17% is exported from other countries including the United States, Australia, Mexico, and Kazakhstan. From the period 2013–2014 to 2017–2018, Italy, Algeria, Morocco, United States, and Japan were the top five importing countries for Canadian durum wheat (International Grains Council [IGC], 2020).

In Australia, durum wheat represents a relatively small component of the Australian wheat crop. Durum production is largely confined to rainfed production in southeastern Australia (South Australia, Victoria, and New South Wales and a small part of southern Queensland) and small pockets of irrigated production. Wheat production in eastern Australia has averaged ∼16 MT per annum over the last 10 years (ABARES 2017), whereas durum production averages ∼4,00,000 tonnes (T) but has fluctuated substantially between ∼50,000 and 8,00,000 t over the last decade (Kniep, 2008; Ranieri et al., 2012); and represents on average 3% of the eastern Australia wheat crop. Australian durum wheat production is relatively stable, and the annual supply of grain is split equally between domestic consumption and export markets, which is somewhat dependent on seasonal fluctuations (Ranieri, 2015).

In the United States, durum is produced primarily in two regions, the desert southwest region under irrigated regimes, and in the central region of the northern Great Plains under rainfed conditions. The largest area planted to durum is in North Dakota (**Table 1**), followed by Montana, Arizona, and California. There is also minor production in Idaho (irrigated), South Dakota and Minnesota. In the last two decades, there has been a substantial reduction in the area planted to durum in the United States. Most of this can be attributed to reductions in North Dakota, where there has been a reduction of nearly 800,000 hectares over the period. Farmers in North Dakota have opted to grow other crops due to the challenges of meeting the high quality standards to reach the top grade and the lack of financial incentives relative to other crops including hard red spring wheat, which has somewhat less stringent quality requirements. Lastly, as with other regions, the lack of genetic

TABLE 1 | Durum production trends in the United States over the last 20 years (USDA National Agricultural Statistic Service database).

State	Area planted in 2018	Average annual growth or reduction in area planted since 1998†	Average yield over of the period	Average annual yield increase since 1998‡
	(ha)	(ha year^{-1})	(*t* ha^{-1})	(*t* ha^{-1} year^{-1})
Arizona	28,745	801	6.74	0.031
California	16,194	−338	6.74	−0.037
Montana	340,081	5,176	1.85	0.021
North Dakota	445,344	−39,734	2.12	0.040

†*Derived from the linear regression of hectares planted, 1998 through 2018.*
‡*Derived from the linear regression of average statewide yield, 1998 through 2018.*

resistance to Fusarium head blight (FHB) caused by *Fusarium graminearum* Schwabe [telomorph: Gibberella zeae Schwein (Petch)] has shifted hectarage away from durum to wheat crops with higher levels of resistance.

Market Access

A recent challenge for producers is slumping durum exports to Italy, the world's largest pasta maker. As reported in April 2018, despite the high exceptional quality of Canadian durum and as Italy's biggest durum supplier, the company Barilla has cut back its imports from Canada by 35%. Italy has expressed concerns over Maximum Residue Limits (MRLs) for glyphosate in Canadian durum as the reason for the blockade. In the northern Great Plains, where durum is managed within no-tillage, soil conservation system, glyphosate is widely used as a pre-plant application to control weeds. Glyphosate applied to durum after it reaches physiological maturity is also approved as a pre-harvest weed control in the United States and Canada. It is most commonly used when environmental conditions prolong the process of drying after maturity. Glyphosate applied prior to harvest has been found to only marginally hastens the rate of dry down in fully mature wheat (Darwent et al., 1994), but it does an excellent job of controlling any green tillers or weeds that might produce seed and increase dockage at grain terminals. Although the MRL is five parts per million for wheat according to Health Canada's current guidelines, the accepted limit of glyphosate established by the Italian pasta industry is under 10 parts per billion. Exporters argue the limits are too low because glyphosate is already used within acceptable limits of the herbicide in the grain production system.

The expanded use of glyphosate as a preharvest aid may be associated with the expanded use of fungicides for FHB control, as plants treated with this late fungicide application tend to remain green longer as the leaves and stems are not affected by late-season diseases that might naturally desiccate them. Glyphosate is highly mobile in the plant and moves to areas of high metabolism. When applied prior to maturity, it can accumulate in developing kernels in sufficient quantity that it can affect germination and therefore is not currently approved for use in seed production or on barley intended for malt (Jenks et al., 2019). Even when applied after physiological maturity, however, it is possible to detect traceable amounts of glyphosate in the grain (Cessna et al., 1994). When applied according to the manufacturer's label, glyphosate does not affect milling and baking characteristic of spring wheat (Manthey et al., 2004) or the functional quality of durum (Zollinger et al., 1999). From a toxicological standpoint, glyphosate is considered to be one of the safest herbicides available, as its mode of action is directed towards an enzyme not found in mammals. Most countries consider pesticide residues in their food based on the guidelines found in the Codex Alimentarius which has a permitted residues limit of 30 ppm of glyphosate in wheat. Some countries like Japan and Canada have lower permitted limits of 5 ppm. Nevertheless, after the International Agency for Research on Cancer concluded that glyphosate may be linked to cancer, many end users have become concerned with any detected levels of glyphosate in the grain. There are alternative chemicals registered for pre-harvest weed control in durum, but they are primarily effective on broadleaf weeds and would have minimal impact on the desiccation of the durum crop (Jenks et al., 2019). Reverting to the practice of swathing the crop and leaving the cut grain in windrows to dry before combining may also be challenging. This practice is labor-intensive and increases operational input costs. Grain quality is also at greater risk as there is less potential airflow permeating through the windrow such that, if there is a rain event, there is increased risk of quality downgrading from weathering and sprouting.

Yield Constraints and Emerging Issues for Durum Wheat Production

Attainable Yield and the Global Yield Gap Atlas

Our ability to fully harness the genetic yield potential of newly deployed genetics and traits in the latest durum cultivars is determined by site-specific parameters and local management practices. Achieving food security and protecting carbon-rich and biodiverse natural ecosystems from massive conversion to cropland ultimately depend on our ability to sustainably increase crop yields on currently cultivated land (Cassman, 1999). Until recently, for most of the world, including data-rich regions such as the United States Corn (*Zea mays* L) Belt and Europe, there were no reliable data on yield potential—the maximum attainable yield as determined by climate and soil in the absence of nutrient deficiencies and biotic stresses (Evans and Fischer, 1999). To help meet this need for data and help producers achieve these potential yield gains, researchers from University of Nebraska-Lincoln (United States) and Wageningen University (WU) began development of the Global Yield Gap Atlas (GYGA) (GYGA,[1]) in 2011, with the goal of defining regions' exploitable yield

[1]www.yieldgap.org

gaps (**Figure 1**). In this figure, yield potential is determined by temperature, precipitation, solar radiation, carbon dioxide, and, in the case of rainfed crops, also by water supply and soil properties influencing soil water balance (Evans, 1993; Van Ittersum et al., 2013). The attainable yield is about 80% of the yield potential (Cassman et al., 2003; Lobell et al., 2009). Achieving yield gain above the attainable yield is difficult as the extra investment on inputs and labor is not cost effective. The exploitable yield gap represents the difference between average on-farm or actual yield and the attainable yield.

The GYGA provides a web-based platform for estimating yield gaps that is transparent, accessible, reproducible, geospatially explicit, agronomically robust, and applied in a consistent manner throughout the world. A standard protocol for assessing yield gaps was developed by leading scientists, which is based on a strong focus on understanding the local farming system and making use of the best available data sources (Grassini et al., 2015; Van Bussel et al., 2015). A number of studies have been published on these topics using the GYGA approach (Hochman et al., 2013; Merlos et al., 2015; Van Oort et al., 2015; Espe et al., 2016; Marin et al., 2016; Van Ittersum et al., 2016; Timsina et al., 2018). Currently, wheat yield gaps have been developed for the EU, AUS, South America, and India; however, yield gaps for Canada and the United States have yet to be developed by GYGA and remain unknown. A project to obtain estimates for Canada and the United States in a joint collaborative project utilizing GYGA methodology was initiated in 2019. In GYGA, yield potential is simulated based on long-term weather data, local soil, cropping system data, current crop sequences, and dominant management practices such as sowing date, plant density, and cultivar maturity (Grassini et al., 2015).

An alternative approach was used to investigate wheat yield trends, attainable yields and yield gaps for the 10 largest wheat producing countries in the world and more localized yield statistics at the state or county level. These data were assembled from available government sources. Attainable yield was determined using an upper quadrant analysis to define the upper frontier or yields over the period of record and yield gaps calculated as the difference between attainable yield and actual yield for each year. In all countries, the attainable yield increase over time was larger than the yield trend indicating the technological advances in genetics and agronomic practices were increasing attainable yield. Average wheat yield gap using this method report that Australia is about 24%, spring wheat in Canada (Saskatchewan and Alberta) is between 21 and 24%. France, Germany, Mexico, and United States are approximately 14, 9, 10, and 12%, respectively. Observations across all the wheat production regions shows no apparent trend in closing the yield gap. The average yield gap using this method is a reflection in the variation in weather among growing season and management of soil water followed by enhanced agronomy. A series of challenges exist as we attempt to assemble regional data with much finer resolution data to be able to elucidate specific soil differences and weather/climate data. The country level data is useful for looking at trends but is not robust enough for managing yield gap solutions as an accurate measure of the yield gap due to management. Therefore, more country partners with finer scale data are needed to make progress on the quantification of yield gaps, the factors which constrain growers achieving water limited yields, and potential strategies for reducing the yield gap with true G × E × M synergies (Hatfield and Beres, 2019). For example, the most recent study on yield gaps of wheat in Australia conducted in-silico experiments to determine the impact on grain yield of sub-optimal practices and tested these against emerging best practices as developed by agronomists (Hochman and Horan, 2018). With this approach, it was possible to quantify reasons for the yield gap. Average national losses were due to growers not applying enough nitrogen, failure to adopt conservation tillage techniques, suboptimal weed control during summer fallow, low seed densities, and delayed sowing. Moreover, the GYGA atlas determined that the yield gap that is related to management for Australia is around 50%. Irrespective of the method to define yield gaps, progress toward closure in the future will require local producers to adopt practices that increase their climate resilience in wheat production systems (Hatfield and Beres, 2019).

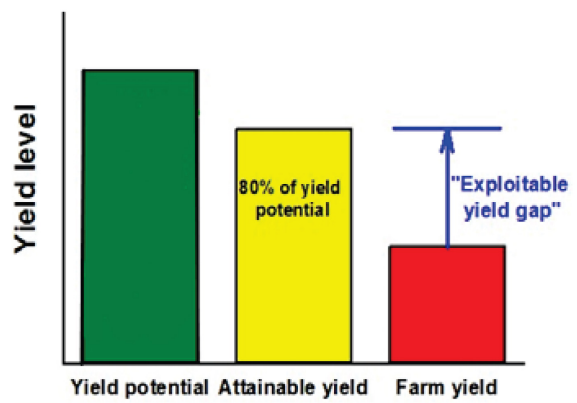


FIGURE 1 | Representation of yield potential, attainable yield, farm yield, and yield gap (Van Ittersum et al., 2013). Appropriate permissions have been obtained from the copyright holder(s) of this work.

Specific to durum yield gaps and for the purposes of this paper, the GYGA methodology was utilized recently along with existing data to calculate durum yield gaps for specific durum producing regions (**Figure 2**). France and Mexico represent regions with the smallest yield gap, with some regions in France nearly achieving yield gap closure. The higher yield potential in Mexico is a reflection of production under irrigation, which can largely eliminate the limitation caused by insufficient water supply from precipitation and stored soil moisture at sowing. Regions within Italy, Greece, and Cyprus with the highest yield potential also displayed actual yields that were not measurably different from the lowest yield potential regions, which creates a wide yield gap exceeding 50%. This suggests large room at the farm level to improve yield *via* improved agronomic management.

Global Challenges for Yield Attainment – Biotic Threats

While insect threats such as wheat stem sawfly *Cephus cinctus* Norton (*Hymenoptera: Cephidae*) and orange blossom

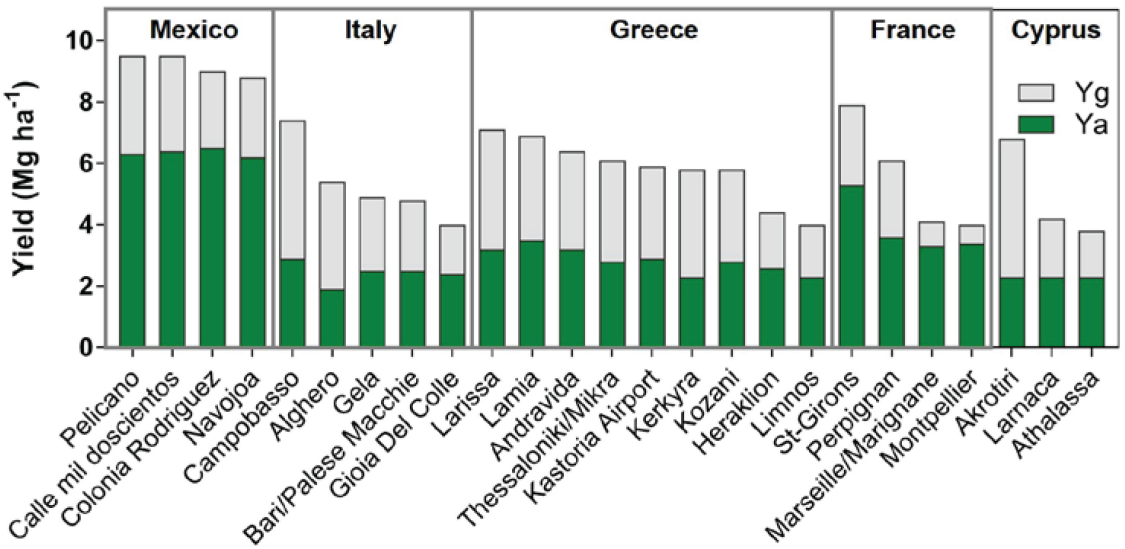


FIGURE 2 | Yield potential for durum wheat in different producing countries. The green and gray portions of the bar correspond to the actual durum yield (Ya) and the yield gap (Yg), respectively. Wheat was rainfed in all cases except for Mexico, where it was managed with irrigation. Source: Global Yield Gap Atlas. Original work.

wheat midge (*Sitodiplosis mosellana*) can be major production constraints, their cyclical nature includes years to a decade where their presence is below economic thresholds. Fungal diseases, however, and most notably FHB caused by *Fusarium graminearum*, have had the greatest impact on the durum industry in the Canadian Prairies and the Great Plains of United States as well as other wheat-growing regions (Beres et al., 2018). Durum is also particularly vulnerable to crown rot (*Fusarium pseudograminearum, Fp*), tan spot (*Drechslera tritici-repentis*), Septoria leaf blotch (*Mycosphaerella graminicola*), bacterial leaf streak (*Xanthomonas translucens*), leaf rust (*Puccinia triticina*), and stem rust (*Puccinia graminis*), and stripe rust or yellow rust (*Puccinia striiformis*) which causes the most damage when pustules develop early on the flag leaf and disrupt photosynthesis (Ransom et al., 2017). While many diseases can potentially reduce the yield and quality of durum, integrated management strategies exist that can help reduce the effects of these diseases. However, genetic progress to develop cultivars resistant to diseases lags behind other wheat classes in both North America (Ransom et al., 2017) and Australia (Hollaway et al., 2013). In Australia, durum cultivars with a high level of resistance to *Fp* have not been identified. This has reduced durum production to farming systems with lower levels of crown rot. Farming practices such as retaining more cereal residue and increasing the intensity of wheat rotations can increase the amount of inoculum in the soil, which can further increase the likelihood of infection (Summerell et al., 1989). With key durum growing regions of Australia predicted to become hotter and drier, yield losses due to crown rot disease are expected to increase. While the development of varieties with improved crown rot resistance has been shown to reduce inoculum levels in soils, improved management solutions must focus on limited exposure to water stress as a key G × E × M strategy to reduce yield loss from crown rot in Australia. One proposed strategy is the development of durum cultivars with improved frost tolerance and subsequently amenable to earlier planting, which would minimize exposure to terminal drought and water stress during grain fill.

Weeds are one of the largest contributors to wheat yield loss. Weed interference can result in spring wheat yield loss ranging from 5% to greater than 80% (Harker et al., 2017). Currently, a paucity of information is available highlighting the magnitude of weed-induced yield loss in durum compared with other wheat classes. However, the information that is available suggests that durum and bread wheat may have similar tolerance to weed interference. One exception showed greater grain yield loss induced by increasing densities of wild oat (*Avena fatua* L.) in a lath-house experiment in a bread wheat cultivar compared with a durum wheat cultivar; suggesting that durum wheat could have greater tolerance to weed interference (Henson and Jordan, 1982). In their study, densities of wild oat equivalent to 162, 406, and 812 plants m^{-2} reduced grain yield of durum wheat on average by 34, 53, and 67%, respectively, when durum wheat was sown at a density equivalent to 162 plants m^{-2}. A crop rotation study in a rainfed Mediterranean environment of central Italy concluded that weed control and nitrogen supply are among the most important factors impacting durum wheat yield and grain quality, and that these factors are of greater importance in years with excess rainfall and low temperatures during reproductive development of durum wheat in organic production systems (Campiglia et al., 2015).

Although the use of herbicides has enabled the adoption of conservation cropping systems, the increased reliance on herbicides for weed control has led to the evolution of herbicide resistance (Walsh and Powles, 2007). It has been estimated that overall, the total cost of weeds for growers with resistance can

increase by as much as \$55 ha^{-1} higher than those without resistance (Llewellyn et al., 2016). The major weed constraint to durum wheat production in Australia is annual ryegrass (*Lolium rigidum* L.). In South Australia, annual ryegrass has evolved considerable resistance to all post-emergent herbicides used in cereal crops (Boutsalis et al., 2012). Similarly in Canada, a recurring theme around weed dynamics in durum fields is the occurrence of herbicide-resistant biotypes of these species (Heap, 2019). In Alberta and Saskatchewan, Canada, separate green foxtail (*Setaria viridis* L.) populations have been found with resistance to acetyl-CoA carboxylase (ACCase)-inhibiting (group A/1), acetolactate synthase (ALS)-inhibiting (group B/2) and microtubule-inhibiting (group K1/3) herbicides. Wild oat populations with resistance to ACCase inhibitors, ALS inhibitors, or lipid synthesis (group N/8) inhibitors are also present in these areas. The increasing frequency of multiple herbicide-resistant wild oats in the Canadian prairies has raised concern over effective management of these populations in cereal and pulse crops. In Canadian durum, no post-emergence herbicide options exist currently for management of wild oat populations with multiple resistance to both the ACCase and ALS inhibitors, assuming cross-resistance to all active ingredients within this herbicide modes-of-action. This has increased grower reliance on pre-emergence herbicide options.

Global Challenges for Yield Attainment – Abiotic Threats

Abiotic stress is a phenomenon encountered in all durum production regions, but Australia is particularly vulnerable as it appears most likely to first experience climate change that will challenge yield and quality targets for durum. Agronomic and genetic solutions have been successful in mitigating yield lost to some degree but challenges remain; therefore, a considerable research effort is still required. Durum yield losses in Australia are primarily due to water stress (Liu et al., 2015) similar to well-published studies in wheat from lack of rainfall during spring, which causes a mild water deficit stress prior to anthesis, moderate stress at anthesis and becomes more severe during grain fill (French and Schultz, 1984). Breeding for genotypes adapted to pre- and post-anthesis water-deficit stress has been a major target of Australian breeders (Liu et al., 2015) and will need to remain a significant priority (Alahmad et al., 2019) given that durum production is in Mediterranean areas most likely influenced by climate change, and wheat yields in Australia are beginning to stall (Hochman et al., 2017).

Frost events in spring are also common, which causes yield loss by directly reducing the number of grains via sterility induced by the combined effect of cold, desiccation and freezing damage to the floral organs and developing grain (Al-Khatib and Paulsen, 1984; Boer et al., 1993; Fuller et al., 2007; Zheng et al., 2015). A large effort has been undertaken in Australia to screen for improved frost tolerance; however, Australia durum cultivars have a significantly higher susceptibility to frost damage than bread wheat and barley (*Hordeum vulgare* L.) (Cocks and Cullis, 2019). Equally high temperatures and heat shock during sensitive reproductive growth stages and grain fill can also result in a yield penalty (Gomez-Macpherson and Richards, 1995). A key study by Zubaidi et al. (2000) concluded the critical issue with durum wheat in South Australia was not its yield potential per se, but its ability to maintain yield when challenged by environmental stress. Agronomic solutions to maintain yield and reduce potential exposure to water stress, heat and frost can be achieved by manipulating plant development and sowing date (Bassu et al., 2009). A series of experiments in South Australia compared the relative performance of durum and bread wheat when exposed to the same level of abiotic stresses. When flowering at a similar time the yield gap ranged from 0.6 to 1.5 t/ha due to increased sterility from increased sensitivity of durum to abiotic stresses such as heat and reproductive frost (McCallum et al., 2019). The best strategy available to growers under these conditions is to ensure that flowering occurs during the optimum flowering period (Flohr et al., 2017) when the combined stresses of frost, drought and heat risk (Flohr et al., 2017) are minimized. A more recent environmental constraint, and a further example of potential negative effects of climate change, is the fact that autumn rainfall required to establish crops in Australia has diminished (Cai et al., 2012). Increased effort is therefore needed to develop genetics with diversity in crop phenology patterns to ensure growers can respond and sow earlier or later than currently practiced, and still achieve flower on time in order to overcome the yield decline (Kirkegaard and Hunt, 2010).

DESIGNING A RESILIENT DURUM CROPPING SYSTEM

Improvements to the "G" in the G × E × M Paradigm in Canada

The rate of genetic gain in yield of durum wheat in Canada averaged 0.63% (approximately 21.5 kg ha^{-1}) per year from 1963 to 2017 (**Figure 3**). There was an increase in the number of new cultivars brought into production during the last decade, reflecting the major increase in funding for breeding programs, principally by farmers. De Vita et al. (2007) reported a gain of 19.9 kg ha^{-1} $year^{-1}$ for durum yield in Italy from 1900 to 1990. Royo et al. (2007) reported a rate of gain of 0.36 to 0.44% per year in a historic set of Italian and Spanish durum cultivars. Investigation of rates of gain in a more recent period (1980–2009) in Spain showed a similar gain of 0.44% per year up to 2003, with little change thereafter (Chairi et al., 2018). Declining rates of genetic yield gain, much less doubling current rates, could be a concern in future, and widening the genetic diversity of crossing programs is necessary, and is a priority of breeding programs unless new genetic diversity is discovered.

Retrospective studies of the factors underlying genetic improvement in grain yield identified key contributing traits. For example, increased harvest index (the ratio of grain weight to total plant biomass) has contributed to yield improvement in durum (De Vita et al., 2007) and bread wheat (Beche et al., 2014). This is largely associated with the introduction of semi-dwarf cultivars, now grown exclusively in many countries, although less so in the northern Plains area of North America, particularly in

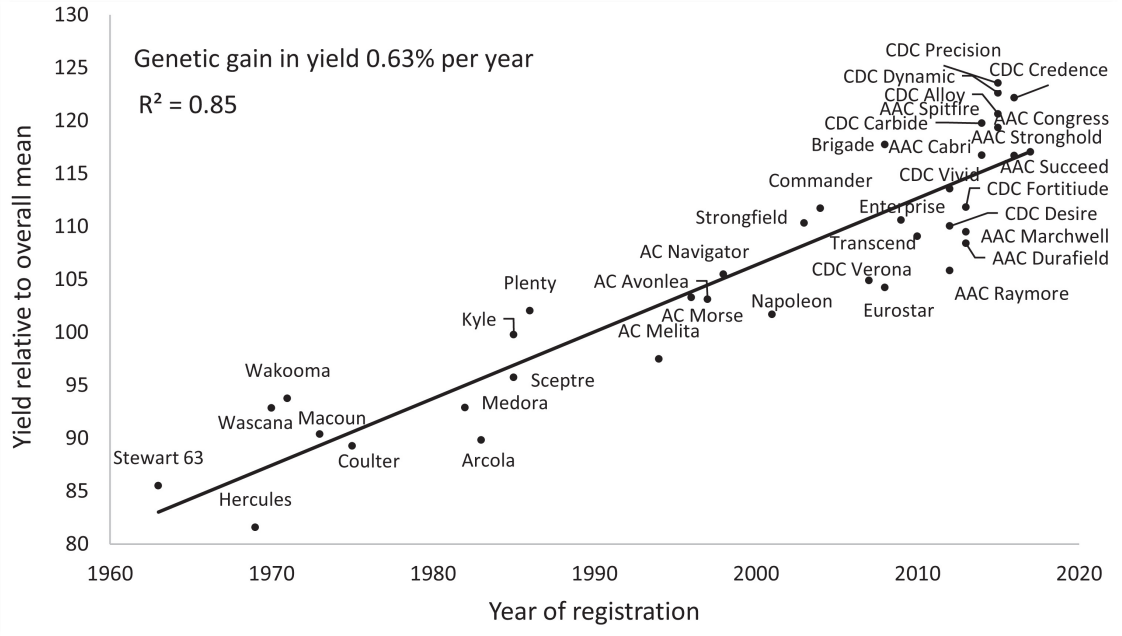


FIGURE 3 | Durum wheat genetic gain of grain yield from 1963 to 2017 in Canada. Source: Prairie Recommending Committee for Wheat, Rye, and Triticale (http://pgdc.ca/committees_wrt_pd.html). Appropriate permissions have been obtained from the copyright holder(s) of this work.

the case of durum. Future increases in harvest index are unlikely because it is impractical to further shorten straw length due to issues of mechanical harvest, so yield increases must come from improved overall biomass, coupled with resistance to lodging. An exhaustive long list of possible selection targets for potential yield and methods for their selection, have accumulated (Fischer and Rebetzke, 2018). Yield increases over time were associated with improvements in plant physiological processes such as higher stomatal conductance (leading to lower leaf canopy temperature) and photosynthetic rates (Fischer et al., 1998; De Vita et al., 2007; Beche et al., 2014) and associated traits such as leaf chlorophyll content (Beche et al., 2014). Manipulation of factors related to photosynthetic efficiency might contribute significantly to future yield gains (Martin et al., 2011), but the probability of success and associated timeline are highly uncertain. Up to the present, physiological traits were improved indirectly through selection for grain yield because direct measurement of such traits was limited by the technology available to measure them. However, research suggests co-selection of physiological traits with yield *per se* can maximize genetic gain, and the availability of new tools associated with ground-based or aerial platforms is driving new interest in assessing crop canopy traits such as transpiration, chlorophyll content, and leaf area, in addition to plant phenology, in high throughput phenotyping systems (White et al., 2012). This can now be done at the scale required for breeding programs (Araus and Cairns, 2014). Combining these data with genomic selection, made possible by next-generation sequencing and SNP genotyping technologies, might help increase the rate of genetic gain compared to traditional breeding approaches (Rutkoski et al., 2016; Crain et al., 2018).

Resistance to disease continues to be a major factor in the maintenance or improvement of durum wheat yields. Leaf rust, yellow or stripe rust and stem rust are globally important biotic constraints to wheat production (Eversmeyer and Kramer, 2000). Other leaf diseases such as tan spot (caused by *Pyrenophora tritici-repentis*) and Stagonospora nodorum blotch and *Septoria tritici* blotch limit yields, which can be mitigated by genetic resistance. For example, Fernandez et al. (2010) demonstrated that a 16% reduction in tan spot symptoms increased durum grain yield by 17%. FHB is a major disease of durum wheat, capable of causing yield reduction and loss of marketability of infected grain. Both genetic resistance and crop management can mitigate the effects and contribute to incremental increases in yield potential (Beres et al., 2018).

Going forward, the challenge is to vastly increase rates of genetic gain in wheat yield. The underlying key to achieving genetic gain for grain yield is having appropriate genetic variability from which to select. Past success is largely based on recombining alleles within elite germplasm and is now supported by advances in genomic selection (Haile et al., 2018; Montesinos-López et al., 2019) and high-throughput phenotyping (Condorelli et al., 2018). Hybrid wheat breeding is another approach to improving durum wheat, with indications of hybrids with 10% greater yield than the mid-parental value (Gowda et al., 2010). However, an effective pollination control system for hybrid seed production remains elusive. Thus, real progress will require an infusion of new genetic diversity into breeding programs. This can be done by crossing with related landraces or wild relatives, recognizing strategies to minimize linkage drag of undesirable traits will be required. Introgressions from wild

relatives have been used successfully for disease resistance, although this has usually entailed many cycles of backcrossing and selection to develop agronomically suitable cultivars. New genomic technologies may be able to assist in the introgression process and shorten the time to cultivar release (Dempewolf et al., 2017), particularly if coupled with approaches to speed generation acceleration (Alahmad et al., 2018). Genome editing is a promising technology that will allow precise generation of new allelic variants for use in breeding. The clustered regularly interspersed short palindromic repeats (CRISPR/Cas9) system was demonstrated to increase grain size in durum and common wheat (Zhang et al., 2016). The CRISPR/Cas9 system requires precise determination of the desirable allele for mutation, which in turn requires accurate genome sequence annotation. Annotation of wild emmer (*Triticum dicoccoides*) and durum wheat genome assemblies will assist in this process (Avni et al., 2017; Maccaferri et al., 2019). However, routine application of gene editing requires extensive research to link gene annotation with phenotypic function (Barabaschi et al., 2016; Adamski et al., 2020).

The development of strong international collaborations during the last decade, such as the Wheat Initiative[2] and associated projects such as the 10+ Wheat Genome project[3] has significantly increased our understanding of the wheat genome. This knowledge paves the way for future research that will link genes to phenotypic function. Maccaferri et al. (2019) recently demonstrated the linkage for the example trait, grain cadmium concentration. Understanding multi-gene (quantitative) traits such as grain yield will be more challenging, but ultimately achievable.

Examples of Innovations in the "M" of the G × E × M Paradigm

Seeding Systems – A Re-think on Sowing Density

The role of management is to limit or overcome losses due to abiotic and biotic factors with the aim to achieve the genetic potential for grain yield. A fundamental step is the consideration of planting density, which is often under-utilized in the context of optimizing genetic yield potential, imparting yield stability and improving crop uniformity and competitive ability. Decisions around optimal sowing density are often influenced by convenience, past practice, equipment, and the cost of purchasing seed of new cultivars. This decision-making process can lead to less than ideal seeding rates as seed input costs may be perceived as cost-prohibitive if yield potential is not achieved. Thus, despite increases in genetic yield potential, sowing densities during the 1970s through to the 1990s was fairly static at a rate of fewer than 200 seeds m^{-2} (Grant et al., 1974; Fowler, 1982; Entz and Fowler, 1991). However, research in wheat (Beres et al., 2010a,b, 2011) indicates higher rates are needed similar to experiences in other crops such as corn where increases in sowing density were needed before grain yield of corn hybrids could reach potential and eclipse older conventional corn cultivars (Duvick, 2005). In wheat, Beres et al. (2011) reported a positive linear improvement to grain yield in durum at rates as high as 450 seeds m^{-2}, which was more than double the rate of standard practices at the time. The potential of higher seeding rates to exploit the yield of new durum genetics was confirmed in follow-up studies and is now the norm in western Canada (Beres et al., 2010b, 2011; Nilsen et al., 2016; Isidro-Sánchez et al., 2017; Ye et al., 2017).

The adoption of high seeding rates also reduces the potential for weed competition and can be an important tool for disease management as crop uniformity can influence the prevalence of some diseases (Beres et al., 2018). Variation in crop uniformity generally prolongs the flowering period of durum wheat for the most sensitive time for FHB infection (Buerstmayr et al., 2019), resulting in greater disease severity in the crop. Higher sowing densities in wheat reduces tillering and results in more main stems in a field, which will shorten flowering and days to maturity. Greater uniformity would also improve fungicide efficacy as crop staging to determine appropriate application windows would be facilitated greatly if the crop was less variable (Beres et al., 2018).

Seeding Systems – Seed Treatments for Abiotic Stress Resistance

Seed-applied fungicidal or insecticidal applications are traditionally based on mitigating biotic stress caused by insects or seed/soil-borne pathogens (Hewett and Griffiths, 1986; Mathre et al., 2001). McMullen and Stack (2009) observed a substantial improvement in germination and seedling vigor even when diseased seeds are used. Vernon et al. (2009) reported improved wheat stand with imidacloprid, in the presence of wireworms. Recently, the importance of seed treatments has been proposed not only in managing the biotic stressors but in mitigating abiotic stressors like heat stress, drought, wind desiccation, and frost, which usually arise from cold ambient and soil temperatures at planting and emergence.

A study of soybean (*Glycine max* L.) seed treated with thiamethoxam reported accelerated germination and larger seedlings concomitant with buffering against the negative effects of water deficit (Cataneo et al., 2010). A Canadian study on eastern Canadian spring wheat reported that a dual fungicide (difenoconazole and metalaxyl) and an insecticide (thiamethoxam) enhanced the freezing tolerance of seedlings (Larsen and Falk, 2013). Ford et al. (2010) established that neonicotinoids such as imidacloprid and clothianidin induce salicylic acid-associated responses, which elicit plant protection to pathogens such as powdery mildew concomitant with abiotic stress tolerance.

Recent work in wheat in western Canada with dual fungicide/insecticidal seed treatments have all reported improvements to crop stand establishment and grain yield in wheat production systems (Beres et al., 2016; Turkington et al., 2016; Ye et al., 2017). The same seed treatments also improved crop vigor and yield stability when integrated into a system with factors related to seed size and sowing density. The most notable responses occurred in the weakest agronomic system with thinner seeds and low seeding rates. An economic analysis supported that a dual (fungicide/insecticidal) seed treatment provides greater gross returns (CAN+$31 ha^{-1}) as well as

[2]https://www.wheatinitiative.org

[3]http://www.10wheatgenomes.com/

improved net returns (+\$22 ha^{-1}) (Beres et al., 2016). Moreover, comparisons between spring and winter growth habits indicate greater responses in winter wheat. This suggests the prolonged period of abiotic stress associated with winter growth habits can be mitigated effectively with seed treatments to overcome the negative aspects of the "E".

Seeding Systems – "Ultra-Early" for Life Cycle Synchrony

Planting date and environment play an impactful role in durum wheat agronomic and end-use characteristics, regardless of cultivar grown (Forster et al., 2017). Historically, the recommended planting date for wheat on the Canadian Prairies has been early- to mid-May in southern regions and a targeted deadline of June 10 in the northern latitudes of the Parkland region. These dates were prescribed largely to meet crop insurance deadlines, or, are a product of on-farm logistics. For example, He et al. (2012) indicated that average commercial spring wheat planting dates were May 9, May 14, and May 15, for Swift Current, Saskatoon and Melfort, SK, respectively. These dates appear late when compared with earlier planting dates they predicted using a DSSAT-CSM model. This is largely to ensure the crop has developed sufficient biomass so as to capture maximum radiation by June 21 when photoperiod tends to peak on the Prairies. Early planting, therefore, is an important integrated crop management strategy designed to optimize genetics and to fully synchronize crop phenology with maximum solar radiation and environmental conditions that achieve high attainable grain yield and quality (Beres and Wang, 2018). Durum wheat yield losses in Australia are primarily due to water stress (Liu et al., 2015) similar to well-published studies in wheat from lack of rainfall during spring which causes a mild water deficit stress prior to anthesis, moderate stress at anthesis and becomes more severe during grain fill (French and Schultz, 1984). Agronomic solutions to maintain yield and reduce potential exposure to water stress, heat and frost can be achieved by manipulating plant development and sowing date (Bassu et al., 2009; Collier et al., 2020). To facilitate the timing of optimal flowering, growers need to match a genotype development speed with sowing date. An emerging trend in Australia has been the shift towards earlier planting systems to overcome the negative impacts of climate change (Hunt et al., 2019b). However, there is less variation in flowering time of the commercially available durum wheat cultivars relative to bread wheat in Australia, which is restricting a grower's ability to adopt this strategy for durum production.

Forster et al. (2017) conducted a study on durum wheat in North Dakota, United States and concluded early planting improved grain yield regardless of cultivar or environmental conditions during vegetative and reproductive growth. Early planting helps to enhance soil moisture usage, which is an important factor for growing durum and can be profitable for the environments with low and high rainfall as well. They also concluded that quality diminishes in late plantings regarding semolina extraction, gluten index, and wet gluten values and a significant reduction in test weight and grain yield on different cultivars with delayed planting. Conversely, protein and kernel yellow pigment contents, vitreous kernels, and falling number were more related to cultivar and did not depend on planting date and environment.

Research in Canada has reported a number of benefits associated with an "ultra-early" planting regime whereby soil temperature in the top 5 cm of the soil surface dictates when planting should occur instead of a traditional calendar date. The authors report that grain yield was not compromised and often maximized when seeding occurred at around 2^oC soil temperature. A greater reduction in grain yield was observed when planting was delayed until soil reach 10^oC, despite extreme environmental conditions after initial seeding, including air temperatures as low as -10.2^oC and as many as 37 nights with air temperatures below 0^oC (Collier et al., 2020). An opportunity associated with such a system that results in earlier maturity is a refined integrated pest management strategy. For example, one major pest of durum wheat is the orange wheat blossom midge, which attacks wheat during anthesis. An ultra-early planting strategy using soil temperature as a trigger facilitates an asynchrony between the vulnerable host plant phenological stage and the primary window of pest infestation (Collier et al., 2020). The same management strategy is used by producers to mitigate FHB infection such that earlier flowering coincides with less optimal environment conditions for fungal infection. Early planting coupled with an early-maturity cultivar managed with practices that optimize crop uniformity might lessen the probability of FHB infection and reduce the risks associated with early fall frost events (Beres et al., 2018).

CASE STUDIES OF G × E × M

Synergies of genetic gain and agronomic management are critically important in achieving higher yields, including the adoption of the new high-yield cultivars, applying precision farming, optimizing the nutrient application, zero tillage, appropriate seeding rates, and irrigation management. Overall, the interaction of genetics and agronomy management strategies has resulted in up to 1.2% increase in yield per year over time in Canada (**Figure 4**), approximately double of the rate of the genetic gains alone (Clarke et al., 2010). While there is no evidence yet that rates of genetic gain have begun to decline, sustaining current rates long-term, or further increasing rates of gain, will require new genetic diversity, breeding technologies, and strategies. To double overall production, assuming the current balance of genetic and agronomic improvement is maintained, the rate of genetic gain in grain yield will have to be increased to greater than 1% per year.

Fusarium Head Blight Management – Experiences in Canada and the United States

Fusarium head blight is a devastating disease of wheat and a serious production and marketing constraint for durum production. It can cause significant yield losses concomitant with reductions to durum quality via development of mycotoxins in the seed if the fungal pathogen fully develops without

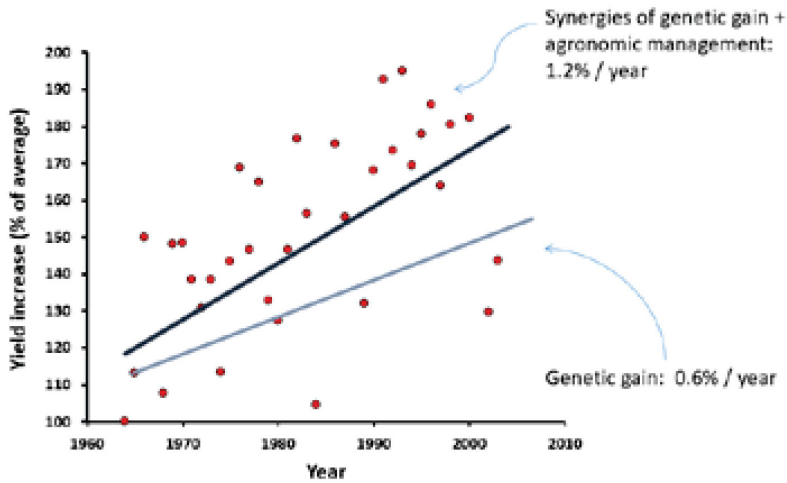


FIGURE 4 | The role of agronomy management and genetics on yield increase over time (Clarke et al., 2010). Appropriate permissions have been obtained from the copyright holder(s) of this work.

intervention. The mycotoxin, deoxynivalenol or DON, is present in the grain and causes serious downgrades when delivered to market. Moreover, once DON levels exceed 2 ppm, markets that would accept the grain are limited. *Fusarium graminearum*, the causal organism of FHB, overwinters on infected residues of small grains and corn. Synchrony between the release of spores and flowering of the subsequent wheat crop phase ensures a high rate of infection within the spike/spikelet, leading to high rates of seed infection. The spread of conidia to the spike is facilitated with conditions that transport spores upward from the colonized stubble (splashing), and/or through release of ascospores into the air where windborne dispersal occurs to neighboring fields. The cycle perpetuates through the annual colonization of host tissues and when infected seed is planted without mitigation tactics. Conditions most favorable for the development of FHB are high humidity, frequent rainfall and relatively warm night temperatures at heading, especially in regions where host crop residues are present. Therefore, inoculum is rarely considered a limiting factor in the development of this disease. Currently, the recommendations for FHB control include growing durum after a broadleaf crop, using an approved fungicide at flowering, and grow resistant varieties if available. Though there are notable differences between varieties in their susceptibility to FHB, there are currently no varieties that are considered resistant or even moderately resistant (Beres et al., 2018).

Since 2000, there has been significant investment in breeding FHB resistant cultivars. In Canada, cultivars are rated along a resistance continuum as follows: "Susceptible" (S) (e.g., Strongfield), "Moderately Susceptible" (MS) (e.g., CDC Credence, Brigade, and Transcend), "Intermediate" (I), "Moderately Resistant" (MR), and "Resistant" (R). To date, the best cultivar rating remains at the level of "MS" (Beres et al., 2018). In the Canadian Prairies, Manitoba is the most vulnerable region due to frequent occurrences of FHB outbreaks; therefore, little to no durum is grown in that region. Some progress has occurred with respect to FHB resistance as MS

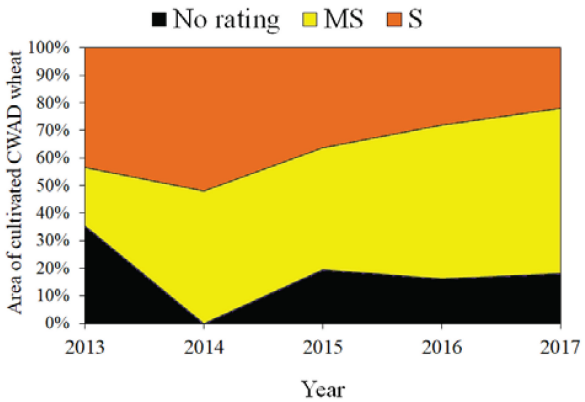


FIGURE 5 | Durum cultivar adoption in the Canadian Prairies based on FHB rating. Original work.

cultivars have now been released. Since 2015, there has been an increase in the cultivated area of the MS varieties compared to other varieties. The rate of adoption of cultivars and respective ratings for FHB in the Canadian Prairies during 2013 to 2017 is presented in **Figure 5**. Producers have displaced susceptible cultivars with improved resistance, such that in 2017, 78% of cultivated area of durum wheat were planted with cultivars rated "MS" (Agriculture Financial Services Corporation [AFSC], 2020; Manitoba Agricultural Services Corporation [MASC], 2020; The Western Producer, 2020). Obviously, improvements to resistance that results in the release of I or MR cultivars is critically important to stabilize or expand both the production area and market access.

Given the lack of genetic control options, greater attention to management strategies is required, particularly when durum seed is sourced from a seed lot with greater than 10% Fusarium damaged kernels. Mitigation strategies are most successful when

the disease cycle is interrupted at or prior to the flowering stage to prevent spore dispersal and/or host infection. While the main management strategy for FHB mostly focusses on in-crop fungicide applications to control FHB, seed-applied fungicides have been recognized as an effective method to improve yield and stand establishment when FHB stress exists (Ye et al., 2017). Durum wheat seed treatments that apply difenconazole, triticonazole, maneb, or fludioxonil have been found to significantly improve germination and reduce Fusarium seedling blight in field trials with 5–45 % levels of infection, but with no significant improvements in yield (Jørgensen et al., 2012). Other wheat trials have reported that seed treatments increased emergence and yield when levels of seed infection were high (>50%), with no improved emergence and grain yield in low levels of seed infection (≤10%), and increased emergence with no change in grain yield in moderate levels of infection (25–35%) (May et al., 2010).

The lack of genetic resistance in durum varieties to FHB and poor efficacy with management practices has led to a reduction of durum production in North Dakota and the migration of durum to the drier parts of the state. In order to improve efficacy of management practices, research was initiated in North Dakota that created a G × M context by exploring genetic aspects with the inclusion of both susceptible cultivars and a breeding line with some level of known resistance. Experiments were conducted in four environments in 2017 through 2019. These experiments consisted of five or six genotypes (released varieties and one advanced line) that were either treated with fungicide or not. Yield and DON levels were measured after harvest. The released varieties are commonly grown in North Dakota and their current rating of FHB tolerance is listed. The fungicide treatment consisted of applying prothioconazole plus tebuconazole when most of the spikes had reached 50% flowering (**Table 2**).

The response of genotype and fungicide on DON levels varied considerably between environments. In Carrington 2017, DON was affected by the interactions of wheat variety and fungicide (**Table 2**). This was due to the more susceptible genotypes having high levels of DON (more than 6 pm) when no fungicide was applied, and all genotypes having similarly low levels of DON (less than 1 ppm) when treated with fungicides. This location was in contrast to two environments where DON levels when averaged across genotypes, were reduced by only 0.4 to 0.6 ppm DON with the application of a fungicide but similar to the 2019 data. Genotypes differed significantly across all environments for DON levels. Within the released varieties, no one variety consistently had the lowest DON. Though there were significant interactions between variety and fungicide treatment, the moderately resistant varieties were still responsive to fungicides, showing the value of integrating control practices. The yield was only measured at two locations. Yield increased dramatically with the application of fungicide at the Carrington 2017 location, but only modestly at the Prosper location. Only at the location exhibiting the highest FHB infection was there a genotype by fungicide interaction for yield.

While fungicide-based options are the primary management response, other components should be considered. When designing crop rotations, the "1 in 4" principle is particularly important for FHB mitigation (Kutcher et al., 2011, 2013), whereby a susceptible crop phase occurs only once in every 4 crop years. Moreover, sequencing durum after a broadleaf crop offers advantages such as: potential increases in yield; reductions in disease pressure such as tan spot and Septoria; a reduction in inoculum load of *Fusarium graminearum*; preventative buildup of soil-borne pathogens; rotation of herbicide chemistry diversity, which helps delay the buildup of herbicide-resistant weeds (Knox, 2018). The need for uniformity has been discussed in the context of the influencing factors of seeding rate and seed treatments; however, crop nutrition such as appropriate levels of phosphorous contribute greatly to crop uniformity (Hooker et al., 2016).

Durum is particularly vulnerable to FHB infection when produced under irrigated conditions. However, alterations to irrigation management, particularly during pre- and post-anthesis, is an effective strategy for controlling FHB. Flowering is an important phenological stage for durum wheat, which typically begins 3 days after the head has fully emerged and lasts for about 3–5 days. This period also coincides with maximum water uptake requirements by the crop. It is recommended that irrigation should be terminated for 8–10 days during flowering to reduce humidity in the durum canopy. Reducing water availability during this critical phase of crop development could compromise yield potential. However, durum grown on loam or clay loam soils may tolerate 10 dry days without significantly impacting yield, if the 50–100 cm of soil in the root zone is at field capacity just before flowers emerge (McKenzie and Woods, 2018). Thus, intensification of irrigation management is required in order to mitigate FHB without compromising yield.

The over-arching principle for FHB management is the manipulation of agronomic factors that facilitate completion of critical crop developmental phases, such as flowering, while doing so rapidly and uniformly, as a consequence of early sowing and increased seeding rates. Experiences in Canada and the United States illustrate the importance of linking together multiple management factors. For example, a management strategy of foliar fungicide and/or ST+foliar fungicide generally produced higher yields with greater stability, particularly for susceptible cultivars in high FHB environments (Ye et al., 2017). These strategies and the adoption of practices involving proper fungicide selection, and optimal application timings and methods will lead to improved yield stability and quality in high risk environments (Beres et al., 2018; Newlands, 2018). This is critically important for durum as it would help to overcome the necessary use of moderately susceptible cultivars until such time that breeding can catch up and deploy genetics with improved levels of resistance.

Integrated Approaches to Weed Management

Cultural weed management plays an important role in managing herbicide-resistant weeds. Durum cultivar selection and agronomic management (G × M) can improve the ability for durum to compete with weeds and thus reduce selection

TABLE 2 | Effect of fungicides on DON levels and yield in four environments in North Dakota, 2017 and 2018.

	Carrington 2017				Carrington 2018		Prosper 2018				Prosper 2019	
	DON		Yield		DON		DON		Yield		DON	
	None	Fung.	None	Fung.	None	Fung.	None	Fung.	None	Fung.	None	Fung.
	(ppm)		(t ha^{-1})		(ppm)		(ppm)		(t ha^{-1})		(ppm)	
Carpio (MS)†	6.7	0.8	3.57	4.88	0.7	0.6	3.3	4.1	3.46	3.95	9.0	9.0
D09555 (–)	6.4	0.5	3.50	5.71	1.7	0.7	3.2	3.0	3.49	3.80	11.1	8.6
Divide (MS)	3.9	0.6	3.63	4.68	2.4	0.9	3.2	3.5	3.76	3.75	10.5	4.9
Joppa (MS)	6.3	0.6	3.44	4.95	2.1	1.4	4.9	2.5	3.29	3.93	11.1	4.3
Mountrail (S)	6.2	0.9	3.19	4.36	2.0	1.0	2.6	3.3	3.50	3.80	15.7	5.0
Tioga (MS/S)	6.4	0.4	3.78	4.93	0.6	1.0	3.8	2.4	3.72	4.07		
Mean	6.0	0.6	3.52	4.92	1.6	0.9	3.5	3.1	3.54	3.88	11.5	6.4
Fung.‡	**†		**		*		NS		**		*	
Variety	**		**		**		**		**		*	
V × F	**		0.06		**		0.08		NS		*	

†*MS, moderately susceptible, S, susceptible.*
‡*Level of statistical significance of the factor within the row. * = 0.05, ** = 0.01, NS = not significant.*

pressure for herbicide resistance. Several physiological traits are linked to the ability of wheat to compete with weeds, including early-season vigor, tillering, leaf area index, plant height, and allelopathic potential, among others. Fewer physiological traits impact the ability of durum wheat cultivars to compete with weeds compared with bread wheat cultivars (Zerner et al., 2008). This is likely due to differences in growth habit among durum and bread wheat. In general, durum cultivars exhibit reduced tillering and leaf area index; two traits that are often associated with early canopy closure. The ability for durum to compete with weeds is associated closely with cultivar height, where taller cultivars are generally more competitive (Zerner et al., 2008; Beres et al., 2010b; Giambalvo et al., 2010). For example, Zerner et al. (2008) studied how the height of near-isogenic wheat lines (NILs) contributed to the ability for wheat to compete with oats (*Avena sativa* L.). In their study, tall bread wheat and durum wheat NILs reduced oat seed production by 26 and 41%, respectively. Similarly, a study of three durum wheat cultivars in the presence and absence of a surrogate weed [barley (*Hordeum vulgare* L.)] showed differences in the ability to compete with weeds among the durum cultivars (Giambalvo et al., 2010). In their study, the ability for durum wheat to compete with the surrogate weed was associated with the stature of the cultivar (i.e., taller plant stature contributed to greater ability to compete with weeds) and the nutrient uptake ability of the cultivar, which reduces nutrients available to competing weeds. Much less is known about allelopathy of durum wheat compared with bread wheat, and the contribution of allelopathy to genotypic differences in the ability for durum wheat to suppress weeds is a current knowledge gap (Oueslati, 2003; Fragasso et al., 2013).

Experiences in North America

Wheat is an important rotational crop, often grown in Canada as the dominant cereal in crop rotations with other commodities including pulses and oilseeds. Pulses like field pea (*Pisum sativum* L.) or lentil (*Lens culinaris* Medic.) are sensitive to many herbicides, leaving limited options for chemical weed control. Many of the herbicides that are effective in these crops (e.g., the ALS inhibitors) have been rendered ineffective on several weed species with evolved resistance to active ingredients within this herbicide mode-of-action. Alberta and Saskatchewan, Canada, alone is home to 18 weed species with confirmed resistance to ALS-inhibiting herbicides (Beckie et al., 2017; Heap, 2019). The cereal phase in the crop rotation provides an excellent opportunity to manage ALS inhibitor-resistant weeds and facilitate inclusion of high-value pulses in crop rotations. In durum, post-emergent application of synthetic auxins (group O/4), photosystem II inhibitors (nitriles) (group C3/6), or (4-hydroxyphenylpyruvate dioxygenase) HPPD inhibitors (group F2/27) can be an excellent rotational weed management strategy, however re-cropping restrictions must be considered when using some of these active ingredients (Geddes et al., 2018). Durum wheat is unique from common wheat as it can be more sensitive to certain herbicides such as pyroxasulfone (Soltani et al., 2012).

Optimizing plant spatial arrangement can help improve the ability for durum to compete with weeds. Some reports indicate that narrower row spacing (5 cm vs. 15 cm vs. 25 cm) improves the weed competitive ability of durum over increased seeding densities (De Vita et al., 2017). Moreover, row widths >25 cm reduced yield of wheat. The greater impact of row spacing compared with seeding density is likely due to reduced tillering in durum compared with bread wheat cultivars, which could limit the extent to which durum occupies inter-row niche space compared with bread wheat classes. However, increased durum seeding rates may be a beneficial management tool in zero- or minimum-tillage production systems, where moderate (20–30 cm) row spacings are commonplace. Indeed, increased durum seeding densities are correlated positively with increased leaf area index (Isidro-Sánchez et al., 2017), which can hasten crop canopy closure and reduce the quality of light available to weeds beneath

the crop canopy. Optimal plant spatial arrangement by reducing row spacing and using high durum seeding densities (≥400 seeds m^{-2}) can improve the ability for durum to compete with weeds and thus reduce selection pressure for herbicide resistance in durum production.

Recent advances in multispectral or hyperspectral imaging technologies hold promise for real-time and site-specific weed management. Several problematic weeds in durum, including wild oat and annual ryegrass, can be successfully discriminated from durum wheat using multispectral imaging in the 400–900 nm range (Lopez-Granados et al., 2008). These methods hold promise for mapping of weed patches in durum, allowing for cost-effective herbicide application including additional effective sites-of-action on herbicide-resistant weed patches.

According to weed surveys of Alberta and Saskatchewan, Canada conducted between 2009 and 2017, the most abundant weeds found in durum wheat were green foxtail, wild oat, volunteer canola [*Brassica napus* L.), stinkweed (*Thlaspi arvense* L.), and wild buckwheat (*Fallopia convolvulus* (L.) Á. Löve] (Geddes et al., 2018). Within these weeds, green foxtail, wild oat, and wild buckwheat have been known as the most abundant weeds in all crops in the prairies since the 1970s. Both green foxtail and wild oat are spring annuals that cause yield loss in durum depending on emergence timing of the weeds and weather conditions (Douglas et al., 1985; Beckie et al., 2012). Green foxtail exhibits seed dormancy at maturity for approximately 10 weeks and wild oat exhibits it for almost one year, resulting in moderately persistent seed banks. Green foxtail is highly responsive to nitrogen fertilizer and can be managed using suitable fertilization techniques like mid- or side-row banding. Volunteer canola is an annual weed, with seeds in the soil seed bank maintaining viability for about three years. Presence of this weed is often caused by large seed losses at canola harvest, resulting in glyphosate, glufosinate, or imidazolinone resistant populations in subsequent crops (Gulden et al., 2008). Disturbing soil shortly after canola harvest can reduce the density of volunteer canola originating from crop harvest losses (Geddes and Gulden, 2017). Wild buckwheat is a spring annual weed with seeds germinating within the first year in the seed bank, and surviving up to five years. It can cause significant crop lodging, grain sample contamination, and harvest difficulties. It has the ability to climb the canopy to capture light (Hume et al., 1983). Stinkweed has the ability to grow like a spring or winter annual, with viability for over 20 years in the soil and causing yield losses in many crops (Warwick et al., 2002).

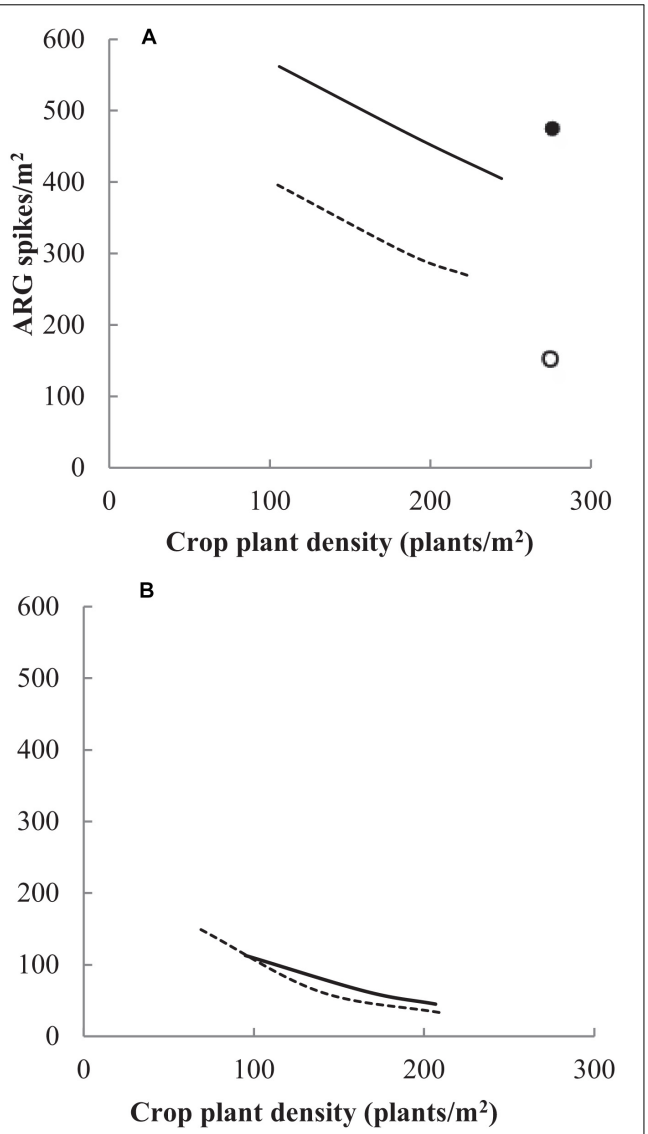


FIGURE 6 | (A) No pre-emergent herbicide and **(B)** treated with pre-emergent herbicide. The effect of the management combination of genotype, herbicide management, and plant density on the survival and density of annual ryegrass spikes (ARG) at Tarlee 2012 (LSD 5% = 55 spikes/m^2). Two durum varieties Saintly (○) and Yawa (●) in combination without a pre-emergent herbicide treatment (A) and treated with the pre-emergent herbicide BoxerGold at 2.5 L/ha incorporated by sowing (B) at three crop plant density levels. Original work.

Experiences in Australia

The evolution of herbicide resistance highlights the need for alternative forms of weed control including improving the competitiveness of the durum crop. The main aim is to increase the competitive ability of the crop against weed species, this not only means lower yield losses from weed competition but also greater suppression of weed growth and seed production (Bajwa et al., 2017). Agronomic or genetic intervention to improve crop competition works particularly well when the weed and crop are phenotypically similar, such as in the case of ryegrass and wheat (Lemerle et al., 2001).

Durum has typically reduced early vigor and crop competition with weeds when compared to other cereal crop options in Australia. A comprehensive study of bread wheat and durum wheat genotypes from all over the world found significant variation in the competitiveness of genotypes on annual ryegrass (Lemerle et al., 1996). Overall, bread wheat was more competitive against ryegrass than durum wheat. In Australia, there has also been improvements in plant breeding for crop vigor

in durum, the introduction of new durum cultivar DBA Aurora has demonstrated a similar ability to reduce ryegrass seed set through competition when compared to bread wheat (Goss and Wheeler, 2014).

To encourage integrated G × E × M approaches in South Australia, Porker and Wheeler (2012) demonstrated that the management combination of variety, seeding rate, and herbicide all play a significant role in the success in managing annual ryegrass (**Figure 6**). The pre-emergent herbicide combination of prosulfocarb plus S-metolachlor remains the main factor providing the greatest proportion of weed control. However combining this with more vigorous varieties such as Saintly and increasing seeding rates from 100 to 300 seeds m^{-2} resulted in increased yield, improved weed competition and reduced weed seed set (**Figure 6**). This is consistent with bread wheat examples and the recommendation for higher wheat seed rates as part of an integrated weed management strategy is now strongly promoted to farmers (Lemerle et al., 2004; Bajwa et al., 2017).

Sowing time also plays a key role in weed control. The benefit of earlier sowing across most of the southeast Australian wheat belt is now well-established (Flohr et al., 2018; Hunt et al., 2019b). It is often tempting for growers to delay sowing in order to achieve greater control of weeds with additional tillage; however, these practices have been shown to significantly contribute to the yield gap in Australian wheat (Hochman and Horan, 2018). Sowing earlier in warmer soils also enables wheat a competitive advantage against weeds (Gomez-Macpherson and Richards, 1995) and, when combined with effective pre-emergent herbicides, can provide adequate control of annual ryegrass and reduced seed production without compromising yield (Preston et al., 2017). Broader adoption of alternative, non-chemical management options in-conjunction with herbicide weed management programs is essential for sustainable durum crop production in Australia.

CONCLUSION

Globally, durum is still considered a minor wheat crop and typically much of the research effort on durum often is conducted in conjunction with studies of bread wheat. Moreover, management practices are often borrowed from research conducted on common wheat. While we have sought to present synergies possible with a G × E × M paradigm, it is clear a coordinated effort and cross-disciplinary approach is yet to be fully realized, which underscores the need for transformational research that exemplifies the G × E × M paradigm (Beres et al., 2020). In many areas of the world, climate change will drive future research and innovation, which requires a re-think of how we link together all the genetic and management components when designing a resilient cropping system (Fletcher et al., 2020). Rather than commonplace 'reactive agronomy', transformational research will be needed where there is greater emphasis on integration and optimization of the overall G × E × M system by the agronomist (Hunt et al., 2019a). Agronomists must be the leaders, the translators and the communicators, accessing the best discipline-based knowledge and expertise where relevant to deliver transformational change. Kirkegaard and Hunt (2010) demonstrated how a systems approach can transform a wheat-based cropping system averaging 1.6 Mg ha^{-1} with inherent sustainability issues into a resilient system that increased attainable yield by 3×. This represents the type of modified paradigm needed to address the very real threat of declining rates of yield growth (Fischer and Connor, 2018). A major component when defining resiliency will be the need for systems to be sustainable. Moreover, the sustainable intensification of cropping systems is considered the only feasible path to meet future food demands (durum no exception) as it is recognized that further conversions of natural ecosystems to farmland is not tenable (Cassman and Grassini, 2020). The Wheat Initiative[4] could provide the framework in charting this path forward as it brings together wheat experts from around the world to participate in cross-disciplinary teams and working groups to establish priorities for global wheat research.

Designing robust agronomic systems for durum demands scientific creativity and foresight based on a deep understanding of constitutive components and their innumerable interactions with each other and the environment. Advances in individual technologies (e.g., genetic improvements, new pesticides, seeding technologies) are of little benefit until they are melded creatively and thoughtfully into resilient Genotype × Environment × Management (G × E × M) systems that will flourish in the field under unpredictable conditions of prairie farmlands.

[4] www.wheatinitiative.org

AUTHOR CONTRIBUTIONS

BB and ER prepared and edited the original manuscript with input, revisions, and editorial contributions provided by all co-authors. JC, CG, PG, KP, and JR provided major contributions to the sections. All authors contributed to the article and approved the submitted version.

FUNDING

Funding for this publication was provided, in part, by the International Engagement Office of Agriculture and Agri-Food Canada. In-kind support provided by each author affiliation.

ACKNOWLEDGMENTS

We dedicate this review article to the memory of JC. JC was internationally recognized for his long list of accomplishments in the area of durum breeding and crop physiology. He was a world-class mentor and friend to many of us in the wheat world. JC and his work serve as an inspiration to many of us and we will miss him dearly.

REFERENCES

Adamski, N. M., Borrill, P., Brinton, J., Harrington, S. A., Marchal, C., Bentley, A. R., et al. (2020). A roadmap for gene functional characterisation in crops with large genomes: lessons from polyploid wheat. *eLife* 9:e55646.

Agriculture Financial Services Corporation [AFSC] (2020). *Yield Alberta*. Available online at: https://afsc.ca/resources/5684-2/?gclid=EAIaIQobChMI35XH1ZzP6QIVMQnnCh2dhA68EAAYASAAEgL-ufD_BwE (accessed May 22, 2020).

Alahmad, S., Dinglasan, E., Leung, K. M., Riaz, A., Derbal, N., Voss-Fels, K. P., et al. (2018). Speed breeding for multiple quantitative traits in durum wheat. *Plant Methods* 14:36.

Alahmad, S., El Hassouni, K., Bassi, F. M., Dinglasan, E., Youssef, C., Quarry, G., et al. (2019). A major root architecture QTL responding to water limitation in durum wheat. *Front. Plant Sci.* 10:436. doi: 10.3389/fpls.2019.00436

Al-Khatib, K., and Paulsen, G. M. (1984). Mode of high temperature injury to wheat during grain development. *Physiol. Plant.* 61, 363–368. doi: 10.1111/j.1399-3054.1984.tb06341.x

Araus, J. L., and Cairns, J. E. (2014). Field high-throughput phenotyping: the new crop breeding frontier. *Trends Plant Sci.* 19, 52–61. doi: 10.1016/j.tplants.2013.09.008

Avni, R., Nave, M., Barad, O., Baruch, K., Twardziok, S. O., Gundlach, H., et al. (2017). Wild emmer genome architecture and diversity elucidate wheat evolution and domestication. *Science* 357, 93–97.

Bajwa, A. A., Walsh, M. J., and Chauhan, B. S. (2017). Weed management using crop competition in Australia. *Crop Protect.* 95, 8–13. doi: 10.1016/j.cropro.2016.08.021

Barabaschi, D., Tondelli, A., Desiderio, F., Volante, A., Vaccino, P., Valè, G., et al. (2016). Next generation breeding. *Plant Sci.* 242, 3–13.

Bassu, S., Asseng, S., Motzo, R., and Giunta, F. (2009). Optimising sowing date of durum wheat in a variable Mediterranean environment. *Field Crops Res.* 111, 109–118. doi: 10.1016/j.fcr.2008.11.002

Beche, E., Benin, G., Da Silva, C. L., Munaro, L. B., and Marchese, J. A. (2014). Genetic gain in yield and changes associated with physiological traits in Brazilian wheat during the 20th century. *Eur. J. Agron.* 61, 49–59. doi: 10.1016/j.eja.2014.08.005

Beckie, H., Shirriff, S., and Leeson, J. (2017). *Saskatchewan Weed Survey of Herbicide-Resistant Weeds in 2014-2015. Weed Survey Series Publication 17-1*. (Saskatoon: Agriculture and Agri-Food Canada).

Beckie, H. J., Francis, A., and Hall, L. M. (2012). The biology of Canadian Weeds. 27. *Avena fatua* L. (updated). *Can. J. Plant Sci.* 92, 1329–1357. doi: 10.4141/cjps2012-005

Beres, B. L., Brûlé-Babel, A. L., Ye, Z., Graf, R. J., Turkington, T. K., Harding, M. W., et al. (2018). Exploring Genotype × Environment × Management synergies to manage fusarium head blight in wheat. *Can. J. Plant Pathol.* 40, 179–188. doi: 10.1080/07060661.2018.1445661

Beres, B. L., Cárcamo, H. A., Yang, R. C., and Spaner, D. M. (2011). Integrating spring wheat sowing density with variety selection to manage wheat stem sawfly. *Agron. J.* 103, 1755–1764. doi: 10.2134/agronj2011.0187

Beres, B. L., Clayton, G. W., Harker, K. N., Stevenson, C. F., Blackshaw, R. E., and Graf, R. J. (2010a). A sustainable management package to improve winter wheat production and competition with weeds. *Agron. J.* 102, 649–657. doi: 10.2134/agronj2009.0336

Beres, B. L., Harker, K. N., Clayton, G. W., Bremer, E., Blackshaw, R. E., and Graf, R. J. (2010b). Weed competitive ability of spring and winter cereals in the Northern Great Plains. *Weed Technol.* 24, 108–116. doi: 10.1614/wt-d-09-00036.1

Beres, B. L., Hatfield, J. L., Kirkegaard, J. A., Eigenbrode, S. D., Pan, W. L., Lollato, R. P., et al. (2020). Towards a better understanding of Genotype × Environment × Management interactions – a global Wheat Initiative agronomic research strategy. *Front. Plant Sci.* 11:828. doi: 10.3389/fpls.2020.00828

Beres, B. L., Turkington, T. K., Kutcher, H. R., Irvine, B., Johnson, E. N., O'donovan, J. T., et al. (2016). Winter wheat cropping system response to seed treatments, seed size, and sowing density. *Agron. J.* 108, 1101–1111. doi: 10.2134/agronj2015.0497

Beres, B. L., and Wang, Z. (2018). "Planting date and seeding rate strategies to optimize genetics and overcome environmental challenges," in *Sustainable Production of Durum Wheat in Canada*, (Bannockburn, IL: Barilla America Inc), 23–28.

Boer, R., Campbell, L., and Fletcher, D. (1993). Characteristics of frost in a major wheat-growing region of Australia. *Aust. J. Agric. Res.* 44, 1731–1743.

Boutsalis, P., Gill, G. S., and Preston, C. (2012). Incidence of Herbicide Resistance in Rigid Ryegrass (*Lolium rigidum*) across Southeastern Australia. *Weed Technol.* 26, 391–398. doi: 10.1614/wt-d-11-00150.1

Buerstmayr, M., Steiner, B., and Buerstmayr, H. (2019). Breeding for Fusarium head blight resistance in wheat—Progress and challenges. *Plant Breed.* 139, 429–454. doi: 10.1111/pbr.12797

Cai, W., Cowan, T., and Thatcher, M. (2012). Rainfall reductions over Southern Hemisphere semi-arid regions: the role of subtropical dry zone expansion. *Sci. Rep.* 2:702.

Campiglia, E., Mancinelli, R., De Stefanis, E., Pucciarmati, S., and Radicetti, E. (2015). The long-term effects of conventional and organic cropping systems, tillage managements and weather conditions on yield and grain quality of durum wheat (*Triticum durum* Desf.) in the Mediterranean environment of Central Italy. *Field Crops Res.* 176, 34–44. doi: 10.1016/j.fcr.2015.02.021

Cassman, K. G. (1999). Ecological intensification of cereal production systems: yield potential, soil quality, and precision agriculture. *Proc. Natl. Acad. Sci. U.S.A.* 96, 5952–5959. doi: 10.1073/pnas.96.11.5952

Cassman, K. G., Dobermann, A., Walters, D. T., and Yang, H. (2003). Meeting cereal demand while protecting natural resources and improving environmental quality. *Annu. Rev. Environ. Resour.* 28, 315–358.

Cassman, K. G., and Grassini, P. (2020). A global perspective on sustainable intensification research. *Nat. Sustain.* 3, 262–268. doi: 10.1038/s41893-020-0507-8

Cataneo, A. C., Ferreira, L. C., Carvalho, J. C., Andréo-Souza, Y., Corniani, N., Mischan, M. M., et al. (2010). Improved germination of soybean seed treated with thiamethoxam under drought conditions. *Seed Sci. Technol.* 38, 248–251. doi: 10.15258/sst.2010.38.1.27

Cessna, A. J., Darwent, A. L., Kirkland, K. J., Townley-Smith, L., Harker, K. N., and Lefkovitch, L. P. (1994). Residues of glyphosate and its metabolite AMPA in wheat seed and foliage following preharvest applications. *Can. J. Plant Sci.* 74, 653–661. doi: 10.4141/cjps94-117

Chairi, F., Vergara-Diaz, O., Vatter, T., Aparicio, N., Nieto-Taladriz, M. T., Kefauver, S. C., et al. (2018). Post-green revolution genetic advance in durum wheat: the case of Spain. *Field Crops Res.* 228, 158–169. doi: 10.1016/j.fcr.2018.09.003

Clarke, J. M., Clarke, F. R., and Pozniak, C. J. (2010). Forty-six years of genetic improvement in Canadian durum wheat cultivars. *Can. J. Plant Sci.* 90, 791–801. doi: 10.4141/cjps10091

Cocks, N., and Cullis, B. (2019). *ANFP Wheat MET Analysis 2010-2017. Centre for Bioinformatics and Biometrics (CBB) Technical Report Series*. Keiraville: Institute for Applied Statistics Research Australia.

Collier, G. R., Spaner, D. M., Graf, R. J., and Beres, B. L. (2020). The integration of spring and winter wheat genetics with agronomy for ultra-early planting into cold soils. *Front. Plant Sci.* 11:89. doi: 10.3389/fpls.2020.00089

Condorelli, G. E., Maccaferri, M., Newcomb, M., Andrade-Sanchez, P., White, J. W., French, A. N., et al. (2018). Comparative aerial and ground based high throughput phenotyping for the genetic dissection of NDVI as a proxy for drought adaptive traits in durum wheat. *Front. Plant Sci.* 9:893. doi: 10.3389/fpls.2018.00893

Crain, J., Mondal, S., Rutkoski, J., Singh, R. P., and Poland, J. (2018). Combining high-throughput phenotyping and genomic information to increase prediction and selection accuracy in wheat breeding. *Plant Genome* 11, 1–14. doi: 10.3835/plantgenome2017.05.0043

CRP-WHEAT (2016). *Wheat Agri-Food Systems Proposal 2017-2022". Research Program on Wheat, (CGIAR)*. Available online at: https://cgspace.cgiar.org/handle/10947/4421?show=full

Darwent, A. L., Kirkland, K. J., Townley-Smith, L., Harker, K. N., Cessna, A. J., Lukow, O. M., et al. (1994). Effect of preharvest applications of glyphosate on the drying, yield and quality of wheat. *Can. J. Plant Sci.* 74, 221–230. doi: 10.4141/cjps94-046

De Vita, P., Colecchia, S. A., Pecorella, I., and Saia, S. (2017). Reduced inter-row distance improves yield and competition against weeds in a semi-dwarf durum wheat variety. *Eur. J. Agron.* 85, 69–77. doi: 10.1016/j.eja.2017.02.003

De Vita, P., Nicosia, O. L. D., Nigro, F., Platani, C., Riefolo, C., Di Fonzo, N., et al. (2007). Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. *Eur. J. Agron.* 26, 39–53. doi: 10.1016/j.eja.2006.08.009

Dempewolf, H., Baute, G., Anderson, J., Kilian, B., Smith, C., and Guarino, L. (2017). Past and future use of wild relatives in crop breeding. *Crop Sci.* 57, 1070–1082. doi: 10.2135/cropsci2016.10.0885

Douglas, B. J., Morrison, I. N., Thomas, A. G., and Maw, M. G. (1985). THE BIOLOGY OF CANADIAN WEEDS.: 70. *Setaria viridis* (L.) Beauv. *Can. J. Plant Sci.* 65, 669–690. doi: 10.4141/cjps85-089

Duvick, D. (2005). Genetic progress in yield of United States maize (*Zea mays* L.). *Maydica* 50, 193–202.

Entz, M. H., and Fowler, D. B. (1991). Agronomic performance of winter versus spring wheat. *Agron. J.* 83, 527–532. doi: 10.2134/agronj1991.00021962008300030002x

Espe, M. B., Cassman, K. G., Yang, H., Guilpart, N., Grassini, P., Van Wart, J., et al. (2016). Yield gap analysis of US rice production systems shows opportunities for improvement. *Field Crops Res.* 196, 276–283. doi: 10.1016/j.fcr.2016.07.011

Evans, L. T. (1993). "Processes, genes, and yield potential," in *International Crop Science I*, eds D. R. Buxton, R. Shibles, R. A. Forsberg, B. L. Blad, K. H. Asay, G. M. Paulsen, et al. (Madison, WI: Crop Science Society of America).

Evans, L. T., and Fischer, R. A. (1999). Yield potential: its definition, measurement, and significance. *Crop Sci.* 39, 1544–1551. doi: 10.2135/cropsci1999.3961544x

Eversmeyer, M. G., and Kramer, C. L. (2000). Epidemiology of wheat leaf and stem rust in the central Great Plains of the USA. *Annu. Rev. Phytopathol.* 38, 491–513. doi: 10.1146/annurev.phyto.38.1.491

FAO (2017). *The Future of Food and Agriculture - Trends and Challenges*. Rome: Food and Agriculture of the United Nations.

Fernandez, M. R., Clarke, F. R., Knox, R. E., Clarke, J. M., and Singh, A. K. (2010). Quantification of effects of leaf spotting diseases on grain yield and market quality of durum wheat using near-isogenic lines. *Can. J. Plant Pathol.* 32, 177–187. doi: 10.1080/07060661003740025

Fischer, R., and Connor, D. (2018). Issues for cropping and agricultural science in the next 20 years. *Field Crops Res.* 222, 121–142.

Fischer, R., and Rebetzke, G. (2018). Indirect selection for potential yield in early-generation, spaced plantings of wheat and other small-grain cereals: a review. *Crop Pasture Sci.* 69, 439–459. doi: 10.1071/cp17409

Fischer, R. A., Rees, D., Sayre, K. D., Lu, Z. M., Condon, A. G., and Larque Saavedra, A. (1998). Wheat yield progress associated with higher stomatal conductance and photosynthetic rate, and cooler canopies. *Crop Sci.* 38, 1467–1475. doi: 10.2135/cropsci1998.0011183x003800060011x

Fletcher, A. L., Chen, C., Ota, N., Lawes, R. A., and Oliver, Y. M. (2020). Has historic climate change affected the spatial distribution of water-limited wheat yield across Western Australia? *Clim. Change* 159, 347–364. doi: 10.1007/s10584-020-02666-w

Flohr, B. M., Hunt, J. R., Kirkegaard, J. A., and Evans, J. R. (2017). Water and temperature stress define the optimal flowering period for wheat in south-eastern Australia. *Field Crops Res.* 209, 108–119. doi: 10.1016/j.fcr.2017.04.012

Flohr, B. M., Hunt, J. R., Kirkegaard, J. A., Evans, J. R., Trevaskis, B., Zwart, A., et al. (2018). Fast winter wheat phenology can stabilise flowering date and maximise grain yield in semi-arid Mediterranean and temperate environments. *Field Crops Res.* 223, 12–25. doi: 10.1016/j.fcr.2018.03.021

Ford, K. A., Casida, J. E., Chandran, D., Gulevich, A. G., Okrent, R. A., Durkin, K. A., et al. (2010). Neonicotinoid insecticides induce salicylate-associated plant defense responses. *Proc. Natl. Acad. Sci. U.S.A.* 107, 17527–17532. doi: 10.1073/pnas.1013020107

Forster, S. M., Ransom, J. K., Manthey, F. A., Rickertsen, J. R., and Mehring, G. H. (2017). Planting date, seeding rate, and cultivar impact agronomic traits and semolina of durum wheat. *Am. J. Plant Sci.* 8, 2040–2055. doi: 10.4236/ajps.2017.89137

Fowler, D. B. (1982). Date of seeding, fall growth, and winter survival of winter wheat and rye. *Agron. J.* 74, 1060–1063. doi: 10.2134/agronj1982.00021962007400060030x

Fragasso, M., Iannucci, A., and Papa, R. (2013). Durum wheat and allelopathy: toward wheat breeding for natural weed management. *Front. Plant Sci.* 4:375. doi: 10.3389/fpls.2013.00375

French, R., and Schultz, J. (1984). Water use efficiency of wheat in a Mediterranean-type environment. I. The relation between yield, water use and climate. *Aust. J. Agric. Res.* 35, 743–764. doi: 10.1071/ar9840743

Fuller, M. P., Fuller, A. M., Kaniouras, S., Christophers, J., and Fredericks, T. (2007). The freezing characteristics of wheat at ear emergence. *Eur. J. Agron.* 26, 435–441. doi: 10.1016/j.eja.2007.01.001

Geddes, C. M., and Gulden, R. H. (2017). Early autumn soil disturbance decreases persistence of volunteer summer-annual oilseed rape (*Brassica napus*). *Weed Res.* 57, 182–192. doi: 10.1111/wre.12248

Geddes, C. M., Tidemann, B. D., Wolf, T., and Johnson, E. N. (2018). "Integrated weed management to minimize yield losses," in *Sustainable Production of Durum Wheat in Canada*, (Bannockburn, IL: Barilla America Inc), 50–57.

Giambalvo, D., Ruisi, P., Di Miceli, G., Frenda, A. S., and Amato, G. (2010). Nitrogen use efficiency and nitrogen fertilizer recovery of durum wheat genotypes as affected by interspecific competition. *Agron. J.* 102, 707–715. doi: 10.2134/agronj2009.0380

Gomez-Macpherson, H., and Richards, R. A. (1995). Effect of sowing time on yield and agronomic characteristics of wheat in south-eastern Australia. *Aust. J. Agric. Res.* 46, 1381–1399. doi: 10.1071/ar9951381

Goss, S., and Wheeler, R. (2014). *Results; Weed Competition Determining Best Management Practices in Durum Wheat [Online]. Hart Field Site*. Available: http://www.hartfieldsite.org.au/media/2014%20Trial%20Results/2014_Results_Weed_competition_determining_best_management_preactices_in_durum_wheat.pdf (accessed May 09, 2019).

Gowda, M., Kling, C., Würschum, T., Liu, W., Maurer, H. P., Hahn, V., et al. (2010). Hybrid breeding in durum wheat: Heterosis and combining ability. *Crop Sci.* 50, 2224–2230. doi: 10.2135/cropsci2009.10.0637

Grant, M. N., Pittman, U. J., Horricks, J. S., Holmes, N. D., Anderson, D. T., and Smith, A. D. (1974). *Winter Wheat Production in Western Canada*. Ottawa, ON: Agriculture and Agri-Food Canada.

Grassini, P., Van Bussel, L. G. J., Van Wart, J., Wolf, J., Claessens, L., Yang, H., et al. (2015). How good is good enough? Data requirements for reliable crop yield simulations and yield-gap analysis. *Field Crops Res.* 177, 49–63. doi: 10.1016/j.fcr.2015.03.004

Gulden, R. H., Warwick, S. I., and Thomas, A. G. (2008). The Biology of Canadian Weeds. 137. *Brassica napus* L. and *B. rapa* L. *Can. J. Plant Sci.* 88, 951–996. doi: 10.4141/cjps07203

Haile, J. K., N'diaye, A., Clarke, F., Clarke, J., Knox, R., Rutkoski, J., et al. (2018). Genomic selection for grain yield and quality traits in durum wheat. *Mol. Breed.* 38:75.

Harker, K. N., O'donovan, J., and Tidemann, B. (2017). "Integrated weed management in wheat cultivation," in *Achieving Sustainable Cultivation of Wheat*, ed. P. Langridge (Cambridge: Burleigh Dodds Science Publishing Limited), 23.

Hatfield, J. L., and Beres, L. B. (2019). Yield gaps in wheat: path to enhancing productivity. *Front. Plant Sci.* 10:1603. doi: 10.3389/fpls.2019.01603

He, Y., Wang, H., Qian, B., Mcconkey, B., and Depauw, R. (2012). How early can the seeding dates of spring wheat be under current and future climate in Saskatchewan, Canada? *PLoS One* 7:e45153. doi: 10.1371\%2Fjournal.pone.0045153

Heap, I. (2019). *The International Survey of Herbicide Resistant Weeds*. Available online at: www.weedscience.org (accessed June 11, 2019).

Henson, J. F., and Jordan, L. S. (1982). Wild Oat (*Avena fatua*) competition with wheat (*Triticum aestivum* and *T. turgidum* durum) for Nitrate. *Weed Sci.* 30, 297–300. doi: 10.1017/s004317450004056x

Hewett, P. D., and Griffiths, D. C. (1986). "Biology of seed treatment," in *Seed Treatment*, ed. K. A. Jeffs (Surrey: BCPC Publications, Thornten Health), 7–12.

Hochman, Z., Gobbett, D., Holzworth, D., Mcclelland, T., Van Rees, H., Marinoni, O., et al. (2013). Reprint of "Quantifying yield gaps in rainfed cropping systems: a case study of wheat in Australia". *Field Crops Res.* 143, 65–75. doi: 10.1016/j.fcr.2013.02.001

Hochman, Z., Gobbett, D. L., and Horan, H. (2017). Climate trends account for stalled wheat yields in Australia since 1990. *Glob. Change Biol.* 23, 2071–2081. doi: 10.1111/gcb.13604

Hochman, Z., and Horan, H. (2018). Causes of wheat yield gaps and opportunities to advance the water-limited yield frontier in Australia. *Field Crops Res.* 228, 20–30. doi: 10.1016/j.fcr.2018.08.023

Hollaway, G. J., Evans, M. L., Wallwork, H., Dyson, C. B., and Mckay, A. C. (2013). Yield loss in cereals, caused by *Fusarium culmorum* and *F. pseudograminearum*, is related to fungal DNA in soil prior to planting, rainfall, and cereal type. *Plant Dis.* 97, 977–982. doi: 10.1094/pdis-09-12-0867-re

Hooker, D. C., Tenuta, A., Tamburic-Ilincic, L., and Schaafsma, A. W. (2016). "Advances in agronomic management of FHB in 20 years," in *Proceedings of the 3rd Canadian Wheat Symposium*, Ottawa, ON.

Hume, L., Martinez, J., and Best, K. (1983). The Biology of Canadian Weeds.: 60. *Polygonum convolvulus* L. *Can. J. Plant Sci.* 63, 959–971. doi: 10.4141/cjps83-121

Hunt, J. R., Kirkegaard, J., Celestina, C., and Porker, K. (2019a). "Transformational agronomy: restoring the role of agronomy in modern agricultural research," in *Australian Agriculture in 2020: From Conservation to Automation*, eds J. Pratley and J. Kirkegaard (Wagga Wagga: Agronomy Australia and Charles Sturt University), 373–388.

Hunt, J. R., Lilley, J. M., Trevaskis, B., Flohr, B. M., Peake, A., Fletcher, A., et al. (2019b). Early sowing systems can boost Australian wheat yields despite recent climate change. *Nat. Clim. Change* 9, 244–247. doi: 10.1038/s41558-019-0417-9

International Grains Council [IGC] (2020). *World Grain Statistics 2016*. Available: https://www.igc.int/en/subscriptions/subscription.aspx (accessed May 21, 2020).

Isidro-Sánchez, J., Perry, B., Singh, A. K., Wang, H., Depauw, R. M., Pozniak, C. J., et al. (2017). Effects of seeding rate on durum crop production and physiological responses. *Agron. J.* 109, 1981–1990. doi: 10.2134/agronj2016.09.0527

Jenks, B., Christoffers, M., Dalley, C., Endres, G., Gramig, G., Howatt, K., et al. (2019). *2019 North Dakota Weed Control Guide*. Fargo, ND: North Dakota State University.

Jørgensen, L. N., Nielsen, L. K., and Nielsen, B. J. (2012). Control of seedling blight in winter wheat by seed treatments – impact on emergence, crop stand, yield and deoxynivalenol. *Acta Agric. Scand. B Soil Plant Sci.* 62, 431–440.

Kirkegaard, J. A., and Hunt, J. R. (2010). Increasing productivity by matching farming system management and genotype in water-limited environments. *J. Exp. Bot.* 61, 4129–4143. doi: 10.1093/jxb/erq245

Kniep, J. (2008). *Durum Wheat Production*. Orange: NDOP Industries.

Knox, R. E. (2018). "Disease management to minimize crop loss and maximize quality," in *Sustainable Production of Durum Wheat in Canada*, (Bannockburn, IL: Barilla America Inc), 58–62.

Kutcher, H. R., Brandt, S. A., Smith, E. G., Ulrich, D., Malhi, S. S., and Johnston, A. M. (2013). Blackleg disease of canola mitigated by resistant cultivars and four-year crop rotations in western Canada. *Can. J. Plant Pathol.* 35, 209–221. doi: 10.1080/07060661.2013.775600

Kutcher, H. R., Johnston, A. M., Bailey, K. L., and Malhi, S. S. (2011). Managing crop losses from plant diseases with foliar fungicides, rotation and tillage on a Black Chernozem in Saskatchewan, Canada. *Field Crops Res.* 124, 205–212. doi: 10.1016/j.fcr.2011.05.018

Larsen, R. J., and Falk, D. E. (2013). Effects of a seed treatment with a neonicotinoid insecticide on germination and freezing tolerance of spring wheat seedlings. *Can. J. Plant Sci.* 93, 535–540. doi: 10.4141/cjps2012-127

Lemerle, D., Cousens, R. D., Gill, G. S., Peltzer, S. J., Moerkerk, M., Murphy, C. E., et al. (2004). Reliability of higher seeding rates of wheat for increased competitiveness with weeds in low rainfall environments. *J. Agric. Sci.* 142, 395–409. doi: 10.1017/s002185960400454x

Lemerle, D., Gill, G. S., Murphy, C. E., Walker, S. R., Cousens, R. D., Mokhtari, S., et al. (2001). Genetic improvement and agronomy for enhanced wheat competitiveness with weeds. *Aust. J. Agric. Res.* 52, 527–548. doi: 10.1071/ar00056

Lemerle, D., Verbeek, B., Cousens, R. D., and Coombes, N. E. (1996). The potential for selecting wheat varieties strongly competitive against weeds. *Weed Res.* 36, 505–513. doi: 10.1111/j.1365-3180.1996.tb01679.x

Liu, H., Searle, I. R., Mather, D. E., Able, A. J., and Able, J. A. (2015). Morphological, physiological and yield responses of durum wheat to pre-anthesis water-deficit stress are genotype-dependent. *Crop Pasture Sci.* 66, 1024–1038. doi: 10.1071/cp15013

Llewellyn, R., Ronning, D., Clarke, M., Mayfield, A., Walker, S. R., and Ouzman, J. (2016). *Impact of Weeds in Australian Grain Production*. Canberra: Grains Research and Development Corporation.

Lobell, D. B., Cassman, K. G., and Field, C. B. (2009). Crop yield gaps: their importance, magnitudes, and causes. *Annu. Rev. Environ. Resour.* 34, 179–204. doi: 10.1146/annurev.environ.041008.093740

Lopez-Granados, F., Pena-Barragan, J. M., Jurado-Exposito, M., Francisco-Fernandez, M., Cao, R., Alonso-Betanzos, A., et al. (2008). Multispectral classification of grass weeds and wheat (*Triticum durum*) using linear and nonparametric functional discriminant analysis and neural networks. *Weed Res.* 48, 28–37. doi: 10.1111/j.1365-3180.2008.00598.x

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51, 885–895.

Manitoba Agricultural Services Corporation [MASC] (2020). *Yield Manitoba*. Available online at: https://www.masc.mb.ca/masc.nsf/mmpp_index.html (accessed May 22, 2020).

Manthey, F. A., Chakraborty, M., Peel, M. D., and Pederson, J. D. (2004). Effect of preharvest applied herbicides on breadmaking quality of hard red spring wheat. *J. Sci. Food Agric.* 84, 441–446. doi: 10.1002/jsfa.1676

Marin, F. R., Martha, G. B. Jr., Cassman, K. G., and Grassini, P. (2016). Prospects for increasing sugarcane and bioethanol production on existing crop area in Brazil. *Bioscience* 66, 307–316. doi: 10.1093/biosci/biw009

Martin, A. J. P., Reynolds, M., Salvucci, M. E., Raines, C., Andralojc, P. J., Zhu, X.-G., et al. (2011). Raising yield potential of wheat. II. Increasing photosynthetic capacity and efficiency. *J. Exp. Bot.* 62, 453–467. doi: 10.1093/jxb/erq304

Mathre, D. E., Johnston, R. H., and Grey, W. E. (2001). *Small Grain Cereal Seed Treatment. The American Phyotpathological Society*. Available online at: http://www.apsnet.org/edcenter/advanced/topics/Pages/CerealSeedTreatment.aspx (accessed October 23, 2019).

May, W. E., Fernandez, M. R., and Lafond, G. P. (2010). Effect of fungicidal seed treatments on the emergence, development, and grain yield of Fusarium graminearum-infected wheat and barley seed under field conditions. *Can. J. Plant Sci.* 90, 893–904. doi: 10.4141/cjps09173

McCallum, M., Porker, K., and Peice, C. (2019). "What drives the yield gap between durum and bread wheat?" in *Proceedings of the 19th Australian Agronomy Conference*, (Wagga Wagga: Wagga Australian Society of Agronomy).

McKenzie, R. H., and Woods, S. A. (2018). "Irrigation management of durum wheat," in *Sustainable Production of Durum Wheat in Canada*, (Bannockburn, IL: Barilla America Inc), 41–47.

McMullen, M., and Stack, R. (2009). *Improving Wheat and Durum Seedling Stands and Vigor by Fungicide Seed Treatments*. Fargo, ND: North Dakota State University.

Merlos, F. A., Monzon, J. P., Mercau, J. L., Taboada, M., Andrade, F. H., Hall, A. J., et al. (2015). Potential for crop production increase in Argentina through closure of existing yield gaps. *Field Crops Res.* 184, 145–154. doi: 10.1016/j.fcr.2015.10.001

Montesinos-López, O. A., Montesinos-López, A., Tuberosa, R., Maccaferri, M., Sciara, G., Ammar, K., et al. (2019). Multi-trait, multi-environment genomic prediction of durum wheat with genomic best linear unbiased predictor and deep learning methods. *Front. Plant Sci.* 10:1311. doi: 10.3389/fpls.2019.01311

Newlands, N. K. (2018). Model-based forecasting of agricultural crop disease risk at the regional scale, integrating airborne inoculum, environmental, and satellite-based monitoring data. *Front. Environ. Sci.* 6:63. doi: 10.3389/fenvs.2018.00063

Nilsen, K. T., Clarke, J. M., Beres, B. L., and Pozniak, C. J. (2016). Sowing density and cultivar effects on pith expression in solid-stemmed durum wheat. *Agron. J.* 108, 219–228. doi: 10.2134/agronj2015.0298

Oueslati, O. (2003). Allelopathy in two durum wheat (*Triticum durum* L.) varieties. *Agric. Ecosyst. Environ.* 96, 161–163. doi: 10.1016/s0167-8809(02)00201-3

Porker, K. D., and Wheeler, R. (2012). *Improving Grass Weed Control in Durum*. Adelaide: SA Research and Development Institute.

Preston, C., Kleemann, S. G. L., Noack, S., Hooper, P., and Gill, G. S. (2017). *Using Time of Sowing of Wheat and Pre-Emergent Herbicides to Control Annual Ryegrass (Lolium rigidum)*. Warragul: Australian Society of Agronomy Inc.

Ranieri, R. (2015). Geography of the durum wheat crop. *Pastaria Int.* 6, 24–36.

Ranieri, R., Worden, G., Seghezzo, M. L., and Mills, C. (2012). "CHAPTER 15 - Marketing Perspectives in the Durum Wheat Trade," in *Durum Wheat*, 2nd Edn, eds M. Sissons, J. Abecassis, B. Marchylo, and M. Carcea (Saint Paul, MN: AACC International Press), 279–286. doi: 10.1016/b978-1-891127-65-6.50020-9

Ransom, J. K., Endres, G., Forster, S., Friskop, A. J., Franzen, D. W., Zollinger, R. K., et al. (2017). *Field Guide to Sustainable Production of High-Quality Durum Wheat in North Dakota*. Fargo, ND: North Dakota State University.

Reynolds, M. P., Quilligan, E., Aggarwal, P. K., Bansal, K. C., Cavalieri, A. J., Chapman, S. C., et al. (2016). An integrated approach to maintaining cereal productivity under climate change. *Glob. Food Sec.* 8, 9–18.

Royo, C., Álvaro, F., Martos, V., Ramdani, A., Isidro, J., Villegas, D., et al. (2007). Genetic changes in durum wheat yield components and associated traits in Italian and Spanish varieties during the 20th century. *Euphytica* 155, 259–270. doi: 10.1007/s10681-006-9327-9

Rutkoski, J., Poland, J., Mondal, S., Autrique, E., Pérez, L. G., Crossa, J., et al. (2016). Canopy temperature and vegetation indices from high-throughput phenotyping improve accuracy of pedigree and genomic selection for grain yield in wheat. *G3* 6, 2799–2808. doi: 10.1534/g3.116.032888

Soltani, N., Shropshire, C., and Sikkema, P. H. (2012). Response of spring planted cereals to pyroxasulfone. *Int. Res. J. Plant Sci.* 3, 113–119.

Sopiwnyk, E. (2018). "Durum production and consumption, a global perspective," in *Sustainable Production of Durum Wheat in Canada*, (Bannockburn, IL: Barilla America Inc), 5–9.

Statistics Canada (2019). *September Estimates of Production of Principal Field Crops, Canada*. Available: https://www150.statcan.gc.ca/n1/daily-quotidien/191206/t001b-eng.htm (accessed May 15, 2020).

Summerell, B. A., Burgess, L. W., and Klein, T. A. (1989). The impact of stubble management on the incidence of crown rot of wheat. *Aust. J. Exp. Agric.* 29, 91–98. doi: 10.1071/ea9890091

The Western Producer (2020). *Yield Saskatchewan*. Available online at: https://www.producer.com/pub/yield-saskatchewan/ (accessed May 22, 2020).

Timsina, J., Wolf, J., Guilpart, N., Van Bussel, L. G. J., Grassini, P., Van Wart, J., et al. (2018). Can Bangladesh produce enough cereals to meet future demand? *Agric. Syst.* 163, 36–44. doi: 10.1016/j.agsy.2016.11.003

Turkington, T. K., Beres, B. L., Kutcher, H. R., Irvine, B., Johnson, E. N., O'donovan, J. T., et al. (2016). Winter wheat yields are increased by seed treatment and fall-applied fungicide. *Agron. J.* 108, 1379–1389. doi: 10.2134/agronj2015.0573

Van Bussel, L. G. J., Grassini, P., Vn Wart, J., Wolf, J., Claessens, L., Yang, H., et al. (2015). From field to atlas: upscaling of location-specific yield gap estimates. *Field Crops Res.* 177, 98–108. doi: 10.1016/j.fcr.2015.03.005

Van Ittersum, M. K., Cassman, K. G., Grassini, P., Wolf, J., Tittonell, P., and Hochman, Z. (2013). Yield gap analysis with local to global relevance — a review. *Field Crops Res.* 143, 4–17. doi: 10.1016/j.fcr.2012.09.009

Van Ittersum, M. K., Van Bussel, L. G. J., Wolf, J., Grassini, P., Van Wart, J., Guilpart, N., et al. (2016). Can sub-Saharan Africa feed itself? *Proc. Natl. Acad. Sci. U.S.A.* 113, 14964–14969.

Van Oort, P. A. J., Saito, K., Tanaka, A., Amovin-Assagba, E., Van Bussel, L. G. J., Van Wart, J., et al. (2015). Assessment of rice self-sufficiency in 2025 in eight African countries. *Glob. Food Sec.* 5, 39–49. doi: 10.1016/j.gfs.2015.01.002

Vernon, R. S., Van Herk, W. G., Clodius, M., and Harding, C. (2009). Wireworm management I: stand protection versus wireworm mortality with wheat seed treatments. *J. Econ. Entomol.* 102, 2126–2136. doi: 10.1603/029.102.0616

Walsh, M. J., and Powles, S. B. (2007). Management strategies for herbicide-resistant weed populations in Australian dryland crop production systems. *Weed Technol.* 21, 332–338. doi: 10.1614/wt-06-086.1

Warwick, S. I., Francis, A., and Susko, D. J. (2002). The biology of Canadian weeds. 9. *Thlaspi arvense* L. (updated). *Can. J. Plant Sci.* 82, 803–823. doi: 10.4141/p01-159

White, J. W., Andrade-Sanchez, P., Gore, M. A., Bronson, K. F., Coffelt, T. A., Conley, M. M., et al. (2012). Field-based phenomics for plant genetics research. *Field Crops Res.* 133, 101–112. doi: 10.1016/j.fcr.2012.04.003

Ye, Z., Brûlé-Babel, A. L., Graf, R. J., Mohr, R., and Beres, B. L. (2017). The role of genetics, growth habit, and cultural practices in the mitigation of Fusarium head blight. *Can. J. Plant Sci.* 97, 316–328.

Zerner, M. C., Gill, G. S., and Vandeleur, R. K. (2008). Effect of height on the competitive ability of wheat with oats. *Agron. J.* 100, 1729–1734. doi: 10.2134/agronj2008.0068

Zhang, Y., Liang, Z., Zong, Y., Wang, Y., Liu, J., Chen, K., et al. (2016). Efficient and transgene-free genome editing in wheat through transient expression of CRISPR/Cas9 DNA or RNA. *Nat. Commun.* 7:12617.

Zheng, B., Chapman, S. C., Christopher, J. T., Frederiks, T. M., and Chenu, K. (2015). Frost trends and their estimated impact on yield in the Australian wheatbelt. *J. Exp. Bot.* 66, 3611–3623. doi: 10.1093/jxb/erv163

Zollinger, R. K., Manthey, F. A., and Fitterer, S. A. (1999). "Effect of preharvest herbicides on durum wheat quality," in *Proceedings 52nd Western Society of Weed Science*, Colorado Spring, CO, 8–11.

Zubaidi, A., Mcdonald, G. K., and Hollamby, G. J. (2000). Shoot growth, root growth and grain yield of bread and durum wheat in South Australia. *Aust. J. Exp. Agric.* 39, 709–720. doi: 10.1071/ea98184

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer HK declared a shared affiliation, with no collaboration, with several of the authors, JC and CP, to the handling editor at the time of review.

frontiers in Plant Science

ORIGINAL RESEARCH
published: 21 December 2020
doi: 10.3389/fpls.2020.569905

The Global Durum Wheat Panel (GDP): An International Platform to Identify and Exchange Beneficial Alleles

*Elisabetta Mazzucotelli[1†], Giuseppe Sciara[2†], Anna M. Mastrangelo[3,4], Francesca Desiderio[1], Steven S. Xu[5], Justin Faris[5], Matthew J. Hayden[6,7], Penny J. Tricker[8], Hakan Ozkan[9], Viviana Echenique[10], Brian J. Steffenson[11], Ron Knox[12], Abdoul A. Niane[13], Sripada M. Udupa[13], Friedrich C. H. Longin[14], Daniela Marone[3], Giuseppe Petruzzino[3], Simona Corneti[2], Danara Ormanbekova[2], Curtis Pozniak[15], Pablo F. Roncallo[10], Diane Mather[8], Jason A. Able[8], Ahmed Amri[13], Hans Braun[15], Karim Ammar[16], Michael Baum[13], Luigi Cattivelli[1], Marco Maccaferri[2], Roberto Tuberosa[2] and Filippo M. Bassi[13]**

[1] Council for Agricultural Research and Economics-Research Centre for Genomics and Bioinformatics, Fiorenzuola d'Arda, Italy, [2] Department of Agricultural and Food Sciences, University of Bologna, Bologna, Italy, [3] Council for Agricultural Research and Economics-Research Centre for Cereal and Industrial Crops, Foggia, Italy, [4] Council for Agricultural Research and Economics-Research Centre for Cereal and Industrial Crops, Bergamo, Italy, [5] Cereal Crops Research Unit, Edward T. Schafer Agricultural Research Center, United States Department of Agriculture, Agricultural Research Service, Fargo, ND, United States, [6] Agriculture Victoria, Agribio, Centre for AgriBiosciences, Bundoora, VIC, Australia, [7] School of Applied Systems Biology, La Trobe University, Bundoora, VIC, Australia, [8] School of Agriculture, Food and Wine, Faculty of Sciences, Waite Research Institute, The University of Adelaide, Adelaide, SA, Australia, [9] Department of Field Crops, Faculty of Agriculture, Çukurova University, Adana, Turkey, [10] Centro de Recursos Naturales Renovables de la Zona Semiárida, Departamento de Agronomía, Universidad Nacional del Sur-Consejo Nacional de Investigaciones Científicas y Técnicas, Bahía Blanca, Argentina, [11] Department of Plant Pathology, University of Minnesota, St. Paul, MN, United States, [12] Swift Current Research and Development Centre, Agriculture and Agri-Food Canada, Swift Current, SK, Canada, [13] International Center for Agricultural Research in the Dry Areas, Beirut, Lebanon, [14] State Plant Breeding Institute, University of Hohenheim, Stuttgart, Germany, [15] Plant Sciences and Crop Development Center, University of Saskatchewan, Saskatoon, SK, Canada, [16] International Maize and Wheat Improvement Center, Texcoco de Mora, Mexico*

OPEN ACCESS

Edited by:
Soren K. Rasmussen,
University of Copenhagen, Denmark

Reviewed by:
Monica Rodriguez,
University of Sassari, Italy
Benjamin Kilian,
Global Crop Diversity Trust, Germany

***Correspondence:**
Filippo M. Bassi
f.bassi@cgiar.org

[†] These authors have contributed equally to this work

Specialty section:
This article was submitted to Plant Breeding, a section of the journal Frontiers in Plant Science

Received: 05 June 2020
Accepted: 24 November 2020
Published: 21 December 2020

Citation:
Mazzucotelli E, Sciara G, Mastrangelo AM, Desiderio F, Xu SS, Faris J, Hayden MJ, Tricker PJ, Ozkan H, Echenique V, Steffenson BJ, Knox R, Niane AA, Udupa SM, Longin FCH, Marone D, Petruzzino G, Corneti S, Ormanbekova D, Pozniak C, Roncallo PF, Mather D, Able JA, Amri A, Braun H, Ammar K, Baum M, Cattivelli L, Maccaferri M, Tuberosa R and Bassi FM (2020) The Global Durum Wheat Panel (GDP): An International Platform to Identify and Exchange Beneficial Alleles. Front. Plant Sci. 11:569905. doi: 10.3389/fpls.2020.569905

Representative, broad and diverse collections are a primary resource to dissect genetic diversity and meet pre-breeding and breeding goals through the identification of beneficial alleles for target traits. From 2,500 tetraploid wheat accessions obtained through an international collaborative effort, a Global Durum wheat Panel (GDP) of 1,011 genotypes was assembled that captured 94–97% of the original diversity. The GDP consists of a wide representation of *Triticum turgidum* ssp. *durum* modern germplasm and landraces, along with a selection of emmer and primitive tetraploid wheats to maximize diversity. GDP accessions were genotyped using the wheat iSelect 90K SNP array. Among modern durum accessions, breeding programs from Italy, France and Central Asia provided the highest level of genetic diversity, with only a moderate decrease in genetic diversity observed across nearly 50 years of breeding (1970–2018). Further, the breeding programs from Europe had the largest sets of unique alleles. LD was lower in the landraces (0.4 Mbp) than in modern germplasm (1.8 Mbp) at $r^2 = 0.5$. *ADMIXTURE* analysis of modern germplasm defined a minimum of 13 distinct genetic clusters (*k*), which could be traced to the breeding program of origin. Chromosome regions putatively subjected to strong selection pressure were identified from fixation index (F_{st}) and diversity reduction index (*DRI*) metrics in pairwise comparisons among

decades of release and breeding programs. Clusters of putative selection sweeps (PSW) were identified as co-localized with major loci controlling phenology (*Ppd* and *Vrn*), plant height (*Rht*) and quality (gliadins and glutenins), underlining the role of the corresponding genes as driving elements in modern breeding. Public seed availability and deep genetic characterization of the GDP make this collection a unique and ideal resource to identify and map useful genetic diversity at loci of interest to any breeding program.

Keywords: durum wheat, genetic diversity, selection sweep, breeding history, wheat initiative

INTRODUCTION

Durum wheat [*Triticum turgidum* L. ssp. *durum* (Desf.) Husn.] is the 10th most important crop worldwide with an annual production of over 40 million tons (Sall et al., 2019). It provides the raw material for semolina, pasta, couscous, burghul and several other dishes of the Mediterranean tradition (Oliveira et al., 2012). Durum wheat evolved from domesticated emmer wheat, *T. turgidum* ssp. *dicoccum* (Schrank ex Schübl.) Thell., which originated from wild emmer wheat, *T. turgidum* ssp. *dicoccoides* (Körn. ex Asch. & Graebn.) Thell. in the Fertile Crescent approximately 10,000 years ago (Ozkan et al., 2002; Dubcovsky and Dvorak, 2007). Thus, three distinct phases can be identified in the human-driven tetraploid wheat evolution process: (*i*) domestication (from wild to domesticated emmer wheat), (*ii*) continued evolution under domestication (from domesticated emmer wheat to durum wheat landraces) and (*iii*) improvements achieved by modern breeding (from landraces to modern durum wheat varieties) (Maccaferri et al., 2019). As a consequence of this evolution, four mega-germplasm groups of tetraploid wheat can be defined: tetraploid wild relatives, tetraploid primitive wheats (domesticated and cultivated), durum wheat landraces and modern durum wheat varieties. During the second evolution phase, the transition from the domesticated form of emmer to durum landraces underwent strong selection pressure by ancient farmers (Tanksley and McCouch, 1997). Modern breeding has accelerated this process by artificially crossing "best by best" and selecting for "the best" with impressive genetic gains being realized, resulting in the development of improved varieties accumulating beneficial alleles (Slafer et al., 1994; Borrelli and Trono, 2016; van Ginkel and Ortiz, 2018). Genetic gain is typically quantified as the slope of the regression between yield and year of release of varieties. A genetic gain of 0.3–1.2% per year has been recorded for durum wheat over the last century in different growing regions (e.g., Giunta et al., 2007; Royo et al., 2008; Clarke et al., 2010; Bassi and Nachit, 2019; Mondal et al., 2020) and often associated with variations in morpho-physiological traits, such as a shift toward earlier flowering and a reduced plant height, with a corresponding increase in harvest index (e.g., De Vita et al., 2007; Royo et al., 2007; Isidro et al., 2011; Bassi and Nachit, 2019). However, the positive yield trend has often been reached at the cost of eroding genetic diversity within elite gene pools (Fernie et al., 2006; Bassi and Nachit, 2019). The limited number of landraces that were used as founder lines of the modern gene pool (e.g., the first modern durum breeding program spearheaded by Nazareno Strampelli in 1910; Scarascia Mugnozza, 2005; Dexter, 2008; Royo et al., 2009; Taranto et al., 2020) and the "best × best" strategy traditionally used by breeders to drive the genetic gain (Hoisington et al., 1999; Maccaferri et al., 2003; van Ginkel and Ortiz, 2018) are the two main causes of this phenomenon. Genetic erosion of the durum wheat cultivated gene-pool in comparison with wild relatives and landraces has been reported, analogously to other crop species (Tanksley and McCouch, 1997; Gur and Zamir, 2004; Raman et al., 2010; Royo et al., 2010; Laidò et al., 2013; Kabbaj et al., 2017; Maccaferri et al., 2019), and it represents a real concern for breeders as it might lead to a lack of novel beneficial alleles for selection, yield stagnation, and/or increased susceptibility to biotic and abiotic stresses. Therefore, breeders are devoting increasing resources and effort to identify beneficial alleles and traits from novel germplasm sources to reinvigorate their programs. Indeed, pre-breeding activities have been pursued by international programs at ICARDA (Zaïm et al., 2017; Bassi et al., 2019; Robbana et al., 2019; El Haddad et al., 2020) and CIMMYT (Singh et al., 2018; Ledesma-Ramírez et al., 2019), and by national research institutes to introgress beneficial alleles from landraces and wild relatives, in parallel to international initiatives which aim to identify, collect, conserve and use the wild cousins of some of the most important food crops, as the CWR project "Adapting Agriculture to Climate Change: Collecting, Protecting and Preparing Crop Wild Relatives[1]. Population structure and genetic diversity have been studied in several modern and landrace collections of durum wheat. Many studies have focused on panels from a restricted country/area such as landraces from Southern Italy (Marzario et al., 2018), Iran (Talebi and Fayaz, 2016), Spain (Giraldo et al., 2016), Tunisia (Robbana et al., 2019; Slim et al., 2019), Turkey and Syria (Baloch et al., 2017), Palestine, Jordan and Israel (Abu-Zaitoun et al., 2018), or specific breeding programs (N'Diaye et al., 2018). Others have considered durum wheat collections of wider origin encompassing a few hundred entries. Among the earliest studies reporting on assembling international and diverse panels of mainly elite durum lines and cultivars, Maccaferri et al. (2005, 2006, 2010, 2011), Reimer et al. (2008) and Laidò et al. (2013) all reported on the genome-wide molecular diversity and LD-decay rate estimated with SSR and DArT^{TM} markers. More recently, germplasm collections have been characterized with the Illumina iSelect 90K SNP (Maccaferri et al., 2016; Mangini et al., 2018; Saccomanno et al., 2018) and

[1]https://www.cwrdiversity.org/

subjected to GWAS for response to diseases, root morphology, canopy traits related to phenology, photosynthesis and grain yield potential (e.g., Maccaferri et al., 2010, 2016; Canè et al., 2014; Condorelli et al., 2018). Similarly, Kabbaj et al. (2017) used a mixed set of modern lines and landraces to define the genetic diversity and origin of modern durum wheat as well as to identify loci controlling resistance to insect pests and tolerance to heat stress (Bassi et al., 2019; El Hassouni et al., 2019). The largest study to date considered a collection of 429 USDA-ARS durum entries including cultivars and landraces from 64 countries. This collection was analyzed with 6,538 polymorphic SNPs (Chao et al., 2017) from the Illumina iSelect wheat 9K array (Cavanagh et al., 2013). More recently, a deeper study of genetic diversity was carried out for the Tetraploid wheat Global Collection (TGC) consisting of 1,856 single-seed purified gene bank entries chosen to comprehensively explore the diversity in tetraploid wheat from durum landraces through domesticated and wild emmer (Wang et al., 2014) in combination with the availability of the reference genome assembly of the cultivar 'Svevo' (Maccaferri et al., 2019).

Genetic diversity is not necessarily considered as relevant *per se*. Rather, with advances in genetics, genomics and functional genomics (Tuberosa and Pozniak, 2014), researchers and breeders are increasingly targeting specific genomic regions known to be relevant, with the objective to improve the *exploitable* and *useful* diversity (Kabbaj et al., 2017; N'Diaye et al., 2018). Accordingly, developing a detailed knowledge at the molecular level of historical loss of diversity events, together with the identification of successful allelic combinations progressively accumulated over repeated breeding cycles, are instrumental for a more effective management of breeding programs (Pfeiffer et al., 2001).

With this aim, the international durum wheat research community met in Bologna, Italy, in 2015 under the umbrella of the Expert Working Group on Durum Wheat Genomics and Breeding, as part of the Wheat Initiative[2], to take joint action toward the identification of beneficial alleles and to make them available for breeding programs and pre-breeding efforts. The result of this international call to action is presented here under the name of the Global Durum wheat Panel (GDP). This panel was designed with the aim of capturing most of the readily *exploitable* genetic diversity, sharing it freely to facilitate research discoveries, and ultimately providing a rapid mean to exchange useful alleles worldwide. This article describes the germplasm composition and genetic structure of the GDP to provide the basic knowledge needed to support its international phenotypic characterization and exploitation.

MATERIALS AND METHODS

Plant Materials

A total of 2,503 accessions of tetraploid wheat were obtained from 25 worldwide partners representing institutions, universities, gene banks and private companies (**Supplementary Table S1**), all exchanged under the Standard Material Transfer Agreement (SMTA, Noriega et al., 2019) to allow full exploitation for breeding and research. This initial set of germplasm was defined as the Durum Wheat Reference Collection (DWRC, **Supplementary Table S2**) and grown in the 2015–2016 season at the ICARDA experimental farm in Terbol, Lebanon. The DWRC included 1,541 *T. turgidum* ssp. *durum* modern breeding accessions (cultivars, varieties and elite lines) from 49 countries/programs, an evolutionary population set from INRA France of 180 entries (Evolutionary Pre-breeding pOpulation, EPO, David et al., 2014), 416 *T. turgidum* ssp. *durum* landraces obtained from 48 countries, and 366 wild and primitive tetraploids from 37 countries (*T. turgidum* ssp. *dicoccoides* and *dicoccum*, *turgidum*, *turanicum, polonicum, carthlicum*, respectively). Each entry was planted in two rows of 2 m in length under supplemental irrigation. Fungicide and fertilizer were provided in-season, following optimal local management practices. From each plot a single tiller was selected and tagged at flowering based on spike size, phenology and shape to be representative of most plants within the same plot. From this tiller, a leaf sample was collected for initial molecular screening. At maturity, the spike of the tagged tiller was harvested and used for advancement. In the 2016–2017 season at the same field station, 10 seeds from each spike were planted in rows of 0.5 m in length. Irrigation and chemical treatments were used to maximize productivity. Using the initial molecular data, a subset of approximately 1,000 entries were selected and defined as the Global Durum wheat Panel (GDP). The whole row was bulk-harvested and used for further advancement. In the 2017–2018 season, each entry was planted in plots of 6 m^2 at the American University of Beirut (AUB) experimental farm in Lebanon. Fungicide, irrigation and fertilizer were applied in order to maximize productivity. Plots were visually inspected for homogeneity and off-types were manually rouged.

From this first multiplication, a total of 762 entries produced enough seed for distribution to 28 collaborators under the name of GDP version 1 (GDPv1-19), which substantially included all *T. durum* lines (modern, EPO, and landraces germplasm) (**Supplementary Table S3**). In the 2018–2019 season, a second and final multiplication cycle was conducted to produce enough seed of 976 entries to generate sets of 50 seeds per entry, ready to sow by 21 requesting partners. These sets were distributed under the name of GDP version 2 (GDPv2-20) (**Supplementary Table S3**). Unfortunately, some entries were lost during multiplication due to excessive susceptibility to yellow rust races in Lebanon. Additional sets remain available for request and distribution under SMTA at this link: http://indms.icarda.org/. Furthermore, 42 additional entries were included in GDPv2-20, mostly representing recently released European varieties and *T. durum* lines carrying introgressions of *Fhb1* developed by Boku University (Prat et al., 2017; **Supplementary Table S3**).

DNA Extraction and Genotyping

The initial molecular screening of the DWRC was performed by sending one leaf from each selected tiller to LGC Genomics (United Kingdom) for DNA extraction and subsequent analyses. Ninety-four KASP® markers (**Supplementary Table S2**) were

[2]www.wheatinitiative.org

selected because evenly distributed along the genome and highly polymorphism (Kabbaj et al., 2017), including markers tagging important loci: *PpdA1*, *VrnA1*, and *RhtB1*. Accessions with more than 50% missing data were discarded, as well as markers which were monomorphic or detected multiple loci (gene calls with multiple allelic classes and heterozygous calls at high frequency).

Lines selected to be part of the GDP were genotyped using the Illumina iSelect 90K SNP array technology (Wang et al., 2014) at the USDA-ARS Small Grain Genotyping Laboratory, Fargo, ND, United States. A pool of three seeds originating from the single spike selected in 2015–2016 were sown in Jiffy pots; 10 days old leaves were collected and DNA extracted using the NucleoSpin Plant II kit (Macherey-Nagel) according to manufacturer's instructions. The raw data (Theta/R) from single genotyping experiments was exported from GenomeStudio software (Illumina Ltd.) and jointly analyzed for cluster assignment and genotype calling using a custom script as described in Maccaferri et al. (2019). The script parameters were $d = 3$, to call samples only within three standard deviations from a known cluster position, and $r = 0.8$, minimum confidence score that the sample belonged to the cluster to which it was assigned versus the next closest cluster. Stepwise data curation was conducted on polymorphic SNP markers. First, markers with minor alleles present in fewer than three genotypes were discarded. Second, the remaining markers were filtered to retain SNPs with a unique map position in the available genetic maps (Maccaferri et al., 2015, 2019), and with the marker sequences aligned to a single position along the Svevo reference genome RefSeq V1.0 (Maccaferri et al., 2019). Third, those markers showing multiple hits along the genome were checked for linkage disequilibrium (LD) against the hypothetical nearby mapped markers, and assigned a unique position based on the highest r^2 (above a 0.3 threshold) with the putatively contiguous markers. SNP imputation was performed using *Beagle 5* software using default parameters (Browning et al., 2018). The imputation accuracy was measured at 98.6% by running 1,000 replicates of randomly masked 1% of the called genotypes (Nothnagel et al., 2009; Hancock et al., 2012). Using the software PLINK (Chang et al., 2015), redundant markers were pruned based on genome wide *linkage disequilibrium* set at $r^2 = 0.99$ and merged into one unique SNP call. Moreover, three additional pruned hapmaps were produced selecting a single SNP among those with r^2 of 0.8, 0.5 and 0.3 to run the population structure analysis.

Genetic Diversity Within the GDP and Putative Signal of Selection Sweeps

Genetic diversity and population differentiation within the GDP, both at the genome-wide and at the single-locus level, were assessed within and between populations defined according to passport data provided by contributors or retrieved from GRIS (Genetic Resources Information System for Wheat and Triticale) through www.wheatpedigree.net. Accessions of wild emmer, primitive cultivated sub-species, and durum landraces were classified on the basis of the country of collection, whereas modern durum germplasm (cultivars, varieties and elite lines) were grouped based on the breeding program of origin and decade of release (five decades considered: '70–'80, '81–'90, '91–'00, '01–'10, and '11–'18). Because the year of release was not available for elite lines included in the GDP, the year in which the cross was performed was used to estimate the year of release by adding 10 years. Polymorphic SNP datasets were selected according to the set filtering for minor allele frequency (MAF) > 5% and pruning at $r^2 < 0.99$.

Genetic diversity among and within populations was calculated by AMOVA, fixation index (F_{st}, Wright, 1965) and the polymorphism information content (PIC, Botstein et al., 1980). The within populations total number of polymorphic loci (N), Nei's gene diversity (Nei, 1973), and mean number of pairwise differences were calculated, and significance was determined based on LSD at $P < 0.05$. Population differentiation was assessed based on Nei's genetic distance (Nei, 1972) and population pairwise F_{st}. All values were derived using the Arlequin 3.5 software (Excoffier and Lischer, 2010), and significance levels for variance components and F_{st} statistics were estimated based on 10,000 and 1,000 permutations, respectively.

Furthermore, single locus analyses of genetic diversity across the whole genome were conducted to identify genomic regions putatively affected by human-driven selection sweeps. Signals of putative selection sweeps were assessed using a hapmap pruned for $r^2 < 0.99$ calculating two different indices: F_{st} was estimated by Arlequin 3.5 software, and the diversity reduction index (*DRI*) was calculated using the modified *ROD* formula presented in Maccaferri et al. (2019). To reduce spurious signals due to different coalescence time between SNPs, the raw single SNP-based results were smoothed by averaging with a sliding window of 15 SNPs with a one-marker step. Significance of selection signals was assessed in a two-step procedure. In the first step, signal peaks falling in the top 10% percentile of the distribution were identified. Additional neighboring signals were merged into the one representing the highest value, considering as neighbors loci falling within a physical distance lower than the LD. After merging adjacent peaks, the index distribution (Jordan et al., 2015) was re-calculated and the 95th percentile was chosen as the index-specific significance threshold.

Population Structure Analysis and Selection of the GDP Collection

A preliminary population stratification analysis was carried out on the DWRC panel using a curated set of 88 KASP$^{(R)}$ markers. The GDP set was then re-stratified using the Illumina 90k SNP genotyping data and three possible pruned hapmaps (r^2 set at 0.3, 0.5, and 0.8) were considered in order to optimize the trade-off between uniformity of genomic sampling and informativeness. Based on the analysis results, the pruned SNP-set at $r^2 = 0.5$ was used for all subsequent population structure analyses. For both the DWRC and GDP, the population structure was estimated by the model-based likelihood method *ADMIXTURE* optimized using the block relaxation algorithm and the quasi-Newton convergence acceleration method and $q = 3$ secants (Alexander et al., 2009), as well as by means of Ward's clustering of Nei's genetic distances, using the

poppr v. 2.8.3 and *adegenet* packages of R (Jombart, 2008; Kamvar et al., 2014; R Core Team, 2016). For both methods, the sub-population membership was defined for *k* values increasing from 2 to 20. The parameters used to define the optimal number of clusters were *ADMIXTURE*'s cross-validated error rate and minimum group size. Lines with strong admixture were defined as those showing less than 30% identity (membership) with any ancestry in the model-based likelihood analysis. Because the GDP is a selected sub-set of the initial DWRC panel, the population stratification was first used to define the most representative DWRC entries to be included in the GDP, and secondly to define what degree of genetic diversity was lost because of the sub-sampling process. Pairwise similarity estimated as identity-by-state (IBS) was also calculated for the DWRC population to filter for duplicated/highly similar entries using TASSEL5 software (Bradbury et al., 2007). To select the subset of DWRC entries that composed the GDP the following procedures were followed. First, genotypes representing historical founders, parents of mapping populations, or known germplasm carrying interesting alleles/phenotypic traits were included, while the name and pedigree were inspected and compared to the similarities defined at the molecular level (IBS-GS matrix) to discard duplicated entries with >0.95 similarity (only one entry was retained per group). The remaining entries were classified into six groups, five of which were defined by genetic structure at *k* of 5, and one extra split to incorporate the EPO set, which was clearly differentiated from the other groups. The GDP collection was then assembled through a *stratified-sampling method*, therefore choosing representative entries from each main Ward's cluster and sub-clusters, depending on each subgroup/subspecies being considered and chosen in order to maximize the number of sub-clusters being considered for GDP sampling. Genotypes with low average genetic similarity to other entries (rare haplotypes) were also chosen. The genetic diversity level present in the two collections was compared to confirm that no major genetic diversity losses occurred after sampling the GDP from the DWRC. The Shannon-Wiener's diversity index, Nei's expected heterozygosity, allelic evenness (Shannon, 1948; Nei, 1978; Smith and Wilson, 1996), MAF, and the site frequency spectrum (SFS) distribution were assessed at the *locus* level both in the DWRC and GDP based on the 88 KASP markers. Diversity indexes analyses were conducted using the "*locus_table*" and "*poppr*" function of the *poppr R* package (Kamvar et al., 2014).

LD Decay

Pairwise marker correlations (r^2 values) were calculated on the SNP dataset of the GDP for each chromosome using TASSEL5 (Bradbury et al., 2007). LD decay curves were fitted using the non-linear model described in Rexroad and Vallejo (2009). Critical parameters of marker distances at $r^2 = 0.3$ and 0.5 were extrapolated from the fitted regression curves. The r^2 of unlinked markers (background noise) was estimated as the 95th quantile of r^2 values of markers on different chromosomes (unlinked set). To estimate the local LD value along chromosomes, each marker LD was calculated using the mean r^2 with the 50 nearest markers, and then smoothed as one value using the step-sliding window.

Identification and Clustering of Putative Selection Sweep (PSW) Signals

Detection of putative selection sweep (PSW) signals was based on genome-wide F_{st} and *DRI* metrics calculated for *modern* vs. *landraces* and for pairwise groups of entries classified by decade or breeding program. PSW clusters were defined as two significant signals on the same chromosomal region in a single pair/comparison or among pairs/comparisons. Moreover, signals also partially overlapping were grouped into one cluster. The catalog of PSW was integrated with data from the literature that included major genes cloned in wheat, known QTL and the comprehensive catalog (a.k.a. *QTLome*) defined in Maccaferri et al. (2019).

RESULTS

From the Durum Wheat Reference Collection to the Global Durum Wheat Panel

The original DWRC was comprised of 2,503 accessions that were genotyped with 94 KASP$^{(R)}$ markers (**Supplementary Table S2**). The curation process yielded a final set of 2,493 accessions (99.1%), each with 88 (93.6%) reliable KASP$^{(R)}$ marker profiles. Population structure assessed by *ADMIXTURE* (**Supplementary Figure S1**) highlighted three subsets at $k = 3$: (*i*) a group including *T. turgidum* spp. *dicoccum* and *dicoccoides,* (*ii*) a second group including modern durum wheat germplasm and (*iii*) a third group comprising modern North American germplasm together with most durum landraces and accessions of the primitives *T. turgidum* spp. *turgidum*, *turanicum* and *polonicum* as durum-related sub-species. At $k = 4$ the North American modern germplasm was separated from landraces and the mentioned primitive subspecies. Finally, at $k = 5$ the group of the modern durum wheat germplasm was further subdivided in two groups: the first one tracing its ancestry to the CIMMYT breeding program, and the second one composed of the Southern European germplasm and those entries with ancestry from the ICARDA breeding program. The structure of the population was confirmed using bootstrapped Ward's clustering (**Supplementary Figure S2**).

A total of 398 genotypes represented identical entries contributed by multiple partners. The remaining entries were divided into six groups: five defined by genetic structure at $k = 5$ and one additional group to incorporate the EPO set. When each of these subsets was subjected to population structure assessment based on Ward's clustering, the sub-clustering concurred with the clustering computed on the whole DWRC and a detailed picture of group differentiation based on geographic origin was revealed. The entries to be included in the GDP were then identified based on the Ward's clustering using a *stratified-sampling method*. Following the criteria defined in Material and Methods, three groups of durum wheat modern germplasm were selected (**Supplementary Figure S3**): (*i*) CIMMYT- and ICARDA-derived genetic materials, and modern semi-dwarf and vernalization-insensitive

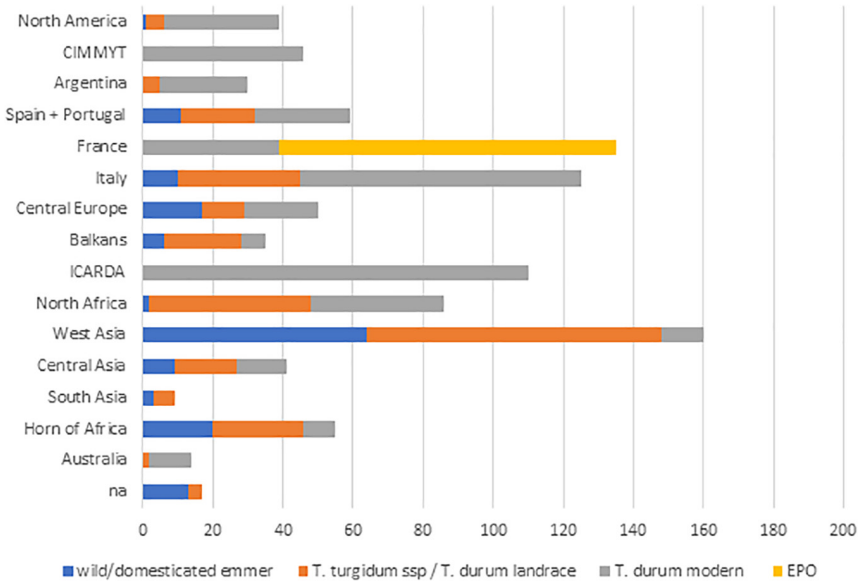


FIGURE 1 | Distribution of the geographic origin of the GDP accessions used for genetic diversity analysis. Countries of origin are grouped as follows: *Central Europe*: Austria, Hungary, Ukraine, Sweden, Poland, United Kingdom, and Germany; *Balkans*: Serbia, Bosnia, and Herzegovina, Bulgaria, Romania, Greece, and Crete; *North Africa*: Egypt, Libya, Algeria, Tunisia, and Morocco; *West Asia*: Turkey, Syria, Lebanon, Israel, Jordan, Iran, Iraq, Armenia, Azerbaijan, Georgia, Oman, Yemen, and Saudi Arabia; *Central Asia*: Kazakhstan, Afghanistan, Russia, Uzbekistan, and China; *Horn of Africa*: Ethiopia, Eritrea, and Kenya.

lines mostly adapted to the Mediterranean environment for a total of 288 genotypes; (*ii*) 96 elite semi-dwarf durum wheat lines with photoperiod and/or vernalization sensitivity mainly developed in Canada, France, Italy, and Central Europe; (*iii*) 96 non-semi-dwarf durum wheat lines of different origins. Three additional groups were selected to incorporate more genetic diversity including; (*iv*) 96 EPO lines (**Supplementary Figure S4**); (*v*) 192 durum wheat landraces representing the geographical distribution of the original collection (**Supplementary Figure S5**); and (*vi*) a final group including domesticated emmer lines (96, **Supplementary Figure S6**), wild emmer accessions and other tetraploid primitives (96, **Supplementary Figures S7, S8**, respectively). A seventh group of 42 entries including recently registered European varieties and durum lines carrying *Fhb1* introgressions developed at the Boku University (Austria) was also included. The final GDP selection consisted of 1,028 accessions, 976 of which were multiplied in sufficient quantity and quality for seed re-distribution by ICARDA, while 42 among European varieties and accessions with *Fhb1* introgressions are available from University of Bologna and Boku University, respectively (**Supplementary Table S3**) for a total of 1,018 entries available as seed stocks. **Figure 1** shows the geographic origin of the GDP accessions.

To assess the extent of the genetic diversity loss in the sampling process from the DWRC to GDP, different indices were calculated based on the KASP data for the two panels. Locus level correlations between DWRC and GDP values resulted in Pearson's coefficients of 0.94 for the MAF, 0.95 for allelic evenness, 0.96 for expected heterozygosity and 0.97 for Shannon-Wiener's diversity index (**Supplementary Figure S9**), indicating that the sampling process that originated the GDP caused a 3–6% loss of the initial DWRC diversity. The SFS (**Supplementary Figure S10**) showed that the distribution of the allele frequencies in the GDP is comparable to that observed in the initial DWRC, except for an appreciable decrease in three rare allele frequency classes (MAF: 0.05–0.10, 0.10–0.15, and 0.35–0.40) and a corresponding increase for three high frequency classes (MAF: 0.15–0.20, 0.30–0.35, and 0.45–0.50).

Deep Genotyping of the Global Durum Wheat Panel (GDP)

Genotyping of the GDP with the iSelect 90K wheat SNP array generated 42,520 polymorphic SNPs. After several quality filtering steps, a total of 16,633 SNP markers were retained and imputed for missing data. Both datasets are available at the repositories GrainGenes[3] and T3/Wheat[4]. The tetraploid genome was thus probed by a mean of 1,188 SNP markers per chromosome with an average density of 1.7 SNPs per Mbp or 6.3 SNPs per cM (**Table 1**). Almost one third (4,119) of the consecutive SNPs were located within 0.5 Kbp of each other, possibly due to the redundancy of the Illumina 90K SNP design, and 4,938 SNPs were located at various interlocus distances between 1 and 100 Kbp. The remaining 7,259 SNPs mapped at distances from >0.1 to 5 Mbp, and only 302 SNPs mapped at distances >5 Mbp (**Supplementary Figure S11A**). The genome coverage calculated as a percent of the physical genome length probed by SNP markers was almost complete with an average of 0.998% (**Table 1**). The marker density

[3]https://wheat.pw.usda.gov/GG3/global_durum_genomic_resources
[4]https://wheat.triticeaetoolbox.org/breeders_toolbox/protocol/158

TABLE 1 | Genome coverage by the SNP marker dataset expressed for each chromosome and genome.

Chromosome	N° SNP	Mean SNP/Mb	Chromosome coverage (%)
1A	1,138	1.9	0.998
1B	1,638	2.4	0.997
2A	1,096	1.4	0.999
2B	1,800	2.3	0.997
3A	979	1.3	0.999
3B	1,250	1.5	0.999
4A	750	1.0	0.996
4B	863	1.3	0.998
5A	962	1.4	0.997
5B	1,419	2.0	0.999
6A	974	1.6	0.998
6B	1,267	1.8	0.993
7A	1,248	1.7	0.998
7B	1,249	1.7	0.997
Genome A	7,147	1.5	0.998
Genome B	9,486	1.9	0.997
Total	*16,633*	*1.7*	*0.997*
Mean	*1,188*	*1.7*	*0.998*

along the chromosomes was higher in proximal and distal portions compared to pericentromeric regions (**Supplementary Figure S11B**), and the opposite for the interlocus distances (**Supplementary Figure S11C**).

After excluding six accessions due to failed genotyping, filtering carried out at the accession level based on IBS_GS matrix (**Supplementary Table S4**) allowed for the identification of 10 accessions whose genotypic data were not relevant (misclassified accessions or contaminated DNA) that were discarded from further analysis. High-density genotyping data are therefore available for a final set of 1,011 accessions, while for a total of 1001 accessions both seed stock and genotypic data are provided (**Supplementary Table S3**).

Genetic Diversity Analysis

Genotyping data allowed to characterize the GDP for genetic diversity and differentiation within and among groups defined on the base of passport data (**Supplementary Table S3**). GDP entries were classified according to the following criteria. The introduction of the semi-dwarf *RhtB1b* allele from CIMMYT durum lines (Motzo and Giunta, 2007; Ortiz et al., 2007) represents the origin of the post green revolution germplasm, so all entries generated from crosses carried out after 1970 were considered as *modern* germplasm. North American varieties and breeding materials released after 1970 were also included in the *modern* set, even though these did not carry the *RhtB1b* allele, which is not beneficial in the northern semi-arid prairie environment. All durum lines pre-dating 1970 were considered as *landraces*, although in a few cases these were obtained through breeding selection of populations or voluntary hybridization among landraces. Notably, the characterization of genetic diversity could not clearly distinguish *T. turgidum* spp. *durum* landraces from other *T. turgidum* sub-species related to durum like *T. turgidum* ssp. *turgidum*, *turanicum* and *polonicum* (Maccaferri et al., 2019). Therefore, the genetic diversity analyses reported hereafter were carried out including all durum- related *T. turgidum* sub-species accessions as *landraces* and grouped according to the country of origin. The *EPO population* was considered as a separate group based on its highly distinct genetic structure.

The primary objective was to describe the pattern of genetic diversity across the history of durum wheat evolution and breeding so these groups composed as above described were considered: (*i*) *modern* germplasm, (*ii*) *landraces* and (*iii*) emmer (*T. turgidum* ssp. *dicoccum*) accessions, for a total of 861 genotypes. AMOVA highlighted a moderate level (23%) of genetic variance distinguishing the three groups (**Table 2A**), with a larger portion still existing within groups (77%). Reduction of overall diversity was observed in *modern* lines with respect to both *T. turgidum* ssp. *dicoccum* and *landraces*. Durum *landraces* showed a level of genetic diversity even higher than that of *T. turgidum* ssp. *dicoccum* accessions included in the GDP, perhaps due to ascertainment bias associated to the type of genotyping array used for the analysis, originally developed to maximize polymorphism among modern bread and durum breeding lines. However, in pairwise differentiation analysis F_{st} value was higher in the comparison *landraces* vs. *dicoccum* (F_{st} = 0.2688) with respect to the comparison *landraces* vs. *modern* lines (F_{st} = 0.1378) (**Figure 2A**). The EPO population, which was bred by INRA based on a composite cross to introduce diversity from wild and primitive accessions of *T. turgidum* subspecies, showed a relatively high level of diversity (David et al., 2014). Considering the *all durum* dataset (885 entries and 8,802 polymorphic SNPs), AMOVA results across the three main groups (*modern* lines, *landraces* and EPO accessions) showed that the highest proportion of molecular variance (86.94%) was observed within clusters rather than among clusters (13.06%) (**Table 2B**). *Landraces* showed the highest value of Nei's genetic diversity (0.358), followed by *modern* germplasm (0.292) and EPO (0.288) (**Table 2B**). As to among-population comparisons, the highest differentiation was found for *landrace* vs. *modern* comparisons (F_{st} = 0.127), while an F_{st} of 0.1 was calculated for the EPO vs. *modern* comparison (**Figure 2B**). This result is also confirmed by comparable values of PIC and F_{st} calculated for *landraces* (0.282 and 0.101, respectively, **Table 2C**) and *modern* lines (0.278 and 0.117, respectively, **Table 2E**).

Durum *landraces* (282) were grouped into 14 sub-populations according to the country of origin. This clustering process accounted only for 10.1% of the variance, while the vast majority of diversity still remained unclustered within sub-populations (**Table 2C**). Nei's gene diversity values ranged from 0.280 (United States–Canada) to 0.374 (Turkey–Transcaucasian).

To analyze the changes in diversity within the *modern* germplasm over time and across breeding groups, the totality of 473 cultivars and elite lines were divided into sub-groups based on two different criteria: (*i*) decade of release from 1970 to 2018; and (*ii*) country of registration/release, which roughly defines the main groups of breeding programs.

TABLE 2 | AMOVA and gene diversity for five germplasm sub-sets defined according to passport data: **(A)** GDP without the wild accessions, with grouping based on historical selection steps: *T. dicoccum* accessions, *T. durum* germplasm sub-sets landraces, *T. durum* germplasm sub-sets cultivars; **(B)** all *T. durum* germplasm sub-sets; groups are EPO, *T. durum* germplasm sub-sets landraces, modern lines; **(C)** all landraces grouped according to country of origin; **(D)** all *T. durum* germplasm sub-sets modern lines, classified according to decade of release; **(E)** all *T. durum* germplasm sub-sets modern lines, classified based on breeding program.

(A)

Source of variation	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**
Among populations	2	222,822.13	443.9	22.98
Within populations	859	1,278,090.16	1,487.88	77.02
Total	861	1,500,912.3	1,931.79	
F_{st}		0.23		
***T. durum* groups**	**N° accessions**	**N° polymorphic loci over 10173**	**Nei's gene diversity**	**Mean number of pairwise differences**
LANDRACE	286	10,154	0.332	3,375.08
EMMER	103	9,901	0.317	3,220.49
MODERN	473	10,010	0.264	2,681.76
Mean value			0.304	3,092.45
Lsd (p = 0.0005)			0.002	17.7

(B)

Source of variation	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**
Among populations	2	103,704.61	207.33	13.06
Within populations	852	1,175,524.24	1,379.72	86.94
Total	854	1,279,228.85	1,587.05	
F_{st}		0.13		
PIC		0.273	range (0.09–0.375)	
***T. durum* groups**	**N° accessions**	**N° polymorphic loci over 8802**	**Nei's gene diversity**	**Mean number of pairwise differences**
LANDRACE	286	8,796	0.358	3,151.85
MODERN	473	8,781	0.292	2,567.05
EPO	96	8,213	0.288	2,538.16
Mean value			0.313	2,752.35
Lsd (p = 0.0005)			0.002	17.6

(C)

Source of variation	**d.f.**	**Sum of squares**	**Variance components**	**Percentage of variation**
Among populations	13	62,147.3	169.12	10.1
Within populations	270	403,266.13	1,504.72	89.9
Total	281	465,413.43	1,673.84	
F_{st}		0.101		
PIC		0.282	range (0.09–0.375)	
Landrace group	**N° accessions**	**N° polymorphic loci over 9414**	**Nei's gene diversity**	**Mean number of pairwise differences**
Turkey-Transcaucasian	29	9,253	0.374	3,521.67
Central Asia	18	9,035	0.356	3,348.15
Arabian Peninsula	9	7,790	0.349	3,285.50
Iberian Peninsula	21	9,048	0.346	3,253.83
Central Europe	18	8,547	0.341	3,206.41
South Asia	6	6,640	0.329	3,094.53
Greece	16	8,539	0.327	3,082.52
Italy	34	8,656	0.303	2,874.58
Ethiopia	26	8,174	0.302	2,843.20
North Africa	47	9,059	0.301	2,839.57
Argentina	5	6,107	0.300	2,829.40

(Continued)

TABLE 2 | Continued

Landrace group	N° accessions	N° polymorphicloci over 9414	Nei's gene diversity	Mean number of pairwise differences
Levant	46	8,496	0.289	2,722.53
North America	5	5,340	0.280	2,640.00
Australia	2	3,136	0.333	3,136.00
Mean value			0.323	3,048.42
Lsd (p = 0.05)*			0.005	48.1
Lsd (p = 0.001)*			0.008	75.8

(D)

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	4	13,737.04	29.36	2.95
Within populations	444	429,609.05	967.59	97.05
Total	448	443,346.08	996.95	
F_{st}		0.029		

Breeding decade	N° lines	N° polymorphicloci over 5685	Nei's gene diversity	Mean number of pairwise differences
70–80	19	5,334	0.357	2,027.88
81–90	62	5,668	0.364	2,069.90
91–00	93	5,679	0.348	1,979.88
01–10	132	5,681	0.337	1,914.88
11–18	143	5,675	0.326	1,855.32
Mean value			0.346	1,969.57
Lsd (p = 0.0005)			0.003	33.2

Breeding group	70–80	81–90	91–00	01–10	11–18	Total
Australia	0	1	2	5	4	12
Central Asia	2	1	3	2	1	9
Central Europe	0	1	4	2	12	19
CIMMYT	3	2	5	6	29	45
Spain	3	4	10	6	1	24
Ethiopia	0	0	1	1	3	5
France	0	7	12	13	7	39
ICARDA	0	8	12	45	40	105
Italy	9	12	22	23	14	80
North America	2	8	13	5	5	33
South America	0	7	1	7	10	25
South Mediterranean	0	11	8	17	17	53
Total	19	62	93	132	143	

(E)

Source of variation	d.f.	Sum of squares	Variance components	Percentage of variation
Among populations	11	60,189.98	121.56	11.67
Within populations	460	423,162.29	919.92	88.33
Total	471	483,352.27	1,041.48	
F_{st}		0.117		
PIC		0.278	range (0.09–0.375)	

(Continued)

TABLE 2 | Continued

Breeding group	N° lines	N° polymorphic loci over 5918	Nei's gene diversity	Mean number of pairwise differences
Italy	80	5,885	0.343	2,031.16
Central Asia	14	5,463	0.339	2,006.78
France	39	5,768	0.339	2,006.03
South America	25	5,535	0.335	1,984.11
Spain	27	5,591	0.332	1,963.08
Central Europe	25	5,466	0.321	1,902.33
South Mediterranean	53	5,776	0.312	1,848.12
Ethiopia	8	4,253	0.297	1,755.71
North America	33	5,316	0.296	1,749.61
ICARDA	110	5,813	0.294	1,741.78
CIMMYT	46	5,046	0.256	1,513.59
Australia	12	4,208	0.255	1,506.68
Mean value			0.310	1,834.08
Lsd (p = 0.05)			0.005	29.7
Lsd (p = 0.001)			0.008	46.9

Geographic area of origin has been assigned to GDP landraces based on country where they have been sampled, as follows: North America: Canada + United States; Arabian Peninsula: Oman, Yemen, and Saudi Arabia; Central Europe: Bulgaria, Serbia, Poland, Ukraine, Hungary, Romania, United Kingdom, Germany, Sweden, Bosnia, and Herzegovina; Iberian peninsula: Spain and Portugal; Levantine: Iran, Iraq, Jordan, Syria, and Israel; North Africa: Egypt, Tunisia, Algeria, Morocco, Libya, and Malt; Turkey-Transcaucasia: Turkey, Armenia, Azerbaijan, and Georgia; South Asia: Pakistan and India; Central Asia: Kazakhstan, Russia, and Afghanistan. GDP modern lines have been assigned to breeding program based on the country where lines have been developed, as follows: South Mediterranean: Egypt, Algeria, Tunisia, Syria, Lebanon, Jordan, and Morocco; North America: Canada + United States (including Desert durum); Central Europe: Austria, Serbia, Hungary, Ukraine, and Bulgaria; Central Asia: Kazakhstan and Uzbekistan. For each summary, the first table reports results of AMOVA, the second table contains computations about gene diversity within groups/populations. AMOVA was significant at $p < 10^{-6}$ upon 10,000 permutations. LSD for within group gene diversity was calculated at the indicated p-value considering n = the least number of group genotypes, except for the landrace group where n = 5 was chosen. For decade of release, a third table reports the composition of each decade group with respect to the breeding programs.

Thus, five decades ('70–'80, '81–'90, '91–'00, '01–'10, '11–'18) and 12 breeding program groups (Australia, North America, Central Europe, Central Asia, France, Italy, South America, Spain, South Mediterranean, Ethiopia, ICARDA, CIMMYT) (**Supplementary Table S3**) were considered. For temporal groups (decades), AMOVA analysis revealed a very low, even if statistically significant, percentage of variation among groups (2.95%, **Table 2D**), attributing the near totality of variance to individuals within groups. Nei's gene diversity showed a constant decreasing trend starting from the decade ('81–'90) to the most recently released (2011–2018), with limited but significant variation. The mean number of pairwise differences within a decade (**Figure 2C**), and pairwise F_{st} among groups confirmed the trend; the highest difference in F_{st} values was observed in the comparison between the '70–'80s and the 2011–2018 decades, confirming a progressive and generalized shift toward the enrichment of fewer successful haplotypes during breeding history (**Figure 2C**).

The last analysis considered the *modern* germplasm, clustered according to breeding groups. AMOVA attributed the highest proportion of molecular variance (88.33%, **Table 2E**) to individuals within breeding programs, while variation between populations accounted for the remaining portion (11.67%). Moderate levels of diversity were observed for Australia and CIMMYT showing the lowest values (0.255 and 0.256, respectively), followed by ICARDA (0.294), North America (0.296), and Ethiopia (0.297), up to highest values calculated for Italy (0.343), Central Asia and France (0.339), and South America (0.335) (**Table 2E**). As for among-population comparisons, the Italian *modern* group showed generally lower pairwise F_{st} values as compared to all the other groups, with relatively higher values against the Northern programs and lower values against the other Mediterranean groups (**Figure 2D**). A reverse pattern of differentiation was evident for the French breeding programs, showing stronger similarities with the Northern programs. Low F_{st} values were calculated for pairwise comparisons among Central Europe, North America and Central Asia programs. Likewise, both CIMMYT and ICARDA showed the highest F_{st} values in the comparison with these breeding groups and the lowest F_{st} values with the Mediterranean groups. Between them, ICARDA and CYMMIT showed a F_{st} = 0.09. Analogously, low F_{st} values evidenced known interactions of international breeding programs with national programs, like ICARDA vs. Ethiopia and North African countries. The Australian breeding program appeared to stand as a separate group.

LD Decay

Genome-wide LD decay was calculated for the two major *T. turgidum* ssp. *durum* groups of the GDP collection: *modern* and *landraces*. As expected, LD was lower in *landraces* than in *modern* lines (**Figure 3**). The critical r^2 values of 0.3 and 0.5 were reached at a distance of 0.9–0.4 Mbp in *landraces*, and at distances of 4.2–1.8 Mbp in *modern*. Overall, 95% of unlinked markers showed a r^2 value <0.09 in *landraces* and 0.04 in *modern*. These r^2 values corresponded to distances of 4.2 Mbp in *landraces* and of 42.3 Mbp in *modern*. **Supplementary Figure S12**

A Tetraploid by historical selection steps

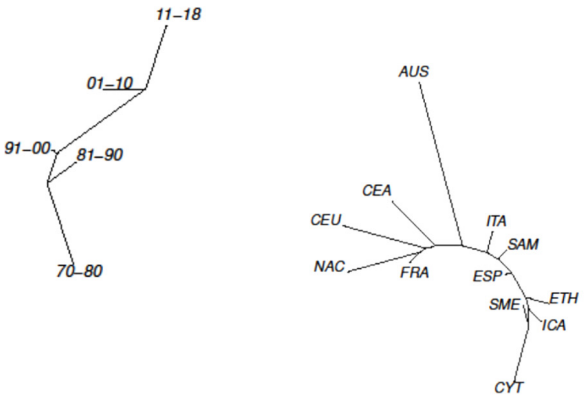
Historical tetraploid groups (n. acc)	EMMER	LANDRACE	MODERN
EMMER (103)	3220.5	4523.9	4791.5
LANDRACE (286)	0.2688	3375.1	3498.4
MODERN (473)	0.3988	0.1378	2681.8

B *T. durum* by main groups

T. durum main groups (n. acc)	EPO	MODERN	LANDRACE
EPO (96)	2567.0	2842.8	3311.0
MODERN (473)	0.1017	2538.2	3284.2
LANDRACE (286)	0.1395	0.1272	3151.9

C *T. durum* modern lines by breeding decades

T. durum breeding decades (n. acc)	1970-1980	1981-1990	1991-2000	2001-2010	2011-2018
1970-1980 (19)	2027.9	*2062.2*	2053.7	2053.6**	2082.5
1981-1990 (62)	*0.0060*	2069.9	2070.6	2076.9	2063.3
1991-2000 (93)	0.0250	0.0223	1979.9	1971.5	1966.7
2001-2010 (132)	0.0421	0.0416	0.0123	1914.9	1932.9
2011-2018 (143)	0.0717	0.0503	0.0253	0.0247	1855.3



D *T. durum* modern lines by breeding programs

T. durum breeding programs (n.acc)	Australia	North America	Central Europe	Central Asia	France	Italy	South America	Spain	North Africa	Ethiopia	ICARDA	CIMMYT
Australia (12)	1506.7	2096.7	2230.1	2209.8	2191.8	2201.1	2133.1	2104.4	2050.1	2055.0	2059.7	1979.0
North America (33)	0.2149	1749.6	2155.3	2175.0	2039.7	2225.6	2167.1	2187.2	2236.9	2213.6	2224.9	2226.8
Central Europe (25)	0.2267	0.1544	1902.3	2246.0	2150.7	2280.7	2273.9	2275.2	2342.5	2291.1	2324.8	2380.1
Central Asia (14)	0.1986	0.1407	0.1304	2006.8	2231.7	2298.2	2266.9	2257.9	2274.6	2235.8	2258.9	2296.0
France (39)	0.1818	0.0787	0.0903	0.1010	2006.0	2127.3	2168.3	2165.7	2209.3	2194.4	2202.3	2197.4
Italy (80)	0.1737	0.1458	0.1347	0.1208	0.0509	2031.2	2119.3	2072.0	2058.2	*2048.5*	2050.6	2009.1
South America (25)	0.1705	0.1401	0.1452	0.1200	0.0797	0.0521	1984.1	2035.2	2029.8	*2000.4*	2015.8	1994.3
Spain (27)	0.1641	0.1523	0.1502	0.1214	0.0834	0.0354	0.0303	1963.1	1940.2	*1958.7*	1957.7	1855.3
North Africa (53)	0.1681	0.1946	0.2007	0.1567	0.1287	0.0568	0.0574	0.0186	1848.1	*1867.0*	1829.6	1748.0
Ethiopia (8)	0.2112	0.2085	0.1964	0.1519	0.1326	0.0649	0.0587	0.0448	0.0317	1755.7	*1811.7*	1808.3
ICARDA (110)	0.1987	0.2156	0.2224	0.1806	0.1544	0.0812	0.0807	0.0573	0.0194	0.0353	1741.8	1795.5
CIMMYT (46)	0.2366	0.2700	0.2925	0.2491	0.2014	0.1128	0.1297	0.0670	0.0379	0.1076	0.0903	1513.6

FIGURE 2 | Population differentiation calculated as pairwise F_{st} and average number of pairwise differences between groups/populations defined according to passport data for: **(A)** evolution from domesticated emmer, to landraces, to modern lines; **(B)** all *T. durum* groups of EPO, landraces, modern lines; **(C)** *T. durum* modern lines classified according to decade of release; **(D)** *T. durum* modern lines classified based on breeding program. In each matrix, above diagonal elements (shades of green) contain the average number of pairwise differences, while below diagonal elements (shades of blue) report pairwise F_{st} values. Diagonal elements (shades of red) contain gene diversity within groups calculated as mean number of pairwise differences. Significance was assessed upon 1000 permutations. All values are significant at $p < 0.001$, except values marked with ** which were significant at $p < 0.01$, or values in italics that were not significant. Relative Neighbor-Joining phylogenetic tree based on Nei's distance are also reported for panels **(C,D)**.

reports LD calculated for each chromosome and for *modern* and *landraces*, independently.

Detection of Putative Selection Signals in Durum Wheat Groups

Considering the durum *modern* germplasm and its whole MAF-unfiltered SNP dataset of 16,633 SNPs, 889 *unique* breeding program-specific alleles were found (5.4% of the total, **Supplementary Table S5**). "Unique" is used to define a minor allele that occurs only in the germplasm of one breeding program and not in any other. The groups with the largest set of unique alleles were Central Europe, Central Asia, and Italy, with 289, 208, and 102 unique alleles, respectively (**Table 3**). Ethiopia and Australia were characterized by the lowest number of unique alleles with 13 and 9, respectively. It was then possible to identify rare alleles (with MAF less than 0.05) within the group of unique alleles. In particular, *rare unique alleles* were observed in all of the breeding groups except Australia, South America, and Ethiopia, ranging between 39 and 100% of the unique alleles. It was interesting to note that for CIMMYT and ICARDA, 100% of unique alleles were also rare, similarly to Italy (99%). Among the remaining unique alleles, none was a frequent allele in the target breeding group, and most (64%) had frequency from 5 to 10%. However, 53 SNPs showed higher frequency, suggesting a role in a specific breeding target or for adaptation to the corresponding environmental conditions.

Fixation of loci controlling traits of interest by intense selection during the breeding process may result in steep increases in allele frequency, reduced variation (reported as a *selective sweep*), and therefore divergence in allele frequency in the proximity of the selected loci. Low-resolution genomic scans can be used to identify regions containing loci and causative genes with a putative major influence on breeding processes. Scans for PSW between *modern* and *landraces* (**Supplementary Table S6**) identified 53 PSW clusters, based on F_{st} only (24) or

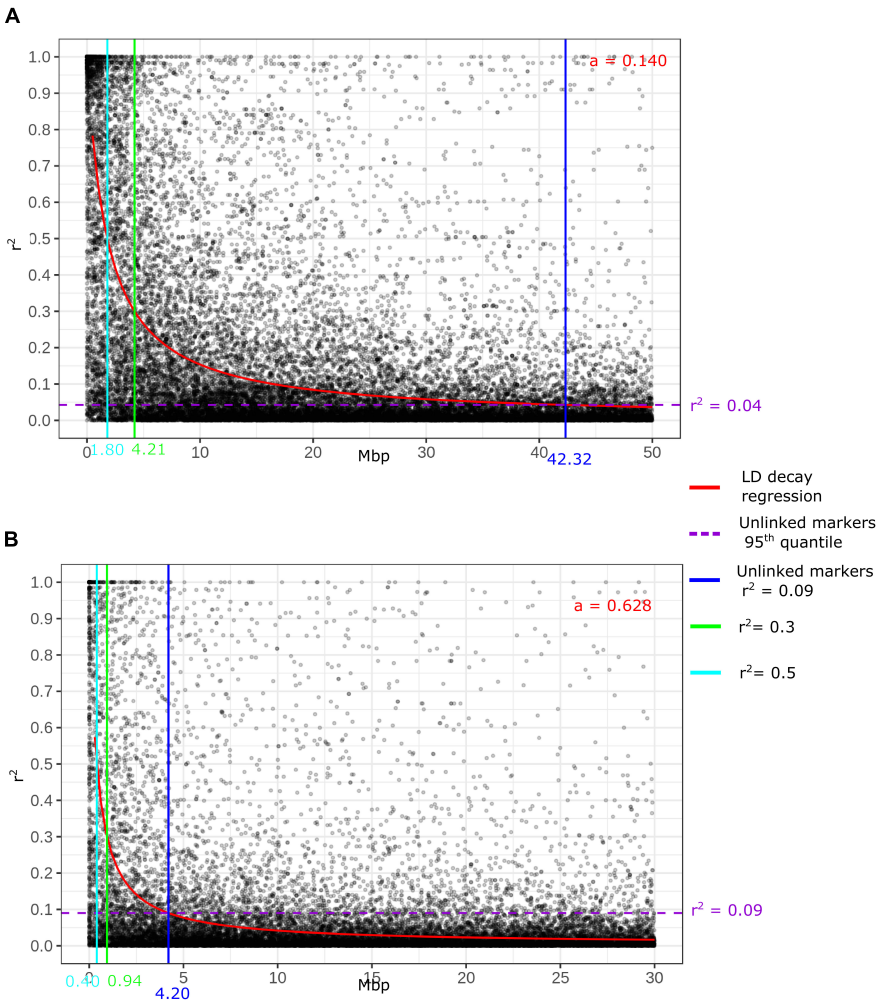


FIGURE 3 | Genome wide linkage disequilibrium (LD) decay in respect to physical distance in the two main groups of the GDP collection: **(A)** *modern* germplasm, **(B)** *landraces*.

on both indices, F_{st} and DRI (8). Most clusters (73%) extended for less than 50 Mbp, but three extended for >150 Mbp. All chromosomes were found to carry PSW clusters, with chromosome 1B being the most targeted by breeders' selection. Promising putative candidate genes were found to co-locate with eleven PSW clusters, for instance the genes *Rht1-B* and *Ppd-A1* on chromosomes 4B and 2A, respectively (**Supplementary Table S6**). Considering four subsequent decades of release, 62 putative signal clusters were highlighted across all six pairwise comparisons between the four decades (**Supplementary Table S7**). Chromosome 2B showed the highest number (9) of PSW clusters, whereas only two clusters per chromosome were identified on chromosomes 4A, 4B and 5B. Considering the five decades comparisons separately, 92 putative signals were found for *DRI*, 74 for F_{st}, and 46 were confirmed by both methods. The signals were distributed across the four comparisons: 30 were found for the '70–'80 vs. '81–'90 decades, 33 for both the comparisons '81–'90 vs. '91–'00 and '91–'00 vs. '01–'10, and 24 for '01–'10 vs. '11–'18. Most clusters were identified for two different decade comparisons (32, 10, and 2 PSW clusters, respectively), while 18 PSW clusters were detected in a single comparison. PSW clusters physical size extended from 11 Mbp for cluster *Cls-chr3B.1* to 386 Mbp for Cls-*chr6A.4*, with an average of 52 Mbp (**Supplementary Table S7**). As expected, the largest clusters were predominantly located in centromeric and peri-centromeric regions. Promising putative candidate genes were found to co-locate with nine PSW clusters (**Supplementary Table S7**).

Further pairwise comparisons were carried out for breeding groups that contributed more than 30 entries to the GDP (**Figure 4**). This investigation included modern *T. durum* genotypes from CIMMYT, ICARDA, Italy, France and North America, for a total of 10 pairwise comparisons. In total, 126 PSW clusters were identified (**Supplementary Table S8**), 59 of them supported by both indices, 40 based on *DRI* only, and 28 by *Fst* only. PSW cluster size ranged between 11 and 468 Mbp, with an average of 45.7 Mbp, and most clusters (81%) extending for less than 50 Mbp. Clusters were found in two or more comparisons (54), and only five were pair-specific. For 19 clusters

TABLE 3 | Unique alleles in the different breeding groups.

Breeding groups	N° of unique alleles	N° of unique alleles with MAF > 5% in the target group	Average frequency of the unique alleles in the target group
Central Europe	289	122	0.11
Central Asia	208	208	0.07
Italy	102	1	0.03
ICARDA	60	0	0.01
North Africa	57	2	0.02
North America	49	9	0.04
France	46	9	0.03
South America	23	14	0.09
Spain	22	5	0.05
CIMMYT	21	0	0.02
Ethiopia	13	13	0.14
Australia	9	9	0.16
Tot	899	392	0.06

a possible correspondence with a putative candidate gene could be proposed. The North American breeding group had the lowest number of PSW clusters (79), followed by CIMMYT with 88 clusters and the French breeding program with 100 PSW clusters. ICARDA and the Italian breeding programs had the highest numbers, 105 and 110, respectively. Considering pair-specific PSW clusters, CIMMYT and French groups showed the lowest number of specific PSW clusters (9), while Italy and ICARDA presented 12 and 11, respectively, and North America showed the highest number of specific PSW clusters (18).

GDP Stratification Analysis

Population stratification was conducted based on both Ward's clustering and admixture sub-population membership from $k = 2$ up to $k = 20$ based on the SNP dataset pruned at $r^2 = 0.5$. Results of these analysis are shown in **Figure 4C** while **Supplementary Table S9** reports sub-population memberships for each genotype and *K* value based on the two analyses. Applying SNP pruning with $r^2 = 0.8$ outperformed the other two in terms of cross-validated group assignment (**Figure 4A**), although pruning at $r^2 = 0.5$ provided comparable results. Grouping statistics, in particular the minimum group size (**Figure 4B**), stabilized at $k > 11$, despite the fact that cross-validated assignment error steadily decreased at higher *k* values (**Figure 4A**) and meaningful differences were still observed up to *k* values of 20. At $k = 2$, most accessions of *T. turgidum* spp. *dicoccum* (98%), *dicoccoides* (98%), *carthlicum* (92%) and *turgidum* (77%) clustered together (reported as dark yellow *Q* membership bars in **Figure 4C**), separated from all the durum wheat entries (reported as dark blue *Q* membership bars in **Figure 4C**). Notably, a small group of 33 (4%) of landraces from Ethiopia and the Arabian Peninsula clustered in the former group, showing appreciable genetic kinship with emmer from the Fertile Crescent. At $k = 5$, the emmer group was split in two main branches, one grouping wild emmer together with European and Fertile Crescent domesticated emmers, and the second having domesticated emmers from the Fertile Crescent together with Ethiopian durum and *T. turgidum* ssp. *carthlicum* entries. At $k = 20$, emmer accessions were further split between central Asian domesticated emmer (subp. 11), European domesticated emmer (subp. 12) and wild emmer (subp. 13).

At $k = 2$, the second mega-cluster included most *T. turgidum* ssp. *durum* (96%), *T. turgidum* ssp. *turanicum* and most of *T. turgidum* ssp. *polonicum* (67%). Separation between durum *modern* and *landraces* started at $k = 3$. At $k = 6$, durum landraces and primitive tetraploids were split into two main groups: Asian and North African landraces. Further meaningful landrace sub-groups were split at higher *k* values. The group including mainly Ethiopian accessions was split in two sub-groups: the first one contained accessions of *T. turgidum* spp. *carthlicum*, *polonicum* and *durum* landraces, while the second one was mainly *T. turgidum* ssp. *dicoccum* accessions, which might represent the founder group of Ethiopian durums.

Durum landraces and primitive tetraploids were grouped into subpopulations as follows: Central Mediterranean landraces (subp. 5), a mixed group of other Mediterranean landraces and old Italian cultivars such as the breeding germplasm founder Cappelli, and (subp. 6) more recent Italian cultivars directly related to landraces (subp. 7), Ethiopian durum landraces and emmers plus *T. turgidum* ssp. *carthlicum* (subp. 8), Central Asia durum landraces and all *T. turgidum* ssp. *turanicum* (subp. 10). Notably, sub-population 9 included a group of ICARDA founder cultivars belonging to the Om Rabi set, which were derived from crossing the Syrian landrace Haurani to the CIMMYT cultivar Jori (Kabbaj et al., 2017).

The modern durum germplasm was first split at $k = 4$ separating photoperiod sensitive accessions from northern countries (North America, France, Austria and the EPO entries) and Mediterranean-adapted photoperiod insensitive accessions. $K = 10$ was the minimum *k* value at which both Ward's clustering and *ADMIXTURE* clearly separated the modern durum entries originating from the two main CGIAR (CIMMYT and ICARDA) breeding programs. At $k = 13$, modern durum entries were already divided in four sub-sets corresponding to French origin and EPO (subp. 1), CIMMYT (subp. 2), ICARDA (subp. 3), North American and Austrian (subp. 4). At $k = 18$ the group containing mainly CIMMYT durum wheat modern lines was further split in three sub-groups: the first one contained CIMMYT and other modern lines with different origins, the second one included CIMMYT and Egyptian germplasm, and the third one only modern germplasm from the Mediterranean countries. Only at $k = 20$ was the EPO set split into two groups.

The GDP phylogenetic tree estimated through Neighbor-Joining clustering for all accessions is reported in **Figure 5** and **Supplementary Table S9**. Bootstrap values indicating branches' consistency are reported in detail in **Supplementary Figure S13**. Overall, good correlation was observed between population stratification analysis performed through admixture and the position on the Neighbor-Joining tree. Three main branches were grouped: (*i*) wild and domesticated emmers and *T. turgidum* ssp. *carthlicum*, (*ii*) durum landraces including the founders of modern germplasm and (*iii*) modern durums. Among durum landraces, one of the two sub-branches included

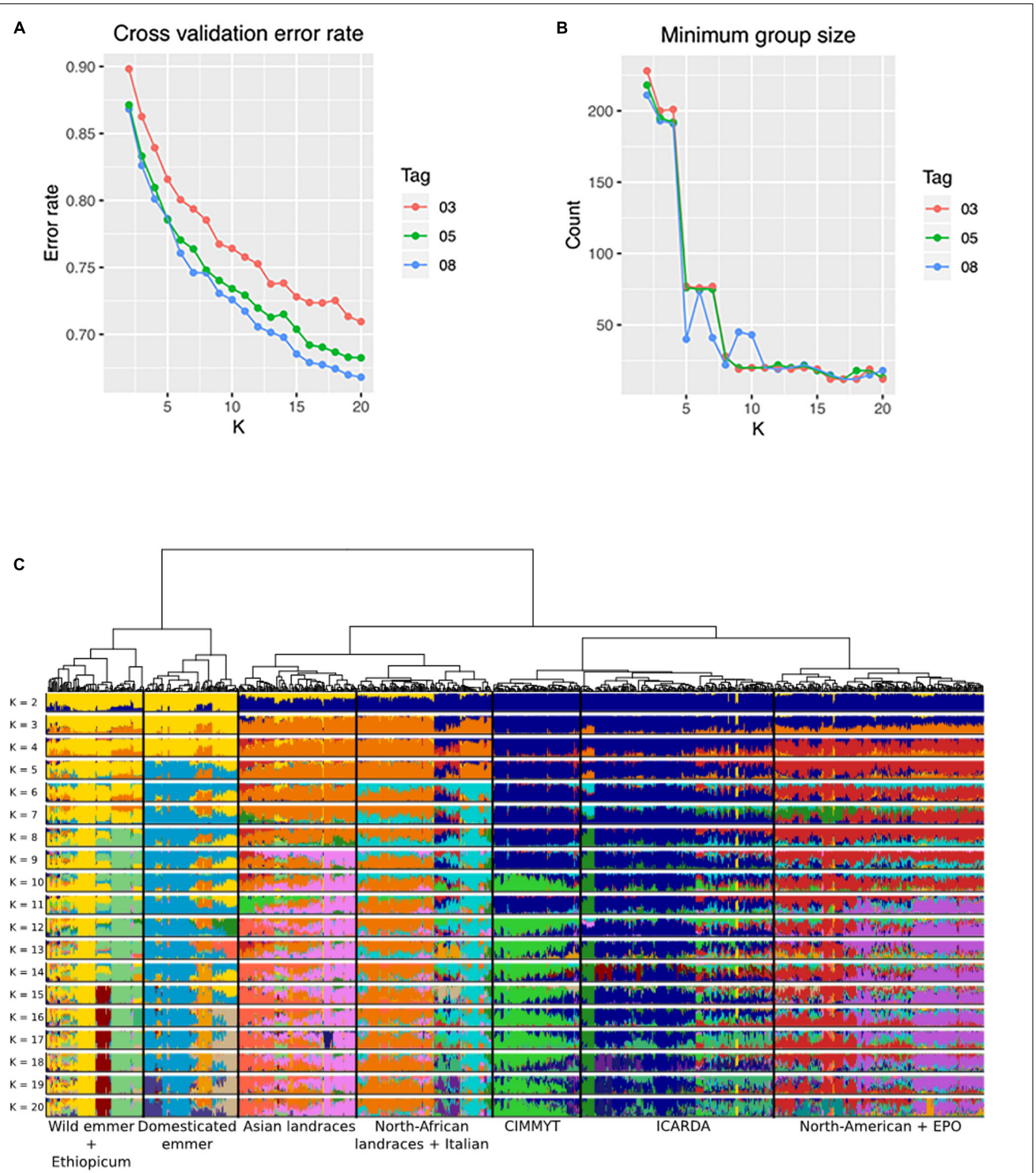


FIGURE 4 | *ADMIXTURE*'s grouping statistics: **(A)** cross validation error rate, and **(B)** minimum group size, from $k = 2$ to $k = 20$ for three LD pruned SNP datasets ($r^2 = 0.3$, $r^2 = 0.5$, $r^2 = 0.8$); **(C)** population structure of the GDP collection based on Ward's clustering and *ADMIXTURE* (SNP dataset at $r^2 = 0.5$); membership from $k = 2$ to $k = 20$.

North African/Southern European landraces and pioneering durum cultivars obtained from landrace selection and landrace intercrossing, such as Senatore Cappelli (selection from a landrace) and Capeiti8 (cross between Cappelli and a Syrian landrace selection). The second group included durum landraces from West Asia including Haurani, well-known as the most

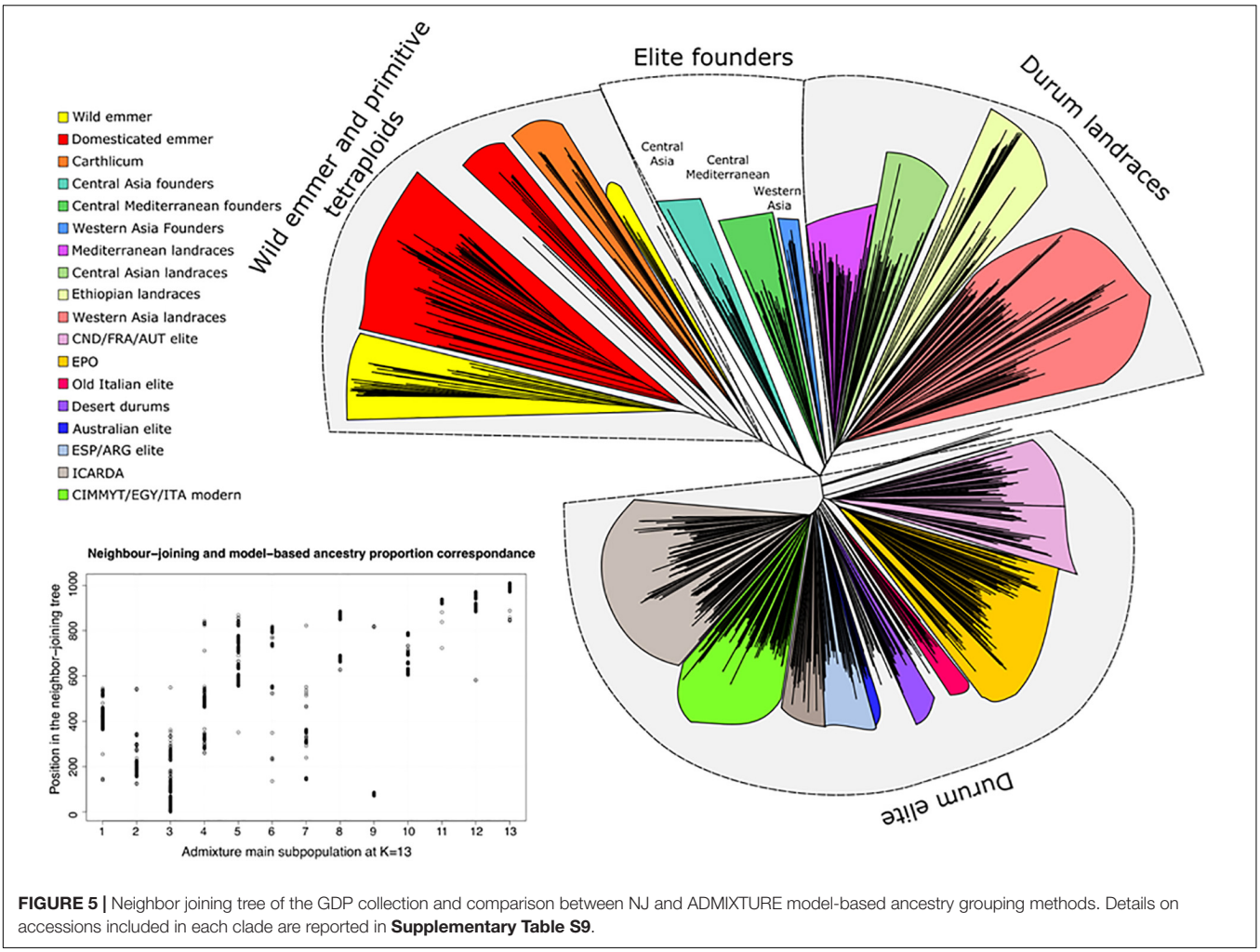


FIGURE 5 | Neighbor joining tree of the GDP collection and comparison between NJ and ADMIXTURE model-based ancestry grouping methods. Details on accessions included in each clade are reported in **Supplementary Table S9**.

widely cultivated landrace population in its area of origin, showing developmental and morphological traits relevant for adaptation to low water availability and high temperatures, widely exploited by the ICARDA durum program since its inception (Elings and Nachit, 1991; Pagnotta et al., 2005). Another small group of interest is that composed of Central-Asian durum landraces that were included phylogenetically within the emmer clade. This group was found to lie between the main emmer clades and the modern durum, supporting a possible role of its members as founders of the Northern breeding programs (Paulsen and Shroyer, 2008).

DISCUSSION

Genetic Diversity and Population Structure in GDP and Breeding Groups

The GDP builds on several studies that have investigated the diversity and phylogeny of durum wheat by assembling these into one panel. The two-step approach deployed here started by gathering entries representing nearly all genetic diversity studies ever conducted for durum wheat within the DWRC. In the second step, 1,011 entries were selected from the DWRC to capture most of this diversity (94–97%), with the strongest reduction affecting some rare alleles.

In the GDP, the mean PIC values of 0.27 for *landraces* and 0.28 for *modern* lines and ranging from 0.09 to 0.38 (Table 2B) indicated a generally higher or similar level of genetic diversity captured within the GDP compared to previously studied collections. Recent studies reported PIC values of 0.26 for durum modern germplasm (Chao et al., 2017), 0.19 for a set of both landraces and modern lines (Ren et al., 2012), and 0.18 in a collection of 168 durum wheat accessions of different origins (Roncallo et al., 2019). Analogously, AMOVA on clusters within GDP based on geography and breeding program of origin showed that only 13% of the total genetic variance could be captured among groups, while most diversity remained among individuals within clusters. These results concur with those reported by Soriano et al. (2016) with 172 landraces from 21 countries, by Roncallo et al. (2019) with a panel of 168 durum accessions and by N'Diaye et al. (2018) with a panel of Canadian durum cultivars where only 10% of variation was captured among groups. Other studies considering similar panels reported capturing over 30% of the total genetic variance by clustering germplasm based

on kinship matrix, but using relatively higher *k* values (Kabbaj et al., 2017; Robbana et al., 2019). Our study aimed primarily at evaluating the historical diversity based on passport information, rather than on clusters derived from population structure. It is therefore evident that the passport information alone, while of great historical interest, is unable to capture the true genetic diversity of durum wheat worldwide. AMOVA on stratified groups may reveal much more variance among sub-populations, as indeed reported by other authors (Kabbaj et al., 2017; Roncallo et al., 2019). The moderate diversification among breeding groups (11.67% of the total variance) and very little among decades of release (2.95% of the total variance) revealed by AMOVA on the 473 modern durum wheat accessions (Tables 2D,E) was probably due to the wide and frequent exchange of parents among durum breeders worldwide. This was clearly evidenced in the Italian breeding programs, characterized by an overall higher level of diversity and lower differentiation against most of the other breeding programs, thus reflecting the necessity to breed for the many different agro-ecological zones that exist in Italy (Fischer et al., 2012). Overall, the results presented here suggest that good genetic diversity remains available within the breeding groups for direct exploitation, and there is even greater potential when considering exchanges between breeding groups.

The EPO is an evolutionary durum wheat pre-breeding population obtained through initial crossing of modern French varieties with various tetraploid wheat subspecies (David et al., 2014). When compared to *landraces* and *modern* durum lines, EPO lines showed the same level of genetic diversity in terms of mean number of pairwise differences and expected heterozygosity of *modern* lines, indicating that the genetic background of EPO lines is relatively homogeneous while being enriched in exotic alleles.

Substantial agreement between NJ, ADMIXTURE and Ward's clustering indicated a complex, still well-defined stratification of the population, driven by historical, geographical and environmental factors. Phylogenetic analysis (**Figure 5**) highlighted three well-defined landrace groups of geographically distinct origin, holding a pivotal role as founders of different breeding programs. These included landraces from North Africa, West Asia and Central Asia as founders of modern breeding, in particular of ICARDA and Italy (Kabbaj et al., 2017; Soriano et al., 2018), while Central Asian landraces have played a critical role in the foundation of the North American modern durum germplasm via the early introduction from Russia and Turkey by Mennonite immigrants (Moon, 2008; Paulsen and Shroyer, 2008). The identification of these founders concurs with the results reported by Kabbaj et al. (2017), Maccaferri et al. (2019), and Taranto et al. (2020), supporting the validity of the phylogeny studies conducted for the GDP.

Putative Signature of Selection Across the Breeding History and the Breeding Groups

Intense breeding in the past decades led to the development of superior cultivars for a broad range of edaphic environments. Current varieties exhibit increased yield potential, spike fertility, pasta quality and are resistant to widespread diseases such as rusts. The process of selection has evidently resulted in "signatures" being incorporated into the durum wheat genome, specific to each breeder's targets and selection procedures, as well as shared preferences across breeding programs. The large set of *unique* alleles in the germplasm of historical breeding groups from Central Europe, Central Asia and Italy appear as a function of the longer effort to improve adaptation compared to more recent breeding groups. The large set of unique alleles, a high proportion of which were rare in Central Europe (58%) and Italy (99%), is consistent with extended selection for a particular environment. Studies aiming to describe allele fixation and genetic diversity are of great importance to guide breeders in planning their crosses and introgressions (Kabbaj et al., 2017; Taranto et al., 2020). In this regard, unique alleles can be seen as strategic targets for capturing exploitable genetic variability when linked to important traits.

The influence of selection on the genome was reflected in the diversity reduction index (*DRI*) and F_{st} metrics. Overall putative selection signals were found throughout the entire genome, including the centromeric regions. The average signal size of 50 Mbp suggested strong selection pressure. Several PSW clusters identified in this study co-located with known loci relevant to durum wheat breeding, thus demonstrating the predictive validity of the genome-wide search method. Expected signals associated with the transition from *landraces* to *modern* were related to the control of traits strongly selected in the post Green Revolution period causing the almost complete fixation of such loci in the *modern* subpopulations. As an example, *Cls-chr4B.2* included the widely used *Rht1-B* (Khush, 2001; Evenson and Gollin, 2003; Borojevic and Borojevic, 2005). This locus has also been identified as a putative signal of selection when comparing the '70–'80 and '81–'90 decades (*Cls-chr4B.1*, **Supplementary Table S7**) as well as when contrasting North American germplasm (tall cultivars) vs. Italy/France (semi-dwarf), and ICARDA (mix tall and semi-dwarf) vs. Italy (all semi-dwarf) breeding programs. Phenology is also a trait under strong and constant selection pressure, supported by the PSW cluster in the *landraces* vs. *modern* germplasm (**Supplementary Table S6**) that co-located with the photoperiod insensitive gene *Ppd-A1* (Beales et al., 2007; Maccaferri et al., 2008; Wilhelm et al., 2009; Bentley et al., 2011). The signal marked the transition from landraces to modern cultivars since the photoperiod insensitive allele was widely and positively selected, as already reported by Motzo and Giunta (2007). Following the Green Revolution, selection for photoperiod insensitivity continued as shown by the inclusion of both *PPD* homeologs on chromosomes 2A and 2B in cluster signals. PSW signals for the *Ppd-A1* and *Ppd-B1* regions were identified from comparisons of the Italian, French and ICARDA breeding groups vs. CIMMYT and North America groups, respectively (clusters *Cls_clv-chr2A.1* and *Cls_clv-chr2A.1*, *Cls-chr2B.1;* **Supplementary Table S8**), indicating a generalized selection strategy to fine tune the photoperiod insensitive alleles

to match the ideal phenology for the targeted environment (Maccaferri et al., 2008).

Another important class of genes known to have undergone strong selective pressure in bread wheat are the *VRN*. In contrast to *PPD*, the PSW signal for *VRN* loci was much weaker in the GDP. For instance, no PSW cluster included *Vrn1*-5A (Yan et al., 2003), while *Vrn3*-7A (Yan et al., 2006) generated PSW signals in both A and B sub-genomes. For example, *Cls-chr7A.4* was identified in the North American group vs. ICARDA, CIMMYT and Italy; *Cls-chr7A.5* was identified for the comparisons of CIMMYT vs. ICARDA and Italy; and *Cls-chr7B.1* corresponded to *Vrn3-7B* for the comparisons of CIMMYT vs. France and Italy (**Supplementary Table S8**). Mild vernalization requirements are still present in *modern* cultivars for the Mediterranean areas where wheat is cultivated as a fall-sown cereal, and distinctions at these loci might depend on the breeder's target of extending or reducing the overall cycle in different agro-ecologies. Lastly, among the earliness *per se* genes, *ELF3-A1* (Zikhali et al., 2016) appears the most likely candidate for the PSW cluster *Cls-chr1A.8*, which differentiated both France and North America *modern* germplasm when comparing ICARDA and Italy (**Supplementary Table S8**).

PSW clusters could also be related to selection for increased spike fertility and grain yield potential, particularly in the *landrace* to *modern* comparisons (**Supplementary Table S6**). This is the case of *Cls-chr3B.2* and *Cls-chr7A.2* whose intervals include the determinant of grain weight identified in bread wheat *TaCKX6* (cytokinin oxidase/dehydrogenase, Zhang et al., 2012) and *TaTGW-7A* (Hu et al., 2016), respectively. Additionally, *Cls-chr2A.4* and *Cls-chr2B.3* overlapped with the recently cloned gene related to floret fertility *GNI-A1* (Sakuma et al., 2019), while in some comparisons among breeding groups (**Supplementary Table S8**) a PSW cluster (*Cls-chr2A.3*) overlaps with *TaSus2* (sucrose synthase), a main driver of starch accumulation in wheat found to be associated with strong changes in haplotype frequency in bread wheat (Hou et al., 2014). Considering nitrogen metabolism and grain protein content, an important quality trait for durum wheat, the *landraces* vs. *modern* contrast co-located *Cls-chr2A.5* and *Cls-chr2B.5* with genes encoding for glutamine synthase *GS2-2A* and *GS2-2B* (**Supplementary Table S8**). Both these genes play a key role in high protein content (Gadaleta et al., 2011). Clusters could be related to selection for quality of grain proteins as shown by *Cls-chr1B.4* and *Cls-chr6A.1* overlapping with genes for glutenins (*Glu-B1,* Xu et al., 2008) and gliadins (*Gli-6A,* Gu et al., 2004), respectively. In particular, *Cls-chr6A.1* was detected for *landraces* vs. *modern* and for three breeding programs pairwise comparisons (i.e., ICARDA, CIMMYT and Italy vs. North America and France) (**Supplementary Tables S6, S9**), while *Cls-chr1A.1* was identified in three decade pairwise comparisons and in ICARDA vs. CIMMYT (**Supplementary Tables S7, S9**). The co-localization between PSW clusters and glutenin and gliadin alleles is not unexpected given the influence of these genes on pasta quality, which is a major target of selection. Convincingly, three chromosomes, 1A, 1B, and 6A, involved in seed storage proteins were represented in the PSW clusters: *Cls-chr1A.1* (PSW found for decade and breeding program pairwise comparisons, co-locating with *Glu-A3* and gliadins*), Cls-chr1B.4 (Glu-B1), Cls-chr6A.1*, *Cls-chr6A.2* and *Cls-chr6A.3*, with the last three PSW partially overlapping and co-locating with *Gli-6A* (**Supplementary Tables S7, S9**).

Lastly, presence of gene candidates was observed for three strong PSW clusters that occurred in chromosome 7B (*Cls-chr7B.3*, centromeric and *Cls-chr7B.12*, distal) and in chromosome 5B (*Cls-chr5B.5*) and that are putatively related to grain quality. The two signals in chromosome 7B were associated to a strong QTL for grain yellow pigment content (reviewed in Colasuonno et al., 2019). The phytoene synthase, *Psy-B1*, a major gene responsible for yellow pigment content in the wheat grain and a common target of modern durum breeding for semolina color is a strong candidate (Pozniak et al., 2007). A signal for this locus emerged from the comparison of *landraces* vs. *modern* lines and North America (*Cls-chr7B.12*) vs. French and ICARDA breeding groups (**Supplementary Tables S6, S9**). The signal also appeared for three decade pairwise comparisons (**Supplementary Table S7**). suggesting a common historical selection for yellowness based on a number of co-located QTL clusters (Roncallo et al., 2012; Giraldo et al., 2016; Colasuonno et al., 2019) associated to specific *Psy-B1* alleles (reviewed in Colasuonno et al., 2019).

A recent study Taranto et al. (2020), aiming to define PSW among Italian cultivars and landraces also identified several of the selection sweeps proposed here, including the major loci controlling phenology and quality characteristics.

In summary, the report of PSW clusters in this manuscript is a first attempt to carry out such analysis across breeding programs from different countries. Although the causative genes of the PSW clusters remain to be verified, several plausible candidates have been proposed. The GDP provides then an unprecedented opportunity for international collaborations to more effectively harness and exploit the diversity identified here.

CONCLUSION

In the present study, a very large and diverse durum wheat panel referred to as the GDP has been assembled and made publicly available to drive further discovery and deployment of beneficial alleles. The GDP is maintained and distributed by ICARDA Genbank[5] under Terms and Conditions of SMTA. The genotypic datasets (both raw data and upon quality filtering and imputing) can be found in the online repositories GrainGenes (see text footnote 3), and T3/Wheat (see text footnote 4). The genetic characterization of this panel increases the knowledge of genetic relationships and population structure of worldwide durum wheat, while facilitating the identification of the optimal sources of genetic diversity for a given target locus. The entire durum community is now empowered to use this panel to discover novel and useful alleles via GWAS. Finally, since the GDP is an open resource available to the whole community, the discovery of useful alleles can be immediately incorporated

[5]http://indms.icarda.org/

in breeding activities irrespective of the country or research group that makes the discovery. This is particularly true now that a number of genomic resources are available for wheat, including the reference sequence of the durum wheat genome (Maccaferri et al., 2019). We believe that this international effort is a great example of how a whole community can come together to support breeders in their efforts to adapt and develop more resilient durum wheat varieties able to withstand climate change and ensure a great future for this important crop.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories: GrainGenes https://wheat.pw.usda.gov/GG3/global_durum_genomic_resources, and T3/Wheat https://wheat.triticeaetoolbox.org/breeders_toolbox/protocol/158.

AUTHOR CONTRIBUTIONS

FB, RT, MM, LC, JA, KA, and SX designed this initiative. EM, DM, SC, SX, JF, and MH produced the genotypic data and all authors supported the genotyping. EM, GS, AM, FD, GP, MM, RT, LC, and FB analyzed the data. EM, GS, AM, MM, and FB developed the first draft. All authors reviewed and approved the final version of this manuscript.

FUNDING

The work of the Global Durum wheat Panel was financially made possible by several international and national donors: the Wheat Initiative – Expert Working Group in Durum Wheat Genomics and Breeding supported the meetings and interactions of the durum wheat research community; CRP WHEAT (CIMMYT) supported the genotyping with KASP of the DWRC collection; "Adapting Agriculture to Climate Change: Collecting, Protecting and Preparing Crop Wild Relatives," which is supported by the Government of Norway, managed by the Global Crop Diversity Trust with the Millennium Seed Bank of the Royal Botanic Gardens, Kew supported the field work for seed purification and multiplication; CRP WHEAT CoA 3.2 supported the seed distribution to partners; import of seeds and respect of quarantine procedure was supported by USDA-ARS; the high resolution genotyping work was conducted under: H2020-MSCA-RISE 2015 EXPOSEED (ID: 691109) "Exploring the molecular control of seed yield in crops, PICT-2015-1401 ANPCyT- Argentina "Análisis de la estructura del genoma y mapeo por asociación para caracteres de calidad y rendimiento en trigo candeal," Genome Canada – Genome Prairie (Saskatchewan Ministry of Agriculture), FAO/ITPGRFA (W3B-PR-21 Morocco), Premier's Research and Industry Fund (Government of South Australia – IRGP15), GRDC (DAN00163), USDA-ARSUSDA-ARS (3060-21000-038-00D), Lieberman-Okinow Endowment at University of Minnesota, Grains Research and Development Corporation (GRDC) and the University of Adelaide, International Funding Initiative of Agriculture and Agri-Food of Canada, Saskatchewan Wheat Development commission, SeCan, and the Saskatchewan Ministry of Agriculture; data analysis was partially supported by: PRIMA2019 CEREALMED "Enhancing diversity in Mediterranean cereal farming systems," H2020 InnoVar Project "Next generation variety testing for improved cropping on European farmland," APSOV-UNIBO 2020–2022 Research Agreement "Identification of loci and markers of agronomic interest in wheat" and MIPAAF Italy Systemic-1063 "An integrated approach to the challenge of sustainable food systems: adaptive and mitigatory strategies to address climate change and malnutrition." Several partners dedicated time and effort in kind to ensure the good outcome of this initiative.

ACKNOWLEDGMENTS

The authors wish to thank Mr. Abu Nakad Rukoz, M.me Nada Saghbini, and M.me Hoda Abou Younes for maintaining, multiplying, and distributing the pure seeds of the GDP collection in Lebanon. Recognition goes also to the several germplasm donors that supported this international initiative.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2020.569905/full#supplementary-material

Figure S1 | Population structure of the DWRC collection based on ADMIXTURE analysis.

Figure S2 | Population structure of the DWRC collection based on bootstrapped Ward's clustering.

Figure S3 | Bootstrapped Ward's clustering of the DWRC subgroup of *T. durum* cultivars, varieties and élite lines.

Figure S4 | Bootstrapped Ward's clustering of DWRC subgroup EPO.

Figure S5 | Bootstrapped Ward's clustering of DWRC subgroup of *T. durum* landraces.

Figure S6 | Bootstrapped Ward's clustering of DWRC subgroup of *T. dicoccum* accessions.

Figure S7 | Bootstrapped Ward's clustering of DWRC subgroup of *T. dicoccoides* accessions.

Figure S8 | Bootstrapped Ward's clustering of DWRC subgroup of *T. turgidum* subspecies *carthlicum, aethiopicum, polonicum, turanicum, turgidum*.

Figure S9 | Sampling effect on genetic diversity between DWRC and GDP: correlation of diversity indexes between GDP and DWRC for Shannon-Wiener index, expected heterozygosity, evenness, and minor allele frequency.

Figure S10 | Site frequency spectrum of loci in GDP and DWRC.

Figure S11 | Distribution of the SNPs along the chromosome and inter SNP distances. **(A)** Average number of SNPs per classes of interlocus distances, across all chromosomes; **(B)** number of SNPs per each chromosome segment, from proximal (1) to distal (10) regions, mediated across all chromosomes; **(C)**: interlocus distances in each chromosome segment, from proximal (1) to distal regions (10), presented for all chromosomes combined.

Figure S12 | Local LD of *T. durum* landraces and modern lines presented for each chromosome.

Figure S13 | Bootstrap neighbor joining phylogenetic tree of GDP.

Table S1 | List of private companies, institutions, international organizations which contributed tetraploid wheat germplasm to the initial DWRC.

Table S2 | DWRC: **(A)** list of accessions constituting the DWRC; **(B)** scoring of DWRC accessions based on KASP marker set; **(C)** KASP markers list used for the DWRC genotyping.

Table S3 | List of accession constituting the GDP, with passport data. The categories based on passport data used to classify accessions for the diversity analyses are also reported, as well as available data about flowering habit and allele status at some known genes (*Rht*, *Ppd*, *Cdu*, *Vrn*, etc.).

Table S4 | Genetic distance matrix of the GDP.

Table S5 | List of unique alleles within the breeding groups of the GDP.

Table S6 | PSWs between *modern* and *landraces:* **(A)** list of clusters of PSWs with position on the Svevo reference genome, metrics detecting PSWs, and the candidate gene; **(B)** significant values for each metrics (F_{st}, DRI) for each SNP sliding window.

Table S7 | PSWs for modern, between different decades: **(A)** list of clusters of PSWs between different decades, with position on the Svevo reference genome, metrics detecting PSWs in each comparison, and the candidate gene; **(B)** significant values for each metrics (F_{st}, DRI) for each SNP sliding window in each comparison.

Table S8 | PSWs for *modern*, between different breeding groups: **(A)** list of clusters of PSWs between different breeding programs, with position on the Svevo reference genome, metrics detecting PSWs for each comparison, and the candidate gene; **(B)** significant values for each metrics (F_{st}, DRI) for each SNP sliding window in each comparison.

Table S9 | Stratification analysis of GDP: **(A)** grouping on the base of the main model-based ancestry estimation and neighbor joining tree position. Accessions are sorted on the base of their position on the NJ tree of Figure 5 and colors correspond to those of groups highlighted in the same Figure 5; **(B)** Ward's clustering of GDP from K2 to K20; **(C)** membership value for each GDP accession at K = 13 based on ADMIXTURE ancestry estimation.

REFERENCES

Abu-Zaitoun, S. Y., Chandrasekhar, K., Assili, S., Shtaya, M. J., Jamous, R. M., Mallah, O. B., et al. (2018). Unlocking the genetic diversity within a middle-east panel of durum wheat landraces for adaptation to semi-arid climate. *Agronomy* 8:233. doi: 10.3390/agronomy8100233

Alexander, D. H., Novembre, J., and Lange, K. (2009). Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* 19, 1655–1664. doi: 10.1101/gr.094052.109

Baloch, F. S., Alsaleh, A., Shahid, M. Q., Çiftçi, V., Sáenz De Miera, L. E., Aasim, M., et al. (2017). A whole genome DArTseq and SNP analysis for genetic diversity assessment in durum wheat from central fertile crescent. *PLoS One* 12:e0167821. doi: 10.1371/journal.pone.0167821

Bassi, F. M., Brahmi, H., Sabraoui, A., Amri, A., Nsarellah, N., Nachit, M. M., et al. (2019). Genetic identification of loci for Hessian fly resistance in durum wheat. *Mol. Breed.* 39:24. doi: 10.1007/s11032-019-0927-1

Bassi, F. M., and Nachit, M. M. (2019). Genetic gain for yield and allelic diversity over 35 years of durum wheat breeding at ICARDA. *Crop Breed. Genet. Genomics* 1, 1–19. doi: 10.20900/cbgg20190004

Beales, J., Turner, A., Griffiths, S., Snape, J. W., and Laurie, D. A. (2007). A pseudo-response regulator is misexpressed in the photoperiod insensitive Ppd-D1a mutant of wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 115, 721–733. doi: 10.1007/s00122-007-0603-4

Bentley, A. R., Turner, A. S., Gosman, N., Leigh, F. J., Maccaferri, M., Dreisigacker, S., et al. (2011). Frequency of photoperiod-insensitive Ppd-A1a alleles in tetraploid, hexaploid and synthetic hexaploid wheat germplasm. *Plant Breed.* 130, 10–15. doi: 10.1111/j.1439-0523.2010.01802.x

Borojevic, K. K., and Borojevic, K. K. (2005). Historic role of the wheat variety akakomugi in Southern and Central European wheat breeding programs. *Breed. Sci.* 55, 253–256. doi: 10.1270/jsbbs.55.253

Borrelli, G., and Trono, D. (2016). Molecular approaches to genetically improve the accumulation of health-promoting secondary metabolites in staple crops - a case study: the lipoxygenase-B1 genes and regulation of the carotenoid content in pasta products. *Int. J. Mol. Sci.* 17:1177. doi: 10.3390/ijms17071177

Botstein, D., White, R. L., Skolnick, M., and Davis, R. W. (1980). Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am. J. Hum. Genet.* 32, 314–331.

Bradbury, P. J., Zhang, Z., Kroon, D. E., Casstevens, T. M., Ramdoss, Y., and Buckler, E. S. (2007). TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* 103, 338–348. doi: 10.1016/J.AJHG.2018.07.015

Canè, M. A., Maccaferri, M., Nazemi, G., Salvi, S., Francia, R., Colalongo, C., et al. (2014). Association mapping for root architectural traits in durum wheat seedlings as related to agronomic performance. *Mol. Breed.* 34, 1629–1645. doi: 10.1007/s11032-014-0177-1

Cavanagh, C. R., Chao, S., Wang, S., Huang, B. E., Stephen, S., Kiani, S., et al. (2013). Genome-wide comparative diversity uncovers multiple targets of selection for improvement in hexaploid wheat landraces and cultivars. *Proc. Natl. Acad. Sci. U.S.A.* 110, 8057–8062. doi: 10.1073/pnas.1217133110

Chang, C. C., Chow, C. C., Tellier, L. C. A. M., Vattikuti, S., Purcell, S. M., and Lee, J. J. (2015). Second-generation PLINK: rising to the challenge of larger and richer datasets. *Gigascience* 4:7. doi: 10.1186/s13742-015-0047-8

Chao, S., Rouse, M. N., Acevedo, M., Szabo-Hever, A., Bockelman, H., Bonman, J. M., et al. (2017). Evaluation of genetic diversity and host resistance to stem rust in USDA NSGC durum wheat accessions. *Plant Genome* 10:plantgenome2016.07.0071. doi: 10.3835/plantgenome2016.07.0071

Clarke, J. M., Clarke, F. R., and Pozniak, C. J. (2010). Forty-six years of genetic improvement in Canadian durum wheat cultivars. *Can. J. Plant Sci.* 90, 791–801. doi: 10.4141/cjps10091

Colasuonno, P., Marcotuli, I., Blanco, A., Maccaferri, M., Condorelli, G. E., Tuberosa, R., et al. (2019). Carotenoid pigment content in durum wheat (*Triticum turgidum* L. var durum): an overview of quantitative trait loci and candidate genes. *Front. Plant Sci.* 10:1347. doi: 10.3389/fpls.2019.01347

Condorelli, G. E., Maccaferri, M., Newcomb, M., Andrade-Sanchez, P., White, J. W., French, A. N., et al. (2018). Comparative aerial and ground based high throughput phenotyping for the genetic dissection of NDVI as a proxy for drought adaptive traits in durum wheat. *Front. Plant Sci.* 9:893. doi: 10.3389/fpls.2018.00893

David, J., Holtz, Y., Ranwez, V., Santoni, S., Sarah, G., Ardisson, M., et al. (2014). Genotyping by sequencing transcriptomes in an evolutionary pre-breeding durum wheat population. *Mol. Breed.* 34, 1531–1548. doi: 10.1007/s11032-014-0179-z

De Vita, P., Nicosia, O. L. D., Nigro, F., Platani, C., Riefolo, C., Di Fonzo, N., et al. (2007). Breeding progress in morpho-physiological, agronomical and qualitative traits of durum wheat cultivars released in Italy during the 20th century. *Eur. J. Agron.* 26, 39–53. doi: 10.1016/j.eja.2006.08.009

Dexter, J. (2008). "The history of durum wheat breeding in Canada and summaries of recent research at the Canadian grain commission on factors associated with durum wheat processing," in *Presented at Bosphorus 2008 ICC (International Cereal Congress) International Conference*, Istanbul.

Dubcovsky, J., and Dvorak, J. (2007). Genome plasticity a key factor in the success of polyploid wheat under domestication. *Science* 316, 1862–1866. doi: 10.1126/science.1143986

El Haddad, N., Kabbaj, H., Zaïm, M., El Hassouni, K., Sall, A. T., Azouz, M., et al. (2020). Crop wild relatives in durum wheat breeding: drift or thrift? *Crop Sci.* 1–18. doi: 10.1002/csc2.20223

El Hassouni, K., Belkadi, B., Filali-Maltouf, A., Tidiane-Sall, A., Al-Abdallat, A., Nachit, M., et al. (2019). Loci controlling adaptation to heat stress occurring at the reproductive stage in durum wheat. *Agronomy* 9:414. doi: 10.3390/agronomy9080414

Elings, A., and Nachit, M. M. (1991). Durum wheat landraces from Syria. I. Agro-ecological and morphological characterization. *Euphytica* 53, 211–224. doi: 10.1007/BF00023273

Evenson, R. E., and Gollin, D. (2003). Assessing the impact of the green revolution, 1960 to 2000. *Science* 300, 758–762. doi: 10.1126/science.1078710

Excoffier, L., and Lischer, H. E. L. (2010). Arlequin suite ver 3.5: a new series of programs to perform population genetics analyses under Linux and Windows. *Mol. Ecol. Resour.* 10, 564–567. doi: 10.1111/j.1755-0998.2010.02847.x

Fernie, A. R., Tadmor, Y., and Zamir, D. (2006). Natural genetic variation for improving crop quality. *Curr. Opin. Plant Biol.* 9, 196–202. doi: 10.1016/j.pbi.2006.01.010

Fischer, G., Nachtergaele, F. O., Prieler, S., Teixeira, E., Toth, G., van Velthuizen, H., et al. (2012). *Global Agro-ecological Zones (GAEZ v3.0)- Model Documentation*. Laxenburg: IIASA.

Gadaleta, A., Nigro, D., Giancaspro, A., and Blanco, A. (2011). The glutamine synthetase (GS2) genes in relation to grain protein content of durum wheat. *Funct. Integr. Genomics* 11, 665–670. doi: 10.1007/s10142-011-0235-2

Giraldo, P., Royo, C., González, M., Carrillo, J. M., and Ruiz, M. (2016). Genetic diversity and association mapping for agro-morphological and grain quality traits of a structured collection of durum wheat landraces including subsp. *durum*, *turgidum* and *diccocon*. *PLoS One* 11:e0166577. doi: 10.1371/journal.pone.0166577

Giunta, F., Motzo, R., and Pruneddu, G. (2007). Trends since 1900 in the yield potential of Italian-bred durum wheat cultivars. *Eur. J. Agron.* 27, 12–24. doi: 10.1016/j.eja.2007.01.009

Gu, Y. Q., Crossman, C., Kong, X., Luo, M., You, F. M., Coleman-Derr, D., et al. (2004). Genomic organization of the complex alpha-gliadin gene loci in wheat. *Theor. Appl. Genet.* 109, 648–657. doi: 10.1007/s00122-004-1672-2

Gur, A., and Zamir, D. (2004). Unused natural variation can lift yield barriers in plant breeding. *PLoS Biol.* 2:e245. doi: 10.1371/journal.pbio.0020245

Hancock, D. B., Levy, J. L., Gaddis, N. C., Bierut, L. J., Saccone, N. L., Page, G. P., et al. (2012). Assessment of genotype imputation performance using 1000 genomes in african american studies. *PLoS One* 7:e50610. doi: 10.1371/journal.pone.0050610

Hoisington, D., Khairallah, M., Reeves, T., Ribaut, J.-M., Skovmand, B., Taba, S., et al. (1999). Plant genetic resources: what can they contribute toward increased crop productivity? *Proc. Natl. Acad. Sci. U.S.A.* 96, 5937–5943. doi: 10.1073/pnas.96.11.5937

Hou, J., Jiang, Q., Hao, C., Wang, Y., Zhang, H., and Zhang, X. (2014). global selection on sucrose synthase haplotypes during a century of wheat breeding. *Plant Physiol.* 164, 1918–1929. doi: 10.1104/pp.113.232454

Hu, M.-J., Zhang, H.-P., Liu, K., Cao, J.-J., Wang, S.-X., Jiang, H., et al. (2016). Cloning and characterization of TaTGW-7A gene associated with grain weight in wheat via SLAF-seq-BSA. *Front. Plant Sci* 7:1902. doi: 10.3389/fpls.2016.01902

Isidro, J., Álvaro, F., Royo, C., Villegas, D., Miralles, D. J., García, et al. (2011). Changes in duration of developmental phases of durum wheat caused by breeding in Spain and Italy during the 20th century and its impact on yield. *Ann. Bot.* 107, 1355–1366. doi: 10.1093/aob/mcr063

Jombart, T. (2008). Adegenet: a R package for the multivariate analysis of genetic markers. *Bioinformatics* 24, 1403–1405. doi: 10.1093/bioinformatics/btn129

Jordan, K. W., Wang, S., Lun, Y., Gardiner, L.-J., MacLachlan, R., Hucl, P., et al. (2015). A haplotype map of allohexaploid wheat reveals distinct patterns of selection on homoeologous genomes. *Genome Biol.* 16:48. doi: 10.1186/s13059-015-0606-4

Kabbaj, H., Sall, A. T., Al-Abdallat, A., Geleta, M., Amri, A., Filali-Maltouf, A., et al. (2017). Genetic Diversity within a global panel of durum wheat (*Triticum durum*) landraces and modern germplasm reveals the history of alleles exchange. *Front. Plant Sci.* 8:1277. doi: 10.3389/fpls.2017.01277

Kamvar, Z. N., Tabima, J. F., and Grünwald, N. J. (2014). Poppr: an R package for genetic analysis of populations with clonal, partially clonal, and/or sexual reproduction. *Peer J.* 2:e281. doi: 10.7717/peerj.281

Khush, G. S. (2001). Green revolution: the way forward. *Nat. Rev. Genet.* 2, 815–822. doi: 10.1038/35093585

Laidò, G., Mangini, G., Taranto, F., Gadaleta, A., Blanco, A., Cattivelli, L., et al. (2013). Genetic diversity and population structure of tetraploid wheats (*Triticum turgidum* L.) estimated by SSR, DArT and pedigree data. *PLoS One* 8:e67280. doi: 10.1371/journal.pone.0067280

Ledesma-Ramírez, L., Solís-Moya, E., Iturriaga, G., Sehgal, D., Reyes-Valdes, M. H., Montero-Tavera, V., et al. (2019). GWAS to identify genetic loci for resistance to yellow rust in wheat pre-breeding lines derived from diverse exotic crosses. *Front. Plant Sci.* 10:1390. doi: 10.3389/fpls.2019.01390

Maccaferri, M., El-Feki, W., Nazemi, G., Salvi, S., Canè, M. A., Colalongo, M. C., et al. (2016). Prioritizing quantitative trait loci for root system architecture in tetraploid wheat. *J. Exp. Bot.* 67, 1161–1178. doi: 10.1093/jxb/erw039

Maccaferri, M., Harris, N. S., Twardziok, S. O., Pasam, R. K., Gundlach, H., Spannagl, M., et al. (2019). Durum wheat genome highlights past domestication signatures and future improvement targets. *Nat. Genet.* 51, 885–895. doi: 10.1038/s41588-019-0381-3

Maccaferri, M., Ricci, A., Salvi, S., Milner, S. G., Noli, E., Martelli, P. L., et al. (2015). A high-density, SNP-based consensus map of tetraploid wheat as a bridge to integrate durum and bread wheat genomics and breeding. *Plant Biotechnol. J.* 13, 648–663. doi: 10.1111/pbi.12288

Maccaferri, M., Sanguineti, M. C., Corneti, S., Ortega, J. L., Salem, M. B., Bort, J., et al. (2008). Quantitative trait loci for grain yield and adaptation of durum wheat {(Triticum} durum Desf.) across a wide range of water availability. *Genetics* 178, 489–511. doi: 10.1534/genetics.107.077297

Maccaferri, M., Sanguineti, M. C., Demontis, A., El-Ahmed, A., Garcia del Moral, L., Maalouf, F., et al. (2011). Association mapping in durum wheat grown across a broad range of water regimes. *J. Exp. Bot.* 62, 409–438. doi: 10.1093/jxb/erq287

Maccaferri, M., Sanguineti, M. C., Donini, P., and Tuberosa, R. (2003). Microsatellite analysis reveals a progressive widening of the genetic basis in the elite durum wheat germplasm. *Theor. Appl. Genet.* 107, 783–797. doi: 10.1007/s00122-003-1319-8

Maccaferri, M., Sanguineti, M. C., Mantovani, P., Demontis, A., Massi, A., Ammar, K., et al. (2010). Association mapping of leaf rust response in durum wheat. *Mol. Breed.* 26, 189–228. doi: 10.1007/s11032-009-9353-0

Maccaferri, M., Sanguineti, M. C., Natoli, V., Ortega, J. L. A., Salem, M. B., Bort, J., et al. (2006). A panel of elite accessions of durum wheat (*Triticum durum* Desf.) suitable for association mapping studies. *Plant Genet. Resour.* 4, 79–85. doi: 10.1079/pgr2006117

Maccaferri, M., Sanguineti, M. C., Noli, E., and Tuberosa, R. (2005). Population structure and long-range linkage disequilibrium in a durum wheat elite collection. *Mol. Breed.* 15, 271–290. doi: 10.1007/s11032-004-7012-z

Mangini, G., Nigro, D., Margiotta, B., De Vita, P., Gadaleta, A., Simeone, R., et al. (2018). Exploring SNP diversity in wheat landraces germplasm and setting of a molecular barcode for fingerprinting. *Cereal Res. Commun.* 46, 377–387. doi: 10.1556/0806.46.2018.033

Marzario, S., Logozzo, G., David, J. L., Zeuli, P. S., and Gioia, T. (2018). Molecular genotyping (SSR) and agronomic phenotyping for utilization of durum wheat (*Triticum durum* Desf.) ex situ collection from Southern Italy: a combined approach including pedigreed varieties. *Genes* 9, 1–20. doi: 10.3390/genes9100465

Mondal, S., Dutta, S., Crespo-Herrera, L., Huerta-Espino, J., Braun, H. J., and Singh, R. P. (2020). Fifty years of semi-dwarf spring wheat breeding at CIMMYT: grain yield progress in optimum, drought and heat stress environments. *Field Crops Res.* 250:107757. doi: 10.1016/j.fcr.2020.107757

Moon, D. (2008). In the Russians' steppes: the introduction of Russian wheat on the great plains of the United States of America. *J. Glob. Hist.* 3, 203–225. doi: 10.1017/S1740022808002611

Motzo, R., and Giunta, F. (2007). The effect of breeding on the phenology of Italian durum wheats: from landraces to modern cultivars. *Eur. J. Agron.* 26, 462–470. doi: 10.1016/j.eja.2007.01.007

N'Diaye, A., Haile, J. K., Nilsen, K. T., Walkowiak, S., Ruan, Y., Singh, A. K., et al. (2018). Haplotype loci under selection in Canadian durum wheat germplasm over 60 years of breeding: association with grain yield, quality traits, protein loss, and plant height. *Front. Plant Sci.* 9:1589. doi: 10.3389/fpls.2018.01589

Nei, M. (1972). Genetic distance between populations. *Am. Nat.* 106, 283–292. doi: 10.1086/282771

Nei, M. (1973). Analysis of gene diversity in subdivided populations. *Proc. Natl. Acad. Sci. U.S.A.* 70, 3321–3323. doi: 10.1073/pnas.70.12.3321

Nei, M. (1978). Estimation of average heterozygosity and genetic distance from a small number of individuals. *Genetics* 89, 583–590.

Noriega, I. L., Halewood, M., Abberton, M., Amri, A., Angarawai, I. I., Anglin, N., et al. (2019). CGIAR operations under the plant treaty framework. *Crop Sci.* 59, 819–832. doi: 10.2135/cropsci2018.08.0526

Nothnagel, M., Ellinghaus, D., Schreiber, S., Krawczak, M., and Franke, A. (2009). A comprehensive evaluation of SNP genotype imputation. *Hum. Genet.* 125, 163–171. doi: 10.1007/s00439-008-0606-5

Oliveira, H. R., Campana, M. G., Jones, H., Hunt, H. V., Leigh, F., Redhouse, D. I., et al. (2012). Tetraploid wheat landraces in the Mediterranean basin: taxonomy, evolution and genetic diversity. *PLoS One* 7:e37063. doi: 10.1371/journal.pone.0037063

Ortiz, R., Trethowan, R., Ortiz Ferrara, G., Iwanaga, M., Dodds, J. H., Crouch, J. H., et al. (2007). High yield potential, shuttle breeding and a new international wheat improvement strategy. *Euphytica* 157, 365–384. doi: 10.1007/s10681-007-9375-9

Ozkan, H., Brandolini, A., Schafer-Pregl, R., and Salamini, F. (2002). AFLP analysis of a collection of tetraploid wheats indicates the origin of emmer and hard wheat domestication in southeast Turkey. *Mol. Biol. Evol.* 19, 1797–1801. doi: 10.1093/oxfordjournals.molbev.a004002

Pagnotta, M. A., Impiglia, A., Tanzarella, O. A., Nachit, M. M., and Porceddu, E. (2005). Genetic variation of the durum wheat landrace Haurani from different agro-ecological regions. *Genet. Resour. Crop Evol.* 51, 863–869. doi: 10.1007/s10722-005-0775-1

Paulsen, G. M., and Shroyer, J. P. (2008). The early history of wheat improvement in the Great Plains. *Agron. J.* 100, 70–78. doi: 10.2134/agronj2006.0355c

Pfeiffer, W. H., Sayre, K. D., Reynolds, M. P., and Payne, T. S. (2001). "Increasing yield potential and yield stability in durum wheat," in *Wheat in a Global Environment*, eds Z. Bedö and L. Láng (Dordrecht: Springer), 569–577. doi: 10.1007/978-94-017-3674-9_76

Pozniak, C. J., Knox, R. E., Clarke, F. R., and Clarke, J. M. (2007). Identification of QTL and association of a phytoene synthase gene with endosperm colour in durum wheat. *Theor. Appl. Genet.* 114, 525–537. doi: 10.1007/s00122-006-0453-5

Prat, N., Guilbert, C., Prah, U., Wachter, E., Steiner, B., Langin, T., et al. (2017). QTL mapping of Fusarium head blight resistance in three related durum wheat populations. *Theor. Appl. Genet.* 130, 13–27. doi: 10.1007/s00122-016-2785-0

Raman, H., Stodart, B. J., Cavanagh, C., Mackay, M., Morell, M., Milgate, A., et al. (2010). Molecular diversity and genetic structure of modern and traditional landrace cultivars of wheat (*Triticum aestivum* L.). *Crop Pasture Sci.* 61:222. doi: 10.1071/CP09093

R Core Team (2016). *R: A Language and Environment for Statistical Computing*. Vienna: R Foundation for Statistical Computing. Available online at: https://www.R-project.org/

Reimer, S., Pozniak, C. J., Clarke, F. R., Clarke, J. M., Somers, D. J., Knox, R. E., et al. (2008). Association mapping of yellow pigment in an elite collection of durum wheat cultivars and breeding lines. *Genome* 51, 1016–1025. doi: 10.1139/G08-083

Ren, Y., He, X., Liu, D., Li, J., Zhao, X., Li, B., et al. (2012). Major quantitative trait loci for seminal root morphology of wheat seedlings. *Mol. Breed.* 30, 139–148. doi: 10.1007/s11032-011-9605-7

Rexroad, C. E., and Vallejo, R. L. (2009). Estimates of linkage disequilibrium and effective population size in rainbow trout. *BMC Genet.* 10:83. doi: 10.1186/1471-2156-10-83

Robbana, C., Kehel, Z., Ben Naceur, M., Sansaloni, C., Bassi, F., and Amri, A. (2019). Genome-Wide genetic diversity and population structure of tunisian durum wheat landraces based on DArTseq technology. *Int. J. Mol. Sci.* 20, 1352. doi: 10.3390/ijms20061352

Roncallo, P. F., Beaufort, V., Larsen, A. O., Dreisigacker, S., and Echenique, V. (2019). Genetic diversity and linkage disequilibrium using SNP (KASP) and AFLP markers in a worldwide durum wheat (*Triticum turgidum* L. var *durum*) collection. *PLoS One* 14:e0218562. doi: 10.1371/journal.pone.0218562

Roncallo, P. F., Cervigni, G. L., Jensen, C., Miranda, R., Carrera, A. D., Helguera, M., et al. (2012). QTL analysis of main and epistatic effects for flour color traits in durum wheat. *Euphytica* 185, 77–92. doi: 10.1007/s10681-012-0628-x

Royo, C., Álvaro, F., Martos, V., Ramdani, A., Isidro, J., Villegas, D., et al. (2007). Genetic changes in durum wheat yield components and associated traits in Italian and Spanish varieties during the 20th century. *Euphytica* 155, 259–270. doi: 10.1007/s10681-006-9327-9

Royo, C., Elias, E. M., and Manthey, F. A. (2009). "Durum wheat breeding," in *Cereals*, ed. M. J. Carena, 199–226.

Royo, C., Maccaferri, M., Álvaro, F., Moragues, M., Sanguineti, M. C., Tuberosa, R., et al. (2010). Understanding the relationships between genetic and phenotypic structures of a collection of elite durum wheat accessions. *Field Crop Res.* 119, 91–105. doi: 10.1016/j.fcr.2010.06.020

Royo, C., Martos, V., Ramdani, A., Villegas, D., Rharrabti, Y., and García del Moral, L. F. (2008). Changes in yield and carbon isotope discrimination of Italian and Spanish durum wheat during the 20th Century. *Agron. J.* 100, 352–360. doi: 10.2134/agronj2007.0060

Saccomanno, A., Matny, O., Marone, D., Laidò, G., Petruzzino, G., Mazzucotelli, E., et al. (2018). Genetic mapping of loci for resistance to stem rust in a tetraploid wheat collection. *Int. J. Mol. Sci.* 19:3907. doi: 10.3390/ijms19123907

Sakuma, S., Golan, G., Guo, Z., Ogawa, T., Tagiri, A., Sugimoto, K., et al. (2019). Unleashing floret fertility in wheat through the mutation of a homeobox gene. *Proc. Natl. Acad. Sci. U.S.A.* 116, 5182–5187. doi: 10.1073/pnas.1815465116

Sall, A., Chiari, T., Legesse, W., Seid-Ahmed, K., Ortiz, R., van Ginkel, M., et al. (2019). Durum wheat (*Triticum durum* Desf.): origin, cultivation and potential expansion in Sub-Saharan Africa. *Agronomy* 9:263. doi: 10.3390/agronomy9050263

Scarascia Mugnozza, G. T. (2005). *The Contribution of Italian Wheat Geneticists: from Nazareno Strampelli to Francesco D'Amato*. Rome: Accademia Nazionale delle Scienze, 53–75.

Shannon, C. E. (1948). A mathematical theory of communication. *Bell Syst. Tech. J.* 27, 623–656. doi: 10.1002/j.1538-7305.1948.tb00917.x

Singh, S., Vikram, P., Sehgal, D., Burgueño, J., Sharma, A., Singh, S. K., et al. (2018). Harnessing genetic potential of wheat germplasm banks through impact-oriented-prebreeding for future food and nutritional security. *Sci. Rep.* 8:12527. doi: 10.1038/s41598-018-30667-4

Slafer, G. A., Satorre, E. H., and Andrade, F. H. (1994). "Increases in grain yield in bread wheat from breeding and associated physiological changes," in *Genetic Improvement of Field Crops: Current Status and Development*, ed. G. A. Slafer (New York, NY: Marcel Dekker, Inc.), 1–68.

Slim, A., Piarulli, L., Kourda, H. C., Rouaissi, M., Robbana, C., Chaabane, R., et al. (2019). Genetic structure analysis of a collection of Tunisian durum wheat germplasm. *Int. J. Mol. Sci.* 20:3362. doi: 10.3390/ijms20133362

Smith, B., and Wilson, J. B. (1996). A consumer's guide to evenness indices. *Oikos* 76, 70. doi: 10.2307/3545749

Soriano, J. M., Villegas, D., Aranzana, M. J., García Del Moral, L. F., and Royo, C. (2016). Genetic structure of modern durum wheat cultivars and Mediterranean landraces matches with their agronomic performance. *PLoS One* 11:e0160983. doi: 10.1371/journal.pone.0160983

Soriano, J. M., Villegas, D., Sorrells, M. E., and Royo, C. (2018). Durum wheat landraces from east and west regions of the Mediterranean basin are genetically distinct for yield components and phenology. *Front. Plant Sci.* 9:80. doi: 10.3389/fpls.2018.00080

Talebi, R., and Fayaz, F. (2016). Geographical diversity pattern in Iranian landrace durum wheat (*Triticum turgidum*) accessions using start codon targeted polymorphism and conserved DNA-derived polymorphism markers. *Environ. Exp. Biol.* 14, 63–68. doi: 10.22364/eeb.14.09

Tanksley, S. D., and McCouch, S. R. (1997). Seed banks and molecular maps: unlocking genetic potential from the wild. *Science* 277, 1063–1066. doi: 10.1126/science.277.5329.1063

Taranto, F., D'Agostino, N., Rodriguez, M., Pavan, S., Minervini, A. P., Pecchioni, N., et al. (2020). Whole genome scan reveals molecular signatures of divergence and selection related to important traits in durum wheat germplasm. *Front. Genet.* 11:217. doi: 10.3389/fgene.2020.00217

Tuberosa, R., and Pozniak, C. (2014). Durum wheat genomics comes of age. *Mol. Breed.* 34, 1527–1530. doi: 10.1007/s11032-014-0188-y

van Ginkel, M., and Ortiz, R. (2018). Cross the best with the best, and select the best: HELP in breeding selfing crops. *Crop Sci.* 58, 17–30. doi: 10.2135/cropsci2017.05.0270

Wang, S., Wong, D., Forrest, K., Allen, A., Chao, S., Huang, B. E., et al. (2014). Characterization of polyploid wheat genomic diversity using a high-density 90˜000 single nucleotide polymorphism array. *Plant Biotechnol. J.* 12, 787–796. doi: 10.1111/pbi.12183

Wilhelm, E. P., Turner, A. S., and Laurie, D. A. (2009). Photoperiod insensitive Ppd-A1a mutations in tetraploid wheat (*Triticum durum* Desf.). *Theor. Appl. Genet.* 118, 285–294. doi: 10.1007/s00122-008-0898-9

Wright, S. (1965). The interpretation of population structure by F-statistics with special regard to systems of mating. *Evolution* 19, 395–420. doi: 10.2307/2406450

Xu, Q., Xu, J., Liu, C. L., Chang, C., Wang, C. P., You, M. S., et al. (2008). PCR-based markers for identification of HMW-GS at Glu-B1x loci in common wheat. *J. Cereal Sci.* 47, 394–398. doi: 10.1016/j.jcs.2007.05.002

Yan, L., Fu, D., Li, C., Blechl, A., Tranquilli, G., Bonafede, M., et al. (2006). The wheat and barley vernalization gene VRN3 is an orthologue of FT. *Proc. Natl. Acad. Sci. U.S.A.* 103, 19581–19586. doi: 10.1073/pnas.0607142103

Yan, L., Loukoianov, A., Tranquilli, G., Helguera, M., Fahima, T., and Dubcovsky, J. (2003). Positional cloning of the wheat vernalization gene VRN1. *Proc. Natl. Acad. Sci. U.S.A.* 100, 6263–6268. doi: 10.1073/pnas.0937399100

Zaïm, M., El Hassouni, K., Gamba, F., Filali-Maltouf, A., Belkadi, B., Sourour, A., et al. (2017). Wide crosses of durum wheat (*Triticum durum* Desf.) reveal good disease resistance, yield stability, and industrial quality across Mediterranean sites. *F. Crop. Res.* 214, 219–227. doi: 10.1016/j.fcr.2017.09.007

Zhang, L., Zhao, Y. L., Gao, L. F., Zhao, G. Y., Zhou, R. H., Zhang, B. S., et al. (2012). TaCKX6-D1, the ortholog of rice OsCKX2, is associated with grain weight in hexaploid wheat. *New Phytol.* 195, 574–584. doi: 10.1111/j.1469-8137.2012.04194.x

Zikhali, M., Wingen, L. U., and Griffiths, S. (2016). Delimitation of the Earliness per se D1 (Eps-D1) flowering gene to a subtelomeric chromosomal deletion in bread wheat (*Triticum aestivum*). *J. Exp. Bot.* 67, 287–299. doi: 10.1093/jxb/erv458

Conflict of Interest: The authors declare that the research was conducted in the absence of any commercial or financial relationships that could be construed as a potential conflict of interest.

The reviewer MR declared a past co-authorship with several of the authors DM and FL to the handling editor.

Advantages of publishing in Frontiers

OPEN ACCESS
Articles are free to read for greatest visibility and readership

FAST PUBLICATION
Around 90 days from submission to decision

HIGH QUALITY PEER-REVIEW
Rigorous, collaborative, and constructive peer-review

TRANSPARENT PEER-REVIEW
Editors and reviewers acknowledged by name on published articles

Frontiers
Avenue du Tribunal-Fédéral 34
1005 Lausanne | Switzerland

Visit us: www.frontiersin.org
Contact us: frontiersin.org/about/contact

REPRODUCIBILITY OF RESEARCH
Support open data and methods to enhance research reproducibility

DIGITAL PUBLISHING
Articles designed for optimal readership across devices

FOLLOW US
@frontiersin

IMPACT METRICS
Advanced article metrics track visibility across digital media

EXTENSIVE PROMOTION
Marketing and promotion of impactful research

LOOP RESEARCH NETWORK
Our network increases your article's readership