DEVELOPMENT OF HEALTHY AND NUTRITIOUS CEREALS: RECENT INSIGHTS ON MOLECULAR ADVANCES IN BREEDING

EDITED BY: Mallikarjuna Swamy, Balram Marathi, Ana I. F. Ribeiro-Barros and Felipe Klein Ricachenevsky

PUBLISHED IN: Frontiers in Genetics

frontiers Research Topics

frontiers

ISSN 1664-8714
ISBN 978-2-88966-755-0
DOI 10.3389/978-2-88966-755-0

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

DEVELOPMENT OF HEALTHY AND NUTRITIOUS CEREALS: RECENT INSIGHTS ON MOLECULAR ADVANCES IN BREEDING

Topic Editors:
Mallikarjuna Swamy, International Rice Research Institute (IRRI), Philippines
Balram Marathi, Professor Jayashankar Telangana State Agricultural University, India
Ana I. F. Ribeiro-Barros, University of Lisbon, Portugal
Felipe Klein Ricachenevsky, Federal University of Rio Grande do Sul, Brazil

Citation: Swamy, M., Marathi, B., Ribeiro-Barros, A. I. F., Ricachenevsky, F. K., eds. (2021). Development of Healthy and Nutritious Cereals: Recent Insights on Molecular Advances in Breeding. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-88966-755-0

Table of Contents

frontiers in Genetics

EDITORIAL
published: 18 March 2021
doi: 10.3389/fgene.2021.635006

Editorial: Development of Healthy and Nutritious Cereals: Recent Insights on Molecular Advances in Breeding

B. P. Mallikarjuna Swamy[1]*, *Balram Marathi*[2]*, *Ana I. F. Ribeiro-Barros*[3]* and *Felipe Klein Ricachenevsky*[4,5]*

[1] International Rice Research Institute, Los Baños, Philippines, [2] Department of Genetics and Plant Breeding, Professor Jayashankar Telangana State Agricultural University, Hyderabad, India, [3] Forest Research Centre (CEF), Instituto Superior de Agronomia, Universidade de Lisboa, Lisbon, Portugal, [4] Departamento de Botânica, Instituto de Biociências, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil, [5] Programa de Pós-Graduação em Biologia Celular e Molecular, Centro de Biotecnologia, Universidade Federal do Rio Grande do Sul, Porto Alegre, Brazil

Keywords: cereals, nutrition, breeding, omics, genome editing, biofortification

Editorial on the Research Topic

Development of Healthy and Nutritious Cereals: Recent Insights on Molecular Advances in Breeding

Worldwide more than 2 billion people are affected by micronutrient deficiencies and most of them are residing in the developing countries of Asia, Africa and Latin America (Kennedy et al., 2002). Malnutrition is linked with heavy dependence on monotonous cereal staples without much dietary diversification or nutrient supplementation. Even though significant efforts have been made over the last six decades to improve production and productivity in most food crops, it lacked associated nutritional improvement (Bouis and Welch, 2010). So, the modern varieties do not have enough variability for several nutrients, making poor rural populations vulnerable to micronutrient deficiencies. More than two dozen mineral elements, vitamins, antioxidants, and health beneficial compounds must be supplied in optimal quantities daily for normal growth and development of humans. Biofortification of cereals with elevated levels of essential micronutrients, vitamins, and reduced levels of toxic elements help to address malnutrition and is a cost-effective approach in reaching target groups, especially rural populations (Bouis and Saltman, 2017). The sustainable development goals and the Lancet Commission Report have emphasized the need for promoting nutritious diets to eradicate malnutrition (Willet et al., 2019; https://sustainabledevelopment.un.org). Among these, deficiencies of iron (Fe), zinc (Zn), and vitamin A are major global health problems. As successful examples, one high Fe rice and several high Zn rice varieties have been successfully released for commercial cultivation (Palanog et al., 2019).

Presently we have a better understanding of the genetic, physiological, and molecular basis, as well as the influence of environmental factors on nutrients accumulation in cereal grains (Swamy et al., 2016; Garcia-Oliveira et al., 2018; Ludwig and Slamet-Loedin, 2019). However, there is a need to integrate our understanding to achieve the goals of biofortification and review the current progress and the prospects for nutritious crops. In this Research Topic, we selected manuscripts on various aspects of nutritional improvement in cereals. Fourteen articles published in our special editorial topic, five of them provided updated review of cereals nutritional enhancement, and nine of them were original research articles on understanding the molecular basis of different grain nutrients and grain quality traits in cereals.

OPEN ACCESS

Edited and reviewed by:
Ahmed El-Sohemy,
University of Toronto, Canada

***Correspondence:**
B. P. Mallikarjuna Swamy
m.swamy@irri.org
Balram Marathi
balumarathi@gmail.com
Ana I. F. Ribeiro-Barros
aribeiro@isa.ulisboa.pt
Felipe Klein Ricachenevsky
felipecruzalta@gmail.com

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *29 November 2020*
Accepted: *23 February 2021*
Published: *18 March 2021*

Citation:
Swamy BPM, Marathi B, Ribeiro-Barros AIF and Ricachenevsky FK (2021) Editorial: Development of Healthy and Nutritious Cereals: Recent Insights on Molecular Advances in Breeding. Front. Genet. 12:635006. doi: 10.3389/fgene.2021.635006

Focusing on improving grain protein quality, Chandran et al. successfully pyramided Lysine, Tryptophan, and Provitamin A into Maize varieties based on *opaque-2 and* β-*carotene* through marker assisted selection (MAS). The improved lines possessed high lysine, tryptophan, and β-carotene content, but they had only slight yield reduction. Even though these lines can be used as genetic resources for maize improvement, they are not yet commercially viable, since successful biofortified crops should have similar or even higher yield along with the other desirable traits. Also aiming at improving protein content in rice, Jang et al. identified multiple genomic regions responsible for amino acid content (AAC) and protein content (PC). They identified two novel *loci* qAAC6.1 and qAAC7.1 and several transgressive segregants for both traits. These loci can be used for quantitative trait loci (QTL) pyramiding programs to develop rice lines with high protein content.

It is quite interesting to note pleotropic effect of heading-date genes on protein content of rice. Xie et al. reported that in three nearly isogenic lines (NIL), the rice florigen genes RFT1 have a strong negative effect on the amino acids content governed by the *Zhenshan97* allele with the genomic region consisting of 14 QTLs located in proximity to Hd3a. Bhuvaneswari et al. characterized 93 aromatic Chakhao rice germplasm from Manipur province of India. Wider variations were observed for the agro-morphological, grain quality and nutraceutical traits. The total anthocyanin content ranged from 29.8 to 275.8 mg.100g^{-1} DW, while total phenolics ranged from 66.5 to 700.3 mg GAE.100g^{-1} DW. The germplasm with higher levels of anthocyanin compounds such as cyanidin-3-O-glucoside (C3G) and peonidin-3-O-glucoside (P3G) are useful for improving the antioxidant properties in rice.

Focusing on micronutrient biofortification, Ashokkumar et al. comprehensively reviewed recent advances in breeding for improved folate, provitamin A, and carotenoids content in rice, wheat, maize, and pearl millet. They discussed in detail the genetic variation, trait discovery, genes/QTL identification for nutritional traits and their introgressions into elite genetic backgrounds. Prasanna et al. carried out a detailed global analysis of molecular breeding for nutritional improvement in maize, a species where systematic efforts have been made to develop and deploy cultivars biofortified with quality protein maize (QPM), provitamin A, and kernel zinc. The limited germplasm characterization, lack of genetic variability, and diagnostic markers for some of the mineral elements is a constraint for breeding. Broadening the genetic base through exploitation of landraces and wild species, use of genomics technologies, market-driven breeding strategies, strengthening of seed systems, and collaborative interdisciplinary efforts were emphasized. Genetic Engineering (GE) and Genome Editing (GEd) technologies are the way forward for improving the traits with no variability and to achieve the target levels of multiple nutrients in cereals.

Babu et al. characterized 40 rice genotypes for agronomic, yield and micronutrient traits. They identified stable high Zn donor lines and genome wide association analysis resulted in identification three *loci* on chromosomes 3 and 7, which were linked to new, uncharacterized putative candidate genes. Bollinedi et al. identified 18 novel marker-trait associations (MTAs) for grain Fe and Zn in brown and milled rice using 192 Indian rice germplasm accessions and found strong association between Zn concentrations in brown and milled rice. Fe concentration in brown rice, however, was not associated with Fe concentration in milled rice, highlighting the need for enriching the Fe concentration of rice endosperm. They have also identified four accessions with grain Zn concentration in milled rice with >28 mg/kg and one accession (IC-2127) with >12 mg/kg Fe, a target set by the HarvestPlus program for rice biofortification which will be used to develop high Fe and Zn varieties.

Focusing on biofortification of not-so-typical grains, Bekkering and Tian provide an overview of how other cereals and non-grass pseudo-cereals can contribute to biofortification. The monocots species, broomcorn millet (*Panicum miliaceum* L.), canary seed (*Phalaris canariensis* L.), teff [*Eragrostis tef* (Zuccagni) Trotter], and the pseudo-cereals (i.e., seeds similar to cereals, but not from the Poaceae family), amaranth (*Amaranthus* spp.), buckwheat (*Fagopyrum esculentum* Moench.), chia (*Salvia hispanica* L.), and quinoa (*Chenopodium quinoa* Willd), are richer in protein and lipids, and have lower starch compared to the major staples. Other characteristics include a more balanced protein composition, higher levels of micronutrients and vitamins. Authors delineate possible avenues in which we should invest to improve and fully utilize these species to provide nutritious grains for consumption, from marketing to breeding to genomics/functional genetics.

In the same line, Rodríguez et al., provided an extensive review on finger millet (*Eleusine coracana*) and foxtail millet (*Setaria italica*) and the pseudocereals quinoa, amaranth and buckwheat. These species can contribute to produce healthy grains, especially in harsh, stressful environments, for which these plants tend to be more resilient. Authors compare nutritional profiles, processes to increase bioavailability of nutrients and decrease anti-nutrients, and thoroughly review the current status of genome resources and molecular markers available to drive the efforts to improve these species, allowing their use in marginal lands to produce nutritious food to combat hidden hunger. Renganathan et al. focus on barnyard millet from the genus *Echinochloa*, including the cultivated Indian barnyard millet (*Echinochloa frumentacea*), Japanese barnyard millet (*Echinochloa esculenta*) and other wild species, which are used for human consumption and livestock feed, and also show tolerance to multiple stresses. The taxonomy, morphological and genetic diversity, genomic and genetic resources, and the potential of these species to contribute for food security and human nutrition are discussed. Altogether, these reviews are an excellent starting point for the biofortification community to launch efforts into including new species in our toolbox to increase human access to sufficient nutrients.

From a genomic perspective, Butardo et al. addressed the structural, regulatory and nutrition roles of Starch Synthase IIa in *O. sativa* ssp. *japonica* rice endosperm, highlighting the importance of this key enzyme in seed morphology, starch granules size and distribution, amylopectin structure, amylose content and glycemic index. Using a different

approach, Kishor et al. used whole-genome next generation sequencing to characterize traditional varieties of basmati rice and identified millions of SNP markers useful for genetic analysis. Finally, Bhuvaneswari et al., combined genomics and metabolomics to bring forward the nutraceutical importance of aromatic glutinous rice through the characterization of a set of 93 landraces for their agro-morphological traits, grain pigmentation, antioxidant properties, and molecular genetic variation. Altogether these three papers provide a solid platform of molecular markers for taxonomy, breeding and conservation programs.

Our Research Topic combines different approaches on biofortification and provides a comprehensive collection of the efforts to improve grain nutrient quality and to increase human nutrition.

AUTHOR CONTRIBUTIONS

All the authors equally contributed to prepare this editorial manuscript.

ACKNOWLEDGMENTS

The authors would like to thank Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) and Fundação de Amparo à Pesquisa do Rio Grande do Sul (FAPERGS) in Brazil; Fundação para a Ciência e a Tecnologia from Portugal, through the research unit UIDB/00239/2020 (CEF). Bill and Melinda Gates Foundation (BMGF) for funding Healthier Rice Project at International Rice Research Institute (IRRI).

REFERENCES

Bouis, H. E., and Saltman, A. (2017). Improving nutrition through biofortification: a review of evidence from HarvestPlus, 2003 through 2016. *Global Food Security* 12, 49–58. doi: 10.1016/j.gfs.2017.01.009

Bouis, H. E., and Welch, R. M. (2010). Biofortification—a sustainable agricultural strategy for reducing micronutrient malnutrition in the global South. *Crop Sci.* 50, 20–32. doi: 10.2135/cropsci2009.09.0531

Garcia-Oliveira, A. L., Subhash, C., Rodomiro, O., Abebe, M., and Melaku, G. (2018). Genetic basis and breeding perspectives of grain iron and zinc enrichment in cereals. *Front. Plant Sci.* 9:937. doi: 10.3389/fpls.2018.00937

Kennedy, G., Burlingame, B., and Nguyen, V. N. (2002). "Nutritional contribution of rice and impact of biotechnology and biodiversity in rice-consuming countries," *20th International Rice Commission Bangkok Thailand 23–26 July 2002* (Bangkok).

Ludwig, Y., and Slamet-Loedin, I. H. (2019). Genetic biofortification to enrich rice and wheat grain iron: from genes to product. *Front. Plant Sci.* 10:833. doi: 10.3389/fpls.2019.00833

Palanog, A. D., Calayugan, M. I. C., Descalsota-Empleo, G. I., Amparado, A., Inabangan-Asilo, M. A., Arocena, E. C., et al. (2019). Zinc and iron nutrition status in the Philippines population and local soils: a review. *Front. Nutr.* 6:81. doi: 10.3389/fnut.2019.00081

Swamy, B. P. M., Rahman, M. A., Inabangan-Asilo, M. A., Amparado, A., Manito, C., Chadha-Mohanty, P., et al. (2016). Advances in breeding for high Zinc in Rice. *Rice* 9, 49–65. doi: 10.1186/s12284-016-0122-5

Willet, W., Rockstrom, J., Loken, B., Springman, M., Lang, T., Vermeulen, S., et al. (2019). Food in the Aantropocene: the EAT-Lancet commission on healthy diets from sustainable food systems. *Lancet* 393, 447–492. doi: 10.1016/S0140-6736(18)31788-4

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 Genetics

ORIGINAL RESEARCH
published: 26 September 2019
doi: 10.3389/fgene.2019.00859

Marker-Assisted Selection to Pyramid the *Opaque-2 (O2)* and β-Carotene *(crtRB1)* Genes in Maize

*Sarankumar Chandran[1], Bharathi Pukalenthy[1], Karthikeyan Adhimoolam[2], Dhasarathan Manickam[2], Vellaikumar Sampathrajan[2], Vanniarajan Chocklingam[1], Kokiladevi Eswaran[3], Kavithapushpam Arunachalam[4], Laishram Joikumar meetei[5], Ravikesavan Rajasekaran[6], Vignesh Muthusamy[7], Firoz Hossain[7] and Senthil Natesan[8]**

[1] Department of Plant Breeding and Genetics, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai, India, [2] Department of Biotechnology, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai, India, [3] Department of Plant Biotechnology, Center for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India, [4] Department of Soil Science and Agricultural Chemistry, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Killikulam, India, [5] Department of Tree Improvement and Plant Breeding and Genetics, Central Agricultural University, Imphal, India, [6] Department of Millet, Center for Plant Breeding and Genetics, Tamil Nadu Agricultural University, Coimbatore, India, [7] Division of Genetics, ICAR—Indian Agricultural Research Institute, New Delhi, India, [8] Department of Plant Molecular Biology and Bioinformatics, Center for Plant Molecular Biology and Biotechnology, Tamil Nadu Agricultural University, Coimbatore, India

OPEN ACCESS

Edited by:
Mallikarjuna Swamy, International Rice Research Institute, Philippines

Reviewed by:
Melaku Gedil, International Institute of Tropical Agriculture, Nigeria
Manje Gowda, The International Maize and Wheat Improvement Center (CIMMYT), Kenya

***Correspondence:**
Senthil Natesan
senthil_natesan@tnau.ac.in

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *02 June 2019*
Accepted: *16 August 2019*
Published: *26 September 2019*

Citation:
Chandran S, Pukalenthy B, Adhimoolam K, Manickam D, Sampathrajan V, Chocklingam V, Eswaran K, Arunachalam K, Joikumar meetei L, Rajasekaran R, Muthusamy V, Hossain F and Natesan S (2019) Marker-Assisted Selection to Pyramid the Opaque-2 (O2) and β-Carotene (crtRB1) Genes in Maize. Front. Genet. 10:859. doi: 10.3389/fgene.2019.00859

Maize is an excellent nutritional source and is consumed as a staple food in different parts of the world, including India. Developing a maize genotype with a combination of higher lysine and tryptophan, along with β-carotene, can help alleviate the problem of protein-energy malnutrition (PEM) and vitamin A deficiency (VAD). This study is aimed at improving lysine and tryptophan content by transferring *opaque-2* (*o2*) gene from donor HKI163 to β-*carotene*-rich inbred lines *viz.,* UMI1200β$^+$ and UMI1230β$^+$. For this purpose, F_1, BC_1F_1, BC_2F_1, BC_2F_2, and BC_2F_3 plants were developed using an o2 line HKI163 and two β-carotene-rich inbred lines, UMI1200β$^+$ and UMI1230β$^+$, as the parents. Foreground selection using the associated marker umc1066 for the *o2* gene and the marker *crtRB1* 3′TE for the *crtRB1* gene was used to select the target genes. A total of 236 simple sequence repeat (SSR) markers distributed evenly across the maize genome were employed for the background selection. To fix the *crtRB1* allele in the BC_1F_1 stage, individual plants homozygous at the *crtRB1* locus and heterozygous at the *o2* locus were selected and used for backcrossing to produce BC_2F_1 plants. Furthermore, the selected heterozygous BC_2F_1 plants from both crosses were selfed to obtain the BC_2F_2 plants, which were then selected for the target gene and selfed to generate the BC_2F_3 lines. From each cross, five improved lines with homozygous marker alleles for the *crtRB1* and *o2* genes with a recurrent parent genome (RPG) recovery ranging from 86.75 to 91.21% in UMI1200β$^+$×HKI163 and 80.00 to 90.08% in UMI1230β$^+$×HKI163 were identified. The improved lines had good agronomic performance and possessed high lysine (0.294–0.332%), tryptophan (0.073–0.081%), and β-*carotene* (6.12–7.38 μg/g) content. These improved lines can be used as genetic resources for maize improvement.

Keywords: β-carotene, *crtRB1*, marker-assisted backcross breeding, *opaque-2*, quality protein maize

INTRODUCTION

Maize (*Zea mays* L.) is a staple food crop and currently grown in more than 150 countries, with a total harvest area of approximately 187 million hectares, producing 1138 million tonnes worldwide (FAOSTAT, 2018). It has good nutritional value, that is, 68.5% carbohydrates, 8% fat, 4% ash, 3% crude fiber, and 16.5% protein (Ullah et al., 2010). In addition, maize carotenoids contain both provitamin A (α-carotene, β-carotene, and β-cryptoxanthin) and non-provitamin A (lutein and zeaxanthin) components. Maize, therefore, is of special importance for the nutrition of people from many countries in Africa, Asia, and Latin America, where protein-energy malnutrition (PEM) and vitamin A deficiency (VAD) affect more than a billion people. The demand for maize has steadily increased over the past decades and is expected to continue to rise in the forthcoming years, at least up until 2050 (Rosegrant et al., 2009). However, normal maize protein possesses low nutritional significance to humans because of very limited amounts of major amino acids, such as lysine (1.6–2.6%) and tryptophan (0.2–0.6%) (Moro et al., 1996), which is less than half of the recommended dose specified for human nutrition. Over the past three decades, many natural maize mutants associated to quality protein maize (QPM) with higher lysine and tryptophan content have been identified, that is, *opaque-2 (o2)* in chromosome 7 (Mertz et al., 1964), *floury-2 (fl2)* in chromosome 8 (Nelson et al., 1965), *opaque-7 (o7)* in chromosome 8 (Ma and Nelson, 1975), *opaque-6 (o6)* in chromosome 8 (McWhirter, 1971), and *floury-3 (fl3)* in chromosome 8 (Ma and Nelson, 1975). Among them, *o2* mutant has been more popular and widely utilized in breeding programs for the improvement of protein quality. The recessive *o2* allele improves the endosperm lysine and tryptophan levels by nearly two-fold. The gene-linked simple sequence repeat (SSR) markers umc1066, phi112, and phi057 have been used to identify the *o2* gene (Yang et al., 2005; Gupta et al., 2013; Surender et al., 2017).

VAD is one of the serious health issues in developing and low-income countries and critically affects over 7 million pregnant women and 125 million children (Giuliano, 2017). *β-Carotene* is the best provitamin A (vitamin A precursor), and maize is a predominant source of *β-carotene*; however, very few maize varieties are rich in β-carotene, and many exhibiting varieties are inherently deficient in β-carotene (Muthusamy et al., 2014). Yan et al. (2010) revealed that *crtRB1* is a major gene responsible for the β-carotene content in maize. This gene is positioned at chromosome 10 and encodes β-carotene hydroxylase, which is responsible for the biosynthesis of lycopene. Association mapping approach led to the identification of three polymorphisms, 5'TE (in the 5'-Untranslated Region), InDel4 (in the coding region), and 3'TE (spanning the sixth exon and 3'-Untranslated Region), in the *crtRB1* gene that were significantly influencing the β-carotene content. Since then, polymerase chain reaction (PCR)-based codominant markers were developed based on these polymorphisms, and these markers aided breeders to identify and develop higher β-carotene content lines using marker-assisted selection (MAS). Moreover, Yan et al. (2010) reported the 3'TE favorable allele (allele 1, 543 bp) that is responsible for reduced transcript expression of the gene associated with higher β-carotene content, with an average increase of 6.50 μg/g in the maize endosperm in comparison with the unfavorable allelic class. Recently, this allele-based marker was successfully used to detect the *crtRB1* gene in diverse maize lines (Muthusamy et al., 2014; Zunjare et al., 2018; Sagare et al., 2019).

To date, numerous maize hybrids with either provitamin A or QPM have been released and commercialized, but genotypes with both the nutritional traits are very limited. This situation necessitates developing maize genotypes with the combination of QPM and provitamin A. Our previous attempts have led to the development of two β-carotene-rich inbred lines *viz.,* UMI1200β$^+$ and UMI1230β$^+$. In this study, our objective was aimed to introgress the *o2* gene from HKI163 into UMI1200β$^+$ and UMI1230β$^+$. We, therefore, applied marker-assisted backcross (MAB) breeding using gene-specific markers for foreground selection and polymorphic SSRs for background selection. Our goal was to obtain innovative breeding materials with high β-carotene, lysine, and tryptophan contents.

MATERIALS AND METHODS

Plant Genetic Materials

HKI163 is an inbred line containing the opaqueness gene *(o2)*. Its grain lysine content is 0.340% in protein, and its tryptophan content is 0.082% in protein (Zunjare et al., 2018). It was obtained from Chaudhary Charan Singh Haryana Agricultural University, Uchani, India. UMI1200β$^+$ and UMI1230β$^+$ are improved inbred lines containing the β-carotene-associated gene *crtRB1*, with a grain lysine content of 0.130 and 0.150%, respectively, and tryptophan content of 0.024 and 0.029%, respectively. These β-carotene-rich inbred lines were developed by transferring *crtRB1* gene from donor HP46715 (CIMMYT, Mexico) to local popular inbred lines *viz.,* UMI1200 and UMI1230. The β-carotene contents of UMI1200β$^+$ and UMI1230β$^+$ were 9.073 and 9.232 μg/g, respectively.

Development of Backcross Progenies

MAB breeding scheme that includes crossing, backcrossing, and selfing was undertaken as mentioned in **Figure 1**. Backcross progenies were developed by crossing UMI1200β$^+$ and UMI1230β$^+$ (recurrent parents) with HKI163 (donor parent) following two cycles of backcrosses during 2016 to 2019. UMI1200β$^+$ and UMI1230β$^+$ were used as recurrent parents and crossed with HKI163 (donor) for developing F_1 plants. Then, F_1 plants were confirmed by foreground selection with *crtRB1* and *o2*-linked markers. These F_1 plants were used as the male parents to develop the BC_1F_1s. Likewise, another round of backcross was followed for UMI1200β$^+$×HKI163 and UMI1230β$^+$×HKI163 to develop BC_2F_1s using MAB breeding to reduce the linkage drag and to increase the recurrent parent genome percentage. Furthermore, selected BC_2F_1 plants that were heterozygous at the *o2* loci and homozygous at the *crtRB1* loci were self-pollinated to produce BC_2F_2 plants and BC_2F_3 plants.

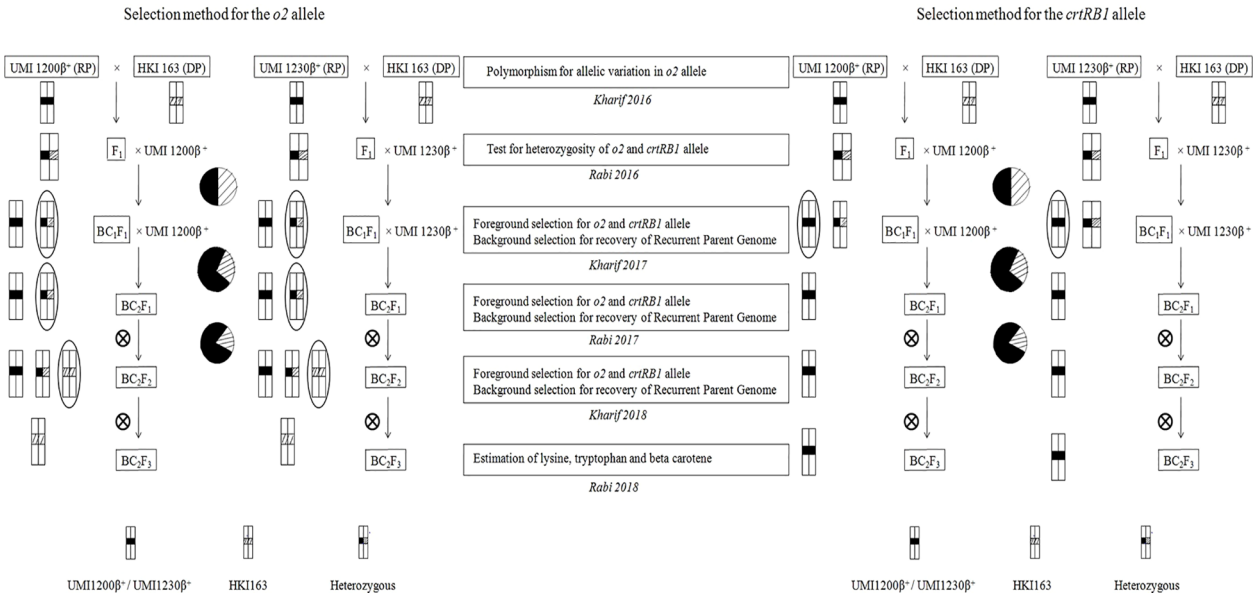


FIGURE 1 | Scheme for the development of *o2* and *crtRB1* genes-derived improved lines using marker-assisted foreground and background selection.

Genomic DNA Isolation and PCR Analysis

Young leaf tissues from two-week-old plants were ground into powder using liquid nitrogen and stored at -80°C. Genomic DNA was isolated using the cetyl trimethylammonium bromide (CTAB) method (Murray and Thompson, 1980). The DNA was checked for its quantity and quality on a 0.8% agarose gel. The PCR for *crtRB1* 3′TE gene-specific and SSR primers and agarose gel electrophoresis were carried out following the method by Muthusamy et al. (2014) and Pukalenthy et al. (2019).

Foreground and Background Selection

o2 gene- and *crtRB1 gene*-linked markers were used for the foreground selection in backcross and selfed lines (**Table 1**). Based on marker polymorphism between donor and recurrent parents, three SSR markers, umc1066, phi 112, and phi057, linked to the *o2* gene and *crtRB1* 3′TE, which is linked to the *crtRB1*, were employed for foreground selection. For the background selection, a total of 236 SSR markers distributed on all 10 chromosomes of maize genome were used to identify polymorphic markers between the donor and recurrent parents. Furthermore, the SSR markers that showed polymorphism among the parents were used in the background selection to determine the recurrent parent genome (RPG) recovery percentage at each backcross generation. All of the SSR primer sequences used in background selection were obtained from the maize genome database (www.maizegdb.org) and synthesized by Eurofin Ltd, Bangalore, India.

Kernel Modification

The parents and heterozygous plant (*O2/o2*) seeds from backcrossed and selfed progenies (BC_1F_1, BC_2F_1, BC_2F_2, and BC_2F_3) were harvested and examined for the kernel modification using a standard light box screening method (Vasal et al., 1980). Maize kernels were categorized into five types *viz.,* type 1, not opaque; type 2, 25% opaqueness; type 3, 50% opaqueness; type 4, 75% opaqueness; and type 5, 100% opaqueness (Vivek et al., 2008). In all of the generations, the kernels with 25% opaqueness were selected and forwarded to the next generation to fix the *o2* allele in its homozygous recessive state and to reduce the undesirable traits caused by the modifier genes acting in the maize endosperm.

Investigation of Morphological Traits in Improved Lines

For the BC_2F_3 improved lines, observations for 15 morphological traits that were categorized and presented chronologically

TABLE 1 | Sequence information of the markers used for polymorphic studies and foreground screening.

S. No.	Marker name	Forward sequence (5'-3')	Reverse sequence (5'-3')	Annealing
1	*phi112*	TGCCCTGCAGGTTCACATTGAGT	AGGAGTACGCTTGGATGCTCTTC	53°C
2	*umc1066*	ATGGAGCACGTCATCTCAATGG	AGCAGCAGCAACGTCTATGACACT	60°C
3	*phi 057*	CTCATCAGTGCCGTCGTCCAT	CAGTCGCAAGAAACCGTTGCC	63°C
4	*crtRB1*	ACACCACATGGACAAGTTCG	ACACTCTGGCCCATGAACAC (R1)	62°C -54°C[a]
			ACAGCAATACAGGGGACCAG (R2)	54°C

[a]*In total of 19 cycles, reduction of temperature by 0.5°C on each cycle starting from the initial 54–62°C.*

according to the plant stage data were taken using standard maize descriptors formulated by the International Board for Plant Genetic Resources (IBPGR) (Anonymous, 1991). Morphological traits *viz.,* days to tasselling (days), days to silking (days), plant height (in centimeters), ear height (in centimeters), tassel length (in centimeters), number of tassel branches, leaf length (in centimeters) and leaf width (in centimeters), cob length (in centimeters), cob girth (in centimeters), number of kernel rows per cob, number of kernels per row, cob weight (in grams), single plant yield (in grams), and 100-kernel weight (in grams) were taken.

Estimation of Lysine, Tryptophan, and β-Carotene Contents

The lysine, tryptophan, and β-carotene contents were estimated in seeds of BC_2F_3 improved lines. The shelled seeds taken for estimation were shade dried and stored at 22–26°C before the analysis. Lysine and tryptophan contents in the endosperm were estimated according to the method described by Galicia et al. (2008). The estimations were done with two replications consisting of two blanks, four checks, and the samples using the spectrophotometer V- 770 UV-VIS-NIT (Japan). The absorbances of lysine and tryptophan were recorded at 390 and 560 nm, respectively. The estimated lysine and tryptophan values were measured with the unit (in percent) (Moro et al., 1996). β-Carotene extraction was done according to the method described by Kurilich and Juvik (1999). The β-carotene content was estimated by high-performance liquid chromatography (HPLC), and samples were eluted by C30 column (5 µm, 4.6 × 250 mm). The mobile phase was composed of acetonitrile:dichlo romethane:methanol (75:20:5). The retention and the spectrum of the carotenoid compounds were found to have a flow rate of 0.4 ml/min and were compared to those of the standard (β-carotene standard-M/s Sigma Aldrich, India). Furthermore, it was reconstituted in the acetonitrile mixture in three different concentrations (1, 10, and 100 ppm).

Statistical Analysis

In BC_1F_1, BC_2F_1, and BC_2F_2 generations, the segregation distortion was studied by chi-square analysis for the deviation from the expected Mendelian ratio. In the background selection, the amplicons were scored as A for recurrent parent, B for donor parent, and H for heterozygous plants. The recovery percentage of the recurrent genome was calculated using the formula RPG (%) = [A + (0.5H)/(A + B + H)] × 100 (Benchimol et al., 2005).

RESULTS

Development of Maize Inbred Lines With the *O2* and *crtRB1* Genes

Three SSR markers, umc1066, phi112, and phi057, located within the *o2* gene were investigated for their polymorphisms among the donor HKI163 and the two recurrent parents *viz.,* UMI1200β$^+$ and UMI1230β$^+$. Among them, umc1066 was found to be polymorphic between the donor and each of the two recurrent parents. This informative SSR marker was further used for the foreground selection. F_1 progenies were produced from two independent crosses of UMI1200β$^+$×HKI163 and UMI1230β$^+$×HKI163. BC_1F_1 progenies were obtained by backcrossing the F_1 plants with UMI1200β$^+$ and UMI1230β$^+$ as the recurrent parents. In the BC_1F_1 generation, individual plants homozygous at the *crtRB1* and heterozygous at the *o2* locus were identified using the *crtRB1* and *o2*-gene specific markers and utilized for next backcrossing with the recurrent parent. Furthermore, BC_2F_1 progenies were obtained from the selected BC_1F_1 plants based on the dual-selection procedure involving foreground selection and light box screening. Applying similar selection procedures and selfing, progenies of BC_2F_1 generation were advanced to BC_2F_2 (**Figure 2**) and BC_2F_3. Finally, from each cross, five BC_2F_3 lines with homozygous marker alleles for the *CrtRB1* and *o2* genes were developed (**Figure 3**). The segregation patterns of backcross progenies are presented in **Table 2**.

SSR-Based Genetic Background Analysis of Improved Lines

A set of 236 SSR markers distributed uniformly across the maize genome was used in polymorphism screening to select polymorphic markers between donor and recurrent parents. Among them, 104 and 107 SSR markers showed polymorphism between UMI1200β$^+$ and HKI163 and UMI1230β$^+$ and HKI163, respectively. The polymorphism percentage was recorded as 44.6 and 49.57%, respectively. Furthermore, these polymorphic markers were employed to screen the progenies derived from backcross and selfed generation for the recovery of RPG (**Figure 4**). In BC_1F_1 generation, a total of 22 and 18 foreground positive plants from UMI1200β$^+$×HKI163 and UMI1230β$^+$×HKI163 were screened, which showed a recovery of RPG of 54.81% and 53.21%, respectively. Furthermore, the recovery of RPG increased in subsequent generations. The 22 and 31 selected positive plants from UMI1200β$^+$×HKI163 and UMI1230β$^+$×HKI163 in BC_2F_1 showed 82.21 and 79.81%, of RPG, respectively. Eight and six positive plants from UMI1200β$^+$×HKI163 and UMI1230β$^+$×HKI163 in BC_2F_2 showed 87.48 and 86.51% recovery of RPG, respectively. The highest recoveries of RPG 91.21% and 90.08% were obtained in each of five BC_2F_3 plants from UMI1200β$^+$×HKI163 and UMI1230β$^+$×HKI163.

Kernel Modification

Opaqueness is the indicator for the presence of *o2* allele, it is also tightly linked to the *o2* gene, selecting the kernels along with the least opaqueness from generation to generation ensures that the *o2* gene is fixed in its homozygous recessive state. Thus, we observed the opaqueness in selected foreground positive progenies from backcrossed and selfed progenies, along with HKI163, UMI1200β$^+$, and UMI1230β$^+$ for kernel modification. HKI163 kernels showed 25 and 50% opaqueness, whereas UMI1200β$^+$ and UMI1230β$^+$ exhibited 0% opaqueness. BC_1F_1, BC_2F_1, and BC_2F_2 progenies showed 0–100% opaqueness. Among them, progenies showing 25% were further selected and advanced to next generation, whereas the remainder were rejected. In maize, the endosperm modifier genes play a major

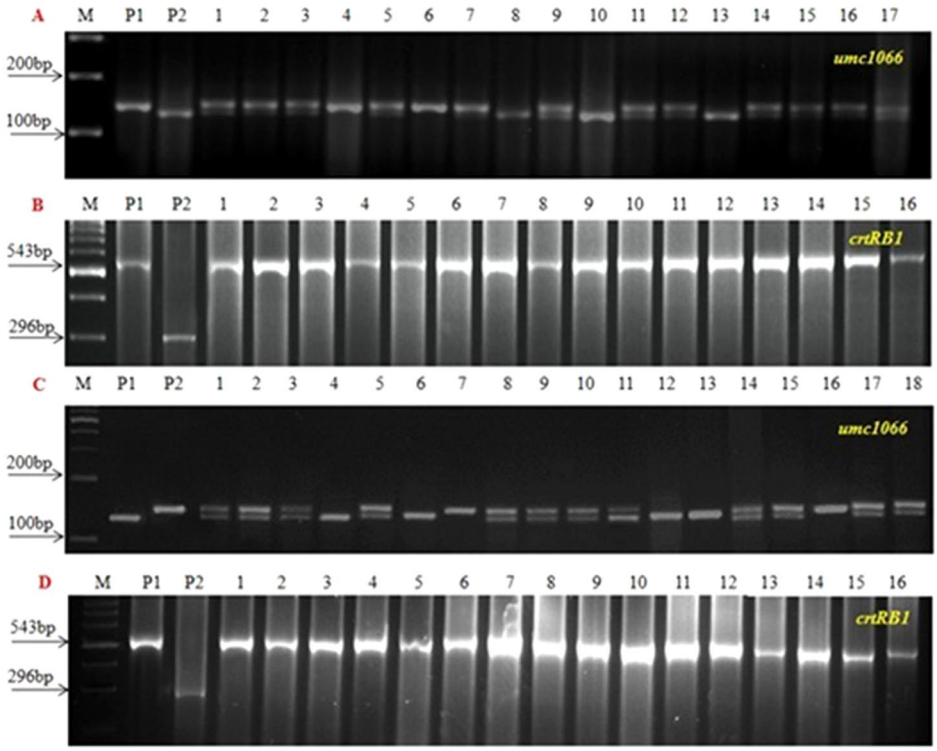


FIGURE 2 | Foreground screening in BC_2F_2 progenies. **(A** and **B)** UMI1200β⁺xHKI163 **(C** and **D)** UMI1230β⁺xHKI163, (M) Marker 100bp, (P_1) UMI1200β⁺/UMI1230β⁺, (P_2) HKI163 and (1-16) BC_2F_2 progenies.

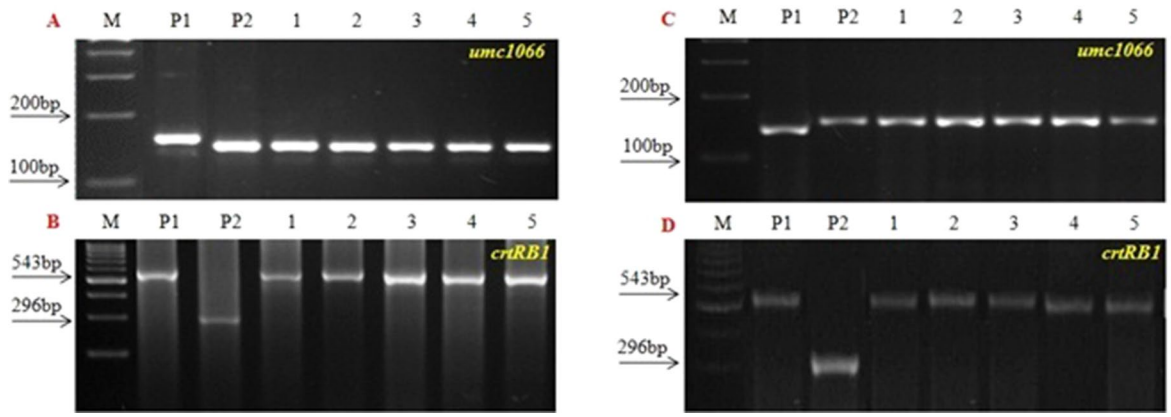


FIGURE 3 | PCR screening of five improved BC2F3 lines using *o2* SSR and *crtRB1* gene specific marker. (A and B) UMI1200β⁺ × HKI163 and (C and D) UMI1230β⁺ × HKI163. (M) Marker 100 bp, (P1) UMI1200β⁺/UMI1230β⁺, (P2) HKI163. (1-5) Improved lines from UMI1200β⁺×HKI163 (1-DBT6-1-5/25-8/25-4/25-4/25, 2- DBT6-1-5/25-8/25-9/25-9/25, 3- DBT6-1-5/25-10/25-15/25-15/25, 4- DBT6-1-5/25-10/25-17/25-17/25, 5- DBT6-1-5/25-14/25-11/25-11/25) and UMI1230β⁺×HKI163 (1- DBT7-1-6/25-9/25-37/25-37/25, 2- DBT7-1-6/25-9/25-57/25-57/25, 3- DBT7-1-6/25-12/25-23/25-23/25, 4- DBT7-1-6/25-27/25-3/25-3/25, 5- DBT7-1-6/25-27/25-67/25-67/25).

role to produce undesirable characteristics, which affect the crop yield. Thus, we selected the progenies with 25% opaqueness to reduce the effect of the *o2* modifier gene action. As a result, the recessive allele of *o2/o2* was fixed in maize kernels and all of the BC_2F_3 lines showed 25% opaqueness **(Figure 5)**.

Morphological Characteristics of Improved Lines

The morphological traits of the improved lines along with their donor and recurrent parents were presented in **Table 3**. Five UMI1200β⁺×HKI163-based improved lines showed phenotypic

TABLE 2 | Segregation pattern of *o2* allele in the backcross and selfed generation of UMI1200β⁺×HKI163 and UMI1230β⁺×HKI163.

S. No.	Cross	Generation	No. of plants genotyped	Genotypic class			χ^2	P value
				No. of A	No. of H	No. of B		
1	UMI1200β⁺ × HKI163	BC_1F_1	197	74	123	0	12.18782**	0.000481
		BC_2F_1	194	84	110	0	03.484536^{ns}	0.061945
		BC_2F_2	126	33	082	11	19.143**	0.000069
2	UMI1230β⁺ × HKI163	BC_1F_1	175	40	135	0	51.5714286**	0.00006903
		BC_2F_1	100	43	057	0	01.96^{ns}	0.161513
		BC_2F_2	106	36	040	30	07.566^{ns}	0.0294

***Significant different ($p < 0.01$), the markers were significantly distorted from the expected segregation ratio (1:2:1); ^{ns}non-significant ($p \geq 0.01$), the markers were non-significant and fitted to the normal Mendalian ratio.*

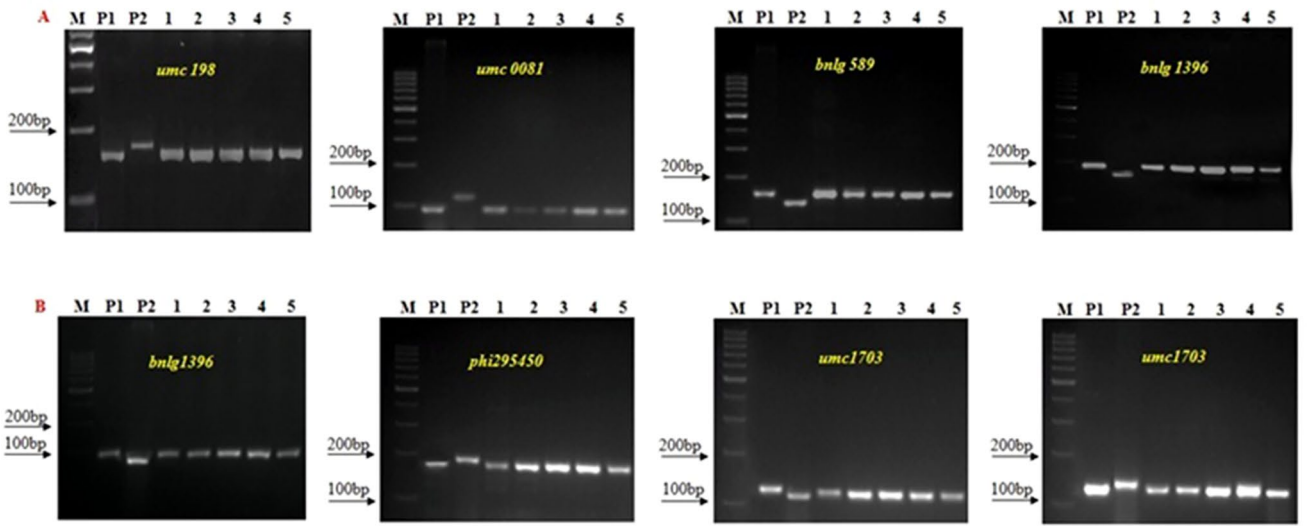


FIGURE 4 | Background Screening of improved lines using SSR makers. **(A)** 1200β⁺xHKI163, **(B)** UMI1230β⁺xHKI163. (M) Marker 100bp, (P_1) UMI 1200β⁺/UMI 1230β⁺, (P_2) HKI163, (1-5) Improved lines from 1200β⁺xHKI163 (1-DBT6-1-5/25-8/25-4/25-4/25, 2-DBT6-1-5/25-8/25-9/25-9/25, 3- DBT6-1 5/25-10/25-15/25-15/25, 4-DBT6-1-5/25-10/25-17/25-17/25, 5- DBT6-1-5/25-14/25-11/25-11/-25) and 1230β⁺xHKI163(1- DBT7-1-6/25-9/25-37/25-37/25, 2-DBT6-1-6/25-9/25-57/25-57/25, 3- DBT7-1-6/25-12/25-23/25-23/25, 4- DBT7-1-6/24-27/25-3/25-3/25, 5-DBT7-1-6/25-27/25-67/25-67/25).

resemblance ranging from 71.43% (number of kernels per row) to 98.38% (days to silking). Among them, DBT6-1-5/25-10/25-17/25-17/25 and DBT6-1-5/25-14/25-11/25-11/25 possessed high phenotypic resemblance with the recurrent parent. For instance, days to tasselling, days to silking, plant height, ear height, tassel length, cob girth, cob weight, 100-kernel weight, and single plant yield showed more than 90% similarity to UMI1200β⁺. Likewise, five UMI1230β⁺×HKI163-based improved lines showed phenotypic resemblance ranging from 82.43% (ear height) to 99.66% (100-kernel weight). DBT7-1-6/25-12/25-23/25-23/25 showed maximum similarity with the recurrent parent, followed by DBT7-1-6/25-27/25-67/25-67/25. These two lines exhibited more than 90% similarity to UMI1230β⁺ for the traits days to tasselling, days to silking, plant height, ear height, tassel length, leaf length, cob length, cob girth, cob weight, 100-kernel weight, and single plant yield **(Figure 6)**.

Analysis of Lysine, Tryptophan, and *β-Carotene*

All of the improved lines showed that lysine and tryptophan content increased many-fold over for both the β-carotene-rich parents *viz.,* UMI1200β⁺ and UMI1230β⁺. Lysine and tryptophan contents varied from 0.294 to 0.332% and 0.073 to 0.081%, with an average of 0.314 and 0.077%. Among the improved lines, DBT 7-1-6/25-27/25-3/25-3/25 from UMI1230β⁺×HKI163 possessed higher levels of lysine (0.332%) and tryptophan (0.081%). Furthermore, accumulation of the β-carotene content was estimated in improved lines, which ranged from 6.127 to 7.387 μg/g with an average of 6.80 μg/g and higher than the QPM parent HKI163. DBT6-1-5/25-8/25-4/25-4/25 from UMI1200β⁺×HKI163 was found to have a high β-carotene content of 7.387 μg/g among improved lines. The lysine, tryptophan, and β-carotene contents of the improved lines are presented in **Table 4**.

DISCUSSION

The Value of the Pyramiding *O2* and *crtRB1* Genes

Lysine, tryptophan, and β-carotene are the key nutritional traits in maize. The genetic nature and environmental factors

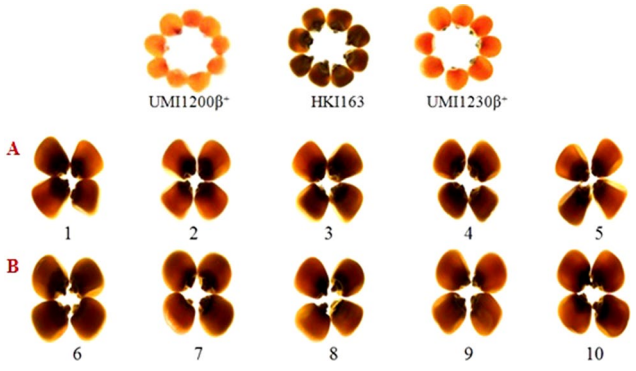


FIGURE 5 | Endosperm modification of BC_2F_3 **(A** and **B)** of UMI1200β+xHKI163 and UMI1230β+xHKI163. **(A)** Five improved lines form the cross UMI1200β+xHKI163(1-DBT6-1-5/25-8/25-4/25-4/25, 2-DBT6-1-5/25-8/25-9/25-9/25, 3-DBT6-1-5/25-10/25-15/25-15/25, 4-DBT6-15/25-10/25-17/25-17/25, 5-DBT6-1-5/25-14/25-11/25-11/25) **(B)** Five improved lines from the cross UMI1230β+XHKI163(6-DBT7-1/25-9/25-37/25-37/25, 7-DBT7-1-6/25-9/25-57/25-57/25, 8-DBT7-1-6/25-12/25-23/25-23/25, 9-DBT7-1-6/25-27/25-3/25-3/25, 10-DBT7-1-6/25-27/25-67/25-67/25).

have an influence on these traits. *crtRB1* and *o2* genes present on chromosomes 10 and 7 (Mertz et al., 1964; Vasal, 2000; Yang et al., 2004) provide increased β-carotene, lysine, and tryptophan contents. Molecular markers linked to these genes are available to facilitate direct selection in the breeding process. In this study, we successfully pyramided the *o2* and *crtRB1* genes in maize by MAS and several generations of backcrossing. The β-carotene content of the improved lines was increased by five- to six-fold for both crosses when compared to the QPM parent. The lysine and tryptophan contents of the improved lines were increased by two- and seven-fold for both crosses compared to the β-carotene parents. Thus, *o2* and *crtRB1* genes can work together in the same genetic background to control the content of lysine, tryptophan, and β-carotene.

Development of Improved Lines Through MAB Breeding

Parental polymorphism screening revealed that recurrent parents UMI1200β+ and UMI1230β+ were clearly distinguishable with *o2* gene and *CrtRB1* gene-linked markers *umc1066* and *crtRB1* 3′TE from the donor line HKI163 and thus were used for foreground selection in the F_1, BC_1F_1, BC_2F_1, BC_2F_2, and BC_2F_3 generations. In foreground selection, F_1 and BC_1F_1 generations screening with *crtRB1* allele indicated that all of the genotypes were heterozygous in the F_1 generation and the segregation distortion in the BC_1F_1 generation (Babu et al., 2013). From the BC_1F_1 generation, the lines were fixed for the *crtRB1* allele by selecting the plants with favorable allele (543bp) and rejecting the heterozygous plants with both allele (543bp+296bp). Therefore, no segregation existed for *crtRB1* allele in the forwarded generations. Screening for the *o2* gene revealed that BC_2F_1 of UMI1200β+×HKI163 and BC_2F_1 and BC_2F_2 of UMI1230β+×HKI163 showed approximately 50% of heterozygous plants with respect to the expected Mendalian ratio (1:1) in the backcross generations and (1:2:1) in the selfed generations. However, segregation distortion was observed in BC_1F_1 and BC_2F_2 of UMI1200β+×HKI163 and BC_1F_1 of UMI1230β+×HKI163. These results are in accordance with the previous reports (Liu et al., 2015; Tripathy et al., 2017; Goswami et al., 2019; Sagare et al., 2019). Furthermore, background analysis using genome-wide SSR markers revealed 91.21 and 90.08% recovery of RPG in each of the five BC_2F_3 plants from UMI1200β+×HKI163 and UMI1230β+×HKI163 and coupled with the earlier findings (Feng et al., 2015; Sarika et al., 2018).

Characteristics of Improved Lines

In addition to the background selection, phenotypic characterization is also useful to find the recovery percentage of recurrent parents (Manna et al., 2005; Gunjaca et al., 2008; Choudhary et al., 2014; Hossain et al., 2018). Phenotypic characterization among the parents and the improved lines showed more than 90% of recovery of the recurrent parents in morphological traits. Among them, DBT6-1-5/25-10/25-17/25-17/25 and DBT6-1-5/25-14/25-11/25-11/25 from UMI1200β+×HKI163 and DBT7-1-6/25-12/25-23/25-23/25 and DBT7-1-6/25-27/25-67/25-67/25 from UMI1230β+×HKI 163 possessed high phenotypic resemblance (90%) with their recurrent parents. Previously, several studies also reported more than 90% recovery of the recurrent parent characteristics in MAS-derived lines (Surender et al., 2017; Pukalenthy et al., 2019; Sagare et al., 2019). The lysine and tryptophan contents of the improved lines ranged from 0.294 to 0.331% and 0.073 to 0.080% for the cross UMI1200β+×HKI163 and 0.298 to 0.332% and 0.073 to 0.081% for the cross UMI1230β+×HKI163. On the average, lysine and tryptophan contents of the improved lines were 0.314 and 0.077%; they are at par with the QPM parent, three and seven-fold increases from the recurrent parents. Likewise, the average β-carotene contents of the improved lines for UMI1200β+×HKI163 and UMI1230β+×HKI163 were 6.846 and 6.766 μg/g, respectively, which were comparable to the β-carotene parents, six-fold higher than the QPM parent. Similar results were obtained by various studies (Muthusamy et al., 2014; Zunjare et al., 2018; Goswami et al., 2019). Overall, the improved inbred lines gained lysine and tryptophan contents but a slight reduction in β-carotene content (>2 ug) and grain yield. We followed the dual-selection procedure of molecular and light box screening to fix the *o2* allele, which is the reason behind increasing lysine and tryptophan contents. We selected the progenies based on the good agronomic performance (>90%) and β-carotene content, even though some of the progenies recorded β-carotene contents at par with the recurrent parents with less agronomic performance. Thus, a slight reduction was observed in β-carotene content (>2 ug) of improved inbred lines. Moreover, introgression of *o2* and *crtRB1* genes caused a reduction in the grain yield. It is reported that QPM lines have some undesirable characteristics because of the modifier gene action in the endosperm. Thus, we used dual selection procedure to select the progenies and developed the improved inbred lines with less undesirable traits along with *o2* and *crtRB1* genes. However, it is not possible to stop the modifier gene (*o2)* activity and remove the undesirable traits completely. It might influence

TABLE 3 | Comparison of BC_2F_3 improved lines from UMI1200β⁺×HKI163 and UMI1230β⁺×HKI163 with the parents for the recovery percentage of morphological traits.

Morphological traits	**UMI1200β⁺ (Recurrent parent)**	**HKI163 (Donor parent)**	**Identified positive lines**					**Recovery percentage (%) for morphological trait**				
(UMI1200β⁺ × HKI163)			**DBT6-1-5/25-8/25-4/25-4/25**	**DBT6-1-5/25-8/25-9/25-9/25**	**DBT6-1-5/25-10/25-15/25-15/25**	**DBT6-1-5/25-10/25-17/25-17/25**	**DBT6-1-5/25-14/25-11/25-11/25**	**DBT6-1-5/25-8/25-4/25-4/25**	**DBT6-1-5/25-8/25-9/25-9/25**	**DBT6-1-5/25-10/25-15/25-15/25**	**DBT6-1-5/25-10/25-17/25-17/25**	**DBT6-1-5/25-14/25-11/25-11/25**
Days to tasseling (days)	59.00	63.00	56.00	57.00	58.00	58.00	57.00	94.91	82.60	98.30	98.30	96.61
Days to silking (days)	62.00	65.00	58.00	60.00	61.00	60.00	59.00	93.54	96.77	98.38	96.77	95.16
Plant height (cm)	126.88	110.00	118.00	121.40	123.50	119.90	122.70	93.00	95.68	97.33	94.49	96.70
Ear height (cm)	71.00	62.00	68.00	67.70	69.30	68.30	68.00	95.77	95.35	97.60	96.19	95.77
Tassel length (cm)	21.00	26.20	18.00	19.20	18.10	19.50	20.30	85.71	91.42	86.19	92.38	96.66
Number of tassel branches	12.00	16.00	10.00	11.00	10.00	11.00	10.00	83.33	91.66	83.33	91.66	83.33
Leaf length (cm)	61.65	63.00	58.20	59.40	59.20	58.00	54.70	94.40	96.35	96.02	94.07	88.72
Leaf width (cm)	06.10	06.00	05.80	05.80	05.70	05.60	05.00	95.08	95.08	93.44	91.80	81.96
Cob length (cm)	15.20	17.00	13.00	13.80	14.00	13.00	12.40	85.52	90.78	92.10	85.52	81.57
Cob girth(cm)	13.80	13.00	12.00	11.50	13.00	12.70	13.10	86.95	83.33	94.20	92.02	94.92
Number of kernel rows per cob	14.00	12.00	12.00	10.00	12.00	10.00	10.00	85.71	85.71	71.43	85.71	71.43
Number of kernels per row	21.00	26.00	18.00	17.00	17.00	18.00	17.00	85.71	80.95	80.95	85.71	94.92
Cob weight (g)	108.00	112,00	98.00	92.80	94.30	101.00	97.90	90.74	85.92	87.31	93.51	90.64
100-kernel weight (g)	24.00	22.00	23.00	22.30	22.80	23.10	22.80	95.83	92.91	95.00	96.25	95.00
Single plant yield (g)	90.67	57.00	82.11	83.07	84.11	86.21	87.32	90.55	91.61	92.76	95.08	96.31
(UMI1200β⁺ × HKI163)	**UMI1230β⁺ (Recurrent parent)**	**HKI163 (Donor parent)**	**DBT7-1-6/25-9/25-37/25-37/25**	**DBT7-1-6/25-9/25-57/25-57/25**	**DBT7-1-6/25-12/25-23/25-23/25**	**DBT7-1-6/25-27/25-3/25-3/25**	**DBT7-1-6/25-27/25-67/25-67/25**	**DBT7-1-6/25-9/25-37/25-37/25**	**DBT7-1-6/25-9/25-57/25-57/25**	**DBT7-1-6/25-12/25-23/25-23/25**	**DBT7-1-6/25-27/25-3/25-3/25**	**DBT7-1-6/25-27/25-67/25-67/25**
Days to tasseling (days)	61.00	63.00	58.00	58.00	59.00	60.00	58.00	95.08	95.08	96.72	98.36	95.08
Days to silking (days)	63.00	65.00	61.00	60.00	62.00	62.00	60.00	96.82	95.23	98.41	98.41	95.23
Plant height (cm)	112.00	110.00	110.50	109.70	104.60	101.00	105.20	98.66	97.94	93.39	90.17	93.92
Ear height (cm)	74.00	62.00	61.00	72.00	71.00	72.10	72.90	82.43	97.29	95.94	97.43	98.51
Tassel length (cm)	24.10	26.20	23.40	22.80	23.70	23.90	22.80	97.09	94.60	98.34	99.17	94.60
Number of tassel branches	14.00	16.00	13.00	12.00	13.00	12.00	12.00	92.85	85.71	92.85	85.71	85.71
Leaf length (cm)	61.00	63.00	58.40	60.20	59.00	59.20	58.90	95.73	98.68	96.72	97.04	96.55
Leaf width (cm)	06.50	06.00	06.30	6.10	05.70	05.80	05.90	96.92	93.84	87.69	89.23	90.76
Cob length (cm)	16.00	17.00	14.00	15.60	14.90	15.80	15.30	87.50	97.50	93.12	98.75	95.62
Cob girth (cm)	12.50	13.00	11.00	11.80	12.10	11.60	12.00	88.00	94.40	96.80	92.80	96.00
Number of kernel rows per cob	16.00	12.00	14.00	14.00	14.00	14.00	14.00	87.50	87.50	87.50	87.50	87.50
Number of kernels per row	23.00	26.00	22.00	22.00	22.00	21.00	19.00	95.65	95.65	95.65	91.30	82.60
Cob weight (g)	124.70	112.00	122.60	118.30	113.80	120.90	121.00	98.31	94.86	91.25	96.95	97.03
100-kernel weight (g)	29.50	22.00	27.40	28.10	28.90	29.40	28.70	92.88	95.25	97.96	99.66	97.28
Single plant yield (g)	76.98	57.00	72.11	69.43	71.67	72.11	73.31	93.67	90.19	93.10	93.67	95.23

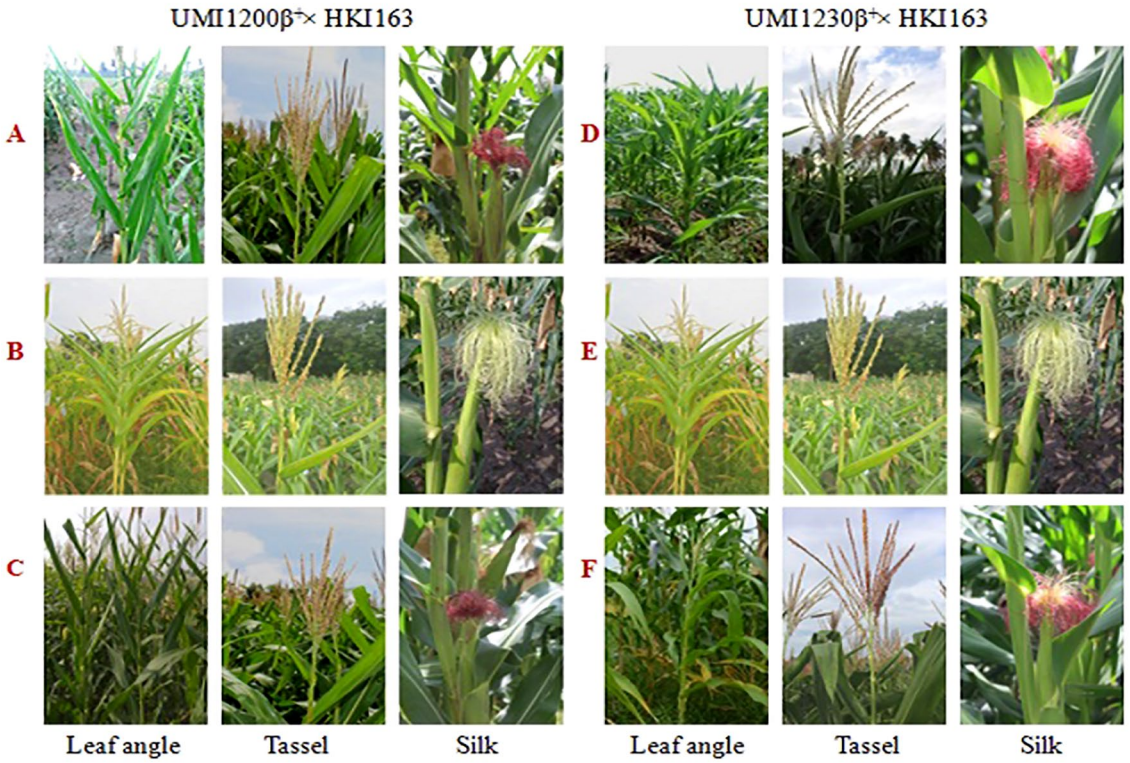


FIGURE 6 | Morphological traits of parents and improved lines of UMI1200β+xHKI163 and UMI1230β+xHKI163. **(A)** UMI1200β+, **(B)** HK163, **(C)** DBT6-1-5/25-10/25-17/25- 17/25, **(D)** UMI1230β+, **(E)** HKI163 **(F)** DBT7-1-6/25-12/25-23/25-23/25.

TABLE 4 | Lysine and tryptophan concentrations in parents and BC_2F_3 improved lines from UMI1200β+×HKI163 and UMI1230β+×HKI163.

Carotenoid/ Amino acids	UMI1200β+	UMI1230β+	HKI163	(UMI1200β+ × HKI163)					(UMI1230β+ × HKI163)				
				DBT6-1-5/25-8/25-4/25-4/25	DBT6-1-5/25-8/25-9/25-9/25	DBT6-1-5/25-10/25-15/25-15/25	DBT6-1-5/25-10/25-17/25-17/25	DBT6-1-5/25-14/25-11/25-11/25	DBT7-1-6/25-9/25-37/25-37/25	DBT7-1-6/25-9/25-57/25-57/25	DBT7-1-6/25-12/25-23/25-23/25	DBT7-1-6/25-27/25-3/25-3/25	DBT7-1-6/25-27/25-67/25-67/25
β- carotene (µg/g)	9.073	9.232	0.800	7.387	6.127	6.665	7.187	6.865	7.123	7.265	6.321	6.812	6.312
Lysine (%)	0.130	0.150	0.340	0.294	0.312	0.299	0.322	0.331	0.298	0.317	0.308	0.332	0.327
Tryptophan (%)	0.024	0.029	0.082	0.077	0.079	0.080	0.073	0.080	0.073	0.077	0.079	0.081	0.075

the yield attributing traits and reduces the yield performances. Thus, we obtained a reduction in the grain yield similar to a previous study (Lauderdale, 2000).

In the present study, using MAB breeding approach, we successfully pyramided the *o2* and *crtRB1* genes and developed the nutrition-rich inbreds, but introgression of multiple genes caused a slight reduction in the yield. To utilize these newly developed inbred lines effectively, our future research focus is on conducting multilocation trial (MLT) in various maize-growing regions and identifying the superior inbred lines to develop new hybrids. In addition, these inbred lines can be used as genetic resources for maize biofortification programs.

DATA AVAILABILITY

The datasets generated for this study are available on request to the corresponding author.

AUTHOR CONTRIBUTIONS

SN, FH, and VM designed the methods and experiments. SC, BP, DM, and RR developed backcross progenies and managed fieldwork. LJ and VC provided suggestions on experiments and monitored the work. SC, BP, DM, KAd, and KE conducted

phenotype and genotype analysis. SC, VS, and KAr performed biochemical analysis. SC, BP, and DM analyzed the data. SC, KAd, and SN drafted the manuscript.

FUNDING

Financial support of Department of Biotechnology (DBT), Government of India (GOI), through project entitled 'Enrichment of nutritional quality in maize through molecular breeding' (BT/PR10922AGII/106/9442014 dt.25.3.2015) and Ministry of Science and Technology, Department of Biotechnology, Government of India DBT's Twinning program for the NE for the project entitled 'Marker assisted introgression of LyCE gene for enhanced ProA in maize' (BT/166/NE/TBP/2011 dt 12.12.2011) is acknowledged. The funders had no role in the work design, data collection and analysis, or decision and preparation of the manuscript.

ACKNOWLEDGMENTS

The authors acknowledged the Centre of Innovation, Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai, for providing lab and instrumental facilities.

REFERENCES

Anonymous (1991). "IBPGR. Descriptors for Maize," in *International Wheat and Maize Improvement Center* (Mexico City: International Board for Plant Genetic).

Babu, R., Rojas, N. P., Gao, S., Yan, J., and Pixley, K. (2013). Validation of the effects of molecular marker polymorphisms in LcyE and CrtRB1 on provitamin A concentrations for 26 tropical maize populations. *Theor. Appl. Genet.* 126, 389–399. doi: 10.1007/s00122-012-1987-3

Benchimol, L., Souza, C. L. D., Jr., and Souza, A. P. D. (2005). Microsatellite-assisted backcross selection in maize. *Genet. Mol. Biol.* 28, 789–797. doi: 10.1590/S1415-47572005000500022

Choudhary, M., Muthusamy, V., Hossain, F., Thirunavukkarasu, N., Pandey, N., Jha, S. K., et al. (2014). Characterization of β-carotene rich MAS-derived maize inbreds possessing rare genetic variation in β-carotene hydroxylase gene. *Indian J. Genet.* 74, 620–623. doi: 10.5958/0975-6906.2014.00900.6

Feng, F., Wang, Q., Liang, C., Yang, R., and Li, X. (2015). Enhancement of tocopherols in sweet corn by marker-assisted backcrossing of ZmVTE4. *Euphytica* 206, 513–521. doi: 10.1007/s10681-015-1519-8

FAOSTAT. (2018). Crop prospects and food situation. Rome: Food and Agriculture Organization of the United Nation. Available from: http://faostat.fao.org (Accessed July 06, 2018).

Galicia, L., Nurit, E., Rosales, A., and Palacios-Rojas, N. (2008). Maize nutrition quality and plant tissue analysis laboratory. *Lab. Protoc.* 2008. 10–17.

Giuliano, G. (2017). Provitamin A biofortification of crop plants: a gold rush with many miners. *Curr. Opin. Biotechnol.* 44, 169–180. doi:10.1016/j.copbio.2017.02.001

Goswami, R., Zunjare, R. U., Khan, S., Baveja, A., Muthusamy, V., and Hossain, F. (2019). Marker-assisted introgression of rare allele of β-carotene hydroxylase (*crtRB1*) gene into elite quality protein maize inbred for combining high lysine, tryptophan and provitamin A in maize. *Plant Breed.* 138, 174–183. doi: 10.1111/pbr.12676

Gunjaca, J., Buhinicek, I., Jukic, M., Sarcevic, H., Vrag olovic, A., Kozic, Z., et al. (2008). Discriminating maize inbred lines using molecular and DUS data. *Euphytica* 161, 165–172. doi: 10.1016/j.copbio.2017.02.001

Gupta, H. S., Raman, B., Agrawal, P. K., Mahajan, V., Hossain, F., and Thirunavukkarasu, N. (2013). Accelerated development of quality protein maize hybrid through marker-assisted introgression of opaque-2 allele. *Plant Breed.* 132, 77–82. doi: 10.1111/pbr.12009

Hossain, F., Muthusamy, V., Pandey, N., Vishwakarma, A. K., Baveja, A., Zunjare, R. U., et al. (2018). Marker-assisted introgression of opaque2 allele for rapid conversion of elite hybrids into quality protein maize. *J. Genet.* 97, 287–298. doi: 10.1007/s12041-018-0914-z

Kurilich, A. C., and Juvik, J. A. (1999). Quantification of Carotenoid and Tocopherol Antioxidants in. *Zea mays. J. Agric. Food Chem.* 47, 1948–1955. doi: 10.1021/jf981029d

Lauderdale, J. (2000). Issues regarding targeting and adoption of quality protein maize (QPM) in *CIMMYT Economics Working Paper 00–02*. Mexico D.F. CIMMYT, pp. iv–31.

Liu, L., Jeffers, D., Zhang, Y., Ding, M., Chen, W., Kang, M. S., et al. (2015). Introgression of the crtRB1 gene into quality protein maize inbred lines using molecular markers. *Mol. Breed.* 35, 154. doi: 10.1007/s11032-015-0349-7

Ma, Y., and Nelson, O. E. (1975). Amino acid composition and storage proteins in two new high-lysine mutants in maize. *Cereal Chem.* 52, 412–419.

Manna, R., Okello, O. K., Imanywoha, J., Pixley, K., and Edema, R. (2005). Enhancing introgression of the opaque-2 trait into elite maize lines using simple sequence repeats. *Afr. Crop Sci. J.* 13, 215–226.

McWhirter, K. S. (1971). A floury endosperm, high lysine locus on chromosome 10. *Maize Genet. Coop. Newsletter* 45, 184.

Mertz, E. T., Bates, L. S., and Nelson, O. E. (1964). Mutant gene that changes protein composition and increases lysine content of maize endosperm. *Science* 145, 279–280. doi: 10.1126/science.145.3629.279

Moro, G. L., Habben, J. E., Hamaker, B. R., and Larkins, B. A. (1996). Characterization of the variability in lysine content for normal and opaque2 maize endosperm. *Crop Sci.* 36, 1651–1659. doi: 10.2135/cropsci1996.0011183X003600060039x

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4326. doi: 10.1093/nar/8.19.4321

Muthusamy, V., Hossain, F., Thirunavukkarasu, N., Choudhary, M., Saha, S., Bhat, J. S., et al. (2014). Development of β-carotene rich maize hybrids through marker-assisted introgression of β-carotene hydroxylase allele. *PLoS One* 9, e113583. doi: 10.1371/journal.pone.0113583

Nelson, O. E., Mertz, E. T., and Bates, L. S. (1965). Second mutant gene affecting the amino acid pattern of maize endosperm proteins. *Science* 150, 1469–1470. doi: 10.1126/science.150.3702.1469

Pukalenthy, B., Manickam, D., Chandran, S., Adhimoolam, K., Sampathrajan, V., Rajasekaran, R., et al. (2019). Incorporation of opaque-2 into 'UMI 1200', an elite maize inbred line, through marker-assisted backcross breeding. *Biotechnol. Biotechnol. Equip.* 1–10. doi: 10.1080/13102818.2018.1556121

Rosegrant, M. W., Ringler, C., Sangi, S. M., Sulser, T. B., Zhu, T., and Cline, S. A. (2009). *Model description international model for policy analysis of agricultural commodities and trade.* International food policy research institute. Washington. D.C. 12–13.

Sagare, D. B., Shetti, P., Surender, M., and Reddy, S. S. (2019). Marker-assisted backcross breeding for enhancing β-carotene of QPM inbreds. *Mol. Breed.* 39, 31. doi: 10.1007/s11032-019-0939-x

Sarika, K., Hossain, F., Muthusamy, V., Zunjare, R. U., Baveja, A., Goswami, R., et al. (2018). Marker-assisted pyramiding of opaque2 and novel opaque16 genes for further enrichment of lysine and tryptophan in sub-tropical maize. *Plant Sci.* 272, 142–152. doi: 10.1016/j.plantsci.2018.04.014

Surender, M., Shetti, P., Sagare, D. B., Durga Rani, C. V., Jabeen, F., Sudarshan, M. R., et al. (2017). Development of QPM version of DHM117 maize hybrid using marker assisted selection. *Int. J. Curr. Microbiol. App. Sci.* 6, 3275–3289. doi: 10.20546/ijcmas.2017.610.384

Tripathy, S. K., Lenka, D., Ithape, D. M., and Maharana, M. (2017). Development of opaque-2 introgression line in maize using marker assisted backcross breeding. *Plant Gene* 10, 26–30. doi: 10.1016/j.plgene.2017.03.005

Ullah, I., Ali, M., and Farooqi, A. (2010). Chemical and nutritional properties of some maize (Zea mays L.) varieties grown in NWFP, Pakistan. *Pak. J. Nutr.* 9, 1113–1117. doi: 10.3923/pjn.2010.1113.1117

Vasal, S. K. (2000). The quality protein maize story. *Food Nutr. Bull.* 21, 445–450. doi: 10.1177/156482650002100420

Vasal, S. K., Villegas, E., Bjarnason, M., Gelaw, B., and Goertz, P. (1980). Genetic modifiers and breeding strategies in developing hard endosperm opaque-2 materials. *CAB Direct*. 33, 37–73.

Vivek, B. S., Krivanek, A. F., Palacios Rojas, N., Twumasi Afriyie, S., Diallo, A. O. (2008). Breeding Quality Protein Maize (QPM): *Protocols for Developing QPM Cultivars*. Mexico, D.F. CIMMYT. 8-11

Yan, J., Kandianis, C. B., Harjes, C. E., Bai, L., Kim, E. H., Yang, X., et al. (2010). Rare genetic variation at Zea mays crtRB1 increases β-carotene in maize grain. *Nat. Genet.* 42, 322. doi: 10.1038/ng.551

Yang, R., Yan, Z., Wang, Q., Li, X., and Feng, F. (2018). Marker-assisted backcrossing of lcyE for enhancement of proA in sweet corn. *Euphytica* 214, 130. doi: 10.1007/s10681-018-2212-5

Yang, W., Zheng, Y., Ni, S., and Wu, J. (2004). Recessive allelic variations of three microsatellite sites within the O2 gene in maize. *Plant Mol. Boil. Rep.* 22, 361–374. doi: 10.1007/BF02772679

Yang, W., Zheng, Y., Zheng, W., and Feng, R. (2005). Molecular genetic mapping of a high-lysine mutant gene (opaque-16) and the double recessive effect with opaque-2 in maize. *Mol. Breed.* 15, 257–269. doi: 10.1007/s11032-004-5947-8

Zunjare, R. U., Hossain, F., Muthusamy, V., Baveja, A., Chauhan, H. S., Bhat, J. S., et al. (2018). Development of biofortified maize hybrids through marker-assisted stacking of β-carotene hydroxylase, lycopene-ε-cyclase and opaque2 genes. *Front. Plant Sci.* 9, 178. doi: 10.3389/fpls.2018.00178

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

frontiers in Genetics

MINI REVIEW
published: 20 December 2019
doi: 10.3389/fgene.2019.01289

Thinking Outside of the Cereal Box: Breeding Underutilized (Pseudo) Cereals for Improved Human Nutrition

Cody S. Bekkering and Li Tian *

Department of Plant Sciences, University of California, Davis, Davis, CA, United States

Cereal grains have historically played a critical role in sustaining the caloric needs of the human population. The major cereal crops, wheat, rice, and maize, are widely cultivated and have been subjected to biofortification to enhance the vitamin and mineral nutrient content of grains. In contrast, grains of several other cereals as well as non-grass pseudocereals are naturally rich in micronutrients, but have yet to be explored for broad-scale cultivation and consumption. This mini review focuses on the micronutrient and phytochemical profiles of a few emerging (pseudo)cereals and examines the current constraints of their integration into the global food system. Prospects of leveraging whole genome sequence information and modern breeding technologies to promote the breeding and accessibility of these crops are also discussed.

Keywords: cereal, pseudocereal, nutrition, micronutrient, phytochemical, biofortification, phytonutrient

OPEN ACCESS

Edited by:
Felipe Klein Ricachenevsky, Universidade Federal de Santa Maria, Brazil

Reviewed by:
Khurram Bashir, RIKEN, Japan
Simon Gerben Strobbe, Ghent University, Belgium

***Correspondence:**
Li Tian
ltian@ucdavis.edu

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *02 September 2019*
Accepted: *22 November 2019*
Published: *20 December 2019*

Citation:
Bekkering CS and Tian L (2019) Thinking Outside of the Cereal Box: Breeding Underutilized (Pseudo) Cereals for Improved Human Nutrition. Front. Genet. 10:1289. doi: 10.3389/fgene.2019.01289

INTRODUCTION

Among staple foods, cereal grains are advantaged for high starch content, relatively long-term storage capacity, and values as seed stocks. Wheat, rice, and maize constitute the major cereal crops that sustain over 50% of the caloric demand of the world population. Although these cereal grains make up a critical portion of many diets, they lack substantial amounts of micronutrients (vitamins and minerals) and phytonutrients (nutraceuticals and phytomedicines). Consequently, the hidden hunger due to micronutrient deficiency reportedly affects approximately 2 billion people globally (FAO, 2013), which raises the health concern regarding our heavy reliance on major cereal crops. To this end, multifaceted approaches including fortification, biofortification, and nutrient supplements have been deployed to ensure access to nutritious food, an important pillar of food security. On the other hand, some micronutrient and/or phytonutrient-rich (pseudo)cereal crops have historically taken on the role of a staple crop across many cultures, but are currently underutilized—having only percolated into small niches in the global food system. The present review examines the nutritional characteristics, cultivation, and germplasm collections of seven underutilized (pseudo)cereals. The limitations and opportunities for breeding and marketing these (pseudo)cereal grains for improving human nutrition are also discussed.

LOOKING BEYOND STAPLE CEREALS AND INTO THE NUTRIENT-DENSE UNDERUTILIZED (PSEUDO)CEREALS

Like wheat, rice, and maize, broomcorn millet (*Panicum miliaceum* L.), canary seed (*Phalaris canariensis* L.), and teff [*Eragrostis tef* (Zuccagni) Trotter] are monocotyledonous plants in the family of Poaceae (grasses) (**Figure 1A**; **Table 1**). Amaranth (*Amaranthus* spp.), buckwheat (*Fagopyrum esculentum* Moench.), chia (*Salvia hispanica* L.), and quinoa (*Chenopodium quinoa* Willd), despite having seeds resembling the cereal grains, do not belong to Poaceae and are considered pseudocereals (**Figure 1A**; **Table 1**). Currently, these (pseudo) cereals or grain products are used as breakfast cereals, snacks, additions to salads, processed foods, flour, and beverages, etc., but not a substantial source of calories.

Although these underutilized grains contain similar or lower starch contents than the staple cereal grains, they possess comparable or higher caloric values because decreases in carbohydrate content are offset by higher protein and lipid content (**Table 1**). The anatomy of the underutilized grains differs from the staple cereals in that they contain less endosperm (accumulating starch) and a higher proportion of embryos (accumulating proteins and lipids) (Prego et al., 1998; Valdivia-López and Tecante, 2015). It is noteworthy that higher caloric content, while a drawback in food systems of developed nations, is an asset in developing regions of the world where calorie deficiencies are a prevalent issue. Additionally, the higher protein content and more balanced amino acid composition of these underutilized grains is desirable. For instance, amaranth and quinoa grains are abundant in essential amino acids and showed a near optimal protein composition—one resembling that of cow milk (National Academy of Science, 1984). Bioactive peptides have also been found in amaranth and chia grains (Silva-Sánchez et al., 2008; Grancieri et al., 2019). Furthermore, the lack of gluten in these grains make them suitable for consumption by patients with coeliac disease.

With some exceptions, mineral nutrient content (potassium, phosphorus, magnesium, zinc, calcium, iron) of these underutilized grains is generally higher than that of their staple counterparts (**Table 1**). This discrepancy is as high as an order of magnitude in some cases (e.g., calcium in chia, amaranth, and teff). These underutilized grains are also more abundant in vitamins than white rice. In particular, grains of quinoa, canary seed, broomcorn millet, and amaranth are remarkably rich in folate (**Table 1**). Besides micronutrients, these underutilized grains also accumulate significant quantities of phenolic acids and flavonoids (phytonutrients) with antioxidant activities (Li et al., 2011; Gebremariam et al., 2014; Martínez-Cruz and Paredes-López, 2014; Zhang et al., 2014; Singh and Sharma, 2017; Tang and Tsao, 2017). However, some phytochemicals found in cereal grains (often located in husks), such as phytate, saponins, and tannins, are deemed as antinutrients because they tend to interfere with nutrient absorption and/or utilization.

CURRENT LIMITATIONS FOR DEVELOPING UNDERUTILIZED (PSEUDO)CEREALS

In spite of the potential advantages of more extensively leveraging these underutilized grain crops, several factors hinder the widespread incorporation of these crops into food systems and breeding regimes—factors that are bolstered by a relative lack of research into these crops (**Figure 1B**). These factors range from agronomical (growth acreage, yield potential), technological (trait improvement), social (knowledge diffusion), and economic (market buy-in), and have stark similarities regardless of the underutilized grain in question.

The agronomic potential of these underutilized grain crops is thus far poorly characterized. Grain crops grown outside of the plots of developed nations, such as quinoa, teff, chia, and amaranth, do not benefit from the high-input agriculture customary in the cultivation of major staple grains. As such, our knowledge of the yield and quality of these underutilized crops comes largely from low-input systems, limiting our ability to gauge their potential alongside major staple grains. Of the underutilized grains detailed here that benefit from high-input agriculture, such as broomcorn millet, buckwheat, and canary seed, their use is often constrained to that of a secondary crop—one grown to replace destroyed fields of staple crops or as a quick alternative to summer fallow. The short and less-than-optimal growth season allocated to these grains, while a sensible decision for a grower, hinders our ability to compare their yield and quality to their staple grain counterparts. This pattern of usage manifests in the low acreage of planting allocated to underutilized grains, magnitudes lower than major cereal crops (**Figure 1C**). Nevertheless, there was a gradual increase in quinoa production during the last few decades (**Figure 1C**).

Genetic limitations exist for some underutilized (pseudo) cereal crops. For instance, buckwheat is naturally cross-pollinated and exhibits self-incompatibility (Ueno et al., 2016). As such, it is necessary to develop self-compatible buckwheat lines for breeding and trait improvement. In addition, pipelines for mutagenesis and transformation are yet undeveloped and/or require optimization, resulting a reliance on natural variations for breeding in these grain crops. Currently, the intersection between genomics and breeding is also limited or nonexistent for these underutilized grains. Overall, underutilized grain crops are at present constrained by a lack of concerted breeding efforts committed to expanding their use in high-input agricultural systems.

Although cultivation and breeding knowledge exists in local communities for many of the underutilized grains noted here, the diffusion of this knowledge is often barred from reaching the broader global community of growers. In the case of chia, amaranth, teff, quinoa, and even buckwheat, both traditional knowledge and modern research trickles slowly across the language barrier into common languages used in global science and agriculture. This is especially pronounced for teff, where relevant information is commonly displayed only in Amharic.

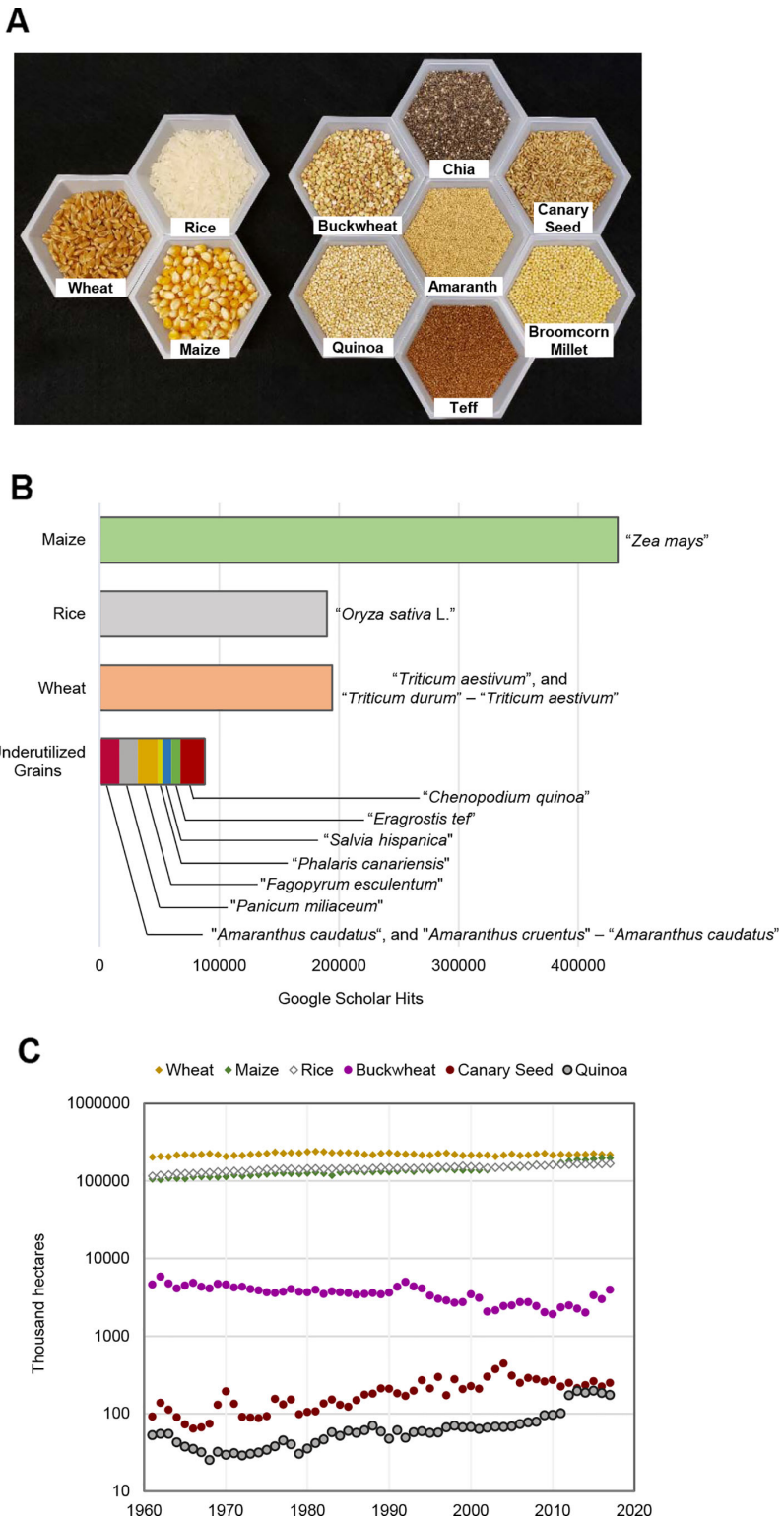


FIGURE 1 | (A) Image of major staple cereal grains and seven selected underutilized (pseudo)cereal grains. **(B)** Google scholar hits from 2018 and before using the search terms indicated in the panel. When two search terms were used for a (pseudo)cereal grain, the not operator (-) was used in conjunction with the second search term to exclude results that also contain the first search term; the hits from the two searches were added. **(C)** Global growth acreages of major staple cereal grains and three selected underutilized (pseudo)cereal grains from 1961 to 2017. Data shown are Food and Agriculture Organization of the United Nations (FAO) aggregated estimates.

TABLE 1 | Comparison of nutritional data and general characteristics of major staple cereal grains and seven selected underutilized (pseudo)cereal grains.

	Wheat	Maize	Rice	Broomcorn millet	Canary seed	Teff	Amaranth	Buckwheat	Chia	Quinoa
Nutritional data (per 100 g grain or grain flour)										
FDC ID	169761	170288	169756	169702	–	169747	170682	170687	170554	168874
Form Consumed	Wheat flour, white, all–purpose, unenriched	Whole grain, yellow	White, long–grain, unenriched	Whole grain	Whole groats	Whole grain	Whole grain	Whole–groat flour	Whole grain	Whole grain
Calories (kcal)	364	365	365	378	399	367	371	335	486	368
Carbohydrate (g)	76.31	74.26	79.95	72.85	68.7	73.13	65.25	70.59	42.12	64.16
Protein (g)	10.33	9.42	7.13	11.02	21.3	13.3	13.56	12.62	16.54	14.12
Total lipid (g)	0.98	4.74	0.66	4.22	6.7	2.38	7.02	3.1	30.74	6.07
Dietary Fiber (g)	2.7	7.3	1.3	8.5	6.2	8	6.7	10	34.4	7
Vitamin A, IU	0	214	0	0	–	9	2	0	54	14
Vitamin B–6 (mg)	0.044	0.622	0.164	0.384	–	0.482	0.591	0.582	–	0.487
Vitamin C (mg)	0	0	0	0	–	–	4.2	0	1.6	–
Vitamin E (mg)	0.06	0.49	0.11	0.05	–	0.08	1.19	0.32	0.5	2.44
Folate (µg)	26	19	8	85	100	–	82	54	49	184
Phosphorus (mg)	108	210	115	285	664	429	557	337	860	457
Potassium (mg)	107	287	115	195	400	427	508	577	407	563
Iron (mg)	1.17	2.71	0.8	3.01	6.6	7.63	7.61	4.06	7.72	4.57
Calcium (mg)	15	7	28	8	32	180	159	41	631	47
Zinc (mg)	0.7	2.21	1.09	1.68	3.7	3.63	2.87	3.12	4.58	3.1
Magnesium (mg)	22	127	25	114	216	184	248	251	335	197
General characteristics										
Group	Monocot	Monocot	Monocot	Monocot	Monocot	Monocot	Dicot	Dicot	Dicot	Dicot
Family	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Poaceae	Amaranthaceae	Polygonaceae	Lamiaceae	Amaranthaceae
Center of origin	Middle East	Southern Mexico	Asia	Northern China	Mediterranean	East Africa	Mexico and Central America	Central Asia and Siberia	Guatemala and southern Mexico	Peru and Bolivia
Photosynthesis	C_3	C_4	C_3	C_4	C_3	C_4	C_4	C_3	–	C_3
Sequenced genome	Yes (Appels et al., 2018)	Yes (Schnable et al., 2009)	Yes (Yu et al., 2002)	Yes (Zou et al., 2019)	No	Yes (Vanburen et al., 2019)	Yes (Clouse et al., 2016)	Yes (Zhang et al., 2017)	No	Yes (Jarvis et al., 2017)

Nutrition data in the uncooked, most commonly consumed form of the grains were obtained from the United States Department of Agriculture Food Data Central (https://fdc.nal.usda.gov/) with the exception of canary seed data, which were extracted from the nutritional factsheets published by Canaryseed Development Commission of Saskatchewan (https://www.canaryseed.ca/). Dashes indicate information not yet reported in the literature. FDC ID, Food Data Central identification number; IU, international unit.

While language barriers do not exist in excess for broomcorn millet and canary seed, the diffusion of information about their cultivation is inhibited by their niche in the market. As broomcorn millet is used as birdseed outside of East Asia, and canary seed almost ubiquitously so, their cultivation has been restricted to growers with connections to distributers in the birdseed market—a market already possessing a limited demand.

Except for quinoa, the noted underutilized grains have thus far received little media and market attention. Without a considerable marketing effort, investment in a farm-to-fork pipeline for underutilized grains may prove unfruitful. Quinoa serves as an example of a marketing success in this regard. The endorsement of quinoa as a functional grain crop by high-visibility public figures contributed to its global spike in cultivation—a spike that was aided by the integration of growers into the global marketplace (**Figure 1C**) (Bazile et al., 2016). Other underutilized grains would need to overcome their marketing constraints to bring their cultivation and consumption out of obscurity and to establish a stronger foothold in the global market.

PROMISES AND POTENTIAL FOR DEVELOPING UNDERUTILIZED (PSEUDO)CEREALS

The above-mentioned limitations present a clear avenue for development that could bring with it many fruitful possibilities —an avenue with promise substantiated by ongoing scientific progress on these underutilized grains. Although the yields of underutilized grains are generally lower than staple grains, this could at least be partially attributable to the fact that these grains are often grown on less arable land with fewer inputs (e.g., teff, quinoa, amaranth, chia) or are briefly grown as cover crop to avoid summer fallow (e.g., buckwheat, broomcorn millet). Therefore, allotment of suitable cropland and growing seasons to underutilized grain crops can uncover their yield potential relative to the grains that serve as the cornerstone of global research, development, and consumption.

The classic breeding methods remain applicable and valuable to these underutilized grains. Except for the limited germplasm collections for chia (Bochicchio et al., 2015) and canary seed (Cogliatti et al., 2011), there are over 3,000 accessions reported for quinoa (FAO, 2011; FAO, 2013), 5,000 accession for teff (Assefa et al., 2015), more than 10,000 accessions for buckwheat (Zhou et al., 2018), over 29,000 accessions for broomcorn millet (Vetriventhan et al., 2019), and at least 61 collection centers for amaranth (Das, 2016). Comprehensive evaluation and characterization of these germplasm collections will provide critical resources for breeding high-yield, elite crop varieties. To this end, next-generation sequencing technologies can be utilized to examine the genetic diversity of germplasms that have been adapted to different regions and production environments. Additionally, whole-genome sequencing (WGS) data of the germplasm collections encompass a broad range of genomic variants and can boost the power of genomic prediction (Hickey et al., 2017). Besides natural variations, the genetic diversity of the breeding population for the underutilized grains can be further enhanced through physical and chemical mutagenesis.

The emergence of genomic information for buckwheat, broomcorn millet, quinoa, amaranth, and teff pave the way for the development of breeding pipelines for desirable traits in the post-genomic era—pipelines that can integrate the emerging omics, phenotyping, and genome editing technologies (**Table 1**). These available reference genomes facilitate not only WGS in genotyping, but also discovery of genes, single nucleotide polymorphisms (SNPs), and genomic structural variants. High throughput genotyping coupled with high throughput phenotyping (phenomics) and crop modeling will enable acquisition of valuable trait data to assist in breeding. The genome sequences also allow precise and effective genome editing of target genes (Chen et al., 2019). With the exception of canary seed, there have been reports on genetic transformation and regeneration of these underutilized (pseudo)cereal plants (Jofre-Garfias et al., 1997; Eisa et al., 2005; Plaza-Wüthrich and Tadele, 2012; Gebre et al., 2013; Marconi et al., 2013; Suvorova, 2016). Although the efficiency of plant transformation remains to be optimized, it enables delivery of the genome-editing system to these crops. The underutilized (pseudo)cereals are reportedly tolerant/resistant to biotic and abiotic stresses that threaten crop production, such as insects, pathogens, weeds, drought, high temperature, UV-B radiation, heavy metal contamination, as well as salinity, alkalinity, acidic, or low fertility in soil (Assefa et al., 2015; Habiyaremye et al., 2017; Hinojosa et al., 2018). Genomic analyses have already begun to associate stress tolerance/resistance to molecular and physiological responses. Understanding the underlying mechanisms of stress tolerance in these underutilized cereals will also be useful for breeding other agronomically and economically important crops.

There is promise for these underutilized (pseudo)cereals in the marketing sector as well. The success of quinoa in being marketed as a functional grain crop with a rich history has laid the groundwork for other grain marketers to follow suit. Even outside the grains, functional foods are increasingly sought after in global markets, with clear parallels being visible in the western markets of avocado, kale, pomegranate, and wine. Globally, marketing these grain crops as a nutritious source of carbohydrates could promote their import to developing regions—a treatment that even quinoa could benefit from. Marketing and subsequent supply chain reconfiguration should of course proceed such that local demand for the traditional crops is still satisfied, as a leading criticism of the rapid adoption of quinoa was the resulting lapse in quinoa consumption by the locals that had depended on it for generations (Friedman-Rudovsky, 2012).

Promises in marketing are substantiated by the fact that most of these grains are sold and consumed either as whole grains or whole grain products (**Table 1**), a form that retains the nutrient content (without losing it to postharvest processing) and fits easily into the functional food space of western markets. The exceptions to this are buckwheat, canary seed, and broomcorn

millet. Buckwheat is almost exclusively sold as dehulled grouts and flour produced from dehulled groats. Canary seed's potential in the human diet has been elucidated thus far for dehulled groats also (Abdel-Aal et al., 2011; Mason et al., 2018), while broomcorn millet is occasionally sold as white flour. The additional processing steps reduced the mineral nutrient content in dehusked buckwheat grains as observed similarly in the major cereals (Oghbaei and Prakash, 2013; Pandey et al., 2015). On the other hand, these additional steps in processing could have unstudied roles in removing antinutritional phytochemicals from these grains as well, much like the saponin removal steps in quinoa production (Jarvis et al., 2017). Thorough examination of the role that postharvest processing could have for antinutrient mitigation in other underutilized grain crops could aid in their wider application while simultaneously providing yet another selling point to leverage in marketing.

PERSPECTIVES

Although it is not envisioned that the underutilized grains will play a major role as food staple in the near future, an expansion of their cultivation and utilization will build nutritional synergy with the major cereal grains. Climates non-conducive to staple crop cultivation such as hot semi-arid, subtropical highland, and arid subtropical could be leveraged for food production, contingent on investment in the biology and marketing of these underutilized crops. A diversity of photosynthetic modes in the underutilized grains substantiates this potential for broader cultivation—with the existence of C_4 species removing the need for extensive engineering efforts such as those carried out in rice (**Table 1**).

By leveraging available germplasm collections and expanding genetic resources, climate-adapted elite varieties can be developed for the underutilized grain crops. Increased understanding of the genetic underpinnings of many plant traits such as lodging resistance, seed size, grain shattering, and stress tolerance/resistance, as well as development of advanced techniques such as mechanical harvesting, food processing, and postharvest storage, will bring into focus clear avenues for improvement of underutilized grains that can be pursued through strategic plant breeding. In addition, making innovative breeding technologies and integrated plant breeding platforms accessible to local breeders and small farmers is essential for implementation of these breeding strategies. Furthermore, international collaborations and partnerships, such as the African Orphan Crops Consortium (AOCC) (Hendre et al., 2019), will accelerate the development of the climate-resilient (pseudo)cereals. Overall, complementary to biofortification of major cereal grains, better utilization of underutilized grains in diet will have far-reaching impact on alleviating the burden of the hidden hunger crisis.

AUTHOR CONTRIBUTIONS

CB and LT conceived and wrote the review.

FUNDING

We thank USDA-NIFA (grant 2017-67013-26164 to LT) for supporting our work on provitamin A biofortification of wheat grains.

ACKNOWLEDGMENTS

We thank Kevin Abuhanna for his assistance in preparing the image of grains and careful reading of the nutrient table.

REFERENCES

Abdel-Aal, E.-S. M., Hucl, P., Shea Miller, S., Patterson, C. A., and Gray, D. (2011). Microstructure and nutrient composition of hairless canary seed and its potential as a blending flour for food use. *Food Chem.* 125, 410–416. doi: 10.1016/j.foodchem.2010.09.021

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

Assefa, K., Cannarozzi, G., Girma, D., Kamies, R., Chanyalew, S., Plaza-Wüthrich, S., et al. (2015). Genetic diversity in tef [*Eragrostis tef* (Zucc.) Trotter]. *Front. Plant Sci.* 6, 177. doi: 10.3389/fpls.2015.00177

Bazile, D., Jacobsen, S.-E., and Verniau, A. (2016). The global expansion of quinoa: trends and limits. *Front. Plant Sci.* 7, 622. doi: 10.3389/fpls.2016.00622

Bochicchio, R., Philips, T. D., Lovelli, S., Labella, R., Galgano, F., Di Marisco, A., et al. (2015). "Innovative crop productions for healthy food: the case of chia (*Salvia hispanica* L.)," in *The sustainability of agro-food and natural resource systems in the Mediterranean basin*. Ed. A. Vastola (Cham: Springer International Publishing), 29–45.

Chen, K., Wang, Y., Zhang, R., Zhang, H., and Gao, C. (2019). CRISPR/Cas genome editing and precision plant breeding in agriculture. *Annu. Rev. Plant Biol.* 70, 667–697. doi: 10.1146/annurev-arplant-050718-100049

Clouse, J. W., Adhikary, D., Page, J. T., Ramaraj, T., Deyholos, M. K., Udall, J. A., et al. (2016). The amaranth genome: genome, transcriptome, and physical map assembly. *Plant Genome* 9, 1–14. doi: 10.3835/plantgenome2015.070062

Cogliatti, M., Bongiorno, F., Dalla Valle, H., and Rogers, W. J. (2011). Canaryseed (*Phalaris canariensis* L.) accessions from nineteen countries show useful genetic variation for agronomic traits. *Can. J. Plant Sci.* 91, 37–48. doi: 10.4141/CJPS09200

Das, S. (2016). "Distribution and maintenance of amaranth germplasm worldwide," in *Amaranthus: a promising crop of future* (Singapore: Springer Singapore), 99–106.

Eisa, S., Koyro, H. W., Kogel, K. H., and Imani, J. (2005). Induction of somatic embryogenesis in cultured cells of *Chenopodium quinoa*. *Plant Cell Tissue Organ Cult.* 81, 243–246. doi: 10.1007/s11240-004-4793-z

FAO. (2011). "Genetic diversity, varieties and germplasm banks," in *Quinoa, an ancient crop to contribute to world food security* (Santiago, Chile: Regional Office for Latin America and the Caribbean), 16–20.

FAO. (2013). The state of food and agriculture. *Food Agric. Organ.*

Friedman-Rudovsky, J. (2012). "Quinoa: the dark side of an Andean superfood," in *Time*. (New York City, New York: Time Inc.).

Gebre, E., Gugsa, L., Schlüter, U., and Kunert, K. (2013). Transformation of tef (*Eragrostis tef*) by Agrobacterium through immature embryo regeneration system for inducing semi-dwarfism. *S. Afr. J. Bot.* 87, 9–17. doi: 10.1016/j.sajb.2013.03.004

Gebremariam, M. M., Zarnkow, M., and Becker, T. (2014). Teff (*Eragrostis tef*) as a raw material for malting, brewing and manufacturing of gluten-free foods and beverages: a review. *J. Food Sci. Technol.* 51, 2881–2895. doi: 10.1007/s13197-012-0745-5

Grancieri, M., Martino, H. S. D., and Gonzalez De Mejia, E. (2019). Chia seed (*Salvia hispanica* L.) as a source of proteins and bioactive peptides with health benefits: a review. *Compr. Rev. Food Sci. F.* 18, 480–499. doi: 10.1111/1541-4337.12423

Habiyaremye, C., Matanguihan, J. B., D'alpoim Guedes, J., Ganjyal, G. M., Whiteman, M. R., Kidwell, K. K., et al. (2017). Proso millet (*Panicum miliaceum* L.) and its potential for cultivation in the Pacific Northwest, U.S.: a review. *Front. Plant Sci.* 7, 1961. doi: 10.3389/fpls.2016.01961

Hendre, P. S., Muthemba, S., Kariba, R., Muchugi, A., Fu, Y., Chang, Y., et al. (2019). African Orphan Crops Consortium (AOCC): status of developing genomic resources for African orphan crops. *Planta* 250, 989–1003. doi: 10.1007/s00425-019-03156-9

Hickey, J. M., Chiurugwi, T., Mackay, I., Powell, W.Implementing Genomic Selection In, C.B.P.W.P., , Hickey, J. M., et al. (2017). Genomic prediction unifies animal and plant breeding programs to form platforms for biological discovery. *Nat. Genet.* 49, 1297. doi: 10.1038/ng3920

Hinojosa, L., González, J. A., Barrios-Masias, F. H., Fuentes, F., and Murphy, K. M. (2018). Quinoa abiotic stress responses: a review. *Plants* 7, 106. doi: 10.3390/plants7040106

Jarvis, D. E., Ho, Y. S., Lightfoot, D. J., Schmöckel, S. M., Li, B., Borm, T. J. A., et al. (2017). The genome of *Chenopodium quinoa*. *Nature* 542, 307–312. doi: 10.1038/nature21370

Jofre-Garfias, A. E., Villegas-Sepúlveda, N., Cabrera-Ponce, J. L., Adame-Alvarez, R. M., Herrera-Estrella, L., and Simpson, J. (1997). Agrobacterium-mediated transformation of *Amaranthus hypochondriacus*: light- and tissue-specific expression of a pea chlorophyll a/b-binding protein promoter. *Plant Cell Rep.* 16, 847–852. doi: 10.1007/s002990050332

Li, W., Qiu, Y., Patterson, C. A., and Beta, T. (2011). The analysis of phenolic constituents in glabrous canaryseed groats. *Food Chem.* 127, 10–20. doi: 10.1016/j.foodchem.2010.12.033

Marconi, P., López, M., Meester, J. D., Bovjin, C., and Alvarez, M. (2013). *In vitro* establishment of *Salvia hispanica* L. plants and callus. *Biotecnología Vegetal.* 13, 203–207.

Martínez-Cruz, O., and Paredes-López, O. (2014). Phytochemical profile and nutraceutical potential of chia seeds (*Salvia hispanica* L.) by ultra high performance liquid chromatography. *J. Chromatogr. A* 1346, 43–48. doi: 10.1016/j.chroma.2014.04.007

Mason, E., L'hocine, L., Achouri, A., and Karboune, S. (2018). Hairless canaryseed: a novel cereal with health promoting potential. *Nutrients* 10, 1327. doi: 10.3390/nu10091327

National Academy of Science. (1984). *Amaranth: modern prospects for an ancient crop* (Washington DC: National Academy Press).

Oghbaei, M., and Prakash, J. (2013). Effect of fractional milling of wheat on nutritional quality of milled fractions. *Trends Carbohyd. Res.* 5, 53–58.

Pandey, S., Senthil, A., and Fatema, K. (2015). Effect of hydrothermal treatment on the nutritional and functional properties of husked and dehusked buckwheat. *J. Food Process. Technol.* 6, 7. doi: 10.4172/2157-7110.1000461

Plaza-Wüthrich, S., and Tadele, Z. (2012). Millet improvement through regeneration and transformation. *Biotech. Mol. Biol. Rev.* 7, 48–61. doi: 10.5897/BMBR12.001

Prego, I., Maldonado, S., and Otegui, M. (1998). Seed structure and localization of reserves in *Chenopodium quinoa*. *Ann. Bot.* 82, 481–488. doi: 10.1006/anbo.19980704

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 326, 1112–1115. doi: 10.1126/science.1178534

Silva-Sánchez, C., De La Rosa, A. P. B., León-Galván, M. F., De Lumen, B. O., De León-Rodríguez, A., and De Mejía, E. G. (2008). Bioactive peptides in Amaranth (*Amaranthus hypochondriacus*) seed. *J. Agric. Food Chem.* 56, 1233–1240. doi: 10.1021/jf072911z

Singh, A., and Sharma, S. (2017). Bioactive components and functional properties of biologically activated cereal grains: a bibliographic review. *Crit. Rev. Food Sci. Nutr.* 57, 3051–3071. doi: 10.1080/10408398.2015.1085828

Suvorova, G. (2016). "Buckwheat tissue cultures and genetic transformation," in *Molecular breeding and nutritional aspects of buckwheat*. Eds. M. Zhou, I. Kreft, S.-H. Woo, N. Chrungoo and G. Wieslander (Cambridge, MA, USA: Academic Press), 365–375.

Tang, Y., and Tsao, R. (2017). Phytochemicals in quinoa and amaranth grains and their antioxidant, anti-inflammatory, and potential health beneficial effects: a review. *Mol. Nutr. Food Res.* 61, 1600767. doi: 10.1002/mnfr.201600767

Ueno, M., Yasui, Y., Aii, J., Matsui, K., Sato, S., and Ota, T. (2016). "Genetic analyses of the heteromorphic self-incompatibility (S) locus in buckwheat," in *Molecular breeding and nutritional aspects of buckwheat*. Eds. M. Zhou, I. Kreft, S.-H. Woo, N. Chrungoo and G. Wieslander (Cambridge, MA: Academic Press), 411–421.

Valdivia-López, M.Á., and Tecante, A. (2015). Chia (*Salvia hispanica*): a review of native Mexican seed and its nutritional and functional properties. *Adv. Food Nutr. Res.* 75, 53–75. doi: 10.1016/bs.afnr.2015.06.002

Vanburen, R., Wai, C. M., Pardo, J., Yocca, A. E., Wang, X., Wang, H., et al. (2019). Exceptional subgenome stability and functional divergence in allotetraploid teff, the primary cereal crop in Ethiopia. *bioRxiv* 580720. doi: 10.1101/580720

Vetriventhan, M., Azevedo, V. C. R., Upadhyaya, H. D., and Naresh, D. (2019). Variability in the global proso millet (*Panicum miliaceum* L.) germplasm collection conserved at the ICRISAT genebank. *Agriculture* 9, 112. doi: 10.3390/agriculture9050112

Yu, J., Hu, S., Wang, J., Wong, G. K., Li, S., Liu, B., et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science* 296, 79–92. doi: 10.1126/science.1068037

Zhang, L., Liu, R., and Niu, W. (2014). Phytochemical and antiproliferative activity of proso millet. *PLoS One* 9, e104058. doi: 10.1371/journal.pone.0104058

Zhang, L., Li, X., Ma, B., Gao, Q., Du, H., Han, Y., et al. (2017). The tartary buckwheat genome provides insights into rutin biosynthesis and abiotic stress tolerance. *Mol. Plant* 10, 1224–1237. doi: 10.1016/j.molp.2017.08.013

Zhou, M., Tang, Y., Deng, X., Ruan, C., Kreft, I., Tang, Y., et al. (2018). "Overview of buckwheat resources in the world," in *Buckwheat germplasm in the world*. Eds. M. Zhou, I. Kreft, G. Suvorova, Y. Tang and S. H. Woo (Cambridge, MA, USA: Academic Press), 1–7.

Zou, C., Li, L., Miki, D., Li, D., Tang, Q., Xiao, L., et al. (2019). The genome of broomcorn millet. *Nat. Commun.* 10, 436. doi: 10.1038/s41467-019-08409-5

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 Genetics

ORIGINAL RESEARCH
published: 31 January 2020
doi: 10.3389/fgene.2020.00013

Pleiotropic Effects of Rice Florigen Gene *RFT1* on the Amino Acid Content of Unmilled Rice

Li-Hong Xie†, Yu-Jun Zhu†, Shao-Qing Tang, Xiang-Jin Wei, Zhong-Hua Sheng, Gui-Ai Jiao, Pei-Song Hu and Jie-Yun Zhuang**

State Key Laboratory of Rice Biology/Chinese National Center for Rice Improvement, China National Rice Research Institute, Hangzhou, China

OPEN ACCESS

Edited by:
Balram Marathi, Professor Jayashankar Telangana State Agricultural University, India

Reviewed by:
Ashok Kumar Singh, Indian Agricultural Research Institute (ICAR), India
Kazuki Matsubara, Institute of Crop Science (NARO), Japan

***Correspondence:**
Pei-Song Hu
qualityh@163.com
Jie-Yun Zhuang
zhuangjieyun@caas.cn

†*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *22 July 2019*
Accepted: *06 January 2020*
Published: *31 January 2020*

Citation:
Xie L-H, Zhu Y-J, Tang S-Q, Wei X-J, Sheng Z-H, Jiao G-A, Hu P-S and Zhuang J-Y (2020) Pleiotropic Effects of Rice Florigen Gene RFT1 on the Amino Acid Content of Unmilled Rice. Front. Genet. 11:13. doi: 10.3389/fgene.2020.00013

In rice, the contents of protein and amino acids are the major parameters of nutritional quality. Co-localization of quantitative trait loci (QTLs) for heading date and protein content were reported, but pleiotropism of heading-date genes on protein contents has not been investigated. Here, we reported that rice florigen gene *RFT1* plays an important role in controlling amino acid contents of rice grain. Firstly, 73 QTLs for the contents of 17 amino acids in unmilled rice were detected using recombinant inbred lines (RILs) of the *indica* rice cross Zhenshan 97 (ZS97)/Milyang 46 (MY46). Then, the effect of the largest cluster consisting of 14 QTLs, located in proximity to the rice florigen genes *RFT1* and *Hd3a*, was validated using three populations consisting of near isogenic lines (NILs) that only segregated a region covering the target QTL. The first and second NIL populations were derived from a residual heterozygote identified from the ZS97/MY46 RIL population, consisting of homozygous lines that were only segregated in a 29.9-kb region covering the two florigen genes and a 1.7-kb region for *RFT1*, respectively. The third NIL population was segregated for the $RFT1^{ZS97}$ transgene in the background of *japonica* rice cultivar Zhonghua 11. In all the three NIL populations, *RFT1* was shown to have a strong effect on the contents of most amino acids, with the ZS97 allele always having the reducing effects. By comparing QTLs for amino acid contents detected in the ZS97/MY46 RIL population and genes/QTLs previously identified for heading date difference between ZS97 and MY46, possible pleiotropism on amino acid contents was also shown for other key heading-date genes including *Hd1*, *Ghd7*, and *OsGI*.

Keywords: amino acid content, heading date, near isogenic line, pleiotropic effect, quantitative trait locus, *Oryza sativa* L.

INTRODUCTION

Rice sustains about half of the world's population, providing a source of energy and protein. Protein content (PC) of the rice grain is influenced by both genotype and growing environment. The PC values in the un-milled (brown) and milled rice of a large collection of rice cultivars were found to range as 5.6–11.2 and 6.0–15.7%, respectively, with a high correlation coefficient of 0.96 between the two measurements (Chen et al., 2006). In addition to the quantity of total protein, contribution of

different amino acids is also an important factor determining the nutritional value of rice grain (Wang et al., 2008). An understanding of the genetic basis underlying the variation of the total grain protein content and the contribution of individual amino acids has the potential to facilitate the breeding of rice cultivars having nutritionally superior grain.

A number of attempts have been made to identify the genetic architecture of the spectrum of grain amino acids in rice by means of quantitative trait locus (QTL) analysis. Using recombinant inbred lines (RILs) derived from a cross between rice cultivars Zhenshan 97 (ZS97) and Nanyangzhan, 18 QTL clusters for 19 components of the amino acid content (AAC) in milled rice were identified (Wang et al., 2008). In two other RIL populations derived from crosses using ZS97 as the female parent, ZS97/Minghui 63 (Lu et al., 2009) and ZS97/Delong 208 (Zhong et al., 2011), 5 and 29 QTL regions for 10 and 17 components of AAC in milled rice were detected, respectively. Many more studies examined the total protein content without analysis on individual amino acids. A region covering the *Wx* locus on the short arm of chromosome 6 was frequently found to be associated with PC (Tan et al., 2001; Aluko et al., 2004; Wada et al., 2006; Yu et al., 2009; Kashiwagi and Munakata, 2018), but its influence on AAC was only occasionally observed (Wang et al., 2008; Lu et al., 2009; Zhong et al., 2011). Whether the *Wx* gene itself or other linked genes are involved in the genetic control of PC and AAC remains to be determined.

In rice, the short arm of chromosome 6 is a region harboring multiple genes that play critical roles in the regulation of heading date (HD), including *Hd1*, *Hd3a*, *RFT1*, and *Hd17*/*Hd3b* (Hori et al., 2016). In a number of segregating populations that derived from intra-subspecies or inter-species crosses, negative correlation between HD and PC was observed (Wada et al., 2006; Kwon et al., 2011; Yun et al., 2016), which could be partially ascribed to a QTL region on the short arm of chromosome 6 that affected both HD and PC with opposite allelic directions (Wada et al., 2006; Yun et al., 2016). These results suggest that one or more heading-date genes located in this region may have pleiotropic effects on the contents of proteins and amino acids. In the present study, QTL analysis for 17 components of AAC in unmilled rice was performed using the ZS97/Milyang 46 (MY46) RIL population, followed by the validation of a QTL cluster on the short arm of chromosome 6 using three populations of near isogenic lines (NILs) in either *indica* or *japonica* backgrounds. A total of 73 QTLs were detected, and the *RFT1* gene was found to have a strong and stable pleiotropic effect on AAC.

MATERIAL AND METHODS

Plant Materials

Four segregating populations of rice (*Oryza sativa* L.) were used in this study. One was a primary mapping population consisting of 247 RILs developed from a cross between *indica* rice cultivars ZS97 and MY46. The other three were NIL populations segregating a region involving the *RFT1* gene in an isogenic background, all of which have been reported by Zhu et al. (2017). Two of the NIL populations, namely TF6-15 and R1, were derived from a residual heterozygote (RH) of ZS97/MY46. An F_{10} plant was selected, which was heterozygous in a 29.9-kb region covering the *RFT1* and *Hd3a* loci and homozygous in other regions. The S_1 plants were assayed with DNA markers located in the segregating region. Homozygous plants were selfed to produce NILs. The TF6-15 population was established, consisting of 10 lines of ZS97 homozygotes and 10 lines of MY46 homozygotes differing in the 29.9 kb region. New RHs were identified from heterozygous progeny of the F_{10} plant. An F_{14} plant that was heterozygous at the *RFT1* locus only was selected. The R1 population was constructed, comprising 20 lines of ZS97 homozygotes and 20 lines of MY46 homozygotes differing for the *RFT1* gene only. The remaining NIL population consisted of 28 homozygous transgenic lines in the genetic background of *japonica* rice cultivar Zhonghua 11 (ZH11), of which 14 lines carried the $RFT1^{ZS97}$ transgene and the others carried no transgene.

Field Experiments

The four populations were planted in the single-cropping rice season at the China National Rice Research Institute in Hangzhou, Zhejiang, China. The RIL population was raised in 2015, 2016, and 2017, and the other three populations were grown in 2017 only. Each line was represented by 12 plants per row, with an inter-plant spacing of 16.7 cm and an inter-row spacing of 26.7 cm. The RILs were grown without replication, while the other populations were represented by two replicates. Plants were managed using standard agricultural practice.

Determination of Amino Acid Content

Grain was bulked from the middle ten plants of each row and dried to a moisture content of ~12%. De-hulling was achieved using THU-35A testing husker (Satake Engineering Co. Ltd., Hiroshima, Japan). The dehusked grain was ground using a Cyclotec 1093 Sample Mill (Tecator, Hoganas, Sweden) and the resulting flour was passed through a 0.42 mm sieve. A 250 mg batch of each flour sample was sealed in a vial containing 10 ml 6.0 M HCl and held at 110°C for 24 h. The resulting hydrolysate was diluted to 50 ml with deionized water and filtered. A 0.2 ml aliquot of the filtrate was transferred to a 2 ml tube and evaporated down to ~0.1 ml by bubbling nitrogen gas. After the addition of 2 ml 20 mM HCl, the solution was passed through a 0.2 μm Acrodisc membrane (Pall Corp., Port Washington, NY, USA). The amino acid content of each sample was acquired using a L8900 amino acid auto analyzer (Hitachi, High-Technology Corporation, Tokyo, Japan). Percentage contributions of each of the 17 amino acids were obtained using Ezchrom Elite software (High-Technology Corporation, Tokyo, Japan).

The amino acids studied were aspartic acid (Asp), threonine (Thr), serine (Ser), glutamic acid (Glu), glycine (Gly), alanine (Ala), cystine (Cys), valine (Val), methionine (Met), isoleucine (Ile), leucine (Leu), tyrosine (Tyr), phenylalanine (Phe), lysine (Lys), histidine (His), arginine (Arg), and proline (Pro). Each sample was analyzed in triplicate. Every analytical batch included

a blank (20 mM HCl only) and a GBW (E) 100010 reference sample (Chinese National Research Center for Certified Reference Materials, Beijing, China), except that the reference samples for Asp and Cys were not available. Concentrations in the standard samples of Ile, Leu, Tyr, Phe, Lys, His, Arg, Thr, Ser, Glu, Pro, Gly, Ala, Met, and Val measured based on six replicates were 0.38 ± 0.060, 0.74 ± 0.040, 0.30 ± 0.020, 0.52 ± 0.068, 0.28 ± 0.012, 0.22 ± 0.012, 0.49 ± 0.048, 0.30 ± 0.014, 0.49 ± 0.014, 3.28 ± 0.18, 0.95 ± 0.010, 0.40 ± 0.048, 0.503 ± 0.026, 0.165 ± 0.046, and 0.50 ± 0.024%, respectively, matching well with the Certified Reference Material's values. The day-to-day reproducibility of the assay checked over three days was satisfactory, having the relative standard deviation lower than 5.0% for all the fifteen amino acids.

Quantitative Trait Locus Mapping

Genetic map of the ZS97/MY46 RIL population was previously constructed, consisting of 256 markers and spanning 1,814.7 cM (Wang et al., 2017b). This map was applied for QTL analysis using composite interval mapping (CIM) and multiple interval mapping (MIM) in Windows QTL Cartographer v2.5 (Wang et al., 2011). A candidate QTL was identified with CIM using a threshold of logarithm of odds (LOD) > 2.0 and then evaluated with MIM using the Bayesian Information Criterion c (n) = ln (n). A putative QTL was claimed if it satisfied both criteria. QTLs were designated as proposed by McCouch and CGSNL (2008). Two-way analysis of variance (ANOVA) was conducted to test the differences between the two genotypic groups in each of the three NIL populations, using a general linear model (GLM) of the SAS Program as described by Dai et al. (2008).

RESULTS

Variation of Amino Acid Contents in the ZS97/MY46 Recombinant Inbred Line Population

Variation of the 17 components of AAC in unmilled rice of the ZS97/MY46 RIL population is summarized in **Table 1**. There was a strong evidence for transgressive segregation in both directions for all the amino acids except His of which the contents of the RILs were all lower than the high-parental value in 2016. It was also shown that differences between the two parental lines varied greatly across the 3 years. Among the contents of the 17 amino acids, ZS97 was found to be the high-value parent for two amino acids only (Phe and His) in 2015, but ZS97 was the high-value parent for seven amino acids (Ser, Glu, Leu, Tyr, Phe, Lys, and His) in 2016 and for 14 amino acids (Asp, Thr, Ser, Glu, Gly, Ala, Val, Ile, Leu, Tyr, Phe, Lys, His, and Arg) in 2017.

Pearson correlation coefficients between the 17 components of AAC were calculated using mean values over 3 years and data of each year. Family error rates were controlled by dividing the P value of 0.05 by 17, thus a threshold of $P < 0.003$ was used for declaring a significant correlation. It was found that a large majority of the correlations were positively significant. Of the 136 estimates produced from the mean values, 114 were positively significant, 4 were negatively significant, and 18 were non-significant (**Table 2**). The four negative correlations occurred between Ala and Gly, Cys, Met, and Pro. The 18 non-significant correlations included nine between Ala and others (Ser, Glu, Val, Ile, Leu, Tyr, Phe, Lys, and Arg), 8 between Cys and others (Asp, Thr, Gly, Val, Met, Lys, His, and Pro), and 1 between Tyr and Met. Common occurrence of significantly positive correlations between different AAC components were also observed when data of each year were used (**Supplementary Table S1**). In 2015, 101 correlations were positively significant and the other 35 were non-significant. In 2016, 122 correlations were positively significant and the other 14 were non-significant. In 2017, 118 correlations were positively significant, three were negatively significant, and 15 were non-significant.

Quantitative Trait Loci for Amino Acid Contents Detected in the ZS97/MY46 Recombinant Inbred Line Population

A total of 73 QTLs were detected based on 3-year's data of the 17 components of AAC in the ZS97/MY46 RIL population (**Supplementary Table S2**). Of these QTLs, seven were identified in all the 3 years, eight were found in 2 years, and the others were detected in 1 year only. The number of QTLs detected for each amino acid ranged from two to eight, with the proportion of the variance explained (R^2) by a single QTL ranging from 2.2 to 35.9%. These QTL were distributed over all the 12 rice chromosomes except chromosome 5 (**Figure 1**). Most of them were located in cluster, with chromosomes 1, 6, and 7 harboring the highest number of loci. Of the 15 QTLs detected in 2 or 3 years, 12, 2, and 1 were located in chromosomes 6, 7, and 11, respectively. Except *qThr11*, allelic directions of these QTLs all remained consistent across different years.

Fourteen QTLs were located in the RM190–RM6917 region on the short arm of chromosome 6, forming the largest cluster in terms of QTL number. Included were seven QTLs detected in three years (*qAsp6*, *qSer6*, *qGly6*, *qLeu6*, *qPhe6*, *qHis6,* and *qArg6*), five QTLs detected in 2 years (*qThr6*, *qGlu6*, *qVal6*, *qMet6,* and *qTyr6*), and two QTLs detected in 1 year (*qLys6* and *qPro6.1*). Enhancing alleles of these QTLs were all derived from the male parent MY46, with *qGly6* having the highest R^2 of 33.9%. Two other QTLs (*qAla6* and *qPro6.2*) were detected in nearby intervals RM253–RM276 and RZ667–RM19784, respectively, of which the enhancing alleles were both derived from the female parent ZS97. Altogether, 16 QTLs were detected on chromosome 6.

The second largest cluster consisting of nine QTLs was located in the RM3325–RM3859 region on the short arm of chromosome 7. Included were two QTLs detected in 2 years (*qAsp7* and *qArg7*) and seven QTLs detected in 1 year (*qThr7*, *qSer7*, *qGly7.1*, *qAla7*, *qVal7*, *qTyr7.1*, and *qPhe7*). Enhancing alleles of these QTLs were all derived from ZS97, with *qAsp7* having the highest R^2 of 19.8%. Five other QTLs were clustered in

TABLE 1 | Phenotype performance of 17 components of amino acid content in the ZS97/MY46 recombinant inbred line (RIL) population.

Trait[a]		Parental mean		RIL Population				
		ZS97	MY46	Mean	Range	SD	Skewness	Kurtosis
Asp	2015	0.848	1.048	0.904	0.632–1.302	0.101	0.359	0.794
	2016	0.819	0.851	0.849	0.629–1.012	0.072	–0.053	–0.050
	2017	1.013	0.904	0.859	0.424–1.163	0.096	–0.341	2.049
Thr	2015	0.354	0.403	0.372	0.265–0.520	0.042	–0.087	0.385
	2016	0.341	0.357	0.359	0.273–0.433	0.029	0.009	0.127
	2017	0.427	0.386	0.369	0.243–0.490	0.034	0.133	0.877
Ser	2015	0.500	0.544	0.517	0.353–0.741	0.062	0.087	0.369
	2016	0.490	0.481	0.482	0.367–0.585	0.039	–0.080	–0.020
	2017	0.540	0.481	0.467	0.295–0.637	0.046	0.141	1.037
Glu	2015	1.782	1.877	1.812	1.181–2.451	0.225	0.034	0.099
	2016	1.757	1.683	1.644	1.203–1.990	0.161	–0.094	–0.565
	2017	1.741	1.554	1.619	1.075–2.242	0.161	0.358	1.111
Gly	2015	0.233	0.273	0.235	0.170–0.360	0.027	0.0004	0.372
	2016	0.236	0.247	0.216	0.156–0.262	0.018	–0.249	0.194
	2017	0.218	0.199	0.269	0.161–0.585	0.128	1.344	0.005
Ala	2015	0.218	0.249	0.247	0.155–0.331	0.027	0.531	1.183
	2016	0.221	0.221	0.240	0.183–0.354	0.020	0.874	4.783
	2017	0.218	0.199	0.422	0.171–0.665	0.122	–0.905	–0.520
Cys	2015	0.396	0.522	0.561	0.245–0.862	0.105	0.058	0.347
	2016	0.556	0.702	0.507	0.125–0.849	0.253	–0.542	–1.530
	2017	0.268	0.272	0.302	0.213–0.423	0.035	0.730	0.580
Val	2015	0.420	0.483	0.458	0.304–0.618	0.062	0.068	–0.351
	2016	0.425	0.482	0.441	0.331–0.542	0.038	0.084	–0.249
	2017	0.439	0.406	0.422	0.296–0.548	0.041	0.354	0.587
Met	2015	0.030	0.068	0.053	0.000–0.203	0.031	1.086	2.576
	2016	0.033	0.110	0.060	0.005–0.172	0.029	0.640	0.571
	2017	0.072	0.103	0.029	0.000–0.121	0.031	0.852	–0.298
Ile	2015	0.256	0.308	0.287	0.198–0.440	0.040	0.788	1.229
	2016	0.257	0.299	0.290	0.220–0.356	0.027	–0.036	–0.205
	2017	0.275	0.248	0.269	0.184–0.363	0.028	0.464	0.628
Leu	2015	0.734	0.815	0.762	0.534–1.085	0.086	0.452	0.452
	2016	0.743	0.724	0.703	0.502–0.856	0.065	–0.148	–0.121
	2017	0.670	0.615	0.658	0.420–0.927	0.071	0.449	0.958
Tyr	2015	0.352	0.353	0.273	0.133–0.415	0.075	–0.117	–1.103
	2016	0.359	0.282	0.272	0.090–0.362	0.033	0.010	0.407
	2017	0.184	0.175	0.211	0.144–0.351	0.034	0.950	1.460
Phe	2015	0.592	0.578	0.546	0.367–0.802	0.077	0.548	0.711
	2016	0.598	0.491	0.488	0.336–0.629	0.049	0.023	0.298
	2017	0.411	0.396	0.434	0.292–0.670	0.060	0.784	1.660
Lys	2015	0.369	0.378	0.326	0.216–0.471	0.052	0.652	0.129
	2016	0.370	0.356	0.340	0.258–0.419	0.030	0.251	0.192
	2017	0.306	0.293	0.319	0.250–0.433	0.034	0.641	0.712
His	2015	0.305	0.298	0.237	0.139–0.393	0.048	1.236	1.654
	2016	0.304	0.222	0.221	0.156–0.288	0.023	0.290	0.235
	2017	0.235	0.215	0.225	0.180–0.292	0.021	0.435	0.041
Arg	2015	0.617	0.769	0.658	0.421–0.993	0.077	0.480	1.355
	2016	0.627	0.659	0.631	0.458–0.774	0.057	0.108	0.276
	2017	0.657	0.559	0.585	0.443–0.782	0.065	0.323	–0.159
Pro	2015	0.402	0.468	0.274	0.098–0.499	0.093	0.445	–0.740
	2016	0.189	0.234	0.269	0.006–0.536	0.076	0.730	1.248
	2017	0.216	0.226	0.253	0.020–0.539	0.094	0.931	–0.025

[a]*Contents of the amino acids are presented as % in unmilled rice. Asp, aspartic acid; Thr, threonine; Ser, serine; Glu, glutamic acid; Gly, glycine; Ala, alanine; Cys, cystine; Val, valine; Met, methionine; Ile, ilsoleucine; Leu, leucine; Tyr, tyrosine; Phe, phenylalanine; Lys, lysine; His, histidine; Arg, arginine and Pro, proline.*

the RZ471–RZ395 region on the long arm of this chromosome. The enhancing alleles were derived from ZS97 at *qTyr7.2* and *qPro7.1*, and from MY46 at *qGly7.2*, *qCys7*, and *qPro7.2*. The five QTLs had high R^2 ranging from 10.7 to 35.9%. Altogether, 14 QTLs were detected on chromosome 7.

The third largest cluster consisting eight QTLs (*qSer1*, *qVal1*, *qIle1*, *qLeu1*, *qTyr1*, *qLys1.1*, *qHis1*, and *qPro1.1*) was located in the RM283–RM3746 region on the short arm of chromosome 1. Enhancing alleles of these QTLs were all derived from ZS97, with *qLeu1* having the highest R^2 of 15.0%. Three other QTLs (*qAsp1*, *qGly1*, and *qPro1.2*) were located in the pericentromeric region of chromosome 1. They had high R^2 ranging from 17.0 to 32.8%, and the enhancing alleles were all derived from ZS97. Two sparsely-distributed QTLs (*qPro1.3* and *qLys1.2*) were located

TABLE 2 | Pearson correlation coefficients between 17 components of AAC in the ZS97/MY46 recombinant inbred line (RIL) population.

Trait	Asp	Thr	Ser	Glu	Ala	Gly	Cys	Val	Met	Ile	Leu	Tyr	Phe	Lys	His	Arg
Thr	0.96*															
Ser	0.94*	0.94*														
Glu	0.88*	0.88*	0.94*													
Ala	0.23*	0.30*	0.19	0.17												
Gly	0.32*	0.30*	0.31*	0.34*	–0.62*											
Cys	0.13	0.06	0.20*	0.25*	–0.22*	0.04										
Val	0.86*	0.87*	0.91*	0.92*	0.11	0.38*	0.17									
Met	0.30*	0.29*	0.37*	0.36*	–0.24*	0.37*	0.16	0.48*								
Ile	0.84*	0.85*	0.88*	0.90*	0.15	0.33*	0.19*	0.95*	0.43*							
Leu	0.88*	0.86*	0.94*	0.95*	0.15	0.31*	0.27*	0.92*	0.37*	0.94*						
Tyr	0.42*	0.33*	0.44*	0.44*	–0.10	0.20*	0.45*	0.39*	0.15	0.47*	0.57*					
Phe	0.77*	0.75*	0.86*	0.87*	0.04	0.33*	0.30*	0.86*	0.40*	0.86*	0.93*	0.58*				
Lys	0.79*	0.78*	0.78*	0.73*	0.14	0.29*	0.13	0.74*	0.27*	0.77*	0.78*	0.51*	0.80*			
His	0.75*	0.77*	0.78*	0.75*	0.25*	0.22*	0.08	0.73*	0.25*	0.72*	0.76*	0.47*	0.76*	0.84*		
Arg	0.88*	0.87*	0.93*	0.90*	0.13	0.33*	0.24*	0.91*	0.39*	0.89*	0.93*	0.54*	0.87*	0.80*	0.81*	
Pro	0.40*	0.36*	0.38*	0.42*	–0.37*	0.66*	0.16	0.37*	0.24*	0.38*	0.41*	0.47*	0.41*	0.41*	0.33*	0.37*

*$P < 0.003$. The correlation coefficients were calculated based on mean values over 3 years.

in lower regions of the long arm. Altogether, 13 QTLs were detected on chromosome 1.

Among the remaining 30 QTLs, two single QTL were located on chromosomes 8 and 10, respectively, and the others were distributed on chromosomes 2, 3, 4, 9, 11, and 12 with 2–6 QTLs per chromosome. The six QTLs on chromosome 2 were all located in the lower region of the long arm; the two QTLs on chromosome 3 were tightly linked; the six QTLs on chromosome 4 involved two pairs of tightly-linked QTLs with two nearby QTLs; the five QTLs on chromosome 9 may be viewed as one cluster and two single QTL; the five QTLs on chromosome 11 was separated into two clusters; and the four QTLs on chromosome 12 included one cluster and one single QTL.

Effect of *RFT1* on Amino Acid Contents Detected Between NIL^{ZS97} and NIL^{MY46}

As described above, the largest QTL cluster for the 17 components of AAC detected in the ZS97/MY46 RIL population was located in the RM190–RM6917 region on the short arm of chromosome 6. This region covered the two florigen genes of rice, *RFT1* and *Hd3a* (Hori et al., 2016), suggesting a possible involvement of *RFT1* and/or *Hd3a* in controlling AAC of rice grain. This assumption was firstly tested using the NIL population TF6-15 segregating a 29.9-kb interval covering both *RFT1* and *Hd3a*. Significant differences ($P < 0.05$) between the 10 homozygous lines of NIL^{ZS97} and 10 homozygous lines of NIL^{MY46} were detected on 15 of the 17 components of AAC (**Table 3**). The R^2 for individual components ranged from 16.7 to 61.2%. The enhancing alleles were all derived from MY46, which is in agreement with the effects detected in the ZS97/MY46 RIL population.

Then, QTL analysis was performed using the NIL population R1 that was homozygous at the *Hd3a* locus but segregated for the *RFT1* gene. Significant differences ($P < 0.05$) between the 20 homozygous lines of NIL^{ZS97} and 20 homozygous lines of NIL^{MY46} were detected on 15 of the 17 components of AAC (**Table 4**). The R^2 for individual components ranged from 9.5 to 63.2%. Again, the enhancing alleles were all derived from MY46. It is also noted that the two components showing no significant difference between NIL^{ZS97} and NIL^{MY46} were commonly found to be Met and Pro in the TF6-15 and R1 populations. These results indicate that the *RFT1* gene has a strong and stable effect on most components of AAC in unmilled rice.

Effect of the *RFT1*^{ZS97} Transgene on Amino Acid Contents in a *Japonica* Rice Background

The effect of *RFT1* on AAC in unmilled rice was further tested using a transgenic population segregating the *RFT1*^{ZS97} transgene in the genetic background of *japonica* cultivar ZH11. Significant differences ($P < 0.05$) between the 14 lines of NIL^{ZS97} carrying homozygous transgenes and 14 lines of NIL^{ZH11} carrying no transgene were detected on 13 of the 17 components of AAC, with R^2 ranging from 8.1 to 38.0% (**Table 5**). Integration of the *RFT1*^{ZS97} transgene into the genome of ZH11 reduced the contents of the amino acids. In addition, the two components showing no significant difference between NIL^{ZS97} and NIL^{MY46} in the TF6-15 and R1 populations, Met and Pro, were included in the four components having no significant difference between NIL^{ZS97} and NIL^{ZH11} in the transgenic population. These results indicate that the effects of *RFT1* on AAC of unmilled rice are consistent in the genetic background of different subspecies of Asian cultivated rice.

DISCUSSION

Heading date, grain yield, and grain quality are three basic traits influencing the commercial utilization of a rice cultivar. The regional and seasonal adaptation is mostly determined by heading date, the productivity is measured by grain yield, and whether the product can meet the demand of end-users is mainly characterized by grain quality. A number of key genes for flowering regulation in rice have been found to play important roles in the genetic control of yield traits, including *Ghd7* (Xue et al., 2008; Weng et al., 2014), *DTH8*/*Ghd8* (Wei et al., 2010; Yan

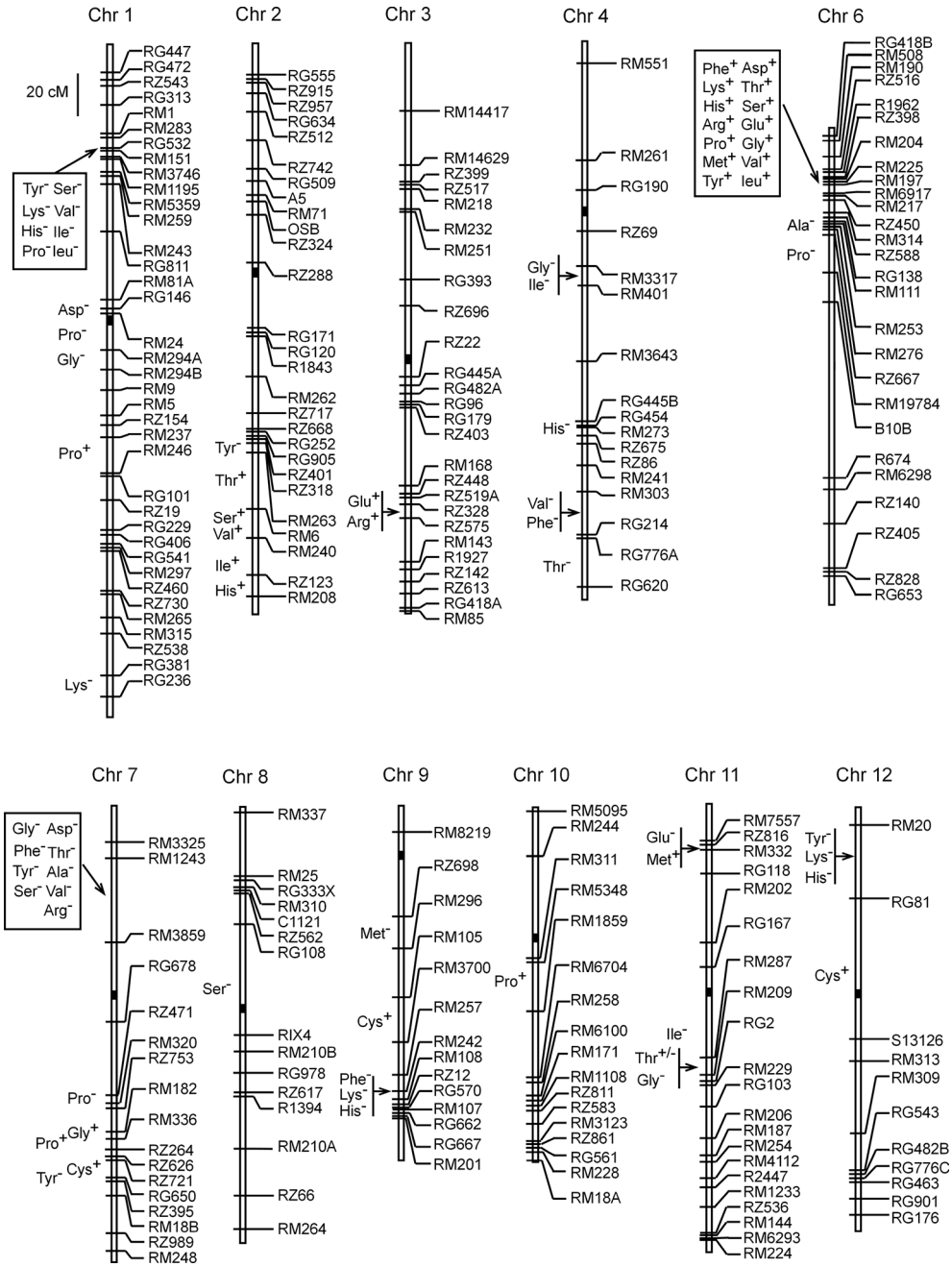


FIGURE 1 | Chromosomal locations of quantitative trait loci for 17 components of amino acid contents detected in the recombinant inbred line population of Zhenshan 97/Milyang 46.

et al., 2011), *Hd1* (Zhang et al., 2012; Zhang et al., 2015; Ye et al., 2018), *Ghd7.1* (Yan et al., 2013), and *RFT1* (Zhu et al., 2017). On the other hand, no study has been reported for the pleiotropic effects of heading-date genes on grain quality in rice. Among traits in the four primary categories of rice grain quality (i.e., milling, appearance, eating and cooking, and nutritional qualities), PC and AAC are the major parameters of nutritional quality (Wang et al., 2008; Wang et al., 2017a). In the present study, a total of 73 QTLs for AAC of unmilled rice were detected using the ZS97/MY46 RIL population, and the largest QTL cluster was validated to be responsible by the *RFT1* gene on the short arm of chromosome 6. It is also evident that the effect of *RFT1* is consistent across different genetic backgrounds. In accordance with the common occurrence of significantly

TABLE 3 | Effects of the *RFT1-Hd3a* region on 17 components of amino acid content (AAC) detected in the TF6-15 population.

Trait	Phenotype (mean±SE)[a]		*P*	*A*[b]	R^2 (%)[c]
	NIL^{ZS97}	NIL^{MY46}			
Asp	0.721±0.006	0.762±0.010	0.0007	0.020	39.8
Thr	0.308± 0.003	0.326±0.004	0.0009	0.009	37.6
Ser	0.427± 0.004	0.452±0.005	0.0005	0.013	38.8
Glu	1.543±0.008	1.626±0.015	0.0003	0.042	42.0
Gly	0.183± 0.001	0.193±0.001	0.0002	0.005	50.0
Ala	0.421±0.004	0.447±0.007	<0.0001	0.013	46.3
Cys	0.273±0.003	0.286±0.006	0.0034	0.006	28.9
Val	0.331±0.005	0.359±0.008	0.0033	0.014	26.3
Met	0.047±0.003	0.051±0.005	0.4121		
Ile	0.202±0.004	0.221±0.005	0.0063	0.009	22.6
Leu	0.577±0.004	0.614±0.006	<0.0001	0.019	47.1
Tyr	0.165±0.002	0.177±0.004	0.0295	0.006	16.7
Phe	0.382±0.003	0.413±0.005	<0.0001	0.015	54.9
Lys	0.231±0.003	0.250±0.005	<0.0001	0.010	52.1
His	0.167± 0.002	0.180±0.003	<0.0001	0.007	51.3
Arg	0.499±0.003	0.544±0.007	<0.0001	0.023	61.2
Pro	0.204±0.011	0.224±0.005	0.0534		

[a] *NIL^{ZS97} and NIL^{MY46} are near isogenic lines carrying homozygous alleles from ZS97 and MY46, respectively.*
[b] *Additive effect of replacing a ZS97 allele with a MY46 allele.*
[c] *Proportion of phenotypic variance explained by the QTL effect.* $R^2 = V_G/V_P \times 100$, *in which* V_G *is the variance between the two genotypic groups, and* V_P *the phenotypic variance.*

TABLE 4 | Effects of the *RFT1* gene on 17 components of amino acid content (AAC) detected in the R1 population.

Trait	Phenotype (mean ± SE) [a]		*P*	*A*[b]	R^2(%)[c]
	NIL^{ZS97}	NIL^{MY46}			
Asp	0.733 ± 0.006	0.819 ± 0.007	<0.0001	0.043	58.7
Thr	0.328 ± 0.002	0.362 ± 0.002	<0.0001	0.017	62.8
Ser	0.432 ± 0.003	0.480 ± 0.004	<0.0001	0.024	63.2
Glu	1.495 ± 0.012	1.660 ± 0.012	<0.0001	0.083	59.8
Gly	0.173 ± 0.001	0.192 ± 0.002	<0.0001	0.010	56.5
Ala	0.440 ± 0.004	0.485 ± 0.003	<0.0001	0.023	53.6
Cys	0.321 ± 0.003	0.334 ± 0.004	0.0009	0.006	10.4
Val	0.377 ± 0.003	0.415 ± 0.004	<0.0001	0.019	46.6
Met	0.066 ± 0.002	0.067 ± 0.002	0.6776		
Ile	0.238 ± 0.002	0.261 ± 0.002	<0.0001	0.012	45.4
Leu	0.595 ± 0.005	0.662 ± 0.005	<0.0001	0.033	58.2
Tyr	0.202 ± 0.003	0.217 ± 0.005	0.0026	0.007	9.5
Phe	0.392 ± 0.004	0.445 ± 0.004	<0.0001	0.026	57.1
Lys	0.266 ± 0.003	0.292 ± 0.002	<0.0001	0.013	42.6
His	0.181 ± 0.002	0.202 ± 0.002	<0.0001	0.011	55.5
Arg	0.500 ± 0.005	0.570 ± 0.006	<0.0001	0.035	52.5
Pro	0.185 ± 0.006	0.202 ± 0.006	0.0582		

[a] *NIL^{ZS97} and NIL^{MY46} are near isogenic lines carrying homozygous alleles from ZS97 and MY46, respectively.*
[b] *Additive effect of replacing a ZS97 allele with a MY46 allele.*
[c] *Proportion of phenotypic variance explained by the QTL effect.* $R^2 = V_G/V_P \times 100$, *in which* V_G *is the variance between the two genotypic groups, and* V_P *the phenotypic variance.*

TABLE 5 | Effects of $RFT1^{ZS97}$ transgene on 17 components of amino acid content (AAC).

Trait	Phenotype (mean ± SE)[a]		*P*	*A*[b]	R^2 (%)[c]
	NIL^{ZS97}	NIL^{ZH11}			
Asp	0.745 ± 0.009	0.789 ± 0.008	<0.0001	–0.022	28.1
Thr	0.315 ± 0.004	0.328 ± 0.003	<0.0001	–0.006	14.8
Ser	0.423 ± 0.004	0.440 ± 0.004	<0.0001	–0.008	19.6
Glu	1.380 ± 0.017	1.423 ± 0.014	0.0191	–0.021	8.1
Gly	0.184 ± 0.002	0.192 ± 0.002	0.0002	–0.004	17.2
Ala	0.441 ± 0.005	0.457 ± 0.005	0.0005	–0.008	13.9
Cys	0.264 ± 0.002	0.280 ± 0.003	0.0014	–0.008	22.7
Val	0.369 ± 0.005	0.378 ± 0.005	0.1036		
Met	0.075 ± 0.003	0.086 ± 0.003	0.0975		
Ile	0.221 ± 0.004	0.226 ± 0.003	0.1427		
Leu	0.558 ± 0.006	0.574 ± 0.005	0.0008	–0.008	11.5
Tyr	0.166 ± 0.002	0.175 ± 0.003	0.0007	–0.005	13.8
Phe	0.380 ± 0.004	0.397 ± 0.004	0.0003	–0.009	17.8
Lys	0.273 ± 0.003	0.292 ± 0.003	<0.0001	–0.010	36.0
His	0.187 ± 0.002	0.199 ± 0.001	<0.0001	–0.006	38.0
Arg	0.533 ± 0.006	0.563 ± 0.005	<0.0001	–0.015	32.7
Pro	0.182 ± 0.005	0.183 ± 0.006	0.9528		

[a] *NIL^{ZS97} and NIL^{ZH11} are T_3 transgenic lines carrying* $RFT1^{ZS97}$ *homozygous transgene and no transgene in the genetic background of ZH11, respectively.*
[b] *Additive effect of replacing a ZH11 allele with a ZS97 allele.*
[c] *Proportion of phenotypic variance explained by the QTL effect.* $R^2 = V_G/V_P \times 100$, *in which* V_G *is the variance between the two genotypic groups, and* V_P *the phenotypic variance.*

positive correlations between different components of AAC (**Table 2**; **Supplementary Table S1**), most of the QTLs were located in cluster and different QTL in a given region usually had the same allelic direction (**Figure 1**; **Supplementary Table S2**).

RFT1 protein is the florigen for promoting the flowering of rice under long-day (LD) conditions (Tsuji et al., 2011). As compared to the ZS97 allele of *RFT1*, the MY46 and ZH11 alleles were shown to promote heading in rice populations grown under natural LD conditions in Hangzhou (Zhu et al., 2017). Replacing a ZS97 allele by a MY46 allele in the R1 population promoted flowering by 11.63 to 15.61 days over 3 years; and replacing a ZS97 allele by a ZH11 allele promoted flowering by 3.91 and 6.12 days in the transgenic population in 2 years. In the present study, replacement of the ZS97 allele of *RFT1* with the MY46 and ZH11 alleles resulted in increasing the contents of most amino acids (**Supplementary Table S2**; **Tables 3–5**). Obviously, the effects of *RFT1* on HD and AAC have opposite allelic directions, which is in accordance with the opposite allelic directions between QTLs for HD and PC located on the short arm of chromosome 6 (Wada et al., 2006; Yun et al., 2016).

Two other QTLs for AAC (*qAla6* and *qPro6.2*) were detected on the short of chromosome 6 in the ZS97/MY46 RIL population. They were located in the intervals RM253–RM276 and RZ667–RM19784 that are closer to the centromere region than is *RFT1*. The alleles for increasing AAC were both derived from ZS97 (**Supplementary Table S2**). At the *Hd1* locus that is tightly linked to RM19784, the functional $Hd1^{ZS97}$ allele acted to decrease HD as compared to the non-functional $Hd1^{MY46}$ allele in the ZS97 background (Zhang et al., 2012). These results suggest that the *Hd1* gene also have pleiotropic effects on HD and AAC with opposite directions.

The second and third largest QTL clusters detected in the ZS97/MY RIL population were located in the RM3325–RM3859 and RM283–RM3746 regions on the short art of chromosomes 7 and 1, covering the heading-date genes *Ghd7* (Xue et al., 2008) and *OsGI* (Hayama et al., 2003), respectively. For all the QTLs included in the two clusters, the alleles for increasing AAC were derived from

ZS97 (**Supplementary Table S2**). In previous studies using the same rice cross, the ZS97 alleles in the two regions were found to decrease HD (Zhang et al., 2011; Zhang et al., 2016). These results suggest that the pleiotropic effects of major heading-date genes on AAC with opposite directions could be a common occurrence.

Similar to previous results reported by other groups (Wang et al., 2008; Lu et al., 2009; Zhong et al., 2011), our study found that it is common that a QTL region affected most components of AAC. However, it is unlikely that a gene can affect the biosynthesis of most amino acids in rice. Given that all the major QTL regions affecting AAC detected in this study were located in approximate to genes/QTLs controlling flowering time, these regions would have large effects on all traits which were influenced by heading date. It is possible that the influence of these regions on most components of AAC could be caused by indirect effects of heading date genes rather than by the direct control of these genes on the biosynthesis of amino acids. It is possible that a heading-date gene is involved in controlling nutrient transportation and accumulation in rice, either by direct involvement in the regulating network or by environmental influences on the nutrition uptake and transport due to heading date variation.

Utilization of the pleiotropic effects of heading-date genes on AAC and PC could help to meet the diverse requirements of protein for human consumption. High contents of protein and amino acids are favorable for enhancing the nutritional value, but unfavorable for eating quality (Martin and Fitzgerald, 2002; Kwon et al., 2011; Yun et al., 2016) and undesirable for some uses such as wine-making (Yoshida et al., 2002) and certain types of diet (Zhang et al., 2008). Alleles for promoting HD and increasing AAC could be selected for developing rice varieties with high nutritional values; and alleles for delaying heading and reducing AAC may be applied for developing high-yielding varieties with good eating quality. In this regard, more efforts are needed to establish a better understanding on the pleiotropism of heading-date genes on multiple traits for grain quality.

DATA AVAILABILITY STATEMENT

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

AUTHOR CONTRIBUTIONS

J-YZ and P-SH conceived and designed the experiments. L-HX, Y-JZ, S-QT, X-JW, Z-HS, and G-AJ performed the experiments. L-HX, Y-JZ and J-YZ analysed the data. L-HX, Y-JZ and J-YZ wrote the manuscript.

FUNDING

This work was funded by the National Key R&D Program of China (Grant number 2017YFD0100300), and the National Natural Science Foundation of China (Grant number 31571637).

ACKNOWLEDGMENTS

We thank Mr. Hongfa Chen for rice material preparation and Mrs. Lingli Sun for preparation of amino acid content analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00013/full#supplementary-material

REFERENCES

Aluko, G., Martinez, C., Tohme, J., Castano, C., Bergman, C., and Oard, J. H. (2004). QTL mapping of grain quality traits from the interspecific cross *Oryza sativa* × *O. glaberrima*. *Theor. Appl. Genet.* 109, 630–639. doi: 10.1007/s00122-004-1668-y

Chen, N., Luo, Y. K., Xie, L. H., Zhu, Z. W., Duan, B. W., and Zhang, L. P. (2006). Protein content and its correlation with other quality parameters of rice in China. *Acta Agron. Sinica*. 32, 1193–1196. (in Chinese with English abstract). doi: 10.3321/j.issn:0496-3490.2006.08.015

Dai, W. M., Zhang, K. Q., Wu, J. R., Wang, L., Duan, B. W., Zheng, K. L., et al. (2008). Validating a segment on the short arm of chromosome 6 responsible for genetic variation in the hull silicon content and yield traits of rice. *Euphytica* 160, 317–324. doi: 10.1007/s10681-007-9501-8

Hayama, R., Yokoi, S., Tamaki, S., Yano, M., and Shimamoto, K. (2003). Adaptation of photoperiodic control pathways produces short-day flowering in rice. *Nature* 422, 719–722. doi: 10.1038/nature01549

Hori, K., Matsubara, K., and Yano, M. (2016). Genetic control of flowering time in rice: integration of Mendelian genetics and genomics. *Theor. Appl. Genet.* 129, 2241–2252. doi: 10.1007/s00122-016-2773-4

Kashiwagi, T., and Munakata, J. (2018). Identification and characteristics of quantitative trait locus for grain protein content, *TGP12*, in rice (*Oryza sativa* L.). *Euphytica* 214, 165. doi: 10.1007/s10681-018-2249-5

Kwon, S. W., Cho, Y. C., Lee, J. H., Suh, J. P., Kim, J. J., Kim, M. K., et al. (2011). Identification of quantitative trait loci associated with rice eating quality traits using a population of recombinant inbred lines derived from a cross between two temperate *japonica* cultivars. *Mol. Cells* 31, 427–445. doi: 10.1007/s10059-011-0289-y

Lu, K., Li, L., Zheng, X., Zhang, Z., Mou, T., and Hu, Z. (2009). Genetic dissection of amino acid content in rice grain. *J. Sci. Food Agric.* 89, 2377–2382. doi: 10.1002/jsfa.3731

Martin, M., and Fitzgerald, M. A. (2002). Proteins in rice influence cooking properties. *J. Cereal Sci.* 36, 285–294. doi: 10.1006/jcrs.2001.0465

McCouch, S. R.CGSNL (Committee on Gene Symbolization, Nomenclature and Linkage, Rice Genetics Cooperative) (2008). Gene nomenclature system for rice. *Rice* 1, 72–84. doi: 10.1007/s12284-008-9004-9

Tan, Y. F., Sun, M., Xing, Y. Z., Hua, Z. P., Sun, X. L., Zhang, Q. F., et al. (2001). Mapping quantitative trait loci for milling quality, protein content and color characteristics of rice using a recombinant inbred line population derived from an elite rice hybrid. *Theor. Appl. Genet.* 103, 1037–1045. doi: 10.1007/s001220100665

Tsuji, H., Taoka, K. I., and Shimamoto, K. (2011). Regulation of flowering in rice: two florigen genes, a complex gene network, and natural variation. *Curr. Opin. Plant Biol.* 14, 45–52. doi: 10.1016/j.pbi.2010.08.016

Wada, T., Uchimura, Y., Ogata, T., Tsubone, M., and Matsue, Y. (2006). Mapping of QTLs for physicochemical properties in *japonica* rice. *Breed. Sci.* 56, 253–260. doi: 10.1270/jsbbs.56.253

Wang, L. Q., Zhong, M., Li, X. H., Yuan, D. J., Xu, Y. B., Liu, H. F., et al. (2008). The QTL controlling amino acid content in grains of rice (*Oryza sativa*) are co-localized with the regions involved in the amino acid metabolism pathway. *Mol. Breed.* 21, 127–137. doi: 10.1007/s11032-007-9141-7

Wang, S., Basten, C., and Zeng, Z. (2011). *Windows QTL Cartographer 2.5* (Raleigh, NC: North Carolina State University).

Wang, X., Pang, Y., Zhang, J., Wu, Z., Chen, K., Ali, J., et al. (2017a). Genome-wide and gene-based association mapping for rice eating and cooking characteristics and protein content. *Sci. Rep.* 7, 17203. doi: 10.1038/s41598-017-17347-5

Wang, Z., Chen, J. Y., Zhu, Y. J., Fan, Y. Y., and Zhuang, J. Y. (2017b). Validation of *qGS10*, a quantitative trait locus for grain size on the long arm of chromosome 10 in rice (*Oryza sativa* L.). *J. Integr. Agr.* 16, 16–26. doi: 10.1016/S2095-3119(16)61410-7

Wei, X., Xu, J., Guo, H., Jiang, L., Chen, S., Yu, C., et al. (2010). *DTH8* Suppresses flowering in rice, influencing plant height and yield potential simultaneously. *Plant Physiol.* 153, 1747–1758. doi: 10.1104/pp.110.156943

Weng, X., Wang, L., Wang, J., Hu, Y., Du, H., Xu, C., et al. (2014). *Grain number, plant height, and heading date7* is a central regulator of growth, development, and stress response. *Plant Physiol.* 164, 735–747. doi: 10.1104/pp.113.231308

Xue, W., Xing, Y., Weng, X., Zhao, Y., Tang, W., Wang, L., et al. (2008). Natural variation in *Ghd7* is an important regulator of heading date and yield potential in rice. *Nat. Genet.* 40, 761–767. doi: 10.1038/ng.143

Yan, W. H., Wang, P., Chen, H. X., Zhou, H. J., Li, Q. P., Wang, C. R., et al. (2011). A major QTL, *Ghd8*, plays pleiotropic roles in regulating grain productivity, plant height, and heading date in rice. *Mol. Plant* 4, 319–330. doi: 10.1093/mp/ssq070

Yan, W., Liu, H., Zhou, X., Li, Q., Zhang, J., Lu, L., et al. (2013). Natural variation in *Ghd7.1* plays an important role in grain yield and adaptation in rice. *Cell Res.* 23, 969–971. doi: 10.1038/cr.2013.43

Ye, J., Niu, X., Yang, Y., Wang, S., Xu, Q., Yuan, X., et al. (2018). Divergent *Hd1*, *Ghd7*, and *DTH7* alleles control heading date and yield potential of *japonica* rice in Northeast China. *Front. Plant Sci.* 9, 35. doi: 10.3389/fpls.2018.00035

Yoshida, S., Ikegami, M., Kuze, J., Sawada, K., Hashimoto, Z., Ishii, T., et al. (2002). QTL analysis for plant and grain characters of *sake*-brewing rice using a double haploid population. *Breed. Sci.* 52, 309–317. doi: 10.1270/jsbbs.52.309

Yu, Y. H., Li, G., Fan, Y. Y., Zhang, K. Q., Min, J., Zhu, Z. W., et al. (2009). Genetic relationship between grain yield and the contents of protein and fat in a recombinant inbred population of rice. *J. Cereal Sci.* 50, 121–125. doi: 10.1016/j.jcs.2009.03.008

Yun, Y. T., Chung, C. T., Lee, Y. J., Na, H. J., Lee, J. C., Lee, S. G., et al. (2016). QTL Mapping of grain quality traits using introgression lines carrying *Oryza rufipogon* chromosome segments in *japonica* rice. *Rice* 9, 62. doi: 10.1186/s12284-016-0135-0

Zhang, W., Bi, J., Chen, L., Zheng, L., Ji, S., Xia, Y., et al. (2008). QTL mapping for crude protein and protein fraction contents in rice (*Oryza sativa* L.). *J. Cereal Sci.* 48, 539–547. doi: 10.1016/j.jcs.2007.11.010

Zhang, Z. H., Guo, L., Zhu, Y. J., Fan, Y. Y., and Zhuang, J. Y. (2011). Mapping of quantitative trait loci for heading date and plant height in two populations of *indica* rice. *Sci. Agric. Sin.* 44, 3069–3077. (in Chinese with English abstract). doi: 10.3864/j.issn.0578-1752.2011.15.001

Zhang, Z. H., Wang, K., Guo, L., Zhu, Y. J., Fan, Y. Y., Cheng, S. H., et al. (2012). Pleiotropism of the photoperiod-insensitive allele of *Hd1* on heading date, plant height and yield traits in rice. *PloS One* 7, e52538. doi: 10.1371/journal.pone.0052538

Zhang, J., Zhou, X., Yan, W., Zhang, Z., Lu, L., Han, Z., et al. (2015). Combinations of the *Ghd7*, *Ghd8* and *Hd1* genes largely define the ecogeographical adaptation and yield potential of cultivated rice. *New Phytol.* 208, 1056–1066. doi: 10.1111/nph.13538

Zhang, Z. H., Cao, L. Y., Chen, J. Y., Zhang, Y. X., Zhuang, J. Y., and Cheng, S. H. (2016). Effects of *Hd2* in the presence of the photoperiod-insensitive functional allele of *Hd1* in rice. *Biol. Open* 5, 1719–1726. doi: 10.1242/bio.021071

Zhong, M., Wang, L. Q., Yuan, D. J., Luo, L. J., Xu, C. G., and He, Y. Q. (2011). Identification of QTL affecting protein and amino acid contents in rice. *Rice Sci.* 18, 187–195. doi: 10.1016/S1672-6308(11)60026-7

Zhu, Y. J., Fan, Y. Y., Wang, K., Huang, D. R., Liu, W. Z., Ying, J. Z., et al. (2017). Rice *Flowering Locus T 1* plays an important role in heading date influencing yield traits in rice. *Sci. Rep.* 7, 4918. doi: 10.1038/s41598-017-05302-3

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 Genetics

ORIGINAL RESEARCH
published: 21 February 2020
doi: 10.3389/fgene.2020.00086

Evaluation of Whole-Genome Sequence, Genetic Diversity, and Agronomic Traits of Basmati Rice (*Oryza sativa* L.)

D.S. Kishor[1†], *Jeonghwan Seo*[1†], *Joong Hyoun Chin*[2*] *and Hee-Jong Koh*[1*]

[1] *Department of Plant Science, Plant Genomics and Breeding Institute, and Research Institute of Agriculture and Life Science, Seoul National University, Seoul, South Korea,* [2] *Department of Integrative Bio-industrial Engineering, Sejong University, Seoul, South Korea*

OPEN ACCESS

Edited by:
Felipe Klein Ricachenevsky, Universidade Federal de Santa Maria, Brazil

Reviewed by:
Gireesh Channappa, Indian Institute of Rice Research (ICAR), India
Raju Bheemanahalli Rangappa, Kansas State University, United States

***Correspondence:**
Joong Hyoun Chin
jhchin@sejong.ac.kr
Hee-Jong Koh
heejkoh@snu.ac.kr

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *10 October 2019*
Accepted: *27 January 2020*
Published: *21 February 2020*

Citation:
Kishor DS, Seo J, Chin JH and Koh H-J (2020) Evaluation of Whole-Genome Sequence, Genetic Diversity, and Agronomic Traits of Basmati Rice (Oryza sativa L.). Front. Genet. 11:86. doi: 10.3389/fgene.2020.00086

Basmati is considered a unique varietal group of rice (*Oryza sativa* L.) because of its aroma and superior grain quality. Previous genetic analyses of rice showed that most of the Basmati varieties are classified into the *aromatic* group. Despite various efforts, genomic relationship of Basmati rice with other varietal groups and genomic variation in Basmati rice are yet to be understood. In the present study, we resequenced the whole genome of three traditional Basmati varieties at a coverage of more than 25X using Illumina HiSeq2500 and mapped the obtained sequences to the reference genome sequences of Nipponbare (*japonica* rice), Kasalath (*aus* rice), and Zhenshan 97 (*indica* rice). Comparison of these sequences revealed common single nucleotide polymorphisms (SNPs) in the genic regions of three Basmati varieties. Analysis of these SNPs revealed that Basmati varieties showed fewer sequence variations compared with the *aus* group than with the *japonica* and *indica* groups. Gene ontology (GO) enrichment analysis indicated that SNPs were present in genes with various biological, molecular, and cellular functions. Additionally, functional annotation of the Basmati mutated gene cluster shared by Nipponbare, Kasalath, and Zhenshan 97 was found to be associated with the metabolic process involved in the cellular aromatic compound, suggesting that aroma is an important specific genomic feature of Basmati varieties. Furthermore, 30 traditional Basmati varieties were classified into three different groups, *aromatic* (22 varieties), *aus* (four varieties), and *indica* (four varieties), based on genome-wide SNPs. All 22 *aromatic* Basmati varieties harbored the fragrant-inducing *Badh2* allele. We also performed comparative analysis of 13 key agronomic and grain quality traits of Basmati rice and other rice varieties. Three traits including length-to-width ratio of grain (L/W ratio), panicle length (PL), and amylose content (AC) showed significant ($P < 0.05$ and $P < 0.01$) differences between the *aromatic* and *indica*/*aus* groups. Comparative analysis of genome structure, based on genome sequence variation and GO analysis, revealed that the Basmati genome was derived mostly from the *aus* and *japonica* groups. Overall, whole-genome sequence data and genetic diversity information obtained in this study will

serve as an important resource for molecular breeding and genetic analysis of Basmati varieties.

Keywords: Basmati rice, aromatic, SNPs, NGS, gene ontology

INTRODUCTION

Rice (*Oryza sativa* L.) is an important cereal crop and represents the staple food of more than half of the global population (Wang and Li, 2005). *O. sativa* is classified into two distinct subspecies, *japonica* and *indica* (Kato, 1928), and into five groups including *indica*, *aus*, *aromatic*, temperate *japonica*, and tropical *japonica* (Garris et al., 2005). *O. sativa* was domesticated more than 10,000 years ago from Asian wild rice species, *O. rufipogon* and *O. nivara* (Kovach et al., 2007; Sang and Ge, 2007; Chen et al., 2019). Both *japonica* and *indica* rice have undergone significant phenotypic changes compared with *O. rufipogon* (proto-*japonica*) and *O. nivara* (proto-*indica*), respectively, and have expanded their geographical distribution during domestication (Fuller et al., 2010).

Basmati rice is considered a unique varietal group because of its aroma and superior grain quality (Ahuja et al., 1995; Siddiq et al., 2012). These unique varietal group occupies a special status among the consumers due to its unique quality traits such as extra-long slender grain, lengthwise excessive kernel elongation upon cooking, soft and fluffy texture after cooking, and aroma. Therefore, Basmati varieties are designated as the most highly produced and economically successful group (Civáň et al., 2019). The term Basmati is derived from two Sanskrit words, "*Vas*" meaning "*aroma*" and "*matup*" meaning "possessing." The combination of the two Sanskrit words, "Vaasmati," is pronounced as "Basmati" (Siddiq et al., 2012). Studies suggest that Basmati rice varieties represent the *aromatic* group from *indica* and *japonica* subspecies (Glaszmann, 1987; Garris et al., 2005).

From the decades, less attention has given at the origin of Basmati group. This is mainly due to the conflicting phylogenetic relationships were observed among Basmati and other rice groups (Choi et al., 2017). Furthermore, genome-wide polymorphism analysis in Asian cultivated rice showed that Basmati rice varieties share a close phylogenetic relationship with *japonica* varieties (Huang et al., 2012; Wang et al., 2018). Recent findings of Choi et al. (2018) and Civáň et al. (2019) providing more evidence that Basmati genome was genetically close to *japonica* and *aus* rice. However, these studies were carried out using single Basmati genome, which has limited information on Basmati genome variation. Although some progresses have been made in understanding of origin of Basmati genome, further study is needed to identify the Basmati-specific genome features and genome variation by assembling the traditional Basmati varieties compared with *japonica*, *indica*, and *aus* groups. Next-generation sequencing (NGS) technologies are important for genomic analysis and molecular breeding (Chen et al., 2014), and enable the identification of functional genomic variation, and unique SNPs, and insertion-deletion polymorphisms (InDels) across the genome, which offer an exciting opportunity to genetic diversity studies in the crop plants (Jimenez et al., 2013; Serba et al., 2019).

In Basmati rice, molecular mapping and cloning of the *fgr* gene, which encodes betaine aldehyde dehydrogenase homologue 2 (*Badh2*), revealed an 8-bp deletion and three single nucleotide polymorphisms (SNPs) in the 7th exon, resulting in the fragrant trait (Bradbury et al., 2005). Haplotype analysis of the *Badh2* gene showed that the 8-bp deletion in the majority of fragrant Basmati varieties causes a loss-of-function mutation, which enhances the biosynthesis of 2-acetyl-1-pyroline (2-AP); this haplotype is identical to the ancestral *japonica* haplotypes, suggesting that introgression between *japonica* accessions and Basmati varieties is responsible for the fragrant trait in Basmati rice (Kovach et al., 2009). A recent study by Daygon et al. (2017) reported that four other amine heterocycles: 6-methyl, 5-oxo-2,3,4,5-tetrahydropyridine (6M5OTP), 2-acetylpyrrole, pyrrole, and 1-pyrroline, that correlate strongly with the production of 2AP, and are present in consistent proportions in a collection of recombinant inbred lines derived from Basmati-type rice, and these compounds were also co-localized with a single QTL that harbors the *fgr* gene. Although genetic basis of fragrant trait in Basmati rice seems to be complicated, most researchers proposed that grain aroma in Basmati rice is controlled by a single recessive gene (*Badh2*) (Bradbury et al., 2005; Kovach et al., 2009). However, some researchers also think that fragrant trait in Basmati rice is controlled by major and minor-effective genes (Daygon et al., 2017), and by several QTLs (Amarawathi et al., 2008; Pachauri et al., 2014; Vemireddy et al., 2015). Overall, the molecular genetic mechanism of fragrant trait is not clearly understood, more studies is needed on the functional allelic variation of aroma gene and number of genes controlling the grain aroma in Basmati rice.

In this study, we analyzed the differences between Basmati rice genome *vs. indica*, *japonica*, and *aus* rice genomes through whole-genome sequencing and marker analysis. The main objective is to identify the genomic features and genetic variation in Basmati rice that can be utilized for genetic studies and marker development for breeding. We also identified unique SNPs and Indel marker sets, and evaluated the key agronomic and grain quality traits of Basmati rice with other rice groups for varietal improvement.

MATERIALS AND METHODS

Plant Materials

A total of 60 rice varieties belonging to *indica*, *aus*, *aromatic*, temperate *japonica*, and tropical *japonica* groups were used in this study (**Table 1**). Among the 60 rice varieties, seeds of 30 traditional Basmati varieties [International Rice GenBank Collection (IRGC) designated] were obtained from the International Rice Research Institute (IRRI), while the other 30 rice varieties were from the Crop Molecular Breeding Lab, Seoul

TABLE 1 | List of rice varieties used in this study.

No.	Varieties	Origin	Accession no. [a]	Subgroup[b]	*Badh2* allele[c]
1	Nipponbare	Japan	981704	Temperate *japonica*	WT
2	Koshihikari	Japan	981581	Temperate *japonica*	WT
3	Yukara	Japan	981584	Temperate *japonica*	WT
4	Ilpumbyeo	South Korea	981585	Temperate *japonica*	WT
5	Jinheungbyeo	South Korea	981576	Temperate *japonica*	WT
6	Dongjinbyeo	South Korea	981626	Temperate *japonica*	WT
7	Hopumbyeo	South Korea	980403	Temperate *japonica*	WT
8	Tong 88-7	South Korea	980609	Temperate *japonica*	WT
9	MS 11	Philippines	981589	Temperate *japonica*	WT
10	Samnambyeo	South Korea	981579	Tropical *japonica*	WT
11	Malagkit Sinaguing	Philippines	961354	*Admixture*	WT
12	B581A6	Philippines	921648	Tropical *japonica*	WT
13	CP-SLO	USA	970083	Tropical *japonica*	WT
14	Azucena	Philippines	971155	Tropical *japonica*	WT
15	Reket Abang	Indonesia	260004	Tropical *japonica*	WT
16	Dawn	USA	981564	Tropical *japonica*	WT
17	Milyang 23	South Korea	981599	*Indica*	WT
18	Dasanbyeo	South Korea	981598	*Indica*	WT
19	Taichung Native 1	Taiwan	981570	*Indica*	WT
20	IR 64	Philippines	981566	*Indica*	WT
21	IR 72	Philippines	18053	*Indica*	WT
22	Chinsurah Boro 2	Bangladesh	851453	*Aus*	WT
23	Dular	India	980384	*Aus*	WT
24	Bina Dhan 10	Bangladesh	961192	*Indica*	WT
25	IR 24	Philippines	18049	*Indica*	WT
26	IR 8	Philippines	981596	*Indica*	WT
27	Minghui 63	China	981601	*Indica*	WT
28	N 22	India	970030	*Aus*	WT
29	Swarna	India	961181	*Indica*	WT
30	Basmati Dhan	Nepal	IRGC 23814	*Aromatic*	*badh2.1*
31	Dheradun Basmati	Nepal	IRGC 23861	*Aus*	WT
32	Basmati Nahan 381	Pakistan	IRGC 27786	*Aromatic*	*badh2.1*
33	Basmati Sufaid 100	Pakistan	IRGC 27791	*Aromatic*	*badh2.1*
34	Basmati 140	Pakistan	IRGC 27813	*Aus*	WT
35	Basmati 370	Pakistan	IRGC 27820	*Aromatic*	*badh2.1*
36	Basmati 372	Pakistan	IRGC 27823	*Aromatic*	*badh2.1*
37	Basmati 377	Pakistan	IRGC 27829	*Aromatic*	*badh2.1*
38	Deraduni Basmati 321	Pakistan	IRGC 27907	*Aromatic*	*badh2.1*
39	Kamoh Basmati 392	Pakistan	IRGC 28000	*Aromatic*	*badh2.1*
40	Sathi Basmati	Pakistan	IRGC 28230	*Aromatic*	*badh2.1*
41	Basmati Sal	India	IRGC 52411	*Aus*	WT
42	Basmati Kunar	Afghanistan	IRGC 58272	*Aromatic*	*badh2.1*
43	Basmati Kunduz	Afghanistan	IRGC 58273	*Indica*	WT
44	Basmati Anpjhutte	Nepal	IRGC 58879	*Aromatic*	*badh2.1*
45	Basmati Gola	Nepal	IRGC 58880	*Aromatic*	*badh2.1*
46	Basmati Lamo	Nepal	IRGC 58881	*Aromatic*	*badh2.1*
47	Basmati Masino	Nepal	IRGC 58883	*Indica*	WT
48	Basmati Nokhi	Nepal	IRGC 58884	*Indica*	WT
49	Basmati Pahade	Nepal	IRGC 58885	*Aromatic*	*badh2.1*
50	Basmati Red	Nepal	IRGC 58886	*Aromatic*	*badh2.1*
51	Basmati White	Nepal	IRGC 58887	*Aromatic*	*badh2.1*
52	Basmati Uzarka	Nepal	IRGC 58888	*Aromatic*	*badh2.1*
53	Kalo Basmati	Nepal	IRGC 59054	*Aromatic*	*badh2.1*
54	Rato Basmati	Nepal	IRGC 59205	*Aromatic*	*badh2.1*
55	Basmati Mwea	Kenya	IRGC 61183	*Aromatic*	*badh2.1*
56	Dahrdun Basmati	India	IRGC 67705	*Indica*	WT
57	Basmatiya	India	IRGC 67734	*Aus*	WT
58	Pakistani Basmati	India	IRGC 67746	*Aromatic*	*badh2.1*
59	Karnal Basmati	Pakistan	IRGC 76362	*Aromatic*	*badh2.1*
60	Kasalath	India	980341	*Indica*	WT

[a] *IRGC, International Rice GenBank Collection.*
[b] *Subgroup was determined based on 190 SNP markers.*
[c] *Badh2 genotype was determined based on the* fgr*-specific InDel marker developed by Sakthivel et al. (2009) and WT indicates wild type allele.*

National University. Among the 30 traditional Basmati varieties, Basmati 370, Rato Basmati, and Dahrdun Basmati were selected for whole-genome resequencing, based on their geo-location (**Figure 1**). Seeds from each accession were surface sterilized and sown in pots containing wet soil. The pots were placed in an experimental greenhouse for 30 days. Then, 30-day-old seedlings were transplanted in an experimental field at Seoul National University.

Genome Sequencing

Supplementary Figure 1 provides an overview of the work plan used in this study. To perform whole-genome resequencing, shotgun DNA libraries were prepared from high molecular weight genomic DNA of three traditional Basmati varieties using the NEXTflex™ Rapid DNA-Seq kit (Bioo Scientific Corporation, Austin, TX, USA). Then, the libraries were used for cluster generation and sequenced for 250 cycles on the Illumina HiSeq2500 platform (Illumina, San Diego, CA, USA), according to the manufacturer's instructions, at the National Instrumentation Center for Environmental Management (NICEM) of Seoul National University.

Mapping and SNP Discovery

Raw sequence reads were subjected to quality trimming using FastQC v0.11.3 (http://www.bioinformatics.babraham.ac.uk/projects/fastqc/), and reads with a Phred quality (Q) score <20 were discarded. Adapter trimming was carried out by using

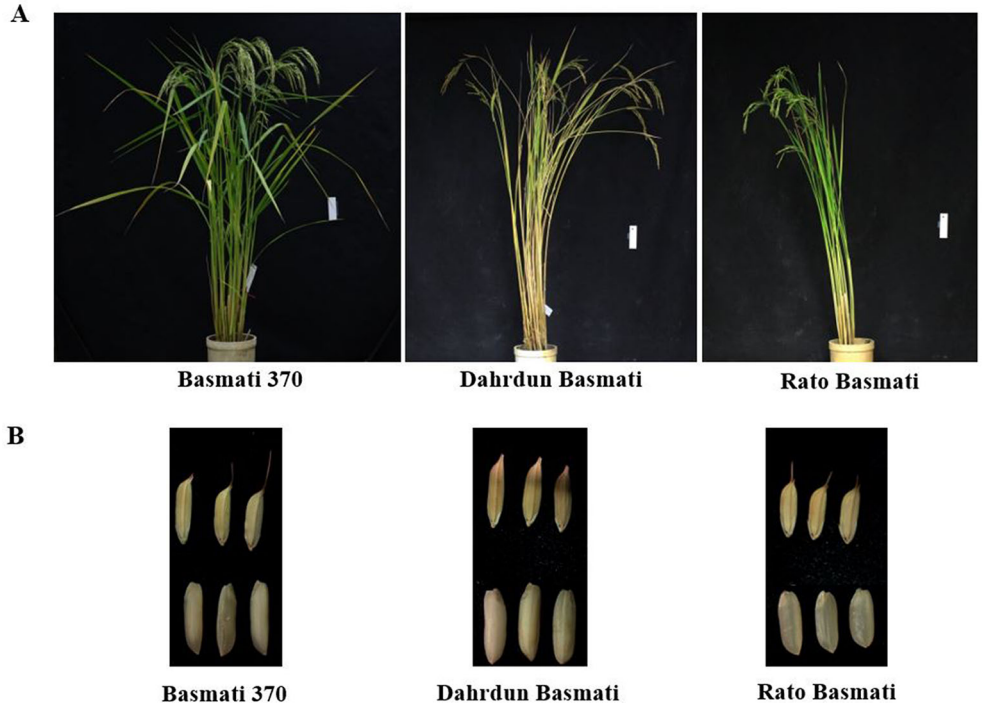


FIGURE 1 | Phenotypic comparison of Basmati 370, Dahrdun Basmati, and Rato Basmati. **(A)** Plant phenotype. **(B)** Spikelet and mature grain.

Trimmomatic (http://www.usadellab.org/cms/?page=trimmomatic). The clean reads were mapped to the reference genomes of the temperate *japonica* cultivar Nipponbare (Os-Nipponbare-Reference-IRGSP-1.0; Kawahara et al., 2013), *indica* cultivar Zhenshan 97 (Os-Zhenshan 97-Reference; Zhang et al., 2016), and *aus* cultivar Kasalath (Os-Kasalath-Reference; Sakai et al., 2014) using the Burrows–Wheeler Aligner (BWA) program (Li and Durbin, 2010). The alignment results were merged and converted into binary alignment map (BAM) files (Barnett et al., 2011). The BAM files were used to calculate the sequencing depth and to identify SNPs and InDels using the GATK program, with default parameters (McKenna et al., 2010).

Genomic Analysis

The genic and intergenic distribution of SNPs and InDels was determined relative to Nipponbare, Zhenshan 97, and Kasalath reference genomes. The distribution of genic SNPs and InDels common to the three Basmati genomes were presented using Circos (Krzywinski et al., 2009).

In silico analysis was performed to identify Basmati-specific SNPs and InDels using resequencing data of 54 diverse rice varieties in the Crop Molecular Breeding Lab, Seoul National University database (unpublished data) and Rice Variation Map v2.0 public database (http://ricevarmap.ncpgr.cn/v2/). InDel in nine traditional Basmati varieties and 11 *indica*, *aus*, and *japonica* check varieties were verified by gel electrophoresis, based on *in silico* analysis, using primers designed with Primer-3 (http://bioinfo.ut.ee/primer3-0.4.0/).

GO Analysis

The annotated Nipponbare, Zhenshan 97, and Kasalath reference genes were classified based on the pattern of common SNPs in the three Basmati genomes. Functional annotation of genes was investigated with "*Oryza sativa*" as the background species. GO analysis was performed using the BLAST2GO software (www.blast2go.com) (Conesa et al., 2005). Whole-genome orthologous gene comparison, annotation, and clustering were performed using the Orthovenn program (Wang et al., 2015).

DNA Extraction and Genome-Wide SNP Marker Analysis

Genomic DNA was isolated from the leaf tissues of plants at the 3–4 leaf stage using the modified cetyltrimethylammonium bromide (CTAB) method (McCouch et al., 1988). DNA concentration and quality were determined using the NanoDrop spectrophotometer (Thermo Scientific, Wilmington, NC, USA).

On the basis of differences in DNA sequences between *indica* and *japonica* genomes, 190 subspecies-specific SNP markers, representing all 12 rice chromosomes, were developed in the Crop Molecular Breeding Lab, Seoul National University (unpublished data). SNP genotyping was conducted on Fluidigm 96.96 Dynamic Arrays using the BioMark HD System (Fluidigm Corp, San Francisco, CA), according to the manufacturer's instructions, and genotypes were determined using the Fluidigm SNP Genotyping Analysis software.

Phylogenetic and Population Structure Analyses

Phylogenetic analysis was performed using PowerMarker v3.25 (Liu and Muse, 2005). Cavalli-Sforza and Edwards (1967) genetic distance was used to construct an unweighted pair group method with an arithmetic average (UPGMA) dendrogram, which was visualized in Molecular Evolutionary Genetics Analysis 7 (MEGA7) (Kumar et al., 2016). The population structure of 60 rice varieties was determined using a model-based approach available in the STRUCTURE 2.3.4 software (Falush et al., 2003). The number of genetically distinct populations (*K*) was adjusted from 1 to 10, and the model was repeated three times for each *K*. The burn in-period was adjusted with 100,000 iterations, followed by 100,000 Markov Chain Monte Carlo (MCMC) per run. The best *K* value was determined based on delta *K* (ΔK) using the Evanno method in the web-based python program, STRUCTURE HARVESTER (Earl and Vonholdt, 2012).

Badh2 Marker Analysis

All 60 rice varieties were classified as *badh2.1* and wild *Badh2* allele harboring genotypes by PCR-based genotyping of the *Badh2* InDel marker using the forward primer 5′-TGTTTTCTGTT AGGTTGCATT-3′ and reverse primer 5′-ATCCACAGAAA TTTGGAAAC-3′ (Sakthivel et al., 2009). PCR was conducted using the following conditions: initial denaturation at 94°C for 2 min, followed by 35 cycles of denaturation at 95°C for 20 s, annealing at 54°C for 30 s, and extension at 72°C for 30 s, and a final extension at 72°C for 1 min. The amplified products were separated by electrophoresis on 3.5% agarose gel.

Agronomic and Grain Quality Trait Analyses

Passport data on 13 agronomic and grain quality traits of 30 traditional Basmati varieties, including days to heading (DH), leaf width (LW), days to maturity (DM), culm length (CL), culm number (CN), culm diameter (CD), grain length (GL), grain width (GW), length-to-width ratio of grain (L/W ratio), 1,000 grain weight (KGW), panicle length (PL), spikelet fertility count (SFC), and amylose content (AC), were obtained from Genesys (https://www.genesys-pgr.org). Cluster analysis and Student's *t*-test were performed using SPSS 16.0 (https://www.ibm.com/analytics/spss-statistics-software).

RESULTS

Basmati Genome Sequencing

High-throughput sequencing of three traditional Basmati varieties was performed to facilitate downstream analysis. A total of 43,024,210 reads were generated from Basmati 370; 43,263,296 reads from Dahrdun Basmati; and 44,099,730 reads from Rato Basmati, each corresponding to more than 10 GB read length, and more than 90% of these reads were clean reads (**Table 2**). The clean reads were mapped to the reference genomes of Nipponbare (*japonica* rice), Zhenshan 97 (*indica* rice), and Kasalath (*aus* rice). The mapping results indicated that all genomes were sequenced at a depth ranging from 26.02X to 30.75X, with more than 90% coverage.

The number of SNPs in each Basmati variety were determined relative to each reference genome. Compared with Nipponbare, we identified 1,544,399 SNPs in Basmati 370; 2,105,019 SNPs in Dahrdun Basmati; and 1,229,155 SNPs in Rato Basmati. Similarly, comparison with the Kasalath reference genome revealed 1,453,259 SNPs in Basmati 370; 1,336,541 in Dahrdun Basmati; and 1,627,481 SNPs in Rato Basmati, whereas comparison with the Zhenshan 97 reference genome revealed 1,409,129 SNPs in Basmati 370; 793,929 SNPs in Dahrdun Basmati; and 1,659,254 SNPs in Rato Basmati. Thus, Dahrdun Basmati showed the highest number of SNPs compared with Nipponbare and the lowest number of SNPs compared with Zhenshan 97 (**Table 3**).

In comparison with the Nipponbare reference genome, relatively high numbers of SNPs were detected on chromosomes 1, 3, 6, and 11 in Basmati genomes, while the lowest numbers of SNPs were detected on chromosomes 9 and 5. Compared with Kasalath, Basmati varieties showed a high proportion of SNPs on chromosomes 1, 2, 3, 6, and 7, and the lowest numbers of SNPs on chromosome 10. Compared with Zhenshan 97, we found a high proportion SNPs on chromosomes 1, 2, 6, and 7 in Basmati varieties and lower SNP numbers on chromosomes 5 and 9. The distribution of SNPs on all 12 chromosomes of the three Basmati varieties in comparison with all three reference genomes is summarized in **Supplementary Table 1**.

Furthermore, we also determined the number of SNPs and InDels in each Basmati variety against the three reference genomes. Accordingly, in **Supplementary Table 2**. In comparison with Nipponbare, InDels were abundant on chromosomes 3 and 6 in Basmati varieties, while the number of SNPs was the highest on chromosome 1. In comparison with Kasalath and Zhenshan 97 reference genomes, chromosomes 1, 2, and 3 of Basmati varieties contained a high proportion of InDels, while chromosomes 1, 2, and 6 showed the highest number of substitutions.

Distribution of Common SNPs and InDels in Genic Regions

Common SNPs in genic regions, functional SNPs [non-synonymous SNPs and SNPs in untranslated regions (UTRs)],

TABLE 2 | Data generated from whole-genome resequencing of three Basmati varieties.

Varieties	Raw reads		Clean reads		Coverage (%)
	Read number	Read length (bp)	Read number	Read length (bp)	
Basmati 370	43,568,684	10,935,739,684	43,024,210	10,117,316,665	92.52
Dahrdun Basmati	43,936,332	11,028,019,332	43,263,296	9,971,121,538	90.42
Rato Basmati	44,616,386	11,198,712,886	44,099,730	10,236,348,518	91.41

TABLE 3 | Mapping and SNP summary of three traditional Basmati varieties.

Reference varieties	Basmati varieties	Mapping information					SNP data						
		Raw reads	Mapped reads	Unmapped reads	Average depth	Coverage (%)	Non-synonymous	Synonymous	Intron	5′UTR	3′UTR	Intergenic	Total
Nipponbare	Basmati 370	43,024,210	40,389,926	578,354	27.06	91.04	35,381	30,767	25,440	15,887	31,829	1,405,095	1,544,399
	Dahrdun Basmati	43,263,296	40,402,732	498,038	26.52	90.07	45,978	40,181	31,729	21,029	43,401	1,922,701	2,105,019
	Rato Basmati	44,099,730	42,127,262	405,096	27.57	92.07	28,484	24,751	19,335	12,643	25,573	1,118,369	1,229,155
Kasalath	Basmati 370	43,024,210	38,696,238	1,149,044	30.11	96.33	14,004	10,905	31,589	1,876	3,935	1,390,950	1,453,259
	Dahrdun Basmati	43,263,296	39,076,602	988,102	29.08	96.63	13,409	10,759	30,844	1,875	4,099	1,275,555	1,336,541
	Rato Basmati	44,099,730	40,369,808	969,672	30.75	96.08	14,796	11,394	33,791	1,929	4,225	1,561,346	1,627,481
Zhenshan 97	Basmati 370	43,024,210	38,936,652	1,673,302	27.06	91.79	95,093	62,744	89,502	23,344	44,958	1,093,488	1,409,129
	Dahrdun Basmati	43,263,296	39,734,380	1,442,176	26.02	94.90	54,599	36,577	50,150	13,471	25,394	613,738	793,929
	Rato Basmati	44,099,730	40,356,836	1,510,382	27.80	90.33	111,918	73,592	103,813	28,359	53,212	1,288,360	1,659,254

and InDels (5–30 bp) in genic regions were identified by comparing all three Basmati genomes with all three reference genomes. The total number of common SNPs identified in Basmati varieties were 52,204 compared with Nipponbare; 19,207 compared with Kasalath; and 73,219 compared with Zhenshan 97. The extracted common SNPs were plotted within the Nipponbare (**Figure 2A**), Kasalath (**Figure 2B**), and Zhenshan 97 (**Figure 2C**) reference genomes.

In addition, *in silico* analysis using resequencing data of 54 varieties revealed 20 novel unique SNPs in genic regions of the Basmati genomes. These unique SNPs were also confirmed using the public rice database (http://ricevarmap.ncpgr.cn/v2/). Additionally, we identified 11 unique InDels in the Basmati genomes. The unique SNPs and InDels, and the functions of genes containing these polymorphisms, are listed in **Supplementary Table 3**. PCR amplification of 289 bp fragments using gene-specific primers (forward primer, 5′-CTGTTTATACGTAGTACGGGTTG-3′; reverse primer, 5′-TGTTTGTAGGGGGATGCAAT-3′), which confirmed that the 25 bp insertion in the intron of the gene involved in seed development regulation (*Os10g0139300*; IRGSP-1.0; position: 2,425,049 bp) was only specific to the Basmati and *aus* groups, and could be discriminated among 20 rice varieties (**Supplementary Figure 2**). We further examined the spatiotemporal expression pattern of *Os10g0139300* in the RiceXpro database (Sato et al., 2011); this gene showed high expression in the embryo and endosperm after flowering, indicating a possible role in seed development during ripening.

GO Analysis of Basmati Varieties

We investigated the functions of genes containing common SNPs and InDels among the three Basmati genomes *via* GO analysis. Genes were assigned to three categories, namely, biological process (BP), molecular function (MF), and cellular component (CC). The major GO associations were found in metabolic process, cellular process, biological regulation for BP terms (**Figure 3A**). For the MF terms, binding and catalytic activity (**Figure 3B**). Whereas, cell, cell part, and membrane were associated with CC terms (**Figure 3C**).

Furthermore, we analyzed genome-wide orthologous clusters of genes from Basmati varieties using common SNPs by comparison with Nipponbare, Zhenshan 97, and Kasalath reference genomes. The analysis revealed 5,395 orthologous clusters based on protein sequences of the three reference genomes (**Figure 4A**). The Venn diagram showed that 1,132 gene clusters were shared by all three reference genomes, suggesting their conservation in the lineage after speciation (**Figure 4B**). Additionally, 348, 354, and 51 clusters specific to Nipponbare, Zhenshan 97, and Kasalath reference genomes, respectively, were identified. Additionally, cluster analysis of the mutated genes in the three Basmati varieties revealed 4,415 clusters in comparison with Nipponbare; 2,721 clusters in comparison with Kasalath; and 4,033 clusters in comparison with Zhenshan 97 reference genomes. The presence of 2,721 clusters in comparison with Kasalath suggests that Basmati varieties show less genetic variation compared with the *aus* group.

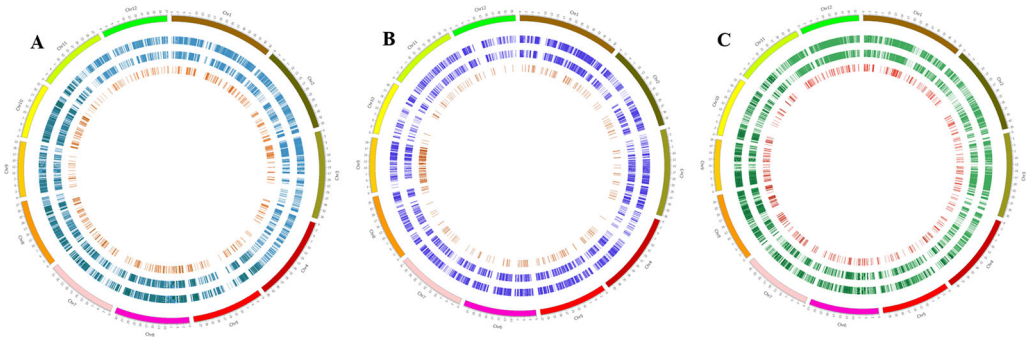


FIGURE 2 | Circos plots showing the distribution pattern of SNPs and InDels in the genic regions of three Basmati varieties. **(A–C)** Distribution of SNPs and InDels in Basmati varieties in comparison with Nipponbare **(A)**, Kasalath **(B)**, and Zhenshan 97 **(C)** reference genomes. The outermost circle represents 12 chromosomes of the rice genome. The second circle from the outside represents common SNPs. The third circle from the outside represents functional SNPs. The innermost circle with red bars shows the distribution of InDels ranging in size from 5 to 30 bp.

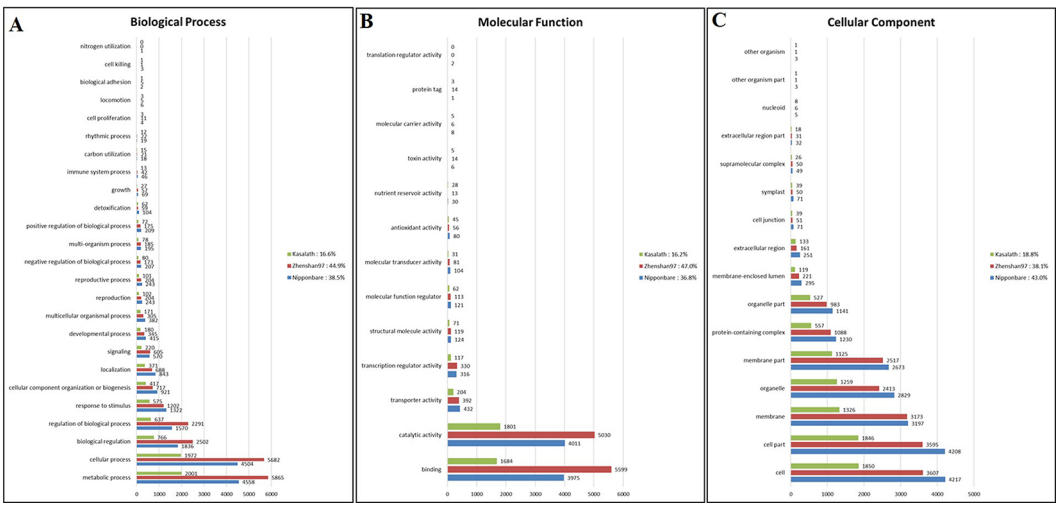


FIGURE 3 | Gene ontology (GO) analysis of Basmati genomes in comparison with Zhenshan 97, Kasalath, and Nipponbare genomes. **(A–C)** GO categories including biological process **(A)**, molecular function **(B)**, and cellular component **(C)** are shown.

In phylogenetic studies, the identification of single-copy orthologs is critical in any species (Creevey et al., 2011). Orthologous cluster analysis revealed 792 clusters representing single-copy genes, which were shared by all three reference genomes, suggesting that the single-copy status of genes was maintained during evolution after species divergence.

Furthermore, 1,132 gene clusters shared by Nipponbare, Kasalath, and Zhenshan 97 reference genomes harbored unique SNPs from all three Basmati varieties, and functional annotation of the genes harboring these unique SNPs showed that the majority of these genes were involved in biological regulation, metabolic process, and cellular process (**Figure 5A**); binding and catalytic activity (**Figure 5B**); and membrane, cell parts, and cellular component (**Figure 5C**). We also detected mutated gene clusters associated with the metabolic process involved in the cellular aromatic compound (**Figure 5A**). Further, a total of 35 genes including *Badh2* gene were found to be involved in aromatic compound biosynthesis based on biological process and molecular functional annotation. While, genomic regions from three Basmati varieties compared to Nipponbare reference genome showed functional variation across the 35 genes involved in aromatic compound biosynthesis (data not shown). Whereas, *in sillico* analysis of 35 genes using Rice Variation Map v2.0 revealed that only nine genes including *Badh2* gene having alternative alleles in 96 varieties of aromatic group with more than 80% of frequency.

SNP Genotyping and Genetic Relationship

To determine the genetic relationship of 30 traditional Basmati varieties, including three resequenced Basmati varieties, with rice varieties belonging to other groups, a total of 60 varieties were genotyped with two sets of 96-plex *indica/japonica* SNPs. Two of these SNP markers were excluded from the analysis because of

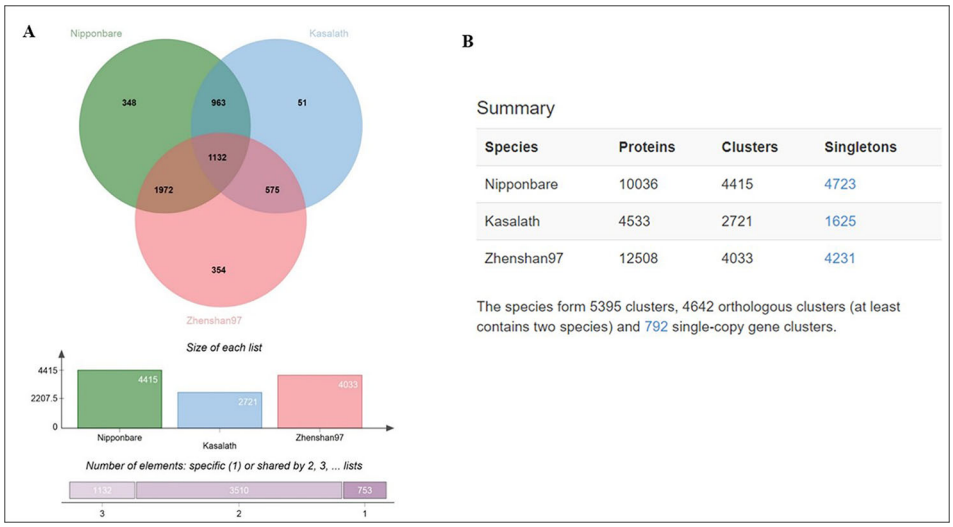


FIGURE 4 | Ortho Venn diagram. **(A)** Venn diagram showing the distribution of shared gene families among Nipponbare, Kasalath, and Zhenshan 97. Specific gene clusters are indicated within the three reference genomes. **(B)** Counts of clusters in each genome.

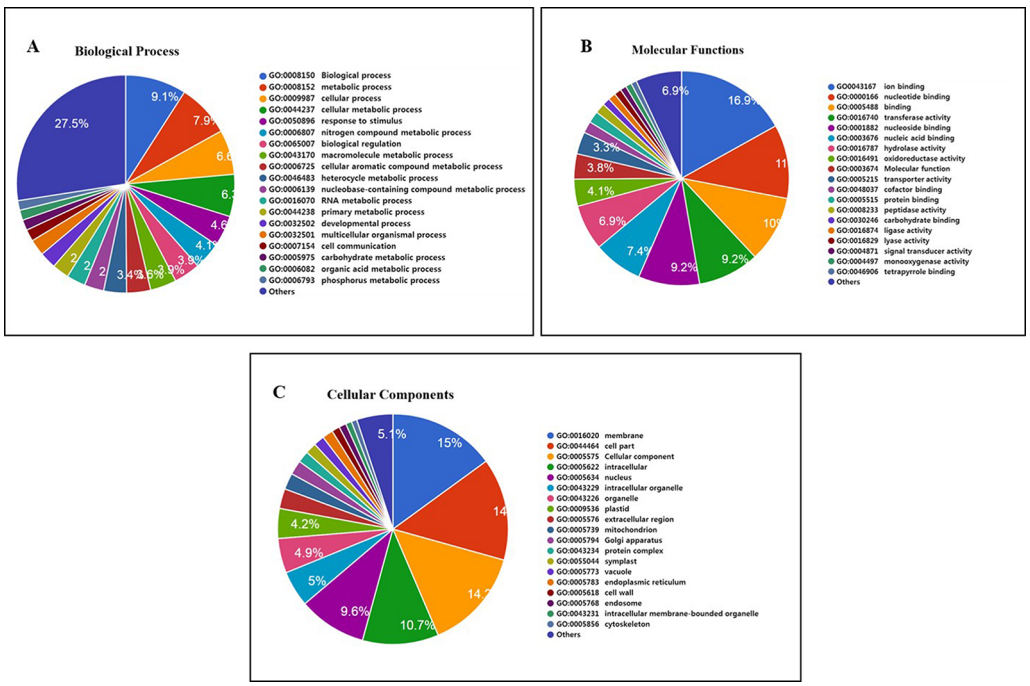


FIGURE 5 | Functional annotation of 1,132 mutated gene cluster of Basmati genome shared by Nipponbare, Kasalath, and Zhenshan 97 reference genomes. **(A)** Biological process. **(B)** Molecular function. **(C)** Cellular component.

their low quality. The number of SNP markers, average physical interval between SNPs per chromosome, and coverage percentage are summarized in **Supplementary Table 4**.

All 190 SNP markers were biallelic between *indica* and *japonica* varieties, and the average allele number was 2.12. In addition, the average value of major allele frequency (MAF) was 0.681, and almost all SNPs showed no heterozygosity (average heterozygosity = 0.020). Consistent with these data, the average polymorphic information content (PIC) was 0.33 (**Supplementary Table 5**).

The UPGMA dendrogram based on Cavalli-Sforza and Edwards (1967) genetic distance (**Figure 6A**) classified all 60 varieties into two subspecies, *indica* and *japonica*. Additionally, the *japonica*

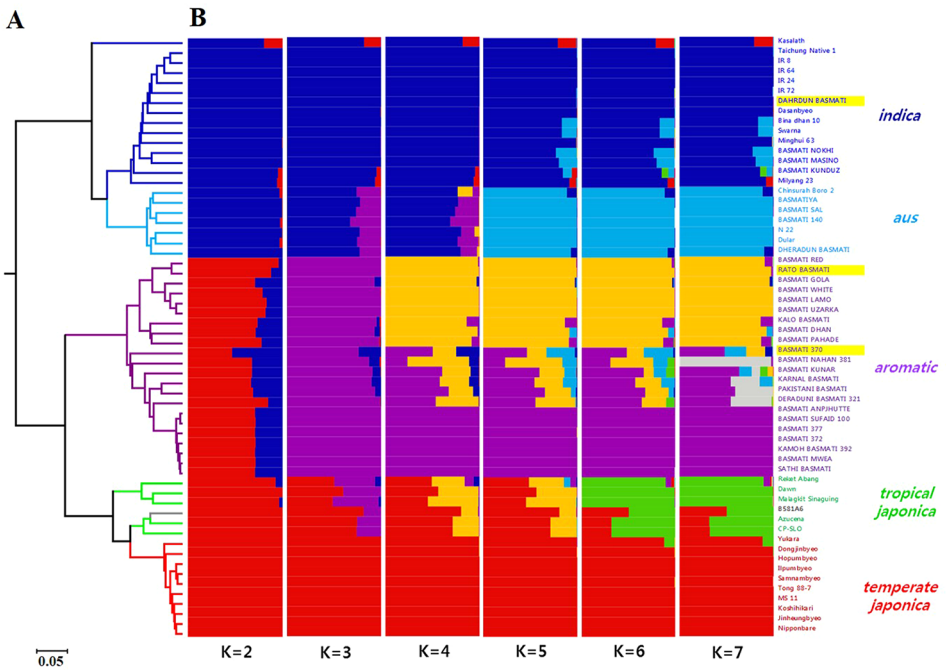


FIGURE 6 | Genetic diversity analysis of 60 rice varieties using 190 SNPs. **(A)** UPGMA dendrogram. The branches are colored according to the subpopulation assessment in **(B)** based on $K = 6$, except for the *aromatic* group, which is based on $K = 3$. Gray branches indicate admixture. **(B)** Population structure analysis using the STRUCTURE software for K values ranging from 2 to 7. Three varieties used for genomic analysis are highlighted in yellow.

group showed two distinct subgroups, *aromatic* and *japonica*. The 30 Basmati varieties were divided into two groups, *indica* (comprising Dahrdun Basmati) and *japonica* (comprising Rato Basmati and Basmati 370). To identify the population structure of all 60 rice varieties, STRUCTURE analysis was carried out. The value of delta K was maximum at $K = 2$ (**Supplementary Figure 3**). At $K = 2$, 60 varieties were classified into *indica* and *japonica*, as expected based on marker characteristics; however, more than half of the varieties in the *japonica* group showed admixture with *indica* ancestry. At $K = 3$, the *aromatic* group along with Rato Basmati and Basmati 370 grouped at the *japonica* group, and at $K = 4$, the *aromatic* group was divided into two clear subgroups and one admixed group. All nine varieties, including Rato Basmati, in the upper yellow subgroup within the *aromatic* group (**Figure 6B**), were from Nepal. At $K = 6$, five subgroups were evident among the 60 varieties including *indica*, *aus*, *aromatic*, tropical *japonica*, and temperate *japonica*, except one variety, which showed less than 65% of estimated ancestry derived from any single subgroup (**Figure 6B**). The results of phylogenetic and population structure analyses were consistent. Among the 30 traditional Basmati varieties, four varieties, including Dahrdun Basmati, were classified into the *indica* group; four into the *aus* group; and 22, including Rato Basmati and Basmati 370, into the *aromatic* group.

Badh2 Marker Analysis

Among 60 rice varieties, 30 traditional Basmati varieties were further investigated on the basis of the 8 bp deletion in the *Badh2* gene to classified into *badh2.1* and wild *Badh2* allele harboring genotypes. PCR-based genotyping of the *Badh2* InDel marker divided the traditional Basmati varieties into two groups: *badh2.1* (22 varieties; 95 bp PCR product) and wild *Badh2* allele carrying genotypes (8 varieties; 103 bp PCR product) (**Supplementary Figure 4**, **Table 1**). The remaining 30 non-Basmati rice varieties were classified in the wild *Badh2* allele group (**Table 1**).

Agronomic and Grain Quality Trait Analysis

The mean performance of 13 agronomic and grain quality traits of 30 traditional Basmati varieties is presented in **Supplementary Table 6**. The coefficient of variation of CN was the highest (25.49), followed by that of SFC (24.87). Comparison of the mean performance between *aromatic* and *indica/aus* groups revealed significant differences in only L/W ratio, PL, and AC; the *aromatic* group showed significantly longer panicles, longer and slender grains, and lower AC than the *indica/aus* group (**Supplementary Table 7**).

Next, hierarchical cluster analysis was performed to elucidate the relationship among the 30 traditional Basmati varieties. These varieties were divided into two major clusters (I and II), and each cluster was further divided into three subclusters (**Supplementary Figure 5**, **Supplementary Table 6**). Cluster I contained 20 moderate duration varieties from diverse geographical regions with superior grain quality. Cluster II consisted of ten late duration varieties, mostly from Nepal, with poor grain quality; thus cluster II showed less genetic diversity than cluster I.

DISCUSSION

Basmati rice varieties, considered a unique varietal group, have been generally classified into the *aromatic* group (Glaszmann, 1987; Garris et al., 2005; Civáň et al., 2015). Recent findings suggest that Basmati rice was derived mostly from *aus* and *japonica* varietal groups (Civáň et al., 2015). Recently, the genome assembly of Basmati rice was performed using "Basmati Surkh 89-15," an improved cultivar from Pakistan (Zhao et al., 2018). However, a higher level of introgression from other rice populations in improved varieties of Basmati makes it difficult to define the genome structure. A latest preprint of phylogenomic analysis involving "Basmati 334" proposed admixture events between Basmati rice, *aus*, and *O. rufipogon*; this study concluded that Basmati rice has a hybrid origin and is closely related to both *japonica* and *aus* rice (Choi et al., 2018). However, phylogenomic analysis using a single genome cannot provide detailed information about the Basmati genome structure, when referring to the entire Basmati group, irrespective of the *Badh2* allele type. Therefore, defining Basmati-specific genome features is important to understand the domestication of Asian rice.

In this study, we performed whole-genome resequencing and analysis of three traditional Basmati varieties. The identification of genome-wide nucleotide polymorphisms, including SNPs and InDels, using NGS has gained importance in the rice genome (Markkandan et al., 2018) and has enabled researchers to identify genome-specific features in rice varieties. Therefore, we performed NGS data analysis of Basmati 370, Dahrdun Basmati, and Rato Basmati to characterize the Basmati genome in detail. We found that millions of SNPs in all Basmati varieties in comparison with Nipponbare, Kasalath, and Zhenshan 97 reference genomes (**Table 3**), thus providing an opportunity to identify Basmati-specific features. Additionally, the genome-wide common SNPs and InDels identified in this study would serve as a useful resource for the development of SNP and InDel markers for the Basmati genome, specific to *japonica*, *aus*, and *indica* varietal groups (**Figures 2A–C**). Similarly, *in silico* analysis of the three Basmati rice genomes along with 54 rice varieties revealed high-quality Basmati-specific features.

Basmati rice varieties showed less genomic variation compared with the *aus* group and was phylogenetically close to the *japonica* group; these results are consistent with those of previous studies (Choi et al., 2018; Civáň et al., 2019). GO enrichment analysis also showed less genomic variation between the Basmati genome and the *aus* group in terms of GO categories. Most of the genes assigned to the three GO categories were mainly involved in metabolic process, cellular process, binding, catalytic activity, cell, and cell part. This functional annotation of genes is consistent with previous findings in rice (Kim et al., 2014; Liu et al., 2017). Additionally, our data showed that the metabolic process involved in the cellular aromatic compound was associated with the common mutated gene cluster (**Figure 5A**) and further analyses revealed that nine genes including *Badh2* gene having alternative allele's among aromatic group of rice varieties with more than 80% of frequency (**Supplementary Table 8**). However, possible involvement of these genes except *Badh2* remains to be determined for cellular aromatic biosynthesis.

A recent genomic analysis of a population of over 1,000 wild and cultivated rice accessions using genome-wide polymorphisms showed that Basmati rice arose from hybridization between *japonica* and wild rice related to the *aus* group (Civáň et al., 2019). Similarly, our comparative analysis of genome structure, based on genomic variation and GO analysis, showed that the Basmati genome is probably derived mostly from the *aus* and *japonica* groups.

Previously, it was shown that the recessive *fgr* allele encoding *Badh2* carries an 8 bp deletion and three SNPs in the seventh exon, resulting in the fragrant trait in Basmati varieties (Bradbury et al., 2005). Recently, haplotype analysis of the *Badh2* gene and analysis of 2-AP using 242 rice accessions classified two Basmati varieties harboring the wild *Badh2* allele under the *aus* and *indica* groups (Kovach et al., 2009). In this study, our comparative analysis found that both Basmati 370 and Rato Basmati carrying the *badh2.1* allele was consistent with the *badh2.1* allele reported by Kovach et al. (2009). Further, we genotyped the *Badh2* allele in the 30 Basmati varieties using the *Badh2* InDel maker developed by Sakthivel et al. (2009). The results indicated that 22 of the 30 traditional Basmati varieties belonging to the *aromatic* group carry the fragrant-inducing *badh2.1* allele and are more closely related to the *japonica* group. However, eight of the 30 Basmati varieties were harboring the wild *Badh2* allele under the *aus* and *indica* groups (**Supplementary Figure 4**, **Table 1**). Thus, the results of *Badh2* allele genotyping were consistent with those of phylogenetic analysis. We propose that classification of these wild *Badh2* allele carrying Basmati varieties under the *indica* and *aus* groups might results from either natural selection or human error during varietal diversification or germplasm collection.

The success of any crop breeding program depends on the magnitude of genetic variability within the germplasm (Kishor et al., 2016). In this study, although efforts were made to evaluate the agronomic and grain quality traits of 30 traditional Basmati varieties in the experimental field of Seoul National University, most of the Basmati varieties failed to flower in the rice growing season of the temperate region. By contrast, Basmati 370 and a few other wild *Badh2* allele carrying Basmati varieties were flowered, and their agronomic traits were evaluated for further studies in temperate regions (data not shown). Furthermore, agronomic and grain quality trait passport data obtained from the public database Genesys showed wide variation in most of the traits among the 30 traditional Basmati varieties (**Supplementary Table 6**). These finding are in agreement with previous genetic diversity studies in Basmati varieties (Lingaiah et al., 2014; Nirmaladevi et al., 2015). Most of the agronomic and grain quality traits, except L/W ratio, PL, and AC, did not show significant differences among Basmati varieties belonging to the *aromatic* and *indica/aus* groups (**Supplementary Table 7**). AC is an important factor affecting the palatability and grain quality of cooked rice (Tian et al., 2009). Rice grains with low AC (12–20%) are usually glossy, soft, and sticky after cooking, whereas those

with high AC (> 25%), generally found in Basmati varieties belonging to the *indica* group, exhibit a dry texture, remain separate, and are less tender upon cooking and become hard upon cooling (Bao et al., 2006).

Hierarchical cluster analysis revealed two major clusters (I and II) among the 30 traditional Basmati varieties, based on agronomic and grain quality traits (**Supplementary Figure 5**). Cluster I comprised varieties from diverse geographical regions, with moderate duration and superior grain qualities. By contrast, cluster II comprised of varieties with late duration and poor grain qualities. These findings are in accordance with a previous study where traditional Basmati varieties with superior agronomic and grain quality traits were grouped in a separate cluster (Roy et al., 2012). Accessions in cluster I with superior agronomic and grain quality could be exploited for the development of improved Basmati varieties in breeding programs.

In conclusion, our study provides a detailed analysis of the Basmati genome structure in comparison with *indica*, *japonica*, and *aus* genomes *via* whole-genome resequencing and genome-wide SNP marker analysis. This data will serve as an important resource for molecular breeding and genetic studies in Basmati rice.

DATA AVAILABILITY STATEMENT

All relevant raw sequence data are available in the NCBI Short Read Archive (SRA) database under the following BioProject accession numbers: Basmati [PRJNA551546], Dahrdun Basmati [PRJNA551547], and Rato Basmati [PRJNA551548].

AUTHOR CONTRIBUTIONS

DK and JS conceptualized the study, conducted formal analysis, determined the software for data analysis, and performed data visualization. DK and H-JK curated the data and determined the methodology and resources for this study. JS and JC performed data validation. H-JK acquired the funding and supervised the study. DK, JS, and JC wrote the first draft of the manuscript. All authors reviewed, edited, and approved the final manuscript.

FUNDING

This research was funded by the Rural Development Administration through the Next-generation BioGreen 21 Program (Grant No. PJ013165).

ACKNOWLEDGMENTS

We thank the Crop Molecular Breeding Lab at Seoul National University for their technical assistance.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00086/full#supplementary-material

REFERENCES

Ahuja, S. C., Panwar, D. V. S., Uma, A., and Gupta, K. R. (1995). *Basmati rice: the scented pearl* (Hisar: Directorate of Publications, CCS Haryana Agricultural University). doi: 10.13140/RG.2.2.26828.87681

Amarawathi, Y., Singh, R., Singh, A. K., Singh, V. P., Mohapatra, T., Sharma, T. R., et al. (2008). Mapping of quantitative trait loci for basmati quality traits in rice *(Oryza sativa* L.*)*. *Mol. Breed.* 21, 49–65. doi: 10.1007/s11032-007-9108-8

Bao, J., Shen, S., Sun, M., and Corke, H. (2006). Analysis of genotypic diversity in the starch physicochemical properties of nonwaxy rice: apparent amylose content, pasting viscosity and gel texture. *Starch - Stärke* 58 (6), 259–267. doi: 10.1002/star.200500469

Barnett, D. W., Garrison, E. K., Quinlan, A. R., Strömberg, M. P., and Marth, G. T. (2011). BamTools: a C++ API and toolkit for analyzing and managing BAM files. *Bioinformatics* 27 (12), 1691–1692. doi: 10.1093/bioinformatics/btr174

Bradbury, L. M. T., Fitzgerald, T. L., Henry, R. J., Jin, Q., and Waters, D. L. E. (2005). The gene for fragrance in rice. *Plant Biotechnol. J.* 3 (3), 363–370. doi: 10.1111/j.1467-7652.2005.00131.x

Cavalli-Sforza, L. L., and Edwards, A. W. (1967). Phylogenetic analysis. models and estimation procedures. *Am. J. Hum. Genet.* 19 (3 Pt 1), 233–257.

Chen, H., Xie, W., He, H., Yu, H., Chen, W., Li, J., et al. (2014). A high-density SNP genotyping array for rice biology and molecular breeding. *Mol. Plant* 7 (3), 541–553. doi: 10.1093/mp/sst135

Chen, E., Huang, X., Tian, Z., Wing, R. A., and Han, B. (2019). The genomics of Oryza species provides insights into rice domestication and heterosis. *Annu. Rev. Plant Biol.* 70, 639–665. doi: 10.1146/annurev-arplant-050718-100320

Choi, J. Y., Platts, A. E., Fuller, D. Q., Hsing, Y.-I., Wing, R. A., and Purugganan, M. D. (2017). The rice paradox: multiple origins but single domestication in Asian rice. *Mol. Biol. Evol.* 34 (4), 969–979. doi: 10.1093/molbev/msx049

Choi, J. Y., Groen, S. C., Zaaijer, S., and Purugganan, M. D. (2018). Nanopore sequence-based genome assembly of the basmati rice. *bioRxiv*, 396515. doi: 10.1101/396515

Civáň, P., Craig, H., Cox, C. J., and Brown, T. A. (2015). Three geographically separate domestications of Asian rice. *Nat. Plants* 1, 15164–15164. doi: 10.1038/nplants.2015.164

Civáň, P., Ali, S., Batista-Navarro, R., Drosou, K., Ihejieto, C., Chakraborty, D., et al. (2019). Origin of the aromatic group of cultivated rice (*Oryza sativa* L.) traced to the Indian Subcontinent. *Genome Biol. Evol.* 11 (3), 832–843. doi: 10.1093/gbe/evz039

Conesa, A., Gotz, S., Garcia-Gomez, J. M., Terol, J., Talon, M., and Robles, M. (2005). Blast2GO: a universal tool for annotation, visualization and analysis in functional genomics research. *Bioinformatics* 21 (18), 3674–3676. doi: 10.1093/bioinformatics/bti610

Creevey, C. J., Muller, J., Doerks, T., Thompson, J. D., Arendt, D., and Bork, P. (2011). Identifying single copy orthologs in Metazoa. *PloS Comput. Biol.* 7 (12), e1002269. doi: 10.1371/journal.pcbi.1002269

Daygon, V. D., Calingacion, M., Forster, L. C., et al. (2017). Metabolomics and genomics combine to unravel the pathway for the presence of fragrance in rice. *Sci. Rep.* 7, 8767. doi: 10.1038/s41598-017-07693-9

Earl, D. A., and Vonholdt, ,. B. M. (2012). Structure harvester: a website and program for visualizing STRUCTURE output and implementing the Evanno method. *Cons Genet. Res.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Falush, D., Stephens, M., and Pritchard, J. K. (2003). Inference of population structure using multilocus genotype data: linked loci and correlated allele frequencies. *Genetics* 164 (4), 1567–1587.

Fuller, D. Q., Sato, Y.-I., Castillo, C., Qin, L., Weisskopf, A. R., Kingwell-Banham, E. J., et al. (2010). Consilience of genetics and archaeobotany in the entangled history of rice. *Archaeol. Anthropological Sci.* 2 (2), 115–131. doi: 10.1007/s12520-010-0035-y

Garris, A. J., Tai, T. H., Coburn, J., Kresovich, S., and McCouch, S. (2005). Genetic structure and diversity in *Oryza sativa* L. *Genetics* 169 (3), 1631–1638. doi: 10.1534/genetics.104.035642

Glaszmann, J. C. (1987). Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet.* 74 (1), 21–30. doi: 10.1007/bf00290078

Huang, X., Kurata, N., Wei, X., Wang, Z.-X., Wang, A., Zhao, Q., et al. (2012). A map of rice genome variation reveals the origin of cultivated rice. *Nature* 490, 497. doi: 10.1038/nature11532

Jimenez, F. R., Maughan., P. J., Alvarez., A., Kietlinski., K. D., Smith., S. M., Pratt, D. B., et al (2013). Assessment of genetic diversity in Peruvian Amaranth (*Amaranthus caudatus* and *A. hybridus*) germplasm using single nucleotide polymorphism markers. *Crop Sci.* 53, 532–541. doi: 10.2135/cropsci2012.07.0413

Kato, S. (1928). On the affinity of rice varieties as shown by fertility of hybrid plants. *Bull. Sci. Fac. Agric. Kyushu Univ.* 3, 132–147.

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 (1), 4. doi: 10.1186/1939-8433-6-4

Kim, B., Kim, D.-G., Lee, G., Seo, J., Choi, I.-Y., Choi, B.-S., et al. (2014). Defining the genome structure of 'Tongil' rice, an important cultivar in the Korean "Green Revolution". *Rice (New York N.Y.)* 7 (1), 22–22. doi: 10.1186/s12284-014-0022-5

Kishor, D. S., Krishnan, A., Kebbahalli Jayaramu, Y., Kumar, V., Yashoda,, and Jong Koh, H. (2016). Genotypic variation among okra *(Abelmoschus esculentus (L.) Moench)* germplasms in South India. *Plant Breed. Biotechnol.* 1, 234–241. doi: 10.9787/PBB.2016.4.2.234

Kovach, M. J., Sweeney, M. T., and McCouch, S. R. (2007). New insights into the history of rice domestication. *Trends In Genet.* 23 (11), 578–587. doi: 10.1016/j.tig.2007.08.012

Kovach, M. J., Calingacion, M. N., Fitzgerald, M. A., and McCouch, S. R. (2009). The origin and evolution of fragrance in rice (*Oryza sativa* L.). *Proc. Natl. Acad. Sci. U S A* 106 (34), 14444–14449. doi: 10.1073/pnas.0904077106

Krzywinski, M., Schein, J., Birol, I., Connors, J., Gascoyne, R., Horsman, D., et al. (2009). Circos: an information aesthetic for comparative genomics. *Genome Res.* 19 (9), 1639–1645. doi: 10.1101/gr.092759.109

Kumar, S., Stecher, G., and Tamura, K. (2016). MEGA7: molecular evolutionary genetics analysis version 7.0 for bigger datasets. *Mol. Biol. Evol.* 33 (7), 1870–1874. doi: 10.1093/molbev/msw054

Li, H., and Durbin, R. (2010). Fast and accurate long-read alignment with Burrows-Wheeler transform. *Bioinformatics* 26 (5), 589–595. doi: 10.1093/bioinformatics/btp698

Lingaiah, N., Venkanna, V., and Cheralu, C. (2014). Genetic variability analysis in rice (*Oryza sativa* L.). *Int. J. Pure App. Biosci.* 2 (5), 203–204.

Liu, K., and Muse, S. V. (2005). PowerMarker: an integrated analysis environment for genetic marker analysis. *Bioinformatics* 21 (9), 2128–2129. doi: 10.1093/bioinformatics/bti282

Liu, W., Ghouri, F., Yu, H., Li, X., Yu, S., Shahid, M. Q., et al. (2017). Genome wide re-sequencing of newly developed Rice Lines from common wild rice (*Oryza rufipogon* Griff.) for the identification of NBS-LRR genes. *PloS One* 12 (7), e0180662. doi: 10.1371/journal.pone.0180662

Markkandan, K., Yoo, S.-I., Cho, Y.-C., and Lee, D. (2018). Genome-wide identification of insertion and deletion markers in Chinese commercial rice cultivars, based on next-generation sequencing data. *Agronomy* 8 (4), 36. doi: 10.3390/agronomy8040036

McCouch, S. R., Kochert, G., Yu, Z. H., Wang, Z. Y., Khush, G. S., Coffman, W. R., et al. (1988). Molecular mapping of rice chromosomes. *Theor. Appl. Genet.* 76 (6), 815–829. doi: 10.1007/bf00273666

McKenna, A., Hanna, M., Banks, E., Sivachenko, A., Cibulskis, K., Kernytsky, A., et al. (2010). The genome analysis toolkit: a MapReduce framework for analyzing next-generation DNA sequencing data. *Genome Res.* 20 (9), 1297–1303. doi: 10.1101/gr.107524.110

Nirmaladevi, G., Padmavathi, G., Kota, S., and Babu, V. (2015). Genetic variability, heritability and correlation coefficients of grain quality characters in rice (*Oryza sativa L.*). *SABRAO J. Breed. Genet.* 47 (4), 424–433.

Pachauri, V., Mishra, V., Mishra, P. P., Singh, A., Singh, S., Singh, R., et al. (2014). Identification of candidate genes for rice grain aroma by combining QTL mapping and transcriptome profiling approaches. *Cereal Res. Commun.* 42 (3), 376–388. doi: 10.1556/CRC.42.2014.3.2

Roy, S., Banerjee, A., K. Senapati, B., and Sarkar, G. (2012). Comparative analysis of agro-morphology, grain quality and aroma traits of traditional and Basmati-type genotypes of rice, Oryza sativa L. *Plant Breed*. 131, 486–492. doi: 10.1111/j.1439-0523.2012.01967.x

Sakai, H., Kanamori, H., Arai-Kichise, Y., Shibata-Hatta, M., Ebana, K., Oono, Y., et al. (2014). Construction of pseudomolecule sequences of the aus rice cultivar Kasalath for comparative genomics of Asian cultivated rice. *DNA Res.* 21 (4), 397–405. doi: 10.1093/dnares/dsu006

Sakthivel, K., Shobha Rani, N., Pandey, M. K., Sivaranjani, A. K. P., Neeraja, C. N., Balachandran, S. M., et al. (2009). Development of a simple functional marker for fragrance in rice and its validation in Indian Basmati and non-Basmati fragrant rice varieties. *Mol. Breed.* 24 (2), 185–190. doi: 10.1007/s11032-009-9283-x

Sang, T., and Ge, S. (2007). Genetics and phylogenetics of rice domestication. *Curr. Opin. In Genet. Dev.* 17 (6), 533–538. doi: 10.1016/j.gde.2007.09.005

Sato, Y., Antonio, B. A., Namiki, N., Takehisa, H., Minami, H., Kamatsuki, K., et al. (2011). RiceXPro: a platform for monitoring gene expression in japonica rice grown under natural field conditions. *Nucleic Acids Res.* 39 (Database issue), D1141–D1148. doi: 10.1093/nar/gkq1085

Serba, D., Muleta, K., Amand, P., Bernardo, A., Perumal, R., and Bashir, E. (2019). Genetic diversity, population structure, and linkage disequilibrium of pearl millet. *Plant Genome.* 12, 180091. doi: 10.3835/plantgenome2018.11.0091

Siddiq, E. A., Vemireddy, L. R., and Nagaraju, J. (2012). Basmati rices: genetics, breeding and trade. *Agric. Res.* 1 (1), 25–36. doi: 10.1007/s40003-011-0011-5

Tian, Z., Qian, Q., Liu, Q., Yan, M., Liu, X., Yan, C., et al. (2009).). Allelic diversities in rice starch biosynthesis lead to a diverse array of rice eating and cooking qualities. *Proc. Natl. Acad. Sci. U S A* 106 (51), 21760–21765. doi: 10.1073/pnas.0912396106

Vemireddy, L. R., Noor, S., Satyavathi, V., et al. (2015). Discovery and mapping of genomic regions governing economically important traits of Basmati rice. *BMC Plant Biol.* 15, 207. doi: 10.1186/s12870-015-0575-5

Wang, Y., and Li, J. (2005). The plant architecture of rice (Oryza sativa). *Plant Mol. Biol.* 59 (1), 75–84. doi: 10.1007/s11103-004-4038-x

Wang, Y., Coleman-Derr, D., Chen, G., and Gu, Y. Q. (2015). OrthoVenn: a web server for genome wide comparison and annotation of orthologous clusters across multiple species. *Nucleic Acids Res.* 43 (W1), W78–W84. doi: 10.1093/nar/gkv487

Wang, W., Mauleon, R., Hu, Z., Chebotarov, D., Tai, S., Wu, Z., et al. (2018). Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature* 557 (7703), 43–49. doi: 10.1038/s41586-018-0063-9

Zhang, J., Chen, L.-L., Sun, S., Kudrna, D., Copetti, D., Li, W., et al. (2016). Building two indica rice reference genomes with PacBio long-read and Illumina paired-end sequencing data. *Sci. Data* 3, 160076. doi: 10.1038/sdata.2016.76

Zhao, Q., Feng, Q., Lu, H., Li, Y., Wang, A., Tian, Q., et al. (2018). Pan-genome analysis highlights the extent of genomic variation in cultivated and wild rice. *Nat. Genet.* 50 (2), 278–284. doi: 10.1038/s41588-018-0041-z

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 Genetics

REVIEW
published: 21 February 2020
doi: 10.3389/fgene.2019.01392

Molecular Breeding for Nutritionally Enriched Maize: Status and Prospects

Boddupalli M. Prasanna[1], Natalia Palacios-Rojas[2], Firoz Hossain[3], Vignesh Muthusamy[3], Abebe Menkir[4], Thanda Dhliwayo[2], Thokozile Ndhlela[5], Felix San Vicente[2], Sudha K. Nair[6], Bindiganavile S. Vivek[6], Xuecai Zhang[2], Mike Olsen[1] and Xingming Fan[7]*

[1] International Maize and Wheat Improvement Center (CIMMYT), Nairobi, Kenya, [2] CIMMYT, Texcoco, Mexico, [3] ICAR-Indian Agricultural Research Institute (IARI), New Delhi, India, [4] International Institute of Tropical Agriculture (IITA), Ibadan, Nigeria, [5] CIMMYT, Harare, Zimbabwe, [6] CIMMYT, ICRISAT, Hyderabad, India, [7] Institute of Crop Sciences, Yunnan Academy of Agricultural Sciences (YAAS), Kunming, China

OPEN ACCESS

Edited by:
Mallikarjuna Swamy, International Rice Research Institute, Philippines

Reviewed by:
Mahalingam Govindaraj, International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India
Senthil Natesan, Tamil Nadu Agricultural University, India

***Correspondence:**
Boddupalli M. Prasanna
b.m.prasanna@cgiar.org

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *31 October 2019*
Accepted: *19 December 2019*
Published: *21 February 2020*

Citation:
Prasanna BM, Palacios-Rojas N, Hossain F, Muthusamy V, Menkir A, Dhliwayo T, Ndhlela T, San Vicente F, Nair SK, Vivek BS, Zhang X, Olsen M and Fan X (2020) Molecular Breeding for Nutritionally Enriched Maize: Status and Prospects. Front. Genet. 10:1392. doi: 10.3389/fgene.2019.01392

Maize is a major source of food security and economic development in sub-Saharan Africa (SSA), Latin America, and the Caribbean, and is among the top three cereal crops in Asia. Yet, maize is deficient in certain essential amino acids, vitamins, and minerals. Biofortified maize cultivars enriched with essential minerals and vitamins could be particularly impactful in rural areas with limited access to diversified diet, dietary supplements, and fortified foods. Significant progress has been made in developing, testing, and deploying maize cultivars biofortified with quality protein maize (QPM), provitamin A, and kernel zinc. In this review, we outline the status and prospects of developing nutritionally enriched maize by successfully harnessing conventional and molecular marker-assisted breeding, highlighting the need for intensification of efforts to create greater impacts on malnutrition in maize-consuming populations, especially in the low- and middle-income countries. Molecular marker-assisted selection methods are particularly useful for improving nutritional traits since conventional breeding methods are relatively constrained by the cost and throughput of nutritional trait phenotyping.

Keywords: biofortification, quality protein maize, provitamin A, kernel zinc, vitamin E

INTRODUCTION

Maize and its products constituted 30% of the food supply in the Americas, 38% in Africa, and 6.5% in Asia, and thus, is a major source of food security and economic development. Maize is major staple food and the most important energy source in sub-Saharan Africa (SSA) with intakes ranging from 50 to >330 g/person/day, and providing daily energy, protein, and micronutrients. In Latin America, maize consumption ranges from 50 to >300 g/person/day. Additionally, maize is part of the livestock-to-meat cycle across the world (Tanumihardjo et al., 2019). In addition to calories, maize is a source of micronutrients and phytochemicals, such as phenolics, carotenoids (yellow and orange maize), anthocyanins (blue, purple, and black maize), phlobaphenes (red maize), insoluble and soluble dietary fiber, and polar and non-polar lipids, providing health benefits and helping prevent diseases (**Figure 1**).

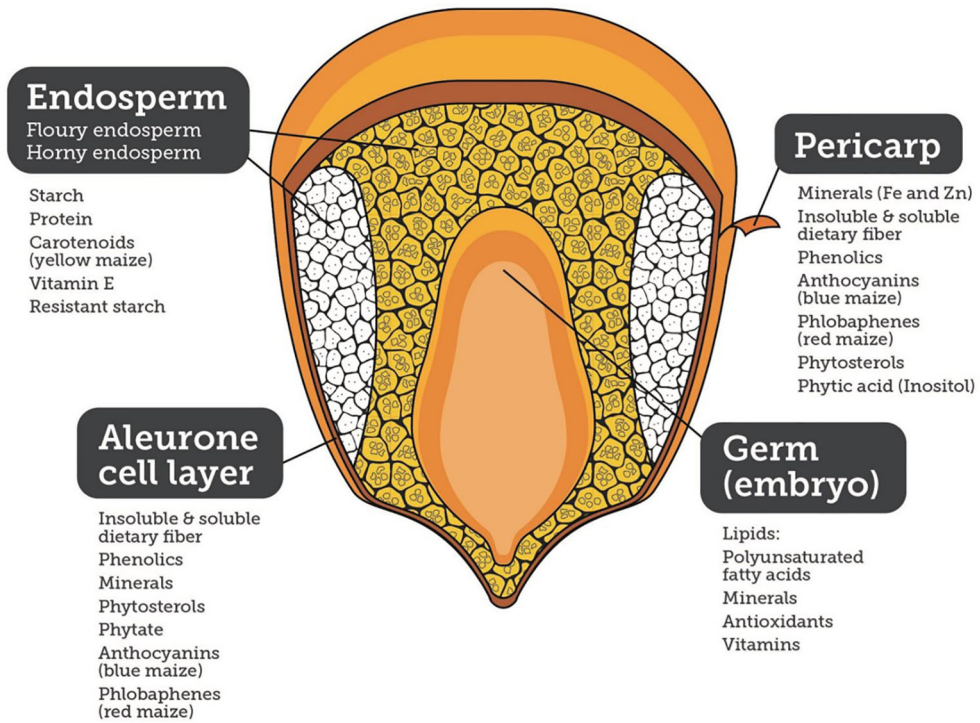


FIGURE 1 | Nutritional quality of different components of a maize kernel.

High per capita consumption of maize and limited diet diversification in several countries across Africa, Latin America, and Asia signifies that the greater part of people's diets in these countries lack some of the essential micronutrients, such as zinc (Zn), vitamin A, and vitamin E, as well as essential amino acids, such as lysine and tryptophan. More than two billion people in Asia, Africa, and Latin America suffer from one or more micronutrient deficiencies that commonly lead to retarded physical growth, impaired cognitive development, complications during pregnancy, diminished work and income earning capacity, and increased risk of morbidity and mortality (Bailey et al., 2015; Rautiainen et al., 2016).

The concentrations of various nutrients in maize kernels depend on the genetic background or the genotype, agronomic management, interaction between genotype and the environment, and post-harvest handling (Ekpa et al., 2019). Although the concentrations of most micronutrients in commonly used maize worldwide are not enough to have a nutritional impact on consumers, there is large genetic variation in maize that allows development of improved cultivars with higher concentrations of certain micronutrients, through biofortification (Bouis and Saltzman, 2017). Additional or complementary technologies in crop management and food science can also contribute to enhancing the nutritional impact of maize-based diets (Nuss and Tanumihardjo, 2011; Ekpa et al., 2018).

Maize scientists have been developing improved cultivars with enhanced nutritional value, such as quality protein maize (QPM) rich in two essential amino acids (Prasanna et al., 2001; Atlin et al., 2011); orange maize biofortified with provitamin A carotenoids (Pixley et al., 2013); and high-Zn-enhanced maize (Andersson et al., 2017). Higher content of lysine and tryptophan, kernel Zn, and provitamin A have been successfully increased in maize through conventional breeding. In SSA, the spread of QPM cultivars has been faster than in Asia, mainly because maize is predominantly used as food in SSA. In Asian countries including India, maize grains are utilized more as poultry feed (~60–70% of the produce), and synthetic lysine and tryptophan are added as supplements. Besides grain corn, nutritional enrichment of sweet corn is another possibility. Sweet corn has emerged as an important source of income for farmers in Asia; US$1,034 million worth of preserved sweet corn is imported globally, while the same for frozen sweet corn was $423 million (FAOSTAT, 2017).

There are now significant opportunities for more effectively developing nutritionally enriched cultivars of both grain and specialty corn, due to various factors, including availability of large genetic diversity for the target traits, advances in understanding key biochemical pathways for metabolite biosynthesis, analytical tools for screening germplasm for quality traits, and the possibilities to utilize molecular markers and genome editing approaches to accelerate product development (Reynolds et al., 2019). In this review, we have highlighted the recent advances in breeding for nutritionally enriched maize, especially in the tropics.

MAIZE WITH ENHANCED PROTEIN QUALITY

Extent of Quality Protein Deficiency

Essential amino acids such as lysine and tryptophan not only act as building blocks of proteins but also serve as neurotransmitters. The recommended daily allowance of lysine is 30 mg kg^{-1} body weight for adults, and 35 mg kg^{-1} body weight for children. As regards tryptophan, the daily requirements are 4 and 4.8 mg kg^{-1} body weight per day for adults and children, respectively (WHO/FAO/UNU, 2007). Deficiency of these amino acids leads to reduced appetite, delayed growth, impaired skeletal development, and aberrant behavior (Tome and Bos, 2007).

Several smallholders and their families in the low- and middle-income countries in SSA and Latin America are dependent on maize not only for their calorie requirement but also for dietary protein. Such populations, along with their monogastric livestock, run the risk of incurring health problems associated with amino acid deficiency, as maize is deficient in two essential amino acids, lysine and tryptophan. Improving the quality of maize endosperm protein by increasing its lysine and tryptophan content has, therefore, been a goal of maize breeding programs of the International Maize and Wheat Improvement Center (CIMMYT), the International Institute of Tropical Agriculture (IITA), and several national programs. This has led to development of an array of QPM cultivars having approximately twice the content of tryptophan (0.07–0.08% in flour) and lysine (0.25–0.40% in flour) compared to conventional maize cultivars (tryptophan: 0.03–0.04% in flour; lysine: 0.15–0.20% in flour), and consequently having greatly improved nutritional quality (Bressani, 1992; Bjarnason and Vasal, 1992; Zarkadas et al., 1995; Prasanna et al., 2001).

Quality Protein Maize Genetics and Breeding

Development of QPM cultivars involves manipulating three distinct genetic systems (Bjarnason and Vasal, 1992; Prasanna et al., 2001): 1) the simple recessive allele of *opaque2* (*o2*) gene in homozygous condition; 2) modifiers/enhancers of the *o2o2*-containing endosperm to confer higher lysine and tryptophan; and 3) genes that modify the *opaque2*-induced soft endosperm to hard endosperm. Phenotypic selection followed by biochemical analysis are required to select desirable genotypes that combine the three systems. Genomic screening methods may be used to enhance the efficiency of this selection process (Babu et al., 2005).

The deployment of *o2* located on chromosome 7 along with the endosperm modifiers led to the successful commercialization of diverse QPM hybrids worldwide with enhancement of both lysine (from 1.6–2.6 to 2.7–4.5% in protein) and tryptophan (from 0.2–0.6 to 0.5–1.1% in protein) (Vivek et al., 2008). The search for a novel mutation that can be successfully utilized to develop high lysine maize continued in the new millennium until Yang et al. (2005) reported another recessive mutant from Robertson's Mutator stocks and named it as *opaque16* (*o16*). Simple sequence repeats (SSRs), *umc1141* and *umc1149* were identified as the closely linked markers to *o16* (located on chromosome 8) using F_2 mapping population developed between Chinese inbreds, QCL3024 (*o16*) and QCL3010 (wild type). *o2o2/o16o16* was reported to increase lysine by 30% over *o2o2* or *o16o16* alone (Yang et al., 2005). Sarika et al. (2017) studied two F_2 populations derived by crossing wild type (CML533 and CML537) and *o16*-donor line (QCL3024). Genotypes with *o16o16* possessed on average nearly two-fold more lysine (0.247% in flour) and tryptophan (0.072% in flour) compared to normal maize (lysine 0.125% and tryptophan 0.035%, in flour), although the wide variation for the two traits across populations (lysine: 0.111–0.376% in flour; tryptophan: 0.027–0.117% in flour) suggested the possible influence of modifier loci.

Yang et al. (2013) reported that average lysine in *o16o16*–based BC_2F_4 seeds was 0.352%; some of the segregants (*o16o16*) possessed comparable lysine and tryptophan usually observed in *o2o2* genotypes, thereby suggesting that *o16* can be used as replacement to *o2* in the QPM breeding program. Sarika et al. (2018a) reported that the seed endosperm of *o16o16* was vitreous and phenotypically similar to wild type (*O16O16*). The mutant did not influence the degree of kernel opaqueness in *o2o2* genetic background as opaqueness in *o2o2/O16O16* and *o2o2/o16o16* was similar. Grain hardness of *o16o16* was comparable with the normal and QPM maize. The pattern of microscopic organization of proteinaceous matrix and starch granules, and zein profiling of the storage protein in *o16o16* were found to be similar with normal maize endosperm, but distinct from the *o2o2*-soft genotype.

Molecular Breeding for Developing Improved Quality Protein Maize Cultivars

A strong recognition of the relevance of QPM came through the award of World Food Prize to Surinder Vasal and Evangelina Villegas in the year 2000, leading to a resurgence in QPM breeding and release of cultivars [both open-pollinated varieties (OPVs) and hybrids] across SSA, Asia, and Latin America. In Asia, more than 40 QPM cultivars have been developed and released through conventional breeding, with India, China, Indonesia, and Vietnam having the highest number of releases. Several of the QPM cultivars were initially identified through CIMMYT International Trials distributed to partners, while national programs in India and China further released an array of QPM cultivars with introgression of the QPM trait in commercially relevant genetic backgrounds.

In India, marker-assisted backcross breeding (MABB) for *o2* led to the development and release of a single-cross QPM hybrid, 'Vivek QPM-9' in 2008 (Gupta et al., 2013). It possessed 41% more tryptophan and 30% more lysine over the original hybrid (Vivek Hybrid-9). Later, *o2* allele was introgressed into the parental inbreds of three popular non-QPM hybrids. These hybrids *viz.*, 'Pusa HM-4 Improved', 'Pusa HM-8 Improved', and 'Pusa HM-9 Improved' have been released in 2017 for commercial cultivation in India (Hossain et al., 2018a), possessing nearly double the concentrations of lysine and tryptophan as compared to normal maize. These four hybrids

are in the flint background, and did not show any yield penalty over the original hybrids (Yadava et al., 2018). At CIMMYT several inbreds were converted to QPM versions (e.g., CML244Q, CMl246Q, CML349Q, and CML354Q) using MABB; the grain yield performance of these QPM versions were at par with the original versions. Several institutions under the Indian Council of Agricultural Research (ICAR) and State Agricultural Universities (SAUs) are now using marker-assisted selection (MAS)-based breeding methods for developing new QPM cultivars (Hossain et al., 2019).

The Institute of Crop Science, Chinese Academy of Agricultural Sciences (CAAS) developed diverse QPM inbred lines by marker-assisted backcrossing in different genetic backgrounds (Tian et al., 2004; Jiang et al., 2005). Zhang et al. (2010) combined *o2* and *o16*, and reported an enhancement of 23% lysine in *o2o2/o16o16* progeny over the *o2o2* inbred comparison. Zhang et al. (2013) further pyramided *o2* and *o16* in waxy genetic background, and reported higher accumulation of lysine (0.616% in flour) in the pyramided lines compared to *o2o2* segregants (0.555% in flour). Yang et al. (2013) introgressed the *o16* allele from QCL3024 into two Chinese waxy lines, QCL5019 and QCL5008 using MAS. The *o16o16*-based waxy inbreds possessed 16–27 and 18–28% higher lysine than the waxy parents, respectively. Liu et al. (2016) developed the functional marker qγ27 linked with endosperm modification, which provides an important technical support in breeding QPM, although laboratory analysis is still needed for quantification of lysine and tryptophan.

Marker-assisted breeding has also been initiated in India to combine both *o2* and *o16* for developing QPM hybrids with enhanced levels of lysine and tryptophan (Sarika et al., 2018b; Chand et al., 2019). Sarika et al. (2018b) introgressed *o16* into the parental lines of four commercial QPM hybrids (HQPM-1, HQPM-4, HQPM-5, and HQPM-7) released in India, using MABB. Reconstituted hybrids showed an average enhancement of 49 and 60% in lysine and tryptophan over the original hybrids, with highest enhancement amounting 64 and 86%, respectively. The degree of endosperm modification in these hybrids was similar to the original QPM hybrids. The grain yield potential of these *o2o2/o16o16* hybrids was also at par with the original *o2o2*-based released hybrids, indicating that *o16* could be a novel genetic resource for enhancing lysine and tryptophan without influencing the degree of kernel opaqueness and grain yield potential.

PROVITAMIN A-ENRICHED MAIZE

Extent of Vitamin a Deficiency

Vitamin A plays a vital role in vision, and lack of it causes night blindness and partial or even complete loss of eyesight in humans. An adult non-pregnant and non-lactating woman requires 500 μg day^{-1} of vitamin A, while children of 4–6 years require 275 μg day^{-1} (Andersson et al., 2017). SSA has recorded the highest rates of vitamin A deficiency or VAD (48%), followed by South Asia (42%) (UNICEF, 2016). Pregnant women, breast-feeding mothers, and their children younger than 5 years of age are at the highest risk of having VAD (Bailey et al., 2015). In some countries in SSA, VAD in the most vulnerable populations is associated with reduced immune response, which can lead to increased infections, such as diarrhea, measles, or respiratory infections, which may either decrease provitamin A intake through reduced appetite or deplete existing vitamin A stores through excessive metabolism (Alvarez et al., 1995; Mitra et al., 1998).

Breeding Provitamin A-Enriched Maize

Maize has been targeted as one of the major food crops for provitamin A (PVA) enrichment and delivery under the HarvestPlus Program (Pfeiffer and McClafferty, 2007; Bouis and Welch, 2010). The main objective of the PVA enrichment in maize breeding program has been to develop high-yielding, provitamin A-enriched maize cultivars that are profitable to farmers and acceptable to the consumers, and with proven effectiveness in reducing vitamin A deficiency (Bouis and Welch, 2010). Yellow maize naturally accumulates PVA carotenoids, including α-carotene, β-carotene (BC), and β-cryptoxanthin (BCX), which can be metabolically converted to active vitamin A in the human body (Asson-Batres and Rochette-Egly, 2016). However, kernels of yellow maize cultivars commonly grown by farmers contain less than 2 μg g^{-1} of PVA (Ortiz-Monasterio et al., 2007; Pixley et al., 2013), which is insufficient to meet the recommended daily requirement in a diet (Institute of Medicine, 2012). Considerable efforts have thus been made to increase the concentrations of PVA carotenoids in maize through conventional and molecular marker-assisted breeding (Pixley et al., 2013; Giuliano, 2017; Andersson et al., 2017; Menkir et al., 2017).

The breeding target for PVA in maize deemed sufficient to impact human health was set at 15 μg g^{-1} of BC equivalents, by a multidisciplinary team that included plant breeders, plant geneticists, biochemists, nutritionists, and food processing specialists (Hotz and McClafferty, 2007). PVA content was estimated as one part of BC plus one and a half part of BCX, based on the number of unmodified β-rings and number of retinol molecules that can be derived from them (Von Lintig, 2010; Wurtzel et al., 2012). The first step in developing PVA maize was screening of more than 1,500 genotypes for their carotenoid profiles. The majority of inbred lines surveyed had BCX and BC content averaging from 1 to 2 μg g^{-1} and PVA content of 2–3 μg g^{-1} (Ortiz-Monasterio et al., 2007; Menkir et al., 2008), but a few temperate lines had PVA levels approaching 15 μg g^{-1} (Ortiz-Monasterio et al., 2007). Nonetheless, the levels of PVA carotenoids in adapted tropical and sub-tropical maize inbred lines were far below the breeding target of 15 μg g^{-1} of PVA set for maize (Bouis et al., 2011), emphasizing the need for accessing and mining novel sources of favorable alleles to boost PVA concentration to new levels. The less complex nature of control of provitamin A content, high heritability, mode of inheritance regulated primarily by additive genetic effects, and statistically non-significant correlation between PVA and agronomic performance suggested that

concurrent improvements of PVA carotenoids and grain yield would be possible (Suwarno et al., 2014; Menkir et al., 2018; Ortiz-Covarrubias et al., 2019). Through systematic breeding efforts, significant improvement has been made in enhancing provitamin A content of tropical maize inbred lines developed at CIMMYT (**Figure 2**).

Selection for PVA content has focused mostly on increasing BC; however, new evidence suggests that BCX may be more bioavailable (Howe and Tanumihardjo, 2006; Schmaelzle et al., 2014; Sugiura et al., 2014) and less susceptible to degradation than BC. Recent studies (Dhliwayo et al., 2014; Ortiz et al., 2016; Taleon et al., 2017; Sowa et al., 2017) suggested that breeding for PVA carotenoids should aim to increase BCX more than BC due to the increasing evidence of low stability of BC, higher BCX bioavailability compared to BC, and BCX's similar bioconversion and bioefficacy to BC (Schmaelzle et al., 2014), in addition to the genetic diversity found for BCX (Suwarno et al., 2015; Menkir et al., 2018). Breeding efforts have started in this regard and inbred lines have been developed and are currently being used in the development of new hybrids and synthetics. Research focused on minimizing carotenoid degradation is also needed.

Molecular Breeding for Provitamin A Enrichment

The carotenoid biosynthesis pathway is well-studied and genes controlling key steps in the pathway have been cloned and characterized. Allelic variation in key genes has been exploited to identify and develop DNA markers associated with BC and BCX, the main carotenoids with PVA activity. Use of MAS in combination with high-performance liquid chromatography (HPLC)/ultra-performance liquid chromatography (UPLC) analysis has been more effective than biochemical screening alone (Zhang et al., 2012; Simpungwe et al., 2017).

Most of the breeding work until about 2011 was based on phenotypic selection and biochemical analysis was used to quantify PVA levels in selected lines (Pixley et al., 2013). Polymorphisms were identified in two genes: *β-carotene hydroxylase1* (*CrtRB1*) which catalyzes the hydroxylation of BC

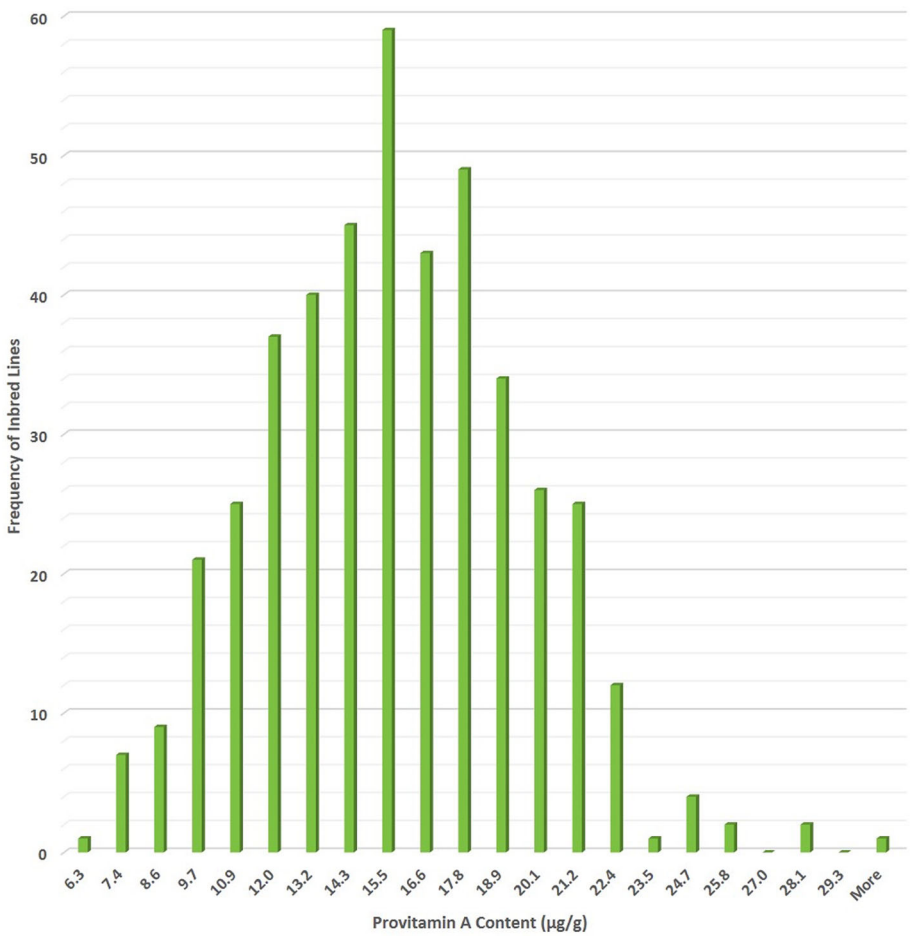


FIGURE 2 | Extent of genetic variation for provitamin A content (µg/g) in International Maize and Wheat Improvement Center (CIMMYT) maize inbred lines developed under HarvestPlus-maize biofortification program.

to BCX (Yan et al., 2010), and *lycopene epsilon cyclase* (*LcyE*), which converts lycopene to zeta (ζ)-carotene and ultimately to α-carotene (Harjes et al., 2008). The favorable allele at *CrtRB1* reduces hydroxylation of BC into BCX whereas *LcyE* reduces flux into the α-branch of the pathway (Von Lintig, 2010). Both alleles increase PVA content, but the favorable *CrtRB1* allele is more effective in increasing PVA content than the favorable *Lcyε* allele (Yan et al., 2010; Babu et al., 2013).

Three functional polymorphisms were thus identified in the two genes: CrtRB1-5′TE, CrtRB1-3′TE, and LcyE 3′Indel. Babu et al. (2013) evaluated these three polymorphisms in CIMMYT maize germplasm and reported more than two-fold increase in BC associated with the favorable allele at CrtRB1-3′TE, irrespective of the genotype at CrtRB1-5′TE. The favorable allele LcyE 3′Indel reduced the ratio of α to β-branch carotenoids by up to 30%, without a notable increase in BC or PVA content. Effectiveness of the CrtRB1-3′TE polymorphism in increasing PVA content was also demonstrated by Zunjare et al. (2018). The maximum PVA content that can be achieved by selecting for both the CrtRB1-3'TE and LcyE 3′Indel polymorphisms is unknown, partly because of a large genetic background effect or epistasis (Babu et al., 2013). Nevertheless, donor germplasm with >20 μg g^{-1} have been developed using MAS targeting the favorable alleles for the two polymorphisms in a pedigree breeding scheme. Foreground MAS selection was done in segregating F_2 or F_3 populations, selecting seeds that had the homozygous favorable alleles at the CrtRB1-3′ TE and LcyE-3′Indel polymorphisms. However, combining foreground and background MAS may be preferable to reduce linkage drag considering that most of the PVA maize donor germplasm is agronomically inferior (Zunjare et al., 2018; Ortiz-Covarrubias et al., 2019).

During 2011 to 2015, CIMMYT's PVA breeding program at Mexico was manually chipping and genotyping 10,000 seeds per season and selecting homozygous individuals (~20–25%) for planting. This process saved land and labor resources by ensuring that only individual plants that carried the favorable alleles were planted while enabling development of inbred lines that reached or exceeded the breeding target coupled with good yield potential, disease resistance, grain quality traits, and other agronomic traits. However, the gel-based assays were expensive and the manual seed chipping method was laborious and time-intensive. Scientists at CIMMYT and IITA developed the single nucleotide polymorphism (SNP) markers, which have been used to develop an automated high-throughput assay with an external service provider. Currently, in SSA, MAS for the favorable CrtRB1-3′TE is done in combination with MAS for *maize streak virus 1* (MSV1) that is associated with resistance to the viral disease maize streak virus (Nair et al., 2015).

HIGH-ZINC MAIZE

Extent of Zinc Deficiency

According to the World Health Organization, Zn deficiencies affected 17% of the global population (www.harvestplus.org). Zn deficiency is widespread and particularly prevalent in Africa, the eastern Mediterranean, and South and Southeast Asia (Caufield et al., 2004). In young children, it increases the risk of diarrheal disease, pneumonia, malaria, and mortality from those diseases. Based on the estimated average requirement (EAR) of 1,860 μg day^{-1} of Zn in maize, the breeding target established by HarvestPlus for maize was 33 μg g^{-1} (Bouis and Welch, 2010). Andersson et al. (2017) indicated a new EAR of 2,960 μg day^{-1} for women and 1,390 μg day^{-1} for children. The baseline content for Zn in maize is about 20 μg g^{-1}; assuming 90% retention of Zn after processing and 25% bioavailability, at least 33 ppm of Zn needs to be accumulated in the maize grains. Thus, an increase of at least 13 μg g^{-1} is targeted by breeding, which is achievable due to significant genetic variation for kernel-Zn concentration in tropical maize germplasm. With the availability of high-Zn tropical maize genotypes, studies on nutrient retention, bioavailability, and efficacy are being planned.

Breeding High-Zn Maize Cultivars

Biofortification of maize kernels with high-Zn has been undertaken at CIMMYT and IITA, in active partnership with public and private sector partners. Substantial genetic variation for kernel-Zn (4–96 ppm) was found in tropical maize germplasm (Bänziger and Long, 2000; Ortiz-Monasterio et al., 2007; Menkir, 2008; Chakraborti et al., 2011; Prasanna et al., 2011; Hindu et al., 2018), including landraces, inbreds, hybrids, and OPVs. Unlike crops like wheat and rice (Guzmán et al., 2014), in maize no statistically significant correlation was observed between kernel Zn and Fe contents. In fact, high Zn maize normally contain between 18 and 20 ppm Fe, which is the average content in maize kernel. In addition, factors like limited genetic variation and the high target levels needed to reach nutritional impact in the consumers affected Fe biofortification in maize using breeding strategies.

Breeding efforts at CIMMYT and IITA were initiated to meet the target level of 33 ppm kernel-Zn (dry weight) (Bouis et al., 2011). The initial focus has been on Latin American countries including Guatemala, Nicaragua, Honduras, and Colombia, and Western African countries, including Ghana, Benin, and Nigeria. Zn biofortification breeding at CIMMYT has utilized white-endosperm and high kernel-Zn lines with QPM background. Three QPM CIMMYT Maize Lines (CMLs)—CML176, CML491, and CML492—were found to be particularly important for improving kernel Zn in tropical maize, and have been used extensively as founder lines in pedigree-based selection. These elite QPM lines were derived from CIMMYT maize Population 62 (white flint QPM) and Population 63 (white dent QPM), which underwent several cycles of intra-population recurrent selection in the 1980s (CIMMYT, 1998). Population 62 traces back to ETO composite, whose main components were the tropical Colombian landraces, Comun and Chococeño, and the Venezuelan landraces, Puya and Cubano Amarillo. Population 63 traces back to Tuxpeño-1 composite, whose main component was the Mexican landrace Tuxpeño (CIMMYT, 1998). Adapted yellow and white endosperm maize inbred lines derived from broad-based populations, bi-parental crosses, and backcrosses with high Zn content have been used for generating pedigree populations at CIMMYT and IITA to develop new high Zn inbred lines.

Interestingly, above-average concentration of Zn in the kernels was found in the QPM germplasm as compared to non-QPM/normal maize germplasm (Chakraborti et al., 2009; Chakraborti et al., 2011). However, not all QPM germplasm is high in kernel Zn, and it is possible to have some non-QPM germplasm with high kernel Zn. Zn plays an important role in tryptophan biosynthesis, which is increased in QPM. Using 923 lines to conduct genome-wide association studies (GWAS) for kernel Zn, Hindu et al. (2018) reported that only 31 were QPM or had QPM background and 33.3% had Zn values higher than 30 μg g^{-1}on dry weight (DW) basis. In contrast, out of the 892 non-QPM used in the panel, 19.9% had values higher than 30 μg g^{-1}DW, and about 6% of them had values higher than the breeding target (33 μg g^{-1}DW). Taken together, these results indicate great potential to develop high Zn maize alone or in combination with better protein quality in biofortification programs.

Molecular Breeding for Kernel-Zn Enrichment

Quantitative trail locus (QTL) mapping studies confirmed that kernel Zn accumulation is under the control of several genetic loci (Qin et al., 2012; Baxter et al., 2013). Complexity of the trait is further increased due to higher environment and genotype × environment interaction (GEI) effects. Over the past few years, genomic regions influencing Zn concentration have been detected through GWAS and biparental QTL mapping in maize. GWAS of 923 tropical/sub-tropical CIMMYT maize inbred lines, phenotyped at three locations in Mexico and genotyped using high density genotyping by sequencing (GBS), identified a total of 20 SNPs significantly associated with kernel-Zn. This effort constitutes the first large-scale screening of tropical/sub-tropical public germplasm for kernel Zn. A set of 11 SNPs identified in GWAS have been subsequently validated in independent biparental populations using single factor QTL analysis, and some of these SNPs explained a relatively high proportion of variance (Hindu et al., 2018).

Qin et al. (2012) identified three stable QTLs for kernel-Zn concentrations in two populations across two environments. Šimić et al. (2012) identified two QTLs on chromosomal bin 3.05 and 4.08 which explained small percentage of the variation in a temperate biparental population phenotyped at two locations. Jin et al. (2013) conducted a meta-QTL study with QTL mapping studies published for kernel Zn and many related minerals, and identified nine meta-QTLs across the maize chromosomes that could have an influence in kernel-Zn concentration. Earlier studies identified some major QTLs for kernel Zn trait in chromosomal bins 3.04 (Qin et al., 2012), 4.06, 5.04 (Jin et al., 2013), and 9.06-07 (Qin et al., 2012; Jin et al., 2013); significant SNP markers were identified through CIMMYT GWAS and validation studies. To assess their utility in breeding program, these SNPs were further analyzed in parental lines selected for high kernel-Zn at CIMMYT, including analysis of the frequency of favorable alleles in the breeding pool. The favorable allele frequency ranged between 0.09 and 0.94 in CIMMYT's breeding lines for these SNPs/haplotypes. Three haplotypes were selected on chromosomes 5, 7, and 9 based on their favorable allele frequency and effect size of favorable alleles. In a set of 1,880 breeding lines from pedigree crosses entering stage 1 testing in the kernel-Zn breeding pipeline, it was seen that the selection for the three favorable haplotypes increased the population mean kernel-Zn content by 16.3% (Sudha K Nair, unpublished).

Genomic selection (GS) has been demonstrated as an effective approach to accelerate genetic gain in maize breeding for improvement of complex traits (Zhang et al., 2015; Cao et al., 2017; Yuan et al., 2019). The genomic prediction accuracy for kernel Zn content in maize has been estimated at CIMMYT in different types of maize populations with repeat amplification sequencing and GBS markers. Moderate to high prediction accuracies, ranging from 0.35 to 0.65, were observed across different types of populations and genotyping platforms (Xuecai Zhang, unpublished). Thus, kernel Zn content in maize could be improved by implementing MAS and GS in a stepwise fashion, where the SNPs/haplotypes detected and validated in the association mapping and linkage mapping analyses can be used in forward breeding at an early generation when there are larger numbers of selection candidates, followed by genomic selection at advanced stages of breeding or the SNPs/haplotypes can be fitted as fixed effects in GS models to improve prediction accuracy.

Low Phytate Maize Genotypes for Enhancing Kernel-Zn Bioavailability

Breeding for high kernel Zn has been a challenge to the maize breeders worldwide, primarily due to involvement of large number of loci with minor effects and existence of very high GEI (Colangelo and Guerinot, 2006). Moreover, bioavailability of Zn in maize grains is only 20% in the human gut (Andersson et al., 2017). The major impediment of low bioavailability of Zn has been the presence of phytic acid/phytate that constitutes nearly 75–80% of the total phosphorus in maize grains (Raboy, 2001). Maize kernels generally contains ~3.2 mg day^{-1} of phytic acid with a range of 2.4 to 4.1 mg day^{-1} (Lorenz et al., 2007). Phytate being negatively charged has a strong tendency to chelate positively charged metal ions, such as Zn, thereby resulting in highly insoluble salts with poor bioavailability of the nutrient (Zhou and Erdman, 1995). Monogastric animals including humans, poultry, and swine cannot digest phytic acid in their gut, since they lack phytic acid hydrolyzing enzyme phytase. Phytate is thus expelled directly to the environment through excreta posing a serious concern as continuous expulsion of high phosphorus load causes pollution in the nearby water bodies (Jorquera et al., 2008). Hence bringing down the phytate in maize could be an important strategy for Zn biofortification.

Low phytic acid (*lpa*) mutants are available in maize (Raboy et al., 2000). These mutants produce seeds that have normal levels of total phosphorus but greatly reduced levels of phytic acid phosphorus. The *lpa* mutations do not affect the ability of a plant to uptake phosphorus and its transportation to a developing seed; instead, they block the ability of a seed to synthesize phosphorus into phytic acid (Pilu et al., 2003). Recessive *lpa1-1* mutation causes up to 55–65% reduction of phytic acid in maize grain and is due to a mutation in trans-membrane transporter protein (MRP). The

lpa2-1 mutation causes 50% reduction in phytic acid, and is due to a mutation in inositol phosphate kinase (IPK) enzyme (Raboy et al., 2000). *lpa241* mutation which reduces phytic acid up to 90% originated from mutation in myo-inositol(3)P1 synthase (MIPS) enzyme (Pilu et al., 2005).

Ertl et al. (1998) developed low phytate maize genotypes *via* backcross method. Beavers et al. (2015) generated low phytic acid maize population without negatively affecting seed quality through three rounds of selections in broad-based synthetic populations. Though *lpa241* showed 30% reduction in germination (Pilu et al., 2005), no negative effects on germination, initial establishment, growth, and response to pests and diseases were observed for *lpa2-1* and *lpa1-1* (Raboy et al., 2000). Considering this, *lpa2-1* and *lpa1-1* have been used in various breeding programs. In India, *lpa2-2* allele was successfully introgressed into well-adapted and productive elite inbred lines *viz.*, UMI-395 and UMI-285 through MAS (Sureshkumar et al., 2014; Tamilkumar et al., 2014). Recessive *lpa1-1* and *lpa2-1* mutants have also been combined with high provitamin-A and QPM traits in elite genetic backgrounds (Bhatt et al., 2018). Though there is no report of release of low phytate maize cultivars, the *lpa* genotypes developed by different research groups hold promise for their deployment to alleviate Zn deficiency. It is important to integrate low phytate program with breeding for high-Zn content, as the benefits of *lpa* mutant can be best realized in genotypes with higher kernel Zn. In addition to genetic interventions, phytate content can also be reduced through maize processing methods like lime-cooking and fermentation (Ekpa et al., 2018).

VITAMIN E-ENRICHED MAIZE

Extent of Vitamin E Deficiency

Vitamin-E is an essential micronutrient in human body, and plays vital role in scavenging of various reactive oxygen species (ROS) and free radicals, quenching of singlet oxygen (high energy oxygen), and providing membrane stability by protecting polyunsaturated fatty acids (PUFA) from lipid peroxidation (Fryer, 1992). Vitamin-E helps in preventing Alzheimer's disease, neurological disorders, cancer, cataracts, age-related macular degeneration, and inflammatory disease (Bramley et al., 2000). Supplementation of vitamin-E in the feed also ensures enhanced quality and prolongs stability of animal meat (De Winne and Dirinck, 1996; Sanders et al., 1997).

Recommended dietary allowance (RDA) for vitamin-E is 4 mg day^{-1} for 0–6 month's old child, while the same for ≥14 years old is 15 mg day^{-1} for both males and females (Institute of Medicine, 2000). It is estimated that over 20% of the examined people both in developed and developing countries possess plasma α-tocopherol lower than the recommended level (Li et al., 2012). In a study in South Korea, two-third of adults were reported to have suboptimal vitamin-E level and almost one fourth are deficient (Kim and Cho, 2015). About one-third of the pregnant women in rural Nepal are severely affected by vitamin E deficiency (VED) (Jiang et al., 2005). In Bangladesh, VED is more critical, as about two-third of women in early pregnancy were found to be severely vitamin-E deficient (Shamim et al., 2015).

Molecular Breeding for Vitamin E Enrichment

Wide genetic variation in tocopherol components has been observed in maize (Rocheford et al., 2002; Egesel et al., 2003; Li et al., 2012; Feng et al., 2015; Muzhingi et al., 2017; Das et al., 2019a). Estimation of tocopherol fractions (α-, β-, γ-, and δ-) is simple and fast (~15 min/sample) using HPLC, but it involves high cost (US$25–30 per sample). Selection of key genes that enhance tocopherol in maize, therefore, provides cost-effective (US$0.5/sample of PCR) solution. Several earlier studies (Wong et al., 2003; Shutu et al., 2012; Feng et al., 2013; Lipka et al., 2013; Diepenbrock et al., 2017) reported QTLs for higher accumulation of kernel α-tocopherol, γ-tocopherol, and total tocopherol in maize. The pathway for vitamin-E biosynthesis is also well-characterized (DellaPenna and Pogson, 2006). Several genes *viz.*, homogentisate phytyltransferase (*VTE2*), homogentisate geranylgeranyl transferase (HGGT), methyl transferase (*VTE3*), tocopherol cyclase (*VTE1*), phytol kinase (*VTE5*), and γ-tocopherol methyl transferase (*VTE4*), play important role in regulating the pathway. Among these genes, *VTE4* was identified as the key gene that enhances the accumulation of α-tocopherol by converting γ-tocopherol (Li et al., 2012). Two insertion/deletions (InDel7 and InDel118) within the gene *VTE4* were found to significantly affect the level of α-tocopherol. InDel118, located 9-bp upstream of the putative transcription start site, controls α-tocopherol content by regulating *VTE4* transcript level, whereas InDel7 affects translation efficiency. Association of *VTE4* with higher accumulation of α-tocopherol was also reported by Lipka et al. (2013). Das et al. (2019b) later identified one SNP (G to A), and three InDels (14 and 27 bp) in the *VTE4* gene comprising a favorable haplotype (0/0) which can differentiate low and high α-tocopherol accumulating maize lines. These newly identified SNP and InDels in addition to the previously reported InDel118 and InDel7 can be useful in selection of favorable genotypes with higher α-tocopherol in maize.

Das et al. (2019c) screened large number of maize inbreds of diverse pedigree using gene-based markers specific to InDel118 and InDel7 of *VTE4* and identified inbreds with favorable haplotype (0/0: deletion at InDel118 and InDel7) of *VTE4*. CML560 and CML496 possess the most favorable haplotype (0/0) for *VTE4* (**Figure 3**). Das et al. (2019b) developed hybrids using inbreds possessing the favorable haplotype of *VTE4*, and reported higher mean α-tocopherol (mean: 21.37 ppm) than the check hybrids (mean: 11.16 ppm). In some of the hybrids *viz.*, MHVTE-2, MHVTE-18, MHVTE-28, MHVTE-10, and MHVTE-3, α-tocopherol constituted ≥50% of the total tocopherol.

Considering the major effect of *VTE4* in accumulating higher α-tocopherol, Feng et al. (2015) transferred the favorable allele from a suitable donor parent (SY999) to four Chinese *shrunken2*-based sweet corn lines (M01, M14, K140, and K185) through MABB. Average increment of 7.73 ppm of α-tocopherol was observed among the MABB-derived progenies, with α-tocopherol as high as 15.99 ppm, compared to 3.14 ppm in recurrent parents. In India, favorable allele of *VTE4* has been

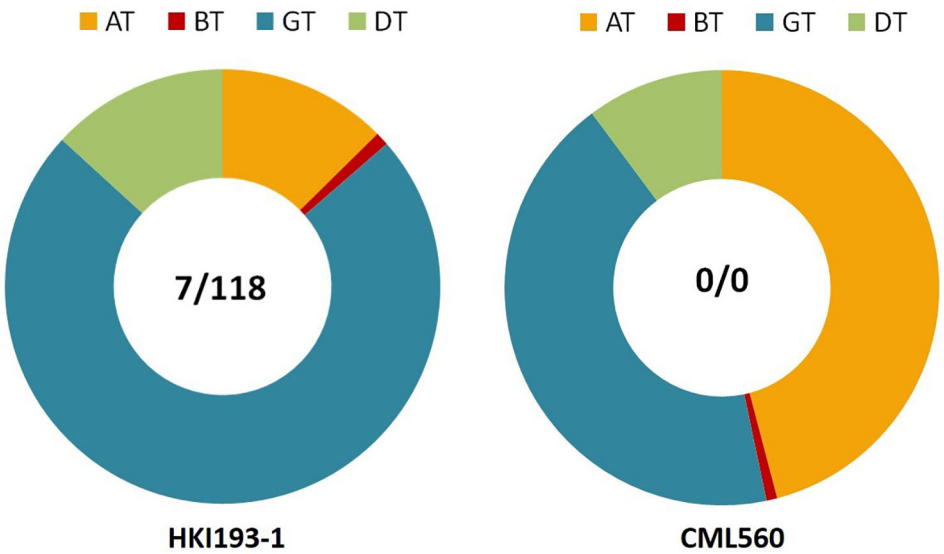


FIGURE 3 | HKI193-1, an elite maize inbred with an unfavorable haplotype (7/118) for *VTE4*, showing low AT (12.6%) *versus* CML560 with a favorable haplotype (0/0) showing high AT (45.9%). AT, α-tocopherol, BT, β-tocopherol, GT, γ-tocopherol, DT, δ-tocopherol.

introgressed into four provitamin-A rich QPM elite inbreds using MABB that led to the increase of α-tocopherol to 15.2 ppm over 8.0 ppm in the original inbreds (Hossain et al., 2018b). Though α-tocopherol (>30 ppm) has been reported in maize germplasm, the extent of increase due to introgression of favorable allele of *VTE4* depends on the base level expression of the other important genes in the pathway, and further interaction with background genome (Das et al., 2019c).

MAIZE NUTRITIONAL QUALITY ANALYSIS

High-throughput analytical methodologies are critical for integration of biofortification strategies into mainstream plant breeding as well as for assurance of the nutritional quality in the seed that reaches farmers and consumers (Guild et al., 2017). Given the fact that several of the nutritional traits are not visible to a naked eye, it is extremely important that these are measured in an appropriate laboratory where such methodologies are well-established. To achieve the level of precision required while optimizing costs, different analytical methods may need to be employed during the process of breeding for the target traits, as well as quality assessment/quality control (QA/QC) of seed of parental lines (Gowda et al., 2017) post-cultivar release (**Table 1**).

Near infrared spectroscopy and X-ray fluorescence methodologies can be effectively used for early breeding generations to identify and reject materials that have low values of PVA. However, for advanced generation materials or commercial samples, it is always recommended to conduct the wet chemistry analysis, using chromatography and inductively coupled plasma optical emission spectroscopy (ICP-OES).

Garcia-Oliveira et al. (2018) pointed out that apart from the actual germplasm and environmental influence on kernel-Zn concentration, there could be numerous possibilities for introducing variation in the results of different studies, ranging from sensitivity of the method used for the quantification of Zn contents, improper postharvest handling of the samples, and the significant variability among microenvironments for Zn. Spectroscopic methods, such as ICP-OES and atomic absorption spectroscopy (AAS), are therefore used to provide robust and accurate results on kernel Zn content in maize. Limits of detection span a wide analytical range from percentage by weight to parts per billion (ppb) levels. Sample digestion can be done chemically or by using micro-wave methods, which increase the processing throughput prior to analysis. For early stages of high-Zn

TABLE 1 | Analytical methods used at different breeding stages for the targeted maize biofortification traits at International Maize and Wheat Improvement Center (CIMMYT).

Trait	Germplasm screening/breeding stage[a]	Analytical methods[b]	Cost per sample (USD)
Provitamin A	Landraces; new germplasm; stages 1–2; QA/QC of seed post-cultivar release	NIRS (for total carotenoids)	2.8
Kernel zinc	Stage 3; variety release and promotion	HPLC/UPLC	38.3
	Landraces; new germplasm; stages 1–2;	XRF	5.8
QPM	QA/QC of seed post-cultivar release		
	Stage 3; variety release and promotion	ICP-OES	13.7
	Landraces; new germplasm; stages 1–2; QA/QC of seed post-cultivar release	NIRS	2.8
	Stage 3; cultivar release and promotion	Colorimetric methods	18.9

[a]*QA/QC, quality assessment/quality control.*
[b]*NIRS, near-infrared spectroscopy; HPLC, high performance liquid chromatography; UPLC, ultra performance liquid chromatography; XRF, X-ray fluorescence; ICP-OES, inductively coupled plasma–optical emission spectroscopy.*

maize breeding, where large number of samples are screened, X-ray fluorescence (XRF) is extensively used (Hindu et al., 2018; Palacios-Rojas et al., 2018). The cost of analysis per sample by XRF is typically lower [approx. 6 US dollars (USD)] than ICP-OES (approx. 14 USD) (**Table 1**). Additional benefits of the method include simplicity, avoiding the use of hazardous chemicals, and less costly sample preparation.

Unlike organic micronutrients, such as PVA carotenoids, degradation is not an issue for high-Zn maize grains; however, the potential for contamination during the sample manipulation and analysis is high due to the high abundance of Zn in the environment. In the case of maize, use of flour is required due to the heterogeneity in grain shape; use of flour also improves data reproducibility and accuracy compared with whole grain analysis. It is important to ensure that grinding is performed using a grinder with non-contaminating material such as zirconium. When ICP-OES is performed, aluminum determination is used as an indicator of sample contamination (Guild et al., 2017).

NUTRITIONALLY ENRICHED MAIZE CULTIVAR RELEASES AND COMMERCIALIZATION

Tropical maize inbred lines with enhanced nutritional quality (especially QPM, PVA, and high-Zn) and other desirable agronomic and adaptive traits developed at CIMMYT and IITA have been used in several countries to develop agronomically competitive hybrids or OPVs or synthetics (Menkir et al., 2018). **Figure 4** shows the varietal release status with regard to provitamin A and high-Zn maize in SSA, Asia, and Latin America.

Extensive multi-location evaluations of PVA-enriched OPVs and hybrids in collaboration with public and private sector partners in Zambia, Zimbabwe, Tanzania, Ghana, Nigeria, Mali, Malawi, and the Democratic Republic of Congo (DRC) led to the selection of promising hybrids and synthetics for further evaluation in national performance trials (NPTs) and farmer participatory on-farm trials. These testing schemes led to the release of 47 PVA-enriched hybrids and synthetics that meet 50–80% of the current PVA breeding target in nine countries, including Zambia, Zimbabwe, Malawi, Tanzania, DRC, Ghana, Mali, and Nigeria (Menkir et al., 2018). In addition, the high PVA inbred lines developed at CIMMYT and IITA have been effectively used as PVA donors by the national maize breeding programs in Asia (e.g., China and India) and Latin America (e.g., Brazil and Panama) to develop improved, nutritionally enriched maize cultivars (Muthusamy et al., 2014; Liu et al., 2015; Zunjare et al., 2018; Goswami et al., 2019; Listman et al., 2019).

The first PVA-enriched maize hybrids and synthetics released in SSA in 2012 had an average PVA content of about 6.0 to 7.5 μg g^{-1}, or 40–50% of the breeding target. Since then several hybrids exceeding 10 μg g^{-1} PVA have been released (Andersson et al., 2017). A hybrid with >90% of the PVA breeding target (14.1 μg g^{-1}) was released in Malawi in 2016. In each of the target countries, PVA-enhanced cultivars are released either by the national research institutes or by private seed companies working in those countries.

HarvestPlus and national governments in SSA have invested significant effort in creating awareness and consumer demand in the target countries where market preference is for white maize. These efforts seem to be paying off; Zambia now has more than 500 tons of certified seed of PVA maize cultivars, covering an estimated 200,000 ha. Further, linkages were established with

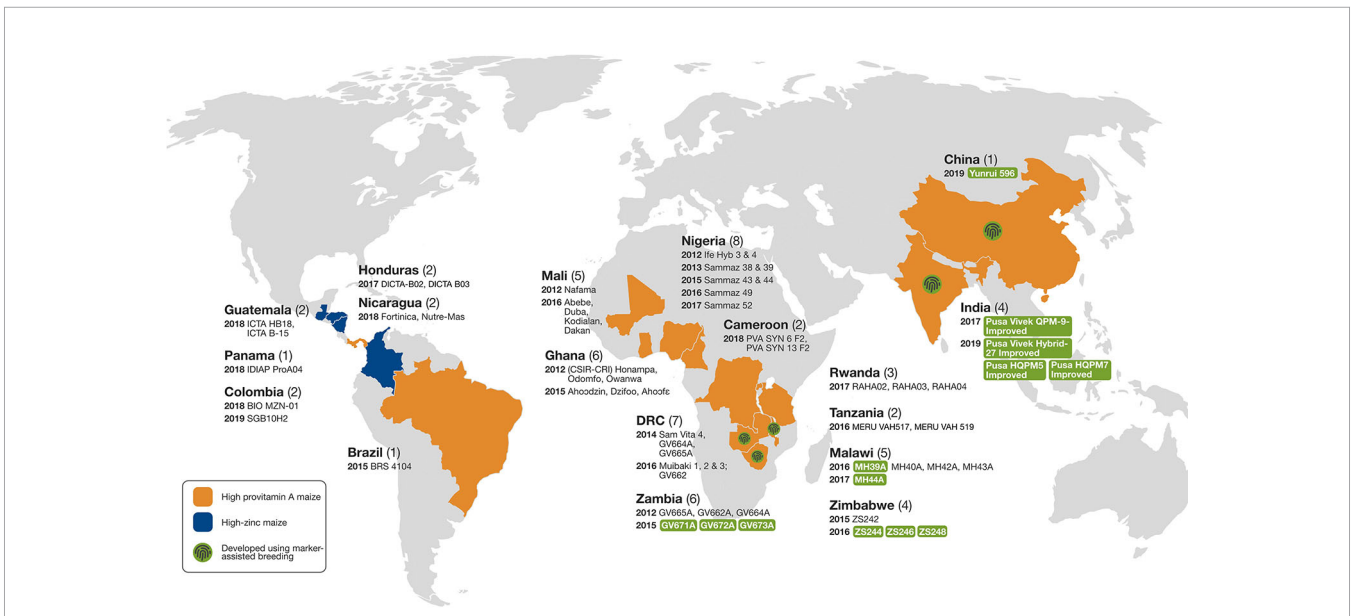


FIGURE 4 | Provitamin A-enriched and high-Zn maize cultivars developed using conventional and molecular marker-assisted breeding and released for commercial cultivation in sub-Saharan Africa, Asia, and Latin America (modified from Listman et al., 2019).

millers and food processing companies to ensure that farmers have access to markets for their excess grain. Despite the awareness campaigns, school feeding schemes that are funded by non-government organizations (NGOs) and the government remain a substantial part of the market for high PVA maize, raising questions about what will happen when that support is not sustained. Influencing national policy on balanced nutrition and health appears to be the most viable way to create and sustain demand for PVA-enhanced maize because most governments in the target countries provide subsidies to farmers and buy most of the grain as part of a national food security strategy (Simpungwe et al., 2017).

CIMMYT, in collaboration with public and private sector partners in Mexico, Guatemala, Nicaragua, El Salvador, Honduras, and Colombia, has been working on development and deployment of elite high-Zn maize cultivars in Latin America. Extensive multi-location trials showed relative yield parity and similar performance for other agronomic traits, relative to commercial checks, indicating the competitiveness of the these products in the lowland tropics of Latin America. So far, four high-Zn maize cultivars (two hybrids and two synthetics) have been released in Latin America. These cultivars have 90 to 110% of the target kernel Zn content set under HarvestPlus and are competitive for grain yield and other adaptive traits with the commercial checks (Listman et al., 2019).

A schematic depiction of decision tree for molecular marker-assisted breeding workflows to accelerate progress toward nutritional trait targets is shown in **Figure 5**. The workflow decisions are based on relative eliteness and adaptation of initial nutritional trait donor breeding parent, genetic complexity of the target nutritional trait, and relative cost of phenotypic assays, trait-linked marker assays, and genome profiling. Efficacy of a converted line *via* MABB or gene editing refers to the nutrient level of the converted line relative to the target. Equivalency of a converted line refers to performance of the new version relative to the elite recurrent parent for important agronomic and adaptive traits.

In recent years, elite maize inbred lines and hybrids with multiple nutritional quality traits, especially combinations of QPM, provitamin A, low phytate, and vitamin E traits, have been developed in India and China, through molecular marker-assisted

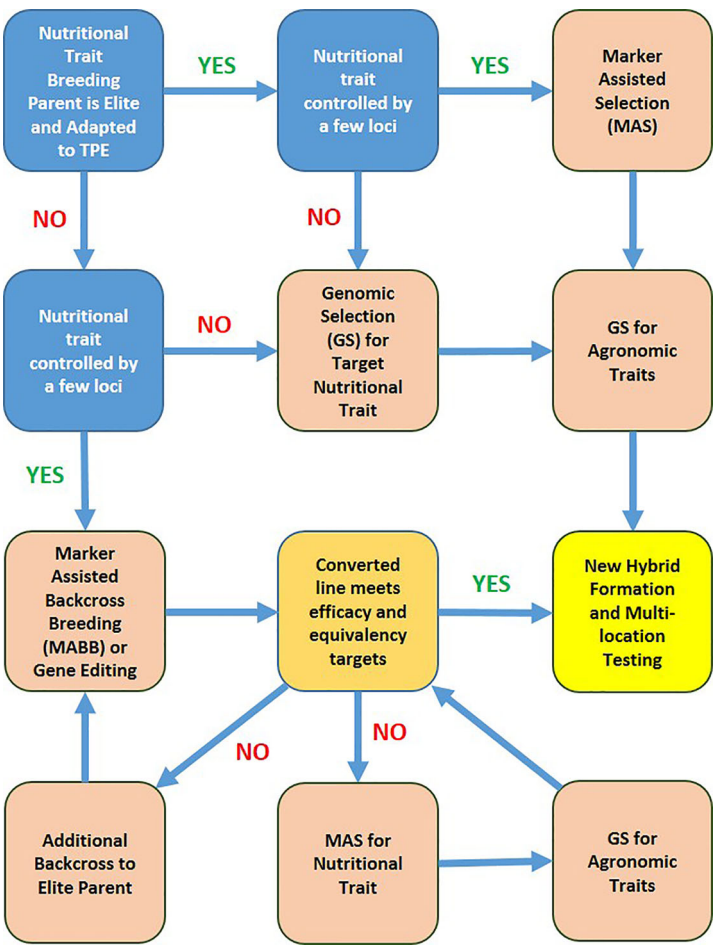


FIGURE 5 | A schematic depicting the strategy with decision tree for molecular marker-assisted breeding workflows to accelerate progress toward nutritional trait targets. TPE refers to the target population of environments of the product profile with the target nutritional trait(s).

breeding (**Table 2**). Stacking multiple traits essentially requires raising of large population size in MABB program compared to a population segregating for one target gene. Since traits like QPM, provitamin A, vitamin E, and kernel Zn are not associated with yield penalty, they can be generally stacked together. One of the elite hybrids with a stack of nutritional quality traits (QPM and Provitamin A), Pusa Vivek QPM-9 Improved, developed by ICAR-Indian Agricultural Research Institute has been officially released in India in 2017 for commercial cultivation. Three MABB-derived hybrids, including one provitamin A hybrid, Pusa Vivek Hybrid-27 Improved, and two QPM + provitamin A-enriched hybrids, Pusa HQPM5 Improved and Pusa HQPM7 Improved, have been approved for release in India in 2019. High provitamin A trait has been introgressed in both traditional and QPM genetic backgrounds (Muthusamy et al., 2014; Andersson et al., 2017). In addition, the germplasm resources developed through stacking of provitamin A and vitamin E could be potentially used for analysis of carotenoid stability.

Genome editing represents a powerful technology for enhancing and stacking nutritional quality traits in maize. Liang et al. (2014) reported editing of the genes involved in phytic acid synthesis (*ZmIPK1A*, *ZmIPK*, and *ZmMRP4*) in maize. Since several of the important genes influencing nutritional quality traits, such as QPM and provitamin A, have been well-characterized, gene editing could provide a powerful system for stacking these traits in agronomically superior genetic backgrounds. For example, if one wants to stack a PVA trait in an elite QPM line, editing the elite line for *CrtRB1* could potentially simplify the breeding process. Similarly, it should be possible to design a PVA target and edit two to four key carotenoid biosynthetic pathway genes in elite climate-resilient maize germplasm.

CONCLUDING REMARKS

Challenges in agriculture and food production keep evolving. In the 1950s, the world needed a large boost in food production to combat famine, in the face of rapid population growth and recurring natural disasters. In the 21st century, the challenge is not only to produce enough to feed the growing population, but also providing nutritionally balanced diets. Moving away from overemphasis on calorie security, the food today on everyone's plates must be of appropriate quantity, nutritious, and produced in an environmentally, economically, and socially sustainable manner. The EAT-Lancet Commission Report (Willett et al., 2019) highlighted the importance of promoting diets that are nutritious and which can reduce the environmental impact of food systems. The sum of different agricultural and nutrition sensitive strategies could contribute to sustainable and nutritional food systems.

Although not a silver bullet solution, biofortification has proven to be an efficient strategy to combat malnutrition. CIMMYT, IITA, and national partners [especially in Africa, Asia, and Latin America (LatAm)] have employed conventional breeding and molecular tools, to successfully develop and release several nutritious maize cultivars without compromising grain yield levels or other important agronomic and adaptive traits. Many of these biofortified maize cultivars are currently grown by farmers and accepted by consumers in many countries (Talsma et al., 2017; Manjeru et al., 2019). Advances in phenotyping coupled with molecular breeding facilitated achievement of the breeding targets for various nutrients in maize. Going forward, the focus should be on mainstreaming breeding for nutrient enrichment into maize breeding efforts to deliver high-performing climate-resilient maize cultivars with improved nutritional quality to farmers and consumers. Efforts need to be

TABLE 2 | Some examples of stacking of nutritional quality traits in maize using molecular marker-assisted breeding.

Traits	Gene combination	Nutritional trait values	Improved genotypes developed	References
QPM + provitamin A	*CrtRB1* in a QPM hybrid (Vivek QPM9)	8.16 μg g^{-1} PVA; 0.74% Trp; and 2.67% Lys (Lys and Trp as % endosperm protein)	Pusa Vivek QPM9 Improved	Muthusamy et al. (2014); Yadava et al. (2018)
	CrtRB1 in QPM inbreds (CML161 and CML171)	5.25 to 8.14 μg g^{-1} PVA; 0.35% Lys in endosperm flour	Provitamin A-enriched elite QPM inbreds CML161 and CML171	Liu et al. (2015)
	CrtRB1, *LcyE*, and *o2* in QPM inbreds	9.25–12.88 μg g^{-1} PVA; 0.334% Lys and 0.080% Trp (Lys and Trp estimated in endosperm flour)	Pusa HQPM-5 Improved; Pusa HQPM-7 Improved	Zunjare et al. (2018)
	CrtRB1 in a QPM inbred	10.75 μg g^{-1} PVA; 0.303% Lys and 0.080% Trp (Lys and Trp estimated in endosperm flour)	Provitamin A-enriched elite QPM inbred HKI1128Q (parent of Pusa HM9 Improved, HM10Q, and HM11Q)	Goswami et al. (2019)
QPM + provitamin A + vitamin E	*CrtRB1, LcyE*, and *VTE4* in QPM background	16.8 μg g^{-1} alpha-tocopherol; 11.5 μg g^{-1} PVA; 0.367% Lys and 0.085% Trp (Lys and Trp estimated in endosperm flour)	Improved versions of QPM and provitamin A rich hybrids (HQPM-1-PV, HQPM-4-PV, HQPM-5-PV, and HQPM-7-PV)	Hossain et al. (2018b)
QPM + provitamin A + low phytate	*lpa1-1* and *lpa2-1* in provitamin A-enriched QPM lines	8.3–11.5 μg g^{-1} PVA; 0.323–0.372% Lys and 0.081–0.087% Trp (Lys and Trp estimated in endosperm flour); 30–40% reduction in phytic acid P	Improved versions of elite inbreds (HKI161-PV, HKI163-PV, HKI193-1-PV, and HKI193-2-PV)	Bhatt et al. (2018)

made to develop elite maize cultivars stacked with multiple nutrients to address multiple nutrient deficiencies that are prevalent especially in SSA, LatAm, and Asia, as well as in combining biofortification with complementary strategies like dietary diversification, and enhancement of nutrients through agronomic and/or food processing interventions.

Nutrition and health have become the main factors influencing people's diet. The combinations of nutritional quality traits, including QPM, PVA, high-Zn, etc. in both maize grain and in fresh corn has consumer appeal, and contributes to national initiatives and sustainable development goals for enhancing nutrition. The rapid advances that have been made in understanding the genetic control of many macro- and micro-nutrients in maize grains, coupled with the availability of new tools/technologies such as genomic selection, will accelerate the rate of genetic gain for improved nutrient content in maize.

Specialty maize including sweet corn, waxy corn, and popcorn has experienced tremendous growth during the last 30 years, and improving its nutritional quality will stimulate further development of this market. The key to the future development of biofortified specialty maize cultivars is to continue to increase the level of heterosis, seed production, and diversification of the products to meet the changing dietary needs and consumer preferences.

Interdisciplinary work and more effective integration of national and international research efforts are key for enhanced development and dissemination of biofortified crops. For biofortified maize cultivars to succeed in the market, it is important to understand the market dynamics. Value chains that effectively link the farmers to the processors and the consumers need to be improved. Only through such linkages can the value of biofortified crop cultivars can be fully exploited, malnutrition alleviated, and new markets opened.

AUTHOR CONTRIBUTIONS

BP developed the outline for the review, and synthesized the manuscript, based on the contributions from all the co-authors (NP-R, FH, VM, AM, TD, TN, FV, SN, BV, XZ, MO and XF).

FUNDING

The work presented in this review article was supported by various projects, especially the HarvestPlus Program, and the CGIAR Research Program on Maize (MAIZE). MAIZE receives Windows 1&2 support from the Governments of Australia, Belgium, Canada, China, France, India, Japan, Korea, Mexico, the Netherlands, New Zealand, Norway, Sweden, Switzerland, the UK, the USA, and the World Bank. The work on maize biofortification in India has been supported by the Indian Council of Agricultural Research (ICAR); Department of Biotechnology (DBT), Government of India; SERB-Department of Science and Technology (DST), Government of India; and ICAR-All-India Coordinated Research Project (Maize).

REFERENCES

Alvarez, J. O., Salazar-Lindo, E., Kohatsu, J., Miranda, P., and Stephensen, C. B. (1995). Urinary excretion of retinol in children with acute diarrhea. *Am. J. Clin. Nutr.* 61, 1273–1276. doi: 10.1093/ajcn/61.6.1273

Andersson, M. S., Saltzman, A., Virk, P. S., and Pfeiffer, W. H. (2017). Progress update: crop development of biofortified staple food crops under Harvestplus. *Afr. J. Food Agric. Nutr. Dev.* 17, 11905–11935. doi: 10.18697/ajfand.78.HarvestPlus05

Asson-Batres, M. A., and Rochette-Egly, C. (2016). *The Biochemistry of retinoid Signaling II: The Physiology of Vitamin A-Uptake, Transport, Metabolism and Signaling* (Berlin: Springer). doi: 10.1007/978-94-024-0945-1

Atlin, G. N., Palacios, N., Babu, R., Das, B., Twumasi-Afriyie, S., Friesen, D. K., et al. (2011). Quality protein maize: progress and prospects. *Plant Breed. Rev.* 34, 83–130. doi:10.1002/9780470880579.ch3

Bänziger, M., and Long, J. (2000). The potential for increasing the iron and zinc density of maize through plant-breeding. *Food Nutr. Bull.* 21, 397–400. doi:10.1177%2F156482650002100410

Babu, R., Nair, S. K., Kumar, A., Venkatesh, S., Sekhar, J., Singh, N. N., et al. (2005). Two-generation marker-aided backcrossing for rapid conversion of normal maize lines to quality protein maize (QPM). *Theor. Appl. Genet.* 111, 888–897. doi:10.1007/s00122-005-0011-6

Babu, R., Rojas, N. P., Gao, S., Yan, J., and Pixley, K. (2013). Validation of the effects of molecular marker polymorphisms in LcyE and CrtRB1 on provitamin A concentrations for 26 tropical maize populations. *Theor. Appl. Genet.* 126, 389–399. doi: 10.1007/s00122-012-1987-3

Bailey, R. L., West, K. P.Jr., and Black, R. E. (2015). The epidemiology of global micronutrient deficiencies. *Ann. Nutr. Metab.* 66, 22–33. doi:10.1159/000371618

Baxter, I. R., Gustin, J. L., Settles, A. M., and Hoekenga, O. A. (2013). Ionomic characterization of maize kernels in the intermated B73 Mo17 population. *Crop Sci.* 53, 208–220. doi: 10.2135/cropsci2012.02.0135

Beavers, A. W., Goggi, A. S., Reddy, M. B., Lauter, A. M., and Scott, M. P. (2015). Recurrent selection to alter grain phytic acid concentration and iron bioavailability. *Crop Sci.* 55, 2244–2251. doi: 10.2135/cropsci2014.12.0807

Bhatt, V., Muthusamy, V., Jha, S., Zunjare, R. U., Baveja, A., Dosad, S., et al. (2018).Development of low phytic acid maize through marker assisted introgression of lpa1-1 and lpa2-1 genes, in: *Abstracts: 13th Asian Maize Conference on and Expert Consultation on Maize for Food, Feed, Nutrition and Environmental Security, Ludhiana, India, October 8-10, 2018* (Mexico: CIMMYT). pp. 143–144.

Bjarnason, M., and Vasal, S. K. (1992). Breeding for Quality Protein Maize (QPM). *Plant Breed. Rev.* 9, 181–216. doi: 10.1002/9780470650363.ch7

Bouis, H. E., and Saltzman, A. (2017). Improving nutrition through biofortification: a review of evidence from HarvestPlus, 2003 through 2016. *Glob. Food Sec.* 12, 49–58. doi: 10.1016/j.gfs.2017.01.009

Bouis, H. E., and Welch, R. M. (2010). Biofortification – A sustainable agricultural strategy for reducing micronutrient malnutrition in the global South. *Crop Sci.* 50, S20–S32. doi: 10.2135/cropsci2009.09.0531

Bouis, H. E., Hotz, C. H., McClafferty, B., Meenakshi, J. V., and Pfeiffer, W. (2011). Biofortification: a new tool to reduce micronutrient malnutrition. *Food Nutr. Bull.* 32, S31–S40. doi: 10.1177/15648265110321S105

Bramley, P. M., Elmadfa, I., Kafatos, A., Kelly, F. J., Manios, Y., Roxborough, H. E., et al. (2000). Vitamin E. *J. Sci. Food Agric.* 80, 913–938. doi: 10.1002/(SICI)1097-0010(20000515)80:7%3C913::AID-JSFA600%3E3.0.CO;2-3

Bressani, R. (1992). "Nutritional value of high-lysine maize in humans," in *Quality Protein Maize*. Ed. E. T. Mertz (St. Paul, Minnesota, USA: American Association of Cereal Chemists).

Cao, S., Loladze, A., Yuan, Y., Wu, Y., Zhang, A., Chen, J., et al. (2017). Genome-wide analysis of Tar Spot Complex resistance in maize using genotyping-by-sequencing SNPs and whole-genome prediction. *Plant Genome* 10 (2), 1–14. doi: 10.3835/plantgenome2016.10.0099

Caulfield, L. E., de Onis, M., Blössner, M., and Black, R. E. (2004). Undernutrition as an underlying cause of child deaths associated with diarrhea, pneumonia, malaria, and measles. *Am. J. Clin. Nutr.* 80, 193–198. doi:10.1093/ajcn/80.1.193

Chakraborti, M., Prasanna, B. M., Hossain, F., Singh, A. M., and Guleria, S. K. (2009). Genetic evaluation of kernel Fe and Zn concentrations and yield performance of selected maize (*Zea mays* L.) genotypes. *Range Manag. Agroforest.* 30, 109–114.

Chakraborti, M., Prasanna, B. M., Hossain, F., and Singh, A. M. (2011). Evaluation of single cross quality protein maize (QPM) hybrids for kernel iron and zinc concentrations. *Indian J. Genet.* 71, 312–319.

Chand, G., Muthusamy, V., Allen, T., Singh, B., Sarika, K., Zunjare, R., et al. (2019).Effect of opaque2 and opaque16 on endosperm modification in MAS-derived quality protein maize (QPM) genotypes, in: *Abstracts: XIV Agricultural Science Congress, NASC, Pusa, New Delhi, India, 20-23 February, 2019* (New Delhi: NAAS, ICAR, IARI) p. 459.

CIMMYT (1998). *A Complete Listing of Improved Maize Germplasm from CIMMYT. Maize Program Special Report* (Mexico, D.F.: CIMMYT).

Colangelo, E. P., and Guerinot, M. L. (2006). Put the metal to the petal: metal uptake and transport throughout plants. *Curr. Opin. Plant Biol.* 9, 322–330. doi:10.1016/j.pbi.2006.03.015

Das, A. K., Muthusamy, V., Zunjare, R. U., Chauhan, H. S., Sharma, P. K., Bhat, J. S., et al. (2019a). Genetic variability, genotype × environment interactions and combining ability analyses of kernel tocopherols among maize genotypes possessing novel allele of γ-tocopherol methyl transferase (ZmVTE4). *J. Cereal Sci.* 86, 1–8. doi: 10.1016/j.jcs.2018.12.018

Das, A. K., Chhabra, R., Muthusamy, V., Chauhan, H. S., Zunjare, R. U., and Hossain, F. (2019b). Identification of novel SNP and InDel variations in the promoter and 5′ untranslated regions of γ-tocopherol methyl transferase (ZmVTE4) affecting higher accumulation of α-tocopherol in maize kernel. *Crop J.* 7, 469–479. doi: 10.1016/j.cj.2019.01.004

Das, A. K., Jaiswal, S. K., Muthusamy, V., Zunjare, R. U., Chauhan, H. S., Chand, G., et al. (2019c). Molecular diversity and genetic variability of kernel tocopherols among maize inbreds possessing favourable haplotypes of γ-tocopherol methyl transferase (γ-VTE4). *J. Plant Biochem. Biotechnol.* 28, 253–262. doi: 10.1007/s13562-018-0470-x

De Winne, A., and Dirinck, P. (1996). Studies on vitamin E and meat quality. 2. Effect of feeding high vitamin E levels on chicken meat quality. *J. Agric. Food Chem.* 44, 1691–1696. doi:10.1021/jf9506848

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

Dhliwayo, T., Palacios-Rojas, N., Crossa, J., and Pixley, K. V. (2014). Effects of S_1 recurrent selection for provitamin A carotenoid content for three open-pollinated maize cultivars. *Crop Sci.* 54, 2449–2460. doi: 10.2135/cropsci2013.11.0764

Diepenbrock, C. H., Kandianis, C. B., Lipka, A. E., Magallanes-Lundback, M., Vaillancourt, B., Góngora-Castillo, E., et al. (2017). Novel loci underlie natural variation in vitamin E levels in maize grain. *Plant Cell* 29, 2374–2392. doi: 10.1105/tpc.17.00475

Egesel, C. O., Wong, J. C., Lambert, R. J., and Rocheford, T. R. (2003). Combining ability of maize inbreds for carotenoids and tocopherols. *Crop Sci.* 43, 818–823. doi: 10.2135/cropsci2003.8180

Ekpa, O., Palacios-Rojas, N., Kruseman, G., Fogliano, V., and Linnemann, A. (2018). Sub-Saharan African maize-based foods: technological perspectives to increase the food and nutrition security impacts of maize breeding programmes. *Global Food Secur.* 17, 48–56. doi: 10.1016/j.gfs.2018.03.007

Ekpa, O., Palacios-Rojas, N., Kruseman, G., Fogliano, V., and Linnemann, A. (2019). Sub-Saharan African maize-based foods: processing practices, challenges and opportunities. *Food Rev. Int.* 35, 609–639. doi: 10.1080/87559129.2019.1588290

Ertl, D. S., Young, K. A., and Raboy, V. (1998). Plant genetic approaches to phosphorus management in agricultural production. *J. Env. Qual.* 272, 299–304. doi: 10.2134/jeq1998.00472425002700020008x

FAOSTAT. (2017). FAO, Rome. Available at: www.faostat.org (accessed 22 October 2019).

Feng, F., Deng, F., Zhou, P., Yan, J., Wang, Q., Yang, R., et al. (2013). QTL mapping for the tocopherols at milk stage of kernel development in sweet corn. *Euphytica* 193, 409–417. doi: 10.1007/s10681-013-0948-5

Feng, F., Wang, Q., Liang, C., Yang, R., and Li, X. (2015). Enhancement of tocopherols in sweet corn by marker-assisted backcrossing of ZmVTE4. *Euphytica* 206, 513–521. doi: 10.1007/s10681-015-1519-8

Fryer, M. J. (1992). The antioxidant effects of thylakoid vitamin-E (alphatocopherol). *Plant Cell Environ.* 15, 381–392. doi: 10.1111/j.1365-3040.1992.tb00988.x

Garcia-Oliveira, A. L., Chander, S., Ortiz, R., Menkir, A., and Gedil, M. (2018). Genetic basis and breeding perspectives of grain iron and zinc enrichment in cereals. *Front. Plant Sci.* 9, 937. doi: 10.3389/fpls.2018.00937

Giuliano, G. (2017). Provitamin A biofortification of crop plants: a gold rush with many miners. *Curr. Opin. Biotechnol.* 44, 169–180. doi: 10.1016/j.copbio.2017.02.001

Goswami, R., Zunjare, R., Khan, S., Baveja, A., Muthusamy, V., and Hossain, F. (2019). Marker-assisted introgression of rare allele of crtRB1 gene into elite quality protein maize inbred for combining high lysine, tryptophan and provitamin A in maize. *Plant Breed.* 138, 174–183. doi: 10.1111/pbr.12676

Gowda, M., Worku, M., Nair, S. K., Palacios-Rojas, N., and Prasanna, B. M. (2017). *Quality Assurance/Quality Control (QA/QC) in Maize Breeding and Seed Production: Theory and Practice* (Nairobi: CIMMYT).

Guild, G., Parkes, E., Nutti, M., Palacios-Rojas, N., and Stangoulis, J. (2017). High-throughput measurement methodologies for developing nutrient-dense crops. *Afr. J. Food Agric. Nutr. Dev.* 17, 11941–11954.

Gupta, H. S., Raman, B., Agrawal, P. K., Mahajan, V., Hossain, F., and Nepolean, T. (2013). Accelerated development of quality protein maize hybrid through marker-assisted introgression of opaque-2 allele. *Plant Breed.* 132, 77–82. doi: 10.1111/pbr.12009

Guzmán, C., Medina-Larqué, A. S., Velu, G., González-Santoyo, H., Singh, R. P., Huerta-Espino, J., et al. (2014). Use of wheat genetic resources to develop biofortified wheat with enhanced grain zinc and iron concentrations and desirable processing quality. *J. Cereal Sci.* 60, 617–622. doi: 10.1016/j.jcs.2014.07.006

Harjes, C. E., Rocheford, T. R., Bai, L., Brutnell, T. P., Vallabhaneni, R., Williams, M., et al. (2008). Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319, 330–333. doi: 10.1126/science.1150255

Hindu, V., Palacios-Rojas, N., Babu, R., Suwarno, W. B., Rashid, Z., Usha, R., et al. (2018). Identification and validation of genomic regions influencing kernel zinc and iron in maize. *Theor. Appl. Genet.* 131, 1443–1457. doi: 10.1007/s00122-018-3089-3

Hossain, F., Muthusamy, V., Pandey, N., Vishwakarma, A. K., Baveja, A., Zunjare, R., et al. (2018a). Marker-assisted introgression of opaque2 allele for rapid conversion of elite hybrids into quality protein maize. *J. Genet.* 97, 287–298. doi: 10.1007/s12041-018-0914-z

Hossain, F., Muthusamy, V., and Zunjare, R. U. (2018b).Molecular Breeding for Development of Biofortified Maize Hybrids in India, in: *Extended Summaries: 13th Asian Maize Conference on and Expert Consultation on Maize for Food, Feed, Nutrition and Environmental Security, Ludhiana, India, October 8-10, 2018*, (Mexico: CIMMYT). pp. 220–230.

Hossain, F., Sarika, K., Muthusamy, V., Zunjare, R. U., and Gupta, H. S. (2019). "Quality Protein Maize for Nutritional Security," in *Quality Breeding in Field Crops*. Ed. A. M. I. Qureshi, Z. A. Dar and S. H. Wani (India: Springer), 217–237. doi: 10.1007/978-3-030-04609-5_11

Hotz, C., and McClafferty, B. (2007). From harvest to health: challenges for developing biofortified staple foods and determining their impact on micronutrient status. *Food Nutr. Bull.* 28, S271–S279. doi: 10.1177/15648265070282s206

Howe, J. A., and Tanumihardjo, S. A. (2006). Carotenoid-biofortified maize maintains adequate vitamin a status in Mongolian gerbils. *J. Nutr.* 136, 2562–2567. doi: 10.1177/156482650002100410

Institute of Medicine (2000). "Vitamin E," in *Dietary Reference Intakes for Ascorbic acid, Vitamin E, Selenium, and Carotenoids. Food and Nutrition Board, Institute of Medicine* (Washington DC: National Academy Press), 186–283. Chapter 6.

Institute of Medicine (2012). *Dietary reference intake (DRIs): Estimated average requirements* (Washington, DC: Food and Nutrition Board, Institute of Medicine). http://www.iom.edu/Activities/Nutrition/summaryDRIs/media/Files/Activity%20Files/Nutrition/DRIs/5_Summary%2Table%/20Tables%201-4pdf.

Jiang, T., Christian, P., Khatry, S. K., Wu, L., and West, K. P.Jr. (2005). Micronutrient deficiencies in early pregnancy are common, concurrent, and

vary by season among rural Nepali pregnant women. *J. Nutr.* 135, 1106–1112. doi: 10.1093/jn/135.5.1106

Jin, T., Zhou, J., Chen, J., Zhu, L., Zhao, Y., and Huang, Y. (2013). The genetic architecture of zinc and iron content in maize grains as revealed by QTL mapping and meta-analysis. *Breed. Sci.* 63, 317–324. doi: 10.1270/2Fjsbbs.63.317

Jorquera, M., Martinez, O. Z., Maruyama, F., Marschner, P., and de la Luz Mora, M. (2008). Current and future biotechnological applications of bacterial phytases and phytase-producing bacteria. *Microbes Environ.* 23, 182–191. doi: 10.1264/jsme2.23.182

Kim, Y. N., and Cho, Y. O. (2015). Vitamin E status of 20-to 59-year-old adults living in the Seoul metropolitan area of South Korea. *Nutr. Res. Pract.* 9, 192–198. doi: 10.4162%2Fnrp.2015.9.2.192

Li, Q., Yang, X., Xu, S., Cai, Y., Zhang, D., Han, Y., et al. (2012). Genome-wide association studies identified three independent polymorphisms associated with α-tocopherol content in maize kernels. *PLoS One* 7, e36807. doi: 10.1371/journal.pone.0036807

Liang, Z., Zhang, K., Chen, K., and Gao, C. (2014). Targeted mutagenesis in Zea mays using TALENs and the CRISPR/Cas system. *J. Genet. Genomics* 41, 63–68. doi: 10.1016/j.jgg.2013.12.001

Lipka, A. E., Gore, M. A., Magallanes-Lundback, M., Mesberg, A., Lin, H., Tiede, T., et al. (2013). Genome-wide association study and pathway level analysis of tocochromanol levels in maize grain. *G3 (Bethesda)* 3, 1287–1299. doi: 10.1534/g3.113.006148

Listman, G. M., Guzman, C., Palacios-Rojas, N., Pfeiffer, W., San Vicente, F., and Govidan, V. (2019). Improving nutrition through biofortifortication: preharvest and postharvest technologies. *Cereal Food World* 64, 3. doi: 10.1094/CFW-64-3-0025

Liu, L., Jeffers, D., Zhang, Y., Ding, M., Chen, W., Kang, M. S., et al. (2015). Introgression of the crtRB1 gene into quality protein maize inbred lines using molecular markers. *Mol. Breed.* 35 (8), 154. doi: 10.1007/s11032-015-0349-7

Liu, H., Shi, J., Sun, C., Gong, H., Fan, X., Qiu, F., et al. (2016). Gene duplication confers enhanced expression of 27-kDa γ-zein for endosperm modification in quality protein maize. *Proc. Nat. Acad. Sci.* 113, 4964–4969. doi: 10.1073/pnas.1601352113

Lorenz, A. J., Scott, M. P., and Lamkey, K. R. (2007). Quantitative determination of phytate and inorganic phosphorus for maize breeding. *Crop Sci.* 47, 600–604. doi: 10.2135/cropsci2006.03.0177

Manjeru, P., Van Biljon, A., and Labuschagne, M. (2019). The development and release of maize fortified with provitamin A carotenoids in developing countries. *Crit. Rev. Food Sci. Nutr.* 59, 1284–1293. doi: 10.1080/10408398.2017.1402751

Menkir, A., Liu, W., White, W. S., Mazlya-Dixon, B., and Rocheford, T. (2008). Carotenoid diversity in tropical-adapted yellow maize inbred lines. *Food Chem.* 109, 521–529. doi: 10.1016/j.foodchem.2008.01.002

Menkir, A., Maziya-Dixon, B., Mengesha, W., Rocheford, T., and Alamu, E. O. (2017). Accruing genetic gain in pro-vitamin A enrichment from harnessing diverse maize germplasm. *Euphytica* 213, 105. doi: 10.1007/s10681-017-1890-8

Menkir, A., Palacios-Rojas, N., Alamu, O., Dias, P., Maria, C., Dhliwayo, T., et al. (2018). *Vitamin A Biofortified Maize. Exploiting Native Genetic Variation for Nutrient Enrichment. Biofortification No.2 (February, 2018)*, (Bonn, Germany: CIMMYT, IITA, EMBRAPA, HarvestPlus and Crop Trust).

Menkir, A. (2008). Genetic variation for grain mineral content in tropical-adapted maize inbred lines. *Food Chem.* 110, 454–464. doi: 10.1016/j.foodchem.2008.02.025

Mitra, A. K., Alvarez, J. O., Guay-Woodford, L., Fuchs, G. J., Wahed, M. A., and Stephensen, C. B. (1998). Urinary retinol excretion and kidney function in children with shigellosis. *Am. J. Clin. Nutr.* 68, 1095–1103. doi: 10.1093/ajcn/68.5.1095

Muthusamy, V., Hossain, F., Thirunavukkarasu, N., Choudhary, M., Saha, S., Bhat, J. S., et al. (2014). Development of β-carotene rich maize hybrids through marker-assisted introgression of β-carotene hydroxylase allele. *PLoS One* 9, e113583. doi: 10.1371/journal.pone.0113583

Muzhingi, T., Palacios-Rojas, N., Miranda, A., Cabrera, M. L., Yeum, K. J., and Tang, G. (2017). Genetic variation of carotenoids, vitamin E and phenolic compounds in provitamin A biofortified maize. *J. Sci. Food Agric.* 97, 793–801. doi: 10.1002/jsfa.7798

Nair, S. K., Babu, R., Magorokosho, C., Mahuku, G., Semagn, K., Beyene, Y., et al. (2015). Fine mapping of *Msv1*, a major QTL for resistance to maize streak virus leads to development of production markers for breeding pipelines. *Theor. Appl. Genet.* 128, 1839–1854. doi: 10.1007/s00122-015-2551-8

Nuss, E. T., and Tanumihardjo, S. A. (2011). Quality protein maize for Africa: closing the protein inadequacy gap in vulnerable populations. *Adv. Nutr.* 2, 217–224. doi: 10.3945/an.110.000182

Ortiz, D., Rocheford, T., and Ferruzzi, M. G. (2016). Influence of temperature and humidity on the stability of carotenoids in biofortified maize (*Zea mays* L.) genotypes during controlled post-harvest storage. *J. Agric. Food Chem.* 64, 2727–2736. doi: 10.1021/acs.jafc.5b05698

Ortiz-Covarrubias, Y., Dhliwayo, T., Palacios-Rojas, N., Ndhlela, T., Magorokosho, C., Aguilar-Rincon, V. H., et al. (2019). Effects of drought and low nitrogen stress on provitamin A carotenoid content of biofortified maize hybrids. *Crop Sci.* 59, 2521–2532. doi: 10.2135/cropsci2019.02.0100

Ortiz-Monasterio, J. I., Palacios-Rojas, N., Meng, E., Pixley, K., Trethowan, R., and Pena, R. J. (2007). Enhancing the mineral and vitamin content of wheat and maize through plant breeding. *J. Cereal Sci.* 46, 293–307. doi: 10.1016/j.jcs.2007.06.005

Palacios-Rojas, N., Twumasi-Afriyie, S., Friesen, D., Teklewold, A., Wegary, D., De Groote, H., et al. (2018). *Lineamientos para el control de calidad de semilla y grano de maíz de alta calidad proteica (QPM): Experiencia en el desarrollo y promoción de QPM en Latinoamérica* (Mexico, D.F.: CIMMYT).

Pfeiffer, W. H., and McClafferty, B. (2007). HarvestPlus: breeding crops for better nutrition. *Crop Sci.* 47, 88–105. doi: 10.2135/cropsci2007.09.0020IPBS

Pilu, R., Panzeri, D., Gavazzi, G., Rasmussen, S. K., Consonni, G., and Nielsen, E. (2003). Phenotypic, genetic and molecular characterization of a maize low phytic acid mutant (lpa241). *Theor. Appl. Genet.* 107, 980–987. doi: 10.1007/s00122-003-1316-y

Pilu, R., Landoni, M., Cassani, E., Doria, E., and Nielsen, E. (2005). The maize lpa241 mutation causes a remarkable variability of expression and some pleiotropic effects. *Crop Sci.* 45, 2096–2105. doi: 10.2135/cropsci2004.0651

Pixley, K., Palacios, N. R., Babu, R., Mutale, R., Surles, R., and Simpungwe, E. (2013). "Biofortification of maize with provitamin A carotenoids," in *Carotenoids in Human Health*. Ed. S. A. Tanumihardo (New York: Springer Science and Business Media), 271–292.

Prasanna, B. M., Vasal, S. K., Kassahun, B., and Singh, N. N. (2001). Quality protein maize. *Curr. Sci.* 81, 1308–1319.

Prasanna, B. M., Mazumdar, S., Chakraborti, M., Hossain, F., Manjaiah, K. M., Agrawal, P. K., et al. (2011). Genetic variability and genotype × year interactions for kernel iron and zinc concentration in maize (*Zea mays* L.). *Indian J. Agric. Sci.* 81, 704–711. https://repository.cimmyt.org/xmlui/bitstream/handle/10883/1886/96904.pdf

Qin, H., Cai, Y., Liu, Z., Wang, G., Wang, J., Guo, Y., et al. (2012). Identification of QTL for zinc and iron concentration in maize kernel and cob. *Euphytica* 187, 345–358. doi: 10.1007/s10681-012-0692-2

Raboy, V., Gerbasi, P. F., Young, K. A., Stoneberg, S. D., Pickett, S. G., Bauman, A. T., et al. (2000). Origin and seed phenotype of maize low phytic acid 1-1 and low phytic acid 2-1. *Plant Physiol.* 124, 355–368. doi: 10.1104/pp.124.1.355

Raboy, V. (2001). Genetics and breeding of seed phosphorus and phytic acid. *J. Plant Physiol.* 158, 489–497. doi: 10.1078/0176-1617-00361

Rautiainen, S., Manson, J. E., Lichtenstein, A. H., and Sesso, H. D. (2016). Dietary supplements and disease prevention – a global overview. *Nat. Rev. Endocrinol.* 12, 407–420. doi: 10.1038/nrendo.2016.54

Reynolds, M., Borrell, A., Braun, H., Edmeades, G., Flavell, R., Gwyn, J., et al. (2019). Translational research for climate resilient, higher yielding crops. *Crop Breed. Genet. Genomics* 1, e190016. doi: 10.20900/cbgg20190016

Rocheford, T. R., Wong, J. C., Egesel, C. O., and Lambert, R. J. (2002). Enhancement of vitamin E levels in corn. *J. Am. Coll. Nutr.* 21, 191S–198S. doi: 10.1080/07315724.2002.10719265

Sanders, S. K., Morgan, J. B., Wulf, D. M., Tatum, J. D., Williams, S. N., and Smith, G. C. (1997). Vitamin E supplementation of cattle and shelf-life of beef for the Japanese market. *J. Animal Sci.* 75, 2634–2640. http://jas.fass.org/content/75/10/2634

Sarika, K., Hossain, F., Muthusamy, V., Baveja, A., Zunjare, R., Goswami, R., et al. (2017). Exploration of novel opaque16 mutation as a source for high-lysine and -tryptophan in maize endosperm. *Indian J. Genet.* 77, 59–64. doi: 10.5958/0975-6906.2017.00008.6

Sarika, K., Hossain, F., Muthusamy, V., Zunjare, R., Goswami, R., Thirunavukkarasu, N., et al. (2018a). opaque16, a high lysine and

tryptophan mutant, does not influence the key physico-biochemical characteristics in maize kernel. *PloS One* 13, e0190945. doi: 10.1371/journal

Sarika, K., Hossain, F., Muthusamy, V., Zunjare, R. U., Baveja, A., Goswami, R., et al. (2018b). Marker-assisted pyramiding of opaque2 and novel opaque16 genes for further enrichment of lysine and tryptophan in sub-tropical maize. *Plant Sci.* 272, 142–152. doi: 10.1016/j.plantsci.2018.04.014

Schmaelzle, S., Gannon, B., Crawford, S., Arscott, S. A., Goltz, S., Palacios-Rojas, N., et al. (2014). Maize genotype and food matrix affect the provitamin A carotenoid bioefficacy from staple and carrot-fortified feeds in Mongolian gerbils (*Meriones unguiculatus*). *J. Agric. Food Chem.* 62, 136–143. doi: 10.1021/jf403548w

Shamim, A. A., Schulze, K. J., Merrill, R. D., Kabir, A., Christian, P., Shaikh, S., et al. (2015). First trimester plasma tocopherols are associated with risk of miscarriage in rural Bangladesh. *Am. J. Clin. Nutr.* 101, 294–301. doi: 10.3945/ajcn.114.094920

Shutu, X., Dalong, Z., Ye, C., Yi, Z., Shah, T., Ali, F., et al. (2012). Dissecting tocopherols content in maize (Zea mays L.) using two segregating populations and high-density single nucleotide polymorphism markers. *BMC Plant Biol.* 12 (1), 201. http://www.biomedcentral.com/1471-2229/12/201

Šimić, D., Drinić, S. M., Zdunić, Z., Jambrović, A., Ledenčan, T., Brkić, J., et al. (2012). Quantitative trait loci for biofortification traits in maize grain. *J. Hered.* 103, 47–54. doi: 10.1093/jhered/esr122

Simpungwe, E., Dhliwayo, T., Palenberg, M., Taleon, V., Birol, E., Oparinde, A., et al. (2017). Orange maize in Zambia: crop development and delivery experience. *Afr. J. Food Agric. Nutr. Dev.* 17, 11973–11999. doi: 10.18697/ajfand.78.HarvestPlus08

Sowa, M., Yu, J., Palacios-Rojas, N., Goltz, S. R., Howe, J. A., Davis, C. R., et al. (2017). Retention of carotenoids in biofortified maize flour and β-cryptoxanthin-enhanced eggs after household cooking. *ACS Omega* 10, 7320–7328. doi: 10.1021/acsomega.7b01202

Sugiura, M., Ogawa, K., and Yano, M. (2014). Comparison of bioavailability between beta-cryptoxanthin and beta-carotene and tissue distribution in its intact form in rats. *Biosci. Biotechnol. Biochem.* 78, 307–310. doi: 10.1080/09168451.2014.878220

Sureshkumar, S., Tamilkumar, P., Senthil, N., Nagarajan, P., Thangavelu, A. U., Raveendran, M., et al. (2014). Marker assisted selection of low phytic acid trait in maize (*Zea mays* L.). *Hereditas* 151, 20–27. doi: 10.1111/j.1601-5223.2013.00030.x

Suwarno, W. B., Pixley, K. V., Palacios-Rojas, N., Kaeppler, S. M., and Babu, R. (2014). Formation of heterotic groups and understanding genetic effects in a provitamin A biofortified maize breeding program. *Crop Sci.* 54, 14–24. doi: 10.2135/cropsci2013.02.0096

Suwarno, W., Pixley, K. V., Palacios-Rojas, N., Kaeppler, S., and Babu, R. (2015). Genome-wide association analysis reveals new targets for carotenoid biofortification in maize. *Theor. Appl. Genet.* 128, 851–864. doi: 10.1007%2Fs00122-015-2475-3

Taleon, V., Mugode, L., Cabrera-Soto, L., and Palacios-Rojas, N. (2017). Carotenoid retention in biofortified maize using different post-harvest storage and packaging methods. *Food Chem.* 232, 60–66. doi: 10.1016/j.foodchem.2017.03.158

Talsma, E. F., Melse-Boonstra, A., and Brouwer, I. D. (2017). Acceptance and adoption of biofortified crops in low- and middle-income countries: a systematic review. *Nutr. Rev.* 75, 798–829. doi: 10.1093/nutrit/nux037

Tamilkumar, P., Senthil, N., Sureshkumar, S., Thangavelu, A. U., Nagarajan, P., Vellaikumar, S., et al. (2014). Introgression of low phytic acid locus (lpa2-2) into elite maize (*Zea mays* L.) inbred through marker assisted backcross breeding. *Aust. J. Crop Sci.* 8, 1224–1231.

Tanumihardjo, S., McCulley, L., Roh, R., Lopez-Ridaura, S., Palacios-Rojas, N., and Gunaratna, N. (2019). Maize agro-food systems to ensure food and nutrition security in reference to the Sustainable Development Goals. *Global Food Sec.* doi: 10.1016/j.gfs.2019.100327

Tian, Q. Z., Li, X. H., Li, M. S., Jiang, W., and Zhang, S. H. (2004). Molecular markers assisted selection to quality protein maize. *J. Maize Sci.* 12, 108–110.

Tome, D., and Bos, C. (2007). Lysinerequirement through the human life cycle. *J. Nutr.* 137, 1642–1645. doi: 10.1093/jn/137.6.1642S

UNICEF (2016). Annual Report 2016. (New York: UNICEF). https://www.unicef.org/publications/index_96412.html.

Vivek, B. S., Krivanek, A. F., Palacios-Rojas, N., Twumasi-Afriyie, S., and Diallo, A. O. (2008). *Breeding Quality Protein Maize (QPM): Protocols for Developing QPM Cultivars* (Mexico, D.F.: CIMMYT).

Von Lintig, J. (2010). Colors with functions: elucidating the biochemical and molecular basis of carotenoid metabolism. *Annu. Rev. Nutr.* 30, 35–56. doi: 10.1146/annurev-nutr-080508-141027

WHO/FAO/UNU. (2007). *Protein and amino acid requirements in human nutrition. Report of a Joint WHO/FAO/UNU Expert Consultation, WHO Technical Report Series, No. 935*, (Geneva: Swizerkand).

Willett, W., Rockström, J., Loken, B., Springmann, M., Lang, T., Vermeulen, S., et al. (2019). Food in the Anthropocene: the EAT–Lancet Commission on healthy diets from sustainable food systems. *Lancet* 393, 447–492. doi: 10.1016/S0140-6736(18)31788-4

Wong, J. C., Lambert, R. J., Tadmor, Y., and Rocheford, T. R. (2003). QTL associated with accumulation of tocopherols in maize. *Crop Sci.* 43, 2257–2266. doi: 10.2135/cropsci2003.2257

Wurtzel, E. T., Cuttriss, A., and Vallabhaneni, R. (2012). Maize provitamin a carotenoids, current resources, and future metabolic engineering challenges. *Front. Plant Sci.* 3, 29. doi: 10.3389/fpls.2012.00029

Yadava, D. K., Hossain, F., and Mohapatra, T. (2018). Nutritional security through crop biofortification in India: status and future prospects. *Indian J. Med. Res.* 148, 621–631. doi: 10.4103/ijmr.IJMR_1893_18

Yan, J., Kandianis, C. B., Harjes, C. E., Bai, L., Kim, E.-H., Yang, X., et al. (2010). Rare genetic variation at Zea mays crtRB1 increases β-carotene in maize grain. *Nat. Genetics* 42, 322–329. doi: 10.1038/ng.551

Yang, W., Zheng, Y., Zheng, W., and Feng, R. (2005). Molecular genetic mapping of a high-lysine mutant gene (opaque-16) and the double recessive effect with opaque-2 in maize. *Mol. Breed.* 15, 257–269. doi: 10.1007/s11032-004-5947-8

Yang, L., Wang, W., Yang, W., and Wang, M. (2013). Marker-assisted selection for pyramiding the waxy and opaque16 genes in maize using cross and backcross schemes. *Mol. Breed.* 31, 767–775. doi: 10.1007/s11032-012-9830-8

Yuan, Y., Cairns, J. E., Babu, R., Gowda, M., Makumbi, D., Magorokosho, C., et al. (2019). Genome-wide association mapping and genomic prediction analyses reveal the genetic architecture of grain yield and flowering time under drought and heat stress conditions in maize. *Front. Plant Sci.* 9, 1919. doi: 10.3389/fpls.2018.01919

Zarkadas, C. G., Yu, Z., Hamilton, R. I., Pattison, P. L., and Rose, N. G. W. (1995). Comparison between the protein quality of northern adapted cultivars of common maize and quality protein maize. *J. Agric. Food Chem.* 43, 84–93. doi: 10.1021/jf00049a016

Zhang, W. L., Yang, W. P., Chen, Z. W., Wang, M. C., Yang, L. Q., and Cai, Y. L. (2010). Molecular marker-assisted selection for o2 introgression lines with o16 gene in corn. *Acta Agron. Sin.* 36, 1302–1309. doi: 10.1016/S1875-2780(09)60067-5

Zhang, X., Pfeiffer, W. H., Palacios-Rojas, N., Babu, R., Bouis, H., and Wang, J. (2012). Probability of success of breeding strategies for improving provitamin A content in maize. *Theor. Appl. Genet.* 125, 235–246. doi: 10.1007/s00122-012-1828-4

Zhang, W., Yang, W., Wang, M., Wang, W., Zeng, G., Chen, Z., et al. (2013). Increasing lysine content of waxy maize through introgression of opaque2 and opaque16 genes using molecular assisted and biochemical development. *PLoS One* 8, 1–10. doi: 10.1371/journal.pone.0056227

Zhang, X., Perez-Rodriguez, P., Semagn, K., Beyene, Y., Babu, R., Lopez-Cruz, M. A., et al. (2015). Genomic prediction in biparental tropical maize populations in water-stressed and well-watered environments using low-density and GBS SNPs. *Heredity* 114, 291–299. doi: 10.1038/hdy.2014.99

Zhou, J.R., and Erdman, J. W. (1995). Phytic acid in health and disease. *Food Sci. Nutr.* 35, 495–508. doi: 10.1080/10408399509527712

Zunjare, R. U., Hossain, F., Muthusamy, M., Baveja, A., Chauhan, H. S., Bhat, J. S., et al. (2018). Development of biofortified maize hybrids through marker-assisted stacking of β-carotene hydroxylase, lycopene-ϵ-cyclase and opaque2 genes. *Front. Plant Sci.* 9, 178. doi: 10.3389/fpls.2018.00178

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 Genetics

REVIEW
published: 27 February 2020
doi: 10.3389/fgene.2020.00049

Healthy and Resilient Cereals and Pseudo-Cereals for Marginal Agriculture: Molecular Advances for Improving Nutrient Bioavailability

Juan Pablo Rodríguez[†], *Hifzur Rahman*[†], *Sumitha Thushar and Rakesh K. Singh*[*]

Crop Diversification and Genetics Program, International Center for Biosaline Agriculture, Dubai, United Arab Emirates

OPEN ACCESS

Edited by:
Felipe Klein Ricachenevsky, Universidade Federal de Santa Maria, Brazil

Reviewed by:
Joel B. Mason, Tufts University School of Medicine, United States
Scott Aleksander Sinclair, Ruhr University Bochum, Germany

***Correspondence:**
Rakesh K. Singh
r.singh@biosaline.org.ae;
rksinghirri@gmail.com

[†]*These authors have contributed equally to this work*

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *07 November 2019*
Accepted: *16 January 2020*
Published: *27 February 2020*

Citation:
Rodríguez JP, Rahman H, Thushar S and Singh RK (2020) Healthy and Resilient Cereals and Pseudo-Cereals for Marginal Agriculture: Molecular Advances for Improving Nutrient Bioavailability. Front. Genet. 11:49. doi: 10.3389/fgene.2020.00049

With the ever-increasing world population, an extra 1.5 billion mouths need to be fed by 2050 with continuously dwindling arable land. Hence, it is imperative that extra food come from the marginal lands that are expected to be unsuitable for growing major staple crops under the adverse climate change scenario. Crop diversity provides right alternatives for marginal environments to improve food, feed, and nutritional security. Well-adapted and climate-resilient crops will be the best fit for such a scenario to produce seed and biomass. The minor millets are known for their high nutritional profile and better resilience for several abiotic stresses that make them the suitable crops for arid and salt-affected soils and poor-quality waters. Finger millet (*Eleucine coracana*) and foxtail millet (*Setaria italica*), also considered as orphan crops, are highly tolerant grass crop species that grow well in marginal and degraded lands of Africa and Asia with better nutritional profile. Another category of grains, called pseudo-cereals, is considered as rich foods because of their protein quality and content, high mineral content, and healthy and balance food quality. Quinoa (*Chenopodium quinoa*), amaranth (*Amaranthus* sp.), and buckwheat (*Fagopyrum esculentum*) fall under this category. Nevertheless, both minor millets and pseudo-cereals are morphologically different, although similar for micronutrient bioavailability, and their grains are gluten-free. The cultivation of these millets can make dry lands productive and ensure future food as well as nutritional security. Although the natural nutrient profile of these crop plant species is remarkably good, little development has occurred in advances in molecular genetics and breeding efforts to improve the bioavailability of nutrients. Recent advances in NGS have enabled the genome and transcriptome sequencing of these millets and pseudo-cereals for the faster development of molecular markers and application in molecular breeding. Genomic information on finger millet (1,196 Mb with 85,243 genes); *S. italica*, a model small millet (well-annotated draft genome of 420 Mb with 38,801 protein-coding genes); amaranth (466 Mb genome and 23,059 protein-coding genes); buckwheat (genome size of 1.12 Gb with 35,816 annotated genes); and quinoa (genome size of 1.5 Gb containing 54,438 protein-coding genes) could pave the way for the genetic improvement of these grains. These genomic resources are an important first step toward genetic improvement of these crops. This

review highlights the current advances and available resources on genomics to improve nutrient bioavailability in these five suitable crops for the sustained healthy livelihood.

Keywords: orphan crops, nutrient-rich crops, pseudo-cereals, molecular profiles, food and nutritional security, healthy crops, marginal environment

INTRODUCTION

All foods have uniqueness in their composition, with a specific range of macro- and micronutrients in different combinations. Some foods are rich in carbohydrate (rice, wheat, maize, etc.), protein (pulses, spinach), fat (oilseeds, groundnut), or mineral (pearl millet, etc.), while some are nutrient-dense and have optimum combinations of nutrients with good digestibility (most of the minor millets, quinoa, etc.). These nutrient-dense foods with a proper mix of nutrients and high bioavailability are sometimes designated as superfoods, but the irony is that most of these are also considered as orphan crops because of lower cropped area, low demand, and low consumption. Minor millets and pseudo-cereals come under the category of underused and neglected crops.

Every day, arable non-stress areas are being converted into marginal lands at an alarming rate, which may further increase under the climate change scenario if business as usual (RCP8.5) continues without arresting the trend of global warming as per the IPCC AR5 (Porter et al., 2014). Recent studies based on compiled data from 53 countries with serious, alarming, or seriously alarming conditions clearly show that more than 50% (27 countries) had decreased consumed calories due to mean climate change (Ray et al., 2019). Although the productivity of major crops has been predicted to be lower under marginal environments, the good news is that the nutrient-rich underused neglected crops are very resilient to harsh environments (drought, salinity, and extreme temperature) and yield well with limited resources (Mabhaudhi et al., 2019). As per FAO estimates (FAO, 2017), most of the food needed for 2 billion more mouths by 2050 has to come from marginal environments. A "Next-Generation Green Revolution" is required to achieve food security, which is a much broader and systems-based approach, and this has to come from the areas that were left out of the first Green Revolution, to achieve future food security in a more sustainable way across all spectrums of society (Nusslein et al., 2016; Dhankher and Foyer, 2018). Huge scope exists for the genetic improvement of these crops but, unfortunately, only limited research programs are undertaking focused research on such crops worldwide. Small and complex flower shape has compounded the difficulty to handle these crops; hence, few genetic studies are being undertaken. This is one of the major factors for slowdown of the research on developing molecular markers for the agronomically important traits to be used in breeding programs. Most genomics studies focus on population structures, grouping, and evolution. The current paper is a review of the scientific work on neglected but nutrient-rich crops such as quinoa, amaranth, finger millet, foxtail millet, and buckwheat. In recent years, progress in addressing all forms of malnutrition has seen a declining trend, alarmingly slow, with >150 million children still stunted (Global Nutrition Report, 2018), but the movement has increased awareness among people and hence increased health-conscious people demanding healthy food relatively more often. This review will deal with five underused but nutritionally important and rich crops: finger millet (*Eleucine coracana*), foxtail millet (*Setaria italica*), quinoa (*Chenopodium quinoa*), amaranth (*Amaranthus* sp.), and buckwheat (*Fagopyrum esculentum*).

BIOAVAILABILITY AND HEALTHY FOOD

Bioavailability, in simple terms, refers to whatever is absorbed out of ingested food and goes into the bloodstream. Many factors could affect this process, but the prominent ones are the original profile of food, processing of food, and digestion efficiency. Nutritional composition is the mix of macronutrients (carbohydrate, fat, and protein) and micronutrients (minerals and vitamins) within a product, but their absorption efficiency in the body depends on the ability of food to be digested easily. Food that is absorbed easily maintains or improves health and energy status by providing appropriate macro- and micronutrients in balanced form and is thus called healthy food. Neglected or underused crops such as minor millets and pseudo-cereals were part of the common diet of ancient cultures but slowly, after the Green Revolution, the higher availability and accessibility of rice, wheat, and maize overtook these neglected crops and started providing >60% of the calorific intake through these three crops only, thus starting to create a nutrient-imbalanced diet. Not many studies exist on the bioavailability of the nutrients provided by nutrient-dense underused crops.

BIOAVAILABILITY AND IMPROVEMENT IN MOLECULAR GENETICS

Finger Millet

Millets serve as a good food source of carbohydrates, proteins, minerals, and vitamins. Although they are the basic food ingredient in the diets of millions of people living in the semi-arid and arid regions of the world, they are still sometimes referred to as orphan crops or even lost crops. These neglected crops are mostly cultivated in developing countries and their world production statistics show low volumes vis-à-vis other popular food crops. These neglected crops are important because of their contribution to biodiversity and climatic resilience, their rich nutrition profile, and their means of livelihood of the poor in various parts of the world (Belton and Taylor, 2004). Finger millet is one of the most efficient crops for nitrogen use efficiency (NUE) and it can grow well with less water requirement, hence

well suited to semi-arid climates (Gupta et al., 2017). It is also responsive to nutrients but has the ability to do well under limited resources (Gull et al., 2014). The most important part is its excellent storing capacity without deterioration even with significant insect and pest attacks. This has earned it the popular name of "famine crop" as it can resist storage pests for as long as 10 years, ensuring a year-round food supply (Mgonja et al., 2007).

Among the millet crops, six crops are called minor millets due to their small size: finger millet (*Eleusine coracana* (L.) Gaertn.), foxtail millet (*S. italica* (L.) P. Beauv.), kodo millet (*Paspalum scrobiculatum* L.), proso millet (*Panicum miliaceum* L.), barnyard millet (*Echinochloa* spp.), and little millet (*Panicum sumatrense* Roth). All these small minor millets are known for their unique nutritional composition and resilience (Kumar et al., 2018). Only foxtail and finger millet from this group will be discussed here. Finger millet is relatively popular in India and many countries in Africa because of its resilience and nutrient-dense grain profile. It has been promoted in Africa to reduce anemic incidence in children (Tripathi and Patel, 2010; Udeh et al., 2017). The nutraceutical importance of finger millet lies in its high content of calcium (0.38%), protein (6–13%), dietary fiber (10–18%), carbohydrate (65–75%), and minerals (2.5–3.5%). Another quality of finger millet is that it is gluten free with low glycemic index (GI) hence suitable for the people suffering from gluten intolerance/coeliac disease as well as diabetes (**Tables 1** and **2**). It is rich in ergocalciferol (vitamin D) and essential amino acids (EAA) such as valine, phenyl-alanine, leucine, and histidine (**Tables 3** and **4**). Besides these important nutrients, it has phytates (0.48%), tannins (0.61%), phenolic compounds (0.3–3.0%), and trypsin inhibitory factors that affect the bioavailability of nutrients; that is why proper processing is important to exploit its positive nutritional qualities (Devi et al., 2014) (**Tables 5** and **6**). Chauhan and Sarita (2018) have reported increased bioavailability of minerals such as Fe and P and vitamins upon grain processing before consumption. Phytates and tannins have negative effects on the bioavailability of nutrients, but processing at germination and little fermentation of grains increase the availability of minerals, amino acids, and free sugars, along with digestibility (Sripriya et al., 1997). Finger millet is mostly consumed as flour, but its processing through germinating and

TABLE 2 | Mineral composition in grains (mg/100 g).

		Finger millet	Foxtail millet	Quinoa	Amaranth	Buckwheat
Calcium	Ca	364	15.27	198	181	18
Copper	Cu	0.67	0.26	0.48	0.81	1.10
Iron	Fe	4.62	2.34	7.51	9.33	2.20
Magnesium	Mg	146	122	119	325	231
Manganese	Mn	3.19	0.33	1.77	5.29	1.30
Phosphorus	P	210	101	212	374	347
Potassium	K	443	94	474	433	460
Sodium	Na	4.75	3.35	4.50	2.70	1.00
Zinc	Zn	2.53	1.65	3.31	2.66	2.40

Compiled from Johnson and Croissant (1985); Gopalan et al. (1989); Longvah (2017); Dayakar et al. (2017); USDA National Nutrient Database for Standard Reference. https://fdc.nal.usda.gov/(accessed on October 28, 2019).

TABLE 3 | Amino acid profile (g/100 g protein).

		Finger millet	Foxtail millet	Quinoa	Amaranth	Buckwheat
Alanine	ALA	6.71	11.00	4.35	4.26	4.50
Arginine	ARG	4.33	3.18	7.85	7.77	9.70
Aspartic acid	ASP	6.40	5.61	8.40	12.57	11.30
Glutamic acid	GLU	20.22	18.25	13.75	16.12	18.60
Glycine	GLY	3.59	3.12	4.80	8.50	6.30
Proline	PRO	5.42	7.33	5.67	3.76	3.80
Serine	SER	4.81	5.50	4.56	7.79	4.70
Tyrosine	TYR	3.37	3.87	1.98	2.85	2.10
Histidine	HIS	2.37	2.14	2.98	1.86	2.70
Isoleucine	ILE	3.70	4.55	3.75	2.82	3.80
Leucine	LEU	8.86	11.96	6.08	4.83	6.40
Lysine	LYS	2.83	1.42	5.55	5.45	6.10
Methionine	MET	2.74	2.69	2.24	1.86	2.50
Cystine	CYS	1.48	1.92	1.85	1.60	1.60
Phenyl-alanine	PHE	5.70	6.27	4.35	3.98	4.80
Threonine	THR	3.84	3.89	3.01	3.02	3.90
Tryptophan	TRP	0.91	1.32	1.25	1.05	2.00
Valine	VAL	5.65	5.49	4.55	4.34	4.70

Compiled from Pomeranz and Robbins (1972); Johnson and Croissant (1985); Gopalan et al. (1989); Ikeda and Kishida (1993); Longvah (2017); Dayakar et al. (2017); USDA National Nutrient Database for Standard Reference. https://fdc.nal.usda.gov/(accessed on October 28, 2019).

TABLE 1 | Proximate component profile of grains.

	Crude protein	Total fat	Dietary fiber			Carbohydrate	Gluten presence	GI
			Total	Insoluble	Soluble			
Finger millet (*Eleucine coracana*)	7.16	1.92	11.18	9.51	1.67	66.82	No	Low
Foxtail millet (*Setaria italica*)	8.92	2.55	6.39	4.29	2.11	66.19	No	Low
Quinoa (*Chenopodium quinoa*)	13.11	5.50	14.66	10.21	4.46	53.65	No	Low
Amaranth (*Amaranthus* spp.)	14.59	5.74	7.02	5.76	1.26	59.98	No	High
Buckwheat (*Fagopyrum esculentum*)	13.25	3.40	10.00	–	–	71.50	No	Low

All values are expressed in percentage of edible portion. Hyphens (-) in the tables represent either below detectable limit or not reported. Compiled from Johnson and Croissant (1985); Gopalan et al. (1989); Longvah (2017); Dayakar et al. (2017); USDA National Nutrient Database for Standard Reference. https://fdc.nal.usda.gov/(accessed on October 28, 2019). GI, glycemic index; (low <55; intermediate 55–70; and high >70).

TABLE 4 | Vitamins in minor millets and pseudo-cereals.

		Unit	Finger millet	Foxtail millet	Quinoa	Amaranth	Buckwheat
Fat soluble	α-Ergocalciferol (vit. D)	µg	41.46	–	–	0.04	–
	α-Tocopherol (vit. E)	mg	0.16	–	2.08	1.92	0.32
	Phylloquinones (vit. K1)	µg	3.00	–	2.00	–	7.00
Water soluble	Thiamine (vit. B_1)	mg	0.37	0.59	0.83	0.04	0.42
	Riboflavin (vit. B_2)	mg	0.17	0.11	0.22	0.04	0.19
	Niacin (vit. B_3)	mg	1.34	3.20	1.70	0.45	6.15
	Pantothenic acid (vit. B_5)	mg	0.29	0.82	0.62	0.24	0.44
	Vit. B_6	mg	0.05	–	0.21	0.50	0.58
	Biotin (vit. B_7)	µg	0.88	–	0.62	1.92	–
	Folates (vit. B_9)	µg	34.66	–	1.73	27.44	54.00

All values are expressed per 100 g edible portion; all blank spaces (-) in the tables represent either below detectable limit or not reported. Compiled from Johnson and Croissant (1985); Gopalan et al. (1989); Longvah (2017); Dayakar et al. (2017); USDA National Nutrient Database for Standard Reference. https://fdc.nal.usda.gov/(accessed on October 28, 2019).

fermenting makes its iron content much higher in grains (Tatala et al., 2007). Seed germination of finger millet can increase the bioavailability of iron from 0.75 to 1.25 mg/100 g and is a potential alternative to mitigate anemia (Tatala et al., 2007). The pre-processing of grains can drastically reduce the impact of anti-nutrients and can improve iron bioavailability and bioactive compounds, which are confirmed by several scientific studies (Tatala et al., 2007; Hithamani and Srinivasan, 2014; Udeh et al., 2017) (**Tables 5** and **6**).

Singh et al. (2018) showed that traditional knowledge practiced by farmers to roast finger millet decreases phytochemical composition, moisture, protein, and antioxidant action, but increases fat, ash, and fiber and improves the bioavailability of iron and calcium. The improvement of iron bioavailability to reduce anemia happens *via* the biochemical changes in fortification with ferrous fumarate, ferric pyrophosphate (6 mg/kg), and zinc oxide (50 mg/kg) in finger millet flour (Tripathi and Patel, 2010). The diversity of finger millet offers a rich source of several antioxidants and calcium in the grains as polyphenols (0.3–3.0%) that possess hypoglycemic, hypercholesterolemic, and anti-ulcerative properties (Chethan and Malleshi, 2007). Growing research interest exists in finger millet that could be attributed to several bioactive compounds such as ferulic acid-rich arabinoxylans, ferulic acid, caffeic acid, quercetin, and flavonoids, which are bio-accessible and have multiple therapeutic effects (Udeh et al., 2017). Hithamani and Srinivasan (2017) demonstrated that the bioavailability of phenolic compounds extracted from finger millet grain and co-administered to rats with piperine had a therapeutic benefit. The wide spectrum of phenolic compounds greatly enhances the nutraceutical potential of finger millet. Unprocessed and processed finger millet flour use in wafer and vermicelli (a fine noodle) have shown that bio-accessibility of Fe, Zn, and Ca through *in vitro* digestibility of starch (IVSD) and protein (IVPD) and bioactive polyphenols and flavonoids could be changed just by processing (Oghbaei and Prakash, 2012).

Genomics of Finger Millet

Finger millet [*E. coracana* (L.) Gaertn.] is a self-pollinated allotetraploid (2n = 4x = 36, AABB) species with a genome size of 1.593 Gb (**Table 7**). The 2C DNA amount in *E. coracana* is 3.36–3.87 picogram (pg) (Mysore and Baird, 1997). It is an annual C4 herbaceous cereal crop belonging to family Poaceae and sub-family Chloridiodeae and exhibits morphological similarity to *E. coracana* subsp. *africana* and *E. indica*. Cytological studies, isozyme chloroplast DNA, and genomic *in situ* hybridization (GISH) have shown that the maternal diploid genome (AA) of *E. coracana* originated from *E. indica* whereas *E. floccifolia* is supposed to be the donor of the B genome to the polyploid species *E. coracana* (Bisht and Mukai, 2001). Because of its resilient nature, it is widely grown in arid and semi-arid areas of India and Africa.

Molecular Markers, Genetic Diversity, and Phylogenetic Studies in Finger Millet

Despite the nutritional benefits and climate-resilient nature of finger millet, the available genomic resources are limited, which has slowed the pace of genetic improvement of this crop (Saha et al., 2016). Immense morphological diversity is present in finger millet with a range of seed color correlated with protein and calcium content, time to maturity, and drought and salinity tolerance (Vadivoo et al., 1998; Tsehaye et al., 2006). Very few reports exist on the use of molecular markers for studying genetic diversity in finger millet, although the development and use of molecular markers in genomic studies of finger millet started a decade ago. Arya et al. (2009) developed 31 expressed sequence tag simple sequence repeats (EST-SSRs), out of which 17 were amplified and nine were found to be polymorphic between 11 elite germplasm accessions of finger millet of Indian and African origin. Reddy et al. (2012) identified 132 EST-based SSRs and developed 30 SSR primers for assessing genetic diversity in 15 finger millet accessions. Out of 30 EST-SSRs, 20 primers showed polymorphism and 13 primers were found to have polymorphism information content (PIC) value above 0.5. Using transcriptome data, Selvam et al. (2015) identified several SSRs and designed and validated 12 SSR primers on 23 finger millet accessions, where the primers showed an average PIC value of 0.67. Dida et al. (2007), using random *HindIII*, *PstI*, and *SalI* libraries, developed 82 genomic SSR markers and developed the first genetic map of finger millet using genomic SSRs, restriction fragment length polymorphism (RFLP), amplified fragment length polymorphism (AFLP), and EST

TABLE 5 | Effects, mechanism, and process of increasing bioavailability of cereals and pseudo-cereal grains.

Crop	Effects	Mechanism	Process to increase bioavailability	References
Finger millet	Reduction in viscosity of weaning food	NA	Malting	Seenappa (1988)
	Eliminate stickiness of cooked millet	NA	Parboiling	Desikachar (1975)
	Flour quality can be increased	NA	Decortication	Geervani and Eggum (1989)
	Loss of protein, mineral, and fiber content	NA	Dehulling, soaking, and cooking	Panwal and Pawar (1989)
	Increase in *in vitro* protein digestibility (IVPD)	NA	Dehulling of seeds	Ramachandra et al. (1977)
	Effective removal of polyphenols and phytates	NA	Dehulling followed by soaking	Pawar and Parlikar (1990)
	Improve recovery of soluble protein and its digestibility *in vitro*	NA		
Foxtail millet	Significant increase in extractability of calcium, phosphorus, iron, zinc, and copper	NA	Roasting	Gahlawat and Sehgal (1995)
	Digestibility and biological values increased	NA	Fortified with lysine	Ganapathy et al. (1957)
	Highest concentration of thiamine, vitamin E, and stearic and linoleic acid	NA	NA	Bandyopadhyay et al. (2017)
	Loss of protein, mineral, and fiber content	NA	Dehulling/soaking/ cooking	Pawar and Machewad (2006)
	Increase in percentage of ionizable iron and soluble zinc	By the removal of polyphenols and breaking down of polyphenols-protein-minerals		
	Two types of fatty acid patterns observed	Glutinous and non-glutinous varieties	NA	Taira (1984)
	High amount of protein (11%) and fat (4%). The protein fractions are represented by albumins and globulins (13%), prolamins (39.4%), and glutelins (9.9%). It is thus recommended as an ideal food for diabetics.	NA	NA	Saleh et al. (2013)
Quinoa	Higher lysine and methionine content	NA	NA	Bhargava et al. (2003)
	Increased protein efficiency ratio (PER)	NA	Cooking	Mahoney et al. (1975)
	Increased *in vitro* digestibility	NA	Cooking, autoclaving, drum drying	Ruales and Nair (1993a)
	Changes in total dietary fiber content	NA	Thermal treatment	
	Decreased oil absorption capacity of quinoa flour	NA	Adding salt	Ogungbenle (2003)
	Rich source of antioxidants	NA	NA	Debski et al. (2013)
	Considered as golden grain because of its nutritional properties. Thus, NASA integrated this into the food of astronauts.	NA	NA	Rojas et al. (2010)
	Helps to reduce fatty acid uptake and esterification in adipocyte	NA	NA	Foucault et al. (2012)
	Significant impact on the chemical profile of quinoa flour	NA	Extrusion and roasting	Brady et al. (2007)
	Helps to degrade phytate in flour	Degradation of phytate in pseudo-cereal flours may depend on the activation of endogenous phytase and on the production of exogenous phytase by starter culture	Fermentation	Castro-Alba et al. (2019)
	Improved mineral availability of flours	Fermentation with *Lactobacillus plantarum*	Fermentation	
	Higher level of phytate degradation in quinoa grains	NA	Abrasion process to eliminate saponins	

(Continued)

TABLE 5 | Continued

Crop	Effects	Mechanism	Process to increase bioavailability	References
	Rich source of phytoecdysteroids	NA	NA	Kumpun et al. (2011)
	Anabolic, performance enhancing, anti-osteoporotic, wound-healing properties	Phytoecdysteroids	NA	Graf et al. (2014)
	Reduction in phytate content	NA	Germination, cooking, and fermentation	Valencia et al. (1999)
	Increased iron solubility	NA	Soaking and germination	
Amaranth	Reduces bioavailability of calcium and magnesium Presence of antinutritional factors	Oxalates	Cooking/popping	Arêas et al. (2016)
	Reduces bioavailability of carbohydrates	Inhibition of amylases contributing to the reduction of glucose levels in blood		
	Reduction in blood cholesterol level	Decrease the solubility of cholesterol micelles by Amaranth oil		
	High-protein amaranth flour (HPAF)	enzymatic hydrolysis	Liquefaction/ saccharification	Guzmán-Maldonado and Paredes-López (1998)
	Improves grain nutrient profile	NA	Malting/germination	Hejazi et al. (2016)
	Increases availability of proteins as well as free amino acid components	NA	Sprouting	Paredes-Lopez and Mora-Escobedo (1989)
	Reduction in antinutrient content, increases amino acids, carbohydrates, fibers, polyphenol content, and antioxidant potential	NA	Germination	Gamel et al. (2006)
	Best way to maintain (and even improve) amaranth nutritional values	NA	Germinated flour at 30°C during 78 h of germination	Perales-Sánchez et al. (2014)
	Quick digestion of starch content and increase in glycemic index	NA	Grinding/roasting	Capriles et al. (2008)
Buck wheat	Increases acceptability score of biscuits	Addition of buck wheat flour	NA	Baljeet et al. (2010)
	Rich source of nutraceutical compounds	NA	NA	Li and Zhang (2001)
	Higher lysine, iron, copper, and magnesium content	NA	NA	Ikeda and Yamashita (1994)
	Antioxidant potential	NA	NA	Oomah and Mazza (1996)
	Reduced starch digestibility, lowering of glycemic index, anticholesterolemic properties of protein fraction, well-balanced amino acid composition, and good source of dietary fiber and minerals,	NA	NA	Pomeranz and Robbins (1972); Kayashita et al. (1997); Skrabanja and Kreft (1998); Tomotake et al. (2000); Skrabanja et al. (2001); Steadman et al. (2001a; 2001b); (Ikeda et al. (2006);
	Reducing high blood pressure, lowering cholesterol, controlling blood sugar, and preventing cancer risk	NA	NA	Fabjan et al. (2003)
	Improved capillary fragility, retarded development of diabetes, anti-lipoperoxidant activities, anti-cancer activity, anti-hyperglycemic effect, protective effects against hemoglobin oxidation, a mitigation effect on cardiovascular diseases, anti-oxidative property, anti-mutagenic activity, anti-inflammatory activity, mitigation of diabetes, suppression of protein	NA	NA	Griffith et al. (1944); He et al. (1985); Odetti et al. (1990); Nègre-Salvayre et al. (1991); Deschner et al. (1991); Wang et al. (1992); Grinberg et al. (1994); Oomah and Mazza (1996); Aheme and O'Brien (1999);

(Continued)

TABLE 5 | Continued

Crop	Effects	Mechanism	Process to increase bioavailability	References
	glycation, anti-platelet formation property, anti-angiogenic effect, neuroprotective effect			Guardia et al. (2001); Je et al. (2002); Nagasawa et al. (2003); Sheu et al. (2004); Guruvayoorappan and Kuttan (2007); Pu et al. (2007)
	Thiamin-binding proteins (TBP) isolated from buckwheat	Serve as B_1 vitamin transporters in the plant and stabilize it during technological processing	NA	Mitsunaga et al. (1986)
	Improvement of true digestibility	NA	Hypothermal transformations	Christa and Soral-Śmietana (2008)
	Increased antioxidative potential	NA	Honey obtained from buckwheat flowers	Gheldof et al. (2003)
	Induced apoptosis in leukemia cells (0.5–100 μg/ml, *in vitro*), induced apoptosis in human solid tumor cells (6.25–50.00 μg/ml)	Buckwheat trypsin inhibitor	NA	Park and Ohba (2004); Wang et al. (2007)
	Coarse type of flour (mainly responsible for producing acceptable flavor) and a fine type of flour (responsible for binding particles to each other that are present in the buckwheat flour) are produced	Important for preparing buckwheat noodles with high palatability and acceptability rather than modern milling with a roll milling machine	Traditional stone milling	Ikeda and Ikeda (2016)
	Increased resistant starch contents	NA	Cooking	Kreft and Skrabanja (2002)
	Reduced glycemic index	Formation of amylase-resistant starch produced by heating	Cooking	Skrabanja et al. (2000)

markers. The map covered 721 centiMorgan (cM) on genome A and 787 cM on genome B and consisted of 18 linkage groups. Phylogenetic studies using 45 genomic SSRs on 79 finger millet accessions showed that finger millet was domesticated in Africa first and was then introduced to India (Dida et al., 2008). Apart from work by Dida et al. (2007), 49 new polymorphic genomic SSR markers were developed by Musia (2013) using next-generation sequencing (NGS) data. Gimode et al. (2016) sequenced two genotypes of finger millet (KNE755 and KNE796) using Roche 454 and Illumina technologies and identified 10,327 SSRs and 23,285 single nucleotide polymorphism (SNP) and tested 101 of each across a diverse set of wild and cultivated finger millet accessions. Several other research groups identified EST-SSRs using sequences deposited in the NCBI database (Arya et al., 2009; Reddy et al., 2012; Babu et al., 2014b). The mean PIC value for 49 polymorphic SSRs tested was 0.42, whereas the mean PIC value for 80 polymorphic SNPs was 0.29. Genetic diversity analysis using molecular markers has reported low polymorphism showing a narrow genetic pool of cultivated finger millet genotypes (Muza et al., 1995). Using 14 polymorphic genomic SSRs and three genic SSRs, Arya et al. (2013) showed that African accessions have higher genetic diversity than Indian finger millet accessions. Apart from SSR markers, Kumar et al. (2016) identified 23,000 SNPs using genotyping by sequencing of 113 finger millet genotypes.

Marker Trait Associations in Finger Millet

Finger millet has been reported to contain 5–30 times higher calcium than other cereals (National Research Council, 1996) and 44.7% of the essential amino acids (Mbithi-Mwikya et al., 2000). Genetic diversity analysis of 103 finger millet genotypes using 36 EST-SSRs associated with *opaque2* modifiers and 20 SSR primers associated with calcium transporters and calmodulin genes differentiated the finger millet genotypes based on protein and calcium content (Nirgude et al., 2014). Cereal endosperm proteins lack essential amino acids such as lysine and tryptophan and *opaque 2* modifier (a bZIP transcription factor) is involved in regulating the accumulation of lysine and tryptophan in seed. A set of 67 functional SSR markers was developed and genetic diversity analysis in a global finger millet genotype collection for *opaque2* modifier genes classified the genotypes into three clusters with high, medium, and low tryptophan content with few exceptions

(Babu et al., 2014c). Association mapping studies identified markers associated with various agronomic traits such as days to flowering, tiller number, plant height, blast resistance, finger number, etc. (Bharathi, 2011; Babu et al., 2014a; Babu et al., 2014b). Association mapping studies for nutritional quality traits identified two QTLs associated with tryptophan content and one QTL associated with protein content, and the marker associated with tryptophan content showed an inverse relationship with protein content in finger millet (Babu et al., 2014c). Further, nine markers were identified to be associated with calcium content (Kumar et al., 2015b). Apart from SSR and SNP markers, the orthologous genes for amino acid composition and calcium content in grains of finger millet were identified and the SSR variations within these genes among the accessions differing in protein and calcium content were used for developing gene-specific functional SSR markers (Reddy et al., 2011; Nirgude et al., 2014). Kumar et al. (2015a) carried out transcriptome analysis in developing spikes of finger millet and identified SSR motifs in the genes encoding calcium transporters and seed storage proteins. Despite the markers and genetic materials identified, progress in marker-assisted selection for genetic improvement of finger millet has lagged because of poor understanding about the complex traits to be transferred to the genotypes of interest.

Transcriptomes and Genomes of Finger Millet

A few transcriptomics studies have been carried out to have a better understanding about the complexity of traits in finger millet. Salinity-responsive transcriptome profiling using a next-generation sequencing platform (Ion Proton) in contrasting finger millet genotypes led to the identification of the genes/pathways involved in an improved salt tolerance mechanism (Rahman et al., 2014). To understand the underlying mechanism of calcium accumulation in grains, Kumar et al. (2015a) carried out transcriptome analysis in the developing spikes of two finger millet genotypes (GPHCPB 45, a high-calcium genotype, and GPHCPB 1, a low-calcium genotype) using Illumina Hiseq-2000. It has been hypothesized that the accumulation of calcium in different tissues and genotypes of finger millet varies due to the differential expression of the genes involved in uptake, translocation, and accumulation of calcium in different tissues. Mirza et al. (2014) carried out expression analysis of the genes involved in calcium translocation and storage in two contrasting finger millet genotypes for calcium content and observed a two-pore channel (TPC1) and $Ca^{(2+)}$ ATPase that might be involved in calcium uptake and translocation, respectively, due to their strong expression in root, stem, and developing spikes; whereas $Ca^{(2+)}/H^{(+)}$ antiporter (CAX1) might be involved in calcium accumulation in seeds due to its over-expression in developing spikes. The correlation between expression of these genes and calcium accumulation shows that these genes can be further used for a biofortification program (**Table 8**).

To unravel and understand the complex genome of finger millet, two independent research programs were started for carrying out whole-genome sequencing of finger millet. These attempts were carried out by the Indo-Swiss collaborative program funded by the Department of Biotechnology, Ministry of Science and Technology, India, and coordinated by the University of Agricultural Sciences, Bangalore, India, in partnership with Functional Genomics Center Zurich (FGCZ), University of Zurich. The researchers sequenced the genome and transcriptome of a drought-tolerant and blast-resistant finger millet genotype (ML-365) using Illumina and SOLiD sequencing technologies (Hittalmani et al., 2017). The sequenced genome consisted of 1,196 Mb covering ~82% of the total genome size. Genome analysis revealed the presence of 85,243 genes and 49.92% of the genome was found to be consisting of repetitive DNA. Hatakeyama et al. (2018) carried out whole-genome sequencing and assembly of finger millet (cv. PR-202) using Illumina NextSeq500 and PacBio RS II systems. The assembled genome was found to be of 1,189 Mb, estimated to be covering 78.2% of the finger millet genome. Genome analysis resulted in the identification of 62,348 genes, of which 91% were functionally annotated. The whole-genome information on ML-365 and PR-202 can be used for candidate genes and marker identification, which can be further used in marker-assisted breeding programs for the genetic improvement of finger millet.

Foxtail Millet

Foxtail millet (*S. italica* (L.) P. Beauv.), also called Italian, German, Hungarian, or Siberian millet, is a nutritional and natural staple used in many countries of East Asia for a long time. It has been cultivated in India, China, and many countries in Southeast Asia and Africa for many millennia and is quite popular in arid zones (Austin, 2006; Cheng and Dong, 2010; Mal et al., 2010). The grain of foxtail millet is quite rich in protein, fiber, and phosphorus compared with that of other minor millets (Muthamilarasan and Prasad, 2015). These grains are very rich in vitamin B_1, B_3, and B_5 (**Table 4**), and also contain a much higher amount than other minor millets and common cereals such as rice, wheat, and maize (Cheng and Dong, 2010). Foxtail millet also contains all the EAA but has isoleucine, leucine, methionine, phenylalanine, threonine, tryptophan, and valine in significantly higher amounts (**Table 3**). This makes foxtail millet a highly important nutritive crop with rich genetic diversity for glutinous and non-glutinous grains with different lipid composition (Taira and Miyahara, 1983). Significant phenotypic variations provide ample opportunity for allele mining and the use of molecular markers to supplement breeding programs. This crop is also one of the most important C4 panicoid crops known for its small genome size (~490 Mb), short life cycle, and self-pollinating nature, which make it an excellent model crop for evolutionary studies within the panicoid grass system (Lata and Shivhare, 2017).

Foxtail millet, vis-à-vis other gramineous crops, is a naturally drought-tolerant crop (Goron and Raizada, 2015). It is a quite resilient crop and is better adapted to marginal environments, especially arid regions. It is an extremely suitable food for type 2 diabetics due to its low glycemic index (GI) as its starch digestibility is much lower than that of wheat (**Table 1**).

TABLE 6 | Anti-nutrients and processes to decrease anti-nutritional activity in minor-millets and pseudo-cereals.

Crop	Anti-nutrients	Processes to decrease effects	References
Finger millet	Phytate content	Seed germination decreases phytic acid High temperature short time (HTST) process reduces anti nutrients	Nirmala et al. (2000); Kumar et al. (2018)
	Tannins	Germination (leaching and soaking) reduces tannins Boiling and pressure cooking can reduce tannins	Nirmala et al. (2000); Platel et al. (2010); Platel and Srinivasan (2016); Kumar et al. (2018)
	Polyphenols and phytates	Dehulling followed by soaking	Pawar and Parlikar (1990)
	Phytates, phenols, tannins	Decortications, milling, soaking, malting, germination, fermentation, popping and cooking	Shibairo et al. (2014)
	Polyphenols	Thermal/hydrothermal treatments, germination, decortication and fermentation	Subba Rao and Muralikrishna (2002)
	Phytates, tannins	Germination, fermentation	Hemalatha et al. (2007a); Platel and Srinivasan (2016)
	Phytate	Treatment with fungal phytase	Hemalatha et al. (2007b)
	Tannin, phytic acid, oxalic acid	Popping	Mishra et al. (2014)
	Phytates, polyphenols, tannins	Soaking, germination, steaming, fermentation	Sripriya et al. (1997)
	Phytates, polyphenols, tannins	Malting Decortication	Platel et al. (2010) Krishnan et al. (2012)
	Phytates	Germination	Tatala et al. (2007)
	Phenolic acids	Sprouting, pressure cooking, open pan-boiling, microwave heating	Hithamani and Srinivasan (2014)
Foxtail millet	Phytate, tannin, polyphenols	Dehulling, soaking, cooking	Pawar and Machewad (2006)
	Phytate	Thermal processing, mechanical processing (decortication, milling and sieving), soaking, fermentation, germination, malting,	Hotz and Gibson (2007)
	Phytate	Germination, soaking, puffing, fermentation, enzymatic hydrolysis	Saleh et al. (2013)
	Polyphenols	Germination and steaming	Mounika et al. (2017)
	Phenolic compounds	Fermentation, malting and steeping, thermal processing	Kaur et al. (2019)
	Phytic acid	Roasting	Gahlawat and Sehgal (1994)
Quinoa	Saponins	Repeated washing or dehulling after harvest. Extrusion and roasting techniques. Wet technique: Strong washing in cold alkaline water. Dry technique: heat treatment, extrusion, roasting, mechanic abrasion Wet technique is recommended as dry technique by abrasive peeling can allow the loss of protein, vitamins and minerals	Koziol (1990); Mastebroek et al. (2000); Dini et al. (2001); Brady et al. (2007); Comai et al. (2007); Spehar (2007); Farro (2008); Paśko et al. (2009); Jancurová et al. (2009); Ruiz et al. (2017); Maradini et al. (2017)
	Phytate compound in grains	Soaking, germination, fermentation and cooking decrease phytate compound	Ruales and Nair, 1993a; Hurrell, 2004; Umeta et al., 2005; Repo-Carrasco-Valencia et al., 2010; Lazarte et al., 2015; Iglesias-Puig et al., 2015
	Phytic acid	Similar process as for reducing saponin as brushing and rinsing can reduce around 30%. Fermentation and cooking can degrade the activation of phytase	Ruales and Nair (1993b); Oliveira et al. (2003); Khattab and Arntfield (2009)
	Tannins	Cleaning and rinsing in water Adequate washing (to cook) can reduce the harmful effect	Chauhan et al. (1992); Jancurová et al. (2009); Borges et al. (2010)
	Trypsin inhibitor	Heat treatment, boiling, roasting, domestic techniques employed for food preparation can reduce trypsin inhibitor concentration	Ruales and Nair (1993b); Jancurová et al. (2009); Borges et al. (2010)
	Nitrates	Found in leaves mainly. If the consumption in amount is higher can be harmful, but nitrate content is lower	Lopes et al. (2009)
	Oxalates	Contained in leaves, stem roots and hypocotyl seeds Quinoa has lower levels of Oxalic acid	Lopes et al. (2009)
Amaranth	Phytate, Phenolic compounds, Trypsin inhibitors, Chymotrypsin inhibitors, Amylase inhibitor	Seeds cooked, popped/extrusion, germinated seeds at 30, 60 and 90°C	Gamel et al. (2006); Ferreira and Arêas (2010)
Buckwheat	Trypsin activity inhibitor, alpha-amylase activity inhibitor, Poly-phenol content, Phytic acid content	Steaming, baking, boiling treatment in seeds and seedlings sprouting after 24, 48, and 72 h	Kreft (1983); Ikeda et al. (1986); Ikeda et al. (1991); Ookubo (1992); Campbell (1997); Amarowicz et al. (2008); Zhang et al. (2015)
	Tannins	Germination and sprouting can reduce protease inhibitors and increase protein digestibility. Further, dehulling grains and roasting	Kreft (1983); Ikeda et al. (1986); Ikeda et al. (1991); Ookubo (1992); Campbell (1997); Handoyo et al. (2006)

TABLE 7 | Details of genome organization and other characteristics of underused crops.

Species	Ploidy level	Chr. no.	Approx. genome size	Genes annotated
Eleusine coracana	Allotetraploid	36	1.6 Gb	85,243
Setaria italica	Diploid	18	513 Mb	38,801
Chenopodium quinoa	Allotetraploid	36	1.5 Gb	62,512
Amaranthus sp.	Diploid	32, 34	466 Mb	23,059
Fagopyrum esculentum	Diploid	16	540 Mb	35,816
Fagopyrum tataricum	Diploid	16	540 Mb	33,366

However, its nutrient bioavailability could be further increased if it were processed by boiling or steaming (Pawar and Machewad, 2006; Ren et al., 2016) (**Tables 5** and **6**). Genotypic differences exist for antioxidant activities and some specific varieties such as SiA-2593 were identified as therapeutic and functional foods (Shejawale et al., 2016).

Foxtail millet starch is of great interest among entrepreneurs owing to its flexibility to make gels using foxtail millet flour and divalent cations such as $CaCl_2$ and $FeSO_4$ (Nagaprabha and Bhattacharya, 2016). *S. italica* is considered one of the best minor millet crops for anemic and diabetic people, thus reinforcing nutrition security besides food security (Bandyopadhyay et al., 2017). Processed protein from *S. italica* has been reported as a potential source of food additive (Mohamed et al., 2009). Ample diversity is available in India, China, France, Japan, Kenya, and Mexico gene banks for foxtail millet (Dwivedi et al., 2012), which can be exploited for various nutritional traits, including protein, as the protein content is higher than that of other selected small millets (Rao et al., 2011). Besides this, its rich composition of beta-glucans (42.6%) present in the fiber helps to enhance sugar and cholesterol metabolism, which ultimately prevents diabetes and cardiovascular diseases (Kumari and Thayumanavan, 1997; Itagi et al., 2012; Muthamilarasan and Prasad, 2015). Efforts on genomics are in progress to identify the underlying factors for the mechanisms that improve nutritional factors in foxtail millet (Zhang et al., 2012; Muthamilarasan and Prasad, 2015). Traditional simple as well as advanced food processing helps to improve the bioavailability of micronutrients (iron, zinc, and proteins) and renders them in a form that is easy to assimilate by the body, along with a significant decrease in anti-nutrients (polyphenols and phytate) (Pawar and Machewad, 2006; Saleh et al., 2013). Anti-nutrients exist but grain processing reduces them drastically; however, omics can contribute a lot to improving the bioavailability of micronutrients and reducing anti-nutrients such as phytic acid, polyphenols, and tannins, to decrease processing cost and time.

Genomics of Foxtail Millet

Foxtail millet (*S. italica* (L.) P. Beauv.) is one of the oldest domesticated cereals in the Old World. The genus *Setaria* belongs to the subtribe Cenchrinae and tribe Paniceae within the subfamily Panicoideae (Kellogg, 2015). *Setaria* is the largest genus in the Cenchrinae, consisting of ≈100 species and all possessing the C4 photosynthetic pathway. The small diploid genome (2n = 2x = 18; 513 Mb), self-pollination behavior, and short life cycle (6 weeks) have made *Setaria* an ideal model plant for functional genomics studies in millets as well as in cereals (**Table 7**).

TABLE 8 | Details of important QTLs/genes linked with accumulation of nutritional/anti-nutritional factors in underused crops.

Crop	QTL/gene/enzymes	Trait	Reference
Finger millet	Two-pore channel (*TPC1*), Ca^{2+} ATPase, Ca^{2+}/H^{+} antiporter (*CAX1*)	Calcium content	Mirza et al. (2014); Kumar et al. (2015a)
	27-kDa c-zein gene of opaque 2 modifier	Tryptophan content	Babu et al. (2014b)
	bZIP transcription factor RISBZ1	Protein content	Babu et al. (2014b)
Foxtail millet	Granule-bound starch synthase 1	Amylose content	Fukunaga et al. (2002); Bai et al. (2013)
Quinoa	Diaminopimelate aminotransferase, diaminopimelate epimerase	Lysine content	Zou et al. (2017)
	Pyridoxal 5′-phosphate synthase, dihydrofolate synthase, tetrahydrofolate synthase	Vitamin B_6 and vitamin E content	Zou et al. (2017)
	Triterpene saponin biosynthesis activating regulator (*TSAR1* and *TSAR2*)	Saponin content	Jarvis et al. (2017)
	Granule-bound starch synthase I	Amylose content	Brown et al. (2015)
Amaranth	Cytochrome P450, 4,5-DOPA dioxygenase extradiol 1	Betalain content	Lightfoot et al. (2017)
	Aspartate kinase 1 and dihydrodipicolinate synthase	Lysine content	Clouse et al. (2016)
Buckwheat	Granule-bound starch synthase 1	Amylose content	Yasui et al. (2016)

Molecular Markers, Genetic Diversity, and Phylogenetic Studies in Foxtail Millet

Among millets, *Setaria* is the most deeply studied genus at both the genetic and molecular level. Various types of molecular markers used for genetic diversity and phylogenetic studies in *S. italica*, including RFLP (Wang et al., 1998; Fukunaga et al., 2002), random amplified polymorphic DNA (RAPD) (Schontz and Rether, 1999), AFLP (d'Ennequin et al., 2000), and SSRs (Lin, 2012; Trivedi et al., 2018), showed that foxtail millet genotypes differed genetically between different regions and Chinese landraces were found to be the most variable compared with landraces from other places. Wang et al. (1998) first reported the RFLP-based map consisting of 160 loci using an intervarietal cross of foxtail millet (Longgu 25 × Pagoda Flower Green), which was later used by Devos et al. (1998) to construct a comparative genetic map of foxtail millet and rice. Seeing the importance of SSR markers, Jia et al. (2007) developed 26 EST-SSRs. Later, Jia et al. (2009), using two genomic libraries enriched for (GA)n and (CA)n, identified 100 polymorphic SSR markers and developed a linkage map using 81 SSRs and 20 RFLP-anchored markers. Later, Lin et al. (2011) developed 45

polymorphic SSR markers from a RAPD-enriched library and used them for genetic diversity analysis as well as proving their cross-species transferability. Gupta et al. (2011; 2012) developed 98 intron-length polymorphic (ILP) and 147 genomic SSR markers and further showed the high-level cross-species transferability. Considering the importance and ease of microsatellite markers in MAS because of their high reproducibility, co-dominant nature, multiallelic variation, and abundance in genome work, *Setaria* was used to mine the genome wide SSRs by analyzing the genome sequence information. The genome-wide analysis of foxtail millet resulted in the identification of 28,342 microsatellite repeat-motifs spanning 405.3 Mb of the genome. Among the identified microsatellites, primers for 21,294 were designed and 15,573 markers were mapped on nine chromosomes to develop a high-density physical map of foxtail millet (Pandey et al., 2013). The validation of 159 developed markers in eight accessions of *Setaria* sp. showed 67% polymorphism and 89.3% cross-genera transferability across millet and non-millet species. Muthamilarasan et al. (2013) developed 5,123 ILP markers and proved their applicability in germplasm characterization, phylogenetic studies, transferability across species, and comparative mapping in millets and bioenergy grass species.

Genomes and Transcriptomes of Setaria

A milestone in the area of *Setaria* genomics was the release of the reference genome of foxtail millet cultivar "Zhang gu" using whole-genome shotgun sequencing combined with the Illumina second-generation sequencer covering ≈86% of the genome (Zhang et al., 2012). The sequencing resulted in the generation of 16,903 contigs and 439 scaffolds covering a total length of 423 Mb. Further sequence analysis identified 38,801 annotated genes. Apart from "Zhang gu," a photo-thermo-sensitive male sterile line (A2) was sequenced and comparison of both genomes resulted in the identification of many SNPs (542,322), small insertions/deletions (33,587), and structural variants (10,839) between the two cultivars. A linkage map was constructed using an F_2 population derived from "Zhang gu" and "A2" using 759 markers consisting of 118 SNPs and 641 structural variants (Zhang et al., 2012). *S. italica* accession "Yugu1" and *Setaria viridis* accession "A10" were sequenced using the ABI3730xl capillary sequencer (Bennetzen et al., 2012). The assembled genome was found to be 396.7 Mb covering 80% of the genome and genome analysis identified 24,000–29,000 expressed genes. To date, foxtail millet is the only millet whose genome assembly is available to a chromosomal scale.

Because of the abiotic stress-tolerant nature of foxtail millet, several functional genomics tools have been applied to dissect its stress-tolerant nature (Zhang et al., 2007; Jayaraman et al., 2008; Lata et al., 2010; Lata and Prasad, 2011; Lata et al., 2011; Puranik et al., 2011; Puranik et al., 2013; Qi et al., 2013; Tang et al., 2017). Besides understanding the molecular basis of abiotic stress tolerance, GWAS studies have been carried out for mapping the QTLs underlying various agronomic traits in foxtail millet (Jia et al., 2013; Gupta et al., 2014; Jaiswal et al., 2019).

However, despite the enormous health benefits, few proper attempts have been made to understand the genetics and genomics of nutritional traits in foxtail millet. Resequencing of waxy landrace Shi-Li-Xiang (SLX) and fine mapping using an F_2 population derived from SLX (waxy) × Yugu1 (non-waxy) identified a waxy locus harboring starch synthase-encoding *GBSS 1* gene. Sequence analysis of GBSS 1 showed transposable elements confirming its waxy nature (Bai et al., 2013). To dissect the genetics and genomics of nutritional traits, genome and transcriptome data can be used to select genes associated with various pathways involved in biosynthesis and regulation of storage compounds and these can be exploited for understanding their role in the accumulation of various nutritional compounds.

Quinoa

Quinoa (*C. quinoa* Willd.), pronounced "keenwa," is a dicotyledonous plant originated from South America, hence called an Andean grain. For seven millennia, Andean cultures from pre-Columbian time have been eating it and it has been part of their diet. It is actually a pseudo-cereal due to its morphological grain shape like grass crops. It is classified into five ecotypes, based on geographic adaptation in the center of diversity (Hinojosa et al., 2018):

- Valley = grown at 2,000 to 3,500 m.a.s.l. in Colombia, Ecuador, Peru, and Bolivia;
- Altiplano = grown at high altitudes of more than 3,500 m.a.s.l. around Titicaca Lake on the border of Bolivia and Peru;
- Salares = grown in the salt flats of Bolivia and Chile and having a high tolerance of salinity;
- Sea-level = grown in the low-altitude areas of southern and central Chile;
- Subtropical or yungas = grown in the low-altitude humid valleys of Bolivia and including late-flowering genotypes.

Quinoa is a highly resilient crop that not only has superior nutritional quality vis-à-vis common cereals but can also withstand environments where other crops have difficulty to grow (Choukr-Allah et al., 2016; Nanduri et al., 2019). Since 2013, when UN/FAO designated 2013 as the International Year of Quinoa, interest in this crop has increased markedly worldwide due to awareness, mainly because of its balanced nutritional profile and its potential as an alternative to feed the growing world population in a sustainable manner, especially when more food has to come from marginal environments (Jacobsen et al., 2013; Zurita-Silva et al., 2014). The natural selection process of quinoa cultivars took place under severe adverse conditions of the Andes, such as limited rainfall and extreme aridity (Martínez et al., 2009), and in salt-affected soils (Ruiz-Carrasco et al., 2011). That explains quinoa's built-in abiotic stress tolerance of aridity, salinity, highland, and frost; hence, it is well suited to marginal environments. Although quinoa is drought and salinity tolerant, it is sensitive to high-temperature stress. It can tolerate a wide range of temperatures (from −8°C to 35°C), but high temperature above 35°C during flowering results in a significant reduction in seed set and ultimately yield (Jacobsen et al., 2005). For example, studies in Italy (Pulvento et al., 2010), Morocco (Hirich et al., 2014),

Germany (Präger et al., 2018), Portugal (Pires, 2017), India (Bhargava et al., 2006a), Egypt (Eisa et al., 2017), Mauritania (Bazile et al., 2016), and the United States (Peterson and Murphy, 2015; Walters et al., 2016) have reported that high temperatures reduce quinoa seed yield.

Quinoa has a unique balance between oil (4–9%), protein (averaging 16%, with high nutritional relevance due to the ideal balance of its essential amino acid content), and carbohydrate (64%) (Schlick and Bubenheim, 1996; Bhargava et al., 2006a; Vega-Gálvez et al., 2010) (**Table 1**). Lysine, one of the essential amino acids, which is usually much less in plant-based diets, is relatively high in quinoa, indeed very close to the standard set by FAO for human nutrition (**Table 3**). Because of its high starch content (51–61%), it can be used in the same way as cereals for flour production (Mastebroek et al., 2000; Ogungbenle, 2003; Repo-Carrasco et al., 2003; Bhargava et al., 2006a; Stikic et al., 2012). In addition, quinoa is a good source of vitamins, oil (high in omega 3, linoleic and linolenic acids, 55–66% of the lipid fraction), and natural antioxidants such as α and γ tocopherol, and it has more minerals such as Ca, Fe, K, Mg, Cu, and Mn than other cereals (**Table 2**) (Repo-Carrasco et al., 2003; Vega-Gálvez et al., 2010; Fuentes and Bhargava, 2011; Stikic et al., 2012). The International Center for Biosaline Agriculture (ICBA), based in Dubai, has been working on the suitability of quinoa for marginal environments since 2006 and has found high Fe content vis-à-vis major cereals such as wheat, rice, and maize. Five improved quinoa genotypes from ICBA (Q1 to Q5) have been analyzed for Fe content and it ranged from 49.55 to 133 ppm depending upon growing conditions, showing high genotype-by-environment-by-management (G × E × M) interaction. Quinoa is highly suitable for eating by diabetics due to its low glycemic index (GI) (**Table 1**). Quinoa seeds release important compounds such as phytoecdysteroids and 20-hydroxyecdysone (20HE) while germinating. These released bioactive phytochemicals from the seeds can be an excellent staple for developing anti-diabetic food products as these bioactive phytochemicals can decrease the glucose level in the blood and are a potential source to treat obesity and hyperglycemia (Graf et al., 2014). Quinoa grains are also gluten-free and are considered as complete protein as they contain all nine essential amino acids that the human body cannot produce itself and they have very high lysine content overall vis-à-vis other cereals (Jacobsen, 2003; Albugoch and Lilian, 2009; Alvarez-Jubete et al., 2010). Quinoa's exceptional nutritional qualities led NASA to include it as part of its astronauts' diet on long space missions. A NASA technical paper mentions that while no single food can supply all the essential life-sustaining nutrients, quinoa comes as close as any other in the plant or animal kingdom (Schlick and Bubenheim, 1996).

Despite these nutritional qualities, some antinutritional factors (triterpenoid glycoside) are present in quinoa. Saponins, when present in the seeds, confer bitterness. Natural occurrence of saponins in quinoa grain is usually higher but some native varieties have low saponin as well. Even though saponins can be removed by repeated washing or dehulling, this consumes additional resources on postharvest processing (**Table 6**). Enough genetic variation has been reported in saponin content in quinoa, varying from 0.2 g/kg in sweet genotypes to 11.3 g/kg in bitter genotypes based on dry matter (Mastebroek et al., 2000). Saponin is present in the seed coat and washed saponin solution from seeds could be used as a by-product in biopesticide and therapeutic compounds (Ruiz et al., 2017). Reducing saponin content could broaden quinoa production globally in a more economically sustainable manner. It is reportedly controlled by a recessive gene (*TSARL1*) and genotypes with low saponin could be developed using conventional breeding techniques with the help of MAS (Jarvis et al., 2017).

Additional mineral inhibitor can influence the bioavailability and bio-accessibility of minerals of quinoa. Plant-based diets usually have a low bioavailability of minerals (mainly zinc, iron, and calcium) due to the presence of inhibitors such as phytates and tannins that reduce absorption. Phytate (myo-inositol-6-phosphate) is the main inhibitor of zinc, iron, and calcium. Degradation of phytate is important to allow the bio-assimilation of minerals. Bioavailability of iron is reduced if the phytate/Fe molar ratio is higher than 1, and, for bioavailability of zinc, this can affect the relation of the phytate/Zn molar ratio when it is higher than 5 (Iglesias-Puig et al., 2015).

Quinoa is a rich source of minerals and it has much higher zinc (2.73–5.01 mg/100 g), iron (4.82–7.19 mg/100 g), and calcium (77.10–211.90 mg/100 g) than other cereals based on results from six varieties of quinoa from Chile (Miranda et al., 2012). Unfortunately, the levels of phytates were quite high in quinoa vis-à-vis other cereals. It is better if the phyate:zinc molar ratio (Phy: Zn) is <15 (Bindra et al., 1986), phytate:iron (Phy : Fe) ratio <1 (Hurrell, 2004), and phytate:calcium (Phy : Ca) ratio <0.17 (Umeta et al., 2005). Ratios above the desirable values indicate low bioavailability of the mineral in the grain. Therefore, phytate can affect the bio-assimilation of important minerals in food if the molar ratios are high above the critical values. However, germination, soaking, cooking, and fermentation decrease the phytate compound in grains of quinoa and allow the bio-assimilation of iron, although much washing to some extent reduces the vitamin and mineral content as well (Ruales and Nair, 1993a; Valencia et al., 1999; Lazarte et al., 2015) (**Tables 5** and **6**). Cooking does not affect the amount of soluble iron; rather, it increases 2–4 times after soaking and germination, 3–5 times after fermentation, and 5–8 times after fermentation of the germinated flour combined with reduced phytates (Valencia et al., 1999). Soluble iron thus could be available to anemic populations in processed products based on germinated quinoa and quinoa sprouts (Vega-Gálvez et al., 2010). Brady et al. (2007) compared quinoa flour when processed by steam pre-conditioning, extrusion, and roasting. Steam pre-conditioning had the least effect on the chemical profile of quinoa while extrusion and roasting changed the chemical profile a lot compared to raw quinoa. The extrusion and roasting techniques can reduce saponin and the bitter taste

(**Table 6**). Bioactive polyphenols and flavonols in quinoa grains and sprouts can help to prevent oxidative stress (Paśko et al., 2009).

Genomics of Quinoa

C. quinoa is an annual pseudo-cereal. It is an allotetraploid with $2n = 4x = 36$ chromosomes having an estimated haploid genome size (C-value) of 1.005–1.596 pg (Bennett and Smith, 1991; Stevens et al., 2006; Bhargava et al., 2007a; Palomino et al., 2008; Kolano et al., 2012) (**Table 7**). Quinoa has mostly smaller metacentric chromosomes of 0.94–1.60 μm (Bhargava et al., 2006b; Palomino et al., 2008). Several studies on molecular understanding and genomics of quinoa have been carried out considering its nutritional importance. Most of the genetics and genomics studies have focused on either understanding its genetic diversity or evolutionary history whereas limited attempts have been made for genetic improvement of quinoa through molecular approaches. Most of the breeding has been carried out through mass selection for selecting high-yielding, early maturing quinoa varieties with low saponin content as well as tolerance of biotic and abiotic stresses.

Molecular Markers, Genetic Diversity, and Phylogenetic Studies in Quinoa

The genetic diversity and phylogenetic relationship studies have been carried out between cultivated species of quinoa and their wild relatives using various types of phenotypic, biochemical, and molecular markers. Wilson (1988a) used morphological and biochemical markers to study the genetic relationships between quinoa ecotypes and classified them into two broad groups: coastal types and Andean types (above 1,800 m.a.s.l.). The Andean ecotypes were further classified into northern and southern Andean quinoa. Further, phylogenetic study showed that the Altiplano was the center of origin of quinoa (Wilson, 1988b). Rojas et al. (2000) classified the 1,512 accessions of the Bolivian National Quinoa Collection into seven distinct groups using various morphological and agronomic traits. Bhargava et al. (2007b) studied genetic diversity in quinoa using morphological and quality traits and showed a high level of genetic variability among the accessions.

Random amplified polymorphic DNA (RAPD) markers were the first markers used to detect DNA polymorphisms among different quinoa accessions (Fairbanks et al., 1993; Ruas et al., 1999; del Castillo et al., 2007). Using RAPD markers, very low intraspecific variations were observed within *C. quinoa* (Ruas et al., 1999). del Castillo et al. (2007) studied the hierarchical structure of the genetic variation present in eight quinoa field populations from Bolivia and found that population structure was related to three major biogeographic zones: the northern and central Altiplano, the inter-Andean valley, and the southern Salar. Apart from RAPD, AFLP markers have been used to study genetic diversity in quinoa. Rodríguez and Isla (2009) used AFLP markers along with 20 phenotypic markers to characterize 14 accessions of quinoa and concluded that Chilean lowland germplasm might be genetically more diverse, and the germplasm clustered together into highland and lowland/coastal as earlier reported by Ruas et al. (1999). Using SSR-enriched library sequencing, 208 polymorphic SSR markers were identified for quinoa (Mason et al., 2005). Recently, because of their reproducibility and co-dominant nature, SSRs have been used to study genetic diversity in quinoa (Mason et al., 2005; Jarvis et al., 2008). Christensen et al. (2007) used 152 accessions of USDA and CIP-FAO quinoa collections for genetic diversity study using 35 SSR markers and found that the accessions from lowlands and highlands clustered together and identified the group of accessions that appears to be hybrid between lowland and highland. Later, the use of multiplex fluorescent SSR markers to understand the genetic diversity and phylogenetics of 59 quinoa accessions resulted in the separation of highland and lowland genotypes into two separate clusters (Fuentes et al., 2009). Recently, Zhang, T., et al (2017) carried out whole-genome resequencing of 11 quinoa accessions and identified various SSR, SNP, and Insertion/Deletion (InDel) markers. They further used the identified SSR and InDel markers to assess the genetic diversity of 129 quinoa accessions from the USDA collection. These studies using various types of markers showed that a strong population structure and huge genetic diversity exist among quinoa germplasm.

Linkage Maps of Quinoa

Seeing the importance of quinoa in food and nutritional security, several breeding programs for improving grain yield, earliness, disease resistance, and drought tolerance and reducing saponin content have begun. Molecular markers and linkage maps help in QTL mapping, which further helps in marker-assisted selection (MAS) for speeding up the breeding process. The first linkage map of quinoa was developed by Maughan et al. (2004) using 19 SSR, 6 RAPD, and 230 AFLP markers spanning 1,020 cM covering 60% of the genome, consisting of 35 linkage groups with an average marker density of 4.0 cM per marker. Later, Jarvis et al. (2008) developed 216 new SSR markers and a more enriched linkage map for quinoa by using new SSR and 75 AFLP markers consisting of 41 linkage groups covering 913 cM. Further, using two RIL populations, Maughan et al. (2012) mapped 511 SNPs across 29 linkage groups of quinoa spanning 1,404 cM with a marker density of 3.1 cM per marker. Recently, Jarvis et al. (2017) developed a high-density linkage map of quinoa consisting of 6,403 SNP markers through genotyping by sequencing (GBS) covering 2,034 cM on 18 linkage groups. Further, apart from molecular markers, several other genomic resources such as bacterial artificial chromosome (BAC) libraries and stress-responsive or tissue-specific transcriptome sequencing have been used, which has helped in gene discovery as well as identifying molecular markers. The first EST library of quinoa consisting of 424 ESTs was developed from seed and floral tissues (Coles et al., 2005). Stevens et al. (2006) developed a BAC library of quinoa using *BamHI* and *EcoRI* restriction enzymes consisting of 26,880 and 48,000 clones, respectively. Using the same BAC library, Balzotti et al. (2008) identified and characterized 11S globulin and 2S albumin seed storage proteins of quinoa, which they predicted were responsible for the relatively high protein content and ideal balance of amino acids in quinoa. Maughan et al. (2009) isolated and characterized the *Salt Overly Sensitive 1 (SOS1)*

gene using a BAC library and reported that *SOS1* is constitutively expressed in quinoa, unlike in other cereals in which mainly it is either stress-inducible or shows tissue-specific expression. Later, Walsh et al. (2015) carried out phylogenetic analysis in quinoa based on the sequence information of two introns of *SOS1* and identified two distinct polyploid lineages.

Transcriptomes and Genomes of Quinoa

The advances in next-generation sequencing technology and computational bioinformatics have accelerated genomics and transcriptomics research in quinoa. In the past few years, transcriptome studies have been carried out in quinoa to understand the molecular basis of drought and salinity tolerance. A drought-responsive transcriptome analysis carried out in two genotypes of quinoa using the Illumina HiSeq 2000 platform led to the identification of the genes involved in imparting drought tolerance to quinoa (Raney, 2012). Morales et al. (2017) performed transcriptome analysis in drought-tolerant Chilean quinoa genotype R49 and identified the drought-induced genes and pathways involved in providing drought stress tolerance in quinoa.

Using next-generation sequencing platforms, three research groups have independently completed quinoa genome sequencing (Yasui et al., 2016; Jarvis et al., 2017; Zou et al., 2017). Using two next-generation sequencing platforms, Illumina HiSeq 2500 (for short high-quality reads) and PacBio RSII (for longer reads and gap filling), Yasui et al. (2016) sequenced and assembled the draft genome of a quinoa inbred (Kd). The draft genome was found to be of 1.1 Gb size consisting of 24,847 scaffolds. The annotated genome of "Kd" was found to consist of 62,512 protein-coding genes around 535.5 Mb, leaving 49.2% of the genome as repetitive sequences. Further, a freely accessible Quinoa Genome DataBase (QGDB; http://quinoa.kazusa.or.jp/) was developed.

Zou et al. (2017) developed another draft genome sequence of an inbred line of quinoa (Real) using Illumina HiSeq 2500 and PacBio RSII platforms. The estimated genome size was 1.49 Gb covering 90.2% of the nuclear genome. Annotation of the draft genome sequence resulted in the identification of 54,438 protein-coding genes, of which 95.3% were functionally annotated. Approximately 65.5% of Real's genome consists of repeat sequences, 85.6% of which were found to be transposable elements comprising mostly retrotransposons. Further, to investigate protein quality, researchers analyzed comparative lysine, phenylalanine, and isoleucine content in three protein families [albumin, globulin, and late embryo abundant (LEA) proteins] and found that the lysine content of quinoa was significantly higher in all three protein families than in the other cereals such as wheat, rice, or maize. The high lysine content was not only at the free amino acid level but also for amino acid usage in seed protein sequences due to the presence of a high copy number of genes encoding the enzymes involved in converting aspartate into lysine. They also found that the high vitamin B and vitamin E content in quinoa is due to the presence of a high copy number of gene-encoding enzymes (*pyridoxal 5′-phosphate synthase, dihydrofolate synthase, tetrahydrofolate synthase*) involved in vitamin B_6 and dihydrofolate biosynthesis (**Table 8**). Further, Zou et al. (2017) reported that the genes involved in ion sequestration, ABA homeostasis, and signaling are responsible for enhancing abiotic stress tolerance in quinoa. Based on transcriptome analysis, Zou et al. (2017) proposed a model for salt accumulation in salt bladders. Since few genes were found to be differentially expressed in epidermal salt bladders, this suggests that most of the transporter genes are constitutively active in salt sequestration in the bladders and that regulation of ion transport in bladder cells in response to salinity occurs at the protein level through protein phosphorylation (Zou et al., 2017).

Jarvis et al. (2017) published a more complete genome sequence of coastal Chilean quinoa accession PI 614886 (QQ74) using PacBio RSII and Illumina HiSeq 2500 sequencing platforms. The genome was assembled into 3,486 scaffolds spanning 1.39 Gb consisting of 44,776 genes and approximately 64% repetitive sequences. Along with quinoa accession PI 614886 (QQ74), the authors resequenced 15 other quinoa accessions along with five accessions of *C. berlandieri* and two of *C. hircinum*, which are supposed to be immediate tetraploid ancestors of quinoa. They further produced 18 pseudo-chromosomes of quinoa consisting of 1.18 Gb (i.e. ~80% of the predicted ~1.45 Gb haploid genome size of quinoa). Jarvis et al. (2017) identified two genes encoding basic Helix-Loop-Helix (bHLH) transcription factors involved in regulating the triterpenoid biosynthetic pathway associated with production in quinoa. They identified the *triterpene saponin biosynthesis activating regulator like (TSARL1* and *TSARL2)* genes, of which *TASRL2* was expressed in roots but not in flowers or immature seeds, whereas *TSARL1* was over-expressed exclusively in seeds of bitter lines vis-à-vis sweet lines, suggesting that *TSARL1* might be the functional TSAR ortholog involved in regulating biosynthesis of saponin in bitter quinoa genotypes (**Table 8**). Further, sequence analysis showed that alternative splicing on *TSARL1* results in a premature stop codon, thereby translating a truncated protein with a compromised functional ability in forming homodimer to bind DNA for its activity.

Amaranth

Amaranth (*Amaranthus* sp.) is an ancient crop whose domestication and cultivation date back to around 8,000 years ago in Mayan civilization of South and Central America. There is no concrete evidence for the origin of *Amaranthus*. It was used as a staple food along with corn (maize) and beans in Mexico starting 1,400 years ago; however, its production declined after the collapse of Central American culture (Alvarez-Jubete et al., 2010). *Amaranthus* sp. is a highly nutritive pseudo-cereal, rich in proteins, vitamins, and minerals. Because of its nutraceutical value and climate resilience, amaranth has been relaunched and is being promoted as a suitable crop for food and nutritional security (Lakshmi and Vimala, 2000). Amaranth's leaves are consumed as vegetables and its grains as cereal. The amaranth family is divided into two sections: Amaranthus saucer and Blitopsis dumort (Allen, 1961). Based on its use, it has been grouped into grain amaranth, vegetable amaranth, and ornamental and weedy amaranth (Sauer, 1967). Grain amaranth has four species, *A. hypochondriacus, A. cruentus, A caudatus,* and *A. edulis* (Martínez-Cruz et al., 2014), whereas vegetable amaranth

belongs to the section Blitopsis and has two major species, *A. tricolor* and *A. lividis* (Pal and Khoshoo, 1972; Madhusoodanan and Pal, 1981). As evident from recent past publications, amaranth has been attracting the attention of food technologists for exploring its functional aspects after the United States National Academy of Sciences showed its high nutritional value and agronomic potential (Monteros et al., 1994; Ulbricht et al., 2009).

Amaranth is a climate-resilient, fast-growing cereal-like plant. Despite its potential nutritional value, it has been underexploited. Protein content in the grain of amaranth species ranges from 13.1% to 21.0% with an average of 15%, which is comparatively higher than that of other cereals (**Table 1**) (Mlakar et al., 2009). Its amino acid profile makes it an attractive protein source as it contains significantly higher lysine content (4.9–6.1%), which is an essential amino acid limiting in most cereals (**Table 3**). Amaranth protein is also richer in sulfur-containing amino acids (≈4.4%), which are limiting in pulse crops (Bressani, 1989). It is considered that if it is consumed along with other cereals, it will provide a "balanced" source of protein (Saunders and Becker, 1984). The balanced amino acid content in amaranth grain protein is close to the optimum protein reference pattern in the human diet as recommended by FAO/WHO (O' Brien and Price, 2008). The balanced nature of amino acid composition in amaranth grain protein is due to the presence of ≈65% of the protein in the embryo and only ≈35% in the perisperm, whereas in other grains ≈85% of the protein is present in endosperm, which is poorer in essential amino acids (Senft, 1979; Betschart et al., 1981). Mainly three amino acids (leucine, isoleucine, and valine) are limiting amino acid in amaranth grain but are present in excess in other grains. Amaranth flour has high glycemic index, therefore it is suitable for blending with other cereals (**Table 1**). Amaranth and wheat flour in 25:75 ratio bring down the GI quite low with good nutritional balance. Similarly amaranth and maize flour in a ratio of 1:1 nearly reaches the perfect score of 100 on the nutritionist's scale and also the combination of amaranth in wheat flour improves the nutritional value of baked products (Saunders and Becker, 1984; Bressani, 1989). Apart from the balanced profile of amino acid content, amaranth protein has high digestibility (≈90%), which further improves the bioavailability of amino acids when ingested. Apart from having good amino acid composition, amaranth protein is gluten-free, making it a choice of food for patients with coeliac disease incidence.

Besides protein, amaranth grain has higher oil content (5–10%) than other cereals (**Table 1**). Amaranth oil contains 76–77% unsaturated fatty acids consisting of primarily linoleic acid (25–62%), oleic acid (19–35%), palmitic acid (12–25%), stearic acid (2.0–8.6%), and linolenic acid (0.3–2.2%). The ratio of saturated to unsaturated fatty acid ranged from 0.26 to 0.31 in oil of amaranth grain. Amaranth lipid is unique due to the presence of high biologically active compounds such as squalene (2–7%), tocopherols (≈2%), phospholipids (up to 10%), and phytosterols (up to 2%) (Becker, 1994; León-Camacho et al., 2001; Berghofer and Schoenlechner, 2002). Squalene is an unsaturated hydrocarbon and is known as an obligatory precursor of sterols and has been reported to have antibacterial properties and anti-oxidative and anti-tumor effects in carcinogenesis.

Amaranth grain contains ≈60–65% of carbohydrate, of which starch constitutes ≈57% and total dietary fiber constitutes ≈8–16% (**Table 1**). Amaranth starch mostly consists of amylopectins (≈89.0–99.9%) and is therefore known as "waxy starch" with unique characteristics of high viscosity and high gelatinization temperature. Despite high amylopectin content, the starch granules of amaranth are smaller (0.8–2.5 μm) than those of other grains (3–34 μm), providing them with high water-binding capacity, higher swelling power, lower gelatinization temperature, and high resistance to amylases, making amaranth a preferential source of starch in the food industry (Williams and Brenner, 1995; Pal et al., 2002) as well as providing high solubility and digestibility. The dietary fiber content in pale-colored amaranth seeds was 8%, whereas it was 16% in black-colored amaranth grains (Mlakar et al., 2009). In the leaves of vegetable amaranth, fiber content ranged from 6.95% to 9.65%, with an average of 8.39% ± 0.1% (Shukla et al., 2006).

Amaranth grains are rich sources of the minerals iron (72–174 mg/kg), calcium (1,300–2,850 mg/kg), phosphorus (455 mg/kg), sodium (160–480 mg/kg), magnesium (2,300–3,360 mg/kg), and zinc (36.2–40.0 mg/kg) as well as the vitamin riboflavin (0.19–0.23 mg/100 g of flour), ascorbic acid (4.5 mg/100 g), niacin (1.17–1.45 mg/100 g), and thiamine (0.07–0.10 mg/100 g) (Becker et al., 1981; Rastogi and Shukla, 2013) (**Tables 2** and **4**). Apart from the seeds, the leaves of vegetable amaranth (*A. tricolor*) are rich in minerals and have an average content of potassium of 3.7 ± 0.26 g/kg, calcium 1.7 ± 0.04 g/kg, magnesium 2.90 ± 0.01 g/kg, zinc 791.7 ± 28.98 mg/kg, iron 1,233.8 ± 50.02 mg/kg, manganese 108.1 ± 3.82 mg/kg, and nickel 222.6 ± 9.51 mg/kg (Shukla et al., 2006). Amaranth has all the necessary daily required vitamins up to a significant level and is an excellent source for reducing vitamin deficiency (Graebner et al., 2004). The leaves of *A. tricolor* contain 0.83 ± 0.02 mg/kg vitamin A (carotenoids) and 112.3 ± 5.0 mg/kg ascorbic acid (vitamin C) (Shukla et al., 2006). Despite these nutritional benefits, amaranth grain contains some antinutritional factors that can limit its food application. Amaranth grain contains growth inhibitors such as phytic acid (0.3–0.6%), saponins (0.09–0.10%), and tannins, etc.

In the recent past, amaranth has become a crop of interest for its high nutritive value and great potential as a functional food given its cholesterol-lowering effect observed in animal models (Plate and Arêas, 2002; Mendonça et al., 2009). Despite the 75% *in vitro* digestibility of amaranth protein, net protein use ranges from 33.5% to 46% (Aguilar et al., 2015). Hejazi et al. (2016) showed that, after germination treatment at 28°C for 48 h, the nutritional quality and digestibility of amaranth resulted in not only improved protein digestibility (84%) but also decreased phytic acid and oxalate. Gamel et al. (2006) and Ferreira and Arêas (2010) demonstrated that amaranth extrusion increased calcium bioavailability in rats and suggested that amaranth can be a complementary source of dietary calcium once its bioavailability is favorably modified by the extrusion process (**Tables 5** and **6**). Whittaker and Ologunde (1990) demonstrated that iron supplementation through amaranth increases hemoglobin content because of the more absorbable form of

iron in amaranth. Subramanian and Gupta (2016) reported that administration of amaranth extract through an oral dose increases the level of NO^{3-} and NO^{2-} in plasma as well as in saliva. The bio-accessibility of phenolic compounds was accessed in five wild (*Amaranthus hybridus*, *Brachiaria brizantha*, *Panicum maximum*, *Rottboellia cochinchinensis*, and *Sorghum arundinaceum*) and two domesticated cereal (*Eleusine corocana* and a red variety of *Sorghum bicolor*) grains found in Zimbabwe and showed that *Amaranthus hybridus* had the highest intestinal bio-accessibility (95.4 ± 0.01%) vis-à-vis the other cereals tested (Chitindingu et al., 2015). Serna-Saldivar et al. (2015) recommended the use of *A. caudatus* grain flour in maize tortillas to improve the bioavailability of nutrients and to reduce diabetes as well as an anthelmintic. Apart from these, no published studies have directly compared the relative bioavailability of other nutritional components of amaranth.

Genomics of *Amaranthus* Species

Amaranth is a C4 pseudo-cereal and can be cultivated in a wide range of environments since it has good tolerance of drought and salinity (Kong et al., 2009; Teng et al., 2016).

Molecular Markers, Genetic Diversity, and Phylogenetic Studies in Amaranth

The species of *Amaranthus* are difficult to characterize taxonomically because of the similarity existing among many of the species, small diagnostic parts, intermediate forms, and the broad geographic distribution (Mujica and Jacobsen, 2003). The taxonomic and phylogenetic position of the genus was investigated using various phenotypic and genotypic evaluations (Sauer, 1955; Stetter and Schmid, 2017). The most recent taxonomic classification divided it into three subgenera: *A. albersia, A. acnida*, and *A. amaranthus* (Mosyakin and Robertson, 1996; Costea and DeMason, 2001). Popa et al. (2010) used nuclear ribosomal DNA (rDNA) Internal Transcribed Spacer (ITS) regions for analyzing 92 accessions of *Amaranthus* and identified 12 out of 92 as weed. Xu and Sun (2001) reported low ITS divergence and poorly resolved phylogeny among the closely related taxa *A. cruentus, A. caudatus,* and *A. hypochondriacus,* and their putative wild progenitors *A. hybridus, A. quitensis,* and *A. powellii*. Chan and Sun (1997) used RAPD markers and reported that grain amaranths originated from a common ancestor, *A. hybridus.* Studies using isozyme markers have demonstrated a high degree of polymorphism between the populations, through which it may be possible to identify the intermediate stages of domestication.

Molecular markers are the essential tools in modern plant breeding research programs. The first attempt to develop and characterize the molecular markers for amaranth was carried out by Mallory et al. (2008), who sequenced 1,457 clones from microsatellite-enriched libraries and identified 353, out of which 179 were found to be polymorphic across the accessions of three grain amaranths. Maughan et al. (2009) applied a genomic reduction strategy with next-generation sequencing and identified 27,658 SNPs among four diverse amaranth accessions. *Amaranthus hypochondriacus* showed the maximum genetic diversity in terms of the number of polymorphic SNPs, whereas *A. cruentus* showed the lowest genetic diversity with 35 polymorphic SNPs only (Maughan et al., 2009). The reduced genetic diversity in *A. cruentus* is consistent with other studies carried out using different marker systems: SSRs, RFLPs, isozymes, and AFLPs (Chan and Sun, 1997; Xu and Sun, 2001; Mallory et al., 2008). Suresh et al. (2014) studied genetic diversity and population structure in 348 amaranth accessions belonging to 33 species using 11 SSR markers. The accessions were grouped into seven different clusters independent of species or geographic origin. The overall PIC value ranged from 0.436 to 0.898, with an average of 0.657, with observed heterozygosity from 0.056 to 0.876. The variation in 29 grain amaranth accessions was studied using 27 phenotypic and 16 RAPD markers, resulting in grouping of these accessions into five clusters at 87.5% similarity coefficient (Akin-Idowu et al., 2016). The molecular phylogeny of 94 accessions representing 35 *Amaranthus* species evaluated using SNP markers through genotyping by sequencing (GBS) showed that most accessions from the same species clustered together and they were classified into three subgenera with a few highly differentiated groups (Stetter and Schmid, 2017). It was also hypothesized that *A. hybridus* might be the ancestor of all three crop amaranth species and *A. quitensis* might be an intermediate between *A. hybridus* and *A. caudatus*.

Transcriptomes and Genomes of Amaranth

Despite the great potentiality of amaranth as a source of nutritional food, few efforts have been made to explore its genomics. Stress-responsive transcriptome analysis of *A. hypochondriacus* was reported to understand the mechanism and identify the genes involved in adapting the species to survive against environmental stresses (Délano-Frier et al., 2011). The first attempt to sequence the draft genome of *A. hypochondriacus used Illumina Genome Analyzer IIx with >100× coverage of estimated genome size of 466 Mb (*Sunil et al., 2014*). The annotation of A. hypochondriacus* resulted in the identification of 24,829 protein-coding genes and 13.76% of the genome was found to consist of repeat elements. They further hypothesized that high lysine content in *Amaranthus* is due to the presence of only one ortholog of *aspartate kinase 1* gene and high expression of the *dihydrodipicolinate synthase* (*DHDPS*) gene in seeds. Recently, a high-quality whole-genome sequencing of agronomically important amaranth cultivar "Plainsman" (*A. hypochondriacus*) was carried out using the Illumina HiSeq platform (Clouse et al., 2016). The assembled genome consisted of 377 Mb in 3,518 scaffolds covering 80.9% of the estimated genome size. Further, the genome consisted of 48% of the repeat sequences, of which Copia-like retrotransposons were predominantly present. Annotation of the genome led to the identification of 23,059 protein-coding genes. Clouse et al. (2016) further resequenced seven accessions of grain amaranths (*A. hypochondriacus, A. caudatus,* and *A. cruentus*) along with *A. hybridus.* SNP-based phylogenetic analysis confirmed that *A. hybridus* is the progenitor species of grain amaranths. Further, the researchers generated a physical map spanning 340 Mb of the whole genome of *A. hypochondriacus.* In addition, Lightfoot et al. (2017) performed single-molecule, real-time sequencing using

the PacBio RSII system to close assembly gaps and used chromatin interaction mapping (Hi-C) to scaffold contigs and thereby improved their own previously reported Illumina-based assembly to a chromosome-scale assembly. The 16 largest scaffolds contain 98.0% of the assembly and represent the haploid chromosome number (n = 16). These researchers further produced physical and genetic maps and identified the betalain locus consisting of *CYP76AD1* (*cytochrome P450*) and *DODA1* (*4,5-DOPA dioxygenase extradiol 1*) involved in betalain biosynthesis, which controls stem color (**Table 8**).

Buckwheat

Buckwheat (*F. esculentum* Monch) is a versatile dicotyledon annual crop and has been grown for centuries for its grains as well as greens to be used as food, feed, vegetable, and fodder, although it was neglected during the 20th century in western countries because of the increased yield of wheat during the Green Revolution (Cawoy et al., 2008). Nonetheless, it is recognized as a potential functional food source in China, Japan, and Taiwan. It is considered as a pseudo-cereal because of its usage as a conventional cereal and its chemical composition (Campbell, 1997). There are several species of buckwheat, out of which only nine species have agricultural and nutritional importance, among which only two are used for food purposes: common buckwheat (*F. esculentum)* and tartary buckwheat (*F. tataricum)*. Common buckwheat is also known as "sweet buckwheat," which is taller, has a thicker stem, and is a widely cultivated species in Asia, Europe, North America, and South Africa, whereas tartary buckwheat is mainly confined to the highlands of southwest China and the Himalayas (Ohnishi, 1993; Zhou et al., 2016).

In the recent past, buckwheat production has increased because of its nutraceutical properties and potential for use in the preparation of functional foods (Li and Zhang, 2001; Bonafaccia et al., 2003a; Bonafaccia et al., 2003b). Buckwheat is rich in nutrients and its seed contains 100–125 mg/g of protein, 650–750 mg/g of starch, 20–25 mg/g of fat, and 20–25 mg/g of mineral (Li and Zhang, 2001).

The protein content in buckwheat flour is higher than in commonly used cereals such as rice, wheat, millet, sorghum, maize, etc. Buckwheat protein primarily contains albumin (180 mg/g), globulin (430 mg/g), prolamin (8 mg/g), and glutelin (230 mg/g) (Javornik and Kreft, 1984; Ikeda et al., 1991; Ikeda and Asami, 2000). The amino acid profile of buckwheat protein is balanced compared with that of other cereals. Buckwheat protein is one of the richest in lysine (EAA) and arginine, which are generally limiting in other cereals (**Table 3**). The amino acid score for buckwheat protein is 100, which is highest among plant sources (Ikeda, 2002). Lys/Arg and Met/Gly ratios determine cholesterol-lowering effects and are lower in buckwheat than in most plants, suggesting that buckwheat protein should have cholesterol-lowering effects (Huff and Carroll, 1980; Sugiyama et al., 1985; Carroll and Kurowska, 1995). Consumption of food products derived from buckwheat could reduce the concentration of cholesterol in blood serum and glycemic and insulin indices (Skrabanja et al., 2001; Stokić et al., 2015). Buckwheat protein contains either no or very low gluten and is thus considered as gluten-free and is recommended to people suffering from coeliac disease. Despite these benefits, the major problem is the low digestibility of buckwheat protein in humans and animals due to the presence of anti-nutritional factors such as protease inhibitors (trypsin inhibitor) and tannins (Ikeda et al., 1986; Ikeda et al., 1991; Campbell, 1997). However, sprouting/germination of buckwheat seed reduces the protease inhibitors, thereby increasing protein digestibility (Kreft, 1983; Ookubo, 1992) (**Table 6**).

Starch is the major carbohydrate in buckwheat grain and it varies from 59% to 70% of the dry mass depending on climatic and cultivation conditions (Qian and Kuhn, 1999). Buckwheat seed starch contains 24% amylose and 76% amylopectin (Praznik et al., 1999) but it's digestibility is much slower (low GI) than other cereals like wheat, therefore good to be consumed by diabetic patients. Buckwheat flour is commonly used during fasting days in India.

The total dietary fiber in seeds varies from 5% to 11% (Izydorczyk et al., 2004). Buckwheat is the richest source of soluble carbohydrates such as fagopyrin and fagopyritols, which have been studied for their use in treating type II diabetes (Kawa et al., 2003). Buckwheat grains also contain from 1.5% to 4.0% lipids (Steadman et al., 2001a). Buckwheat is rich in minerals such as potassium, magnesium, and phosphorus (**Table 2**). Apart from this, buckwheat contains higher contents of zinc, copper, and manganese than other cereals (Ikeda et al., 1999; Steadman et al., 2001b). The bioavailability of zinc, copper, and potassium from buckwheat is high and 100 g of buckwheat flour can provide 13–89% of the recommended dietary allowance (RDA) for Zn, Cu, Mg, and Mn (Bonafaccia et al., 2003b). Buckwheat grains contain higher contents of vitamin B_1 (thiamine), B_2 (riboflavin), B_3 (niacin and niacinamide), K (phylloquinones), and B_6 and B_6 (folates) than other cereals (**Table 4**). Tartary buckwheat contains higher vitamin B_1, B_2, and B_3, whereas common buckwheat contains higher vitamin E (Ikeda et al., 2006). The contents of vitamin C, B_1, and B_6 can be further increased by germinating buckwheat. It has been reported that the content of vitamin C can be increased by up to 0.25 mg/g in buckwheat sprouts (Lintschinger et al., 1997; Kim et al., 2004). Buckwheat also contains trace elements such as selenium (0.0099–0.1208 mg/g), which provides resistance against cancer and AIDS (Shi et al., 2011). Besides being rich in high-quality protein and minerals, buckwheat is rich in many rare components such as flavones, flavonoids, phytosterols, fagopyrins, and thiamin-binding proteins, which are known to have healing effects against some chronic diseases. The presence of flavonoids such as rutin, quercetin, orientin, homoorientin, vitexin, and isovitexin in its leaves, flowers, seeds, and sprouts imparts buckwheat with nutraceutical properties (Zielińska et al., 2012; Raina and Gupta, 2015). Tartary buckwheat has higher flavonoid content (19.54 mg/g) than common buckwheat (0.28 mg/g) (Jiang et al., 2007). Buckwheat is the only grain crop with rutin content that is known to have antioxidant, anti-inflammatory, and anti-

carcinogenic property and it reduces the fragility of blood vessels related to hemorrhagic diseases and hypertension in humans (Oomah and Mazza, 1996; Baumgertel et al., 2003). Whole buckwheat contains 2–5 times more phenolic compounds, whereas buckwheat hull and bran contain 2–7 times more antioxidant than oat and barley (Holasova et al., 2002; Zduńczyk et al., 2006). Because of its good nutritional profile and presence of various nutraceutical compounds with unique medicinal properties and adaptability toward harsh climatic conditions, buckwheat is a suitable alternative for food and nutritional security for marginal environments.

Buckwheat Genomics

Buckwheat (*Fagopyrum* sp.), with its origin in China, is a pseudo-cereal belonging to family Polygonaceae. The genus has 19 species, of which *F. esculentum* (sweet buckwheat) and *F. tataricum* (bitter buckwheat) are the two species in the buckwheat genepool being cultivated predominantly in most parts of the world (Krkošková and Mrazova, 2005). Most of the species in the genus are diploid (2n = 2x = 16), except for *F. cymosum* and *F. gracilipes* being tetraploid (2n = 4x = 32) (Chrungoo et al., 2012; Farooq et al., 2016), with an estimated genome size of 540 Mb (Nagano et al., 2000) (**Table 7**). Despite having so many species, the genetic diversity in buckwheat has become depleted during the past few decades due to changing cropping patterns and food habits. Approximately 5,000 accessions of buckwheat have been collected in Southeast Asia, representing about 52% of the world's collection. China has the largest collection of buckwheat (2800), followed by Russia (2116), Ukraine (1600), and India (1050) (Zhou and Zhang, 1995; IPGRI-APO, 1999; Zhang et al., 2004; Zhou et al., 2018). The number of germplasm collections might be overrepresented due to duplications because of the exchange of germplasm between organizations within and between countries. It is an underused crop but holds tremendous potential due to its short life cycle (≈60 days), ability to grow and survive at higher altitude, and high-quality protein content.

Molecular Markers, Genetic Diversity, and Phylogenetic Studies in Buckwheat

A significant amount of research has been carried out to understand the properties of buckwheat proteins, flavonoids, flavones, phytosterols, thiamin-binding proteins, and other rare compounds (Tomotake et al., 2002; Kreft et al., 2006; Zielinski et al., 2009). However, little effort has been made in the development of molecular markers and genomic resources for buckwheat. RAPD has been used to study the relationship and estimate the genetic diversity between different accessions of *Fagopyrum* species (Sharma and Jana, 2002a; Sharma and Jana, 2002b) as well as to identify molecular markers linked with the *homostylar (Ho)* gene associated with self-compatibility (Aii et al., 1998). Nagano et al. (2001) used AFLP markers to identify markers tightly linked to the homostylar region and designed two primers to develop locus-specific markers for further use in marker-assisted breeding. Using the AFLP marker system, five markers linked to shattering habit in buckwheat were identified and a linkage map around the *sht1* locus was constructed (Matsui et al., 2004). An interspecific linkage map was developed using *F. esculentum* and *F. homotropicum* covering 548.8 cM on eight linkage groups. Microsatellite markers showed comparatively higher polymorphism and expected heterozygosity than AFLP, RAPD, and RFLP markers (Powell et al., 1996). Konishi and Ohnishi (2006) developed 180 SSR markers by sequencing 2,785 clones, out of which 54 were found to be polymorphic. Ma et al. (2009) developed 136 new SSR markers for *F. esculentum* and used them for diversity analysis in related species of the genus *Fagopyrum*. However, out of the 136 SSRs, only 10 were found to be polymorphic on 41 accessions. Kishore et al. (2012) assessed genetic diversity and phylogenetic relationship in 75 accessions using 15 SSR markers and the estimated average gene diversity was 0.2098. Li et al. (2007) attempted to develop microsatellite markers for tartary buckwheat by constructing a genomic library enriched with (gT)n repeats. Apart from these markers, a couple of BAC libraries were constructed for *F. homotropicum* (Nagano et al., 2001) and *F. esculentum* (Yasui et al., 2008), which can be further used for identifying useful genes as well as developing markers. More efforts have been made to carry out research in molecular genetics and plant breeding in common buckwheat, but only fragmentary research efforts have been made in tartary buckwheat. Genetic maps have been developed for both *F. esculentum* (Yasui et al., 2004; Pan and Chen, 2010) and *F. tataricum* (Du et al., 2013) and QTLs governing photosensitivity (Hara et al., 2011) and stem length (Yabe et al., 2014) have been identified. However, these maps remain underused for mapping genes/QTLs due to the lack of enough informative markers distributed across all linkage groups. The whole genome *de novo* sequencing of "HeiFeng No. 1," an elite tartary buckwheat cultivar, using Illumina HiSeq 2000 resulted in 204,340 contigs and 348 Mb of assembled sequences. Further, the sequence survey resulted in the identification of 24,505 SSR motifs. The researchers further carried out SSR fingerprinting of 64 accessions and predicted the population structure of tartary buckwheat (Hou et al., 2016).

Transcriptomes and Genomes of Buckwheat

Transcriptome analysis for floral structure (Logacheva et al., 2011), aluminum toxicity (Chen et al., 2017; Xu et al., 2017) and salt tolerance (Wu et al., 2017) has been carried out to understand the gene regulatory mechanism in buckwheat. Transcriptome analysis at 12 different developmental stages of seed development from fertilization to maturation led to the identification of 11,676 differentially expressed genes in tartary buckwheat (Huang et al., 2017). The candidate genes identified through various transcriptomic studies provide a rich genomic resource for further functional characterization for various traits in buckwheat through a reverse genetics approach.

The whole-genome sequencing of buckwheat (*F. esculentum* Moench) using Illumina HiSeq 2000 resulted in assembly of the genome in 387,594 scaffolds consisting of 1.17 Gb of data. Gene prediction analysis identified 35,816 annotated genes and developed the Buckwheat Genome DataBase (BGDB; http://buckwheat.kazusa.or.jp) (Yasui et al., 2016). Researchers further showed the applicability of genome sequence information in identifying genes involved in flavonoid, 2S albumin-type allergens biosynthesis, and *granule bound starch*

synthases (GBSSs), and in controlling heteromorphic self-incompatibility. Zhang, L., et al (2017) sequenced the whole genome of tartary buckwheat (*F. tataricum* cv. Pinku1) using multiple Illumina NGS platforms and SMRT sequencing technology. A total 8,778 contigs assembled into 489.3 Mb (N50 = 550.7 kb), with maximum contig length of 6.64 Mb. The annotation of the genome resulted in the identification of 33,366 protein-coding genes. Further, the reference genome helped in identifying predicted genes involved in aluminum stress tolerance and abiotic stress responses as well as rutin biosynthesis and regulation. Rutin is a flavonoid with antioxidant property known for its ability to strengthen blood vessels, aiding in the usage of vitamin C and production of collagen. Identification of the genes involved in rutin biosynthesis will further help in increasing the rutin content in buckwheat by using modern genetic and genomics tools.

SUMMARY AND CONCLUSIONS

Improving the climate resilience of crops is crucial to the future food and nutritional security of marginal areas. Climate change is no longer a mirage but a reality that is evident from rising global temperature and more frequent drought, hot, cold, and rain spells worldwide. Every day, a large chunk of normal areas is being transformed into marginal lands, thus significantly decreasing productivity and endangering food security. More than 2 billion people that depend on major staple crops such as wheat, maize, and rice are suffering from "hidden hunger," which has resulted in either malnutrition, due to deficiency of minerals, vitamins, and essential amino acids, or obesity, due to the surfeit of energy-rich carbohydrates. All of these major cereals are unable to tolerate climatic aberrations and marginality because of significant abiotic stresses. The need of the hour is to identify the crops and their varieties that can offer robust resistance against the harsh conditions of marginal environments and sustain food and nutritional security. Indeed, the targets for both globally healthy diets and sustained food production have to be met for 10 billion people on this planet by 2050. Several underused "neglected crops" are considered as "minor crops" and have received less importance globally in terms of both production and research. The five crops (finger millet, quinoa, foxtail millet, buckwheat, and amaranth) reviewed in this manuscript are nutrient-dense and are rich sources of macro- and micro-elements (e.g. protein with balanced amino acids, essential amino acids, vitamins, minerals, etc.). Small amounts of the minor millets and pseudo-cereals in the daily diet of people can ensure "no malnutrition." Apart from being nutritionally rich, these underused crops are climate-resilient and suitable for cultivation in marginal environments. These neglected crops have huge potential for food and nutritional security through sustainable agriculture in marginal areas. This will help farmers to maintain productivity against a backdrop of rising temperatures, higher salinity, and increasing water scarcity and provide a practical and sustainable way of adapting to climate change. Unfortunately, the depth of scientific research on the yield and quality improvement of these minor millets and pseudo-cereals is extremely inadequate vis-à-vis that on the major staples; hence, international funding is extremely important to support research programs on genetic improvement for yield and nutritional traits for these crops.

AUTHOR CONTRIBUTIONS

JR and ST contributed on bioavailability part while HR contributed for genomics part. RS conceptualized, organized and finally polished the manuscript. All authors listed have made a substantial, direct, and intellectual contribution to the work, and approved for publication.

FUNDING

The work was supported by the International Center for Biosaline Agriculture (ICBA) Dubai, UAE.

ACKNOWLEDGMENTS

The authors acknowledge support from Dr. Ismahane Elouafi, Director General, and Ms. Seta Tutundjian, Director Programs, ICBA. We thank the reviewers for helpful comments and suggestions.

REFERENCES

Aguilar, E. G., de Jesús Albarracín, G., Uñates, M. A., Piola, H. D., Camiña, J. M., and Escudero, N. L. (2015). Evaluation of the nutritional quality of the grain protein of new amaranths varieties. *Plant Foods Hum. Nutr.* 70, 21–26. doi: 10.1007/s11130-014-0456-3

Aheme, S. A., and O'Brien, N. M. (1999). Protection by the flavonoids myricetin, quercetin, and rutin against hydrogen peroxide-induced DNA damage in Caco-2 HepG2 cells. *Nutr. Cancer* 34, 160–166. doi: 10.1207/S15327914NC3402_6

Aii, J., Nagano, M., Penner, A. G., Campbell, G. C., and Adachi, T. (1998). Identification of RAPD markers linked to the homostylar (*Ho*) gene in buckwheat. *Japanese J. Breed.* 48, 59–62. doi: 10.1270/jsbbs1951.48.59

Akin-Idowu, P., Gbadegesin, M., Orkpeh, U., Ibitoye, D., and Odunola, O. (2016). Characterization of grain amaranth (*Amaranthus* spp.) germplasm in south west Nigeria using morphological, nutritional, and random amplified polymorphic DNA (RAPD) analysis. *Resources* 5, 6. doi: 10.3390/resources5010006

Albugoch, J., and Lilian, E. (2009). "Chapter 1: Quinoa (*Chenopodium quinoa* Willd.): Composition, Chemistry, Nutritional, and Functional Properties," in *Advances in Food and Nutrition Research*, vol. 58. (Academic Press), 1–31. doi: 10.1016/S1043-4526(09)58001-1

Allen, P. (1961). *Die Amaranthaceen Mittleleuropas Carl Verlag* (Munchen).

Alvarez-Jubete, L., Auty, M., Arendt, E. K., and Gallagher, E. (2010). Baking properties and microstructure of pseudocereal flours in gluten-free bread formulations. *Eur. Food Res. Technol.* 230, 437. doi: 10.1007/s00217-009-1184-z

Amarowicz, R., Dykes, G. A., and Pegg, R. B. (2008). Antibacterial activity of tannin constituents from *Phaseolus vulgaris, Fagoypyrum esculentum, Corylus avellana* and Juglans nigra. *Fitoterapia* 79 (3), 217–219. doi: 10.1016/j.fitote.2007.11.019

Arêas, J. A. G., Carlos-Menezes, A. C. C. C., Soares, R. A. M., Caballero, B., Finglas, P. M., and Toldrá, F. (2016). *Encyclopedia of Food and Health* Vol. Amaranth (Oxford: Academic Press), 135–140. doi: 10.1016/B978-0-12-384947-2.00025-8

Arya, L., Verma, M., Gupta, V. K., and Karihaloo, J. L. (2009). Development of EST SSRs in finger millet (*Eleusine coracana* ssp. *coracana*) and their transferability to pearl millet (*Pennisetum glaucum*). *J. Plant Biochem. Biotechnol.* 18, 97–100. doi: 10.1007/BF03263303

Arya, L., Verma, M., Gupta, V. K., and Seetharam, A. (2013). Use of genomic and genic SSR markers for assessing genetic diversity and population structure in Indian and African finger millet (*Eleusine coracana* (L.) Gaertn.) germplasm. *Plant Syst. Evol.* 299, 1395–1401. doi: 10.1007/s00606-013-0822-x

Austin, D. F. (2006). Fox-tail millets (Setaria: Poaceae)—abandoned food in two hemispheres. *Economic Bot.* 60, 143–158. doi: 10.1663/0013-0001(2006)60[143:FMSPFI]2.0.CO;2

Babu, B. K., Agrawal, P. K., Pandey, D., Jaiswal, J. P., and Kumar, A. (2014a). Association mapping of agro-morphological characters among the global collection of finger millet genotypes using genomic SSR markers. *Mol. Biol. Rep.* 41, 5287–5297. doi: 10.1007/s11033-014-3400-6

Babu, B. K., Agrawal, P. K., Pandey, D., and Kumar, A. (2014b). Comparative genomics and association mapping approaches for *opaque2* modifier genes in finger millet accessions using genic, genomic and candidate gene-based simple sequence repeat markers. *Mol. Breed.* 34, 1261–1279. doi: 10.1007/s11032-014-0115-2

Babu, B. K., Dinesh, P., Agrawal, P. K., Sood, S., Chandrashekara, C., Bhatt, J. C., et al. (2014c). Comparative genomics and association mapping approaches for blast resistant genes in finger millet using SSRs. *PloS One* 9, e99182. doi: 10.1371/journal.pone.0099182

Bai, H., Cao, Y., Quan, J., Dong, L., Li, Z., Zhu, Y., et al. (2013). Identifying the genome-wide sequence variations and developing new molecular markers for genetics research by re-sequencing a landrace cultivar of foxtail millet. *PloS One* 8, e73514. doi: 10.1371/journal.pone.0073514

Baljeet, S. Y., Ritika, B. Y., and Roshan., L. Y. (2010). Studies on functional properties and incorporation of buckwheat flour for biscuit making. *Int. Food Res. J.* 17, 1067–1076.

Balzotti, M. R., Thornton, J. N., Maughan, P. J., McClellan, D. A., Stevens, M. R., Jellen, E. N., et al. (2008). Expression and evolutionary relationships of the *Chenopodium quinoa* 11S seed storage protein gene. *Int. J. Plant Sci.* 169, 281–291. doi: 10.1086/523874

Bandyopadhyay, T., Mehanathan, M., and Prasad, M. (2017). Millets for next generation climate-smart agriculture. *Front. Plant Sci.* 8, 1266. doi: 10.3389/fpls.2017.01266

Baumgertel, A., Grimm, R., Eisenbeiß, W., and Kreis, W. (2003). Purification and characterization of a flavonol 3-O-β-heterodisaccharidase from the dried herb of *Fagopyrum esculentum* Moench. *Phytochemistry* 64, 411–418. doi: 10.1016/S0031-9422(03)00418-7

Bazile, D., Jacobsen, S. E., and Verniau, A. (2016). The global expansion of quinoa: trends and limits. *Front. In Plant Sci.* 7, 622. doi: 10.3389/fpls.2016.00622

Becker, R., Wheeler, E. L., Lorenz, K., Stafford, A. E., Grosjean, O. K., Betschart, A. A., et al. (1981). A compositional study of amaranth grain. *J. Food Sci.* 46, 1175–1180. doi: 10.1111/j.1365-2621.1981.tb03018.x

Becker, R. (1994). "Amaranth oil: composition, processing, and nutritional qualities," in *Amaranth biology, chemistry and technology*. Ed. L. O. Paredes (México: Instituto Politécnico Nacional), 133–139.

Belton, P. S., and Taylor, J. R. (2004). Sorghum and millets: protein sources for Africa. *Trends In Food Sci. Technol.* 15, 94–98. doi: 10.1016/j.tifs.2003.09.002

Bennett, M. D., and Smith, J. B. (1991). Nuclear DNA amounts in angiosperms. *Philos. Trans.: Biol. Sci.*, 309–345. doi: 10.1098/rstb.1991.0120

Bennetzen, J. L., Schmutz, J., Wang, H., Percifield, R., Hawkins, J., Pontaroli, A. C., et al. (2012). Reference genome sequence of the model plant *Setaria*. *Nat. Biotechnol.* 30, 555. doi: 10.1038/nbt.2196

Berghofer, E., and Schoenlechner, R. (2002). "Grain Amaranth," in *Pseudocereals and less common cereals, grain properties and utilization potential*. Eds. P. Belton and J. Taylor (Springer-Verlag), 219–260.

Betschart, A. A., Irving, D. W., Shepherd, A. D., and Saunders, R. M. (1981). *Amaranthus cruentus*: milling characteristics, distribution of nutrients within seed components, and the effects of temperature on nutritional quality. *J. Food Sci.* 46, 1181–1187. doi: 10.1111/j.1365-2621.1981.tb03019.x

Bharathi, A. (2011). *Phenotypic and Genotypic Diversity of Global Finger Millet (Eleusine coracana (L.) Gaertn.) Composite Collection. (dissertation/Ph.D. thesis)* (Tamil Nadu Agricultural University).

Bhargava, A., Shukla, S., and Ohri, D. (2003). Genetic variability and heritability of selected traits during different cuttings of vegetable *Chenopodium*. *Indian J. Genet. Plant Breed.* 63 (4), 359–360.

Bhargava, A., Shukla, S., and Ohri, D. (2006a). Chenopodium quinoa—an Indian perspective. *Ind. Crops Prod.* 23, 73–87. doi: 10.1016/j.indcrop.2005.04.002

Bhargava, A., Shukla, S., and Ohri, D. (2006b). Karyotypic studies on some cultivated and wild species of *Chenopodium* (Chenopodiaceae). *Genet. Resour. Crop Evol.* 53, 1309–1320. doi: 10.1007/s10722-005-3879-8

Bhargava, A., Shukla, S., and Ohri, D. (2007a). Genome size variation in some cultivated and wild species of *Chenopodium* (Chenopodiaceae). *Caryologia* 60, 245–250. doi: 10.1080/00087114.2007.10797943

Bhargava, A., Shukla, S., Rajan, S., and Ohri, D. (2007b). Genetic diversity for morphological and quality traits in quinoa (*Chenopodium quinoa* Willd.) germplasm. *Genet. Resour. Crop Evol.* 54, 167–173. doi: 10.1007/s10722-005-3011-0

Bindra, G. S., Gibson, R. S., and Thompson, L. U. (1986). [Phytate][calcium]/[zinc] ratios in Asian immigrant lacto-ovo vegetarian diets and their relationship to zinc nutriture. *Nutr. Res.* 6, 475–483. doi: 10.1016/S0271-5317(86)80101-4

Bisht, M. S., and Mukai, Y. (2001). Genomic *in situ* hybridization identifies genome donor of finger millet (*Eleusine coracana*). *Theor. Appl. Genet.* 102, 825–832. doi: 10.1007/s001220000497

Bonafaccia, G., Gambelli, L., Fabjan, N., and Kreft, I. (2003a). Trace elements in flour and bran from common and tartary buckwheat. *Food Chem.* 83, 1–5. doi: 10.1016/S0308-8146(03)00228-0

Bonafaccia, G., Marocchini, M., and Kreft, I. (2003b). Composition and technological properties of the flour and bran from common and tartary buckwheat. *Food Chem.* 80, 9–15. doi: 10.1016/S0308-8146(02)00228-5

Borges, J. T. S., Bonomo, R. C., Paula, C. D., Oliveira, L. C., and Cesário, M. C. (2010). Physicochemical and nutritional characteristics and uses of Quinoa (*Chenopodium quinoa* Willd.). *Temas Agrários* 15 (1), 9–23.

Brady, K., Ho, C.-T., Rosen, R. T., Sang, S., and Karwe, M. V. (2007). Effects of processing on the nutraceutical profile of quinoa. *Food Chem.* 100 (3), 1209–1216. doi: 10.1016/j.foodchem.2005.12.001

Bressani, R. (1989). The proteins of grain amaranth. *Food Rev. Int.* 5, 13–38. doi: 10.1080/87559128909540843

Brown, D. C., Cepeda-Cornejo, V., Maughan, P. J., and Jellen, E. N. (2015). Characterization of the granule-bound starch synthase I gene in Chenopodium. *Plant Genome* 8, 1–12. doi: 10.3835/plantgenome2014.09.0051

Campbell, C. G. (1997). *Buckwheat Fagopyrum esculentum Moench. promoting the conservation and use of underutilized and neglected crops 19* (Rome, Italy: Institute of Plant Genetics and Crop Plant Research; Gatersleben/International Plant Genetic Resources Institute).

Capriles, V. D., Lopes Almeida, E., Ferreira, E. R., Gomes Arêas, J. A., Steel, C. J., and Chang, Y. K. (2008). Physical and sensory properties of regular and reduced-fat pound cakes with added amaranth flour. *Cereal Chem.* 85 (5), 614–618. doi: 10.1094/CCHEM-85-5-0614

Carroll, K. K., and Kurowska, E. M. (1995). Soy consumption and cholesterol reduction: review of animal and human studies. *J. Nutr.* 125, 594S–597S. doi: 10.1093/jn/125.3_Suppl.594S

Castro-Alba, V., Lazarte, C. E., Perez-Rea, D., Carlsson, N.-G., Almgren, A., Bergenståhl, B., et al. (2019). Fermentation of pseudocereals quinoa, cañihua, and amaranth to improve mineral accessibility through degradation of phytate. *J. Sci. Food Agric.* 99, 5239–5248. doi: 10.1002/jsfa.9793

Cawoy, V., Kinet, J. M., and Jacquemart, A. L. (2008). Morphology of nectaries and biology of nectar production in the distylous species *Fagopyrum esculentum*. *Ann. Bot.* 102, 675–684. doi: 10.1093/aob/mcn150

Chan, K. F., and Sun, M. (1997). Genetic diversity and relationships detected by isozyme and RAPD analysis of crop and wild species of *Amaranthus*. *Theor. Appl. Genet.* 95, 865–873. doi: 10.1007/s001220050637

Chauhan, E., and Sarita, S. (2018). Effects of processing (germination and popping) on the nutritional and anti-nutritional properties of finger millet (*Eleusine coracana*). *Curr. Res. In Nutr. Food Sci.* 6 (2), 566–572. doi: 10.12944/CRNFSJ.6.2.30

Chauhan, G. S., Eskin, N. A. M., and Tkachuk, R. (1992). Nutrients and antinutrients in quinoa seed. *Cereal Chem.* 69 (1), 85–88.

Chen, W., Xu, J., Jin, J., Lou, H., Fan, W., and Yang, J. (2017). Genome-wide transcriptome analysis reveals conserved and distinct molecular mechanisms of Al resistance in buckwheat (*Fagopyrum esculentum* Moench) leaves. *Int. J. Mol. Sci.* 18, 1859. doi: 10.3390/ijms18091859

Cheng, R., and Dong, Z. (2010). Breeding and production of foxtail millet in China. In: Ed. Z. He APA Cereals in China. (CIMMYT, Mexico, pp 87–95.

Chethan, S., and Malleshi, N. (2007). Finger millet polyphenols: characterization and their nutraceutical potential. *Am. J. Food Technol.* 2 (7), 582–592. doi: 10.3923/ajft.2007.582.592

Chitindingu, K., Benhura, M. A., and Muchuweti, M. (2015). In vitro bioaccessibility assessment of phenolic compounds from selected cereal grains: a prediction tool of nutritional efficiency. *LWT-Food Sci. Technol.* 63, 575–581. doi: 10.1016/j.lwt.2015.02.026

Choukr-Allah, R., Rao, N. K., Hirich, A., Shahid, M., Alshankiti, A., Toderich, K., et al. (2016). Quinoa for marginal environments: toward future food and nutritional security in MENA and Central Asia regions. *Front. Plant Sci.* 7, 346. doi: 10.3389/fpls.2016.00346

Christa, K., and Soral-Śmietana, M. (2008). Buckwheat grains and buckwheat products: nutritional and prophylactic value of their components – a review. *Czech J. Food Sci.* 26, 153–162. doi: 10.17221/1602-cjfs

Christensen, S. A., Pratt, D. B., Pratt, C., Nelson, P. T., Stevens, M. R., Jellen, E. N., et al. (2007). Assessment of genetic diversity in the USDA and CIP-FAO international nursery collections of quinoa (*Chenopodium quinoa* Willd.) using microsatellite markers. *Plant Genet. Resour.* 5, 82–95. doi: 10.1017/S1479262107672293

Chrungoo, N. K., Kreft, I., Sangma, S. C., Devadasan, N., Dohtdong, L., and Chetri, U. (2012). "Genetic diversity in Himalayan buckwheats: a perspective for use in crop improvement programmes," in *Proceedings of the 12th International Symposium on Buckwheat* (Pernica), 198–211.

Clouse, J. W., Adhikary, D., Page, J. T., Ramaraj, T., Deyholos, M. K., Udall, J. A., et al. (2016). The amaranth genome: genome, transcriptome, and physical map assembly. *Plant Genome* 9, 1–14. doi: 10.3835/plantgenome2015.07.0062

Coles, N. D., Coleman, C. E., Christensen, S. A., Jellen, E. N., Stevens, M. R., Bonifacio, A., et al. (2005). Development and use of an expressed sequenced tag library in quinoa (*Chenopodium quinoa* Willd.) for the discovery of single nucleotide polymorphisms. *Plant Sci.* 168, 439–447. doi: 10.1016/j.plantsci.2004.09.007

Comai, S., Bertazzo, A., Bailoni, L., Zancato, M., Costa, C. V. L., and Allegri, G. (2007). The content of proteic and nonproteic (free and protein-bound) tryptophan in quinoa and cereal flours. *Food Chem.* 100 (4), 1350–1355. doi: 10.1016/j.foodchem.2005.10.072

Costea, M., and DeMason, D. A. (2001). Stem morphology and anatomy in *Amaranthus* L. (Amaranthaceae), taxonomic significance. *J. Torrey Bot. Soc.* 128, 254–281. doi: 10.2307/3088717

Délano-Frier, J. P., Avilés-Arnaut, H., Casarrubias-Castillo, K., Casique-Arroyo, G., Castrillón-Arbeláez, P. A., Herrera-Estrella, L., et al. (2011). Transcriptomic analysis of grain amaranth (*Amaranthus hypochondriacus*) using 454 pyrosequencing: comparison with *A. tuberculatus*, expression profiling in stems and in response to biotic and abiotic stress. *BMC Genomics* 12, 363. doi: 10.1186/1471-2164-12-363

d'Ennequin, M. L. T., Panaud, O., Toupance, B., and Sarr, A. (2000). Assessment of genetic relationships between *Setaria italica* and its wild relative *S. viridis* using AFLP markers. *Theor. Appl. Genet.* 100, 1061–1066. doi: 10.1007/s001220051387

Dayakar, R. B., Bhaskarachary, K., Arlene, C. G. D., Sudha, D. G., and Vilas, A. T. (2017). *Nutritional and Health Benefits of Millets* (Rajendranagar, Hyderabad: ICAR_Indian Institute of Millets Research (IIMR)), 112.

Debski, B., Gralak, M. A., Bertrandt, J., and Kłos, A. (2013). Minerals and polyphenols content of quinoa (*Chenopodium quinoa* Willd.) Plant. *Probl. Hig. Epidemiol.* 94 (2), 300–304.

del Castillo, C., Winkel, T., Mahy, G., and Bizoux, J.-P. (2007). Genetic structure of quinoa (*Chenopodium quinoa* Willd.) from the Bolivian altiplano as revealed by RAPD markers. *Genet. Resour. Crop Evol.* 54, 897–905. doi: 10.1007/s10722-006-9151-z

Deschner, E. E., Ruperto, J., Wong, G., and Newmark, H. L. (1991). Quercetin and rutin as inhibitors of azoxymethanol-induced colonic neoplasia. *Carcinogenesis* 12, 1193–1196. doi: 10.1093/carcin/12.7.1193

Desikachar, H. S. R. (1975). Processing of maize, sorghum and millets for food uses. *J. Sci. Ind. Res.* 34, 231–237.

Devi, P. B., Vijayabharathi, R., Sathyabama, S., Malleshi, N. G., and Priyadarisini, V. B. (2014). Health benefits of finger millet (*Eleusine coracana* L.) polyphenols and dietary fiber: a review. *J. Food Sci. Technol.* 51, 1021–1040. doi: 10.1007/s13197-011-0584-9

Devos, K. M., Wang, Z. M., Beales, J., Sasaki, T., and Gale, M. D. (1998). Comparative genetic maps of foxtail millet (*Setaria italica*) and rice (*Oryza sativa*). *Theor. Appl. Genet.* 96, 63–68. doi: 10.1007/s001220050709

Dhankher, O. P., and Foyer, C. H. (2018). Climate resilient crops for improving global food security and safety. *Plant Cell Environ.* 41, 877–884. doi: 10.1111/pce.13207

Dida, M. M., Ramakrishnan, S., Bennetzen, J. L., Gale, M. D., and Devos, K. M. (2007). The genetic map of finger millet, *Eleusine coracana*. *Theor. Appl. Genet.* 114, 321–332. doi: 10.1007/s00122-006-0435-7

Dida, M. M., Wanyera, N., Dunn, M. L. H., Bennetzen, J. L., and Devos, K. M. (2008). Population structure and diversity in finger millet (*Eleusine coracana*) germplasm. *Trop. Plant Biol.* 1, 131–141. doi: 10.1007/s12042-008-9012-3

Dini, I., Schettino, O., Simioli, T., and Dini, A. (2001). Studies on the constituents of Chenopodium quinoa seeds: Isolation and characterization of new triterpene saponins. *J. Agric. Food Chem.* 49 (2), 741–746. doi: 10.1021/jf000971y

Du, X., Zhang, Z., Wu, B., Li, Y., and Yu, A. H. (2013). Construction of genetic linkage map in tartary buckwheat (*Fagopyrum tataricum*) using SSR. *Chin. Agric. Sci. Bull.* 29 (21), 61–65.

Dwivedi, S. L., Upadhyaya, H. D., Senthilvel, S., Hash, C. T., Fukunaga, K., Diao, X., et al. (2012). Millets: genetic and genomic resources, in Plant Breeding Reviews, Ed. J. Janick. *Plant Breed. Rev.* (Hoboken, NJ: John Wiley and Sons, Inc.), 247–374.

Eisa, S. S., Eid, M. A., Abd El-Samad, E. H., Hussin, S. A., Abdel-Ati, A. A., El-Bordeny, N. E., et al. (2017). *Chenopodium quinoa* Willd.: a new cash crop halophyte for saline regions of Egypt. *Aust. J. Crop Sci.* 11, 343. doi: 10.21475/ajcs.17.11.03.pne316

Fabjan, N., Rode, J., Kozir, I. J., Wang, Z., Zhang, Z., and Kreft, I. (2003). Tartary buckwheat (*Fagopyrum tartaricum* Gaertn.) as a source of dietary rutin and quercitrin. *J. Agric. Food Chem.* 51 (22), 6452–6455. doi: 10.1021/jf034543e

Fairbanks, D. J., Waldrigues, A., Ruas, C. F., Ruas, P. M., Maughan, P. J., Robison, L. R., et al. (1993). Efficient characterization of biological diversity using field DNA extraction and random amplified polymorphic DNA markers. *Rev. Bras. Genet* 16, 11–11.

FAO. (2017). *The future of food and agriculture–Trends and challenges* (Rome: Food and Agriculture Organisation).

Farooq, S., Rehman, R. U., Pirzadah, T. B., Malik, B., Dar, F. A., and Tahir, I. (2016). "Cultivation, agronomic practices, and growth performance of buckwheat," in *Molecular breeding and nutritional aspects of buckwheat* (Elsevier), 299–319.

Farro, P. C. A. (2008). *Desenvolvimento de filmes biodegradáveis a partir de derivados do grão de quinoa (Chenopodium quinoa Willdenow) da variedade "Real." Thesis, Faculdade de Engenharia de Alimentos* (Campinas, Brazil: Universidade Estadual de Campinas). 303 pp.

Ferreira, T. A., and Arêas, J. A. G. (2010). Calcium bioavailability of raw and extruded amaranth grains. *Food Sci. Technol.* 30, 532–538. doi: 10.1590/S0101-20612010000200037

Foucault, A.-S., Mathé, V., Lafont, R., Even, P., Dioh, W., Veillet, S., et al. (2012). Quinoa extract enriched in 20-hydroxyecdysone protects mice from diet-induced obesity and modulates adipokines expression. *Obesity* 20 (2), 270–277. doi: 10.1038/oby.2011.257

Fuentes, F., and Bhargava, A. (2011). Morphological analysis of quinoa germplasm grown under lowland desert conditions. *J. Agron. Crop Sci.* 197, 124–134. doi: 10.1111/j.1439-037X.2010.00445.x

Fuentes, F., Martinez, E. A., Hinrichsen, P. V., Jellen, E. N., and Maughan, P. J. (2009). Assessment of genetic diversity patterns in Chilean quinoa (*Chenopodium quinoa* Willd.) germplasm using multiplex fluorescent microsatellite markers. *Conserv. Genet.* 10, 369–377. doi: 10.1007/s10592-008-9604-3

Fukunaga, K., Wang, Z., Kato, K., and Kawase, M. (2002). Geographical variation of nuclear genome RFLPs and genetic differentiation in foxtail millet, *Setaria italica* (L.) P. Beauv. *Genet. Resour. Crop Evol.* 49, 95–101. doi: 10.1023/A:1013852007770

Gahlawat, P., and Sehgal, S. (1994). Protein and starch digestibility and iron availability in developed weaning foods as affected by roasting. *J. Hum. Nutr. Dietetics* 7, 121–126. doi: 10.1111/j.1365-277X.1994.tb00419.x

Gahlawat, P., and Sehgal, S. (1995). Improvement in HCl-extractability of minerals in home made weaning foods. *Plant Foods Hum. Nutr.*, 47, 173–179. doi: 10.1007/BF01089267

Gamel, T. H., Jozef, P. L., Ahmed, S. M., Ahmed, A. D., and Lila, A. S. (2006). Seed treatments affect functional and antinutritional properties of amaranth flours. *J. Sci. Food Agric.* 86 (7), 1095–1102. doi: 10.1002/jsfa.2463

Ganapathy, N. S., Chitra, R. G., and Gokhale, S. K. (1957). The effect of the protein of Italian millet (*Setaria italica*) on nitrogen retention in albino rats. *Indian J. Med. Res.* 45, 395–399.

Geervani, P., and Eggum, B. O. (1989). Nutrient composition and protein quality of minor millets. *Plant Foods Hum. Nutr. (Netherlands)* 39, 201. doi: 10.1007/BF01091900

Gheldof, N., Wang, X.-H., and Engeseth, N. J. (2003). Buckwheat honey increases serum antioxidant capacity in humans. *J. Agric. Food Chem.* 51, 1500–1505. doi: 10.1021/jf025897t

Gimode, D., Odeny, D. A., de Villiers, E. P., Wanyonyi, S., Dida, M. M., Mneney, E. E., et al. (2016). Identification of SNP and SSR markers in finger millet using next generation sequencing technologies. *PloS One* 11, e0159437. doi: 10.1371/journal.pone.0159437

Global Nutrition Report(2018). https://globalnutritionreport.org/reports/global-nutrition-report-2018/burden-malnutrition/.

Gopalan, C., Rama Sastri, B. V., and Balasubramanian, S. C. (1989). Nutritive Value of Indian Foods (NVIF), Revised & Updated by B G. Deosthala, and K. C. Pant (Reprinted 2007, 2011). National Institute of Nutrition, Indian Council of Medical Research (ICMR) Hyderabad-500007, India.

Goron, T. L., and Raizada, M. N. (2015). Genetic diversity and genomic resources available for the small millet crops to accelerate a New Green Revolution. *Front. In Plant Sci.* 6, 157. doi: 10.3389/fpls.2015.00157

Graebner, I. T., Siqueira, E. M., Arruda, S. F., and de Souza, E. M. (2004). Carotenoids from native Brazilian dark-green vegetables are bioavailable: a study in rats. *Nutr. Res.* 24, 671–679. doi: 10.1016/j.nutres.2003.10.012

Graf, B. L., Poulev, A., Kuhn, P., Grace, M. H., Lila, M. A., and Raskin, I. (2014). Quinoa seeds leach phytoecdysteroids and other compounds with anti-diabetic properties. *Food Chem.* 163, 178–185. doi: 10.1016/j.foodchem.2014.04.088

Griffith, J. Q., Couch, J. F., and Lindauer., M. A. (1944). Effect of rutin on increased capillary fragility in man. *Proc. Soc. Exp. Biol. Med.* 55, 228–229. doi: 10.3181/00379727-55-14532

Grinberg, L. N., Rachmilewitz, E. A., and Newmark, H. (1994). Protective effects of rutin against haemoglobin oxidation. *Biochem. Pharmacol.* 48, 643–649. doi: 10.1016/0006-2952(94)90040-x

Guardia, T., Rotelli, A. E., Juarez, A. O., and Pelzer, L. E. (2001). Anti-inflammatory properties of plant flavonoids: effects of rutin, quercetin and hesperidin on adjuvant arthritis in rat. *Farmaco* 56, 683–687. doi: 10.1016/S0014-827X(01)01111-9

Gull, A., Jan, R., Nayik, G. A., Prasad, K., and Kumar, P. (2014). Significance of finger millet in nutrition, health and value added products: a review. *Magnesium (mg)* 130, 120.

Gupta, S., Kumari, K., Das, J., Lata, C., Puranik, S., and Prasad, M. (2011). Development and utilization of novel intron length polymorphic markers in foxtail millet (*Setaria italica* (L.) P. Beauv.). *Genome* 54, 586–602. doi: 10.1139/g11-020

Gupta, S., Kumari, K., Sahu, P. P., Vidapu, S., and Prasad, M. (2012). Sequence-based novel genomic microsatellite markers for robust genotyping purposes in foxtail millet [*Setaria italica* (L.) P. Beauv.]. *Plant Cell Rep.* 31, 323–337. doi: 10.1007/s00299-011-1168-x

Gupta, S., Kumari, K., Muthamilarasan, M., Parida, S. K., and Prasad, M. (2014). Population structure and association mapping of yield contributing agronomic traits in foxtail millet. *Plant Cell Rep.* 33 (6), 881–893. doi: 10.1007/s00299-014-1564-0

Gupta, S. M., Arora, S., Mirza, N., Pande, A., Lata, C., Puranik, S., et al. (2017). Finger millet: a "certain" crop for an "uncertain" future and a solution to food insecurity and hidden hunger under stressful environments. *Front. In Plant Sci.* 8, 643. doi: 10.3389/fpls.2017.00643

Guruvayoorappan, C., and Kuttan, G. (2007). Antiangiogenic effect of rutin and its regulatory effect on the production of VEGF, IL-1 and TNF-a in tumor associated macrophages. *J. Biol. Sci.* 7, 1511–1519. doi: 10.3923/jbs.2007.1511.1519

Guzmán-Maldonado, S. H., and Paredes-López, O. (1998). Production of high-protein flours as milk substitutes. In. *Funct. Prop. Proteins Lipids* 708, 66–79. doi: 10.1021/bk-1998-0708.ch005. ACS Symposium Series 708. American Chemical Society.

Handoyo, T., Meda, T., Urisu, A., Adachi, T., and Morita, M. (2006). Hypoallergenic buckwheat flour preparation by Rhizopus oligosporus and its application to soba noodle. *Food Res. Int.* 39, 598–605. doi: 10.1016/j.foodres.2005.12.003

Hara, T., Iwata, H., Okuno, K., Matsui, K., and Ohsawa, R. (2011). QTL analysis of photoperiod sensitivity in common buckwheat by using markers for expressed sequence tags and photoperiod-sensitivity candidate genes. *Breed. Sci.* 61, 394–404. doi: 10.1270/jsbbs.61.394

Hatakeyama, M., Aluri, S., Balachadran, M. T., Sivarajan, S. R., Patrignani, A., Grüter, S., et al. (2018). Multiple hybrid *de novo* genome assembly of finger millet, an orphan allotetraploid crop. *DNA Res.* 25, 39–47. doi: 10.1093/dnares/dsx036

He, J., Klag, M. J., Whelton, P. K., Mo, J. P., Chen, J. Y., Qian, M. C., et al. (1985). Oats and buckwheat intakes and cardiovascular disease risk factors in an ethnic minority of China. *Am. J. Clin. Nutr.* 61, 366–372. doi: 10.1093/ajcn/61.2.366

Hejazi, S. N., Orsat, V., Azadi, B., and Kubow, S. (2016). Improvement of the *in vitro* protein digestibility of amaranth grain through optimization of the malting process. *J. Cereal Sci.* 68, 59–65. doi: 10.1016/j.jcs.2015.11.007

Hemalatha, S., Platel, K., and Srinivasan, K. (2007a). Zinc and iron content and their bioaccessibility in cereals and pulses consumed in India. *Food Chem.* 102, 1328–1336. doi: 10.1016/j.foodchem.2006.07.015

Hemalatha, S., Platel, K., and Srinivasan, K. (2007b). Influence of germination and fermentation on bioaccessibility of zinc and iron from food grains. *Eur. J. Clin. Nutr.* 61, 342–348. doi: 10.1038/sj.ejcn.1602524

Hinojosa, L., González, J. A., Barrios-Masias, F. H., Fuentes, F., and Murphy, K. M. (2018). Quinoa abiotic stress responses: a review. *Plants* 7, 106. doi: 10.3390/plants7040106

Hirich, A., Choukr-Allah, R., and Jacobsen, S. E. (2014). Quinoa in Morocco–effect of sowing dates on development and yield. *J. Agron. Crop Sci.* 200, 371–377. doi: 10.1111/jac.12071

Hithamani, G., and Srinivasan, K. (2014). Effect of domestic processing on the polyphenol content and bio-accessibility in finger millet (*Eleusine coracana*) and pearl millet (*Pennisetum glaucum*). *Food Chem.* 164, 55–62. doi: 10.1016/j.foodchem.2014.04.107

Hittalmani, S., Mahesh, H. B., Shirke, M. D., Biradar, H., Uday, G., Aruna, Y. R., et al. (2017). Genome and transcriptome sequence of finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. *BMC Genomics* 18, 465. doi: 10.1186/s12864-017-3850-z

Holasova, M., Fiedlerova, V., Smrcinova, H., Orsak, M., Lachman, J., and Vavreinova, S. (2002). Buckwheat—the source of antioxidant activity in functional foods. *Food Res. Int.* 35, 207–211. doi: 10.1016/S0963-9969(01)00185-5

Hotz, C., and Gibson, R. S. (2007). Traditional Food-Processing and Preparation Practices to Enhance the Bioavailability of Micronutrients in Plant-Based Diets. *J. Nutr.* 137 (4), 1097–1100. doi: 10.1093/jn/137.4.1097

Hou, S., Sun, Z., Linghu, B., Xu, D., Wu, B., Zhang, B., et al. (2016). Genetic diversity of buckwheat cultivars (*Fagopyrum tartaricum* Gaertn.) assessed with SSR markers developed from genome survey sequences. *Plant Mol. Biol. Rep.* 34, 233–241. doi: 10.1007/s11105-015-0907-5

Huang, J., Deng, J., Shi, T., Chen, Q., Liang, C., Meng, Z., et al. (2017). Global transcriptome analysis and identification of genes involved in nutrients accumulation during seed development of rice tartary buckwheat (*Fagopyrum tararicum*). *Sci. Rep.* 7, 11792. doi: 10.1038/s41598-017-11929-z

Huff, M. W., and Carroll, K. K. (1980). Effects of dietary protein on turnover, oxidation, and absorption of cholesterol, and on steroid excretion in rabbits. *J. Lipid Res.* 21, 546–548.

Hurrell. (2004). Phytic acid degradation as a means of improving iron absorption. *Int. J. Vitam. Nutr. Res.* 74, 445–452. doi: 10.1024/0300-9831.74.6.445

Iglesias-Puig, E., Monedero, V., and Haros, M. (2015). Bread with whole quinoa flour and bifidobacterial phytases increases dietary mineral intake and bioavailability. *LWT-Food Sci. Technol.* 60, 71–77. doi: 10.1016/j.lwt.2014.09.045

Ikeda, K., and Asami, Y. (2000). Mechanical characteristics of buckwheat noodles. *Fagopyrum* 17, 67–72.

Ikeda, K., and Ikeda, S. (2016). "Chapter fifteen - Factors important for structural properties and quality of buckwheat products," in *Molecular Breeding and Nutritional Aspects of Buckwheat*. Eds. M. Zhou, I. Kreft, S.-H. Woo, N. Chrungoo and G. Wieslander (Academic Press), 193–202. doi: 10.1016/B978-0-12-803692-1.00015-8

Ikeda, K., and Kishida, M. (1993). Digestibility of proteins in buckwheat seed. *Fagopyrum* 13, 21–24.

Ikeda, S., and Yamashita, Y. (1994). Buckwheat as a dietary source of zinc, copper and manganese. *Fagopyrum* 14, 29–34.

Ikeda, K., Oku, M., Kusano, T., and Yasumoto, K. (1986). Inhibitory potency of plant antinutrients towards the *in vitro* digestibility of buckwheat protein. *J. Food Sci.* 511527 - (6), 1530. doi: 10.1111/j.1365-2621.1986.tb13851.x

Ikeda, K., Sakaguchi, T., Kusano, T., and Yasumoto, K. (1991). Endogenous factors affecting protein digestibility in buckwheat. *Cereal Chem.* 68, 424–427.

Ikeda, S., Yamashita, Y., and Kreft, I. (1999). Mineral composition of buckwheat by-products and its processing characteristics to konjak preparation. *Fagopyrum* 16, e94.

Ikeda, S., Yamashita, Y., Tomura, K., and Kreft, I. (2006). Nutritional comparison in mineral characteristics between buckwheat and cereals. *Fagopyrum* 23, 5.

Ikeda, K. (2002). Buckwheat composition, chemistry, and processing. *Adv. in Food Nutr. Res.* 44, 395–434. doi: 10.1016/S1043-4526(02)44008-9

IPGRI-APO. (1999). *Status reports on genetic resources of buckwheat.* (Serdang, Malaysia; IPGRI regional office for Asia, the Pacific and Oceania). 73.

Itagi, S., Naik, R., Bharati, P., and Sharma, P. (2012). Readymade foxtail millet mix for diabetics. *Int. J. Sci. Nat.* 3, 131–134.

Izydorczyk, M. S., You, S., and Campbell, C. (2004). "Molecular characteristics of amylose and amylopectin in buckwheat starches," in *American Association of Cereal Chemists 2004 meeting.* (St. Paul MN, USA. American Association of Cereal Chemists) Abstract available online: https://www.cerealsgrains.org/meetings/Documents/Pre2009Abstracts/2004Abstracts/a04ma17.htm.

Jacobsen, S. E., Monteros, C., Christiansen, J. L., Bravo, L. A., Corcuera, L. J., and Mujica, A. (2005). Plant responses of quinoa (*Chenopodium quinoa* Willd.) to frost at various phenological stages. *Eur. J. Agron.* 22, 131–139. doi: 10.1016/j.eja.2004.01.003

Jacobsen, S.-E., Sørensen, M., Pedersen, S. M., and Weiner, J. (2013). Feeding the world: genetically modified crops versus agricultural biodiversity. *Agron. Sustain. Dev.* 33, 651–662. doi: 10.1007/s13593-013-0138-9

Jacobsen, S.-E. (2003). The worldwide potential for quinoa (*Chenopodium quinoa* Willd.). *Food Rev. Int.* 19, 167–177. doi: 10.1081/FRI-120018883

Jaiswal, V., Gupta, S., Gahlaut, V., Muthamilarasan, M., Bandyopadhyay, T., Ramchiary, N., et al. (2019). Genome-wide association study of major agronomic traits in foxtail millet (*Setaria italica* L.) using ddRAD sequencing. *Sci. Rep.* 9, 5020. doi: 10.1038/s41598-019-41602-6

Jancurová, M., Minarovičová, L., and Dandár, A. (2009). Quinoa: A review. *Czech J. Food Sci.* 27 (2), 71–79. doi: 10.17221/32/2008-CJFS

Jarvis, D. E., Kopp, O. R., Jellen, E. N., Mallory, M. A., Pattee, J., Bonifacio, A., et al. (2008). Simple sequence repeat marker development and genetic mapping in quinoa (*Chenopodium quinoa* Willd.). *J. Genet.* 87, 39–51. doi: 10.1007/s12041-008-0006-6

Jarvis, D. E., Ho, Y. S., Lightfoot, D. J., Schmöckel, S. M., Li, B., Borm, T. J. A., et al. (2017). The genome of *Chenopodium quinoa*. *Nature* 542, 307–312. doi: 10.1038/nature21370

Javornik, B., and Kreft, I. (1984). Characterization of buckwheat proteins. *Fagopyrum* 4, 30–38.

Jayaraman, A., Puranik, S., Rai, N. K., Vidapu, S., Sahu, P. P., Lata, C., et al. (2008). cDNA-AFLP analysis reveals differential gene expression in response to salt stress in foxtail millet (*Setaria italica* L.). *Mol. Biotechnol.* 40, 241–251. doi: 10.1007/s12033-008-9081-4

Je, H. D., Shin, C. Y., Park, S. Y., Yim, S. H., Kum, C., Huh, I. H., et al. (2002). Combination of vitamin c and rutin on neuropathy and lung damage of diabetes mellitus rats. *Arch. Pharm. Res.* 25, 184–190. doi: 10.1007/BF02976561

Jia, X. P., Shi, Y. S., Song, Y. C., Wang, G. Y., Wang, T. Y., and Li, Y. (2007). Development of EST-SSR in foxtail millet (*Setaria italica*). *Genet. Resour. Crop Evol.* 54, 233–236. doi: 10.1007/s10722-006-9139-8

Jia, X. P., Zhang, Z., Liu, Y., Zhang, C., Shi, Y., Song, Y., et al. (2009). Development and genetic mapping of SSR markers in foxtail millet [*Setaria italica* (L.) P. Beauv.]. *Theor. Appl. Genet.* 118, 821–829. doi: 10.1007/s00122-008-0942-9

Jia, G., Huang, X., Zhi, H., Zhao, Y., Zhao, Q., Li, W., et al. (2013). A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nat. Genet.* 45, 957. doi: 10.1038/ng.2673

Jiang, P., Burczynski, F., Campbell, C., Pierce, G., Austria, J. A., and Briggs, C. J. (2007). Rutin and flavonoid contents in three buckwheat species *Fagopyrum esculentum, F. tataricum*, and *F. homotropicum* and their protective effects against lipid peroxidation. *Food Res. Int.* 40, 356–364. doi: 10.1016/j.foodres.2006.10.009

Johnson, D. L., and Croissant, R. (1985). *Quinoa production in Colorado. Service in action, no. 112* (Colorado, USA: Colorado State University, Cooperative Extension, Fort Collins).

Kaur, P., Purewal, S. S., Sandhu, K. S., Maninder, K., and Salar, R. K. (2019). Millets: a cereal grain with potent antioxidants and health benefits. *Food Meas.* 13, 793–806. doi: 10.1007/s11694-018-9992-0

Kawa, J. M., Taylor, C. G., and Przybylski, R. (2003). Buckwheat concentrate reduces serum glucose in streptozotocin-diabetic rats. *J. Agric. Food Chem.* 51, 7287–7291. doi: 10.1021/jf0302153

Kayashita, J., Shimaoka, I., Nakajoh, M., Yamazaki, M., and Kato, N. (1997). Consumption of buckwheat protein lowers plasma cholesterol and raises fecal neutral sterols in cholesterol-fed rats because of its low digestibility. *J. Nutr.* 127, 1395–1400. doi: 10.1093/jn/127.7.1395

Kellogg, E. A. (2015). *Flowering plants. Monocots. Poaceae* (Springer International Publishing). Vol. 13, 416 pp. doi: 10.1007/978-3-319-15332-2

Khattab, R. Y., and Arntfield, S. D. (2009). Nutritional quality of legume seeds as affected by some physical treatments. 2. Antinutritional factors. *LWT Food Sci. Technol.* 42 (6), 1113–1118. doi: 10.1016/j.lwt.2009.02.004

Kim, S. L., Kim, S. K., and Park, C. H. (2004). Introduction and nutritional evaluation of buckwheat sprouts as a new vegetable. *Food Res. Int.* 37, 319–327. doi: 10.1016/j.foodres.2003.12.008

Kishore, G., Gupta, S., and Pandey, A. (2012). Assessment of population genetic diversity of Fagopyrum tataricum using SSR molecular marker. *Biochem. Syst. Ecol.* 43, 32–41. doi: 10.1016/j.bse.2012.02.018

Kolano, B., Siwinska, D., Pando, L. G., Szymanowska-Pulka, J., and Maluszynska, J. (2012). Genome size variation in *Chenopodium quinoa* (Chenopodiaceae). *Plant Syst. Evol.* 298, 251–255. doi: 10.1007/s00606-011-0534-z

Kong, X., Bao, J., and Corke, H. (2009). Physical properties of Amaranthus starch. *Food Chem.* 113, 371–376. doi: 10.1016/j.foodchem.2008.06.028

Konishi, T., and Ohnishi, O. (2006). A linkage map for common buckwheat based on microsatellite and AFLP markers. *Fagopyrum* 23, 1–6.

Koziol, M. J. (1990). "Composición química," in *Quinua Hacia su Cultivo Comercial, Chap. VIII*. Eds. C. Wahli (Brazil: Quito, Latinreco SA), 137–159.

Kreft, I., and Skrabanja, V. (2002). Nutritional properties of starch in buckwheat noodles. *J. Nutr. Sci. Vitaminol.* 48, 47–50. doi: 10.3177/jnsv.48.47

Kreft, I., Fabjan, N., and Yasumoto, K. (2006). Rutin content in buckwheat (*Fagopyrum esculentum* Moench) food materials and products. *Food Chem.* 98, 508–512. doi: 10.1016/j.foodchem.2005.05.081

Kreft, I. (1983). "Buckwheat breeding perspectives," in *Buckwheat Research 1983: Proceedings of the 2nd International Symposium on Buckwheat held in Miyazaki, Japan, Sept. 7–10, 1983*. Eds. T. Nagatomo and T. Adachi (Miyazaki, Japan: Kuroda-Toshado Printing Co., Japan), 3–12.

Krishnan, R., Dharmaraj, U., and Malleshi, N. G. (2012). Influence of decortication, popping and malting on bioaccessibility of calcium, iron and zinc in finger millet. *LWT - Food Sci. Technol.* 48 (2), 169–174. doi: 10.1016/j.lwt.2012.03.003

Krkošková, B., and Mrazova, Z. (2005). Prophylactic components of buckwheat. *Food Res. Int.* 38, 561–568. doi: 10.1016/j.foodres.2004.11.009

Kumar, A., Gaur, V. S., Goel, A., and Gupta, A. K. (2015a). De novo assembly and characterization of developing spikes transcriptome of finger millet (*Eleusine coracana*): a minor crop having nutraceutical properties. *Plant Mol. Biol. Rep.* 33, 905–922. doi: 10.1007/s11105-014-0802-5

Kumar, A., Yadav, S., Panwar, P., Gaur, V. S., and Sood, S. (2015b). Identification of anchored simple sequenceF repeat markers associated with calcium content

in finger millet (*Eleusine coracana*). *Proc. Natl. Acad. Sci.* 85, 311–317. doi: 10.1007/s40011-013-0296-1

Kumar, A., Sharma, D., Tiwari, A., Jaiswal, J. P., Singh, N. K., and Sood, S. (2016). Genotyping-by-sequencing analysis for determining population structure of finger millet germplasm of diverse origins. *Plant Genome* 9, 1–15. doi: 10.3835/plantgenome2015.07.0058

Kumar, A., Tomer, V., Kaur, A., Kumar, V., and Gupta, K. (2018). Millets: a solution to agrarian and nutritional challenges. *Agric. Food Secur.* 7, 31. doi: 10.1186/s40066-018-0183-3

Kumari, S. K., and Thayumanavan, B. (1997). Comparative study of resistant starch from minor millets on intestinal responses, blood glucose, serum cholesterol and triglycerides in rats. *J. Sci. Food Agric.* 75, 296–302. doi: 10.1002/(SICI)1097-0010(199711)75:3<296::AID-JSFA877>3.0.CO;2-X

Kumpun, S., Maria, A., Crouzet, S., Evrard-Todeschi, N., Girault, J.-P., and Lafont, R. (2011). Ecdysteroids from *Chenopodium quinoa* Willd., an ancient Andean crop of high nutritional value. *Food Chem.* 125, 1226–1234. doi: 10.1016/j.foodchem.2010.10.039

Lakshmi, B., and Vimala, V. (2000). Nutritive value of dehydrated green leafy vegetable powders. *J. Food Sci. Technol.* 37, 465–471.

Lata, C., and Prasad, M. (2011). Role of DREBs in regulation of abiotic stress responses in plants. *J. Exp. Bot.* 62, 4731–4748. doi: 10.1093/jxb/err210

Lata, C., and Shivhare, R. (2017). "Genetic Determinants of Abiotic Stress Tolerance in Foxtail Millet," in *The Foxtail Millet Genome* (Springer), 85–104.

Lata, C., Sahu, P. P., and Prasad, M. (2010). Comparative transcriptome analysis of differentially expressed genes in foxtail millet (*Setaria italica* L.) during dehydration stress. *Biochem. Biophys. Res. Commun.* 393, 720–727. doi: 10.1016/j.bbrc.2010.02.068

Lata, C., Jha, S., Dixit, V., Sreenivasulu, N., and Prasad, M. (2011). Differential antioxidative responses to dehydration-induced oxidative stress in core set of foxtail millet cultivars [*Setaria italica* (L.)]. *Protoplasma* 248, 817–828. doi: 10.1007/s00709-010-0257-y

Lazarte, C. E., Carlsson, N. G., Almgren, A., Sandberg, A.-S., and Granfeldt, Y. (2015). Phytate, zinc, iron and calcium content of common Bolivian food, and implications for mineral bioavailability. *J. Food Compos. Anal.* 39, 111–119. doi: 10.1016/j.jfca.2014.11.015

León-Camacho, M., García-González, D. L., and Aparicio, R. (2001). A detailed and comprehensive study of amaranth (*Amaranthus cruentus* L.) oil fatty profile. *Eur. Food Res. Technol.* 213, 349–355. doi: 10.1007/s002170100340

Li, S., and Zhang, Q. H. (2001). Advances in the development of functional foods from buckwheat. *Crit. Rev. In Food Sci. Nutr.* 41, 451–464. doi: 10.1080/20014091091887

Li, Y. Q., Shi, T. L., and Zhang, Z. W. (2007). Development of microsatellite markers from tartary buckwheat. *Biotechnol. Lett.* 29, 823–827. doi: 10.1007/s10529-006-9293-2

Lightfoot, D. J., Jarvis, D. E., Ramaraj, T., Lee, R., Jellen, E. N., and Maughan, P. J. (2017). Single-molecule sequencing and Hi-C-based proximity-guided assembly of amaranth (*Amaranthus hypochondriacus*) chromosomes provide insights into genome evolution. *BMC Biol.* 15, 74. doi: 10.1186/s12915-017-0412-4

Lin, H. S., Chiang, C. Y., Chang, S. B., and Kuoh, C. S. (2011). Development of simple sequence repeat (SSR) markers in *Setaria italica* (Poaceae) and cross-amplification in related species. *Int. J. Mol. Sci.* 12, 7835–7845. doi: 10.3390/ijms12117835

Lin, H. S. (2012). Genetic diversity in the foxtail millet (*Setaria italica*) germplasm as determined by agronomic traits and microsatellite markers. *Aust. J. Crop Sci.* 6, 342–349.

Lintschinger, J., Fuchs, N., Moser, H., Jäger, R., Hlebeina, T., Markolin, G., et al. (1997). Uptake of various trace elements during germination of wheat, buckwheat and quinoa. *Plant Foods Hum. Nutr.* 50, 223–237. doi: 10.1007/BF02436059

Logacheva, M. D., Kasianov, A. S., Vinogradov, D. V., Samigullin, T. H., Gelfand, M. S., Makeev, V. J., et al. (2011). De novo sequencing and characterization of floral transcriptome in two species of buckwheat (*Fagopyrum*). *BMC Genomics* 12, 30. doi: 10.1186/1471-2164-12-30

Longvah, T. (2017). *Indian Food Composition Tables* (National Institute of Nutrition).

Lopes, C. O., Dessimoni, G. V., Da Silva, M. C., Vieira, G., and Pinto, N. A. V. D. (2009). Aproveitamento, composição nutricional e antinutricional da farinha de quinoa (*Chenopodium quinoa*). *Alime. Nutr.* 20 (4), 669–675.

Ma, K. H., Kim, N. S., Lee, G. A., Lee, S. Y., Lee, J. K., Yi, J. Y., et al. (2009). Development of SSR markers for studies of diversity in the genus *Fagopyrum*. *Theor. Appl. Genet.* 119, 1247–1254. doi: 10.1007/s00122-009-1129-8

Mabhaudhi, T., Chimonyo, V. G. P., Hlahla, S., Massawe, F., Mayes, S., Nhamo, L., et al. (2019). Prospects of orphan crops in climate change. *Planta*, 250, 695–708. doi: 10.1007/s00425-019-03129-y

Madhusoodanan, K. J., and Pal, M. (1981). Cytology of vegetable amaranths. *Bot. J. Linn. Soc.* 82 (1), 61–68. doi: 10.1111/j.1095-8339.1981.tb00950.x

Mahoney, A. W., Lopez, J. G., and Hendricks, D. G. (1975). An evaluation of the protein quality of quinoa. *J. Agric. Food Chem.* 23, 90. doi: 10.1021/jf60198a035

Mal, B., Padulosi, S., and Ravi, S. B. (2010). *Minor millets in South Asia: learnings from IFAD-NUS Project in India and Nepal* (Chennai, India: Bioversity International, Maccarese, Rome, Italy, and the M.S. Swaminathan Research Foundation), 1–176.

Mallory, M. A., Hall, R. V., McNabb, A. R., Pratt, D. B., Jellen, E. N., and Maughan, P. J. (2008). Development and characterization of microsatellite markers for the grain amaranths. *Crop Sci.* 48, 1098–1106. doi: 10.2135/cropsci2007.08.0457

Maradini, A. M., Ribeiro, M., Da Silva, J. T., Pinheiro, H. M., Paes, J. B., and Dos Reis, J. S. (2017). Quinoa: Nutritional, Functional and Antinutritional Aspects. *Crit. Rev. In Food Sci. Nutr.* 57 (8), 1618–1630. doi: 10.1080/10408398.2014.1001811

Martínez, E. A., Jorquera-Jaramillo, C., Veas, E., and Chia, E. (2009). El futuro de la quínoa en la región árida de Coquimbo: Lecciones y escenarios a partir de una investigación sobre su biodiversidad en Chile para la acción con agricultores locales. *Rev. Geogr. Valparaíso*, 42, 95–111.

Martínez-Cruz, O., Cabrera-Chávez, F., and Paredes-López, O. (2014). Biochemical characteristics, and nutraceutical and technological uses of amaranth globulins. In: Eds. S. D. Milford and S. D. Globulins *Biochemistry, Production and Role in Immunity*. (Nova Science Publishers, Inc), 41–70.

Mason, S. L., Stevens, M. R., Jellen, E. N., Bonifacio, A., Fairbanks, D. J., Coleman, C. E., et al. (2005). Development and use of microsatellite markers for germplasm characterization in quinoa (*Chenopodium quinoa* Willd.). *Crop Sci.* 45, 1618–1630. doi: 10.2135/cropsci2004.0295

Mastebroek, H. D., Limburg, H., Gilles, T., and Marvin, H. J. P. (2000). Occurrence of sapogenins in leaves and seeds of quinoa (*Chenopodium quinoa* Willd). *J. Sci. Food Agric.* 80, 152–156. doi: 10.1002/(SICI)1097-0010(20000101)80:1<152::AID-JSFA503>3.0.CO;2-P

Matsui, K., Kiryu, Y., Komatsuda, T., Kurauchi, N., Ohtani, T., and Tetsuka, T. (2004). Identification of AFLP makers linked to non-seed shattering locus (sht1) in buckwheat and conversion to STS markers for marker-assisted selection. *Genome* 47, 469–474. doi: 10.1139/g04-007

Maughan, P. J., Bonifacio, A., Jellen, E. N., Stevens, M. R., Coleman, C. E., Ricks, M., et al. (2004). A genetic linkage map of quinoa (*Chenopodium quinoa*) based on AFLP, RAPD, and SSR markers. *Theor. Appl. Genet.* 109, 1188–1195. doi: 10.1007/s00122-004-1730-9

Maughan, P. J., Turner, T. B., Coleman, C. E., Elzinga, D. B., Jellen, E. N., Morales, J. A., et al. (2009). Characterization of *Salt Overly Sensitive 1* (*SOS1*) gene homoeologs in quinoa (*Chenopodium quinoa* Willd.). *Genome* 52, 647–657. doi: 10.1139/G09-041

Maughan, P. J., Smith, S. M., Rojas-Beltran, J. A., Elzinga, D., Raney, J. A., Jellen, E. N., et al. (2012). Single nucleotide polymorphism identification, characterization, and linkage mapping in quinoa. *Plant Genome* 5, 114–125. doi: 10.3835/plantgenome2012.06.0011

Mbithi-Mwikya, S., Ooghe, W., Van Camp, J., Ngundi, D., and Huyghebaert, A. (2000). Amino acid profiles after sprouting, autoclaving, and lactic acid fermentation of finger millet (*Eleusine coracana*) and kidney beans (*Phaseolus vulgaris* L.). *J. Agric. Food Chem.* 48, 3081–3085. doi: 10.1021/jf0002140

Mendonça, S., Saldiva, P. H., Cruz, R. J., and Arêas, J. A. (2009). Amaranth protein presents cholesterol-lowering effect. *Food Chem.* 116, 738–742. doi: 10.1016/j.foodchem.2009.03.021

Mgonja, M. A., Lenne, J. M., Manyasa, E., and Sreenivasaprasad, S. (2007). Finger millet blast management in East Africa: creating opportunities for improving production and utilization of finger millet. *Int. Crops Res. Inst. Semi-Arid Trop.* 196.

Miranda, M., Vega-Gálvez, A., Quispe-Fuentes, I., Rodríguez, M. J., Maureira, H., and Martínez, E. A. (2012). Nutritional aspects of six quinoa (*Chenopodium*

quinoa Willd.) ecotypes from three geographical areas of Chile. *Chil. J. Agric. Res.* 72, 175.

Mirza, N., Taj, G., Arora, S., and Kumar, A. (2014). Transcriptional expression analysis of genes involved in regulation of calcium translocation and storage in finger millet *(Eleusine coracana* L. Gartn.). *Gene* 550, 171–179. doi: 10.1016/j.gene.2014.08.005

Mishra, G., Joshi, D. C., and Panda, B. K. (2014). Popping and puffing of cereal grains: a review. *J. Grain Process. Storage* 1 (2), 34–46.

Mitsunaga, T., Matsuda, M., Shimizu, M., and Iwashima, A. (1986). Isolation and properties of thiamine-binding protein from bckwheat seed. *Cereal Chem.* 4, 332–335.

Mlakar, S. G., Turinek, M., Jakop, M., Bavec, M., and Bavec, F. (2009). Nutrition value and use of grain amaranth: potential future application in bread making. *Agricultura* 6: 43–53.

Mohamed, T. K., Zhu, K., Issoufou, A., Fatmata, T., and Zhou, H. (2009). Functionality, *in vitro* digestibility and physicochemical properties of two varieties of defatted foxtail millet protein concentrates. *Int. J. Mol. Sci.* 10, 5224–5238. doi: 10.3390/ijms10125224

Monteros, C. J., Nieto, C. C., Caicedo, C. V., Rivera, M. M., and Vimos, C. N. (1994). INIAP-Alegría, primera variedad mejorada de amaranto para la sierra ecuatoriana. Boletin Divulgativo No. 245, Ecuador

Morales, A., Zurita-Silva, A., Maldonado, J., and Silva, H. (2017). Transcriptional responses of Chilean quinoa (*Chenopodium quinoa* Willd.) under water deficit conditions uncovers ABA-independent expression patterns. *Front. Plant Sci.* 8, 216. doi: 10.3389/fpls.2017.00216

Mosyakin, S. L., and Robertson, K. R. (1996). New infrageneric taxa and combinations in *Amaranthus* (Amaranthaceae). *Ann. Botan. Fenn. (JSTOR)*, 33, 275–281.

Mounika, M., Devi, K. U., and Devi, S. S. (2017). Bioavailability of iron from foxtail millet and sorghum millet recipes. *Int. J. Agric. Sci. Res. (IJASR)* 7 (4), 703–708. doi: 10.20546/ijcmas.2017.610.332

Mujica, A., and Jacobsen, S. E. (2003). The genetic resources of Andean grain amaranths (*Amaranthus caudatus* L., *A. cruentus* L. and *A. hypochondriacus* L.) in America. *Plant Genet. Resour. Newslett.* 133, 41–44.

Musia, G. D. (2013). *Identification of microsatellite markers for finger millet [Eleusine coracana (L.) Gaertn.] by analysis of Roche 454 GS-FLX Titanium sequence data. Dissertation/master's thesis* (Nairobi, Kenya: Kenyatta University).

Muthamilarasan, M., and Prasad, M. (2015). Advances in *Setaria* genomics for genetic improvement of cereals and bioenergy grasses. *Theor. Appl. Genet.* 128, 1–14. doi: 10.1007/s00122-014-2399-3

Muthamilarasan, M., Venkata Suresh, B., Pandey, G., Kumari, K., Parida, S. K., and Prasad, M. (2013). Development of 5123 intron-length polymorphic markers for large-scale genotyping applications in foxtail millet. *DNA Res.* 21, 41–52. doi: 10.1093/dnares/dst039

Muza, F. R., Lee, D. J., Andrews, D. J., and Gupta, S. C. (1995). Mitochondrial DNA variation in finger millet (*Eleusine coracana* L. Gaertn). *Euphytica* 81, 199–205. doi: 10.1007/BF00025434

Mysore, K. S., and Baird, V. (1997). Nuclear DNA content in species of *Eleusine* (Gramineae): a critical re-evaluation using laser flow cytometry. *Plant Syst. Evol.* 207, 1–11. doi: 10.1007/BF00985206

Nègre-Salvayre, A., Affany, A., Hariton, C., and Salvayre, R. (1991). Additional antilipoperoxidant activities of alpha-tocopherol and ascorbic acid on membrane-like systems are potentiated by rutin. *Pharmacology* 42, 262–272. doi: 10.1159/000138807

Nagano, M., Aii, J., Campbell, C., Kawasaki, S., and Adachi, T. (2000). Genome size analysis of the genus *Fagopyrum*. *Fagopyrum* 17, 9.

Nagano, M., Aii, J., Kuroda, M., Campbell, C., and Adachi, T. (2001). Conversion of AFLP markers linked to the *Sh* allele at the *S locus* in buckwheat to a simple PCR based marker form. *Plant Biotechnol.* 18, 191–196. doi: 10.5511/plantbiotechnology.18.191

Nagaprabha, P., and Bhattacharya, S. (2016). Textural characterization of foxtail millet gels: effect of cations and hydrocolloids. *J. Food Sci. Technol.* 53 (1), 257–268. doi: 10.1007/s13197-015-2046-2

Nagasawa, T., Tabata, N., Ito, Y., Aiba, Y., Nishizawa, N., and Kitts, D. D. (2003). Dietary G-Rutin suppresses glycation in tissue proteins of streptozotocin-induced diabetic rats. *Mol. Cell. Biochem.* 252, 141–147. doi: 10.1023/A:1025563519088

Nanduri, K. R., Hirich, A., Salehi, M., Saadat, S., and Jacobsen, S. E. (2019). "Quinoa: a new crop for harsh environments," in *Sabkha Ecosystems* (Springer), 301–333.

National Research Council. (1996). "Lost crops of Africa," in *Grains. Board on Science and Technology for International Development*, vol. I. (Washington, D.C: National Academy Press).

Nirgude, M., Babu, B. K., Shambhavi, Y., Singh, U. M., Upadhyaya, H. D., and Kumar, A. (2014). Development and molecular characterization of genic molecular markers for grain protein and calcium content in finger millet (*Eleusine coracana* (L.) Gaertn.). *Mol. Biol. Rep.* 41, 1189–1200. doi: 10.1007/s11033-013-2825-7

Nirmala, M., Subba Rao, M.V.S.S.T., and Muralikrishna, G. (2000). Carbohydrates and their degrading enzymes from native and malted finger millet (Ragi, *Eleusine coracana*, Indaf-15). *Food Chem.* 69 (2), 175–180. doi: 10.1016/S0308-8146(99)00250-2

Nusslein, K., Parkash Dhankher, O., Xing, B., Smith-Doerr, L., Sacco, T., Maathuis, F., et al. (2016). Project management: food security needs social-science input. *Nature* 535, 37–37. doi: 10.1038/535037d

O' Brien, G. K., and Price, M. L. (2008). Amaranth: grain and vegetable types. *Echo Tech. Note*. 1–15.

Odetti, P. R., Borgoglio, A., De Pascale, A., Rolandi, R., and Adezati, L. (1990). Prevention of diabetes-increased aging effect on rat collagen-linked fluorescence by aminoguanidine and rutin. *Diabetes* 39, 796. doi: 10.2337/diab.39.7.796

Oghbaei, M., and Prakash, J. (2012). Bioaccessible nutrients and bioactive components from fortified products prepared using finger millet (*Eleusine coracana*). *J. Sci. Food Agric.* 92, 2281–2290. doi: 10.1002/jsfa.5622

Ogungbenle, H. N. (2003). Nutritional evaluation and functional properties of quinoa (*Chenopodium quinoa*) flour. *Int. J. Food Sci. Nutr.* 54 (2), 153–158. doi: 10.1080/0963748031000084106

Ohnishi, O. (1993). Population genetics of cultivated common buckwheat, *Fagopyrum esculentum* Moench. VIII. Local differentiation of land races in Europe and the silk road. *Japanese J. Genet.* 68, 303–316. doi: 10.1266/jjg.68.303

Oliveira, A. C., Reis, S. M. P. M., Carvalho, E. M., Pimenta, F. M. V., Rios, K. R., Paiva, K. C., et al. (2003). Adições crescentes de ácido fitico à dieta não interferiram na digestibilidade da caseína e no ganho de peso em ratos. *Rev. Nutr.* 16 (2), 211–217. doi: 10.1590/S1415-52732003000200008

Ookubo, K. (1992). "Nutrition and functionality of soybean," in *Science of Soybean*. Eds. F. Yamauchi and K. Ookubo (Tokyo: Asakura-Shoten Press), 57–75.

Oomah, B. D., and Mazza, G. (1996). Flavonoids and antioxidative activities in buckwheat. *J. Agric. Food Chem.* 44 (7), 1746–1750. doi: 10.1021/jf9508357

Paśko, P., Bartoń, H., Zagrodzki, P., Gorinstein, S., Fołta, M., and Zachwieja, Z. (2009). Anthocyanins, total polyphenols and antioxidant activity in amaranth and quinoa seeds and sprouts during their growth. *Food Chem.* 115, 994–998. doi: 10.1016/j.foodchem.2009.01.037

Pal, M., and Khoshoo, T. N. (1972). Evolution and improvement of cultivated amaranths: V. Inviability, weakness, and sterility in hybrids. *J. Hered.* 63 (2), 78–82. doi: 10.1093/oxfordjournals.jhered.a108234

Pal, J., Singhal, R. S., and Kulkarni, P. R. (2002). Physicochemical properties of hydroxypropyl derivative from corn and amaranth starch. *Carbohydr. Polym.* 48, 49–53. doi: 10.1016/S0144-8617(01)00209-0

Palomino, G., Hernández, L. T., and de la Cruz Torres, E. (2008). Nuclear genome size and chromosome analysis in *Chenopodium quinoa* and *C. berlandieri* subsp. *nuttalliae*. *Euphytica* 164, 221. doi: 10.1007/s10681-008

Pan, S. J., and Chen, Q. F. (2010). Genetic mapping of common buckwheat using DNA, protein and morphological markers. *Hereditas* 147, 27–33. doi: 10.1111/j.1601-5223.2009.02116.x

Pandey, G., Misra, G., Kumari, K., Gupta, S., Parida, S. K., Chattopadhyay, D., et al. (2013). Genome-wide development and use of microsatellite markers for large-scale genotyping applications in foxtail millet [*Setaria italica* (L.)]. *DNA Res.* 20, 197–207. doi: 10.1093/dnares/dst002

Panwal, J. H., and Pawar, V. D. (1989). Effect of soaking pearl millet in acid on the bleaching and functional properties of flours. *J. Food Sci. Technol.* 26 (2), 79–82.

Paredes-Lopez, O., and Mora-Escobedo, R. (1989). Germination of amaranth seeds: effects on nutrient composition and color. *J. Food Sci.* 54, 761–762. doi: 10.1111/j.1365-2621.1989.tb04702.x

Park, S.-S., and Ohba, H. (2004). Suppressive activity of protease inhibitors from buckwheat seeds against human T-acute lymphoblastic leukemia cell lines. *Appl. Biochem. Biotechnol.* 117, 65. doi: 10.1385/ABAB:117:2:065

Pawar, V. D., and Machewad, G. M. (2006). Processing of foxtail millet for improved nutrient availability. *J. Food Process. Preserv.* 30 (3), 269–279. doi: 10.1111/j.17454549.2006.00064.x

Pawar, V. D., and Parlikar, G. S. (1990). Reducing the polyphenols and phytate and improving the protein quality of pearl millet by dehulling and soaking. *J. Food Sci. Technol.* 27 (3), 140–143. doi: 10.4236/fns.2011.28122

Perales-Sánchez, J. X., Reyes-Moreno, C., Gómez-Favela, M. A., Milán-Carrillo, J., Cuevas-Rodríguez, E. O., Valdez-Ortiz, A., et al. (2014). Increasing the antioxidant activity, total phenolic and flavonoid contents by optimizing the germination conditions of amaranth seeds. *Plant Foods Hum. Nutr.* 69, 196–202. doi: 10.1007/s11130-014-0430-0

Peterson, A., and Murphy, K. (2015). Tolerance of lowland quinoa cultivars to sodium chloride and sodium sulfate salinity. *Crop Sci.* 55, 331–338. doi: 10.2135/cropsci2014.04.0271

Pires, J. L. (2017). *Avaliação do Comportamento Agronómico da Quinoa (Chenopodium quinoa Willd.), em Diferentes Regimes Hídricos e Níveis de Fertilização Azotada, nas Condições Agroecológicas de Trás-os-Montes. Master"s Thesis* (Bragança, Portugal: Instituto Politecnico de Bracanca Escola Superior Agraria).

Plate, A. Y., and Arêas, J. A. (2002). Cholesterol-lowering effect of extruded amaranth (*Amaranthus caudatus* L.) in hypercholesterolemic rabbits. *Food Chem.* 76, 1–6. doi: 10.1016/S0308-8146(01)00238-2

Platel, K., and Srinivasan, K. (2016). Bioavailability of Micronutrients from Plant Foods: An Update. *Crit. Rev. in Food Sci. Nutr.* 56 (10), 1608–1619. doi: 10.1080/10408398.2013.781011

Platel, K., Eipeson, S. W., and Srinivasan, K. (2010). Bioaccessible Mineral Content of Malted Finger Millet (*Eleusine coracana*), Wheat (*Triticum aestivum*), and Barley (*Hordeum vulgare*). *J. Agric. Food Chem.* 58 (13), 8100–8103. doi: 10.1021/jf100846e

Pomeranz, Y., and Robbins, G. S. (1972). Amino acid composition of buckwheat. *J. Agric. Food Chem.* 20 (2), 270–274. doi: 10.1021/jf60180a029

Popa, G., Cornea, C. P., Ciuca, M., Babeanu, N., Popa, O., and Marin, D. (2010). Studies on genetic diversity in *Amaranthus* species using the RAPD markers. *Tom* 17 (2), 280–285.

Porter, J. R., Xie, L., Challinor, A. J., Cochrane, K., Howden, S. M., Iqbal, M. M., et al. (2014). *Food security and food production systems in: Climate Change 2014: Impacts, Adaptation, and Vulnerability. Part A: Global and Sectoral Aspects. Contribution of Working Group II to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. [Eds. C. B. Field, V. R. Barros, D. J. Dokken, K. J. Mach, M. D. Mastrandrea, T. E. Bilir, M. Chatterjee, K. L. Ebi, Y. O. Estrada, R. C. Genova, B. Girma, E. S. Kissel, A. N. Levy, S. MacCracken, P. R. Mastrandrea and L. L. White]. (Cambridge, United Kingdom and New York, NY, USA: Cambridge University Press), 485–533.

Powell, W., Morgante, M., Andre, C., Hanafey, M., Vogel, J., Tingey, S., et al. (1996). The comparison of RFLP, RAPD, AFLP and SSR (microsatellite) markers for germplasm analysis. *Mol. Breed.* 2, 225–238. doi: 10.1007/BF00564200

Präger, A., Munz, S., Nkebiwe, P., Mast, B., and Graeff-Hönninger, S. (2018). Yield and quality characteristics of different quinoa (*Chenopodium quinoa* Willd.) cultivars grown under field conditions in Southwestern Germany. *Agronomy* 8, 197. doi: 10.3390/agronomy8100197

Praznik, W., Mundigler, N., Kogler, A., Pelzl, B., Huber, A., and Wollendorfer, M. (1999). Molecular background of technological properties of selected starches. *Starch-Stärke* 51, 197–211. doi: 10.1002/(SICI)1521-379X(199906)51:6<197::AID-STAR197>3.0.CO;2-K

Pu, F., Mishima, K., Irie., Motohashi, K., Tanaka, Y., Orito, K., et al. (2007). Neuroprotective effects of quercetin and rutin on spatial memory impairment in an 8-arm radial maze task and neuronal death induced by repeated cerebral ischemia in rats. *J. Pharmacol. Sci.* 104 (4), 329–334. doi: 10.1254/jphs.FP0070247

Pulvento, C., Riccardi, M., Lavini, A., d'Andria, R., Iafelice, G., and Marconi, E. (2010). Field trial evaluation of two *Chenopodium quinoa* genotypes grown under rain-fed conditions in a typical Mediterranean environment in South Italy. *J. Agron. Crop Sci.* 196, 407–411. doi: 10.1111/j.1439-037X.2010.00431.x

Puranik, S., Jha, S., Srivastava, P. S., Sreenivasulu, N., and Prasad, M. (2011). Comparative transcriptome analysis of contrasting foxtail millet cultivars in response to short-term salinity stress. *J. Plant Physiol.* 168, 280–287. doi: 10.1016/j.jplph.2010.07.005

Puranik, S., Sahu, P. P., Mandal, S. N., Parida, S. K., and Prasad, M. (2013). Comprehensive genome-wide survey, genomic constitution and expression profiling of the NAC transcription factor family in foxtail millet (*Setaria italica* L.). *PloS One* 8, e64594. doi: 10.1371/journal.pone.0064594

Qi, X., Xie, S., Liu, Y., Yi, F., and Yu, J. (2013). Genome-wide annotation of genes and noncoding RNAs of foxtail millet in response to simulated drought stress by deep sequencing. *Plant Mol. Biol.* 83, 459–473. doi: 10.1007/s11103-013-0104-6

Qian, J., and Kuhn, M. (1999). Physical properties of buckwheat starches from various origins. *Starch-Stärke* 51, 81–85. doi: 10.1002/(SICI)1521-379X(199903)51:2<81::AID-STAR81>3.0.CO;2-I

Rahman, H., Jagadeeshselvam, N., Valarmathi, R., Sachin, B., Sasikala, R., Senthil, N., et al. (2014). Transcriptome analysis of salinity responsiveness in contrasting genotypes of finger millet (*Eleusine coracana* L.) through RNA-sequencing. *Plant Mol. Biol.* 85, 485–503. doi: 10.1007/s11103-014-0199-4

Raina, A. P., and Gupta, V. (2015). Evaluation of buckwheat (*Fagopyrum species*) germplasm for rutin content in seeds. *Indian J. Plant Physiol.* 20, 167–171. doi: 10.1007/s40502-015-0147-6

Ramachandra, G., Virupaksha, T. K., and Shadaksharaswamy, M. (1977). Relation between tannin levels and *in vitro* protein digestibility in finger millet (*Eleusine coracana* Gaertn.). *J. Agric. Food Chem.* 25, 1101–1104. doi: 10.1021/jf60213a046

Raney, J. A. (2012). Transcriptome analysis of drought induced stress in *Chenopodium quinoa*. *Am. J. Plant Sci.* 05 (03), 338–357. doi: 10.4236/ajps.2014.53047

Rao, B., Nagasampige, M. K., and Ravikiran, M. (2011). Evaluation of nutraceutical properties of selected small millets. *J. Pharm. BioAllied Sci.* 3 (2), 277–279. doi: 10.4103/0975-7406.80775

Rastogi, A., and Shukla, S. (2013). Amaranth: a new millennium crop of nutraceutical values. *Crit. Rev. in Food Sci. Nutr.* 53, 109–125. doi: 10.1080/10408398.2010.517876

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

Reddy, I. N. B. L., Reddy, D. S., Narasu, M. L., and Sivaramakrishnan, S. (2011). Characterization of disease resistance gene homologues isolated from finger millet (*Eleusine coracana* L. Gaertn). *Mol. Breed.* 27, 315–328. doi: 10.1007/s11032-010-9433-1

Reddy, I. N. B. L., Narasu, M. L., and Sivaramakrishnan, S. (2012). Identification and characterization of EST-SSRs in finger millet (*Eleusine coracana* (L.) Gaertn.). *J. Crop Sci. Biotechnol.* 15, 9–16. doi: 10.1007/s12892-011-0064-9

Ren, X., Chen, J., Molla, M. M., Wang, C., Diao, X., and Shen, Q. (2016). In vitro starch digestibility and *in vivo* glycemic response of foxtail millet and its products. *Food Funct.* 7 (1), 372–379. doi: 10.1039/C5FO01074H

Repo-Carrasco, R., Espinoza, C., and Jacobsen, S.-E. (2003). Nutritional value and use of the Andean crops quinoa (*Chenopodium quinoa*) and kañiwa (*Chenopodium pallidicaule*). *Food Rev. Int.* 19, 179–189. doi: 10.1081/FRI-120018884

Repo-Carrasco-Valencia, R. A., Encina, C. R., Binaghi, M. J., Greco, C. B., and Ronayne de Ferrer, P. A. (2010). Effects of roasting and boiling of quinoa, kiwicha and kañiwa on composition and availability of minerals *in vitro*. *J. Sci. Food Agric.* 90, 2068–2073. doi: 10.1002/jsfa.4053

Rodríguez, L. A., and Isla, M. T. (2009). Comparative analysis of genetic and morphologic diversity among quinoa accessions (*Chenopodium quinoa* Willd.) of the South of Chile and highland accessions. *J. Plant Breed. Crop Sci.* 1, 210–216.

Rojas, W., Barriga, P., and Figueroa, H. (2000). Multivariate analysis of the genetic diversity of Bolivian quinua germplasm. *Plant Genet. Resour. Newslett.* 122, 16–23.

Rojas, W., Pinto, M., Soto, J. L., and Alcocer, E. (2010). "Valor Nutricional, Agroindustrial y Funcional de Los Granos Andinos," in *Granos Andinos: Avances, Logros y Experiencias Desarrolladas en Quinua, Cañahua y Amaranto en Bolivia* (Rome, Italy: Bioversity International), 151–164.

Ruales, J., and Nair, B. M. (1993a). Content of fat, vitamins and minerals in quinoa (*Chenopodium quinoa*, Willd) seeds. *Food Chem.* 48, 131–136. doi: 10.1016/0308-8146(93)90047-J

Ruales, J., and Nair, B. M. (1993b). Saponins, phytic acid, tannis and protease inhibitors in quinoa (*Chenopodium quinoa* Willd) seeds. *Food Chem.* 48 (2), 137–143. doi: 10.1016/0308-8146(93)90048-K

Ruas, P. M., Bonifacio, A., Ruas, C. F., Fairbanks, D. J., and Andersen, W. R. (1999). Genetic relationship among 19 accessions of six species of *Chenopodium* L. by random amplified polymorphic DNA fragments (RAPD). *Euphytica* 105, 25–32. doi: 10.1023/A:1003480414735

Ruiz, K. B., Khakimov, B., Søren, B. E., Søren, B., Stefania, B., and Sven-Erik, J. (2017). Quinoa seed coats as an expanding and sustainable source of bioactive compounds: an investigation of genotypic diversity in saponin profiles. *Ind. Crops Prod.* 104, 156–163. doi: 10.1016/j.indcrop.2017.04.007

Ruiz-Carrasco, K., Antognoni, F., Coulibaly, A. K., Lizardi, S., Covarrubias, A., Martínez, E. A., et al. (2011). Variation in salinity tolerance of four lowland genotypes of quinoa (*Chenopodium quinoa* Willd.) as assessed by growth, physiological traits, and sodium transporter gene expression. *Plant Physiol. Biochem.* 49, 1333–1341. doi: 10.1016/j.plaphy.2011.08.005

Saha, D., Gowda, M. C., Arya, L., Verma, M., and Bansal, K. C. (2016). Genetic and genomic resources of small millets. *Crit. Rev. In Plant Sci.* 35, 56–79. doi: 10.1080/07352689.2016.1147907

Saleh, A. S. M., Zhang, Q., Chen, J., and Shen, Q. (2013). Millet grains: nutritional quality, processing, and potential health benefits. *Compr. Rev. In Food Sci. Food Saf.* 12 (3), 281–295. https://onlinelibrary.wiley.com/doi/abs/10.1111/1541-4337.12012.

Sauer, J. (1955). Revision of the dioecious amaranths. *Madroño* 13, 5–46.

Sauer, J. D. (1967). The grain amaranths and their relatives: a revised taxonomic and geographic survey. *Ann. Missouri Bot. Garden* 54, 103–137. doi: 10.2307/2394998

Saunders, R. M., and Becker, R. (1984). Amaranthus: a potential food and feed resource. *Adv. In Cereal Sci. Technol. (USA)*, 6, 357–396.

Schlick, G., and Bubenheim, D. L. (1996). Quinoa: Candidate crop for NASA's Controlled Ecological Life Support Systems. 632–640 In: Ed. J. Janick, *Janick Progress in new crops*. (Arlington, VA: ASHS Press).

Schontz, D., and Rether, B. (1999). Genetic variability in foxtail millet, *Setaria italica* (L.) P. Beauv.: identification and classification of lines with RAPD markers. *Plant Breed.* 118, 190–192. doi: 10.1046/j.1439-0523.1999.118002190.x

Seenappa, M. (1988). "Sorghum and Millets in East Africa with Reference to Their Use in Weaning Foods," in *Proceedings of a Workshop Held in Nairobi, Kenya, 12-16 October 1987, 39–54. Proceedings Series/IDRC* (Nairobi, Kenya: International Development Research Centre, IDRC).

Selvam, J. N., Muthukumar, M., Rahman, H., Senthil, N., and Raveendran, M. (2015). Development and validation of SSR markers in finger millet (*Eleusine coracana* Gaertn). *Int. J. Trop. Agric.* 33, 2055–2066.

Senft, J. P. (1979). "Protein quality of amaranth grain," in *Proceeding of second amaranth conference* (Emmaus, PA: Rodale Press), 43–47.

Serna-Saldivar, S. O., Gutiérrez-Uribe, J. A., and García-Lara, S. (2015). "Phytochemical Profiles and Nutraceutical Properties of Corn and Wheat Tortillas," in *Tortillas*. Eds. L. W. Rooney and S. O. Serna-Saldivar (Elsevier Inc: (AACC) International Press), 65–96. doi: 10.1016/B978-1-891127-88-5.50003-7

Sharma, T., and Jana, S. (2002a). Species relationships in *Fagopyrum* revealed by PCR-based DNA fingerprinting. *Theor. Appl. Genet.* 105, 306–312. doi: 10.1007/s00122-002-0938-9

Sharma, T. R., and Jana, S. (2002b). Random amplified polymorphic DNA (RAPD) variation in *Fagopyrum tataricum* Gaertn. accessions from China and the Himalayan region. *Euphytica* 127, 327–333.

Shejawale, D. D., Hymavathi, T. V., Manorama, K., and Zabeen, F. (2016). Effect of processing on nutraceutical properties of foxtail millet (*Setaria italica*) varieties grown in India. *J. Food Meas. Charact.* 10 (1), 16–23. doi: 10.1007/s11694-015-9271-2

Sheu, J. R., Hsiao, G., Chou, P. H., Shen, M. Y., and Chou, D. S. (2004). Mechanisms involved in the antiplatelet activity of rutin, a glycoside of the flavonol quercetin, in human platelets. *J. Agric. Food Chem.* 52 (14), 4414–4418. doi: 10.1021/jf040059f

Shi, Z., Han, C., and Huang, K. (2011). Research on Se content of different tartary buckwheat genotypes. *Agric. Sci. Technol.-Hunan* 12, 102–156.

Shibairo, S. I., Nyongesa, O., Onwonga, R., and Ambuko, J. (2014). Variation of nutritional and anti-nutritional contents in finger millet (*Eleusine coracana* (L.) Gaertn) genotypes. *J. Agric. Vet. Sci.* 7 (11), 6–12. doi: 10.9790/2380-071110612

Shukla, S., Bhargava, A., Chatterjee, A., Srivastava, J., Singh, N., and Singh, S. P. (2006). Mineral profile and variability in vegetable amaranth (*Amaranthus tricolor*). *Plant Foods Hum. Nutr.* 61, 21–26. doi: 10.1007/s11130-006-0004-x

Singh, N., David, J., Thompkinson, D. K., Seelam, B. S., Rajput, H., and Morya, S. (2018). Effect of roasting on functional and phytochemical constituents of finger millet (*Eleusine coracana* L.). *Pharma Innovation J.* 7, 414–418.

Skrabanja, V., and Kreft, I. (1998). Resistant starch formation following autoclaving of buckwheat (*Fagopyrum esculentum* Moench) groats: an *in vitro* study. *J. Agric. Food Chem.* 46 (5), 2020–2023. doi: 10.1021/jf970756q

Skrabanja, V., Helle, N. L., and Kreft, I. (2000). Protein-polyphenol interactions and *in vivo* digestibility of buckwheat groat proteins. *Pflügers Archiv - Eur. J. Physiol.* 440, 129–131. doi: 10.1007/s004240000033

Skrabanja, V., Liljeberg Elmståhl, H. G., Kreft, I., and Björck, I. M. (2001). Nutritional properties of starch in buckwheat products: studies *in vitro* and *in vivo*. *J. Agric. Food Chem.* 49, 490–496. doi: 10.1021/jf000779w

Spehar, C. R. (2007). *Quinoa: Alternativa Para a Diversificação Agrícola e Alimentar*. Eds. D. F. Planaltina (Embrapa Cerrados, Brasil: Técnico). 103 pp.

Sripriya, G., Antony, U., and Chandra, T. S. (1997). Changes in carbohydrate, free amino acids, organic acids, phytate and HCl extractability of minerals during germination and fermentation of finger millet (*Eleusine coracana*). *Food Chem.* 58 (4), 345–350. doi: 10.1016/S0308-8146(96)00206-3

Steadman, K. J., Burgoon, M. S., Lewis, B. A., Edwardson, S. E., and Obendorf, R. L. (2001a). Buckwheat seed milling fractions: description, macronutrient composition and dietary fibre. *J. Cereal Sci.* 33, 271–278. doi: 10.1006/jcrs.2001.0366

Steadman, K. J., Burgoon, M. S., Lewis, B. A., Edwardson, S. E., and Obendorf, R. L. (2001b). Minerals, phytic acid, tannin and rutin in buckwheat seed milling fractions. *J. Sci. Food Agric.* 81, 1094–1100. doi: 10.1002/jsfa.914

Stetter, M. G., and Schmid, K. J. (2017). Analysis of phylogenetic relationships and genome size evolution of the *Amaranthus* genus using GBS indicates the ancestors of an ancient crop. *Mol. Phylogenet. Evol.* 109, 80–92. doi: 10.1016/j.ympev.2016.12.029

Stevens, M. R., Coleman, C. E., Parkinson, S. E., Maughan, P. J., Zhang, H. B., Balzotti, M. R., et al. (2006). Construction of a quinoa (*Chenopodium quinoa* Willd.) BAC library and its use in identifying genes encoding seed storage proteins. *Theor. Appl. Genet.* 112, 1593–1600. doi: 10.1007/s00122-006-0266-6

Stikic, R., Glamoclija, D., Demin, M., Vucelic-Radovic, B., Jovanovic, Z., Milojkovic-Opsenica, D., et al. (2012). Agronomical and nutritional evaluation of quinoa seeds (*Chenopodium quinoa* Willd.) as an ingredient in bread formulations. *J. Cereal Sci.* 55, 132–138. doi: 10.1016/j.jcs.2011.10.010

Stokić, E., Mandić, A., Sakač, M., Mišan, A., Pestorić, M., Šimurina, O., et al. (2015). Quality of buckwheat-enriched wheat bread and its antihyperlipidemic effect in statin treated patients. *LWT-Food Sci. Technol.* 63, 556–561. doi: 10.1016/j.lwt.2015.03.023

Subba Rao, M.V.S.S.T., and Muralikrishna, G. (2002). Evaluation of the antioxidant properties of free and bound phenolic acids from native and malted finger millet (Ragi, *Eleusine coracana* Indaf-15). *J. Agric. Food Chem.* 50 (4), 889–892. doi: 10.1021/jf011210d

Subramanian, D., and Gupta, S. (2016). Pharmacokinetic study of amaranth extract in healthy humans: a randomized trial. *Nutrition* 32, 748–753. doi: 10.1016/j.nut.2015.12.041

Sugiyama, K., Kushima, Y., and Muramatsu, K. (1985). Effects of sulfur-containing amino acids and glycine on plasma cholesterol level in rats fed on a high cholesterol diet. *Agric. Biol. Chem.* 49, 3455–3461.

Sunil, M., Hariharan, A. K., Nayak, S., Gupta, S., Nambisan, S. R., Gupta, R. P., et al. (2014). The draft genome and transcriptome of *Amaranthus hypochondriacus*: a C4 dicot producing high-lysine edible pseudo-cereal. *DNA Res.* 21, 585–602. doi: 10.1093/dnares/dsu021

Suresh, S., Chung, J. W., Cho, G. T., Sung, J. S., Park, J. H., Gwag, J. G., et al. (2014). Analysis of molecular genetic diversity and population structure in *Amaranthus* germplasm using SSR markers. *Plant Biosyst. Int. J. Dealing All Aspects Plant Biol.* 148, 635–644. doi: 10.1080/11263504.2013.788095

Taira, H., and Miyahara, T. (1983). *Amylose content of foxtail millet [Setaria italica] starch* (Japan: Report of National Food Research Institute).

Taira, H. (1984). Lipid content and fatty acid composition of nonglutinous and glutinous varieties of foxtail millet. *J. Agric. Food Chem.* 32, 369–371. doi: 10.1021/jf00122a047

Tang, S., Li, L., Wang, Y., Chen, Q., Zhang, W., Jia, G., et al. (2017). Genotype-specific physiological and transcriptomic responses to drought stress in *Setaria italica* (an emerging model for Panicoideae grasses). *Sci. Rep.* 7, 10009. doi: 10.1038/s41598-017-08854-6

Tatala, S., Ndossi, G., Ash, D., and Mamiro, P. (2007). Effect of germination of finger millet on nutritional value of foods and effect of food supplement on nutrition and anaemia status in Tanzania children. *Tanzania Health Res. Bull.* 9, 77–86. doi: 10.4314/thrb.v9i2.14308

Teng, X. L., Chen, N., and Xiao, X.-G. (2016). Identification of a catalase-phenol oxidase in betalain biosynthesis in red amaranth (*Amaranthus cruentus*). *Front. Plant Sci.* 6, 1228. doi: 10.3389/fpls.2015.01228

Tomotake, H., Shimaoka, I., Kayashita, J., Yokoyama, F., Nakajoh, M., and Kato, N. (2000). A buckwheat protein product suppresses gallstone formation and plasma cholesterol more strongly than soy protein isolate in hamsters. *J. Nutr.* 130 (7), 1670–1674. doi: 10.1093/jn/130.7.1670

Tomotake, H., Shimaoka, I., Kayashita, J., Nakajoh, M., and Kato, N. (2002). Physicochemical and functional properties of buckwheat protein product. *J. Agric. Food Chem.* 50, 2125–2129. doi: 10.1021/jf011248q

Tripathi, B., and Patel, K. (2010). Finger millet (*Eleucine coracana*) flour as a vehicle for fortification with zinc. *J. Trace Elem. In Med. Biol.* 24 (1), 46–51. doi: 10.1016/j.jtemb.2009.09.001

Trivedi, A. K., Arya, L., Verma, S. K., Tyagi, R. K., Hemantaranjan, A., Verma, M., et al. (2018). Molecular profiling of foxtail millet (*Setaria italica* (L.) P. Beauv) from Central Himalayan Region for genetic variability and nutritional quality. *J. Agric. Sci.* 156, 333–341. doi: 10.1017/S0021859618000382

Tsehaye, Y., Berg, T., Tsegaye, B., and Tanto, T. (2006). Farmers' management of finger millet (*Eleusine coracana* L.) diversity in Tigray, Ethiopia and implications for on-farm conservation. *Biodivers. Conserv.* 15, 4289–4308. doi: 10.1007/s10531-005-3581-3

Udeh, H., Gyebi, D., and Jideani, A. (2017). Finger millet bioactive compounds, bioaccessibility, and potential health effects: a review. *Czech J. Food Sci.* 35, 7–17. doi: 10.17221/206/2016-CJFS

Ulbricht, C., Abrams, T., Conquer, J., Costa, D., Grimes Serrano, J. M., Taylor, S., et al. (2009). An evidence-based systematic review of amaranth (*Amaranthus* spp.) by the natural standard research collaboration. *J. Diet. Suppl.* 6, 390–417. doi: 10.3109/19390210903280348

Umeta, M., West, C. E., and Fufa, H. (2005). Content of zinc, iron, calcium and their absorption inhibitors in foods commonly consumed in Ethiopia. *J. Food Compos. Anal.* 18, 803–817. doi: 10.1016/j.jfca.2004.09.008

USDA National Nutrient Database for Standard Reference(2019). https://fdc.nal.usda.gov/. (accessed on 28 Oct 2019).

Vadivoo, A. S., Joseph, R., and Ganesan, N. M. (1998). Genetic variability and diversity for protein and calcium contents in finger millet (*Eleusine coracana* (L.) Gaertn.) in relation to grain color. *Plant Foods Hum. Nutr.* 52, 353–364. doi: 10.1023/A:1008074002390

Valencia, U. S., Sandberg, A. S., and Ruales, J. (1999). Processing of quinoa (*Chenopodium quinoa*, Willd): effects on *in vitro* iron availability and phytate hydrolysis. *Int. J. Food Sci. Nutr.* 50 (3), 203–211. doi: 10.1080/096374899101247

Vega-Gálvez, A., Miranda, M., Vergara, J., Uribe, E., Puente, L., and Martínez, E. A. (2010). Nutrition facts and functional potential of quinoa (*Chenopodium quinoa* willd.), an ancient Andean grain: a review. *J. Sci. Food Agric.* 90, 2541–2547. doi: 10.1002/jsfa.4158

Walsh, B. M., Adhikary, D., Maughan, P. J., Emshwiller, E., and Jellen, E. N. (2015). Chenopodium polyploidy inferences from *Salt Overly Sensitive 1 (SOS1)* data. *Am. J. Bot.* 102, 533–543. doi: 10.3732/ajb.1400344

Walters, H., Carpenter-Boggs, L., Desta, K., Yan, L., Matanguihan, J., and Murphy, K. (2016). Effect of irrigation, intercrop, and cultivar on agronomic and nutritional characteristics of quinoa. *Agroecol. Sustain. Food Syst.* 40, 783–803. doi: 10.1080/21683565.2016.1177805

Wang, J., Liu, Z., Fu, X., and Run, M. A. (1992). A clinical observation on the hypoglycaemic effect of Xinjiang buckwheat. In: Proceedings of the 5th International Symposium on Buckwheat. Eds. R. Lin, M. Zhou, Y. Tao, J. Li and Z. Zhang. (Beijing, China: Agriculture Publishing House), 465–467

Wang, Z. M., Devos, K. M., Liu, C. J., Wang, R. Q., and Gale, M. D. (1998). Construction of RFLP-based maps of foxtail millet, *Setaria italica* (L.) P. Beauv. *Theor. Appl. Genet.* 96, 31–36. doi: 10.1007/s001220050705

Wang, Z. H., Gao, L., Li, Y. Y., Zhang, Z., Yuan, J. M., Wang, H. W., et al. (2007). Induction of apoptosis by buckwheat trypsin inhibitor in chronic myeloid leukemia K562 cells. *Biol. Pharm. Bull.* 30 (4), 783–786. doi: 10.1248/bpb.30.783

Whittaker, P., and Ologunde, M. O. (1990). Study on iron bioavailability in a native Nigerian grain amaranth cereal for young children, using a rat model. *Cereal Chem.* 67, 505–508.

Williams, J. T., and Brenner, D. (1995). "Grain amaranth (*Amaranthus* species)," in *Underutilized Crops: Cereals and Pseudocereals*. Eds. J. T. Williams (London, UK: Chapman & Hall), 129–187.

Wilson, H. D. (1988a). Allozyme variation and morphological relationships of *Chenopodium hircinum*. *Syst. Bot.* 13, 215–228. doi: 10.2307/2419100

Wilson, H. D. (1988b). Quinua biosystematics I: domesticated populations. *Economic Bot.* 42, 461–477. doi: 10.1007/BF02862791

Wu, Q., Bai, X., Zhao, W., Xiang, D., Wan, Y., Yan, J., et al. (2017). De novo assembly and analysis of tartary buckwheat (*Fagopyrum tataricum* Gaertn.) transcriptome discloses key regulators involved in salt-stress response. *Genes* 8, 255. doi: 10.3390/genes8100255

Xu, F., and Sun, M. (2001). Comparative analysis of phylogenetic relationships of grain amaranths and their wild relatives (*Amaranthus*; Amaranthaceae) using internal transcribed spacer, amplified fragment length polymorphism, and double-primer fluorescent intersimple sequence repeat markers. *Mol. Phylogenet. Evol.* 21, 372–387. doi: 10.1006/mpev.2001.1016

Xu, J. M., Fan, W., Jin, J. F., Lou, H. Q., Chen, W. W., Yang, J. L., et al. (2017). Transcriptome analysis of Al-induced genes in buckwheat (*Fagopyrum esculentum* Moench) root apex: new insight into Al toxicity and resistance mechanisms in an Al accumulating species. *Front. Plant Sci.* 8, 1141. doi: 10.3389/fpls.2017.01141

Yabe, S., Hara, T., Ueno, M., Enoki, H., Kimura, T., Nishimura, S., et al. (2014). Rapid genotyping with DNA micro-arrays for high-density linkage mapping and QTL mapping in common buckwheat (*Fagopyrum esculentum* Moench). *Breed. Sci.* 64, 291–299. doi: 10.1270/jsbbs.64.291

Yasui, Y., Wang, Y., Ohnishi, O., and Campbell, C. G. (2004). Amplified fragment length polymorphism linkage analysis of common buckwheat (*Fagopyrum esculentum*) and its wild self-pollinated relative *Fagopyrum homotropicum*. *Genome* 47, 345–351. doi: 10.1139/g03-126

Yasui, Y., Mori, M., Matsumoto, D., Ohnishi, O., Campbell, C. G., and Ota, T. (2008). Construction of a BAC library for buckwheat genome research. *Genes Genet. Syst.* 83, 393–401. doi: 10.1266/ggs.83.393

Yasui, Y., Hirakawa, H., Oikawa, T., Toyoshima, M., Matsuzaki, C., Ueno, M., et al. (2016). Draft genome sequence of an inbred line of *Chenopodium quinoa*, an allotetraploid crop with great environmental adaptability and outstanding nutritional properties. *DNA Res.* 23, 535–546. doi: 10.1093/dnares/dsw037

Zduńczyk, Z., Flis, M., Zieliński, H., Wróblewska, M., Antoszkiewicz, Z., and Juśkiewicz, J. (2006). *In vitro* antioxidant activities of barley, husked oat, naked oat, triticale, and buckwheat wastes and their influence on the growth and biomarkers of antioxidant status in rats. *J. Agric. Food Chem.* 54, 4168–4175. doi: 10.1021/jf060224m

Zhang, Z., Upadhayay, M. P., Zhao, Z., Lin, R., Zhou, M. D., Rao, V. R., et al. (2004). "Conservation and use of buckwheat biodiversity for the livelihood development," in *Proceedings of the 9th International Symposium on Buckwheat, Prague, Czech Republic, August 18-22, 2004*, Czech Republic: Research Institute of Crop Production 442–431.

Zhang, J., Liu, T., Fu, J., Zhu, Y., Jia, J., Zheng, J., et al. (2007). Construction and application of EST library from *Setaria italica* in response to dehydration stress. *Genomics* 90, 121–131. doi: 10.1016/j.ygeno.2007.03.016

Zhang, G., Liu, X., Quan, Z., Cheng, S., Xu, X., Pan, S., et al. (2012). Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat. Biotechnol.* 30, 549. doi: 10.1038/nbt.2195

Zhang, G., Xu, Z., Gao, Y., Huang, X., Zou, Y., and Yang, T. (2015). Effects of germination on the nutritional properties, phenolic profiles, and antioxidant activities of buckwheat. *J. Food Sci.* 80 (5), H1–H9. doi: 10.1111/1750-3841.12830

Zhang, L., Li, X., Ma, B., Gao, Q., Du, H., Han, Y., et al. (2017). The tartary buckwheat genome provides insights into rutin biosynthesis and abiotic stress tolerance. *Mol. Plant* 10, 1224–1237. doi: 10.1016/j.molp.2017.08.013

Zhang, T., Gu, M., Liu, Y., Lv, Y., Zhou, L., Lu, H., et al. (2017). Development of novel InDel markers and genetic diversity in *Chenopodium quinoa* through whole-genome re-sequencing. *BMC Genomics* 18, 685. doi: 10.1186/s12864-017-4093-8

Zhou, M., and Zhang, Z. (1995). *Directory of germplasm collections – Buckwheat* (Beijing, 26 pp: International Plant Genetic Resources Institute).

Zhou, M., Kreft, I., Suvorova, G., Chrungoo, N., Tang, Y., and Woo, S. H. (Eds.) (2018). *Buckwheat germplasm in the world* (Academic Press).

Zhou, M., Kreft, I., Woo, S. H., Chrungoo, N., and Wieslander, G. (2016). *Molecular breeding and nutritional aspects of buckwheat* (Academic Press). 482 p.

Zielińska, D., Turemko, M., Kwiatkowski, J., and Zieliński, H. (2012). Evaluation of flavonoid contents and antioxidant capacity of the aerial parts of common and tartary buckwheat plants. *Molecules* 17, 9668–9682. doi: 10.3390/molecules17089668

Zielinski, H., Michalska, A., Amigo-Benavent, M., del Castillo, M. D., and Piskula, M. K. (2009). Changes in protein quality and antioxidant properties of buckwheat seeds and groats induced by roasting. *J. Agric. Food Chem.* 57, 4771–4776. doi: 10.1021/jf900313e

Zou, C., Chen, A., Xiao, L., Muller, H. M., Ache, P., Haberer, G., et al. (2017). A high-quality genome assembly of quinoa provides insights into the molecular basis of salt bladder-based salinity tolerance and the exceptional nutritional value. *Cell Res.* 27, 1327–1340. doi: 10.1038/cr.2017.124

Zurita-Silva, A., Fuentes, F., Zamora, P., Jacobsen, S.-E., and Schwember, A. R. (2014). Breeding quinoa (*Chenopodium quinoa* Willd.): potential and perspectives. *Mol. Breed.* 34, 13–30. doi: 10.1007/s11032-014-0023-5

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 Genetics

ORIGINAL RESEARCH
published: 17 March 2020
doi: 10.3389/fgene.2020.00240

Mapping and Validation of QTLs for the Amino Acid and Total Protein Content in Brown Rice

Su Jang[1], Jae-Hyuk Han[2], Yoon Kyung Lee[1], Na-Hyun Shin[2], Yang Jae Kang[3,4], Chang-Kug Kim[5] and Joong Hyoun Chin[2]**

[1] Department of Plant Science, Plant Genomics and Breeding Institute, Research Institute of Agriculture and Life Sciences, Seoul National University, Seoul, South Korea, [2] Department of Integrative Bio-Industrial Engineering, Sejong University, Seoul, South Korea, [3] Division of Applied Life Science (BK21 Plus Program), Plant Molecular Biology and Biotechnology Research Center, Gyeongsang National University, Jinju, South Korea, [4] Division of Life Science, Gyeongsang National University, Jinju, South Korea, [5] Genomics Division, National Institute of Agricultural Sciences, Jeonju, South Korea

OPEN ACCESS

Edited by:
Mallikarjuna Swamy, International Rice Research Institute, Philippines

Reviewed by:
Hyun-Sook Lee, Chungnam National University, South Korea
O-Young Jeong, Rural Development Administration, South Korea

***Correspondence:**
Chang-Kug Kim
chang@korea.kr
Joong Hyoun Chin
jhchin@sejong.ac.kr; joonghyoun.chin@gmail.com

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *28 December 2019*
Accepted: *28 February 2020*
Published: *17 March 2020*

Citation:
Jang S, Han J-H, Lee YK, Shin N-H, Kang YJ, Kim C-K and Chin JH (2020) Mapping and Validation of QTLs for the Amino Acid and Total Protein Content in Brown Rice. Front. Genet. 11:240. doi: 10.3389/fgene.2020.00240

Highly nutritious rice production will be benefited with the improvement of amino acid content (AAC) and protein content (PC). The identification of quantitative trait loci (QTLs) associated with the PC and AAC of rice grains could provide a basis for improving the nutritional value of rice grains. Here, we conducted QTL analyses using recombinant inbred lines from the cross between *indica* (Milyang 23 or M23) and *japonica* (Tong 88-7 or T887) rice varieties, afterward employing genotyping-by-sequencing to obtain a high-density genetic map. A total of 17 and 3 QTLs were detected for AAC and PC, respectively. Among them, two QTLs associated with more than 10 AACs, *qAAC6.1* and *qAAC7.1*, were identified for the first time in this study. Each favorable allele that increased the AAC of the two QTLs was derived from M23 and T887, respectively. Allelic combination of $qAAC6.1^{M23}$ and $qAAC7.1^{T887}$ showed significantly higher content of associated amino acids (AAs) than other allelic combinations. Near-isogenic line (NIL) possessing $qAAC7.1^{T887}$ with M23 genetic background had significantly higher AACs than both parents. These results indicate that the pyramiding of QTLs would be useful in developing brown rice with a high AA and protein content.

Keywords: rice (*Oryza sativa* L.), amino acid content, protein content, quantitative trait loci, brown rice, genotyping-by-sequencing

INTRODUCTION

Rice has been a staple food for thousands of years, especially for Asians and Africans. The global consumption of rice is gradually increasing as its dietary benefits in human nutrition has been more factually accepted in many countries. In most rice growing countries, rice eating qualities have evolved through highly polished rice by utilizing advanced milling technologies. In the process, rice bran, which is the most nutritional component of rice, is easily removed. Common polished rice loses various phytochemicals, vitamins, and minerals (Vetha-Varshini et al., 2013). In healthy rice breeding programs, the promotion for brown rice consumption can be a powerful option. Furthermore, brown rice can be a better option for agriculture policy makers and economists to secure higher rice production.

The protein content (PC) and amino acid content (AAC) may not be *indica–japonica* specific traits. The positive effects of increasing PC and AAC could be derived either from *japonica* or from *indica*. Although the identification of the associated quantitative trait loci (QTLs)/genes for AAC is important, there are very few studies on the trait identification on AAC. Wang et al. (2007) reported the QTLs for AAC in milled rice, using recombinant inbred lines (RILs) from a cross between Zhenshan 97 (*indica*) and Nanyangzhan (*japonica*). In all amino acids (AAs), the AAC of Zhenshan 97 was larger than that of Nanyangzhan in the 2-year experiments. In the study, several QTL clusters for AAC were identified on chromosomes 1, 2, 3, 4, 7, 8, and 10. There were some large effect QTLs for Asp, Thr, Gly, and Ala on RM472–RM104 on chromosome 1 (phenotypic variance explained, PVE = 17.5–33.3%), and Gly, Arg, Met, and Pro on RM125–RM214 on chromosome 7 (PVE = 4.68–7.69%). Another study on the QTL identification for AAC in brown rice was reported (Yoo, 2017). In the study, the QTL for AAC including Ala, Val, Leu, Ile, and Phe was the only QTL cluster on chromosome 3 in Dasanbyeo (*indica*) and TR22183 (*japonica*) RILs. In the same study, another QTL for Lys was identified on chromosome 3. In both the studies, the precise linkage mapping analysis was limited due to lack of markers in the QTL regions.

On the other hand, several QTLs for PC were identified. The rice PC varied from 4.9–19.3% in *indica* and 5.9–16.5% in *japonica* (Lin et al., 1993). The earliest study on QTL mapping for PC was conducted using Zhenshan 97/Minghui 63 RILs (Tan et al., 2001). Two QTLs on chromosomes 6 and 7 were identified using the RFLP marker system. The *Wx* marker associated with amylose content (AC) was linked to QTLs on chromosome 6. The QTL for PC linked to ACs were identified repeatedly in the following studies. The QTL near *Wx* (RM190) on chromosome 6 explained 19.3% PVE in the other study (Yu et al., 2009). Two major QTLs for PC were identified on chromosomes 6 (RM588–RM540) and 7 (RM5436–RM6776) (Lou et al., 2009). Eight QTLs for PC included two QTLs on chromosome 6 (Kinoshita et al., 2017) and a new QTL for PC and AC on chromosome 8 was reported.

Due to the complex genetic structure in the study of AAC and PC, very few genes were cloned and functionally investigated. A major QTL for grain PC (GPC), *qPC1* (a putative AA polymerase, *OsAAP6*) on chromosome 1, was cloned and characterized (Peng et al., 2014). *qPC1* is a QTL cluster exhibiting the highest PVE (32.4%), conferred by the Zhenshan 97 allele.

Currently, the genes/QTLs for protein turnover and complicated AA catabolism have not been characterized well – especially in brown rice. The linked QTLs for AA should be mapped within narrow regions on the chromosomes, and some candidate genes should be suggested for future studies.

In this study, other *indica–japonica* RILs have been analyzed for AAC and PC in brown rice. Followed by the identification of thousands of single nucleotide polymorphic loci (SNPs) using the genotyping-by-sequencing (GBS) technique, several QTLs were identified in high-resolution genetic maps. Furthermore, validation of major QTL, *qAAC7.1*, were conducted using backcross progeny. The QTLs identified in this study and their closely linked markers could be used to develop high nutritional rice breeding program.

MATERIALS AND METHODS

Plant Materials

The F_{15} RILs, derived from a cross between Milyang 23 (M23, *indica*) and Tong88-7 (T887, *japonica*), were developed at Seoul National University in Korea (Jiang et al., 2008). The plant materials, together with the parental lines, were grown at the Seoul National University Experimental Farm in Suwon. The plants were transplanted to one seedling per hill at a planting density of 30 × 15 cm. Plants were cultivated under normal fertilizer conditions ($N–P_2O_5–K_2O$ = 100–80–80 kg/ha). The rice field was regularly irrigated.

Quantification of Amino Acid Content and Total Crude Protein Content

Unhulled rice grains for assay were harvested at 50 days after heading, and were naturally dried on drying beds with transparent roofs for 14 days. Unhulled grains were stored at 10°C, until all samples for assay were ready.

Dehulling and grinding were conducted simultaneously for all samples. Before grinding, the brown rice grain was cleaned by removing broken, green, and immature grain. A total 30 g of brown rice grain was ground at a time and the aliquots of powder were used to each assay.

AAC was determined using the high-performance liquid chromatography (HPLC) method. One hundred milligrams of powdered milled brown rice samples were collected using more than 100 seed samples of different plants of each RIL. Afterward, 40 mL of 6 mol/L HCl was added, and hydrolysis was conducted for 24 h at 110°C. Samples were analyzed by HPLC (Dionex Ultimate 3000, Thermo Fisher Scientific) after pre-injection derivatization (Henderson et al., 2000). The primary AAs were reacted first with *o*-phthalaldehyde (OPA) and the secondary AAs were derivatized with fluorenylmethyl chloroformate (FMOC) before injection. The contents of each AA in each hydrolysate were calculated in reference to the standard AA and expressed as mg/g rice powder. The levels of the 16 AAC traits of rice grains (Ala, alanine; Arg, arginine; Asx, aspartic acid + asparagine; Glx, glutamic acid + glutamine; Gly, glycine; Ile, isoleucine; Leu, leucine; Lys, lysine; Met, methionine; Pro, proline; Phe, phenylalanine; Val, valine; Tyr, tyrosine; His, histidine; Thr, threonine; Ser, Serine) were obtained. The total nitrogen of crude protein was determined by using the micro-Kjeldahl method. The PC was calculated using a conversion factor of 5.95 (Jones, 1931).

Genotyping-by-Sequencing Library Preparation and Genotyping Analysis

The whole genomic DNA for each RIL and parents were isolated from 1-month-old leaf tissues using the CTAB extraction method (Murray and Thompson, 1980). The quality and quantity of DNA were measured using PicoGreen (Invitrogen, Carlsbad, CA, United States), before diluting it to a concentration of

10 ng/μL for GBS. A GBS library was prepared with the restricted enzyme *Ape*KI, as described by Elshire et al. (2011). The quantity and quality of the GBS library was determined through the Bioanalyzer Kit (Agilent Genomics, Santa Clara, CA, United States). For accurate variation calling from our RIL population of inter subspecific cross, we prepared subset of rice reference genomes with *Ape*KI restriction sites, which are target genomic regions for GBS genotyping with IRGSP 1.0 pseudomolecule. We *in silico*-searched the *Ape*KI sites in the reference genome sequences, and split it into fragments. The subset of genome fragments smaller than 2 kb were collected. The genome subsets were indexed and the GBS short reads were mapped using BWA software (Li and Durbin, 2009). The genotypes of the RIL population were determined into homo- and heterozygous genotypes by the software SAMtools, with default parameters. For cross-validation, the parents, genomic DNA of M23 and T887, were fully sequenced using the HiSeq 2000 platform (Illumina, Inc., San Diego, CA, United States) to compare allele contents for each locus with GBS results. Only the called SNPs from the GBS results which matched with resequencing results of the parents, M23 and T887, were chosen for the following analyses.

QTL Mapping and Validation

Genetic map construction and QTL analyses were carried out using the software ICIMapping 4.1 (Meng et al., 2015). All 26,424 SNPs discovered through GBS were binned using the BIN function allowing only SNPs which have less than 10% missing rate of individuals. A total of 1,327 bin markers were used to construct the linkage map, and the ordering of the markers within each chromosome was based on the physical position of the Nipponbare reference genome, IRGSP 1.0 (Sakai et al., 2013). The recombination distance was calculated using the Kosambi mapping function (Kosambi, 1943). An inclusive composite interval mapping (ICIM) method was used to detect additive QTLs. The logarithm of odds (LOD) values greater than 2.5 was used to declare significant QTLs. Backcross population for selecting *qAAC7.1* recombinants was derived from BC_3F_1 plants with *qAAC6.1* allele of recurrent parents (M23) and heterozygous *qAAC7.1*. *qAAC6.1* allele were discriminated using two insertion/deletion (InDel) markers, id3428419 and id4203043. Primers used in recombinants selection were listed in **Supplementary Table S1**. Near-isogenic line (NIL) was developed by repeated backcrossing with the M23 as recurrent parent. The backcross progeny was selected by marker-assisted foreground selection using two InDel markers, id4638739 and id5227020 (**Supplementary Table S1**). PCR amplification was performed in 20 μL reaction mixture containing 100 ng of DNA, 0.5 U Prime Taq polymerase (GeNet Bio, South Korea), 1× PCR buffer, 0.5 μM of each primer, 0.5 mM of dNTPs. PCR condition was 94°C for 5 min, followed by 30 cycles of 94°C for 30 s, 58°C for 30 s, 72°C for 30 s, and final extension at 72°C for 10 min. A total 176 SNP type assays (Fluidigm, San Francisco, CA, United States) were employed to select genetic background of NIL. All assays were designed based on polymorphism between *indica* and *japonica* (unpublished). All statistical analyses were performed using R version 3.4.3.

RESULTS

Phenotypic Analysis

The 16 AACs and PC of brown rice in both parents and RILs were analyzed (**Figure 1**). Unlike the other AAs, the values of Met, Val, Leu, Lys, and Pro were higher in T887 than in M23. The phenotypic variation of total PC and most AACs showed a continuous and normal distribution in RILs, based on the Kolmogorov–Smirnov normality test. However, Pro content were asymmetrically distributed with a long tail to higher values, showing a positive skewness ($D = 0.1$, $p < 0.01$). Transgressive segregations in both directions were observed for PC and all AAC traits in the RIL population showed values higher or lower than that of both parents. Significant positive correlations were found between PC and most AACs, except for Pro (**Supplementary Figure S1**). However, positive correlation between PC and two AAs, Lys and Pro, were relatively weak (correlation coefficient <0.2). Each AAC was significantly positively correlated with the other AAC traits.

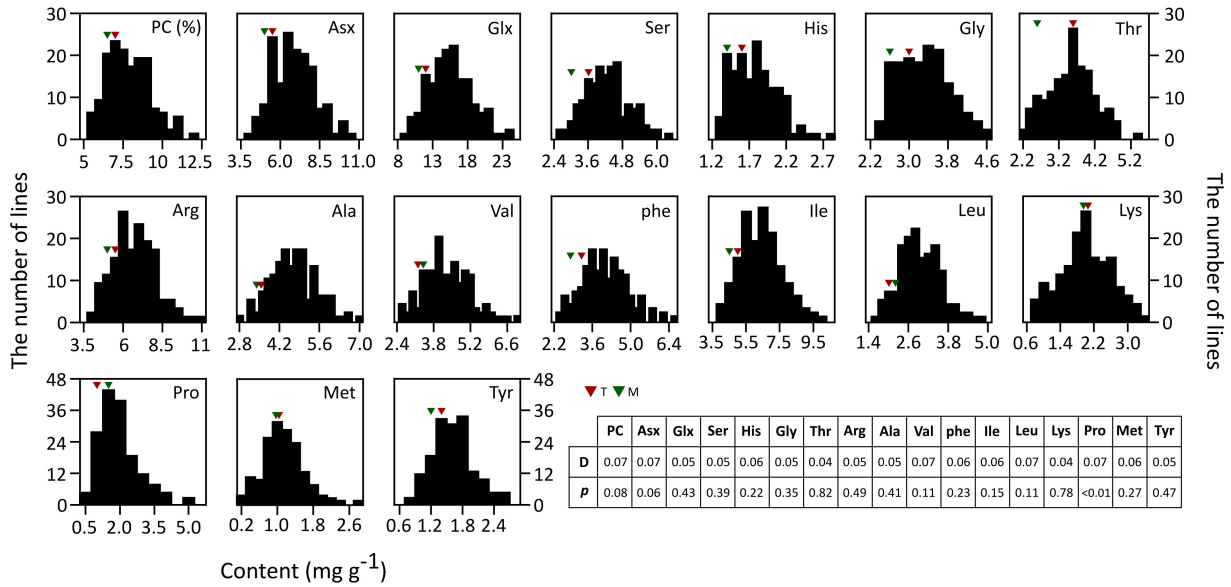
GBS Analysis and Genetic Map Construction

A total of 2,468,603 and 256,571 variants were found in M23 and T887 genome sequence, compared to Nipponbare reference genome (IRGSP 1.0), respectively. Only the called SNPs from the GBS results of RILs which matched with resequencing results of the parents were chosen for the following analyses. A total of 26,424 SNPs were called successfully from the GBS result. All the SNPs showing the same allele calls across the lines were filtered and binned. Consequently, 1,327 SNPs were regarded as the representative markers of each linkage block in the population (**Supplementary Figure S2**). The SNPs identified from GBS were non-uniformly distributed, resulting in large gaps across all chromosomes, especially in chromosome 4 (66.1–111.34 cM), chromosome 6 (3–26 cM), chromosome 8 (7.5–38.5 cM), and chromosome 12 (12.9–36.3 cM and 50.7–77.2 cM). The total length of the genetic map was 1,421.6 cM. The average physical distance between markers were 3.56 Mb.

QTL Analysis for PC and AACs

A total of three QTLs for PC and 17 QTLs for AACs were identified on six chromosomes: 1, 2, 3, 6, 7, and 8 (**Figure 2**). Three QTLs for PC, including *qPC1.1*, *qPC1.2*, and *qPC7.1*, were co-located with QTLs for Ser (*qAAC1.3*), His (*qAAC1.6*) and Thr *(qAAC7.2*), respectively. The positive QTL effects of *qPC1.1* and *qPC7.1* were derived from T887 (**Supplementary Table S2**), whereas the positive effects of *qPC1.2* derived from M23. Among them, *qPC1.2* had the largest additive effect, explaining 18.1% of the phenotypic variance in PC.

Except for proline, the QTLs for all the other AAs were identified in this study (**Figure 2**). Five QTLs (*qAAC6.1*, *qAAC6.2*, *qAAC7.1*, *qAAC7.2*, and *qAAC8.2*) were linked with the essential AAC. *qAAC2.2* for Ser showed the highest LOD score greater than 51.7, explaining 29.5% of phenotypic variance (**Supplementary Table S2**). QTLs for Ser were the most



	PC	Asx	Glx	Ser	His	Gly	Thr	Arg	Ala	Val	phe	Ile	Leu	Lys	Pro	Met	Tyr
D	0.07	0.07	0.05	0.05	0.06	0.05	0.04	0.05	0.05	0.07	0.06	0.06	0.07	0.04	0.07	0.06	0.05
p	0.08	0.06	0.43	0.39	0.22	0.35	0.82	0.49	0.41	0.11	0.23	0.15	0.11	0.78	<0.01	0.27	0.47

FIGURE 1 | Phenotypic distributions of 16 AACs and PC in brown rice of 155 RILs. Each AAC and PC is represented in percentage and mg/g, respectively. The result of Kolmogorov–Smirnov test used to test normality of distribution is shown in table ($p \leq 0.05$).

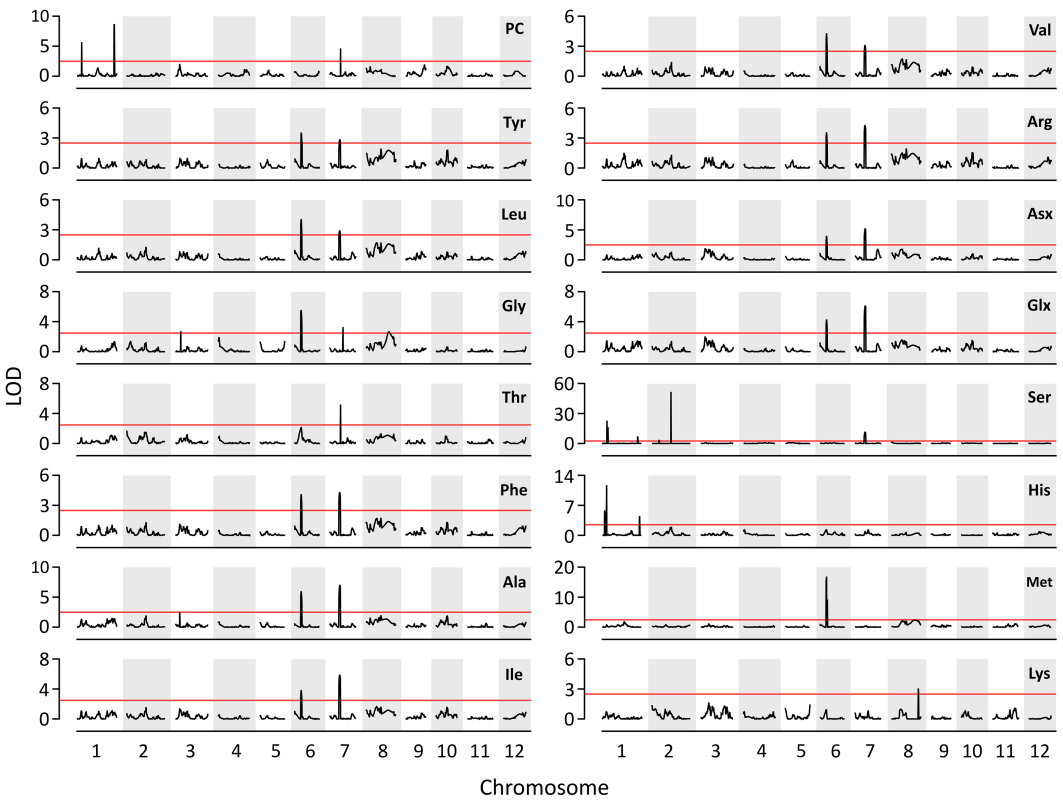


FIGURE 2 | The results of QTL analysis for AAC and PC. Red horizontal line indicates LOD threshold (>2.5).

frequently detected in RILs, with a total of six loci including *qAAC1.3*, *qAAC1.4*, *qAAC1.5*, *qAAC2.1*, *qAAC2.2*, and *qAAC7.1*.

Two major QTLs (*qAAC6.1* and *qAAC7.1*) affected several AACs simultaneously. *qAAC6.1* affected a total of 11 AACs (Ala, Arg, Asx, Glx, Gly, Ile, Leu, Met, Phe, Tyr, and Val), explaining 7.8–13.8% phenotypic variance (**Figure 2** and **Supplementary Table S2**). *qAAC7.1* had an effect to 10 AAC traits, showing 3.2–16.3% PVE. This QTL included two LOD peaks. There was one LOD peak for Arg, Ile, Leu, and Phe, detected between 4.86 and 5.02 Mb, and another LOD peak for Ala, Asx, Glx, Ser, Tyr, and Val, detected in the 5.02–5.21 Mb interval. Since each LOD peak area overlapped at around 5.02 Mb, the two detected loci were designated as one QTL, *qAAC7.1*. The lines carrying the *qAAC6.1* allele derived from M23 ($qAAC6.1^{M23}$) showed significantly higher contents of 11 AAs than lines with $qAAC6.1^{T887}$ (**Figure 3A**), while lines with $qAAC7.1^{T887}$ showed significantly higher contents of 10 AAs than lines with $qAAC7.1^{M23}$ (**Figure 3B**). Allelic combinations of two additive QTLs, *qAAC6.1* and *qAAC7.1*, additively affected most AACs, showing no epistatic interaction (**Figure 3C**). Notably, the combination of $qAAC6.1^{M23}$ and $qAAC7.1^{T887}$ exhibited significantly higher contents of 11 AAs, including Tyr, Leu, Thy, Phe, Ser, Val, Ala, Ile, Arg, Asx, and Glx than the allelic combinations of both parents ($qAAC6.1^{M23}$/$qAAC7.1^{M23}$ and $qAAC6.1^{T887}$/$qAAC7.1^{T887}$).

Validation of qAAC7.1

In order to confirm genetic effect of *qAAC7.1* detected in primary mapping (**Figure 4A**), 154 BC_3F_2 populations, which were derived from BC_3F_1 individual plant with *qAAC6.1* allele of recurrent parents (M23) and heterozygous *qAAC7.1* in the genetic background of M23, were used. Five recombinant plants were selected with two InDel markers, id4638739 and id5227020

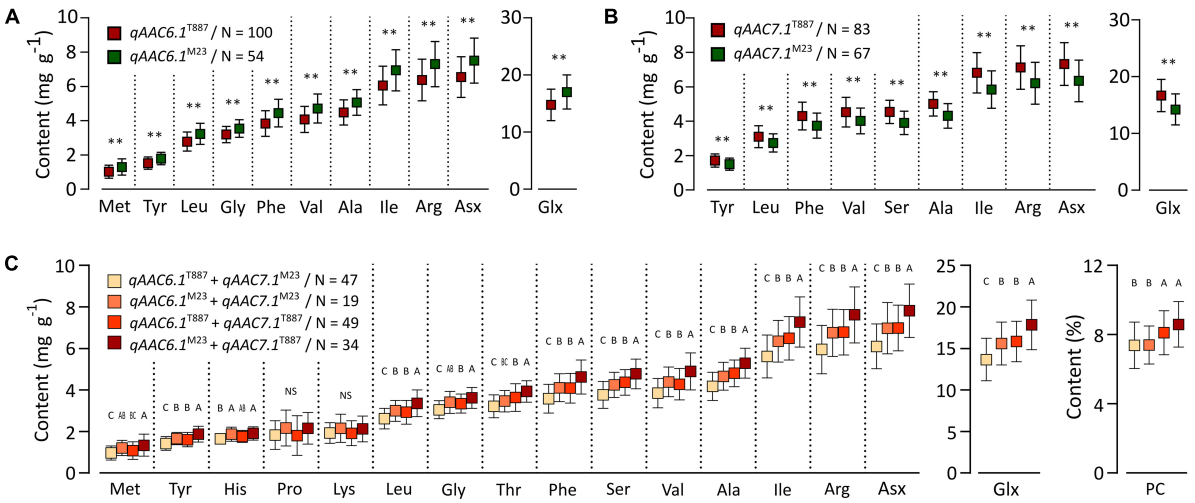


FIGURE 3 | Allelic effects for AACs in 155 RILs. Allelic effects of *qAAC6.1* **(A)** and *qAAC7.1* **(B)**. Asterisk represents statistically significant difference (**$p \leq 0.01$). Red box and green box represent the average AAC of each line containing the alleles originated from T887 and M23, respectively. **(C)** Effect of different allelic combination of *qAAC6.1* and *qAAC7.1*. Different letters indicate significant differences between groups ($p \leq 0.05$, Tukey's HSD test). N, the number of lines.

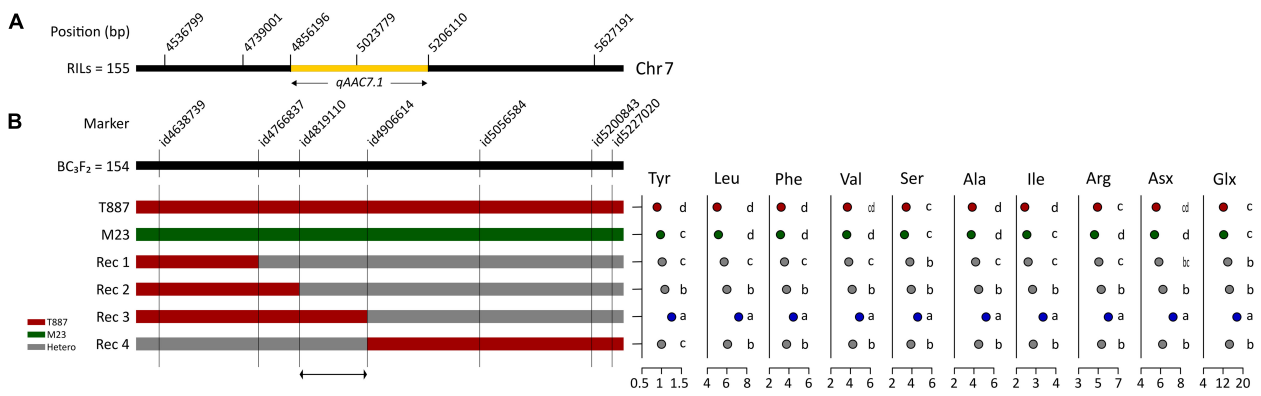


FIGURE 4 | Selection of *qAAC7.1* recombinants. **(A)** *qAAC7.1* region detected by primary mapping in RILs. **(B)** Selection of *qAAC7.1* recombinants using BC_3F_2 population. Different letters indicate significant differences between groups ($p \leq 0.05$, Tukey's HSD test). Red and green bar indicate the homozygotes with T887 and M23 allele, respectively. The gray bar indicates the heterozygotes.

(**Figure 4B**). Five additional markers were used to narrow down target region and, consequently, a recombinant plant with homozygous *qAAC7.1*^{T887} between id4819110 and id4906614 markers was selected (**Figure 4B**). NIL carrying homozygous *qAAC7.1*^{T887} with M23 genetic background (M23-*qAAC7.1*^{T887}) had genomic composition with recovery of recurrent genotype at 163 out of 176 loci (**Figure 5A**). M23-*qAAC7.1*^{T887} possessed a combination of *qAAC6.1*^{M23} and *qAAC7.1*^{T887}, which showed higher AACs than the other allelic combinations in RILs. As expected, this line had significantly higher contents of *qAAC7.1*-related AAs than both parents. In addition, M23-*qAAC7.1*^{T887} had also increased PC and the others AACs, including Met, His, Lys, and Gly than both parents (**Figure 5B**).

Comparison of QTLs for PC and AACs in the Present and Previous Studies

The consensus map of QTLs between present and previous studies were constructed to compare the loci where several QTLs were co-located. Marker positions of previous studies were determined using the Nipponbare reference genome (IRGSP 1.0). A total of five QTL clusters, on chromosome 1, 2, 7, and 8, were found to overlap the previously reported QTLs (**Figure 6** and **Supplementary Table S3**). In particular, the 1.5–5.19 Mb region on chromosome 1 included two QTLs for PC, J-*qPC1.1* (this study) Q-*qPC1.1* (Qin et al., 2009), and four QTLs for AACs. Two major QTLs, *qAAC6.1* and *qAAC7.1* were novel, only detected in this study.

DISCUSSION

Highly nutritional rice may contain high PC with desirable AA constitutions. Detecting the QTL which controls PC and AAC of the rice grain would provide a basis for improving nutritional qualities of the rice grain. In this study, 155 RILs and 1,327 markers were used to identify QTL for PC and AAC in grains. The genetic map showed an average of a 280.9 kb distance between

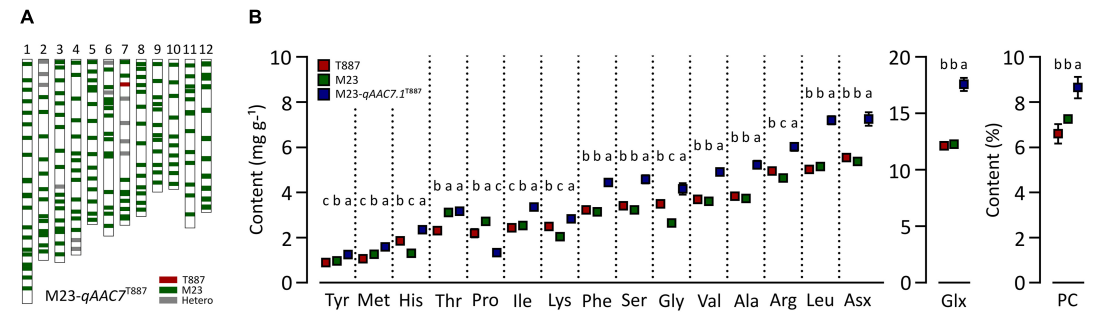


FIGURE 5 | Validation of *qAAC7.1* effect. **(A)** Genetic background of M23-*qAAC7.1*^{T887}. Red and green bar indicate the homozygous alleles from T887 and M23 allele, respectively. The gray bars indicate the heterozygous allele. **(B)** Allelic effect of *qAAC7.1*^{T887} for AACs and PC. Different letters indicate significant differences between groups ($p \leq 0.05$, Tukey's HSD test).

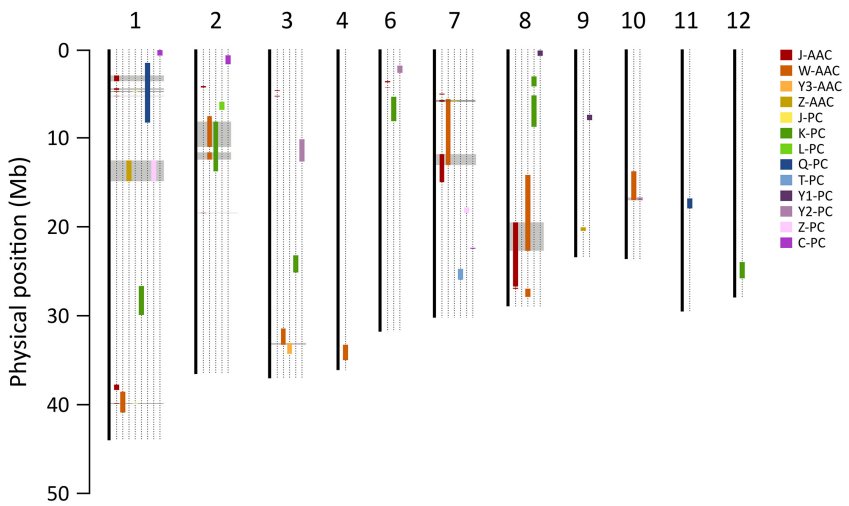


FIGURE 6 | Comparison of QTLs for PC and AAC in the present and previous studies. Color bars on chromosomes represent QTL regions for AAC and PC reported by the present and previous studies. Gray blocks indicate the loci where two more QTLs were co-located. Each prefix of trait indicates the previous study (J, this study; C, Chattopadhyay et al., 2019; K, Kinoshita et al., 2017; L, Lou et al., 2009; Q, Qin et al., 2009; T, Tan et al., 2001; W, Wang et al., 2007; Y1, Yun et al., 2014; Y2, Yu et al., 2009; Y3, Yoo, 2017; Z, Zhong et al., 2011).

the markers and an average of 0.93 markers within a 1-cM distance. Most of the studies for AAC and PC employed mapping populations from the crosses between *indica* and *japonica*. In the inter-subspecific crosses of rice, hybrid barriers, such as segregation distortion, hybrid sterility, and hybrid breakdown occur (Harushima et al., 2001; Jiang et al., 2008; Chin et al., 2011; Reflinur et al., 2014; Kim et al., 2017). Due to these reasons, several linkage gaps are shown in some chromosomal regions. Similarly, several large gaps existed in the genetic map constructed in this study. Lacking the proper number of markers or loci in the regions, the identification of some important QTLs might have not yet been identified and characterized.

A significant positive correlation among the majority of AAs were found in M23/T887 RILs (**Supplementary Figure S1**). This result is consistent with a previous study using Dasanbyeo (*indica*)/TR22183 (*japonica*) RILs (Yoo, 2017). In addition, it has been reported that several AACs were controlled by a single QTL in previous studies (Wang et al., 2007; Lu et al., 2009; Yoo, 2017). Similar results for several AACs have been found in this study (**Figure 2**). A total of 11 and 10 AACs, for instance, were affected simultaneously by *qAAC6.1* and *qAAC7.1*, respectively. These results suggest that these QTLs are related to the upstream regulator of AA biosynthesis, which would simultaneously affect contents of several AAs.

To validate the effects of *qAAC7.1*, M23-*qAAC7.1*^{T887}, possessing homozygous *qAAC7.1*^{T887} in the M23 genetic background, was selected (**Figure 5A**). This line showed improved contents of most AACs, except Thr, Pro, and Lys (**Figure 5B**). In addition to AACs affected by *qAAC7.1* in the primary QTL analysis, PC and other AACs, including Met, His, Lys, and Gly, were also significantly higher than those of both parents, indicating that *qAAC7.1*^{T887} could confer high AACs and PC consistently.

Transgressive segregations were observed for all AAC and PC traits, appearing values higher or lower than both parents (**Figure 1**). This result indicates that both parental lines have alleles that contribute to AACs and PC at several loci, and progenies that able to have higher contents by containing favorable alleles derived from each other donor parent.

The favorable alleles of QTLs for AAC and PC were evenly detected in both *japonica* and *indica* parents. For example, among three positive alleles of QTLs for His contents, *qAAC1.1* and *qAAC1.6* (total PVE = 11%) were derived from M23, while *qAAC1.2* (PVE = 13.7%) originated from T887 (**Supplementary Table S2**). Similarly, for *qPC1.1*^{T888}, *qPC1.2*^{M23}, and *qPC7.1*^{M23}, positive alleles for PC were also derived from each other parent. These results imply that pyramiding favorable alleles by inter-subspecies cross were an effective strategy to develop highly nutritional rice, improving AAC and PC of the grain. The QTLs for AACs and PC identified in this study would be useful in developing high nutritional rice.

DATA AVAILABILITY STATEMENT

The GBS data is available at NCBI under the accession numbers PRJNA601019 and PRJNA264250. Raw genotyping data of each RIL is available in a **Supplementary Data Sheet S1**.

AUTHOR CONTRIBUTIONS

SJ, C-KK, and JC designed the experiments. SJ, J-HH, YL, and N-HS conducted the experiments and collected the data. SJ, C-KK, YK, and JC analyzed the data. J-HH and N-HS collected the phenotypic data. SJ and JC wrote the manuscript. All authors read and approved the manuscript.

FUNDING

This work was supported by grants from the "Research Program for Agricultural Science and Technology Development" (Project No. PJ013485) in Rural Development Administration, South Korea.

ACKNOWLEDGMENTS

We specially thank Jeehyoung Shim for the contribution in the technical assistance in agronomic trait observation and material preparation for amino acid analysis, and Yeisoo Yu (DNACare Co., Ltd.) for the first-round informatics analysis on GBS results.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00240/full#supplementary-material

REFERENCES

Chattopadhyay, K., Behera, L., Bagchi, T. B., Sardar, S. S., Moharana, N., Patra, N. R., et al. (2019). Detection of stable QTLs for grain protein content in rice (*Oryza sativa* L.) employing high throughput phenotyping and genotyping platforms. *Sci. Rep.* 9, 1–16. doi: 10.1038/s41598-019-39863-2

Chin, J. H., Chu, S.-H., Jiang, W., Cho, Y.-I., Basyirin, R., Brar, D. S., et al. (2011). Identification of QTLs for hybrid fertility in inter-subspecific crosses of rice (*Oryza sativa* L.). *Genes Genom* 33, 39–48. doi: 10.1007/s13258-010-0100-z

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A Robust, Simple Genotyping-by-Sequencing (GBS) approach for high diversity species. *PLoS One* 6:e19379. doi: 10.1371/journal.pone.0019379

Harushima, Y., Nakagahra, M., Yano, M., Sasaki, T., and Kurata, N. (2001). A genome-wide survey of reproductive barriers in an intraspecific hybrid. *Genetics* 159, 883–892.

Henderson, J. W., Ricker, R. D., Bidlingmeyer, B. A., and Woodward, C. (2000). Rapid, accurate, sensitive, and reproducible HPLC analysis of amino acids. Agilent Technologies Technical Note 5980-1193E.

Jiang, W., Chu, S.-H., Piao, R., Chin, J.-H., Jin, Y.-M., Lee, J., et al. (2008). Fine mapping and candidate gene analysis of hwh1 and hwh2, a set of complementary genes controlling hybrid breakdown in rice. *Theor. Appl. Genet.* 116, 1117–1127. doi: 10.1007/s00122-008-0740-4

Jones, D. B. (1931). *Factors for Converting Percentages of Nitrogen in Foods and Feeds Into Percentages of Proteins*. Rome: *FAO*.

Kim, C.-K., Chu, S.-H., Park, H. Y., Seo, J., Kim, B., Lee, G., et al. (2017). Identification of heterosis QTLs for yield and yield-related traits in indica-japonica recombinant inbred lines of rice (*Oryza sativa* L.). *Plant Breed. Biotech.* 5, 371–389. doi: 10.9787/PBB.2017.5.4.371

Kinoshita, N., Kato, M., Koyasaki, K., Kawashima, T., Nishimura, T., Hirayama, Y., et al. (2017). Identification of quantitative trait loci for rice grain quality and yield-related traits in two closely related *Oryza sativa* L. subsp. japonica cultivars grown near the northernmost limit for rice paddy cultivation. *Breed. Sci.* 67, 191–206. doi: 10.1270/jsbbs.16155

Kosambi, D. D. (1943). The Estimation of Map Distances from Recombination Values. *Annals of Eugenics* 12, 172–175. doi: 10.1111/j.1469-1809.1943.tb02321.x

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Lin, R., Luo, Y., Liu, D., and Huang, C. (1993). *Determination and Analysis on Principal Qualitative Characters of Rice Germplasm. Rice Germplasm Resources in China*. Beijing: Agricultural Science and Technology Publisher of China, 83–93.

Lou, J., Chen, L., Yue, G., Lou, Q., Mei, H., Xiong, L., et al. (2009). QTL mapping of grain quality traits in rice. *J. Cereal Sci.* 50, 145–151. doi: 10.1016/j.jcs.2009.04.005

Lu, K., Li, L., Zheng, X., Zhang, Z., Mou, T., and Hu, Z. (2009). Genetic dissection of amino acid content in rice grain. *J. Sci. Food Agricult.* 89, 2377–2382. doi: 10.1002/jsfa.3731

Meng, L., Li, H., Zhang, L., and Wang, J. (2015). QTL IciMapping: integrated software for genetic linkage map construction and quantitative trait locus mapping in biparental populations. *Crop J.* 3, 269–283. doi: 10.1016/j.cj.2015.01.001

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4326. doi: 10.1093/nar/8.19.4321

Peng, B., Kong, H., Li, Y., Wang, L., Zhong, M., Sun, L., et al. (2014). OsAAP6 functions as an important regulator of grain protein content and nutritional quality in rice. *Nat. Commun.* 5:5847. doi: 10.1038/ncomms5847

Qin, Y., Kim, S.-M., and Sohn, J.-K. (2009). QTL analysis of protein content in double-haploid lines of rice. *Korean J. Crop Sci.* 54, 165–171.

Reflinur, Kim, B., Jang, S. M., Chu, S.-H., Bordiya, Y., Akter, M. B., et al. (2014). Analysis of segregation distortion and its relationship to hybrid barriers in rice. *Rice* 7:3. doi: 10.1186/s12284-014-0003-8

Sakai, H., Lee, S. S., Tanaka, T., Numa, H., Kim, J., Kawahara, Y., et al. (2013). Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. *Plant Cell Physiol.* 54:e6. doi: 10.1093/pcp/pcs183

Tan, Y. F., Xing, Y. Z., Sun, X. L., Zhang, Q. F., Corke, H., Sun, M., et al. (2001). Mapping quantitative trait loci for milling quality, protein content and color characteristics of rice using a recombinant inbred line population derived from an elite rice hybrid. *TAG Theor. Appl. Genet.* 103, 1037–1045. doi: 10.1007/s001220100665

Vetha-Varshini, P. A., Azhagusundaram, K., and Vijay Parveen, P. (2013). Brown rice—hidden nutrients. *J. Biosci. Technol.* 4, 503–507.

Wang, L., Zhong, M., Li, X., Yuan, D., Xu, Y., Liu, H., et al. (2007). The QTL controlling amino acid content in grains of rice (*Oryza sativa*) are co-localized with the regions involved in the amino acid metabolism pathway. *Mol. Breed.* 21, 127–137. doi: 10.1007/s11032-007-9141-7

Yoo, S.-C. (2017). Quantitative trait loci controlling the amino acid content in rice (*Oryza sativa* L.). *J. Plant Biotechno.* 44, 349–355. doi: 10.5010/JPB.2017.44.4.349

Yu, Y.-H., Li, G., Fan, Y.-Y., Zhang, K.-Q., Min, J., Zhu, Z.-W., et al. (2009). Genetic relationship between grain yield and the contents of protein and fat in a recombinant inbred population of rice. *J. Cereal Sci.* 50, 121–125. doi: 10.1016/j.jcs.2009.03.008

Yun, B.-W., Kim, M.-G., Handoyo, T., and Kim, K.-M. (2014). Analysis of rice grain quality-associated quantitative trait loci by using genetic mapping. *AJPS* 05, 1125–1132. doi: 10.4236/ajps.2014.59125

Zhong, M., Wang, L., Yuan, D., Luo, L., Xu, C., and He, Y. (2011). Identification of QTL affecting protein and amino acid contents in rice. *Rice Sci.* 18, 187–195. doi: 10.1016/S1672-6308(11)60026-7

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 handling Editor declared a past co-authorship with one of the authors, JC.

frontiers in Genetics

ORIGINAL RESEARCH
published: 02 April 2020
doi: 10.3389/fgene.2020.00289

Functional Genomic Validation of the Roles of *Soluble Starch Synthase IIa* in *Japonica* Rice Endosperm

Vito M. Butardo Jr.[1,2], Jixun Luo[1], Zhongyi Li[1], Michael J. Gidley[3], Anthony R. Bird[4], Ian J. Tetlow[5], Melissa Fitzgerald[6], Stephen A. Jobling[1] and Sadequr Rahman[1,7]*

[1] CSIRO Agriculture and Food, Canberra, ACT, Australia, [2] Department of Chemistry and Biotechnology, Faculty of Science, Engineering and Technology, Swinburne University of Technology, Hawthorn, VIC, Australia, [3] Centre for Nutrition and Food Sciences, The University of Queensland, St Lucia, QLD, Australia, [4] Nutrition and Health, CSIRO, Adelaide, SA, Australia, [5] Department of Molecular and Cellular Biology, College of Biological Science, University of Guelph, Guelph, ON, Canada, [6] School of Agriculture and Food Sciences, Faculty of Science, University of Queensland, St Lucia, QLD, Australia, [7] School of Science and the Tropical Medicine and Biology Platform, Monash University, Bandar Sunway, Malaysia

The enzyme starch synthase IIa (SSIIa) in cereals has catalytic and regulatory roles during the synthesis of amylopectin that influences the functional properties of the grain. Rice endosperm SSIIa is more active in *indica* accessions compared to *japonica* lines due to functional SNP variations in the coding region of the structural gene. In this study, downregulating the expression of *japonica*-type SSIIa in Nipponbare endosperm resulted in either shrunken or opaque grains with an elevated proportion of A-type starch granules. Shrunken seeds had severely reduced starch content and could not be maintained in succeeding generations. In comparison, the opaque grain morphology was the result of weaker down-regulation of *SSIIa* which led to an elevated proportion of short-chain amylopectin (DP 6-12) and a concomitant reduction in the proportion of medium-chain amylopectin (DP 13-36). The peak gelatinization temperature of starch and the estimated glycemic score of cooked grain as measured by the starch hydrolysis index were significantly reduced. These results highlight the important role of medium-chain amylopectin in influencing the functional properties of rice grains, including its digestibility. The structural, regulatory and nutritional implications of down-regulated *japonica*-type *SSIIa* in rice endosperm are discussed.

Keywords: starch synthase, amylopectin, amylose, rice, RNA silencing

OPEN ACCESS

Edited by:
Mallikarjuna Swamy, International Rice Research Institute, Philippines

Reviewed by:
Jinsong Bao, Zhejiang University, China
Naoko Crofts, Akita Prefectural University, Japan

****Correspondence:***
Vito M. Butardo Jr.
vbutardo@swin.edu.au

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *24 January 2020*
Accepted: *10 March 2020*
Published: *02 April 2020*

Citation:
Butardo VM Jr, Luo J, Li Z, Gidley MJ, Bird AR, Tetlow IJ, Fitzgerald M, Jobling SA and Rahman S (2020) Functional Genomic Validation of the Roles of Soluble Starch Synthase IIa in Japonica Rice Endosperm. Front. Genet. 11:289. doi: 10.3389/fgene.2020.00289

INTRODUCTION

Starch synthases (SS) play a critical role during starch biosynthesis, elongating glucan chains by the addition of glucose from the substrate ADP-glucose (Denyer et al., 1995; Denyer et al., 1996). These elongated glucan chains act as substrates for branching enzymes and debranching enzymes (Tetlow and Emes, 2014; Pfister and Zeeman, 2016). Cereals possess some isoforms of SS with varying substrate affinities and catalytic activities and of these, the endosperm-specific enzyme SSIIa has a particularly important role in starch biosynthesis. The gene for starch synthase IIa (*SSIIa*) codes for a major starch synthase enzyme isoform involved in the elongation of short-chain amylopectin in the cereal endosperm, which is important in distinguishing starch properties (Hannah and James, 2008; Jeon et al., 2010; Tetlow and Emes, 2017). In rice, the *SSIIa* gene [also

known as *acl(t)*, *alk*, *gel(t)* and *SS2-3* gene in other studies] was originally mapped to the *alk* locus located on chromosome 6 (Umemoto et al., 2002; Gao et al., 2003; Umemoto et al., 2004) and is highly expressed during grain development (Hirose and Terao, 2004; Ohdan et al., 2005). *SSIIa* alleles determine the peak gelatinization temperature (GT) of rice, an essential trait in predicting cooking and eating qualities (Umemoto and Aoki, 2005; Waters et al., 2006).

Several single nucleotide polymorphisms (SNPs) located along the SSIIa coding gene have been linked with varietal differences in GT due to variations in amylopectin chain length distribution (CLD) (Umemoto et al., 2002; Umemoto and Aoki, 2005; Bao et al., 2006; Waters et al., 2006; Bao et al., 2009; Cuevas et al., 2010). Using the SSIIa sequence from the *indica* line Kasalath as the canonical protein sequence, previous research has demonstrated that substitution of either Valine-737 with Methionine (due to SNP3, which is G in Kasalath and A in Nipponbare at 2209 bp from the start codon), or Leucine-781 with Phenylalanine (due to SNP4, which is GC in Kasalath and TT in Kinmaze at 2340–2341 bp from the start codon) consequently lead to a reduction in specific activity of less than 10% compared to that of the *indica* sequence (Nakamura et al., 2005; Umemoto and Aoki, 2005). Substitutions of Methionine and Valine are common in the *japonica* lines. Thus, higher proportions of shorter amylopectin chain (S-type) is common among *japonica* rice lines such as Nipponbare because its SSIIa is weakly active, making the enzyme less efficient in catalyzing the elongation of short amylopectin chains, which leads to low GT (Umemoto et al., 1999; Nakamura et al., 2002, 2005; Umemoto and Aoki, 2005; Waters et al., 2006; Cuevas et al., 2010). In contrast, a higher proportion of longer amylopectin chain (L-type) is common among *indica* rices such as IR64 because its SSIIa enzyme is catalytically active (Nakamura et al., 2005). This results in elongation of short amylopectin chains and hence the increase in GT observed in the grain starch of *indica* rice accessions. Complementation of *japonica* S-type amylopectin (weakly active) in Nipponbare by *indica SSIIa* (active) from Kasalath produced *indica* L-type amylopectin (Nakamura et al., 2005).

Another potential consequence of amino acid substitutions due to SNP3 and SNP4 is on the ability of rice SSIIa to associate with starch granules (Umemoto and Aoki, 2005). Rice grains belonging to *japonica* types (haplotypes 3 and 4) have similar levels of SSIIa in the soluble phase but reduced levels in the starch associated protein fraction compared to those belonging to *indica* types (haplotypes 1 and 2) (Umemoto et al., 2004; Umemoto and Aoki, 2005; Waters et al., 2006; Bao et al., 2009). Furthermore, rice grains belonging to *indica* types are also observed to have higher amounts of starch-associated starch branching enzyme IIb (SBEIIb) compared to *japonica* types (Umemoto and Aoki, 2005) due to the SSIIa isoform present. This observation was confirmed by the association of *SSIIa* alleles with the relative distribution of SBEIIb and SSI between the starch granule and amyloplast stroma of rice (Luo J. et al., 2015). Additionally, following the multi-enzyme starch biosynthetic complex model in cereal endosperm proposed by Liu F.S. et al. (2012) and Tetlow et al. (2015), it is believed that SSIIa plays a scaffolding role in the formation of the complex (Umemoto et al., 2004; Nakamura et al., 2005; Umemoto and Aoki, 2005). Additionally, the presence of SSIIa appears to be important for the association of other proteins such as starch synthase I (SSI) and SBEIIb in starch granules (Liu F.S. et al., 2012). Clearly, therefore, SSIIa has diverse roles in starch biosynthesis by virtue of its enzymatic, scaffolding and stromal distribution functions during starch biosynthesis in cereals (Miura et al., 2018). All these findings highlight the importance of SSIIa in determining fine amylopectin structure and the resulting functional properties of rice grain.

In rice, an SSIIa mutant from the *japonica* line Kinmaze was identified completely devoid of *SSIIa* expression. This SSIIa null mutant, generated using *N*-methyl-*N*-nitrosourea, had a 4% increased amylose content, and a reduction of about 6°C in the gelatinization temperature (Miura et al., 2018). The 4% increase of amylose in the SSIIa null *japonica* rice resulted in the same levels of amylose comparable to that of indica rice lines. Loss of SSIIa in barley produced shrunken grains with high amylose and elevated resistant starch contents and reduced starch digestibility (Morell et al., 2003; Topping et al., 2003). It will be interesting to determine the phenotypic outcome of gradual reductions in the amount of weakly active SSIIa in rice grains with *japonica* background such as Nipponbare using transgenic approaches. Gradually reducing SSIIa expression by RNA silencing may aid in the further clarification of its roles during starch biosynthesis. In this study, the already low expression of weakly active *japonica*-type SSIIa in Nipponbare endosperm was further down-regulated to determine any impact on its possible roles in amylopectin biosynthesis and starch structure. Our results, using a different *japonica* rice line (Nipponbare) and a different technique (RNA silencing), are in broad agreement with the results obtained by Miura et al. (2018).

MATERIALS AND METHODS

Downregulating *SSIIa* Expression

A hairpin RNA (hp-SSIIa) was constructed to downregulate the expression of SSIIa in the rice endosperm of *japonica* rice (Nipponbare cultivar) using a technique effective in partially downregulating the expression of *SBEIIb* in rice endosperm (Butardo et al., 2011). Total RNA was extracted from Nipponbare rice endosperm and used as the template for the synthesis of total cDNA (see next section). A 443 bp target sequence was amplified from the cDNA using forward (5′-GCTACCTCTGGGAGCTGAAGACGACGGAG-3′) and reverse primers (5′-GGGTGGGGTTCTCGGTGAAGA-3′) targeting position 1628-2071 of *japonica* rice SSIIa based on Nipponbare reference genome. The PCR fragment was cloned in pGEM-T Easy and subcloned into pBX17 and pVec8 using previously published methods (Butardo et al., 2011). The generated constructs were driven by wheat high molecular weight glutenin (wHMWG) promoter to ensure seed-specificity as previously demonstrated (Butardo et al., 2011). The generated constructs were verified by restriction digestion and DNA sequencing after every subcloning and transformation steps to ensure that the correct hairpin RNA sequences

were maintained in the correct orientation. In addition, two artificial microRNAs (TTACAAAACAGAATCGTGGGC and TTAAGCGATATTATGTATCAC), also driven by wHMWG promoter were constructed using a technique which was previously shown to be more effective than hp-RNA in completely down-regulating the expression of SBEIIb in rice endosperm (Butardo et al., 2011). The Nipponbare rice calli were transformed with the silencing constructs using *Agrobacterium tumefaciens* AGL1 as previously described (Butardo et al., 2011). Tissue culture transformation, regeneration and selection were also as previously described (Butardo et al., 2011).

Regenerated transformed plants and negative controls were grown in pots partially submerged in water-filled tanks to simulate irrigated conditions inside a biosafety glasshouse. The temperature was maintained at 22°C during the night for 8 h and 29°C during the day time for 16 h of natural light. The actual average daily temperature was observed to be 26.5 ± 3.5°C and is consistent with our previous study (Butardo et al., 2011). Initial screening for putative transformants was done by PCR detection of a fragment of hygromycin resistance gene from genomic DNA of 1-month-old leaves extracted by FastDNA Kit (Q-BIOgene). The putative transformants were verified using gene-specific primers that amplify a hybrid PCR fragment containing a portion of the wheat high molecular weight glutenin (wHMWG) and a portion of the forward hairpin fragment. PCR amplification was carried out using HotStar Taq (Qiagen) using Hyper Ladder IV (Bio Line) as molecular weight standards.

Gene Expression Analyses

RNA was extracted from grains at mid-development (15 dpa) using NucleoSpin RNA Plant (Macherey-Nagel). A total of 5 μg RNA template was used to synthesize cDNA using SuperScript III reverse transcriptase (Invitrogen). Quantitative real-time PCR (qRT-PCR) was done in a Rotor-Gene 6000 (Corbett) using 100 ng cDNA templates amplified utilizing previously published SSI and SSIIa primer pairs (Hirose and Terao, 2004; Ohdan et al., 2005; Yamakawa et al., 2007). The real-time PCR amplification was conducted using Platinum Taq DNA polymerase (Invitrogen) and Sybr Green I (Invitrogen) reporter dye. Comparative quantitation was conducted using α-tubulin as a reference gene (Toyota et al., 2006), with data validation and melt curve analysis done using the Real-Time Rotary Analyzer Software (Corbett).

Protein Expression Analyses

Native soluble proteins from developing rice grains of selected homozygous plants (15 dpa) were extracted as previously described (Regina et al., 2006). This developmental stage ensures the maximal expression of *SSIIa* in wild-type rice grains, which was used to screen for SSIIa-downregulated lines. A total of 100 μg protein, quantified using Coomassie Protein Assay Reagent (Bio-Rad), was resolved in 4–10% precast gradient gels (Invitrogen). Total proteins were extracted from 5 to 20 dpa developing grains using a previously published method to track *SSIIa* expression across several developmental time points. In addition, soluble, granule-associated and granule-bound proteins from starch granules of mature grains (32 dpa) were also extracted as previously described (Butardo et al., 2012). For granule-bound proteins, 4 mg of starch for each sample was used for the extraction of proteins as described by Luo J.X. et al. (2015). Mature rice grains of selected SSIIa-downregulated lines were used to verify the silencing of the gene. Gels were blotted onto nitrocellulose membrane to detect SSI, SSIIa, and GBSSI using the appropriate antisera at 1:2000 dilution. Preliminary screening of SSIIa in total soluble protein extracts of T1 rice endosperms was conducted using anti-wheat polyclonal antibodies at 1:500 dilution but succeeding western blots for SSIIa were done using rice anti-SSIIa polyclonal antibodies due to improved specificity (Butardo et al., 2012). The immunoreactive proteins were probed by goat anti-rabbit or anti-mouse immunoglobulins conjugated to horseradish peroxidase (Bio-Rad). Antibody detection was carried out using ECL Western Blotting Detection Reagents (GE Healthcare Life Sciences) and Hyperfilm ECL chemiluminescence film (Amersham Biosciences). The film was developed using a CP 1000 automatic film processor (Agfa). Bands that were immunoreactive to anti-rice SSI and anti-wheat GBSSI antiserum were used to verify protein normalization in all Western blot experiments. Electroblotted gels were also stained with Sypro to verify the exact location of protein bands (see **Supplementary Figure 2** as an example). IR64, an *indica* rice variety with catalytically active and highly expressed SSIIa were used as the positive control. This control rice line was grown with the Nipponbare transgenic and wild-type rice lines at the same time and in the same phytotron glasshouse to ensure uniformity of response.

Grain Screening and Characterization

Only grains that have the hairpin RNA insert detected by PCR and downregulated SSIIa as detected by western blot as described above, were selected and subsequently characterized. Mature panicles of transgenic plants were harvested and dried at 37°C overnight. The opaque seeds were carefully threshed by hands, and then dehulled and polished by machine using standard methods. Shrunken seeds were also manually threshed and dehulled but they were not polished because they shatter into powder after passing through the milling machine. Consequently, all succeeding starch structural and functional analyses were done in polished (white) and wholemeal flours (brown) for opaque lines and only in wholemeal flours for shrunken lines. Photomicrographs of whole rice grain samples were obtained using either a LEITZ M8 stereomicroscope or by scanning in Image Scanner III (GE Healthcare Health Sciences). Some seeds were set aside for planting in subsequent generations.

Carbohydrate and Digestibility Analyses

Peak GT from rice flour was measured by differential scanning calorimetry (Cuevas et al., 2010). Total starch content was determined using a 96-well plate format of a Megazyme assay procedure (AACC Method 76.13). The apparent amylose content (AAC) of flour samples was determined by iodine colorimetry using an AutoAnalyser3 Digital Colorimeter (Brann + Luebbe, United States). Data analysis was performed using the colorimeter's Automated Analyser Control and

Evaluation Software. Rice flour samples of IR24, IR64 and IR8 were used as amylose calibration standards and check controls.

The resistant starch (RS) content and estimated glycemic score (EGS) of freshly cooked polished rice grains were estimated using an *in vitro* starch hydrolysis index (HI) method which mimics the oral and gastrointestinal phases of carbohydrate assimilation in humans (Butardo et al., 2011). The HI method was comprehensively validated against *in vivo* clinical measurements in adult volunteers (Fitzgerald et al., 2011). A total of 50 and 500 mg of available carbohydrates were used to predict EGS and RS, respectively. For EGS prediction, aliquots of supernatant from starch hydrolyzates were sampled at designated regular time intervals for up to 5 h and glucose concentration determined using an automated electrochemical procedure (Butardo et al., 2011; Fitzgerald et al., 2011). Because the rice grains generated in this study were transgenic, which required special biosafety clearance for human consumption, the digestibility values obtained were used as highly correlated proxy measures for glycemic impact upon rice grain consumption based on previous work (Butardo et al., 2011; Fitzgerald et al., 2011).

Starch Structural Determination

Determination of chain length distribution (CLD) of debranched amylopectin by fluorescence-assisted capillary electrophoresis (FACE) was performed based on a previous method (O'Shea and Morell, 1996). Molecular weight distribution (MWD) of debranched starch was determined by size-exclusion chromatography (SEC) as previously described (Castro et al., 2005). Precisely 10.0 mg of flour sample was gelatinized and debranched with isoamylase (Megazyme, Ireland) at 50°C. Immediately after debranching, samples were spun down using a microcentrifuge at room temperature. Aliquots for CE (50 μL) were obtained from the debranched supernatant and dried using a speed vacuum at 50°C for 2 h or until the pellet was completely dried. In addition, aliquots for SEC (750 μL) were obtained and desalted using AG 501-X8 (D) resin (Bio-Rad) for 30 min in 50°C water bath with occasional mixing by inversion every 10 min. Each dried pellet for CE analysis was labeled for 16 h with the 3.5 μL 0.2M 9-aminopyrene-1,4,6-trisulfonate (APTS) and analyzed in Beckman P/ACE system as previously described (Cuevas et al., 2010). On the other hand, 40 μL of each desalted sample was loaded into Alliance 2695 SEC machine (Waters, United States), resolved using an Ultrahydrogel 250 column (Waters, United States) using 0.05M NH4OAc pH 4.75 with 0.02% sodium azide as mobile phase and detected using 2414 Refractive Index detector (Waters, United States). For CE, the chain length distribution was determined from the peak area by converting to the velocity area to give *N(X)* (Demorest and Dubrow, 1991). For SEC, molecular weight distribution (MWD) was estimated from elution time using pullulan standards (Shodex P-82) calibrated with the Mark-Houwink-Sakaruda equation and universal calibration (Castro et al., 2005). A waxy rice was used to differentiate the debranched amylose and amylopectin regions with a cut-off at DP 120 (Butardo et al., 2011; Butardo et al., 2017). The following four MWD regions from debranched starch were identified: true amylose chains (DP > 1,000), long-chain amylopectin (DP 121-1000), medium-chain amylopectin (DP 37-120) and short-chain amylopectin (DP 6-36) as previously defined (Butardo et al., 2011; Butardo et al., 2017).

Starch Granule Analyses

Cross-sections of transgenic rice grains that differed significantly from the controls were observed uncoated with an environmental scanning electron microscope (ESEM, Zeiss EVO LS15) under variable-pressure mode. Images of starch granules were taken with a back-scattered electron detector. Starch granules were also isolated and viewed under a polarized light microscope after iodine staining to check for birefringence. Granule size distribution (by volume) of the starch slurries was determined using a laser diffraction particle size analyzer (Mastersizer 2000, Malvern Instruments, Malvern, United Kingdom). The percentage of small starch granules was determined using a cut-off diameter of 1.9 μm (refer to the results section). The diameter of large starch granules (>1.9 μm) was calculated as the diameter of starch granules at the peak of large starch granules. Characterization of starch crystallinity by x-ray diffraction (XRD) was carried out on a Panalytical X'Pert Pro diffractometer using the crystal defect method based on our previous publication (Lopez-Rubio et al., 2008). Solid-state ^{13}C cross polarization/magic angle spinning (CP/MAS) nuclear magnetic resonance (NMR) experiments were performed at a ^{13}C frequency of 75.46 MHz on a Bruker MSL-300 spectrometer also as previously described (Butardo et al., 2011).

Sampling and Statistical Analyses

Three biological replicates from at least two independent transformed lines were used for every analysis whenever applicable during the screening of hp-SSIIa-shr and hp-SSIIa-op. Mature grains of hp-SSIIa-op were used for whole grain and starch analyses for five generations (T1 to T5), planted and harvested at different seasons to assess stability and replicability of traits. Gene and protein expression profiling was performed on later generations (T3 to T5). A comparison of results was done using Nipponbare negative segregant as control. Statistical analyses (one-way analysis of variance with Tukey post-test, two-way analysis of variance with Bonferroni post-test, and unpaired *T*-test) were done using GraphPad Prism Version 6. The standard error of the mean (SEM) was used to representing error values. Statistical significance was defined at least as $P < 0.05$.

RESULTS

Screening of Transgenic Lines

A total of 35 independent transformants were obtained from hp-SSIIa lines, of which 31 (88%) harbored the hygromycin resistance gene based on marker screening by PCR. A total of 5 out of 31 lines (16%) of hp-SSIIa were selected at T1 generation based on corroborative results from the seed appearance (**Figure 1**) and PCR screening using hygromycin resistance gene marker. Three of the selected segregating hp-SSIIa lines (9.6%) had shrunken seeds (SS9, SS17, SS28), while the other two lines (6.4%) had chalky to opaque seeds (SS3

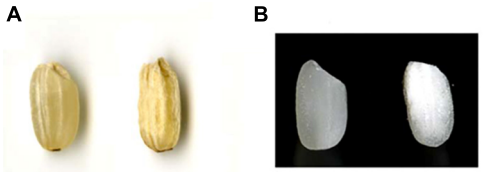


FIGURE 1 | Seed morphology of representative grain samples harboring hp-SSIIa construct compared with the control. **(A)** Unpolished grain of a negative segregant (left) compared with hp-SSIIa-shr line (SS28), showing shrunken seed phenotype. **(B)** Polished grain of a negative segregant (left) compared with an hp-SSIIa-op line (SS5), showing opaque phenotype.

and SS5). These selected lines were designated as hp-SSIIa-shr and hp-SSIIa-op, respectively, with representative seed samples shown in **Figure 1**. The hp-SSIIa-op grains are amenable to gentle polishing as they tend to fissure and break using standard milling methods. The two hp-SSIIa-op lines were viable in succeeding generations and they were stably maintained beyond T5 generation without losing the observed phenotypic traits. In contrast, grains of hp-SSIIa-shr cannot be polished because the shrunken seeds are brittle and shatter into powder during milling. In addition, the shrunken seeds of the hp-SSIIa-shr lines were either sterile or they did not fully mature in the succeeding generations. Those that grew at T2 stage had low germination rates and they needed to be grown by tissue culture. All the hp-SSIIa-shr lines were sterile beyond the T2 generation, with most spikelets empty or filled with transparent watery grain. As both the hp-SSIIa-shr and hp-SSIIa-op phenotypes were generated from multiple independent transgenic events, it is highly unlikely that the phenotype is due to the chance perturbation of a gene critical for the grain development. We demonstrate below that the phenotype is due to the reduction of SSIIa.

A total of 23 independent transformants were obtained that harbored artificial microRNAs targeted toward SSIIa (ami-SSIIa). Three ami-SSIIa lines cloned using the osa-miR528 backbone using the pWBVec8 expression vector were selected at T1 generation based on grain appearance and PCR screening. However, the lines could not be maintained beyond T2 generations. Consequently, the results for hairpin RNA downregulation of *SSIIa* expression in Nipponbare rice endosperm are the only ones reported in this study.

Characterization of Starch Granules

The starch granules of hp-SSIIa-shr T1 seeds were severely distorted during development compared with the controls. The granules were rounded and had lost their compound structure (**Figure 2**). Furthermore, the starch granules located toward the middle region of the grain were very small and their development appeared to be aborted (**Figures 2A,B**). Starch synthesis appears severely hampered based on seed morphology (**Figure 1**), starch granule appearance (**Figures 2A,B**) and reduced total starch content (**Table 1**). The starch granules of some less shriveled seeds have formed into a complex quaternary structure but they still lost their angularity (data not shown).

In contrast to the phenotypic consequence observed in hp-SSIIa-shr seeds, most of the starch granules of hp-SSIIa-op are rounded but they have maintained their compound structure (**Figures 2C,D**). The distribution of small starch granules in these opaque samples was reminiscent of the hp-SSIIa shrunken starch grain phenotype described above (**Figures 2A,B**). To provide more quantitative results, one line (SS5) was tested using a particle size analyzer which confirmed that it has a higher proportion of smaller starch granules compared to Nipponbare, approximately 1–4 μm in size (**Figure 3**). In addition, it also has higher proportion of bigger starch granules at 10–11 μm in size compared to its parental line. In contrast to the unimodal starch granule size distribution of hp-SSIIa-op, Nipponbare had a bimodal distribution, with a higher proportion of bigger starch granules 4–10 μm in size (**Figure 3**). T4 grains of hp-SSIIa-op (SS5) were obtained from the selfed panicles in order to verify that the observed starch granular organization was maintained. The small granules were still more pronounced toward the middle of the grain (data not shown).

The particle size distribution of starch granules in hp-SSII-shr was not determined due to the sample size limitation. However, based on the starch granule morphology (**Figures 2A,B**), it is expected that the proportion of its smaller-sized starch granules would be elevated even when compared to hp-SSIIa-op.

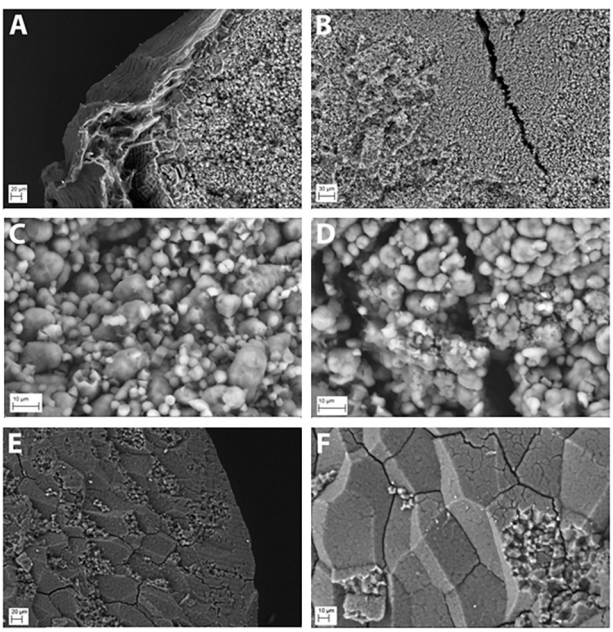


FIGURE 2 | Starch granule morphology of representative hp-SSIIa-shr shrunken (SS28) and hp-SSIIa-op opaque (SS5) lines compared with the control. The starch granules of hp-SSIIa-shr, **(A)** showing the edge and **(B)** the middle section of an unpolished grain, demonstrating that severely distorted seed development is due to aborted starch granule formation. The starch granules of hp-SSIIa-op, **(C)** showing mixture of rounded compound and simple starch granules, **(D)** with some very small starch granules in the middle of the grain. The parent Nipponbare is used as a control, **(E)** showing the edge of a polished grain **(F)** and the middle showing compact and angular compound starch granules.

TABLE 1 | Functional properties of hp-SSIIa lines compared with the controls.

Lines	Apparent Amylose (%)	Total Starch (%)	Peak GT (°C)	Resistant Starch (%)**	Glycemic Index (predicted)
hp-SSIIa-shr*	6.0 ± 0.8^{a}	73.0 ± 2.1^{a}	71.8 ± 0.0^{a}	ND	ND
Nipponbare Brown	7.2 ± 0.0^{a}	80.7 ± 1.4^{c}	ND	ND	ND
hp-SSIIa-op	13.8 ± 0.6^{b}	92.9 ± 2.4^{b}	70.4 ± 1.1^{a}	0.2 ± 0.1	63.5 ± 3.5^{a}
Nipponbare Polished	15.2 ± 0.8^{b}	90.4 ± 0.9^{b}	73.9 ± 0.6^{b}	0.2 ± 0.0	85.2 ± 1.2^{b}

**Unpolished rice grains were used because shrunken grains are brittle and not amenable to polishing. **Using CSIRO Method. ND, not determined. Data marked with the different letters are significantly different at $P < 0.05$.*

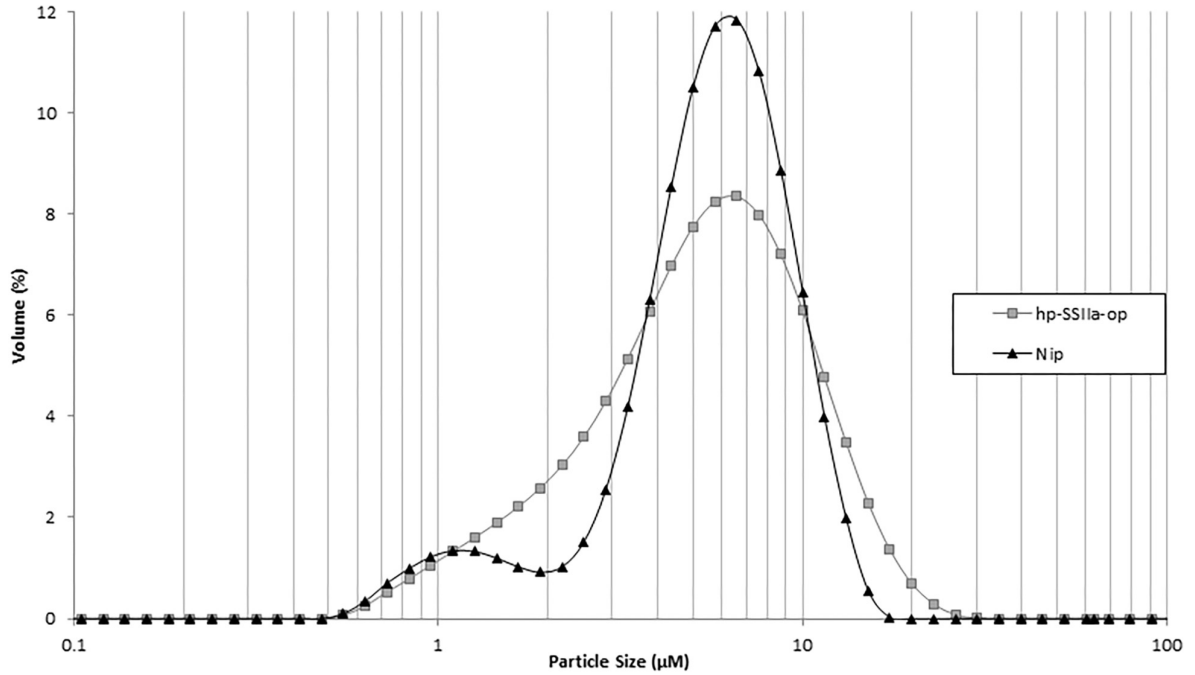


FIGURE 3 | Particle size analysis of representative hp-SSIIa-op (SS5) in comparison with the control Nipponbare (Nip). The *y*-axis represents the amount of each size of starch granules as a percentage of total starch. The *x*-axis represents the size of starch granules. The names of lines are labeled on the right of the graph.

Despite pronounced alterations in starch granule organization, the hp-SSIIa-op lines retained their A-type crystalline polymorph similar to its wild-type parent Nipponbare (**Supplementary Figure 1**). The crystallinity of hp-SSIIa-op as estimated by ^{13}C CP/MAS NMR also did not vary significantly from that of Nipponbare (**Supplementary Table 1**). Lastly, all the starch granules exhibited normal birefringence similar to that of the control (data not shown).

Characterization of Grain and Starch Functional Properties

Table 1 summarizes the results of the functional assays of hp-SSIIa lines compared with the controls. Analyses for hp-SSIIa-shr lines were done on unpolished T1 grains as they were brittle and shattered upon polishing, while the analyses of hp-SSIIa-op were done on polished grains for reasons described earlier, although they were more prone to over milling due to softer grains. The apparent amylose content of polished hp-SSIIa-op lines was not significantly different from that of the control. Similarly, the starch from hp-SSIIa-shr lines had percentage amylose comparable with that of unpolished control grains. The total starch content including the soluble glucan of hp-SSIIa-shr was reduced by 8% compared to Nipponbare brown rice control (P-value < 0.0001), while that of hp-SSIIa-op had no reduction compared to polished Nipponbare control. The peak gelatinization temperature of our rice lines are quite high compared to other reference values for Nipponbare in literature, which is 66.5°C in Umemoto et al. (2008), but closer to 68.0°C in Waters et al. (2006). We attribute the difference in GT values to variations in DSC machines and calibrations methods employed. Nonetheless, the peak gelatinization temperature of the shrunken and opaque lines of hp-SSIIa was reduced by at least 2°C compared to Nipponbare, and this was found to be statistically significant (P-value = 0.0320). Lastly, the resistant starch content of the hp-SSIIa-op lines was unaltered while the predicted glycemic index by *in vitro* hydrolysis index was significantly reduced (P-value = 0.0005). The hp-SSIIa-shr lines were not included in the estimation of resistant starch content and *in vitro* hydrolysis

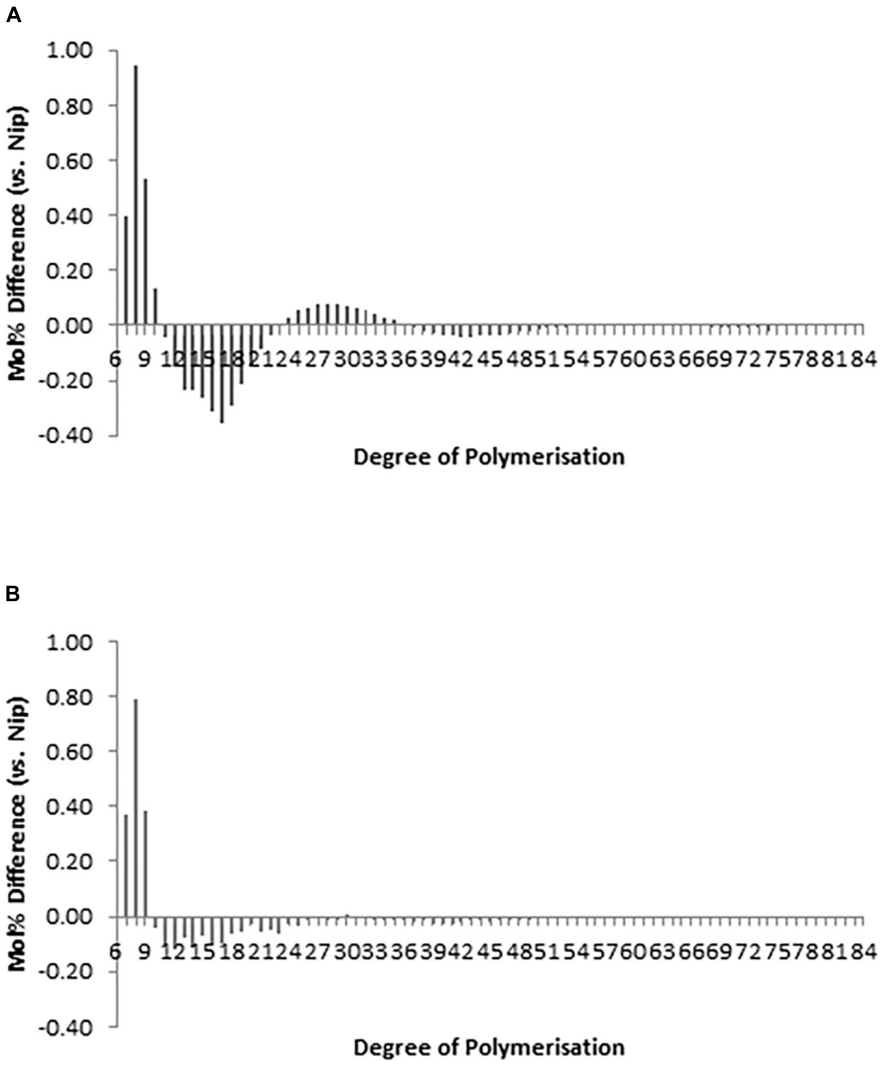


FIGURE 4 | Chain length distribution (CLD) profile of debranched amylopectin from representative **(A)** hp-SSIIa-op (SS5) and **(B)** hp-SSIIa-shr (SS28) lines at T1 generation, showing significant elevation in short amylopectin A-chains from DP 6-10 and reduction from DP 12-24 ($P = 0.0004$). The results are average of three trials. Comparison of CLD is expressed as a difference plot (percentage of molar molecules) between each mutant and control Nipponbare.

index due to sample size limitations and also because they were not amenable to polishing.

Starch Structural Characterizations

The chain-length distribution (CLD) characterization of debranched amylopectin from selected seeds based on appearance, western blot and PCR screening of hp-SSIIa-op opaque lines (**Figure 4A**) and hp-SSIIa-shr shrunken lines (**Figure 4B**) revealed a significant increase in the proportion of short amylopectin chains (DP 6–12). In addition, a small decrease in longer chains of DP 12–24 was detected in both hp-SSIIa-op and hp-SSIIa-shr lines (**Table 1**). This CLD is corroborated by molecular size distribution analysis of normalized SEC traces from debranched starch which revealed significant elevation of short chains of apparent DP 6–12 in shrunken (3.2 ± 0.1 mol%) and opaque seeds (2.9 ± 0.2 mol%) compared with the control (2.6 ± 0.1 mol%) (**Figures 5A,B**). This shift in the proportion of short-chain amylopectin is accompanied by a significant reduction in medium-chain amylopectin of apparent DP 13–36 for shrunken (52.4 ± 1.7 mol%) and opaque seeds (54.9 ± 0.6%) compared with the control (56.6 ± 0.2%) (**Figure 5C**). There was no change observed in the proportion of true amylose chains (apparent DP > 1000) and long-chain amylopectin (DP 121-1000). Similar shifts in chain length and molecular size distribution profiles of amylopectin were further confirmed in homozygous generations of hp-SSIIa-op SS5 line at the T4 generation (data not shown).

Determination of SSIIa Protein Inside Starch Granules

T4 grains of hp-SSIIa-op SS5 were further characterized to determine whether SSIIa was detectable inside the starch granules

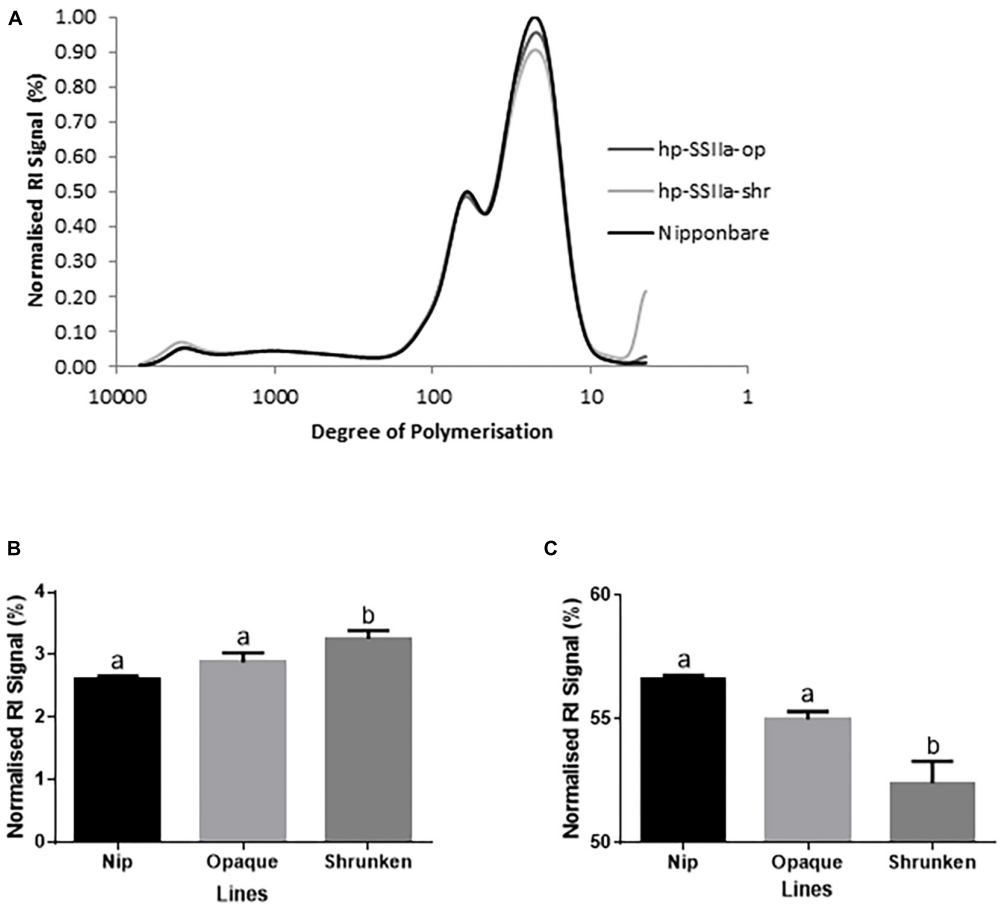


FIGURE 5 | Molecular size distribution profile of debranched starch in T1 generation by SEC **(A)** for representative lines of hp-SSIIa-op (SS3 and SS5) and hp-SSIIa-shr (SS17 and SS28) compared with Nipponbare control. **(B)** Short chain amylopectin (DP 6–12) were elevated (P-value < 0.05) and **(C)** medium chain amylopectin (DP 13–36) were reduced (P-value < 0.02) in the transgenic lines compared to Nipponbare control. The profiles show the distribution of normalized refractive index (RI) signals of debranched starches from each sample. Mean comparison with control using Bonferroni multiple comparison test revealed that only the shrunken lines have statistically significant result. Bars with different letters indicate significantly different mean. The identity of samples is indicated on the right side of the figure.

after *SSIIa* downregulation by hairpin RNA silencing. Western blot detection of granule-bound proteins of SS5 revealed that SSIIa was undetectable (**Figure 6A**). In contrast, a small amount of SSIIa was detected in Nipponbare, while its concentration in the *indica* control (IR64) was very high. As expected, the amount of GBSSI was higher in IR64 compared to Nipponbare and SS5, while SSI appeared to be similar in the granule-bound proteins of the three lines tested. SSIIa was not detected in the soluble protein fraction from flour samples of mature desiccated grains of the three lines tested (**Figure 6B**). The amount of SSI appeared similar in the three lines tested, indicating proper protein normalization (**Supplementary Figure 2**). The amount of SBEIIb and GBSSI inside the starch granules appeared to be higher in IR64 compared to that of the other two lines (**Supplementary Figure 2**).

Comparison of the amounts of SSIIa present in the total protein extracts of developing endosperms of SS5 (5, 10, 15, 20 dpa) confirmed that the level of SSIIa in the total protein extract was partially down-regulated compared to their corresponding developmental stages of Nipponbare control, as well as compared to the 15 dpa IR64 used as a reference band (**Figure 7A**). The amount of GBSSI in SS5 appeared up-regulated compared to Nipponbare (**Figure 7B**). Western-blotting analyses (**Figure 6B**) and transcript analyses of 15 dpa grains by qRT-PCR (**Supplementary Figure 3**) revealed that the amount of SSIIa in SS5 was half of that in Nipponbare, and the amount of SSI was two-fold of that in Nipponbare.

DISCUSSION

Our results revealed that endosperm down-regulation of *japonica*-type SSIIa resulted in distinct seed and starch granule morphologies, elevation in the proportion of short-chain amylopectin, and the reduction of peak gelatinization temperature and estimated glycemic score of cooked grains. Nipponbare is a *japonica* variety with an SSIIa belonging to haplotype 3 (Waters et al., 2006). This *japonica* variety is believed

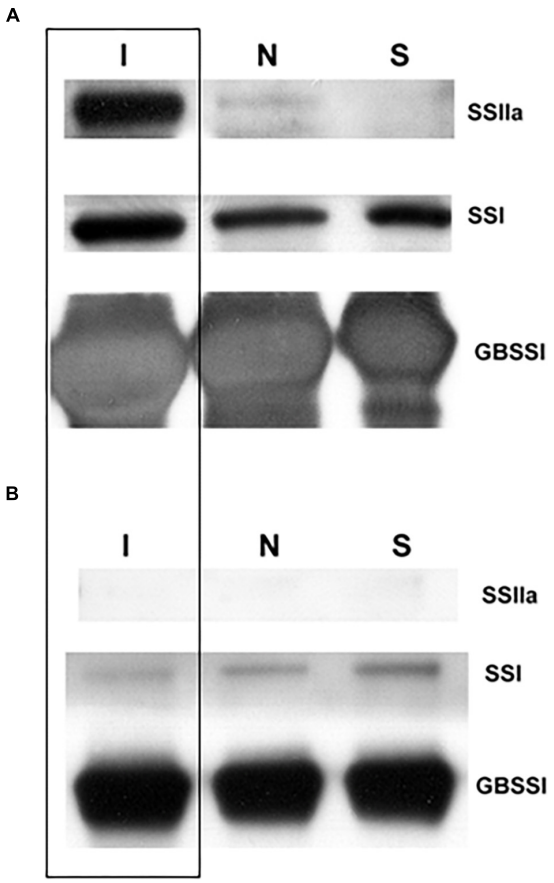


FIGURE 6 | Western blot analysis of SSIIa, SSI and GBSSI in **(A)** granule-bound and **(B)** soluble proteins of hp-SSIIa-op (SS5) (S) compared with its parent Nipponbare (N) and an *indica* positive control IR64 (I) (boxed) using mature grains.

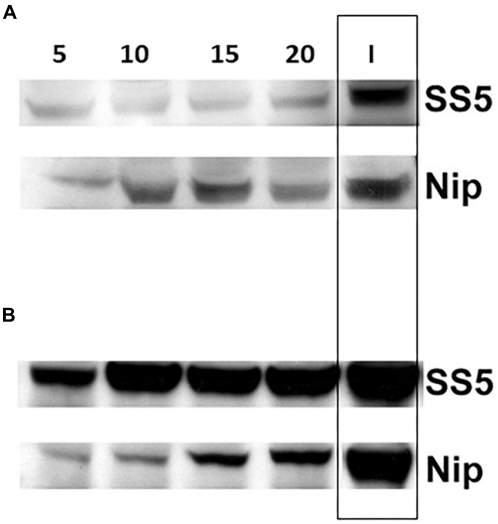


FIGURE 7 | Western blot analysis of **(A)** SSIIa and **(B)** GBSSI protein expression in the total protein extract from developing grains of hp-SSIIa-op (SS5) compared with Nipponbare (Nip) using 5–20 DPA grains. 15 DPA IR64 (I) was used as a reference positive control band (boxed). Days of each sample is labeled above the lanes (5, 10, 15, and 20).

to have a weakly active SSIIa enzyme (Umemoto et al., 2004; Nakamura et al., 2005). However, functional genomic validation presented here showed that the downregulation of SSIIa in Nipponbare led to pronounced changes in seed appearance, starch granule morphology, amylopectin fine structure, as well as thermal and digestibility properties (**Table 1**). These results demonstrate that SSIIa in Nipponbare either has a critical residual catalytic activity or possesses additional functions that are affected at these low concentrations as will be elaborated below.

Effect of Downregulating SSIIa on Starch Granule Structure and Organization

The proportion of debranched amylopectin at DP 13-36 was further reduced in the SSIIa down-regulated lines compared to Nipponbare. It is highly likely that the alteration in seed appearance, starch granule morphology, and starch granule size distribution are the phenotypic consequences of this perturbation in amylopectin structure. These observations are in agreement with previous studies in null SSIIa mutants of wheat, barley and maize (Liu F. et al., 2012; Luo J. et al., 2015). The endosperm starch from these SSIIa-null mutants all exhibited increased proportions of short amylopectin chains (approximately DP 6–12) and decreased levels of medium chains (approximately DP 13–36) compared to their respective parents. Similar changes in the chain length distribution profile were observed in Nipponbare compared with NILs(*Alk*), a near-isogenic line of Nipponbare with introgressed *indica SSIIa* (Haplotype 1) from Kasalath (Umemoto et al., 2004) and a recent study of an SSIIa null mutation in *japonica* rice (Miura et al., 2018). The 8% reduction observed in the total starch content of the shrunken lines and the slight reduction in peak gelatinization temperature (at least 2°C) of all shrunken and opaque lines tested also support the hypothesis that the residual SSIIa in Nipponbare has functional roles because its downregulation leads to significantly altered starch biosynthesis in the rice grain. The reduction of the peak gelatinization temperature was consistent with the earlier report by Miura et al. (2018), which observed a reduction of about 6°C in the gelatinization temperature of the null SSIIa line.

SSIIa is involved in the formation of multi-enzyme complexes as demonstrated in other cereal endosperms where it may provide a central structural scaffold for other enzymes participating in amylopectin biosynthesis (Liu F.S. et al., 2012). More recent work has shown that similar protein complexes occur in rice (Crofts et al., 2015; Hayashi et al., 2018). Although SSIIa is enzymatically less active in *japonica* lines, and expressed at lower levels compared to *indica* lines, it may still be crucial for starch biosynthesis by providing the structural scaffold for the formation of enzyme complexes (Tetlow et al., 2015). Our results are consistent with the conclusion that the formation of the scaffold is affected when the expression of SSIIa is abolished (probably in the case of hp-SSIIa-shr) or when the amount of SSIIa is significantly reduced (as in the case of hp-SSIIa-op).

It is probable that the opaque seed phenotype is due to a milder SSIIa down-regulation, and that the more severe shrunken seed phenotype showed the further impact of the complete loss of SSIIa. We speculate that RNA silencing of SSIIa in rice endosperm is lethal to rice grain development due to pleotropic effects not detected in this study because it is accompanied by severe distortion of seed and starch granule morphology and reduction of starch content, which ultimately result in non-viable seed in succeeding generations. Unfortunately, it was not possible to directly test the expression of SSIIa in the homozygous shrunken lines due to sterility in succeeding generations. The three SSIIa down-regulated lines with shrunken grains (SS9, SS17, and SS28) were shown to have severely distorted starch granules whose formation appeared to have been aborted early during grain development. These lines also showed reduced total starch content and lower proportions of amylopectin chains of DP 12–36. These results suggest severely reduced activity of other enzymes involved in amylopectin biosynthesis upon reduction of SSIIa protein, and will be the focus of future studies. Such pleiotropic effects are consistent with the generally accepted idea of functional multi-enzyme complexes carrying out starch biosynthesis in cereals, which can be disrupted in the complete absence of SSIIa. Interestingly, the effects of down-regulation of SSIIa may be different in different genetic backgrounds, or under different environmental or growth conditions. A similar study showed field-grown plants lacking SSIIa able to flower and set seed (Miura et al., 2018). The study by Miura et al. (2018) also noted increases in apparent amylose, which were not observed in the present study. Further experiments need to be conducted to elucidate the factors responsible for causing the differences in growth and seed viability between the two SSIIa-deficient *japonica* rice lines.

The NMR data showed very similar levels of molecular order (double helix and single helix contents) between isolated granules from Nipponbare and the two hp-SSIIa lines tested. Double helical molecular order arises from non-enzymic inter-twining of adjacent glucan chains in the branched amylopectin structure, provided they are longer than about six residues (Gidley and Bulpin, 1987). It cannot, however, be determined which chain lengths are directly responsible for molecular order, and a change in branch length profile may or may not lead to a change in the intertwining of chains depending on the detailed architecture of the amylopectin molecules. It is clear, however, that even slight perturbations in amylopectin structure can influence starch granule packing, which in turn can have an impact on the translucency of the grain. This observation is supported by *SBEIIb* transgenic mutants in rice, which shifted from A to C to B-type crystalline polymorph and is accompanied by chalky to opaque grain phenotypes (Butardo et al., 2011).

Effect of Down-Regulating SSIIa on Grain Starch Digestibility

Down-regulating the expression of *SSIIa* in *japonica* line, Nipponbare did not produce high amylose and high resistant starch phenotype as occurs in barley (Morell et al., 2003). The increase in amylose and reduced starch content in the SSIIa mutant in barley is accompanied by a significant increase in resistant starch and non-starch polysaccharide contents, as well as substantial reductions in total starch content and glycemic index (Morell et al., 2003; Topping et al., 2003). Another noteworthy SSIIa phenotype is found in the *sugary-2* mutation in maize where amylose levels doubled from ∼20% to ∼40% (Zhang et al., 2004). It is possible that down-regulating the expression of *SSIIa* in Nipponbare did not produce a high amylose phenotype because it is in a Wx^b background, where the GBSSI allele has low transcriptional efficiency (Sano, 1984; Hirano et al., 1998). More recently, a null mutation of SSIIa in Kinmaze (*japonica*) lead to a 4% elevation of amylose content in the *japonica* background (Miura et al., 2018). Considerable amounts of SSI and SBEIIb remain in the starch granule fraction of endosperms, although *SSIIa* has been significantly down-regulated, which was also reported by Miura et al. (2018). Our previous study on *SSIIa* mutants of different cereals suggested that only when SSI, SSIIa, and SBEIIb are all absent in the starch granules, the amylose content will be significantly elevated (Luo J. et al., 2015). It would be interesting to determine in future studies whether down-regulating or genome editing of *SSIIa* in *indica* lines where GBSSI is more active can lead to elevated amounts of amylose or long-chain amylopectin and whether simultaneously down-regulating *SSI*, *SSIIa* and *SBEIIb* lead to high amylose rice grains even in *japonica* background. However, it has been reported that null mutations of *ss1* and *be1*, or *ss1* and *be2b* resulted in sterile rice plants (Abe et al., 2014).

Interestingly, despite the lack of elevation in amylose and resistant starch contents, downregulating the *SSIIa* expression in Nipponbare reduced the predicted GI value, which is similar to the SSIIa null mutant in barley (Topping et al., 2003). The estimated glycemic score as measured by *in vitro* starch hydrolysis index (HI) was significantly reduced from 85 to 63 GI units (or 25%) in two hp-SSIIa-op lines (SS3 and SS5) based on biologically replicated assays. This reduction was still not as much as the reduction in glycemic index estimate for the ami-SBEIIb down-regulated lines (44 ± 1). Nonetheless, it is still noteworthy that the slight reduction in medium-chain amylopectin content in the hp-SSIIa-op line was linked with a lower glycemic index estimate even though this was not accompanied by an increase in amylose chains or long-chain amylopectin as was previously observed in the SBEIIb mutant line, IR36ae (Butardo et al., 2012) and down-regulated lines harboring ami-BEIIb (Butardo et al., 2011). Two candidate genes in rice for improving the GI response have been identified so far: SBEIIb (Butardo et al., 2011) and GBSSI (Fitzgerald et al., 2011). It appears that fine-tuning of amylopectin structure by SSIIa down-regulation provides a novel mechanism of reducing digestibility in rice by introducing subtle reductions in the proportion of medium-chain amylopectin (this study). Lowering the glycemic index without increasing the amylose and resistant starch contents is important in producing slowly digestible rice grains with acceptable cooking and eating qualities (Butardo et al., 2017). It is also possible, however, that other changes in the starch granules or non-starch constituents of the grain, which were not tested in this study, are responsible for the observed glycemic effects. For instance, the interaction of seed storage proteins and lipids with starch in rice can have a profound

influence on grain digestibility (Butardo and Sreenivasulu, 2016). Other structural carbohydrates that act as dietary fiber such as hemicelluloses can also have an impact on digestibility (Butardo and Sreenivasulu, 2016).

In summary, down-regulating the expression of *SSIIa* in Nipponbare produced rice grains with distinct phenotypes, including alterations in seed appearance, starch granule morphology, starch granule size distribution, and amylopectin fine structure. Unlike barley, further reduction in the amount of SSIIa was not accompanied by an increase in the amylose and resistant starch contents in a *japonica* background. However, a slight reduction in the proportion of medium-chain amylopectin (DP 13–36) in the SSIIa down-regulated lines is associated with a reduction in the glycemic index estimates. Thus, in addition to increasing amylose content and increasing the proportion of long-chain amylopectin, a reduction in the proportion of medium chains of amylopectin is also associated with the reduction of GI. Lastly, this study also supports the data of Crofts et al. (2015) and Hayashi et al. (2018), which indicate the structural and regulatory role of SSIIa in multimeric enzyme complex formation in amylopectin biosynthesis in developing rice endosperms.

DATA AVAILABILITY STATEMENT

All datasets generated for this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

VB conducted all the genetic engineering, functional genomic validation, gene and protein expression, and starch biochemistry experiments. JL conducted the protein expression experiment of developing grains. ZL determined the size distribution of the starch granules. MG determined the crystallinity and x-ray diffraction patterns of starch granules. AB determined the estimated glycemic score of transgenic and control lines. SR assisted in rice tissue culture and transformation experiments. MG, IT, MF, SJ, and SR provided technical supervision, intellectual guidance, and multi-disciplinary expertise during the conduct of all experiments. VB wrote the manuscript with the assistance of all co-authors, especially IT and SR who helped in the several revisions of the manuscript.

FUNDING

VB was supported by the Australian Aid (AusAid) Leadership Awards Scholarship and Commonwealth Scientific and Industrial Research Organisation (CSIRO) Top-Up Scholarship.

ACKNOWLEDGMENTS

We would like to acknowledge Dr. Bernadine Flanagan for her assistance in running the NMR and XRD experiments. We also acknowledge Dr. Mark Talbot for his assistance in running the environmental scanning electron microscopy experiments, Dr. Oscar Larroque and Dr. Adoracion Resurreccion for their assistance in running the debranched starch SEC and CE, and Dr. Christine Konik-Rose and Dr. Rosa Paula Cuevas for their assistance in running the DSC experiments.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00289/full#supplementary-material

REFERENCES

Abe, N., Asai, H., Yago, H., Oitome, N. F., Itoh, R., Crofts, N., et al. (2014). Relationships between starch synthase I and branching enzyme isozymes determined using double mutant rice lines. *BMC Plant Biol.* 14:80. doi: 10.1186/1471-2229-14-80

Bao, J.-S., Corke, H., and Sun, M. (2006). Nucleotide diversity in starch synthase IIa and validation of single nucleotide polymorphisms in relation to starch gelatinization temperature and other physicochemical properties in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 113, 1171–1183. doi: 10.1007/s00122-006-0355-6

Bao, J.-S., Xiao, P., Hiratsuka, M., Sun, M., and Umemoto, T. (2009). Granule-bound SSIIa protein content and its relationship with amylopectin structure and gelatinization temperature of rice starch. *Starch* 61, 431–437. doi: 10.1002/star.200800115

Butardo, V. Jr., and Sreenivasulu, N. (2016). "Chapter two - tailoring grain storage reserves for a healthier rice diet and its comparative status with other cereals," in *International Review of Cell and Molecular Biology*, ed. K. W. Jeon (Cambridge, MA: Academic Press), 31–70. doi: 10.1016/bs.ircmb.2015.12.003

Butardo, V. M., Fitzgerald, M. A., Bird, A. R., Gidley, M. J., Flanagan, B. M., Larroque, O., et al. (2011). Impact of down-regulation of starch branching enzyme IIb in rice by artificial microRNA- and hairpin RNA-mediated RNA silencing. *J. Exp. Bot.* 62, 4927–4941. doi: 10.1093/jxb/err188

Butardo, V. M. Jr., Anacleto, R., Parween, S., Samson, I., de Guzman, K, Alhambra, C. M., et al. (2017). Systems genetics identifies a novel regulatory domain of amylose synthesis. *Plant Physiol.* 173, 887–906. doi: 10.1104/pp.16.01248

Butardo, V. M. Jr., Daygon, V. D., Colgrave, M. L., Campbell, P. M., Resurreccion, A., Cuevas, R. P., et al. (2012). Biomolecular analyses of starch and starch granule proteins in the high-amylose rice mutant Goami 2. *J. Agric. Food Chem.* 60, 11576–11585. doi: 10.1021/jf303205p

Castro, J. V., Ward, R. M., Gilbert, R. G., and Fitzgerald, M. A. (2005). Measurement of the molecular weight distribution of debranched starch. *Biomacromolecules* 6, 2260–2270. doi: 10.1021/bm050041t

Crofts, N., Abe, N., Oitome, N. F., Matsushima, R., Hayashi, M., Tetlow, I. J., et al. (2015). Amylopectin biosynthetic enzymes from developing rice seed form enzymatically active protein complexes. *J. Exp. Bot.* 66, 4469–4482. doi: 10.1093/jxb/erv212

Cuevas, R. P., Daygon, V. D., Corpuz, H. M., Reinke, R. F., Waters, D. L. E., and Fitzgerald, M. A. (2010). Melting the secrets of gelatinisation temperature in rice. *Funct. Plant Biol.* 37, 439–447. doi: 10.1071/FP09258

Demorest, D., and Dubrow, R. (1991). Factors influencing the resolution and quantitation of oligonucleotides separated by capillary electrophoresis on a gel-filled capillary. *J. Chromatogr.* 559, 43–56. doi: 10.1016/0021-9673(91)80057-n

Denyer, K., Clarke, B., Hylton, C., Tatge, H., and Smith, A. M. (1996). The elongation of amylose and amylopectin chains in isolated starch granules. *Plant J.* 10, 1135–1143. doi: 10.1046/j.1365-313x.1996.10061135.x

Denyer, K., Hylton, C., Jenner, C. F., and Smith, A. M. (1995). Identification of multiple isoforms of soluble and granule-bound starch synthase in developing wheat endosperm. *Planta* 196, 256–265.

Fitzgerald, M. A., Rahman, S., Resurreccion, A. P., Concepcion, J. C., Daygon, V. D., Dipti, S. S., et al. (2011). Identification of a major genetic determinant of glycaemic index in rice. *Rice* 4, 66–74. doi: 10.1007/s12284-011-9073-z

Gao, Z., Zeng, D., Cui, X., Zhou, Y., Yan, M., Huang, D., et al. (2003). Map-based cloning of the Alk gene, which controls the gelatinization temperature of rice. *Sci. Chin. Ser. C Life Sci.* 46, 661–668. doi: 10.1360/03yc0099

Gidley, M. J., and Bulpin, P. V. (1987). Crystallisation of malto-oligosaccharides as models of the crystalline forms of starch: minimum chain-length requirement for the formation of double helices. *Carbohydr. Res*. 161, 291–300. doi: 10.1016/S0008-6215(00)90086-7

Hannah, L. C., and James, M. G. (2008). The complexities of starch biosynthesis in cereal endosperms. *Curr. Opin. Biotechnol.* 19, 160–165. doi: 10.1016/j.copbio.2008.02.013

Hayashi, M., Crofts, N., Oitome, N. F., and Fujita, N. (2018). Analyses of starch biosynthetic protein complexes and starch properties from developing mutant rice seeds with minimal starch synthase activities. *BMC Plant Biol.* 18:59. doi: 10.1186/s12870-018-1270-0

Hirano, H.-Y., Eiguchi, M., and Sano, Y. (1998). A single base change altered the regulation of the *Waxy* gene at the post-transcriptional level during the domestication of rice. *Mol. Biol. Evol.* 15, 978–987. doi: 10.1093/oxfordjournals.molbev.a026013

Hirose, T., and Terao, T. (2004). A comprehensive expression analysis of the starch synthase gene family in rice (*Oryza sativa* L.). *Planta* 220, 9–16. doi: 10.1007/s00425-004-1314-6

Jeon, J. S., Ryoo, N., Hahn, T. R., Walia, H., and Nakamura, Y. (2010). Starch biosynthesis in cereal endosperm. *Plant Physiol. Biochem.* 48, 383–392. doi: 10.1016/j.plaphy.2010.03.006

Liu, F., Ahmed, Z., Lee, E. A., Donner, E., Liu, Q., Ahmed, R., et al. (2012). Allelic variants of the *amylose extender* mutation of maize demonstrate phenotypic variation in starch structure resulting from modified protein-protein interactions. *J. Exp. Bot.* 63, 1167–1183. doi: 10.1093/jxb/err341

Liu, F. S., Romanova, N., Lee, E. A., Ahmed, R., Evans, M., Gilbert, E. P., et al. (2012). Glucan affinity of starch synthase IIa determines binding of starch synthase I and starch-branching enzyme IIb to starch granules. *Biochem. J.* 448, 373–387. doi: 10.1042/BJ20120573

Lopez-Rubio, A., Flanagan, B. M., Gilbert, E. P., and Gidley, M. J. (2008). A novel approach for calculating starch crystallinity and its correlation with double helix content: a combined XRD and NMR study. *Biopolymers* 89, 761–768. doi: 10.1002/bip.21005

Luo, J., Ahmed, R., Kosar-Hashemi, B., Larroque, O., Butardo, V. M. Jr., Tanner, G. J., et al. (2015). The different effects of starch synthase IIa mutations or variation on endosperm amylose content of barley, wheat and rice are determined by the distribution of starch synthase I and starch branching enzyme IIb between the starch granule and amyloplast stroma. *Theor. Appl. Genet.* 128, 1407–1419. doi: 10.1007/s00122-015-2515-z

Luo, J. X., Jobling, S. A., Millar, A., Morell, M. K., and Li, Z. Y. (2015). Allelic effects on starch structure and properties of six starch biosynthetic genes in a rice recombinant inbred line population. *Rice* 8:15. doi: 10.1186/s12284-015-0046-5

Miura, S., Crofts, N., Saito, Y., Hosaka, Y., Oitome, N. F., Watanabe, T., et al. (2018). Starch Synthase IIa-deficient mutant rice line produces endosperm starch with lower gelatinization temperature than japonica rice cultivars. *Front. Plant Sci.* 9:645. doi: 10.3389/fpls.2018.00645

Morell, M. K., Kosar-Hashemi, B., Cmiel, M., Samuel, M. S., Chandler, P., Rahman, S., et al. (2003). Barley *sex6* mutants lack starch synthase IIa activity and contain a starch with novel properties. *Plant J.* 34, 172–184.

Nakamura, Y., Francisco, P. B., Hosaka, Y., Sato, A., Sawada, T., Kubo, A., et al. (2005). Essential amino acids of starch synthase IIa differentiate amylopectin structure and starch quality between japonica and indica rice varieties. *Plant Mol. Biol.* 58, 213–227. doi: 10.1007/s11103-005-6507-2

Nakamura, Y., Sakurai, A., Inaba, Y., Kimura, K., Iwasawa, N., and Nagamine, T. (2002). The fine structure of amylopectin in endosperm from Asian cultivated rice can be largely classified into two classes. *Starch* 54, 117–131. doi: 10.1002/1521-379x(200204)54:3/4<117::aid-star117>3.0.co;2-2

Ohdan, T., Francisco, P. B., Sawada, T., Hirose, T., Terao, T., Satoh, H., et al. (2005). Expression profiling of genes involved in starch synthesis in sink and source organs of rice. *J. Exp. Bot.* 56, 3229–3244. doi: 10.1093/jxb/eri292

O'Shea, M. G., and Morell, M. K. (1996). High resolution slab gel electrophoresis of 8-amino-1,3, 6-pyrenetrisulfonic acid (APTS) tagged oligosaccharides using a DNA sequencer. *Electrophoresis* 17, 681–686. doi: 10.1002/elps.1150170410

Pfister, B., and Zeeman, S. C. (2016). Formation of starch in plant cells. *Cell. Mol. Life Sci.* 73, 2781–2807. doi: 10.1007/s00018-016-2250-x

Regina, A., Bird, A. R., Topping, D. L., Bowden, S., Freeman, J., Barsby, T. L., et al. (2006). High-amylose wheat generated by RNA interference improves indices of large-bowel health in rats. *Proc. Natl. Acad. Sci. U.S.A.* 103, 3546–3551. doi: 10.1073/pnas.0510737103

Sano, Y. (1984). Differential regulation of waxy gene expression in rice endosperm. *Theor. Appl. Genet.* 68, 467–473. doi: 10.1007/BF00254822

Tetlow, I. J., and Emes, M. J. (2014). A review of starch-branching enzymes and their role in amylopectin biosynthesis. *Iubmb Life* 66, 546–558. doi: 10.1002/iub.1297

Tetlow, I. J., and Emes, M. J. (2017). Starch biosynthesis in the developing endosperms of grasses and cereals. *Agronomy* 7:81. doi: 10.3390/agronomy7040081

Tetlow, I. J., Liu, F., and Emes, M. J. (2015). "Protein-protein interactions during starch biosynthesis," in *Starch: Metabolism and Structure*, ed. Y. Nakamura (Tokyo: Springer), 291–313. doi: 10.1007/978-4-431-55495-0_8

Topping, D. L., Morell, M. K., King, R. A., Li, Z., Bird, A. R., and Noakes, M. (2003). Resistant starch and health - himalaya 292, a novel barley cultivar to deliver benefits to consumers. *Starch* 55, 539–545. doi: 10.1002/star.200300221

Toyota, K., Tamura, M., Ohdan, T., and Nakamura, Y. (2006). Expression profiling of starch metabolism-related plastidic translocator genes in rice. *Planta* 223, 248–257. doi: 10.1007/s00425-005-0128-5

Umemoto, T., and Aoki, N. (2005). Single-nucleotide polymorphisms in rice *starch synthase IIa* that alter starch gelatinisation and starch association of the enzyme. *Funct. Plant Biol.* 32, 763–768.

Umemoto, T., Aoki, N., Lin, H., Nakamura, Y., Inouchi, N., Sato, Y., et al. (2004). Natural variation in rice *starch synthase IIa* affects enzyme and starch properties. *Funct. Plant Biol.* 31, 671–684.

Umemoto, T., Horibata, T., Aoki, N., Hiratsuka, M., Yano, M., and Inouchi, N. (2008). Effects of variations in starch synthase on starch properties and eating quality of rice. *Plant Prod. Sci.* 11, 472–480. doi: 10.1626/pps.11.472

Umemoto, T., Terashima, K., Nakamura, Y., and Satoh, H. (1999). Differences in amylopectin structure between two rice varieties in relation to the effects of temperature during grain-filling. *Starch* 51, 58–62. doi: 10.1002/(sici)1521-379x(199903)51:2<58::aid-star58>3.0.co;2-j

Umemoto, T., Yano, M., Satoh, H., Shomura, A., and Nakamura, Y. (2002). Mapping of a gene responsible for the difference in amylopectin structure between japonica-type and indica-type rice varieties. *Theor. Appl. Genet.* 104, 1–8. doi: 10.1007/s001220200000

Waters, D. L. E., Henry, R. J., Reinke, R. F., and Fitzgerald, M. A. (2006). Gelatinization temperature of rice explained by polymorphisms in *starch synthase*. *Plant Biotechnol. J.* 4, 115–122. doi: 10.1111/j.1467-7652.2005.00162.x

Yamakawa, H., Hirose, T., Kuroda, M., and Yamaguchi, T. (2007). Comprehensive expression profiling of rice grain filling-related genes under high temperature using DNA microarray. *Plant Physiol.* 144, 258–277. doi: 10.1104/pp.107.098665

Zhang, X. L., Colleoni, C., Ratushna, V., Sirghle-Colleoni, M., James, M. G., and Myers, A. M. (2004). Molecular characterization demonstrates that the *Zea mays* gene *sugary2* codes for the starch synthase isoform SSIIa. *Plant Mol. Biol.* 54, 865–879. doi: 10.1007/s11103-004-0312-1

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 Genetics

ORIGINAL RESEARCH
published: 22 April 2020
doi: 10.3389/fgene.2020.00213

Genome-Wide Association Study Reveals Novel Marker-Trait Associations (MTAs) Governing the Localization of Fe and Zn in the Rice Grain

*Haritha Bollinedi[1], Ashutosh Kumar Yadav[1], K. K. Vinod[1], S. Gopala Krishnan[1], Prolay Kumar Bhowmick[1], M. Nagarajan[2], C. N. Neeraja[3], Ranjith Kumar Ellur[1] and Ashok Kumar Singh[1]**

[1] Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India, [2] ICAR-Indian Agricultural Research Institute, Rice Breeding and Genetics Research Centre, Aduthurai, India, [3] ICAR-Indian Institute of Rice Research, Hyderabad, India

Micronutrient malnutrition due to Fe and Zn, affects around two billion people globally particularly in the developing countries. More than 90% of the Asian population is dependent on rice-based diets, which is low in these micronutrients. In the present study, a set of 192 Indian rice germplasm accessions, grown at two locations, were evaluated for Fe and Zn in brown rice (BR) and milled rice (MR). A significant variation was observed in the rice germplasm for these micronutrients. The grain Fe concentration was in the range of 6.2–23.1 ppm in BR and 0.8–12.3 ppm in MR, while grain Zn concentration was found to be in the range of 11.0–47.0 ppm and 8.2–40.8 ppm in the BR and MR, respectively. Grain Fe exhibited maximum loss upon milling with a mean retention of 24.9% in MR, while Zn showed a greater mean retention of 74.2% in MR. A genome-wide association study (GWAS) was carried out implementing the FarmCPU model to control the population structure and kinship, and resulted in the identification of 29 marker-trait associations (MTAs) with significant associations for traits viz. FeBR (6 MTAs), FeMR (7 MTAs), ZnBR (11 MTAs), and ZnMR (5 MTAs), which could explain the phenotypic variance from 2.1 to as high as 53.3%. The MTAs governing the correlated traits showed co-localization, signifying the possibility of their simultaneous improvement. The robust MTAs identified in the study could be valuable resource for enhancing Fe and Zn concentration in the rice grain and addressing the problem of Fe and Zn malnutrition among rice consumers.

Keywords: Fe, Zn, biofortification, GWAS, donors, rice, milled rice, brown rice

OPEN ACCESS

Edited by:
Balram Marathi, Professor Jayashankar Telangana State Agricultural University, India

Reviewed by:
Jieyun Zhuang, Zhejiang University, China
Zhao Hu, Huazhong Agricultural University, China

****Correspondence:***
Ashok Kumar Singh
aks_gene@yahoo.com

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *12 September 2019*
Accepted: *24 February 2020*
Published: *22 April 2020*

Citation:
Bollinedi H, Yadav AK, Vinod KK, Gopala Krishnan S, Bhowmick PK, Nagarajan M, Neeraja CN, Ellur RK and Singh AK (2020) Genome-Wide Association Study Reveals Novel Marker-Trait Associations (MTAs) Governing the Localization of Fe and Zn in the Rice Grain. Front. Genet. 11:213. doi: 10.3389/fgene.2020.00213

INTRODUCTION

In the human body, iron (Fe) and zinc (Zn) are the two most abundant trace minerals. An average adult human with a body weight of 65 kg has about 3–4 g of Fe and 1.5–2.5 g of Zn (King et al., 2006; Wood et al., 2006). Fe is essential for the synthesis of oxygen-transporting proteins *viz.* hemoglobin and myoglobin and is also an integral part of the enzymes involved in energy production and the maintenance of immune functions (Roeser, 1986; Stoltzfus, 2001). Zn is a pre-requisite for biological functions like gene expression, cell division, cell development, reproduction,

and immunity (Brown et al., 2001, 2004). Fe and Zn are commonly viewed together in human nutrition, as they can be obtained through common dietary sources, and their absorption or inhibition is believed to be affected by similar factors (Lim et al., 2013). Fe deficiency in humans is related to increased risk of maternal mortality, anemia, premature births, low birth weight, and impaired cognitive and motor development (Bouis, 2003). Severe or clinical Zn deficiency is a condition associated with short stature, immune system dysfunction, hypogonadism, skin disorders, anorexia, delayed wound healing, skeletal abnormalities, and cognitive dysfunction (Prasad, 1991; Salgueiro et al., 2000).

Micronutrient malnutrition popularly known as "hidden hunger," primarily attributed to Zn and Fe deficiency, is evolving as a global pandemic with serious health effects. The situation is more grim in developing countries and has become a fundamental limitation in achieving the Sustainable Development Goals (SDGs) in these countries. Mineral deficiency disorders caused by Fe and Zn deficiency are the primary reasons for the reduced work productivity in developing countries, which further affects the gross national product. Fe deficiency is recognized as the most widespread nutritional disorder with about 1.6 billion people suffering from iron deficiency anemia (IDA) globally. Additionally, Zn deficiency is ranked as the 5th major health risk factor in developing countries, affecting nearly two billion people worldwide (Zhang et al., 2018). The loss incurred due to Zn deficiency amounts to almost 16 million global disability-adjusted life years (DALYs) (Caulfield et al., 2004; Black et al., 2008).

The dietary habit of the majority of the population in developing countries as well as of a significant portion of the world populace is cereal based. In general, cereal foods are low in micronutrient content, especially when grains are consumed after extensive processing. Further, when cereal based diets are not balanced with other supplementary food items, there can be a severe deficiency of micronutrient nutrition leading to hidden hunger. Although the grain Fe and Zn concentrations in the brown rice (BR) of the modern-day high yielding rice cultivars are in the range of 6.3–24.4 mg/kg and 15.3–58.4 mg/kg, respectively (Gregorio et al., 2000), much of which are lost upon polishing, retaining a maximum of 2 mg/kg of Fe and 12 mg/kg of Zn in the grains that are consumed (Pradhan et al., 2020). Therefore, genetic fortification of rice grain becomes the most viable solution to address the challenge of micronutrient malnutrition. To begin with, identification of molecular players affecting the mineral uptake and transport, and understanding the genes and pathways governing mineral homeostasis and localization, would assist in devising targeted breeding strategies for biofortification.

A genome-wide association study (GWAS) provides an opportunity to utilize the tremendous genetic diversity of rice preserved among the landraces, to unravel the molecular mechanisms of complex traits, like mineral uptake and their grain accumulation. The method uses numerous historic recombination in a large natural population that offers the potential to localize the trait genetic determinants effectively to a narrower region. Further, application of single nucleotide polymorphism (SNP) in a GWAS study provides dense coverage of the entire genome, which would help identify the functional variation governing the trait. We used a diverse set of 192 rice germplasm accessions including landraces, popular varieties, Basmati accessions, and breeding lines collected from all over India in a GWAS to analyze the morphological variation for grain Fe and Zn content across two diverse locations, and to identify the marker-trait associations (MTAs)/genomic regions associated with Fe and Zn content in rice grain, and elite parents that can serve as donors in rice biofortification programs.

MATERIALS AND METHODS

Plant Material and Field Experiment

The study comprised of an association mapping panel of 192 rice accessions collected from different parts of India and maintained at the Division of Genetics, ICAR-Indian Agricultural Research Institute (ICAR-IARI), New Delhi, India. The complete set of genotypes were evaluated at two locations *viz*. the ICAR-IARI farm (28.04°N; 77.12°E, 228 m) in New Delhi during the wet season (WS) of 2017, and the Rice Breeding and Genetics Research Centre (RBGRC) farm (11° 00′N; 79° 28′E, 19.5 m) in Aduthurai, Tamil Nadu during the dry season (DS) of 2017–2018. The soils from the IARI farm are of a sandy loam type with a pH of 7.4–7.8 with an average available Fe and Zn content of 12.3 and 3.3 mg/kg, respectively, while the soils from Aduthurai are typical haplusterts with a pH of 7.2 and mean available Fe and Zn content of 12.5 and 0.7 mg/kg, respectively. An augmented randomized complete block design was adopted for field evaluation. Similar agronomic practices were followed at both locations. The seeds were germinated on a raised seedbed to ensure uniform germination, and 28-day old seedlings were transplanted into a puddled rice field. Each genotype was planted in two rows of 3 m length with a 20 cm spacing between rows and 15 cm between plants. The recommended package of practices was adopted to ensure good crop stand. At maturity, plants were selected from the middle of the rows, and the grains were harvested, threshed, and dried to a moisture content of $\leq$13% and stored in zip lock bags at room temperature until analysis.

Analysis of Grain Fe and Zn Concentration

Fe and Zn content of the samples was estimated in parts per million units (mg/kg) from both BR and milled rice (MR) using an energy dispersive X-ray fluorescence spectrometer (ED-XRF) (X-Supreme 8000, Oxford Instruments, CA, USA). Grain samples from each genotype was dehusked using a palm de-husker, and the BR was cleaned thoroughly and analyzed directly in ED-XRF. The samples were further milled in a non-ferrous, non-zinc rice polisher (Mini Lab Rice Polisher Model K-710, Krishi International, Hyderabad, India) to avoid contamination with the metals during polishing. The milled samples were further cleaned with a non-shredding tissue paper to remove all the residual bran, and whole grains were used for analysis in ED-XRF. Three independent samples drawn from the bulk sample were analyzed for Fe and Zn.

DNA Isolation and SNP Genotyping

Genomic DNA was isolated from the 192 germplasm accessions using young leaves at 30 days after transplanting by adopting the CTAB method (Murray and Thompson, 1980). The quality of the DNA was assessed on a 0.8% agarose gel and was further quantified using a nano-drop spectrophotometer (NanoDrop^{TM} 2000/2000c, Thermo Fisher Scientific, DE, United States). High throughput genotyping was carried out using a custom-designed 50 K SNP chip. The chip was based on single-copy genes, covering all 12 rice chromosomes with an average interval of fewer than 1 kb between adjacent SNP markers (Singh et al., 2015). DNA amplification, fragmentation, chip hybridization, single-base extension through DNA ligation, and signal amplification were carried out as detailed in Singh et al. (2015).

Population Structure and Linkage Disequilibrium Analyses

The number of subgroups in the association mapping panel was estimated using both a model-based approach using STRUCTURE 2.3.4 software (Pritchard et al., 2000) and principal component (PC) analysis. For the model-based analysis, a Bayesian model approach using an ancestry model of ADMIXTURE (Alexander et al., 2009) and a frequency model assuming correlated allele frequencies among the sub-populations was used. For this, a subset of 5,000 genome-wide SNP markers with minor allele frequency (MAF) $\geq$0.05 and missing data <20% was used. The analysis was run with an assumed number of subgroups (K) ranging from 1 to 10 and with each K replicated 10 times. For each run, 100,000 burn-in steps followed by 100,000 Markov Chain Monte Carlo simulations were implemented. The optimum number of K was determined according to Evanno et al. (2005) embedded in STRUCTURE HARVESTER (Earl and vonHoldt, 2012) by plotting the *ad hoc* statistic ΔK against the natural logarithms of probability data [LnP(K)]. PCA analysis, which was incorporated in the package "GAPIT" (genomic association and prediction integrated tool) running under *R* environment (Lipka et al., 2012) was used. A significant number of PCs that are adequate to explain the population structure were determined based on the scree plot generated by GAPIT. Tukey's multiple comparison test was implemented to assess the significance of the difference in means of sub-populations. The extent of linkage disequilibrium between the SNP markers was analyzed by calculating the r^2 values in TASSEL v5.2.20 (Bradbury et al., 2007). Only r^2 values with $p < 0.05$ within each chromosome were considered for linkage disequilibrium (LD) decay analysis. Marker pairs were clustered into 5 kb bins, and the average r^2 value of each bin was estimated and plotted against the distance. The physical distance at which the r^2 value dropped to half of its average maximum value was considered an LD decay rate (Huang et al., 2010).

GWAS Analysis

For GWAS analysis, the data obtained from 50,000 SNP markers was filtered for MAF $\geq$0.05 and maximum missing sites per SNP <20% and maximum missing sites per genotype <20%. A total of 31,132 SNPs remained after filtering, and were used for further analysis. One accession was dropped from the analysis due to the poor quality of the genotypic data. The data was analyzed using multiple statistical models *viz*. general linear model (GLM), mixed linear model (MLM), and fixed and random model circulating probability unification (FarmCPU; Liu, 2015; Liu et al., 2016). The efficiency of these models in controlling the familial relatedness and population structure was assessed by comparing the quantile–quantile (Q–Q) plots obtained through plotting observed p-values against the expected values. In each case, significant MTAs were identified after a modified Bonferroni multiple test correction calculated from the reciprocal of total number of markers used for analysis [$p < 3.21E\text{-}05$; $-\log_{10}(p) > 5.79$]. The percentage of phenotypic variance explained (PVE) by individual SNP was calculated through the single-marker analysis.

Assessment of the Novelty of Identified MTAs

The novelty of MTAs identified in the study was determined through the comparison of the physical positions, with those of previously reported quantitative trait loci (QTLs) for Fe and Zn. In addition to the literature survey, the gramene QTL database[1] and QTL annotation rice online database (QTARO)[2] were searched to identify the physical locations of the previously reported QTLs. The candidate genes in the vicinity of the MTA region were determined using the genome browser of the Rice Annotation Project Database (RAP-DB)[3].

RESULTS

Population Structure

Model-based simulation of population structure showed a sharp peak at $K = 3$ when the number of clusters was plotted against ΔK, depicting the presence of three sub-populations in the panel (**Figure 1A**). The sub-populations are denoted as POP1, POP2, and POP3. **Figure 1B** shows a representative picture of the population structure in which each individual is indicated by a vertical bar, which is divided into segments based on its estimated membership fractions in sub-populations. POP2 was the largest and constituted 73.5% of the panel numbering 139 accessions, of which 74 were pure types (includes accessions with <5% of admixture) and 65 were admixture types. POP1 consisted of 35 accessions including 13 pure types and 22 admixture types. Unique landraces from the Jammu and Kashmir region like Begum, BalaKoan, Buta Baber, Mehvan, etc., were all included in POP1. In addition, it also consisted of aromatic rice accessions including the traditional Basmati varieties like Basmati 370, Super Basmati, Type 3 and short grain aromatic accessions like Kalanamak, Chittimutyalu, etc. All the evolved varieties like MTU1001, ADT 39, IR 70, MAS946-1, Improved Sabarmati, and breeding lines like PRR 109, PRR

[1]http://archive.gramene.org/qtl/
[2]http://qtaro.abr.affrc.go.jp/
[3]https://rapdb.dna.affrc.go.jp/

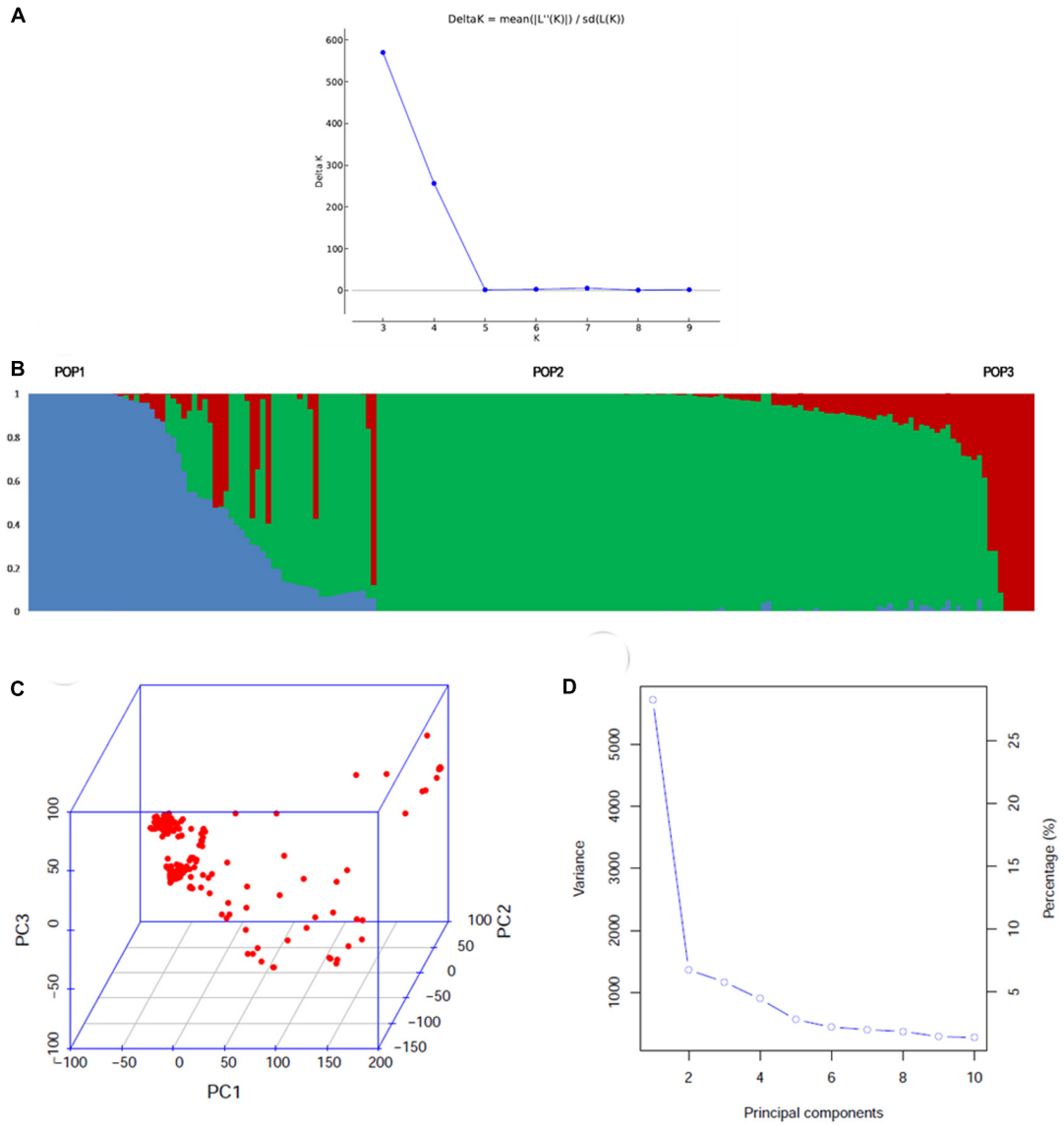


FIGURE 1 | Population structure of the association mapping panel comprising of 192 genotypes analyzed by model based and PCA based approaches. **(A)** ΔK plot depicting three subgroups in the population by Evanno's method. The highest ΔK was 580 at $K = 3$, **(B)** the bar plot showing the three sub-populations identified. POP2 was the largest and showed remarkable admixture with POP1. POP3 was the smallest group, which showed less admixing with POP1 **(C)** 3D graph depicting the distribution of accessions along the three PCs **(D)** scree plot depicting the number of significant PCs. There were three PCs that explained a cumulative variation of ∼40%.

127, etc., were grouped into POP2. POP3 is the smallest, with 15 accessions, of which six were pure types and the rest were admixture types. Two accessions, PDKV-Chinoor 2 and SAF 1221-83, were admixtures and were therefore not included in any of the sub-groups. The fixation index (F_{st}) was 0.74, 0.74, and 0.68 for the sub-populations POP1, POP2, and POP3, respectively, and the expected heterozygosity or the average distance between individuals within a sub-population was 0.15, 0.18, and 0.17. The allelic frequency divergence of POP1 from POP2 was 0.42, and from POP3 was 0.34, while between POP2 and POP3 it was 0.39. The PCA based lookout for the presence of structure in the population also indicated significance of three PCs in the population, explaining about 39.6% of total variation (**Figure 1C**). The scree plot (**Figure 1D**) showed that the first PC explained the highest variation of 26.7% followed by second and third PCs explaining 6.8 and 6.0% of the total variation, respectively. Admixture analysis revealed that about 53.60% of the accessions (103/192) showed 0–5% of admixture, 12% of the accessions (23/191) showed 5–20% of the admixture, while the remaining 34% of the accessions depicted >20% of the admixture. LD analysis in the association panel revealed a highest average r^2 of 0.7 at shorter

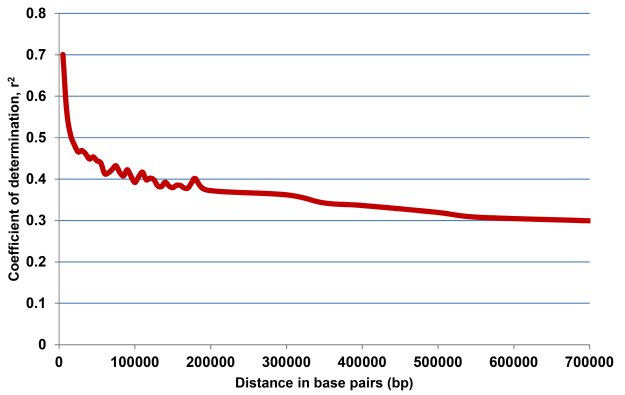


FIGURE 2 | Linkage disequilibrium (LD) decay plot of the association mapping panel derived from 31,132 SNPs, plotted against physical distance in base pairs (bp) and co-efficient of determination (r^2).

distance of 5 kb, and it decreased to half its value at around 350 kb (**Figure 2**).

Variation for Grain Mineral Micronutrients in the Association Panel

The grain Fe and Zn contents exhibited significant variation in the association panel (**Table 1**). In New Delhi, during WS 2017, Fe concentration in BR (FeBR) varied from 6.5 (Jayati) to 23.1 (Shah Pasand) mg/kg with a mean of 12.7 ± 0.2 mg/kg, while in milled rice (FeMR) it ranged from 0.8 (Arupathaam Kuruvai) to 12.3 (IC-2127) mg/kg with an average of 3.6 ± 0.1 mg/kg. In Aduthurai, the variation for FeBR was similar to that of New Delhi and ranged from 6.2 (Khara Munga) to 21.4 (PRR 109) mg/kg with a mean value of 10.6 ± 0.2 mg/kg. The FeMR exhibited a comparatively narrow range of 0.9 (Ramachandi) to 5.6 (PRR 109) mg/kg with a mean of 2.6 ± 0.1 mg/kg. Alternatively, grain Zn content in BR (ZnBR) varied from 13.0 (Sagar samba) to 46.2 (Karuppunel) mg/kg in the New Delhi location, while in Aduthurai its range was 11.0 (Samanta and OYR 128) to 47 (Karuppunel) mg/kg. In MR, however, the mean value for grain Zn (ZnMR) was 18.0 ± 0.4 mg/kg in New Delhi with a range between 8.2 (Sagar Samba) and 40.9 (Karuppunel) mg/kg, while in Aduthurai, it varied between 8.5 (Bhadrakali) and 40.8 (Karuppunel). The average grain ZnMR in Aduthurai was 16.5 ± 0.4 mg/kg, which was much closer to that observed in the New Delhi conditions.

The association panel depicted significant variation for percent retention of mineral micronutrients upon polishing. A substantial quantity of Fe was lost upon polishing and its retention on milling ranged from as low as 5.4% (Aziz Beoul) to 89.1% (OYR 128) with a mean retention of 24.9% in the MR. When compared to grain Fe, less polishing loss of grain Zn was evident with a retention ranging from 45.8% (OYC 183) to as high as 97.3% (Kalanamak). Of the 192 accessions analyzed, as many as 75 genotypes exhibited a retention of ≥80% of grain Zn.

The sub-populations do not differ significantly for grain Fe and Zn concentration except for the POP2 having slightly lower mean values than POP1 and POP3 (**Table 2**). POP3 showed a significant difference for the traits between the locations with highest values recorded at the New Delhi location (**Figure 3**). Population structure accounted for a variation in Fe and Zn concentration ranging from 2.7 (FeMR) to 34.2% (ZnBR) at the New Delhi location and 2.1 (FeBR) to 32.4% (ZnMR) at the Aduthurai location.

TABLE 1 | Basic statistics of the grain Fe and Zn content in the association mapping panel used in the study, at two diverse locations, New Delhi and Aduthurai.

Trait	New Delhi			Aduthurai		
	Mean ± SE	Range	Variance	Mean ± SE	Range	Variance
FeBR	12.7 ± 0.2	6.5–23.1	10.6	10.6 ± 0.2	6.2–21.4	7.9
FeMR	3.6 ± 0.1	0.8–12.3	3.3	2.6 ± 0.1	0.9–5.6	1.0
ZnBR	22.9 ± 0.4	13.0–46.2	32.9	21.7 ± 0.4	11–47	32.5
ZnMR	18.0 ± 0.4	8.2–40.9	28.9	16.5 ± 0.4	8.5–40.8	25.9

FeBR, Fe content in brown rice; FeMR, Fe content in milled rice; ZnBR, Zn content in brown rice; ZnMR, Zn content in milled rice; SE, standard error.

Associations Among the Mineral Micronutrients

Grain mineral micronutrient content between two locations were significantly correlated ($p > 0.001$). Strong positive correlations were observed for ZnMR ($r = 0.78$), ZnBR ($r = 0.77$), and FeBR ($r = 0.67$) while FeMR ($r = 0.30$) indicated a moderate association. Trait wise within location correlation analysis indicated that grain Zn in BR was significantly correlated with Zn concentration in MR with high positive values in both New Delhi ($r = 0.95$; $p > 0.0001$) and Aduthurai ($r = 0.90$; $p > 0.0001$). However, FeBR showed non-significant association with FeMR. Nevertheless, FeBR exhibited significant positive correlation with ZnBR as well as with ZnMR (**Figure 4**).

Detection of Stable and Environment-Specific MTAs by GWAS

Association analysis was performed separately for the two locations. The Q–Q plots generated through the FarmCPU model depicted less deviation of the observed *p*-values from the expected *p*-values and was therefore chosen as the best fit (**Figure 5**). A total of 29 QTLs located along 9 out of 12 chromosomes were identified for the four traits analyzed in two environments (**Table 3**). Eleven were major effect QTLs with a PVE more than 20%, 10 were moderate effect QTLs (PVE 20–10%), while the remaining seven were minor effect QTLs (PVE < 10%). One MTA, *qZnMR1.1* showed moderate effect at the New Delhi location and minor effect at the Aduthurai location. Chromosome 3 had the highest number of 11 MTAs, followed by chromosome 1 (six MTAs), chromosomes 2, 6, and 7 (two MTAs each), and chromosomes 8, 9, and 10 (one MTA each). No MTAs affecting Fe and/or Zn concentration in rice grains were found on chromosome 5, 11, and 12. Manhattan plots depicting the significant SNP markers above the modified Bonferroni threshold are provided in **Figure 5**.

TABLE 2 | Mean and range values for Fe and Zn content among the three sub-populations of the association mapping panel for the two locations, New Delhi and Aduthurai.

Traits	POP1		POP2		POP3		R^2
	Mean ± SD	Range	Mean ± SD	Range	Mean ± SD	Range	
New Delhi							
FeBR	15.9 ± 3.9	8.8–23.1	11.6 ± 2.3	6.5–21.2	15.4 ± 2.5	11.6–20.5	31.6
FeMR	3.4 ± 1.4	1.1–6.5	3.5 ± 1.5	0.8–11.8	3.8 ± 1.9	1.6–8.5	2.7
ZnBR	28.3 ± 6.5	16.6–40.1	20.9 ± 3.8	13–32.5	28.7 ± 6.6	19.2–46.2	34.2
ZnMR	22.8 ± 5.1	12.7–33.4	16.2 ± 4.1	8.2–28.1	22.7 ± 6.5	16–40.8	30.0
Aduthurai							
FeBR	13.5 ± 4.2	7.8–20.7	9.7 ± 2.0	6.2–21.4	12.1 ± 3.2	9.1–20.1	25.9
FeMR	2.8 ± 0.9	0.9–4.8	2.4 ± 1.0	0.9–5.6	2.7 ± 1.2	1–4.7	2.1
ZnBR	27.6 ± 6.6	15.8–42.6	20.0 ± 3.8	11–31.4	26.7 ± 9.0	11.0–47.0	30.2
ZnMR	21.4 ± 5.5	11.0–32.5	14.8 ± 3.5	8.5–26.4	21.5 ± 6.7	16–40.8	32.4

R^2 indicates the percentage of total phenotypic variation explained by sub-populations.

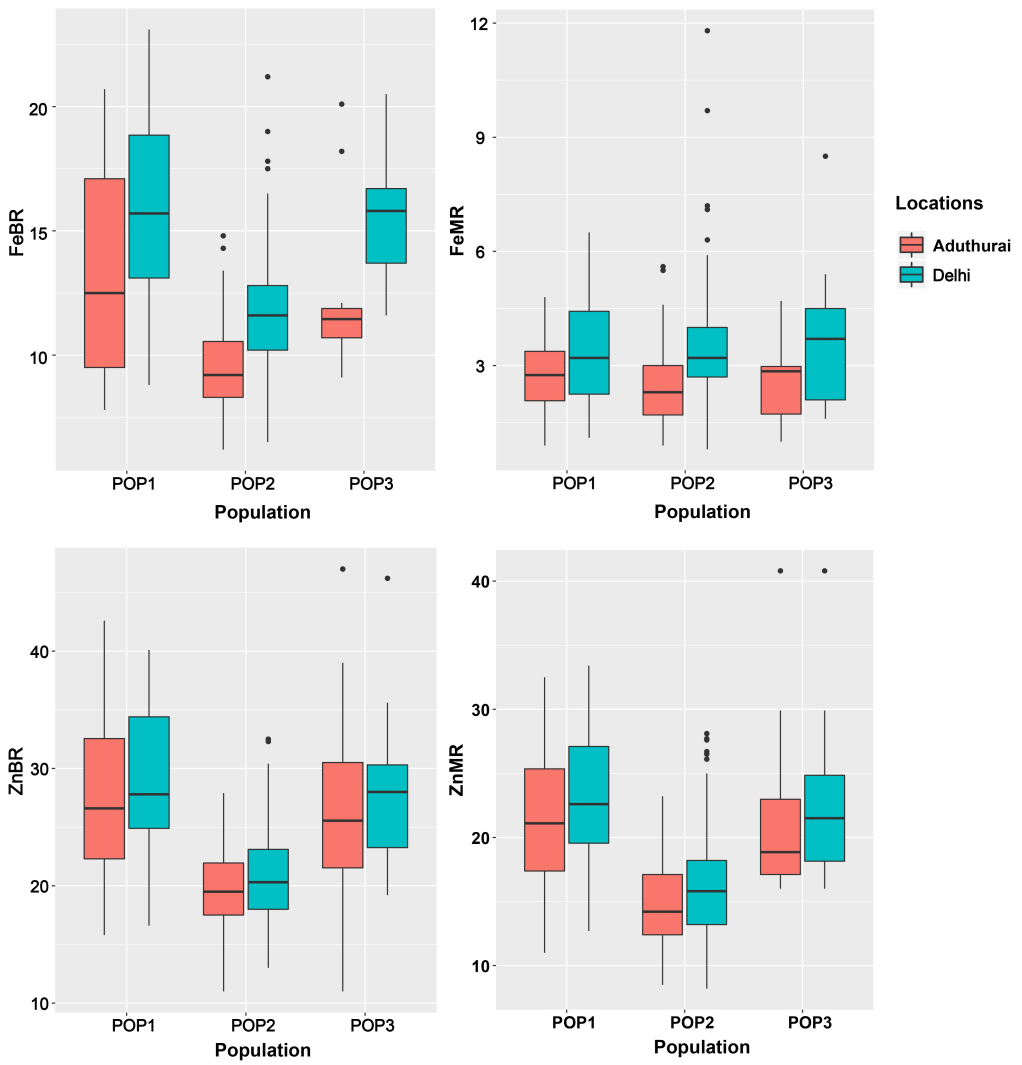


FIGURE 3 | Boxplots showing distribution of grain Fe and Zn content among the sub-populations POP1, POP2, and POP3. FeBR, Fe concentration in brown rice; FeMR, Fe concentration in milled rice; ZnBR, Zn concentration in brown rice; ZnMR, Zn concentration in milled rice.

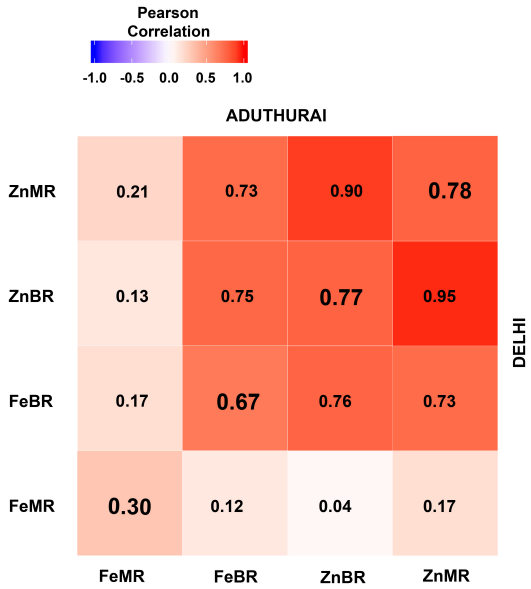


FIGURE 4 | Correlations of Fe and Zn between the locations (diagonal elements) and independent locations at New Delhi (upper diagonal) and Aduthurai, Tamil Nadu (lower diagonal). FeBR, Fe concentration in brown rice; FeMR, Fe concentration in milled rice; ZnBR, Zn concentration in brown rice; ZnMR, Zn concentration in milled rice.

FeBR: A total of six MTAs affecting FeBR were identified in the study. Two MTAs were identified at the New Delhi location one each on chromosomes 2 and 6. The MTA on chromosome 6 had a major effect with a PVE of 45.2%, and its favorable allele A at the SNP marker locus AX-95956929 had an additive effect of 2.7 mg/kg on the trait. The MTA on chromosome 2 was a minor effect one with a PVE of only 6.8%. Four MTAs for FeBR were identified at the Aduthurai location, two of them were major effect associations with a PVE as high as 53.3% (*qFeBR3.1*) and 41.5% (*qFeBR7.1*) and an additive effect of 4.1 and 4.7, respectively.

FeMR: For FeMR, no MTAs were detected at Bonferroni threshold. However, by adopting a false discovery rate (FDR) threshold, seven MTAs were identified across two locations, of which six were moderate effect MTAs with a PVE ranging from 12.3 to 13.6%. All of these were detected at the Aduthurai location and were found clustered around the 32 Mb physical location on chromosome 3. Only one minor effect MTA was identified at the New Delhi location which explained 5.3% variation.

ZnBR: For this trait, a total of 11 MTAs were identified across two locations, eight under New Delhi conditions and three under Aduthurai conditions. The MTAs detected in New Delhi explained a PVE ranging between 12.4 (*qZnBR1.1*) and 47.6% (*qZnBR4.1*), of which six were major MTAs. The additive effect described by the favorable alleles of the MTAs was in the range of 3.4 to 6.6 mg/kg. At Aduthurai, three MTAs were recorded for ZnBR, one each on chromosomes 6, 7, and 10. The MTA on chromosome 6 was a major effect MTA and explained a PVE of 44.9%, which also recorded the highest additive effect of 7.4 mg/kg. Of the remaining two MTAs, the MTA on chromosome 7 was a moderate effect MTA with a PVE of 16.5%, while the one on chromosome 10 was a minor MTA with a PVE of 6.8%.

ZnMR: For the Zn concentration in MR, a total of five MTAs were identified, two each in the New Delhi and Aduthurai locations. One MTA, *qZnMR1.1* was commonly detected in both the sites with a PVE of 9.3 and 10.3% in New Delhi and Aduthurai, respectively. The PVE explained by the MTAs identified under the New Delhi condition ranged between 42.7 (*qZnMR4.1*) and 9.3% (*qZnMR1.1*), while those at the Aduthurai location explained a PVE between 18.3 (*qZnMR3.1*) and 7.9% (*qZnMR9.1*).

Co-localization of MTAs

The MTAs, *qZnBR1.2a* and *qZnBR1.2b* identified for ZnBR at the New Delhi location were located in the same LD block and were co-localized on chromosome 1 at the Aduthurai location, six MTAs for FeMR namely *qFeMR3.1a*, *qFeMR3.1b*, *qFeMR3.1c*, *qFeMR3.1d*, *qFeMR3.1e*, and *qFeMR3.1f* were found to be in tight linkage on chromosome 3. Further, for the highly correlated traits such as ZnBR and ZnMR, co-localized MTAs were detected on chromosome 4 (*qZnBR4.1* and *qZnMR4.1*) as well as on chromosome 1 (*qZnBR1.1* and *qZnMR1.1*) under the New Delhi environment. Besides, *qZnBR4.1* and *qZnMR4.1* were found to share a common peak SNP. Between the environments, *qZnBR3.1* identified under New Delhi conditions for the trait ZnBR was found co-localized with eight other MTAs affecting the traits ZnMR (*qZnMR3.1*), FeBR (*qFeBR3.1*), and FeMR (*qFeMR3.1a*, *qFeMR3.1b*, *qFeMR3.1c*, *qFeMR3.1d*, *qFeMR3.1e*, and *qFeMR3.1f*) identified under Aduthurai conditions. Further, the *qFeBR6.1* identified at the New Delhi location was found co-localized with the *qZnBR6.1* identified at the Aduthurai location for the trait ZnBR. Additionally, the MTAs governing ZnBR, *qZnBR7.1* from the Aduthurai location and *qZnBR7.2* from the New Delhi location have also appeared co-localized on chromosome 7.

It was interesting to note that 12 of the 29 MTAs identified in this study were found to be in the vicinity of previously reported QTL regions and candidate genes. The MTA, *qFeBR3.1* was in close proximity to the known candidate genes *OsMIT*, *OsNAS1*, and *OsNAS2* on chromosome 3 and were located on the same LD block. Further, *qZnBR4.1* and *qZnMR4.1*, the co-localized MTAs associated with grain Zn concentration, were found linked to the reported candidate gene *OsZIP3*, while *qFeBR6.1* was found to be near the gene *OsNRAMP3* on chromosome 6.

DISCUSSION

Mineral micronutrient malnutrition is a widespread malady among the rice-eating populations, particularly in developing nations who cannot afford dietary diversity. Biofortification of popular rice varieties with mineral micronutrients, especially Fe and Zn, is a sustainable solution to tackle hidden hunger. However, polygenic inheritance of grain micronutrient accumulation in rice (Huang et al., 2015; Descalsota et al., 2018; Swamy et al., 2018) makes it relatively difficult to map

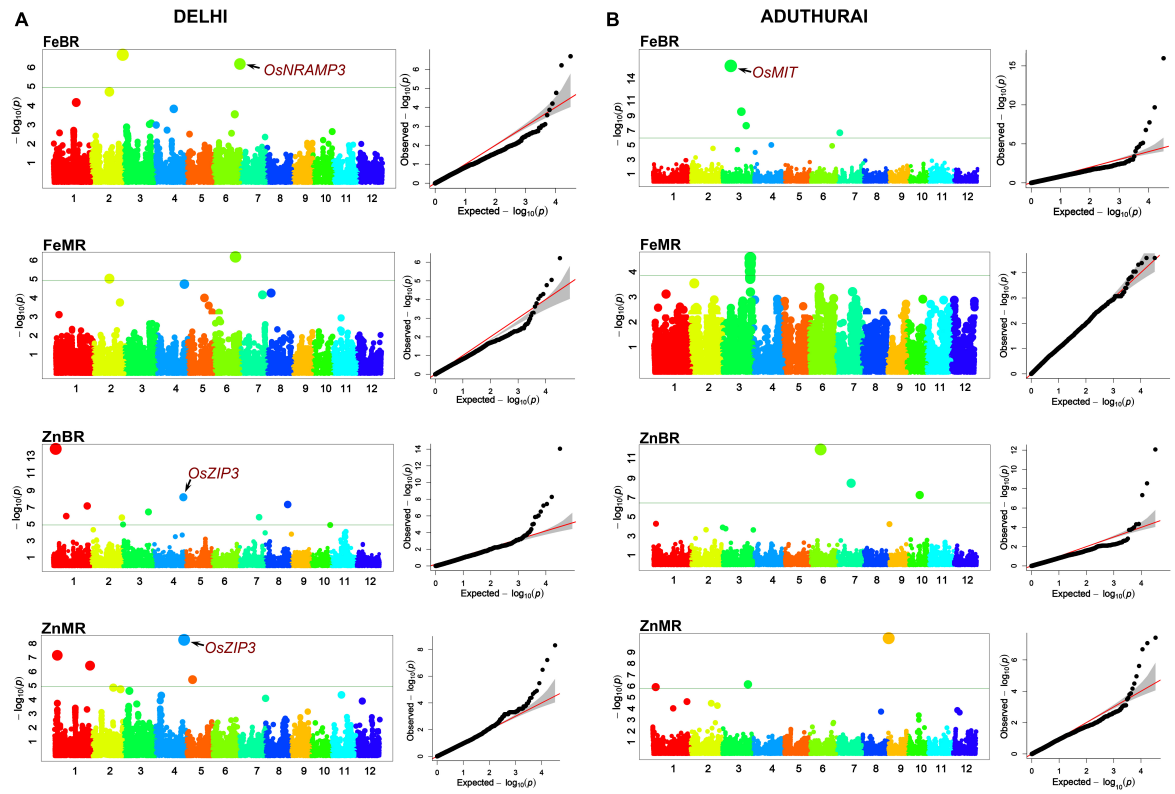


FIGURE 5 | Manhattan plots and Q–Q plots derived through FarmCPU model depicting the significant marker trait associations on 12 chromosomes of rice for the traits analyzed at New Delhi **(A)** and Aduthurai **(B)** locations.

multiple genes using a biparental population, when the genes have low individual effects and are sparsely distributed in the gene pool. A GWAS therefore offers a dual advantage of analyzing the extensive trait variation among the germplasm lines and identifying several genomic regions affecting the trait. In the present study, we analyzed a set of 192 rice germplasm accessions indigenous to different parts of India, for two important mineral micronutrients *viz.* Fe and Zn, in both BR and MR grown at two different locations, and found that the accessions possessed more extensive genetic diversity for both traits. Wide variability in rice accessions especially involving several landraces, has previously been reported particularly in BR (Anuradha et al., 2012b; Maganti et al., 2019). In this study, additionally, we demonstrated the existence of variation in MR which is more commonly consumed. In this case also, as reported earlier, we observed a pattern of polygenic inheritance of the Fe and Zn concentration from their normal distribution under both locations. This is in line with previous studies that demonstrated complex and multi-factorial inheritance of Fe and Zn in rice grain (Huang et al., 2015; Descalsota et al., 2018; Swamy et al., 2018).

A significant positive correlation between Fe and Zn in BR had been reported earlier (Stangoulis et al., 2007; Anuradha et al., 2012a) and together this study affirms the possibility of the simultaneous improvement of Fe and Zn in rice grain. As mentioned earlier, in rice grains, Fe and Zn are accumulated in the bran to the tune of about 30 and 6 mg/g, respectively (Bhosale and Vijayalakshmi, 2015), which accounts for the significant proportion of the mineral content in the grain. Additionally, the study also depicted strong association between Zn concentrations in BR and MR. It ascertains that the Zn concentration in BR is a fair indicator of its level in MR and can therefore be used as a selection criterion for the quick and non-destructive evaluation of Zn in the breeding populations targeted to improve Zn content in MR. In contrast, Fe concentration in BR was not found to be associated with Fe concentration in MR, signifying the practical need for enriching the Fe concentration of endosperm independent of its BR Fe concentration. When quantified, the percent retention upon polishing was found to be significantly low for Fe, in line with the previous reports that established localization of Fe in the embryo and aleurone layer (Prom-u-Thai et al., 2003; Choi et al., 2007; Kyriacou et al., 2014). Conversely, retention of Zn was significantly higher in the endosperm with moderate to minimal losses on polishing. However, examining carefully on an individual basis, the retention level varied across the genotypes, opening up the possibility of identifying genotypes with better nutrient retention in milled grains. Further studies are needed to investigate if this variation in percent retention upon polishing is due to differences in the ability of genotypes to translocate the mineral elements from

TABLE 3 | Details of the marker-trait associations (MTAs) mapped for grain Fe and Zn content at two different locations, New Delhi and Aduthurai.

SNP	Allele		Chromosome	Position	Probability	Trait	Site	QTL	R^2	a_i	Previous report
	Favorable allele	Alternate allele									
AX-95918814	A	G	1	3,565,733	6.19E-08	ZnMR	DEL	qZnMR1.1	9.3	1.7	*
AX-95918814	A	G	1	3,565,733	6.19E-08	ZnMR	ADT	*qZnMR1.1*	10.4	1.8	*
AX-95934119	T	C	1	3,568,378	8.72E-15	ZnBR	DEL	*qZnBR1.1*	12.4	3.4	*
AX-95919006	A	C	1	15,673,604	8.88E-07	ZnBR	DEL	*qZnBR1.2a*	45.5	4.8	Stangoulis et al., 2007
AX-95917932	G	A	1	39,382,522	5.63E-08	ZnBR	DEL	*qZnBR1.2b*	40.1	4.3	*
AX-95918225	G	T	1	41,121,295	3.33E-07	ZnMR	DEL	*qZnMR1.2*	17.7	4.2	*IRO2*
AX-95921433	C	A	2	34,958,759	2.01E-07	FeBR	DEL	*qFeBR2.1*	6.8	1.5	Swamy et al., 2018
AX-95965057	A	G	2	35,093,342	1.35E-06	ZnBR	DEL	*qZnBR2.1*	26.4	5.0	Swamy et al., 2018
AX-95922070	A	G	3	10,164,543	1.06E-16	FeBR	ADT	*qFeBR3.1*	53.3	4.1	*MIT*, *NAS1*, *NAS2*
AX-95962496	A	C	3	22,193,098	2.03E-10	FeBR	ADT	*qFeBR3.2*	2.1	0.2	*
AX-95921838	G	T	3	27,810,157	1.71E-08	FeBR	ADT	*qFeBR3.3*	2.1	0.1	*
AX-95923364	C	T	3	29,493,448	2.95E-07	ZnBR	DEL	*qZnBR3.1*	39.9	4.0	Anuradha et al., 2012a
AX-95921738	T	G	3	30,176,449	5.86E-07	ZnMR	ADT	*qZnMR3.1*	18.3	3.3	*
AX-95935621	G	A	3	32,326,592	4.79E-05	FeMR	ADT	*qFeMR3.1a*	12.8	0.4	*
AX-95950999	T	G	3	32,335,075	4.13E-05	FeMR	ADT	*qFeMR3.1b*	12.9	0.4	*
AX-95935460	G	A	3	32,374,286	8.90E-05	FeMR	ADT	*qFeMR3.1c*	12.3	0.4	*
AX-95924055	C	G	3	32,380,341	8.90E-05	FeMR	ADT	*qFeMR3.1d*	12.3	0.4	*
AX-95923317	T	C	3	32,380,432	2.61E-05	FeMR	ADT	*qFeMR3.1e*	13.7	0.4	*
AX-95923159	G	A	3	32,380,964	2.61E-05	FeMR	ADT	*qFeMR3.1f*	13.7	0.4	*
AX-95951158	A	G	4	32,811,874	5.24E-09	ZnBR	DEL	*qZnBR4.1*	47.6	6.3	*ZIP3*
AX-95951158	A	G	4	32,811,874	4.86E-09	ZnMR	DEL	*qZnMR4.1*	42.7	5.6	*ZIP3*
AX-95927387	A	G	6	11,840,203	8.42E-13	ZnBR	ADT	*qZnBR6.1*	44.9	7.4	*
AX-95928882	C	A	6	24,918,156	6.05E-07	FeMR	DEL	*qFeMR6.1*	3.2	0.6	*
AX-95956929	A	G	6	30,278,880	6.11E-07	FeBR	DEL	*qFeBR6.1*	45.2	2.7	*NRAMP3*
AX-95915606	A	T	7	2,636,599	1.66E-07	FeBR	ADT	*qFeBR7.1*	41.5	4.7	*
AX-95929962	T	G	7	15,616,980	2.71E-09	ZnBR	ADT	*qZnBR7.1*	16.5	3.4	Anuradha et al., 2012a
AX-95929638	T	A	7	22,727,786	1.23E-06	ZnBR	DEL	*qZnBR7.2*	33.4	6.6	Anuradha et al., 2012a
AX-95930744	C	A	8	25,586,491	3.74E-08	ZnBR	DEL	*qZnBR8.1*	2.3	0.6	Garcia-Oliveira et al., 2009
AX-95959928	T	C	9	849,821	4.76E-11	ZnMR	ADT	*qZnMR9.1*	7.9	2.2	*
AX-95932094	A	G	10	12,685,215	4.51E-08	ZnBR	ADT	*qZnBR10.1*	6.8	1.7	*

SNP, single nucleotide polymorphic marker; DEL, New Delhi; ADT, Aduthurai; PVE, percentage variation explained; a_i, additive effect of the favorable allele; FeBR, Fe content in brown rice; FeMR, Fe content in milled rice; ZnBR, Zn content in brown rice; ZnMR, Zn content in milled rice; asterisk indicates MTAs not reported so far.

aleurone to endosperm, or due to differences in the thickness of the aleurone layer.

GWAS Identified Significant MTAs for Biofortification

Assessment of genetic diversity and population structure is an important pre-requisite in a GWAS and the presence of three sub-populations in our panel has been depicted by both PCA and STRUCTURE analyses. Variations observed among the three sub-populations for the Fe and Zn in both BR and MR at two locations, implied that grain micronutrients can only be subtly influenced by genetic grouping. The genotypes in the POP3 showed specific adaptation to location, with higher mean values recorded at the New Delhi location compared to the Aduthurai location. Huang et al. (2015) also reported the importance of population structure in determining the variation for the mineral micronutrients Fe and Zn and heavy metals like Pb, Cd, and Se in whole grain rice.

It has been proven that the GWAS accelerated the speed and accuracy of detecting QTLs and candidate genes in comparison to biparental linkage mapping. Several statistical models like GLM and MLM have been developed to control false positives and false negatives that arise due to familial relatedness and population structure. Ever since the publication of MLM, it has been popularly adopted for GWAS in crops, particularly, in rice (Zhang et al., 2014; Ya-fang et al., 2015; Wang et al., 2016). Nevertheless, MLM being a single locus method that allows testing of one marker locus at a time, had an inherent limitation in matching the real genetic architecture of the complex traits that are under the influence of multiple loci acting simultaneously (Kaler and Purcell, 2019). Multi-locus models like FASTmrEMMAa (Zhang et al., 2018), LASSO (Xu et al., 2017), BLASSO (Tamba et al., 2017), FarmCPU, pLARmEB (Zhang et al., 2018), and pKWmEB (Ren et al., 2018) are being

used to overcome the limitation above. A few recent studies on plant height and flowering time (Wallace et al., 2016), ear traits (Zhu et al., 2018), and starch pasting properties (Xu et al., 2018) in maize, yield-related features in wheat (Ward et al., 2019), stem rot resistance in soybean (Wei et al., 2017), agronomic traits in foxtail millet (Jaiswal et al., 2019), and panicle architecture in sorghum (Zhou et al., 2019), have demonstrated the power of the FarmCPU model that uses both fixed effect and random effect models iteratively to effectively control the false discovery. In the present study, a comparison of Q–Q plots obtained through different models revealed FarmCPU as a best-fit model with improved power of test statistics. Our research forms the first report in rice that demonstrates the power of the FarmCPU over MLM.

Population size is another critical factor that determines the power of detection of QTLs in a GWAS (Garcia et al., 2005). In the present study, we identified significant MTAs even after adopting the highly stringent Bonferroni threshold, implying that the population size of 192 individuals is just enough for a GWAS in rice. Hoang et al. (2019) used a set of 180 individuals in a GWAS to identify QTLs affecting tolerance to water deficit while Descalsota et al. (2018) used a set of 144 MAGIC lines as a GWAS panel and mapped significant QTLs for agronomic and biofortification traits in rice. As rice is an autogamous species, it carries haplotype blocks larger than allogamous crops like corn. Yonemaru et al. (2012) demonstrated that size of the haplotype blocks in rice could vary widely, with a mean of ∼50 kb, while in corn it was ∼1 kb (Maldonado et al., 2019). Because of this, in crops like rice, a relatively smaller number of genotypes are required to cover a more significant proportion of the evolutionarily conserved genomic regions. Throughout the current decade, the use of GWAS has been particularly widespread in rice, helping to map QTLs related to traits such as agronomic characters (Zhang et al., 2014), plant height and grain yield (Ma et al., 2016), grain traits (Edzesi et al., 2016; Feng et al., 2016; Wang et al., 2016), panicle traits (Zhang et al., 2014), and milling quality (Qiu et al., 2015) to mention a few. Nevertheless, only a couple of studies reported a GWAS for mineral elements either in BR (Norton et al., 2014; Yang et al., 2018) or in MR (Descalsota et al., 2018).

In the present study, we have identified a total of 29 QTLs affecting Fe and Zn concentration in both BR and MR, holding significance in rice biofortification programs. Interestingly, 18 of the 29 MTAs were different from previous reports and can therefore be considered novel MTAs. The correlated traits like FeBR, ZnBR, and ZnMR shared common MTAs while uncorrelated features such as FeBR and FeMR did not share any co-localized MTAs. These co-localized MTAs can be targeted for the simultaneous improvement of Fe and Zn in rice grain. Further, the MTA, *qFeBR3.1* reported in this study on chromosome 3 was found in close proximity (222 kb) to the candidate gene, the mitochondrial iron transporter (*OsMIT*). Using knockdown mutants, *MIT* gene expression was shown to affect Fe localization in rice seeds (Bashir et al., 2013). Additionally, *qFeBR3.1* also shares the same LD block with the nicotianamine synthase (NAS) family protein *OsNAS1* (760 Kb) and *OsNAS2* (763 Kb). NAS family members catalyze the biosynthesis nicotinamine (NA) that acts a chelator of metal cations like Fe^{2+} and Fe^{3+} and play an essential role in both the short-distance as well as long-distance transport of metal cations (von Wiren et al., 1999; Takahashi et al., 2003). Overexpression of NAS genes resulted in several folds increase in Fe, Zn, and NA in rice grain (Lee et al., 2009; Johnson et al., 2011). The MTAs, *qZnBR4.1*, and *qZnMR4.1* on chromosome 4 were found to be co-localized with the *OsZIP3* gene, one of the members of the Zn-regulated transporter family proteins in rice. Initially, it was reported that the cation transporter proteins ZIP1 and ZIP3 might be involved in the uptake of Zn from soil (Ramesh et al., 2003; Bashir et al., 2012). Subsequently, Sasaki et al. (2015) observed that *OsZIP3* is localized in the nodes, and they further demonstrated that the OsZIP3 protein is responsible for unloading the Zn from the xylem of enlarged vascular bundles in nodes. In addition to these vital candidate genes, 7 of the 29 MTAs were within the intervals of previously mapped QTLs (**Table 3**).

Potential Donor Germplasm Identified

We have identified four accessions (**Table 4**) with grain Zn concentration in MR > 28 mg/kg, a target set by the HarvestPlus program for rice bio-fortification (Bouis and Saltzman, 2017). These accessions showed consistent high grain Zn in both the environments tested. Additionally, several specifically adapted germplasm were also identified in the study. However, for the Fe concentration in MR, only one accession (IC-2127) showed >12 mg/kg Fe, as per the HarvestPlus target, specifically in the New Delhi location. Our data is in agreement with the previous studies that reported an average of only 2 mg/kg Fe in MR (Johnson et al., 2011; Yuan et al., 2013),

TABLE 4 | The average grain content (both brown and milled rice) of Fe and Zn (mg/kg) in elite accessions identified from the association mapping panel to be used as donors in bio-fortification programs.

Lines	New Delhi				Aduthurai			
	FeBR	FeMR	ZnBR	ZnMR	FeBR	FeMR	ZnBR	ZnMR
Karuppunel	16.2	5.4	46.2	40.9	11.9	2.6	47.0	40.8
Budgi	17.8	1.4	35.6	31.6	17.4	2.0	35.4	30.2
Mehvan green	20.5	1.6	35.6	28.5	20.1	2.4	39.0	29.9
Mehvan purple	21.7	4.2	37.0	33.4	17.1	3.3	33.8	29.6

FeBR, Fe content in brown rice; FeMR, Fe content in milled rice; ZnBR, Zn content in brown rice; ZnMR, Zn content in milled rice.

indicating the limited scope for the improvement of Fe concentration in rice using conventional breeding approaches. Genetic engineering, especially the recent techniques of genome editing like CRISPR/Cas9 and transcription activator-like effector nucleases, offers viable alternatives to traditional methods for Fe bio-fortification in rice grain (Goto et al., 1999; Wirth et al., 2009; Masuda et al., 2012; Trijatmiko et al., 2016).

DATA AVAILABILITY STATEMENT

The SNP data has been deposited into a publicly accessible repository held under ICAR: https://krishi.icar.gov.in/jspui/handle/123456789/31947.

AUTHOR CONTRIBUTIONS

AS conceptualized the idea. HB, PB, SG, MN, and RE conducted the field experiments. HB, AY, and CN carried out phenotyping of the germplasm accessions. HB and KV carried out the data analysis. HB and KV prepared the manuscript. HB, AS, KV, and SG edited the manuscript.

FUNDING

The authors acknowledge the Indian Council of Agricultural Research (ICAR) for funding and supporting the work under the CRP-Biofortification project.

REFERENCES

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

Anuradha, K., Agarwal, S., Batchu, A. K., Babu, P. A., Swamy, B. P. M., Longvah, T., et al. (2012a). Evaluating rice germplasm for iron and zinc concentration in brown rice and seed dimensions. *J. Phytol.* 4, 19–25.

Anuradha, K., Agarwal, S., Rao, Y. V., Rao, K. V., Viraktamath, B. C., and Sarla, N. (2012b). Mapping QTLs and candidate genes for iron and zinc concentrations in unpolished rice of Madhukar× Swarna RILs. *Gene* 508, 233–240. doi: 10.1016/j.gene.2012.07.054

Bashir, K., Ishimaru, Y., and Nishizawa, N. K. (2012). Molecular mechanisms of zinc uptake and translocation in rice. *Plant Soil* 361, 189–201. doi: 10.1007/s11104-012-1240-5

Bashir, K., Nozoye, T., Ishimaru, Y., Nakanishi, H., and Nishizawa, N. K. (2013). Exploiting new tools for iron bio-fortification of rice. *Biotechnol. Adv.* 31, 1624–1633. doi: 10.1016/j.biotechadv.2013.08.012

Bhosale, S., and Vijayalakshmi, D. (2015). Processing and nutritional composition of rice bran. *Curr. Res. Nutr. Food Sci.* 3, 74–80. doi: 10.12944/CRNFSJ.3.1.08

Black, R. E., Allen, L. H., Bhutta, Z. A., Caulfield, L. E., de Onis, M., Ezzati, M., et al. (2008). Maternal and child undernutrition: global and regional exposures and health consequences. *Lancet* 371, 243–260. doi: 10.1016/S0140-6736(07)61690-0

Bouis, H. E. (2003). Micronutrient fortification of plants through plant breeding: can it improve nutrition in man at low cost? *Proc. Nutr. Soc.* 62, 403–411. doi: 10.1079/pns2003262

Bouis, H. E., and Saltzman, A. (2017). Improving nutrition through biofortification: a review of evidence from HarvestPlus, 2003 through 2016. *Glob. Food. Sec.* 12, 49–58. doi: 10.1016/j.gfs.2017.01.009

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 indiverse samples. *Bioinformatics* 23, 2633–2635. doi: 10.1093/bioinformatics/btm308

Brown, K. H., Rivera, J. A., Bhutta, Z., Gibson, R. S., King, J. C., Lönnerdal, B., et al. (2004). International Zinc Nutrition Consultative Group (IZiNCG). Assessment of the risk of zinc deficiency in populations and options for its control. *Food Nutr. Bull.* 25(Suppl.2), S99–S203.

Brown, K. H., Wuehler, S. E., and Peerson, J. M. (2001). The importance of zinc in human nutrition and estimation of the global prevalence of zinc deficiency. *Food Nutr. Bull.* 22, 113–125. doi: 10.1177/156482650102200201

Caulfield, L. E., Richard, S. A., and Black, R. E. (2004). Under nutrition as an underlying cause of malaria morbidity and mortality in children less than five years old. *Am. J. Trop. Med. Hyg.* 71, S55–S63.

Choi, E. Y., Graham, R., and Stangoulis, J. (2007). Semi-quantitative analysis for selecting Fe-andZn-dense genotypes of staple food crops. *J. Food Compos. Anal.* 7, 496–505. doi: 10.1016/j.jfca.2007.01.004

Descalsota, G. I. L., Swamy, M., Zaw, H., Asilo, M. A., Amparado, A., Mauleon, R. P., et al. (2018). Genome-wide association mapping in a rice MAGIC plus population detects QTLs and genes useful for biofortification. *Front. Plant Sci.* 9:1347. doi: 10.3389/fpls.2018.01347

Earl, D. A., and vonHoldt, B. (2012). Structure harvester: a website and program for visualizing structure output and implementing the Evanno method. *Cons. Genet. Res.* 4, 359–361. doi: 10.1007/s12686-011-9548-7

Edzesi, W. M., Dang, X., Liang, L., Liu, E., Zaid, I. U., and Hong, D. (2016). Genetic diversity and elite allele mining for grain traits in rice (*Oryza sativa* L.) by association mapping. *Front. Plant Sci.* 7:787. doi: 10.3389/fpls.2016.00787

Evanno, G., Regnaut, S., and Goudet, J. (2005). Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol. Ecol.* 14, 2611–2620. doi: 10.1111/j.1365-294x.2005.02553.x

Feng, Y., Lu, Q., Zhai, R., Zhang, M., Xu, Q., Yang, Y., et al. (2016). Genome wide association mapping for grain shape traits in indica rice. *Planta* 244, 819–830. doi: 10.1007/s00425-016-2548-9

Garcia, F. S. A., Thuillet, A. C., Yu, J., Pressoir, G., Romero, S. M., Mitchell, S. E., et al. (2005). Maize association population: a high-resolution platform for quantitative trait locus dissection. *Plant J.* 44, 1054–1064. doi: 10.1111/j.1365-313X.2005

Garcia-Oliveira, A. L., Tan, L., Fu, Y., and Sun, C. (2009). Genetic identification of quantitative trait loci for contents of mineral nutrients in rice grain. *J. Integr. Plant Biol.* 51, 84–92. doi: 10.1111/j.1744-7909.2008.00730.x

Goto, F., Yoshihara, T., Shigemoto, N., Toki, S., and Takaiwa, F. (1999). Iron fortification of rice seed by the soybean ferritin gene. *Nat. Biotechnol.* 17, 282–286. doi: 10.1038/7029

Gregorio, G. B., Senadhira, D., Htut, H., and Graham, R. D. (2000). Breeding for trace mineraldensity in rice. *Food Nutr. Bull.* 21, 382–386. doi: 10.1177/156482650002100407

Hoang, G. T., Gantet, P., Nguyen, K. H., Phung, N. T. P., Ha, L. T., Nguyen, T. T., et al. (2019). Genome-wide association mapping of leaf mass traits in a Vietnamese rice landrace panel. *PLoS One* 14:e0219274. doi: 10.1371/journal.pone.0219274

Huang, X., Wei, X., Sang, T., Zhao, Q., Feng, Q., Zhao, Y., et al. (2010). Genome-wide association studies of 14 agronomic traits in rice landraces. *Nat. Genet.* 42, 961–967. doi: 10.1038/ng.695

Huang, Y., Sun, C., Min, J., Chen, Y., Tong, C., and Bao, J. (2015). Association mapping of quantitative trait loci for mineral element contents in whole grain rice (*Oryza sativa* L.). *J. Agric. Food. Chem.* 63, 10885–10892. doi: 10.1021/acs.jafc.5b04932

Jaiswal, V., Bandyopadhyay, T., Gahlaut, V., Gupta, S., Dhaka, A., Ramchiary, N., et al. (2019). Genome-wide association study (GWAS) delineates genomic loci for ten nutritional elements in foxtail millet (*Setariaitalica* L.). *J. Cereal Sci.* 85, 48–55. doi: 10.1016/j.jcs.2018.11.006

Johnson, A. A., Kyriacou, B., Callahan, D. L., Carruthers, L., Stangoulis, J., Lombi, E., et al. (2011). Constitutive overexpression of the OsNAS gene family reveals single-gene strategies for effective iron- and zinc-biofortification of rice endosperm. *PLoS One* 6:e24476. doi: 10.1371/journal.pone.0024476

Kaler, A. S., and Purcell, L. C. (2019). Estimation of a significance threshold for genome-wide association studies. *BMC Genomics* 20:618. doi: 10.1186/s12864-019-5992-7

King, J., Cousins, R., and Shils, M. (eds) (2006). *Modern Nutrition in Health and Disease*, 10th Edn. Philadelphia, PA: Lippincott Williams and Wilkins, 271–285.

Kyriacou, B., Alexander, J. A. T., Howard, L. D., Paterson, D., Lombi, E., Stangoulis, J., et al. (2014). Localization of iron in rice grain using synchrotron X-ray fluorescence microscopy and high resolution secondary ion mass spectrometry. *J. Cereal Sci.* 59, 173–180. doi: 10.1016/j.jcs.2013.12.006

Lee, S., Jeon, U. S., Lee, S. J., Kim, Y. K., Persson, D. P., Husted, S., et al. (2009). Iron fortification of rice seeds through activation of the nicotianamine synthase gene. *Proc. Natl. Acad. Sci. U.S.A.* 106, 22014–22019. doi: 10.1073/pnas.0910950106

Lim, K., Riddell, L., Nowson, C., Booth, A., and Szymlek-Gay, E. (2013). Iron and zinc nutrition in the economically-developed world: a review. *Nutrients* 5, 3184–3211. doi: 10.3390/nu5083184

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444

Liu, X., Huang, M., Fan, B., Buckler, E. S., and Zhang, Z. (2016). Iterative usage of fixed and random effect models for powerful and efficient genome-wide association studies. *PLoS Genet.* 12:e1005767. doi: 10.1371/journal.pgen.1005767

Liu, X. L. (2015). *Development of an Iterative Usage of Fixed Effect and Random Effect Models for Powerful and Efficient Genome-Wide Association Studies*. Master's thesis, Huazhong agricultural University, Wuhan.

Ma, X., Feng, F., Wei, H., Mei, H., Xu, K., Chen, S., et al. (2016). Genome-wide association study for plant height and grain yield in rice under contrasting moisture regimes. *Front. Plant Sci.* 7:1801. doi: 10.3389/fpls.2016.01801

Maganti, S., Swaminathan, R., and Parida, A. (2019). Variation in Iron and Zinc content in traditional rice genotypes. *Agric. Res.* doi: 10.1007/s40003-019-00429-3

Maldonado, C., Mora, F., Scapim, C. A., and Coan, M. (2019). Genome-wide haplotype-based association analysis of key traits of plant lodging and architecture of maize identifies major determinants for leaf angle: hapLA4. *PLoS One* 14:e0212925. doi: 10.1371/journal.pone.0212925

Masuda, H., Ishimaru, Y., Aung, M. S., Kobayashi, T., Kakei, Y., Takahashi, M., et al. (2012). Iron biofortification in rice by the introduction of multiple genes involved in iron nutrition. *Sci. Rep.* 2:543. doi: 10.1038/srep00543

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325. doi: 10.1093/nar/8.19.4321

Norton, G. J., Douglas, A., Lahner, B., Yakubova, E., Guerinot, M. L., Pinson, S. R., et al. (2014). Genome wide association mapping of grain Arsenic, Copper, Molybdenum and Zinc in rice (*Oryza sativa* L.) grown at four international field sites. *PLoS One* 9:e89685. doi: 10.1371/journal.pone.0089685

Pradhan, S. K., Pandit, E., Pawar, S., Naveenkumar, R., Barik, S. R., Mohanty, S. P., et al. (2020). Linkage disequilibrium mapping for grain Fe and Zn enhancing QTLs useful for nutrient dense rice breeding. *BMC Plant Biol.* 20:57. doi: 10.1186/s12870-020-2262-4

Prasad, A. S. (1991). Discovery of human zinc deficiency and studies in an experimental human model. *Am. J. Clin. Nutr.* 53, 403–412. doi: 10.1093/ajcn/53.2.403

Pritchard, J. K., Stephens, M., and Donnelly, P. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Prom-u-Thai, C., Dell, B., Thomson, G., and Rerkasem, B. (2003). Easy and rapid detection of iron in rice grain. *Sci. Asia* 29, 203–207. doi: 10.2306/scienceasia1513-1874

Qiu, C., Zhao, M., Decker, E. A., and McClements, D. J. (2015). Influence of anionic dietary fibers (xanthan gum and pectin) on oxidative stability and lipid digestibility of wheat protein-stabilized fish oil-in-water emulsion. *Food Res. Int.* 74, 131–139. doi: 10.1016/j.foodres.2015.04.022

Ramesh, S. A., Shin, R., Eide, D. J., and Schachtman, D. P. (2003). Differential metal selectivity and gene expression of two zinc transporters from rice. *Plant Physiol.* 133, 126–134. doi: 10.1104/pp.103.026815

Ren, W. L., Wen, Y. J., Dunwell, J. M., and Zhang, Y. M. (2018). pKWmEB: integration of Kruskal-Wallis test with empirical Bayes under polygenic background control for multi-locus genome-wide association study. *Heredity* 120, 208–218. doi: 10.1038/s41437-017-0007-4

Roeser, H. P. (1986). Iron. *J. Food Nutr.* 42, 82–92.

Salgueiro, M. J., Zubillaga, M., Lysionek, A., Sarabia, M. I., Caro, R., De Paoli, T., et al. (2000). Zinc as an essential micronutrient: a review. *Nutr. Res.* 20, 737–755. doi: 10.1016/S0271-5317(00)00163-9

Sasaki, A., Yamaji, N., Mitani-Ueno, N., Kashino, M., and Ma, J. F. (2015). A node-localized transporter OsZIP3 is responsible for the preferential distribution of Zn to developing tissues in rice. *Plant J.* 84, 374–384. doi: 10.1111/tpj.13005

Singh, N., Jayaswal, P. K., Panda, K., Mandal, P., Kumar, V., Singh, B., et al. (2015). Single-copy gene based 50 K SNP chip for genetic studies and molecular breeding in rice. *Sci. Rep.* 5:11600. doi: 10.1038/srep11600

Stangoulis, J. C., Huynh, B. L., Welch, R. M., Choi, E. Y., and Graham, R. D. (2007). Quantitative trait loci for phytate in rice grain and their relationship with grain micronutrient content. *Euphytica* 154, 289–294. doi: 10.1007/s10681-006-9211-7

Stoltzfus, R. J. (2001). Iron-deficiency anemia: reexamining the nature and magnitude of the public health problem Summary: implications for research and programs. *J. Nutr.* 131, 697S–700S. doi: 10.1093/jn/131.2.697S

Swamy, B. P. M., Kaladhar, K., Anuradha, K., Batchu, A. K., Longvah, T., and Sarla, N. (2018). QTL analysis for grain iron and zinc concentrations in two *O. nivara* derived backcross populations. *Rice Sci.* 25, 197–207. doi: 10.1016/j.rsci.2018.06.003

Takahashi, M., Terada, Y., Nakai, I., Nakanishi, H., Yoshimura, E., Mori, S., et al. (2003). Role of nicotianamine in the intracellular delivery of metals and plant reproductive development. *Plant Cell* 15, 1263–1280. doi: 10.1105/tpc.010256

Tamba, C. L., Ni, Y. L., and Zhang, Y. M. (2017). Iterative sure independence screening M-Bayesian LASSO algorithm for multi-locus genome-wide association studies. *PLoS Comput. Biol.* 13:e1005357. doi: 10.1371/journal.pcbi.1005357

Trijatmiko, K. R., Dueñas, C., Tsakirpaloglou, N., Torrizo, L., Arines, F. M., Adeva, C., et al. (2016). Biofortifiedindica rice attains iron and zinc nutrition dietary targets in the field. *Sci. Rep.* 6:19792. doi: 10.1038/srep19792

von Wiren, N., Klair, S., Bansal, S., Briat, J. F., Khodr, H., Shioiri, T., et al. (1999). Nicotianamine chelates both FeIII and FeII. Implications for metal transport in plants. *Plant Physiol.* 119, 1107–1114. doi: 10.1104/pp.119.3.1107

Wallace, J. G., Zhang, X., Beyene, Y., Semagn, K., Olsen, M., Prasanna, B. M., et al. (2016). Genome-wide association for plant height and flowering time across 15 tropical maize populations under managed drought stress and well-watered conditions in Sub-Saharan Africa. *Crop Sci.* 56, 2365–2378. doi: 10.2135/cropsci2015.10.0632

Wang, Y., Zheng, Y., Cai, Q., Liao, C., Mao, X., Xie, H., et al. (2016). Population structure and association analysis of yield and grain quality traits in hybrid rice primal parental lines. *Euphytica* 212, 261–273. doi: 10.1007/s10681-016-1766-3

Ward, B. P., Brown-Guedira, G., Kolb, F. L., Van Sanford, D. A., Tyagi, P., Sneller, C. H., et al. (2019). Genome-wide association studies for yield-related traits in soft red winter wheat grown in Virginia. *PLoS One* 14:e0208217. doi: 10.1371/journal.pone.0208217

Wei, W., Mesquita, A. C. O., Figueiró, A. D. A., Wu, X., Manjunatha, S., Wickland, D. P., et al. (2017). Genome-wide association mapping of resistance to a Brazilian isolate of Sclerotinia sclerotiorum in soybean genotypes mostly from Brazil. *BMC Genomics* 18:849. doi: 10.1186/s12864-017-4160-1

Wirth, J., Poletti, S., Aeschlimann, B., Yakandawala, N., Drosse, B., Osorio, S., et al. (2009). Rice endosperm iron biofortification by targeted and synergistic action of nicotianamine synthase and ferritin. *Plant Biol. J.* 7, 631–644. doi: 10.1111/j.1467-7652.2009.00430.x

Wood, R., Ronnenberg, A., and Shils, M. (eds) (2006). *Modern nutriTion in Health and Disease*, 10th Edn. Philadelphia, PA: Lippincott Williams and Wilkins, 248–270.

Xu, Y., Xu, C., and Xu, S. (2017). Prediction and association mapping of agronomic traits in maize using multiple omic data. *Heredity* 119, 174–184. doi: 10.1038/hdy.2017.27

Xu, Y., Yang, T., Zhou, Y., Yin, S., Li, P., Liu, J., et al. (2018). Genome-wide association apping of starch pasting properties in maize using single-locus and multi-locus models. *Front. Plant Sci.* 9:1311. doi: 10.3389/fpls.2018.01311

Ya-fang, Z., Yu-yin, M. A., Zong-xiang, C., Jie, Z., Tian-xiao, C., Qian-qian, L., et al. (2015). Genome-wide association studies reveal new genetic targets for

five panicle traits of international rice varieties. *Rice Sci.* 22, 217–226. doi: 10.1016/j.rsci.2015.07.001

Yang, M., Lu, K., Zhao, F. J., Xie, W., Ramakrishna, P., Wang, G., et al. (2018). Genome-wide association studies reveal the genetic basis of ionomic variation in rice. *Plant Cell* 30, 2720–2740. doi: 10.1105/tpc.18.00375

Yonemaru, J. I., Yamamoto, T., Ebana, K., Yamamoto, E., Nagasaki, H., Shibaya, T., et al. (2012). Genome-wide haplotype changes produced by artificial selection during modern rice breeding in Japan. *PLoS One* 7:e32982. doi: 10.1371/journal.pone.0032982

Yuan, L., Wu, L., Yang, C., and Lv, Q. (2013). Effects of iron and zinc foliar applications on rice plants and their grain accumulation and grain nutritional quality. *J. Sci. Food Agric.* 93, 254–261. doi: 10.1002/jsfa.5749

Zhang, C. M., Zhao, W. Y., Gao, A. X., Su, T. T., Wang, Y. K., Zhang, Y. Q., et al. (2018). How could agronomic biofortification of rice be an alternative strategy with higher cost-effectiveness for human iron and zinc deficiency in China? *Food Nutr. Bull.* 39, 246–259. doi: 10.1177/0379572117745661

Zhang, P., Liu, X., Tong, H., Lu, Y., and Li, J. (2014). Association mapping for important agronomic traits in core collection of rice (*Oryza sativa* L.) with SSR markers. *PLoS One* 9:e111508. doi: 10.1371/journal.pone.0111508

Zhou, B., Zhao, H., Puig, X., Xiao, T., Fidler, S., Barriuso, A., et al. (2019). Semantic understanding of scenes through the ade20k dataset. *Int. J. Comput. Vis.* 127, 302–321. doi: 10.1007/s11263-018-1140-0

Zhu, X. M., Shao, X. Y., Pei, Y. H., Guo, X. M., Li, J., Song, X. Y., et al. (2018). Genetic diversity and genome-wide association study of major ear quantitative traits using high-density SNPs in maize. *Front. Plant Sci.* 9:966. doi: 10.3389/fpls.2018.00966

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 Genetics

REVIEW
published: 29 May 2020
doi: 10.3389/fgene.2020.00414

Genomics-Integrated Breeding for Carotenoids and Folates in Staple Cereal Grains to Reduce Malnutrition

Kaliyaperumal Ashokkumar[1], Mahalingam Govindaraj[2], Adhimoolam Karthikeyan[3], V. G. Shobhana[2] and Thomas D. Warkentin[4]*

[1] Crop Improvement, Cardamom Research Station, Agricultural University, Pampadumpara, India, [2] Crop Improvement program, International Crops Research Institute for the Semi-Arid Tropics, Hyderabad, India, [3] Subtropical Horticulture Research Institute, Jeju National University, Jeju, South Korea, [4] Department of Plant Sciences, College of Agriculture and Bioresources, University of Saskatchewan, Saskatoon, SK, Canada

Globally, two billion people suffer from micronutrient deficiencies. Cereal grains provide more than 50% of the daily requirement of calories in human diets, but they often fail to provide adequate essential minerals and vitamins. Cereal crop production in developing countries achieved remarkable yield gains through the efforts of the Green Revolution (117% in rice, 30% in wheat, 530% in maize, and 188% in pearl millet). However, modern varieties are often deficient in essential micronutrients compared to traditional varieties and land races. Breeding for nutritional quality in staple cereals is a challenging task; however, biofortification initiatives combined with genomic tools increase the feasibility. Current biofortification breeding activities include improving rice (for zinc), wheat (for zinc), maize (for provitamin A), and pearl millet (for iron and zinc). Biofortification is a sustainable approach to enrich staple cereals with provitamin A, carotenoids, and folates. Significant genetic variation has been found for provitamin A (96–850 μg and 12–1780 μg in 100 g in wheat and maize, respectively), carotenoids (558–6730 μg in maize), and folates in rice (11–51 μg) and wheat (32.3–89.1 μg) in 100 g. This indicates the prospects for biofortification breeding. Several QTLs associated with carotenoids and folates have been identified in major cereals, and the most promising of these are presented here. Breeding for essential nutrition should be a core objective of next-generation crop breeding. This review synthesizes the available literature on folates, provitamin A, and carotenoids in rice, wheat, maize, and pearl millet, including genetic variation, trait discovery, QTL identification, gene introgressions, and the strategy of genomics-assisted biofortification for these traits. Recent evidence shows that genomics-assisted breeding for grain nutrition in rice, wheat, maize, and pearl millet crops have good potential to aid in the alleviation of micronutrient malnutrition in many developing countries.

Keywords: biofortification, nutri-genomics, cereal, folate, provitamin A, lutein, zeaxanthin, human nutrition

OPEN ACCESS

Edited by:
Mallikarjuna Swamy, International Rice Research Institute, Philippines

Reviewed by:
Neeraja N. C., Indian Institute of Rice Research (ICAR), India
Rajneesh Paliwal, International Institute of Tropical Agriculture, Nigeria
Maryke T. Labuschagne, University of the Free State, South Africa

****Correspondence:***
Mahalingam Govindaraj
m.govindaraj@cgiar.org

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *07 October 2019*
Accepted: *01 April 2020*
Published: *29 May 2020*

Citation:
Ashokkumar K, Govindaraj M, Karthikeyan A, Shobhana VG and Warkentin TD (2020) Genomics-Integrated Breeding for Carotenoids and Folates in Staple Cereal Grains to Reduce Malnutrition. Front. Genet. 11:414. doi: 10.3389/fgene.2020.00414

INTRODUCTION

Micronutrient and vitamin-deficiency-induced malnutrition is widely prevalent in South Asia and sub-Saharan Africa, affecting approximately two billion people worldwide. In the human diet, more than 50% of total calories come from major cereals, including rice, wheat, and maize, in developing countries and more than 70% in Southeast Asia and Africa. The green revolution contributed

to remarkable increases in grain yield in these crops, which helped to prevent starvation in developing countries (Bouis and Welch, 2010). It is well known that cereal grains supply enough calories; however, these grains are inherently low in essential micronutrients, including carotenoids and folates (Bouis and Welch, 2010). The global production of rice is 769.4 m tons (from 167.2 m ha), wheat is 771.7 m tons (from 218.5 m ha), maize is 1134.7 m tons (from 197.2 m ha), and millet is 28.4 m tons (from 31.2 m ha) (Food and Agriculture Organization [FAO] et al., 2017), i.e., these crops play a critical role in food systems. Therefore, enhancing the nutritional quality of staple cereal crops is important for human health, particularly for resource-poor people in developing countries. Globally, 792.5 million people are malnourished, of which 780 million people live in developing countries (McGuire, 2015). Globally, two billion people suffer from hidden hunger due to inadequacies of micronutrients in their daily diet (Muthayya et al., 2013). Although major attention has been given to iron and zinc, in this review we also report on breeding efforts to improve concentrations of provitamin A, folate, and carotenoids.

Carotenoids are the second largest group of naturally occurring lipophilic pigments, following flavonoids, and at least 50 of them occur in plants. The most important carotenoids in food crops are β-carotene, α-carotene, β-cryptoxanthin, lutein, zeaxanthin, and lycopene. These carotenes are metabolized and converted to provitamin A (Davey et al., 2009). Humans are incapable of carotenoid biosynthesis, and we therefore depend on dietary carotenoid sources from plant-based foods (Fraser and Bramley, 2004). More than three million children in developing countries are affected by xerophthalmia, and 250,000–500,000 people become blind each year because of vitamin A deficiency (Food and Agriculture Organization [FAO] et al., 2017). The Recommended Dietary Allowance (RDA) of vitamin A for men and women is 900 and 700 μg Retinol Activity Equivalents (RAE)/day, respectively. For dietary provitamin A carotenoids, β-carotene, α-carotene, and β-cryptoxanthin RAEs have been set at 12, 24, and 24 μg, respectively (Institute of Medicine Food and Nutrition Board, 1998).

Folates act as cofactors in several metabolic functions, including the biosynthesis of nucleic acids and methylation of hormones, lipids, and proteins (Forges et al., 2007). Among many naturally occurring folates, cereal and pulse grains largely contain tetrahydrofolic acid (THF), 5-methyl-THF (5-MTHF), 10-formyl-THF (10-FTHF), and pteroylpolyglutamates (Jha et al., 2015; Ashokkumar et al., 2018b). Folate deficiency is a major problem for people from developing countries and can cause severe health issues, including impaired cognitive function, neural tube defects, and cardiovascular diseases (Ramos et al., 2005; McCully, 2007) as well as low birth weight, preterm delivery, and fetal growth retardation (Scholl and Johnson, 2000). Over 300,000 birth defects occur each year worldwide due to folate-deficiency-induced neural tube defects (Flores et al., 2014). Consumption of a folate-rich diet, fortification of foods with folic acid, and folic acid supplements can increase folate concentration in humans (Hefni et al., 2010). The RDA of folates is 400 μg for adults, 500 μg for lactating women, and 600 μg for pregnant women (Institute of Medicine Food and Nutrition Board, 1998).

Biofortification of staple crops through plant breeding and genomics integrated approaches is an effective strategy for delivering vitamins and nutrients to reduce micronutrient deficiencies in developing countries (Bouis, 2002; Welch and Graham, 2005). As urban development increasingly occupies fertile lands, the achievable agricultural production will be pushed toward marginal lands in developing countries. Enhancement of the nutritional value of staple crops through biofortification breeding might have a substantial impact on with their increased consumption worldwide. Increasing the availability of biofortified crops is a relatively straightforward approach to reach low-income people with limited access to healthy diets. Biofortification is a long-term, cost-effective, and sustainable approach to fight malnutrition in developing countries (Meenakshi et al., 2010). In the upcoming decades, the human population will increase in developing countries, and, with the altering climate conditions, food security will pose an increasing challenge (Das et al., 2013; Smith and Myers, 2018). Currently, the most common targeted micronutrients through biofortification breeding are iron, zinc, and carotenoids since these micronutrient deficiencies are common in children under the age of five and in pregnant and lactating women (Bouis and Welch, 2010). The World Health Organization (WHO) and Consultative Group on International Agricultural Research (CGIAR) aim to develop biofortified crops with enhanced nutrition (Bouis, 2000). To date, 36 biofortified varieties have been developed in maize, and these have reached 126,000 households in Zambia (Saltzman et al., 2017). The hybrid Pusa Vivek QPM nine Improved is the first biofortified maize variety in India with enhanced provitamin A. It was released in 2017 and is suitable for cultivation in nearly all states of India. Developing countries have included biofortification in their national agricultural nutrition strategies. For instance, India is the first country to prioritize biofortification and has set minimum standards for the release of pearl millet cultivars of 420 and 320 μg/100 g for iron and zinc, respectively. In this review, first major food sources and traits associated with carotenoids and folates have been discussed. In the next section, genetic variation and breeding strategies for enhancing the carotenoids and folates in major cereals (i.e., rice, wheat, maize, and pearl millet) have been summarized and discussed. In the final section we have discussed genomics integrated breeding and biofortification for carotenoids and folates as well as research gaps and future research directions.

IMPORTANT FOOD SOURCES OF CAROTENOIDS AND FOLATES

Folate is also referred to as vitamin B_9 and is involved in DNA and RNA synthesis. It is required to produce healthy red blood cells and is critical during periods of rapid growth, such as during pregnancy and fetal development. Carotenoids are essential for protecting eyes and bones and protecting against various types of cancer. Regular consumption of naturally available food sources can give a substantial quantity of folates, β-carotene, and macular carotenoids (lutein and zeaxanthin). However, the

availability and affordability of such food sources are not possible in rural, poor, and remote areas in developing countries. The top ten food sources that are rich in (per 100 g) folates and carotenoids from earlier published reports and international food databases are summarized (**Tables 1, 2** and **Figure 1**). **Table 1** summarizes the percentage of recommended dietary allowance (% RDA) of folate, which is calculated based on a 100 g serving of each crop type expressed for adults, pregnant women, and lactating women. β-carotene is the precursor of provitamin A, and it is predominantly accumulated in fruits and vegetables (Ashokkumar et al., 2018a). Ten major food crops with the highest concentration of β-carotene are presented in **Table 2**. Among them, kale or leafy cabbage, sweet potato, and carrot have the greatest concentration of β-carotene. Continuous availability and accessibility of these sources at affordable prices is challenging; improving the nutritional value of locally produced and available foods is an appropriate way to address this issue.

TRAITS ASSOCIATED WITH CAROTENOIDS AND FOLATES

The growing food markets pay close attention to grain nutritional quality due to the mounting health concerns among consumers. Yellow to orange pigmented grain types are positively correlated with carotenoid concentration in maize (da Silva Messias et al., 2014). Carotenoids are located in amyloplasts in maize. Lutein is the major carotenoid present in the grains of wheat (Ramachandran et al., 2010), pulses (Ashokkumar et al., 2014, 2015), oilseeds (McGraw et al., 2001), and spices (Ashokkumar et al., 2020). The seeds of wild-type maize chiefly accumulate lutein, followed by zeaxanthin, xanthophyll, and trace amounts of β-carotene (Janick-Buckner et al., 1999). Lutein and zeaxanthin are the major carotenoids in millets, with lutein being the predominant in white millet, while zeaxanthin is the main carotenoid in red millet (McGraw et al., 2001). Similarly, the

TABLE 1 | Folate-rich food sources available worldwide.

Sl. No.	Food source	Concentration (μg/100 g)	% RDA[§]			References
			Adult	Pregnant	Lactating women	
1.	Mung bean, raw	626.0	156.5	104.3	125.2	USDA–ARS (2012)
2.	Chickpea, raw	470.7	117.7	78.5	94.1	Jha et al. (2015)
3.	Common bean, raw	191.7	47.9	32.0	38.3	Jha et al. (2015)
4.	Lentil, green, raw	156.5	39.1	26.1	31.3	Jha et al. (2015)
5.	Soybean, green, raw	165.0	41.3	27.5	33.0	USDA–ARS (2012)
6.	Spinach, cooked	146.0	36.5	24.3	29.2	USDA–ARS (2012)
7.	Broccoli, cooked	108.0	27.0	18.0	21.6	USDA–ARS (2012)
8.	Bread wheat, raw	85.0	21.3	14.2	17.0	USDA–ARS (2012)
9.	Rice, pigmented, raw	51.0	12.8	8.5	10.2	Ashokkumar et al. (2018b)
10.	Corn, sweet, white, raw	46.0	11.5	7.7	9.2	USDA–ARS (2012)

[§]The percentage of recommended dietary allowance (RDA) of folate concentration was calculated based on the serving of 100 g of each species. The United States (U.S.), Food and Nutrition Board, RDAs required 400 μg/day, 600 μg/day, and 500 μg/day for adult, pregnant and lactating women, respectively.

TABLE 2 | Rich food sources of provitamin A (μg/100 g)[a].

Sl. No.	Food source	β-carotene (μg/100 g)	RAE (μg/day)[‡]	% RDA[§]			
				Children (1–3 years)	Children (4–8 years)	Men (>19 years)	Women (>19 years)
1.	Kale or leaf cabbage, raw	9226	768.8	256.3	192.2	85.4	109.8
2.	Sweet potato, raw	9180	765.0	255.0	191.3	85.0	109.3
3.	Carrot, raw	8836	736.3	245.4	184.1	81.8	105.2
4.	Squash, winter, butternut, raw	4226	352.2	117.4	88.0	39.1	50.3
5.	Collards, raw	3323	276.9	92.3	69.2	30.8	39.6
6.	Pepper, sweet, red, raw	2379	198.3	66.1	49.6	22.0	28.3
7.	Melon, cantaloupe, raw	1595	132.9	44.3	33.2	14.8	19.0
8.	Lettuce, romaine, raw	1272	106.0	35.3	26.5	11.8	15.1
9.	Apricots	664	55.3	18.4	13.8	6.1	7.9
10.	Peas, green, raw	432	36.0	12.0	9.0	4.0	5.1

[a]Source: USDA–NCC Carotenoid Database for the US Foods-1998 published by Holden et al., 1999. [§]Recommended dietary allowance (RDA) for vitamin A was calculated based on daily value (DV) of retinol activity equivalents (RAE) μg/day from 100 g serving of each species. The United States (U.S.), RDAs required RAE 300 μg/day and 400 μg/day for children aged 1–3 years and 4–8 years, respectively; 900 μg/day and 700 μg/day for adult men and women, respectively. [‡]RAE was calculated by 12 μg dietary β-carotene converted to 1 μg retinol (REA ratio 12:1).

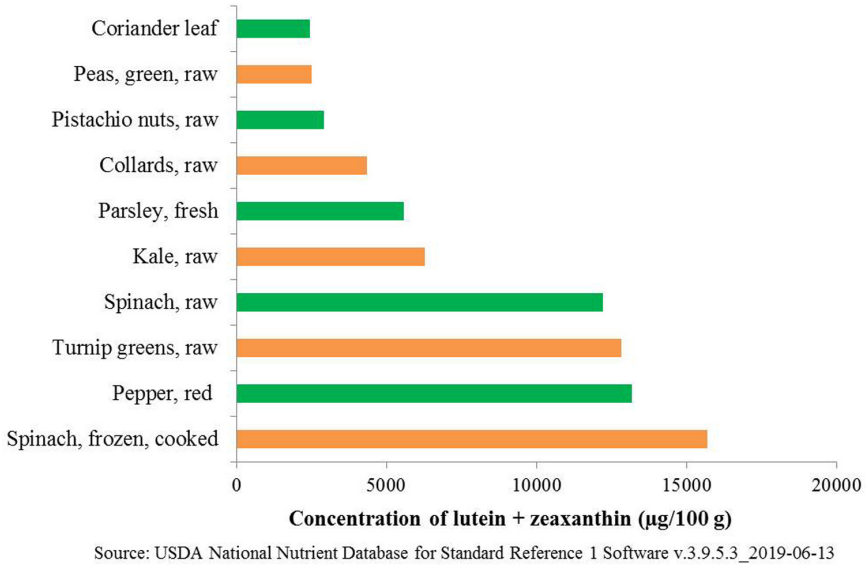


FIGURE 1 | Top 10 food sources providing macular carotenoids (lutein + zeaxanthin).

TABLE 3 | Genetic control of carotenoid concentration in major cereal grains.

Crop	Trait	Gene effects	References
Wheat	Yellow pigment concentration	Additive	Clarke et al. (2006)
Wheat	β-carotene	Digenic epistasis (additive × dominance)	Santra et al. (2005)
Maize	β-carotene	Incompletely dominant	Hauge and Trost (1928)
Maize	Carotenoids	Additive	Kandianis et al. (2013)
Maize	β-carotene	Additive	Jittham et al. (2017)
Maize	Provitamin A	Non-additive	Halilu et al. (2016)
Pearl millet	β-carotene	Non-additive	Khangura et al., 1980
Sorghum	β-carotene	Additive	Fernandez et al. (2008)

yellow kernel color of maize was positively correlated with non-provitamin carotenoids lutein and zeaxanthin (Muthusamy et al., 2015). A red-pigmented rice grain variety accumulated two-fold higher folate concentration than that found in white rice grains (Ashokkumar et al., 2018b). Abscisic acid (ABA) accumulation in grains is one of the important traits associated with carotenoid concentration (Maluf et al., 1997). Sometimes, reducing the antinutrient factors, such as phytic acid, may enhance the nutritional quality and bioavailability of cereals (Bohn et al., 2008; Tamanna et al., 2013). This approach has been effectively used to enhance the nutrition of maize grown for animal feed (Raboy, 1996). The highest accumulation of total carotenoids in wheat grain was reported at 12–15 days after anthesis and thereafter the level of accumulation declined (Graham and Rosser, 2000).

GENETICS AND GENETIC VARIATION OF CAROTENOIDS AND FOLATES IN CEREALS

Genetic analysis of carotenoids offers expedient directions to breeders initiating further breeding events. However, limited information is available on the genetic control of carotenoid concentration in staple cereal crops (**Table 3**). Yellow pigment concentration (YPC) in wheat and β-carotene, α-carotene, β-cryptoxanthin, and provitamin A in maize endosperms are largely controlled by additive genetic variance (Elouafi et al., 2001; Halilu et al., 2016). These complex traits may be linked to genotype-dependent and environmental factors. Grain yield and carotenoid concentration were predominantly controlled by non-additive gene actions in maize (Halilu et al., 2016). Furthermore, earlier investigations reported that carotenoids and its related compounds were controlled by both additive and non-additive gene action in maize endosperm (Chander et al., 2008). Babu et al. (2013) noticed that partial dominant and partial recessive gene action was in play in maize for the genes *LCYE-50TE* and *crtrB1-30TE*, respectively. The superiority of additive gene action and non-additive gene action suggested the application of recurrent selection and heterosis breeding followed genetic improvement of a particular trait in cereal crops.

Heritability estimates are mainly used for the determination of genotypic proportion of the trait, which favors the estimation of the effect of selection. If a particular trait has a higher heritability value, that trait might be modified by proper selection

methods. Conversely, lower heritability values indicated that those selection methods are not suitable for that particular trait. However, various researchers remarked that low to high heritability values were observed for carotenoids in maize. The heritability of YPC ranged from low (11%) to high (69%) in wheat (Elouafi et al., 2001; Clarke et al., 2006). Broad-sense heritability (H^2) was observed for lutein (61.49%), zeaxanthin (58.91%), and β-carotene (67.37%) in maize. Studies also noted narrow sense heritability (h^2) for lutein (19.00%) and zeaxanthin (18.09%) (Halilu et al., 2016). However, higher broad sense heritability was detected for lutein and zeaxanthin (Chander et al., 2008), and medium heritability values were observed for provitamin A (Wong et al., 2004). Genetic studies for gene action and heritability estimates are essential before initiating biofortification breeding programs for provitamin A and folates since heritability and gene action could be varied for different plant materials and environmental factors. Additionally, the investigation of gene action is imperative to design breeding programs.

In order to breed varieties with enhanced carotenoid and folate concentrations, information on the magnitude of genetic variation for carotenoids and folate in rice, wheat, maize, and pearl millet is needed. The variability for carotenoid and folate concentrations that have been recorded in the available genetic resources is summarized in **Table 4**. Genetic variation for β-carotene ranged from 96–850 μg/100 g in wheat, and 0.0–1780 μg/100 g in maize (Santra et al., 2005; Badakhshan et al., 2013; Muthusamy et al., 2014, 2015). In a study of 100 maize inbred lines, lutein and zeaxanthin concentrations ranged in the order of 20–1130 μg/100 g and 20–2000 μg/100 g, respectively. The highest lutein (1130 μg/100 g) and zeaxanthin (2000 μg/100 g) contents were recorded in two maize genotypes, namely, HP180-25 and CML161. According to Ortiz-Monasterio et al. (2007), 5–30% total carotenoids were provitamin A carotenoids while, β-carotene and β-cryptoxanthin were around 21 and 27% of the total concentrations of kernel carotenoids of yellow maize genotypes, respectively (Suwarno et al., 2014). These studies show that substantial genetic variability is present in the maize genetic resources for provitamin A and non-provitamin A concentrations of carotenoids, which could be used for the development of biofortified maize varieties/hybrids.

Pearl millet has limited concentrations of β-carotene, but a few accessions were identified with higher levels. For instance, genotype PT 6129 was high in β-carotene (241.7 μg/100 g), and such a line is useful to breed carotenoid rich varieties (Aarthy et al., 2011). Additionally, genetic variability is being explored in sorghum through the yellow endosperm lines which are available in the germplasm collections of the International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Hyderabad, India. The β-carotene concentrations of sorghum lines ranged from 56–113 μg/100 g, six lines, namely, IS 7684, IS 7776, IS 24703, IS 24868, IS 24883, and IS 26886, having an average of 85 μg (Reddy et al., 2005).

In terms of carotenoids in wheat seeds, lutein, zeaxanthin, and β-cryptoxanthin were predominant in the germ, while the endosperm had predominantly lutein, followed by β-cryptoxanthin and zeaxanthin (**Table 5**). For instance, the lutein concentration of wheat endosperm and germ varied significantly from 15.5 to 70.7 μg and 43.1 to 193.7 μg in 100 g, respectively (Adom et al., 2005; Masisi et al., 2015). The total carotenoids in wheat ranged from 170.1 to 227 μg/100 g in endosperm and 945 to 1029 μg/100 g in bran or germ (Ndolo and Beta, 2013). Interestingly, maize endosperm had substantial

TABLE 4 | Range of carotenoid and folate concentrations in the available genetic resources of major cereal grains.

Crop	Genotypes evaluated	Nutrient trait	Concentration (μg/100 g)	References
Rice	4 genotypes	Folate	11.0–51.0	Ashokkumar et al. (2018b)
Wheat	130 winter wheat genotypes	Folate	36.4–77.4	Piironen et al. (2008)
Wheat	20 spring wheat genotypes	Folate	32.3–74.1	Piironen et al. (2008)
Wheat	10 durum wheat	Folate	63.7–89.1	Piironen et al. (2008)
Wheat	82 wheat accessions	β-carotene	96.0–169.0	Badakhshan et al. (2013)
Wheat	5 genotypes	β-carotene	300.0–850.0	Santra et al. (2005)
Maize	12 inbred lines	Provitamin A	738.0–1359.0	Zunjare et al. (2018)
Maize	111 inbred lines	Total carotenoid	650.0–6730.0	Sivaranjani et al. (2013)
Maize	105 inbred lines	Lutein	20.0–1130.0.0	Muthusamy et al. (2015)
Maize	105 inbred lines	Zeaxanthin	20.0–2000.0	Muthusamy et al. (2015)
Maize	105 inbred lines	β-carotene	0.0–1500.0	Muthusamy et al. (2015)
Maize	105 inbred lines	β-cryptoxanthin	10.0–330.0	Muthusamy et al. (2015)
Maize	27 inbred lines	β-carotene	130.0–1780.0	Muthusamy et al. (2014)
Maize	64 inbred lines	Total carotenoids	558.0–390.0	Safawo et al. (2010)
Maize	64 inbred lines	β-carotene	12.0–474.0	Safawo et al. (2010)
Pearl millet	10 F_5 progeny lines	β-carotene	129.0–173.0	Jiji et al. (2017)
Pearl millet	10 F_5 progeny lines	Total carotenoids	329.0–810.0	Jiji et al. (2017)
Sorghum	11 genotypes	β-carotene	56.0–113.0	Reddy et al. (2005)
Sorghum	121 RILs	Lutein	8.0–63.0	Fernandez et al. (2008)
Sorghum	121 RILs	Zeaxanthin	6.1–102.0	Fernandez et al. (2008)

concentrations of zeaxanthin (1367.1 μg/100 g) and total carotenoids (1417.1–3135.2 μg/100 g).

Few studies have been conducted for the evaluation and identification of plant genetic resources for folate enhancement in cereal grains. This is likely due to the complexity, stability, and cost of folate concentration assays. Folate concentration was double in red-pigmented rice (Nootripathu) compared to non-pigmented rice genotypes (IR 20, N 22, and Pusa Basmati-1), and it ranged from 11 to 51 μg/100 g (Ashokkumar et al., 2018b). Piironen et al. (2008), assessed the total folate concentration in 160 genotypes of winter, spring, and durum wheat, and it ranged from 32.3 to 89.1 μg/100 g, with the greatest range evident in durum (63.7–89.1 μg/100 g). Their growing environments significantly influenced total folate concentration of winter wheat genotypes, more so than the genetic factors (Kariluoto et al., 2010). Variation for folates in maize, sorghum and pearl millet was not reported among the available genetic resources. Hence, further studies are needed to investigate the folate concentrations in grains of those major cereals. The rich sources of germplasm and their use for the genetic improvement and grain localization of carotenoids and in major cereal grains are described (**Table 6**).

BREEDING FOR INCREASED CAROTENOID AND FOLATE CONCENTRATION

The breeding strategies that are widely used to improve the carotenoid and folate concentrations in cereals are presented in **Figure 2**. Rice does not contain adequate amounts of

TABLE 5 | Grain localization of carotenoids and folates in major cereals.

Crop	Genotypes evaluated	Nutrient trait	Concentration (μg/100 g)			References
			Endosperm	Germ	Aleurone	
Wheat	5 genotypes	Lutein	36.9–70.7	164.1–191.7	–	Adom et al. (2005)
Wheat	5 genotypes	Zeaxanthin	1.6–2.7	19.4–26.2	–	Adom et al. (2005)
Wheat	5 genotypes	β-cryptoxanthin	3.5–4.4	8.91–10.0	–	Adom et al. (2005)
Wheat	1 genotype	Lutein	15.5	43.1	2.2	Masisi et al. (2015)
Wheat	1 genotype	Zeaxanthin	0.7	21.5	21.2	Masisi et al. (2015)
Wheat	4 genotypes	Total carotenoids	171.0.–227.1	845.1–987.1	–	Ndolo and Beta (2013)
Maize	1 genotype	Lutein	136.9	7.2	16.1	Masisi et al. (2015)
Maize	1 genotype	Zeaxanthin	1367.1	98.9	35.8	Masisi et al. (2015)
Maize	4 genotypes	Total carotenoids	1417.1–3135.2	33.3–53.6	–	Ndolo and Beta (2013)

TABLE 6 | Available genetic resources for carotenoids and folate improvement in major cereal grains.

Crop	Genotype	Nutrient trait	Concentration (μg/100 g)	RAE (μg/day)‡	% RDA§,†				References
					Children (1–3 years)	Children (4–8 years)	Men (>19 years)	Women (>19 years)	
Rice	Nootripathu	Folates	51.0		34.0	25.5	12.8	12.8	Ashokkumar et al. (2018b)
Maize	HP704-22	Provitamin A	1605.0	133.8	44.6	33.4	14.9	19.1	Zunjare et al. (2018)
Maize	HP704-23	Provitamin A	1528.0	127.3	42.4	31.8	14.1	18.2	Zunjare et al. (2018)
Maize	HP465-41	Provitamin A	1550.0	129.2	43.1	32.3	14.4	18.5	Muthusamy et al. (2015)
Maize	HP465-30	Provitamin A	1510.0	125.8	41.9	31.5	14.0	18.0	Muthusamy et al. (2015)
Maize	HP180-25	Lutein	1130.0	–		–		–	Muthusamy et al. (2015)
Maize	CML161	Zeaxanthin	2000.0	–		–		–	Muthusamy et al. (2015)
Maize	HPLET-03-36	Total carotenoid	6730.0	–		–		–	Sivaranjani et al. (2013)
Maize	HPLET-03-37	Total carotenoid	6320.0	–		–		–	Sivaranjani et al. (2013)
Maize	HPLET-03-35	Total carotenoid	5990.0	–		–		–	Sivaranjani et al. (2013)
Maize	BLSB-RIL17	Total carotenoid	5700.0	–		–		–	Sivaranjani et al. (2013)
Maize	BLSB-RIL43	Total carotenoid	5670.0	–		–		–	Sivaranjani et al. (2013)
Maize	HPLET-03-41	Total carotenoid	5610.0	–		–		–	Sivaranjani et al. (2013)
Maize	BLSB-RIL95	Total carotenoid	5090.0	–		–		–	Sivaranjani et al. (2013)
Maize	UMI176	β-carotene	580.0	48.3	16.1	12.1	5.4	6.9	Safawo et al. (2010)
Pearl millet	PT 6129	β-carotene	241.7	20.1	6.7	5.0	2.2	2.9	Aarthy et al. (2011)
Pearl millet	PT 6129	Total carotenoid	899.0	–		–		–	Jiji et al. (2017)
Sorghum	PI 585351	Total carotenoid	234.3	–		–		–	Shen et al. (2017)

‡*RAE was calculated by converting 12 μg of dietary β-carotene into 1 μg of retinol (REA ratio 12:1);* §*Recommended dietary allowance (RDA) for vitamin A was calculated based on daily value (DV) of retinol activity equivalents (RAE) μg/day from 100 g serving of each species. The United States (U.S.), RDAs required RAE 300 μg/day and 400 μg/day for children aged 1–3 years and 4–8 years, respectively; 900 μg/day and 700 μg/day for adults of both men and women, respectively.*

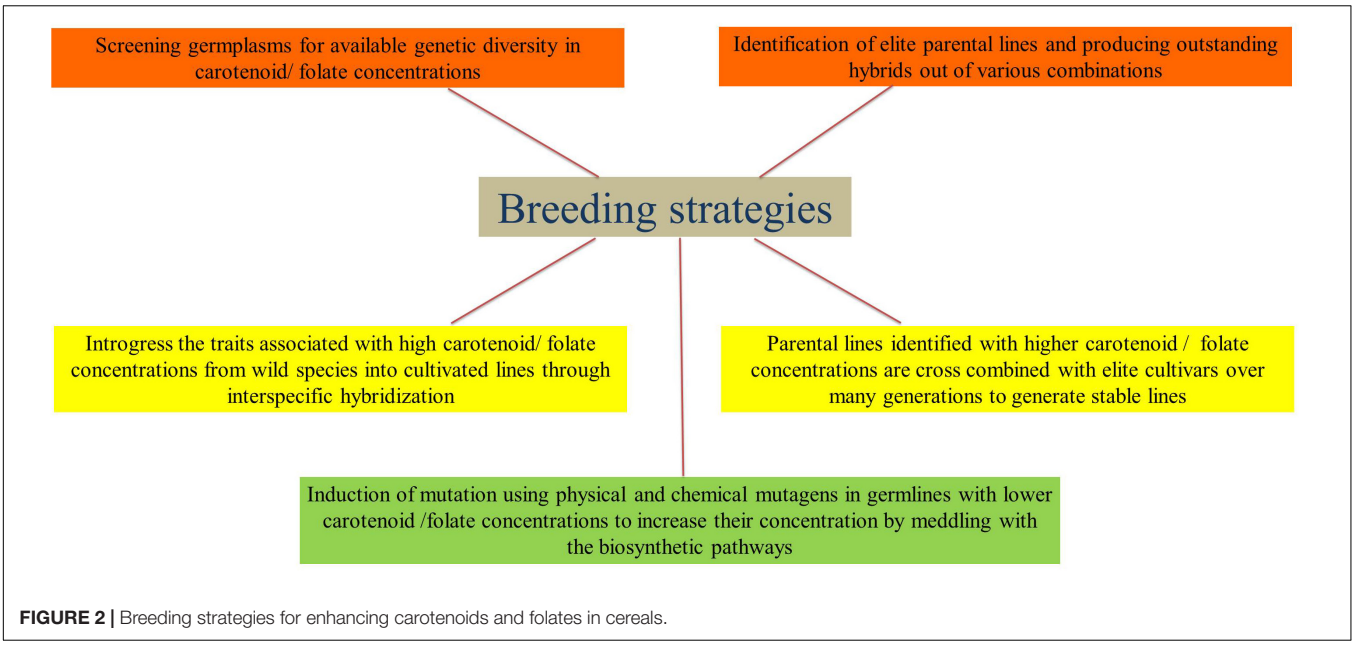


FIGURE 2 | Breeding strategies for enhancing carotenoids and folates in cereals.

carotenoids (i.e., β-carotene), which the human body could convert into vitamin A. Conventional breeding strategies have not been successful in increasing the β-carotene contents in rice endosperm. This is due to the fact that there is no genotype/cultivar that can synthesize carotenoid in the endosperm of the seed and the available contents are very low. Tan et al. (2004) showed that brown rice contains carotene and/or lutein, but the polishing process considerably reduces its nutritional value. In this respect, genetic engineering offers opportunities to improve the levels of provitamin A in rice grain. The control of the expression of ferritin through its control on the glutelin promoter has been successful in increasing nutritional levels in the whole and polished grains of rice. Similar principles have been used in the development of golden rice (Datta et al., 2007; Paine et al., 2005; Ye et al., 2000). Currently, no rice genotype has been enhanced for β-carotene content through traditional breeding strategies. It is obvious that there is huge potential in the exploitation of genetic variability of the carotenoid content in rice grains. However, the bioavailability of β-carotene should be studied in greater depth. In the case of folates, very few attempts were made to characterize the folate profile in rice by screening the germplasm. Blancquaert et al. (2015) screened 12 rice cultivars and found a two-fold difference (up to 70 μg/100 g) in the total grain folate content. The natural range of folate concentrations was determined in 78 rice varieties and both in milled (up to 78 mg/100 g) and whole grains (up to 111 μg/100 g), the contents exhibited an eight-fold difference (Yu and Tian, 2018). In all diverse accessions of rice germplasms around the world, an even more extensive screening for folate would bring out higher levels of variation in folate contents. This could be utilized in breeding programs for enhancing the folate contents in rice.

Natural genetic variability is very low for β-carotene contents in wheat grains. Lutein is the most common carotenoid in tetraploid wheat grains, whereas hexaploid wheat grains contain minimal levels of total carotenoids (Abdel-Aal el et al., 2007; Lachman et al., 2013). The durum wheat variety HI 8627 with high provitamin A was released by IARI, India, in 2005. The "Yellow pigment" is primarily caused by lutein, which is one of the significant factors in the enhancement of quality traits. Both lutein and anthocyanins are antioxidants in nature, which provokes a lot of interest in the research community. Black grained wheat cultivars and colored wheat cultivars are already being exploited in many breeding programs around the globe and they are rich in protein and selenium (Li et al., 2006). The purple wheat cultivar Indigo, which was released in Austria in 2006 (Eticha et al., 2011), the purple wheat cultivar PS Karkulka of Slovakia in 2014, and purple, blue, and black white lines of India in 2017 (Garg et al., 2016) are major sources of carotenoids in wheat breeding. Poutanen et al. (2008) evaluated the genetic variation for folates in the Health Grain wheat diversity screen with whole and milled wheat grains. Around 150 varieties of hexaploid, diploid, and tetraploid wheat showed two-fold variation in folate content (up to 77 μg/100 g) in whole grains (Piironen et al., 2008; Ward et al., 2008). Environmental effects cause variations in folate contents indicating low heritability and high G × E interactions in diverse varieties (Shewry et al., 2010). Induced chemical or physical mutagens could be utilized to identify mutants with greater folate contents.

Most of the breeding programs targeted to improve the provitamin A in maize aims at developing high yielding, provitamin A-enriched maize cultivars that fetch profit for the farmers and also promise customer preference and may ensure the effective reduction of vitamin A deficiency (Bouis and Welch, 2010). The simultaneous improvement of provitamin A carotenoids and grain yield is easily attainable. This is due to the weak correlation between provitamin A and agronomic performance. Other factors, like the relatively high heritability

of the trait, the mode of inheritance (additive genetic effects), and the genetic control of provitamin A, are also accountable (Suwarno et al., 2014; Menkir et al., 2018; Ortiz-Covarrubias et al., 2019). So far, the enhancement of provitamin A is mostly focused on the selection of β-carotene content. A target of 1500 μg/100 g of β carotene equivalents was set for breeders beyond which there occurs an increasingly marked effect on the human health (Hotz and McClafferty, 2007). Around 1,500 genotypes were screened for their carotenoid contents by various researchers, resulting in about 200–300 μg/100 g in their profiles (Ortiz-Monasterio et al., 2007). Among these germplasms, only a few lines of the temperate zones contained target level in their seeds (Menkir et al., 2008). In the meantime, the tropical and sub-tropical inbred lines possessed very low levels of provitamin A when compared with the breeding target in maize (Bouis et al., 2011). It demands the necessity and the initiation of searching for novel sources of favorable alleles to boost provitamin A concentration to new levels. Taleon et al. (2017) and Sowa et al. (2017) emphasized the application of breeding for provitamin A carotenoids that would increase β-cryptoxanthin rather than β-carotene, as β-carotene has lower stability, while β-cryptoxanthin ensures higher bioavailability and bioefficacy to β-carotene (Schmaelzle et al., 2014; Menkir et al., 2018). Breeding programs with this vision have already been initiated, resulting in inbreds that are being used in the improvement of new hybrids and synthetics. So far, most of the pearl millet breeding programs are targeted for improving grain iron and zinc concentration and yield related traits. Limited breeding efforts have been made thus far to explore the genetic variation of carotenoids and folates in pearl millet. Current circumstances demand carotenoid- and folate-rich donor lines for pearl millet breeding, and large numbers of germplasms must therefore be screened.

Typically, plant breeders use bi-parental populations for identification of QTL and development of varieties for the traits of interest. Many varieties developed of rice, wheat, maize, and pearl millet are based on single crosses between two parents. However, a higher number of parents and initial crosses will lead to a better dissection of complex traits. Thus, breeders recently introduced new experimental design namely multiparent populations, which provide significant benefits for genetic and QTL studies in plants. One of the most popular multiparent populations is the multiparent advanced generation intercross (MAGIC) population. The major goal of constructing MAGIC populations is to encourage intercrossing and shuffling of the genome into a single line (Huang et al., 2012; Holland, 2015). It is a diverse population with high recombination, thus providing excellent breeding materials to genetic and QTL mapping studies for complex traits such as carotenoids and folates. MAGIC populations have identified multiple loci and demonstrated the genetic complexity of the grain micronutrients (Fe and Zn), cooking quality, and agronomic traits (Holland, 2015; Descalsota et al., 2018; Ponce et al., 2018) in rice. Similarly, genetic properties of the MAGIC populations have also been detected in maize and wheat, and their benefits in detecting the complex traits have been confirmed by many researchers (Huang et al., 2012; Verbyla et al., 2014; Holland, 2015; Chen et al., 2016; Butrón et al., 2019). However, no study has been published that investigates carotenoids and folates using a MAGIC population design in cereals. ICRISAT has been developing a MAGIC population for various traits, including grain micronutrients (unpublished). Thus, it is a highly prioritized research area in which to work in the future for cereal-based national and international research organizations.

GENOMICS-ENABLED BREEDING APPROACHES FOR IMPROVING CAROTENOIDS AND FOLATES

Genomics research in cereals has substantially improved our knowledge of the QTLs/genes and biochemical pathways involved in carotenoids and folates in cereals (**Figure 3**). Second- and third-generation sequencing technologies have been game changers for genomics research and contributed to completion of the reference genome sequences for major cereal crops including rice (Yu et al., 2002), wheat (Brenchley et al., 2012), maize (Schnable et al., 2009), and pearl millet (Varshney et al., 2017). This genomic revolution has led to a pronounced increase in our knowledge of cereal genomics and our understanding of the structure and behavior of the cereal genomes. So far, an impressive number of genomic resources including detailed high-density genetic maps, cytogenetic stocks, contig-based physical maps, and deep coverage, and large-insert libraries have been developed in cereal crops (Muthamilarasan and Prasad, 2016). More interestingly, the genomic resources from a model or major cereal species (i.e., rice, maize, and wheat) also have potential that can be exploited for the development of minor cereals through comparative genomics approaches (Varshney et al., 2006). The transfer of genomic information and techniques from model or major to minor cereals provides detailed information about the genetic diversity of the crop and assists in the identification of the potentially beneficial variants in minor cereals. It also provides a greater chance for the identification of favorable alleles and the cloning and transfer of favorable alleles within the species.

Genomics offers tools to improve the contents of carotenoids and folates in cereals through advanced breeding techniques. Refining the breeding strategies through marker-assisted selection (MAS) is significantly improving the effectiveness of breeding for the enhancement of carotenoids and folates in cereals. The availability of the whole genome sequence data of major cereals enables the development of molecular markers. Among the different types of molecular markers, simple sequence repeats (SSRs) and single nucleotide polymorphisms (SNPs) markers are considered to be the markers of choice for a variety of applications, mainly in marker-assisted breeding (MAB). Besides, the genic or functional markers developed from the transcribed regions of the genome also act as ideal markers for MAB and as a major resource for assessing the functional variation in natural or breeding populations, in cereals. Expressed sequence tags (ESTs) or gene sequences have also been used to find SSRs/SNPs, and genic molecular markers have also been developed in cereals (McCouch et al., 2002). Fortunately, the information about SSR and SNP markers are available in the public domain for crop such as rice (McCouch et al., 2002),

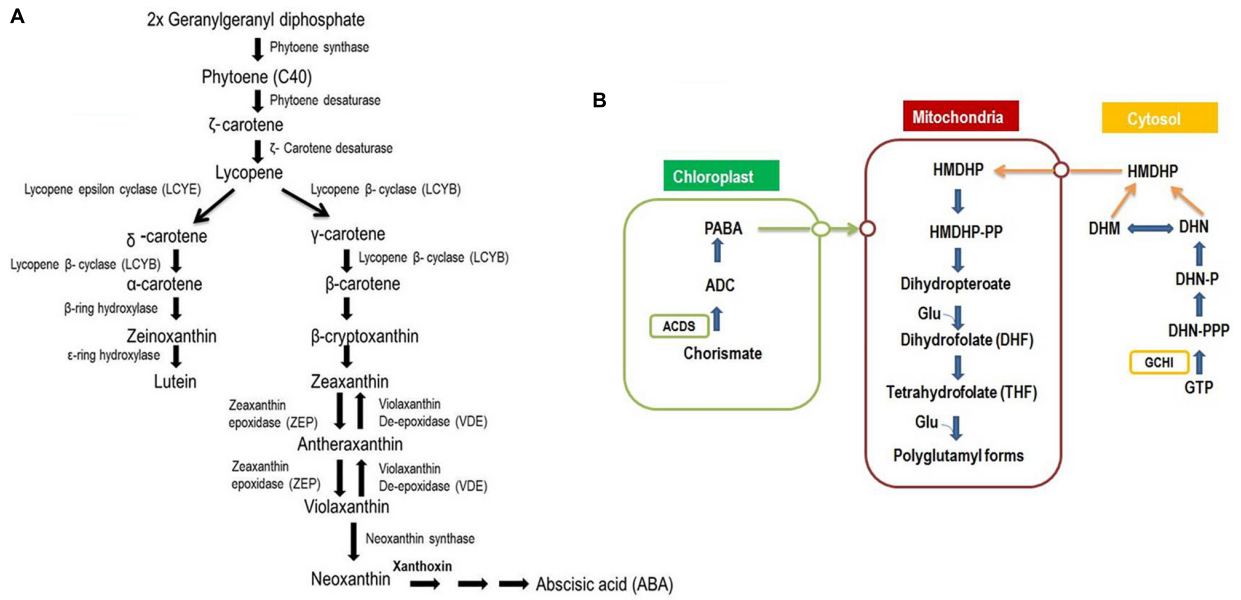


FIGURE 3 | Biosynthesis of carotenoids and Folates in plants. **(A)** Carotenoids biosynthesis and subsequent influential of phytohormones and provitamins. Footnotes: The first committed step in carotenoid biosynthesis is the condensation of two molecules of Geranylgeranyl diphosphate (GGPP) by phytoene synthase (*PSY*) to form phytoene (C40). The colorless phytoene is subsequently desaturated to give zeta-carotene and lycopene. Desaturation of phytoene occurs by two enzymes, phytoene desaturase (*PDS*) and zeta-carotene desaturase (*ZDS*), which are required to form lycopene. A major branch point occurs after lycopene synthesis when cyclization mediated by the enzymes lycopene-b-cyclase (*LCYB*) and lycopene-3-cyclase (*LCYE*) gives rise to α-carotene and β-carotene. α-carotene is acted upon by a β-ring hydroxylase to form zeinoxanthin, which is then hydroxylated by a ε-ring hydroxylase to produce lutein. β-carotene can be hydroxylated β-carotene hydroxylase (*CRTRB*) in a two-step reaction to zeaxanthin, with β-cryptoxanthin as an intermediate product. Zeaxanthin can be epoxidized to violaxanthin, and a set of light- and dark-controlled reactions, known as the xanthophyll cycle, rapidly optimize the concentration of violaxanthin and zeaxanthin in the cell through the action of zeaxanthin epoxidase (*ZEP*) and violaxanthin de-epoxidase (*VDE*), respectively, via antheraxanthin. Violaxanthin undergoes synthesis by the enzyme neoxanthin synthase to form neoxanthin and as precursor of the plant hormone abscisic acid. **(B)** Biosynthetic pathway of folates (Adapted by DellaPenna, 2007). Footnotes: The pteridine pathway leading to hydroxymethyldihydropterin (HMDHP) is shown in blue, the pathway leading to p-aminobenzoate is shown in green, and steps localized in the mitochondria are in black. Open circles indicate possible transporters. Red arrows indicate the two enzymes GTP-cyclohydrolase I (*GCHI*) and aminodeoxychorismate synthase (*ADCS*). DHN, dihydroneopterin; -P, monophosphate; -PP, pyrophosphate; -PPP, triphosphate; DHM, dihydromonapterin.

wheat (Jaiswal et al., 2017), maize (Sharopova et al., 2002) and pearl millet (Senthilvel et al., 2008). Thus, MAB in cereals has become standard procedure and many researchers to improve the levels of carotenoids and folates in cereals are pursuing these markers. Capitalizing on the genome-wide marker data, linkage-map-based QTL mapping, genome-wide association studies (GWAS), and genomic selection (GS) have become powerful tools to dissect the QTL and investigate trait-allele associations in cereals. To date, several QTLs/genes associated with carotenoids and folates in cereals were identified using linkage-map-based QTL mapping and GWAS. In particular, GWAS effectively pinpoints the genes that play a key role in the biosynthesis of carotenoids and their accumulation, and to find out the variation in the alleles at the concerned loci that are related to the biosynthesis of carotenoids in maize and wheat (Yan et al., 2010; Colasuonno et al., 2017). However, the nutritional traits like carotenoids and folates are quantitative and governed by minor QTLs that are responsible for the large phenotypic variation, including epistatic interactions. In this case, GS can capture both minor effects of QTL and epistatic interaction effects, so it could be a highly useful strategy in trait genetic gains of crop breeding programs. GS determines the genetic potential of an individual based upon the genomic estimated breeding values (GEBVs) instead of identifying the specific QTL (Robertsen et al., 2019). In the process of enhancing various complex traits, genomic selection has been used in cereals and other commercial crops. Still, the benefits of GS have not been utilized for the improvement of folates and carotenoids in cereals. This should provoke interest among the researchers working for upgrading the nutritional status of the major cereal crops in developing countries. So far, many QTLs/genes associated with carotenoids (provitamin A, lutein, and zeaxanthin) and folates have been identified in major cereals (i.e., rice, maize, and wheat). However, all of these QTLs/genes are not equally effective in the production of carotenoids and folates. Therefore, some of the important QTLs/genes that are identified so far are summarized and discussed here.

QTLs AND CANDIDATE GENES FOR CAROTENOIDS AND FOLATES

(a) Rice

QTL and genes have recently been identified for folate contents in rice through mapping studies, but in the case of carotenoids, no such information is available. In experiments with recombinant

inbred lines and backcrossed lines of milled rice, several major QTLs were identified to be associated with a higher level of folate. Dong et al. (2013) identified three QTLs, *qQTF-3-1, qQTF-32*, and *qQTF-3-3*, located on chromosome 3, which contributed 7.8, 11.1–15.8, and 25.3% of the variation in folate concentration. Three genes are associated with these QTLs, i.e., a rice homologue of plastidial folate transferase of *Arabidopsis*, a rice homologue of human folate hydrolase, and the serine hydroxymethyl transferase gene. When these newly identified QTLs associated with high folate are used in the synthesis of commercial varieties with high folate concentrations, there will be a larger wealth of knowledge about folate and its metabolism, regulation, and accumulation in grains (Yu and Tian, 2018).

(b) Wheat

Genetic analyses based on molecular markers have mapped major QTL for YPC on chromosome 7. Minor QTLs, associated with YPC, were detected on almost all chromosomes of the wheat genome. Some of these QTLs are stable, and they may be suitable for MAS in breeding programs. Two major QTLs were on chromosomes 3A and 7A, with 13 and 60% of the phenotypic variance, respectively (Parker et al., 1998). The QTL on chromosome 7A found to be closely related to an AFLP marker *Xwua26-7A.4* (Parker and Langridge, 2000), which was later transformed into an STS marker. Further, QTL that controls YP concentration of the kernels was detected on chromosome 7A with 12.9–37.6% of phenotypic variance in five different locations (Zhang et al., 2006). The YPC genes that encode phytoene synthase (*Psy*) have been mapped on the homologous groups of chromosomes 7 and 5 in wheat (Pozniak et al., 2007). There is an association between the loci of *Psy-B1*, which co-segregated with a QTL for endosperm color on 7B. Through *in silico* cloning, He et al. (2008) have categorized the association of the YPC in wheat grain across the full-length of the sequence of the genomic DNA sequence of a *Psy-A1*, which is linked to the SSR marker, *Xwmc809*, on the long arm of chromosome 7A with 20–28% of the phenotypic variance for the YP concentrations. Zhang et al. (2009) identified four QTLs namely, *QYpc-1A, QYpc-1B, QYpc-4A*, and *QYpc-7A*, for the YP concentration on chromosomes 1A, 1B, 4A, and 7A, which explained 1.5–33.9% of the phenotypic variance. Blanco et al. (2011) investigated the recombinant inbred line population arising from wheat cultivars Latino and Primadur, and they found that the QTLs linked with the concentration of YP and individual carotenoid compounds, namely, lutein, α-carotene zeaxanthin, β-cryptoxanthin, and β-carotene, were present on the same genomic regions of chromosomes 2A, 3B, 5A, and 7A. A single locus *called Lute*, controlling the lutein esterification on the short arm of chromosome 7D in wheat (Ahmad et al., 2015). The syntenic region of the rice genome contained a GDSL-like lipase gene. The sequences of wheat that are similar to this gene were mapped at the same locus of *Lute*. Folate variation in wheat accessions is very limited; almost no information is available on the folate QTLs and genes in wheat. Thus, researchers are trying

TABLE 7 | QTLs/Genes associated with carotenoids and folate concentrations in rice, wheat, maize, and pearl millet.

Crop	Nutrient	QTL/gene	Chromosome	References
Rice	Folate	*qQTF-3-1, qQTF-3-2* and *qQTF-3-3*	3	Dong et al. (2013)
Wheat	*Carotenoid*	*Lute*	7	Ahmad et al. (2015)
Wheat	Provitamin A	*Psy-B1*	7	Pozniak et al. (2007)
		Psy-A1	7	He et al. (2008)
		QYpc-1A, Qypc-1B, Qypc-4A, and *Qypc-7A*	1A, 1B, 4A, and 7A	Zhang et al. (2009)
		TaZds-A1	2A	Dong et al. (2012)
		AO1, *AO2*, and *AO3*	2, 5, and 7	Colasuonno et al. (2017)
Maize	Folate	*q5-FTHFa* and *q5-FTHFb*	5	Guo et al. (2019)
Maize	Provitamin A	*lcyE*	8	Harjes et al. (2008)
		crtRB1	10	Yan et al. (2010)
		crtRB3	2	Vallabhaneni and Wurtzel (2009); Zhou et al. (2012)
		Y1/PSY1	6	Buckner et al. (1996)
		PDS	1	Li et al. (1996).
		ZDS		
		(ZISO)/y9 locus	7	Chen et al. (2010); Matthews et al. (2003)
		qbc1-1, qbc5-1, qbc6-1, and *qbc10-1*	1, 5, 6, and 10	Jittham et al. (2017)
Maize	Lutein	qtl^{l}/*umc1447–umc1692–umc2373*	5	Chander et al. (2008)
		Qtl^{l} /*phi091–atf2*	7	Chander et al. (2008)
		qlut1-1 and *qlut6-1*	1 and 6	Jittham et al. (2017)
Maize	Zeaxanthin	qtl^{z}/*phi30870–umc1553*	1	Chander et al. (2008)
		qtl^{z}/*phi115–umc1735*	8	Chander et al. (2008)
		ZEP1		Vallabhaneni and Wurtzel (2009); Zhou et al. (2012)
		qzea6-1, *qzea8-1*, and *qzea10-1*	6, 8, and 10	Jittham et al. (2017)
		PS1/LCYB	5	Singh et al. (2003)
Sorghum	β-carotene	*Bc-1.1, Bc-2.1, Bc-2.2, Bc-2.3, Bc-10b.1*	1,2, 10b	Fernandez et al. (2008)

to identify novel QTLs and markers that are closely associated with folate for marker-assisted breeding in wheat.

(c) Maize

Several QTLs and genes related to carotenoids (provitamin A, lutein, and zeaxanthin) and folates have been reported in maize using different mapping approaches. *Yellow* 1 (*Y1*) gene encoding *PSY1* (phytoene synthase1) and is positioned on chromosome 6 in maize (Buckner et al., 1996). The gene *PSY1* was studied through association mapping in two different populations of maize. This gene has two alleles that are responsible for the differences in total carotenoids. Further, QTL mapping was carried out in one segregating population and lines that are polymorphic for genomic regions within *PSY1* were studied for expression analysis. Two functional sites that are concerned with the total carotenoid concentration of maize contributed 7 and 8% of the genetic variation (Fu et al., 2013).

Phytoene desaturase (*PDS*) and zeta-carotene desaturase are the enzymes that desaturate phytoene into lycopene. Lycopene is the first pigment that is produced in maize (Li et al., 1996). PDS is associated with viviparous 5 (*vp5*) that was mapped on chromosome 1. It was found that ζ−carotene isomerase (*Z−ISO*) was encoded by locus *y9* (Chen et al., 2010) and located on chromosome 7. Without the presence of Z−ISO, no provitamin A carotenoids could be synthesized in the endosperm (Matthews et al., 2003; Chen et al., 2010). Furthermore, 30 QTLs for carotenoid composition were also identified (Wong et al., 2004; Chander et al., 2008). A few of these are tightly linked to the biosynthetic pathway of *y1* or *y9* (Li et al., 2007) and are also associated with β-carotene, zeaxanthin, and lutein in maize. Lycopene epsilon cyclase (*lcyE*) on chromosome 8 (Harjes et al., 2008) and β-carotene hydroxylase enzyme (*crtRB1*) also known as *BCH2* and *HYD3*, on chromosome 10 (Yan et al., 2010) have the most significant effect on provitamin A concentrations in the maize grains. As per Harjes et al. (2008), the gene *LcyE,* causes different variation in of concentration of carotenoids because of its four alleles affect ß−branches of the biosynthesis pathway of carotenoids. Three polymorphisms were identified in the gene *crtRB1,* which controlled the variations in carotenoids (Yan et al., 2010). There was a 5.2-fold increase in the carotenoid concentrations in the haplotypes, which possessed the favorable alleles of *crtRB1-50 TE* and *crtRB1-30 TE*. The gene *crtRB1* was identified to have a much greater effect on the concentrations of provitamin A than that of *LcyE* (Babu et al., 2013).

The gene *crtRB3*, which encodes the α-carotene hydroxylase enzyme (also called *BCH1*), is a major role player in the metabolic pathway of carotenoids in maize (Vallabhaneni and Wurtzel, 2009; Zhou et al., 2012). On chromosome 2, there is a QTL locus cluster that is associated with carotenoids (**Table 7**). The gene, *crtRB3,* was mapped on this QTL locus cluster. Eighteen polymorphic sites within *crtRB3* that are closely linked to the QTL cluster were found through candidate−gene association analysis using 126 diverse inbred lines of yellow maize. Significant effects on the level of α−carotene were noticed (from 8.7 to 34.8%) among the two

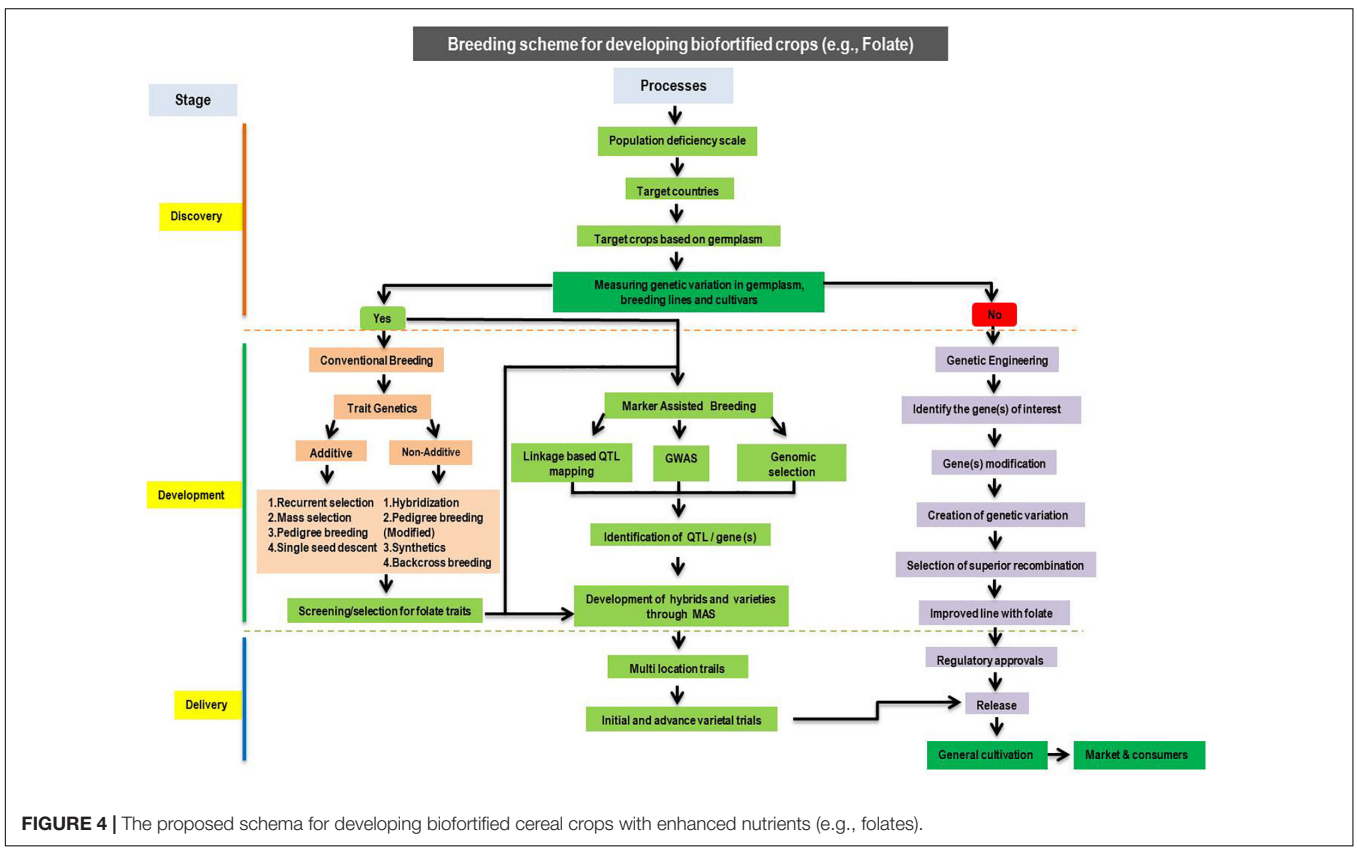


FIGURE 4 | The proposed schema for developing biofortified cereal crops with enhanced nutrients (e.g., folates).

SNPs, SNP1343 (in the 5' untranslated region) and SNP2172 (in the second intron), with 1.7- to 3.7-fold differences. Recently, four QTLs namely, *qbc1-1, qbc5-1, qbc6-1*, and *qbc10-1* were mapped by Jittham et al. (2017) on three chromosomes (1, 5, and 6) of maize for β-carotene with 5.04 to 17.03 % phenotypic variation.

Zeaxanthin and lutein are the other major carotenoids that are found in maize. But, only a few QTLs/genes that are associated with zeaxanthin and lutein have been identified so far. The *Ps1* locus located on chromosome 5 was encoding *LCYB*. This locus is considered essential for the accumulation of zeaxanthin in maize (Singh et al., 2003). *ZEP1* is one of the major genes in the metabolic pathway of carotenoids in maize (Zhou et al., 2012). It controls the gene zeaxanthin epoxidase (Vallabhaneni and Wurtzel, 2009). Three QTLs, namely, *qzea6-1*, *qzea8-1*, and *qzea10-1*, explaining 12.5%, 6.7%, and 19.4% of phenotypic variation in zeaxanthin, are found on chromosomes 6, 8, and 10 of maize, respectively (Jittham et al., 2017). Jittham et al. (2017) identified two lutein QTLs that are mapped on chromosomes 1 and 6. They are designated as *qlut1-1* and *qlut6-1*, explaining 9.1 and 28.9 of phenotypic variation.

Folates are quantitative or polygenic traits typically controlled by several small effect QTLs. However, two major effect QTLs namely, *q5−F−THFa* and *q5−F−THFb,* explaining 26.7 and 14.9% of the folate variation were identified in maize (Guo et al., 2019) on chromosome 5 by whole−exome sequencing and F_3 kernel−folate profiling. A unique correlation between the folate and the expression of the conserved genes of folate biosynthesis and metabolism was reported in the kernels of maize (Lian et al., 2015). Naqvi et al. (2009) and Liang et al. (2019) stated that the molecular understanding of the genetic networks of folates in grains is unclear even when successful increments have been made through transgenic experiments in maize.

(d) Pearl Millet

The Pearl Millet inbred Germplasm Association Panel (PMiGAP) contains around 1000 accessions, cultivars, and landraces of pearl millet that have been collected from three major pearl millet growing continents (Sehgal et al., 2015). *Iniari* germplasms that are found on the landraces of West Africa have already been collected and stored along with other Indian landraces and cultivars by ICRISAT (Yadav et al., 1999). The USDA National Plant Germplasm System Pearl Millet Collection, presented at the Plant Genetic Resources Conservation Unit in Griffin, GA, United States, preserves about 1297 unique germplasm lines from around 31 countries. However, no efforts have been made to explore the genetic variation of carotenoids and folates in these germplasm collections. Thus, marker-assisted breeding to improve the carotenoids and folates suffer due to the lack of donors. There also appears to be a significant gap in the literature, as no genes/QTLs/markers have been discovered with an association with carotenoid and folate concentrations in pearl millet. The headway toward the genetic enhancement of pearl millet is still painstaking due to the lack of PCR-based co-dominant markers. It is hoped that the recently released reference genome of pearl millet will facilitate the discovery of markers/QTLs/genes associated with carotenoid and folate concentration. Also, researchers should consider synteny studies with other cereal crops to improve pearl millet nutritional breeding programs.

BIOFORTIFICATION OF CAROTENOIDS AND FOLATES IN CEREAL GRAINS

Biofortification is the process of increasing the natural content of bioavailable nutrients in plants. It is a successful and cost-effective method that associates nutritious agriculture with human health, can be efficient and more maintainable than the delivery of food supplements. Major tools in biofortification include conventional breeding, modern biotechnology, and agronomic practices (**Figure 4**). As mentioned above, carotenoids and folates are essentials for the human diet. Thus, biofortification of major cereal crops with carotenoids and folates may assist in easing micronutrient deficiencies in humans. Existing evidence recommends that genetic biofortification by breeding and modern biotechnology could be appropriate for increasing folates and pro-vitamin A carotenoids, and an agronomic strategy could be effective for Zn. Conventional breeding-based biofortification is the most successful approach to develop micronutrients rich crops, and several important food crops have been targeted for fortification by conventional breeding. So far, many more studies have been conducted to improve the provitamin A concentration and a few targeted at folate. Biofortification in maize has been attempted in many different ways. For instance, improvement of single or group of micronutrient (s) (single biofortification) and diverse micronutrients (double biofortification), including (i) the incorporation of favorable alleles of *crtRB1* and *lcyE* into popular elite genotype by MAB and transgenic approaches to increase the amount of provitamin A concentration and (ii) the development of genotypes with *crtRB1* and *lcyE* and *o2* alleles to increase the essential amino acids and provitamin A concentration by MAB (Hossain et al., 2019). In the recent decades, CIMMYT, Mexico, and IITA, Nigeria, developed and released many provitamin A varieties and hybrids (i.e., GV662A, GV664A, and GV665A, Ife maize hyb-3, and Ife maize hyb-4, Sammaz 38, Sammaz 39, and CSIR-CRI Honampa) in African countries (Dhliwayo et al., 2014; Simpungwe et al., 2017; Andersson et al., 2017). IARI released four provitamin A hybrids viz., HM4, HM8, and Vivek Hybrid-27 [which possessed provitamin A as high as 2170 μg/100 g (in freshly harvested grains) with a 8.5-fold maximum change] in India. The hybrid, "Pusa Vivek QPM 9 Improved," which was developed through MAB, contains higher provitamin A (815 μg/100 g) even after storing for 2 months with higher levels of tryptophan, 0.74% and lysine, 2.67% (Muthusamy et al., 2014; Yadava et al., 2017). This hybrid was developed by the introgression of the *crtRB1* allele into a *o2*-based hybrid. In a similar manner, four popular QPM hybrids namely HQPM1, HQPM4, HQPM5, and HQPM7 were developed by

pyramiding *crtRB1* and *lcyE* to improve the concentration of provitamin A (Hossain et al., 2019). Despite the success stories, conventional or marker-assisted breeding suffers due to the lack of genetic variation in micronutrient traits within the species or closely related species. In this context, transgenic technologies are an alternative to conventional breeding and useful to improve the genotypes by creating variations in targeted metabolic pathways. The concentration of provitamin A in rice was improved through transgenic methods. Over-expression of *PSY, CrtI* and β-*lcy* from daffodil, *Erwinia uredovora* and maize facilitated an increase in provitamin A concentration in rice lines (Ye et al., 2000; Beyer et al., 2002; Paine et al., 2005). In particular, *PSY* from maize increased provitamin A concentration up to 3700μg/100 g (Paine et al., 2005). Similarly, the contents of β-carotene increased to 1000 μg/100 g in the Hi-II maize line through the over-expression of *crtB* and *crtI* genes from *Erwinia herbicola* (Aluru et al., 2008). Likewise, five genes, namely, *psy1*, *crtI, lycb*, *bch*, and *crtW*, were used to develop transgenic maize genotypes that contained 6000 μg/100 g of β-carotene (Zhu et al., 2008; Naqvi et al., 2009). The over expression of *psy1* from maize and *crtI* or *CrtB* from the bacteria enhanced provitamin A to 496 μg/100 g and 321 μg/100 g of seed dry weight in wheat (Cong et al., 2009; Wang et al., 2014). Likewise, five genes, namely, *psy1*, *c rtI, lycb*, *bch*, and *crtW*, were used to develop transgenic maize genotypes that contained 6000 μg/100 g of β-carotene (Zhu et al., 2008; Naqvi et al., 2009). Despite the success of transgenic technologies, the main drawback to biofortified transgenic crops is their public acceptance and extensive regulatory processes required before they get clearance for cultivation and consumption by humans.

CONCLUSION AND FUTURE PROSPECTS

The growing world population requires many key nutrients and vitamins that can be delivered through staple foods. Advancing genomic tools can play an important role in accelerating genetic enhancement of these vitamins and minerals through biofortification in major cereal grains. Bioavailable vitamins or nutrients bred into varieties can be made available to resource-poor people generation after generation by their cultivation and regular consumption. The surplus production brings better livelihoods through marketing to other regions. Crop breeding requires substantial genetic variability and diagnostic markers to handle traits in segregating early generations. Nutrient-dense germplasm resources are essential to the breeding of adequate carotenoids and folates for fulfilling daily dietary requirements. National and international organizations have made excellent research progress in this direction to incorporate carotenoids into cereal crops. High throughput phenotyping tools (XRF, HPLC, and LC-MS/MS) are being developed and will be made accessible to partners at various organizations. These methods are cost effective for analyzing large sets of germplasm. The diagnostic markers play a key role in discarding low vitamin/nutrient materials. Integrating MAB creates the opportunity to introduce/track the QTL that are associated with nutritional quality into popular varieties. A survey of wild and cultivated accessions demonstrated noticeable variations in the carotenoid and folate concentration and the possibility to identify novel sources for alleles to be used to broaden the present gene pool. So far, substantial genetic variation has been exhibited only in the genetic resources of maize for provitamin A. Other major cereals, like rice, wheat, and pearl millet commercial or elite lines, lack sufficient concentration of provitamin A to achieve global target levels. Almost no folate research has been done in major cereal crops. Biofortification based breeding has been demonstrated as a successful of enhance the micronutrients in cereals. However, new breeding designs, such as MAGIC populations and GS, also need to be explored on parallel to maximize the genetic understanding and identification of QTLs and genes for complex traits such as carotenoids and folates. Hence, greater prospects await with the use of these technologies in nutrition breeding. On the other hand, where inadequate genetic variability exists within the cultivated germplasms and primary gene pools, then the transgenic technology may be an option for enhancing carotenoids and folates in cereals but has limited scope for acceptance in most of the developing countries. Genetic gain for yield alone may not be appropriate to feed the growing population, but concurrently achieving nutrition traits genetic gains is a sustainable approach. Government programs are required to create public awareness for the adoption of biofortified varieties by farmers through increased consumer acceptance. Moreover, research coordination is required between agriculture and nutritional experts for strengthening the target level of carotenoids and folates, their retention after cooking, storage, processing, and consumption of prospective concentrations in the target population. Therefore, with the available genetic resources and genomic tools, breeding investment should be made and optimized for increasing vitamins and nutrients in staple food crops besides increasing sustainable yields.

AUTHOR CONTRIBUTIONS

MG and KA conceptualized the manuscript. KA, MG, AK, and VS wrote the manuscript. MG and TW edited and updated the manuscript.

FUNDING

This work was partially supported by HarvestPlus program of CGIAR to the corresponding author (MG).

ACKNOWLEDGMENTS

The authors thank editors and reviewers for their detailed suggestions for improving the manuscript.

REFERENCES

Aarthy, T., Sumathi, P., Senthil, N., Raveendran, M., Vellaikumar, S., and Veerabadhiran, P. (2011). "Screening of high carotene and its segregating generation through HPLC for developing mapping population for enhanced carotene," in *Proceedings of the World Congress on Biotechnology*, Hyderabad, 381.

Abdel-Aal el, S. M., Young, J. C., Rabalski, I., Hucl, P., and Fregeau-Reid, J. (2007). Identification and quantification of seed carotenoids in selected wheat species. *J. Agric. Food Chem.* 55, 787–794. doi: 10.1021/jf062764p

Adom, K. K., Sorrells, M. E., and Liu, R. H. (2005). Phytochemicals and antioxidant activity of milled fractions of different wheat varieties. *J. Agric. Food Chem.* 53, 2297–2306. doi: 10.1021/jf048456d

Ahmad, F. T., Mather, D. E., Law, H., Li, M., Yousif, S., Chalmers, K. J., et al. (2015). Genetic control of lutein esterification in wheat (*Triticum aestivum* L.) grain. *J. Cereal Sci.* 64, 109–115.

Aluru, M., Xu, Y., Guo, R., Wang, Z., Li, S., White, W., et al. (2008). Generation of transgenic maize with enhanced provitamin A content. *J. Exp. Bot.* 59, 3551–3562. doi: 10.1093/jxb/ern212

Andersson, M. S., Saltzman, A., Virk, P. S., and Pfeiffer, W. (2017). Progress update: crop development of biofortified staple food crops under Harvestplus. *Afr. J. Food Agric. Nutr. Dev.* 17, 11905–11935. doi: 10.18697/ajfand.78.HarvestPlus05

Ashokkumar, K., Diapari, M., Jha, A. B., Tar'an, B., Arganosa, G., and Warkentin, T. D. (2015). Genetic diversity of nutritionally important carotenoids in 94 pea and 121 chickpea accessions. *J. Food Compos. Anal.* 43, 49–60. doi: 10.1016/j.jfca.2015.04.014

Ashokkumar, K., Pandiyan, A., Murugan, M., Dhanya, M. K., Sathyan, T., Sivakumar, P., et al. (2018a). Simple and rapid extraction method for determination of carotenoids in the edible parts of *Vitis vinifera, Vaccinum* sect. *cyanococcus, Ipomoea batatas* and *Capsicum annum*. *Adv. Res.* 17, 1–8. doi: 10.9734/air/2018/45132

Ashokkumar, K., Pandian, A., Murugan, M., Dhanya, M. K., Sathyan, T., Sivakumar, P., et al. (2020). Profiling bioactive flavonoids and carotenoids in select south Indian spices and nuts. *Nat. Prod. Res.* 34, 1306–1310. doi: 10.1080/14786419.2018.1557179

Ashokkumar, K., Sivakumar, P., and Saradha Devi, M. (2018b). Identification and determination of naturally occurring folates in grains of rice (*Oryza sativa* L.) by UPLC-MS/MS analysis. *Nat. Prod. Res.* 32, 1733–1737. doi: 10.1080/14786419.2017.1392957

Ashokkumar, K., Tar'an, B., Diapari, M., Arganosa, G., and Warkentin, T. D. (2014). Effect of cultivar and environment on carotenoid profile of pea and chickpea. *Crop Sci.* 54, 2225–2235. doi: 10.2135/cropsci2013.12.0827

Babu, R., Rojas, N. P., Gao, S., Yan, J., and Pixley, K. (2013). Validation of the effects of molecular marker polymorphisms in LcyE and CrtRB1 on provitamin A concentrations for 26 tropical maize populations. *Theor. Appl. Genet.* 126, 389–399. doi: 10.1007/s00122-012-1987-3

Badakhshan, H., Morad, N., Mohammadzadeh, H., and Zakeri, M. R. (2013). Genetic variability analysis of grains Fe, Zn and β-carotene concentration of prevalent wheat varieties in Iran. *Int. J. Agri. Crop Sci.* 6, 57–62.

Beyer, P., Al-Babili, S., Ye, X., Lucca, P., Schaub, P., Welsch, R., et al. (2002). Golden rice Introducing the β-Carotene biosynthesis pathway into rice endosperm by genetic engineering to defeat vitamin a deficiency. *J. Nutr.* 132, 506S–510S. doi: 10.1093/jn/132.3.506S

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.

Blancquaert, D., Van Daele, J., Strobbe, S., Kiekens, F., Storozhenko, S., De Steur, H., et al. (2015). Improving folate (vitamin B9) stability in biofortified rice through metabolic engineering. *Nat. Biotechnol.* 33, 1076–1078. doi: 10.1038/nbt.3358

Bohn, L., Meyer, A. S., and Rasmussen, S. (2008). Phytate: impact on environment and human nutrition. A challenge for molecular breeding. *J. Zhejiang Univ. Sci. B* 9, 165–191. doi: 10.1631/jzus.B0710640

Bouis, H. E. (2000). Enrichment of food staples through plant breeding: a new strategy for fighting micronutrient malnutrition. *Nutrition* 16, 701–704. doi: 10.1016/s0899-9007(00)00266-5

Bouis, H. E. (2002). Plant breeding: a new tool for fighting micronutrient malnutrition. *J. Nutrition* 132, 491S–494S. doi: 10.1093/jn/132.3.491S

Bouis, H. E., Hotz, C. H., McClafferty, B., Meenakshi, J. V., and Pfeiffer, W. (2011). Biofortification: a new tool to reduce micronutrient malnutrition. *Food Nutr. Bull.* 32, S31–S40. doi: 10.1177/15648265110321S105

Bouis, H. E., and Welch, R. M. (2010). Biofortification– a sustainable agricultural strategy for reducing micronutrient malnutrition in the global South. *Crop Sci.* 50, S20–S32.

Brenchley, R., Spannagl, M., Pfeifer, M., Barker, G. L., D'Amore, R., Allen, A. M., et al. (2012). Analysis of the bread wheat genome using whole−genome shotgun sequencing. *Nature* 491, 705–710. doi: 10.1038/nature11650

Buckner, B., Miguel, P. S., Janickbuckner, D., and Bennetzen, J. L. (1996). The Y1 gene of maize codes for phytoene synthase. *Genetics* 143, 479–488.

Butrón, A., Santiago, R., Cao, A., Samayoa, L. F., and Malvar, R. A. (2019). QTLs for resistance to *Fusarium* ear rot in a multi-parent advanced generation inter-cross (MAGIC) maize population. *Plant Dis.* 103, 897–904.

Chander, S., Guo, Y. Q., Yang, X. H., Zhang, J., Lu, X. Q., Yan, J. B., et al. (2008). Using molecular markers to identify two major loci controlling carotenoid concentrations in maize grain. *Theor. Appl. Genet.* 116, 223–233. doi: 10.1007/s00122-007-0661-7

Chen, J., Zhang, L., Liu, S., Li, Z., Huang, R., Li, Y., et al. (2016). The genetic basis of natural variation in kernel size and related traits using a four-way cross population in maize. *PLoS One* 11:e0153428. doi: 10.1371/journal.pone.0153428

Chen, Y., Li, F., and Wurtzel, E. T. (2010). Isolation and characterization of the Z-ISO gene encoding a missingcomponent of carotenoid biosynthesis in plants. *Plant Physiol.* 153, 66–79. doi: 10.1104/pp.110.153916

Clarke, F. R., Clarke, J. M., McCaig, T. N., Knox, R. E., and DePauw, R. M. (2006). Inheritance of yellow pigment concentration in four durum wheat crosses. *Can. J. Plant Sci.* 86, 133–141.

Colasuonno, P., Marcotuli, I., Lozito, M. L., Simeone, R., Blanco, A., and Gadaleta, A. (2017). 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., Chen, L., Liu, H., Yang, G., and He, G. (2009). Expression of phytoene synthase1 and carotene desaturase crtI genes result in an increase in the total carotenoids content in transgenic elite wheat (*Triticum aestivum* L.). *J. Agric. Food Chem.* 57, 8652–8660. doi: 10.1021/jf9012218

da Silva Messias, R., Galli, V., Dos Anjos, E., Silva, S. D., and Rombaldi, C. V. (2014). Carotenoid biosynthetic and catabolic pathways: gene expression and carotenoid content in grains of maize landraces. *Nutrients* 6, 546–563. doi: 10.3390/nu6020546

Das, J. K., Kumar, R., Salam, R. A., and Bhutta, Z. A. (2013). Systematic review of zinc fortification trials. *Ann. Nutr. Metab.* 62, 44–56. doi: 10.1159/000348262

Datta, S. K., Datta, K., Parkhi, V., Rai, M., Baisakh, N., Sahoo, G., et al. (2007). Golden rice: introgression, breeding, and field evaluation. *Euphytica* 154, 271–278.

Davey, M. W., Mellidou, I., and Keulermans, W. (2009). Considerations to prevent the breakdown and loss of fruit carotenoids during extraction and analysis in Musa. *J. Chromatogr. A* 1216, 5759–5762. doi: 10.1016/j.chroma.2009.06.022

DellaPenna, D. (2007). Biofortification of plant-based food: enhancing folate levels by metabolic engineering. *Proc. Natl. Acad. Sci. U.S.A.* 104, 3675–3676. doi: 10.1073/pnas.0700640104

Descalsota, G. I. L., Swamy, B. P. M., Zaw, H., Inabangan-Asilo, M. A., Amparado, A., Mauleon, R., et al. (2018). Genome-wide association mapping in a rice MAGIC plus population detects QTLs and genes useful for biofortification. *Front. Plant Sci.* 9:1347. doi: 10.3389/fpls.2018.01347

Dhliwayo, T., Palacios-Rojas, N., Crossa, J., and Pixley, K. (2014). Effects of S1 recurrent selection for provitamin A carotenoid content for three open-pollinated maize cultivars. *Crop Sci.* 54, 2449–2460.

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.

Dong, W., Cheng, Z., Xu, J., Zheng, T., Wang, X., Zhang, H., et al. (2013). Identification of QTLs underlying folate concentration in milled rice. *J. Integr. Agr.* 9, 1827–1834.

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

Eticha, F., Grausgruber, H., Siebenhandl-ehn, S., and Berghofer, E. (2011). Some agronomic and chemical traits of blue aleurone and purple pericarp wheat (*Triticum* L.). *J. Agric. Sci. Technol.* 1, 48–58.

Fernandez, M. G. S., Hamblin, M. T., Li, L., Rooney, W. L., Tuinstra, M. R., and Kresovich, S. (2008). Quantitative trait loci analysis of endosperm color and carotenoid concentration in sorghum grain. *Crop Sci.* 48, 1732–1743.

Flores, A. L., Vellozzi, C., Valencia, D., and Sniezek, J. (2014). Global burden of neural tube defects, risk factors, and prevention. *Indian J. Community Health* 26, 3–5.

Food and Agriculture Organization [FAO], International Fund for Agricultural Development, UNICEF, World Food Programme, and WHO (2017). *The State of Food Security and Nutrition in the World 2017: Building Resilience for Peace and Food Security*. Available online at: http://www.fao.org/3/a-i7695e.pdf (accessed August 5, 2019).

Forges, T., Monnier-Barbarino, P., Alberto, J. M., Guéant-Rodriguez, R. M., Daval, J. L., and Guéant, J. L. (2007). Impact of folate and homocysteine metabolism on human reproductive health. *Hum. Reprod. Update* 13, 225–238. doi: 10.1093/humupd/dml063

Fraser, P. D., and Bramley, P. M. (2004). The biosynthesis and nutritional uses of carotenoids. *Prog. Lipid Res.* 43, 228–265. doi: 10.1016/j.plipres.2003.10.002

Fu, Z. Y., Chai, Y. C., Zhou, Y., Yang, X. H., and Warburton, M. L. (2013). Natural variation in the sequence of PSY1 and frequency of favorable polymorphisms among tropical and temperate maize germplasm. *Theor. Appl. Genet.* 126, 923–935. doi: 10.1007/s00122-012-2026-0

Garg, M., Chawla, M., Chunduri, V., Kumar, R., Sharma, S., Sharma, N. K., et al. (2016). Transfer of grain colors to elite wheat cultivars and their characterization. *J. Cereal Sci.* 71, 138–144. doi: 10.1016/j.jcs.2016.08.004

Graham, R. D., and Rosser, J. M. (2000). Carotenoids in staple foods: their potential to improve human nutrition. *Food Nutr. Bull.* 21, 404–409.

Guo, W., Lian, T., Wang, B., Guan, J., Yuan, D., Wang, H., et al. (2019). Genetic mapping of folate QTLs using a segregated population in maize (*Zea mays* L.). *J. Integr. Plant Biol.* 61, 675–690. doi: 10.1111/jipb.12811

Halilu, A. D., Ado, S. G., Aba, A. D., and Usman, I. S. (2016). Genetics of carotenoids for provitamin A biofortification in tropical-adapted maize. *Crop J.* 4, 313–322. doi: 10.1016/j.cj.2016.05.002

Harjes, C. E., Rocheford, T. R., Bai, L., Brutnell, T. P., Kandianis, C. B., Sowinski, S. G., et al. (2008). Natural genetic variation in lycopene epsilon cyclase tapped for maize biofortification. *Science* 319, 330–333. doi: 10.1126/science.1150255

Hauge, S. M., and Trost, J. F. (1928). An inheritance study of the distribution of vitamin A in maize. *J. Biol. Chem.* 80, 107–115.

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

Hefni, M., Ohrvik, V., Tabekha, M. M., and Witthoft, C. (2010). Folate concentration in foods commonly consumed in Egypt. *Food Chem.* 121, 540–545.

Holden, J. M., Eldridge, A. L., Beecher, G. R., Buzzard, I. M., Bhagwat, S., Davis, C. S., et al. (1999). Carotenoid concentration of U.S. foods: an update of the database. *J. Food Compos. Anal.* 12, 169–196.

Holland, J. B. (2015). MAGIC maize: a new resource for plant genetics. *Genome Biol.* 16:163. doi: 10.1186/s13059-015-0713-2

Hossain, F., Muthusamy, V., Zunjare, R. U., and Gupta, H. S. (2019). "Biofortification of maize for protein quality and provitamin-a content," in *Nutritional Quality Improvement in Plants. Concepts and Strategies in Plant Sciences*, eds P. Jaiwal, A. Chhillar, D. Chaudhary, and R. Jaiwal (Cham: Springer), 115–136.

Hotz, C., and McClafferty, B. (2007). From harvest to health: challenges for developing biofortified staple foods and determining their impact on micronutrient status. *Food Nutr. Bull.* 28, S271–S279. doi: 10.1177/15648265070282s206

Huang, B. E., George, A. W., Forrest, K. L., Kilian, A., Hayden, M. J., Morell, M. K., et al. (2012). A multiparent advanced generation inter-cross population for genetic analysis in wheat. *Plant Biotechnol. J.* 10, 826–839. doi: 10.1111/j.1467-7652.2012.00702.x

Institute of Medicine Food and Nutrition Board (1998). *Dietary Reference Intakes, Dietary Reference Intakes for Thiamin, Riboflavin, Niacin, Vitamin B6, Folate, Vitamin B_{12}, Pantothenic Acid, Biotin, and Choline*. Washington, DC: National Academy Press.

Jaiswal, S., Sheoran, S., Arora, V., Angadi, U. B., Iquebal, M. A., Raghav, N., et al. (2017). Putative microsatellite DNA marker-based wheat genomic resource for varietal improvement and management. *Front. Plant Sci.* 8:2009. doi: 10.3389/fpls.2017.02009

Janick-Buckner, D., Hammock, J. D., Johnson, J. M., Osborn, J. M., and Buckner, B. (1999). Biochemical and ultrastructural analysis of the y10 mutant of maize. *J. Hered.* 90, 507–513. doi: 10.1093/jhered/90.5.507

Jha, A. B., Ashokkumar, K., Diapari, M., Ambrose, S. J., Zhang, H., Tar'an, B., et al. (2015). Genetic diversity of folates profiles in seeds of common bean, lentil, chickpea, and pea. *J. Food Comp. Anal.* 42, 134–140. doi: 10.1016/j.jfca.2015.03.006

Jiji, J., Sumathi, P., and Senthil, N. (2017). Identification of high beta carotene lines in f5 generation of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Int. J. Curr. Microbiol. App. Sci.* 6, 120–128. doi: 10.20546/ijcmas.2017.605.015

Jittham, O., Fu, X., Xu, J., Chander, S., Li, J., and Yang, X. (2017). Genetic dissection of carotenoids in maize kernels using high-density single nucleotide polymorphism markers in a recombinant inbred line population. *Crop J.* 5, 63–72. doi: 10.1186/s12870-015-0675-2

Kandianis, C. B., Stevens, R., Liu, W., Palacios, N., Montgomery, K., Pixley, K., et al. (2013). Genetic architecture controlling variation in grain carotenoid composition and concentrations in two maize populations. *Theor. Appl. Genet.* 126, 2879–2895. doi: 10.1007/s00122-013-2179-5

Kariluoto, S., Edelmann, M., and Piironen, V. (2010). Effects of environment and genotype on folate concentrations in wheat in the healthgrain diversity screen. *J. Agric. Food Chem.* 2010, 9324–9331. doi: 10.1021/jf100251j

Khangura, B. S., Gill, K. S., and Phul, P. S. (1980). Combining ability analysis of beta-carotene, total carotenoids and other grain characteristics in pearl millet. *Theor. Appl. Genet.* 56, 91–96. doi: 10.1007/BF00264433

Lachman, J., Hamouz, K., Musilová, J., Hejtmánková, K., Kotíková, Z., Pazderù, K., et al. (2013). Effect of peeling and three cooking methods on the content of selected phytochemicals in potato tubers with various colour of flesh. *Food Chem.* 138, 1189–1197. doi: 10.1016/j.foodchem.2012.11.114

Li, F. Q., Murillo, C., and Wurtzel, E. T. (2007). Maize Y9 encodes a product essential for15-cis-zeta-carotene isomerization. *Plant Physiol.* 144, 1181–1189. doi: 10.1104/pp.107.098996

Li, W., Beta, T., Sun, S., and Corke, H. (2006). Protein characteristics of Chinese black-grained wheat. *Food Chem.* 98, 463–472.

Li, Z. H., Matthews, P. D., Burr, B., and Wurtzel, E. T. (1996). Cloning and characterization of a maize cDNA encoding phytoene desaturase, an enzyme of the carotenoid biosynthetic pathway. *Plant Mol. Biol.* 30, 269–279. doi: 10.1007/bf00020113

Lian, T., Guo, W., Chen, M., Li, J., Liang, Q., Liu, F., et al. (2015). Genome-wide identification and transcriptional analysis of folate metabolism-related genes in maize kernels. *BMC Plant. Biol.* 15:204. doi: 10.1186/s12870-015-0578-2

Liang, Q., Wang, K., Liu, X., Riaz, B., Jiang, L., Wan, X., et al. (2019). Improved folate accumulation in genetically modified maize and wheat. *J. Exp. Bot.* 70, 1539–1551. doi: 10.1093/jxb/ery453

Maluf, M. P., Saab, I. N., Wurtzel, E. T., and Sachs, M. M. (1997). The viviparous 12 maize mutant is deficient in abscisic acid, carotenoids, and chlorophyll synthesis. ?*J. Exp. Bot.* 48, 1259–1268.

Masisi, K., Diehl-Jones, W. L., Gordon, J., Chapman, D., Moghadasian, M. H., and Beta, T. (2015). Carotenoids of aleurone, germ, and endosperm fractions of barley, corn and wheat differentially inhibit oxidative stress. *J. Agric. Food Chem.* 63, 2715–2724. doi: 10.1021/jf5058606

Matthews, P. D., Luo, R., and Wurtzel, E. T. (2003). Maize phytoene desaturase and zeta carotene desaturase catalyze a poly Z desaturation pathway: Implications for genetic engineering of carotenoid concentration among cereal crops. *J. Exp. Bot.* 54, 2215–2230. doi: 10.1093/jxb/erg235

McCouch, S. R., Teytelman, L., Xu, Y., Lobos, K. B., Clare, K., Walton, M., et al. (2002). Development and mapping of 2,240 new SSR markers for rice (*Oryza sativa* L.). *DNA Res.* 9, 199–207.

McCully, K. S. (2007). Homocysteine, vitamins, and vascular disease prevention. *Am. J. Clin. Nutr.* 86, 1563S–1568S. doi: 10.1093/ajcn/87.4.1069

McGraw, K. J., Hill, G. E., Stradi, R., and Parker, R. S. (2001). The influence of carotenoid acquisition and utilization on the maintenance of species-typical plumage pigmentation in male American goldfinches (*Carduelis tristis*) and

northern cardinals (*Cardinalis cardinalis*). *Physiol. Biochem. Zool.* 74, 843–852. doi: 10.1086/323797

McGuire, S. (2015). FAO., IFAD., and WFP. The state of food insecurity in the world 2015: meeting the 2015 international hunger targets: taking stock of uneven progress, Rome: FAO. *Adv. Nutr.* 6, 623–624. doi: 10.3945/an.115.009936

Meenakshi, J. V., Johnson, N. L., Manyong, V. M., DeGroote, H., Javelosa, J., Yanggen, D. R., et al. (2010). How cost-effective is biofortification in combating micronutrient malnutrition? An ex ante assessment. *World Dev.* 38, 64–75.

Menkir, A., Liu, W., White, W. S., Mazlya-Dixon, B., and Rocheford, T. (2008). Carotenoid diversity in tropical-adapted yellow maize inbred lines. *Food Chem.* 109, 521–529. doi: 10.1016/j.foodchem.2008.01.002

Menkir, A., Palacios-Rojas, N., Alamu, O., Dias, P., Maria, C., Dhliwayo, T., et al. (2018). *Vitamin A Biofortified Maize. Exploiting Native Genetic Variation for Nutrient Enrichment.* Biofortification No. 2 (February, 2018). Bonn: CIMMYT.

Muthamilarasan, M., and Prasad, M. (2016). "Role of genomics in enhancing nutrition concentration of cereals," in *Advances in Plant Breeding Strategies: Agronomic, Abiotic and Biotic Stress Traits*, Vol. 2, eds J. M. Al-Khayri, S. M. Jain, and D. V. Johnson (Cham: Springer), 77–96.

Muthayya, A., Rah, J. H., Sugimoto, J. D., Roos, F. F., Kraemer, K., and Black, R. E. (2013). The global hidden hunger indices and maps: an advocacy tool for action. *PLoS One* 8:e67860.1. doi: 10.1371/journal.pone.0067860

Muthusamy, V., Hossain, F., Thirunavukkarasu, N., Choudhary, M., Saha, S., Bhat, J. S., et al. (2014). Development of β-carotene rich maize hybrids through marker-assisted introgression of β-carotene hydroxylase allele. *PLoS One*. 9:e113583. doi: 10.1371/journal.pone.0113583

Muthusamy, V., Hossain, F., Thirunavukkarasu, N., Saha, S., Agrawal, P. K., Guleria, S. K., et al. (2015). Genetic variability and inter-relationship of kernel carotenoids among indigenous and exotic maize (*Zea mays* L.) inbreds. *Cereal Res. Commun.* 43, 567–578.

Naqvi, S., Zhu, C., Farre, G., Ramessar, K., Bassie, L., Breitenbach, J., et al. (2009). Transgenic multivitamin corn through biofortification of endosperm with three vitamins representing three distinct metabolic pathways. *Proc. Natl. Acad. Sci. U.S.A.* 106, 7762–7767. doi: 10.1073/pnas.0901412106

Ndolo, V. U., and Beta, T. (2013). Distribution of carotenoids in endosperm, germ, and aleurone fractionsof cereal grain kernels. *Food Chem.* 139, 663–671. doi: 10.1016/j.foodchem.2013.01.014

Ortiz-Covarrubias, Y., Dhliwayo, T., Palacios-Rojas, N., Ndhlela, T., Magorokosho, C., Aguilar-Rincon, V. H., et al. (2019). Effects of drought and low nitrogen stress on provitamin A carotenoid content of biofortified maize hybrids. *Crop Sci.* 59, 2521–2532. doi: 10.2135/cropsci2019.02.0100

Ortiz-Monasterio, J. I., Palacios-Rojas, N., Meng, E., Pixley, K., Trethowan, R., and Pe˜na, R. J. (2007). Enhancing the mineral and vitamin content of wheat and maize through plant breeding. *J. Cereal Sci.* 46, 293–307.

Paine, J. A., Shipton, C. A., Chaggar, S., Howells, R. M., Kennedy, M. J., Vernon, G., et al. (2005). Improving the nutritional value of golden cice through increased pro-vitamin A content. *Nat. Biotechnol.* 23, 482–487. doi: 10.1038/Nbt108

Parker, G. D., Chalmers, K. J., Rathjen, A. J., and Langridge, P. (1998). Mapping loci associated with flour color in wheat. *Theor. Appl. Genet.* 97, 238–245. doi: 10.1111/pbi.13126

Parker, G. D., and Langridge, P. (2000). Development of a STS marker linked to a major locus controlling flour colour in wheat (*Triticum aestivum* L.). *Mol. Breed.* 6, 169–174.

Piironen, V., Edelmann, M., Kariluoto, S., and Bedo, Z. (2008). Folate in wheat genotypes in the healthgrain diversity screen. *J. Agric. Food Chem.* 56, 9726–9731. doi: 10.1021/jf801066j

Ponce, K. S., Ye, G., and Zhao, X. (2018). QTL identification for cooking and eating quality in indica rice using multi-parent advanced generation intercross (MAGIC) population. *Front. Plant Sci.* 9:868. doi: 10.3389/fpls.2018.00868

Poutanen, K., Shepherd, R., Shewry, P., Delcour, J., Björck, I., and Van der Kamp, J. (2008). Beyond whole grain: the European HEALTHGRAIN project aims at healthier cereal foods. *Cereal Food World* 53, 32–35.

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

Raboy, V. (1996). *Cereal Low Phytic Mutants: A 'Global' Approach to Improving Mineral Nutritional Quality.* Micronutrients and Agriculture, No. 2. Washington, DC: Federation of American Scientists.

Ramachandran, A., Pozniak, C. J., Clarke, J. M., and Singh, A. K. (2010). Carotenoid accumulation during grain development in durum wheat. *J. Cereal Sci.* 52, 30–38. doi: 10.1007/s11103-013-0160-y

Ramos, M. I., Allen, L. H., Mungas, D. M., Jagust, W. J., Haan, M. N., Green, R., et al. (2005). Low folate status is associated with impaired cognitive function and dementia in the Sacramento Area Latino Study on Aging. *Am. J. Clin. Nutr.* 82, 1346–1352. doi: 10.1093/ajcn/82.6.1346

Reddy, B. V. S., Ramesh, S., and Longvah, T. (2005). Prospects of breeding for micronutrients and β-carotene-dense sorghums. *Int. Sorghum Millets Newslett.* 46, 10–14.

Robertsen, D. C., Hjortshøj, L. R., and Janss, L. L. (2019). Genomic selection in cereal breeding. *Agronomy* 9:95. doi: 10.3390/agronomy9020095

Safawo, T., Senthil, N., Raveendran, M., Vellaikumar, S., Ganesan, K. N., Nallathambi, G., et al. (2010). Exploitation of natural variability in maize for β-carotene concentration using HPLC and gene specific markers. *Electron. J. Plant Breed.* 1, 548–555.

Saltzman, A., Birol, E., Oparinde, A., Andersson, M. S., Asare-Marfo, D., Diressie, M. T., et al. (2017). Availability, production, and consumption of crops biofortified by plant breeding: current evidence and future potential. *Ann. N. Y. Acad. Sci. U.S.A.* 1390, 104–114. doi: 10.1111/nyas.13314

Santra, M., Santra, D. K., Rao, V. S., Taware, S. P., and Tamhankar, S. A. (2005). Inheritance of β-carotene concentration in durum wheat (*Triticum turgidum* L. ssp. durum). *Euphytica* 144, 215–221. doi: 10.1007/s10681-005-5815-6

Schmaelzle, S., Gannon, B., Crawford, S., Arscott, S. A., Goltz, S., Palacios-Rojas, N., et al. (2014). Maize genotype and food matrix affect the provitamin A carotenoid bioefficacy from staple and carrot-fortified feeds in Mongolian gerbils (*Meriones unguiculatus*). *J. Agric. Food Chem.* 62, 136–143. doi: 10.1021/jf403548w

Schnable, P. S., Ware, D., Fulton, R. S., Stein, J. C., Wei, F., Pasternak, S., et al. (2009). The B73 maize genome: complexity, diversity, and dynamics. *Science* 20, 1112–1115. doi: 10.1126/science.1178534

Scholl, T. O., and Johnson, W. G. (2000). Folic acid: influence on the outcome of pregnancy. *Am. J. Clin. Nutr.* 71, 1295S–1303S. doi: 10.1093/ajcn/71.5.1295s

Sehgal, D., Skot, L., Singh, R., Srivastava, R. K., Das, S. P., Taunk, J., et al. (2015). Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PLoS One* 10:e0122165. doi: 10.1371/journal.pone.0122165

Senthilvel, S., Jayashree, B., Mahalakshmi, V., Kumar, P. S., Nakka, S., Nepolean, T., et al. (2008). Development and mapping of simple sequence repeat markers for pearl millet from data mining of expressed sequence tags. *BMC Plant Biol.* 8:119. doi: 10.1186/1471-2229-8-119

Sharopova, N., McMullen, M. D., Schultz, L., Schroeder, S., Sanchez-Villeda, H., Gardiner, J., et al. (2002). Development and mapping of SSR markers for maize. *Plant Mol. Biol.* 48, 463–481. doi: 10.1023/a:1014868625533

Shen, Y., Su, X., Rhodes, D., Herald, T., Xu, J., Chen, X., et al. (2017). The pigments of sorghum pericarp are associated with the concentrations of carotenoids and pro-vitamin A. *Int. J. Food Nutri. Sci.* 6, 48–56.

Shewry, P. R., Piironen, V., Lampi, A., Edelmann, M., Kariluoto, S., Nurmi, T., et al. (2010). Effects of genotype and environment on the content and composition of phytochemicals and dietary fiber components in rye in the HEALTHGRAIN diversity screen. *J. Agric. Food Chem.* 58, 9372–9383. doi: 10.1021/jf100053d

Simpungwe, E., Dhliwayo, T., Palenberg, M., Taleon, V., Birol, E., Oparinde, A., et al. (2017). Orange maize in Zambia: crop development and delivery experience. *Afr. J. Food Agric. Nutr. Dev.* 17, 11973–11999.

Singh, M., Lewis, P. E., Hardeman, K., Bai, L., Rose, J. K. C., Mazourek, M., et al. (2003). Activator mutagenesis of the pink scutellum1/viviparous7 locus of maize. *Plant Cell* 15, 874–884. doi: 10.1105/tpc.010249

Sivaranjani, R., Prasanna, B. M., Hossain, F., and Santha, I. M. (2013). Genetic variability for total carotenoid concentration in selected tropical maize (*Zea mays*) inbred lines. *Indian J. Agric. Sci.* 83, 431–436.

Smith, M. R., and Myers, S. S. (2018). Impact of anthropogenic CO_2 emmission on global human nutrition. *Nat. Clim. Change* 8, 834–839.

Sowa, M., Yu, J., Palacios-Rojas, N., Goltz, S. R., Howe, J. A., Davis, C. R., et al. (2017). Retention of carotenoids in biofortified maize flour and β-cryptoxanthin-enhanced eggs after household cooking. *ACS Omega* 2, 7320–7328. doi: 10.1021/acsomega.7b01202

Suwarno, W. B., Pixley, K. V., Palacios-Rojas, N., Kaeppler, S. M., and Babu, R. (2014). Formation of heterotic groups and understanding genetic effects in a provitamin A biofortified maize breeding program. *Crop Sci*. 54, 14–24. doi: 10.2135/cropsci2013.02.0096

Taleon, V., Mugode, L., Cabrera-Soto, L., and Palacios-Rojas, N. (2017). Carotenoid retention in biofortified maize using different post-harvest storage and packaging methods. *Food Chem.* 232, 60–66. doi: 10.1016/j.foodchem.2017.03.158

Tamanna, S., Sayma, P., Sanjay, K., Alak, K. D., Aysha, F. M., Ali, S., et al. (2013). Concentration of some minerals and their bioavailability in selected popular rice varieties from Bangladesh. *Int. J. Curr. Microbiol. Appl. Sci*. 2, 35–43.

Tan, J., Baisakh, N., Oliva, N., Torrizo, L., Abrigo, E., Datta, K., et al. (2004). The screening of rice germplasm including those transgenic rice lines which accumulate b-carotene in their polished seeds for their carotenoid profile. *Int. J. Food Sci. Technol.* 40, 563–569.

USDA–ARS (2012). *USDA National Nutrient Database for Standard Reference, RELEASE 25. Nutrient Data Laboratory Home Page*. Available online at: http://www.ars.usda.gov/ba/bhnrc/ndl (accessed August 5, 2019).

Vallabhaneni, R., and Wurtzel, E. T. (2009). Timing and biosynthetic potential for carotenoid accumulation in genetically diverse germplasm of maize. *Plant Physiol*. 150, 562–572. doi: 10.1104/pp.109.137042

Varshney, R. K., Hoisington, D. A., and Tyagi, A. K. (2006). Advances in cereal genomics and applications in crop breeding. *Trends Biotechnol*. 24, 490–499. doi: 10.1016/j.tibtech.2006.08.006

Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35, 969–976. doi: 10.1038/nbt.3943

Verbyla, A. P., George, A. W., Cavanagh, C. R., and Verbyla, K. L. (2014). Whole-genome QTL analysis for MAGIC. *Theor. Appl. Genet.* 127, 1753–1770. doi: 10.1007/s00122-014-2337-4

Wang, C., Zeng, J., Li, Y., Hu, W., Chen, L., Miao, Y., et al. (2014). Enrichment of provitamin A content in wheat (*Triticum aestivum* L.) by introduction of the bacterial carotenoid biosynthetic genes CrtB and CrtI. *J. Exp. Bot.* 65, 2545–2556. doi: 10.1093/jxb/eru138

Ward, J. L., Poutanen, K., Gebruers, K., Piironen, V., Lampi, A. M., Nyström, L., et al. (2008). The HEALTHGRAIN cereal diversity screen: concept, results, and prospects. *J. Agric. Food Chem*. 56, 9699–9709. doi: 10.1021/jf8009574

Welch, R. M., and Graham, R. D. (2005). Agriculture: the real nexus for enhancing bioavailable micronutrients in food crops. *J. Trace Elem. Med. Biol.* 18, 299–307. doi: 10.1016/j.jtemb.2005.03.001

Wong, J. C., Lambert, R. J., Wurtzel, E. T., and Rocheford, T. R. (2004). QTL and candidate genes phytoene synthase and zeta-carotene desaturase associated with the accumulation of carotenoids in maize. *Theor. Appl. Genet*. 108, 349–359. doi: 10.1007/s00122-003-1436-4

Yadav, R. S., Hash, C. T., Bidinger, F. R., and Howarth, C. J. (1999). "QTL analysis and marker-assisted breeding of traits associated with drought tolerance in pearl millet," in *Proceedings of the Workshop on Genetic Improvement of Rice for Water Limited Environemnts, 1-3 December 1998*, eds O. Ito, J. O'Toole, and B. Hardy (Los Banos: International Rice research Institute), 211–223. doi: 10.1007/s001220200008

Yadava, D. K., Choudhury, P. K., Hossain, F., and Kumar, D. K. (2017). *Biofortified Varieties: Sustainable Way to Alleviate Malnutrition*. New Delhi: Indian Council of Agricultural Research, 7–10.

Yan, J. B., Kandianis, C. B., Harjes, C. E., Bai, L., Kim, E. H., Yang, X. H., et al. (2010). Rare genetic variation at Zea mays crtRB1 increases b-carotene in maize grain. *Nat. Genet.* 42, 322–327. doi: 10.1038/ng.551

Ye, X. S., Al-Babili, A., Kloti, J., Zhang, P., Lucca, P., Beyer, P., et al. (2000). Engineering the provitamin A (b-carotene) biosynthetic pathway into (carotenoid free) rice endosperm. *Science* 287, 303–305. doi: 10.1126/science.287.5451.303

Yu, J., Hu, S., Wang, J., Wong, G. K., Li, S., Liu, B., et al. (2002). A draft sequence of the rice genome (*Oryza sativa* L. ssp. indica). *Science* 296, 79–92. doi: 10.1126/science.1068037

Yu, S., and Tian, L. (2018). Breeding major cereal grains through the lens of nutrition sensitivity. *Mol. Plant* 11, 23–30. doi: 10.1016/j.molp.2017.08.006

Zhang, L. P., Yan, J., Xia, X. C., He, Z. H., and Sutherland, M. W. (2006). QTL mapping for kernel yellow pigment concentration in common wheat. *Acta Agron. Sin.* 32, 41–45.

Zhang, Y. L., Wu, Y. P., Xiao, Y. G., He, Z. H., Zhang, Y., Yan, J., et al. (2009). QTL mapping for flour color components, yellow pigment concentration and polyphenol oxidase activity in common wheat (*Triticum aestivum* L.). *Euphytica* 165, 435–444.

Zhou, Y., Han, Y. J., Li, Z. G., Fu, Y., Fu, Z. Y., Yang, X., et al. (2012). ZmcrtRB3 encodes a carotenoid hydroxylase that affects the accumulation of a-carotene in maize kernel. *J. Integr. Plant Biol.* 54, 260–269. doi: 10.1111/j.1744-7909.2012.01106.x

Zhu, C., Naqvi, S., Breitenbach, J., Gerhard, S., Paul, C., and Teresa, C. (2008). Combinatorial genetic transformation generates a library of metabolic phenotypes for the carotenoid pathway in maize. *Proc. Natl. Acad. Sci. U.S.A.* 105, 18232–18237. doi: 10.1073/pnas.0809737105

Zunjare, R. U., Hossain, F., Muthusamy, V., Baveja, A., Chauhan, H. S., Bhat, J. S., et al. (2018). Development of biofortified maize hybrids through marker-assisted stacking of β-carotene hydroxylase, lycopene cyclase and opaque2 genes. *Front. Plant Sci.* 9:178. doi: 10.3389/fpls.2018.00178

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 Genetics

REVIEW
published: 23 June 2020
doi: 10.3389/fgene.2020.00500

Barnyard Millet for Food and Nutritional Security: Current Status and Future Research Direction

*Vellaichamy Gandhimeyyan Renganathan[1,2], Chockalingam Vanniarajan[1], Adhimoolam Karthikeyan[2] and Jegadeesan Ramalingam[2]**

[1] Department of Plant Breeding and Genetics, Agricultural College & Research Institute, Tamil Nadu Agricultural University, Madurai, India, [2] Department of Biotechnology, Centre of Innovation, Agricultural College & Research Institute, Tamil Nadu Agricultural University, Madurai, India

Barnyard millet (*Echinochloa* species) has become one of the most important minor millet crops in Asia, showing a firm upsurge in world production. The genus *Echinochloa* comprises of two major species, *Echinochloa esculenta* and *Echinochloa frumentacea*, which are predominantly cultivated for human consumption and livestock feed. They are less susceptible to biotic and abiotic stresses. Barnyard millet grain is a good source of protein, carbohydrate, fiber, and, most notably, contains more micronutrients (iron and zinc) than other major cereals. Despite its nutritional and agronomic benefits, barnyard millet has remained an underutilized crop. Over the past decades, very limited attempts have been made to study the features of this crop. Hence, more concerted research efforts are required to characterize germplasm resources, identify trait-specific donors, develop mapping population, and discover QTL/gene (s). The recent release of genome and transcriptome sequences of wild and cultivated *Echinochloa* species, respectively has facilitated in understanding the genetic architecture and decoding the rapport between genotype and phenotype of micronutrients and agronomic traits in this crop. In this review, we highlight the importance of barnyard millet in the current scenario and discuss the up-to-date status of genetic and genomics research and the research gaps to be worked upon by suggesting directions for future research to make barnyard millet a potential crop in contributing to food and nutritional security.

Keywords: barnyard millet, *Echinochloa* species, micronutrients, small millets, genetic and genomic resources, value addition

OPEN ACCESS

Edited by:
Mallikarjuna Swamy, International Rice Research Institute, Philippines

Reviewed by:
Atefeh Sabouri, University of Guilan, Iran
Passoupathy Rajendrakumar, ICAR-Indian Institute of Millets Research (IIMR), India
Lohithaswa Hirenallur Chandappa, University of Agricultural Sciences, India

***Correspondence:**
Jegadeesan Ramalingam
ramalingam.j@tnau.ac.in

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *05 November 2019*
Accepted: *22 April 2020*
Published: *23 June 2020*

Citation:
Renganathan VG, Vanniarajan C, Karthikeyan A and Ramalingam J (2020) Barnyard Millet for Food and Nutritional Security: Current Status and Future Research Direction. Front. Genet. 11:500. doi: 10.3389/fgene.2020.00500

INTRODUCTION

Barnyard millet (*Echinochloa* species) is an ancient millet crop grown in warm and temperate regions of the world and widely cultivated in Asia, particularly India, China, Japan, and Korea. It is the fourth most produced minor millet, providing food security to many poor people across the world. Globally, India is the biggest producer of barnyard millet, both in terms of area (0.146 m ha^{-1}) and production (0.147 mt) with average productivity of 1034 kg/ha during the last 3 years (IIMR, 2018). The details on major areas of cultivation and worldwide production are presented in **Figures 1**, **2**. Barnyard millet is primarily cultivated for human consumption, though it is also used as a livestock feed. Among many cultivated and wild species of barnyard millet, two of the most popular species are *Echinochloa frumentacea* (Indian barnyard millet) and *Echinochloa esculenta* (Japanese barnyard millet) (Sood et al., 2015). Barnyard millet is a

FIGURE 1 | Major cultivation and production of barnyard millet around the world (from IIMR, 2018).

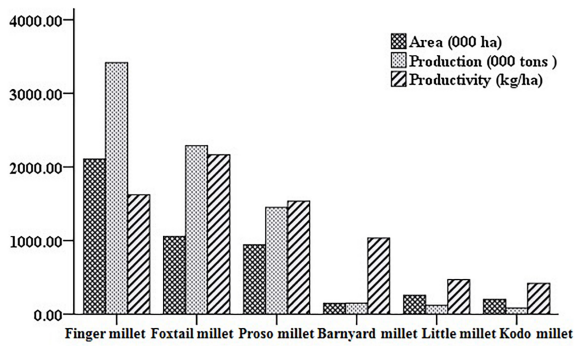


FIGURE 2 | World area, production, and productivity scenario of small millets (from IIMR, 2018).

short duration crop that can grow in adverse environmental conditions with almost no input and can withstand various biotic and abiotic stresses. In addition to these agronomic advantages, the grains are valued for their high nutritional value and lower expense as compared to major cereals like rice, wheat, and maize. It contains a rich source of protein, carbohydrates, fiber, and, most notably, micronutrients like iron (Fe) and zinc (Zn) (Singh et al., 2010; Saleh et al., 2013; Chandel et al., 2014) that are related to numerous health benefits (Saleh et al., 2013; Ugare et al., 2014). All these features make barnyard millet an ideal supplementary crop for subsistence farmers and also as an alternate crop during the failure of monsoons in rice/major crop cultivating areas (Gupta et al., 2009).

Despite barnyard millet's excellent nutritional and agronomic value, the lack of awareness has led this crop to be considered as a neglected and underutilized crop. Over the past decades, efforts made to study the features of barnyard millet are limited compared to other minor millets. So far, most of the studies have been aimed at exploring the knowledge of diversity within the germplasm through morphological (Mehta et al., 2005; Gupta et al., 2009; Nirmalakumari and Vetriventhan, 2009; Sood et al., 2015; Renganathan et al., 2017) and molecular markers (Nozawa et al., 2006; Altop and Mennan, 2011; Prabha et al., 2012; Wallace et al., 2015; Manimekalai et al., 2018; Murukarthick et al., 2019). Also, several studies disclosed the nutritional profile of barnyard millet, particularly the high Fe and Zn content in the grains (Veena et al., 2005; Saleh et al., 2013; Chandel et al., 2014; Renganathan et al., 2017). However, comprehensive research is needed to understand the accurate details of germplasm accessions, identify the trait-specific donors, develop mapping population, and discover the quantitative trait locus (QTLs) and gene. Genomic resources are helpful for the progress of any crop species and they assist effective characterization of germplasm resources and their subsequent use in the discovery of QTL/gene(s) for the crop improvement program. However, genome research in barnyard millet is still in the early stage and far behind the other minor millets. This is mainly due to the complex nature of the genome ($2n = 6x = 54$, hexaploid). Recently, second and third-generation sequencing technologies unlocked several genome sequencing issues and facilitated to identify the genome sequence of wild and transcriptome sequences in cultivated *Echinochloa* species (Li et al., 2013a,b; Yang et al., 2013; Nah et al., 2015; Xu et al., 2015; Guo et al., 2017; Murukarthick et al., 2019). These genome resources facilitated the chance for better genotyping studies such as genetic diversity analysis, development of highly dense linkage maps and accurate physical maps, and detection of QTLs associated with micronutrients and agronomic traits. For instance, Wallace et al. (2015) developed the genome through single nucleotide polymorphism (SNP) markers and analyzed the genetic diversity in the barnyard millet core collection. Murukarthick et al. (2019)

investigated the transcriptional changes between *E. frumentacea* and *E. cru-galli* and discovered genes related to drought and micronutrient content.

Accumulating evidence suggests that the interest in barnyard millet research has increased markedly over recent years; since 2010, more than 350 publications on barnyard millet have been available in the National Center for Biotechnology Information (NCBI) PubMed database[1] (last accessed on December 2019) (**Figure 3A**). This review discusses the origin and taxonomy, nutritional value and health benefits, stress adaptation as well as the current status of genetic and genomics research in barnyard millet. The final section highlights the research gap and future research directions needed to promote barnyard millet as a potential crop for food and nutritional security.

ORIGIN, TAXONOMY AND GENOMIC RELATIONSHIP OF *ECHINOCHLOA* SPECIES

Barnyard millet belongs to the genus *Echinochloa*, the family *Poaceae*, and the sub-family *Panicoideae* (Clayton and Renvoize, 2006). The genus *Echinochloa* consists of approximately 250 annual and perennial species that are widely distributed in the warmer and temperate parts of the world (Bajwa et al., 2015). However, the lack of clarity over the *Echinochloa* species makes it hard to differentiate themselves via the morphological markers due to low interspecific and intraspecific variations in nature and their phenotype plasticity (Chauhan and Johnson, 2011). Despite this challenge, 35 species have been identified to date for their taxa and phylogenetic relationship through morphological, cytological, and molecular marker studies (Yabuno, 1966, 1987; Yuichiro et al., 1999; Yamaguchi et al., 2005). Among them, most of the *Echinochloa* species, including *E. crus-galli* (allohexaploid, $2n = 6x = 54$), *E. colona* (allohexaploid, $2n = 6x = 54$), *E. oryzicola* (allotetraploid, $2n = 4x = 36$), and others, have been designated as problematic weeds in major crop fields (Yabuno, 1966, 1987; Wanous, 1990; Yuichiro et al., 1999; Yamaguchi et al., 2005; Kraehmer et al., 2015). *E. crus-galli* is a predominant weed in rice fields in more than 60 countries, due to its quick germination (even in hypoxic conditions, up to 100 mm deep), rapid growth, mimicking character of rice, broad ecological tolerance, and profuse seed production (Barrett, 1983).

Echinochloa species have very few cultivatable forms and thereby are cultivated as minor millet by marginal farmers in warmer and temperate regions of the world. *E. frumentacea* (Roxb.) Link; syn. *E. colona* var. *frumentacea* (allohexaploid, $2n = 6x = 54$), commonly known as Indian barnyard millet, originated from wild *E. colona* (L.) (Jungle rice), and exhibits a parallel line of evolution in India and Africa. *E. frumentacea* species has four races, namely *stolonifera*, *intermedia*, *robusta*, and *laxa*, that are widely cultivated in Central Africa, India, Malawi, Nepal, Pakistan, and Tanzania (Doggett, 1989; Upadhyaya et al., 2014). Another cultivated allohexaploid species, *E. esculenta* (A. Braun) H. Scholz; syn. *E. utilis* var. *esculenta*; known as Japanese barnyard millet, originated from wild *E. crus-galli* (L.) (Barnyard grass) was domesticated some 4,000 years ago in the temperate regions of Japan (De Wet et al., 1983; Doggett, 1989). *Utilis* and *intermedia* are two races of *E. crus-galli*, widely cultivated in Japan, Korea, China, Russia, and Germany (De Wet et al., 1983; Yabuno, 1987; Upadhyaya et al., 2014). Both wild and cultivated *Echinochloa* species are different from each other in terms of growth habitat, general morphology, and other characteristics (**Table 1**). The interspecific relationship between *Echinochloa* species was unclear till a series of prominent taxonomic reports by Yabuno, (1962, 1984, 1996, 2001). The interspecific hybrids between wild species and its progenitor, i.e., *E. crus-galli* × *E. esculenta* and *E. colona* × *E. frumentacea* produce normal meiotic division (27 bivalents) i.e., fertile. But, interspecific hybrids between two cultivated species and their respective wild counterparts, *E. esculenta* × *E. frumentacea* and *E. crus-galli* × *E. colona*,

[1] http://www.ncbi.nlm.nih.gov/sites/entrez

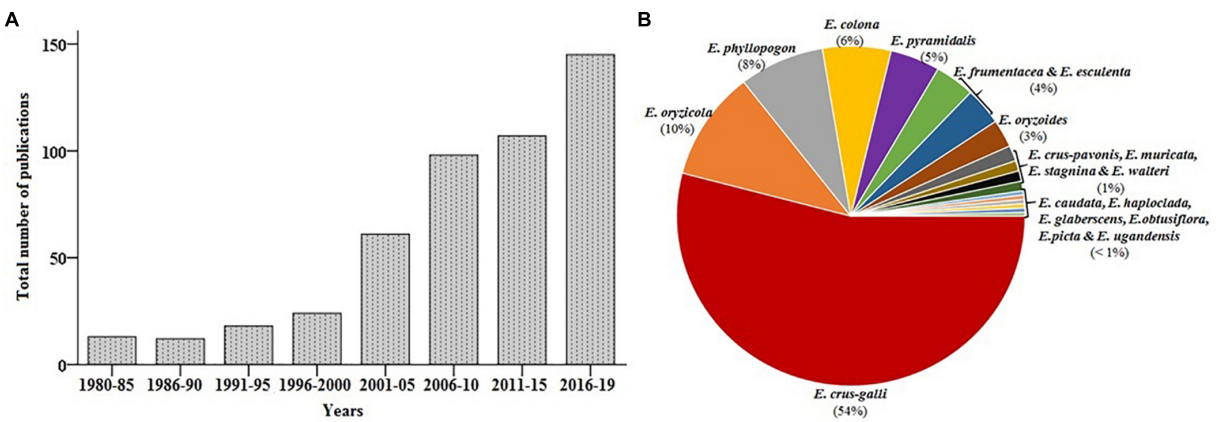


FIGURE 3 | Graphical representations of literature and nucleotide sequence availability. **(A)** PubMed (https://www.ncbi.nlm.nih.gov/pubmed) **(B)** nucleotide sequence (https://www.ncbi.nlm.nih.gov/). (Note: Data verified till December 2019).

TABLE 1 | Morphological differences among wild and cultivated species of *Echinochloa* species.

Traits	*Echinochloa colona*[a]	*Echinochloa crus-galli*[a]	*Echinochloa frumentacea*	*Echinochloa esculenta*
Common name	Jungle rice	Barnyard grass	Indian barnyard millet/sawa millet	Japanese barnyard millet
Synonyms	*Echinochloa colonum*, *Echinochloa crus-galli* subsp. *colona*, *Panicum colonum*, *Panicum cumingianum*, *Panicum zonale*, and *Milium colonum*, *Oplismenus colonus*	*Panicum crus-galli*, *Panicum hispidulum*, *Milium crus-galli*, and *Pennisetum crus-galli*	Billion Dollar grass, sawa millet, sama millet	Japanese millet, marsh millet, Siberian millet, and white millet
Origin	China and Japan	China, Japan, and Korea	India, Pakistan, and Nepal	Eastern Asia, Japan, China, and Korea
Distribution	Widely distributed in South and Southeast Asia, Australia, Africa, Europe, and America	Widely distributed in South and Southeast Asia, Africa, Europe, and America	Widely distributed in Central Africa, Africa, Temperate Asia, Tropical Asia, Australasia, and South America.	Widely distributed in India, Temperate Asia, Australasia, and Pacific.
Habitat	Annual, warmer region	Annual, temperate region	Annual, warmer region	Annual, warm-season
Cultivable range	–	–	Altitude below 1,900 m	Altitude up to 2,500 m
General morphology	Erect to decumbent, 60 cm tall, short leaf length 10-15 cm, red tinges at the basal portion of leaf, leaf blade surface smooth, leaf blades 3–30 cm long and 2–8 mm wide	Erect, 200 cm tall, leaf 10-40 cm long, leaf blade surface smooth, leaf blades 0.5–35 cm long and 6–20 mm wide	Erect, 242 cm tall, leaf length 15–40 cm long and 1–2.5 cm wide, plants mostly green, however, purple tinges also found in vegetative and reproductive parts, leaf blades are smooth and glabrous, culms slender to robust	Robust, 60–122 cm tall, leaf sheath smooth 10–50 cm long and 7–25 mm wide, plants green, however, light to dark purple pigmentation in various plant parts, thicker stem
Inflorescence morphology	Green to purple, inflorescence length 5-15 cm, simple ascending racemes 5–15 numbers, 0.5–3 cm long raceme	Green to purple, inflorescence length 10-25 cm, compound ascending racemes 5–15 numbers, 2–10 cm long, slightly hairy, green to purplish awns 2-5 mm long	Green to purple, usually erect and compact, inflorescence length 1-28 cm long, racemes numerous 20–70, 1–3 cm long, rarely drooping, awnless	Brown to purple, compact inflorescence 12–15 cm long, racemes 5–15 numbers, arcuate to flexuous, 0.5–3 cm long, rarely awned
Spikelets on panicle	Spikelets arranged in 4 uniform rows on the primary rachis, spikelets 1-3 mm long	Open, branched spikelets on the rachis, setae on the primary branches	Compact, non-branched spikelets on the rachis, tightly clustered, 2–4 mm long, acute and awnless .	Compact, branched spikelets, larger spikelets and longer primary branches, dense clustered, 3–4 mm long, shortly cuspidate and rarely awned.
Caryopsis	Straw white, brown	Dark gray, brown	Turgid, whitish	Dull, pale yellow to light brown
Seed dormancy	Present	Absent	Present	Absent
Seed shattering	High	Low	High	Low

[a]*http://www.knowledgebank.irri.org/training/fact-sheets/item/echinochloa. Source: Yabuno, 1987; Muldoon et al., 1982; De Wet et al., 1983; Doggett, 1989; Bandyopadhyay, 1999; Yadav and Yadav, 2013; Renganathan et al., 2015.*

showed irregular meiotic division that leads to sterility (Yabuno, 1966, 1984). Collectively, all the cytological studies reveal the poor genomic affinity among species of *Echinochloa*. Besides, Yamaguchi et al. (2005) confirmed three cross compatible groups identified by Yabuno (1966) using chloroplast DNA sequence analysis, and separated these *Echinochloa* complexes into *E. oryzicola* complex, *E. crus-galli* complex, and *E. colona-frumentacea* complex. The *E. oryzicola* complex consists of two weedy species, *E. oryzicola* and *E. phyllopogon*, and one rare cultivated Mosuo barnyard millet. The *E. crus-galli* complex includes four wild species, *E. crus-galli* var. *crus-galli*, *E. crus-galli* var. *praticola*, *E. crus-galli* var. *formosensis*, and *E. crus-galli* var. *oryzoides*, and one major cultivated species, *E. esculenta* (Japanese barnyard millet). The third, *E. colona-frumentacea* complex, consists of one wild species, *E. colona*, and one major cultivated species, *E. frumentacea* (Indian barnyard millet). Further, through molecular analysis, Aoki and Yamaguchi (2008) reported that, though all these three groups exhibit different cytoplasmic lineages, the nuclear lineage between *E. oryzicola* complex and *E. crus-galli* complex have a higher affinity than *E. colona-frumentacea* complexes proving Yabuno's hypothesis that *E. oryzicola* is the probable paternal parent of *E. crus-galli* (Aoki and Yamaguchi, 2008). However, information regarding ancestors of *E. colona* and their cultivated *E. frumentacea* remains almost unknown. Therefore, analyses the meiotic behavior of inter or intra specific hybrid combinations through advanced cytogenetic techniques like genomic in situ hybridization (GISH)/fluorescent in situ hybridization (FISH) would not only provides their ancestral information but also differentiate many unresolved genomes of *Echinochloa* species.

PLANT ARCHITECTURE, FLORAL BIOLOGY AND SEED TRAITS

The cultivated barnyard millet is an annual, robust, and tall crop that grows up to a height of 220 cm (Denton, 1987; Padulosi et al., 2009). The inflorescence is a terminal panicle with varying shapes (cylindrical, pyramidal, and globose to elliptic),

colors (green, light purple, and dark purple) and compactness (compact, intermediate, and open) (Gupta et al., 2009; Prakash and Vanniarajan, 2013; Sood et al., 2014; Renganathan et al., 2017; Kuraloviya et al., 2019) (**Figure 4**). Racemes are present in one side, two sides, or around the axis of rachis and vary from 22 to 64 numbers per inflorescence (Renganathan et al., 2017). The arrangement of spikelets is either on one side or around the rachis of the raceme. Each spikelet contains two florets in which the lower floret is sterile and consists of lemma with small palea, while the upper floret is bisexual with a shiny lemma that partially encloses palea. Fertile lemma and palea have three stamens varying from a white color to a dark purplish color with stigma plumose and bifid, ranging from white to dark purple (**Figure 4**). The two unequal glumes further enclose the seed kernel (Gupta et al., 2010b; Singh et al., 2010). Anthesis and pollination progress in the direction from top to bottom of the inflorescence in the early morning (5 am) and reaches a maximum during 6 am -7 am, while it closes at 10 am (Sundararaj and Thulasidas, 1976; Jayaraman et al., 1997). Sundararaj and Thulasidas (1976) further reported on the requirement of 10–14 days of duration for the flowering process. Though self-pollination is a strict rule, the reception of stigmatic branches before dehiscence of anther provides some chances for cross-pollination (Seetharam et al., 2003).

Compared to other minor millets like kodo and foxtail millet, barnyard millet grains are less hard. The mature pericarp of the seed consists of two epidermal layers with cells of the inner epidermis completely compressed over the outer epidermis (Singh et al., 2010). The cell wall of the aleuronic layer cutinized (Zee and O'brien, 1971), and also contains a maximum amount of carbohydrate (57–66%), followed by fiber (6.4–12.2%), protein (5–8.5%), fat (3.5–4.6%), and ash (2.5–4.0%) content. Starch granules are simple and are spherical to polygonal shapes with a diameter of 1.2–10 μm, which is larger than other small millets (Kumari and Thayumanavan, 1998). The pericarp color of grain differs among genotypes from straw white to light gray and dark gray (Renganathan et al., 2017; Kuraloviya et al., 2019). The seeds usually germinate easily under proper storage conditions at 12°C and are able to retain their viability for up to 13 months (Kannan et al., 2013) and beyond, although improper or poor storage may lead to loss of viability in both species of barnyard millet. The seed dormancy, a major limiting factor in the cultivation of small millets, has not been studied yet in detail. However, in barnyard millet, both wild and freshly harvested seeds of cultivated species reported to have seed dormancy (Maun and Barrett, 1986; Sung et al., 1987; Manidool, 1992). Although the deep physiological dormancy in *E. crus-galli* grain was the most probable feature for its prolonged existence (Song et al., 2015), the innate dormancy present in cultivated *Echinochloa* species further hinders the evaluation or multiplication of seeds in germplasm conservation centers (Kovach et al., 2010). Despite this, the dormancy breaking treatments in *Echinochloa* also varies with species; some accessions may require light or dark and cold or heat or a combination of both (Kovach et al., 2010). Seed application of 100 ppm of IAA (Indoleacetic acid) improved germination

FIGURE 4 | Genetic diversity of various morphological traits of barnyard millet. **(Panel 1)** Ear head diversity of different barnyard millet accessions. **(Panel 2)** Variation in ear head compactness **(a)** compact, **(b)** intermediate, and **(c)** open. **(Panel 3)** Variation in ear head shape **(a)** cylindrical, **(b)** pyramidal, and **(c)** globose to elliptic. **(Panel 4)** Grain variation in color and size of different barnyard millet accessions **(a)** straw white, medium **(b)** light gray, bold **(c)** gray, medium, and **(d)** dark gray, narrow. **(Panel 5)** variation in ear head color **(a)** green, **(b)** medium purple, and **(c)** dark purple. **(Panel 6)** Anther color variation **(a)** white, **(b)** light purple, and **(c)** dark purple. **(Panel 7)** Variation in tillering ability and pigmentation **(a)** high tillers, green, **(b)** medium tillers, light pigmentation, and **(c)** low tillers, dark pigmentation (from Prakash and Vanniarajan, 2013; Renganathan et al., 2017; Kuraloviya et al., 2019).

percentage (18%), speed of germination (5.58 days earlier), and increased the seed length (11%), dry matter (3.80%), and vigor index (21%) (Sujatha et al., 2013). In another study, barnyard millet seeds treated with *Pseudomonas fluorescens* enhanced the seed germination and seedling growth attributes in barnyard millet (Sridevi and Manonmani, 2016). This is mainly due to the direct suppression of deleterious pathogens or the indirect production of growth hormones that ultimately increases the uptake, solubilization, and translocation of less available minerals (Olanrewaju et al., 2017).

Echinochloa species are generally considered to be a short-day plant (Muldoon, 1985) exhibiting photoperiodism and perform as per the different ranges of photoperiods from short days (8–13 h) to long days (16 h) (Maun and Barrett, 1986; Mitich, 1990). For instance, the variety CO (Kv) 2 reported having variable flowering times in temperate and hot regimes within regions of southern parts of India, with hindered uniform grain yield across the state (Vanniarajan et al., 2018). To alleviate that, the latest variety MDU 1, released in Tamil Nadu, India, has been found to have a short duration along with stable grain yield across the state (Vanniarajan et al., 2018).

RELEVANCE OF BARNYARD MILLET IN CLIMATE CHANGE AND NUTRITIONAL SECURITY

Responses to Biotic and Abiotic Stresses

The *Echinochloa* species generally has potential resistance against various biotic and abiotic stresses. However, cultivated species such as *E. esculenta* and *E. frumentacea* are widely threatened by pest and diseases (i.e., shoot fly, stem borer, grain smut, and loose smut) at different growth stages of the crop (Jain et al., 1997; Jagadish et al., 2008). Aphid's infection at the vegetative stage causes considerable yield reduction to *E. frumentacea*. So far, DHBM 996 and TNEF-204 were found to be resistant genotypes for shoot fly and stem borer (Rawat et al., 2019). Meanwhile, Kim et al. (2008) reported that some *E. frumentacea* accessions have the potential for antifeeding activity against brown plant hopper, which is among the major pests that affect rice production. On the other hand, loose smut (*Ustilago tritici*) and grain smut (*Ustilago panici frumentacea*) are major fungal diseases that affect the grain formation in both the cultivated species of *Echinochloa* (Jain et al., 1997; Gupta et al., 2010a). A heavy infestation of smuts during head formation leads to a significant reduction in grain yield and quality (Gupta et al., 2010a). However, Nagaraja and Mantur (2008) and Gupta et al. (2010a) showed that some of the *E. esculenta* accessions had the immunity against both smut diseases and further provide the chance to breed the resistance lines.

Abiotic stresses are a major threat to important food crops such as rice, wheat, and maize, and cause serious yield loss across the world. However, *Echinochloa* species have a high degree of tolerance to various abiotic stresses (Gupta et al., 2010b; Singh et al., 2010). For instance, a recent investigation from Arthi et al. (2019) showed that among the 89 *Echinochloa* accessions, CO (Kv) 2, MDU 1, PRJ1, TNEf 301, TNEf 204, TNEf 361, TNEf 364, and VL 29 exhibited better germination as compared to the rice variety, White Ponni, at 200 mM NaCl concentration. Similarly, *Echinochloa* species are also the preferable choice of farmers for cultivation in various adverse environments such as those prone to drought or flooding. These features showed that the *Echinochloa* species might have some specialized rhizosphere organizations that can facilitate the uptake and release of oxygen (O_2) from their roots at stressful conditions. Zegada-Lizarazu and Iijima (2005) reported higher water uptake efficiency (deep root) of barnyard millet (*E. utilis*) over other minor millets, including pearl millet, and found that barnyard millet sustained and increased the water use efficiency, leaf area index, and dry matter production in both drought and flooding conditions. Therefore, it is also worth investigating the *Echinochloa* species mechanism behind the tolerance to drought and flooding stress. Further, identification and characterization of genes and pathways associated with resistance to saline, drought, and flooding stress in *Echinochloa* species may not only be useful to develop superior cultivars but also assist in improving the tolerance in a major cereal crop. It is also well known for its excellent nitrogen-use efficiency over cereal crops (Goron and Raizada, 2015) and has been recommended as a natural phyto-extractor in heavy metal (lead, cadmium, and chromium) contaminated soils and sodic soils due to hyper accumulation nature. Since heavy metals are currently of much environmental concern, phyto-based soil reclamation is an alternative, cost-effective, and eco-friendly approach (Subhashini and Swamy, 2014) that needs to be imparted in soil health restoration programs.

Nutritional Significance and Health Benefits

In terms of nutritive value, barnyard millet is superior to major and minor millets. Barnyard millet grains are a rich source of dietary fiber, iron, zinc, calcium, protein, magnesium, fat, vitamins, and some essential amino acids (Singh et al., 2010; Saleh et al., 2013; Chandel et al., 2014). The nutritional composition of barnyard millet is presented in **Supplementary Table S1**. The average carbohydrate content of barnyard millet varies between 51.5 and 62.0 g/100 g (Saleh et al., 2013), which is lower than that of other major and minor millets. Ugare et al. (2014) reported that the crude fiber of barnyard millet is higher than any other cereal, ranging between 8.1 and 16.3%. The high ratio of carbohydrate to crude fiber ensures the slower release of sugars in the blood, and so thus aids in maintaining blood sugar level. The resistant starch in barnyard millet has shown to lower blood glucose, serum cholesterol, and triglycerides in rats (Kumari and Thayumanavan, 1998). In a clinical study with human volunteers, Ugare et al. (2014), confirmed a lower glycemic index (GI) in type 2 diabetic groups during regular consumption of barnyard millet meal. Existing evidence showed that the protein content (11.2–12.7%) in barnyard millet was reasonably higher than other major cereals and millets. Although the total minerals, ash, fat, and amino acid content in barnyard millet were although comparable with other cereals and millets, the iron content in the grain was significantly higher than others.

For instance, the iron content in barnyard millet grain is about 15.6–18.6 mg/100 g (Saleh et al., 2013; Renganathan et al., 2017; Vanniarajan et al., 2018), which is rationally higher than major cereals and millets. In addition, a lower phytate (3.30–3.70 mg/100 g) content in grains (Panwar et al., 2016) followed by the dehulling process has also decreased phytic acids significantly, favoring the bioavailability of minerals. This makes barnyard millet an ideal food not only for people with lifestyle diseases, but also for anemic patients and especially women in developing countries. The polyphenols and carotenoids are known to have several potential benefits to humans, and are twofold higher in barnyard millet than finger millet (Panwar et al., 2016). Similarly, alkaloids, steroids, carbohydrates, glycosides, tannins, phenols, and flavonoids present in barnyard millet have various ethno-medical properties like being antioxidant, anti-carcinogenic, anti-inflammatory, antimicrobial, having a wound healing capacity, biliousness, and alleviating constipation-associated diseases (Kim et al., 2011; Ajaib et al., 2013; Moreno Amador et al., 2014; Borkar et al., 2016; Nguyen et al., 2016; Sharma et al., 2016; Sayani and Chatterjee, 2017). Collectively, all these features make the barnyard millet a suitable and secured food for present-day consumers in their overall physical and nutritional well-being.

Physico-Chemical Properties and Relationship With Cooking and Value Addition

The studies on the physical and mechanical properties of grains are an important criterion to design any processing instruments like dehullers, polishers, sorters, storage, and other processing machineries (Singh et al., 2010). In barnyard millet, grain moisture is the prime criteria playing a key role not only in storage but also in the development of processing machineries. The moisture level of barnyard millet grain highly influences the quality as well as the time of milling and polishing (Lohani et al., 2012). For instance, the 8% moisture content of the grain is better for polishing than at 14% moisture. However, at 14% moisture, the degree of polishing increases grain recovery and decreases the loss of protein, fat, ash, and fiber. Based on this, Lohani et al. (2012) suggested 10% as the optimum moisture level for polishing. Similarly, physical (grain diameter, grain surface area, 1,000 grain mass, true density, dynamic angle of repose, coefficient of internal friction, and coefficient of static friction), aerodynamic (terminal velocity), and mechanical (specific deformation and rupture energy) properties are other major parameters influenced by the moisture content of the grains. All these factors ultimately influence the processing of grains in the machines (Singh et al., 2010). Therefore, a study in detail on these properties should be conducted in order to design and develop better milling, polishing, grading, and sorting machineries for barnyard millet.

The cooking and flour quality of the grain are primary standards to assess consumer acceptability. At the same time, different processing techniques aim to increase the storage time of grain/flour as well as increase the physicochemical accessibility of nutrients with reduced anti-nutrient losses during consumption. Barnyard millet grains are usually parboiled-dehulled-cooked and consumed in a similar way to rice (Surekha et al., 2013). It requires about 12 min to cook. The grain can also transform into flour for the preparation of various food formulation by processing techniques. Nevertheless, different processing methods cause the variations in the functional, nutritional, anti-nutritional, and pasting properties of barnyard millet flour (Nazni and Shobana, 2016). In addition, physical parameters, such as bulk density and porosity are important criteria in flour storage and oxidation related problems. Nazni and Shobana (2016) compared the flour of raw and germinated rice for setting up different processing methods for storage and transportation. The study found that the germinated flour exhibited decreased bulk density and porosity (air spacing) than raw rice flour. Therefore, the germinated flours are comparatively less prone to autoxidation than raw rice due to reduced air space between the flour molecules, and this could prevent the spoilage of flour and facilitate easy packaging and enable long-distance transportation. Apart from that, the germinated flours also have an increased oil absorption capacity that makes the flour suitable for enrichment in flavor and mouth feel. Similarly, value-added food products are not only free of anti-nutritional factors but also increase nutritional compounds, making barnyard millet a good base ingredient for infant food formulas. Besides, the flour is highly amenable for various food preparations such as baby foods, snacks, and dietary foods (Vijayakumar et al., 2009; Anju and Sarita, 2010; Surekha et al., 2013). The flour is also highly compatible to blend with other food flours for making novel or any value-added products without affecting the flavor and taste (Veena et al., 2004; Surekha et al., 2013). For instance, a ready-to-eat snack food can successfully be prepared with barnyard millet, potato mash, and tapioca powder in the ratio of 60:37:3 (Jaybhaye and Srivastav, 2015). However, tannin content (0.21–0.36%) in the grain affects the *in vitro* protein digestibility (IVPD). Although, compared to other small millets (kodo millet and finger millet), it was very low. Therefore, it is suggested that there are many chances for the application of several processing techniques to improve barnyard millet flour quality and nutritional properties, especially in the value-addition strategy of food industries.

Versatile Research and Industrial Applications

The amylose-rich barnyard millet starch has now attracted attention in biodegradable film making industries as an antioxidant packaging material (Cao et al., 2017). The incorporation of borage seed oil in barnyard millet starch increases the elongation range and decreases the tensile strength, water permeability, and moisture content properties of the starch, which makes it suitable for biofilm production. These biofilms are found to be resistant against various microbes and block light transparency and free radical formation in food industries (Cao et al., 2017). Research on Nanoparticle by Kumar et al. (2016) suggested the use of an aqueous extract from aerial parts of *E. colona* plant in the synthesis of silver nanoparticles (AgNPs) as a new eco-friendly approach in bio-synthesizing nanoparticle in plants. Such synthesis of AgNPs from plant extracts could be a

safe and eco-friendly approach with possibilities for application at large-scale in the near future in the field of medicine, engineering, and agriculture. Mosovska et al. (2010) reported that *E. esculenta* extract showed antimutagenicity against 3-(5-nitro-2-furyl) acrylic acid in strains of *Salmonella typhimurium* due to its higher polyphenolic content, thereby playing a major role as antioxidants in scavenging H_2O_2 radicals. Similarly, a novel antifungal peptide, *EcAMP1*, was identified in the seeds of *E. crus-galli*, a unique antimicrobial peptide with a wide spectrum of antifungal activity against phytopathogens such as *Alternaria*, *Botrytis*, *Fusarium*, and *Trichoderma*. This peptide has a unique disulfide stabilized α-helical hairpin structure that intensively binds to the surface of fungal conidia, accumulates in the cytoplasm, and finally inhibits the elongation of hyphae without lysis of the cytoplasmic membrane (Nolde et al., 2011). This property could be exploited in future protein engineering technologies for the synthesis of novel antimicrobials in the agriculture and pharmaceutical industries. Besides all these, Barnyard millet has a higher straw yield and fodder value even at multiple cuttings (Bandyopadhyay, 2009). The fodder yield is about 6.3 tons/ha (Vanniarajan et al., 2018). Fodder contains a good amount of protein (7.6%), digestible fiber (23%), ash (12%), and fat (2.0%). Besides its superior feed quality, higher digestibility and nitrogen concentrations have meant barnyard millet is used as a potential livestock feed crop in the dry areas of the Deccan plateau to the extreme hills of the temperate sub-Himalayan region (Singh and Singh, 2005; Bandyopadhyay, 2009; Yadav and Yadav, 2013; Sood et al., 2015).

GERMPLASM RESOURCES AND UTILIZATION

The major collections of barnyard millet germplasm accessions are housed in India and Japan. Vivekananda Parvatiya Krishi Anusandhan Sansthan (VPKAS), India; Indian Institute of Millets Research (IIMR), India; National Institute of Agrobiological Sciences (NIAS), Japan, and Consultative Group on International Agricultural Research like International Crop Research Institute for the Semi-Arid Tropics (ICRISAT), India are actively working on germplasm evaluation for various agronomic, biotic, and abiotic stresses, grain, and nutritional content traits in barnyard millet. India made a series of collaborative exploration missions (i.e., Indo-Australian missions, Indo-Japanese missions, and Indo-Soviet protocol) for the improvement of barnyard millet and other millets with different countries across the world. So far, six hundred and one exotic barnyard millet accessions have been introduced into India between the period of 1976 and 2007 to increase the food and fodder production (Gomashe, 2017). The major source of introduction was from Australia, Canada, France, Germany, Ghana, Italy, Japan, Kenya, Malawi, the Philippines, Russian Federation, South Africa, Spain, United States of America, and Yugoslavia. During this period, Indian barnyard millet accessions were also introduced in the United States, Canada, and Australia for feed and forage purposes (Gomashe, 2017). At present, 8,000 barnyard millet germplasms have been conserved at different centers throughout the world (**Table 2**).

TABLE 2 | Major organizations across the globe conserving the *Echinochloa* species (till December 2019).

Country	Number of accessions	Organization
India	1888	National Bureau of Plant Genetic Resources, New Delhi
	985	University of Agricultural Sciences, Bangalore
	749	International Crop Research Institute for the Semi-Arid Tropics, Patancheru
	300	Vivekananda Parvatiya Krishi Anusandhan Sansthan, Almora
	1561	Indian Institute of Millets Research, Hyderabad
Japan	3671	National Institute of Agrobiological Sciences, Kannondai
	65	Plant Germplasm Institute, Kyoto University
United States	232	USDA Agricultural Research Service, Washington
	306	National Centre for Genetic Resources Conservation, Collins
	304	North Central Regional Plant Introduction Station, Ames
China	717	Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing
Kenya	208	National Gene bank of Kenya, Crop Plant Genetic Resources Centre, Muguga
Ethiopia	92	International Livestock Research Institute, Addis Ababa
Australia	66	Australian Plant Genetic Resource Information Service, Queensland
Pakistan	50	Plant Genetic Resources Program, Islamabad
Norway	44	Svalbard Gene Bank, Spitsbergen
United Kingdom	44	Millennium Seed Bank Project, Seed Conservation Department, Royal Botanic Gardens, London
Germany	36	Leibniz Institute of Plant Genetics and Crop Plant Research, Gatersleben

Source: http://apps3.fao.org/wiews/germplasm_report.jsp?

To obtain basic knowledge about germplasm, morphological characterization is the preliminary step for characterizing and classifying any collected/introduced materials. This not only provides the heritability of traits but also increases the utility of promising materials in breeding programs. In a study, Gupta et al. (2009) evaluated barnyard millet germplasm at the Himalayan regions of Uttarakhand, India, and identified some promising donors for plant height (<120 and >200 cm), productive tillers (>4), inflorescence length (>28 cm), raceme number (>50), raceme length (>3.1 cm), and grain yield (>16 g). In another study, to enable efficient use of genetic resources and to increase its access for breeders, barnyard millet core collection comprising of 89 accessions had also been established based on phenotypic and genotypic characterization (Upadhyaya et al., 2014). It was revealed that the mean difference between 89 core accessions and 736 whole accessions for most of the agronomic traits were not significant, indicating that the entire genetic variation had been sufficiently preserved in this core collection

(Upadhyaya et al., 2014). The coefficient of variation among core germplasm varied from 0.79 to 36.43% for days to maturity and basal tiller number, while the heritability (broad sense) varied between 70.14 and 99.87% for inflorescence length and days to maturity, respectively.

Further, the multidimensional principal component analysis (PCA)-based phenotypic characterization of these 89 accessions resulted in three different groups for agronomic and other phenotypic traits based on its origin (Sood et al., 2015). The study has also identified some promising genotypes, which could be efficiently used in a breeding program for the improvement of early maturity, grain yield, and yield contributing traits. Similarly, the IIMR, Hyderabad, evaluated the 146 barnyard millet accessions and found a larger variation for grain yield and yield contributing traits, which led to the identification of 18 promising accessions for barnyard millet breeding programs (IIMR, 2016). A comparison of agronomic traits from various trials conducted across India is given in **Supplementary Table S2**, which revealed barnyard millet genotypes to have considerable variation for yield and yield-related traits. For instance, the genotypes with higher grain yield and yield contributing traits (panicle length, number of raceme, and grain yield) were identified in the Southern States of India viz., Telangana and Tamil Nadu. In contrast, early maturing (58–90 days) genotypes were mostly found in the Northern States of India.

On the other hand, registration of trait-specific germplasms in the National Gene Banks (NGB) not only protects the natural resources from Intellectual Property Rights (IPRs) but also facilitates the breeders to access important/valuable genotypes for any crop improvement programs. In crop plants, 60% of the registered traits of germplasms belong to cereals, oilseeds, and legumes for resistance against various biotic and abiotic stresses. With regard to cereals, the maximum number of germplasms were registered in paddy and wheat mainly for biotic stress related traits (Radhamani et al., 2011). However, in millets except for sorghum and pearl millet, most of them were registered with limited traits only, mainly, cytoplasmic male sterile (CMS) line in foxtail millet (Radhamani et al., 2011), waxy trait in proso millet (Santra et al., 2015), and easy de-hulling in barnyard millet. The registered barnyard millet genotype B29 by VPKAS, Almora, showed a 42–146.4% faster de-hulling percentage over other check varieties (Gupta et al., 2014). Therefore, despite the focus on higher grain yields alone, barnyard millet breeding programs should also include the strategy of registration of unique traits that might be conserved in the landraces, germplasms, or rejected entries from the evaluation trials.

The successful utilization of barnyard millet genetic resources resulted in the release of more than 20 varieties and cultures across India (Gomashe, 2017). The first variety of K1 was developed by the pureline selection method from local landraces of Tenkasi, Tamil Nadu, India, released during 1970, which possesses an average state yield of 1,000 kg ha^{-1}. Later, several varieties were released against various pests and diseases across India through pureline selection from local landraces or exotic germplasm accessions. Among these, the notable variety PRJ 1, a direct selection from exotic collections of ICRISAT was released during 2003, by Vivekananda Institute of Hill Agriculture, Almora, Uttarakhand, India, possess a higher grain yield (2,500 kg ha^{-1}) with resistance against various smuts (Upadhyaya et al., 2008). Recently, MDU 1, a variety developed by Agricultural College and Research Institute, Tamil Nadu Agricultural University, Madurai, India, through pureline selection of local landrace of Tamil Nadu possesses the characteristic features of short duration (<100 days) and higher grain yield (2,500 kg ha^{-1}) (Vanniarajan et al., 2018). Besides, this variety also possesses a higher amount of iron content (16 mg/100 g) in the grains with good milling and cooking quality. In Japan, the "Noge-Hie," a low amylose grain-containing cultivar was identified from a local landrace possessing natural deletion in one of three waxy genes (Hoshino et al., 2010). At the same time, "Chojuromochi," a mutant developed through artificial γ-radiation, was completely devoid of Waxy (*Wx*) protein synthesis. In addition, the waxy protein trait was found to be stably inherited. Such glutinous variety in barnyard millet is in huge demand from Japanese consumers and industries for various food preparations similar to rice from paddy.

Hybridization is a difficult task in small millets, however, the hot water-based method followed by the contact method of crossing was found to be effective in finger millet (Raj et al., 1964; Nandini and Fakrudin, 1999) and foxtail millet (Siles et al., 2001). The same method was also applied in barnyard millet hybridization programs (Renganathan et al., 2015; Sood et al., 2015). Prior to pollination (early morning), the panicles which started flowering were selected for emasculation (Renganathan et al., 2015; Sood et al., 2015). The selected panicle was trimmed for 1/3rd portion by removing the opened and immature flowers in the respective upper and lower portions of the racemes, and then the remaining middle portion was immersed in hot water at 52°C in a thermos flask for 1 min. The emasculated panicles were then covered with butter paper bags to avoid contamination. The panicles in which flowering had already commenced were chosen as a pollinator source and panicle to panicle contacts were made by tying them together with a thread. The male and female panicles thus secured together were covered by a butter paper bag to avoid contamination with foreign pollen. However, priority should be given for the development of CMS line in barnyard millet for the better exploitation of variability as followed in foxtail millet.

GENOMIC RESOURCES AND UTILIZATION

Whole genome sequence (WGS) is fundamental to understand the genome composition and gene repertoire of a crop. It helps to identify important genes and pathways related to economically important traits in crops. Recent advances in second and third-generation sequencing technologies have facilitated simple and cost-effective sequencing platforms to generate genome and transcriptome sequences. Among millets, the whole genome sequencing was completed in sorghum, pearl millet, foxtail millet, finger millet, and proso millet by various researchers (Zhang et al., 2012; Mace et al., 2013; Hittalmani et al., 2017; Varshney et al., 2017; Zou et al., 2019). Genomic resources are also considerably

well-defined in sorghum, pearl millet, foxtail millet, and finger millet due to the presence of genetic linkage maps, physical maps, cytogenetic stocks, and large-insert libraries (as reviewed by Varshney et al., 2006; Gomashe, 2017). However, in barnyard millet, very limited attempts have been made to discover the genomic structure and associated downstream processes due to its genome complexity and lack of research funding on this orphan crop.

Chloroplast Genomes and Phylogeny Analysis

The chloroplast genome of *Echinochloa* is highly conserved for its genome structure, organization, and gene order (Ye et al., 2014). So far, chloroplast genomes of seven *Echinochloa* species including *E. crus-galli*, *E. ugandensis*, *E. stagnina*, *E. colona*, *E. esculenta*, *E. frumentacea*, and *E. oryzicola* (Ye et al., 2014; Perumal et al., 2016; Lee et al., 2017; Sebastin et al., 2019) have been sequenced. It revealed their genome structure (quadripartite), identity (99.5%), and size (139,592–139,851 bp) of the chloroplast genomes. The quadripartite structure generally confers a major impact on the evolution of plastome sequences of an organism (Yang et al., 2013). Such a quadripartite genome comprises a pair of inverted repeats (IR) separated by a small single-copy region (SSC) and a large-single copy region (LSC). The comparison of chloroplast genomes of wild and cultivated *Echinochloa* is given in **Table 3**. The size of IR, LSC, and SSC regions varied from 22,618 to 22,748, 81,837 to 82,053, and 12,518 to 12,519 bp, respectively. Similar to other angiosperms, the chloroplast genome of *Echinochloa* species comprises of 38.6% GC regions and 61% AT regions (Sebastin et al., 2019). In contrast, the number of genes varied from 111 to 131 among the species of *Echinochloa*, with the cultivated species exhibiting minimum. This could be mainly due to the reorganization of gene copy number and structure during the course of evolution or speciation. The divergence in copy number of any gene further creates the genetic polymorphism between the species, which contributes a major variation in their genome size and phenotype (Suryawanshi et al., 2016). As reported in angiosperms by Wendel (2015), the high morphological variation among the wild and its cultivated species occurs due to the consequences of genome reorganization during the evolutionary process. The comparative analysis of *Echinochloa* chloroplast genomes revealed that they are closer to the *Panicum virgatum* than other grasses (Ye et al., 2014). Further, molecular divergence clock analysis of grass species revealed that *Echinochloa* species had diverged 21.6 Mya than others. The wild species, such as *E. oryzicola* and *E. crus-galli*, diverged around 3.3 Mya while another wild species, *E. colona*, diverged from *E. oryzicola* and *E. crus-galli* between 2.65 and 3.18 Mya, respectively (Lee et al., 2017). Later, the cultivated species (*E. frumentacea*) diverged from *E. oryzicola* and *E. crus-galli* in and around 1.9–2.7 Mya (Perumal et al., 2016) (**Figure 5**). However, both wild and cultivated *Echinochloa* species have a high sequence identity with *P. virgatum* and *Sorghum bicolor* and low sequence identity with *Triticum aestivum* and *Oryza sativa* (Ye et al., 2014), which further concludes that *Echinochloa* species are more closely related to the *P. virgatum* and *S. bicolor* than *Triticum* and *Oryza.*

TABLE 3 | Summary statistics of chloroplast genomes for wild and cultivated barnyard millet.

Genome composition	*E. crus-galli*[a]	*E. colona*[b]	*E. esculenta*[c]	*E. frumentacea*[d]
Genome size (bp)	139,851	139,592	139,851	139,593
Inverted Repeat size (bp)	22,640	22,618	22,748	22,618
Large-single copy region size (bp)	82,053	81,837	81,837	81,839
Small-single copy region size (bp)	12,518	12,519	12,518	12,518
Number of genes	131	131	111	112
Number of Protein coding genes	88	88	76	77
Number of tRNA	40	40	30	30
Number of rRNA	4	4	4	4
GC contents (%)	38.6	38.6	38.6	38.6

[a]*Ye et al., 2014;* [b]*Lee et al., 2017;* [c]*Sebastin et al., 2019;* [d]*Perumal et al., 2016.*

Transcriptome Analysis

With the advent of next-generation sequencing technologies (NGS), RNA-seq (RNA-sequencing) has now superseded the previous microarray technologies and a huge number of genomic resources are being generated in a cost and time effective manner (Weber, 2015). It not only generates differential genes, but also the functional molecular markers like simple sequence repeats (SSRs) and SNPs in various minor millet species. Enormous transcript profiles have been developed in weedy *Echinochloa* species till date for various traits associated with invasiveness and adaptations such as herbicide resistance, photosynthesis, flooding response, and other homeobox genes (Li et al., 2013a,b; Yang et al., 2013; Nah et al., 2015; Xu et al., 2015; Guo et al., 2017; Gao et al., 2018). Recently, transcriptome sequences developed from cultivated *E. frumentacea* variety CO (Kv) 2, yielded 97,065 transcripts with an average length of 94 Mbp (Murukarthick et al., 2019). Further *de novo* assembly, functional annotation, and comparison to *E. crus-galli* transcripts identified some key genes regulating Fe and Zn accumulation and drought tolerance. In addition, the study also generated 300 SSR primer pairs from 10,881 SSR loci targeting major repeats of trinucleotide (122) followed by dinucleotide (121), tetra-nucleotide (35), penta-nucleotide (20), and hexa-nucleotide (2).

A review of transcriptomes data in the NCBI database has revealed the presence of 952 gene sequences to date, generated from *E. crus-galli* (170), *E. oryzicola* (132), *E. frumentacea* (130), *E. esculenta* (130), *E. colona* (130), *E. ugandensis* (131), and *E. stagnina* (129). The details of transcriptome sequences published in the NCBI database were consolidated and presented in **Table 4**. Most of the genes were related to photosynthesis (PS I, PS II, NADH-plastoquinone *oxidoreductase*, ATP *synthase*), C4 pathways (phosphoenolpyruvate *carboxylase*, *aldolase*, *maturase* K, *kinase*), micronutrient transportation [Fe^{2+} transport protein 2-like protein (*IRT2*) gene, *nicotianamine synthase* 1 (*NAS1*), *nicotianamine synthase* 2 (*NAS2*), *polymerases* (RNA, DNA)],

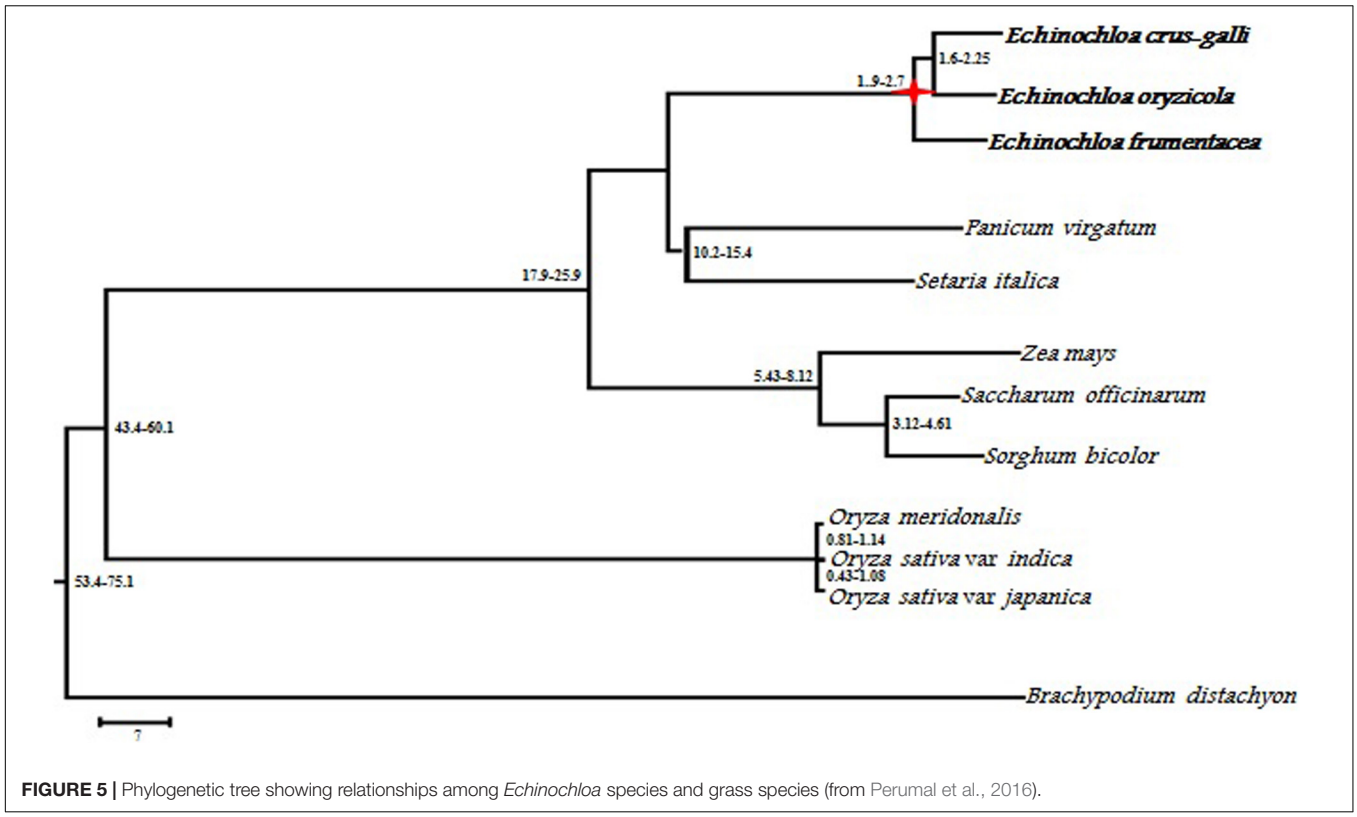


FIGURE 5 | Phylogenetic tree showing relationships among *Echinochloa* species and grass species (from Perumal et al., 2016).

herbicide resistances (*1-aminocyclopropane-1-carboxylic acid synthase* 3, *acetolactate synthase*, calcineurin, cyclophilin 2, cytochrome P450, GH31, *glutathione S-transferase*), flooding tolerances (*enolase*, *alcohol dehydrogenase*), waxy grains (*granule-bound starch synthase*), non-shattering grains (sh_4), ribosomal RNA, and transfer RNAs, etc. The proteomics exploration also revealed that a total of 540 proteins are found to be commonly expressed in *Echinochloa* species, of which most of the annotated protein sequences are tRNA, ribosomal, and other photosystem proteins. In addition, most of the proteins found in the *Echinochloa* species showed orthologs among themselves for proteins of C_4 pathways, calcium binding protein, photosynthesis, *bZIP* transcription factor 1, translational initiation factors, transporters, and hypothetical proteins, etc. (Yang et al., 2013). However, some uniquely expressing proteins were also identified in the *Echinochloa* species. For instance, the maximum expression of quinclorac-resistant proteins, Cu/Zn superoxide *dismutase*, defensin, cadmium tolerant, viral nucleoprotein, and antimicrobial peptides was observed in *E. crus-galli* (Odintsova et al., 2008), multiple-herbicide-resistant proteins in *E. phyllopogan* (Iwakami et al., 2014), and *granule-bound starch synthase* in *E. esculenta* (Ishikawa et al., 2013). However, we require more research/data to draw a valid conclusion on the species-specific expression.

TABLE 4 | Details of transcriptome sequences of *Echinochloa* species.

Species	Parts	Data generated (Gb)	Platform	Accession number
E. frumentacea	Leaf	18.3	ILLUMINA (Illumina HiSeq 2500)	SRX5210765
E. frumentacea	Leaf	48.6	ILLUMINA (NextSeq 500)	SRX3029505
E. esculenta	NA	11.6	ILLUMINA (Illumina HiSeq 2000)	SRX2698648
E. crus-galli var. *zelayensis*	Leaf	07.3	ILLUMINA (Illumina HiSeq 2000)	SRX3574154
E. crus-galli	NA	32.8	PacBio	SRX3081138
*E. colona**	Leaf	11.6	ILLUMINA (Illumina HiSeq 2500)	SRX2588690
E. stagnina	Leaf	02.4	ILLUMINA (Illumina HiSeq 2500)	SRX3330321
E. stagnina	Roots	01.5	ILLUMINA (Illumina HiSeq 2500)	SRX3330365
E. glabrescens	NA	05.0	ILLUMINA (NextSeq 2000)	ERX990971
Barnyard millet core collection	NA	69.0	ILLUMINA (Illumina HiSeq 2500)	SRX734221

**Imazamox (herbicide) treated; NA, not available. Source: https://www.ncbi.nlm.nih.gov/search/all/?term=echinochloa.*

Genome Sequence

Research conducted in China released the whole genome sequence of weedy *E. crus-galli* during 2017 and was annotated successfully for its unique nature of invasiveness and adaptation in the fields of crop plants (Guo et al., 2017). The total sequence length of the genome at a depth of 171× was estimated to the size of 1.27 Gb, representing around 90.7% of the predicted genome size. The genomic libraries range between[2] 160 bp and 20 Kb with a total number of contigs of 4534 with minimum and maximum contigs size of 1 kb and 11.7 Mb, respectively. The gene annotation of *E. crus-galli* further revealed 108,771 protein-coding genes, 785 miRNAs, 514 Mb repetitive elements, and non-coding RNAs. As of 2019, the genomic resources available in the NCBI domain include 1,246 nucleotide sequences, 822 gene sequences, 2,468 protein sequences, 105 short read archive (SRA) sequences, 74 Expressed Sequence Tags (ESTs), and one Gene Expression Omnibus (GEO) dataset related to various species of *Echinochloa*. Among the species *E. crus-galli* (652), *E. oryzicola* (126), *E. phyllopogon* (96), *E. colona* (76), *E. pyramidalis* (46), *E. esculenta* (44), *E. frumentacea* (43), and *E. oryzoides* (32) hold the maximum number of sequences (**Figure 3B**). To date, 54% of nucleotide sequences are available for *E. crus-galli*, while cultivated barnyard millet *E. frumentacea* and *E. esculenta* have only 4%. The comparative scenario of genomic resources among small millets is presented in **Supplementary Table S3**, which further emphasized the need to enrich the cultivated barnyard millet genome in the future.

Molecular Markers and Its Application

Genetic Diversity Analysis

Molecular markers are nucleotide sequences that are widely used for genetic diversity, linkage map construction, and marker assisted selection of crop plants (Muthamilarasan and Prasad, 2014). Early in the period of molecular marker research, Random Amplified Polymorphic DNA (RAPD) markers were utilized to access the genetic diversity and phylogeny among *Echinochloa* species (Hilu, 1994). Hilu (1994) proving that RAPD markers are effective in distinguishing both the cultivated and wild progenitors of the *Echinochloa* species at the genomic level. The genetic diversity among *E. frumentacea* was also found to be more diverse than *E. utilis* populations. However, isozyme marker analysis between these two species revealed that the accessions within the same species formed two different clusters and accessions from different species grouped into the same cluster, creating the possibility of the existence of intergrades and overlaps between the species (Prabha et al., 2010). Previously, Rutledge et al. (2000) obtained 90 polymorphic bands using 21 primer pairs with an average of 4.3 alleles per primer and Ruiz-Santaella et al. (2006) obtained 75 polymorphic bands using 13 primer pairs with an average of 5.8 alleles per primer. This suggests a low exhibition of polymorphism in the germplasm by the RAPD markers. The low level of polymorphism using RAPD markers has also been previously reported in finger millet diversity studies (Muza et al., 1995). Notwithstanding, the Amplified fragment length polymorphism (AFLP) marker system later developed a higher ability in revealing the genetic diversity in *Echinochloa* species compared to RAPD markers (Danquah et al., 2002; Tabacchi et al., 2009), since it generates more alleles per primer. For instance, a total of 166 polymorphic bands were produced in four primer pairs with an average of 41.5 per primer pairs in 28 genotypes. Whereas, the polymorphism information content (PIC) value of markers ranged between 0.44 and 0.52 (Kaya et al., 2014). This was in accordance with the previous report of Tabacchi et al. (2009), where seven primer pairs produced 156 polymorphic bands with an average of 22.3 alleles per primer in 80 genotypes. Similarly, polymerase chain reaction–restriction fragment length polymorphism (PCR-RFLP) techniques were successfully applied for species identification in *E. oryzicola* and *E. crus-galli* (Yamaguchi et al., 2005; Yasuda et al., 2006). Recently, InDel markers for *rbcL*, *matK*, and *ITS* genes identified in *E. colona* have been widely used as a cost effective approach in the DNA barcoding of *E. colona*, *E. oryzicola*, and *E. crus-galli* (Lee et al., 2014). All these studies emphasized that with insufficient sequence information, the RAPD and isozymes markers are the only choice, not only helpful in differentiating *Echinochloa* species, but also in laying the foundation for molecular breeding in barnyard millet.

Later on, a significant development in the sequencing technologies further eliminated the limitations present in the RAPD, RFLP, and AFLP techniques through sequence-based markers such as SSRs, EST-SSRs (Expressed sequence tags-simple sequence repeats), and SNPs. The sequenced-based markers are more desirable in genetic diversity studies due to their co-dominant, reproducible, highly polymorphic, and effective utilization in many crop plants (Lin et al., 2011). The information available on sequence-based markers in barnyard millet is still in its infancy, despite the reports of microsatellite markers related to genetic diversity studies in germplasm accessions gaining attention today. For instance, utilizing five SSR markers, 155 accessions of *Echinochloa* species including *E. esculenta* (49), *E. crus-galli* (94), and *E. esculenta* var. *formosensis* (12) were grouped into three separate clusters (Nozawa et al., 2006). The same study reported that the accessions belonging to *E. esculenta* were less diverse than those of *E. crus-galli* or *E. esculenta* var. *formosensis*. More recently, the ESTs markers also been proven to be a very informative and effective tool for the analysis of genetic diversity in many small millets. Extensive transcriptomics and annotation studies previously performed on herbicide resistant varieties of *E. crus-galli* resulted in 74 ESTs (Li et al., 2013b; Yang et al., 2013). However, those ESTs were limitedly used in the marker development and diversity studies in barnyard millet, since they are weedy ancestors. For instance, the *in silico* mining of *E. crus-galli* ESTs resulted in the identification of 22 pairs of EST-SSR primers (Babu and Chauhan, 2017). The study also reported that frequent SSR repeats were found to be tetra-nucleotide repeat followed by the penta- and hexa- nucleotide repeats. Among the repeats, GT (dimer), AGG and AGA (trimer), CAAA (tetra), TGTTT

[2]https://www.ncbi.nlm.nih.gov/sra?

(penta), and AGACGA (hexa) were the most common type of repeat motifs in barnyard millet. On the other hand, a restriction-site associated DNA (RAD) approach combined with Illumina DNA sequencing strategy was performed in *E. phyllopogon* for the rapid and mass discovery of SSR and SNP markers by Chen et al. (2017). This study yields sequencing reads of 4132 contigs, of which 4710 are annotated to be putative SSRs and 49,179 are probable SNPs. Out of 4710 putative SSR markers, 78 were potentially polymorphic. Besides, the most frequent motif was AT and maximum motif length was dinucleotide type (>82%) followed by tri, tetra, penta, and hexa. The further validation of eight SSRs in four *E. phyllopogon* population resulted in 66 alleles with an average of 3.1–4.8 alleles from locus per population. Moreover, the study also identified a higher percentage of GC (48.9%) content in the genome, proving their successful nature of adaptation against freezing and desiccation, with GC% indicating more stability in an organism. Hence, SSRs and SNPs markers developed from *E. phyllopogon* may be very useful in studying not only the diversity, origin, and distribution of herbicides-resistant population (Osuna et al., 2011; Okada et al., 2013), but also for predicting gene location and molecular breeding in cultivated types. Recently, Manimekalai et al. (2018) and Murukarthick et al. (2019) used EST-SSR markers developed from the cultivated, *E. frumentacea* transcriptome sequence to analyze the genetic diversity of Indian barnyard millet germplasm. Manimekalai et al. (2018) used 51 EST-SSR markers to study the genetic diversity of 61 barnyard millet germplasms. Among 51 EST-SSR markers, 14 were polymorphic and produced 29 alleles with the PIC value ranging between 0.276 and 0.652. Similarly, Murukarthick et al. (2019) identified 10 polymorphic markers from 30 EST-SSRs and showed clear polymorphism in the 30 Indian barnyard millet germplasms.

Apart from SSR markers, a total of 21,000 SNPs were identified and characterized through the whole-genome genotyping-by-sequencing (GBS) method using core germplasm comprising of 95 barnyard millet accession (Wallace et al., 2015). About 10,816 out of 21,000 SNPs were spread across 65 biotypes of *E. colona*, and 8,217 SNPs across 22 biotypes of *E. crus-galli*. The SNPs discriminating among *E. colona* and *E. crus-galli* biotypes were 1,299 and 1,444, respectively. Further, population structure analysis with SNPs strongly separated these two species with four clusters in *E. colona* and three clusters in *E. crus-galli* (Wallace et al., 2015).

Gene/QTL Mapping

The use of molecular markers such as SSRs and SNPs provide opportunities for breeders to identify the QTL/gene(s) for important micronutrient and agronomical traits in barnyard millet. So far, many SSR and SNP markers (Wallace et al., 2015; Chen et al., 2017; Manimekalai et al., 2018; Murukarthick et al., 2019) have been developed to speed up the linkage map construction and QTL mapping in barnyard millet, but no genetic linkage map or QTLs published yet compares to other millets such as foxtail millet and finger millet (**Supplementary Table S4**). To date, two mapping studies only have been published on barnyard millet (**Table 5**). Ishikawa et al. (2013) identified functional SNP markers for waxy traits and found that these waxy traits are controlled by three loci, namely *EeWx1*, *EeWx1*, and *EeWx3*. The plants with functional alleles in all three loci exhibited normal amylose content (wild), while any one of the alleles (natural mutant) and/or completely homozygous mutant alleles (artificial mutant) exhibited low amylose and very low amylose content (waxy), respectively. Besides, the phylogenetic analysis also revealed that the waxy gene sequences are highly conserved among grass species. In another study, bulk segregant analysis (BSA) and 51 EST-SSR markers were used to analyze the F_2 individuals of ACM 331 × MA 10, contrast parents for anthocyanin pigments, and the results showed that the SSR marker, BMESSR 39, was linked with anthocyanin pigments in barnyard millet (Renganathan et al., 2019). Conclusively, progress in barnyard millet genome mapping remains slow and is still in its initial stage. Moreover, these published two reports contain preliminary results only; further experimental investigation is required to apply for marker-assisted selection (MAS).

Comparative Genomics and Synteny Analyses

Comparative genomics studies brought considerable benefit to barnyard millet crop. Mainly, SSR markers obtained from the cereals and millets were successfully utilized to characterize the barnyard millet germplasm. The summary of cross transferable molecular markers developed in other cereals and millet are presented in **Table 6**. Due to the non-availability of whole genome sequencing in barnyard millet, the genomes of rice, maize, finger millet, and foxtail millet have served as essential models to study the marker-based syntenic relationships. The genomic SSR (gSSR) markers developed through *in silico* mining of the foxtail millet sequence showed a high degree of cross-transferability in barnyard millet and other related small millet species. Among the 159 gSSRs, 58 were found to show consistent amplification in barnyard millet, that is, 91.3% cross-species amplification ability (Pandey et al., 2013). Similarly, 106 eSSR

TABLE 5 | Molecular markers associated with waxy and anthocyanin pigment traits in barnyard millet.

Markers	Trait	F_2 Mapping population	Segregation ratio (χ^2 analysis)	References
SNP	Waxy	Wild × Mutant (γ) and Natural mutant × Mutant (γ)	3:1 and 63:1	Ishikawa et al., 2013
ESTSSR	Anthocyanin pigment	Pigmented × Green	1:2:1	Renganathan et al., 2019

SNP, single nucleotide polymorphism; ESTSSR, EST-derived simple sequence repeat, EMS, ethyl methane sulfonate.

TABLE 6 | Transferability details of cross cereal species markers to barnyard millet.

Species	Marker	Number of amplified markers	PIC	Polymorphic marker	Cross-transferability (%)	References
Finger millet	eSSR	15/104	NA	NA	14.2	Arya et al., 2014
	gSSR	20/132	NA	NA	15.5	Arya et al., 2014
	eSSR, gSSR	33/101	NA	NA	55.4	Krishna et al., 2018
	SSR	07/18	0.16–0.53	06	85.7	Babu et al., 2018a
	SSR	32/32	0.27–0.73	29	90.6	Babu et al., 2018b
Pearl millet	SSR	10/32	NA	NA	31.2	Arya et al., 2014
Foxtail	SSR	53/58	NA	NA	91.3	Pandey et al., 2013
	eSSR	106	0.00–0.48	NA	90.6	Kumari et al., 2013
	ILP	94/100	0.03–0.47	NA	94.1	Muthamilarasan and Prasad, 2014
	SSR	46/64	0.02–0.66	NA	73.4	Gupta et al., 2013
	SSR	7/26	NA	2	65.38	Krishna et al., 2018
Sorghum	eSSR	42	NA	NA	80.9	Yadav et al., 2014
Rice	eSSR	102	NA	NA	72.1	Yadav et al., 2014
	SSR	85/120	0.15–0.67	41	48.2	Babu et al., 2018b
Maize	SSR	32/46	0.25–0.73	26	70.0	Babu et al., 2018a

SSR, simple sequence repeat; eSSR, EST-derived simple sequence repeat; gSSR, genomic simple sequence repeat; ILP, intron length polymorphism; PIC, polymorphic information content; NA, not available.

(EST-derived simple sequence repeats) markers from *Setaria* showed consistent amplification in millet and non-millet species and also exhibited high cross species transferability in barnyard millet (90.6%) (Kumari et al., 2013). Muthamilarasan and Prasad (2014) reported that 100 out of 5,000 intron-length polymorphic markers (ILP) mined from the foxtail millet genome showed amplification in various small millets with 94 percentage cross-transferability in barnyard millet. Yadav et al. (2014), also found that the rice genic SSR primers from calcium transporters and calcium kinases group showed 100% and 72.2% cross transferability, respectively, in barnyard millet.

The orthologs and paralogs analyses of the genome *E. crus-galli* against some of the grass family revealed that the approximate divergence times of *Oryza* – *Sorghum* and *Sorghum* – *Echinochloa* were estimated to be 48.5 and 28.5 Mya, respectively, followed by the polyploidization and speciation events by 7.8 Mya (Guo et al., 2017). Three copies of gene clusters related to the biosynthesis of DIMBOA (2,4-dihydroxy-7-methoxy-1,4-benzoxazin-3-one), a unique allelopathic compound reported in maize (Frey et al., 2009), were found in the *E. crus-galli* genome and each of them showed a perfect synteny with segments of BX1-5 and BX8 of the maize genome. Similarly, *E. crus-galli* also exhibited synteny with rice for momilactones, a phytoalexin compound expressed to protect against blast pathogens (Guo et al., 2017). Babu and Chauhan (2017) also found homology of some barnyard millet ESTs against the chromosomal regions of 2, 5, 6, 8, 9, and 12 of rice, the waxy gene of maize, *granule-bound starch synthase* I (GBSSI-S) gene of *Panicum repens*, *Setaria italica*, *Panicum miliaceum*, and super oxide *dismutase* (SOD) gene of *Colletotrichum eremochloae*. On the other hand, Babu et al. (2018a; 2018b) compared rice, maize, and finger millet gSSRs for cross species amplification in barnyard millet and reported that maize and finger millet SSRs exhibited higher PIC values, efficient cross species amplification, and polymorphism percentage than rice SSRs. However, the comparative genetic mapping between rice and barnyard millet showed several putative syntenic regions across the genome that regulated the traits including seed dormancy, plant height, panicle length, spikelet characters, leaf senescence, seed weight/yield-related traits, shattering character, root traits, blast resistance, brown plant hopper (BPH) resistance, and amylose content (Babu et al., 2018b). Eventually, using the available literature in the published reports, we concluded that EST-derived SSR markers had higher cross-genome amplification than genomic SSR markers, indicating higher conservation of the former than the latter across the species of the grass family. Therefore, cross transferability mechanisms could be exploited in barnyard millet for trait-based marker identification.

CONCLUSION AND FUTURE PROSPECTS

Despite its nutritional and agronomic benefits, barnyard millet has remained an underutilized crop and has received very little attention from researchers as well as farmers across the globe. Barnyard millet breeding programs have stagnated due to limited funding from various funding agencies and research organizations. Therefore, considerable efforts are needed to develop varieties or hybrids with farmer/consumer preferred traits. More breeding programs have to be designed in the future for harnessing the genetic variability for high yield potential, yield stability, improved salinity tolerance, pest and disease resistance, as well as enhanced nutritional quality, especially micronutrient composition. However, the progress of barnyard millet breeding programs is very slow due to the lack of genetic and genomic resources. With respect to genetic resources, the size of the core collection in barnyard millet is comparatively less than that of other minor millets (foxtail and finger millet) and so far, breeding

populations have not been developed. Therefore, core and mini-core collections representing maximum diversity as well as bi-parental and multiparent populations have to be established and evaluated for various nutritional and agronomic traits. These resources will be useful to track the genomic regions associated with targeted traits by the linkage-based QTL mapping, genome-wide association study (GWAS), and genomic selection (GS), as well as for the detection of candidate genes.

Despite genome research in barnyard millet being at its infancy and far behind other minor millets, transcriptome sequencing has allowed researchers to develop several genomic resources, including EST-SSRs and SNPs, that could be useful for marker-assisted breeding. However, extensive efforts are needed in the future to develop the reference genome, genome-wide SSR and SNP markers, construction of genetic linkage maps, and physical maps. The recent release of the genome sequence of a weedy ancestor (*E. crus-galli*), together with the genomic resources from major and minor millet crops, offers an initial framework for enriching genomic research in an orphan crop like barnyard millet by comparative genomic approaches. It is also fruitful to use the *E. crus-galli* genome as a reference genome for cultivated barnyard millet species similar to the case in bread wheat. It helps not only to understand the genome composition of cultivated barnyard millet species and increases mapping accuracy, but also helps us to know the effect of variants on protein function.

Barnyard millet is a potential crop for the biofortification of micronutrients. The grains of barnyard millet are rich in micronutrients (Fe and Zn) and hence, the identification of potential genes related to the accumulation of micronutrients (Fe and Zn) will be helpful to transfer these genes to high yielding barnyard millet cultivars or even to other major crops like rice, wheat, maize, etc. Strategies for the improvement of micronutrients in barnyard millet are presented in **Figure 6**, which are also applicable to other agronomic traits. Furthermore, barnyard millet is well adapted to both warm and temperate regions and it is a rich source of genes responsible for stress tolerance. Therefore, understanding the molecular mechanism of plant responses to stress in inherently stress-tolerant crops such as barnyard millet will be useful in developing highly stress-tolerant cultivars. So far, several stress tolerance genes were identified in barnyard millet, but the function of these genes has not been tested by overexpression studies, mainly due to the lack of a genetic transformation system. To date, very limited reports have been published on genetic transformation in barnyard millet. Therefore, there is an immense need to develop an efficient transformation system for barnyard millet in the future so that it also paves the way for functional genomics studies related to tolerance against biotic and abiotic stresses as well as micronutrient traits.

Besides these research gaps, the farming community is still unaware of the true potential of barnyard millet cultivation in terms of nutritional value and productivity. Farmers generally cultivate this crop under marginal areas, but they still depend on low yielding local landraces. Therefore, support from non-government organizations (NGOs) can help in increasing awareness among the farmers, stakeholders, nutritionists, and consumers to adopt and promote barnyard millet cultivation as

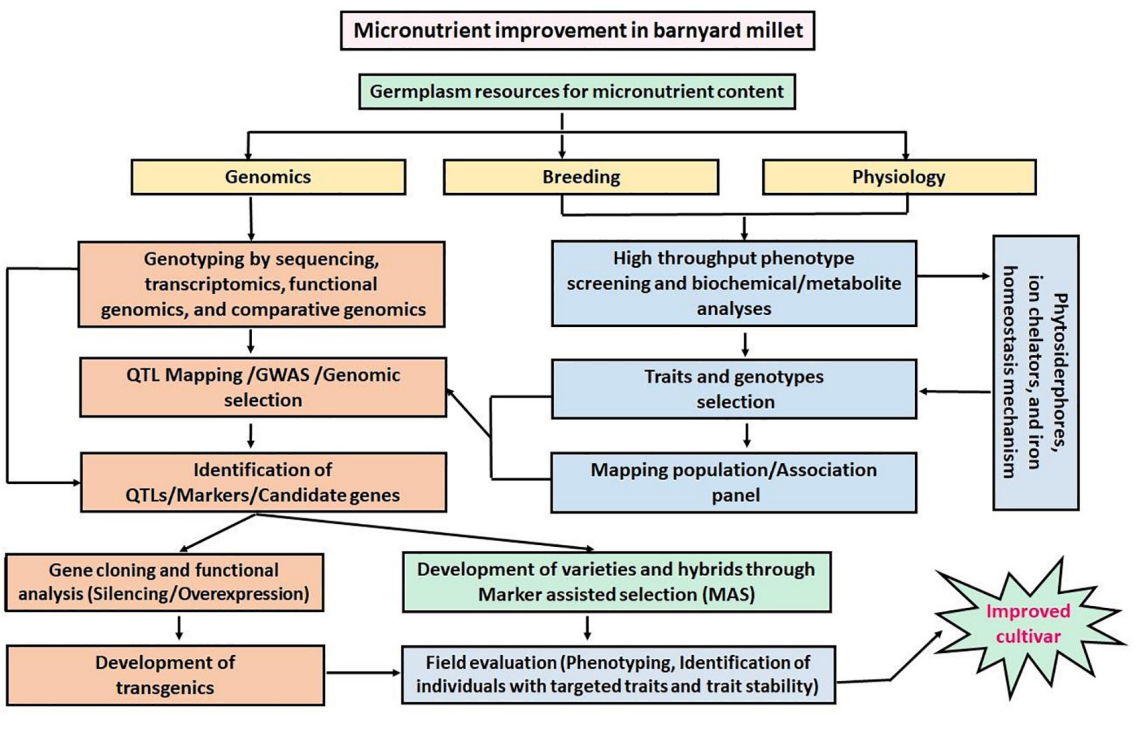


FIGURE 6 | Schematic representation of the proposed strategy for micronutrient improvements in barnyard millet.

well as consumption. Moreover, being a polyploid, ratooning (or) multi-cutting practices have to be standardized, like in sorghum, for better utilization of the growing season for grain and green fodder production. There is also an urgent need for advancements in post-harvest technologies for better processing and value-addition in the barnyard and other minor millets. At the same time, a change in consumer preference toward small millets with simultaneous development of suitable food products, along with an increase in market price, would fetch better returns for farmers and healthier choices for consumers. Finally, when these challenges are overcome, barnyard millet, being nutritionally sound and environmentally hardy, is going to be a promising crop for sustainable food and nutritional security in future climate scenarios.

AUTHOR CONTRIBUTIONS

VR built the layout of the manuscript, collected literature, and wrote the manuscript. CV, JR, and AK provided suggestions. VR and AK revised and prepared the final draft of the manuscript.

ACKNOWLEDGMENTS

This work was supported by the Centre of Excellence – Post Doctoral program, Tamil Nadu Agricultural University to VR. All the authors would like to extend their sincere appreciation to Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu and India. We are also thankful to editors and reviewers for their detailed suggestions in improving the manuscript.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00500/full#supplementary-material

REFERENCES

Ajaib, M., Khan, K. M., Perveen, S., and Shah, S. (2013). Antimicrobial and antioxidant activities of *Echinochloa colona* (Linn.) link and *Sporobolus coromandelianus* (Retz.) kunth. *J. Chem. Soc. Pakistan.* 35, 1384–1398.

Altop, K. E., and Mennan, H. (2011). Genetic and morphologic diversity of *Echinochloa crus*-galli populations from different origins. *Phytoparasitica* 39, 93–102. doi: 10.1007/s12600-010-0135-3

Anju, T., and Sarita, S. (2010). Suitability of foxtail millet (*Setaria italica*) and barnyard millet (*Echinochloa frumentacea*) for development of low glycemic index biscuits. *Malays. J. Nutr.* 16, 361–368.

Aoki, D., and Yamaguchi, H. (2008). Genetic relationship between Echinochloa crus-galli and Echinochloa oryzicola accessions inferred from internal transcribed spacer and chloroplast DNA sequences. *Weed Biol. Manag.* 8, 233–242. doi: 10.1111/j.1445-6664.2008.00303.x

Arthi, N., Rajagopal, B., Geethanjali, S., Nirmalakumari, A., and Natesan, S. (2019). Screening of barnyard millet (*Echinochloa frumentacea*) germplasm for salinity tolerance. *Electron. J. Plant Breed.* 10, 659–666. doi: 10.5958/0975-928X.2019.00083.8

Arya, L., Chauhan, D., Yadav, Y., and Verma, M. (2014). "Transferability of simple sequence repeat (SSR) markers developed in Finger Millet, and Pearl Millet to Kodo Millet and barnyard Millet," in *Innovative Approach in Stem Cell Research, Cancer Biology and Applied Biotechnology*, eds A. K. Johri, and G. C. Mishra, (New Delhi: Excellent Publishing House)60–64.

Babu, B. K., Agrawal, P. K., Pandey, D., Jaiswal, J. P., and Kumar, A. (2014a). Association mapping of agro-morphological characters among the global collection of finger millet genotypes using genomic SSR markers. *Mol. Biol. Rep.* 41, 5287–5297. doi: 10.1007/s11033-014-3400-6

Babu, B. K., Agrawal, P., Pandey, D., and Kumar, A. (2014b). Comparative genomics and association mapping approaches for opaque2 modifier genes in finger millet accessions using genic, genomic and candidate gene-based simple sequence repeat markers. *Mol. Breed.* 34, 1261–1279. doi: 10.1007/s11032-014-0115-2

Babu, B. K., Dinesh, P., Agrawal, P. K., Sood, S., Chandrashekara, C., Bhatt, J. C., et al. (2014c). Comparative genomics and association mapping approaches for blast resistant genes in finger millet using SSRs. *PLoS One* 9:e99182. doi: 10.1371/journal.pone.0099182

Babu, B. K., Pandey, D., Agrawal, P. K., Sood, S., and Kumar, A. (2014d). In-silico mining, type and frequency analysis of genic microsatellites of finger millet (Eleusine coracana (L.) Gaertn.): a comparative genomic analysis of NBS–LRR regions of finger millet with rice. *Mol. Biol. Rep.* 41, 3081–3090. doi: 10.1007/s11033-014-3168-8

Babu, B., and Chauhan, R. (2017). In-Silico Identification of EST Based Microsatellite Markers and SNPs, and Comparative Genomic Analysis of ESTs in Barnyard Millet for their Omics Applications. *Curr. Agric. Res. J.* 5, 279–287. doi: 10.12944/CARJ.5.3.03

Babu, B., Rashmi, C., and Sood, S. (2018a). Cross transferability of finger millet and maize genomic SSR markers for genetic diversity and population structure analysis of barnyard millet. *Indian J. Genet. Plant Breed.* 78, 364–372. doi: 10.31742/IJGPB.78.3.5

Babu, B., Sood, S., Kumar, D., Joshi, A., Pattanayak, A., Kant, L., et al. (2018b). Cross-genera transferability of rice and finger millet genomic SSRs to barnyard millet (*Echinochloa* spp.). *3 Biotech* 8:95. doi: 10.1007/s13205-018-1118-1

Babu, B. K., Joshi, A., Sood, S., and Agrawal, P. K. (2017). Identification of microsatellite markers for finger millet genomics application through cross transferability of rice genomic SSR markers. *Indian J. Genet.* 77, 92–98.

Bai, H., Cao, Y., Quan, J., Dong, L., Li, Z., Zhu, Y., et al. (2013). Identifying the genome-wide sequence variations and developing new molecular markers for genetics research by re-sequencing a landrace cultivar of foxtail millet. *PLoS One* 8:e73514. doi: 10.1371/journal.pone.0073514

Bajwa, A., Jabran, K., Shahid, M., Ali, H. H., Chauhan, B., and Ehsanullah. (2015). Eco-biology and management of *Echinochloa crus*-galli. *Crop Prot.* 75, 151–162. doi: 10.1016/j.cropro.2015.06.001

Bandyopadhyay, B. B. (1999). Genotypic differences in relation to climatic adaptation of two cultivated barnyard millet at Garhwal hills. *Indian J. Genet.* 59, 105–108.

Bandyopadhyay, B. B. (2009). Evaluation of barnyard millet cultivars for fodder yield under single and double cut treatments at higher elevation of hills. *Agric. Sci. Dig.* 29, 66–68.

Barrett, S. H. (1983). Crop mimicry in weeds. *Econ. Bot.* 37, 255–282. doi: 10.1007/bf02858881

Bennetzen, J. L., Schmutz, J., Wang, H., Percifield, R., Hawkins, J., Pontaroli, A. C., et al. (2012). Reference genome sequence of the model plant Setaria. *Nat. Biotechnol.* 30, 555–561. doi: 10.1038/nbt.2196

Borkar, V. S., Senthil Kumaran, K., Senthil Kumar, K. L., Gangurde, H. H., and Chordiya, M. A. (2016). Ethno medical properties of *Echinochloa colona* and *Hydrolea zeylanica*: a review. *World J. Pharmaceut. Res.* 5, 354–360.

Cao, T. L., Yang, S. Y., and Song, K. B. (2017). Characterization of barnyard millet starch films containing borage seed oil. *Coatings* 7:183. doi: 10.3390/coatings7110183

Chandel, G., Meena, R., Dubey, M., and Kumar, M. (2014). Nutritional properties of minor millets: neglected cereals with potentials to combat malnutrition. *Curr. Sci.* 107, 1109–1111.

Chauhan, B., and Johnson, D. E. (2011). Ecological studies on Echinochloa crus-galli and the implications for weed management in direct-seeded rice. *Crop Prot.* 30, 1385–1391. doi: 10.1016/j.cropro.2011.07.013

Chen, G., Zhang, W., Fang, J., and Dong, L. (2017). Identification of massive molecular markers in *Echinochloa phyllopogon* using a restriction-site associated DNA approach. *Plant Divers.* 39, 287–293. doi: 10.1016/j.pld.2017.08.004

Cho, Y. I., Chung, J. W., Lee, G. A., Ma, K. H., Dixit, A., Gwag, J. G., et al. (2010). Development and characterization of twenty-five new polymorphic microsatellite markers in proso millet (*Panicum miliaceum* L.). *Genes Genomics* 32, 267–273. doi: 10.1007/s13258-010-0007-8

Cidade, F. W., de Souza-Chies, T. T., Batista, L. A. R., Dall'agnol, M., Zucchi, M. I., Jungmann, L., et al. (2009). Isolation and characterization of microsatellite loci in Paspalum notatum Flüggé (Poaceae). *Conserv. Genet.* 10, 1977–1980. doi: 10.1007/s10592-009-9872-6

Cidade, F. W., de Souza-Chies, T. T., Souza, F. H. D., Batista, L. A. R., Agnol, M. D., Valls, J. F. M., et al. (2010). Microsatellite loci for Paspalum atratum (Poaceae) and cross-amplification in other species. *Am. J. Bot.* 97, 107–110. doi: 10.3732/ajb.1000207

Cidade, F. W., Vigna, B. B. Z., de Souza, F. H. D., Valls, J. F. M., Dall'Agnol, M., Zucchi, M. I., et al. (2013). Genetic variation in polyploid forage grass: assessing the molecular genetic variability in the *Paspalum* genus. *BMC Genet.* 14:50. doi: 10.1186/1471-2156-14-50

Clayton, W. D., and Renvoize, S. A. (2006). *Genera Graminum: Grasses of the world in Kew Bulletin Additional Series XIII, Royal Botanical Gardens Kew*. Chicago, IL: University of Chicago Press.

Danquah, E. Y., Hanley, S. J., Brookes, R. C., Aldam, C., and Karp, A. (2002). Isolation and characterization of microsatellites in *Echinochloa* (L.) Beauv. spp. *Mol. Ecol. Notes* 2, 54–56. doi: 10.1046/j.1471-8286.2002.00144.x

De Wet, J., Prasada Rao, K., Mengesha, M., and Brink, D. (1983). Domestication of mawa millet (*Echinochloa colona*). *Econ. Bot.* 37, 283–291. doi: 10.1007/BF02858883

Denton, D. C. (1987). "Food crops for waterfowl," in *Fireside Waterfowler: Fundamentals of Duck and Goose Ecology*, eds D. E. Wesley, and W. G. Leitch, (Mechanicsburg, PA: Stackpole Books), 352.

Dida, M. M., Srinivasachary, Ramakrishnan, S., Bennetzen, J. L., Gale, M. D., and Devos, K. M. (2007). The genetic map of finger millet, *Eleusine coracana*. *Theor. Appl. Genet.* 114, 321–332. doi: 10.1007/s00122-006-0435-7

Doggett, H. (1989). "Small millets-a selective overview," in *Small Millets in Global Agriculture*, eds A. Seetharam, K. W. Riley, and G. Harinarayana, (Oxford: Oxford), 3–18.

Fang, X., Dong, K., Wang, X., Liu, T., He, J., Ren, R., et al. (2016). A high density genetic map and QTL for agronomic and yield traits in Foxtail millet [*Setaria italica* (L.) P. Beauv.]. *BMC Genomics* 17:336. doi: 10.1186/s12864-016-2628-z

Frey, M., Schullehner, K., Dick, R., Fiesselmann, A., and Gierl, A. (2009). Benzoxazinoid biosynthesis, a model for evolution of secondary metabolic pathways in plants. *Phytochemistry* 70, 1645–1651. doi: 10.1016/j.phytochem.2009.05.012

Gao, Y., Li, J., Pan, X., Liu, D., Napier, R., and Dong, L. (2018). Quinclorac resistance induced by the suppression of the expression of 1-aminocyclopropane-1-carboxylic acid (ACC) synthase and ACC oxidase genes in *Echinochloa crus-galli* var. zelayensis. *Pestic. Biochem. Physiol.* 146, 25–32. doi: 10.1016/j.pestbp.2018.02.005

Gimode, D., Odeny, D. A., de Villiers, E. P., Wanyonyi, S., Dida, M. M., Mneney, E. E., et al. (2016). Identification of SNP and SSR markers in finger millet using next generation sequencing technologies. *PLoS One* 11:e0159437. doi: 10.1371/journal.pone.0159437

Gomashe, S. S. (2017). Barnyard millet: present status and future thrust areas. *Millets Sorghum Biol. Genet. Improv.* 134, 184–198. doi: 10.1002/9781119130765.ch7

Goron, T. L., and Raizada, M. N. (2015). Genetic diversity and genomic resources available for the small millet crops to accelerate a New Green Revolution. *Front. Plant Sci.* 6:157. doi: 10.3389/fpls.2015.00157

Guo, L., Qiu, J., Ye, C.-Y., Jin, G., Lingfeng, M., Zhang, H., et al. (2017). Echinochloa crus-galli genome analysis provides insight into its adaptation and invasiveness as a weed. *Nat. Commun.* 8:1031. doi: 10.1038/s41467-017-01067-5

Gupta, A., Joshi, D., Mahajan, V., and Gupta, H. (2010a). Screening barnyard millet germplasm against grain smut (Ustilago panici-frumentacei Brefeld). *Plant Genet. Resour.* 8, 52–54. doi: 10.1017/S1479262109990141

Gupta, A., Mahajan, V., and Gupta, H. S. (2010b). "Genetic resources and varietal improvement of small millets for Indian Himalaya," in *Biodiversity Potentials of the Himalaya*, eds L. M. Tewari, Y. P. S. Pangtey, and G. Tewari, (Gorakhpur: Gyanodaya Prakashan press), 305–316.

Gupta, A., Mahajan, V., Kumar, M., and Gupta, H. (2009). Biodiversity in the barnyard millet (*Echinochloa frumentacea* Link. Poaceae) germplasm in India. *Genet. Resour. Crop Evol.* 56, 883–889. doi: 10.1007/s10722-009-9462-y

Gupta, A., Sood, S., Agrawal, P. K., and Bhatt, J. C. (2014). B 29: an Easy Dehulling Barnyard Millet (*Echinochloa frumentacea* Link) Genotype. *Natl. Acad. Sci. Lett.* 38, 21–24. doi: 10.1007/s40009-014-0295-7

Gupta, P., Raghuvanshi, S., and Tyagi, A. (2001). Assessment of the efficiency of various gene promoters via biolistics in leaf and regenerating seed callus of millets, eleusine coracana and *Echinochloa crusgalli*. *Plant Biotechnol.* 18, 275–282. doi: 10.5511/plantbiotechnology.18.275

Gupta, S., Kumari, K., Muthamilarasan, M., Alagesan, S., and Prasad, M. (2013). Development and utilization of novel SSRs in foxtail millet [*Setaria italica* (L.) P. Beauv.]. *Plant Breed.* 132, 367–374. doi: 10.1111/pbr.12070

Hilu, K. (1994). Evidence from RAPD markers in the evolution of *Echinochloa millets* (Poaceae). *Plant Syst. Evol.* 189, 247–257. doi: 10.1007/BF00939730

Hittalmani, S., Mahesh, H. B., Shirke, M. D., Biradar, H., Uday, G., Aruna, Y. R., et al. (2017). Genome and Transcriptome sequence of Finger millet (*Eleusine coracana* (L.) Gaertn.) provides insights into drought tolerance and nutraceutical properties. *BMC Genomics* 18:465. doi: 10.1186/s12864-017-3850-z

Hoshino, T., Nakamura, T., Seimiya, Y., Kamada, T., Ishikawa, G., Ogasawara, A., et al. (2010). Production of a fully waxy line and analysis of waxy genes in the allohexaploid crop, Japanese barnyard millet. *Plant Breed.* 129, 349–355.

Hou, S., Sun, Z., Li, Y., Wang, Y., Ling, H., Xing, G., et al. (2017). Transcriptomic analysis, genic ssr development, and genetic diversity of proso millet (*Panicum miliaceum*;Poaceae). *Appl. Plant Sci.* 5:1600137. doi: 10.3732/apps.1600137

Hunt, H. V., Denyer, K., Packman, L. C., Jones, M. K., and Howe, C. J. (2010). Molecular basis of the waxy endosperm starch phenotype in broomcorn millet (*Panicum miliaceum* L.). *Mol. Biol. Evol.* 27, 1478–1494. doi: 10.1093/molbev/msq040

IIMR, (2016). *Annual Report 2016-17*. Hyderabad: Indian Institute of Millets Research.

IIMR, (2018). *Annual Report 2017-18*. Hyderabad: Indian Institute of Millets Research.

Ishikawa, G., Seimiya, Y., Saito, M., Nakamura, T., and Hoshino, T. (2013). Molecular characterization of spontaneous and induced mutations in the three homoeologous waxy genes of Japanese barnyard millet [*Echinochloa esculenta* (A. Braun) H. Scholz]. *Mol. Breed.* 31, 69–78. doi: 10.1007/s11032-012-9769-9

Iwakami, S., Uchino, A., Kataoka, Y., Shibaike, H., Watanabe, H., and Inamura, T. (2014). Cytochrome P450 genes induced by bispyribac-sodium treatment in a multiple-herbicide-resistant biotype of *Echinochloa phyllopogon*. *Pest Manag. Sci.* 70, 549–558. doi: 10.1002/ps.3572

Jagadish, P. S., Mohapatra, H. K., Chakravarthy, M. K., Srivastava, N., and Nangia, N. (2008). *A Compendium of Insect Pests of Finger Millet and Other Small Millets*. Available online at: http://www.aicrpsm.res.in/Downloads/Publications//A%20compendium%20of%20Insect%20pests%20of%20FM%20&%20other%20SM.pdf

Jain, A. K., Jain, S. K., and Yadava, H. S. (1997). Assessment of yield losses due to grain smut in barnyard millet. *Indian Phytopathol.* 50, 49–52.

Jaiswal, V., Gupta, S., Gahlaut, V., Muthamilarasan, M., Bandyopadhyay, T., Ramchiary, N., et al. (2019). Genome-wide association study of major agronomic traits in foxtail millet (*Setaria italica* L.) Using ddRAD sequencing. *Sci. Rep.* 9:5020. doi: 10.1038/s41598-019-41602-6

Jayaraman, N., Suresh, S., and Nirmala Ganeshan, N. M. (1997). "Genetic enhancement and breeding strategies in small millets," in *Proceedings of the National Seminar on Small Millets: Current Research Trends and Future Priorities as Food Feed and in Processing for Value Addition*, Bangalore, 23–24.

Jaybhaye, R. V., and Srivastav, P. P. (2015). Development of barnyard millet ready-to-eat snack food: Part II. *Food Sci. Res. J.* 6, 285–291. doi: 10.15740/HAS/FSRJ/6.2/285-291

Jia, G., Huang, X., Zhi, H., Zhao, Y., Zhao, Q., Li, W., et al. (2013). A haplotype map of genomic variations and genome-wide association studies of agronomic traits in foxtail millet (*Setaria italica*). *Nat. Genet.* 45, 957–961. doi: 10.1038/ng.2673

Jia, X., Zhang, Z., Liu, Y., Zhang, C., Shi, Y., Song, Y., et al. (2009). Development and genetic mapping of SSR markers in foxtail millet [*Setaria italica* (L.) P. Beauv.]. *Theor. Appl. Genet.* 118, 821–829. doi: 10.1007/s00122-008-0942-9

Johnson, M., Deshpande, S., Vetriventhan, M., Upadhyaya, H. D., and Wallace, J. G. (2019). Genome-Wide population structure analyses of three minor millets: kodo millet, little millet, and proso millet. *Plant Genome* 12:190021. doi: 10.3835/plantgenome2019.03.0021

Joshi, R. P., Jain, A. K., Chauhan, S. S., and Singh, G. (2015). Characterization of barnyard millet (Echinochloa frumentacea (Roxb.) Link.) landraces for agro-morphological traits and disease resistance. *Electron. J. Plant Breed.* 6, 888–898.

Joshi, V. (2013). Assessment of genetic variability and identification of genotypes for different traits in Barnyard millet (*Echinochola* spp.). *Int. J. Agric. Food Sci. Technol.* 4, 65–67.

Kannan, S. M., Thooyavathy, R. A., Kariyapa, R. T., Subramanian, K., and Vijayalakshmi, K. (2013). *Data from: Seed Production Techniques for Cereals and Millets. Seed node of the Revitalizing Rainfed Agriculture Network Centre for Indian Knowledge Systems (CIICS)*. Available at: http://www.ciks.org/downloads/seeds/5.%20Seed%20Production%20Techniques%20for%20Cereals%20and%20Millets.pdf (Accessed August 09, 2019).

Kaya, H., Demirci, M., and Tanyolac, B. (2014). Genetic structure and diversity analysis revealed by AFLP on different *Echinochloa* spp. from northwest Turkey. *Plant Syst. Evol.* 300, 1–11. doi: 10.1007/s00606-013-0965-9

Kim, C. S., Alamgir, K. M., Matsumoto, S., Tebayashi, S. I., and Koh, H. S. (2008). Antifeedants of Indian barnyard millet, *Echinochloa frumentacea* link, against brown planthopper, *Nilaparvata lugens* (Stål). *Sect. C J. Biosci.* 63, 755–760. doi: 10.1515/znc-2008-9-1022

Kim, J. Y., Chang, J. K., Park, B.-R., Han, S.-I., Choi, K.-J., Kim, S.-Y., et al. (2011). Physicochemical and antioxidative properties of selected barnyard millet (*Echinochloa utilis*) species in Korea. *Food Sci. Biotechnol.* 20, 461–469. doi: 10.1007/s10068-011-0064-z

Kovach, D. A., Widrlechner, M., and Brenner, D. (2010). Variation in seed dormancy in Echinochloa and the development of a standard protocol for germination testing. *Seed Sci. Technol.* 38, 559–571. doi: 10.15258/sst.2010.38.3.04

Kraehmer, H., Jabran, K., Mennan, H., and Chauhan, B. (2015). Global distribution of rice weeds - A review. *Crop Prot.* 80, 73–86. doi: 10.1016/j.cropro.2015.10.027

Krishna, T. P. A., Maharajan, T., David, R. H. A., Ramakrishnan, M. S., Ceasar, A., Duraipandiyan, V., et al. (2018). Microsatellite markers of finger millet (*Eleusine coracana* (L.) Gaertn) and foxtail millet (*Setaria italica* (L.) Beauv) provide resources for cross-genome transferability and genetic diversity analyses in other millets. *Biocat. Agrl. Biotechnol.* 16, 493–450.

Kumar, A., Yadav, S., Panwar, P., Gaur, V., and Sood, S. (2015). Identification of anchored simple sequence repeat markers associated with calcium content in finger millet (*Eleusine coracana*). *Proc. Natl. Acad. Sci. India B Biol. Sci.* 85, 311–317. doi: 10.1007/s40011-013-0296-1

Kumar, L. D., Siva Sankar, S., Venkatesh, P., and Hepcy Kalarani, D. (2016). Green synthesis of silver nanoparticles using aerial parts extract of *Echinochloa colona* and their characterization. *Eur. J. Pharm. Med. Res.* 3, 325–328.

Kumari, K., Muthamilarasan, M., Misra, G., Gupta, S., Alagesan, S., Parida, S., et al. (2013). Development of eSSR-Markers in *Setaria italica* and their applicability in studying genetic diversity, cross-transferability and comparative mapping in millet and non-millet species. *PLoS One* 8:e67742. doi: 10.1371/journal.pone.0067742

Kumari, K. S., and Thayumanavan, B. (1998). Characterization of starches of proso, foxtail, barnyard, kodo, and little millets. *Plant Foods Hum. Nutr.* 53:47. doi: 10.1023/A:1008083020810

Kuraloviya, M., Vanniarajan, C., Vetriventhan, M., Babu, C., Kanchana, S., and Sudhagar, R. (2019). Qualitative characterization and clustering of early-maturing barnyard millet (*Echinochloa* spp.) germplasm. *Electron. J. Plant Breed.* 10:535. doi: 10.5958/0975-928x.2019.00067.x

Lata, C., and Prasad, M. (2012). Validation of an allele-specific marker associated with dehydration stress tolerance in a core set of foxtail millet accessions. *Plant Breed.* 132, 496–499. doi: 10.1111/j.1439-0523.2012.01983.x

Lee, J., Kim, C.-S., and Lee, I.-Y. (2017). Discrimination of *Echinochloa colona* (L.) Link from other Echinochloa Species using DNA Barcode. *Weed Turfgrass Sci.* 4, 225–229. doi: 10.5660/WTS.2015.4.3.225

Lee, J., Park, K., Lee, I., Kim, C., Kown, O., and Park, T. (2014). Simple sequence repeat analysis of genetic diversity among Acetyl-CoA carboxylase inhibitor-resistant and -susceptible Echinochloa crus-galli and *E. oryzicola* populations in Korea. *Weed Res.* 55, 90–100. doi: 10.1111/wre.12119

Li, G., Wu, S., Cai, L., Wang, Q., Zhao, X., and Wu, C. (2013a). Identification and mRNA expression profile of glutamate receptor-like gene in quinclorac-resistant and susceptible Echinochloa crus-galli. *Gene* 531, 489–495. doi: 10.1016/j.gene.2013.09.013

Li, G., Wu, S. G., Yu, R. X., Cang, T., Chen, L. P., Zhao, X. P., et al. (2013b). Identification and expression pattern of a glutathione S-transferase in *Echinochloa crus*-galli. *Weed Res.* 53, 314–321. doi: 10.1111/wre.12031

Lin, H.-S., Chiang, C. Y., Chang, S.-B., and Kuoh, C.-S. (2011). Development of simple sequence repeats (SSR) Markers in *Setaria italica* (Poaceae) and cross-amplification in related species. *Int. J. Mol. Sci.* 12, 7835–7845. doi: 10.3390/ijms12117835

Lohani, U., Pandey, J., and Shahi, N. (2012). Effect of degree of polishing on milling characteristics and proximate compositions of barnyard millet (*Echinochloa frumentacea*). *Food Bioprocess Technol.* 5, 1113–1119. doi: 10.1007/s11947-011-0518-6

Lule, D., de Villiers, S., Fetene, M., Odeny, D. A., Rathore, A., Das, R. R., et al. (2018). Genetic diversity and association mapping of Ethiopian and exotic finger millet accessions. *Crop Pasture Sci.* 69, 879–891. doi: 10.1071/CP18175

Mace, E. S., Tai, S., Gilding, E. K., Li, Y., Prentis, P. J., Bian, L., et al. (2013). Whole-genome sequencing reveals untapped genetic potential in Africa's indigenous cereal crop sorghum. *Nat. Commun.* 4:2320. doi: 10.1038/ncomms3320

Manidool, C. (1992). "Echinochloa crus-galli (L.) P. Beauv," in *Plant Resources of South-East Asia*, eds t' Mannetje, and R. M. Jones, (Jodhpur: Pudoc Scientific Publishers), 303.

Manimekalai, M., Dhasarathan, M., Karthikeyan, A., Murukarthick, J., Renganathan, V. G., Thangaraj, K., et al. (2018). Genetic diversity in the barnyard millet (Echinochola frumentacea) germplasms revealed by morphological traits and simple sequence repeat markers. *Curr. Plant Biol.* 14, 71–78. doi: 10.1016/j.cpb.2018.09.006

Maun, M. A., and Barrett, S. C. (1986). The biology of Canadian weeds, *Echinochloa crus- galli* (L.) Beauv. *Can. J. Plant Sci.* 66, 739–759. doi: 10.4141/cjps86-093

Mauro-Herrera, M., Wang, X., Barbier, H., Brutnell, T. P., Devos, K. M., and Doust, A. N. (2013). Genetic Control and Comparative Genomic Analysis of Flowering Time in Setaria (Poaceae). *G3* 3, 283L–295L. doi: 10.1534/g3.112.005207

Mehta, H., Tyagi, P. C., and Mohapatra, K. P. (2005). Genetic diversity in Barnyard millet (*Echinochloa frumentacea* Roxb.). *Indian J. Genet.* 65, 293–295.

Mishra, A. K., Puranik, S., Bahadur, R. P., and Prasad, M. (2012). The DNA-binding activity of an AP2 protein is involved in transcriptional regulation of a stress-responsive gene, SiWD40, in foxtail millet. *Genomics* 100, 252–263. doi: 10.1016/j.ygeno.2012.06.012

Mitich, L. W. (1990). Barnyardgrass. *Weed Technol.* 4, 918–920.

Monteiro, P. V., Sudharshana, L., and Ramachandra, G. (1988). Japanese barnyard millet (*Echinochloa frumentacea*): protein content, quality and SDS-PAGE of protein fractions. *J. Sci. Food Agric.* 43:17e25.

Moreno Amador, M., de Comino, I., and Sousa, C. (2014). Alternative grains as potential raw material for gluten-free food development in the diet of celiac and gluten-sensitive patients. *Austin J. Nutr. Food Sci.* 2:9.

Mosovska, S., Mikulasova, M., Brindzova, L., Valik, L., and Mikusova, L. (2010). Genotoxic and antimutagenic activities of extracts from pseudocereals in the *Salmonella* mutagenicity assay. *Food Chem. Toxicol.* 48, 1483–1487. doi: 10.1016/j.fct.2010.03.015

Muldoon, D. K. (1985). The effect of photoperiod on the growth and development of *Echinochloa* spp. Millets. *Aust. J. Exp. Agric.* 25, 428–433. doi: 10.1071/EA9850428

Muldoon, K. D., Pearson, C. J., and Wheeler, J. (1982). The effect of temperature on growth and development of *Echinochloa* millets. *Ann. Bot.* 50, 665–672. doi: 10.1093/oxfordjournals.aob.a086408

Murukarthick, J., Manimekalai, M., Karthikeyan, A., Perumal, S., Dhasarathan, M., Kandasamy, T., et al. (2019). Transcriptomes of Indian barnyard millet and barnyardgrass reveal putative genes involved in drought adaptation and micronutrient accumulation. *Acta Physiol. Plant.* 41:66. doi: 10.1007/s11738-019-2855-4

Muthamilarasan, M., and Prasad, M. (2014). Advances in Setaria genomics for genetic improvement of cereals and bioenergy grasses. *Theor. Appl. Genet.* 128, 1–14. doi: 10.1007/s00122-014-2399-3

Muza, F. R., Lee, D. J., Andrews, D. J., and Gupta, S. C. (1995). Mitochondrial DNA variation in finger millet [*Eleusine coracana* (L.) Gaertn.]. *Euphytica* 81, 199–205. doi: 10.1007/BF00025434

Nagaraja, A., and Mantur, S. G. (2008). Evaluation of barnyard millet entries for grain smut resistance and yield. *Mysore J. Agric. Sci.* 42, 375–377.

Nah, G., Im, J.-H., Kim, J.-W., Park, H.-R., Yook, M.-J., Yang, T.-J., et al. (2015). Uncovering the differential molecular basis of adaptive diversity in three echinochloa leaf transcriptomes. *PLoS One* 10:e0134419. doi: 10.1371/journal.pone.0134419

Nandini, R., and Fakrudin, B. (1999). Emasculation in finger millet using hot water. *Mysore J. Agric. Sci.* 33, 6–10.

Nazni, P., and Shobana, D. R. (2016). Effect of processing on the characteristics changes in barnyard and foxtail millet. *J. Food Process. Technol.* 7:566. doi: 10.4172/2157-7110.1000566

Nguyen, H. D., Bingtian, Z., Le, D. A. T., Yoon, Y. H., Ko, J. Y., Woo, K. S., et al. (2016). Isolation of lignan and fatty acid derivatives from the grains of *Echinochloa utilis* and Their Inhibition of Lipopolysaccharide-Induced Nitric Oxide Production in RAW 264.7 cells. *J. Agric. Food Chem.* 64, 425–432. doi: 10.1021/acs.jafc.5b05638

Nirgude, M., Babu, B. K., Shambhavi, Y., Singh, U. M., Upadhyaya, H. D., and Kumar, A. (2014). Development and molecular characterization of genic molecular markers for grain protein and calcium content in finger millet (*Eleusine coracana* (L.) Gaertn.). *Mol. Biol. Rep.* 41, 1189–1200. doi: 10.1007/s11033-013-2825-7

Nirmalakumari, A., and Vetriventhan, M. (2009). Phenotypic analysis of anther and pollen in diversified genotype of barnyard millet (*Echinochloa frumentaceae*) floral characters. *ICFAI Univ. J. Genet. Evol.* 2, 12–16.

Nolde, S., Vassilevski, A. A., Rogozhin, E. A., Barinov, N. A., Balashova, T. A., Samsonova, O. V., et al. (2011). Disulfide-stabilized helical hairpin structure and activity of a novel antifungal peptide EcAMP1 from seeds of barnyard grass (*Echinochloa crus*-galli). *J. Biol. Chem.* 286, 25145–25153. doi: 10.1074/jbc.M110.200378

Nozawa, S., Takahashi, M., Nakai, H., and Sato, Y.-I. (2006). Difference in SSR Variations between Japanese Barnyard Millet (*Echinochloa esculenta*) and its Wild Relative *E. crus-galli*. *Breed. Sci.* 56, 335–340. doi: 10.1270/jsbbs.56.335

Obidiegwu, O. N., Obidiegwu, J. E., and Parzies, H. (2013). Development of SSR for foxtail millet (*Setaria italica* (L.) P. Beauv.) and its utility in genetic discrimination of a core set. *Genes Genomics* 35, 609–615. doi: 10.1007/s13258-013-0110-8

Odintsova, T. I., Rogozhin, E. A., Baranov, Y., Musolyamov, A. K., Yalpani, N., Egorov, T. A., et al. (2008). Seed defensins of barnyard grass *Echinochloa crusgalli* (L.) Beauv. *Biochimie* 90, 1667–1673. doi: 10.1016/j.biochi.2008.06.007

Okada, M., Hanson, B. D., Hembree, K. J., Peng, Y., Shrestha, A., Wright, D., et al. (2013). Evolution and spread of glyphosate resistance in *Conyza canadensis* in California. *Evol. Appl.* 6, 761–777. doi: 10.1111/eva.12061

Olanrewaju, O. S., Glick, B. R., and Babalola, O. O. (2017). Mechanisms of action of plant growth promoting bacteria. *World J. Microbiol. Biotechnol.* 33:197. doi: 10.1007/s11274-017-2364-9

Osuna, M., Okada, M., Ahmad, R., Fischer, A., and Jasieniuk, M. (2011). Genetic Diversity and spread of thiobencarb resistant early watergrass (*Echinochloa oryzoides*) in California. *Weed Sci.* 59, 195–201. doi: 10.1614/WS-D-10-00124.1

Padulosi, S., Bhag, M. C., Bala, R. S., Gowda, J. V., Gowda, K. T. K., Shanthakumar, G. C., et al. (2009). Food security and climate change: role of plant genetic resources of minor millets. *Indian J. Plant Genet. Resour.* 22, 1–16.

Pandey, G., Misra, G., Kumari, K., Gupta, S., Kumar Parida, S., Chattopadhyay, D., et al. (2013). Genome-wide development and use of microsatellite markers for large-scale genotyping applications in foxtail millet [*Setaria italica* (L.)]. *DNA Res.* 20, 197–207. doi: 10.1093/dnares/dst002

Panwar, P., Dubey, A., and Verma, A. K. (2016). Evaluation of nutraceutical and antinutritional properties in barnyard and finger millet varieties grown in Himalayan region. *J. Food Sci. Technol.* 53, 2779e2787.

Perumal, S., Jayakodi, M., Kim, D.-S., Yang, T.-J., and Natesan, S. (2016). The complete chloroplast genome sequence of Indian barnyard millet, *Echinochloa frumentacea* (Poaceae). *Mitochondrial DNA* 1, 79–80. doi: 10.1080/23802359.2015.1137832

Prabha, D., Negi, Y. K., and Khanna, V. K. (2010). Morphological and isozyme diversity in the accessions of two cultivated species of barnyard millet. *Nat. Sci.* 8, 71–76.

Prabha, D., Negi, Y. K., and Khanna, V. K. (2012). Assessment of genetic diversity of barnyard millet accessions using molecular markers. *Indian J. Plant Genet. Resour.* 25, 174–179.

Prakash, R., and Vanniarajan, C. (2013). Genetic variability for panicle characters in indigenous and exotic barnyard millet (*Echinochloa frumentacea* (Roxb.) Link) germplasm over environment. *Veg. An Int. J. Plant Res.* 26:297. doi: 10.5958/j.2229-4473.26.2.088

Qie, L., Jia, G., Zhang, W., Schnable, J., Shang, Z., Li, W., et al. (2014). Mapping of Quantitative Trait Locus (QTLs) that contribute to germination and early seedling drought tolerance in the interspecific cross *Setaria italica*×*Setaria viridis*. *PLoS One* 9:e101868. doi: 10.1371/journal.pone.0101868

Radhamani, J., Pandey, A., Srinivasan, K., and Tyagi, V. (2011). Conserving millet genetic resources in India. *Proc. Indian Natl. Sci. Acad.* 77, 295–304.

Raj, S. M., Mahudeswaran, K., and Shanmugasundaran, A. (1964). Observation on the hot water technique of emasculation of ragi flowers (*Eleusine coracana* (L.) Gaertn.). *Madras Agric. J.* 51, 71–75.

Rajput, S. G., and Santra, D. K. (2016). Evaluation of genetic diversity of proso millet germplasm available in the United States using simple-sequence repeat markers. *Crop Sci.* 56, 2401–2409. doi: 10.2135/cropsci2015.10.0644

Rajput, S. G., Santra, D. K., and Schnable, J. (2016). Mapping QTLs for morpho-agronomic traits in proso millet (*Panicum miliaceum* L.). *Mol. Breed.* 36:37. doi: 10.1007/s11032-016-0460-4

Ramakrishnan, M., Antony Ceasar, S., Duraipandiyan, V., Al-Dhabi, N. A., and Ignacimuthu, S. (2016). Assessment of genetic diversity, population structure and relationships in Indian and non-Indian genotypes of finger millet (*Eleusine coracana* (L.) Gaertn) using genomic SSR markers. *Springerplus* 5, 1–11. doi: 10.1186/s40064-015-1626-y

Ramakrishnan, M., Ceasar, S. A., Vinod, K. K., Duraipandiyan, V., Ajeesh Krishna, T. P., Upadhyaya, H. D., et al. (2017). Identification of putative QTLs for seedling stage phosphorus starvation response in finger millet (*Eleusine coracana* L. Gaertn.) by association mapping and cross species synteny analysis. *PLoS One* 12:e0183261. doi: 10.1371/journal.pone.0183261

Rawat, L., Nautiyal, A., Bisht, T. S., Prasad, S., Naithani, D., Makhloga, K., et al. (2019). Screening of barnyard millet germplasm against shoot fly and stem borer damage under field conditions. *Int. J. Curr. Microbiol. Appl. Sci.* 8, 1221–1226. doi: 10.20546/ijcmas.2019.802.142

Renganathan, V. G., Ibrahim, S. M., and Vanniarajan, C. (2015). Combining ability for quantitative traits in Barnyard millet (*Echinochloa frumentaceae* (Roxb.) Link). *Electron. J. Plant Breed.* 6, 389–394.

Renganathan, V. G., Vanniarajan, C., Nirmalakumari, A., Raveendran, M., and Thiyageshwari, S. (2017). Cluster analyses for qualitative and quantitative traits in barnyard millet *Echinochloa frumentacea* (Roxb.Link) germplasm. *Bioscan* 12, 1927–1931.

Renganathan, V. G., Vanniarajan, C., and Ramalingam, J. (2019). "Genetic analysis and identification of molecular markers linked to the anthocyanin pigmentation in barnyard millet [*Echinochloa frumentacea* Roxb (Link)]," in *Proceedings of the Neglected and Underutilized crop species for Food, Nutrition, Energy and Environment*, (New Delhi: NIPGR), 43.

Ruiz-Santaella, P. J., Bastida, F., Franco, A., and Prado, R. (2006). Morphological and molecular characterization of different *Echinochloa* spp. and *Oryza sativa* populations. *J. Agric. Food Chem.* 54, 1166–1172. doi: 10.1021/jf0520746

Rutledge, J., Talbert, R. E., and Sneller, C. H. (2000). RAPD analysis of genetic variation among propanil-resistant and -susceptible Echinochloa crus-galli populations in Arkansas. *Weed Sci.* 48, 669–674. doi: 10.1614/0043-1745(2000)048[0669:raogva]2.0.co;2

Saleh, A., Zhang, Q., Chen, J., and Shen, Q. (2013). Millet grains: nutritional quality, processing, and potential health benefits. *Compr. Rev. Food Sci. Food Saf.* 12, 281–295. doi: 10.1111/1541-4337.12012

Santra, D., Heyduck, R., Baltensperger, D., Graybosch, R., Nelson, L., Frickel, G., et al. (2015). Registration of 'Plateau' Waxy (Amylose-Free) proso millet. *J. Plant Regist.* 9:41. doi: 10.3198/jpr2013.11.0067crc

Sato, K., Mukainari, Y., Naito, K., and Fukunaga, K. (2013). Construction of a foxtail millet linkage map and mapping of spikelet-tipped bristles1 (stb1) by using transposon display markers and simple sequence repeat markers with genome sequence information. *Mol. Breed.* 31, 675–684. doi: 10.1007/s11032-012-9825-5

Sayani, R., and Chatterjee, A. (2017). Nutritional and biological importance of the weed *Echinochloa colona*: a review. *Int. J. of Food Sci. Biotechnol.* 2, 31–37. doi: 10.11648/j.ijfsb.20170202.13

Sebastin, R., Lee, K. J., Cho, G. T., Lee, J. R., Shin, M. J., Kim, S. H., et al. (2019). The complete chloroplast genome sequence of Japanese Millet *Echinochloa esculenta* (A. braun) H. scholz (Poaceae). *Mitochondrial DNA Part B Resour.* 4, 1392–1393. doi: 10.1080/23802359.2019.1598787

Seetharam, A., Gowda, J., and Halaswamy, J. H. (2003). "Small millets," in *Nucleus and Breeder Seed Production Manual*, eds S. K. Chowdhury, and S. K. Lal, (New Delhi: Indian Agricultural Research Institute), 54–67.

Sharma, A., Sood, S., Agrawal, P. K., Kant, L., Bhatt, J. C., and Pattanayak, A. (2016). Detection and assessment of nutraceuticals in methanolic extract of finger (*Eleusine coracana*) and Barnyard Millet (*Echinochloa frumentacea*). *Asian J. Chem.* 28, 1633–1637. doi: 10.14233/ajchem.2016.19790

Sharma, D., Tiwari, A., Sood, S., Jamra, G., Singh, N. K., Meher, P. K., et al. (2018). Genome wide association mapping of agro-morphological traits among a diverse collection of finger millet (*Eleusine coracana* L.) genotypes using SNP markers. *PLoS One* 13:e19444. doi: 10.1371/journal.pone.0199444

Siles, M., Baltensperger, D., and Nelson, L. A. (2001). Technique for Artificial Hybridization of Foxtail Millet [(L.) Beauv.]. *Crop Sci.* 41, 1408–1412. doi: 10.2135/cropsci2001.4151408x

Singh, H. S., and Singh, K. (2005). "Status and needs of pasture and fodder management in Uttaranchal," in *Road Map for Pasture and Fodder Development in NWHR for Livestock Sustenance*, eds J. K. Bisht, and A. K. Srivastava, (Almora: Vivekananda Parvatiya Krishi Anusandhan Sansthan), 39–64.

Singh, K. P., Mishra, H. N., and Saha, S. (2010). Moisture-dependent properties of barnyard millet grain and kernel. *J. Food Eng.* 96, 598–606. doi: 10.1016/j.jfoodeng.2009.09.007

Song, B. Y., Shi, J. X., and Song, S. Q. (2015). Dormancy release and germination of *Echinochloa crus-galli* grains in relation to galactomannan-hydrolysing enzyme activity. *J. Integr. Agric.* 14, 1627–1636. doi: 10.1016/S2095-3119(14)60940-0

Sood, S., Khulbe, R., Kumar, R. A., Agrawal, P. K., and Upadhyaya, H. (2015). Barnyard millet global core collection evaluation in the sub mountain Himalayan region of India using multivariate analysis. *Crop J.* 3, 517–525. doi: 10.1016/j.cj.2015.07.005

Sood, S., Khulbe, R., Saini, N., Gupta, A., and Agrawal, P. K. (2014). Research Note Interspecific Hybrid between Echinochloa esculenta (Japanese barnyard millet) and E. frumentacea (Indian barnyard millet) – A New Avenue for Genetic Enhancement of Barnyard Millet. *Electron. J. Plant Breed.* 5, 248–253.

Sridevi, R., and Manonmani, V. (2016). Seed priming effect on physiological traits of kodo millet and barnyard millet. *Int. J. Agric. Sci. Res.* 6, 187–194.

Sridhar, R., and Lakshminarayana, G. (1992). Lipid class contents and fatty acid composition of small millets: barnyard little (*Panicum sumatrense*), kodo (*Paspalum scrobiculatum*), and barnyard (*Echinochloa colona*). *J. Agric. Food Chem.* 40:2131e2134.

Subhashini, V., and Swamy, A. V. V. S. (2014). Screening Potential of three native grass species for phyto remediation of heavy metals. *Int. J. Eur. Acad. Res.* 4, 5887–5903.

Sujatha, K., Selvarani, K., Vijayalakshmi, V., Vanniarajan, C., and Sivasubramaniam, K. (2013). Seed fortification studies in barnyard millet (*Echinochloa frumentacea*) cv. CO1. *IOSR J. Agric. Vet. Sci.* 5, 22–24. doi: 10.9790/2380-0542224

Sun, J., Luu, N. S., Chen, Z., Chen, B., Cui, X., Wu, J., et al. (2019). Generation and Characterization of a Foxtail Millet (*Setaria italica*) Mutant Library. *Front. Plant Sci.* 10:369. doi: 10.3389/fpls.2019.00369

Sundararaj, D. P., and Thulasidas, G. (1976). *Botany of Field Crops*. New Delhi: Macmillan Publisher.

Sung, S.-J., Leather, G. R., and Hale, M. G. (1987). Development and germination of barnyardgrass (*Echinochloa crus-galli*) seeds. *Weed Sci.* 35, 211–215. doi: 10.1017/S0043174500079078

Surekha, N., Ravikumar, S. N., Mythri, S., and Rohini, D. (2013). Barnyard Millet (*Echinochloa frumentacea* Link) cookies: development, value addition, consumer acceptability, nutritional and shelf life evaluation. *IOSR J. Environ. Sci. Toxicol. Food Technol.* 7, 01–10. doi: 10.9790/2402-0730110

Suryawanshi, V., Talke, I. N., Weber, M., Eils, R., Brors, B., Clemens, S., et al. (2016). Between-species differences in gene copy number are enriched among functions critical for adaptive evolution in *Arabidopsis halleri*. *BMC Genomics* 17:1034. doi: 10.1186/s12864-016-3319-5

Tabacchi, M., Mantegazza, R., Spada, A., and Ferrero, A. (2009). Morphological traits and molecular markers for classification of *Echinochloa* species from Italian rice fields. *Weed Sci.* 54, 1086–1093. doi: 10.1614/WS-06-018R1.1

Trivedi, A. K., Arya, L., Verma, S., Tyagi, R., and Hemantaranjan, A. (2017). Evaluation of barnyard millet diversity in central Himalayan region for environmental stress tolerance. *J. Agric. Sci.* 155, 1–11. doi: 10.1017/S0021859617000545

Ugare, R., Chimmad, B., Naik, R., Bharati, P., and Itagi, S. (2014). Glycemic index and significance of barnyard millet (*Echinochloa frumentacae*) in type II diabetics. *J. Food Sci. Technol.* 51, 392–395. doi: 10.1007/s13197-011-0516-8

Upadhyaya, H., Dwivedi, S. L., Singh, S. K., Sube, S., Vetriventhan, M., and Sharma, S. (2014). Forming core collections in Barnyard, Kodo, and little millets using morphoagronomic descriptors. *Crop Sci.* 54, 2673–2682. doi: 10.2135/cropsci2014.03.0221

Upadhyaya, H., Gowda, C., and Sastry, D. (2008). Plant genetic resources management: collection, characterization, conservation and utilization. *J. SAT Agric. Res.* 6:16.

Vanniarajan, C., Anand, G., Kanchana, S., Arun Giridhari, V., and Renganathan, V. G. (2018). A short duration high yielding culture - Barnyard millet ACM 10145. *Agric. Sci. Dig. A Res. J.* 8, 123–126. doi: 10.18805/ag.D-4574

Varshney, R. K., Hoisington, D. A., and Tyagi, A. K. (2006). Advances in cereal genomics and applications in crop breeding. *Trends Biotechnol.* 24, 490–499. doi: 10.1016/j.tibtech.2006.08.006

Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35, 969–976. doi: 10.1038/nbt.3943

Veena, B., Chimmad, B. V., Naik, R. K., and Shantakumar, G. (2004). Development of barnyard millet based traditional foods. *Karnataka J. Agric. Sci.* 17, 522–527.

Veena, S., Chimmad, B. V., Naik, R. K., and Shanthakumar, G. (2005). Physico-chemical and nutritional studies in Barnyard millet. *Karnataka J. Agric. Sci.* 18:101e105.

Veeranagamallaiah, G., Puli, C., Jyothsnakumari, G., and Sudhakar, C. (2007). Glutamine synthetase expression and pyrroline-5-carboxylate reductase activity influence proline accumulation in two cultivars of foxtail millet (*Setaria italica* L.) with differential salt sensitivity. *Environ. Exp. Bot.* 60, 239–244. doi: 10.1016/j.envexpbot.2006.10.012

Vijayakumar, T. P., Mohankumar, J. B., and Jaganmohan, R. (2009). Quality evaluation of chapati from millet flour blend incorporated composite flour. *Indian J. Nutr. Diet.* 46, 144–155.

Wallace, G. J., Upadhyaya, H., Vetriventhan, M., Buckler, E., Hash Charles, J., and Ramu, P. (2015). The genetic makeup of a global barnyard millet germplasm collection. *Plant Genome* 08, 01-07. doi: 10.3835/plantgenome2014.10.0067

Wang, J., Wang, Z., Du, X., Yang, H., Han, F., Han, Y., et al. (2017). A high-density genetic map and QTL analysis of agronomic traits in foxtail millet [*Setaria italica* (L.) P. Beauv.] using RAD-seq. *PLoS One* 12:e0179717. doi: 10.1371/journal.pone.0179717

Wang, J., Wang, Z. L., Yang, H. Q., Yuan, F., Guo, E. H., Tian, G., et al. (2013). Genetic analysis and preliminary mapping of a highly male-sterile gene in

foxtail millet (*Setaria italica* L. Beauv.) using SSR markers. *J. Integr. Agric.* 12, 2143–2148. doi: 10.1016/S2095-3119(13)60392-5

Wanous, M. K. (1990). Origin, taxonomy and ploidy of the millets and minor cereals. *Plant Var. Seeds* 3, 99–112.

Weber, A. (2015). Discovering new biology through RNA-Seq. *Plant Physiol.* 169:01081.2015. doi: 10.1104/pp.15.01081

Wendel, J. F. (2015). The wondrous cycles of polyploidy in plants. *Am. J. Bot.* 102, 1753–1756. doi: 10.3732/ajb.1500320

Xu, W., Di, C., Zhou, S., Liu, J., Li, L., Liu, F., et al. (2015). Rice transcriptome analysis to identify possible herbicide quinclorac detoxification genes. *Front. Genet.* 6:306. doi: 10.3389/fgene.2015.00306

Yabuno, T. (1962). Cytotaxonomic studies on the two cultivated species and the wild relatives in the genus *Echinochloa*. *Cytologia* 27, 296–305. doi: 10.1508/cytologia.27.296

Yabuno, T. (1966). Biosystematic study of the genus *Echinochloa*. *Jpn. J. Bot.* 19, 277–323.

Yabuno, T. (1984). A biosystematic study on *Echinochloa oryzoides* (Ard.) Fritsch. *Cytologia* 49, 673–678. doi: 10.1508/cytologia.49.673

Yabuno, T. (1987). Japanese barnyard millet (*Echinochloa utilis*. Poaceae) in Japan. *Econ. Bot.* 41, 484–493. doi: 10.1007/BF02908141

Yabuno, T. (1996). "Taxonomy and phylogeny of the genus *Echinochloa*," in *Natural History of Genus Echinochloa*, eds T. Yabuno, and H. Yamaguchi, (Tokyo: Daw Elanco/Dow Chemical Company), 16–28.

Yabuno, T. (2001). "Taxonomy and phylogeny of the genus *Echinochloa*," in *Natural History of Genus Echinochloa*, eds T. Yabuno, and H. Yamaguchi, (Tokyo: Zennokyo Shuppan), 15–30.

Yadav, R., and Yadav, V. (2013). Comparative performance of Indian and Japanese barnyard millet cultivars under varied fertility conditions for dual use in Indian Central Himalaya. *Range Manag. Agrofor.* 34, 175–178.

Yadav, S., Gaur, V., Jaiswal, J. P., and Kumar, A. (2014). Simple sequence repeat (SSR) analysis in relation to calcium transport and signaling genes reveals transferability among grasses and a conserved behavior within finger millet genotypes. *Plant Syst. Evol.* 300, 1–8. doi: 10.1007/s00606-014-0982-3

Yamaguchi, H., Utano, A. Y. A., Yasuda, K., Yano, A., and Soejima, A. (2005). A molecular phylogeny of wild and cultivated *Echinochloa* in East Asia inferred from non-coding region sequences of trnT-L-F. *Weed Biol. Manag.* 5, 210–218. doi: 10.1111/j.1445-6664.2005.00185.x

Yang, X., Yu, X.-Y., and Li, Y. F. (2013). De novo assembly and characterization of the barnyardgrass (*Echinochloa crus-galli*) transcriptome using next-generation pyrosequencing. *PLoS One* 8:e69168. doi: 10.1371/journal.pone.0069168

Yasuda, K., Yano, A., Nakayama, Y., and Yamaguchi, H. (2006). Molecular identification of *Echinochloa oryzicola* Vasing. and *E. crus-galli* (L.) Beauv. using a polymerase chain reaction–restriction fragment length polymorphism technique. *Weed Biol. Manag.* 2, 11–17. doi: 10.1046/j.1445-6664.2002.00041.x

Ye, C.-Y., Lin, Z., Li, G., Wang, Y., Qiu, J., Fu, F., et al. (2014). Echinochloa chloroplast genomes: insights into the evolution and taxonomic identification of two weedy species. *PLoS One* 9:e113657. doi: 10.1371/journal.pone.0113657

Yoshitsu, Y., Takakusagi, M., Abe, A., Takagi, H., Uemura, A., Yaegashi, H., et al. (2017). QTL-seq analysis identifies two genomic regions determining the heading date of foxtail millet. *Setaria italica* (L.) P.Beauv. *Breed. Sci.* 67, 518–527. doi: 10.1270/jsbbs.17061

Yuichiro, M., Shinya, U., and Yamaguchi, H. (1999). Identification of polyploid groups in the genus Echinochloa by isozyme analysis. *J. Weed Sci. Technol.* 44, 205–217. doi: 10.3719/weed.44.205

Zee, S. Y., and O'brien, T. P. (1971). Aleurone Transfer Cells and other Structural Features of the Spikelet of Millet. *Aust. J. Biol. Sci.* 24, 391–396.

Zegada-Lizarazu, W., and Iijima, M. (2005). Deep root water uptake ability and water use efficiency of pearl millet in comparison to other millet species. *Plant Prod. Sci.* 8, 454–460. doi: 10.1626/pps.8.454

Zhang, G., Liu, X., Quan, Z., Cheng, S., Xu, X., Pan, S., et al. (2012). Genome sequence of foxtail millet (*Setaria italica*) provides insights into grass evolution and biofuel potential. *Nat. Biotechnol.* 30, 549–554. doi: 10.1038/nbt.2195

Zhang, K., Fan, G., Zhang, X., Zhao, F., Wei, W., Du, G., et al. (2017). Identification of QTLs for 14 agronomically important traits in *Setaria italica* based on SNPs generated from high-throughput sequencing. *G3* 7, 1587–1594. doi: 10.1534/g3.117.041517

Zhang, S., Tang, C., Zhao, Q., Li, J., Yang, L., Qie, L., et al. (2014). Development of highly polymorphic simple sequence repeat markers using genome-wide microsatellite variant analysis in Foxtail millet [*Setaria italica* (L.) P. Beauv.]. *BMC Genomics* 15:78. doi: 10.1186/1471-2164-15-78

Zhao, W., Lee, G. A., Kwon, S. W., Ma, K. H., Lee, M. C., and Park, Y. J. (2012). Development and use of novel SSR markers for molecular genetic diversity in Italian millet (*Setaria italica* L.). *Genes Genom.* 34, 51–57. doi: 10.1007/s13258-011-0102-5

Zou, C., Li, L., Miki, D., Li, D., Tang, Q., Xiao, L., et al. (2019). The genome of broomcorn millet. *Nat. Commun.* 10:436. doi: 10.1038/s41467-019-08409-5

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 Genetics

ORIGINAL RESEARCH
published: 07 July 2020
doi: 10.3389/fgene.2020.00543

Rice Biofortification With Zinc and Selenium: A Transcriptomic Approach to Understand Mineral Accumulation in Flag Leaves

Faustino Adriano Roda[1,2,3], Isabel Marques[3], Paula Batista-Santos[3], Maria Glória Esquível[3], Alexis Ndayiragije[4,5], Fernando Cebola Lidon[6], B. P. Mallikarjuna Swamy[4,5], José Cochicho Ramalho[3,6] and Ana I. Ribeiro-Barros[3,6]**

[1] Ministério de Agricultura e Segurança Alimentar, Instituto de Investigação Agrária de Moçambique, Centro Zonal Noroeste, Lichinga, Mozambique, [2] Universidade Eduardo Mondlane-Centro de Biotechnologia, Maputo, Mozambique, [3] PlantStress&Biodiversity Lab, Forest Research Center (IM, JCR, AIRB) and Linking, Landscape, Environment, Agriculture and Food (PBS, MGE), Instituto Superior de Agronomia, Universidade de Lisboa, Lisbon, Portugal, [4] International Rice Research Institute, Maputo, Mozambique, [5] International Rice Research Institute, Laguna, Philippines, [6] Unidade de Geobiociências, Geoengenharias e Geotecnologias, Faculdade de Ciências e Tecnologia, Universidade NOVA de Lisboa, Caparica, Portugal

OPEN ACCESS

Edited by:
L. Joseph Su,
University of Arkansas for Medical Sciences, United States

Reviewed by:
David Lin,
Cornell University, United States
Mahalingam Govindaraj,
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

***Correspondence:**
José Cochicho Ramalho
cochichor@mail.telepac.pt
Ana I. Ribeiro-Barros
aribeiro@isa.ulisboa.pt

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: 15 December 2019
Accepted: 05 May 2020
Published: 07 July 2020

Citation:
Roda FA, Marques I, Batista-Santos P, Esquível MG, Ndayiragije A, Lidon FC, Swamy BPM, Ramalho JC and Ribeiro-Barros AI (2020) Rice Biofortification With Zinc and Selenium: A Transcriptomic Approach to Understand Mineral Accumulation in Flag Leaves. Front. Genet. 11:543. doi: 10.3389/fgene.2020.00543

Human malnutrition due to micronutrient deficiencies, particularly with regards to Zinc (Zn) and Selenium (Se), affects millions of people around the world, and the enrichment of staple foods through biofortification has been successfully used to fight hidden hunger. Rice (*Oryza sativa* L.) is one of the staple foods most consumed in countries with high levels of malnutrition. However, it is poor in micronutrients, which are often removed during grain processing. In this study, we have analyzed the transcriptome of rice flag leaves biofortified with Zn (900 g ha^{-1}), Se (500 g ha^{-1}), and Zn-Se. Flag leaves play an important role in plant photosynthesis and provide sources of metal remobilization for developing grains. A total of 3170 differentially expressed genes (DEGs) were identified. The expression patterns and gene ontology of DEGs varied among the three sets of biofortified plants and were limited to specific metabolic pathways related to micronutrient mobilization and to the specific functions of Zn (i.e., its enzymatic co-factor/coenzyme function in the biosynthesis of nitrogenous compounds, carboxylic acids, organic acids, and amino acids) and Se (vitamin biosynthesis and ion homeostasis). The success of this approach should be followed in future studies to understand how landraces and other cultivars respond to biofortification.

Keywords: biofortification, flag leaves, rice, RNASeq, selenium, transcriptomics, zinc

INTRODUCTION

The 2030 Agenda of the United Nations brings forward 17 Sustainable Developing Goals among which agriculture lies at the core. However, according to the last report of the Food and Agriculture Organization (FAO et al., 2019), more than 820 million people in the world face hunger and undernourishment and, thus, poor health, particularly in Africa where greater efforts should be made to achieve the Zero Hunger by 2030. Undernutrition and

micronutrient deficiencies account for three million deaths each year being more widespread problems than energy consumption (Prentice et al., 2008).

Zinc (Zn) and selenium (Se) are essential mineral nutrients (Fairweather-Tait et al., 2011). However, the intake of these minerals is deficient for ca. 30% (Zn) and 15% (Se) of the world population (White and Broadley, 2011; Moreno et al., 2013). Zn deficiency affects growth and the immune system, being especially severe for young children and pregnant women (White and Broadley, 2011). Se deficiency also weakens the immune system and has been associated with cardiovascular diseases, cognitive decline, cancer, and AIDS (Rayman, 2002; Jablonska and Vinceti, 2015). Thus, providing adequate contents of these minerals in food has become an important objective to fight "hiden hunger" in vulnerable populations (Hirschi, 2009). In this context, a supplementary and diversified diet is the most straight forward strategy to mitigate Zn and Se deficiencies. However, such strategies have been inefficient and uneffective to be implemented in developing countries were livelihoods are strongly dependent on small-scale agriculture (Hartikainen, 2005; Forsman et al., 2014). In order to overcome this constraint, biofortification offers a fast, reliable, and sustainable solution solution which has been successfully achieved in several crops, such as rice, wheat, and beans (White and Broadley, 2009, 2011; Li et al., 2018; Ramalho et al., 2020).

Rice is one of the world's most important staple crops (FAO, 2016), constituting one of the most important sources of energy and micronutrients for more than half of the world population (Lucca et al., 2006; Muthayya et al., 2014). Both whole and polished rice grains contain low concentrations of Zn (6–28 mg kg^{-1}) and Se (0.3 mg kg^{-1}) (Walter et al., 2008), and agronomic biofortification has been used to increase these two minerals in the grain (Phattarakul et al., 2012; Ram et al., 2016; Mangueze et al., 2018; Lidon et al., 2017, 2019; Ramalho et al., 2020). Rice is also a model cereal species, and its genome was the first (among crops) to be fully sequenced in 2005 (Sasaki, 2005) and fully unified in 2013 (Kawahara et al., 2013). This crop has been thoroughly used to investigate the molecular mechanisms underlying several biological processes related to plant development, metabolism, and senescence from seedlings to grains (Jackson, 2016; Qian et al., 2018; Wang et al., 2018). The role of flag leaves in photosynthesis and nutrient mobilization to the grains has been highlighted by several earlier reports (Zhou et al., 2007; Pang et al., 2009; Sperotto et al., 2009; Tari et al., 2009; Xu et al., 2011). Thus, the identification of the molecular mechanisms involved in mineral transport from flag leaves to grains is of utmost importance to understand the biochemical processes associated with the absorption, translocation, and fixation of minerals, such as Zn and Se.

High−throughput *Omics* technologies such as *genomics*, *transcriptomics*, *proteomics*, *metabolomics*, *lipidomics*, or *interactomics* are nowadays widely used to understand biological systems as a whole (Sauer et al., 2007). Such extensive and integrated approaches allowed great advances in plant research, namely the elucidation of biological processes, such as plant development, plant-environment interactions, genomics-assisted breeding, or the discovery of phytocompounds with application in agriculture, medicine, and in a wide range of industries (Kole et al., 2015; Sundell et al., 2015; Bode et al., 2016; Nützman et al., 2016). Transcriptome analysis of rice flag leaves confirmed their importance in grain filling, namely, in the biosynthesis and translocation of photoassimilates and minerals to the seeds (Narayanan et al., 2007; Sperotto et al., 2009).

In this study, we have analyzed the transcriptional changes in flag leaves associated with the agronomic biofortification of rice with Zn and/or Se using a next-generation RNAseq approach that uses a high yield cultivar that is able to accumulate high levels of Zn and Se when exposed to biofortification treatments through foliar spraying (Mangueze et al., 2018).

MATERIALS AND METHODS

Plant Material and Biofortification Experiments

Experiments were performed in the experimental fields of the International Rice Research Institute (IRRI), located at the Umbeluzi—Instituto de Investigação Agrária de Moçambique (IIAM) in Boane, Mozambique (Lat 26° 3′3.75″S; Long 32°21′56.48″E; Alt 8.8 m), using one rice cultivar, Makassane, containg 9.8 mg Kg^{-1} Zn and 0 mg Kg^{-1} Se in whole grains, under control conditions (that is, without any biofortification treatment with Zn or Se) (Mangueze et al., 2018). For the trials, three blocks of 95 m^2 (15 × 5) each containing three biological replicates were established. Biofortification experiments were performed through single and combined foliar spraying of Zn and Se, using a back sprayer, at the beginning of grain filling, corresponding to Z51 stage in the Zadoks scale, i.e., at an adequate stage to promote the translocation to the grain (Cakmak, 2008; Cakmak et al., 2010). The following doses were applied: (i) 900 g ha^{-1} Zn (applied as zinc sulfate – $ZnSO_4$ $7H_2O$), (ii) 500 g ha^{-1} Se (applied as sodium selenite – Na_2O_3Se), and (iii) 900 g ha^{-1} Zn together with- 500 g ha^{-1} Se. The Zn and Se doses as well as the use of Se-selenite (instead of Se-selenate) were based on previous reports in rice by Phattarakul et al. (2012), Lidon et al. (2019), and, especially, Mangueze et al. (2018), which used the same cultivars and cropped area. Each element was applied twice (using the same volume of the solution), with an interval of 7 days to reach the desired concentration. Control plants received only water.

The basal field fertilization was carried out with NPK (12:24:12) using 100 kg ha^{-1} 26 and 60 days after sowing. Foliar fertilization was made using 50 kg ha^{-1} of NPK (12:24:12) plus 50 kg ha^{-1} urea (46%) with a total of 200 kg ha^{-1} for the two applications.

For RNAseq analysis, flag leaves from three different plants per treatment were harvested 15 days after spraying, stored immediately in RNA Latter (Thermo Fisher Scientific), and frozen once in the lab.

RNA Whole Transcriptome Deep Sequencing

Total RNA was extracted from 100 mg frozen material of each biological replicate per treatment using the inuPREP extraction Kit (Analytik Jena AG) following the manufacturer's instructions. RNA integrity and purity were first evaluated by visual inspection of RNA bands through electrophoresis in a 1.5% agarose–TBE gel containing GelRed Nucleic Acid Gel Stain (Biotium) and then by an Agilent 2100 Bioanalyzer (Agilent). All samples had an RNA integrity number (RIN) higher than 8.9. Library preparation was performed with the TruSeq RNA Sample Prep Kit v2 (Illumina) and RNA-Seq analyzes by Illumina NovaSeq 6000 of 2 × 100 bp pair-end reads (30 million reads per sample) at Macrogen (South Korea).

Alignment and Analysis of Illumina Reads

Raw reads obtained by sequencing were quality-checked using FastQC version 0.11.5 (Andrews, 2010). To reduce biases, artifacts as low-quality reads, adaptor sequences, or contaminant DNA were removed using Trimmomatic version 0.32 (Bolger et al., 2014). HISAT2 version 2.0.5 was used for the mapping of high-quality filtered reads against the reference genome (Os-Nipponbare-Reference-IRGSP-1.0) downloaded from the Rice Genome Annotation Project Database[1] (Kawahara et al., 2013). Known genes and transcripts were assembled with StringTie version 1.3.3b (Pertea et al., 2015, 2016) based on the reference genome model. After assembly, the abundance of gene transcripts was calculated for each sample and normalized as FPKM (Fragments per Kilobase of transcript per Million Mapped reads) using Cufflinks version 2.2.1 (Trapnell et al., 2010). The similarity between samples was obtained through Pearson's coefficient of the Log_2(FPKM+1) value with a range of $-1 \leq r \leq 1$ (the closer the value is to 1, the more similar the samples are) and graphically depicted using a correlation matrix.

Differentially Expressed Gene Analysis

During data preprocessing, low quality transcripts were filtered. Afterward, log2 transformation of FPKM+1 and quantile normalization were performed. To identify the differentially expressed genes (DEGs) from the dataset, an FDR adjusted *P*-value of <0.05 was set and a fold change (FC) of ≥2 was assigned. For significant lists, a hierarchical clustering analysis was performed to group the similarity of transcripts and samples by expression level of normalized values. Standardized expression patterns were visualized as Z-scores in a heatmap generated by hierarchical clustering (function hclust in R). The significant DEGs found were mapped on the 12 chromosomes of rice using the chromosome map tool in Oryzabase database (Yamazaki et al., 2010), and a map was drawn based on output generated. Following Raza et al. (2019), all DEGs with a gene ontology (GO) function related to cation ion binding/transport, heme binding, Se, Zn ion binding/transport, metal ion binding, transport, and homeostasis were filtered and considered as putative candidate genes (CGs) associated with traits of interest in biofortification.

[1]http://rice.plantbiology.msu.edu/

Functional Annotation, Enrichment and Pathway Analysis

To assign functional categories to the DEGs, a gene-set enrichment analysis was performed using the DAVID 6.7 database for annotation, visualization, and integrated discovery, an online tool for the analysis of the relevant biological annotation of gene lists[2]. The significant DEGs were annotated for GO terms and categorized into biological process (BP), molecular function (MF), and cellular component (CC). For significant DEGs, a gene enrichment test was then performed using the DAVID default background, representing the corresponding genes with at least one annotation in the analyzing categories in the enrichment calculation. *P*-value for enrichment was calculated for each GO term represented and corrected via Bonferroni family-wise error rate (FWER) method. Only the GO terms exhibiting a corrected *P*-value of <0.05 were considered to be significantly enriched for a given set of genes. To investigate which DEGs were activated or suppressed in different class of pathways, gene expression information was mapped using the Kyoto Encyclopedia of Genes and Genomes, KEGG[3]. Pathway images were generated using the online tool KEGG Mapper-Colour Pathway[4]. Raw and processed RNA-sequencing data have been deposited in NCBI.

RESULTS

Overall Transcriptome Profiling and Mapping Statistics

In total, the 24 RNA libraries generated an average of 38 million reads with a GC content of 53.28% (**Table 1**). An average of 1.64% of the reads were removed after being trimmed. The vast majority (98.36%) of the total number of reads was mapped to the reference rice genome demonstrating a high coverage over the transcriptome. From these, an additional average of 7.55% reads could not be mapped into the reference genome. Statistics of each sample are provided in detail in **Supplementary Table S1**. After trimming and cleaning, a total of 34 million reads were analyzed. A high similarity was found between samples through Pearson's coefficient of the Log_2(FPKM+1) value (**Figure 1**).

Differentially Expressed Genes

We identified 3170 genes that were differentially expressed between biofortified and control plants, of which only 224 were significantly different (FDR < 0.05) and had a FC ≥2 (**Figure 2**). Hierarchical clustering analysis of all DEGs showed no specific trends in expression convergence (**Figure 3**), suggesting that biofortification with Se, Zn, and Se-Zn led to different changes in DEGs.

Biofortification with Zn alone triggered a higher number of DEGs (106) than in combination with Se (72) and even less when only Se was used (46) (**Figure 4A**). DEGs were usually

[2]https://david.ncifcrf.gov/home.jsp

[3]http://www.genome.jp/kegg/

[4]http://www.genome.jp/kegg/tool/map_pathway3.html

TABLE 1 | Overview of RNA-Seq data of rice cultivar Makassane (Mak) in control conditions (Ctr) and under different treatments of Selenium (Se500: 500 g ha^{-1}) and Zinc (Zn900: 900 g ha^{-1}) or with these two elements together (Zn-Se).

Genotype	Setup	Total reads	GC (%)	Processed	Mapped	Unmapped
Mak	Ctr	36,821,836	54.03	36,205,938 (98.32%)	33,869,031 (93.55%)	2,336,906 (6.45%)
Mak	Se500	34,451,820	52.81	33,907,595 (98.42%)	30,784,505 (90.79%)	3,123,090 (9.21%)
Mak	Zn900	39,822,173	52.91	39,162,976 (98.36%)	36,396,575 (92.94%)	2,766,401 (7.06%)
Mak	Zn-Se	41,069,693	53.35	40,394,978 (98.42%)	37,382,433 (92.54%)	3,012,544 (7.46%)
Average		38,041,381	53.28	37,417,872 (98.36%)	34,508,136 (92.45%)	2,809,735 (7.55%)

Numbers indicate the average of the three biological replicates concerning the total number of reads after trimming (total reads), GC content (GC%), number of cleaned reads after trimming (processed), number of reads mapped to the reference (mapped), and number of reads that failed to aligned (unmapped). Statistics of each sample are provided in detail in ***Supplementary Table S1****.*

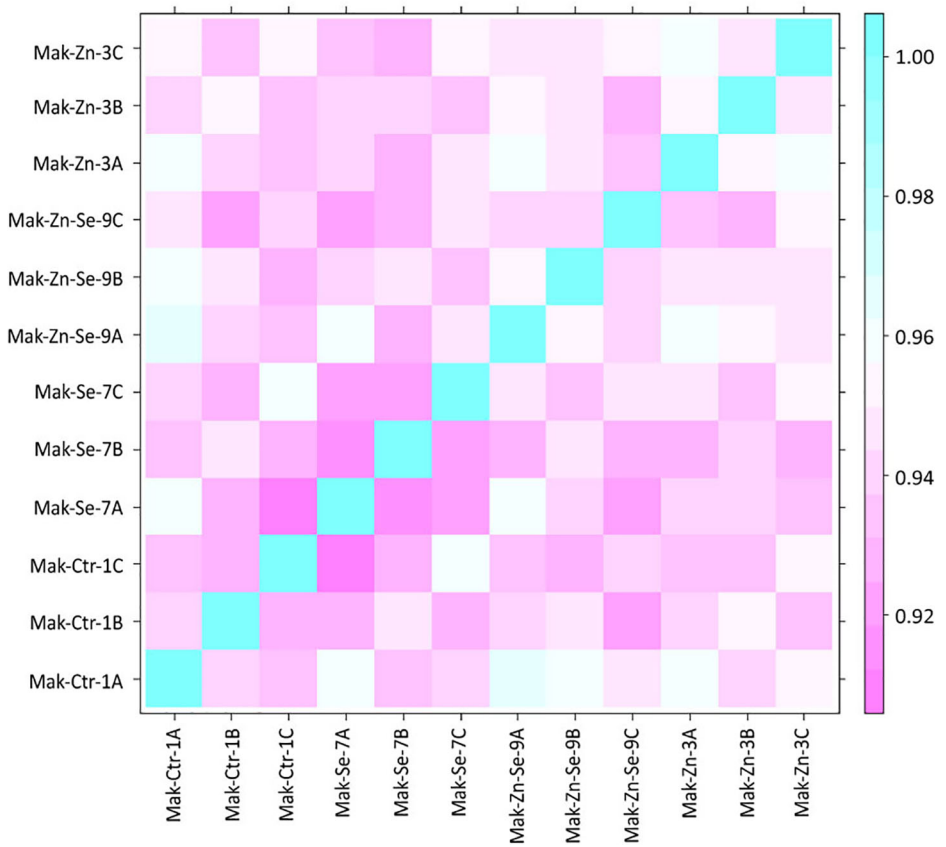


FIGURE 1 | Correlation matrix for all samples obtained through Pearson's coefficient of the Log2(FPKM+1) value. The closer the value is to 1, the more similar the samples are.

upregulated in the Zn-biofortified and the Zn-Se-biofortified pool (61 and 54%, respectively), while Se-biofortified flag-leaves expressed an equal number of up- and downregulated DEGs (23%) (**Figure 4B** and **Supplementary Tables S2–S4**).

The type of DEGs and the level of FC also varied between the three biofortified pools. For instance, distribution trends in terms of FC ranged from *ca.* −10 to 54 for DEGs in the Se-biofortified pool (**Supplementary Table S2**). Two genes were found to be highly enriched with FC 4–5 times higher than for the remaining genes: ataxin-2 C-terminal region family protein (Os03g0180300) and a CBL-interacting protein kinase 16 (Os09g0418000). Meanwhile, GFA2 (Os06g0116800) and a hypothetical protein (Os03g0180300) were the top downregulated genes. From those 46 genes, 7 were putative CGs associated with traits of interest in Se-biofortification: Biotin synthase (Os08g0540100), Similar to inducible alpha-dioxygenase (Os12g0448900), Queuine tRNA-ribosyltransferase (Os09g0469900), Thiazole biosynthetic enzyme 1^{-1} (Os07g0529600), Cytochrome P450 (Os02g0221900), Zinc finger (Os01g0667700), and Protein phosphatase 2C domain (Os05g0358500).

In comparison, FC of DEGs from flag leaves of Zn-biofortified plants varied between *ca.* −13 to 61

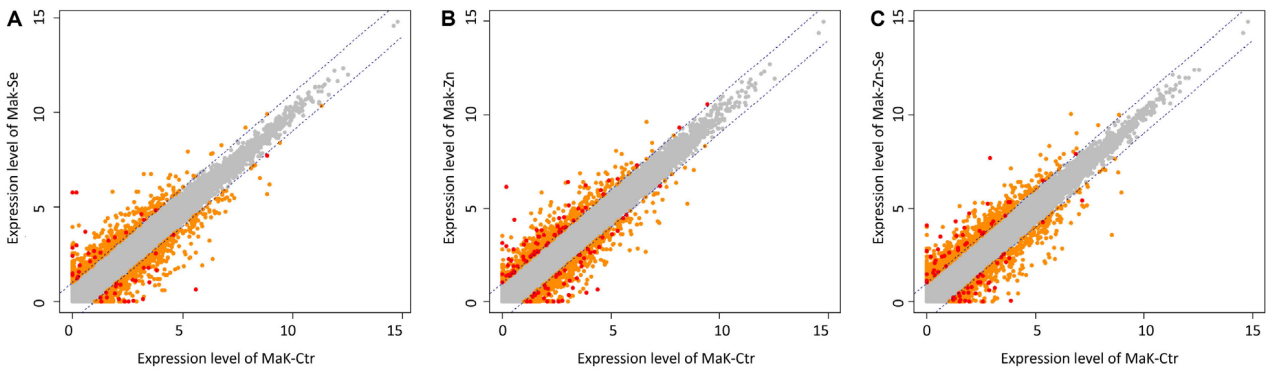


FIGURE 2 | Scatter plots of expression levels between control and the average normalized value of the Zn-biofortified pool **(A)** between control and Se-biofortified pool **(B)** and between control and Zn-Se-biofortified pool **(C)**. Transcripts levels significantly above control levels are upregulated in response to biofortification while those below control levels are downregulated. Each plot contains all identified transcripts, including non-significant (gray dots) as well as specific DEGs that are considered significant under |FC| ≥ 2 (orange dots) and under |FC| ≥ 2 and $P < 0.05$ (red dots).

(**Supplementary Table S3**). Top downregulated included a Resistance protein candidate (Os05g0318700) and a Photosystem I protein-like (Os07g0148900) encoding genes, while a 2-oxoglutarate dehydrogenase E2 subunit gene (Os04g0394200) was found to be highly upregulated, with an FC six times higher than the remaining DEGs. From those 106 genes, eight were considered putative CGs associated with traits of interest in Zn-biofortification: NifU-like protein (Os01g0662600), WD-40 repeat containing protein (Os02g0791800), Laccase-6 (Os01g0850550), Ribose-phosphate pyrophosphokinase 3 (Os01g0723600), C-type lectin domain (Os01g0104000), Homodimeric diiron-carboxylate protein (Os04g0600300), Queuine tRNA-ribosyltransferase (Os09g0469900), and Cytochrome P450 (Os02g0173100).

In the set of plants biofortified with the two elements (Zn-Se), FC of DEGs varied between ca. -13 to 27 (**Supplementary Table S4**). A DRE binding protein 2 (Os01g0733801) was found to be the top downregulated gene, while a cystathionine gamma-synthase (CGS) (Os03g0376100) and an UDP-glucuronosyl/UDP-glucosyltransferase (UGT) family protein genes (Os01g0736100) were the top upregulated genes under Zn-Se biofortification. From those 72 genes, eight could be CGs associated with traits of interest in Zn-biofortification: Adaptin ear-binding coat-associated protein 2 (Os10g0476000), Endoribonuclease Dicer homolog 2a (Os03g0583900), Thioredoxin family Trp26 (Os01g0559000), Cytochrome family protein (Os11g0151400), Syntaxin 6 (Os08g0244100), Cytochrome P450 (Os02g0173100), Queuine tRNA-ribosyltransferase (Os09g0469900), and Secretory carrier membrane protein (Os04g0597000). Interestingly, one gene (Os03g0103300) from the Se-Zn biofortified pool was a quantitative trait loci (QLT) G-3-1 protein, targeted for low-temperature germinability (**Supplementary Table S4**).

Only six DEGs were shared between the three biofortified pools: Eukaryotic initiation factors 3 (Os04g0112300) and 4 (Os04g0112300), Cytochrome P450 (Os02g0173100 and Os02g0221900), UDP-glucuronosyl/UGT (Os01g0736100), and Queuine tRNA-ribosyltransferase (Os09g0469900). Additionally, an Ethylene response factor 2 gene (Os07g0617000) and two Hypothetical proteins (Os01g0358300 and Os07g0536966) were commonly upregulated in the Zn- and in the Se-biofortified pool (**Supplementary Tables S2, S3**). Genes encoding a Prolin-rich protein (Os04g0612500) and a Chitinase-like protein (Os09g0494200) (both downregulated) were found to be commonly expressed between the Zn- and the Zn-Se- biofortified pool (**Figure 4B** and **Supplementary Tables S3, S4**).

DEGs were unevenly distributed among the 12 rice chromosomes being predominant on chromosome 1 (with 32 DEGs), chromosome 3 (with 23 DEGs), and chromosome 4 (with 22 DEGs; **Supplementary Figure S1**). Few DEGs could be mapped on chromosomes 11 (3 DEGs, none from Se-biofortification) and 10 (5 DEGs) and on chromosome 12 (7 DEGs). DEGs from biofortification with Se, Zn, and Se-Zn were predominantly mapped on chromosome 1 (7 DEGs), chromosomes 1 and 4 (13 DEGs), and chromosome 1 (10 DEGs), respectively.

Gene Ontology Annotation of DEGs

Gene ontology categories from the list of significant DEGs were overall upregulated (although at different levels) except for "metal ion binding" in the Zn-biofortified pool (**Figure 5**). GO categories also showed opposite profiles in the three biofortified pools, which corroborates the differential gene regulations (**Figure 5**). BP such as "vitamin metabolic process" and "vitamin biosynthetic process" were enriched in the Se-biofortified pool while "nitrogen compound biosynthetic process," "carboxylic acid biosynthetic process," "organic acid biosynthetic process," "cellular amino acid biosynthetic process," and "chitin metabolic process" were the in the Zn-biofortified pool. In contrast, BP categories such as "carbohydrate catabolic process," "lipid localization," and "lipid transport" were enriched after Zn-Se biofortification. MF such as "cation binding," "ion binding," and "transition metal ion binding" were enriched after Se-biofortification. Categories as "metal ion binding," "cofactor binding," and "coenzyme binding" were enriched after Zn-biofortification, the two latter categories also enriched after biofortification with the two

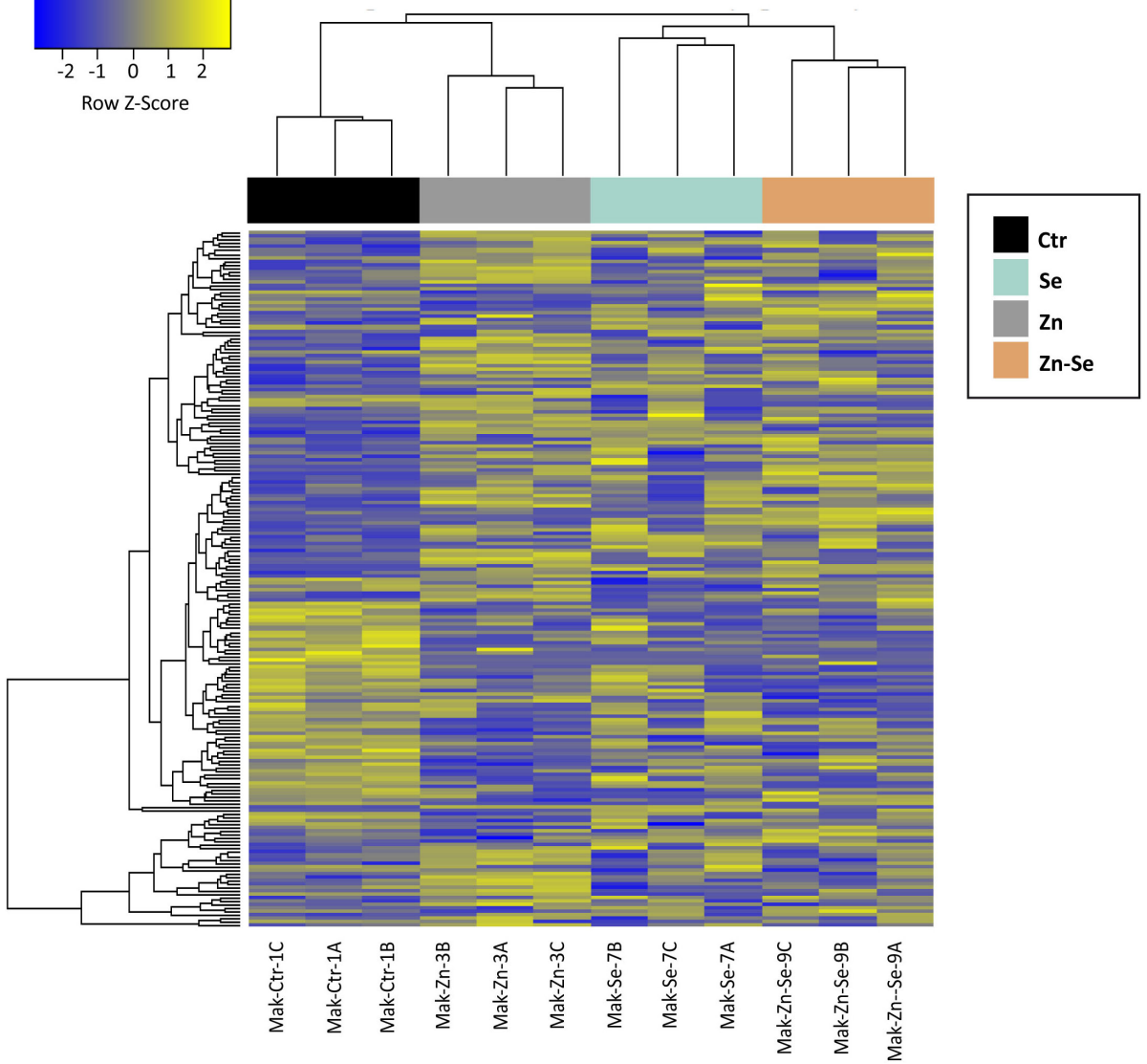


FIGURE 3 | Heatmap of rice cultivar Mak differentially expressed genes after biofortification with single and combined Zn and Se and application. Expression values are depicted as Z-standardized scores for each DEG, where blue represents downregulated DEGs and yellow upregulated DEGs.

elements together (**Figure 5**). Cellular components (CC) such as "cytoplasmic membrane-bounded vesicle," "cytoplasmic vesicle," and "ubiquitin ligase complex" were only enriched after Zn-Se biofortification (**Figure 5**).

Effect of Biofortification on Biological Pathways

Photosynthesis was the only biological pathway involving downregulation of gene expression and only in the biofortification with Zn. Five different genes, all involving Photosystem I (*PsaD*, *PsaE*, *PsaF*, *PsaG*, and *PsaH*), were significantly downregulated after Zn-biofortification ($P < 0.01$; **Figure 6**). By contrast, four different biological pathways involving upregulated DEGs were significantly enriched after biofortification: the citrate cycle (TCA cycle) from the carbohydrate metabolism and the RNA degradation pathway (respectively, $P < 0.001$ and $P < 0.05$), while the vitamin metabolic pathway involving the production of thiamine and biotin were enriched after Se-biofortification ($P < 0.01$ in both pathways; **Figure 7**).

Six different upregulated genes were found to be enriched in the tricarboxylic acid (TCA) cycle: one isocitrate dehydrogenase (1.1.1.42), two 2-oxoglutarate dehydrogenase E1 component (1.2.4.2), 2-oxoglutarate dehydrogenase E1 component (2.3.1.61),

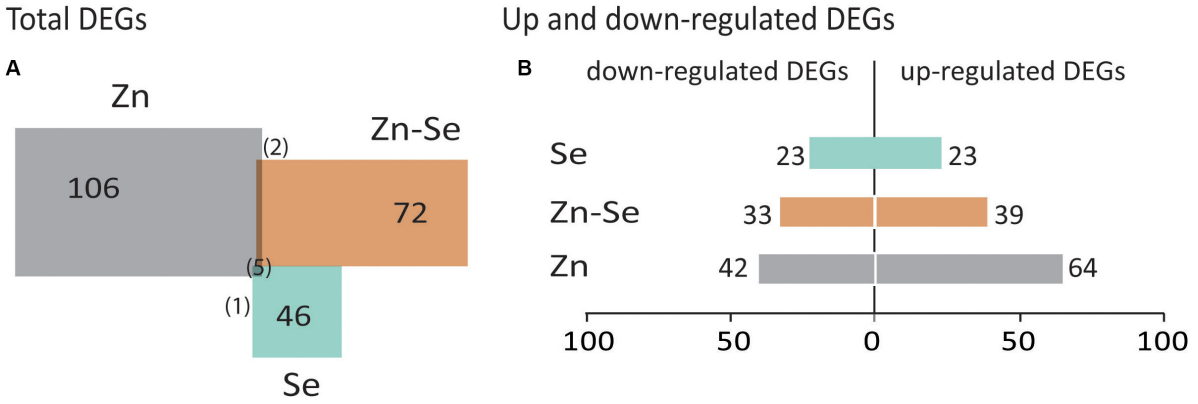


FIGURE 4 | Total number of differentially expressed genes (DEGs). Common DEGs between comparisons are indicated between brackets **(A)**. Up- and downregulated DEGs between comparisons with $|FC| \geq 2$ and $p < 0.05$ **(B)**. Comparisons indicate the number of DEGs found in biofortified flag-leaves with Se, Zn, or Zn-Se in comparison with control conditions. Colors of the biofortified pools are the same as in **Figure 3**.

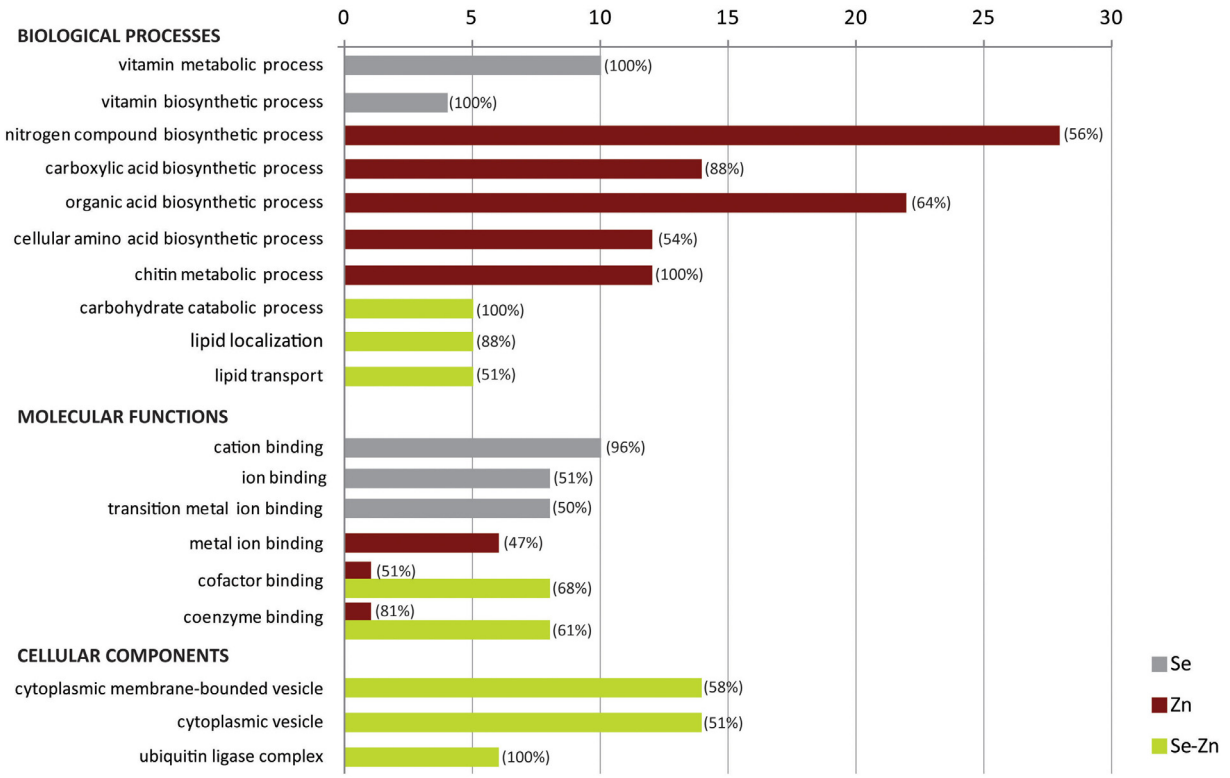


FIGURE 5 | Gene ontology enrichment analysis of biological processes (BP), molecular function (MF), and cellular component (CC) for up- and downregulated genes between biofortified flag leaves with single and combined Zn and Se and application in comparison with control conditions. The percentage of upregulated DEGs found within each category is indicated between brackets.

and two 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductase subunit alpha (1.2.7.11 and 1.2.7.3; **Figure 8**). Three genes associated with the pathway of RNA degradation were also found after biofortification with Zn: ATP-dependent RNA helicase DDX6/DHH1 (DDX6), enhancer of mRNA-decapping protein 3 (EDC3), and enhancer of mRNA-decapping protein 4 (EDC4) (**Supplementary Figure S2**). In comparison, the thiamine metabolism pathway was enriched after Se-biofortification involving the upregulation of nucleoside-triphosphatase (3.6.1.15), thiamine-monophosphate kinase (2.7.4.16), and adenylate kinase (2.7.4.3), while the biotin metabolism pathway was enriched after Se-biofortification directly through the upregulation of biotin synthase (2.8.1.6) (**Supplementary Figure S3**).

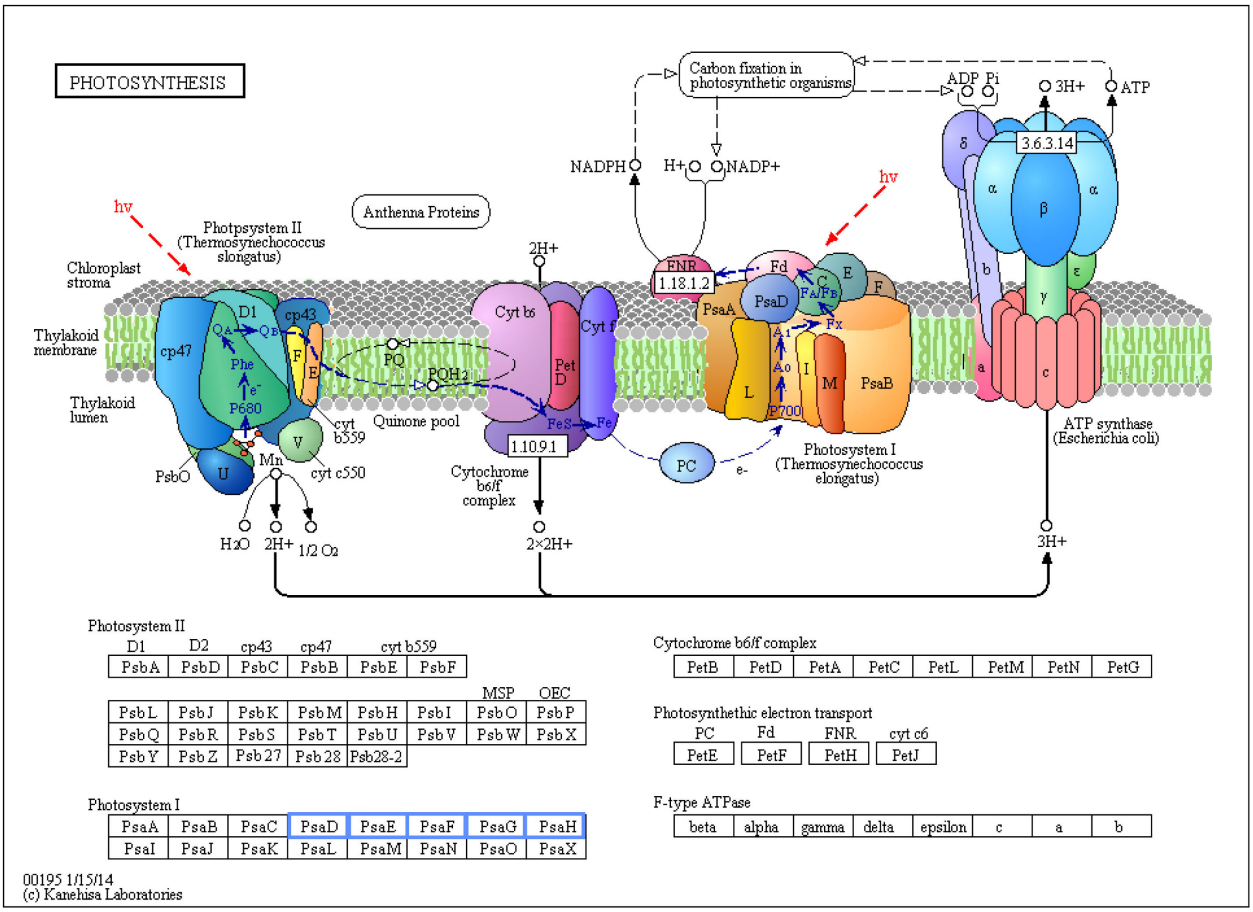
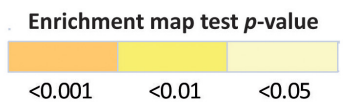


FIGURE 6 | Significant differentially expressed genes (DEGs) involved in photosynthesis of Zn-biofortified rice cultivar Mak ($P < 0.01$). Genes significantly downregulated by Zn-biofortification are shown in blue boxes. White boxes indicate non-responsive genes.



Metabolism	Pathway identifiers	Se	Zn
Carbohydrate metabolism	Citrate cycle		
Vitamin metabolism	Thiamine metabolism		
	Biotin metabolism		
Folding, sorting and degradation	RNA degradation		

FIGURE 7 | Heatmap showing significant KEGG pathways involving upregulated DEGs after Se and Zn biofortification. Colors indicate the level of significance.

DISCUSSION

Rice is both a model plant species and one of the most important staple foods worldwide. It is, however, a poor source of micronutrients, such as Zn and Se, whose deficiency has several impacts on human health and child growth, particularly in developing countries. To compensate nutrient-poor staples, genetic and agronomic biofortification strategies have been widely used to fight hidden hunger (Kondwakwenda et al., 2018; Neeraja et al., 2018; Hefferon, 2019; Zhou et al., 2019).

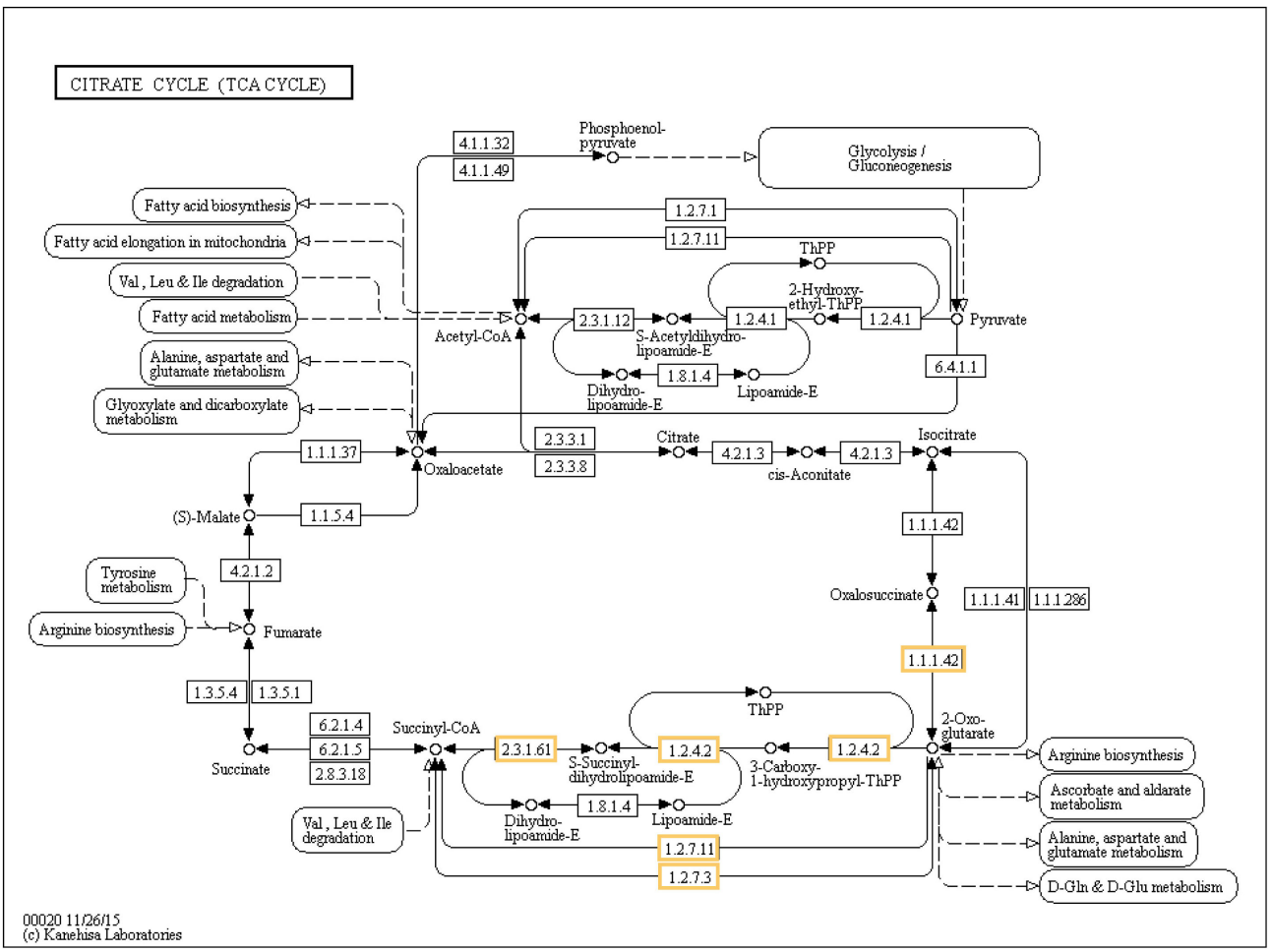


FIGURE 8 | Significant differentially expressed genes (DEGs) involved in the TCA cycle of Zn-biofortified rice cultivar Mak ($P < 0.001$). Genes significantly upregulated by Zn biofortification are shown in yellow boxes. White boxes indicate non-responsive genes.

While genetic biofortification approaches are powerful tools for cereals (including rice) biofortification, the extended time frame and resources needed for conventional breeding as well as the economic, ethic, and legal issues associated with genetically modified crops are among the main limiting factors, and, in both cases, mineral availability is largely dependent on the soil properties (Cakmak, 2008). On the other hand, agronomic biofortification constitute a useful and reliable short-term strategy, complementary to conventional breeding (Cakmak, 2008; White and Broadley, 2009, 2011; Ramalho et al., 2020).

However, the transcriptional basis of the biofortification process at the leaf level is still poorly understood (Sperotto et al., 2013; Neeraja et al., 2018; Tang et al., 2018). In this context, we have analyzed the transcriptome of rice flag leaves to gain insight about their contribution to mineral mobilization to grains. For that we have used a rice cultivar (Makassane) with high-yield (6–7 tonnes per hectare), high grain quality, and resistance to the two major rice diseases (bacterial leaf blight and blast) that has been bred by the IRRI and adapted to the irrigated agro-ecological conditions of Mozambique, Southern Africa (Singh et al., 2013), that have a good perspective for future commercialization. Although a certain degree of Zn-Se antagonism has been referred in some rice cultivars, in this specific genotype co-application of Zn and Se did not interfere with Zn accumulation while also promoting Se accumulation in grains (Mangueze et al., 2018).

Transcriptome analysis of flag leaves from rice plants biofortified with Zinc (Zn; 900 g Zn ha^{-1}), Selenium (Se; 500 g Se ha^{-1}), and both minerals (Zn-Se; 900 g Zn ha^{-1} and 500 g Se ha^{-1}) identified a total of 34 million reads, which is in line with the results obtained from other transcriptome studies on biofortified rice (Neeraja et al., 2018), wheat (Mishra et al., 2019), and maize (Yi et al., 2019). The number of DEGs was *ca.* 3000, which is also in agreement with other transcriptome projects in rice, where *ca.* 1000–3000 DEGs have been reported (e.g., Li et al., 2018; Cao et al., 2019; Chen et al., 2019; Sun et al., 2019). The response of the rice transcriptome is quite variable, depending on the organ, cultivar, and environmental conditions. For instance, in panicles of two Indian landraces (Chittimutyalu, CTT and Kala Jeera Joha, KJJ) and one improved variety (BPT

5204), with differential ability to uptake zinc, the total number of DEGs varied from *ca.* 1400 (BPT 5204) to *ca.* 2200 in CTT and KJJ, of which *ca.* 450–800 with significant FC (Neeraja et al., 2018). On the other hand, in roots of a Chinese rice cultivar (Nipponbare) subjected to Cadmium (Cd) stress ca. 1200 DEGs were identified, 226 with significant FC at low and 1162 at high Cd concentrations, 10 μM and 100 μM, respectively (He et al., 2015). In our study, the rice biofortification of cultivar Makassane triggered a relatively low number of DEGs (224) with significant FC. Such a low number might be anchored in the following arguments: (i) molecular changes associated with the biofortification processes are narrow and limited to specific pathways; and/or (ii) slight (statistically non-significant) changes in gene expression are enough to trigger the pathways involved in mineral absorption and translocation; and/or (iii) post-transcriptionally regulated events.

The number of significant DEGs was higher in flag leaves of Zn-biofortified plants (106) than those from Se- and Zn-Se-biofortified plants (72 and 46, respectively) (**Figure 4**). Differences in DEGs were extensive to FC (−10 to 54 for Se, −13 to 61 for Zn, and −13 to 27 for Zn-Se) as well as to the gene classes (**Supplementary Tables S2–S4**). These differences are in line with the reported variation of the rice transcriptome referred above (He et al., 2015; Neeraja et al., 2018) and might be also related to the fact the mineral concentrations used for spraying were higher for Zn than for Se, triggering greater changes related to more demanding changes during Zn accumulation.

Gene ontology analysis of DEGs showed also different profiles of the three sets of biofortified samples, corroborating the expression patterns of the individual set of DEGs (**Figure 5**). Thus, while, in the case of Se biofortification, vitamin biosynthesis and metabolism were among the main BP, in the case of Zn, the main BP included biosynthesis of nitrogen compounds, carboxylic acids, organic acids, amino acids, as well as chitin metabolism. Additionally, in the case of the combined Zn and Se application, the BP associated with carbohydrate metabolism and lipid dynamics were also enriched. Such distinct GO profiles are likely related to the specific roles of Se and Zn in plant functioning as well as with specificities of Zn and Se biofortification processes, even considering a certain antagonism extent between these two minerals as regards their accumulation potential (Boldrin et al., 2012; Mangueze et al., 2018). This hypothesis is supported by the classification based on the MF, i.e., ion homeostasis in Se-biofortified flag leaves, probably related to the role of this mineral in controlling oxidative stress, the coenzyme/cofactor function of Zn in Zn-biofortified flag leaves, and a rather diverse set of GOs in the double biofortified flag leaves (Sunde, 2018). Nevertheless, the three biofortification processes have in common alterations in cell metabolism inherent of the mineral mobilization and their specific cellular functions (Garg et al., 2018).

Among the top DEGs, two genes encoding an ataxin-2 and a CBL (Calcineurin B-like)-interacting protein kinase (CIPK) were the most abundant in Se-biofortified flag leaves. Ataxin-2 is a CID (CTC-interacting domains) protein highly conserved among eukaryotes (Jiménez-López and Guzmán, 2014). Although the functional characterization of ataxins in plants has not yet been fully addressed, key and evolutionary conserved roles of this group of proteins have been proposed, namely, in post-transcriptional regulatory assembly in many biological processes, including, growth, development, and environmental responses (Jiménez-López et al., 2015; Ostrowski et al., 2017). Thus, it is possible that Se-biofortification in rice is highly regulated post-transcriptionally through Ataxin-2, explaining the low number of DEGs identified in this pool of samples. On the other hand, the activation of the plant-specific CBL-CIPK complex is likely related to the induction of a plant stress signaling cascade (Ligaba-Osena et al., 2017; Zhang et al., 2018; Aslam et al., 2019; Liu et al., 2019) necessary to mobilize Se. In fact, although its role in plants remains controversial, Se is known to act in mechanisms of plant protection against a variety of abiotic stresses, such as cold, drought, desiccation, and metal stress (Gupta and Gupta, 2017). Furthermore, several studies have also found that CBL-CIPK pathways work as regulators in nutrient transport systems, namely, Na^+ (Shi et al., 2002), K^+ (Xu et al., 2006), Mg^{2+} (Tang et al., 2015), and NO_3^- (Ho et al., 2009).

With Se biofortification, the induction of vitamin metabolic pathways related to the production of the thiamine and biotin were enriched (**Figure 7**), the first involving the transcriptional activation of genes encoding a nucleoside-triphosphatase, a thiamine-monophosphate kinase, and an adenylate kinase and the second the upregulation of a biotin synthase (**Supplementary Figure S3**). Consequently, biofortification with Se might reinforce thiamine presence in rice, which is usually very poor, ranging from 0.053 to 3 mg per 100 g of grain. These values become even lower with the elimination of the aleurone layer (where thiamine is predominantly stored) in polished rice as well in cooked rice (Minhas et al., 2018).

With regards to the flag leaves from Zn-biofortified plants, the most abundant transcript corresponded to a gene enconding a 2-oxoglutarate (2-OG) dehydrogenase (2-ODD), the largest family of non-heme oxidizing enzymes (Kawai et al., 2014), with a key function in the TCA cycle that uses 2-OG as an obligatory substrate (Scheible et al., 2000). This will link the TCA cycle (and thus ATP production) to amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis (Araújo et al., 2014). Taking into account the wide Zn roles in plants, which include the composition of proteins and other macromolecules, its action as a functional, structural, or regulatory cofactor of a large number of enzymes, and its role in gene expression control (Brown et al., 1993), the huge activation of 2-ODD might be related to the enhanced cell metabolic activity induced by Zn biofortification. Such large accumulation of 2-ODD transcripts is also in line with the GO analysis and its involvement in several biosynthetic pathways (Araújo et al., 2014; Farrow and Facchini, 2014; Wang et al., 2019).

In line with the GO analysis and with the proposed roles for Zn in plant cells, the analysis of the impact of Zn biofortification in biological pathways confirms the relation with the TCA cycle and carbohydrate metabolism but also with RNA degradation, highlighting its role in cell metabolism and control of gene expression (Brown et al., 1993). The genes encoding enzymes from the TCA cycle included an isocitrate

dehydrogenase, three 2-oxoglutarate dehydrogenases, and two 2-oxoglutarate/2-oxoacid ferredoxin oxidoreductases involved in plant defense against stress (Araújo et al., 2014; Yu et al., 2018). The genes associated with RNA degradation included an ATP-dependent RNA helicase (DDX6), and two enhancers of mRNA-decapping proteins 3 and 4 (EDC3 and EDC4), factors regulating mRNA turnover in plants probably involved in developmental processes (Goeres et al., 2007) as well as in stress responses (Kawa and Testerink, 2017).

On the other hand, a Photosystem I (PSI) gene was among the top downregulated DEGs. A deeper analysis of biological pathways indicates that Zn biofortification resulted in a downregulation of several photosynthesis-related genes, namely, those related with PSI (*PsaD*, *PsaE*, *PsaF*, *PsaG,* and *PsaH*; **Figure 6**). Nevertheless, this downregulation of PSI related genes is likely linked to metabolic adjustments needed for Zn mobilization, since in our experiments Zn biofortification did not negatively affect plant growth and yield (our lab, unpublished data), and because a positive impact on photosynthesis has been frequently reported in Zn fertilized/biofortified cereals (e.g., Lidon et al., 2015, 2017; Liu et al., 2016; Chattha et al., 2017). This is in agreement with the fact that most Zn is taken up by active transport and that the energy demanded by this process is largely supported by photosynthesis (Nakandalage et al., 2016).

In double biofortified plants (Zn and Se), genes enconding a CGS and an UDP-glucuronosyl/UGT were transcriptionally activated. These genes are usually involved in amino acid (methionine) (Zhao et al., 2018) and flavonoid (Yin et al., 2017) biosynthesis, suggesting the activation of these two pathways during double biofortification conditions. CGS is specifically associated with perturbation of the folates pool that serve as donors and acceptors in one-carbon (C1) transfer reactions, which are essential in major metabolic pathways, such as amino acids, nucleic acids, and vitamin B5 (Loizeau et al., 2017). Crops that are able to *de novo* synthesize folates serve as an excellent dietary source, which is not the case of most staple crops, such as rice, potato, and maize (Gorelova et al., 2017). Thus, biofortification with Zn-Se could have a positive impact regarding the enhancement of folate content, further improving the nutritional value of rice, a result that should be tackled in future studies. One DEG (Os03g0103300) from the Se-Zn biofortification was a QLTG-3-1 protein targeted for low-temperature germinability detected on chromosome 3 (**Supplementary Table S4**), emphasizing the importance of this gene for further molecular characterization of biofortification traits. Improvement of cold tolerance at the germination stage is a major determinant for rice cultivation in tropical or subtropical areas causing severe reductions in the final yield and in crop productivity (Jiang et al., 2017). Further identification of major QTLs associated with the molecular basis of micronutrient uptake/homeostasis (Swamy et al., 2016; Raza et al., 2019) will facilitate the identification of genes of interest and its exploration for breeding Se and Zn rich rice varieties. Six common DEGs were shared between the three biofortified pools, interestingly with similar FC trends in Se and Zn biofortified flag leaves and opposite expression patterns in the Zn-Se pool (**Supplementary Tables S2–S4**). Among them, eukaryotic initiation factor (eIEF) 3 and eIEF 4 were, respectively, down- and upregulated in the Se (FC −6.528 and 6.860) and Zn (FC −6.452 and 6.813) set, and up- and downregulated in the Zn-Se set (FC 2.378 and −2.869). eIEFs are large protein complexes that participate in translation initiation (Yahalom et al., 2008). As referred above, rice biofortification significantly activate a specific and limited set of DEGs, and, thus, it is expected that the set of genes involved in translation is also limited. This is supported by the expression patterns of a queuine tRNA-ribosyltransferase, i.e., negative FC of −2.077 in Se, −2.245 in Zn and −2.303 in Zn-Se, involved in the biosynthesis of tRNA subunits (Zallot et al., 2014) and known to be inhibited by Zn (Schomburg and Schomburg, 2010). The expression pattern of a Cytochrome P450 (CYP) gene followed also the same trend of the eIEF genes, i.e., FC of 6.860 in Se, 6.813 in Zn and −2.869 in Zn-Se, highlighting its importance in monooxygenation/hydroxylation reactions in biochemical pathways (Wei and Chen, 2018) of Se and Zn biofortified plants. The fact that an ethylene response factor 2 was commonly upregulated in the Zn- and in the Se-biofortified pool (**Supplementary Tables S2, S3**) may imply that Zn and Se are sensed by the plant as a stress condition, leading to ethylene production, which is similar to what happens in Fe-biofortification (dos Santos et al., 2017). DEGs were randomly distributed among the 12 rice chromosomes being predominant on chromosomes 1, 3, and 4, while few could be mapped on chromosomes 10, 11, and 12. Although more genomic studies are necessary to understand the molecular basis of biofortification across rice chromosomes, a meta-analysis of rice QTLs associated with iron and zinc in grains have already identified 48 meta-QTLs (MQTL) randomly distributed across rice chromosomes, though they are predominant on chromosome 7 (27 QTLs) and scarcely mapped on chromosome 11 (8 QTLs; Raza et al., 2019). Several genes/transcripts (e.g., OsATM3, OsDMAS1, OsFRO2, OsNAS1-3, OsVIT2, OsYSL16, OsZIP3, and OsZIP7) are physically located within or near these MQTL regions and are, as found here, involved in binding, oxidation reduction process, metabolic process, regulation of transcription, and transport (Raza et al., 2019).

CONCLUSION

In conclusion, we have in this work settled the foundations to understand genomic changes in the flag leaves of rice plants biofortified with the single and combined use of Zn and Se, which are mainly based on the activation of a limited number of metabolic pathways related to micronutrient mobilization and to the specific functions of Zn (i.e., its enzymatic co-factor/coenzyme function in the biosynthesis of nitrogenous compounds, carboxylic acids, organic acids and amino acids) and Se (vitamin biosynthesis and ion homeostasis). The success of this approach should be followed in future studies to understand how landraces and other rice cultivars respond to biofortification. For that, we are currently analyzing the transcriptome of a biofortified rice landrace and integrating agronomic and

molecular analyzes to further study the transcriptional patterns of putatively key biofortification genes during grain development.

DATA AVAILABILITY STATEMENT

The datasets GENERATED for this study can be found in the NCBI Sequence Read Archive under Bioproject No. PRJNA629980.

AUTHOR CONTRIBUTIONS

FR, JR, and FL conceptualized the work. FR, JR, FL, AN, and AR-B performed the experimental design. FR, JR, and AN performed the experimental assays. FR, PB-S, and AR-B performed the laboratory work. FR, IM, ME, BS, and AR-B performed the bioinformatic analysis. FR, IM, JR, and AR-B wrote the manuscript draft. ME and BS performed the critical review of the manuscript. FR, IM, JR, and AR-B were responsible for the final version of the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by funds of Camões, Instituto da Cooperação e da Língua, and by Fundação para a Ciência e Tecnologia, Portugal, through the contribution to the International Rice Research Institute, Mozambique, and the research units UIDB/00239/2020 (CEF), UID/04129/2020 (LEAF), and UIDP/04035/2020 (GeoBioTec).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00543/full#supplementary-material

REFERENCES

Andrews, S. (2010). *Fast QC: A Quality Control Tool for High Throughput Sequence Data*. Available online at: http://www.bioinformatics.babraham.ac.uk/projects/fastqc/ (accessed October 31, 2018).

Araújo, W. L., Auxiliadora, O. M., Fernie, A. R., and Tohge, T. (2014). 2-Oxoglutarate: linking TCA cycle function with amino acid, glucosinolate, flavonoid, alkaloid, and gibberellin biosynthesis. *Front. Plant Sci.* 5:552. doi: 10.3389/fpls.2014.00552

Aslam, M., Fakher, B., Jakada, B. H., Zhao, L., Cao, S., Cheng, Y., et al. (2019). Genome-wide identification and expression profiling of CBL-CIPK gene family in pineapple (*Ananas comosus*) and the role of AcCBL1 in abiotic and biotic stress response. *Biomolecules* 9:293. doi: 10.3390/biom9070293

Bode, R., Ivanov, A. G., and Hu, N. P. A. (2016). Global transcriptome analyzes provide evidence that chloroplast redox state contributes to intracellular as well as long distance signalling in response to stress and acclimation in *Arabidopsis*. *Photos. Res.* 128, 287–312. doi: 10.1007/s11120-016-0245-y

Boldrin, P. F., Valdemar, F., Ramos, S. J., Guilherme, L. R. G., Bastos, C., Carvalho, G., et al. (2012). Selenate and selenite on yield and agronomic biofortification with selenium in rice. *Pesq. Agropec. Bras.* 47, 831–837. doi: 10.1590/S0100-204X2012000600014

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

Brown, P. H., Cakmak, I., and Zhang, Q. (1993). "Form and function of zinc plants," in *Zinc in Soils and Plants. Developments in Plant and Soil Sciences 55*, ed. A. D. Robson (Dordrecht: Springer), 93–106. doi: 10.1007/978-94-011-0878-2_7

Cakmak, I. (2008). Enrichment of cereal grains with zinc: agronomic or genetic biofortification? *Plant Soil* 302, 1–17. doi: 10.1007/s11104-007-9466-3

Cakmak, I., Pfeiffer, W. F., and McClafferty, B. (2010). Biofortification of durum wheat with zinc and iron. *Cereal Chem.* 87, 10–20. doi: 10.1094/CCHEM-87-1-0010

Cao, Z. Z., Lin, X. Y., Yang, Y. J., Guan, M. Y., Xu, P., and Chen, M. X. (2019). Gene identification and transcriptome analysis of low cadmium accumulation rice mutant (*lcd1*) in response to cadmium stress using MutMap and RNA-seq. *BMC Plant Biol.* 19:250. doi: 10.1186/s12870-019-1867-y

Chattha, M. U., Hassan, M. U., Khan, I., Chattha, M. B., Mahmood, A., Chattha, M. U., et al. (2017). Biofortification of wheat cultivars to combat zinc deficiency. *Front. Plant Sci.* 8:281. doi: 10.3389/fpls.2017.00281

Chen, X., Tao, Y., Ali, A., Zhuang, Z., Guo, D., Guo, Q., et al. (2019). Transcriptome and proteome profiling of dfferent colored rice reveals physiological dynamics involved in the flavonoid pathway. *Int. J. Mol. Sci.* 20:2463. doi: 10.3390/ijms20102463

dos Santos, R. S., Araujo Junior, A. T., Pegoraro, C., and de Oliveira, A. C. (2017). Dealing with iron metabolism in rice: from breeding for stress tolerance to biofortification. *Genet. Mol. Biol.* 40, 312–325. doi: 10.1590/1678-4685-GMB-2016-0036

Fairweather-Tait, S. J., Bao, Y., Broadley, M. R., Collings, R., Ford, D., and Hesketh, J. E. (2011). Selenium in human health and disease. *Antioxid. Redox Signal.* 14, 1337–1383.

FAO (2016). *The State of Food and Agriculture – Climate Change, Agriculture and Food Security*. Rome: Food and Agriculture Organization of the United Nations.

FAO, IFAD, UNICEF, WFP, and WHO (2019). *The State of Food SECURITY and Nutrition in The World 2019. Safeguarding Against Economic Slowdowns and Downturns*. Rome: FAO.

Farrow, S. C., and Facchini, P. J. (2014). Functional diversity of 2-oxoglutarate/Fe(II)-dependent dioxygenases in plant metabolism. *Front. Plant Sci.* 5:524. doi: 10.3389/fpls.2014.00524

Forsman, C., Milani, P., Schondebare, J. A., Matthias, D., and Guyondet, C. (2014). Rice fortification: a comparative analysis in mandated settings. *Ann. N. Y. Acad. Sci.* 1324, 67–81. doi: 10.1111/nyas.12453

Garg, M., Sharma, N., Sharma, S., Kapoor, P., Kumar, A., Chunduri, V., et al. (2018). Biofortified crops generated by breeding, agronomy, and transgenic approaches are improving lives of millions of people around the world. *Front. Nutr.* 5:12. doi: 10.3389/fnut.2018.00012

Goeres, D. C., Van Norman, J. M., Zhang, W., Fauver, N. A., Spencer, M. L., and Sieburth, L. E. (2007). Components of the *Arabidopsis* mRNA decapping complex are required for early seedling development. *Plant Cell* 19, 1549–1564. doi: 10.1105/tpc.106.047621

Gorelova, V., Ambach, L., Rébeillé, F., Stove, C., and Van der Straeten, D. (2017). Folates in plants: research advances and progress in crop biofortification. *Front. Chem.* 5:21. doi: 10.3389/fchem.2017.00021

Gupta, M., and Gupta, S. (2017). An overview of selenium uptake, metabolism, and toxicity in plants. *Front. Plant Sci.* 7:2074. doi: 10.3389/fpls.2016.02074

Hartikainen, H. (2005). Biogeochemistry of selenium and its impact of food chain quality and human health. *J. Trace Elem. Med. Biol.* 18, 309–318. doi: 10.1016/j.jtemb.2005.02.009

He, F., Liu, Q., Zheng, L., Cui, Y., Shen, Z., and Zheng, L. (2015). RNA-seq analysis of rice roots reveals the involvement of post-transcriptional regulation in response to cadmium stress. *Front. Plant Sci.* 6:1136. doi: 10.3389/fpls.2015.01136

Hefferon, K. (2019). Biotechnological approaches for generating zinc-enriched crops to combat malnutrition. *Nutrients* 11:253. doi: 10.3390/nu11020253

Hirschi, K. D. (2009). Nutrient biofortication of food crops. *Annu. Rev. Nutr.* 29, 401–421. doi: 10.1146/annurev-nutr-080508-141143

Ho, C. H., Lin, S. H., Hu, H. C., and Tsay, Y. F. (2009). CHL1 functions as a nitrate sensor in plants. *Cell* 138, 1184–1194. doi: 10.1016/j.cell.2009.07.004

Jablonska, E., and Vinceti, M. (2015). Selenium and human health: witnessing a Copernican revolution? *J. Environ. Sci. Health Part C* 33, 328–368. doi: 10.1080/10590501.2015.1055163

Jackson, S. A. (2016). Rice: the first crop genome. *Rice* 9:14. doi: 10.1186/s12284-016-0087-4

Jiang, N., Shi, S., Shi, H., Khanzada, H., Wassan, G. M., Zhu, C., et al. (2017). Mapping QTL for seed germinability under low temperature using a new high-density genetic map of rice. *Front. Plant Sci.* 8:1223. doi: 10.3389/fpls.2017.01223

Jiménez-López, D., Bravo, J., and Guzmán, P. (2015). Evolutionary history exposes radical diversification among classes of interaction partners of the MLLE domain of plant poly(A)-binding proteins. *BMC Evol. Biol.* 15:195. doi: 10.1186/s12862-015-0475-1

Jiménez-López, D., and Guzmán, P. (2014). Insights into the evolution and domain structure of ataxin-2 proteins across eukaryotes. *BMC Res. Notes* 7:453. doi: 10.1186/1756-0500-7-453

Kawa, D., and Testerink, C. (2017). Regulation of mRNA decay in plant responses to salt and osmotic stress. *Cell. Mol. Life Sci.* 74, 1165–1176. doi: 10.1007/s00018-016-2376-x

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* 6:4. doi: 10.1186/1939-8433-6-4

Kawai, Y., Ono, E., and Mizutani, M. (2014). Evolution and diversity of the 2-oxoglutarate-dependent dioxygenase superfamily in plants. *Plant J.* 78, 328–343. doi: 10.1111/tpj.12479

Kole, C., Muthamilarasan, M., Henry, R., Edwards, D., Sharma, R., Abberton, M., et al. (2015). Application of genomics-assisted breeding for generation of climate resilient crops: progress and prospects. *Front. Plant Sci.* 6:563. doi: 10.3389/fpls.2015.00563

Kondwakwenda, A., Sibiya, J., Zengeni, R., Musvosvi, C., and Meike, A. S. (2018). Provitamin A maize biofortification in sub-Saharan Africa. *Maydica* 63, 1–9.

Li, N., Liu, H., Sun, J., Zheng, H., Wang, J., Yang, L., et al. (2018). Transcriptome analysis of two contrasting rice cultivars during alkaline stress. *Sci. Rep.* 8:9586. doi: 10.1038/s41598-018-27940-x

Lidon, F. C., Almeida, A. S., Costa, A. R., Bagulho, A. S., Scotti-Campos, P., Semedo, J. N., et al. (2015). Sequential zinc and iron biofortification of bread-wheat grains: from controlled to uncontrolled environments. *Crop Pasture Sci.* 66, 1097–1104. doi: 10.1071/CP14270

Lidon, F. C., Oliveira, K., Galhano, C., Guerra, M., Ribeiro, M. M., Pelica, J., et al. (2019). Selenium biofortification of rice through foliar application with selenite and selenate. *Exp. Agric.* 55, 528–542. doi: 10.1017/S0014479718000157

Lidon, F. C., Ramalho, J. C., Leitão, A. E., Pessoa, M. F., Silva, M. M. A., and Reboredo, F. H. (2017). Essential key points for zinc biofortification: uptake, translocation and accumulation mechanisms in higher plants. *Agric. Res. Tech.* 4, 555–647. doi: 10.19080/ARTOAJ.2017.04.555647

Ligaba-Osena, A., Fei, Z., Liu, J., Xu, Y., Shaff, J., Lee, S. C., et al. (2017). Loss-of-function mutation of the calcium sensor CBL1 increases aluminum sensitivity in *Arabidopsis*. *New Phytol.* 214, 830–841. doi: 10.1111/nph.14420

Liu, H., Gan, W., Rengel, Z., and Zhao, P. (2016). Effects of zinc fertilizer rate and application method on photosynthetic characteristics and grain yield of summer maize. *J. Soil Sci. Plant Nutr.* 16, 550–562. doi: 10.4067/S0718-95162016005000045

Liu, H., Wang, Y. X., Li, H., Teng, R. M., Wang, Y., and Zhuang, J. (2019). Genome-wide identification and expression analysis of calcineurin B-like protein and calcineurin B-like protein-interacting protein kinase family genes in tea plant. *DNA Cell Biol.* 38, 824–839. doi: 10.1089/dna.2019.4697

Loizeau, K., Gambonnet, B., Zhang, G. F., Curien, G., Jabrin, S., Van Der Straeten, D., et al. (2017). Regulation of one-carbon metabolism in *Arabidopsis*: the N-terminal regulatory domain of cystathionine γ-synthase is cleaved in response to folate starvation. *Plant Physiol.* 145, 491–503. doi: 10.1104/pp.107.105379

Lucca, P., Poletti, S., and Sautter, C. (2006). Genetic engineering approaches to enrich rice with iron and vitamin A. *Physiol. Plant.* 126, 291–303. doi: 10.1111/j.1399-3054.2006.00609x

Mangueze, A. V. J., Pessoa, M. F. G., Silva, M. J., Ndayiragije, A., Magaia, H. E., Cossa, V. S. I., et al. (2018). Simultaneous zinc and selenium biofortification in rice. Accumulation, localization and implications on the overall mineral content of the flour. *J. Cereal Sci.* 82, 34–41. doi: 10.1016/j.jcs.2018.05.005

Minhas, A. P., Tuli, R., and Puri, S. (2018). Pathway editing targets for thiamine biofortification in rice grains. *Front. Plant Sci.* 9:975. doi: 10.3389/fpls.2018.00975

Mishra, V. K., Gupta, S., Ghand, R., Yadav, P. S., Singh, S. K., Joshi, A. K., et al. (2019). Unveiling the transcriptome complexity of the high- and low- zinc and iron accumulating Indian wheat (*Triticum aestivum* L.) cultivars. *bioRxiv* [Preprint]. doi: 10.1101/538819 bioRxiv 538819,

Moreno, R. G., Burdock, R., Álvarez, M. C. D., and Crawford, J. W. (2013). Managing the selenium content in soils in semiarid environments through the recycling of organic matter. *Appl. Environ. Soil Sci.* 2013:e283468. doi: 10.1155/2013/283468

Muthayya, S., Sugimoto, J. D., Montgomery, S., and Maberly, G. F. (2014). An overview of global rice production, supply, trade, and consumption. *Ann. N. Y. Acad. Sci.* 1324, 7–14. doi: 10.1111/nyas.12540

Nakandalage, N., Nicolas, M., Norton, R. M., Hirotsu, N., Milham, P. J., and Seneweera, S. (2016). Improving rice zinc biofortification success rates through genetic and crop management approaches in a changing environment. *Front. Plant Sci.* 7:764. doi: 10.3389/fpls.2016.00764

Narayanan, N. N., Vasconcelos, M. W., and Grusak, M. A. (2007). Expression profiling of *Oryza sativa* L. metal homeostasis genes in different rice cultivars using a cDNA macroarray. *Plant Physiol. Biochem.* 45, 277–286. doi: 10.1016/j.plaphy.2007.03.021

Neeraja, C. N., Kulkarni, K. S., Babu, P. M., Rao, D. S., Surekha, K., and Babu, V. R. (2018). Transporter genes identified in landraces associated with high zinc in polished rice through panicle transcriptome for biofortification. *PLoS One* 13:e0196160. doi: 10.1371/journal.pone.0192362

Nützman, H. W., Huang, A., and Osbourn, A. (2016). Plant metabolic clusters-from genetics to genomics. *New Phytol.* 211, 771–789. doi: 10.1111/nph.13981

Ostrowski, L. A., Hall, A. C., and Mekhail, K. (2017). Ataxin-2: from RNA control to human health and disease. *Genes* 8:157. doi: 10.3390/genes8060157

Pang, J., Kobayashi, K., and Zhu, J. (2009). Yield and photosynthetic characteristics of flag leaves in Chinese rice (*Oryza sativa* L.) varieties subjected to free-air release of ozone. *Agric. Ecosyst. Environ.* 132, 3–4. doi: 10.1016/j.agee.2009.03.012

Pertea, M., Kim, D., Pertea, G. M., Leek, J. T., and Salzberg, S. L. (2016). Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* 11, 1650–1667. doi: 10.1038/nprot.2016.095

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33, 290–295. doi: 10.1038/nbt.3122

Phattarakul, N., Rerkasem, B., Li, L. J., Wu, L. H., Zou, C. Q., Ram, H., et al. (2012). Biofortication of rice grain with zinc through zinc fertilization in different countries. *Plant Soil* 361, 131–141. doi: 10.1007/s11104-012-1211-x

Prentice, A. M., Gershwin, M. E., Schaible, U. E., Keusch, G. T., Victora, C. G., and Gordon, J. I. (2008). New challenges in studying nutrition-disease interactions in the developing world. *J. Clin. Invest.* 118, 1322–1329. doi: 10.1172/JCI34034

Qian, L., Liu, Z., Wu, J., Zhang, C., Huang, S., Liu, X., et al. (2018). Transcriptome analysis of the responses involved in the regulation of cadmium stress by exogenous hydrogen in rice (*Oryza sativa* L.). *Pak. J. Bot.* 50, 2123–2129.

Ram, H., Rashid, A., Zhang, W., Duarte, A. P., Phattarakul, N., Simunji, S., et al. (2016). Biofortification of wheat, rice and common bean by applying foliar zinc fertilizer along with pesticides in seven countries. *Plant Soil* 403, 389–401. doi: 10.1007/s11104-016-2815-3

Ramalho, J. C., Roda, F. A., Pessoa, M. F. G., Reboredo, F. H., Pais, I. P., Ndayiragije, A., et al. (2020). "Selenium agronomic biofortification in rice: improving crop quality against malnutrition," in *The Future of Rice Demand: Quality Beyond Productivity, Chapter 8*, eds A. Costa de Oliveira, C. Pegoraro, and V. Ebeling Viana (Cham: Springer Nature), 179–203. doi: 10.1007/978-3-030-37510-2_8

Rayman, M. P. (2002). The argument for increasing selenium intake. *Proc. Nutr. Soc.* 61, 203–215. doi: 10.1079/PNS2002153

Raza, Q., Riaz, A., Sabar, M., Atif, R. M., and Bashir, K. (2019). Meta-analysis of grain iron and zinc associated QTLs identified hotspot chromosomal

regions and positional candidate genes for breeding biofortified rice. *Plant Sci.* 288:110214. doi: 10.1016/j.plantsci.2019.110214

Sasaki, T. (2005). The map-based sequence of the rice genome. *Nature* 436, 793–800. doi: 10.1038/nature03895

Sauer, U., Heinemann, M., and Zamboni, N. (2007). Genetics: getting closer to the whole picture. *Science* 316, 550–551. doi: 10.1126/science.1142502

Scheible, W. R., Krapp, A., and Stitt, M. (2000). Reciprocal diurnal changes of phosphoenolpyruvate carboxylase expression and cytosolic pyruvate kinase, citrate synthase and NADP-isocitrate dehydrogenase expression regulate organic acid metabolism during nitrate assimilation in tobacco leaves. *Plant Cell Environ.* 23, 1155–1167. doi: 10.1046/j.1365-3040.2000.00634.x

Schomburg, D., and Schomburg, I. (2010). *Springer Handbook of Enzymes*. Berlin: Springer.

Shi, H., Quintero, F. J., Pardo, J. M., and Zhu, J. K. (2002). The putative plasma membrane Na^+/H^+ antiporter SOS_1 controls long-distance Na(+) transport in plants. *Plant Cell* 14, 465–477. doi: 10.1105/tpc.010371

Singh, R. K., Murori, R., Ndayiragije, A., Bigirimana, J., Kimani, J. M., Kanyeka, Z. L., et al. (2013). Rice breeding activities in Eastern and Southern Africa. *SABRAO J. Breed. Genet.* 45, 73–83.

Sperotto, R. A., Ricachenevsky, F. K., de, A., Waldow, V., Müller, A. L. H., Dressler, V. L., et al. (2013). Rice grain Fe, Mn and Zn accumulation: how important are flag leaves and seed number? *Plant Soil Environ.* 59, 262–266. doi: 10.17221/841/2012-PSE

Sperotto, R. A., Ricachenevsky, F. K., Duarte, G. L., Boff, T., Lopes, K. L., Sperb, E. R., et al. (2009). Identification of upregulated genes in flag leaves during rice grain filling and characterization of *OsNAC5*, a new ABA-dependent transcription factor. *Planta* 230, 985–1002. doi: 10.1007/s00425-009-1000-9

Sun, L., Wang, J., Song, K., Sun, Y., Qin, Q., and Xue, Y. (2019). Transcriptome analysis of rice (*Oryza sativa* L.) shoots responsive to cadmium stress. *Sci. Rep.* 9:10177. doi: 10.1038/s41598-019-46684-w

Sunde, R. A. (2018). Selenium regulation of selenoprotein enzyme activity and transcripts in a pilot study with Founder strains from the Collaborative Cross. *PLoS One* 13:e0191449. doi: 10.1371/journal.pone.0191449

Sundell, D., Mannapperuma, C., Netotea, S., Delhomme, N., Lin, Y. C., Sjödin, A., et al. (2015). The plant genome integrative explorer resource: PlantGenIE.org. *New Phytol.* 208, 1149–1156. doi: 10.1111/nph.13557

Swamy, B. P. M., Rahman, M. A., Inabangan-Asilo, M. A., Amparado, A., Manito, C., Chadha-Mohanty, P., et al. (2016). Advances in breeding for high grain Zinc in Rice. *Rice* 9, 49. doi: 10.1186/s12284-016-0122-5

Tang, R. J., Zhao, F. G., Garcia, V. J., Kleist, T. J., Yang, L., Zhang, H. X., et al. (2015). Tonoplast CBL-CIPK calcium signaling network regulates magnesium homeostasis in *Arabidopsis*. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3134–3139. doi: 10.1073/pnas.1420944112

Tang, X., Gong, R., Sun, W., Zhang, C., and Yu, S. (2018). Genetic dissection and validation of candidate genes for flag leaf size in rice (*Oryza sativa* L.). *Theor. Appl. Genet.* 131, 801–815. doi: 10.1007/s00122-017-3036-8

Tari, B. D., Gazanchian, A., Pirdashti, A. H., and Nasiri, M. (2009). Flag leaf morphophysiological response to different agronomical treatments in a promising line of rice (*Oryza sativa* L.). *Am. Eur. J. Agric. Environ. Sci.* 5, 403–408.

Trapnell, C., Williams, B. A., Pertea, G., Mortazavi, A., Kwan, G., van Baren, M. J., et al. (2010). Transcript assembly and quantification by RNA-Seq reveals unannotated transcripts and isoform switching during cell differentiation. *Nat. Biotechnol.* 28, 511–515. doi: 10.1038/nbt.1621

Walter, M., Marchezan, E., and Avila, L. A. (2008). Arroz: composição e características nutricionais. *Ciencia Rural* 38, 1184–1192. doi: 10.1590/S0103-84782008000400049

Wang, J., Cappa, J. J., Harris, J. P., Edger, P. P., Wen Zhou, W., Pires, J. C., et al. (2018). Transcriptome-wide comparison of selenium hyperaccumulator and nonaccumulator *Stanleya* species provides new insight into key processes mediating the hyperaccumulation syndrome. *Plant Biotechnol. J.* 16, 1582–1594. doi: 10.1111/pbi.12897

Wang, Z., Wang, S., Wu, M., Li, Z., Liu, P., Li, F., et al. (2019). Evolutionary and functional analyzes of the 2-oxoglutarate-dependent dioxygenase genes involved in the flavonoid biosynthesis pathway in tobacco. *Planta* 249, 543–561. doi: 10.1007/s00425-018-3019-2

Wei, K., and Chen, H. (2018). Global identification, structural analysis and expression characterization of cytochrome P450 monooxygenase superfamily in rice. *BMC Genomics* 19:35. doi: 10.1186/s12864-017-4425-8

White, P. J., and Broadley, M. R. (2009). Biofortification of crops with seven mineral elements often lacking in human diets Iron, zinc, copper, calcium, magnesium, selenium and iodine. *New Phytol.* 182, 49–84. doi: 10.1111/j.1469-8137.2008.02738.x

White, P. J., and Broadley, M. R. (2011). Physiological limits to zinc biofortification of edible crops. *Front. Plant Sci.* 2:80. doi: 10.3389/fpls.2011.00080

Xu, J., Li, H. D., Chen, L. Q., Wang, Y., Liu, L. L., He, L., et al. (2006). A protein kinase, interacting with two calcineurin B-like proteins, regulates K^+ transporter AKT1 in *Arabidopsis*. *Cell* 125, 1347–1360. doi: 10.1016/j.cell.2006.06.011

Xu, W., Yang, R., Li, M., Xing, Z., Yang, W., Chen, G., et al. (2011). Transcriptome phase distribution analysis reveals diurnal regulated biological processes and key pathways in rice flag leaves and seedling leaves. *PLoS One* 6:e17613. doi: 10.1371/journal.pone.0017613

Yahalom, A., Kim, T. H., Roy, B., Singer, R., von Arnim, A. G., and Chamovitz, D. A. (2008). *Arabidopsis* eIF3e is regulated by the COP9 signalosome and has an impact on development and protein translation. *Plant J.* 53, 300–311. doi: 10.1111/j.1365-313X.2007.03347.x

Yamazaki, Y., Sakaniwa, S., Tsuchiya, R., Nonomura, K. I., and Kurata, N. (2010). Oryzabase: an integrated information resource for rice science. *Breed. Sci.* 60, 544–548. doi: 10.1270/jsbbs.60.544

Yi, F., Gu, W., Chen, J., Song, N., Gao, X., Zhang, X., et al. (2019). High temporal-resolution transcriptome landscape of early maize seed development. *Plant Cell* 31, 974–992. doi: 10.1105/tpc.18.00961

Yin, Q., Shen, G., Di, S., Fan, C., Chang, Z., and Pang, Y. (2017). Genome-wide identi?cation and functional characterization of UDP-glucosyltransferase genes involved in flavonoid biosynthesis in *Glycine max*. *Plant Cell Physiol.* 58, 1558–1572. doi: 10.1093/pcp/pcx081

Yu, B., Niu, J., Feng, J., Xu, M., Xie, X., Gu, W., et al. (2018). Regulation of ferredoxin-NADP+ oxidoreductase to cyclic electron transport in high salinity stressed *Pyropia yezoensis*. *Front. Plant Sci.* 9:1092. doi: 10.3389/fpls.2018.01092

Zallot, R., Brochier-Armanet, C., Gaston, K. W., Forouhar, F., Limbach, P. A., Hunt, J. F., et al. (2014). Plant, animal, and fungal micronutrient queuosine is salvaged by members of the DUF2419 protein family ACS. *Chem. Biol.* 9, 1812–1825. doi: 10.1021/cb500278k

Zhang, Y., Yang, L., Jahan, N., Chen, G., Ren, D., and Guo, L. (2018). Sensing of abiotic stress and ionic stress responses in plants. *Int. J. Mol. Sci.* 19:3298. doi: 10.3390/ijms19113298

Zhao, M., Wang, W., Wei, L., Chen, P., Yuan, F., Wang, Z., et al. (2018). Molecular evolution and expression divergence of three key Met biosynthetic genes in plants: CGS, HMT and MMT. *PeerJ* 6:e6023. doi: 10.7717/peerj.6023

Zhou, J., Wang, X., Jiao, Y., Qin, Y., Liu, X., He, K., et al. (2007). Global genome expression analysis of rice in response to drought and high-salinity stresses in shoot, flag leaf, and panicle. *Plant Mol. Biol.* 63, 591–608. doi: 10.1007/s11103-006-9111-1

Zhou, S., Williams, A. P., Berg, A. M., Cook, B. I., Zhang, Y., Hagemann, S., et al. (2019). Land-atmosphere feedbacks exacerbate concurrent soil drought and atmospheric aridity. *Proc. Natl. Acad. Sci. U.S.A.* 116, 18848–18853. doi: 10.1073/pnas.1904955116

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 Genetics

ORIGINAL RESEARCH
published: 11 August 2020
doi: 10.3389/fgene.2020.00763

Stable SNP Allele Associations With High Grain Zinc Content in Polished Rice (*Oryza sativa* L.) Identified Based on ddRAD Sequencing

P. Madhu Babu[1], C. N. Neeraja[1], Santosha Rathod[1], K. Suman[1], G. Anurag Uttam[1], Navajeet Chakravartty[2], V. B. Reddy Lachagari[2], U. Chaitanya[1], Lella V. Subba Rao[1] and Sitapati Rao Voleti[1]*

[1] ICAR-Indian Institute of Rice Research, Hyderabad, India, [2] AgriGenome Labs Pvt. Ltd., Hyderabad, India

OPEN ACCESS

Edited by:
Mallikarjuna Swamy, International Rice Research Institute (IRRI), Philippines

Reviewed by:
Paramesh Hanumanthaiah, Debre Berhan University, Ethiopia
Noraziyah Abd Aziz Shamsudin, National University of Malaysia, Malaysia

***Correspondence:**
C. N. Neeraja
cnneeraja@gmail.com

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: *24 April 2020*
Accepted: *29 June 2020*
Published: *11 August 2020*

Citation:
Babu PM, Neeraja CN, Rathod S, Suman K, Uttam GA, Chakravartty N, Lachagari VBR, Chaitanya U, Rao LVS and Voleti SR (2020) Stable SNP Allele Associations With High Grain Zinc Content in Polished Rice (Oryza sativa L.) Identified Based on ddRAD Sequencing. Front. Genet. 11:763. doi: 10.3389/fgene.2020.00763

Polished rice is widely consumed staple food across the globe, however, it contains limited nutrients especially iron (Fe) and zinc (Zn). To identify promising genotypes for grain Zn, a total of 40 genotypes consisting 20 rice landraces, and 20 released high yielding rice varieties were evaluated in three environments (wet seasons 2014, 2015 and 2016) for nine traits including days to 50% flowering (DFF), plant height (PH), panicle length (PL), total number of tillers (TNT), single plant yield (SPY), Fe and Zn in brown (IBR, ZBR) and polished rice (IPR, ZPR). Additive Main Effect and Multiplicative Interaction (AMMI), Genotype and Genotype × Environment Interaction (GGE) analyses identified genotypes G22 (Edavankudi Pokkali), G17 (Taraori Basmati), G27 (Chittimuthyalu) and G26 (Kalanamak) stable for ZPR and G8 (Savitri) stable for SPY across three environments. Significant negative correlation between yield and grain Zn was reaffirmed. Regression analysis indicated the contribution of traits toward ZPR and SPY and also desirable level of grain Zn in brown rice. A total of 39,137 polymorphic single nucleotide polymorphisms (SNPs) were obtained through double digest restriction site associated DNA (dd-RAD) sequencing of 40 genotypes. Association analyses with nine phenotypic traits revealed 188 stable SNPs with six traits across three environments. ZPR was associated with SNPs located in three putative candidate genes (LOC_Os03g47980, LOC_Os07g47950 and LOC_Os07g48050) on chromosomes 3 and 7. The genomic region of chromosome 7 co localized with reported genomic regions (rMQTL $_{7.1}$) and OsNAS3 candidate gene. SPY was found to be associated with 12 stable SNPs located in 11 putative candidate genes on chromosome 1, 6, and 12. Characterization of rice landraces and varieties in terms of stability for their grain Zn and yield identified promising donors and recipients along with genomic regions in the present study to be deployed rice Zn biofortification breeding program.

Keywords: rice, grain zinc, stable donors, ddRAD sequencing, SNP, candidate genes

INTRODUCTION

Biofortification is one of the promising strategies to address malnutrition through enhancing the nutritional value of crops and more than 200 biofortified crop varieties have been released worldwide (HarvestPlus and FAO, 2019). Rice, being major staple food for half of the world population has been targeted for biofortification for various nutrients since 2000 using genetic engineering and conventional breeding approaches (Mahender et al., 2016; Neeraja et al., 2017; Majumder et al., 2019; Ludwig and Slamet-Loedin, 2019). Zinc (Zn) is one of the most essential nutrients for human health and associated with various metabolic activities (Roohani et al., 2013). Enhancing of Zn content of the rice grains could have a positive impact on human health (Graham et al., 2001; Stein, 2010). Polished rice, the most preferred form for consumption is poor source of Zn with a range of 8 to 12 ppm. Popular rice varieties usually contain lesser micronutrients in grains compared to the traditional cultivars and landraces (Nachimuthu et al., 2015; Pradhan et al., 2020; Rao et al., 2020). Through conventional breeding, enhancement of grain Zn without yield compromise in polished rice has been demonstrated and many rice varieties with increased Zn content have been released in a few Asian countries *viz*., 10 in India, three in Bangladesh, one each in Indonesia and Philippines[1,2] (Palanog et al., 2019). A range of grain Zn from 20 to 28 ppm in polished rice and similar or higher yield levels as popular adopted varieties along with standard quality traits are mandatory for the released of Zn biofortified rice varieties as defined by the government agencies in different countries. Thus, for the development of biofortified rice varieties, high grain Zn content, yield and quality should be combined. For developing such varieties, availability of suitable germplasm is the foremost requirement. Wide genetic variability was reported for grain Zn ranging from in brown (7.3 to 52.7 ppm) and polished (8 to 38 ppm) rice and several donors have been identified (Swamy et al., 2016; Rao et al., 2020). Traditional varieties or landraces are known to be the source of novel genes/alleles for the agronomically important traits including grain Zn content (Neeraja et al., 2018; Bollinedi et al., 2020). Increasing grain Zn content in rice was reported to be feasible by utilizing germplasm in the breeding programs to reach the international standards of grain Zn 28 ppm in polished rice (Rao et al., 2020[1]).

The success of genetic improvement programs depends on the selection of productive and stable genotypes which depends on understanding of the interaction between genotypes and environments (G × E). Breeding for high grain Zn is slow because of low heritability and genetic interactions such as epistasis, environmental-genotype interactions, and polygenic effects (Zhang et al., 2014; Pradhan et al., 2020). To study G × E interactions, multivariate techniques such as biplots; Additive Main effects and Multiplicative Interaction (AMMI) and Genotype main effects and G x E interaction effects (GGE) are being widely adopted (Gauch and Zobel, 1997; Li et al., 2017). Considering the influence of environment on performance of genotypes for high grain Zn, effects of genotype (G), environment (E) and G x E were studied through analysis of variance (ANOVA), AMMI and GGE model in rice (Malosetti et al., 2013; Inabangan-Asilo et al., 2019).

Identification of candidate genes/genomic regions controlling high grain Zn would be an important approach for the marker assisted breeding of biofortified rice varieties. Using bi-parental mapping populations, 22 independent studies have reported 220 QTL for grain Fe and Zn in rice using simple sequence repeat (SSR) markers or candidate gene based markers (Raza et al., 2019). Either a major QTL > 30% phenotypic variance (PV) or pyramiding of a few minor QTL (∼20% PV) can be deployed in the marker assisted breeding of rice Zn biofortified rice varieties. Association mapping of grain Fe and Zn was also studied using SSR and candidate gene based markers in rice (Pradhan et al., 2020). Using next generation sequencing (NGS) approaches, genome-wide single-nucleotide polymorphic (SNP) molecular markers can be generated for the efficient discovery of the genomic regions associated with complex traits. A modified NGS approach based for scoring of SNPs based reduced representation of genome is known as genotyping-by-sequencing (GBS) (Deschamps et al., 2012) and is being applied in mapping of QTL and identification of genes in rice (De Leon et al., 2016; Furuta et al., 2017). Restriction-site associated DNA sequencing (RAD-seq) involves restriction enzyme digestion and NGS of regions adjacent to restriction sites, results in high throughput genetic markers across the genome (Baird et al., 2008; Davey et al., 2011; Peterson et al., 2012). A modified RAD-seq based on two restriction enzymes comprising a rare-cutting and frequently cutting as double-digest RAD-seq (ddRAD-seq) increases the selection of stable and repeatable size regions across samples (Peterson et al., 2012) and has been deployed in several crop plant species (Bodanapu et al., 2019; Gali et al., 2019; Sudan et al., 2019; Tafesse et al., 2020). Recently, SNPs associated with Fe, Zn and selenium (Se) concentration in field pea (*Pisum sativum* L.) were identified using dd RAD-seq in field pea (Dissanayaka et al., under early view). Using SNP markers, genomic regions for grain Zn along with other minerals have also been identified in rice using genome wide association studies (GWAS) (Norton et al., 2010; Zhang et al., 2014; Pradhan et al., 2020; Bollinedi et al., 2020) and also in Multi-parent Advanced Generation Inter-Cross (MAGIC) population (Descalsota et al., 2018).

The objectives of the present study were to identify stable donors for high grain Zn from shortlisted promising landraces in comparison to popular high yielding varieties based on evaluation for three years; correlate high grain Zn with yield and other traits; and identify polymorphic genome wide SNPs and their associations with high grain Zn, yield and other traits in the studied genotypes.

MATERIALS AND METHODS

Plant Material

In our earlier studies during 2013, > 5000 rice genotypes consisting landraces from various parts of India, breeding lines

[1] www.harvestplus.org

TABLE 1 | Details of pedigree for released rice varieties and origin for rice landraces in the study.

S No	Code	Variety Name	Designation	Cross Combination/origin*
1	G 01	Samba Mahsuri	BPT 5204	GEB 24/TN1//Mahsuri
2	G 02	Vijetha	MTU 1001	MTU 5249/MTU 7014
3	G 03	Cottondora Sannalu	MTU 1010	Krishnaveni/IR 64
4	G 04	Swarna	MTU 7029	Vashistha/Mahsuri
5	G 05	IR 64	IR 18348-36-3-3	IR 5657-33-2-1/IR 2061-465-1-5-5
6	G 06	PR 116	PAU 2020-10-3-1	PR 108/PAU 1628//PR 108
7	G 07	Narendradhan 359	NDR 359	BG 90-2-4/OB 677
8	G 08	Savitri	CR 210-1009	Pankaj/Jagannath
9	G 09	Pusa Basmati 1	Pusa 615-140-10-1	Pusa 167/Karnal local
10	G 10	Jaya	12306	TN1/T141
11	G 11	Mahsuri	-	Taichung 65/Mayang Ebos 6080/2
12	G 12	Lalat	ORS26-2014-4	OBS677/IR2071//Vikram/W1263
13	G 13	Sampada	PR4-60-105-6-8-2-5-1B	Vijaya/C14-8
14	G 14	Pushyami	MTU 1075	MTU 2716/MTU 1010
15	G 15	Jalpriya	NDGR 150	IET 4060/Jalmagna
16	G 16	ARB-45	Azucena/Moromutant	Breeding line
17	G 17	Taraori Basmati	HBC 19	Pure line selection from HBC 19
18	G 18	Akut Phou	KD 14-1-9	Lang Phou/IR 1364-37-3-1
19	G 19	Seetasail	Landrace	North India
20	G 20	Tilakkachari	Landrace	North India
21	G 21	High Iron Rice	Landrace	South India
22	G 22	Edavankudi Pokkali	Landrace	South India
23	G 23	Kadamakudy Pokkali	Landrace	South India
24	G 24	DRRDhan 45	IR80463−B−39−3	Areumbyeo/IRRI123
25	G 25	HP-5	IR80463−B−39−3-1	AreumbyeO/IRRI123
26	G 26	Kalanamak	Landrace	North India
27	G 27	Chittimuthyalu	Landrace	South India
28	G 28	Kala Jira Jaha	Landrace	North East Collection
29	G 29	Raga-binni	Landrace	North East Collection
30	G 30	Gopalbhok	Landrace	North East Collection
31	G 31	Hatibandha	Landrace	North East Collection
32	G 32	Mima	Landrace	North East Collection
33	G 33	Jahagipok	Landrace	North East Collection
34	G 34	Jahagisim	Landrace	North East Collection
35	G 35	Kewelhi lolu-R I	Landrace	North East Collection
36	G 36	Nedu	Landrace	North East Collection
37	G 37	Sirarakhong Manui	Landrace	North East Collection
38	G 38	Wungrei	Landrace	North East Collection
39	G 39	Arunachal pradesh-1	Landrace	North East Collection
40	G 40	Sakha	Landrace	North East Collection

and high yielding varieties were evaluated for their grain Zn content (Babu et al., 2014). A few promising landraces and high yielding were shortlisted based on their wide variation of grain yield and zinc content (**Table 1**). The set of 40 genotypes with contrasting grain Zn and yield were evaluated for their agro-morphological and yield traits along with grain Fe and Zn content in brown and polished rice over a period of three wet seasons of 2014, 2015, and 2016.

Field Experimental Details

Field experiments were conducted at research farm, Indian Council of Agricultural Research (ICAR)-Indian Institute of Rice Research (IIRR) (17°19′N and 78°29′E), Hyderabad, Telangana State, India during three consecutive wet seasons (*Kharif*) considered as three environments *viz*., 2014 (E1), 2015 (E2) and 2016 (E3) (**Supplementary Table S1**). The range of experimental soil characteristics across three years were: pH 8.2 – 8.4; non-saline (EC 0.7l – 0.72 dS/m); calcareous (free $CaCO_3$ 5.01 – 5.04%); CEC 44.1 – 45.2 C mol (p +)/kg soil and medium soil organic carbon (0.69 – 0.72%); low soil available nitrogen (228 – 230 kg ha^{-1}); high available phosphorus (105 – 108 kg P_2O_5 ha^{-1}), high available potassium (530 – 540 kg K_2O ha^{-1}), and high available Zn (12.5 –14.0 ppm). One month old seedlings of 40 genotypes were transplanted in a

randomized complete block design with three replications. In each replication, every genotype was grown in one m^2 plot (33.3 plants) with 20 cm row spacing and 15 cm intra-row spacing. Recommended package of rice crop production and protection practices were followed[3].

Trait Measurements

The genotypes were evaluated for days to 50% flowering (DFF), plant height (PH) (cm), number of tillers per plant (TNT), panicle length (PL) (cm), single plant yield (SPY) (g), grain Fe and Zn content in brown and polished rice (IBR, ZBR, IPR, ZPR) (ppm). The observations were recorded for three representative uniform plants from the center of the plot of each genotype. The seeds of each genotype and replication were dehusked using JLGJ4.5 testing rice husker (Jingjian Huayuan International Trade Co., Ltd.) and polisher (Krishi International India Ltd.) with non-ferrous and non-zinc components. Each sample of brown and polished rice (5 g) was subjected to energy dispersive X-ray fluorescent spectrophotometer (ED-XRF) (OXFORD Instruments X-Supreme 8000) at ICAR-IIRR as per standardized protocols (Rao et al., 2014).

Statistical Analysis

Descriptive statistics such as mean, standard error of mean (SEm), skewness, kurtosis and coefficient variations (%) were calculated to understand the characteristics, dispersion and heterogeneity of traits of the study. Graphical representation of summary statistics was depicted in frequency distribution plots and boxplots using R software (R Core Team, 2018). ANOVA was used to compare the variation in agro-morphological and yield traits along with grain Zn and Fe content among 40 genotypes across three years and G × E interactions of 40 genotypes. The performance of genotypes was assessed using stability models *viz*, (1) Additive Main effects and Multiplicative Interaction (AMMI) (Gauch and Zobel, 1997), and (2) GGE Biplot or Site Regression model (Yan and Kang, 2002).

The AMMI model is a combination of ANOVA and principal component analysis (PCA), where the additive (main) effects were estimated using ANOVA and G × E interaction effects (multiplicative effects) using principal components. The AMMI model is expressed as follows;

$$Y_{ij} = \mu + \delta_i + \beta_j + \sum_{k=1}^{K} \lambda_k \delta_{ik} \beta_{jk} + e_{ij} \quad (1)$$

where $\boldsymbol{Y_{ij}}$ is mean of i^{th} genotype in j^{th} environment, μ is the overall mean, $\boldsymbol{\delta_i}$ is the i^{th} genotypic effect, $\boldsymbol{\beta_j}$ is the j^{th} environmental effect, $\boldsymbol{\lambda_k}$ is the Eigen value for PC axis k, $\boldsymbol{\delta_{ik}}$ is the principal component score ithgenotype for k^{th} PC axis, $\boldsymbol{\beta_{jk}}$ is the principal component score j^{th} environment for k^{th} PC axis, $\mathbf{e_{ij}}$ is the error term. AMMI analysis and graphical representation through AMMI biplots was done using 'agricolae' r package (Mendiburu and de Mendiburu, 2019).

The equation for GGE biplots which uses site regression linear bilinear model is depicted as follows;

$$Y_{ij} - \mu_j = \sum_{k=1}^{t} \lambda_k \delta_{ik} \beta_{jk} + e_{ij} \quad (2)$$

GGE biplots were used to determine the best environment (which-won-where pattern) for recommending specific genotypes to specific environments or seasons. GGE biplots helps to determine stable genotype(s) across the locations or seasons and to understand discriminative power among genotypes in target locations or seasons. GGE biplots were plotted using 'GGEBiplot' r package (Dumble, 2017).

Correlation and stepwise regression analysis was carried out in SAS Version 9.3 software available at ICAR-IIRR. The regression model in terms of matrix notation is expressed as follows;

$$Y = X\beta + e \quad (3)$$

Where Y is the response variable, X is the vector of exogenous variables, β is the regression coefficient vector and e is the residuals term assumed to be normally distributed with e∼N(0,σˆ2).

Genotyping by ddRAD-seq

The ddRAD-seq protocol was followed in the present study (Peterson et al., 2012). Genomic DNA was extracted using the DNeasy Plant Mini Kit (Qiagen, Hilden, Germany) from pooling five 14 days old seedlings. DNA was checked for its concentration and quality on agarose gel and Qubit fluorometer (Thermofisher Scientific, United States). Genomic DNA of 40 genotypes was double digested using restriction enzymes *Sph*I and *Mluc*I (NEB, England) and the digested products were fractionated in 2% agarose gel electrophoresis to check product size in 250–400 bp. The Agencourt AMpure XP beads clean-up technology (with Dynabeads, Invitrogen) was used to clean the digested products using standard protocols (Beckman Coulter, United States). Barcoded adapters were ligated to each DNA sample using T4 ligase (NEB, England) followed by indexing with the addition of Index-1 and Index-2 (8 nt long) for multiplexing sequencing library in NGS Illumina. To increase the concentration of sequencing libraries, PCR amplification (8–12 cycles) was performed using PhusionTM PolymeraseKit (Fisher Scientific, United Kingdom). Products of PCR amplification were analyzed in an Agilent Bio-analyzer (Agilent, United States) to quantify molarity and fragment size distribution. The size-selected library was sequenced on an Illumina HiSeq2500 (Illumina, United States). RAD tags were identified in the raw reads, processed for base trimming and removal of the Illumina adapter sequences. The high quality paired end reads were aligned to *Oryza sativa* L. cv. Nipponbare reference genome (MSU7) using.bowtie2 version 2.2.2.6 (Langmead and Salzberg, 2012). The variant calling was performed based on the aligned reads to the reference genome and identified the SNPs using SAMtools version 0.1.19 (Li et al., 2009). The SNP markers were screened based on 90% call rate, locus homozygosity, and minor

[2]www.icar.iirr.org

[3]www.rkmp.co.in

allele frequency (MAF) 0.05. The variant annotation was performed based on rice gene models, using in-house pipelines VARIMAT (AgriGenome Labs Pvt., Ltd., India). The Population structure was determined by setting the number of groups (K) from 1 to 10.

Genome Wide Association Studies

In the present study, Kinship analysis was studied according to Endelman and Jannink (2012) using Centered-IBS matrix value in TASSEL5 (Bradbury et al., 2007). Dendrogram was constructed based on similarity matrix of the SNPs variations using neighbor-joining module in TASSEL 5 program. Out of the total variants, the monomorphic SNPs were filtered out and only polymorphic SNPs in the annotated regions were considered for downstream analyses. Associations between SNPs and phenotypic data were computed using the Genome Association and Prediction Integrated Tool—GAPIT (Lipka et al., 2012) based on the Mixed Linear Model (MLM) that controls the population structure and genetic relatedness among the individuals by incorporating the Q and K matrices. Principal component analysis (PCA) was performed using GAPIT software and genome wide association between traits were calculated (Lipka et al., 2012).

In MLM, X represents the genotype and Y the phenotype (The phenotype comprised three data sets of nine traits recorded during 2014, 2015, and 2016), allowing associated values of each SNP to be calculated (**Supplementary Table S2**). A value of < 0.05 was used as the threshold to determine the existence of a significant association. SNPs associated with nine traits in common identified across three years evaluation were only further analyzed. The amount of phenotypic variation explained by each marker was estimated by r^2. Associations were considered significant when $p \leq 0.01$ or LOD scores greater than 4.0. The functional annotation of the associated SNPs was identified in *Oryza sativa* L. cv. Nipponbare as reference genome[4].

Co-localization

The positions of the associated SNPs of ZPR and SPY of the present study were compared to the genomic positions of the markers from the reported QTL to study the co-localization.

Candidate Genes in the Region of SNPs Associated With ZPR and SPY

Based on earlier studies, the linkage disequilibrium was reported to be between 100 and 300 kbp on average for different subpopulations of rice (Zhao et al., 2011), thus a total of 600 kbp region (300 kbp each side of associated SNP) spanning each of the associated SNPs with ZPR and SPY was surveyed for the putative candidate genes[5].

[4]http://plantbiology.msu.edu/pub/data/Eukaryotic_Projects/o_sativa/annotation_dbs/pseudomolecules/version_7.0/all.dir

[5]http://rice.plantbiology.msu.edu

RESULTS

Agro- Morphological, Yield and Mineral Traits

All the evaluated 40 genotypes have shown a wide range variation for the agro-morphological yield and mineral related traits within an environment and between the environments E1, E2, and E3. DFF and SPY were higher in E3 compared to E2 and E1. The trait values for PH, PL, TNT, IBR, ZBR, IPR, and ZPR were higher in E1 compared to E2 and E3. Wide range was observed for the traits mean of three environments as 84.72 to 114.17 days for DFF, 62.87 to 142.57 cm for PH, 19.24 to 28.10 cm for PL, 4.53 to 13.91 for TNT, 8.92 to 32.45 g for SPY, 4.71 to 13.78 ppm for IBR, 10.03 to 29.66 for ZBR, 1.06 to 4.74 for IPR and 5.20 to 22.65 for ZPR (**Figure 1**). The trait wise descriptive statistic values of three environments were presented in **Supplementary Table S2**. Morphological variation of seed of 40 genotypes and the frequency distribution of nine traits were presented in **Supplementary Figures S1A–C**. Across the three environments G26 (Kalanamak) showed maximum DFF (114 days); G37 (Sirarakhong Manui) showed highest PH (142.5cm); G34 (Jahagisim) showed longest PL (29.1 cm); G13 (Sampada) has shown maximum TNT (13); G8 (Savitri) has shown highest SPY (32.4 g). G28 (Kala Jira Jaha) has shown highest IBR (13.7 ppm); G17 (Taroari Basmati) has shown highest ZBR (29.6 ppm). G17 (Taroari Basmati) has shown highest IPR (4.7 ppm); G22 (Edavankudi Pokkali) has shown highest ZPR (22.6 ppm). The mean loss found to be 7 ppm for Fe and 5 ppm of Zn content during polishing of rice. The maximum loss was found in G32 for Fe (11.1 ppm) (Mima) and in G17 (Taroari Basmati) for Zn (10.3 ppm). Minimum loss was noted in G14 (Pushyami) for Fe (3.6 ppm) and in G9 (Pusa Basmati 1) for Zn (2.5 ppm). There were four genotypes G22 (Edavankudi Pokkali), G18 (Akut Phou), G24 (DRRDhan 45) and G21 (High Fe rice) with maximum ZPR ($>$ 20.8 ppm). G8 (Savitri), G7 (NDR 359), G38 (Wungrei) and G13 (Sampada) were the top four genotypes for single plant yield $>$ 26.4 g. In E1, for SPY, G8 (Savitri) showed the maximum and G26 (Kalanamak) had the minimum, whereas the highest ZPR was observed in G23 (Kadamkudi Pokkali) and the least in G14 (Pushyami). In E2, G7 (NDR 359) had maximum and G26 (Kalanamak) had minimum SPY, while G22 (Edavankudi Pokkali) has shown the highest and G14 (Pushyami) has shown the least ZPR. In E2, G8 (Savitri) had maximum and G26 (Kalanamak) minimum SPY, while G32 (Mima) has shown maximum and G14 has shown minimum ZPR (Pushyami).

Correlations

For the mean values of nine traits across three environments, highly significant and positive correlations were found among Zn, Fe with brown and polished grains *viz*., ZBR, ZPR, IBR and IPR. For the four traits of Zn and Fe in brown and polished grains, highly significant negative correlations with SPY were observed. Similarly, DFF has also showed significant negative correlation with ZBR, IBR and IPR and also with ZPR though not

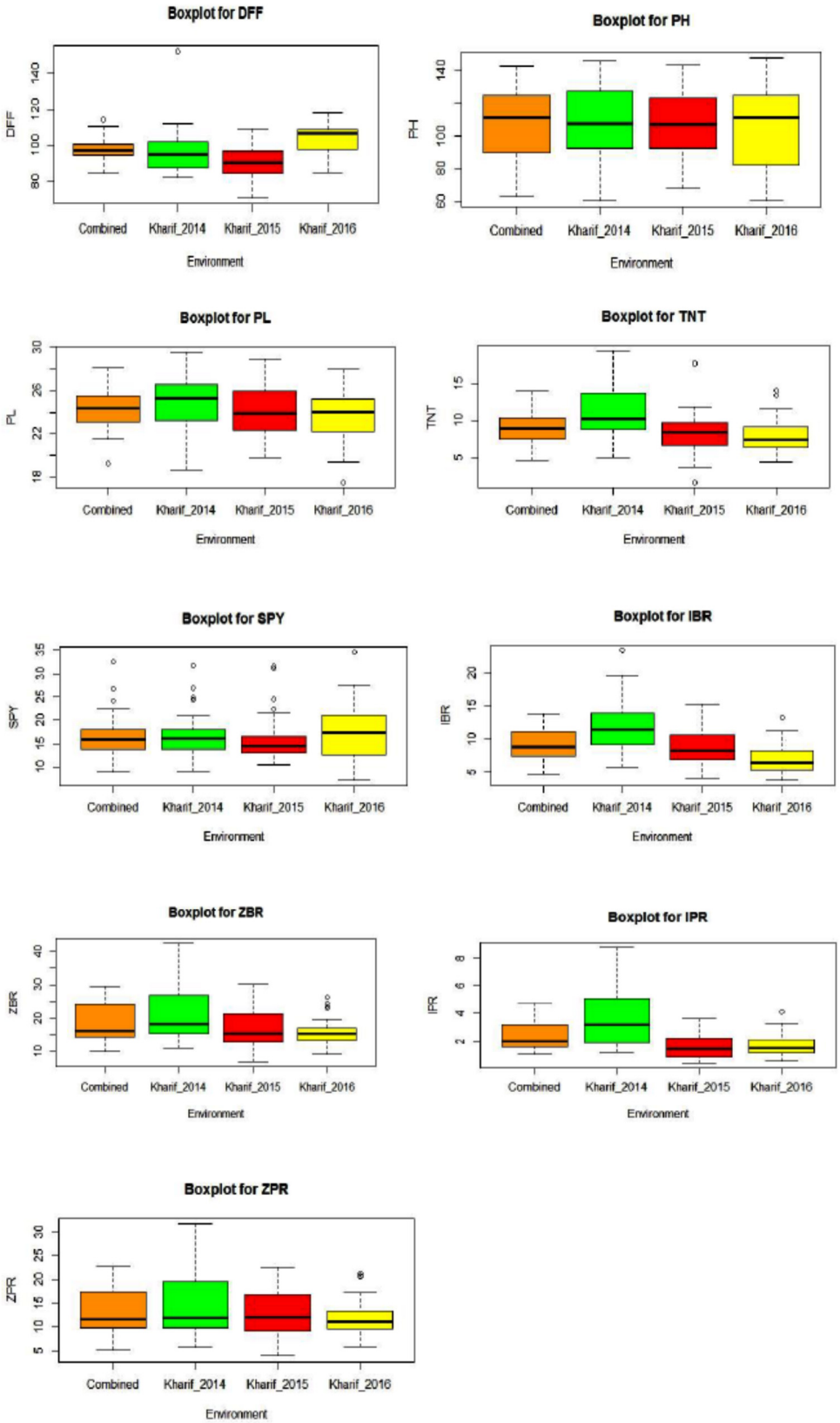


FIGURE 1 | Box plots for nine traits [DFF, Days to fifty percent flowering; PH, Plant Height (cm); PL, Panicle Length (cm); TNT, Total Number of Tillers per plant; SPY, Single Plant Yield (g); IBR, Fe content in Brown Rice (ppm); ZBR, Zn content in Brown Rice (ppm); IPR, Fe content in Polished Rice (ppm); ZPR, Zn content in Polished Rice (ppm)] across three environments E1 – *Kharif* 2014, E2 - *Kharif* 2015, E3 – *Kharif* 2016 and combined across environments.

TABLE 2 | Pearson correlation analysis for mean values of nine traits across three environments among 40 genotypes.

	DFF	PH	PL	TNT	SPY	IBR	ZBR	IPR	ZPR
DFF	1								
PH	−0.246	1							
PL	0.006	0.545**	1						
TNT	−0.232	0.029	−0.125	1					
SPY	−0.086	−0.130	−0.104	−0.043	1				
IBR	−0.056	0.575**	0.263	0.013	−0.420	1			
ZBR	−0.325	0.310	0.116	0.213	−0.431	0.518**	1		
IPR	−0.321	0.391	0.231	0.277	−0.381	0.554**	0.797**	1	
ZPR	−0.293	0.288	0.134	0.093	−0.442	0.560**	0.962**	0.794**	1

***Significant p-Value < 0.001. DFF, Days to fifty percent flowering; PH, Plant Height (cm); PL, Panicle Length (cm); TNT, Total Number of Tillers per plant; SPY, Single Plant Yield (g); IBR, Fe content in Brown Rice (ppm); ZBR, Zn content in Brown Rice (ppm); IPR, Fe content in Polished Rice (ppm); ZPR, Zn content in Polished Rice (ppm).*

TABLE 3 | Analysis of variance of Zn content in Polished Rice (ZPR); Single Plant Yield (SPY) of 40 genotypes across three environments.

Source	DF	SS	MSS	F Value	Pr > F	R-Square	SS(%)
ZPR							
Model	125	10262.72	82.10	443.11	<0.0001	0.995	99.58
Env	2	405.45	202.72	1094.12	<0.0001		3.93
Rep(Env)	6	160.76	26.79	144.60	<0.0001		1.56
Genotypes	39	6870.07	176.16	950.73	<0.0001		66.66
Env*Genotypes	78	2826.44	36.24	195.57	<0.0001		27.43
Error	234	43.36	0.19				0.42
Corrected Total	359	10306.07					100.00
SPY							
Model	125	9486.06	75.89	88.88	<0.0001	0.979	97.94
Env	2	125.73	62.87	73.63	<0.0001		1.30
Rep(Env)	6	180.70	30.12	35.27	<0.0001		1.87
Genotypes	39	7142.86	183.15	214.51	<0.0001		73.75
Env*Genotypes	78	2036.78	26.11	30.58	<0.0001		21.03
Error	234	199.79	0.85				2.06
Corrected Total	359	9685.85					100.00

significantly. Interestingly, both IPR and IBR have shown positive correlation with PH (**Table 2**). Environment wise correlation data was given in **Supplementary Table S3**.

Cluster Analysis Based on Phenotype

Dendrogram of 40 genotypes based on nine traits has shown 15 clusters with cluster 1 consisting 19 genotypes, cluster 2 with four genotypes; cluster 3 with four genotypes; cluster 4 with two accessions and the remaining clustering was as single genotypes (**Supplementary Figure S2**).

Stability of Genotypes Across the Environments

Combined ANOVA for the data of 40 genotypes across three environments indicated significant variance and mean sum of squares for genotype as well as for genotype × environment effect (G × E) for nine traits of the study (**Supplementary Table S4**). As important target traits for the development of biofortified rice varieties, the results were elaborated only for ZPR and SPY.

ZPR

Stability analysis of ZPR across all the environments has shown genotypic (G) effect was 66.66%, environment (E) effect was 3.93%, and genotype and environment (G × E) effect was 27.43% (**Table 3**). Three environments showed almost equal discrimination power, whereas E2 was found as the representative environment (**Figure 2A**); as it falls near to the Average-Environment Axis. The G37 (Sirarakhong Manui) was found near the origin was the less interactive genotype. G23 (Kadamakudi Pokkali) and G18 (Akutphou) were found to be the best in E1, whereas G22 (Edavankudi Pokkali) was found to be the best in E2 and E3, and G32 (Mima) was found to be the best in E3 (**Figure 2B**). Based on the mean vs stability G22 (Edavankudi Pokkali), G17 (Taraori Basmati), G27 (Chittimuthyalu) and G26 (Kalanamak) were found to be stable among the given three environments, whereas G32 (Mima) and G23 (Kadamakudi Pokkali) found more unstable genotypes (**Figure 2C**). The graph Which Won Where/What (**Figure 2D**) clearly explains that the G18 (Akut Phou) won in E1, G22 (Edavankudi Pokkali) won in E2 and E3, G32 (Mima) won in E3. The AMMI biplot shows

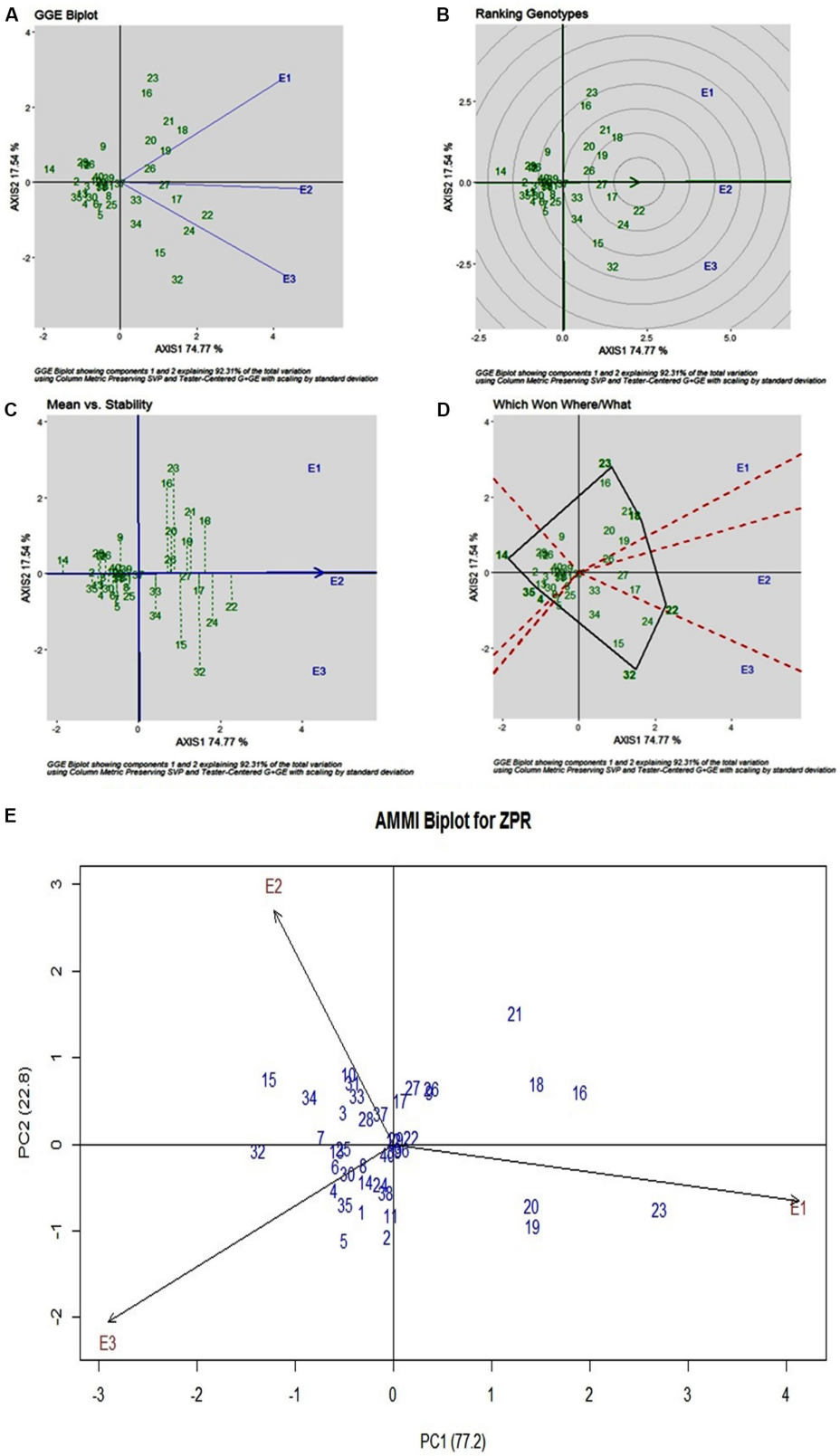


FIGURE 2 | (A) GGE Biplot for Zn content in polished rice (ZPR). **(B)** GGE Biplot for Zn content in polished rice (ZPR)-Ranking of Genotypes. **(C)** GGE Biplot for Zn content in polished rice (ZPR) – Mean v/s Stability. **(D)** Which Won Where/What plot for Zn content in polished rice (ZPR). **(E)** AMMI Biplot for Zn content in polished rice (ZPR).

PC1 contributes 77.2% variability and PC2 contributes 22.8% variability (**Figure 2E**).

SPY

For SPY, across three environments, genotypic (G) effect was 73.75%, environment (E) effect was 1.30%, and genotype and environment (G × E) effect was 21.03% (**Table 3**). All three environments showed almost equal discrimination power, whereas E1 was found as the representative environment (**Figure 3A**), as it falls on the Average-Environment Axis. G24 (DRRDhan 45) was found near the origin and was the less interactive genotype. The G8 (Savitri) was found to be the best in E1, whereas G7 (NDR359) was found to be the best in E2 (**Figure 3B**). Based on the mean *vs* stability G8 (Savitri) was found more stable among three environments, whereas G7 (NDR359) and G20 (Tilakkachari) found more unstable genotypes (**Figure 3C**). The graph Which Won Where/What clearly explains that the G8 (Savitri) won in E1 and E3, whereas G7 (NDR359) won in E2 (**Figure 3D**). The AMMI biplot shows PC1 contributes 95.3% variability and PC2 contributes 4.7% variability (**Figure 3E**).

Stepwise Regression Analysis of ZPR and SPY

Stepwise regression analysis for grain Zn polished rice (ZPR) was carried out to identify the factors influencing ZPR and regression equation for ZPR is depicted as followed.

$$\widehat{ZPR} = 1.068 + 0.73ZBR - 0.27TNT + 0.510IPR \quad (4)$$

Factors like ZBR (92.56%), TNT (1.31%) and IPR (0.4%) explain 94.3% variation in the model (**Table 4a**). For every one-ppm increase in ZBR, there is 0.73 ppm increase in ZPR, thus ZBR has an expected positive effect on ZPR. For every one-tiller increase in TNT, there was a 0.27 ppm decrease in ZPR, thus TNT has negative effect on ZPR. For every one ppm increase in IPR, there was 0.51 ppm increase in ZPR, thus IPR has positive effect on ZPR.

For SPY, the resultant regression equation is as followed

$$\widehat{SPY} = 40.83 - 0.18DFF - 0.52ZPR \quad (5)$$

Factors *viz*., DFF (19.52%) and ZPR (5.7%) explained 25% variation in the model and rest of the variations may be due to other factors which were not considered in this study (**Table 4b**). In terms of individual factors, for every one ppm increase in ZPR, there was a −0.52 g decrease in SPY, implying negative impact of ZPR on SPY. And also for every one day increase in DFF, there was −0.18 g decrease in SPY, indicating that the DFF has negative effect on SPY.

Genotyping by Sequencing ddRAD

Out of 65,670,220 total reads, 52,162,603 reads (90.4%) were aligned with 47,221,247 unique reads (94.7%) (**Supplementary Table S5**). A total of 481,984 variants were called after filtering out duplicated reads with a mean sequence length of 100 bp and phred quality score ≥ 30. Only 452,550 uniquely mapped reads were used for identification of the SNPs. Raw data generated for this study was submitted to the sequence read archive at NCBI under BioProject No: PRJNA626560.

Genome-Wide Discovery of SNPs

Considering variants only at read depth (RD) 10 chromosome wise of 40 genotypes, the range was from 20,531 (chromosome 9) to 43,143 (chromosome 1) for SNPs (**Supplementary Figure S3**) and SNP variants were not uniformly distributed among 12 chromosomes. The variant density was estimated to be 95 SNPs per 100 Kb in comparison with MSU release 7 assembly. Chromosome 2 shows the highest SNP density (111/100 Kb), while the lowest SNP density (81/100 Kb) was found in chromosome 4 (**Supplementary Table S6**). Out of the total 390,346 SNP variants identified, ranged from 7 (G15 Jalpriya) to 34,670 (G10 Jaya) **(Supplementary Table S6**). On pair wise comparisons of polymorphic SNPs, lowest (540) between G24 (DRRDhan 45) and G25 (HP5) and highest between G24 (DRRDhan 45) and G32 (Mima) were observed **(Supplementary Table S7**). A majority of the SNP variants (95.7%) were found to be homozygous. Most of the SNP changes transitions (A/G and C/T) than transversions with Ts/Tv ratio of 1.52 (**Supplementary Table S8**).

Annotation of SNP Variants in the Genomic Regions

The SNP variants were annotated using an in-house pipeline against the gene model provided by MSU release 7. A total of 39,137 polymorphic SNPs variants in the annotated regions were identified across 40 genotypes. Among the total SNPs, 65.95% were intergenic, 16.72% were intronic, 12.95% were in exonic-CDS, 2.82% were in exonic-3UTR, 0.78% were in exonic-5UTR, 0.43% were in intronic-5 splice site and 0.35% in intronic-3 splice site regions (**Supplementary Figure S4**). Across 12 chromosomes, the polymorphic SNPs ranged from 2,320 to 4,388 with a mean of 3,261. The identified SNPs in the exonic-3UTR region were 1,104 and ranged from 34 to 150 with a mean of 92 SNPs per chromosome. In the exonic-5 UTR region, a total of 304 SNPs were identified ranging from 5 to 48 with a mean of 25 SNPs per chromosome. A total of 5,070 SNP variants were identified in the exonic-CDS regions with a range from 309 to 558 with a mean of 423 SNPs per chromosome. A total of 6,544 SNP variants were distributed in intronic regions ranging from 366 to 784 with a mean of 545 SNPs per chromosome. In intronic-3 splice site regions, 138 SNPs were reported and a minimum of 5 SNPs and maximum of 18 SNPs per chromosome were distributed with a mean of 12 SNPs. There were 168 SNPs identified in intronic-5 splice site regions with a mean of 14 SNPs per chromosome which ranged from 19 to 14 SNPs per chromosome. In the intergenic regions, a total of 25,809 SNP variants were identified with minimum of 1,518 SNPs and maximum of 2,902 SNPs per chromosome and with a mean of 2,151 SNPs (**Figure 4** and **Supplementary Table S9**).

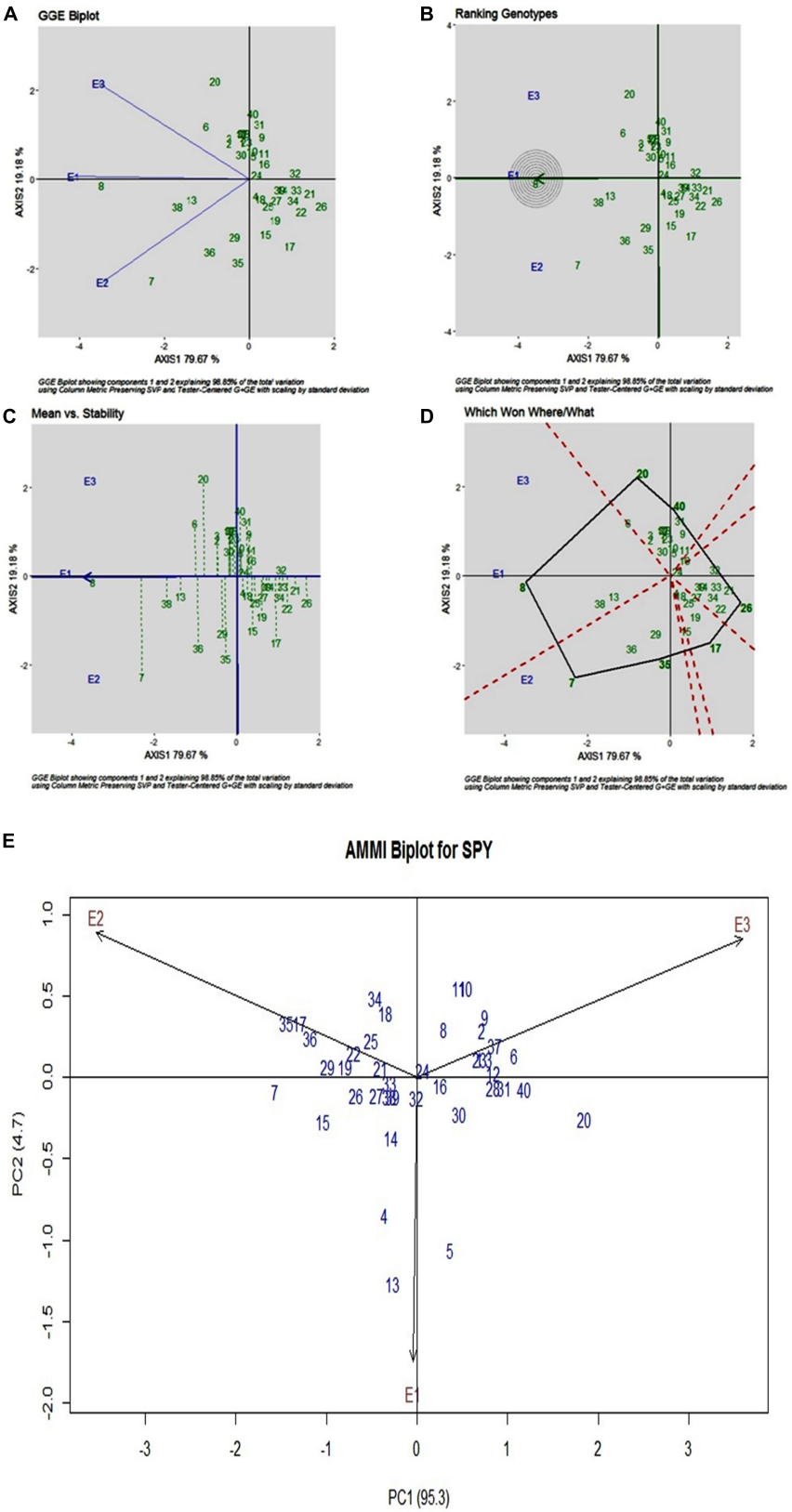


FIGURE 3 | (A) GGE Biplot for Single Plant Yield (SPY). **(B)** GGE Biplot for Single Plant Yield (SPY)-Ranking of Genotypes. **(C)** GGE Biplot for Single Plant Yield (SPY) – Mean v/s Stability. **(D)** Which Won Where/What plot for Single Plant Yield (SPY). **(E)** AMMI Biplot for Single Plant Yield (SPY).

TABLE 4a | Stepwise Regression Analysis of ZPR.

Variable	Parameter Estimate	Standard Error	Probability	Partial R-Square	Model R-Square
Intercept	1.068	0.929	0.2577		0.943
ZBR	0.732	0.053	<0.0001	0.9256	
TNT	−0.278	0.088	0.0033	0.0131	
IPR	0.510	0.299	0.0974	0.0046	

TABLE 4b | Stepwise Regression Analysis of SPY.

Variable	Parameter Estimate	Standard Error	Probability	Partial R-Square	Model R-Square
Intercept	40.83	11.76	0.0013		0.25
DFF	−0.18	0.11	0.1234	0.1952	
ZPR	−0.52	0.15	0.0015	0.0507	

DFF, Days to fifty percent flowering; TNT, Total Number of Tillers per plant; SPY, Single Plant Yield(g); ZBR, Zn content in Brown Rice (ppm); IPR, Fe content in Polished Rice (ppm); ZPR, Zn content in Polished Rice (ppm).

Cluster Analyses Based on Genotyping

Based on 39,137 SNPs, 40 genotypes were separated into three distinct groups, with 19 genotypes as one group, and the other two groups with 16 and five genotypes. Two diverse groups were identified by principal component analysis with well separated lines (**Supplementary Figure S5**). The clustering appears to be coincided with PH *viz.*, group 1 with 93.8 cm); group 2 (124.9 cm) and group 3 (102.8 cm). Modest concurrence of clustering was also observed for ZPR as mean for group1 to be 10.8 ppm, group 2 to be 15 ppm and group 3 to be 15.8 ppm (**Supplementary Table S10**). It was observed that out of 19 genotypes in group 1, 12 genotypes were high yielding varieties.

Association of SNPs With Phenotype

Genome Wide Association Study using 39,137 SNPs from 40 genotypes with phenotype data of three seasons showed that multiple regions are associated with the traits under study ZPR, ZBR, IPR, IBR, SPY, PL, NT, PH, and DFF. Considering the common SNPs across the three years, a total of 188 SNPs ($P < 0.01$) were found and were classified based on their location in exonic, exonic UTR, intronic splice sites, intronic and intergenic regions (**Table 5**). Maximum number of associated SNPs was distributed by intergenic region with 63.8%, followed by intronic region with 18.6%, exonic CDS region with 15.4% and a least of 1.1% SNP in exonic UTR region and intronic-5splice_site region. The associated SNPs with traits and positions are detailed in **Supplementary Table S11**. The chromosome 7 has shown highest number of SNPs (50) and lowest SNPs (2)

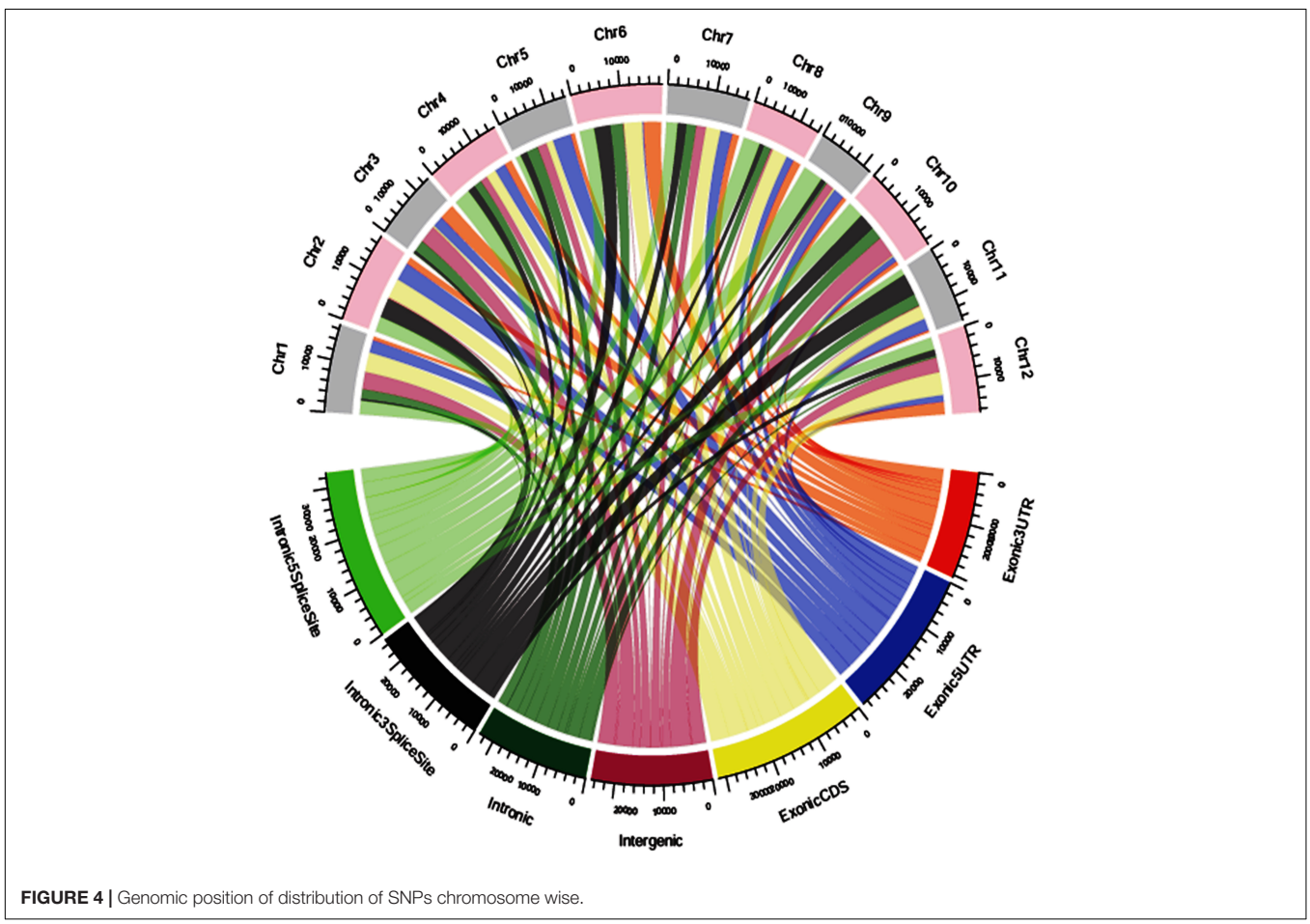


FIGURE 4 | Genomic position of distribution of SNPs chromosome wise.

were identified on chromosomes 5 and 8 with a mean of 15.7 SNPs per chromosome.

Association of SNPs in Putative Candidate Genes With SPY and ZPR

A total of 188 SNPs were found to be associated with six traits across three environments and their genomic positions were detailed in **Supplementary Table S11**. Only the candidate genes associated ZPR and SPY are discussed in detail. The Manhattan and QQ plots of associated SNPs across 40 genotypes for ZPR and SPY were presented in **Figures 5A,B**. Out of three candidate genes found to be associated for ZPR, SNP (A > G) was located in intronic region of LOC_Os03g47980 encoding armadillo like protein and it explained 17% of phenotypic variance (PV). Another SNP (T > C) was located in the exonic CD of LOC_Os07g48050 encoding peroxidase and found to be silent mutation (TGT > TGC) explaining 19% of PV. The third candidate gene LOC_Os07g47950 encoding protein found to be most interesting explaining 23% of PV with three SNPs *viz.*, G > A (GCG > ACG), a missense mutation changing alanine to threonine; G > A (GCG > GCA), a silent mutation and C > T (CTC > TTC), a missense mutation changing leucine into phenylalanine (**Table 6**).

A total of 11 genes has shown association with SPY, from which 6 genes has shown the intronic mutations *viz.* (1) LOC_Os01g64960 -chlorophyll A-B binding protein, putative, expressed gene has intronic mutation of C/A. (2) LOC_Os12g19090-metalloprotease ATP23, putative, expressed has intronic variation C/T. (3) LOC_Os12g19350-expressed protein has intronic mutation of A/G, (4) LOC_Os12g19590 - WD domain, G-beta repeat domain containing protein, has shown intronic mutation of A/T, (5) LOC_Os12g21500-exosome complex exonuclease, putative, expressed has change in intronic region with G/T, and (6) LOC_Os12g22300 - retrotransposon protein, putative, Ty3-gypsy subclass, expressed gene has intronic SNP A/G. Another gene LOC_Os12g19370 encoding DNA polymerase I family protein, has an intronic mutation with A/G and exonic-CD silent mutation of GGC > GGA. A gene LOC_Os01g64890 coding for CorA-like magnesium transporter protein, putative has a silent mutation TCT > TCA in exonic-CD. Three genes have shown missense mutation in exonic-CD region. LOC_Os06g30070 encoding retrotransposon protein, putative, Ty3-gypsy subclass gene has expressed a missense mutation (CCT > CTT) in exonic-CDS region, thus amino acid proline was replaced by leucine. The second gene LOC_Os12g19890 also encodes retrotransposon protein, putative, LINE subclass and has an Exonic-CD region missense mutation (GAG > AAG) changing glutamic acid by lysine. The third gene LOC_Os12g20420 also encoding transposon protein, putative, unclassified, ha also shown a missensed mutation

TABLE 5 | Stable SNPs associated with six traits common across the three seasons.

Trait	Exonic CDS	Exonic UTR	Intronic-5 splice_Site	Intronic	Intergenic	Total
ZPR	4			1	5	10
ZBR	7		1		13	21
IPR	–	–	–	–	1	1
IBR	–	–	–	–	–	0
SPY	6	–	–	11	20	37
PL	4	1	1	11	6	23
NT	–	–	–	–	–	0
PH	8	1	–	12	75	96
DFF	–	–	–	–	–	0
Total SNPs	29	2	2	35	120	188

DFF, Days to 50% flowering; TNT, Total Number of Tillers per plant; SPY, Single Plant Yield(g); ZBR, Zn content in Brown Rice (ppm); IPR, Fe content in Polished Rice (ppm); ZPR, Zn content in Polished Rice (ppm); CDS, coding sequences; UTR, untranslated region.

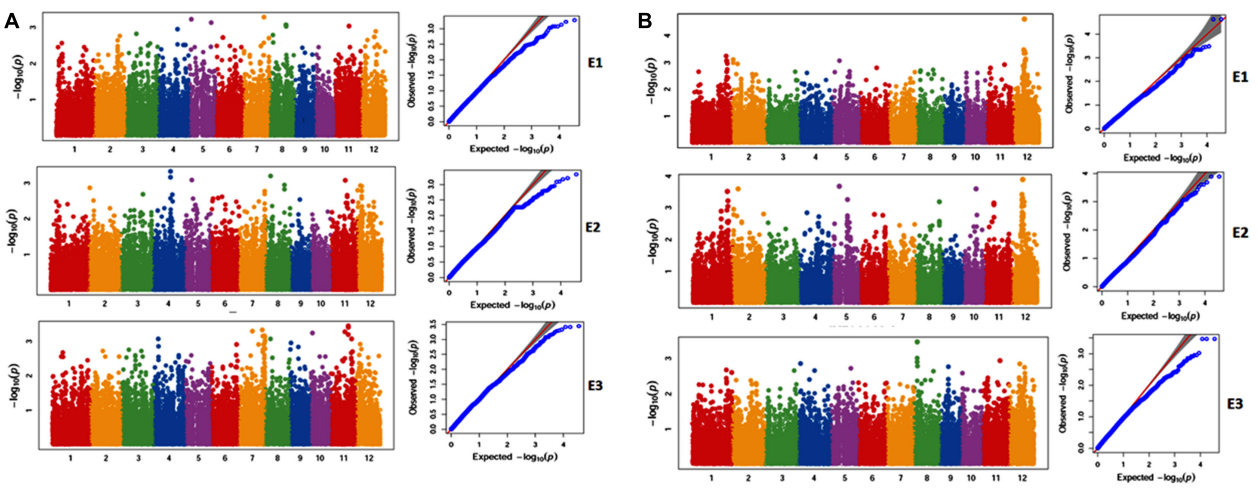


FIGURE 5 | (A) Selected Manhattan and QQ plots of associated SNPs across 40 genotypes for Zn content in polished rice (ZPR). **(B)** Selected Manhattan and QQ plots of associated SNPs across 40 genotypes for Single Plant Yield (SPY).

TABLE 6 | Significant SNPs located in candidate genes associated with ZPR and SPY along with phenotypic variance (PV%).

S No	Gene ID	Chromosome	Marker ID	Genic region	*R*-square	PV %
ZPR						
1	LOC_Os03g47980 (-)	3	C3.27280585	Intronic	0.178590206	17
2	LOC_Os07g47950 (+)	7	C7.28641194	Exonic-CDS	0.24118621	23
		7	C7.28641196	Exonic-CDS	0.24118621	
		7	C7.28641209	Exonic-CDS	0.217452607	
3	LOC_Os07g48050 (+)	7	C7.28688976	Exonic-CDS	0.197142974	19
SPY						
1	LOC_Os01g64960 (+)	1	C1.37698648	Intronic	0.206905872	21
2	LOC_Os01g64890 (+)	1	C1.37665805	Exonic-CDS	0.166038992	17
3	LOC_Os06g30070 (+)	6	C6.17331199	Exonic-CDS	0.183732247	18
4	LOC_Os12g19090 (-)	12	C12.11083190	Intronic	0.254497827	25
5	LOC_Os12g19350 (-)	12	C12.11224184	Intronic	0.247276479	25
6	LOC_Os12g19590 (+)	12	C12.11423711	Intronic	0.186273816	19
7	LOC_Os12g21500 (+)	12	C12.12083507	Intronic	0.247276479	25
8	LOC_Os12g22300 (-)	12	C12.12589709	Intronic	0.247276479	25
9	LOC_Os12g19370 (+)	12	C12.11245053	Intronic	0.247276479	22
		12	C12.11245866	Exonic-CDS	0.185023395	
10	LOC_Os12g19890 (-)	12	C12.11588978	Exonic-CDS	0.362240391	36
11	LOC_Os12g20420 (-)	12	C12.11924838	Exonic-CDS	0.244926128	24

(CGT > TGT) at exonic-CD region changing arginine to cysteine (**Table 6**).

Co-localization of QTL

Out of three associated SNPs located in candidate genes for ZPR, two genes on chromosome 7 were co-localized with earlier reported genomic region with rMQTL7.1 (Raza et al., 2019). The associated candidate gene on chromosome 3 found to be novel. Interestingly, none of the SNPs associated with SPY were found to be co-localized with earlier reported genomic regions.

Genes in the Proximity Regions of (∼600 kbp) Associated SNPs With ZPR and SPY

For ZPR, for the associated SNPs of the candidate genes (LOC_Os07g47950 and LOC_Os07g48050) on chromosome 7, ∼140 genes were noted and for LOC_Os03g47980, 90 genes were found. For SPY, 143 genes on chromosome 1, 82 genes on chromosome 6 and 328 genes on chromosome 12 were observed (**Supplementary Table S12**).

DISCUSSION

Identification of environmentally stable donor lines for high grain Zn would contribute and accelerate the development of biofortified Zn rice varieties. Several landraces have been identified with grain Zn up to 58.4 ppm in brown rice (Gregorio et al., 2000; Norton et al., 2014; Tan et al., 2020) and up to 40.9 ppm in polished rice across world since 2000 (Lee et al., 2008; Bollinedi et al., 2020; Rao et al., 2020). But reports on validation of the identified landraces with high grain Zn for their stability across environments and systematic deployment of identified landraces as donors in breeding of Zn biofortified rice varieties are limited. Stable donor/s identified for the trait of interest can hasten the development of breeding lines/varieties in rice as demonstrated for salinity tolerance from donor 'Pokkali,' submergence tolerance from donor 'FR13A,' phosphorus uptake from 'Kasalath,' aroma and grain quality from 'Basmati' donors (Septiningsih et al., 2009; Chin et al., 2011; Singh et al., 2018; Rana et al., 2019). In the present study, we identified four stable promising donors (Edavankudy Pokkali, Taraori Basmati, Chittimuthyalu and Kalanamak) for high grain Zn > 28 ppm in polished rice for the development of Zn biofortified varieties. Aromatic genotypes like Basmati were reported to be high in grain Zn in earlier studies of International Rice Research Institute (IRRI) (Gregorio et al., 2000). G27 Chittimuthyalu and G26 Kalanamak are included as check genotypes for high grain Zn in the Biofortification trials of rice varietal release program in India through All India Coordinated Rice Improvement Project (AICRIP)[6] (Rao et al., 2020).

In the present study, for released widely adopted popular varieties, the mean value of grain Zn in polished rice was found to be only 14 ppm ranging from 10.2 ppm (G14 Pushyami) to 16.5 ppm (G9 Pusa Basmati 1 and G8 Savitri). Most of the released rice varieties are known to be less in grain Zn content ranging from < 12 to 14 ppm in polished rice (Swamy et al., 2016; Mahender et al., 2016). Akut Phou (G18), a variety included in the present study appears to be already biofortified with polished grain Zn > 24 ppm and reported yield of 4–6 tons/ha. The variety was released from Manipur state of India during 1990 and is only locally popular for its grain quality. Moderate levels of higher grain Zn was also observed for widely adopted rice varieties like Savitri (G8) and Pusa Basmati 1 (G9). Thorough evaluation

[6]http://www.icar-iirr.org/

of released rice varieties for the grain Zn could lead to the identification of base material with proven yield and a possible higher Zn content.

Wide genetic variability was observed for the nine traits of the study considering the diversity of the genetic material comprising released varieties and breeding lines including developed for high grain Zn and landraces from various geographical regions of India (**Table 1**). For IPR, maximum value of 8 ppm was obtained in polished rice, whereas the recommended content is 12 ppm by HarvestPlus[1]. Correlation analyses of the mean values of three environments in the present study corroborated the reported significant positive correlations among Zn and Fe of brown and polished grains and significant negative correlations of grain Zn and Fe with yield (Gao et al., 2006; Norton et al., 2010; Anuradha et al., 2012; Gangashetty et al., 2013; Sathisha, 2013; Inabangan-Asilo et al., 2019). Significant negative correlation of DFF and positive correlation of PH with grain Zn and Fe observed in the study need to be confirmed with large set of genotypes. While there was no significant correlation observed between PL with grain Zn and Fe in our study, Descalsota et al. (2018) reported PL had direct effect on grain Fe and Zn. Significant positive correlation between grain Fe and Zn with plant height and positive correlation between grain Fe and Zn with days to maturity was reported by the same group was also earlier reported by Inabangan-Asilo et al. (2019). The wide range in duration of the genotypes of study, especially of landraces would have contributed toward negative association of DFF with grain Fe and Zn. Stepwise regression analyses clearly depicted the expected positive contribution of ZBR and IPR and negative influence of total number of tillers (R^2 = 0.943) for enhanced ZPR and only ZBR on shown very close association (R^2 = 0.92) (**Tables 4a,b**). The loss of grain Zn during polishing is due to the loss of aleurone layer, which was reported to be significant for mineral content in rice (Sellappan et al., 2009). Previous studies reported a loss to the tune of 5 to 30% of grain Zn during the polishing in germplasm and mapping populations (Lee et al., 2008; Rao et al., 2020). Two significant conclusions from stepwise regression analysis in the present study were the determination of threshold of grain Zn in brown rice *viz*., for fulfilling international threshold value of rice grain Zn content (28 ppm), the developers may need to target $\geq$ 38.5 ppm increase in ZBR. Based on germplasm screening for grain Zn in brown and polished rice, our group suggested germplasm with zinc content $\geq$ 35 mg/kg in brown rice can be promising based on the threshold value of 28 mg/kg by HarvestPlus and 19.0% overall mean loss of zinc during polishing (Rao et al., 2020). Thus, the value of 38.5 ppm of Zn in brown rice obtained through regression analysis in the present study also confirms our earlier study. And the second observation was negative association (R^2 = 0.25) with DFF (R^2 = 0.19) and ZPR (R^2 = 0.5) of SPY suggesting ZPR has only limited negative effect on grain yield and the possibility of simultaneous improvement of yield and grain Zn content as suggested by earlier studies (Swamy et al., 2016; Pradhan et al., 2020). The impact of soil factors on rice grain Zn content was earlier demonstrated through stepwise regression analysis (Pandian et al., 2011).

The influence of plant, soil and climate factors compounds the development of biofortified Zn rice varieties. Many soil criteria like pH, composition, mineral content, biome and agronomic practices like fertilization and irrigation are known to impact the Zn uptake and metabolism in rice plant (Gregorio et al., 2000; Wissuwa et al., 2008; Chandel et al., 2010; Rerkasem et al., 2015). Varying genotypic grain Zn content across locations was widely reported in rice underscoring the importance of identifying stable donors for grain Zn (Norton et al., 2014; Rao et al., 2020). In the present study, based on the stability and G x E interaction analyses of 20 shortlisted genotypes across three years (E1-E3), five promising donors were identified for grain Zn (ZPR) *viz*., G18 (Akut Phou), G22 (Edavankudi Pokkali), G17 (Taraori Basmati), G27 (Chittimuthyalu) and G26 (Kalanamak). For SPY, G8 (Savitri) was identified as stable genotype across environments. AMMI and GGE biplot models suggested the stable performers across the environments and partitioned the total phenotypic variance into individual factors (Gauch, 2006). Both ZPR and SPY showed maximum contribution from genotype because of the genetic material included in the present study. Through Which Won Where/What plot, common winner could not be found for ZPR, whereas G8 (Savitri) won in two environments for yield. Stability and G x E analysis of eight Zn-biofortified rice breeding lines evaluated in four seasons and eight to nine locations also identified different ideal genotypes for yield and Zn in brown rice (Inabangan-Asilo et al., 2019). Significant G $\times$ E interactions were observed among a set of 37 diverse genotypes evaluated in three environments during wet season (kharif), four stable genotypes were identified each for grain Fe and Zn based on regression (Ajmera et al., 2017). Significant G $\times$ E interactions were also reported for grain Fe in a set of 10 genotypes screened at eight environments using the AMMI-biplot and a single stable genotype was identified (Suwarto, and Nasrullah, 2011). Analysis of G x E interaction experiment conducted at same location (ICAR-IIRR) for a set of 14 backcross introgression lines (BILs) for yield related traits during two wet seasons (*Kharif)* and one dry season (*Rabi*) using AMMI, GGE biplot and *Ysi* stastics identified two stable lines with high yield (Balakrishnan et al., 2016). Based on the mean *vs* stability, three landrace genotypes *viz*., G22 (Edavankudi Pokkali), G17 (Taraori Basmati), G27 (Chittimuthyalu) and G26 (Kalanamak) were found to be stable among the given three environments suggesting their utility as donors for high grain Zn. The stable genotype for SPY G8 (Savitri) is one of the mega varieties known for its high yield in several parts of India (Robin et al., 2019). G8 (Savitri) performed well in E1 and E3 when compared to E2 indicating the influence of environmental factors such as temperature, solar radiation and sunshine, relative humidity and wind velocity. G22 (Edavankudi Pokkali) was found stable for E2 and E3 based on GGE biplot and Which Won Where What (W4). Moreover W4 explains G18 (Akut Phou) was stable in E1. The G27 (Chittimuthyalu), G26 (Kalanamak) and G17 (Taraori Basmati) were closer to Mean vs Stability zero line and based on ranking genotypes these genotypes were shown to be relatively stable genotypes. G22, G18, G27, G26 and G17 were stable for ZPR and can be

used as donors in breeding programs for developing high Zn genotypes. For SPY, G8 was stable genotype based on GGE biplot, closer to Mean vs Stability zero line and W4 explains G8 was stable in E1 and common stable genotype in E2 and E3. Thus G8 can be used in future plant breeding programs as a donor for SPY.

With the advent of NGS technologies, development of genome wide SNPs in the target genotypes despite the number has become feasible. ddRAD sequencing facilitates the identification of SNPs by reducing genomic complexity and has been successfully demonstrated in rice and other crops (Yang et al., 2016; Sudan et al., 2019). In the present study, a pair of restriction enzymes digestion *(Sph*I *and MlucI)* was used for the discovery of unique SNPs following the standard procedures *viz*., raw reads de-multiplexing, sequence quality analysis, SNP calling and genotyping (Peterson et al., 2012). A total of 39,137 high quality polymorphic SNPs variants were identified in 40 rice genotypes and annotated with rice genome to identify the region, function and significance. Across chromosomes, ∼1 SNP/1 Kb was observed in the present study, with maximum number of SNPs in intergenic (65.95%), followed by intronic (16.7%), exonic or CDs (12.9%) and intronic-3 splice site (0.35%) regions. Similar trend of SNP localization in the rice genome was earlier reported (Feltus et al., 2004; Ali et al., 2018). The grouping of genotypes based on SNPs showed moderate consensus with the phenotype values of clusters suggesting the contribution of genomic regions/SNPs to phenotypic response as observed in rice introgression lines with tolerance to multiple biotic and abiotic stresses (Ali et al., 2018). Interestingly, five genotypes *viz*., G7 (Narendradhan 359), G8 (Savitri), G16 (ARB-45), G18 (Akut Phou) and G23 (Kadamakudy Pokkali) in cluster 3 found to be promising for DFF, NT, SPY, IBR, ZBR and ZPR, thus they could be deployed in the future breeding program (**Supplementary Table S10**). The information about pairwise combinations between promising donors and popular varieties can be further deployed for mapping the target traits (**Supplementary Table S7**). Association analysis of polymorphic SNPs has shown 188 stable SNPs ($p < 0.01$) significantly associated with six traits *viz*. IPR (1), ZPR (10), ZBR (21), PL (23), SPY (37) and PH (96) across the three environments. Considering only the common SNPs across three environments increased the reliability of identified SNPs and compensated lesser number genotypes of study. While there were a total of 1952 associated SNPs for ZPR across three years (489 in E1; 539 in E2 and 564 in E3), only stable 10 SNPs were selected based on their stability. Similarly, only 37 common stable SNPs were studied from a total of 1319 associated SNPs found (436 in E1; 44 in E2 and 395 in E3 for SPY). Similar studies of 17 marker-phenotype associations were reported for grain Fe and Zn in 102 rice genotypes with 25 gene specific and 75 SSRs (Pradhan et al., 2020). Genotyping of 175 accessions with 155 SSRs revealed 60 marker- trait associations with eight grain elements (Nawaz et al., 2015). Association mapping of 378 accessions with 143 gene specific and SSR markers identified 20 QTL for five grain elements (Huang et al., 2015). GBS analyses of 144 rice MAGIC Plus lines using 14,242 SNP markers identified 57 significant genomic regions associated with various traits including nutritional quality (Descalsota et al., 2018). Several subspecies significant specific association loci for grain mineral concentration were reported through GWAS of *indica* and *japonica* accessions (Zhang et al., 2018; Tan et al., 2020). GWAS of ∼300 accessions with 36,901 SNPs identifies 16 associations for three grain elements as top 1% most significant in at least four of the five field sites (Norton et al., 2014).

Out of 10 stably associated SNPs for ZPR, three markers were located in exonic-CDS region encoding three putative candidate genes C3.27280585 (LOC_Os03g47980), C7.28641194 (LOC_Os07g47950) and C7.28688976 (LOC_Os07g48050). While the association on chromosome 3 appears to be novel, the identified SNPs around ∼28 Mb on chromosome 7 have shown consistent co-localization with genomic regions grain Zn in rice reported through association and biparental mapping (Lu et al., 2008; Huang et al., 2015; Zhang et al., 2018). Recently, a meta QTL rMQTL7.1 has also been detected in this region through meta-analysis (Raza et al., 2019). The identification of consistent co-localized genomic region on chromosome 7 asserts the marker-trait associations in the present study. The identified genomic regions explaining moderate PV ranging from 17 – 23% can be pyramided during the breeding of biofortified Zn rice varieties. Candidate gene analyses in the flanking region (∼300 Kb upstream and downstream) of the mapped SNPs showed several genes and gene families involved in the uptake, transport, and accumulation of Zn in plants along with several hypothetical proteins. The results clearly demonstrate the utility of GBS for identification of significantly associated variant loci might involve in the Zn accumulation mechanism of rice. Twelve stable SNP markers associated with SPY were detected on chromosome 1, 6, and 12 explaining 16 to 36% of PV (**Supplementary Table S12**). Interestingly, co-localized genomic regions could not found out for these 12 associated SNPs for yield suggesting the possibility of their novelty. Pairwise polymorphic SNPs identified between the donors and high yielding varieties in the study are being utilized for mapping of grain Zn in the recombinant inbred lines developed between donors and high yielding varieties (Rao et al., 2020).

CONCLUSION

Deploying AMMI model and GGE biplots for analyses of 20 landraces and 20 popular rice varieties across three environments, four stable landraces (Edavankudi Pokkali, Taraori Basmati, Chittimuthyalu, and Kalanamak) were identified as donors for high grain Zn for rice biofortification breeding program. Using regression analysis, contributing factors for ZPR were identified and threshold levels for ZBR for 28 ppm of ZPR were estimated. Through GBS-dd-RAD analyses, 188 stable SNPs with their locations in the genes across three environments were identified for six traits. Functionality of identified SNPs in the three candidate genes identified for ZPR and 11 candidate genes identified for SPY has been analyzed through their location in the gene. The genomic region for ZPR on chromosome 7 has co-localized with reported metaQTL region ($rMQTL_{7.1}$) and is a promising region for marker assisted introgression.

DATA AVAILABILITY STATEMENT

The data has been submitted to NCBI and the accession number has been provided in the manuscript as submitted to the sequence read archive at NCBI under BioProject No: PRJNA626560.

AUTHOR CONTRIBUTIONS

CN conceptualized the idea. PB, KS, and UC conducted the field experiments. CN, PB, SR, GU, NC, and VL carried out the data analysis. PB and CN prepared the manuscript. CN, LR, and SV edited the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

The authors acknowledge Indian Council of Agricultural Research (ICAR) for funding the study under the project CRP-Biofortification and Department of Biotechnology for funding under the project Rice bio-fortification with enhanced iron and zinc in high yielding non-basmati cultivars through marker assisted breeding and transgenic approaches-Phase II.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.00763/full#supplementary-material

FIGURE S1A | Morphological variation of seed of 40 genotypes in the study.

FIGURE S1B | Frequency distribution plots for DFF-Days to 50% flowering; PH-Plant Height(cm); PL- Panicle Length(cm); NT- Number of Tillers per plant and SPY- Single Plant Yield (g).

FIGURE S1C | Frequency distribution plots for IBR-Iron content in Brown Rice (ppm); ZBR- Zinc content in Brown Rice (ppm); IPR- Iron content in Polished Rice (ppm); ZPR- Zinc content in Polished Rice (ppm).

FIGURE S2 | Cluster dendrogram of 40 genotypes based on nine traits (DFF-Days to fifty percent flowering; PH- Plant Height (cm); PL- Panicle Length (cm); TNT- Total Number of Tillers per plant; SPY- Single Plant Yield (g); IBR-Fe content in Brown Rice (ppm); ZBR- Zn content in Brown Rice (ppm); IPR- Fe content in Polished Rice (ppm); ZPR- Zn content in Polished Rice (ppm).

FIGURE S3 | Chromosome wise total SNPs of 40 genotypes at Read Depth 10.

FIGURE S4 | Distribution of SNP variants in the genomic region.

FIGURE S5 | Dendrogram of 40 genotypes based on 39,137 SNPs.

TABLE S1 | Weather data of three environments of the study.

TABLE S2 | Descriptive statistic values of nine traits across three environments along with combined values for 40 genotypes.

TABLE S3 | Environment wise correlation analysis of nine traits among 40 genotypes.

TABLE S4 | Analysis of variance of nine traits of 40 genotypes across three environments.

TABLE S5 | Alignment statistics of ddRAD reads of 40 genotypes.

TABLE S6 | Genotype wise polymorphic SNPs (Read Depth = 10).

TABLE S7 | Pair wise polymorphic SNPs between the genotypes.

TABLE S8 | Transisionsvstransversions in the SNP variants of the study.

TABLE S9 | Distribution of SNP variants across the chromosome in the genomic regions.

TABLE S10 | Phenotypic values of genotypic groups.

TABLE S11 | Details of associated SNPs with traits and their genomic positions.

TABLE S12 | Genes in the proximate regions of associated SNPs with ZPR and SPY.

REFERENCES

Ajmera, S., Kumar, S. S., and Ravindrababu, V. (2017). Studies on stability analysis for grain yield and its attributes in rice (*Oryza sativa* L.) genotypes. *Int. J. Pure Appl. Biosci.* 5, 892–908. doi: 10.18782/2320-7051.4079

Ali, J., Aslam, U. M., Tariq, R., Murugaiyan, V., Schnable, P. S., Li, D., et al. (2018). Exploiting the genomic diversity of rice (*Oryza sativa* L.): SNP-typing in 11 early-backcross introgression-breeding populations. *Front. Plant Sci.* 9:849. doi: 10.3389/fpls.2018.00849

Anuradha, K., Agarwal, S., Rao, Y. V., Rao, K. V., Viraktamath, B. C., and Sarla, N. (2012). Mapping QTLs and candidate genes for iron and zinc concentrations in unpolished rice of Madhukar× Swarna RILs. *Gene* 508, 233–240. doi: 10.1016/j.gene.2012.07.054

Babu, V. R., Neeraja, C. N., SanjeevaRao, D., Sundaram, R. M., Longvah, T., Usharani, G., et al. (2014). *Bioforotification in Rice*. DRR Technical Bulletin No. 81/2014. Hyderabad: Directorate of Rice Research, 86.

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z. A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PLoS One* 3:e3376. doi: 10.1371/journal.pone.0003376

Balakrishnan, D., Subrahmanyam, D., Badri, J., Raju, A. K., Rao, Y. V., Beerelli, K., et al. (2016). Genotype× environment interactions of yield traits in backcross introgression lines derived from *Oryza sativa* cv. Swarna/*Oryza nivara*. *Front. Plant Sci.* 7:1530. doi: 10.3389/fpls.2016.01530

Bodanapu, R., Vasudevan, S. V., Lekkala, S. P., Chakravartty, N., Lalam, K., Kuriakose, B., et al. (2019). Deciphering the unique SNPs among leading Indian tomato cultivars using double digestion restriction associated DNA sequencing. *bioRxiv* doi: 10.1101/541227

Bollinedi, H., Yadav, A. K., Vinod, K. K., Gopala Krishnan, S., Bhowmick, P. K., Nagarajan, M., et al. (2020). Genome-wide association study reveals novel marker-trait associations (MTAs) governing the localization of Fe and Zn in the rice grain. *Front. Genet.* 11:213. doi: 10.3389/fgene.2020.00213

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

Chandel, G., Banerjee, S., See, S., Meena, R., Sharma, D. J., and Verulkar, S. B. (2010). Effects of different nitrogen fertilizer levels and native soil properties on rice grain Fe, Zn and protein contents. *Rice Sci.* 17, 213–227. doi: 10.1016/S1672-6308(09)60020-2

Chin, J. H., Gamuyao, R., Dalid, C., Bustamam, M., Prasetiyono, J., Moeljopawiro, S., et al. (2011). Developing rice with high yield under phosphorus deficiency: Pup1 sequence to application. *Plant Physiol.* 156, 1202–1216. doi: 10.1104/pp.111.175471

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., and Blaxter, M. L. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi: 10.1038/nrg3012

De Leon, T. B., Linscombe, S., and Subudhi, P. K. (2016). Molecular dissection of seedling salinity tolerance in rice (*Oryza sativa* L.) using a high-density GBS-based SNP linkage map. *Rice* 9:52.

Descalsota, G. I., Swamy, B. P., Zaw, H., Inabangan-Asilo, M. A., Amparado, A., Mauleon, R., et al. (2018). Genome-wide association mapping in a rice MAGIC Plus population detects QTLs and genes useful for biofortification. *Front. Plant Sci.* 9:1347. doi: 10.3389/fpls.2018.01347

Deschamps, S., Llaca, V., and May, G. D. (2012). Genotyping-by-sequencing in plants. *Biology* 1, 460–483. doi: 10.3390/biology1030460

Dumble, S. (2017). *GGEBiplots: GGE Biplots with 'ggplot2'. R package version 0.1.1.*

Endelman, J. B., and Jannink, J. L. (2012). Shrinkage estimation of the realized relationship matrix. *G3* 2, 1405–1413. doi: 10.1534/g3.112.004259

Feltus, F. A., Wan, J., Schulze, S. R., Estill, J. C., Jiang, N., and Paterson, A. H. (2004). An SNP resource for rice genetics and breeding based on subspecies indica and japonica genome alignments. *Genome Res.* 14, 1812–1819. doi: 10.1101/gr.2479404

Furuta, T., Ashikari, M., Jena, K. K., Doi, K., and Reuscher, S. (2017). Adapting genotyping-by-sequencing for rice F2 populations. *G3* 7, 881–893. doi: 10.1534/g3.116.038190

Gali, K. K., Sackville, A., Tafesse, E. G., Lachagari, V. B., McPhee, K. E., Hybl, M., et al. (2019). Genome-wide association mapping for agronomic and seed quality traits of field pea (*Pisum sativum* L.). *Front. Plant Sci.* 10:1538. doi: 10.3389/fpls.2019.01538

Gangashetty, P., Salimath, P. M., and Hanamaratti, N. G. (2013). Association analysis in genetically diverse non-basmati local aromatic genotypes of rice (*Oryza sativa* L). *Mol. Plant Breed.* 4, 31–37.

Gao, X., Hoffland, E., Stomph, T., Grant, C. A., Zou, C., and Zhang, F. (2006). Improving Zn bioavailability in transition from flooded to aerobic rice: a review. *Agron. Sustain. Dev.* 32, 465–478. doi: 10.1007/s13593-011-0053-x

Gauch, H. G. Jr., and Zobel, R. W. (1997). Identifying mega-environments and targeting genotypes. *Crop Sci.* 37, 311–326. doi: 10.2135/cropsci1997.0011183x003700020002x

Gauch, H. G. J. (2006). Statistical analysis of yield trials by AMMI and GGE. *Crop Sci.* 46, 1488–1500. doi: 10.2135/cropsci2005.07-0193

Graham, R. D., Welch, R. M., and Bouis, H. E. (2001). Addressing micronutrient malnutrition through enhancing the nutritional quality of staple foods. *Adv. Agron.* 70, 77–142. doi: 10.1016/s0065-2113(01)70004-1

Gregorio, G. B., Senadhira, D., Htut, H., and Graham, R. D. (2000). Breeding for trace mineral density in rice. *Food Nutr. Bull.* 21, 382–386. doi: 10.1177/156482650002100407

HarvestPlus and FAO (2019). *Biofortification: a Food Systems Solution to Help end Hidden Hunger*. Available online at: www.HarvestPlus.org (accessed December 29, 2019).

Huang, Y., Tong, C., Xu, F., Chen, Y., Zhang, C., and Bao, J. (2015). Variation in mineral elements in grains of 20 brown rice accessions in two environments. *Food Chem.* 192, 873–878. doi: 10.1016/j.foodchem.2015.07.087

Inabangan-Asilo, M. A., Swamy, B. M., Amparado, A. F., Descalsota-Empleo, G. I., Arocena, E. C., and Reinke, R. (2019). Stability and G× E analysis of zinc-biofortified rice genotypes evaluated in diverse environments. *Euphytica* 215:61.

Langmead, B., and Salzberg, S. L. (2012). Fast gapped-read alignment with Bowtie 2. *Nat. Methods* 9, 357–359. doi: 10.1038/nmeth.1923

Lee, J. H., Lee, K. S., Hwang, H. G., Yang, C. I., Lee, S. B., Choi, Y. H., et al. (2008). Evaluation of iron and zinc content in rice germplasms. *Korean J. Breed. Sci.* 40, 101–105.

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

Li, Y., Suontana, M., Burdon, R. D., and Dungey, H. S. (2017). Genotype by environment interactions in forest tree breeding: review of methodology and perspectives on research and application. *Tree Genet. Genomes* 13, 1–18.

Lipka, A. E., Tian, F., Wang, Q., Peiffer, J., Li, M., Bradbury, P. J., et al. (2012). GAPIT: genome association and prediction integrated tool. *Bioinformatics* 28, 2397–2399. doi: 10.1093/bioinformatics/bts444

Lu, K., Li, L., Zheng, X., Zhang, Z., Mou, T., and Hu, Z. (2008). Quantitative trait loci controlling Cu, Ca, Zn, Mn and Fe content in rice grains. *J. Genet.* 87, 305–310. doi: 10.1007/s12041-008-0049-8

Ludwig, Y., and Slamet-Loedin, I. H. (2019). Genetic biofortification to enrich rice and wheat grain iron: from genes to product. *Front. Plant Sci.* 10:833. doi: 10.3389/fpls.2019.00833

Mahender, A., Anandan, A., Pradhan, S. K., and Pandit, E. (2016). Rice grain nutritional traits and their enhancement using relevant genes and QTLs through advanced approaches. *Springerplus* 5:2086.

Majumder, S., Datta, K., and Datta, S. K. (2019). Rice biofortification: high Iron, Zinc, and vitamin-A to fight against "hidden hunger". *Agronomy* 9:803. doi: 10.3390/agronomy9120803

Malosetti, M., Ribaut, V., and Van Eeuwiik, F. A. (2013). The statistical analysis of multi-environment data: modeling genotype by- environment interaction and its genetic basis. *Front. Physiol.* 4:44. doi: 10.3389/fphys.2013.00044

Mendiburu, D. F., and de Mendiburu, M. F. (2019). *Package 'agricolae'. R package version, 1–2*.

Nachimuthu, V. V., Muthurajan, R., Duraialaguraja, S., Sivakami, R., Pandian, B. A., Ponniah, G., et al. (2015). Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in *Oryza sativa*. *Rice* 8:30.

Nawaz, Z., Kakar, K. U., Li, X. B., Li, S., Zhang, B., Shou, H. X., et al. (2015). Genome-wide association mapping of quantitative trait loci (QTLs) for contents of eight elements in brown rice (*Oryza sativa* L.). *J. Agric. Food Chem.* 63, 8008–8016. doi: 10.1021/acs.jafc.5b01191

Neeraja, C. N., Babu, V. R., Ram, S., Hossain, F., Hariprasanna, K., Rajpurohit, B. S., et al. (2017). Biofortification in cereals: progress and prospects. *Curr. Sci.* 113, 1050–1057.

Neeraja, C. N., Kulkarni, K. S., Babu, P. M., Rao, D. S., Surekha, K., and Babu, V. R. (2018). Transporter genes identified in landraces associated with high zinc in polished rice through panicle transcriptome for biofortification. *PLoS One* 13:e0192362. doi: 10.1371/journal.pone.0192362

Norton, G. J., Deacon, C. M., Xiong, L., Huang, S., Meharg, A. A., and Price, A. H. (2010). Genetic mapping of the rice ionome in leaves and grain: identification of QTLs for 17 elements including arsenic, cadmium, iron and selenium. *Plant Soil* 329, 139–153. doi: 10.1007/s11104-009-0141-8

Norton, G. J., Douglas, A., Lahner, B., Yakubova, E., Guerinot, M. L., Pinson, S. R., et al. (2014). Genome wide association mapping of grain arsenic, copper, molybdenum and zinc in rice (*Oryza sativa* L.) grown at four international field sites. *PLoS One* 9:e89685. doi: 10.1371/journal.pone.0089685

Palanog, A. D., Calayugan, M. I., Descalsota-Empleo, G. I., Amparado, A., Asilo, M. A., Arocena, E. C., et al. (2019). Zinc and iron nutrition status in the Philippines population and local soils: a review. *Front. Nutr.* 6:81. doi: 10.3389/fnut.2019.00081

Pandian, S. S., Robin, S., Vinod, K. K., Rajeswari, S., Manonmani, S., Subramanian, K. S., et al. (2011). Influence of intrinsic soil factors on genotype-by-environment interactions governing micronutrient content of milled rice grains. *Aust. J. Crop Sci.* 5, 1737–1744.

Peterson, B. K., Weber, J. N., Kay, E. H., Fisher, H. S., and Hoekstra, H. E. (2012). Double digest RADseq: an inexpensive method for de novo SNP discovery and genotyping in model and non-model species. *PLoS One* 7:e37135. doi: 10.1371/journal.pone.0037135

Pradhan, S. K., Pandit, E., Pawar, S., Naveenkumar, R., Barik, S. R., Mohanty, S. P., et al. (2020). Linkage disequilibrium mapping for grain Fe and Zn enhancing QTLs useful for nutrient dense rice breeding. *BMC Plant Biol.* 20:57. doi: 10.1186/s12870-020-2262-4

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

Rana, M. M., Takamatsu, T., Baslam, M., Kaneko, K., Itoh, K., Harada, N., et al. (2019). Salt tolerance improvement in rice through efficient SNP marker-assisted selection coupled with speed-breeding. *Int. J. Mol. Sci.* 20:2585. doi: 10.3390/ijms20102585

Rao, D. S., Babu, P. M., Swarnalatha, P., Kota, S., Bhadana, V. P., Varaprasad, G. S., et al. (2014). Assessment of grain zinc and iron variability in rice germplasm using energy dispersive X-ray fluorescence spectrophotometer. *J. Rice Res.* 7, 45–52.

Rao, D. S., Neeraja, C. N., Madhu Babu, P., Nirmala, B., Suman, K., Rao, L. V. S., et al. (2020). Zinc biofortified rice varieties: challenges, possibilities, and progress in India. *Front. Nutr.* 7:26. doi: 10.3389/fnut.2020.00026

Raza, Q., Riaz, A., Sabar, M., Atif, R. M., and Bashir, K. (2019). Meta-analysis of grain iron and zinc associated QTLs identified hotspot chromosomal regions and positional candidate genes for breeding biofortified rice. *Plant Sci.* 288:110214. doi: 10.1016/j.plantsci.2019.110214

Rerkasem, B., Jumrus, S., Yimyam, N., and Prom-u-thai, C. (2015). Variation of grain nutritional quality among Thai purple rice genotypes grown at two different altitudes. *Sci. Asia* 41, 377–385.

Robin, S., Jeyaprakash, P., Amudha, K., Pushpam, R., Rajeswari, S., Manonmani, S., et al. (2019). Rice CR1009 sub 1 (IET 22187)-A new flood tolerant rice variety. *Electron. J. Plant Breed.* 10, 995–1004.

Roohani, N., Hurrell, R., Kelishadi, R., and Schulin, R. (2013). Zinc and its importance for human health: an integrative review. *J. Res. Med. Sci.* 18, 144–157.

Sathisha, T. N. (2013). Genetic variation among traditional landraces of rice with specific reference to nutritional quality. *Karnataka J. Agric. Sci.* 26:474.

Sellappan, K., Datta, K., Parkhi, V., and Datta, S. K. (2009). Rice caryopsis structure in relation to distribution of micronutrients (iron, zinc, b-carotene) of rice cultivars including transgenic indica rice. *Plant Sci.* 177, 557–562. doi: 10.1016/j.plantsci.2009.07.004

Septiningsih, E. M., Pamplona, A. M., Sanchez, D. L., Neeraja, C. N., Vergara, G. V., Heuer, S., et al. (2009). Development of submergence-tolerant rice cultivars: the Sub1 locus and beyond. *Ann. Bot.* 103, 151–160. doi: 10.1093/aob/mcn206

Singh, V., Singh, A. K., Mohapatra, T., and Ellur, R. K. (2018). Pusa Basmati 1121–a rice variety with exceptional kernel elongation and volume expansion after cooking. *Rice* 11:19.

Stein, A. J. (2010). Global impacts of human mineral malnutrition. *Plant Soil* 335, 133–154. doi: 10.1007/s11104-009-0228-2

Sudan, J., Singh, R., Sharma, S., Salgotra, R. K., Sharma, V., Singh, G., et al. (2019). ddRAD sequencing-based identification of inter-genepool SNPs and association analysis in *Brassica juncea*. *BMC Plant Biology* 19:594. doi: 10.1186/s12870-019-2188-x

Suwarto, and Nasrullah. (2011). Genotype× environment interaction for iron concentration of rice in central Java of Indonesia. *Rice Sci.* 18, 75–78. doi: 10.1016/s1672-6308(11)60011-5

Swamy, B. P. M., Rahman, M. A., Inabangan-Asilo, M. A., Amparado, A., Manito, C., Chadha-Mohanty, P., et al. (2016). Advances in breeding for high grain zinc in rice. *Rice* 9:49.

Tafesse, E. G., Gali, K. K., Lachagari, V. B., Bueckert, R., and Warkentin, T. D. (2020). Genome-wide association mapping for heat stress responsive traits in field pea. *Int. J. Mol. Sci.* 21:2043. doi: 10.3390/ijms2106 2043

Tan, Y., Sun, L., Song, Q., Mao, D., Zhou, J., Jiang, Y., et al. (2020). Genetic architecture of subspecies divergence in trace mineral accumulation and elemental correlations in the rice grain. *Theor. Appl. Genet.* 133, 529–545. doi: 10.1007/s00122-019-03485-z

Wissuwa, M., Ismail, A. M., and Graham, R. D. (2008). Rice grain zinc concentrations as affected by genotype, native soil-zinc availability, and zinc fertilization. *Plant Soil* 306, 37–48. doi: 10.1007/s11104-007-9368-4

Yan, W., and Kang, M. S. (2002). *GGE Biplot Analysis: A Graphical Tool for Breeders, Geneticists, and Agronomists*. Boca Raton, FL: CRC Press.

Yang, G. Q., Chen, Y. M., Wang, J. P., Guo, C., Zhao, L., Wang, X. Y., et al. (2016). Development of a universal and simplified ddRAD library preparation approach for SNP discovery and genotyping in angiosperm plants. *Plant Methods* 12:39.

Zhang, G.-M., Zheng, T.-Q., Chen, Z., Wang, Y.-L., Wang, Y., Shi, Y.-M., et al. (2018). Joint exploration of favorable haplotypes for mineral concentrations in milled grains of rice (*Oryza sativa* L.). *Front. Plant Sci.* 9:447. doi: 10.3389/fpls.2018.00447

Zhang, M., Pinson, S. R. M., Tarpley, L., Huang, X. Y., Lahner, B., Yakubova, E., et al. (2014). Mapping and validation of quantitative trait loci associated with concentrations of 16 elements in unmilled rice grain. *Theor. Appl. Genet.* 127, 137–165. doi: 10.1007/s00122-013-2207-5

Zhao, K., Tung, C. W., Eizenga, G. C., Wright, M. H., Ali, M. L., Price, A. H., et al. (2011). Genome-wide association mapping reveals a rich genetic architecture of complex traits in *Oryza sativa*. *Nat. Commun.* 2:467. doi: 10.1038/ncomms1467

Conflict of Interest: VL was employed by company AgriGenome Labs Pvt. Ltd.

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 Genetics

ORIGINAL RESEARCH
published: 15 October 2020
doi: 10.3389/fgene.2020.570731

Genetic Architecture and Anthocyanin Profiling of Aromatic Rice From Manipur Reveals Divergence of *Chakhao* Landraces

*S. Bhuvaneswari[1,2], S. Gopala Krishnan[1], Haritha Bollinedi[1], Supradip Saha[3], Ranjith Kumar Ellur[1], K. K. Vinod[1], I. Meghachandra Singh[2], Narendra Prakash[2], Prolay Kumar Bhowmick[1], M. Nagarajan[4], Nagendra Kumar Singh[5] and Ashok Kumar Singh[1]**

[1] Division of Genetics, Indian Council of Agricultural Research (ICAR)-Indian Agricultural Research Institute, New Delhi, India, [2] ICAR-Research Complex for North Eastern Hill Region, Manipur Centre, Imphal, India, [3] Division of Agricultural Chemicals, ICAR-Indian Agricultural Research Institute, New Delhi, India, [4] Rice Breeding and Genetics Research Centre, ICAR-Indian Agricultural Research Institute, Aduthurai, India, [5] ICAR-National Institute of Plant Biotechnology, New Delhi, India

OPEN ACCESS

Edited by:
Balram Marathi,
Professor Jayashankar Telangana State Agricultural University, India

Reviewed by:
Yafang Shao,
Chinese Academy of Agricultural Sciences, China
Amit Baran Das,
Tezpur University, India

***Correspondence:**
Ashok Kumar Singh
aks_gene@yahoo.com

Specialty section:
This article was submitted to Nutrigenomics, a section of the journal Frontiers in Genetics

Received: 08 June 2020
Accepted: 04 September 2020
Published: 15 October 2020

Citation:
Bhuvaneswari S, Gopala Krishnan S, Bollinedi H, Saha S, Ellur RK, Vinod KK, Singh IM, Prakash N, Bhowmick PK, Nagarajan M, Singh NK and Singh AK (2020) Genetic Architecture and Anthocyanin Profiling of Aromatic Rice From Manipur Reveals Divergence of Chakhao Landraces. Front. Genet. 11:570731. doi: 10.3389/fgene.2020.570731

Aromatic rice of Manipur popularly known as *Chakhao* is a speciality glutinous rice, for which protection under geographical indication in India has been granted recently. The agronomic and nutraceutical variability of the *Chakhao* rice germplasm is yet to be genetically characterized. To address this gap, characterization of ninety-three landraces for agro-morphological traits, grain pigmentation, antioxidant properties, and molecular genetic variation was carried out to unravel their population genetic structure. Two major groups were identified based on pericarp color, namely, purple and non-purple, which showed a significant variation for plant height, panicle length, and grain yield. Molecular marker analysis revealed three subpopulations that could be associated with pericarp pigmentation. Deep purple genotypes formed POP3, *japonica* genotypes adapted to hill environment formed POP1, while POP2 comprised of both *indica* and *aus* types. Liquid chromatography–mass spectrometry (LC-MS) analysis revealed two major anthocyanin compounds in pigmented rices, namely, cyanidin-3-O-glucoside (C3G) and peonidin-3-O-glucoside (P3G). The total anthocyanin content among pigmented genotypes ranged from 29.8 to 275.8 mg.100g^{-1} DW. Total phenolics ranged from 66.5 to 700.3 mg GAE.100g^{-1} DW with radical scavenging activity (RSA) varying between 17.7 and 65.7%. Anthocyanins and phenolics showed a direct relationship with RSA implying the nutraceutical benefits of deep pigmented rice such as Manipur black rice. Aromatic rices from Manipur were found to be genetically diverse. Therefore, efforts need to be made for maintaining the geographic identity of these rice and utilization in breeding for region-specific cultivar improvement.

Keywords: Manipur black rice, population structure, antioxidant properties, SSR markers, anthocyanin, *Chakhao* landraces, diversity

INTRODUCTION

Aromatic rice is superior-quality rice having fragrance along with other grain and cooking quality characteristics. Owing to these properties, they are popular among the consumers realizing a higher market value. In the rice gene pool, aromatic rice cultivars form a distinct group (Group V) as revealed by the isozyme analysis (Glaszmann, 1987; Khush, 2000). At the global level, most popular aromatic rices include Basmati rice from the Indo-Gangetic plains of the Indian subcontinent, Jasmine rice from Thailand, and Sadri rice from Iran.

North-eastern India is one of the major agro-biodiversity hotspots in the world, enriched with more than 10,000 diverse indigenous diverse rice cultivars including both aromatic and non-aromatic rice (Mao et al., 2009). Special among these are distinctly scented landraces such as *Joha* cultivars of Assam (Talukdar et al., 2017), *Chakhao* cultivars of Manipur, *Tai* cultivars of Mizoram, and *Kampti* cultivars of Arunachal Pradesh, which are grown and conserved by farmers over ages and distributed across different ecological niches (Durai et al., 2015; Roy et al., 2015). The Manipur state of North-eastern India is an isolated hilly region encircled by nine hill ranges and a central valley having climate varying from tropical to subalpine (GoM, 2018). *Chakhao*, meaning "delicious rice" in *Manipuri*, is the most popular aromatic rice of Manipur which also includes several lesser-known landraces. *Chakhao* cultivars have either pigmented (black, *amubi*) or non-pigmented (white, *angouba*) rice kernels. The cultivars with colored pericarp are distinct from other rice varieties originating from different parts of India (Gayacharan et al., 2018; Tulsiram et al., 2018). Particularly for sociocultural uses, farmers grow several of the *Chakhao* landraces such as *Chakhao Poireiton*, *Chakhao Amubi*, *Chakhao Sempak*, *Ching Chakhao*, *Chakhao Angouba* in local farm holds in smaller areas, covering less than 10% of the holdings. Historical accounts describe that black rice was restricted to the Royals, and the local *Meitei* community used it only during religious festivals and special occasions (Borah et al., 2018). Among the *Chakhao* rice, those with deep-pigmented kernels are popularly called "Manipur black rice." They possess a high anthocyanin content in the pericarp, conferring antioxidant properties. Recognizing their exquisite nutraceutical quality, geographical indication (GI) status has been conferred to Manipur black rice in 2019 by the Government of India, registering it under GI No. 602 in the Geographical Indication Registry[1] (GoI, 2019).

There are several landraces of aromatic rice of Manipur that share the common epithet *Chakhao* but remain seldom characterized. Earlier explorations during different periods conducted in Manipur have collected several such landraces that are conserved in the National Gene Bank (NGB) at the ICAR-National Bureau of Plant Genetic Resources (ICAR-NBPGR), New Delhi (Hore, 2005). In one study, the genetic diversity of 37 *Chakhao* landraces was assessed using 40 microsatellite (SSR) markers to reveal significant gene diversity (0.673) with markers having a PIC value of 0.63. These landraces were found grouped into six classes, having close correlation with farmers' classification (Roy et al., 2014). Significant variation was also reported for yield-related traits in ten black rice genotypes of Manipur (Asem et al., 2019). By biochemical analyses, Asem et al. (2015) showed that the major anthocyanin fraction of black rice genotypes, *Chakhao Poireiton* and *Chakhao Amubi*, was delphinidin 3-galactoside, with *Chakhao Poireiton* having an average anthocyanin content of 740 mg/kg, and total phenolic content ranging from 5 to 6 g/kg of dried flour. Later, the 26 aromatic compounds were reported from *Chakhao Poireiton*, while 11 were reported from *Chakhao Amubi* (Asem et al., 2017). Another study by Chanu et al. (2016) reaffirmed the presence of high levels of anthocyanins, polyphenols, and zinc content having significant antioxidant activity in *Chakhao* landraces. However, the earlier studies suffered from one or other shortcomings, either having been carried out on a limited number of genotypes or having been characterized only for morphological, biochemical, or molecular variation. Therefore, a comprehensive study was felt necessary to assess the variation among several of the *Chakhao* rice including representative landraces, for agro-morphological, biochemical, and molecular diversity. There is no report of anthocyanin profiling and their variation across different black scented landraces cultivated in Manipur. Accordingly, the present study characterizes one of the comprehensive germplasm sets of aromatic rice landraces originating from Manipur including black rice, for phytochemical properties such as pigmentation, anthocyanin content, and antioxidant activities together with agro-morphological, molecular, and grain qualities such as cooking and aroma.

MATERIALS AND METHODS

A total of 93 aromatic rice germplasm accessions collected from different parts of Manipur covering both hill and valley ecosystems were used in the study (**Supplementary Table S1**). Among these, 79 genotypes were sourced from NGB, ICAR-NBPGR, New Delhi; seven were collected from farmers' field in Manipur and seven were sourced from ICAR-Regional Center for North-Eastern Hill Region (ICAR-RC-NEH), Manipur center. The genotypes were initially grouped based on *a priori* information on pericarp pigmentation, aroma, and adaptation ecologies (**Table 1**). The most contrasting feature of the study material was their diversity for spikelet and pericarp pigmentation (**Figure 1**). All the genotypes were initially multiplied at ICAR-Indian Agricultural Research Institute (ICAR-IARI), New Delhi, during *Kharif* 2017. During *Kharif* 2018, the genotypes were grown at the ICAR-RC-NEH Region, Manipur center, in lowland rainfed conditions. Each genotype was grown in three rows of 2.7 m length with a spacing of 20 cm between rows and 15 cm between plants. The field experiment was laid out in augmented design with four blocks and five non-aromatic checks, *viz*., RC Maniphou 7, RC Maniphou 10, RC Maniphou 11, RC Maniphou 12, and RC Maniphou 13. The experimental crop was raised with standard

[1]http://ipindiaservices.gov.in/GirPublic/Application/Details/602

TABLE 1 | Distribution of aromatic genotypes based on the various locations from which they originated in Manipur, included in the panel of genotypes used in the study based on *a priori* information such as pericarp color, aroma, ecosystem, and local names.

Attributes	Classes	Collection ecosystem		
		Hill	Valley	Unknown
Pericarp color	White	8	23	4
	Light brown	15	9	–
	Variegated brown	3	2	–
	Dark Purple	5	21	–
	Dark brown	–	1	–
	Variegated purple	–	2	–
Aroma	Strong	–	21	–
	Mild	22	31	4
	Undetected*	1	14	–
Local name	*Buhman* ¶	16	–	–
	Chakhao¶	6	57	4
	Maklei	4	–	–
	Napnang Hangmei	2	–	–
	The Vumnu	2	–	–
	Ethe Buw	1	–	–
	Black rice	–	1	–

**Those genotypes in which aroma could not be detected are included here.*
¶Genotypes those share a common name in full or in part are grouped together.

agronomic practices to maintain and harvest a good crop. The postharvest grain quality analysis, estimation of anthocyanin compounds, and molecular work were carried out at the Division of Genetics, and Division of Agricultural Chemicals, ICAR-IARI, New Delhi.

Agro-Morphological Characterization

Morphological observations were taken from five randomly selected uniform-looking plants within each line. Data was recorded on 12 quantitative traits including agro-morphological and grain quality traits and six qualitative pigmentation-related traits. The agro-morphological traits included days to fifty per cent flowering (DF), plant height (PH), panicle number (PN), panicle length (PL), 1000 grain weight (GW), and single plant yield (PY), while the grain quality traits included kernel length (KL), kernel breadth (KW), length–breadth ratio (LR), amylose content (AC), alkali spreading value (AS), and gel consistency (GC). Qualitative morphological information such as pigmentation status of basal leaf sheath (BL), ligule (LG), auricle (AU), and collar (CO) were recorded at the tillering stage as presence/absence, while the color of lemma/palea (LP) and pericarp (PC) was recorded at grain maturity on a scale of 1–9 as per the rice distinctness, uniformity, stability (DUS) guidelines (Rani et al., 2006). Grain quality traits such as amylose content (Juliano, 1971), alkali spreading value (Cagampang et al., 1973), and gel consistency (Little et al., 1958) were analyzed following standard evaluation system (SES) for rice (IRRI, 2013).

Estimation of Anthocyanins, Total Phenols, and Antioxidant Activity in Grains

As anthocyanin is accumulated only in pigmented rice, the estimation of anthocyanin content was limited to a subset of thirty pigmented genotypes mostly with *Chakhao* nomenclature having either black, purple, or brown kernels with two cultivars with white rice kernels as non-pigmented checks. For the estimation of compounds, an anthocyanin-rich black rice extract (ABRE) was prepared from decorticated kernels using the method described by Sompong et al. (2011) with slight modification. Briefly, dehusked rice kernels were finely powdered in a mortar by manual grinding and stored at 4°C. About 100 mg of the flour was extracted with 25 ml of acidified methanol (HCl/methanol, 0.14% v/v) for 30 min at 40°C with ultrasonication two to three times to ensure complete color extraction. The extract was centrifuged at 8000 rpm for 5 min, and the supernatant was evaporated totally with a rotary evaporator (Heidolph Laborota 4001 efficient, Germany) at 40°C. The extract was reconstituted in 5 ml acidified distilled water (0.14% v/v concentrated HCl) and stored under refrigeration at −20°C till further analysis.

Identification and Quantification of Anthocyanin Compounds

Identification of the anthocyanin compound in the black rice kernels was carried out with ABRE from *Chakhao Poireiton* (MAR70) by the liquid chromatography–mass spectrometry (LC-MS) system with Synapt G2 high-definition mass spectrometry (Waters Corp., Milford, Massachusetts) at the Advanced Instrumentation Research Facility, Jawaharlal Nehru University, New Delhi. The sample was eluted with water: methanol (90:10, v/v) with a flow rate of 0.1 ml/minute using the BEH C18 column of 2.1 × 100 mm with particle size 1.7 μm with temperature maintained at 25°C. Operating in a single quadrupole mode, LC-MS employed electrospray ionization (ESI). The instrument scanned over the mass (m)/charge (z) range of 100–1100 in the ESI positive ion mode (Lee, 2010).

Based on identified anthocyanin compounds in LC-MS, total anthocyanin content was quantified in different black rice genotypes by high-performance liquid chromatography (HPLC) as described by Lee (2010). The separation of anthocyanin compounds was carried out in reversed-phase separation with a C18 ODS Hypersil column (Thermo Electron Corporation; 250 × 4.6 mm, 5 μ). Chromatographic analysis was performed on the Waters® HPLC system (Alliance 2695 separation module) with quaternary pumps, an autosampler, and a 2996 photodiode array (PDA) detector and driven by Empower 2 software for data recording.

Mobile phases composed of Solvent A containing water, acetonitrile, and trifluoroacetic acid (TFA) in the proportion 53:46:1 and Solvent B containing 0.1% TFA in HPLC-grade water (Singh et al., 2017) with a run time of 20 minutes. The gradient solvent system with Solvent A (20:60:20:20) and Solvent B (80:40:80:80) at 0–7 min, 7–11 min, 11–16 min, and 16–20 min, respectively, was used for maximum resolution. The flow rate

FIGURE 1 | Variation for grain characters in a representative set of aromatic rice germplasm from Manipur. *Chakhao Angouba* (MAR 67), *Napnang Hangmei* (MAR 31), *Kabo Chakhao* (MAR 93), *Buhman* (MAR 62), *Buhman* (MAR 59), *Buhman* (MAR 60), *Buhman Te* (MAR 51), *Chakhao Phou* (MAR 42), *Ethe Buw* (MAR 77), *Chakhao* (MAR 17), *Chakhao Amubi* (MAR 69), *Chakhao Poireiton* (MAR 70), *Chakhao Poireitol* (MAR 6), *Chakhao Poireiton* (MAR 30), *Chakhao* (MAR 15). Note that genotypes with the same names have significant variation in grain types.

was set at 600 μl per minute, and the column temperature was set at 25°C. The elution of the compounds was monitored at 517 nm wavelength, and peak pick was performed by comparing the retention time with the standard compound. The calibration curves were obtained for standard anthocyanin compounds by plotting different concentrations against the peak area in the chromatogram. By comparing the retention time and peak area with that of standard compounds, anthocyanin content in the sample was obtained.

Quantification of Total Phenolics

To determine the total phenolic content in the ABRE, a modified Folin–Ciocalteu assay (Slinkard and Singleton, 1977; Saikia et al., 2012) was used. Briefly, an aqueous solution consisting of 100 μl of ABRE of the sample, 1.50 ml distilled water, and 100 μl of Folin–Ciocalteu reagent (2N) were mixed well. After 5 min., 300 μl of 20% sodium carbonate was added, mixed well, covered with silver foil, and kept at room temperature for 60 min. A blank was prepared similarly but by substituting the sample diluted mix with distilled water. The absorbance was measured at 765 nm using the Epoch 2 microplate reader (Biotech R, United States). A standard curve was prepared using different concentrations of gallic acid (100, 200, 300, 400, 500 $\mu g.ml^{-1}$) from a stock solution of 10 $mg.ml^{-1}$. The total phenolic content was calculated by the formula (CxV)/W, where C is the gallic acid equivalent (GAE) of the sample ($mg.ml^{-1}$) obtained from the standard curve, V is the volume of the extract in ml, and W is the weight of the sample (g). Total phenolic content is expressed as mg GAE per 100g dry weight (DW).

Antioxidant Activity

Antioxidant activity of the ABRE was tested using the 2,2-diphenyl-1-picrylhydrazyl (DPPH) radical scavenging activity (RSA) (Brand-Williams et al., 1995). Fresh DPPH solution (0.066 mM) was prepared by dissolving 0.0026 g in 100 ml of 95% methanol. 100 μl of the sample extract was added to 2.9 ml of freshly prepared DPPH solution and incubated in the dark at room temperature for 30 min. The absorbance was measured using a spectrophotometer at 517 nm against methanol as a blank and 100 μl of 0.1% acidified water in 2.9 ml of DPPH solution as a control. RSA was calculated and expressed in percentage as $[(A_0\text{-}A_s)/A_0] \times 100$, where A_0 is the absorbance of control and A_s is the absorbance of the sample extract.

Pigmentation of the Rice Kernels

L^*, a^*, and b^* color scales were used to determine the pigmentation of decorticated grain samples using the Hunter-Lab Colorimeter system (Miniscan® XE Plus 4500 L, Virginia, United States). The L value indicated the level of darkness (0–50) and lightness (51–100), the a scale of positive value indicated the redness and negative value the greenness, and the b scale indicated yellowness for the positive value and blueness for the negative value. All three values were required to completely describe an object's color (Hunter and Harold, 1987) and color analysis was carried out as described by Murdifin et al. (2015).

Characterization of Molecular Variation Using Microsatellite Markers

A panel of fifty SSR markers recommended by the generation challenge program (GCP) of the Consultative Group for International Agricultural Research (CGIAR) providing genome-wide coverage was used for analyzing the genetic diversity (Ali et al., 2011). Genotypes representing diverse groups of rice, namely, *indica* (IR64), tropical *japonica* (IRGC3764), temperate *japonica* (Taipei 309), *Aus* (Nagina 22), and aromatic (Taraori Basmati) were included as checks along with 93 germplasm accessions to assess their genetic relatedness and clustering. Leaf samples were collected from the individual genotype, and DNA was extracted using the cetyltrimethylammonium bromide (CTAB) method (Murray and Thompson, 1980). Amplification

by polymerase chain reaction (PCR) was carried out with 25 ng template DNA, 5 pmol of each primer, and 2 × ready-to-use PCR master mix (Genei, Bangalore) in a 10 μl reaction mixture. The PCR amplification parameters included initial denaturation at 95°C for 5 min, followed by 35 cycles of thermal profile consisting of 95°C for 30 s, marker-specific annealing temperature for 30 s, and 72°C at 1 min, and a final extension at 72°C for 10 min. The amplicons were resolved on 3.5% metaphor agarose gel stained with ethidium bromide and visualized on GelDoc XR (Bio-Rad Laboratories Inc., United States). A ladder of 50 bp was used for comparison of allele size.

Data Analyses

The quantitative data was tested for descriptive statistics and normality and analyzed for variance pattern, using STAR software (IRRI, 2014), running under R environment. Significant testing is carried out at a minimum probability level of 95%. The data were subjected to association analyses, and the uncorrelated variables were used for resolving morphological diversity using principal component analysis. Marker data were subjected to diversity analysis, using the simple matching coefficient (SMC) as the estimate of genetic distance (Sokal and Michener, 1958). The diversity matrix was subjected to clustering, using the unweighted neighbor joining method with bootstrapping 10,000 times. The diversity pattern was further resolved for population structure (Pritchard et al., 2000) using Structure v.2.3.4 (Prichard et al., 2010), and subpopulation genetic statistics were worked out using GenAlex v.6.5 (Peakall and Smouse, 2012).

RESULTS

Variability in Agro-Morphological Traits

Trait-based frequency analysis of the aromatic rice germplasm is represented in **Figure 2**. Majority of these genotypes (66.7%) were late flowering (DFF: 111–130 days), tall (>130 cm) with a low number (<11) of longer panicles (26–30 cm), and having a medium range (21–25 g) of 1,000 grain weight. Observations on pigmentation on different parts of rice plants (**Figure 3**) revealed that 42 genotypes (45.2%) showed a purple pigmentation of the basal leaf, 36 genotypes (38.7%) possessed a purple ligule, and 44 genotypes (47.3%) produced a purple auricle and collar. Straw-colored lemma and palea were predominant followed by a black color. Grain quality assessment revealed that majority of them were long and slender followed by a long bold category. Further, most of the genotypes had low amylose content combined with soft gel consistency and intermediate alkali spreading value. 26% of the genotypes were found to possess a dark purple pericarp.

The correlation among different agro-morphological traits (**Figure 4**) depicted that PH and PL were significantly positively related and KW significantly influenced GW and GY. Among the cooking quality traits, GC was negatively correlated with AC. All the pigmentation parameters were negatively associated with GW and GY. It was observed that pigmentation of vegetative plant parts (BL, LG, AU, and CO) was significantly correlated with seed pigmentation (LP and PC).

Analysis of variance (ANOVA) for different agro-morphological traits was assessed (**Table 2**) between pigmented (27) versus non-purple color groups (71) and based on the adaptation ecology and valley (58) versus hill (31) regions. A significant difference was observed between the two-color groups for traits like PH, PL, and GY. The deep purple genotypes were taller and produced longer panicles and lower grain yield as compared to non-deep purple genotypes. Among regions, valley genotypes were taller and possessed a longer panicle length as compared to hill genotypes.

Based on principal component analysis (PCA) using all the agronomic, pigmentation, and grain quality traits, the seven most variable phenotypic traits were identified contributing significantly to the total phenotypic variation (**Supplementary Table S2**). Using the subset traits, the first two principal components (PC) accounted for 76% of the total variance (**Table 3**) with the first PC accounting for 55% of the total variation. The main contributing variables to the first PC were pigmentation status of plant organs, namely, LG, CO, AU, PC, LP, and BL. The second principal component contributed 21% of the total variation and had influence from AC, GC, and PH. The PCA biplot (**Figure 5**) dispersed genotypes clearly across PC1 and PC2. Most of the deep purple and white pericarp genotypes were clustered separately in opposite directions along the PC1 axis, whereas other color categories such as light brown, dark brown, variegated purple, and variegated brown were found dispersed in between. Similarly, the tall genotype MAR103 (*Chakhao Amubi*) (177.4 cm) and dwarf genotype MAR43 (*Chakhao Phou*) (98.0 cm) were placed diagonally opposite along the PC2 axis. Genotypes with low amylose were found coupled with high GC values (MAR43, MAR57, MAR58), which were clearly separated from genotypes with high amylose and low GC values (MAR94, MAR105).

Molecular Variation Based on SSR Markers

Among the fifty SSR markers used, two markers (RM 133 and RM 484) were found to be monomorphic across the germplasm and hence excluded from further analysis. The summary statistics of 48 SSR markers is presented in **Table 4**. A total of 171 alleles were identified, with an average of 3.5 alleles per marker while the number of alleles per marker varied from 2 to 7. A maximum number of seven alleles each were detected with RM 413, RM 552, and RM 144. The major allele frequency was lowest for RM 552 (0.206) and highest for RM 454 (0.959) with a mean of 0.664 (**Supplementary Table S3**). The gene diversity or expected heterozygosity ranged from 0.115 (RM125) to 0.826 (RM 552) with a mean value of 0.443. The chromosome level diversity was maximum for chromosome 11 (0.624), while the minimum diversity was observed in chromosome 6 (0.281). The highest PIC value was obtained for RM 552 (0.802) and lowest for RM 454 (0.078) with a mean of 0.394. Seventeen rare alleles with frequencies less than 5% across accessions were identified in this study. Further, there were 11 unique alleles also. Three of these unique alleles were found in genotypes MAR50, MAR51, and MAR58, all of which had *Buhman* as part of its name.

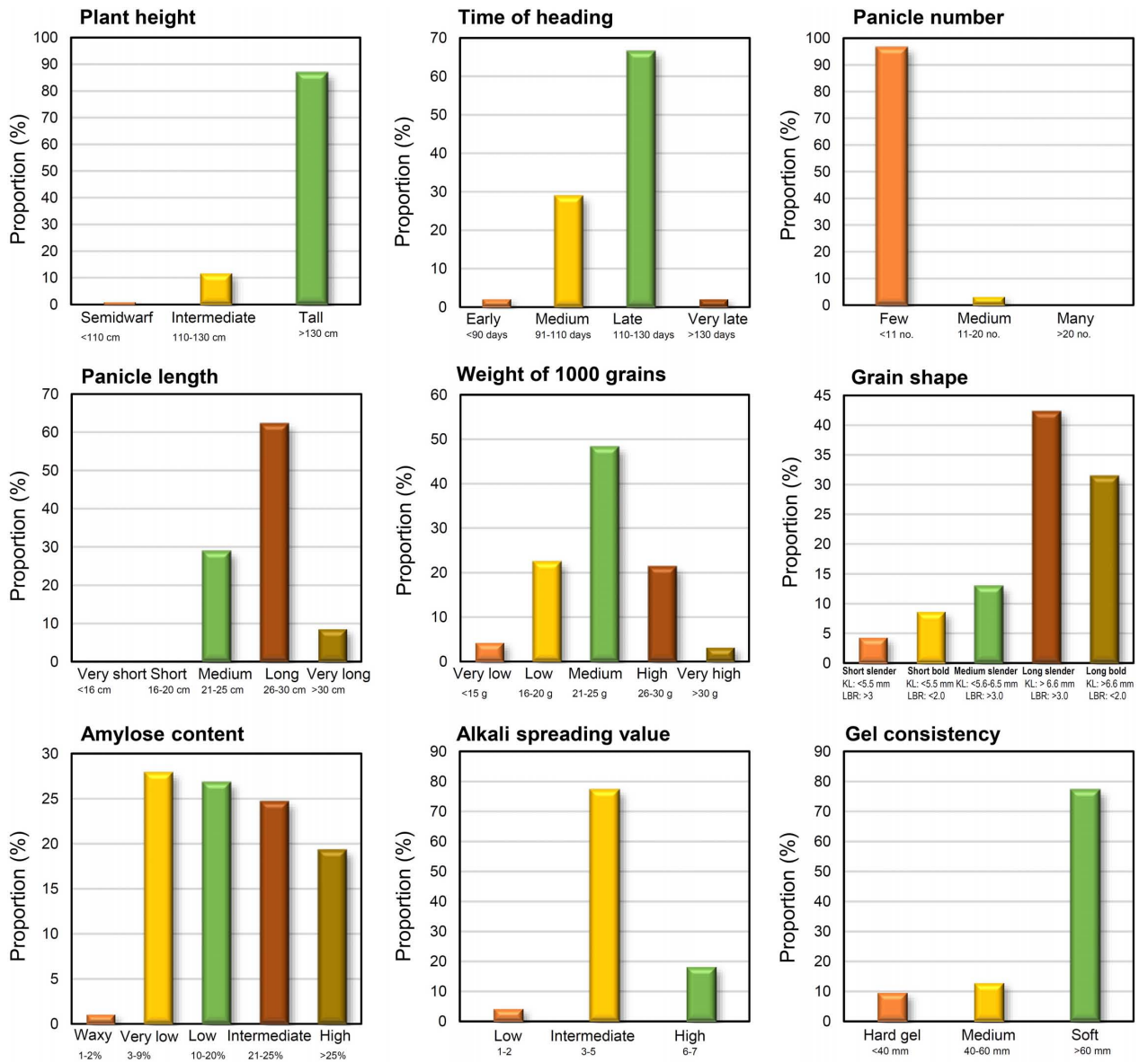


FIGURE 2 | Distribution of aromatic rice genotypes of Manipur, based on agro-morphology and grain quality by standard evaluation system (SES) classes.

Cluster Analysis Using Molecular Data

The genetic distance estimated through SMC, between every pair of genotypes, varied between 0.08 and 0.86 with an average of 0.47. The SMC dissimilarity matrix across 98 genotypes, which included 93 aromatic rice germplasm and five check lines, were used to group the genotypes into three major clusters (**Figure 6**). Cluster I contained 23 genotypes distributed in three subgroups of 12, 5, and 6 genotypes included *japonica*, *Basmati*, and *aus* checks, respectively. Most of the members from Cluster I were from the hill region and possessed a light brown pericarp color. Cluster II included 73 genotypes mostly grown under valley ecology. It included subgroups of pigmented (deep purple, variegated brown, variegated purple), white and light brown pericarp genotypes. Cluster II also included the *indica* check, IR64. Cluster III contained only two *Chakhao* genotypes, MAR10 and MAR11, originating from the valley region.

Population Structure of Aromatic Rice Accessions From Manipur

The Bayesian analysis of the population structure of aromatic rice of Manipur revealed three subpopulations as indicated by the *ad hoc* statistic, ΔK (**Figure 7**). The subpopulations, POP1, POP2, and POP3 included 12, 34, and 52 genotypes, respectively. It was interesting to note that the cluster I subgroups from the earlier analysis were bifurcated into two subpopulations (POP1 and POP2). The members of each

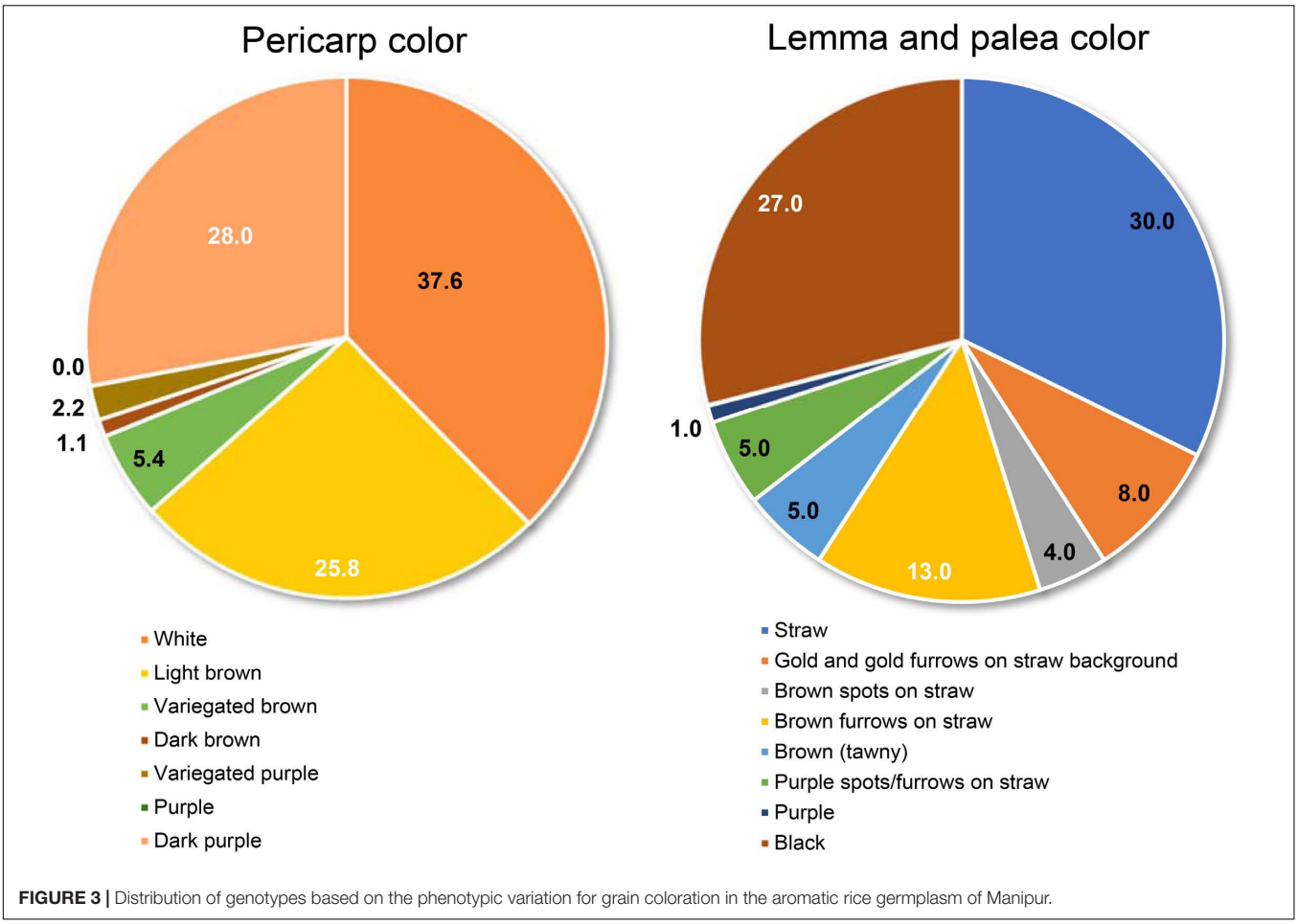


FIGURE 3 | Distribution of genotypes based on the phenotypic variation for grain coloration in the aromatic rice germplasm of Manipur.

subpopulation were further divided as pure or admixed based on inferred ancestry coefficients. Those genotypes with the coefficient of =0.95 were counted as pure types, and those with coefficients < 0.95 were counted as admixtures. Accordingly, POP1 contained nine pure genotypes, which included six of the hill accessions with white and light brown pericarp. The remaining members of this group included a landrace, "*Maklei*," with variegated brown pericarp, the tropical *japonica* check, IRGC3764, and the temperate *japonica* check, Taipei 309. The aromatic group check, Taraori Basmati, was grouped along with *japonica* check, in POP1 as admixture. These hilly genotypes were also grouped as distinct by cluster analysis. They showed a distinctly discernible allelic pattern for the biallelic markers, namely, RM 489, RM 338, RM 161, RM 455, and RM 284, from other genotypes. POP2 included 12 genotypes as pure types and 22 genotypes as admixtures that included *indica* check, IR64, and the *aus* check, Nagina 22. These genotypes either belonged to hill and valley regions or were mostly with white and light brown pericarp. Two *Chakhao* genotypes, MAR10 and MAR11, which formed Cluster III were included in this subpopulation. POP3, the largest subpopulation, included 29 genotypes as pure and 23 genotypes as admixtures. Most of the genotypes (22) in the pure category were dark pigmented (deep purple, variegated purple/brown) from both valley and hill regions.

The analysis of molecular variance (AMOVA) revealed significant variation among and within subpopulations obtained (**Table 5**). Among subpopulations, a variance of 23% was found, however, within subpopulations 77% variance was obtained among individuals. No within-individual variation was found. The population-specific F_{st} of the three subpopulations were 0.260, 0.401, and 0.106, respectively, with an average of 0.256 indicating higher level of genetic differentiation.

Quantification of Anthocyanins, Polyphenols, and Antioxidant Activity

In the ARBE of *Chakhao poireiton* (MAR70), by mass spectrometry, two peaks were detected at m/z of 449.1 and 463.1 corresponding to C3G and P3G, respectively, together with two major peaks at 287.05 and 301.07 (**Figure 8**), which could be identified as cyanidin and peonidin, the aglycons. There were also minor peaks detected which could not be identified due to their very low concentration in the extract. They are quantified subsequently as cyanidin-3-glucoside equivalents (C3GE). The major anthocyanin fractions were further confirmed chromatographically by their retention time (RT) of 11.95 min for

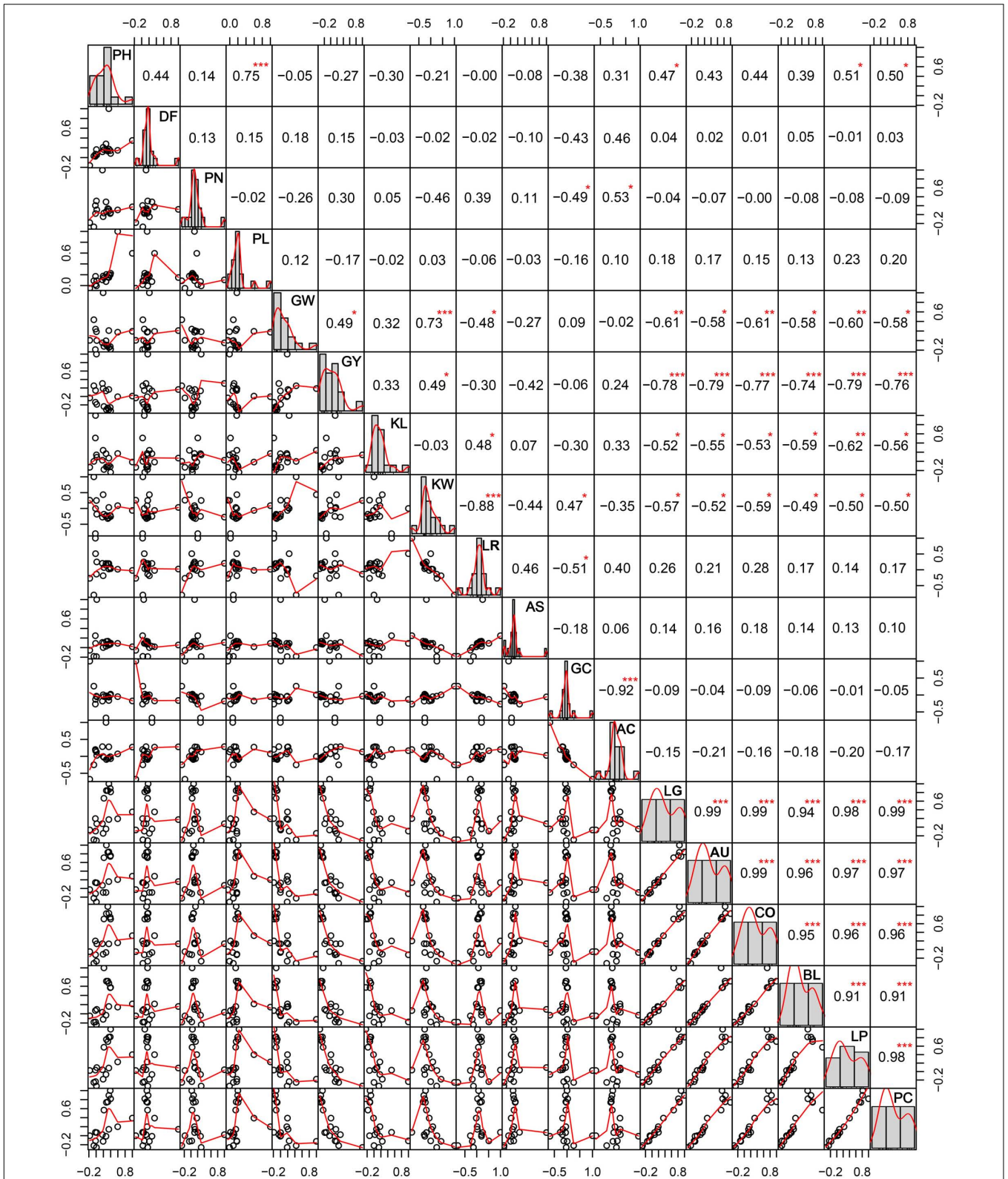


FIGURE 4 | Correlations among different agro-morphological traits in the aromatic rice germplasm of Manipur. PH, plant height in cm; DF, days to 50% flowering; PN, panicle number; PL, panicle length in cm; GW, weight of 1,000 grains in g; GY, grain yield per plant in g; KL, kernel length in mm; KW, kernel width in mm; LR, length to width ratio; AS, alkali spreading value; GC, gel consistency; AC, apparent amylose content in percentage; LG, ligule pigmentation; AU, auricle pigmentation; CO, collar pigmentation; BL, basal leaf color; LP, lemma and palea color; PC, pericarp color. *, **, *** Significant at probability levels of 0.05, 0.01 and 0.001, respectively.

TABLE 2 | Analysis of variance for different agro-morphological and grain quality traits in the aromatic rice germplasm of Manipur over *a priori* classification based on pericarp color and ecosystem adaptation.

Traits	Pericarp color class					Ecosystem class				
	Category	Count	Mean	CV%	*p*-value	Category	Count	Mean	CV%	*p*-value
PH	LB	24	145.5	15.2	0.00	Valley	58	162.6	8.1	0.00
	White	35	154.0	12.0		Hill	31	143.7	13.2	
	DP/VP/DB	34	163.2	6.9						
DF	LB	24	107.7	10.1	0.00	Valley	58	115.1	8.7	0.07
	White	35	116.5	9.2		Hill	31	111.0	9.6	
	DP/VP/DB	34	116.4	7.0						
PN	LB	24	7.2	36.6	0.08	Valley	58	7.3	23.7	0.52
	White	35	6.9	30.7		Hill	31	7.0	36.8	
	DP/VP/DB	34	7.4	20.5						
PL	LB	24	25.5	8.5	0.08	Valley		27.0	8.5	0.03
	White	35	26.3	10.4		Hill		25.8	9.8	
	DP/VP/DB	34	26.9	7.8						
GW	LB	24	23.9	22.6	0.08	Valley	58	25.2	21.6	0.07
	White	35	25.6	21.8		Hill	31	22.9	26.9	
	DP/VP/DB	34	22.9	19.3						
GY	LB	24	12.8	78.5	0.01	Valley	58	11.6	68.5	0.54
	White	35	14.6	54.6		Hill	31	12.7	59.3	
	DP/VP/DB	34	8.9	62.8						
KL	LB	24	6.3	7.8	0.32	Valley	58	6.2	6.9	0.59
	White	35	6.3	8.1		Hill	31	6.3	8.5	
	DP/VP/DB	34	6.1	6.1						
KW	LB	24	2.2	8.2	0.01	Valley	58	2.2	13.1	0.22
	White	35	2.2	13.8		Hill	31	2.1	11.1	
	DP/VP/DB	34	2.1	10.8						
LR	LB	24	2.9	13.3	0.22	Valley	58	2.9	14.2	0.16
	White	35	2.9	13.9		Hill	31	3.0	11.9	
	DP/VP/DB	34	3.0	12.9						
AS	LB	24	2.9	38.4	0.30	Valley	58	5.0	19.3	0.29
	White	35	2.9	41.6		Hill	31	4.8	19.2	
	DP/VP/DB	34	3.0	28.2						
GC	LB	24	113.7	36.6	0.08	Valley	58	91.5	46.2	0.01
	White	35	91.4	42.9		Hill	31	116.2	29.9	
	DP/VP/DB	34	92.7	44.1						
AC	LB	24	14.0	64.7	0.06	Valley	58	17.7	51.7	0.07
	White	35	19.5	45.2		Hill	31	14.2	57.7	
	DP/VP/DB	34	16.6	51.1						

PH, plant height in cm; DF, days to 50% flowering; PN, panicle number; PL, panicle length in cm; GW, weight of 1,000 grains in g; GY, grain yield per plant in g; KL, kernel length in mm; KW, kernel width in mm; LR, length-to-width ratio; AS, alkali spreading value; GC, gel consistency; AC, apparent amylose content in percentage; LB, light brown; DP, dark purple; VP, variegated purple; DB, dark brown.

C3G and 12.9 min for P3G, as obtained from the corresponding standards (**Supplementary Table S4**). There were also minor peaks at lower RT, which are unidentified. Quantification of anthocyanins from thirty genotypes, identified C3G as the major anthocyanin fraction, were followed by P3G (**Figure 8**). There was a significant variation among black pigmented genotypes for total anthocyanins that ranged from 29.8 to 275.8 mg.100g^{-1} (**Table 6**). The genotypes with similar names like *Chakhao Poireiton*, *Chakhao Amubi*, and *Ching Chakhao* collected from different places showed significant variation for total anthocyanin content and its constituent compounds (**Table 6**). The average C3G content (113.3 mg.100g^{-1}) among the genotypes was almost six times higher than the average P3G content (19.7 mg.100g^{-1}) and seven times higher than C3GE (16.2 mg.100g^{-1}). On average, the percent composition of C3G, P3G, and C3GE was 76, 13, and 11, respectively. Total phenolic content was found to be significantly different between white (56.0 mg GAE.100g^{-1} DW) and black (339.2 mg GAE.100g^{-1} DW) rice genotypes. RSA (%) by DPPH assay depicted a significant variation between white (7% on average) and pigmented genotypes assessed (38% on average). Interestingly, among the pigmented genotypes, minimum (17.7) and maximum (65.7) RSA (%) values were

TABLE 3 | Principal components extracted from the most influential phenotypic traits among the aromatic rice germplasm of Manipur.

Components	EV	Var	Contribution of traits (%)								
			LG	AU	CO	BL	LP	PC	PH	GC	AM
PC1	4.93	0.55	18.3	16.8	16.8	12.0	15.6	16.4	3.7	0.4	0.0
PC2	1.93	0.21	0.1	0.6	0.2	0.2	0.3	0.1	17.1	39.2	42.2
PC3	0.68	0.08	0.2	1.4	0.7	19.6	2.8	2.6	56.7	13.3	2.7

EV, eigenvalue; Var, proportion of total variation; LG, ligule pigmentation; AU, auricle pigmentation; CO, collar pigmentation; BL, basal leaf color; LP, lemma and palea color; PC, pericarp color; PH, plant height in cm; GC, gel consistency; AM, amylose content.

observed in deep purple genotypes, *Ching Chakhao Amubi* (MAR1) and *Chakhao* (MAR 91), respectively.

Apparently, color values L^* and b^* were higher in white genotypes as compared to deep black-pigmented genotypes. The b^* value was negative in deeply pigmented genotypes. The parameter a^* depicting redness–greenness was higher in low anthocyanin-containing genotypes (brown) as compared to white and dark purple pigmented genotypes.

Relationship of Pigmentation Features and Nutraceutical Properties

Total anthocyanin content, C3G content, P3G content, and C3GE content were significantly positively associated with total phenolics and RSA (**Figure 9**). Anthocyanin content was high in deep purple genotypes, mostly *Chakhao* and *Chakhao Poireiton*, as compared to dark brown or variegated purple genotypes which were pigmented on the dorsal side of the seed kernel. However, some of the deep purple genotypes (MAR 6, MAR 85) with high anthocyanin content possessed relatively lesser total phenolics as compared to lower anthocyanin content genotypes (MAR 70 and MAR 69) and *vice versa*. However, anthocyanin constituents

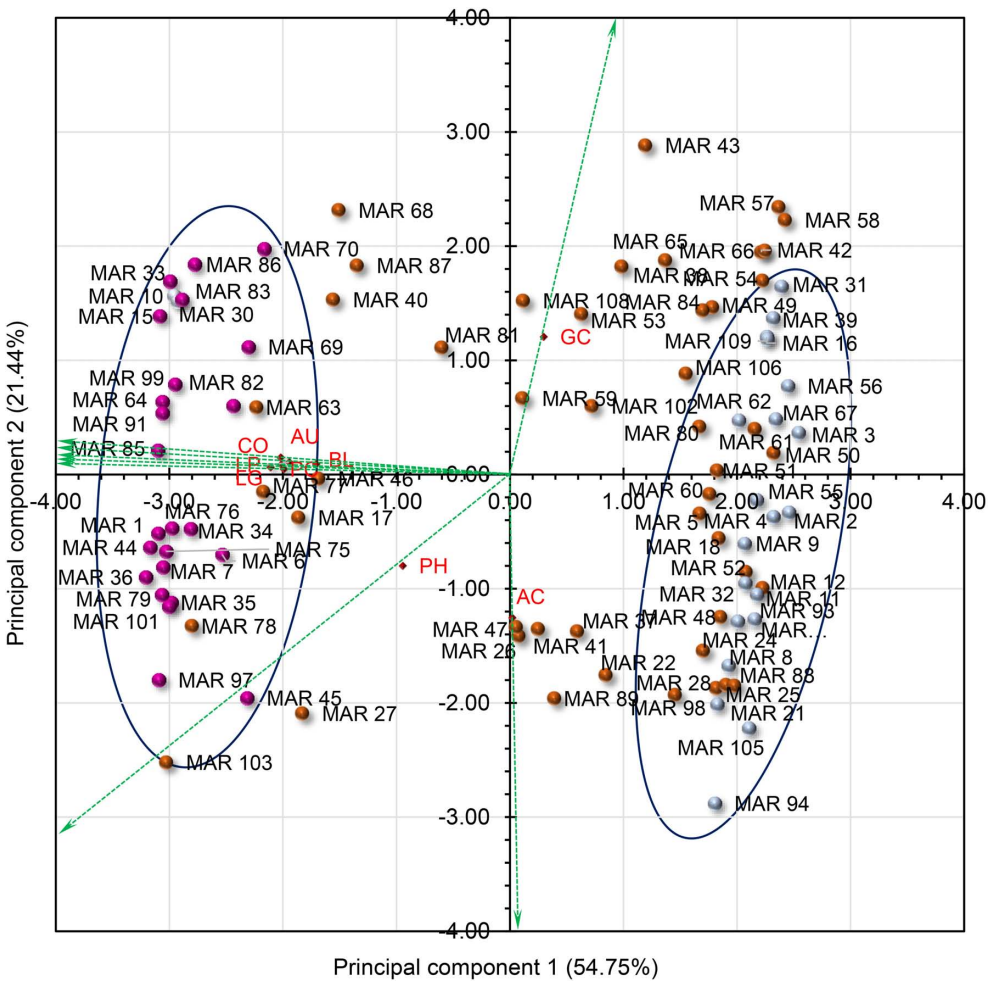


FIGURE 5 | PCA biplot showing the partitioning of Manipur aromatic germplasm using most variable phenotypic traits. There were three groups of genotypes, grouped based on pericarp color. The genotypes with white pericarp (light blue) formed the cluster 1, while deep purple-colored (purple) genotypes formed the second. The third group was intermediate to the first two (brown). The green arrows show the direction of influence of the traits. LG, ligule pigmentation; AU, auricle pigmentation; CO, collar pigmentation; BL, basal leaf color; LP, lemma and palea color; PC, pericarp color; PH, plant height in cm; GC, gel consistency; AC, amylose content.

TABLE 4 | Marker polymorphism and allelic status of GCP panel microsatellite markers among aromatic rice germplasm from Manipur.

Chr	No. of SSR markers	Frequency of multiallelic loci						PIC	
		2	3	4	5	6	7	Mean	Range
Chr 1	8	2	2	3	1	–	–	0.358	0.179–0.447
Chr 2	2	–	1	1	–		–	0.502	0.403–0.600
Chr 3	5	1	4	–	–		–	0.299	0.204–0.383
Chr 4	2	1	–	–	–	1	–	0.430	0.166–0.694
Chr 5	5	1	2	–	–	1	1	0.380	0.157–0.741
Chr 6	3	1	2	–	–	–	–	0.250	0.075–0.386
Chr 7	4	3	–	–	1	–	–	0.258	0.108–0.419
Chr 8	7	2	1	3	1	–	–	0.419	0.182–0.591
Chr 9	3	–	2	1	–	–	–	0.417	0.364–0.498
Chr 10	3	–	1	1	–	1	–	0.488	0.282–0.599
Chr 11	4	–	1	1	–	–	2	0.587	0.233–0.802
Chr 12	2	–	1	1	–	–	–	0.506	0.455–0.556
Overall	48	11	17	11	3	3	3	0.394	0.075–0.802

Chr, chromosome; PIC, polymorphism information content.

C3G and P3G showed a positive association with RSA; some of the genotypes with equal total anthocyanin content (MAR 70, MAR 69, and MAR 36) but with a higher proportion of C3G (MAR 36) showed higher RSA (%). All the color scales (L^*, a^*, b^*) were significantly negatively associated with phytochemical parameters studied. Within color scales, L^* depicting lightness and darkness showed variation within pigmented genotypes wherein dark purple genotypes showed values lower than 20 whereas dark/variegated brown rice possessed between 20 and 25 as compared to white rice genotypes with > 30. The b^* value in dark purple genotypes was either negative or lower value (<3) depicting blueness (−), except *Ching Chakhao Amubi* (MAR 1 with > 3). The white and variegated brown genotypes showed a higher positive value (> 5) depicting yellowness. The value of color scale a^* depicting redness (+) was low in white genotypes as well as deep purple genotypes and higher in dark/variegated brown genotypes. The L^* and b^* scales were strongly positively correlated with each other, whereas the association between L^* and a^* was non-significant. All the color scales were negatively associated with phytochemicals estimated.

Grouping of pigmented genotypes based on the PCA with pigmentation data identified two major PCs accounting for 83.2% of the total variation with PC1 and PC2 explaining 70 and 13% of the total variation, respectively. PC1 was majorly determined by anthocyanins (52%), phenolics (13.66%), and RSA (14.1%). PC2 was mainly influenced by color parameters L^*, b^*, and a^*, contributing 40.5, 34.2, and 9.7%, respectively. The PCA biplot clearly placed the less/non-pigmented genotypes including dark brown (1), variegated brown (4), and white (2) rice categories separately from major deep purple rice genotypes (**Supplementary Figure S1**).

DISCUSSION

Aromatic rice genotypes from north-eastern India, particularly from Manipur, are relatively lesser known cultivars because of their confined cultivation within their geographical adaptation. Recent interest on their culinary and nutraceutical properties, particularly of Manipur black rice, turned the attention of rice scientists in understanding the genetics of these genotypes. Known by their vernacular names, cultivars carrying the epithet *Chakhao—chak* means "rice" and *ahaoba* meaning "delicious" (Dayanidhi et al., 2017)—are a conglomeration of local landraces that share similarity in grain aroma and cooking and taste properties but genetically different. The scented glutinous *Chakhao* with black pericarp are predominantly grown in valley districts of Manipur. The black rice of Manipur has a long history of exclusive adornment of royal cuisine and highly restricted use outside the aristocracy, and their culinary properties were lesser known until recently. Recent studies on black pigmented rice in general, have established their dietary significance especially on nutraceutical properties (Goufo and Trindade, 2014; Samyor et al., 2017). Notwithstanding, the genetic identity of aromatic rices of Manipur, particularly of black rice, needs to be established to realize their potential use. Therefore, there is an immediate need to characterize their genetic diversity and to utilize them in genetic improvement, as well as to conserve them for the posterity of future generations.

Agro-morphological characterization for assessing the genetic diversity in crops draws the foundation for genetic improvement. In rice, plant height, crop duration, and grain yield are three major agronomic parameters used for determining productivity, while secondary characteristics such as panicle length, grain number, grain weight, grain color, and shape are used for classification and identification (Bhandari et al., 2017). Although limited, past research on *Chakhao* rice landraces of Manipur was carried out using fewer genotypes using fewer traits (Roy et al., 2014; Asem et al., 2015; Chanu et al., 2016). Our analyses attempt to address this lacuna, by characterizing a larger set of aromatic landraces for agro-morphological, grain cooking quality, pigmentation, and antioxidant properties as well as at the molecular level. We found no distinct association of

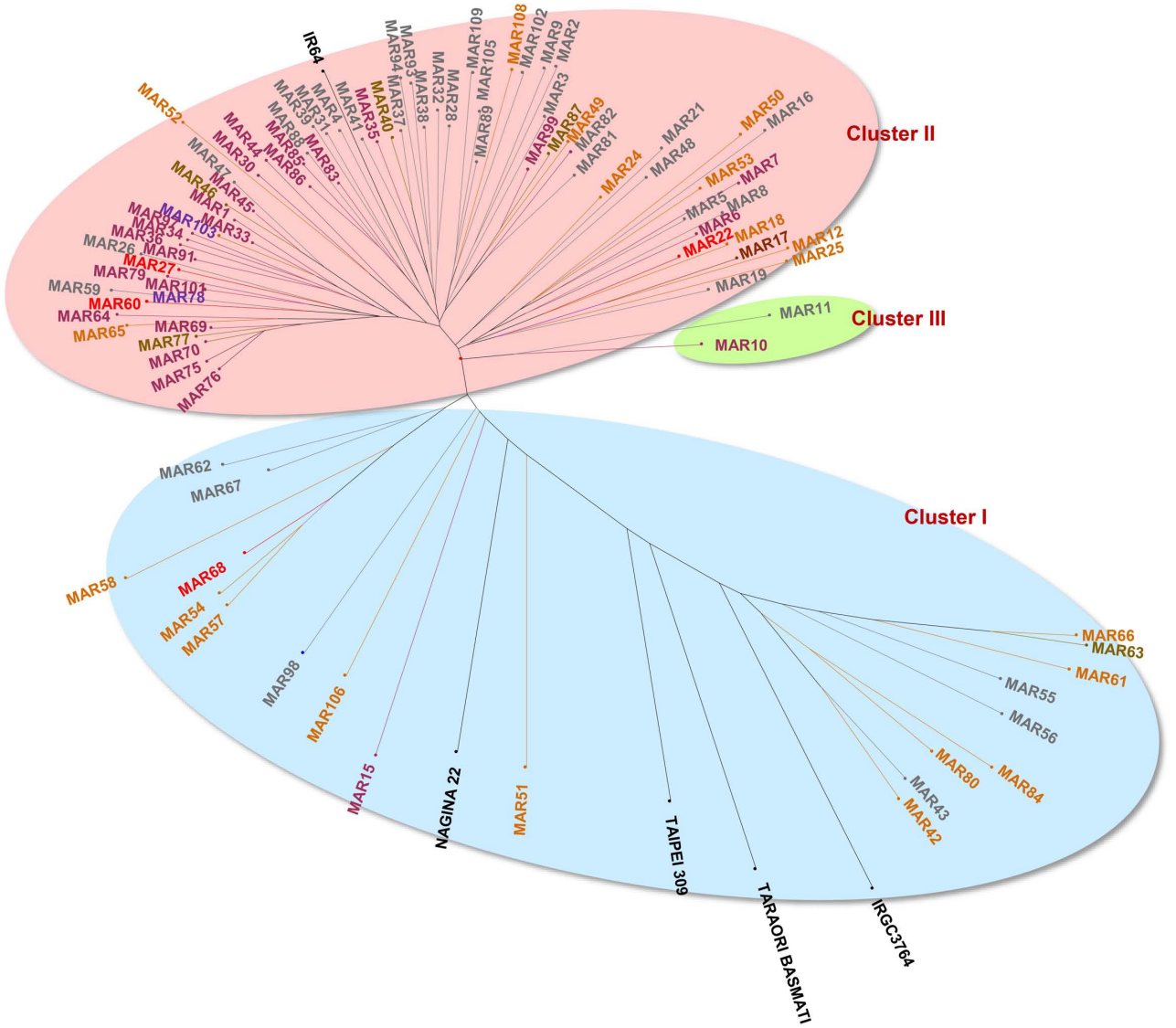


FIGURE 6 | Hierarchical clustering of aromatic rice genotypes from Manipur based on SSR markers by unweighted neighbor joining using simple matching coefficients showing three distinct clusters.

any trait(s) with the landrace names. On the contrary, we could find that landraces sharing the same epithets differed significantly for most of the traits assessed. However, the aromatic genotypes with a deep pericarp color differed significantly for plant height, panicle length, and grain yield per plant. In general, the most widely used epithet was *Chakhao* among Manipur rice. In our study, there were 69 landraces that carried the name *Chakhao*, or its close dialectical variants such as *Chakhao* and *Chahou*. Among these, there were 12 landraces that shared the name *Chakhao Poireiton* or its close resembling name *Chakhao Poireitol*. Twenty-four genotypes, however, were only called as *Chakhao*. Further, there were ten cultivars that carried the suffix *Amubi* and seven that had *Angouba* suffixed. Prominent cultivars such as *Chakhao Poireiton* generally had strongly scented deep purple/black kernel, with tall stature, long panicles, and low grain yield. *Chakhao Angouba* cultivars were white colored with moderate scent, while *Chakhao Amubi* included all categories of pericarp color, from light brown, variegated brown, to deep purple. *Buhman* was the second prominent name among the landraces in the panel, with 16 cultivars carrying this name. While the landraces from valleys mostly carried the prefix "*Chakhao*," the landraces from the hill districts carried the name such as "*Buhman*" (Churachandpur district), "*Maklei*" (Ukhrul district), and "*The Vumnu*" (Chandel district). Majority of the light brown pericarp color genotypes originated from hill districts. Genotypes from hill districts varied significantly from valley genotypes for their plant height and panicle length. Irrespective of the collection ecology and pericarp color, grains of the landraces were of low amylose type, with majority of them having long slender kernels. The traditional

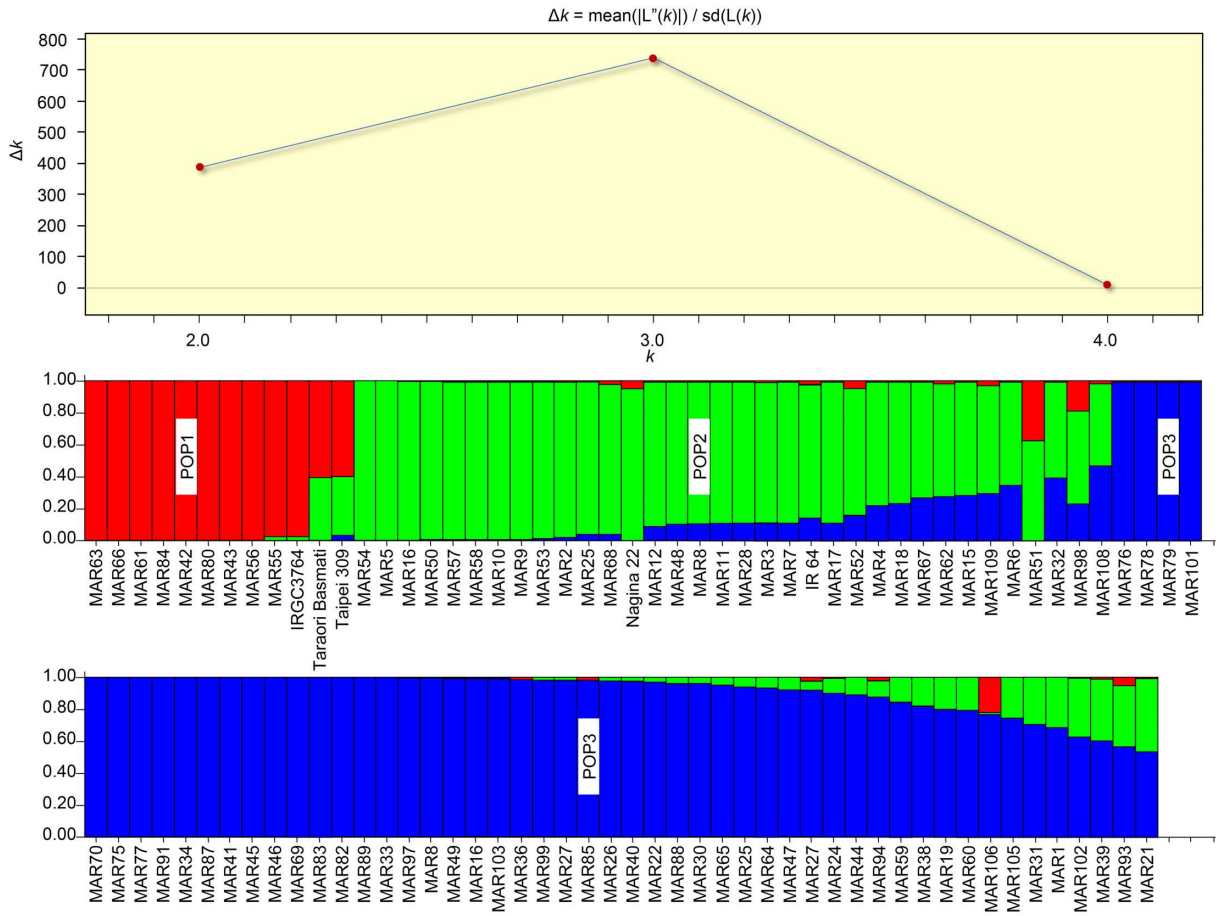


FIGURE 7 | Population structure of aromatic rice germplasm from Manipur. Three subpopulations were resolved by the highest ΔK value (above). Among the subpopulations, POP1 showed more distinctness from other two. POP1 also was the group that did not have genetic admixing. POP2 and POP3 shared a large set of admixtures.

preference of the local community of Manipur is for glutinous rice with soft cooking quality. They are exclusively used in *kheer* preparation called *Chakhao Thongba*. Chanu et al. (2016) also observed low amylose in two major cultivars, namely, *Chakhao Poireiton* and *Chakhao Amubi*. They also found these two cultivars to be nutritionally rich with high dietary fiber, protein, and minerals as compared to a popular rice cultivar, Sona Mashuri, grown in southern India. Another study that evaluated 10 aromatic rice together with three non-aromatic indigenous cultivars of Manipur for grain quality revealed that most of them possessed long bold chalky grains with high ASV and low to intermediate amylose content (Thongbam et al., 2010). Our study revealed that the landraces were low yielding, exhibited high spikelet sterility, were grain shattering, and were photo-sensitive (data not presented). Since most of these landraces are conserved and grown at their respective collection sites, in order to meet the necessities of social rituals rather than for subsistence farming, there has been no serious attempt made for their systematic genetic improvement. Besides, over the decades, there has been a decline in the area of aromatic rice of Manipur with farmers opting not to grow these landraces owing to their very low yield (Borah et al., 2018). If this trend continues, it might lead to genetic erosion of these valuable indigenous landraces. Alternatively, because

TABLE 5 | Analysis of molecular variance (AMOVA) and population differentiation statistics among the subpopulations of Manipur aromatic rice panel.

Source	df	SS	MS	Ve	PVE (%)
Among populations	2	359.5	179.8	2.79	23.0
Among individuals	95	1804.1	19.0	9.49	77.0
Within individuals	98	0.000	0.0	0.00	–
Total	195	2163.6		12.3	100.0
Population statistics	**F_{st}**	**G_{st}**	**G'_{st}**	**G'_{st} (Hed)**	**D_{est}**
POP1–POP2	0.177	0.150	0.261	0.424	0.322
POP1–POP3	0.250	0.227	0.370	0.549	0.417
POP2–POP3	0.069	0.055	0.105	0.121	0.069

df, degrees of freedom; SS, sum of squares; MS, mean square; Ve, estimated variance; PVE, percentage of total variation explained; F_{st}, fixation statistic (Wright, 1951); G_{st}, G statistic (Nei, 1973); G'_{st}, standardized G_{st} (Nei, 1986); G'_{st} (Hed), Hedrick's G_{st} (Hedrick, 2000); D_{est}, Jost's D (Jost, 2008).

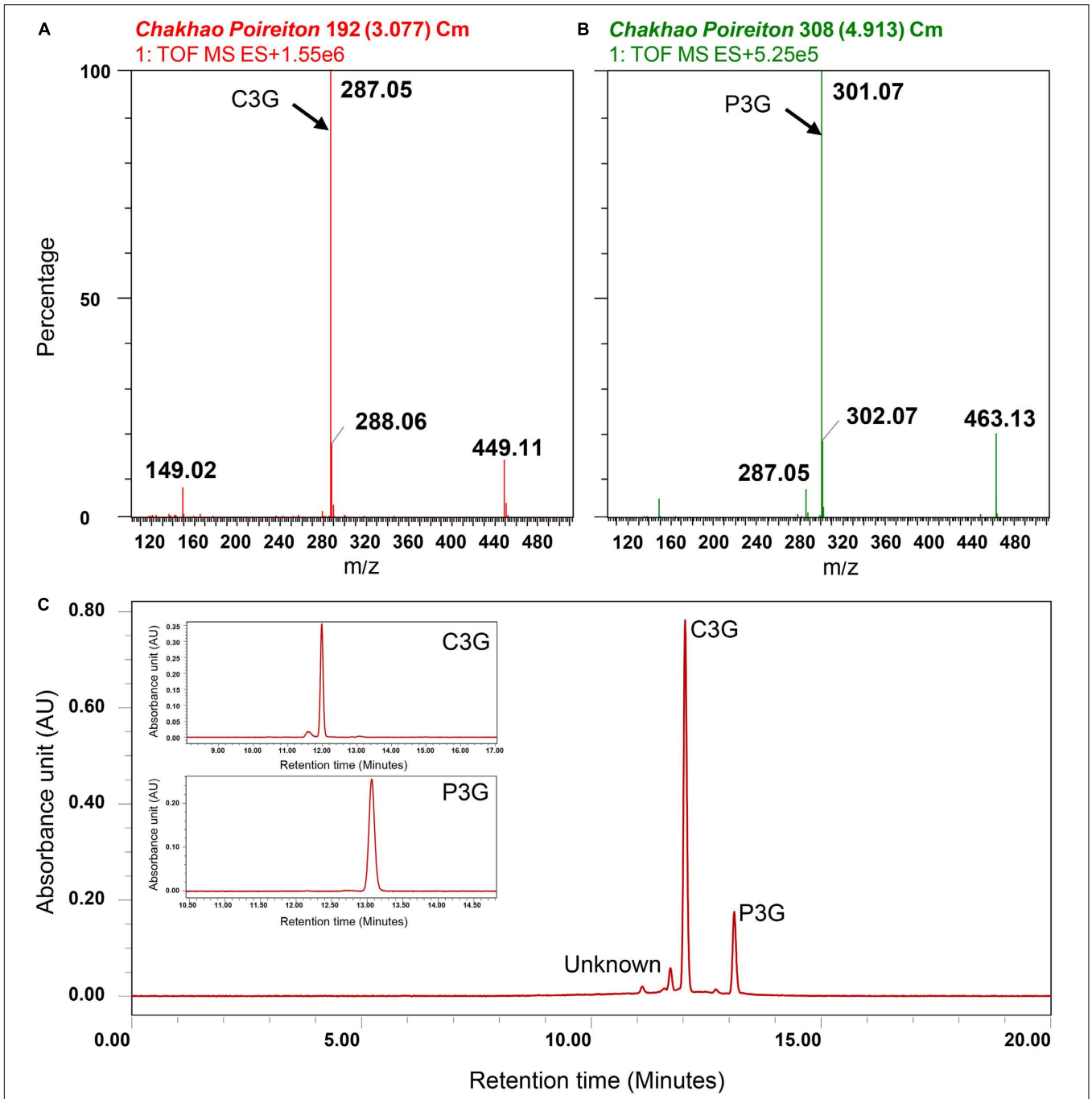


FIGURE 8 | Identification of anthocyanins in Manipur black rice cultivar *Chakhao Poireiton.* MS spectra depicting **(A)** cyanidin-3-o-glucoside (C3G) fragments, **(B)** peonidin-3-o-glucoside (P3G) fragments. Quantification of identified anthocyanins by high-performance liquid chromatography **(C)** Chromatogram of *Chakhao Poireiton*, with standard C3G and P3G peaks as insets.

of the benefits such as nutraceutical properties, pigmentation, glutinous endosperm, and aroma, pigmented rice is gaining popularity among the consumers across the world. Additionally, we found that some of the landraces also possessed desirable traits such as early maturity and relatively higher grain yield, which would provide ample scope for crop improvement. For systematic improvement of Manipur aromatic rice, for both yield and quality, it is desirable to generate a profile of desirable traits within this gene pool, such as kernel pigmentation and antioxidant properties as well as undesirable traits such as photosensitivity, seed shattering, low spikelet fertility, and poor yield.

Analysis of genome-wide variations using molecular markers is one of the means to delineate the evolutionary relationship

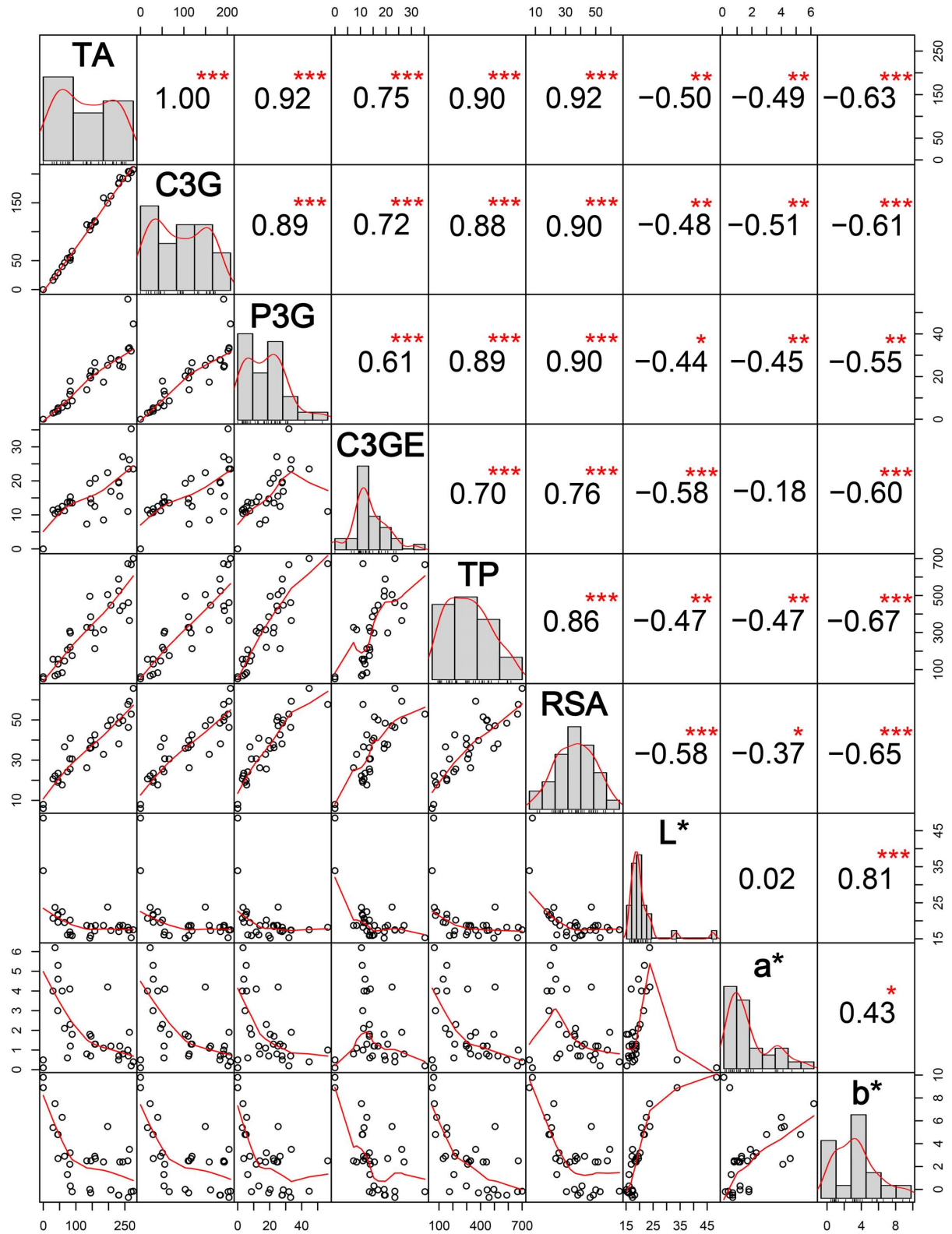


FIGURE 9 | Interrelations among anthocyanin content, nutraceutical properties, and CIELAB color scales among thirty black rice genotypes from Manipur. TA, total anthocyanin content in mg $100g^{-1}$ DW; C3G, cyanidin-3-o-glucoside in mg $100g^{-1}$ DW; P3G, peonidin-3-o-glucoside in mg $100g^{-1}$ DW; C3GE, cyanidin-3-o-glucoside equivalent in mg $100g^{-1}$ DW; TP, total phenolics in mg gallic acid equivalent (GAE) 100 g^{-1} DW; RSA, radical scavenging activity in percentage; L*, a*, and b* are CIELAB color scales. *, **, *** Significant at probability levels of 0.05, 0.01 and 0.001, respectively.

between genotypes that are believed to share a common ancestry. SSRs are highly useful in assessing diversity among closely related rice cultivars (Singh et al., 2013; Yadav et al., 2013) as they provide better resolvability, are multi-allelic, provide genome-wide coverage, are highly reproducible, are easy to score, and are cost-effective (Akagi et al., 1997; Singh and Singh, 2015). There are several studies that used SSR markers for finding the genetic structure of the rice germplasm, as one of the most widely used molecular markers for genetic diversity studies (Nachimuthu et al., 2015; Singh et al., 2016; Aljumaili et al., 2018; Islam et al., 2018a,b; Pathaichindachote et al., 2019; Suvi et al., 2019; Verma et al., 2019). Further, SSR markers are also proven to be efficient in delineating major genetic groups of rice, namely, *indica*, temperate *japonica*, tropical *japonica*, *aus*, and aromatic (Roy et al., 2015, 2016; Wang et al., 2014). In the present study, the panel of 48 SSR markers used could establish the diversity pattern within the Manipur aromatic rice germplasm. Although the gene diversity (0.443) as well as the PIC value (0.394) was lower than those observed from previous studies (Das et al., 2013), this could be attributed to the narrow ecological range from which the lines were sourced, as well as to a greater number of genotypes tested in this study. Notwithstanding, a significant level of genetic variation observed within the lines could be due to their long history of evolution in specific ecologies. An attempt was made in the present study to classify the aromatic rice landraces into rice ecotypes *indica*, *tropical* and *temperate japonica*, and *aus* and their admixtures based on molecular data. The germplasm contained three subpopulations with the checks distributed in the first two subpopulations. As expected, *indica* genotypes (POP2) dominated *japonica* types (POP1) by number and possessed more admixtures. The hilly accessions were more of *japonica* type (POP1) and had very less admixtures. This

TABLE 6 | Variation in total anthocyanins, total phenolics, radical scavenging activity, and color scales in the kernels of thirty pigmented rice landraces in comparison with two white rice genotypes from Manipur.

Code	Cultivar name	TA	C3G	P3G	C3GE	TP	RSA	L*	a*	b*
MAR 21	*Kabo Chakhao*	0.0^{a}	0.0^{a}	0.0^{a}	0.0^{a}	50.6 ± 1.9^{a}	6.0	33.9 ± 2.1^{o}	$0.5 \pm 0.2^{a-d}$	8.9 ± 1.1^{j}
MAR 31	*Chakhao Angouba*	0.0^{a}	0.0^{a}	0.0^{a}	0.0^{a}	61.3 ± 0.8^{b}	8.0	48.5 ± 1.0^{p}	0.14 ± 0.1^{a}	9.8 ± 0.5^{j}
MAR 17	*Chakhao*	29.8 ± 1.9^{b}	16.2 ± 0.3^{b}	2.9 ± 0.1^{b}	11.4 ± 0.3^{cd}	156.3 ± 1.9^{h}	20.7	20.7 ± 0.9^{kl}	4 ± 0.6^{k}	5.4 ± 0.7^{g}
MAR 78	*Black rice*	36.0 ± 1.4^{b}	22.2 ± 1.0^{c}	3.3 ± 0.2^{bc}	10.4 ± 0.2^{c}	66.5 ± 1.5^{c}	22.2	23.7 ± 0.8^{n}	6.2 ± 0.7^{m}	7.5 ± 0.8^{i}
MAR 46	*Chakhao Amubi*	44.4 ± 1.7^{c}	28.9 ± 1.4^{d}	4.6 ± 0.2^{bcd}	10.9 ± 0.4^{cd}	74.6 ± 2.5^{d}	19.0	21.5 ± 0.5^{lm}	3.0 ± 0.8^{j}	4.8 ± 0.9^{g}
MAR 77	*Ethe Buw*	45.2 ± 4.0^{c}	29.1 ± 1.5^{d}	5.3 ± 0.3^{cde}	10.7 ± 0.2^{cd}	156.5 ± 2.5^{h}	19.9	21.8 ± 0.8^{lm}	5.3 ± 0.4^{l}	4.8 ± 0.7^{g}
MAR 83	*Langphou Chakhao*	45.2 ± 3.1^{c}	29.5 ± 1.9^{d}	3.8 ± 1.0^{bc}	11.8 ± 0.3^{cde}	130.4 ± 2.6^{f}	23.5	19.6 ± 0.8^{jk}	4.6 ± 0.7^{k}	2.7 ± 0.6^{def}
MAR 1	*Ching Chakhao Amubi*	57.9 ± 1.3^{d}	39.8 ± 1.0^{e}	5.6 ± 0.2^{cde}	12.5 ± 0.1^{def}	83.4 ± 2.6^{e}	17.7	22.5 ± 0.6^{m}	4.0 ± 0.8^{k}	6.3 ± 0.4^{h}
MAR 15	*Chakhao*	65.2 ± 3.6^{e}	46.3 ± 2.2^{f}	7.6 ± 0.7^{ef}	11.2 ± 1.0^{cd}	146.7 ± 1.7^{g}	36.6	19.5 ± 0.3^{jk}	2.1 ± 0.1^{i}	2.9 ± 0.1^{def}
MAR 76	*The Vumnu*	74.1 ± 0.4^{f}	54.1 ± 0.4^{g}	6.3 ± 0.5^{def}	$13.7 \pm 0.2^{e-h}$	207.4 ± 2.6^{j}	24.2	$16.1 \pm 0.6^{a-e}$	$0.6 \pm 0.1^{a-d}$	1.3 ± 0.6^{c}
MAR 99	*Chakahao Huikap*	81.4 ± 3.4^{g}	55.9 ± 2.4^{g}	11.6 ± 1.2^{g}	13.9 ± 0.4^{fgh}	306.9 ± 4.2^{n}	30.7	$16.9 \pm 0.4^{b-e}$	4.1 ± 0.4^{k}	2.2 ± 0.0^{d}
MAR 40	*Chakhao Amubi*	82.2 ± 3.8^{g}	50.7 ± 4.6^{fg}	17.9 ± 0.8^{h}	13.5 ± 1.7^{efg}	220.7 ± 1.2^{l}	25.6	20.2 ± 0.3^{k}	2.3 ± 0.2^{i}	3.2 ± 0.2^{ef}
MAR 75	*The Vumnu*	82.9 ± 0.3^{g}	54.4 ± 0.7^{g}	13.2 ± 0.4^{g}	15.3 ± 0.1^{ghi}	297.4 ± 2.7^{m}	40.9	$16.1 \pm 0.6^{a-d}$	$1.2 \pm 0.5^{d-h}$	0.3 ± 0.6^{b}
MAR 79	*Buhman*	88.5 ± 2.5^{g}	66.3 ± 2.0^{h}	8.6 ± 0.5^{f}	13.5 ± 0.3^{efg}	175.8 ± 2.5^{i}	30.6	15.9 ± 0.5^{abc}	$1.8 \pm 0.2^{g-i}$	-0.2^{ab}
MAR 34	*Chakhao Amubi*	133.4 ± 3.6^{h}	112.2 ± 2.0^{ijk}	13.8 ± 1.6^{g}	7.3 ± 0.3^{b}	327.4 ± 4.4^{p}	26.2	18.7 ± 0.5^{ij}	$0.8 \pm 0.1^{a-d}$	2.5 ± 0.7^{de}
MAR 101	*Chakhao*	142.7 ± 3.5^{i}	103.2 ± 2.4^{i}	20.3 ± 1.2^{ij}	19.3 ± 0.7^{j}	495.8 ± 7.3^{w}	35.9	15.2 ± 0.7^{a}	$1.8 \pm 0.2^{f-i}$	0.0 ± 0.1^{ab}
MAR 82	*Chakhao Amubi*	145.2 ± 7.6^{i}	111 ± 6.4^{j}	19.4 ± 1.7^{hi}	14.7 ± 0.5^{gh}	384.9 ± 2.7^{r}	39.8	15.9 ± 0.4^{ab}	$0.7 \pm 0.4^{a-d}$	-0.3^{ab}
MAR 45	*Chakhao*	146.3 ± 0.8^{i}	110 ± 3.9^{j}	22.7 ± 2.5^{k}	13.5 ± 0.7^{efg}	326.7 ± 1.9^{p}	36.2	$18.4 \pm 0.4^{g-j}$	$1.7 \pm 0.4^{e-i}$	2.5 ± 0.3^{de}
MAR 7	*Chakhao Poireitol*	157.9 ± 1.3^{j}	118.8 ± 1.3^{l}	26.5 ± 0.1^{lm}	12.5 ± 0.1^{def}	213.6 ± 4.9^{k}	42.6	$17.2 \pm 0.3^{b-g}$	$1.3 \pm 0.1^{c-g}$	2.9 ± 0.3^{def}
MAR 30	*Chakhao Poireiton*	159.7 ± 5.4^{j}	116.6 ± 3.0^{kl}	22.3 ± 0.9^{ijk}	20.7 ± 3.5^{j}	297.6 ± 4.1^{m}	37.7	18.7 ± 0.7^{hij}	$1.3 \pm 0.3^{d-h}$	2.7 ± 0.3^{def}
MAR 35	*Chakhao Poireiton*	184.6 ± 6.0^{k}	158.6 ± 2.2^{n}	17.4 ± 3.2^{h}	8.5 ± 2.3^{b}	315.0 ± 1.6^{o}	32.9	18.7 ± 0.6^{ij}	$1.1 \pm 0.03^{c-g}$	2.4 ± 0.4^{de}
MAR 103	*Chakhao Amubi*	197.3 ± 2.9^{l}	149.5 ± 2.7^{lm}	25.3 ± 0.6^{l}	22.4 ± 1.0^{k}	504.6 ± 3.2^{x}	47.1	23.8 ± 0.6^{n}	4.2 ± 0.6^{k}	5.5 ± 0.3^{gh}
MAR 64	*Maklei*	207.1 ± 3.1^{m}	161.6 ± 2.8^{n}	28.5 ± 0.7^{m}	16.9 ± 0.6^{i}	447.5 ± 2.7^{u}	50.0	$17.3 \pm 0.7^{c-h}$	$1.2 \pm 0.4^{d-g}$	-0.04^{ab}
MAR 70	*Chakhao Poireiton*	230.9 ± 6.9^{n}	183.1 ± 5.0^{o}	27.9 ± 1.8^{m}	19.7 ± 2.0^{j}	589.2 ± 1.6^{z}	48.4	$18.2 \pm 0.5^{f-j}$	$0.5 \pm 0.3^{a-d}$	-0.5^{ab}
MAR 69	*Chakhao Amubi*	231.9 ± 3.6^{n}	184.4 ± 0.5^{o}	27.7 ± 2.5^{m}	19.4 ± 1.2^{j}	525.1 ± 3.4^{y}	38.1	$17.0 \pm 0.8^{b-f}$	$0.7 \pm 0.2^{a-d}$	-0.5^{ab}
MAR 36	*Chakhao Pungdol Angouba*	234.4 ± 9.4^{n}	193.7 ± 7.5^{p}	25.1 ± 1.6^{l}	15.5 ± 0.5^{hi}	418.5 ± 0.9^{s}	51.5	18.7 ± 0.6^{ij}	$1.2 \pm 0.1^{c-g}$	2.4 ± 0.3^{de}
MAR 44	*Chakhao Poireiton*	243.7 ± 7.0^{o}	192 ± 4.5^{p}	24.5 ± 2.6^{kl}	27.1 ± 0.2^{l}	443.0 ± 3.8^{t}	49.7	$18.7 \pm 0.4^{g-j}$	$0.8 \pm 0.1^{a-d}$	2.4 ± 0.3^{de}
MAR 33	*Chakhao Poireiton*	259.2 ± 4.1^{p}	191.8 ± 4.8^{p}	56.3 ± 2.1^{p}	11.0 ± 1.3^{cd}	672.3 ± 2.4^{b}	57.7	$18.2 \pm 0.4^{f-j}$	$1.0 \pm 0.1^{b-e}$	2.5 ± 0.1^{def}
MAR 85	*Chakhao*	260.8 ± 2.2^{p}	204.1 ± 1.0^{q}	33.1 ± 1.4^{n}	23.5 ± 0.5^{k}	461.3 ± 2.7^{v}	46.3	$17.4 \pm 0.5^{d-i}$	$0.7 \pm 0.1^{a-d}$	-0.7^{a}
MAR 6	*Chakhao Poireitol*	264.0 ± 5.2^{pq}	204.2 ± 4.6^{q}	33.5 ± 1.1^{n}	26.2 ± 0.5^{l}	364.5 ± 3.0^{q}	59.3	$17.4 \pm 0.7^{d-i}$	1.9 ± 0.8^{hi}	3.5 ± 0.6^{f}
MAR 86	*Ching Chakhao*	269.4 ± 4.4^{qr}	202 ± 5.3^{q}	32.0 ± 0.2^{n}	35.3 ± 1.1^{m}	667.1 ± 3.2^{a}	52.9	15.3 ± 0.9^{a}	0.2 ± 0.1^{ab}	-0.2^{ab}
MAR 91	*Chakhao*	275.8 ± 5.1^{r}	207.6 ± 3.6^{q}	44.7 ± 1.2^{o}	23.5 ± 0.7^{k}	700.3 ± 2.8^{c}	65.7	$17.5 \pm 0.4^{e-i}$	0.4 ± 0.1^{abc}	-0.16^{ab}

TA, total anthocyanin in mg.100g^{-1} dry weight (DW); C3G, cyanidin-3-O-glucoside in mg.100g^{-1} DW; P3G, peonidin-3-O-glucoside mg.100g^{-1} DW; C3GE, cyanidin-3-glucoside equivalents in mg.100g^{-1} DW; TP, total phenolics in mg gallic acid equivalent (GAE) 100g^{-1}; RSA, radical scavenging activity in percentage; L, a*, b* are CIELAB color scales. Means with different letter(s) indicate significant difference between genotypes at p = 0.05 by Duncan's multiple-range test.*

indicated that even as there was population pressure coming from the majority of *indica* subtypes, the *japonica* types retained their genetic identity among the Manipur aromatic germplasm. Earlier, Roy et al. (2016) reported the predominance of *japonica* and their admixtures among the hill rice of northeast India. POP2 included genotypes such as MAR16, MAR51, MAR54, MAR57, MAR58, MAR68, and MAR98, which possessed most of the rare and unique alleles from both hill and valley regions. Pigmented rice formed a separate large subpopulation (POP3) among the genotypes tested. The origin of black rice of Manipur could be related to introduction of *japonica* rice from China, beginning from second century BC (Tensuba, 1993; Singh and Baghel, 2003; Lalit, 2007). Prominent among the members of POP3 were *Chakhao* landraces that included both pigmented and unpigmented types. Our study also revealed that the pigmented rice showed significant variation in the content of phytochemicals such as anthocyanin and phenolics, implicating that the *Chakhao* landraces had undergone isolated conservation in the local farm holds. As the RSA was found highly correlated with color parameters, these lines also possessed high antioxidant properties in the rice grain.

Earlier, genome analysis of 21 black rice landraces demonstrated that the origin of black rice gene occurred within tropical *japonica*, which later migrated to *indica* and then to temperate *japonica* (Oikawa et al., 2015). Signature of cross subspecific migration of black rice trait could also be seen in the present study, wherein one landrace *Maklei* (dark brown) from the hill district of Ukhrul was grouped along with *japonica* type, whereas the remaining black rice landraces were grouped within the *indica* subgroup cluster. However, this observation needs to be further confirmed by taking more reference entries of respective subgroups and evaluating the diversity with high-density genome-wide markers. Further, reinforcing the theory of cross subspecific migration, based on population structure, we could identify a subpopulation (POP3) with predominance in genotypes with deep purple pericarp as well as for hill and valley adapted genotypes. However, Roy et al. (2014) reported six subgroups such as *Chakhao Poireiton*, *Chakhao Amubi*, *Chakhao, Maklei*, *Buhman*, and *Chakhao Angouba* in 37 *Chakhao* landraces based on 47 random SSR markers, which could be due to a limited number of genotypes used in the study.

Rice pigmentation is attributed to the accumulation of anthocyanins in the pericarp of the grains. Anthocyanins are subgroups of flavonoids which are water-soluble pigments imparting different shades of red, blue, purple, to plant parts. They belong to a class of secondary metabolites of the polyphenol group. In the present study, C3G and P3G were found to be the predominant anthocyanin compounds, while the unidentified fraction was quantified as C3GE compounds. The identification was based on the mass spectral values depicted as m/z ratio. The aglycones, cyanidin and peonidin, have a m/z value of 287.05 and 301.07, respectively (Abdel-Aal et al., 2006; Kim et al., 2008; Lee, 2010), which, on addition of the sugar moiety attachment of glucose (m/z: 162.00), makes the total m/z to 449.1 and 463.1 corresponding to C3G and P3G. In our spectrum, corresponding signals were identified as the prominent peaks confirming the presence of C3G and P3G. The presence of both fractionated and unfractionated residues, served as a reconfirmation of the result. Identification of similar compounds was earlier reported from the black rice cultivars, *Heugjinjubyeo* (Lee, 2010) and *Kilimheugmi* (Ryu et al., 1998) from Korea and black rices of China (Sompong et al., 2011). Hou et al. (2013) identified the constituents of the C3GE fraction from the *japonica* black rice variety, Longjing No. 1 of China, as cyanidin 3,5-diglucoside and cyanidin-3-rutinoside. Among the Japanese black rice, Chen et al. (2012) identified four different anthocyanins such as C3G, P3G, malvidin, and petunidin-3-O-glucoside (Pt-3G). Asem et al. (2015) reported delphinidin 3-galactoside, delphinidin 3-arabinoside, cyanidin 3-galactoside, and C3G from *Chakhao Poireiton*. Further, they had also identified the first three of these compounds in *Chakhao Amubi*. However, the identification of these compounds in their study was based on the comparison of retention time (RT) reported in earlier publications, which is very subjective in nature leading to a possibility of errors due to changes in the solvent system and the instrumentation conditions of the assessment.

The total anthocyanins among the rice landraces in the present study varied widely, imparting different color shades to the kernels. It is interesting to note that genotypes sharing the same epithet(s) as a part of their names like *Chakhao Poireiton*, *Chakhao Amubi*, and *Ching Chakhao* but which were collected from different localities within Manipur significantly varied for total anthocyanin content and its constituents. There was no significant difference in patterns between genotypes from hill districts and valley districts. Although anthocyanin content is reported to be influenced by the environment (Somsana et al., 2013), in the present study, the genotypes were grown uniformly in the experimental plots at a valley region (Imphal), and hence, the variation among the different lines could be purely genetical. Further, pigmentation had no influence on the genotype adaptation to different ecologies. However, any influence of pericarp development, seed coat thickness, grain shape, and weight on the anthocyanin accumulation and color intensity among the pigmented rice of Manipur needs further detailed investigations. Shen et al. (2009) have reported that flavonoid and phenolic contents were positively related to grain shape and negatively related to grain weight. Likewise, we could also observe that the intensity of the pigmentation depended on the proportion and content of pigment compounds, starting from light color to dark. Deeply pigmented rice kernels rather appear black and are popularly traded as black or purple rice. The nutraceutical properties of the pigmented rice and their potential health benefits have been established using cell lines, animal models, and human clinical trials (Pojer et al., 2013; Das et al., 2014; Samyor et al., 2017; Khoo et al., 2017; Seechamnanturakit et al., 2018; Thanuja and Parimalavalli, 2018; Callcott et al., 2019; Limtrakul et al., 2019).

Traditionally, purple or black rice is grown and consumed in many Asian countries like China, Thailand, Sri Lanka, Republic

of Korea, Vietnam, Indonesia, India, Philippines, and Japan. In India, popular black rice varieties include *Chakhao* rice from Manipur and *Kalabhat* from West Bengal. The anthocyanin content of black rice from Korea was as high as 493 $mg.100g^{-1}$, having a large fraction of C3G varying between 80 and 95.3% (Ryu et al., 1998), while Lee (2010) reported C3G content ranging from 52.1 ± 6.3 to 1601.0 ± 8.5 $\mu g.g^{-1}$ and P3G content from 0.0 to 82.6 ± 1.2 $\mu g.g^{-1}$. Anthocyanins being the constituent of polyphenols in plants, the total phenolic content show a significant correlation with anthocyanin content, particularly in pigmented rice (Dai and Mumper, 2010; Deng et al., 2013).

The radical scavenging property of antioxidant compounds provides immense health benefits by scavenging reactive oxygen species (ROS) and other harmful free radicals when oxidative stress occurs (Dröge, 2002). When cellular machinery fails to contain the stress, antioxidants such as anthocyanins can prevent damage either by delaying the oxidative process or by scavenging the excess free radicals. However, antioxidant property differs in different anthocyanins depending upon their molecular structure. The activity increases with the increased number of free hydroxyl group around the pyrone ring (Miguel, 2011). Among the most common anthocyanins, cyanidin is the most active against superoxide after delphinidin. In the present study, C3G was found to be the predominant anthocyanin pigment in *Chakhao* landraces, which could play an important role in scavenging the superoxides. The RSA of the pigmented rice was significantly higher than that of unpigmented genotypes, ranging from more than two to eight times. Chanu et al. (2016) reported an RSA of 72.5% in *Chakhao Poireiton* and 59.0% in *Chakhao Amubi*. Although RSA identified in this study was slightly less than these values, we could observe a maximum activity of 65.7% in our panel. However, the presence of a large variation in free RSA *vis-à-vis* total anthocyanin observed in different deep purple genotypes with similar epithets necessitates defining the baseline values for grain quality and nutraceutical standards aiding improvement and promotion of this speciality rice.

CONCLUSION

In the present study, we elucidated a large variation for the agro-morphological, grain quality, pigmentation, and phytochemical characteristics in the hitherto uncharacterized aromatic rice landraces of Manipur. Although several of them shared common epithets, we found that they are significantly diverse and shared distinct subpopulation memberships based on genetic data. Despite the commonness in names, the average gene diversity of the landraces was 0.443, implying their genetic uniqueness. This necessitates efforts for conservation, documentation, and utilization of this diversity for further improvement. The striking feature of Manipur aromatic rice is the variation for pericarp pigmentation ranging from white to dark purple/black color. Several of the pigmented genotypes are low in amylose with soft cooking quality and high anthocyanin content. Two growing niches in the region *viz*., valley and hill, also show a distinct pattern of genotype characteristics. A detailed investigation with high-throughput markers such as single nucleotide polymorphisms (SNPs) providing high-density genome-wide coverage can help in assessing the genetic relatedness of these genotypes with *indica* and *japonica* subtypes, as well as for mining novel allelic variants that may be present in this gene pool. There is also a need to assess grain quality through metabolic profiling, quantification of aroma, and micronutrient content, which can lay a strong foundation for *in situ* conservation and improvement of these unique landraces from Manipur.

DATA AVAILABILITY STATEMENT

All datasets presented in this study are included in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

AS and SG conceptualized the idea and supervised the experiments. SB, SG, and NS: carried molecular work. SB, HB, and SS carried and monitored quality, phytochemical estimations. RE, KV, and SB carried out the statistical analysis. SB, IS, NP, PB, and MN carried and monitored field experiments. SB, KV, and SG prepared manuscript. All authors read and approved the final manuscript.

FUNDING

This study was supported under the World Bank-Indian Council of Agricultural Research funded National Agricultural Higher Education Project (NAHEP) through its Centre for Advanced Agricultural Science and Technology (CAAST) on 'Genomics Assisted Breeding for Crop Improvement' to ICAR-IARI, New Delhi.

ACKNOWLEDGMENTS

Seed material shared by the National Gene Bank, ICAR-NBPGR, New Delhi, and by the farmers of Manipur for the study is gratefully acknowledged.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2020.570731/full#supplementary-material

REFERENCES

Abdel-Aal, E. M., Young, J. C., and Rabalski, I. (2006). Anthocyanin composition in black, blue, pink, purple, and red cereal grains. *J. Agric. Food Chem.* 54, 4696–4704. doi: 10.1021/jf0606609

Akagi, H., Yokozeki, Y., Inagaki, A., and Fujimura, T. (1997). Highly polymorphic microsatellites of rice consist of AT repeats, and a classification of closely related cultivars with these microsatellite loci. *Theor. Appl. Genet.* 94, 61–67. doi: 10.1007/s001220050382

Ali, M. L., McClung, A. M., Jia, M. H., Kimball, J. A., McCouch, S. R., and Eizenga, G. C. A. (2011). A rice diversity panel evaluated for genetic and agro-morphological diversity between subpopulations and its geographic distribution. *Crop. Sci.* 51, 2021–2035. doi: 10.2135/crosci2010.11.0641

Aljumaili, S. J., Rafii, M. Y., Latif, M. A., Sakimin, S. Z., Arolu, I. W., and Miah, G. (2018). Genetic diversity of aromatic rice germplasm revealed by SSR markers. *Biomed. Res. Int.* 2018:7658032. doi: 10.1155/2018/7658032

Asem, I., Amanda, N., Shaheen, K., Nanita, H., Yurembam, R., Asem, R., et al. (2019). Phenotypic characterization, genetic variability and correlation studies among ten Chakhao (scented) rice of Manipur. *Int. J. Curr. Microbial. Appl. Sci.* 8, 612–618. doi: 10.20546/ijcmas.2019.802.070

Asem, I. D., Imotomba, R. K., and Mazumder, P. B. (2017). "The deep purple color and the scent are two great qualities of the black scented rice (Chakhao) of Manipur," in *Advances in International Rice Research*, ed. J. Q. Li (London: Intech Open), 125–136. doi: 10.5772/67193

Asem, I. D., Imotomba, R. K., Mazumder, P. B., and Laishram, J. M. (2015). Anthocyanin content in the black scented rice (Chakhao): its impact on human health and plant defence. *Symbiosis* 66, 47–54. doi: 10.1007/s13199-015-0329-z

Bhandari, H. R., Bhanu, A. N., Srivastava, K., Singh, M. N., Shreya, and Hemantaranjan, A. (2017). Assessment of genetic diversity in crop plants - an overview. *Adv.. Plants Agric. Res.* 7, 279–286. doi: 10.15406/apar.2017.07.00255

Borah, N., Athokpam, F. D., Semwal, R., and Garkoti, S. C. (2018). Chakhao (Black rice. *Oryza sativa L.*): a culturally important and stress tolerant traditional rice variety of Manipur. *Indian J. Tradit. Know.* 17, 789–794.

Brand-Williams, W., Cuvelier, M. E., and Berset, C. (1995). Use of a free radical method to evaluate antioxidant activity. *Lebenson Wiss. Technol.* 28, 25–30. doi: 10.1016/s0023-6438(95)80008-5

Cagampang, G. B., Perez, C. M., and Juliano, B. O. (1973). A gel consistency test for eating quality of rice. *J. Sci. Food Agric.* 24, 1589–1594. doi: 10.1002/jsfa.2740241214

Callcott, E. T., Blanchard, C. L., Snell, P., and Santhakumar, A. B. (2019). The anti-inflammatory and antioxidant effects of acute consumption of pigmented rice in humans. *Food Funct.* 10, 8230–8239. doi: 10.1039/C9FO0'

Chanu, C. S., Yenagi, N. B., and Math, K. K. (2016). Nutritional and functional evaluation of black rice genotypes. *J. Farm. Sci.* 29, 61–64.

Chen, X. Q., Nagao, N., Itani, T., and Irifune, K. (2012). Anti-oxidative analysis, and identification and quantification of anthocyanin pigments in different coloured rice. *Food Chem.* 135, 2783–2788. doi: 10.1016/j.foodchem.2012.06.098

Dai, J., and Mumper, R. J. (2010). Plant phenolics: extraction, analysis and their antioxidant and anticancer properties. *Molecules* 15, 7313–7352. doi: 10.3390/molecules15107313

Das, B., Sengupta, S., Parida, S. K., Roy, B., Ghosh, M., Prasad, M., et al. (2013). Genetic diversity and population structure of rice landraces from eastern and north eastern states of India. *BMC Genet.* 14:71. doi: 10.1186/1471-2156-14-71

Das, K. R., Medhabati, K., Nongalleima, K. K., and Devi, H. S. (2014). The potential of dark purple scented rice - From staple food to nutraceutical. *Curr. World Environ.* 9:3. doi: 10.12944/CWE.9.3.38

Dayanidhi, H., Singh, N. J., and Maibam, P. (2017). Development and formulation of Chak-hao (Black rice) based noodle. *CAU Res. Newslet.* 8, 14–17.

Deng, G. F., Xu, X. R., Zhang, Y., Li, D., Gan, R. Y., and Li, H. B. (2013). Phenolic compounds and bioactivities of pigmented rice. *Crit. Rev. Food Sci. Nutr.* 53, 296–306. doi: 10.1080/10408398.2010.529624

Dröge, W. (2002). Free radicals in the physiological control of cell function. *Physiol. Rev* 82, 47–95. doi: 10.1152/physrev.00018.2001

Durai, A., Tomar, J. M. S., Devi, P., Devi, P., Arunachalam, A., and Mehta, H. (2015). Rice diversity – The genetic resource grid of North-East India. *Indian J. Plant Genet. Resour.* 28, 205–212. doi: 10.5958/0976-1926.2015.00024.8

Gayacharan, Bisht, I. S., Pandey, A., Yadav, M. C., Singh, A. K., Pandravada, S. R., et al. (2018). Population Structure of some indigenous aromatic rice (*Oryza sativa* L.*) landraces of India. Indian J. Biotechnol.* 17, 110–117.

Glaszmann, J. C. (1987). Isozymes and classification of Asian rice varieties. *Theor. Appl. Genet.* 74, 21–30. doi: 10.1007/BF00290078

GoI (2019). Chak-Hao - GI Application No. 602. *Geogr. Indications J.* 126, 13–23.

GoM (2018). *Economic Survey Manipur 2017-18*. Lamphelpat: Directorate of Economics and Statistics, Government of Manipur.

Goufo, P., and Trindade, H. (2014). Rice antioxidants: phenolic acids, flavonoids, anthocyanins, proanthocyanidins, tocopherols, tocotrienols, γ-oryzol and phytic acid. *Food Sci. Nutr.* 2, 75–104. doi: 10.1002/fsn3.86

Hedrick, P. W. (2000). *Genetics of Populations*, 2nd Edn. Boston: Jones and Bartlett.

Hore, D. K. (2005). Rice diversity collection, conservation and management in northeastern India. *Genet. Resour. Crop Evol.* 52, 1129–1140. doi: 10.1007/s10722-004-6084-2

Hou, Z., Qin, P., Zhang, Y., Cui, S., and Ren, G. (2013). Identification of anthocyanins isolated from black rice (Oryza sativa L.*) and their degradation kinetics. Food Res. Int.* 50, 691–697. doi: 10.1016/j.foodres.2011.07.037

Hunter, R. S., and Harold, R. W. (1987). *The Measurement of Appearance*, 2nd Edn. New York: John Wiley and Sons, Inc.

IRRI (2013). *Standard Evaluation System for Rice*, 5th Edn. Los Baños, PH: International Rice Research Institute.

IRRI (2014). *STAR version 2.0.1. Biometrics and Breeding Informatics*. Los Baños, PH: International Rice Research Institute.

Islam, M. Z., Khalequzzaman, M., Bashar, M. K., Ivy, N. A., Mian, M. A. K., Pittendrigh, B. R., et al. (2018a). Variability assessment of aromatic rice germplasm by pheno-genomic traits and population structure analysis. *Sci. Rep.* 8:9911. doi: 10.1038/s41598-018-28001-z

Islam, M. Z., Khalequzzaman, M., Prince, M. F. R. K., Siddique, M. A., Rashid, E. S. M. H., Ahmed, M. S. U., et al. (2018b). Diversity and population structure of red rice germplasm in Bangladesh. *PLoS One* 13:e0196096. doi: 10.1371/journal.pone.0196096

Jost, L. (2008). GST and its relatives do not measure differentiation. *Mol. Ecol.* 17, 4015–4026. doi: 10.1111/j.1365-294x.2008.03887.x

Juliano, B. O. (1971). A simplified assay for milled rice amylose. *Cereal Sci. Today* 16, 334–340.

Khoo, H. E., Azlan, A., Tang, S. T., and Lim, S. M. (2017). Anthocyanidins and anthocyanins: coloured pigments as food, pharmaceutical ingredients, and the potential health benefits. *Food Nutr. Res.* 61:1361779. doi: 10.1080/16546628.2017.1361779

Khush, G. S. (2000). "Taxonomy and origin of rice," in *Aromatic Rices*, eds R. K. Singh, U. S. Singh, and G. S. Khush (New Delhi: Oxford and IBH Publishing Co. Pvt. Ltd), 5–13.

Kim, M., Kim, H., Koh, K., Kim, H., Lee, Y., and Kim, Y. (2008). Identification and quantification of anthocyanin pigments in colored rice. *Nutr. Res. Pract.* 2, 46–49. doi: 10.4162/nrp.2008.2.1.46

Lalit, P. (2007). *A Brief History (Puwari) of the Meiteis of Manipur*. Available online at: http://themanipurpage.tripod.com/history/puwarimeitei.html. (Accessed April 03, 2020).

Lee, J. H. (2010). Identification and quantification of anthocyanins from the grains of black rice (Oryza sativa L.) *varieties. Food Sci. Biotechnol.* 19, 391–397. doi: 10.1007/s10068-010-0055-5

Limtrakul, P., Semmarath, W., and Mapoung, S. (2019). "Anthocyanins and proanthocyanidins in natural pigmented rice and their bioactivities," in *Phytochemicals in Human Health*, eds V. Rao, D. Mans, and L. Rao (London: IntechOpen), doi: 10.5772/intechopen.86962

Little, R. R., Hilder, G. B., and Dawson, E. H. (1958). Differential effect of dilute alkali on 25 varieties of milled white rice. *Cereal Chem.* 35, 111–126.

Mao, A. A., Hynniewta, T. M., and Sanjappa, M. (2009). Plant wealth of North east India with reference to ethnobotany. *Indian J. Tradit. Know.* 8, 96–103.

Miguel, M. G. (2011). Anthocyanins: antioxidant and/or anti-inflammatory activities. *J. Appl. Pharm. Sci.* 01, 07–15.

Murdifin, M., Pakki, E., Rahim, A., Syaiful, S. A., Ismail, Evary, Y. M., et al. (2015). Physicochemical properties of Indonesian pigmented rice (Oryza sativa Linn.) *varieties from South Sulawesi. Asian J. Plant Sci.* 14, 59–65. doi: 10.3923/ajps.2015.59.65

Murray, M. G., and Thompson, W. F. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8, 4321–4325. doi: 10.1093/nar/8.19.4321

Nachimuthu, V. V., Muthurajan, R., Duraialaguraja, S., Sivakami, R., Pandian, B. A., Ponniah, G., et al. (2015). Analysis of population structure and genetic diversity in rice germplasm using SSR markers: an initiative towards association mapping of agronomic traits in Oryza sativa. *Rice* 8:3030. doi: 10.1186/s12284-015-0062-5

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. (1986). Definition and estimation of fixation indices. *Evolution* 40, 643–645. doi: 10.2307/2408586

Oikawa, T., Maeda, H., Oguchi, T., Yamaguchi, T., Tanabe, N., Ebana, K., et al. (2015). The birth of a black rice gene and its local spread by introgression. *Plant Cell* 27, 2401–2414. doi: 10.1105/tpc.15.00310

Pathaichindachote, W., Panyawut, N., Sikaewtung, K., Patarapuwadol, S., and Muangprom, A. (2019). Genetic diversity and allelic frequency of selected Thai and exotic rice germplasm using SSR markers. *Rice Sci.* 26, 393–403. doi: 10.1016/j.rsci.2018.11.002

Peakall, R., and Smouse, P. E. (2012). GenAlEx 6.5: genetic analysis in Excel. Population genetic software for teaching and research- an update. *Bioinformatics* 28, 2537–2539. doi: 10.1093/bioinformatics/bts460

Pojer, E., Mattivi, F., Johnson, D., and Stockley, C. S. (2013). The case for anthocyanin consumption to promote human health: a review. *Compr. Rev. Food Sci. Food Saf.* 12, 483–508. doi: 10.1111/1541-4337.12024

Prichard, J. K., Wen, X., and Falush, D. (2010). *Documentation for Structure Software: Version 2.3*. Chicago, CL: University of Chicago.

Pritchard, J. K., Stephens, M., and Donnelly, P. J. (2000). Inference of population structure using multilocus genotype data. *Genetics* 155, 945–959.

Rani, N. S., Rao, L. V. S., and Viraktamath, B. C. (2006). *National Guidelines for the Conduct of Tests for Distinctness, Uniformity and Stability: Rice (Oryza sativa L) – Zero Draft*. Hyderabad: Directorate of Rice Research.

Roy, S., Banerjee, A., Mawkhlieng, B., Misra, A. K., Pattanayak, A., Harish, G. D., et al. (2015). Genetic diversity and population structure in aromatic and quality rice (*Oryza sativa* l.) landraces from north-eastern India. *PLoS One* 10:e0129607. doi: 10.1371/journal.pone.0129607

Roy, S., Banerjee, A., Pattanayak, A., Roy, S. S., Rathi, R. S., Misra, A. K., et al. (2014). Chakhao (delicious) rice landraces (*Oryza sativa* L.) of north-east India: collection, conservation, and characterization of genetic diversity. *Plant Genet. Resour.-C.* 12, 264–272. doi: 10.1017/S14179262113000580

Roy, S., Marndi, B. C., Mawkhlieng, B., Banerjee, A., Yadav, R. M., Misra, A. K., et al. (2016). Genetic diversity and structure in hill rice (*Oryza sativa* L.) landraces from the north-eastern Himalayas of India. *BMC Genet.* 17:107. doi: 10.1186/s12863-016-0414-1

Ryu, S. N., Park, S. Z., and Ho, C. T. (1998). High performance liquid chromatographic determination of anthocyanin pigments in some varieties of black rice. *J. Food Drug Anal.* 6, 729–736.

Saikia, S., Dutta, H., Saikia, D., and Mahanta, C. L. (2012). Quality characterisation and estimation of phytochemicals content and antioxidant capacity of aromatic pigmented and nonpigmented rice varieties. *Food Res. Int.* 466, 334–340. doi: 10.1016/j.foodres.2011.12.021

Samyor, D., Das, A. B., and Deka, S. C. (2017). Pigmented rice a potential source of bioactive compounds: a review. *Int. J. Food Sci. Technol.* 52, 1073–1081. doi: 10.1111/ijfs.13378

Seechamnanturakit, V., Karrila, T. T., Sontimuang, C., and Sukhoom, A. (2018). The natural pigments in pigmented rice bran and their relation to human health: a literature review. *Kmutnb Int. J. Appl. Sci. Technol.* 11, 3–13. doi: 10.14416/j.ijast.2018.01.004

Shen, Y., Liang, J., Xiao, P., Yan, L., and Jinsong, B. (2009). Total phenolics, flavonoids, antioxidant capacity in rice grain and their relations to grain color, size and weight. *J. Cereal. Sci.* 49, 106–111. doi: 10.1016/j.jcs.2008.07.010

Singh, B. D., and Singh, A. K. (2015). *Marker-Assisted Plant Breeding: Principles and Practices*. New Delhi: Springer.

Singh, J., Sharma, K., Walia, S., and Saha, S. (2017). Anthocyanin profiling method based on isocratic elution for comparable speed, reproducibility, and quantitation with gradient elution. *Food Anal. Methods* 10, 118–128. doi: 10.1007/s12161-016-0561-z

Singh, N., Choudhury, D. R., Singh, A. K., Kumar, S., Srinivasan, K., Tyagi, R. K., et al. (2013). Comparison of SSR and SNP markers in estimation of genetic diversity and population structure of Indian rice varieties. *PLoS One* 8:e84136. doi: 10.1371/journal.pone.0084136

Singh, N., Choudhury, D. R., Tiwari, G., Singh, A. K., Kumar, S., Srinivasan, K., et al. (2016). Genetic diversity trend in Indian rice varieties: an analysis using SSR markers. *BMC Genet.* 17:127. doi: 10.1186/s12863-016-0437-7

Singh, R. K., and Baghel, S. S. (2003). "Aromatic rices of Manipur," in *A Treatise on the Scented Rices of India*, eds R. K. Singh and U. S. Singh (New Delhi: Kalyani Publishers), 347–354.

Slinkard, S., and Singleton, V. L. (1977). Total phenol analysis: automation and comparison with manual methods. *Am. J. Enol. Viticult.* 28, 49–55.

Sokal, R. R., and Michener, C. D. (1958). A statistical method for evaluating systematic relationships. *Univ. Kans. Sci. Bull.* 28, 1409–1438.

Sompong, R., Siebenhandl-Ehn, S., Linsberger-Martin, G., and Berghofer, E. (2011). Physicochemical and antioxidative properties of red and black rice varieties from Thailand, China and Sri Lanka. *Food Chem.* 124, 132–140. doi: 10.1016/j.foodchem.2010.05.115

Somsana, P., Pattanagul, W., Suriharn, B., and Sanitchon, J. (2013). Stability and genotype by environment interactions for grain anthocyanin content of Thai black glutinous upland rice (Oryza sativa). *SABRAO J. Breed. Genet.* 45, 523–532.

Suvi, W. T., Shimelis, H., Laing, M., Mathew, I., and Shayanowako, A. I. T. (2019). Assessment of the genetic diversity and population structure of rice genotypes using SSR markers. *Acta Agr. Scand. B Soil Plant Sci.*. 70, 76–86. doi: 10.1080/09064710.2019.1670859

Talukdar, P. R., Rathi, S., Pathak, K., Chetia, S. K., and Sarma, R. N. (2017). Population structure and marker-trait association in indigenous aromatic rice. *Rice Sci.* 24, 145–154. doi: 10.1016/j.rsci.2016.08.009

Tensuba, K. C. (1993). *Genesis of Indian tribes: an Approach to the History of Meiteis and Thais*. New Delhi: Inter-India Publications.

Thanuja, B., and Parimalavalli, R. (2018). Role of black rice in health and diseases. *Int. J. Health Sci. Res.* 8, 241–248.

Thongbam, P. D., Durai, A. A., Singh, A., Taorem, B. D., Gupta, S., Mitra, J., et al. (2010). Grain and food quality traits of some indigenous medicinal rice cultivars of Manipur. *India. Int. J. Food Prop.* 13, 1244–1255. doi: 10.1080/10942910903034833

Tulsiram, S. D., Vikash, K., and Bidhan, R. (2018). Grouping of farmers' varieties of rice (*Oryza sativa* L.) collected from West Bengal and adjoining states. *Indian J. Plant Genet. Resour.* 31, 251–259. doi: 10.5958/0976-1926.2018.00029.3

Verma, H., Borah, J. L., and Sarma, R. N. (2019). Variability assessment for root and drought tolerance traits and genetic diversity analysis of rice germplasm using SSR Markers. *Sci. Rep.* 9:16513. doi: 10.1038/s41598-019-52884-1

Wang, C. H., Zheng, X. M., Xu, Q., Yuan, X. P., Huang, L., Zhou, H. F., et al. (2014). Genetic diversity and classification of *Oryza sativa* with emphasis on Chinese rice germplasm. *Heredity* 112, 489–496. doi: 10.1038/hdy.2013.130

Wright, S. (1951). The genetical structure of populations. *Ann. Eugenics* 15, 323–354. doi: 10.1111/j.1469-1809.1949.tb02451.x

Yadav, S., Singh, A., Singh, M. R., Goel, N., Vinod, K. K., Mohapatra, T., et al. (2013). Assessment of genetic diversity in Indian rice germplasm (*Oryza sativa* L.): use of random versus trait-linked microsatellite markers. *J. Genet.* 92, 545–557. doi: 10.1007/s12041-013-0312-5

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.

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