frontiers | Research Topics

Genomics-based breeding of crops for food and nutritional security in 21st century

volume 2

Edited by
Reyazul Rouf Mir, Ajay Kumar, Manish Kumar Pandey and Mahendar Thudi

Published in
Frontiers in Genetics

ISSN 1664-8714
ISBN 978-2-8325-2899-0
DOI 10.3389/978-2-8325-2899-0

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

Genomics-based breeding of crops for food and nutritional security in 21st century - volume 2

Topic editors

Reyazul Rouf Mir — Sher-e-Kashmir University of Agricultural Sciences and Technology, India

Ajay Kumar — North Dakota State University, United States

Manish Kumar Pandey — International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

Mahendar Thudi — Dr. Rajendra Prasad Central Agricultural University, India

Citation

Mir, R. R., Kumar, A., Pandey, M. K., Thudi, M., eds. (2023). *Genomics-based breeding of crops for food and nutritional security in 21st century - volume 2* Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2899-0

Table of contents

frontiers | Frontiers in Genetics

REVIEW
published: 02 June 2022
doi: 10.3389/fgene.2022.914029

Biotechnological Approaches for Host Plant Resistance to Insect Pests

Pritam Kumari[1,2], Poonam Jasrotia[1], Deepak Kumar[2], Prem Lal Kashyap[1], Satish Kumar[1], Chandra Nath Mishra[1], Sudheer Kumar[1] and Gyanendra Pratap Singh[1]*

[1]ICAR-Indian Institute of Wheat and Barley Research, Karnal, India, [2]CCS Haryana Agricultural University, Hisar, India

Annually, the cost of insect pest control in agriculture crosses billions of dollars around the world. Until recently, broad-spectrum synthetic pesticides were considered as the most effective means of pest control in agriculture. However, over the years, the overreliance on pesticides has caused adverse effects on beneficial insects, human health and the environment, and has led to the development of pesticide resistant insects. There is a critical need for the development of alternative pest management strategies aiming for minimum use of pesticides and conservation of natural enemies for maintaining the ecological balance of the environment. Host plant resistance plays a vital role in integrated pest management but the development of insect-resistant varieties through conventional ways of host plant resistance takes time, and is challenging as it involves many quantitative traits positioned at various loci. Biotechnological approaches such as gene editing, gene transformation, marker-assisted selection etc. in this direction have recently opened up a new era of insect control options. These could contribute towards about exploring a much wider array of novel insecticidal genes that would otherwise be beyond the scope of conventional breeding. Biotechnological interventions can alter the gene expression level and pattern as well as the development of transgenic varieties with insecticidal genes and can improve pest management by providing access to novel molecules. This review will discuss the emerging biotechnological tools available to develop insect-resistant engineered crop genotypes with a better ability to resist the attack of insect pests.

Keywords: biotic stress, insect pests, insecticide, biotechnology, resistant variety

OPEN ACCESS

Edited by:
Reyazul Rouf Mir, Sher-e-Kashmir University of Agricultural Sciences and Technology, India

Reviewed by:
Jiban Shrestha, Nepal Agricultural Research Council, Nepal
Jasdeep Chatrath Padaria, Indian Council of Agricultural Research, India
Mohammad Monobrulla, ICAR-Research Complex for Eastern Region, India
Mahadevan Raghuraman, Banaras Hindu University, India

****Correspondence:***
Poonam Jasrotia
poonam.jasrotia@icar.gov.in
poonamjasrotia@gmail.com

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

Received: *06 April 2022*
Accepted: *16 May 2022*
Published: *02 June 2022*

Citation:
Kumari P, Jasrotia P, Kumar D, Kashyap PL, Kumar S, Mishra CN, Kumar S and Singh GP (2022) Biotechnological Approaches for Host Plant Resistance to Insect Pests. Front. Genet. 13:914029. doi: 10.3389/fgene.2022.914029

1 INTRODUCTION

Under the changing climate scenario, the world's population is estimated to increase by 2 billion in the next 30 years, rising from the current 7.7 billion population to 10 billion by 2050 (Zsögön et al., 2022). In this context, there is a continuous need for an increment of food production to fulfill the need of the rising worldwide population. Additionally, it is assessed that total global food demand will increase from 35% in 2010 to 56% in 2050 (van Dijk et al., 2021). To meet these goals it is critical to increase crop yield and reduce pre- and post-harvest losses. Also despite of operating control measures, a large portion of the economically significant harvests experiences a wide range of yield losses. Population extension, depletion of natural resources, environmental change, and developing insect pests are the key limitations that contrarily affect overall agricultural production and productivity (Alemu, 2020). Amongst these production constraints, biotic factors address one of the foremost constraints to crop productivity. Amongst the biotic factors, the insect pests are assessed to cause 25%–30% losses to agricultural production (Joshi et al., 2020; Mateos Fernandez et al., 2021).

TABLE 1 | Transgenic crops carrying *Bt* genes for insect resistance.

S. No.	Target insects	Transgene	Target crop	References
1	*Chilo suppressalis*, *Cnaphalocrocis medinalis*	*cry 1a(b)*	Rice	Fujimoto et al. (1993)
2	*Scirpophaga incertulas* & *Chilo suppressalis*	*cry 1 a(b)*	Rice	Wünn et al. (1996)
3	*Scirpophaga incertulas*, *Cnaphalocrocis medinalis*	*cry 1 a(b)*	Rice	Ghareyazie et al. (1997)
4	*Scirpophaga incertulas*	*cry 1a(c)*	Rice	Nayak et al. (1997)
5	*Scirpophaga incertulas*	*cry 1a(b)/cry1a(c)*	Rice	Tu et al. (2000)
6	*Cnaphalocrocis medinalis*, *Scirpophaga incertulas*	*cry 2a/cry 1a(c)*	Rice	Maqbool et al. (2001)
7	*Helicoverpa armigera*	*cry1Ab + NptII*	Cotton	Khan et al. (2011)
8	*Heliothis* sp.	*cry1Ab*	Cotton	Khan et al. (2013)
9	*Helicoverpa armigera*	*cry2AX*	Cotton	Sakthi et al. (2015)
10	*Helicoverpa armigera*	*cry1AC + cry2Aa*	Pigeon pea	Ghosh et al. (2017)
11	*Helicoverpa armigera*	*cryIIAa*	Chickpea	Sawardekar et al. (2017)
12	*Cnephalocrosis medinalis*	*cry2A*	Rice	Gunasekara et al. (2017)
13	*Tuta absoluta*	*cry1Ac*	Tomato	Selale et al. (2017)
14	*Anthamous grandis*	*cry1Aa*	Cotton	Ribeiro et al. (2017)
15	*Helicoverpa armigera*	*cry2Aa*	Pigeon pea	Baburao and Sumangala (2018)
16	*Helicoverpa armigera*	*cry2Aa*	Pigeon pea	Singh et al. (2018)
17	*Scirpophaga incertulas*, *Cnaphalocrocis medinalis*	*cry2AX1*	Rice	Rajadurai et al. (2018)
18	*Spodoptera litura*	*cry1Aa*	Sweet Potato	Zhong et al. (2019)
19	*Spodoptera litura*	*cry1AC + cry2Ab*	Cotton	Siddiqui et al. (2019)
20	*Holtrichia panallele*	*cry 8 like*	Soyabean	Qin et al. (2019)
21	*Achaea janata, Spodoptera litura*	*cry1AC*	Castor	Muddanuru et al. (2019)
22	*Helicoverpa armigera*	*cry2AX1*	Cotton	Jadhav et al. (2020)
23	*Tuta absoluta*	*cry1Ab*	Tomato	Soliman et al. (2021)

TABLE 2 | Expression of *VIP* genes for insect resistance.

S. No.	Target insects	Transgene	Target crop	References
1	*Heliothis. zea* and *H. virescens*	*Vip3A + cry1Ab*	Cotton	Bommireddy et al. (2011)
2	*Maruca vitrata*	*Vip3Ba1*	Cowpea	Bett et al. (2017)
3	*Helicoverpa armigera*	*Vip3AcAaa (Vip3Aa1+Vip3Ac1)*	Cotton	Chen et al. (2018a)
4	*Chilo infuscatellus*	*Vip3A*	Sugarcane	Riaz et al. (2020)

This suggests that insect pests pose a severe danger to food security and sustainable development, necessitating the development of effective plant protection technologies to prevent and control pest-related crop losses (Oerke et al., 2006). Chemical pesticides provide the first line of defense for farmers against insect pests, and their widespread use has resulted in a number of issues, including the killing of the beneficial insects, environmental pollution (Pedigo and Rice, 2006; Stevens et al., 2012), human and animal health problems and; imparting resistance in pests (Nderitu et al., 2020). To meet these challenges, it is necessary to move towards more sustainable and modern agricultural practices. Moreover, these detrimental non-target effects have motivated researchers around the world to develop novel and environment-friendly, alternative insect pest management strategies. Therefore, host plant resistance can form the backbone of pest management in different agro-ecosystems (Sharma, 2007).

Host plant resistance is the key component of pest management and one of the most appreciated control tactics in advanced agriculture (Horgan et al., 2020; El-Dessouki et al., 2022). It is the consequences of heritable plant characteristics that make a plant to be less damaged than a plant lacking these qualities. Insect-resistant crop varieties reduce the number of insect pests by increasing their tolerance for injury. Three types of resistance determine the relationship between the insect and the plant, e.g. antibiosis, antixenosis (non-preference), and tolerance (Koch et al., 2016; Iqbal et al., 2018) (**Figure 1**). Antibiosis resistance influences the biology of the pest to diminish its population and subsequent damage, resulting in higher mortality or reduced longevity and reproduction of the insect. Antixenosis resistance is defined as non-preference of the pest for a resistant plant and influences the behavioral traits of a pest (Painter, 1951; Smith, 2005). Tolerance is a resistance where a plant can resist or recover from damage caused by the pest population (Smith, 2005).

The development of insect-resistant plants began in 1782 when Havens published an article on a Hessian fly-resistant wheat cultivar. Since that time, several insect-resistant cultivars have been developed by the international and national research centers, the private sector using conventional or biotechnological tools (Jaiswal et al., 2018). The major biochemical principles underlying such resistance and the genes included have been distinguished for their coordinated utilization through biotechnological advancements in the course of the most recent 30 years (Joshi et al., 2020). Furthermore, for global food security and agricultural sustainability, contemporary

TABLE 3 | Expression of lectin genes for insect resistance.

Sr. No.	Target insects	Transgene	Target crop	References
1	Sap sucking pests	Snowdrop lectin (*Galanthus nivalis* agglutinin; GNA)	Rice	Sudhakar et al. (1998)
2	*Sitobion avenae*	Snowdrop lectin (*Galanthus nivalis* agglutinin; GNA)	Wheat	Stoger et al. (1999)
3	*Eoreuma loftini* (Dyar) and *Diatraea saccharalis*	Snowdrop lectin (*Galanthus nivalis* agglutinin; GNA)	Rice, Sugarcane	Sétamou et al. (2002)
4	*Lacanobia oleracea*	GNA-neuropeptide-allatostatin	Tomato	Fitches et al. (2004)
5	*Lacanobia oleracea*	GNA-lepidopteran-specific toxin (ButalT)	Tomato	Trung et al. (2006)
6	*Nilaparvata lugens and Myzus persicae*)	GNA-spider-venom toxin I (SFI1)	Rice and Potato	Down et al. (2006)
7	*Aedes aegypti* eggs and larvae	WSMoL (water-soluble *M. oleifera* lectin)	-	Coelho et al. (2009); Santos et al. (2012), Santos et al. (2020)
8	*Nilaparvata lugens, Sogatella furcifera and Nephotettix nigropictus*	*Allium sativum* leaf agglutinin (ASAL) and *Galanthus nivalis* lectin (GNA)	Rice	Bharathi et al. (2011)
9	*Nasutitermes corniger* workers and soldiers	Endoglucanase, phosphatases, b-glucosidase, and trypsin	-	de Albuquerque et al. (2012)
10	*Sitophilus zeamais*	*M. urundeuva* leaf lectin (MuLL)	Stored grains	Napoleão et al. (2013)
11	*Myzus persicae* and *Sitobion avenae*	Hv1a/GNA	Potato	Nakasu et al. (2014)
12	*Myzus persicae*	*Galanthus nivalis* agglutinin (GNA)	Potato	Mi et al. (2017)
13	*Lipaphis erysimi*	Lentil lectin (LL) and Chickpea protease inhibitor (CPPI) genes	Transgenic *Brassica juncea*	Rani et al. (2017b)
14	*Sitophilus zeamais*	*Schinus terebinthifolius* leaf lectin (SteLL)	Stored grains	de Santana Souza et al. (2018)
15	*Lipaphis erysim*i	*Colocasia esculenta* tuber agglutinin (CEA)+ *Galanthus nivalis* agglutinin (GNA)	Mustard	Das et al. (2018)
16	*Metopolophium dirhodum, Schizaphis graminum, Rhopalosiphum padi,* and *Sitobion avenae*	*Pinellia pedatisecta* agglutinin (PPA)	Wheat	Duan et al. (2018)
17	*Aphis gossypii* and *Spodoptera litura*	Insect gut binding lectin from *Sclerotium rolfsii*	Cotton	Vanti et al. (2018)
18	*Callosobruchus chinensis*	Arcl on APA locus from *Phaeselous vulgaris*	Cowpea	Grazziotin et al. (2020)
19	*Sitophilus. zeamais*	*Microgramma vacciniifolia* rhizome lectin (MvRL)	Stored grains	de Albuquerque et al. (2020)
20	*Sitophilus zeamais*	Water-soluble *Moringa oleifera* lectin (WSMoL)	Stored grains	de Oliveira et al. (2020)
21	*Callosobruchus chinensis*	Arcelin	Common bean	Hilda et al. (2022)
22	*Callosobruchus chinensis*	Arcelin-5, Leucoagglutinin, Erythroagglutinin	common bean	Caroline et al. (2022)
23	*Sitophilus oryzae*	*Polygonum persicaria* L. (PPA) Lectin	Stored grains	Khoobdel et al (2022)

TABLE 4 | Fusion proteins for insect resistance in crop plants.

Sr. No.	Target insects	Transgene	Target crop	References
1	*Scirpophaga incertulas*, *Cnaphalocrocis medinalis*	*cry2AX1 (cry2Aa + cry2Ac)*	Rice	Chakraborty et al. (2016)
2	*Lygus* spp.	*cry51Aa2*	Cotton	Gowda et al. (2016)
3	*Spodoptera exigua, Harmonia axyridis*	*cry1Ab/cry2Aj*	Maize	Chang et al. (2017)
4	*Spodoptera litura, Ostrinia nubialis*	*cry1Be + cry1Fa*	Cotton	Meade et al. (2017)
5	*Lipaphis erysimi*	Lentil lectin (LL) and chickpea protease inhibitor (CPPI) genes	*Brassica juncea*-mustard	Rani et al. (2017b)
6	*Scirpophaga incertulas*, *Cnaphalocrocis medinalis, Nilaparvata lugens*	*cry1AC + ASAL*	Rice	Boddupally et al. (2018)
7	*Ostrinia furnacalis, Cnaphalocrocis medinalis*	*cry1Ab + vip3A*	Rice	Xu et al. (2018)
8	*Chilo suppressalis*	*cry2Aa + cry1Ca*	Rice	Qiu et al. (2019)
9	*Helicoverpa armigera, Spodoptera litura*	*cry2Ab + cry1F + cry1AC*	Cotton	Katta et al. (2020)
10	*Scirphophaga excerptalis*	*cry2Aa + cry1Ca, cry1Ab + cry1Ac*	Sugarcane	Koerniati et al. (2020)

agriculture's primary goal is to enhance yields using existing land and resources. Therefore, innovative technologies have to be exploited to control pests and ensure adequate food availability in the future. Different plant protection innovations have been created to control, prevent and manage these pests with the trend of emphasizing/concentrating on the use of newer and more advance biotechnological approaches that are proven to be most efficient and provide results in a very short time as compared to conventional methods. These approaches can act as the backbone of crop protection against a broad spectrum of insect pests. In this due consideration, modern biotechnology put forward the best possible alternatives for

TABLE 5 | Insect engineered for pest management using CRISPR/Cas9.

Sr. No.	Target insects	Target gene	References
1	*Tribolium castaneum*	E-cadherin gene, *EGFP*	Gilles et al. (2015)
2	*Plutella xylostella*	Abdominal-A homeotic gene (*Pxabd-A*)	Huang et al. (2016)
3	*Drosophila melanogaster*	Chitin synthase 1	Douris et al. (2016)
4	*Agrotis ipsilo*	*Yellow-Y* Gene	Chen et al. (2016)
5	*Locusta migratoria*	Odorant receptor co-receptor (*Orco*) gene	Li et al. (2016)
6	*Spodoptera litura*	Abdominal-A (*Slabd-A*) gene	Bi et al. (2016)
7	*Spodoptera littoralis*	Olfactory receptor co-receptor (*Orco*) gene	Koutroumpa et al. (2016)
8	*Helicoverpa armigera*	*HaCad*	Wang et al. (2016)
9	*Spodoptera exigua*	Ryanodine receptor	Zuo et al. (2017)
10	*Ceratitis capitata*	Eye Pigmentation Gene White Eye (*We*)	Meccariello et al. (2017)
11	*Helicoverpa armigera*	*Tetraspainin*	Jin et al. (2018)
12	*Plutella xylostella*	*PxABCC2, PxABCC3*	Guo et al. (2019)
13	*Helicoverpa armigera*	α-6- nicotinic acetylcholine receptor (*nAchR*)	Zuo et al. (2020)
14	*Rhopalosiphum padi*	*ß-1-3glucanase* in maize	Kim et al. (2020)
15	*Ostrinia furnacalis*	*ABCC2*	Wang et al. (2020a), Wang et al. (2020b)

TABLE 6 | Transgenic crops for insect resistance through RNA interference.

Sr. No.	Target insects	Silenced gene	Target crop	References
1	*Diabrotica virgifera virgifera* LeConte	Suppression of target mRNA	Maize	Baum et al. (2007)
2	*Diabrotica v. virgifera*	hunchback (hb) and brahma (brm) gene	Maize	Khajuria et al. (2015)
3	*Leptinotarsa decemlineata*	*β-actin* gene	Potato	Zhang et al. (2015)
4	Lepidopteran	dsRNA-Spray	Maize	Li et al. (2015)
5	*H. armigera*	Chitinase gene-*HaCHI*	Tomato, Tobacco	Mamta et al. (2016)
6	*C. suppressalis*	Aminopeptidase N genes *APN1+APN2*	Rice	Qiu et al. (2017)
7	*Leguminivora glycinivorella*	SpbP0-dsRNA	Soyabean	Meng et al. (2017)
8	*Helicoverpa armigera*	Juvenile hormone methyl transferase (JHMT)	Cotton	Ni et al. (2017)
9	*Diabrotica virgifera virgifera* LeConte	*Dvvgr, dvbol*	Maize	Niu et al. (2017)
10	*Leptinotarsa decemlineata*	*ECR* gene	Potato	Hussain et al. (2019)
11	*Scirpophaga incertulas*	*AchE*-Acetylcholine esterase	Rice	Kola et al. (2019)
12	*Manduca sexta*	v-*ATPaseA* gene	Tobacco	Burke et al. (2019)
13	*Bemisia tabaci*	*BtACTB gene*	Tobacco	Dong et al. (2020)
14	*Aphis glycines*	*TREH, ATPD, ATPE, CHSI*	Soyabean	Yan et al. (2020)
15	*Bemisia tabaci*	Phenolic glucoside malonyltransferase	Tobacco	Xia et al. (2021)
16	*Spodoptera littoralis*	*SI 102 immune* gene	Tobacco	Di Lelio et al. (2022)

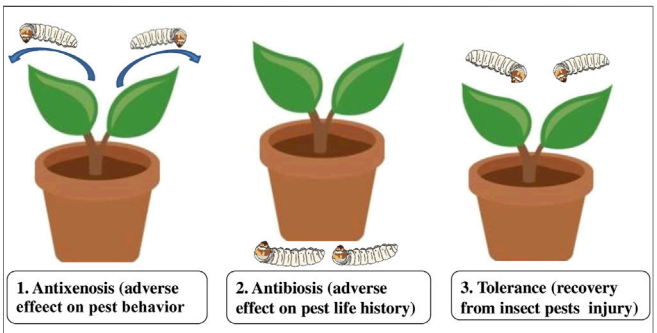


FIGURE 1 | Mechanism of host plant resistance in insects.

diversifying agricultural production by accelerating the development of new insect resistant varieties of cereals, horticultural and even underutilized crops (Abah et al., 2010). Therefore, the main objective of this review is to assess the opportunities provided by these new biotechnological tools in developing various crop plants that are resistant to a wide range of insect pests.

2 BIOTECHNOLOGICAL APPROACHES IN INSECT PEST MANAGEMENT

Biotechnology can be broadly defined as "a method of creating or modifying a product, improving plants or animals, or developing microorganisms for specific purposes by utilising biological systems, living organisms, or derivatives thereof" (Persley, 2000). However, it can be described as the regulated and deliberate manipulation of biological processes to accomplish efficient insect pest control. From insect resistance breeding to transgenic introgression of novel genes, biotechnological interventions in insect pest management to protect crop yield have been enormous. The different biotechnological approaches include gene transformation, genome editing, RNA interference, marker-assisted selection, anther culture, embryo culture,

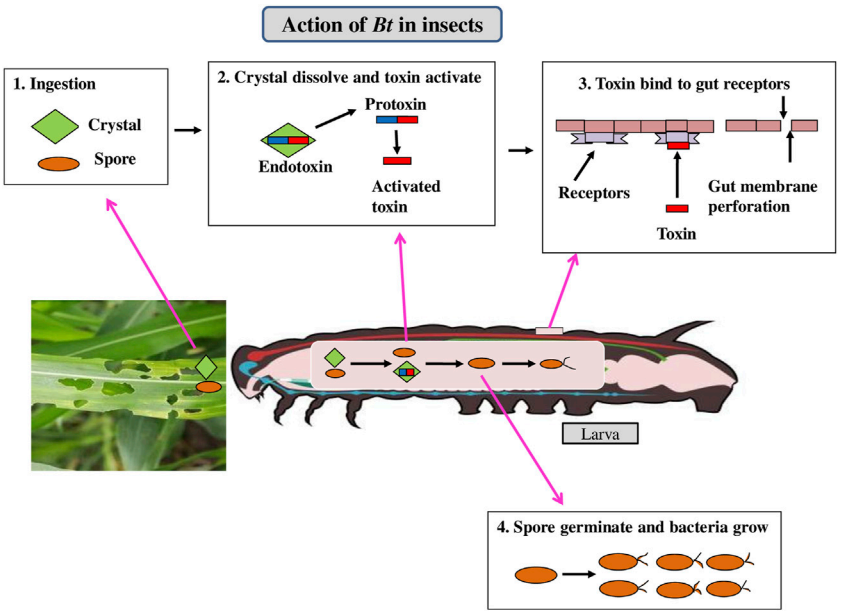


FIGURE 2 | Mechanism of action of *Bt* cry toxin in insects.

protoplast fusion, somaclonal variations etc. (Talakayala et al., 2020). These approaches are described in detail below.

2.1 Gene Transformation

Gene transformation or genetic engineering of crops for insect resistance involves incorporation of specific DNA segment or gene into crop plants to provide resistance against insect pests. The DNA segment which get introduced usually encode a protein with insecticidal activity. Resistance is plant is conferred against specific insect pest through the expression of an insecticidal protein present in the introduced DNA segment (Gatehouse, 2013). This technology has been tested against a wide range of insect pests belonging to orders; Lepidoptera, Coleoptera, and Diptera (Birkett and Pickett, 2014). Genetically modified crops producing insecticidal proteins from *Bacillus thuringiensis*, a soil bacterium, have been widely used in agriculture globally since their introduction in 1996 (Abbas, 2018). The *cry* gene transformation technology involves the transfer of specified DNA sequences or genes into crop plants via *Agrobacterium*-mediated transformation or particle bombardment (Juturu et al., 2015). Apart from this, other strategies to protect plant from insects attack have also been investigated. Lectins, usually found in a number of plants, bind to carbohydrates in the midguts of phytophagous insects disrupting the digestive system (Vandenborre et al., 2011). Transgenic techniques have also been used to deploy protease inhibitors, which are designed to prevent insects from digesting their food (Singh et al., 2020). Similarly, transgenic plants expressing alpha-amylase inhibitors have been produced which are resistant to Lepidopterans, Coleoptrans, Dipterans and Hemipterans insects. Also, insect chitinase and chitinase-like proteins play a significant role in degrading chitin in the exoskeletal and gut linings of insects. These have been successfully cloned into plants and show insecticidal properties. Each of these strategies along with their role in insect pest management is discussed in detail below:

2.1.1 Cry Genes

Bacillus thuringiensis (*Bt*) is a Gram-positive soil bacterium expressing insecticidal crystalline proteins (ICPs) that are exceptionally toxic to specific classes of pests (Panwar et al., 2018). Insecticidal activity in insect-resistant *Bt* crops is expressed by the genes coding for parasporal crystal protoxins (Palma et al., 2014). ICPs produced by transgenic plants have had a significant impact on the successful evolution of insect resistance. The crystal involves a protoxin protein which get solubilized in the larval midgut due to alkaline pH and subsequently cleaved enzymatically to an active toxin. The toxin diffuses through the peritrophic membrane covering the gut and binds to receptors present in the midgut epithelium (Paul and Das, 2020) making pores in the midgut epithelium. The gut gets paralyzed and then the pest stops feeding and dies within 2–3 days (**Figure 2**).

The first-generation *Bt* cotton, Bollgard I (BG I) expressing*cry1Ac* was commercialized and released in 2002 to control the dominant bollworms including *Pectinophora gossypiella, Earias vittella,* and *Helicoverpa armigera* in cotton-cultivating areas of India. After thatBollgard II (BG II) was launched in 2006 as a second-generation *Bt* cotton with pyramided traits expressing *cry1Ac* and *cry2Ab* (MON15985 event), which is now cultivated in 95% of the total India's cotton sowing area. In comparison to BG I containing *cry 1Ac* alone, BG II having multiple toxins, *cry1Ac* and *cry2Ab* have greater ability for pest management (Carrière et al., 2015). Another transgenic cotton *i.e.,* wide strike cotton expressing

cry1Ac + *cry1F* was approved in the USA in the year 2004 by Dow Agro Sciences which improved both the crop yield as well as farmer's income. It is observed that both BG II and wide-strike cotton expressing pyramided traits have a higher potential to suppress a wide range of Lepidopteran, Coleopteran and Dipteran insects than BG I. Another transgenic cotton containing *cry10Aa* exhibited strong resistance to cotton boll weevil (*Anthonomous grandis*) with 100% mortality observed through bioassay testing when the larvae of T_1 generation consumed the leaves of transgenic cotton (Ribeiro et al., 2017). *Agrobacterium*-mediated transformation was employed to produce transgenic cotton expressing pyramided traits, *cry1Ac* and *cry2Ab* cloned in the T-DNA, and the resulting plant demonstrated resistance to *S. litura* with 93 percent larval mortality (Siddiqui et al., 2019). Transgenic rice lines constructed through the expression of the *cry2AX1* gene showed resistance to rice leaf folder (*C. medinalis*) and rice yellow stem borer (*S. incertulas*) (Rajadurai et al., 2018).

Transgenic cotton and brinjal cultivars having resistance to borers were permitted for commercial usage in Bangladesh and in Latin America, insect-resistant Bt soybean expressing cry1Ac + cry1Ab were allowed for production during 2014 (Koch et al., 2015). Another study found that a synthetic *cry1Ab* gene introduced into tomato conferred resistance to the tomato leaf miner, *T. absoluta*, with 100 percent insect mortality at T_0 generation within 4–5 days (Soliman et al., 2021). Rice lines (var. Bg94–1) produced by transferring the insecticidal protein *cry2A* cause mortality in rice leaf folders in 80% of cases (Gunasekara et al., 2017). Similarly, transgenic pigeon pea lines constructed using a combination of *cry1Ac* and *cry2Aa* exhibited resistance to *H. armigera*, resulting in 80%–100% larval mortality (Ghosh et al., 2017). *Cry1Aa* gene expression in sweet potatoes conferred resistance to Lepidopteran insect *i. e, Spodoptera litura* (Zhong et al., 2019). Expression of *cry2AX* gene in transgenic cotton event CH12 showed 88% mortality in *H. armigera* at T_0 generation (Sakthi et al., 2015). et alAn industrially important non-edible castor developed by transferring the *cry1Aa gene* using *Agrobacterium* transformation technique, exhibited strong resistance against two lepidopteran pests, *i.e.*, *Achaea janata* (semi-looper) and *S. litura* (Muddanuru et al., 2019). Transgenic soybean, expressing *cry8*-like gene from *B. thuringenesis* conferred resistance to *Holotrichia parallela*, a Coleopteran pest (Qin et al., 2019). The transgenic cotton event MNH93 carrying *cry1Ab* demonstrated 40–60 percent larval mortality against *H. armigera* with displaying 0.26 percent transformation frequency (Khan et al., 2011). In *H. armigera*, expression of the *cry2AXI* gene in the T_3 generation of the cotton event CH12 resulted in 90% death (Jadhav et al., 2020). Other toxins or proteins, such as *cyt2Aa*, which provides aphid resistance (Chougule et al., 2013) and *cry51Aa2*, increased *Lygus* species mortality in cotton (Gowda et al., 2016). The expression of *cry* gene has been studied in different crop species and is depicted in **Table 1**. Despite the successful deployment of cry gene technologies in crops to achieve resistance against several insect pests, some of the agricultural pests often develop resistance to insecticidal toxins and devastate the crop production. Other problems that limit the usefulness of transgenic crops for insect control include secondary pest outbreak, evolution of new biotypes, effects on non-target organisms, environmental influences on gene expression, biosafety of food from transgenic crops, and socio-economic and ethical issues. Also the research groups should take up the challenges of understanding plant insect interactions to understand the mechanism of resistance development in insects against *cry* genes.

2.1.2 Vegetative Insecticidal Proteins Genes

The *Bacillus thuringiensis* (*Bt*) bacteria found in a variety of ecological habitats, has natural entomo-pesticidal properties against a variety of economically important crop pests due to the secretion of various proteins during different growth phases. One of the best known families of *Bt* proteins is Vip, which are produced during the vegetative growth phase of the plant and are recognized as an excellent toxic candidate due to the sequence homology and receptor sites differences from *cry* proteins. There are three subfamilies of Vip proteins. Vip1 and Vip2 heterodimer toxins, effective against pests belonging to Hemiptera and Coleoptera orders, whereas Vip3, the most extensively studied family of Vip toxins have toxicity toward Lepidopterans et al. Vip proteins are also known as second-generation insecticidal proteins, that can be used either alone or in combination with *cry* proteins to control a number of insect-pests (Gupta et al., 2021). In terms of toxicity potential against susceptible insects, these Vip3 proteins are comparable to *cry* proteins. They are reported to be toxic toward pests, which can't be controlled with *cry* proteins. They reduce the insect pest's population by osmotic lysis which causes swelling and interruption of the midgut epithelial cells. The Vip3 proteins have been successfully pyramided along with *cry* proteins in transgenic crops such as maize and cotton, to overcome resistant pest populations and delay the evolution of resistance (Syed et al., 2020).

Vip genes exhibiting greater resistance against cotton bollworm (*H. armigera*) and tobacco budworm (*Heliothis virescens*) showing that these genes are an excellent option for these insect control. The transgenic cotton containing Vip3A alone and another multitoxin line expressing Vip3A and *cry1Ab* (VipCot) had greater resistance to both the insects, *H. zea* and *H. virescens* (Bommireddy et al., 2011). Transgenic cotton lines expressing Vip3AcAaa demonstrated greater insect resistance, showing that the Vip3AcAaa protein is highly effective in insect control (Chen et al., 2018a). The cowpea lines containing Vip3Ba1 exhibited greater suppression of larval growth and showed further resistance against the pod borer (Bett et al., 2017). A toxin Vip3A transferred in sugarcane showed superior resistance against sugarcane stem borer (*Chilo infuscatellus*) with 100% mortality (Riaz et al., 2020). The effect of the Vip gene in different crop species is depicted in **Table 2**. These proteins are promising candidates for further development of insect resistant plants and show extended ranges of toxicity particularly toward lepidopteran pests. Efforts are underway to use these proteins to induce insect pest resistance.

2.1.3 Lectins

Carbohydrate-binding proteins known as lectins, are entomotoxic proteins with insecticidal properties and are found in many plant species. They prevent predation by being

detrimental to a variety of insects and animals that eat plants. They are mostly found in plants belonging to the families Solanaceae, Fabaceae and Poaceae; especially some of leguminous seeds contain a high concentration of lectins. Plant lectins act as storage proteins and are involved in defense mechanisms against phytophagus insects. Various plants lectins from different sources have already been reported to be toxic towards important members of insects belonging to Lepidoptera (Czapla and Lang, 1990), Coleoptera (Gatehouse et al., 1984; Czapla and Lang, 1990) and Homoptera orders (Powell et al., 1993; Sauvion et al., 1996). The first lectin discovered and commercially available was Concanavalin A; which is now the most extensively studied lectin for insect pest control. The adverse impact of lectins on biological parameters of insects includes, feeding inhibition, reduction in larval weight delays in adult emergence, retardation in total developmental period and increased mortality and reduced fecundity in the first and second generation (Powell et al., 1993).

Insect-resistant plants have emerged in recent years, paving the way for the use of plant lectins in pest management strategies. The use of lectins in transgenic plants has yielded promising results, particularly for crops expressing *Bacillus thuringiensis* (*Bt*) Cry toxins, which show resistance to sap-sucking insects. Furthermore, lectins in artificial diets and their expression in transgenic plants have been shown to reduce performance in insects of various orders, including Lepidoptera, Coleoptera, and Hemiptera.

Plant lectins are carbohydrate-binding proteins that have greater affinity for certain sugar components found in glycoproteins and glycolipids in the cell membrane (Camaroti et al., 2017). Transgenic rice carrying *Allium sativum* leaf agglutinin (ASAL) and *Galanthus nivalis* lectin (GNA) conferred resistance to major sap-sucking insects like brown planthoppers (BPH), white-backed planthoppers (WBPH), and green leafhoppers (GLH) (Bharathi et al., 2011). Transgenic rice (*Oryza sativa* L.) plants expressing an insecticidal protein (the snowdrop lectin, GNA) produced through particle bombardment exhibited the significant levels of resistance against sap-sucking pests (Sudhakar et al., 1998). Enhanced toxicity of GNA-spider-venom toxin I (SFI1) fusion protein to larvae of the tomato moth (*L. oleracea*), rice brown planthopper (*N. lugens*), and the peach potato aphid (*M. persicae*) reported by Fitches et al. (2004) and Down et al. (2006). In an another study, the harmful effects of snowdrop lectin (GNA) expressed in transgenic sugarcane on the life cycle of Mexican rice borer *Eoreuma loftini* (Dyar) and sugarcane borer *Diatraea saccharalis* (F.) was reported by Sétamou et al. (2002). Expression of *Galanthus nivalis* lectin (GNA) gene showed resistance to aphids in potato (Mi et al., 2017). Transgenic *Arabidopsis* expressing recombinant fusion protein (Hv1a/GNA) exhibited resistance to peach potato aphids and grain aphids (Nakasu et al., 2014). A GNA-neuropeptide-allatostatin fusion protein was found to inhibit the feeding and growth of the tomato moth (*L. oleracea*). (Fitches et al., 2004). The larvae of the tomato moth were found to be more toxic to a fusion protein containing a GNA-lepidopteran-specific toxin (ButalT) from the South Indian red scorpion (*Mesobuthus tamulus*). (Trung et al., 2006). The transgenic cotton lines containing insect gut binding lectin demonstrated significant level of resistance to sucking and chewing insects at T_1 generation (Vanti et al., 2018). The lentil lectin (LL) and chickpea protease inhibitor (CPPI) genes transferred to *Brassica juncea* lines and the resulting transgenic plants showed resistance to sap-sucking pests such as aphids (Rani et al., 2017b). Mannose-binding lectin expressed through *Agrobacterium*-mediated transformation, exhibited resistance to wheat aphid in BE104 (Duan et al., 2018). In a bioassay study it was found that the transgenic wheat plants produced through the particle bombardment method expressing the gene encoding snowdrop lectin (*Galanthus nivalis* agglutinin; GNA) showed reduced fecundity of grain aphids *Sitobion avenae* (Stoger et al., 1999). At T_1 and T_2 generations, transgenic *B. juncea* expressed a new lectin gene, *Colocasia esculenta* tuber agglutinin (CEA), as well as a GNA, exhibited enhanced resistance against mustard aphid (*Lipaphis erysimi*) (Das et al., 2018). Transgenic cowpea expressing *arcelin 1* gene from *Phaseolus vulgaris* L. conferred greater resistance to bruchids like *Zabrotes subfasciatus* and *Callosobruchus maculatus*. In addition, against both bruchid insects, the percentage of eggs laid, hatching, adult emergence, and grain mass loss was much lower in transgenic cowpea than in control (Grazziotin et al., 2020). Hilda et al. (2022) demonstrated that arcelin found in the wild accession of common bean is an insecticidal protein that can prevent the bruchid beetle (Bruchidae: Coleoptera) from digesting it. Aside from focusing on coleopteran insects, arcelin was reported to be highly effective against specific insects that belong to Lepidoptera and Hemiptera insect order (Oriani and Lara, 2000; Malaikozhundan et al., 2003). The binding of lectin molecules to glycosylated proteins in the midgut of larvae reduces the efficiency of nutrient uptake and diet utilisation, resulting in a drop in total larval mass and decline in the average survival rate. (Paiva et al., 2012). Caroline et al. (2022) reported that Erythroagglutinin, Arcelin-5, Leucoagglutinin, and a hypothetical seed storage protein are responsible for bruchid resistance which is among the most devastating insect pest of the common bean. Findings of Khoobdel et al. (2022) indicated that a lectin derived from *Polygonum persicaria* L. (PPA), causes oxidative stress in *Sitophilus oryzae* in addition to causing digestive disorders,. The leaf lectin (SteLL) from *Schinus terebinthifolius* did not cause death in *S. zeamais* adults, according to Camaroti et al. (2018) and de Santana Souza et al. (2018), but it did inhibit protease activity and promote amylase activity in the digestive system. MvRL was found to affect the activity of gut endoglucanase, phosphatases, b-glucosidase, and trypsin-like enzymes in *Nasutitermes corniger* workers and soldiers, reported by de Albuquerque et al. (2012).

Plant lectins have been found to be biologically active against a variety of insects. Chitin-binding lectins derived from *Microgramma vacciniifolia* rhizome lectin (MvRL) showed anti-nutritional effects on survival, feeding, and nutrition of *Sitophilus zeamais* adults (de Albuquerque et al., 2020). *Moringa oleifera* seeds containing a water-soluble lectin (WSMoL) negatively impacted the physiology of the pest *Sitophilus zeamais*, which could have long-term consequences

(de Oliveira et al., 2020). According to Napoleão et al. (2013), the leaf lectin (MuLL) derived from *M. urundeuva* adversely affected the activity of digestive enzymes in the stomachs of *S. zeamais* adults, inhibiting digestive processes. Seeds containing the lectin WSMoL (water-soluble *M. oleifera* lectin) have been shown to be insecticidal against *Aedes aegypti* eggs and larvae. (Coelho et al., 2009; Santos et al., 2012, Santos et al., 2020). The survival and nutritional parameters of *S. zeamais* adults are negatively affected by the ingestion of an artificial diet containing a saline concentrate from *Schinus terebinthifolius* Raddi leaves (LE) or its lectin (SteLL, *S. terebinthifolius* leaf lectin). (Camaroti et al., 2018). However, progress in lectin research has been hampered due to concerns about the effects of ingesting snowdrop lectin on higher animals,, although no adverse effect was seen in rats fed for 90 days on transgenic rice containing GNA was seen (Poulsen et al., 2007). Insect resistance was demonstrated by the expression of lectin genes in various crop plants (**Table 3**). Although, lectins have been found to have negative effects on insect pests of various orders and stages of development, preventing growth, survival, nutrition, development, and reproduction (Napoleão et al., 2019). However, because of their known toxicity to mammals and humans, caution should be exercised in their use in transgenic plants.

2.1.4. Fusion Proteins

Bt insecticides are widely used, with *Bt* toxins accounting for up to 90% of microbiological insect control products. As a result, there's a possibility that insects will become resistant to *Bt* toxins. (Tabashnik, 1994; Ferré et al., 1995). The diamondback moth (*Plutella xylostella*) has evolved resistance in some open field populations in response to repeated exposure to foliar sprays containing *Bt* proteins. (Perez and Shelton, 1997), whereas recessive mutant alleles can confer resistance to multiple *Bt* toxins in laboratory selection experiments with other insect pests (Tabashnik et al., 1996). The stacking or pyramiding of multiple transgenes in the same transgenic plant and the use of hybrid toxins against insect pests are two recent strategies to address potential limitations in conventional transgenic insect pest control. et alThus fusion proteins are formed by joining together different insecticidal proteins. In the host plant system after transcription and translation these proteins form a single polypeptide units which are more effective against phytophagous insects. Dow Agro Sciences' transgenic cotton lines carrying a hybrid fusion protein, *cry1Be* + *cry1Fa*, demonstrated increased resistance to *S. litura* and *O. nubilalis* insects (Meade et al., 2017). Transgenic rice containing the *cry2AX1* (derivative of *cry2Aa* and *cry2Ac*) gene was viewed as resistant to lepidopteran insects, as indicated by Chakraborty et al. (2016). Koerniati et al. (2020) reported that the synthetic sugarcane expressing *cry1Ab* + *cry1Ac* fusion protein showed resistance against shoot borer. Agrobacterium-mediated transformation of transgenic *B. juncea* expressing a fusion protein derived from lectin and a protease inhibitor resulted in resistance to phytophagous aphids. (Rani et al., 2017a). In another case study, Chang et al. (2017) demonstrated that transgenic maize containing *cry1Ab*/*cry2Aj* fusion protein in kernel conferred higher mortality in *S. exigua*, a lepidopteran pest and *Harmonia axyridis*, a coleopteran pest. Transgenic rice line expressing *cry1Ac* + ASAL conferred durable and enhanced resistance to major insects such as yellow stem borer, leaf folder,, and brown plant hopper (Boddupally et al., 2018). Pyramiding of *cry1Ab* + vip3A showed resistance against the rice leaf folder and Asiatic rice borer (Xu et al., 2018). A significant level of protein expression was observed in transgenic rice lines carrying *cry2Aa* + *cry1Ca* protein lethal to Asiatic rice borer, *Chilo suppressalis* (Qiu et al., 2019). Katta et al. (2020) investigated that the transgenic plant developed by transferring a triple gene construct containing *cry2Ab* + *cry1F* + *cry1Ac* genes into an elite cotton variety (Narasimha) conferred a significant level of mortality to *H. armigera* and *S. litura* at T_2 generation. An account of studies utilization of the fusion proteins used for insect resistance in crop plants for pest management is depicted in **Table 4**. Pyramiding of two or more genes is a sustainable strategy for achieving good management of lepidopteran, coleopteran and hemipteran insects pests. Furthermore, these innovations could pave the way for development of insect resistant crops by delaying the phenonmenon of resistance development in insects.

2.1.5 Protease Inhibitors

Protease inhibitors (PIs) are plant-derived inhibitors that prevent insect pests from digesting their food by inhibiting the activity of digestive proteases (Haq et al., 2004; Macedo and Freire, 2011; Zhu-Salzman and Zeng, 2015). Insect digestive proteases are known to be inhibited by PIs, through preventing proteolysis and results in decreased fecundity, increased mortality and longer developmental period due to the deficiencies of essential amino acids. The most investigated plant PIs against pests are serpins and cystatins. Serpins, with a molecular mass of approximately 39–43 kDa, are irreversible serious inhibitors of serine proteases. Serine proteases have been discovered in insect orders like, Diptera (flies), Lepidoptera (moths and butterflies), Orthoptera (grasshoppers, locusts), Coleoptera (beetles) and Hymenoptera (bees and wasps) (Irving et al., 2002). Cystatins, a PIs protein with a molecular mass of 12–16 kDa, inhibit the activity of cysteine proteases, which are the primary digesting proteases in Coleopterans and Hemipterans. Several studies have also reported that volatile compounds such as methyl jasmonate, one of the key regulators of plants' defensive response to insect herbivores, inhibit gut protease after wounding. (Singh et al., 2016), cause neighboring unwounded plants to produce proteinase inhibitors, effectively prearming the local population against insect attack (Stevens et al., 2012). Legume trypsin inhibitors inhibit a wide spectrum of proteases and have an insecticidal activity against a variety of key insects (Macedo et al., 2004; Sharma, 2015). Protease inhibitor genes were incorporated in rice cultivars (Duan et al., 1996; Xu et al., 1996) to improve protection against stem borers, and wheat (Altpeter et al., 1999) to protect them against foliage-feeding and storage pests. Protease inhibitors when fed to insect pests either through artificial diet or transgenic plants resulted into increased insect mortality (de Pg Gomes et al., 2005; Gatehouse, 2011) and adversely affected the growth and development of insect larvae from different insect orders (de Pg Gomes et al., 2005; A,; Gatehouse, 2011) (Burgess et al., 1994; De Leo et al., 2001; Outchkourov et al., 2004; Ribeiro et al., 2006; Tamhane

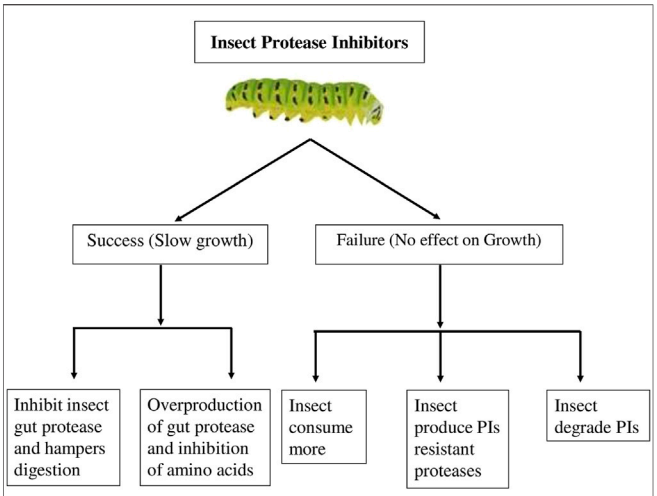


FIGURE 3 | Impact of protease inhibitors (PIs) on insect pest growth.

et al., 2007; Dunse et al., 2010; Schneider et al., 2017). **Figure 3** depicts the success and failure of protease inhibitors in a variety of insect pests. Protease inhibitors may enlighten a new dimension in insect pest management. However, due to lack of understanding of insect physiology and biochemistry, it suffered great failure in recent past. Also PIs turned worthless due to immense adaptive potential of insect pests and its long coevolutionary relationship with host plant. Solving these issues could pave the way for future research.

2.1.6 α-Amylase Inhibitors

α-amylase is a digestive enzyme present in insects for digestion of carbohydrates. An α-amylase inhibitor, affect digestion in insects by inhibiting the activity of α-Amylase enzyme in insects. Various types of α-amylase inhibitors, present in seeds and vegetative organs of plants, was found to control a numbers of phytophagous insects (Chrispeels et al., 1998). Seeds of *Phaseolus vulgaris* expressing an α-amylase inhibitor negatively affected the growth and development of cowpea weevil *Callosobruchus maculatus* and Azuki bean weevil *Callosobruchus chinensis* (Ishimoto and Kitamura, 1989; Shade et al., 1994). Morton et al. (2000) reported that transgenic pea and azuki bean seeds expressing the inhibitor, aAI-1, exhibited enhanced resistance against bruchids, the pea weevil (*Bruchus pisorum*), the cowpea weevil (*Callosobruchus maculatus*) and the azuki bean weevil (*Callosobruchus chinensis*). Kaur et al. (2022) investigated that higher activity of α-amylase inhibitors in the central whorl leaves and stems of maize genotypes might be responsible for inducing resistance against *Chilo partellus* infestation. The multiplication and damage of *Rhyzopertha dominica*, a major pests of stored wheat grains can be effectively controlled by inhibiting the α-amylase enzyme through wheat α-amylase inhibitors (Priya et al., 2010). A gene named aAI-Pc1, encoding an α-amylase inhibitor was isolated from cotyledons of *Phaseolus coccineus* and introduced into coffee plants, confers resistance to coffee berry borer, *Hypothenemus hampei* (de Azevedo Pereira et al., 2006). Recently, a wheat gene encoding an α-amylase inhibitor was expressed in tobacco, resulting in increased resistance to *Spodaptera* spp. and *Agrotis* spp (Jaiswal et al., 2018). Isolation of a novel alpha-amylase inhibitor from papaya seeds (*Carica papaya*) showed increased larval mortality, decreased insect fecundity and adult longevity of cowpea weevil (*Callosobruchus maculates*) (Farias et al., 2007). Therefore these studies indicates the successful utilization of α-amylase inhibitors in insect pests management.

2.1.6.1 Insect Chitinase

Insect chitinases are the hydrolytic enzymes having potential to inhibit or degrade the chitin. In insects, chitin is the main component of the exoskeleton and peritrophic membrane. It provide protection from harsh environmental conditions, external mechanical disruption and natural enemies (Chen et al., 2018b). The hydrolysis of chitin is essential for ecdysis (periodic shedding of the old cuticle). Chitinases are expressed in various organisms including those that lack chitin such as plants to recognize and degrade the chitin in chitin containing insects ((Oyeleye and Normi, 2018). Transgenic expression of chitinase enzyme has been proposed as a crop protection strategy. Role of chitinase enzyme in insect pests management has been studied in several insects such as silkworm *B. mori*, rice brown planthopper *N. lugens*, cotton mealybug *P. solenopsis and* rice striped stem borer *C. suppressalis* (Pan et al., 2012; Xi et al., 2015; Su et al., 2016; Omar et al., 2019). Transgenic maize plants expressing a chitinase gene showed enhanced resistance against corn borer (*Sesamia cretica*) (Osman et al., 2016). Insect chitinases have been established as biopesticides and transgenes in crop protection due to the inhibitory effects on the growth and development of insects. Not much research on insect chitinase has been done and lack of structural information on some insect chitinase has hampered the development of potential agrochemicals targeting insect chitinase. Better understanding of their structure and biochemistry will accelerate their usage in biotechnological processes.

2.2 Genome Editing

Insects acquiring resistance to the *Bt* traits has posed a threat to agricultural productivity, prompting researchers to seek out novel, cost-effective, and environment friendly techniques to insect pest management, as well as ways to combat insect resistance. Nowadays, insect pest management tactic has shifted to gene editing, which is a newer and more advanced method (Anastacia Books, 2019). Gene editing, also called genome editing, is a technique that involves inserting, deleting, or replacing DNA bases in a specific target DNA sequence of the genome for effectively altering the function of a gene by using the cell's natural mechanisms (Bortesi and Fischer 2015). It is one of the most widely used technologies in present-day science that empowers researchers to change/alter a living being's DNA (Belfort and Bonocora, 2014). It is an emerging opportunity increasingly being used in insect pest management through expanding its possibilities and opportunities to enhance plant resistance to insect pests.

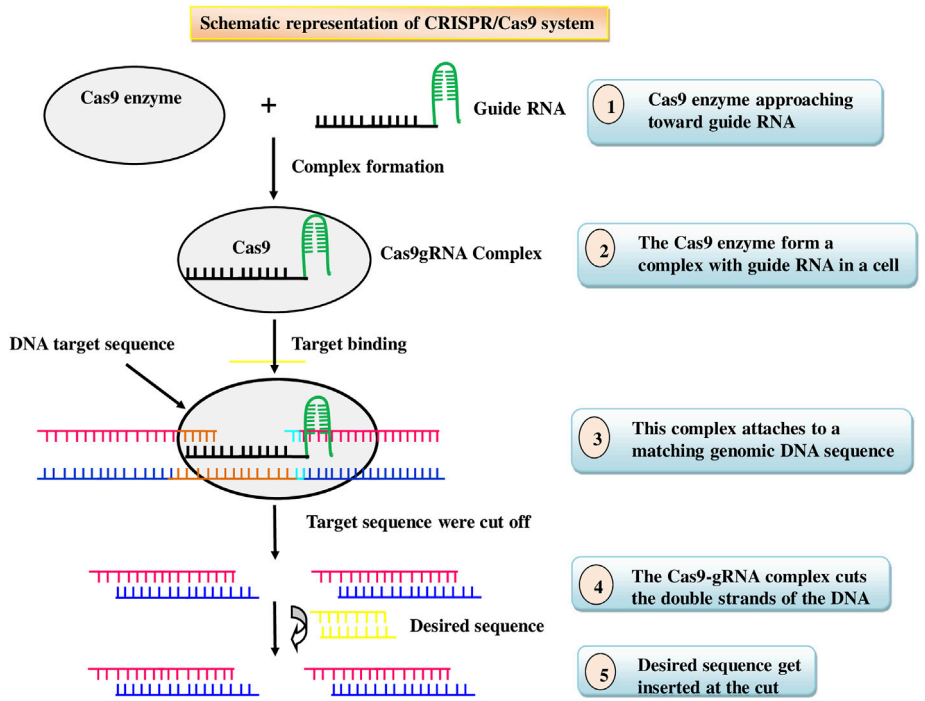


FIGURE 4 | Schematic representation of CRISPR/Cas9 system.

Nucleases are used in these technologies to cut certain genomic target sequences. Two types of genome editing tools, comprising transcriptional activator-like effector nucleases (TALENs) and clustered regularly interspaced short palindromic repeats (CRISPR)/Cas9, are accessible and are frequently applied. Cas9 protein and single-guide RNAs (sgRNAs) which can be easily designed are the two main parts of the CRISPR/Cas9 system, while TALEN requires to be redesigned to target different loci each time. Cas9-mediated genome editing is achieved by a process: DNA cleavage followed by DNA repair (**Figure 4**). Furthermore, the CRISPR-Cas9 technology allows researchers to add, remove, replace, or regulate genes in a variety of animals, resulting in heritable, targeted modifications that were previously difficult to create (Ricroch et al., 2017). One of the most important practical advantages of CRISPR-Cas9 technology is multiplexing, or the introduction of double-stranded breaks (DSBs) at many locations in the genome that can be used to edit multiple genes simultaneously (Li et al., 2013). Thus, in recent times, CRISPR/Cas9 has emerged as a technically simple, newest, most effective, and an effective tool for developing insect pest resistance. It has been successfully used to prevent the accumulation of specific gene products in a variety of crops by either deleting the gene or inducing missense mutations in the target gene (Gao, 2021). Most polyphagous insects use the plant's own volatiles, gustatory signs, visual appearance, oviposition sites, and collaborations to recognise host plants. (Larsson et al., 2004). Genome editing can be utilized to change plant volatile mixtures, which could be an alternative pest management strategy. However, caution should be taken to ensure that the alteration has no negative consequences for the beneficial insect population. In a study, the overproduction of anthocyanin pigmentation caused the transgenic tobacco plant's leaves to turn red which deterred both the herbivores *Spodoptera litura* and *Helicoverpa armigera* due to change in leaf color (Malone et al., 2009). In an anotherstudy, CRISPR/Cas9 was usedto target six loci associated with tomato yield and efficiency in wild tomato *S. pimpinellifolium* (Zsögön et al., 2018). Although this wild tomato shows resistance to a numberof arthropod insects, including spider mites, and produces modest yields (Rakha et al., 2017). Ming et al. (2021) demonstrated that the CRISPR/Cas9 genome editing system is an effective tool for studying the function of SfABCC2, a *Cry1F* gene receptor that confers resistance to *S. frugiperda*. Insect-resistant rice plants with mutations in the cytochrome P450 gene CYP71A, which catalyses the conversion of tryptamine to serotonin, accumulated high levels of salicylic acid but lacked serotonin. (Lu et al., 2018). This could make genome editing more appealing than transgene stacking for the production of next-generation insect-resistant crops. Although CRISPR gene editing is an effective tool to combat insect pest problems as it has the capacity to alter the specific gene of interest. However, commercial use of CRISPR/Cas9 in insect pest management is still in its early stages. It has been extensively reformed for various applications in model animals, which may reveal potential

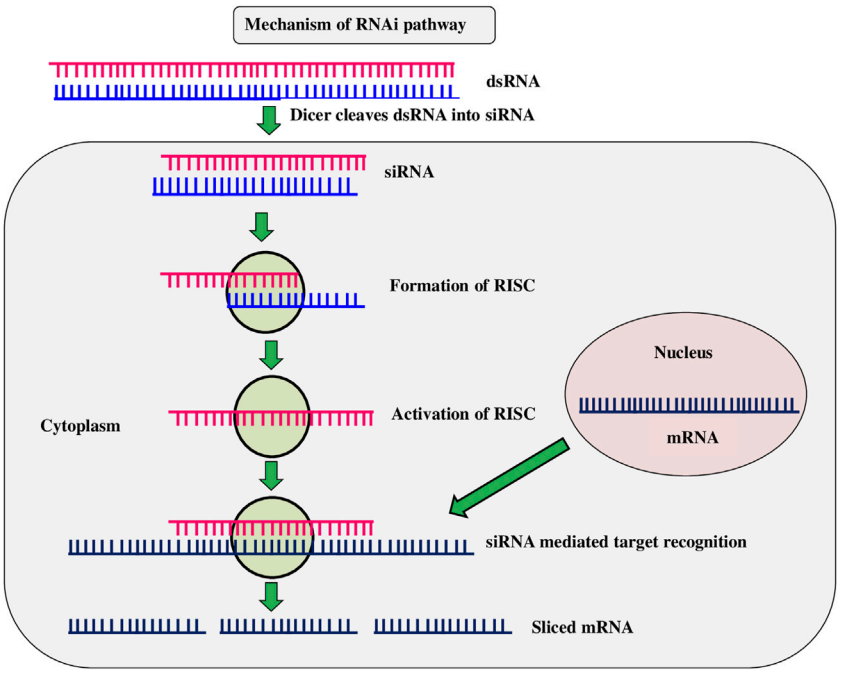


FIGURE 5 | RNA interference (RNAi).

insect applications. The list of insect pests engineered for pest management using CRISPR/Cas9 are given in **Table 5**.

2.3 RNA Interference for Plant Resistance to Insect

RNAi is a method of suppressing gene expression by suppressing specific sequences and is known by various names, including co-concealment, post-transcriptional gene silencing (PTGS), and suppressing. It is an advancement of novel gene silencing mechanisms triggered by double-stranded RNA at the cellular level. (**Figure 5**). When a double-stranded RNA (dsRNA) is injected into a cell, it makes undesirable genes to be repressed (Kamthan et al., 2015). The RNAi strategy for pest control is based on ingestion of double-stranded RNA (dsRNA) into the target pest system. After ingestion, dsRNA expresses either through hairpin or by other different means and spread throughout the insect system (Katoch et al., 2013). Transgenic *Bt* toxins are mostly effective against Lepidoptera and Coleoptera larvae, by acting in the mid-gut of susceptible target insects, leaving other insect orders unmanaged. The RNAi technique was expected to be able to control a wider range of insects, especially sap-sucking insects, which transgenic crops had failed to control. It also opens up new possibilities for eco-friendly insect pest control in agricultural crop plants (Mamta and Rajam 2017). dsRNAs are often utilized in plants to interfere with specific gene silence in order to develop disease resistance through genetic changes. Resistance to *C. suppressalis* was provided by rice knockdown lines TT51 (cryAb and cry1Ac) and T1C-19 (cry1Ac) with two aminopeptidase N genes (APN1 and APN2) (Qiu et al., 2017). Western corn rootworm (*D. virgifera*) fertility and larval feeding were reduced when the *dvvgr* and *dvbol* genes were silenced in maize (Niu et al., 2017). The use of a dsRNA/nano carrier formulation to target the TREH, ATPD, ATPE, and CHSI genes resulted in a greater proportion of soybean aphid (*Aphis glycines*) mortality (Yan et al., 2020). Transgenic cotton lines generated by combining *Bt* toxin with RNAi caused inhibition of juvenile hormone methyl transferase (JHMT) in *H. armigera*. (Ni et al., 2017). Knockdown of acetylcholine esterase gene (*AChE*) in rice lines resulted in reduced larval length and weight of yellow stem borer within 15 days (Kola et al., 2019). Ingestion of double-stranded (ds) RNAs in an artificial diet causes RNA interference in several coleopteran species, including the western corn rootworm (WCR) *Diabrotica virgifera virgifera*, which results in larval stunting and mortality. (Baum et al., 2007). The RNAi mechanism was tested by the spraying of dsRNAs in maize, resulting in gene knockdown and increased insect mortality rates in piercing, sucking, and stem borer insects (Li et al., 2015). Mamta et al. (2016) found that HI-RNAi produced induced death and developmental abnormalities in *H. armigera* larval, pupal, and adult stages when the chitinase gene (HaCHI) was silenced to establish resistance in tobacco and tomato. Tomato plants with a dsRNA targeting a gene encoding a phenolic glucoside malonyltransferase, which detoxifies phenolic glycosides, were recently found to be completely resistant to the tobacco whitefly, *Bemisia tabaci* (Xia et al., 2021). The Colorado potato beetle (Zhang et al., 2015), tobacco whitefly and tobacco hornworm, *Manduca sexta* (Burke et al., 2019), were all killed by chloroplast-expressed dsRNAs (Dong et al., 2020). Di Lelio et al. (2022) reported that when *Spodoptera littoralis* larvae eat tobacco plants expressing a dsRNA targeting the Sl 102 immune gene, the gene

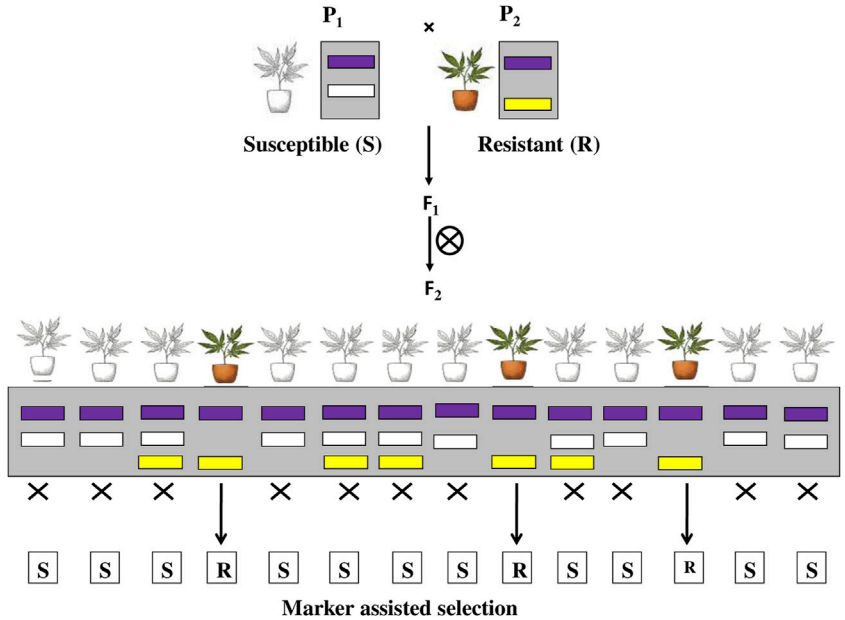


FIGURE 6 | Marker assisted selection and its application in insect pest management.

get silenced (et al.). RNAi technology is effective in knocking down target genes in a variety of insect orders, including *Diabrotica v. virgifera*, maize rootworm larvae (Khajuria et al., 2015). These findings suggest that to generate insect-resistant plants, RNAi is one of the most effective methods. However, the technology is currently being investigated, and its existing limitations make it less viable as an insect pest management strategy (Anastacia Books, 2019).

The most challenging task of this technology is allowing for efficient dsRNA uptake by the insect. The generation and transport of dsRNAs have been proven in two ways. The first is HIGS (host induced gene silencing), which involves the transgenic expression of dsRNAs derived from the crop genome. In this method, insects supposed to feed on the crop will eat the dsRNA. In the second strategy, dsRNAs are synthesized in high concentrations and applied to insect-infested crops as a foliar spray. Spray-induced gene silencing is the name for this method (SIGS). The target genes will be silenced in both approaches in the target species (Christiaens et al., 2020a). Many challenges are still there. The inherent ability of RNA, while ensuring that dsRNAs don't persist in the environment, is destructive in unfavorable environmental conditions and restricts opportunities for SIGS approaches (Bramlett et al., 2020). Similarly, while some insects quickly take up dsRNA, resulting in high death rates, other species have low dsRNA take-up and nuclease degradation, resulting in inefficient outcomes. (Christiaens et al., 2020b; Shaffer, 2020). Success is also dependent on whether sufficient amount of dsRNA accumulate in the tissues on which the insects feed. The development of novel spray formulations, many of which use nanomaterials, is being used to address dsRNA stability and uptake in ongoing research. (Christiaens et al., 2020b). Several studies on insect pest management using the RNAi tool are shown in **Table 6**.

2.4 Marker-Assisted Selection

The use of molecular markers to assist phenotypic selections in crop improvement is known as marker-assisted selection (MAS). It involves selecting individuals based on their marker pattern (genotype) rather than their observable traits (phenotype) as shown in **Figure 6**. There are various types of molecular markers, such as single nucleotide polymorphism (SNP), have been recognised and have shown great promise in enhancing the efficiency and accuracy of conventional plant breeding. Molecular marker techniques are the most advanced method for transferring desired genes into desired crop plants in the required combination. It is the most widely used molecular techniques, and their application is a novel opportunity for increasing the yield of crop (Das et al., 2017). MAS studies showed introgression of *Bph14* and *Bph15* through molecular marker-assisted selection (MAS) to enhance the resistance in Minghui 63 and its derived hybrids against BPH (Hu et al., 2012). Resistance to bacterial blight (BB) and brown planthopper (BPH) was achieved in Yuehui9113 and F1 hybrids by pyramiding one BB resistance gene (*Xa21*) and two BPH resistance genes (*Bph14 and Bph15*) in Yuehui9113 using a marker-assisted backcrossing (MABC) strategy combined with phenotypic selection (He et al., 2019). Rice line, ASD7 expressing a BPH resistance gene *bph2* when crossed to a susceptible cultivar C418, a japonica restorer line and evaluated through marker-assisted selection (MAS) exhibited significantly higher resistance against brown plant hopper *Nilaparvata lugens,* one of the most destructive pests of rice crop (Li-Hong et al., 2006). Liu et al. (2016) investigated that the pyramiding of two brown plant hopper resistance genes

Bph3 and *Bph27* (t), into elite rice cultivars through marker-assisted pyramiding showed significantly enhanced resistance against BPH and reduction in the yield loss caused by BPH. Shabanimofred et al. (2015) developed rice cultivars through marker-assisted selection (MAS) that provided resistance in rice against biotypes 2 and 3 of brown planthopper (BPH). Sharma et al. (2004) used marker-assisted pyramiding to successfully construct the *Bph1* and *Bph2* resistance genes on rice chromosome 12 to provide resistance against rice BPH. et al. As a result, using MAS to improve pest resistance would be very beneficial. There are various advantages of using MAS to enhance selection efficiency of insect resistant plants 1) It can be performed on seedling material, 2) less affected by environmental conditions, 3) MAS may be cost effective and faster than conventional phenotypic assays, 4) multiple markers can be evaluated using the same DNA sample etc. But the potential drawbacks of MAS are 1) Recombination between the marker and the gene of interest may occur, leading to inaccurate results 2) Incorrect estimates of QTL locations and effects may result in slower progress than expected, 3) Markers developed for MAS in one population may not be transferrable to other populations.

2.5 Anther Culture

Anther culture is a technique by which immature pollen is allowed to divide and grow into tissue (either callus or embryonic tissue) with the intention of generating haploids (plants with a N chromosome number). In this process, pollen-containing anthers are separated from a flower and placed in a suitable growing medium. An artificial medium could be used to culture anther or pollen grain *in vitro*. Anther can produce a callus, shoot, root, and eventually the entire plant in an artificial medium. All of the plants that are grown are haploid. It is the most viable and effective way for rapidly producing homozygous haploid plants. This method can hasten the development of a homozygous population of insect-resistant plants. Rice anther culture lines, 952836, 953508, 953509, 953510, 953511, 953527 and 953541 showed moderate level of tolerance to the rice water weevil, *Lissorhoptrus oryzophilus* Kuschel (N'guessan et al., 1994). Park et al. (2014) develop multi-resistant rice lines using anther culture for providing resistance against bacterial blight, rice stripe virus and brown planthopper.

2.6 Embryo Culture

Embryo culture is a technique in which immature or mature zygotic embryo is recovered without injury which normally aborts. These embroys are further cultured on the artificial nutrient media under an aseptic environment to get a vigor and viable plant through a successful ontogeny process. The standardization of the nutrient medium is required for induction of embryogenesis and seedling development. Generally wild species are often more resistant to insect pests. Wide hybridization has been used to transfer genes conferring insect-resistant from wild species to cultivated plants. It has been observed that such hybridization leads to the production of abnormal inter-specific hybrid embryos, which can be rescued using embryo culture technique. Jaiswal et al. (2018) has reported studies of insect resistant genes transfer from wild to cultivated species in wheat, rice, peanut, lettuce and cotton using embryo culture technique.

2.7 Protoplast Fusion

Protoplasts are plant cells of which cell walls are taken out and the cytoplasmic membrane is the peripheral layer. Protoplast can be isolated by digesting the cell wall with specific lytic enzymes. Protoplast fusion is a physical phenomenon, wherein at least two protoplasts come together and stick with each other either spontaneously or in presence of fusion-inducing agents. By protoplast fusion, it is feasible to transfer a few desirable genes from one species to another. For example, pest resistance characteristics may be present in one of two species that cannot be sexually hybridized. In this situation, protoplast fusion may result in the formation of a hybrid between two species. Protoplasts can be cultivated in an artificial medium, and some of them will grow into full-fledged plants. Thus the plants produced may be carrying the resistant traits. Also, it is the only means of combining two cytoplasmically inherited characteristics in a single genotype. Protoplast derived clones produced by Mexican wild species and cultivated potato species using protoplast fusion system expresses a significant level of resistance to both Colorado potato beetle and potato late blight (Chen et al., 2008).

2.8 Somaclonal Variation

Insect-resistant varieties can be selected through somaclonal variation. These can be chosen using the procedures below. 1) High-yielding varieties' calli or cell suspensions were cultured for numerous or long-term cycles, 2) long-term cell lines were regenerated into plants, and 3) the regenerated plants were tested against target insects. In field plots, about 2000 sugarcane seedlings were tested for resistance to the sugarcane borer under artificial and natural infestations and it was found that some somaclones were reported to be resistant to sugarcane borer. The same process was utilized to develop sorghum somaclones, resistant to the autumn armyworm. Diawara et al. (1996) reported that somaclonal lines K-26 [1], K-180 [3]2 and K-128 exhibited improved resistant to celery major insect, *Spodoptera exigua*.

Although the above-mentioned plant tissue culture techniques like anther culture, embryo culture, protoplast fusion, and somaclonal variations proved to be more efficient in the development of plants resistant to various insect pests. However, these techniques are not widely used due to some potential drawbacks, such as high costs, the production of harmful secondary metabolites that kill the desired insect-resistant plants, the medium required for growth is not known, and so on. Also, due to advancements in technology, which are more efficient, quick, and reliable, these methods are no longer used.

3 CONCLUSION AND FUTURE OUTLOOKS

Insects are the major concern for declining the agricultural production. To cope with the problem of insect pests, farmers

are more inclined to the use of chemical insecticides as these provide a quick solution to the problem. The rapidly increasing awareness of the human and animal health issues as well as environmental impacts, of indiscriminate use of pesticide has offered new incentive to the potential alternative pest-control methods. In this perspective, host plant resistance is an environmentally friendly control method that is an important part of IPM (Integrated pest management) programmes. The development of insect-resistant varieties offers a stable and cumulative effect on the pests' population and has no harmful effect on the environment. The identification of insect pest resistant sources in various crops has made significant progress. However, development of insect resistant crop varieties through conventional methods is slow and difficult to attain due to the entanglement of quantitative traits at multiple loci. New opportunities in the form of newer biotechnological tools have opened new ways of pest control and offers great opportunities to develop a sustainable, multi-mechanistic resistance to insect pests. Biotechnological approaches are now being used to develop novel plant resistance characteristics that provided excellent protection against invasive and destructive crop pests in a variety of crops by utilization of novel molecules, exploiting insecticidal genes and changing the level and pattern of expression of genes. Many insect-resistant plants have been developed as a result of biotechnology like corn, rice, cotton, canola, soybean, tobacco, apple, potato etc. With the advent of several tools of biotechnology such as genome editing, genetic transformation, anther culture, embryo culture, protoplast fusion, somaclonal variation, and marker-assisted selection will accelerate the development of insect-resistant crops now and in the future. By expressing bacterial delta-endotoxins, vegetative insecticidal proteins, and other plant qualities like lectins, protease inhibitors, etc., hereditary designing will guide towards the development of insect-resistant crops at much faster rate. Furthermore, RNA interference and genome editing by CRISPR/Cas9 offer novel approach to the production of insect-resistant crops. Therefore, biotechnology have come as a boon in tackling global pest problem, contributing to the development of noval insect resistant crop plants that have proven to be cost effective, pesticide-resistant, and environmentally safe. Despite the utilization of modern technology in crops to achieve resistance to a variety of insect pests, some agricultural pests frequently develop resistance to insecticidal toxins, wreaking havoc on crop productivity. The obstacles of understanding plant-insect interactions should be addressed by the research groups. To develop plants resistant to insects advances like RNAi and CRISPR techniques can be used to silence/edit sensitive or negative regulatory alleles of plant immune genes. New advancements that give more viable solutions for arising pests can improve and supplement the perseverance of plant-resistant elements. However, before the commercialization of an insect pest resistant transgenic crop variety, it is pertinent to study the potential impacts on environment specifically on non-target organisms. Also, the benefits and hazards associated with the adoption of insect-resistant crops, particularly for developing nations and resource-poor smallholder farmers, should be considered prior to carry out such initiatives. No doubt, biotechnology has opened the door to a plethora of novel ways for controlling insect pests, many of these products will necessitate regulatory frameworks that may not currently exist for certain products or in some places. They will also require the support of producers and consumers, which will necessitate open conversations, including the potential for new technology to make a significant contribution to societal change. In short it can be concluded that biotechnology exhibits unique applications of science that can be used for the welfare of society through the development of crops with improved nutritional quality, resistance to pests and diseases, and low cost of production. Biotechnology, in this context, is an aspect of science that, if used with caution and ethics, has the potential to offer substantial benefits.

AUTHOR CONTRIBUTIONS

PK and DK—Collected and compiled the literature. PJ: Conceived the idea and wrote the manuscript. PK, SaK, CM, SuK, and GS: Assisted in manuscript writing.

FUNDING

Funds will be given by ICAR.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.914029/full#supplementary-material

REFERENCES

Abah, J., Ishaq, M. N., and Wada, A. C. (2010). The Role of Biotechnology in Ensuring Food Security and Sustainable Agriculture. *Afr. J. Biotechnol.* 9 (52), 8896–8900.

Abbas, M. S. T. (2018). Genetically Engineered (Modified) Crops (Bacillus Thuringiensis Crops) and the World Controversy on Their Safety. *Egypt. J. Biol. Pest Control* 28 (1), 1–12. doi:10.1186/s41938-018-0051-2

Albuquerque, L. P. D., Procópio, T. F., Guedes, C. C. D. S., Pontual, E. V., Paiva, P. M. G., and Napoleão, T. H. (2020). Antinutritional Effects of the Chitin-Binding Lectin from *Microgramma Vacciniifolia* Rhizome (MvRL) on *Sitophilus Zeamais*. *J. Stored Prod. Res.* 88, 101652. doi:10.1016/j.jspr.2020.101652

Alemu, M. (2020). Trend of Biotechnology Applications in Pest Management: a Review. *Int. J. Appl. Sci. Biotechnol.* 8 (2), 108–131. doi:10.3126/ijasbt.v8i2.28326

Altpeter, F., Diaz, I., McAuslane, H., Gaddour, K., Carbonero, P., and Vasil, I. K. (1999). Increased Insect Resistance in Transgenic Wheat Stably Expressing Trypsin Inhibitor CMe. *Mol. Breed.* 5 (1), 53–63. doi:10.1023/a:1009659911798

Anastacia Books (2019). *New Biotechnological Approaches to Insect Pest Management and Crop Protection; Gene Editing Approach (CRISPR-Cas System)*. Lincoln: University of Nebraska.

Baburao, T. M., and Sumangala, B. (2018). Development and Molecular Characterization of Transgenic Pigeon Pea Carrying cry2Aa for Pod Borer Resistance. *J. Pharm. Phytochem.* 75, 1581–1585.

Baum, J. A., Bogaert, T., Clinton, W., Heck, G. R., Feldmann, P., Ilagan, O., et al. (2007). Control of Coleopteran Insect Pests through RNA Interference. *Nat. Biotechnol.* 25, 1322–1326. doi:10.1038/nbt1359

Belfort, M., and Bonocora, R. P. (2014). Homing Endonucleases: from Genetic Anomalies to Programmable Genomic Clippers. *Homing Endonucleases* 1, 1–26. doi:10.1007/978-1-62703-968-0_1

Bett, B., Gollasch, S., Moore, A., James, W., Armstrong, J., Walsh, T., et al. (2017). Transgenic Cowpeas (*Vigna Unguiculata* L. Walp) Expressing *Bacillus Thuringiensis* Vip3Ba Protein Are Protected against the Maruca Pod Borer (*Maruca Vitrata*). *Plant Cell Tiss. Organ Cult.* 131 (2), 335–345. doi:10.1007/s11240-017-1287-3

Bharathi, Y., Vijaya Kumar, S., Pasalu, I. C., Balachandran, S. M., Reddy, V. D., and Rao, K. V. (2011). Pyramided Rice Lines Harbouring *Allium Sativum* (Asal) and *Galanthus Nivalis* (Gna) *Lectin* Genes Impart Enhanced Resistance against Major Sap-Sucking Pests. *J. Biotechnol.* 152 (3), 63–71. doi:10.1016/j.jbiotec.2011.01.021

Bi, H.-L., Xu, J., Tan, A.-J., and Huang, Y.-P. (2016). CRISPR/Cas9-mediated Targeted Gene Mutagenesis inSpodoptera Litura. *Insect Sci.* 23 (3), 469–477. doi:10.1111/1744-7917.12341

Birkett, M. A., and Pickett, J. A. (2014). Prospects of Genetic Engineering for Robust Insect Resistance. *Curr. Opin. Plant Biol.* 19, 59–67. doi:10.1016/j.pbi.2014.03.009

Boddupally, D., Tamirisa, S., Gundra, S. R., Vudem, D. R., and Khareedu, V. R. (2018). Expression of Hybrid Fusion Protein (Cry1Ac: ASAL) in Transgenic Rice Plants Imparts Resistance against Multiple Insect Pests. *Sci. Rep.* 8 (1), 8458. doi:10.1038/s41598-018-26881-9

Bommireddy, P. L., Leonard, B. R., Temple, J., Price, P., Emfinger, K., Cook, D., et al. (2011). Field Performance and Seasonal Efficacy Profiles of Transgenic Cotton Lines Expressing Vip3A and VipCot against *Helicoverpa Zea* (Boddie) and *Heliothis virescens* (F.). *J. Cotton Sci.* 15, 251–259.

Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 System for Plant Genome Editing and beyond. *Biotechnol. Adv.* 33 (1), 41–52. doi:10.1016/j.biotechadv.2014.12.006

Bramlett, M., Plaetinck, G., and Maienfisch, P. (2020). RNA-based Biocontrols-A New Paradigm in Crop Protection. *Engineering* 6 (5), 522–527. doi:10.1016/j.eng.2019.09.008

Burgess, E. P. J., Main, C. A., Stevens, P. S., Christeller, J. T., Gatehouse, A. M. R., and Laing, W. A. (1994). Effects of Protease Inhibitor Concentration and Combinations on the Survival, Growth and Gut Enzyme Activities of the Black Field Cricket, *Teleogryllus Commodus*. *J. Insect Physiology* 40 (9), 803–811. doi:10.1016/0022-1910(94)90010-8

Burke, W. G., Kaplanoglu, E., Kolotilin, I., Menassa, R., and Donly, C. (2019). RNA Interference in the Tobacco Hornworm, *Manduca Sexta*, Using Plastid-Encoded Long Double-Stranded RNA. *Front. Plant Sci.* 10, 313. doi:10.3389/fpls.2019.00313

Camaroti, J. R. S. L., de Almeida, W. A., do Rego Belmonte, B., de Oliveira, A. P. S., de Albuquerque Lima, T., Ferreira, M. R. A., et al. (2018). *Sitophilus Zeamais* Adults Have Survival and Nutrition Affected by *Schinus Terebinthifolius* Leaf Extract and its Lectin (SteLL). *Industrial Crops Prod.* 116, 81–89. doi:10.1016/j.indcrop.2018.02.065

Camaroti, J. R. S. L., Oliveira, A. P. S., Paiva, P. M. G., Pontual, E. V., and Napoleao, T. H. (20172017). "Phytoinsecticides for Controlling Pests and Mosquito Vectors of Diseases," in *Biocontrol Agents: Types, Applications and Research Insights*. Editor V. Green (New York: Nova Science Publishers), 147e188.

Caroline, N. h. M., Deogracious, P. M., George, M. T., James, R. M., Joel, W. D., and Paul, M. K. (2022). Identification of Potential Seed Storage Protein Responsible for Bruchid Resistance in Common Bean Landraces from Tanzania and Malawi. *Afr. J. Biotechnol.* 21 (1), 35–45. doi:10.5897/ajb2021.17354

Carrière, Y., Crickmore, N., and Tabashnik, B. E. (2015). Optimizing Pyramided Transgenic Bt Crops for Sustainable Pest Management. *Nat. Biotechnol.* 33 (2), 161–168. doi:10.1038/nbt.3099

Chakraborty, M., Reddy, P. S., Mustafa, G., Rajesh, G., Narasu, V. M. L., Udayasuriyan, V., et al. (2016). Transgenic Rice Expressing the cry2AX1 Gene Confers Resistance to Multiple Lepidopteran Pests. *Transgenic Res.* 25 (5), 665–678. doi:10.1007/s11248-016-9954-4

Chang, X., Lu, Z., Shen, Z., Peng, Y., and Ye, G. (2017). Bitrophic and Tritrophic Effects of Transgenic cry1Ab/cry2Aj Maize on the Beneficial, Nontarget *Harmonia axyridis* (Coleoptera: Coccinellidae). *Environ. Entomol.* 46 (5), 1171–1176. doi:10.1093/ee/nvx113

Chen, B., Hu, J., Almeida, R., Liu, H., Balakrishnan, S., Covill-Cooke, C., et al. (2016). Expanding the CRISPR Imaging Toolset with *Staphylococcus aureus* Cas9 for Simultaneous Imaging of Multiple Genomic Loci. *Nucleic Acids Res.* 44 (8), e75. doi:10.1093/nar/gkv1533

Chen, Q., Li, H. Y., Shi, Y. Z., Beasley, D., Bizimungu, B., and Goettel, M. S. (2008). Development of an Effective Protoplast Fusion System for Production of New Potatoes with Disease and Insect Resistance Using Mexican Wild Potato Species as Gene Pools. *Can. J. Plant Sci.* 88 (4), 611–619. doi:10.4141/cjps07045

Chen, W., Liu, C., Lu, G., Cheng, H., Shen, Z., and Wu, K. (2018a). Effects of Vip3AcAa+ Cry1Ac Cotton on Midgut Tissue in *Helicoverpa Armigera* (Lepidoptera: Noctuidae). *J. Insect Sci.* 18 (4), 13. doi:10.1093/jisesa/iey075

Chen, W., Qu, M., Zhou, Y., and Yang, Q. (2018b). Structural Analysis of Group II Chitinase (ChtII) Catalysis Completes the Puzzle of Chitin Hydrolysis in Insects. *J. Biol. Chem.* 293 (8), 2652–2660. doi:10.1074/jbc.ra117.000119

Chougule, N. P., Li, H., Liu, S., Linz, L. B., Narva, K. E., Meade, T., et al. (2013). Retargeting of the *Bacillus Thuringiensis* Toxin Cyt2Aa against Hemipteran Insect Pests. *Proc. Natl. Acad. Sci. U.S.A.* 110 (21), 8465–8470. doi:10.1073/pnas.1222144110

Chrispeels, M. J., Fatima Grossi de Sa, M., and Higgins, T. J. V. (1998). Genetic Engineering with α-amylase Inhibitors Makes Seeds Resistant to Bruchids. *Seed Sci. Res.* 8 (2), 257–264. doi:10.1017/s0960258500004153

Christiaens, O., Petek, M., Smagghe, G., and Taning, C. N. T. (2020a). "The Use of Nanocarriers to Improve the Efficiency of RNAi-Based Pesticides in Agriculture," in *Nanopesticides* (Cham: Springer), 49–68. doi:10.1007/978-3-030-44873-8_3

Christiaens, O., Whyard, S., Vélez, A. M., and Smagghe, G. (2020b). Double-stranded RNA Technology to Control Insect Pests: Current Status and Challenges. *Front. Plant Sci.* 11, 451. doi:10.3389/fpls.2020.00451

Coelho, J. S., Santos, N. D. L., Napoleão, T. H., Gomes, F. S., Ferreira, R. S., Zingali, R. B., et al. (2009). Effect of *Moringa Oleifera* Lectin on Development and Mortality of *Aedes aegypti* Larvae. *Chemosphere* 77 (7), 934–938. doi:10.1016/j.chemosphere.2009.08.022

Czapla, T. H., and Lang, B. A. (1990). Effect of Plant Lectins on the Larval Development of European Corn Borer (Lepidoptera: Pyralidae) and Southern Corn Rootworm (Coleoptera: Chrysomelidae). *J. Econ. Entomol.* 83 (6), 2480–2485. doi:10.1093/jee/83.6.2480

Das, A., Ghosh, P., and Das, S. (2018). Expression of *Colocasia Esculenta* Tuber Agglutinin in Indian Mustard Provides Resistance against *Lipaphis Erysimi* and the Expressed Protein Is Non-allergenic. *Plant Cell Rep.* 37 (6), 849–863. doi:10.1007/s00299-018-2273-x

Das, G., Patra, J. K., and Baek, K.-H. (2017). Insight into MAS: a Molecular Tool for Development of Stress Resistant and Quality of Rice through Gene Stacking. *Front. Plant Sci.* 8, 985. doi:10.3389/fpls.2017.00985

de Azevedo Pereira, R., Nogueira Batista, J. A., Mattar da Silva, M. C., Brilhante de Oliveira Neto, O., Zangrando Figueira, E. L., Valencia Jiménez, A., et al. (2006). An Alpha-Amylase Inhibitor Gene from Phaseolus Coccineus Encodes a Protein with Potential for Control of Coffee Berry Borer (*Hypothenemus Hampei*). *Phytochemistry* 67, 2009–2016. doi:10.1016/j.phytochem.2006.06.029

De Leo, F., Bonadé-Bottino, M., Ceci, L. R., Gallerani, R., and Jouanin, L. (2001). Effects of a Mustard Trypsin Inhibitor Expressed in Different Plants on Three Lepidopteran Pests. *Insect Biochem. Mol. Biol.* 31 (6-7), 593–602. doi:10.1016/s0965-1748(00)00164-8

de Oliveira, A. P. S., Agra-Neto, A. C., Pontual, E. V., de Albuquerque Lima, T., Cruz, K. C. V., de Melo, K. R., et al. (2020). Evaluation of the Insecticidal Activity of *Moringa Oleifera* Seed Extract and Lectin (WSMoL) against *Sitophilus Zeamais*. *J. stored Prod. Res.* 87, 101615. doi:10.1016/j.jspr.2020.101615

de Santana Souza, C., Procópio, T. F., do Rego Belmonte, B., Paiva, P. M. G., de Albuquerque, L. P., Pontual, E. V., et al. (2018). Effects of *Opuntia Ficus-indica*

Lectin on Feeding, Survival, and Gut Enzymes of Maize Weevil, *Sitophilus Zeamais*. *Appl. Biol. Chem.* 61 (3), 337–343. doi:10.1007/s13765-018-0363-7

Di Lelio, I., Barra, E., Coppola, M., Corrado, G., Rao, R., and Caccia, S. (2022). Transgenic Plants Expressing Immunosuppressive dsRNA Improve Entomopathogen Efficacy against *Spodoptera Littoralis* Larvae. *J. Pest Sci.* 95, 1413–1428. doi:10.1007/s10340-021-01467-z

Diawara, M. M., Trumble, J. T., Lacy, M. L., White, K. K., and Carson, W. G. (1996). Potential of Somaclonal Celeries for Use in Integrated Pest Management. *J. Econ. Entomology* 89 (1), 218–223. doi:10.1093/jee/89.1.218

Dong, Y., Yang, Y., Wang, Z., Wu, M., Fu, J., Guo, J., et al. (2020). Inaccessibility to Double-stranded RNAs in Plastids Restricts RNA Interference in *Bemisia Tabaci* (Whitefly). *Pest Manag. Sci.* 76 (9), 3168–3176. doi:10.1002/ps.5871

Douris, V., Steinbach, D., Panteleri, R., Livadaras, I., Pickett, J. A., Van Leeuwen, T., et al. (2016). Resistance Mutation Conserved between Insects and Mites Unravels the Benzoylurea Insecticide Mode of Action on Chitin Biosynthesis. *Proc. Natl. Acad. Sci. U.S.A.* 113 (51), 14692–14697. doi:10.1073/pnas.1618258113

Down, R. E., Fitches, E. C., Wiles, D. P., Corti, P., Bell, H. A., Gatehouse, J. A., et al. (2006). Insecticidal Spider Venom Toxin Fused to Snowdrop Lectin Is Toxic to the Peach-Potato Aphid, *Myzus persicae* (Hemiptera: Aphididae) and the Rice Brown Planthopper, *Nilaparvata Lugens* (Hemiptera: Delphacidae). *Pest. Manag. Sci.* 62 (1), 77–85. doi:10.1002/ps.1119

Duan, X., Hou, Q., Liu, G., Pang, X., Niu, Z., Wang, X., et al. (2018). Expression of *Pinellia Pedatisecta* Lectin Gene in Transgenic Wheat Enhances Resistance to Wheat Aphids. *Molecules* 23 (4), 748. doi:10.3390/molecules23040748

Duan, X., Li, X., Xue, Q., Abo-Ei-Saad, M., Xu, D., and Wu, R. (1996). Transgenic Rice Plants Harboring an Introduced Potato Proteinase Inhibitor II Gene Are Insect Resistant. *Nat. Biotechnol.* 14 (4), 494–498. doi:10.1038/nbt0496-494

Dunse, K. M., Stevens, J. A., Lay, F. T., Gaspar, Y. M., Heath, R. L., and Anderson, M. A. (2010). Coexpression of Potato Type I and II Proteinase Inhibitors Gives Cotton Plants Protection against Insect Damage in the Field. *Proc. Natl. Acad. Sci. U.S.A.* 107 (34), 15011–15015. doi:10.1073/pnas.1009241107

El-Dessouki, W. A., Mansour, M. R. K., and Eryan, N. L. (2022). Effects of Certain Weather, Biotic Factors and Chemical Components on the Population of Aphids in Egyptian Wheat Fields. *Egypt. Acad. J. Biol. Sci. A, Entomology* 15 (1), 1–13. doi:10.21608/eajbsa.2022.212703

Farias, L. R., Costa, F. T., Souza, L. A., Pelegrini, P. B., Grossi-de-Sá, M. F., Neto, S. M., et al. (2007). Isolation of a Novel Carica Papaya α-amylase Inhibitor with Deleterious Activity toward Callosobruchus Maculatus. *Pesticide Biochem. Physiology* 87, 255–260. doi:10.1016/j.pestbp.2006.08.004

Ferré, J., Escriche, B., Bel, Y., and Van Rie, J. (1995). Biochemistry and Genetics of Insect Resistance to *Bacillus Thuringiensis* Insecticidal Crystal Proteins. *FEMS Microbiol. Lett.* 132 (1-2), 1–7.

Fitches, E., Wilkinson, H., Bell, H., Bown, D. P., Gatehouse, J. A., and Edwards, J. P. (2004). Cloning, Expression and Functional Characterisation of Chitinase from Larvae of Tomato Moth (*Lacanobia Oleracea*): a Demonstration of the Insecticidal Activity of Insect Chitinase. *Insect Biochem. Mol. Biol.* 34 (10), 1037–1050. doi:10.1016/j.ibmb.2004.06.012

Fujimoto, H., Itoh, K., Yamamoto, M., Kyozuka, J., and Shimamoto, K. (1993). Insect Resistant Rice Generated by Introduction of a Modified δ-endotoxin Gene of Bacillus Thuringiensis. *Nat. Biotechnol.* 11 (10), 1151–1155. doi:10.1038/nbt1093-1151

Gao, C. (2021). Genome Engineering for Crop Improvement and Future Agriculture. *Cell* 184 (6), 1621–1635. doi:10.1016/j.cell.2021.01.005

Gatehouse, A. M. R., Dewey, F. M., Dove, J., Fenton, K. A., and Pusztai, A. (1984). Effect of Seed Lectins fromPhaseolus Vulgaris on the Development of Larvae ofCallosobruchus Maculatus; Mechanism of Toxicity. *J. Sci. Food Agric.* 35 (4), 373–380. doi:10.1002/jsfa.2740350402

Gatehouse, J. A. (2013). "Genetic Engineering of Crops for Insect Resistance," in *Sustainable Food Production*. Editors P. Christou, R. Savin, B. A. Costa-Pierce, I. Misztal, and C. B. A. Whitelaw (New York: Springer).

Gatehouse, J. A. (2011). Prospects for Using Proteinase Inhibitors to Protect Transgenic Plants against Attack by Herbivorous Insects. *Curr. Protein Pept. Sci.* 12 (5), 409–416. doi:10.2174/138920311796391142

Ghareyazie, B., Alinia, F., Menguito, C. A., Rubia, L. G., de Palma, J. M., Liwanag, E. A., et al. (1997). Enhanced Resistance to Two Stem Borers in an Aromatic Rice Containing a Synthetic cryIA (B) Gene. *Mol. Breed.* 3 (5), 401–414. doi:10.1023/a:1009695324100

Ghosh, G., Ganguly, S., Purohit, A., Chaudhuri, R. K., Das, S., and Chakraborti, D. (2017). Transgenic Pigeonpea Events Expressing Cry1Ac and Cry2Aa Exhibit Resistance to *Helicoverpa Armigera*. *Plant Cell Rep.* 36 (7), 1037–1051. doi:10.1007/s00299-017-2133-0

Gilles, A. F., Schinko, J. B., and Averof, M. (2015). Efficient CRISPR-Mediated Gene Targeting and Transgene Replacement in the beetleTribolium Castaneum. *Development* 142 (16), 2832–2839. doi:10.1242/dev.125054

Gomes, A. D. P. G., Dias, S. C., Bloch, C., Jr, Melo, F. R., Furtado, J. R., Jr, Monnerat, R. G., et al. (2005). Toxicity to Cotton Boll Weevil *Anthonomus grandis* of a Trypsin Inhibitor from Chickpea Seeds. *Comp. Biochem. Physiology Part B Biochem. Mol. Biol.* 140 (2), 313–319. doi:10.1016/j.cbpc.2004.10.013

Gowda, A., Rydel, T. J., Wollacott, A. M., Brown, R. S., Akbar, W., Clark, T. L., et al. (2016). A Transgenic Approach for Controlling *Lygus* in Cotton. *Nat. Commun.* 7 (1), 12213. doi:10.1038/ncomms12213

Grazziotin, M. A. G. D., Cabral, G. B., Ibrahim, A. B., Machado, R. B., and Aragão, F. J. L. (2020). Expression of the Arcelin 1 Gene from *Phaseolus vulgaris* L. In Cowpea Seeds (*Vigna Unguiculata* L.) Confers Bruchid Resistance. *Ann. Appl. Biol.* 176, 268–274. doi:10.1111/aab.12568

Gunasekara, J., Jayasekera, G., Perera, K., and Wickramasuriya, A. (2017). Development of a Sri Lankan Rice Variety Bg 94-1 Harbouring cry2A Gene of *Bacillus Thuringiensis* Resistant to Rice Leaffolder [*Cnaphalocrocis Medinalis* (Guenée)]. *J. Natn. Sci. Found. Sri Lanka* 45 (2), 143. doi:10.4038/jnsfsr.v45i2.8180

Guo, Z., Li, Y., and Ding, S.-W. (2019). Small RNA-Based Antimicrobial Immunity. *Nat. Rev. Immunol.* 19 (1), 31–44. doi:10.1038/s41577-018-0071-x

Gupta, M., Kumar, H., and Kaur, S. (2021). Vegetative Insecticidal Protein (Vip): A Potential Contender from *Bacillus Thuringiensis* for Efficient Management of Various Detrimental Agricultural Pests. *Front. Microbiol.* 12, 1139. doi:10.3389/fmicb.2021.659736

Haq, S. K., Atif, S. M., and Khan, R. H. (2004). Protein Proteinase Inhibitor Genes in Combat against Insects, Pests, and Pathogens: Natural and Engineered Phytoprotection. *Archives Biochem. Biophysics* 431 (1), 145–159. doi:10.1016/j.abb.2004.07.022

He, C., Xiao, Y., Yu, J., Li, J., Meng, Q., Qing, X., et al. (2019). Pyramiding *Xa21, Bph14*, and *Bph15* Genes into the Elite Restorer Line Yuehui9113 Increases Resistance to Bacterial Blight and the Brown Planthopper in Rice. *Crop Prot.* 115, 31–39. doi:10.1016/j.cropro.2018.09.001

Hilda, K., Bhuvaragavan, S., Kamatchi, R., Meenakumari, M., and Janarthanan, S. (2022). Cloning, Expression and Characterization of Arcelin and its Impact on Digestive Enzymes of the Stored Product Insect Pest, *Callosobruchus Maculatus* (F.). *Pesticide Biochem. Physiology* 180, 104982. doi:10.1016/j.pestbp.2021.104982

Horgan, F. G., Garcia, C. P. F., Haverkort, F., de Jong, P. W., and Ferrater, J. B. (2020). Changes in Insecticide Resistance and Host Range Performance of Planthoppers Artificially Selected to Feed on Resistant Rice. *Crop Prot.* 127, 104963. doi:10.1016/j.cropro.2019.104963

Hu, J., Li, X., Wu, C., Yang, C., Hua, H., Gao, G., et al. (2012). Pyramiding and Evaluation of the Brown Planthopper Resistance Genes *Bph14* and *Bph15* in Hybrid Rice. *Mol. Breed.* 29 (1), 61–69. doi:10.1007/s11032-010-9526-x

Huang, Y., Chen, Y., Zeng, B., Wang, Y., James, A. A., Gurr, G. M., et al. (2016). CRISPR/Cas9 Mediated Knockout of the Abdominal-A Homeotic Gene in the Global Pest, Diamondback Moth (*Plutella Xylostella*). *Insect Biochem. Mol. Biol.* 75, 98–106. doi:10.1016/j.ibmb.2016.06.004

Hussain, T., Aksoy, E., Çalışkan, M. E., and Bakhsh, A. (2019). Transgenic Potato Lines Expressing Hairpin RNAi Construct of Molting-Associated EcR Gene Exhibit Enhanced Resistance against Colorado Potato Beetle (*Leptinotarsa decemlineata*, Say). *Transgenic Res.* 28 (1), 151–164. doi:10.1007/s11248-018-0109-7

Iqbal, M. A., Hassan, M. W., Hafiz, M. U. R., Muhammad, A., and Moazzam, J. (2018). Evaluation of Different Wheat Varieties for Resistance against Aphid, *Schizaphis graminum* R. (Homoptera: Aphididae) under Laboratory Conditions. *Asian J. Agric. Biol.* 6, 549–555.

Irving, J. A., Pike, R. N., Dai, W., Brömme, D., Worrall, D. M., Silverman, G. A., et al. (2002). Evidence that Serpin Architecture Intrinsically Supports Papain-like Cysteine Protease Inhibition: Engineering α1-antitrypsin to Inhibit Cathepsin Proteases. *Biochemistry* 41 (15), 4998–5004. doi:10.1021/bi0159985

Ishimoto, M., and Kitamura, K. (1989). Growth Inhibitory Effects of an α-Amylase Inhibitor from the Kidney Bean, *Phaseolus vulgaris* (L.) on Three Species of

Bruchids (Coleoptera: Bruchidae). *Appl. Entomol. Zool.* 24 (3), 281–286. doi:10.1303/aez.24.281

Jadhav, M. S., Rathnasamy, S. A., Natarajan, B., Duraialagaraja, S., and Varatharajalu, U. (2020). Study of Expression of Indigenous *Bt* Cry2ax1 Gene in T_3 Progeny of Cotton and its Efficacy against *Helicoverpa Armigera* (Hubner). *Braz. Archieves Biol. Technol.* 63 (2), 2020. doi:10.1590/1678-4324-2020180428

Jaiswal, D. K., Raju, S. V. S., Kumar, G. S., Sharma, K. R., Singh, D. K., and Vennela, P. R. (2018). Biotechnology in Plant Resistance to Insects. *Indian J. Agric. Allied Sci.* 4, 7–18.

Jin, L., Wang, J., Guan, F., Zhang, J., Yu, S., Liu, S., et al. (2018). Dominant Point Mutation in a Tetraspanin Gene Associated with Field-Evolved Resistance of Cotton Bollworm to Transgenic *Bt* Cotton. *Proc. Natl. Acad. Sci. U.S.A.* 115 (46), 11760–11765. doi:10.1073/pnas.1812138115

Joshi, M. J., Prithiv Raj, V., and Solanki, C. B. (2020). Role of Biotechnology in Insect-Pests Management. *Agric. Food e-Newsletter* 2, 574–576.

Juturu, V. N., Mekala, G. K., and Kirti, P. B. (2015). Current Status of Tissue Culture and Genetic Transformation Research in Cotton (*Gossypium* spp.). *Plant Cell Tiss. Organ Cult.* 120 (3), 813–839. doi:10.1007/s11240-014-0640-z

Kamthan, A., Chaudhuri, A., Kamthan, M., and Datta, A. (2015). Small RNAs in Plants: Recent Development and Application for Crop Improvement. *Front. Plant Sci.* 06, 208. doi:10.3389/fpls.2015.00208

Katoch, R., Sethi, A., Thakur, N., and Murdock, L. L. (2013). RNAi for Insect Control: Current Perspective and Future Challenges. *Appl. Biochem. Biotechnol.* 171 (4), 847–873. doi:10.1007/s12010-013-0399-4

Katta, S., Talakayala, A., Reddy, M. K., Addepally, U., and Garladinne, M. (2020). Development of Transgenic Cotton (Narasimha) Using Triple Gene *Cry2Ab-Cry1F-Cry1Ac* Construct Conferring Resistance to Lepidopteran Pest. *J. Biosci.* 45 (1), 1–11. doi:10.1007/s12038-020-0006-0

Kaur, S., Kaur, K., and Jindal, J. (2022). Status of Phenolic Metabolism and α-amylase Inhibitor in Maize under *Chilo Partellus* Infestation. *Cereal Res. Commun.*, 1–9. doi:10.1007/s42976-021-00230-5

Khajuria, C., Vélez, A. M., Rangasamy, M., Wang, H., Fishilevich, E., Frey, M. L. F., et al. (2015). Parental RNA Interference of Genes Involved in Embryonic Development of the Western Corn Rootworm, *Diabrotica Virgifera Virgifera* LeConte. *Insect Biochem. Mol. Biol.* 63, 54–62. doi:10.1016/j.ibmb.2015.05.011

Khan, G. A., Bakhsh, A., Ghazanfar, M., Riazuddin, S., and Husnain, T. (2013). Development of Transgenic Cotton Lines Harboring a Pesticidal Gene (cry1Ab). *Emir. J. Food Agric.* 25 (6), 434. doi:10.9755/ejfa.v25i6.13133

Khan, G. A., Bakhsh, A., Riazuddin, S., and Husnain, T. (2011). Introduction of *cry1Ab* Gene into Cotton (*Gossypium Hirsutum*) Enhances Resistance against Lepidopteran Pest (*Helicoverpa Armigera*). *Span. J. Agric. Res.* 9, 296–302. doi:10.5424/sjar/20110901-136-10

Khoobdel, M., Rahimi, V., Ebadollahi, A., and Krutmuang, P. (2022). Evaluation of the Potential of a Lectin Extracted from *Polygonumpersicaria* L. As a Biorational Agent against *Sitophilusoryzae* L. *Molecules* 27 (3), 793. doi:10.3390/molecules27030793

Kim, S.-Y., Bengtsson, T., Olsson, N., Hot, V., Zhu, L.-H., and Åhman, I. (2020). Mutations in Two Aphid-Regulated β-1,3-Glucanase Genes by CRISPR/Cas9 Do Not Increase Barley Resistance to Rhopalosiphum Padi L. *Front. Plant Sci.* 11, 1043. doi:10.3389/fpls.2020.01043

Koch, K. G., Chapman, K., Louis, J., Heng-Moss, T., and Sarath, G. (2016). Plant Tolerance: a Unique Approach to Control Hemipteran Pests. *Front. Plant Sci.* 7, 1363. doi:10.3389/fpls.2016.01363

Koch, M. S., Ward, J. M., Levine, S. L., Baum, J. A., Vicini, J. L., and Hammond, B. G. (2015). The Food and Environmental Safety of *Bt* Crops. *Front. Plant Sci.* 06, 283. doi:10.3389/fpls.2015.00283

Koerniati, S., Sukmadjaja, D., and Samudra, I. M. (2020). C Synthetic Gene of CryIAb-CryIAc Fusion to Generate Resistant Sugarcane to Shoot or Stem Borer. *IOP Conf. Ser. Earth Environ. Sci.* 418 (1), 012069. doi:10.1088/1755-1315/418/1/012069

Kola, V. S. R., Pichili, R., Padmakumari, A. P., Mangrauthia, S. K., Balachandran, S. M., and Madhav, M. S. (2019). Knockdown of Acetylcholinesterase (AChE) Gene in Rice Yellow Stem Borer, *Scirpophaga Incertulas* (Walker) through RNA Interference. *Agri Gene* 11, 100081. doi:10.1016/j.aggene.2019.100081

Koutroumpa, F. A., Monsempes, C., François, M. C., de Cian, A., Royer, C., Concordet, J. P., et al. (2016). Heritable Genome Editing with CRISPR/Cas9 Induces Anosmia in a Crop Pest Moth. *Sci. Rep.* 6 (1), 29620–29629. doi:10.1038/srep29620

Larsson, M. C., Domingos, A. I., Jones, W. D., Chiappe, M. E., Amrein, H., and Vosshall, L. B. (2004). Or83b Encodes a Broadly Expressed Odorant Receptor Essential for Drosophila Olfaction. *Neuron* 43 (5), 703–714. doi:10.1016/j.neuron.2004.08.019

Li-Hong, S. U. N., Chun-Ming, W. A. N. G., Chang-Chao, S. U., Yu-Qiang, L. I. U., Hu-Qu, Z. H. A. I., and Jian-Min, W. A. N. (2006). Mapping and Marker-Assisted Selection of a Brown Planthopper Resistance Gene *bph2* in rice (*Oryza sativa* L.). *Acta Genetica Sinica* 33 (8), 717–723.

Li, H., Guan, R., Guo, H., and Miao, X. (2015). New Insights into an RNAi Approach for Plant Defence against Piercing-Sucking and Stem-Borer Insect Pests. *Plant Cell Environ.* 38 (11), 2277–2285. doi:10.1111/pce.12546

Li, J.-F., Norville, J. E., Aach, J., McCormack, M., Zhang, D., Bush, J., et al. (2013). Multiplex and Homologous Recombination-Mediated Genome Editing in Arabidopsis and Nicotiana Benthamiana Using Guide RNA and Cas9. *Nat. Biotechnol.* 31 (8), 688–691. doi:10.1038/nbt.2654

Li, Y., Zhang, J., Chen, D., Yang, P., Jiang, F., Wang, X., et al. (2016). CRISPR/Cas9 in Locusts: Successful Establishment of an Olfactory Deficiency Line by Targeting the Mutagenesis of an Odorant Receptor Co-receptor (Orco). *Insect Biochem. Mol. Biol.* 79, 27–35. doi:10.1016/j.ibmb.2016.10.003

Liu, Y., Chen, L., Liu, Y., Dai, H., He, J., Kang, H., et al. (2016). Marker Assisted Pyramiding of Two Brown Planthopper Resistance Genes, *Bph3* and *Bph27* (T), into Elite Rice Cultivars. *Rice* 9 (1), 1–7. doi:10.1186/s12284-016-0096-3

Lu, H.-P., Luo, T., Fu, H.-W., Wang, L., Tan, Y.-Y., Huang, J.-z., et al. (2018). Resistance of Rice to Insect Pests Mediated by Suppression of Serotonin Biosynthesis. *Nat. Plants* 4 (6), 338–344. doi:10.1038/s41477-018-0152-7

Macedo, M. L., de Sá, C. M., Freire, M. D., and Parra, J. R. (2004). A Kunitz-type Inhibitor of Coleopteran Proteases, Isolated from *Adenanthera pavonina* L. Seeds and its Effect on *Callosobruchus Maculatus*. *J. Agric. Food Chem.* 52 (9), 2533–2540. doi:10.1021/jf035389z

Macedo, M. R., and Freire, M. D. G. (2011). Insect Digestive Enzymes as a Target for Pest Control. *Invertebr. Surviv. J.* 8 (2), 190–198.

Malaikozhundan, B., Suresh, P., Seshadri, S., and Janarthanan, S. (2003). Toxicity Assessment of Wild Bean Seed Protein-Arcelin on Asian Armyworm, *Spodoptera Litura* (Fabricius). *Indian J. Exp. Biol.* 41 (12), 1463–1465.

Malone, L. A., Barraclough, E. I., Lin-Wang, K., Stevenson, D. E., and Allan, A. C. (2009). Effects of Red-Leaved Transgenic Tobacco Expressing a MYB Transcription Factor on Two Herbivorous insects,Spodoptera lituraandHelicoverpa Armigera. *Entomologia Exp. Appl. Entomol. Exp. Appl* 133 (2), 117–127. doi:10.1111/j.1570-7458.2009.00910.x

Mamta, B., and Rajam, M. V. (2017). RNAi Technology: a New Platform for Crop Pest Control. *Physiol. Mol. Biol. Plants* 23 (3), 487–501. doi:10.1007/s12298-017-0443-x

Mamta, Reddy, K. R. K., and Rajam, M. V. (2016). Targeting Chitinase Gene of *Helicoverpa Armigera* by Host-Induced RNA Interference Confers Insect Resistance in Tobacco and Tomato. *Plant Mol. Biol.* 90 (3), 281–292. doi:10.1007/s11103-015-0414-y

Maqbool, S. B., Riazuddin, S., Loc, N. T., Gatehouse, A. M. R., Gatehouse, J. A., and Christou, P. (2001). Expression of Multiple Insecticidal Genes Confers Broad Resistance against a Range of Different Rice Pests. *Mol. Breed.* 7 (1), 85–93. doi:10.1023/a:1009644712157

Mateos Fernández, R., Petek, M., Gerasymenko, I., Juteršek, M., Baebler, Š., Kallam, K., et al. (2021). Insect Pest Management in the Age of Synthetic Biology. *Plant Biotechnol. J.* 20 (1), 25–36. doi:10.1111/pbi.13685

Meade, T., Narva, K., Storer, N. P., Sheets, J. J., Burton, S. L., and Woosley, A. T. (2017). *U.S. Patent No. 9,556,453*. Washington, DC: U.S. Patent and Trademark Office.

Meccariello, A., Monti, S. M., Romanelli, A., Colonna, R., Primo, P., Inghilterra, M. G., et al. (2017). Highly Efficient DNA-free Gene Disruption in the Agricultural Pest *Ceratitis Capitata* by CRISPR-Cas9 Ribonucleoprotein Complexes. *Sci. Rep.* 7, 10061. doi:10.1038/s41598-017-10347-5

Meng, F., Li, Y., Zang, Z., Li, N., Ran, R., Cao, Y., et al. (2017). Expression of the Double-Stranded RNA of the Soybean Pod Borer *Leguminivora Glycinivorella* (Lepidoptera: Tortricidae) Ribosomal Protein Po Gene Enhances the Resistance of Transgenic Soybean Plants. *Pest. Manag. Sci.* 73 (12), 2447–2455. doi:10.1002/ps.4637

Mi, X., Liu, X., Yan, H., Liang, L., Zhou, X., Yang, J., et al. (2017). Expression of the *Galanthus Nivalis Agglutinin* (GNA) Gene in Transgenic Potato Plants Confers Resistance to Aphids. *Comptes Rendus Biol.* 340 (1), 7–12. doi:10.1016/j.crvi.2016.10.003

Ming, H., Li, B., Zhou, L., Goel, A., and Huang, C. (2021). Long Non-coding RNAs and Cancer Metastasis: Molecular Basis and Therapeutic Implications. *Biochimica Biophysica Acta (BBA) - Rev. Cancer* 1875 (2), 188519. doi:10.1016/j.bbcan.2021.188519

Morton, R. L., Schroeder, H. E., Bateman, K. S., Chrispeels, M. J., Armstrong, E., and Higgins, T. J. V. (2000). Bean α-amylase Inhibitor 1 in Transgenic Peas (Pisum Sativum) Provides Complete Protection from Pea Weevil (Bruchus Pisorum) under Field Conditions. *Proc. Natl. Acad. Sci. U.S.A.* 97 (8), 3820–3825. doi:10.1073/pnas.070054597

Muddanuru, T., Polumetla, A. K., Maddukuri, L., and Mulpuri, S. (2019). Development and Evaluation of Transgenic castor (*Ricinus communis* L.) Expressing the Insecticidal Protein Cry1Aa of *Bacillus Thuringiensis* against Lepidopteran Insect Pests. *Crop Prot.* 119, 113–125. doi:10.1016/j.cropro.2019.01.016

N'guessan, F. K., Quisenberry, S. S., and Croughan, T. P. (1994). Evaluation of Rice Anther Culture Lines for Tolerance to the Rice Water Weevil (Coleoptera: Curculionidae). *Environ. Entomol.* 23 (2), 331–336. doi:10.1093/ee/23.2.331

Nakasu, E. Y. T., Edwards, M. G., Fitches, E., Gatehouse, J. A., and Gatehouse, A. M. R. (2014). Transgenic Plants Expressing ω-ACTX-Hv1a and Snowdrop Lectin (GNA) Fusion Protein Show Enhanced Resistance to Aphids. *Front. Plant Sci.* 5, 673. doi:10.3389/fpls.2014.00673

Napoleão, T. H., Albuquerque, L. P., Santos, N. D., Nova, I. C., Lima, T. A., Paiva, P. M., et al. (2019). Insect Midgut Structures and Molecules as Targets of Plant-derived Protease Inhibitors and Lectins. *Pest. Manag. Sci.* 75 (5), 1212–1222. doi:10.1002/ps.5233

Napoleão, T. H., Belmonte, B. D. R., Pontual, E. V., de Albuquerque, L. P., Sá, R. A., Paiva, L. M., et al. (2013). Deleterious Effects of *Myracrodruon Urundeuva* Leaf Extract and Lectin on the Maize Weevil, *Sitophilus Zeamais* (Coleoptera, Curculionidae). *J. stored Prod. Res.* 54, 26–33. doi:10.1016/j.jspr.2013.04.002

Nayak, P., Basu, D., Das, S., Basu, A., Ghosh, D., Ramakrishnan, N. A., et al. (1997). Transgenic Elite Indica Rice Plants Expressing CryIAcƋ-Endotoxin of *Bacillus Thuringiensis* Are Resistant against Yellow Stem Borer (*Scirpophaga Incertulas*). *Proc. Natl. Acad. Sci. U.S.A.* 94 (6), 2111–2116. doi:10.1073/pnas.94.6.2111

Nderitu, P. W., Jonsson, M., Arunga, E., Otieno, M., Muturi, J. J., and Wafula, G. O. (2020). Combining Host Plant Resistance, Selective Insecticides, and Biological Control Agents for Integrated Management of *Tuta Absoluta*. *Adv. Agric.* 2020, 1–8. doi:10.1155/2020/6239491

Ni, M., Ma, W., Wang, X., Gao, M., Dai, Y., Wei, X., et al. (2017). Next-generation Transgenic Cotton: Pyramiding RNAi and *Bt* Counters Insect Resistance. *Plant Biotechnol. J.* 15 (9), 1204–1213. doi:10.1111/pbi.12709

Niu, X., Kassa, A., Hu, X., Robeson, J., McMahon, M., Richtman, N. M., et al. (2017). Control of Western Corn Rootworm (*Diabrotica Virgifera Virgifera*) Reproduction through Plant-Mediated RNA Interference. *Sci. Rep.* 7 (1), 1–13. doi:10.1038/s41598-017-12638-3

Oerke, E. C., Dehne, H. W., Schönbeck, F., and Weber, A. (2006). *Crop Production and Crop Protection: Estimated Losses in Major Food and Cash Crops*. Netherlands. Elsevier Science B.V. doi:10.1017/S0021859605005708

Omar, M. A. A., Ao, Y., Li, M., He, K., Xu, L., Tong, H., et al. (2019). The Functional Difference of Eight Chitinase Genes between Male and Female of the Cotton Mealybug, *Phenacoccus solenopsis*. *Insect Mol. Biol.* 28 (4), 550–567. doi:10.1111/imb.12572

Oriani, M. A. D. G., and Lara, F. M. (2000). Antibiosis Effects of Wild Bean Lines Containing Arcelin on *Bemisia Tabaci* (Genn.) Biotype B (Homoptera: Aleyrodidae). *An. Soc. Entomol. Bras.* 29 (3), 573–582. doi:10.1590/S0301-80592000000300020

Osman, G. H., Assem, S. K., Alreedy, R. M., El-Ghareeb, D. K., Basry, M. A., Rastogi, A., et al. (2016). Erratum: Development of Insect Resistant Maize Plants Expressing a Chitinase Gene from the Cotton Leaf Worm, *Spodoptera Littoralis*. *Sci. Rep.* 6, 20449. doi:10.1038/srep20449

Outchkourov, N. S., De Kogel, W. J., Schuurman-de Bruin, A., Abrahamson, M., and Jongsma, M. A. (2004). Specific Cysteine Protease Inhibitors Act as Deterrents of Western Flower Thrips, *Frankliniella Occidentalis* (Pergande), in Transgenic Potato. *J. Plant Biotechnol.* 2 (5), 439–448. doi:10.1111/j.1467-7652.2004.00088.x

Oyeleye, A., and Normi, Y. M. (2018). Chitinase: Diversity, Limitations, and Trends in Engineering for Suitable Applications. *Biosci. Rep.* 38 (4), 1–21. doi:10.1042/BSR20180323

Painter, R. H. (1951). Insect Resistance in Crop Plants. *Soil Sci.* 72 (6), 481. LWW. doi:10.1097/00010694-195112000-00015

Paiva, P. M., Pontual, E. V., Napoleão, T. H., and Coelho, L. C. B. B. (2012). "Effects of Plant Lectins and Trypsin Inhibitors on Development, Morphology and Biochemistry of Insect Larvae," in *Larvae: Morphology, Biology and Life Cycle* (New York, NY: Science Publishers), 3, 37–55.

Palma, L., Muñoz, D., Berry, C., Murillo, J., and Caballero, P. (2014). *Bacillus Thuringiensis* Toxins: an Overview of Their Biocidal Activity. *Toxins* 6 (12), 3296–3325. doi:10.3390/toxins6123296

Pan, Y., Lü, P., Wang, Y., Yin, L., Ma, H., Ma, G., et al. (2012). In silicoIdentification of Novel Chitinase-like Proteins in the Silkworm,*Bombyx mori*, Genome. *J. Insect Sci.* 12, 1–14. doi:10.1673/031.012.15001

Panwar, B. S., Ram, C., Narula, R. K., and Kaur, S. (2018). Pool Deconvolution Approach for High-Throughput Gene Mining from *Bacillus Thuringiensis*. *Appl. Microbiol. Biotechnol.* 102 (3), 1467–1482. doi:10.1007/s00253-017-8633-6

Park, H.-S., Baek, S.-H., Kim, W.-J., Jeung, J.-U., Lee, J.-H., Ha, K.-Y., et al. (2014). Development of Multi-Resistant Lines to Brown Planthopper, Bacterial Blight, and Rice Stripe Virus Using Anther Culture in Rice. *Korean J. Breed. Sci.* 46 (1), 78–89. doi:10.9787/kjbs.2014.46.1.078

Paul, S., and Das, S. (2020). Natural Insecticidal Proteins, the Promising Bio-Control Compounds for Future Crop Protection. *Nucleus* 64 (1), 7–20. doi:10.1007/s13237-020-00316-1

Pedigo, L. P., and Rice, M. E. (2006). *Economic Decision Levels for Pest Populations*. Columbus, OH: Entomology Pest Management Pearson Prentice Hall, 253–284.

Pereira de Albuquerque, L., Maria de Sá Santana, G., Pontual, E. V., Napoleão, T. H., Breitenbach Barroso Coelho, L. C., and Paiva, P. M. G. (2012). Effect of *Microgramma Vaccinifolia* Rhizome Lectin on Survival and Digestive Enzymes of *Nasutitermes Corniger* (Isoptera, Termitidae). *Int. Biodeterior. Biodegrad.* 75, 158–166. doi:10.1016/j.ibiod.2012.06.030

Perez, C. J., and Shelton, A. M. (1997). Resistance of Plutella Xylostella (Lepidoptera: Plutellidae) to *Bacillus Thuringiensis* Berliner in Central America. *J. Econ. Entomology* 90 (1), 87–93. doi:10.1093/jee/90.1.87

Persley, G. J. (2000). *Agricultural Biotechnology and the Poor: Promethean Science*. Washington, D.C.: Consultative Group on International Agricultural Research (CGIAR), 3–21.

Poulsen, M., Kroghsbo, S., Schrøder, M., Wilcks, A., Jacobsen, H., Miller, A., et al. (2007). A 90-day Safety Study in Wistar Rats Fed Genetically Modified Rice Expressing Snowdrop Lectin *Galanthus Nivalis* (GNA). *Food Chem. Toxicol.* 45 (3), 350–363. doi:10.1016/j.fct.2006.09.002

Powell, K. S., Gatehouse, A. M. R., Hilder, V. A., and Gatehouse, J. A. (1993). Antimetabolic Effects of Plant Lectins and Plant and Fungal Enzymes on the Nymphal Stages of Two Important Rice Pests, *Nilaparvata Lugens* and *Nephotettix Cinciteps*. *Entomologia Exp. Appl.* 66 (2), 119–126. doi:10.1111/j.1570-7458.1993.tb00699.x

Priya, S., Kaur, N., and Gupta, A. K. (2010). Purification, Characterization and Inhibition Studies of α-amylase of Rhyzopertha dominica. *Pesticide Biochem. Physiology* 98 (2), 231–237. doi:10.1016/j.pestbp.2010.06.012

Qin, D., Liu, X.-Y., Miceli, C., Zhang, Q., and Wang, P.-w. (2019). Soybean Plants Expressing the *Bacillus Thuringiensis* Cry8-like Gene Show Resistance to *Holotrichia Parallela*. *BMC Biotechnol.* 19 (1), 1–12. doi:10.1186/s12896-019-0563-1

Qiu, L., Fan, J., Zhang, B., Liu, L., Wang, X., Lei, C., et al. (2017). RNA Interference Knockdown of Aminopeptidase N Genes Decrease the Susceptibility of *Chilo Suppressalis* Larvae to Cry1Ab/Cry1Ac and Cry1Ca-Expressing Transgenic Rice. *J. Invertebr. Pathology* 145, 9–12. doi:10.1016/j.jip.2017.03.001

Qiu, L., Sun, Y., Jiang, Z., Yang, P., Liu, H., Zhou, H., et al. (2019). The Midgut V-ATPase Subunit A Gene Is Associated with Toxicity to Crystal 2Aa and Crystal 1Ca -expressing Transgenic Rice in Chilo Suppressalis. *Insect Mol. Biol.* 28 (4), 520–527. doi:10.1111/imb.12570

Rajadurai, G., Kalaivani, A., Varanavasiyappan, S., Balakrishnan, N., Udayasuriyan, V., Sudhakar, D., et al. (2018). Generation of Insect Resistant

Marker-free Transgenic Rice with a Novel *cry2AX1* Gene. *Electron. Journ. Plan. Breed.* 9 (2), 723–732. doi:10.5958/0975-928x.2018.00086.8

Rakha, M., Bouba, N., Ramasamy, S., Regnard, J.-L., and Hanson, P. (2017). Evaluation of Wild Tomato Accessions (*Solanum* spp.) for Resistance to Two-Spotted Spider Mite (*Tetranychus Urticae* Koch) Based on Trichome Type and Acylsugar Content. *Genet. Resour. Crop Evol.* 64 (5), 1011–1022. doi:10.1007/s10722-016-0421-0

Rani, S., Sharma, V., Hada, A., Bhattacharya, R. C., and Koundal, K. R. (2017b). Fusion Gene Construct Preparation with Lectin and Protease Inhibitor Genes against Aphids and Efficient Genetic Transformation of Brassica Juncea Using Cotyledons Explants. *Acta Physiol. Plant* 39 (5), 1–13. doi:10.1007/s11738-017-2415-8

Rani, S., Sharma, V., Hada, A., and Koundal, K. R. (2017a). Efficient Genetic Transformation of *Brassica Juncea* with Lectin Using Cotyledons Explants. *Int. J. Adv. Biotechnol.* 7, 1–12.

Riaz, S., Nasir, I. A., Bhatti, M. U., Adeyinka, O. S., Toufiq, N., Yousaf, I., et al. (2020). Resistance to *Chilo Infuscatellus* (Lepidoptera: Pyraloidea) in Transgenic Lines of Sugarcane Expressing *Bacillus Thuringiensis* Derived Vip3A Protein. *Mol. Biol. Rep.* 47 (4), 2649–2658. doi:10.1007/s11033-020-05355-0

Ribeiro, A. P. O., Pereira, E. J. G., Galvan, T. L., Picanco, M. C., Picoli, E. A. T., Silva, D. J. H., et al. (2006). Effect of Eggplant Transformed with *Oryzacystatin* Gene on *Myzus persicae* and *Macrosiphum euphorbiae*. *J. Appl. Entomol.* 130 (2), 84–90. doi:10.1111/j.1439-0418.2005.01021.x

Ribeiro, T. P., Arraes, F. B. M., Lourenço-Tessutti, I. T., Silva, M. S., Lisei-de-Sá, M. E., Lucena, W. A., et al. (2017). Transgenic Cotton Expressing Cry10Aa Toxin Confers High Resistance to the Cotton Boll Weevil. *Plant Biotechnol. J.* 15 (8), 997–1009. doi:10.1111/pbi.12694

Ricroch, A., Clairand, P., and Harwood, W. (2017). Use of CRISPR Systems in Plant Genome Editing: toward New Opportunities in Agriculture. *Emerg. Top. Life Sci.* 1 (2), 169–182. doi:10.1042/etls20170085

Sakthi, A. R., Naveenkumar, A., Deepikha, P. S., Balakrishnan, N., Kumar, K. K., Devi, E. K., et al. (2015). Expression and Inheritance of Chimeric *cry2AX1* Gene in Transgenic Cotton Plants Generated through Somatic Embryogenesis. *Vitro Cell.Dev.Biol.-Plant* 51 (4), 379–389. doi:10.1007/s11627-015-9695-8

Santos, N. D. D. L., de Moura, K. S., Napoleão, T. H., Santos, G. K. N., Coelho, L. C. B. B., Navarro, D. M. D. A. F., et al. (2012). Oviposition-stimulant and Ovicidal Activities of *Moringa Oleifera* Lectin on *Aedes aegypti*. *PLoS ONE* 7, e44840. doi:10.1371/journal.pone.0044840

Santos, N. D. L., Napoleão, T. H., Benevides, C. A., Albuquerque, L. P., Pontual, E. V., Oliveira, A. P. S., et al. (2020). Effect of Gamma Irradiation of *Moringa Oleifera* Seed Lectin on its Larvicidal, Ovicidal, and Oviposition-Stimulant Activities against *Aedes aegypti*. *South Afr. J. Bot.* 129, 3–8. doi:10.1016/j.sajb.2018.05.001

Sauvion, N., Rahbé, Y., Peumans, W. J., Damme, E. J. M., Gatehouse, J. A., and Gatehouse, A. M. R. (1996). Effects of GNA and Other Mannose Binding Lectins on Development and Fecundity of the Peach-Potato Aphid *Myzus persicae*. *Entomologia Exp. Appl.* 79 (3), 285–293. doi:10.1111/j.1570-7458.1996.tb00836.x

Sawardekar, S. V., Katageri, I. S., Salimath, P. M., Kumar, P. A., and Kelkar, V. G. (2017). Standardization of *In-Vitro* Genetic Transformation Technique in Chickpea (*Cicer Arietinum* L.) for Pod-Borer Resistance. *Adv. Agric. Res. Technol.* 1 (2), 198.

Schneider, V. K., Soares-Costa, A., Chakravarthi, M., Ribeiro, C., Chabregas, S. M., Falco, M. C., et al. (2017). Transgenic Sugarcane Overexpressing CaneCPI-1 Negatively Affects the Growth and Development of the Sugarcane Weevil *Sphenophorus Levis*. *Plant Cell Rep.* 36 (1), 193–201. doi:10.1007/s00299-016-2071-2

Selale, H., Dağlı, F., Mutlu, N., Doğanlar, S., and Frary, A. (2017). Cry1Ac-mediated Resistance to Tomato Leaf Miner (*Tuta Absoluta*) in Tomato. *Plant Cell Tiss. Organ Cult.* 131 (1), 65–73. doi:10.1007/s11240-017-1262-z

Sétamou, M., Bernal, J. S., Legaspi, J. C., Mirkov, T. E., and Legaspi, B. C., Jr (2002). Evaluation of Lectin-Expressing Transgenic Sugarcane against Stalkborers (Lepidoptera: Pyralidae): Effects on Life History Parameters. *J. Econ. Entomol.* 95 (2), 469–477. doi:10.1603/0022-0493-95.2.469

Shabanimofrad, M., Yusop, M. R., Ashkani, S., Musa, M. H., Adam, N. A., Haifa, I., et al. (2015). Marker-assisted Selection for Rice Brown Planthopper (*Nilaparvata Lugens*) Resistance Using Linked SSR Markers. *Turk J. Biol.* 39 (5), 666–673. doi:10.3906/biy-1406-78

Shade, R. E., Schroeder, H. E., Pueyo, J. J., Tabe, L. M., Murdock, L. L., Higgins, T. J. V., et al. (1994). Transgenic Pea Seeds Expressing the α-Amylase Inhibitor of the Common Bean Are Resistant to Bruchid Beetles. *Nat. Biotechnol.* 12 (8), 793–796. doi:10.1038/nbt0894-793

Shaffer, L. (2020). Inner Workings: RNA-Based Pesticides Aim to Get Around Resistance Problems. *Proc. Natl. Acad. Sci. U.S.A.* 117 (52), 32823–32826. doi:10.1073/pnas.2024033117

Sharma, H. C. (2007). Host Plant Resistance to Insects: Modern Approaches and Limitations. *Industrial J. Plant Prot.* 35 (2), 179–184.

Sharma, K. (2015). Protease Inhibitors in Crop Protection from Insects. *Int. J. Curr. Res. Acad. Rev.* 3, 55–70. doi:10.12983/ijsres-2015-p0411-0419

Sharma, P. N., Torii, A., Takumi, S., Mori, N., and Nakamura, C. (2004). Marker-assisted Pyramiding of Brown Planthopper (*Nilaparvata Lugens* Stål) Resistance Genes *Bph1* and *Bph2* on Rice Chromosome 12. *Hereditas* 140 (1), 61–69. doi:10.1111/j.1601-5223.2004.01726.x

Siddiqui, H. A., Asif, M., Asad, S., Naqvi, R. Z., Ajaz, S., Umer, N., et al. (2019). Development and Evaluation of Double Gene Transgenic Cotton Lines Expressing Cry Toxins for Protection against Chewing Insect Pests. *Sci. Rep.* 9 (1), 1–7. doi:10.1038/s41598-019-48188-z

Singh, A., Singh, S., and Singh, I. K. (2016). Recent Insights into the Molecular Mechanism of Jasmonate Signaling during Insect-Plant Interaction. *Australas. Plant Pathol.* 45 (2), 123–133. doi:10.1007/s13313-015-0392-1

Singh, S., Kumar, N. R., Maniraj, R., Lakshmikanth, R., Rao, K. Y. S., Muralimohan, N., et al. (2018). Expression of Cry2Aa, a *Bacillus Thuringiensis* Insecticidal Protein in Transgenic Pigeon Pea Confers Resistance to Gram Pod Borer, *Helicoverpa Armigera*. *Sci. Rep.* 8 (1), 1–12. doi:10.1038/s41598-018-26358-9

Singh, S., Singh, A., Kumar, S., Mittal, P., and Singh, I. K. (2020). Protease Inhibitors: Recent Advancement in its Usage as a Potential Biocontrol Agent for Insect Pest Management. *Insect Sci.* 27 (2), 186–201. doi:10.1111/1744-7917.12641

Smith, C. M. (Editor) (2005). *Plant Resistance to Arthropods: Molecular and Conventional Approaches* (Dordrecht: Springer Netherlands).

Soliman, H. I. A., Abo-El-Hasan, F. M., El-Seedy, A. S., and Mabrouk, Y. M. (2021). *Agrobacterium*-mediated Transformation of Tomato (*Lycopersicon esculentum* mill.) Using a Synthetic Cry1ab Gene for Enhanced Resistance against *Tuta Absoluta* (Meyrick). *J. Microbiol. Biotechnol. Food Sci.* 7, 67–74. doi:10.15414/jmbfs.2017.7.1.67-74

Stevens, J., Dunse, K., Fox, J., Evans, S., and Anderso, M. (2012). "Biotechnological Approaches for the Control of Insect Pests in Crop Plants," in *Pesticides-Advances in Chemical and Botanical Pesticides*. Editor R. P. Soundararajan (InTech Open), 269–308. doi:10.5772/46233

Stoger, E., Williams, S., Christou, P., Down, R. E., and Gatehouse, J. A. (1999). Expression of the Insecticidal Lectin from Snowdrop (*Galanthus Nivalis Agglutinin*; GNA) in Transgenic Wheat Plants: Effects on Predation by the Grain Aphid *Sitobion Avenae*. *Mol. Breed.* 5 (1), 65–73. doi:10.1023/a:1009616413886

Su, C., Tu, G., Huang, S., Yang, Q., Shahzad, M. F., and Li, F. (2016). Genome-wide Analysis of Chitinase Genes and Their Varied Functions in Larval Moult, Pupation and Eclosion in the Rice Striped Stem borer,Chilo Suppressalis. *Insect Mol. Biol.* 25 (4), 401–412. doi:10.1111/imb.12227

Sudhakar, D., Fu, X., Stoger, E., Williams, S., Spence, J., Brown, D. P., et al. (1998). Expression and Immunolocalisation of the Snowdrop Lectin, GNA in Transgenic Rice Plants. *Transgenic Res.* 7 (5), 371–378. doi:10.1023/a:1008856703464

Syed, T., Askari, M., Meng, Z., Li, Y., Abid, M. A., Wei, Y., et al. (2020). Current Insights on Vegetative Insecticidal Proteins (Vip) as Next Generation Pest Killers. *Toxins (Basel)* 12 (8), 522. doi:10.3390/toxins12080522

Tabashnik, B. E. (1994). Evolution of Resistance to *Bacillus Thuringiensis*. *Annu. Rev. Entomol.* 39 (1), 47–79. doi:10.1146/annurev.en.39.010194.000403

Tabashnik, B. E., Malvar, T., Liu, Y. B., Finson, N., Borthakur, D., Shin, B. S., et al. (1996). Cross-resistance of the Diamondback Moth Indicates Altered Interactions with Domain II of *Bacillus Thuringiensis* Toxins. *Appl. Environ. Microbiol.* 62 (8), 2839–2844. doi:10.1128/aem.62.8.2839-2844.1996

Talakayala, A., Katta, S., and Garladinne, M. (2020). Genetic Engineering of Crops for Insect Resistance: An Overview. *J. Biosci.* 45 (1), 1–12. doi:10.1007/s12038-020-00081-y

Tamhane, V. A., Giri, A. P., Sainani, M. N., and Gupta, V. S. (2007). Diverse Forms of Pin-II Family Proteinase Inhibitors from *Capsicum Annuum* Adversely Affect the Growth and Development of *Helicoverpa Armigera*. *Gene* 403 (1-2), 29–38. doi:10.1016/j.gene.2007.07.024

Trung, N. P., Fitches, E., and Gatehouse, J. A. (2006). A Fusion Protein Containing a Lepidopteran-specific Toxin from the South Indian Red Scorpion (*Mesobuthus Tamulus*) and Snowdrop Lectin Shows Oral Toxicity to Target Insects. *BMC Biotechnol.* 6 (1), 1–12. doi:10.1186/1472-6750-6-18

Tu, J., Zhang, G., Datta, K., Xu, C., He, Y., Zhang, Q., et al. (2000). Field Performance of Transgenic Elite Commercial Hybrid Rice Expressing Bacillus Thuringiensis δ-endotoxin. *Nat. Biotechnol.* 18 (10), 1101–1104. doi:10.1038/80310

van Dijk, M., Morley, T., Rau, M. L., and Saghai, Y. (2021). A Meta-Analysis of Projected Global Food Demand and Population at Risk of Hunger for the Period 2010-2050. *Nat. Food* 2 (7), 494–501. doi:10.1038/s43016-021-00322-9

Vandenborre, G., Smagghe, G., and Van Damme, E. J. M. (2011). Plant Lectins as Defense Proteins against Phytophagous Insects. *Phytochemistry* 72 (13), 1538–1550. doi:10.1016/j.phytochem.2011.02.024

Vanti, G. L., Katageri, I. S., Inamdar, S. R., Hiremathada, V., and Swamy, B. M. (2018). Potent Insect Gut Binding Lectin from *Sclerotium Rolfsii* Impart Resistance to Sucking and Chewing Type Insects in Cotton. *J. Biotechnol.* 278, 20–27. doi:10.1016/j.jbiotec.2018.04.018

Wang, J., Zhang, H., Wang, H., Zhao, S., Zuo, Y., Yang, Y., et al. (2016). Functional Validation of Cadherin as a Receptor of *Bt* Toxin Cry1Ac in *Helicoverpa Armigera* Utilizing the CRISPR/Cas9 System. *Insect Biochem. Mol. Biol.* 76, 11–17. doi:10.1016/j.ibmb.2016.06.008

Wang, J., Ma, H., Zuo, Y., Yang, Y., and Wu, Y. (2020a). CRISPR -mediated Gene Knockout Reveals Nicotinic Acetylcholine Receptor (nAChR) Subunit α6 as a Target of Spinosyns in *Helicoverpa Armigera*. *Pest Manag. Sci.* 76 (9), 2925–2931. doi:10.1002/ps.5889

Wang, X., Xu, Y., Huang, J., Jin, W., Yang, Y., and Wu, Y. (2020b). CRISPR-mediated Knockout of the ABCC2 Gene in *Ostrinia Furnacalis* Confers High-Level Resistance to the *Bacillus Thuringiensis* Cry1Fa Toxin. *Toxins* 12 (4), 246. doi:10.3390/toxins12040246

Wünn, J., Klöti, A., Burkhardt, P. K., Biswas, G. C. G., Launis, K., Iglesias, V. A., et al. (1996). Transgenic Indica Rice Breeding Line IR58 Expressing a Synthetic crylA (B) Gene from *Bacillus Thuringiensis* Provides Effective Insect Pest Control. *Bio/technology* 14 (2), 171–176.

Xi, Y., Pan, P.-L., Ye, Y.-X., Yu, B., Xu, H.-J., and Zhang, C.-X. (2015). Chitinase-like Gene Family in the Brown planthopper,Nilaparvata Lugens. *Insect Mol. Biol.* 24 (1), 29–40. doi:10.1111/imb.12133

Xia, J., Guo, Z., Yang, Z., Han, H., Wang, S., Xu, H., et al. (2021). Whitefly Hijacks a Plant Detoxification Gene that Neutralizes Plant Toxins. *Cell* 184 (7), 1693–1705. doi:10.1016/j.cell.2021.02.014

Xu, C., Cheng, J., Lin, H., Lin, C., Gao, J., and Shen, Z. (2018). Characterization of Transgenic Rice Expressing Fusion Protein Cry1Ab/Vip3A for Insect Resistance. *Sci. Rep.* 8 (1), 1–8. doi:10.1038/s41598-018-34104-4

Xu, D., Xue, Q., McElroy, D., Mawal, Y., Hilder, V. A., and Wu, R. (1996). Constitutive Expression of a Cowpea Trypsin Inhibitor Gene, CpTi, in Transgenic Rice Plants Confers Resistance to Two Major Rice Insect Pests. *Mol. Breed.* 2 (2), 167–173. doi:10.1007/bf00441431

Yan, S., Qian, J., Cai, C., Ma, Z., Li, J., Yin, M., et al. (2020). Spray Method Application of Transdermal dsRNA Delivery System for Efficient Gene Silencing and Pest Control on Soybean Aphid *Aphis Glycines*. *J. Pest Sci.* 93 (1), 449–459. doi:10.1007/s10340-019-01157-x

Zhang, J., Khan, S. A., Hasse, C., Ruf, S., Heckel, D. G., and Bock, R. (2015). Full Crop Protection from an Insect Pest by Expression of Long Double-Stranded RNAs in Plastids. *Science* 347 (6225), 991–994. doi:10.1126/science.1261680

Zhong, Y., Ahmed, S., Deng, G., Fan, W., Zhang, P., and Wang, H. (2019). Improved Insect Resistance against *Spodoptera Litura* in Transgenic Sweetpotato by Overexpressing Cry1Aa Toxin. *Plant Cell Rep.* 38 (11), 1439–1448. doi:10.1007/s00299-019-02460-8

Zhu-Salzman, K., and Zeng, R. (2015). Insect Response to Plant Defensive Protease Inhibitors. *Annu. Rev. Entomol.* 60, 233–252. doi:10.1146/annurev-ento-010814-020816

Zsögön, A., Peres, L. E. P., Xiao, Y., Yan, J., and Fernie, A. R. (2022). Enhancing Crop Diversity for Food Security in the Face of Climate Uncertainty. *Plant J.* 109 (2), 402–414. doi:10.1111/tpj.15626

Zsögön, A., Čermák, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., et al. (2018). De Novo domestication of Wild Tomato Using Genome Editing. *Nat. Biotechnol.* 36 (12), 1211–1216. doi:10.1038/nbt.4272

Zuo, Y., Wang, H., Xu, Y., Huang, J., Wu, S., Wu, Y., et al. (2017). CRISPR/Cas9 Mediated G4946E Substitution in the Ryanodine Receptor of *Spodoptera Exigua* Confers High Levels of Resistance to Diamide Insecticides. *Insect Biochem. Mol. Biol.* 89, 79–85. doi:10.1016/j.ibmb.2017.09.005

Zuo, Y., Xue, Y., Lu, W., Ma, H., Chen, M., Wu, Y., et al. (2020). Functional Validation of Nicotinic Acetylcholine Receptor (nAChR) α6 as a Target of Spinosyns in *Spodoptera Exigua* Utilizing the CRISPR/Cas9 System. *Pest Manag. Sci.* 76 (7), 2415–2422. doi:10.1002/ps.5782

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

REVIEW
published: 14 June 2022
doi: 10.3389/fgene.2022.900897

Genome Editing Targets for Improving Nutrient Use Efficiency and Nutrient Stress Adaptation

Lekshmy Sathee[1†], B. Jagadhesan[1†], Pratheek H. Pandesha[1,2], Dipankar Barman[1], Sandeep Adavi B[1], Shivani Nagar[1], G. K. Krishna[3], Shailesh Tripathi[4], Shailendra K. Jha[4] and Viswanathan Chinnusamy[1*]*

[1]Division of Plant Physiology, ICAR-Indian Agricultural Research Institute, New Delhi, India, [2]Roy and Diana Vagelos Division of Biology and Biomedical Sciences, Washington University in St. Louis, St. Louis, MO, United States, [3]Department of Plant Physiology, College of Agriculture, KAU, Thrissur, India, [4]Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

In recent years, the development of RNA-guided genome editing (CRISPR-Cas9 technology) has revolutionized plant genome editing. Under nutrient deficiency conditions, different transcription factors and regulatory gene networks work together to maintain nutrient homeostasis. Improvement in the use efficiency of nitrogen (N), phosphorus (P) and potassium (K) is essential to ensure sustainable yield with enhanced quality and tolerance to stresses. This review outlines potential targets suitable for genome editing for understanding and improving nutrient use (NtUE) efficiency and nutrient stress tolerance. The different genome editing strategies for employing crucial negative and positive regulators are also described. Negative regulators of nutrient signalling are the potential targets for genome editing, that may improve nutrient uptake and stress signalling under resource-poor conditions. The promoter engineering by CRISPR/dead (d) Cas9 (dCas9) cytosine and adenine base editing and prime editing is a successful strategy to generate precise changes. CRISPR/dCas9 system also offers the added advantage of exploiting transcriptional activators/repressors for overexpression of genes of interest in a targeted manner. CRISPR activation (CRISPRa) and CRISPR interference (CRISPRi) are variants of CRISPR in which a dCas9 dependent transcription activation or interference is achieved. dCas9-SunTag system can be employed to engineer targeted gene activation and DNA methylation in plants. The development of nutrient use efficient plants through CRISPR-Cas technology will enhance the pace of genetic improvement for nutrient stress tolerance of crops and improve the sustainability of agriculture.

Keywords: genome editing, CRISPR-Cas, nutrient stress, nutrient use efficiency, biofortification, abiotic stress

OPEN ACCESS

Edited by:
Reyazul Rouf Mir, Sher-e-Kashmir University of Agricultural Sciences and Technology, India

Reviewed by:
Anuj Kumar, Dalhousie University, Canada
Vennampally Nataraj, ICAR-Indian Institute of Soybean Research, India
Sajid Fiaz, University of Haripur, Pakistan
Narendra Kumar Singh, Govind Ballabh Pant University of Agriculture and Technology, India

***Correspondence:**
Lekshmy Sathee
lekshmynair@gmail.com
Viswanathan Chinnusamy
viswa.chinnusamy@gmail.com

†These authors have contributed equally to this work

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

Received: *21 March 2022*
Accepted: *17 May 2022*
Published: *14 June 2022*

Citation:
Sathee L, Jagadhesan B, Pandesha PH, Barman D, Adavi B S, Nagar S, Krishna GK, Tripathi S, Jha SK and Chinnusamy V (2022) Genome Editing Targets for Improving Nutrient Use Efficiency and Nutrient Stress Adaptation. Front. Genet. 13:900897. doi: 10.3389/fgene.2022.900897

INTRODUCTION

Genetic improvement of crop abiotic stress tolerance and nutrient use efficiency (NtUE) has become indispensable due to the climate change vagaries and the constant challenge of feeding the burgeoning population. The availability of genetically robust resource use efficient genotypes could minimize the input cost and ensure sustainable food sufficiency (Zhu, 2002; Mahajan and Tuteja, 2005; Hirel et al., 2007; Balyan et al., 2016). Mutation led heritable variations are an important

source for genetic improvement of crops (Olsen and Wendel, 2013) which have been exploited since the early phase of crop breeding (Tang et al., 2020). From the 1950s onward, induced mutations gained importance in varietal development. Due to their randomness, the conventional mutagenesis strategy is prone to several challenges including the screening of thousands of mutants to obtain desirable mutations (Miglani, 2017). The availability of genome editing technologies enabled precise and predictable genetic modifications in the plant genome to bring desirable changes of economic and environmental importance (Gaj et al., 2016).

Genome editing is a powerful tool applicable to every branch of life science for knock out, knock down, or alteration of gene expression without disturbing the genetic makeup of the organism (Ahmed et al., 2019). Genome editing can create a loss of function mutant of "susceptibility genes" or shut down a negative regulator of the stress response pathway (Feng et al., 2013), at the same time it can be employed to enhance the expression, activity and stability of positive regulators of the stress response pathway (Zhou et al., 2017). Theoretically, any gene in an organism can be manipulated with the help of sequence-specific DNA nucleases (Takasu et al., 2010). Mega nucleases, Zinc Finger Nucleases (ZFNs), Transcription Activator-like Effector Nucleases (TALENs), and Clustered Regularly Interspaced Short Palindromic Repeat associated Nuclease (Cas), are site-specific DNA nucleases developed over the past few years (Hisano et al., 2021).

Cas nuclease consists of a specificity governing DNA binding domain and a non-specific nuclease domain that creates a double-stranded break (DSB) on the target DNA, thus driving targeted modification of the genome (Zhang F. et al., 2020). The DSB in DNA is repaired by Non-Homologous End-Joining (NHEJ) or the Homologous Recombination (HR) mechanisms of the cell. Errors in NHEJ or changes in the repair template DNA in HDR (Homology Directed Repair) causes mutation (Chen et al., 2013). The NHEJ repair is an error-prone mechanism that fuses the broken DNA with minor additions and deletions of nucleotides (Kato-Inui et al., 2018). The HDR pathway uses a homologous donor DNA as a template to repair the DSBs precisely (Kato-Inui et al., 2018). Most plant genome editing experiments have taken advantage of NHEJ pathway for gene knockout and generating frameshift mutations. Before the discovery of CRISPR as a gene-editing tool, ZFNs were used for genome editing (Chandrasegaran and Carroll, 2016). A pair of ZFNs target a specific site, one recognizes the upstream sequence and the other one identify the downstream sequences of the target site to be altered (Miller et al., 2007; Szczepek et al., 2007). Similar to ZFNs, TALENs involve a fusion of a FokI nuclease domain with a DNA binding domain which is the TALE (Transcription Activator-like Effector) pattern adapted from the virulence factors of the plant bacterial pathogen *Xanthomonas* (Gaj et al., 2016). TALEN's DNA binding domain has multiple repeating units with each unit spanning 33–35 amino acids that can recognize one DNA base pair (Joung and Sander, 2013). Though TALENs are efficient genome editors, it involves protein engineering depending on the target DNA sequence (Gupta and Musunuru, 2014). Recently a groundbreaking gene-targeting tool based on the RNA-guided Cas9 nuclease from the type II prokaryotic Cluster Regularly Interspaced Short Palindromic Repeats (CRISPR)-Cas was developed (Jinek et al., 2012b). The first instance of CRISPR/Cas was observed in 1987, as an adaptive immunity mechanism of bacteria to viral DNA (Ishino et al., 1987). In the natural system, CRISPR loci upon transcription produce CRISPR-RNA (crRNA) and trans-activating CRISPR-RNA (tracrRNA) which upon base pairing make functional guide RNA (gRNA). For genome editing, CRISPR loci are transcribed into a synthetic single-guide RNA (sgRNA). The gRNA/sgRNA forms a functional complex with CRISPR-associated nuclease (Cas9) and directs the nuclease to genomic loci based on the complementarity of a 20-bp stretch spacer sequence of sgRNA, cleaving it upstream of a necessary 5′-NGG Protospacer Adjacent Motif (PAM) (Xie and Yang, 2013). The *Streptococcus pyogenes* (*Sp*) Cas9 and Cas9 isolated from other organisms have different PAMs and have different distances to the active (cleaving) sites. CRISPR/Cas9 system utilizes gRNA that make a complex with Cas9 and make DSB at the target. Unlike ZFNs and TALENs which requires complex protein engineering for changing the DNA target sequence CRISPR/Cas requires a change of only 20 bp (Liu et al., 2017).

In the recent past, several strategies have been utilized to improve the nutrient response of crop plants, such as differential alternative splicing of genes in response to boron deficiency (Gu et al., 2019), overexpression of genes coding for enzymes of the ammonia assimilation (GS/GOGAT) pathway under constitutive promoters for improving nitrogen (N) use efficiency (Wei et al., 2018), exploiting the role of long non-coding RNA (lncRNA) in *nigt1.1 nigt1.2* double mutant for enhanced nitrate uptake under low P stress (Ueda et al., 2020), etc. But these approaches achieved no significant and stable improvement in NtUE. For example, transgenic plants overexpressing N assimilatory genes did not significantly enhance N Utilization Efficiency (Pathak et al., 2008; Lightfoot, 2009). This review outlines the potential of CRISPR-Cas technology in understanding and facilitating NtUE and nutrient stress tolerance.

Genome Editing Targets for Improving Nitrogen Use Efficiency

Plants posess well-developed uptake and signalling systems to cope with soil N fluctuations. Nitrate induces several genes including its transporter families, NRT1 (Nitrate Transporter), NRT2 and its assimilation pathway genes, encoding Nitrate Reductase (NR), Nitrite Reductase (NiR) (Wang Y.-Y et al., 2018). Induction of these genes in NR-deficient mutants in response to nitrate indicates that nitrate *per se* is also a signal molecule (Zhao et al., 2018). Nitrate mediated activation of genes occurs very rapidly (within minutes), without the necessity for protein synthesis, is termed "Primary Nitrate Response" (PNR) (Krouk et al., 2010). PNR is regulated by various kinases, transcription factors, and noncoding RNAs. Phospho-proteome level changes are controlled by nitrate supply, and 773 unique phosphorylated peptides were identified within a short period of nitrate supply (Engelsberger and Schulze, 2012). As a signalling molecule, nitrate plays a crucial role in plant growth and development (**Figure 1**). Several developmental

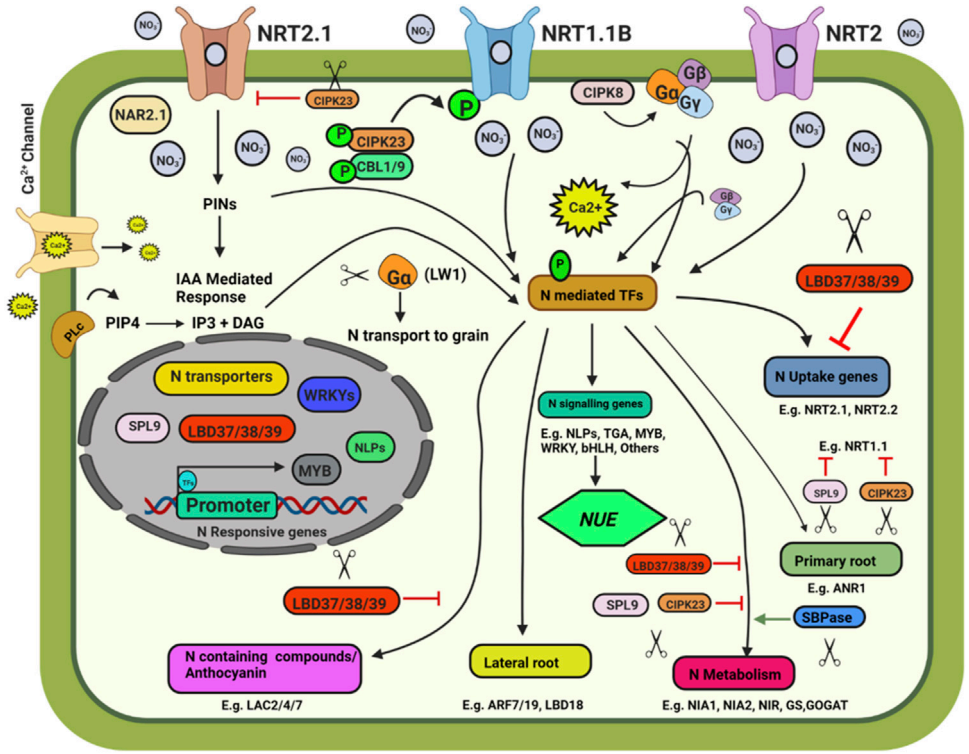


FIGURE 1 | Schematic model depicting plant response to nitrate availability and signaling. Illustration was made with Biorender (https://biorender.com/). The negative regulators for targeted genome editing are indicated with a scissor sign.

processes such as germination (Alboresi et al., 2005), root system architecture (RSA), leaf emergence, leaf growth (Guo et al., 2001; Rahayu et al., 2005), and flowering (Castro Marín et al., 2011) are responsive to nitrate availability.

The Site-Directed Nuclease 1 (SDN1) genome editing approach was found to be promising and can be used to knock out the negative regulators of N uptake and signalling pathways. Some of the potential targets are discussed below. Three members of the *LATERAL ORGAN BOUNDARY DOMAIN (LBD)* gene family of transcription factors (LBD37, LBD38, and LBD39) were identified as repressors of anthocyanin biosynthesis and N availability signals in *Arabidopsis* (Rubin et al., 2009). The overexpression of LBD37/38/39 could repress the expression of nitrate-responsive genes, including *NRT2.1*, *NRT2.2*, *NIA1*, and *NIA2*, indicating the negative role of the three LBD members in nitrate signalling. SQUAMOSA Promoter-Binding Protein-Like 9 (SPL9) and CBL-interacting protein kinase 23 (CIPK23) also negatively affect the nitrate response. SPL9 reduces the transcript levels of genes involved in N uptake, i.e., *NRT1.1*, *NIA2*, and *NIR*, while miR156 negatively regulates SPL9's transcript level (Krouk et al., 2010). CIPK23 negatively regulates the expression of the high-affinity nitrate transporter *NRT2.1*. CIPK23 acts as a regulator for the high-affinity switch of NRT1.1 under low nitrate concentrations to enable NRT1.1 to operate as a high-affinity nitrate transporter by phosphorylating it at T101, thus playing an essential role in N uptake (Ho et al., 2009). MicroRNAs such as miR399 and miR827 play an indispensable role in maintaining N and Pi homeostasis in plants by targeting *NITROGEN LIMITATION ADAPTATION* (*NLA*) and *PHOSPHATE2* (*PHO2*) (Nguyen et al., 2015). Knocking out these miRNAs using CRISPR-Cas will help better understand their role in N and P homeostasis. Knocking out of Sedoheptulose-1,7-Bisphosphatase (SBPase), involved in carbon assimilation, could reduce the expression of NR, GS and GOGAT, and alter N metabolism in tomato plants (Ding et al., 2018). NIN like proteins (NLPs) are transcriptional regulators of nitrate signalling (Jagadhesan et al., 2020). CRISPR-Cas9 knockout mutant *OsNLP4* showed reduced NUE and yield (Wu et al., 2020). *OsNLP1* is rapidly inducted by N deficit. OsNLP1 genome-edited lines showed impaired grain yield and NUE under N limiting conditions (Alfatih et al., 2020). The AtNLP7 promotes lateral root development by binding to *TRYPTOPHAN AMINOTRANSFERASE RELATED 2* (TAR2) promoter region in a nitrate dependent manner (Zhang et al., 2021b). In cotton, the overexpression of Arabidopsis *NLP7* enhanced the NUE and yield (Jan et al., 2022). CRISPR/Cas edited *OsNLP3* lines showed reduced plant growth, grain yield, and NUE under sufficient nitrate conditions, whereas in low nitrate or in ammonium supplementation, *osnlp3* mutants performed at par with the wild type (Zhang et al., 2022).

Map-based cloning followed by CRISPR-Cas9, gene-editing confirmed the involvement of Leaf Width 5 (LW5) in N uptake and utilization (Zhu et al., 2020). The *lw5* mutant showed higher rate of photosynthesis, higher chlorophyll content and higher N

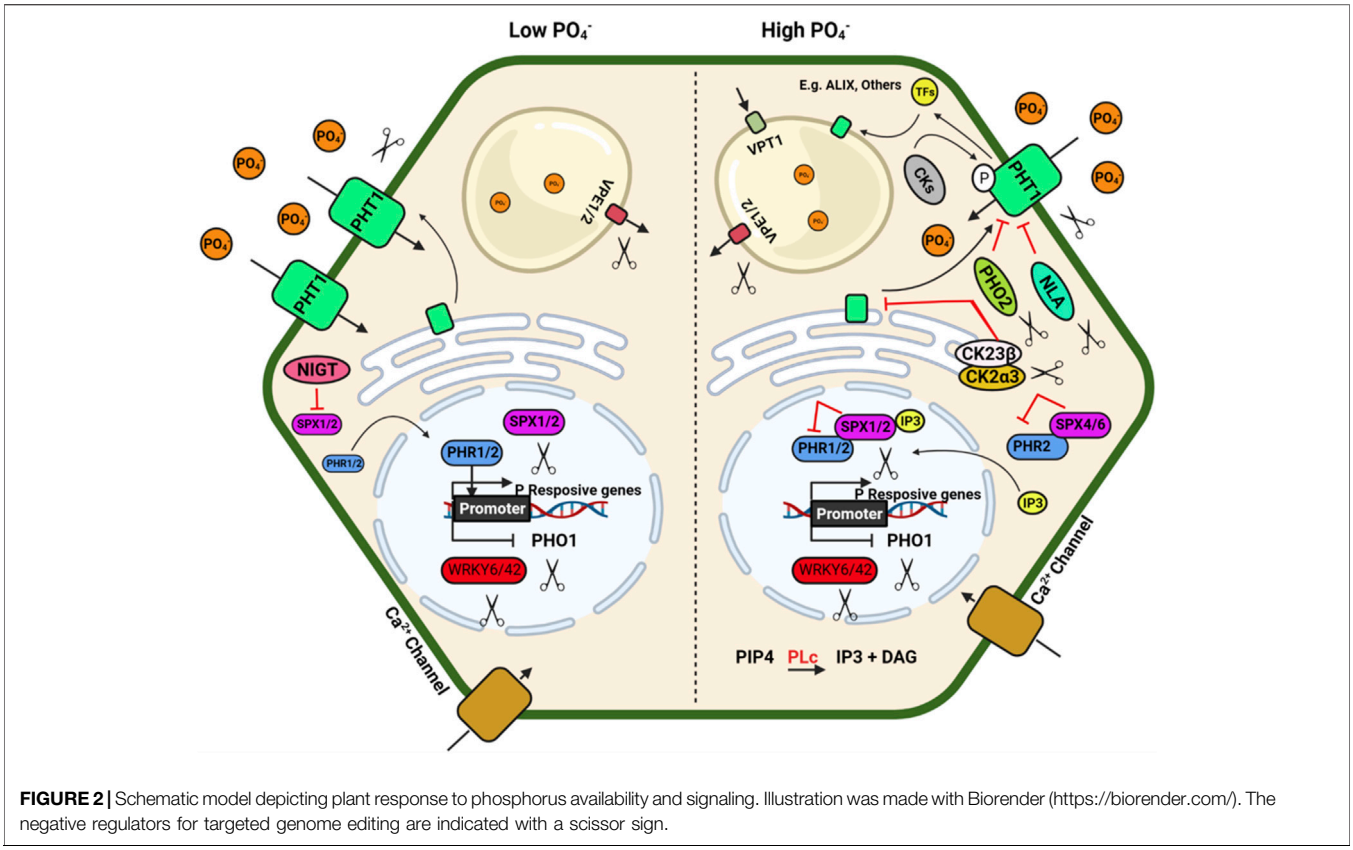


FIGURE 2 | Schematic model depicting plant response to phosphorus availability and signaling. Illustration was made with Biorender (https://biorender.com/). The negative regulators for targeted genome editing are indicated with a scissor sign.

uptake rate, however the reduced grain nutrient remobilization results in small grains. The *abnormal cytokinin response1 repressor 1* (*HvARE1*) was identified as a regulator of NUE in a genome-wide association analysis. ARE1 was identified as a suppressor of plastidic Fd-GOGAT in rice. The *are1* mutant plants of rice, wheat, and barley showed improved yield and higher NUE, making it a worthy genome editing candidate (Wang et al., 2018b, 2021; Zhang et al., 2021a; Karunarathne et al., 2022). In wheat and barley, CRISPR/Cas9 gene-editing of *ARE1* enhanced NUE (Zhang et al., 2021a; Karunarathne et al., 2022). The GROWTH-REGULATING FACTOR 4 (OsGRF4) is a transcriptional regulator of numerous N metabolism genes that work in opposition to the DELLA growth repressor. CRISPR/Cas editing has shown that changing the OsGRF4-DELLA balance by increasing *OsGRF4* abundance improves NUE and grain production (Li et al., 2018a; Gao et al., 2020). Zhang J. et al. (2020) used a multiplex CRISPR/Cas9 vector to modify the *MIR396* gene family in rice, targeting MIR396a, MIR396b, MIR396c, MIR396e, and MIR396f at the same time. Two microRNAs, MIR396e and MIR396f, have been discovered to regulate grain size and plant architecture. The role of *miR396* genes in rice was investigated using CRISPR/Cas9 by knocking down MIR396e and MIR396f. MiR396ef mutants demonstrated a relative increase in grain yield with larger biomass under reduced N conditions and increased grain production under normal N conditions. The DELLA protein gene *SLR1* and nitrate transporter NRT1.1B were targeted using the CRISPR/Cas9 cytidine deaminase enzyme (APOBEC1) base editing system, resulting in point mutations and a C/T substitution (Thr327Met) in NRT1.1B, which boosted NUE in rice (Lu and Zhu, 2017).

Genome Editing Targets for Improving Phosphorus Use Efficiency

Phosphorus (P) serves is a constituent of ATP, pyrophosphate (PPi), as well as an essential structural component of nucleic acid and phospholipids (Chiou and Lin, 2011; Kumar et al., 2018). P deficiency severely limits crop yield; however, only 10–20% of total applied P fertilizer is used by crops, and the remaining P is unavailable to plants (Oelkers and Valsami-Jones, 2008). The P acquisition efficiency (PAE) and physiological P use efficiency (PUE) depend on the uptake, transport and metabolism driven Pi recycling and overall P homeostasis within the plant (Stigter and Plaxton, 2015; Zeeshan et al., 2020; Prathap et al., 2022). Some of the potential targets and strategies to improve PAE and PUE through genome editing tools are discussed in **Figure 2**.

The inorganic P (Pi) is immobile in soil and is generally present in the top layer of soil; its acquisition by plants mainly depend on a dense root system with well-developed lateral roots and root hairs (Lynch, 2007; Nestler et al., 2016). A detailed review of root system architecture under P availability and multiple regulatory checkpoints at the molecular level has

been described by Niu et al. (2013). Under P deficiency, proteoid or cluster roots (specialized bottle brush-like dense lateral roots) are formed in white lupin, efficiently mobilising Pi from Fe–Pi and Ca–Pi sources. Additionally, some plants have developed a mechanism to enhance Pi availability by secreting organic acids that helps to release Pi from different Pi-containing complexes in soil, which is otherwise unavailable to the plants (Reise and Waller, 2009). Organic acids such as citrate, malate, malonate, and oxalate form stable complexes with metal ions like Al, Fe, and Ca compared to Pi and make free Pi available to the plants (Álvarez-Fernández et al., 2014). Al-activated Malate Transporters (ALMTs) family members play a crucial role in P uptake in soybean and help in K and Fe homeostasis in lupin by regulating malate exudation in acid soil (Liang et al., 2013; Zhou et al., 2020).

P acquisition and homeostasis involve five members of phosphate (PHT) transporters viz., PHT1, PHT2, PHT3, PHT4 and PHT5 in plants (Wang and Wu, 2017) which poses a tremendous opportunity to develop genetically edited crops with high P acquisition efficiency. The plasma membrane-localized PHT1 regulates Pi uptake in roots, while the other PHTs regulate Pi translocation in different cellular organelles. Apart from the PHT family, the PHO1 (an SPX-EXS subfamily) transporter regulates Pi transport from roots to shoot, and VPT1 and OsSPX-MFS1-3 (SPX-MFS subfamily) regulate vascular vacuolar Pi transport (Wang et al., 2018a). The CRISPR knockouts of *SlPH O 1.1* in tomato (Zhao et al., 2019) resulted in a Pi starvation response demonstrated as decreased shoot fresh weight, increased root biomass, and a higher root-to-shoot ratio. These mutants also showed a higher anthocyanin accumulation in the shoot and higher root to shoot soluble Pi content. These results confirmed the importance of *SlPHO1;* 1 gene function in Pi transport in the tomato at the seedling stage. Generation of CRISPR-Cas9 genome-edited plants for the six *PHO1* genes in the tomato genome (*SlPH O 1;1-SlPH O 1;6*) will help identify the isoforms' critical role in P uptake and translocation. The R-type MYB transcription factors PHR1(Phosphate Starvation Response) and PHR2 in Arabidopsis and rice regulate phosphate homeostasis and root hair development. PHR binds to the P1BS (PHOSPHATE STARVATION RESPONSE1 binding site) and helps the plant adapt to Pi deficiency in Arabidopsis. Nuclear localized SPX1/2 and cytoplasmic SPX4 are functional inhibitors of AtPHR1/OsPHR2, a positive regulator of Pi signalling and uptake. N availability regulates Pi uptake and starvation signalling through the NIGT1–SPX–PHR cascade (Ueda et al., 2020). NIGT1 and NIGT1.2 inhibit the expression of *SPX* and thus indirectly activate the *PHR* expression (Poza-Carrión and Paz-Ares, 2019; Ueda et al., 2020; Wang et al., 2020). The *nigt1.1 nigt1.2* double mutant developed by CRISPR-Cas9 displayed reduced P uptake and improved N uptake at low Pi conditions. Other potential negative regulators of Pi signalling are *IPS1/2*, *PHO2*, *AtWRKY6,* ATWRKY43, miR827, etc., affecting multiple downstream genes. Consequently, up-regulation of the different positive regulators of Pi signalling such as *AtPHR1*, *At PHR2*, *PHT1*, *PHO1*, *PAP, SQD2*, *OsSPX*, *MFSs*, *VPT*, *Cm-PAP10.1*, *Cm-PAP10.2*, *Cm-RNS1*, *TaALMT1* and *AVP1/AVP1D* etc.) will enhance PUE (Hasan et al., 2016; Wang et al., 2018a).

PHRs also regulate the expression of mi*R399,* which reduces the transcript level of *PHO2*, a negative regulator of Pi deficiency signalling (Bari et al., 2006; Wu and Wang, 2008). CRISPR Cas9 mediated rice mutants *phr1*, *phr2* and *phr3* significantly reduced plant growth. At the same time, double mutants (*phr1/2 and phr2/3*) and triple mutant (*phr1/2/3*) had maximum growth retardation under Pi-sufficient (200 μM Pi) and deficient (10 μM Pi) conditions, respectively. Overexpression of *OsPHR3* improved growth under low-Pi soils and showed average growth under sufficient Pi conditions, implicating positive and diverse regulation of downstream genes for Pi signalling and homeostasis (Guo et al., 2015). In addition to the PHR TFs, several other TFs, namely AtMYB2, AtMYB62, OsMYB2P-1, OsMYB4P, and OsMYB1, are also involved in Pi signalling (Devaiah et al., 2009; Yang et al., 2012, 2014; Gu et al., 2017). Mutation of *MYB1* transcription factor using CRISPR Cas9 system showed an increase in Pi uptake and accumulation (Gu et al., 2017). *MYB1* mutation also altered the expression of multiple genes related to Pi starvation signalling and Pi transporters (*PHT1; 2, PHT1; 9, PHT1; 10, IPS, miR399j,* and *PHO2.1*). A recent study shows that CRISPR-Cas9 double mutant for vacuolar Pi efflux transporters *Osvpe1 Osvpe2* has higher vascular Pi content under low Pi stress than the wild type plants (Xu et al., 2019).

Genome Editing Targets for Improving Potassium use Efficiency

Potassium (K) is an essential macronutrient for diverse physiological activities during plant growth and development. Fluctuation in external K^+ concentration generates chemical and physical signals to cope with the imbalanced state of the cellular K^+ level. These signals directly or indirectly regulate the downstream targets, especially the K^+ channels and transporters. Thus, plants regulate K^+ homeostasis to adapt to the varying K^+ concentration. In plants, K^+ is involved in multiple physiological activities such as osmotic adjustment, maintenance of membrane potential, ionic balance, enzyme activation, stomatal movement and pollen tube growth (Wang and Wu, 2017). The transporter family HAC/KUP/KT are important for K^+ transport in plants (Véry et al., 2014). AtAKT1, AtHAK5, and AtKUP7 are major K^+ transporters in the *Arabidopsis* root system. AtHAK5 is a high-affinity transporter, while AtKUP7 is a low-affinity transporter and mediates K^+ transport into the xylem and shoots (Han et al., 2016; Nieves-Cordones et al., 2016). In rice, three K^+ transporters OsHAK1/5/21 have been characterized to transport K^+ into the root (Yang et al., 2014; Chen et al., 2015). Under deficiency condition, high-affinity transporter OsHAK5 play a role in root and root to shoot K^+ transport which can be a potent target for genome editing mediated up-regulation (Yang et al., 2014). While OsHAK1 is an interesting transporter with low affinity and high-affinity transport activity (Chen et al., 2015) making it an excellent choice for genome editing techniques. OsHAK16 plays a dual

role in K^+ uptake and translocation and maintain shoot K^+ homeostasis in rice (Wang et al., 2019c). Members of CHX (cation/H^+ exchanger), AtCHX14 acts as a K^+/H^+ exchanger, a plasma membrane K^+ efflux transporter. Similarly, other CHXs like AtCHX17 (pre-vacuolar compartment), AtCHX16/18/19, OsCHX14 (ER-localized), etc., are critical regulators of K^+ homeostasis during different developmental stages (Chanroj et al., 2013; Chen et al., 2016a). In rice, the plasma membrane located cation chloride co-transporter (OsCCC1) transports K^+ and maintains ion homeostasis and thus play a significant role in cell elongation (Chen et al., 2016b). An *Arabidopsis* the triple mutant of *kea1/2/3* exhibited significantly stunted growth. This is because of their regulatory role in chloroplast development and regulation of pmf (proton motive force) across the thylakoid membrane as a K^+ transporter (Armbruster et al., 2014; Kunz et al., 2014; Aranda-Sicilia et al., 2016).

Under K^+ deficiency, CBL1/9-CIPK23-AKT1 and CBL4-CIPK6- AKT2 cascades play important role in maintaining K^+ homeostasis in plants (Xu et al., 2006; Behera et al., 2017; Cuin et al., 2018; Saito and Uozumi, 2019). The deficiency of K^+ induces ethylene signalling to generate ROS and regulate the *AtHAK5* transcription level (Schachtman, 2015). Similarly, low K stress decreases cytokinin levels, stimulating ROS production and increasing *AtHAK5* expression (Nam et al., 2012). ROS directly activates Ca^{2+} channels to enhance cytosolic Ca^{2+} levels, activating different K^+ channels and related transcription factors. Rare Cold Inducible gene 3 (*RCI3*), a type III peroxidase, is also involved in ROS production under K^+ deficiency (Kim et al., 2010; Demidchik, 2018; Buet et al., 2019). At flowering, jasmonic acid regulates the *OsCHX14* K^+ transporter gene expression in rice and maintains K^+ homeostasis (Chen et al., 2016a). Under low K^+ stress, *OsmIR399* is upregulated and represses the expression of *LTN1/OsPHO2* and indirectly activates the *OsHAK25* transporter (Chen et al., 2015).

Transcription factors are essential regulators of the K^+ homeostasis in plants. Under low K^+ stress, *AtHAK5* expression increases in Arabidopsis (Gierth et al., 2005). The transcription factor *AtARF2* is a negative regulator of *AtHAK5*. Under sufficient K^+ conditions, *AtARF2* –TF binds to the promoter of the *AtHAK5* gene and suppresses the expression of *AtHAK5*. K^+ deficiency leads to phosphorylation of *AtARF2*, and the phosphorylated *AtARF2* dislodges from the *AtHAK5* promoter region and thus permits the expression of the AtHAK5 gene (Zhao et al., 2016). CBL1-CIPK23-AKT1 cascade phosphorylates AKT1 transporter to activate it in *Arabidopsis*, rice, and Venus flytrap (Li et al., 2014; Chen et al., 2015; Ragel et al., 2015). In Arabidopsis, AtCPK13 reduces stomatal opening by phosphorylating AtKAT1/2, inhibiting the inward K^+ currents and regulating ammonium transporters AtAMT1; 1/1; 2 under high ammonium concentration (Ronzier et al., 2014; Straub et al., 2017). Raf-like MAPKK kinase (*AtILK1*) and rice receptor-like kinase *OsRUPO* also regulate K^+ homeostasis by activating AtHAK5 accumulation on the plasma membrane and inhibiting K^+ proliferation by OsHAKs transporters in Arabidopsis and rice, respectively (Brauer et al., 2016; Liu et al., 2016). Rice homozygous *OsRUPO* knockout lines developed by CRISPR-Cas maintained a higher level of K^+ in pollen than the wild-type plants due to the active OsHAK1. Under normal conditions, RUPO phosphorylates the HAK1 transporter and inhibits the transporter activity (Liu et al., 2016). Besides this TF, some regulatory channel proteins interact with the transporters to modulate their activity. In Arabidopsis channel regulatory subunit AtKC1 and AtCIPK23 synergistically balance K^+ uptake or leakage and modulate AKT1 mediated low K^+ responses (Wang et al., 2016). A regulatory SNARE (soluble N-ethyl maleimide sensitive factor attachment protein receptor) protein VAMP721 targets vesicles to the plasma membrane to suppress the AtAKT1-AtKC1 heteromeric channel protein (Li et al., 2017). Nitrate-dependent shoot K^+ homeostasis is regulated by AtNRT1.5 and OsNPF2.4 in Arabidopsis and rice, respectively (Drechsler et al., 2015; Xia et al., 2015; Meng et al., 2016). Chloroplast localized OsPRX2, a thiol-based peroxidase involved in H_2O_2 homeostasis, is a crucial stomatal closure and K^+ accumulation regulator. Compared to the CRISPR-Cas9 mutated *OsPRX2* lines, *OsPRX2* overexpression lines were more tolerant to K^+ deficiency tolerance. This result shows that *OsPRX2* is a potential target for genome-editing for K deficiency tolerance (Mao et al., 2018). In rice, OsHAK3 is a K^+ transporter to maintain K homeostasis. CRISPR Cas9 mediated mutation of *OsHAK3* resulted in lower K^+ uptake, making them sensitive to low K stress and salinity stress (Zhang L. et al., 2020). By targeted mutation of K^+ transporter OsHAK1, low Cs^+ rice was developed (Nieves-Cordones et al., 2017). Above mentioned regulatory checkpoints from K^+ uptake to transport involving transporters, transcription factors, post-transcriptional and post-translational regulation, etc., are potential targets for improving crop yield under multiple stress conditions and K^+ limited conditions through genome editing (**Figure 3**).

Genome Editing Targets for Improving Iron Use Efficiency

Iron is an essential element found in cofactors associated with electron transfer, hydroxylation, and dehydration reactions. Photosystem I and II, ferredoxins, and a variety of metabolic enzymes are particularly rich in heme and iron-sulphur (FeS) proteins (Tissot et al., 2014).) The oxidised condition of Fe in earth crust makes it inaccessible to plants. Iron deficiency led anaemia is a serious human health problem affecting around 30% of the global population, according to the World Health Organization (WHO) (http://www.who.int/nutrition/topics/ida/en/). The Fe deficiency faced by crop plants will thus impact global food and nutritional security. Researchers have discovered a slew of critical factors that regulate iron uptake and metabolism (**Figure 4**). Traditionally, Fe uptake in plants has been divided into two techniques: strategy I and strategy II, commonly known as reducing and chelating strategies, respectively (Römheld and Marschner, 1986). The fundamental difference is the oxidation state of Fe ion up by the plant: ferrous Fe^{2+} in strategy I and ferric Fe^{3+} in strategy II. In the rhizosphere, Fe is mostly found as Fe^{3+} oxyhydrates with low solubility. Tomato and *Arabidopsis* have

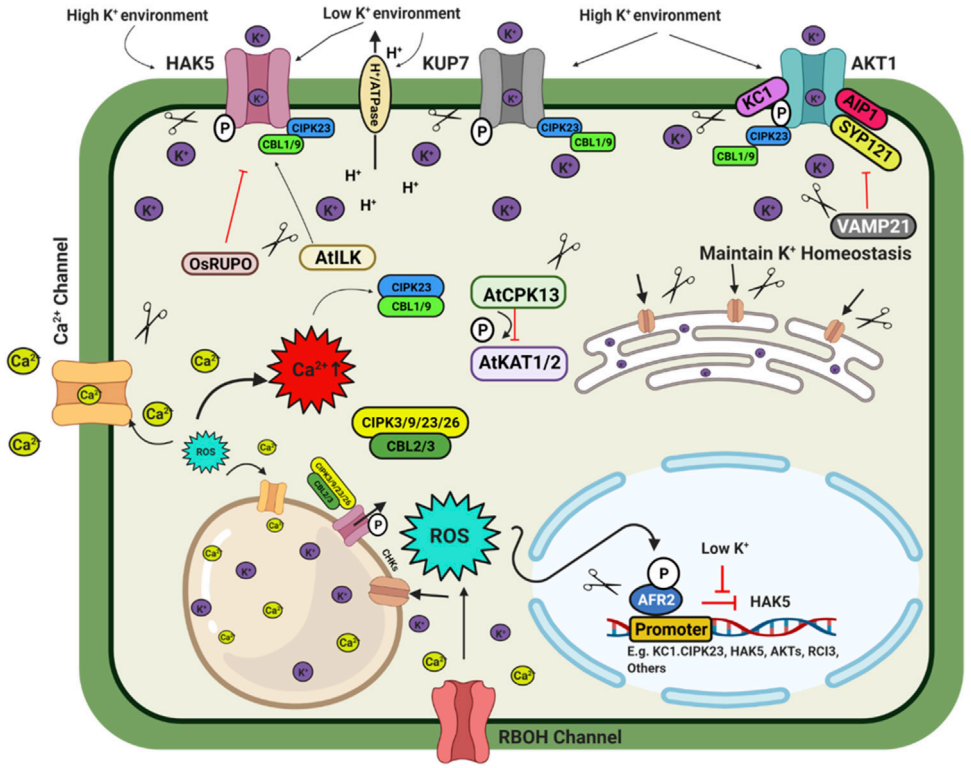


FIGURE 3 | Schematic model depicting plant response to potassium availability and signaling. Illustration was made with Biorender (https://biorender.com/). The negative regulators for targeted genome editing are indicated with a scissor sign.

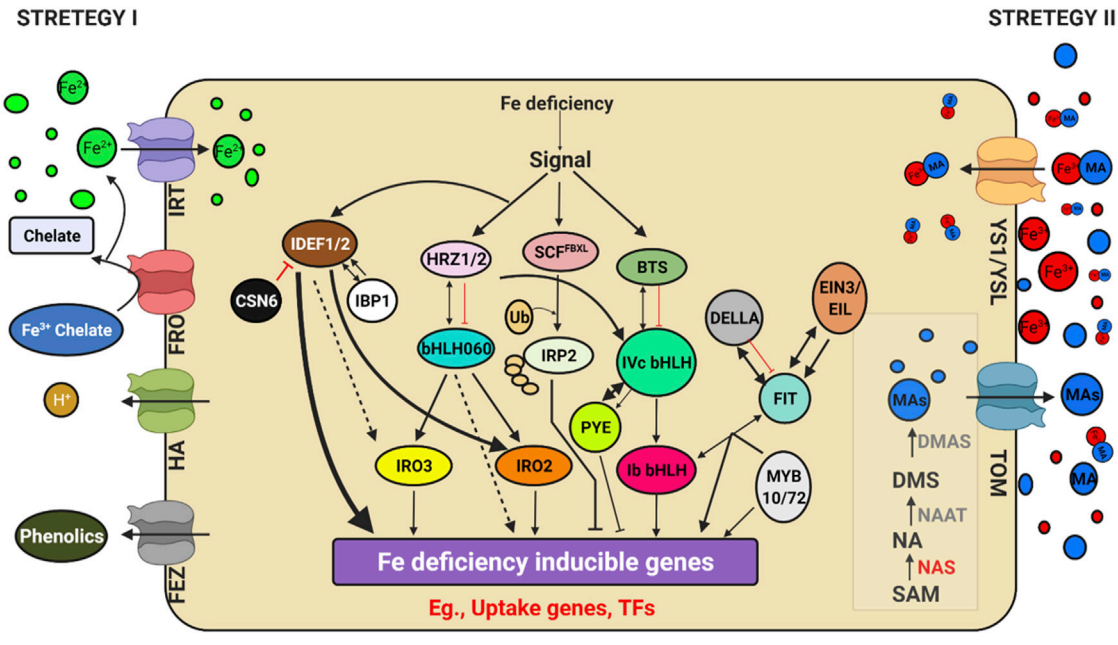


FIGURE 4 | Schematic model depicting Genes regulating iron (Fe) uptake and deficiency response. Illustration was made with Biorender (https://biorender.com/). The possible negative regulators and genome editing targets are indicated with a scissor sign.

been used as models for Strategy I, in which Fe^{3+} is reduced at the plasma membrane by Ferric Reduction Oxidase 2 (FRO2) before being transported across the membrane by Iron-Regulated Transporter 1 (IRT1) (Robinson et al., 1999). AHA2, a plasma membrane proton pump, also aids in acidifying the rhizosphere and increasing Fe^{3+} solubility (Santi and Schmidt, 2009). Grass family (Poaceae) crops were used as a model for strategy II plants, which secrete phytosiderophores (PS), which are tiny organic molecules produced by plants that have a higher affinity for iron molecules (Kanazawa et al., 1994). Deoxymugineic acid is the most abundant phytosiderophore in rice and barley, and it is secreted by TOM1 (Transporter of Mugineic Acid Family Phytosiderophores) (Nozoye et al., 2011). Organic metabolites such as organic acids, flavonoids, phenolics, and flavins were discovered in Strategy I plants. Phenolics were initially thought to help with the solubilization and availability of apoplastic iron until coumarin-related phenolics associated with soil Fe uptake (Lešková et al., 2017). Fraxetin, a major coumarin related phenolic compound, is produced by Feruloyl CoA orthohydroxylase 1 (Fourcroy et al., 2014). Instead of coumarins, alfalfa and sugar beet secrete flavins that aid in the reduction of Fe^{2+} ions (Rodríguez-Celma et al., 2013; Fourcroy et al., 2014). Furthermore, the exudation of putrescine, a polyamine, improved iron mobilisation within the plant cell wall (Zhu et al., 2015). The first report of an oligopeptide transporter (OPT) family protein capable of transporting Fe^{3+}-siderophore complex was identified in maize from a mutation in the Yellow Stripe1-Like (YSL) transporter (Curie et al., 2001). Even though non-grasses do not synthesise or secrete PS, many YSL genes have been discovered and are conserved among all land plants. In monocots and dicots, YSLs played a key role in the long-distance transport of metals bound to nicotiamine (NA), such as Fe, Cu, and Zn (Waters et al., 2006), except for the Fe-PS transporters *ZmYS1* (Curie et al., 2009; Lee et al., 2009) and its rice ortholog, OsYSL15 (Inoue et al., 2009; Kobayashi et al., 2014). Eighteen *YSL* genes (*OsYSLs*) have been discovered in rice (Aoyama et al., 2009), and the *OsYSL15* gene has been identified as a Fe-PS transporter involved in Fe acquisition (Inoue et al., 2009; Kobayashi et al., 2014). Under Fe deficiency, the *OsYSL15* gene was highly expressed in root tissues, particularly the epidermis, as demonstrated by OsYSL15 promoter-GUS transgenic experiments (Inoue et al., 2009). YSL1-8 are the eight members of the YSL family identified in Arabidopsis (Jean et al., 2005; Waters et al., 2006). Metal remobilization from senescent leaves is a major role of *YSL* genes, which also serve as transporters in seed development, reproductive organ growth, and long-distance transport of metal complexes with NA (DiDonato et al., 2004). Iron and Cu regulate *AtYSL2* expression, and *YSL2* is abundantly expressed in the root endodermis, pericycle, and xylem cells (Lei et al., 2014). AtYSL2 is thought to play a function in transporting metals to and from the vasculature based on its location. Under control conditions, *ysl2-1* mutants display no apparent characteristics (Schaaf et al., 2005). The *ysl2-1* mutants showed no visual manifestation even when Fe was depleted, showing that additional YSLs must have redundancy activities in transporting metal-NA complexes. Overall, the mutant analysis revealed that FRO2 and IRT1 are required for the uptake of Fe mobilised by coumarins, and that IRT1 is the primary pathway for Fe uptake in Arabidopsis (Shin et al., 2013; Fourcroy et al., 2014; Boonyaves et al., 2016). Furthermore, rice and barley contain a functional homologue of IRT1, which is responsible for facilitating Fe^{2+} absorption in low-oxygen environments (Ishimaru et al., 2006; Boonyaves et al., 2016). Thus, the distinction between strategy I and II is increasingly blurred. In Nature, two opposing Fe^{2+} and Fe^{3+} uptake methods have been observed in the same organism. Fe^{2+} is transferred by the Divalent Metal Transporter DMT1 and Fe^{3+} is captured by transferrin by the same cells as in mammals (Anderson et al., 2013) and Bacteria. Since the discovery of FER, the first critical TF, in tomato (Ling et al., 2002), a plethora of TFs involved in Fe-deficiency response have been identified in *Arabidopsis thaliana* (Hindt and Guerinot, 2012; Lin et al., 2016) and rice (Kobayashi et al., 2014). Because of their functions as potent regulators of Fe deficiency responses and their Fe-binding properties, Iron Deficiency-responsive Element-binding Factor 1 (IDEF1) and Hemerythrin motif-containing Really Interesting New Gene- and Zinc-finger proteins (HRZs)/BRUTUS (BTS) have recently emerged as candidate iron sensors (Kobayashi et al., 2014; Kobayashi and Nishizawa, 2014). IDEF1 is a transcriptional regulator of graminaceous genes involved in Fe absorption and utilisation, which is especially important in the early stages of Fe shortage. Both graminaceous and non-graminaceous plants have HRZs/BTS, which are E3 ubiquitin ligases and negative regulators of Fe deficiency responses. Furthermore, a recent study found that iron man (IMA)/Fe-uptake inducing peptides (FEPs) are positively regulated in plant Fe-deficiency responses (Grillet and Schmidt, 2019; Kobayashi, 2019). Few of these critical regulators are found in both rice and Arabidopsis, and are thus thought to be conserved across plant species, despite differences in downstream genes involved in Fe uptake. The major participants in the transcriptional control of these conserved pathways are basic helix–loop–helix (bHLH) TFs, such as rice OsbHLH060 (OsPRI1) and Arabidopsis AtbHLH34, 104, 105 (ILR3) and 115(Rodríguez-Celma et al., 2013). POPEYE (PYE), is a bHLH TF involved in pericycle-specific iron deficiency response and regulates growth and development in iron-deficient environments (Long et al., 2010). PYE interacts with PYE homologs, such as IAA-Leu Resistant3 (ILR3), a metal ion homeostasis-related bHLH transcription factor. In addition, ILR3 interacts with BRUTUS (BTS), a putative E3 ligase protein with metal ion binding and DNA binding domains, which inhibits the response to Fe. The Fe sensing and signaling molecules upstream of Fe-deficiency responses in plant cells are unclear (Kobayashi, 2019).

Genome Editing Targets and Case Studies on Improving Salinity and Ion Toxicity Tolerance

Heavy metal toxicity and salinity are major stresses of edaphic origin, detrimental to crop growth and toxic to the animals and

humans that feed on them (Munns et al., 2020). Genome editing is a promising tool to enhance the tolerance of crops to salinity and ion toxicity from heavy metal contamination. The research in this regard is evolving faster and has come up with several successful instances.

Salinity Tolerance

The negative regulators in abiotic stress response pathways are poorly characterised, as a result, there is a scarcity of CRISPR/Cas-mediated studies on abiotic stress tolerance. RECEPTOR FOR ACTIVATED C KINASE 1 (RACK1), a WD40-repeat family protein, is a potent negative regulator of ABA-mediated stress signalling (Guo et al., 2009; Zhang et al., 2013). The constitutive RNAi-mediated knockdown of a RACK1 homolog boosted drought and salt tolerance in soybean plants significantly compared to WT plants (Zhang et al., 2013). The rice RACK1 RNAi lines performed better under drought and salinity stresses, implying that RACK1 is a good target for genome editing. In Arabidopsis, the FARNESYL TRANSFERASE A (FTA) and ENHANCED RESPONSE TO ABA1 (ERA1), which encodes subunits of the enzymatic protein farnesyl transferase that negatively regulate ABA signalling. The knockdown/knockout of both these genes causes ABA hypersensitivity, stomatal closure, and decreased transpiration rates. Mutants of *era1* showed drought resistance and seed dormancy in crops like wheat and rapeseed. These mutants also have enhanced pathogen susceptibility and improved heat tolerance (Wu et al., 2017). Drought-induced down-regulation of the *BnFTA* and *BnERA1* genes in *Brassica napus* has been shown to protect yield under water deficit conditions during blooming while negatively affecting growth under well-irrigated conditions. The ABA hypersensitive Arabidopsis mutants, such as *abh1*, *abo1*, and *cyp85a2*, show improved drought sensitivity; however, less information is available about potential pleiotropic effects or functions in other plant species. Recently it was reported that protein farnesylation reduces the stability of BES1 and negatively regulates the brassinosteroid signalling in *Arabidopsis thaliana* (Zengxiu Feng, 2021). In rice, CRISPR editing of genes associated with carbohydrate metabolism, hormonal homeostasis, and stress signalling gave successful results as compiled in (Ganie et al., 2021).

Ion Toxicity Tolerance

Excessive accumulation of heavy metals is toxic to crop plants. Metal toxicity severely affects plant growth, development, and yield. It causes oxidative stress and disturbs cellular ionic homeostasis leading to cellular damage. CRISPR-Cas9 system may hold potential for improvement in reducing metal toxicity in plants. The plant genes regulating heavy metal uptake and translocation are potential candidate genes for CRISPR/Cas9 genome editing.

Aluminium toxicity (Al^{3+}) is one of the most prevalent toxicity problems in acidic soils. Aluminium inhibits root growth in plants, and the *OsAUX3* (*AUXIN3*) gene plays an important role in the Al sensitivity of rice. *OsAUX3* is an auxin influx transporter that promotes hormone transport acropetally (towards the tips). The *osaux3* mutants generated by CRISPR/Cas9 technology were insensitive to auxin and Al. The defective root growth was less pronounced in the mutants than in the wild type under Al toxicity. The Al accumulation in tissues was also lesser in mutants, as shown by ICP-OES and Morin fluorescence methods (Wang et al., 2019b). Genome editing of positive regulators gave detailed insights into tolerance mechanisms. *ART2* (*Al RESISTANCE TRANSCRIPTION FACTOR 2*) mutant lines in rice were observed to have no pleiotropic effects under normal conditions, but under Al toxicity, the inhibition of elongation was increased (Che et al., 2018). ALMT family of Malate efflux transporters involved in chelating free aluminium in soil solution. The generation of Sl-ALMT9 (*Al-ACTIVATED MALATE TRANSPORTER9*) knockout mutants of tomato resulted in lower root growth and higher Al content in roots, which was evident from hematoxylin staining in root apices (Ye et al., 2017).

The transcription factor OsARM1 (ARSENITE-RESPONSIVE MYB1) is involved in the Arsenic (As)-dependent upregulation of As transport genes in rice. The knockout of OsARM1 resulted in reduced As uptake and tolerance of rice to even higher doses of the heavy metal (Wang et al., 2017). The hydrated diameter of As is similar to P, and hence several phosphate transporters are found to transport the heavy metal. The OsPHF1 (PHOSPHATE TRANSPORTER TRAFFIC FACILITATOR 1) is one such protein having dual specificity. *OsPHF1* mutants generated by EMS treatment led to increased tolerance to As and enhanced biomass even under contamination treatments. The lines had a reduced uptake of both As and Pi (Wu et al., 2011). Thus *OsPHF1* is a promising gene for imparting tolerance against As toxicity via genome editing. Other good targets include *OsPht1;8* (PHOSPHATE TRANSPORTER 1; 8), *Lsi1,* and *Lsi2* (*LOW SILICON RICE 1* and *2*) are significant conduits for As uptake and are early targets for genome editing (Chen et al., 2017). As the knockout of *OsPht1:8*. *Lsi1* and *Lsi2* will influence P and Silicon (Si) uptake, selected alleles specific for As uptake needs to be identified. OsNRAMP1 (Takahashi et al., 2011) transporter responsible for As uptake is also an important genome editing target to reduce AS content in rice.

Os *BET1* (*BORON EXCESS TOLERANT 1*) belongs to the NAC (NAM, ATAF, and CUC) family of transcription factors. It was identified from a RIL derived from IR36 and Nekken 1. The mutated form of the gene led to improved B toxicity tolerance (Ochiai et al., 2011). Another promising target is BnaA9.WRKY47 is a positive transcriptional regulator of B uptake genes in tomatoes. The CRISPR/Cas9 mutants were similar to WT but were sensitive to B deficiency. *Bnaa9.wrky49* mutant line can perform better in B excess soils as the efficient uptake pathway is compromised. The knockout may also lead to such a desirable phenotype of the *BnaA3. NIP5;1* (*NOD26-LIKE INTRINSIC PROTEINS*), the boric acid channel gene downstream of *BnaA9. WRKY47* (Feng et al., 2020).

Iron toxicity is a significant problem in several crops under waterlogged (anaerobic) conditions. Also, Fe, an essential nutrient, must not be prevented from entering the plant. Hence targeting the sequestration and cellular tolerance mechanisms need to be prioritised. Several genes can be targeted to overcome Fe toxicity. *OsFRO1 (FERRIC REDUCTASE OXIDASE 1)* is one such gene that reduces ferric (Fe^{3+}) to ferrous (Fe^{2+}) form. RNAi-mediated knockdown of OsFRO1 imparted enhanced tolerance to Fe toxicity in rice (Li et al., 2019). Other targetable signalling modules include OsHRZ1 (*HEMERYTHRIN MOTIF-CONTAINING INTERESTING NEW GENE AND ZINC-FINGER PROTEIN 1*), *OsPRI1 (POSITIVE REGULATOR OF IRON HOMEOSTASIS 1), OsIRO2/3 (IRON-RELATED BHLH TRANSCRIPTION FACTOR 2 and 3),* and *NAS1/2 (NICOTIANAMINE AMINOTRANSFERASE 1 and 2)* to develop iron toxicity tolerant rice (Zhang et al., 2017).

Cadmium is a major toxic heavy metal that can enter the food chain, especially through staple foods like rice. OsCCX2/OsCTD1 (CATION/Ca EXCHANGER 2 or CADMIUM TOLERANCE 1) is an efflux channel protein expressed in stem nodes and flowers of rice which transports Cd to aerial parts. The knockout lines exhibited a lower Cd content as well as translocation ratio. The noticeable advantage was that Cd accumulation in the grains was significantly reduced in the mutant lines without compromising yield (Hao et al., 2018). Another Cd transport protein being targeted is OsNRAMP5, whose knockout mutants reduced the metal content in shoots and grains. The inhibition in entry occurred at the roots' level, and the phenotypic differences were absent between WT and *osnramp5* lines under normal conditions (Tang et al., 2017). In another study by Songmei et al. (2019), mutant lines were generated for OsNRAMP5 (nramp5 × 7 and nramp5 × 9) using CRISPR/Cas9-mediated mutagenesis. Osnramp5 mutants showed low grain Cd accumulation (<0.06 mg/kg). However, only nramp5×7 showed normal growth and yield. Similar results were obtained from knockout lines of the *OsLCT1* (*LOW-AFFINITY CATION TRANSPORTER 1*) gene, expressed in the uppermost node and vascular bundles in rice. In *oslct1* lines, most of the agronomic and yield traits were similar to WT, with 40% lower Cd accumulation (Songmei et al., 2019).

In Arabidopsis, the targeting of *MYB49* led to reduced Cd accumulation. Also, it was found that the abscisic acid-dependent transcription factor ABI5 (ABSCISSIC ACID INSENSITIVE 5) prevents MYB49 activity by direct interaction. It also explains why ABA reduces the uptake of Cd in plants (Zhang et al., 2019). Recently Jia et al. (2022), reported that the knockout of NtNRAMP3 increases Cd tolerance by reducing cytosolic Cd accumulation in tobacco. Specific promising targets to improve Cd tolerance by genome editing include miR390 from rice, a negative regulator of Cd tolerance. In support of this, it was found that the overexpressed lines of miR390 were more sensitive to Cd treatment (Ding et al., 2016). Another set of promising targets is *IRT1* (*IRON-REGULATED TRANSPORTER 1*), *HIPP22, and HIPP44 (HEAVY METAL-ASSOCIATED ISOPRENYLATED PLANT PROTEINs)* from Arabidopsis (Zhang and Reynolds, 2019). From *Triticum turgidum,* the *TtNRAMP6* was also suggested to be a promising target to raise Cd-free wheat (Wang et al., 2019a). In rice root cells, OsNRAMP5 is a key gene involved in controlling the uptake of Cd, Mn, and other metal ions. Knocking out the metal transporter gene OsNRAMP5 by CRISPR/Cas9 reduced Cd accumulation in rice without significantly affecting yield (Tang et al., 2017). Thus, CRISPR-aided genome engineering holds the potential to develop plants with a high level of nutrients along with the reduced amount of toxic metal and anti-nutritional factors for better health.

Genome Editing Targets and Case Studies on Improving Grain Nutrient Availability and Biofortification

The deficiency of micronutrients or hidden hunger is more pronounced in developing countries. Xie et al. (2015) developed a strategy to produce numerous gRNAs from a single polycistronic gene using an endogenous tRNA-processing system, which precisely cleaves both ends of the tRNA precursor. This system could boost the targeting and multiplex editing capability of the CRISPR/Cas9 system. Using this strategy, three transcription factor genes were targeted (*MADS, MYBR*, and *AP2*) for simplex editing and three other genes (*RPL, PPR*, and *IncRNA*) for multiplex editing. They achieved stable transgenic rice plants with high efficiency (up to 100%). Because tRNA and its processing system are virtually conserved in all living organisms, this method could be broadly used to boost the targeting capability and editing efficiency of CRISPR/Cas9 toolkits. CRISPR-edited plants are free from any foreign DNA; thus, they may have better acceptability compared to traditional GM crops. CRISPR edited mushrooms (Kim and Kim, 2016), false flax with increased oil content, and a drought-tolerant soybean (Waltz, 2018) were able to clear government regulation to reach the market, indicating the potential of CRISPR-edited crops in revolutionizing crop improvement. Developing crop plants with improved nutrient content, reduced toxic metal content, and reduced anti-nutritional factors will benefit better human health. Development of plants with enhanced nutrient content will have a double advantage as improvement in the nutrient content of plants will increase the yield potential, and along with this, it will also increase the flow of nutrients in the food chain as plants are the most important primary producer of our food chain. Genome editing can be exploited to improve the nutritional content of crop plants.

The anti-nutritional factors like phytic acid, protease inhibitors, glucosinolates, lectins, tannins, saponins, amylase inhibitors, reduces bioavailability of nutrients. CRISPR/Cas9 genome editing can be utilized to reduce anti-nutritional factors in plants. In cereal grains, phytate serve as a P store, at the same time, it is considered an anti-nutrient because of its low digestibility. Phytic acid limits the utilization of phosphate as it is present mainly in the organic form of inositol hexakisphosphate (IP6). Hence, low IP6 content can improve crops and grains' phosphate and mineral bioavailability. Inositol trisphosphate 5/6 kinases (ITPK) enzyme catalyzes the phosphorylation of inositol phosphate to inositol hexakisphosphate, a major phosphate storage form in cereal grains. In Barley, homozygous mutants

for the *HvITPK1* gene were generated by CRISPR/Cas9 to elucidate the role of *HvITPK1* in inositol hexakisphosphate synthesis and stress signalling (Vlčko and Ohnoutkova, 2020). The mature grains' phosphate content was variable, from 65 to 174% compared to wild type (WT). Among 11 mutants highest increase in phosphate content, of 74%, was detected in the homozygous deletion mutant itpk1-14. The insertion mutants showed improved tolerance to salinity stress with a concomitant decrease in grain P content. In wheat the CRISPR/Cas9 editing of *TaIPK1* effectively reduced the phytate content and thereby increased the grain Fe and Zinc (Zn) content (Ibrahim et al., 2022).

Nearly three billion people suffer from deficiency of micronutrients and vitamins, which affects their growth, development, and immunity, increasing the risk for infectious illness. Biofortification, the increase in the content of bioavailable micronutrients in edible parts of staple food crops such as rice, wheat, maize, etc., is considered an effective strategy to provide balanced diets with enriched levels of vitamins and minerals for better human health. The deficiency of micronutrients is attributed to lower uptake of nutrients by plants and accumulation in edible plant parts, presence of high levels of inhibitors affecting their absorption. CRISPR/Cas9 genome editing can be utilized as a powerful tool in manipulating these pathways to increase the micronutrient and vitamin content in plants. In rice and banana, there have been successful attempts to increase carotenoid range by CRISPR/Cas9. Kaur et al. (2018) successfully demonstrated that genome editing through CRISPR/Cas9 can be applied as an efficient tool for banana genome modification by creating mutation in phytoene desaturase (RAS-PDS) of banana cv. Rasthali. Further, in 2020, they developed the β-carotene-enriched Cavendish banana cultivar (cv.) Grand Naine (AAA genome). In the carotenoid biosynthesis pathway, lycopene is bifurcated by lycopene cyclases, i.e., lycopene beta-cyclase (LCYβ) and lycopene epsilon-cyclase (LCYε) into β-carotene and α-carotene, respectively. A low activity level of lycopene epsilon-cyclase will increase the availability of lycopene for β-carotene. Thus, the Lycopene Epsilon-Cyclase (LCYε) gene was targeted for CRISPR gene editing. Point mutations/premature termination of LCYε protein multiple types of indels in the genomic region of Grand Naine LCYε (GN-LCYε) were obtained. In edited plants the β-carotene content was improved by 6-fold (~24 μg/g) in comparison to non-edited plants (Dong et al., 2020). These results demonstrate that CRISPR-Cas9 genome editing is a promising strategy for the genetic improvement of rice and other crops. The targeted gene expression strategy can be utilised to insert genes involved in the uptake and translocation of micronutrients to improve their content in plants.

The sgRNA guided genome editing system, using type II CRISPR/Cas9, has been demonstrated by targeting the *Phytoene desaturase* (*ClPDS*) gene to create knockout mutations in *C. lanatus* (Tian et al., 2017). The *ClPDS* gene is essential for chlorophyll biosynthesis and acts as a critical enzyme in the biosynthesis of carotenoids (Tian et al., 2017). The phytates in protein storage bodies chelate with several mineral cations, including $Zn^{2+,}$ Fe^{2+}, Ca^{2+}, and Mg^{2+} in grains. During seed germination to utilize the stored mineral ions in developing seedling, endogenous grain phytase is activated to degrade phytate, releasing myo-inositol, P, and bound mineral cations. The crop plants with low levels of IP6 content may have higher phosphate and mineral bioavailability. The inositol trisphosphate five and inositol trisphosphate six kinases enzymes (ITPK's) participate in the sequential phosphorylation of inositol phosphate to inositol hexakisphosphate, an effective phosphate storage form in cereal grains. In Barley, homozygous mutants for the *HvITPK1* gene were generated by CRISPR/Cas9 to elucidate the role of HvITPK1 in inositol hexakisphosphate synthesis and stress signalling (Vlčko and Ohnoutkova, 2020). The mutation in *HvIPTK1* altered phosphate levels from 65 to 174% in the mature grains compared to wild type content. Among 11 mutants highest increase in phosphate content, of 74%, was detected in the homozygous deletion mutant *itpk1-14*. On the contrary, mutant insertion lines revealed a higher tolerance to salinity stress than deletion mutants and reduced grain P content.

Nucleotide substitutions of *OsITPK6* could significantly reduce rice grain phytic acid content. Jiang et al. (2019) targeted the first exon of *OsITPK6* using the CRISPR/Cas9 method. In the four OsITPK6 mutant lines, one (*ositpk6_1*) is with a 6-bp deletion (no change in frame), and the other three with frameshift mutations (*ositpk6_2, _3, and _4*). In frameshift mutant lines, plant growth and reproduction were severely impaired. At the same time, the effect of the in-frame mutation in *ositpk6_1* was relatively limited. The mutant lines *ositpk6_1* and *6_2* had significantly low phytic acid content and higher inorganic P levels than the WT.

Researchers utilised the CRISPR/Cas9 technique to engineer browning genes in mushrooms, where minimising browning in white truffles led to the extension of their life span, thereby offering a business advantage. The deletion mutant of the polyphenol oxidase gene reduced the oxidation of polyphenols when exposed to air, bringing an appetising and attractive appearance to the mushrooms (Waltz, 2018). Similarly, the mutation in the gene that encoded the polyphenol oxidase (PPO) enzyme in apples led to a reduction in browning in cut apples (Nishitani et al., 2016). Other workers reported the technology application by developing acrylamide free potatoes (Halterman et al., 2016) and low phytic acid with higher available P content corn line (Liang et al., 2014). Rice, a principal food crop, has been the primary target for quality enrichment. Amylose is considered a vital nutritional quality parameter of rice grains. Targeted editing of the starch branching enzymes produced mutant lines with lower amylose content in rice grains (Sun et al., 2017). In soybean, targeted modification of the omega-6 desaturase (*GmFAD2*) gene via the CRISPR/Cas9 method resulted in the higher accumulation of oleic acids via a reduction in the linoleic and a-linolenic acids (Al Amin et al., 2019).

Strategies to Exploit Potent Regulatory Genes for Improving NtUE

Single-nucleotide polymorphisms regulate agronomically important traits, including the determinants of NtUE. Base editing and prime editing techniques are recently developed precise genome editing techniques which can generate/correct a point mutation without inducing a DSB. The base editors are classified as cytosine base editors (CBEs; C: G to T: A) and adenine base editors (ABEs; A: T to G: C). Cytosine base editors consist of dead Cas9 (dCas9)/Cas9 nickase, cytidine deaminases, and a uracil DNA glycosylase inhibitor. In ABEs, cytosine deaminase is replaced by a mutant *Escherichia coli* transfer RNA adenosine deaminase (TadA). The sgRNA guides the CBE system to the target region of DNA and induces R-loop formation. Then, the cytidine deaminase deaminates the cytosine into uracil in the non-target DNA strand, and subsequent DNA repair and replication results in cytosine to thymine base conversion (Jiang et al., 2020). In ABEs, adenine is deaminated into inosine in the non-target strand, and subsequent DNA repair and replication cause adenine to guanine base conversion. The CBE system was first established for plants by creating a point mutation in the nitrate transporter NRT1.1B and Slender Rice 1 (SLR1, a DELLA repressor) gene in rice plants (Lu and Zhu, 2017). Recently a new TadA variant, TadA9 was identified, which has high-efficiency, multiplex adenine base editing and compatible with CRISPR/SpCas9, CRISPR/SpCas9-NG, CRISPR/SpRY and CRISPR/ScCas9 nickase thus enhancing its editing window at diverse PAM sites. The identification of class 2 CRISPR effector proteins, Cpf1 (Clustered regularly interspaced short palindromic repeats from prevotella and francisella 1) lead to the development of an advanced genome editing system, i.e. CRISPR/Cpf1. This new system of base editors (CRISPR/Cpf1) is evolved as more accurate and efficient genome editing as they have the advantages of low molecular weight and reduced off-target activity (Alok et al., 2020). The reduced size of Cpf1 proteins compared to Cas9 makes it possible to design smaller vectors and better transformation efficiency.

Prime editing technique is also known as "search and replace genome" editing technique (Lin et al., 2020). The prime editing guide RNA (pegRNA) containing a desired edits is fused to Cas9 nickase and reverse transcriptase (RT). The pegRNA has a primer binding site and an RT template including edited nucleotides. The Cas9 nickase and RT complex is guided by the pegRNA to the target DNA followed by Cas9 mediated nick formation. The reverse transcriptase then uses the edited pegRNA as a template and starts reverse transcription from the nick site. The newly edited DNA strand replaces the original DNA strand and makes a hetero-duplex of one edited and one unedited DNA strand. In the final stage, the edited DNA strand acts as a template over non-edited DNA strand and the DNA repair mechanism repairs the non-edited DNA strand. Prime editing holds great promise for precision breeding for developing superior crops for traits, such as increasing yield, stress tolerance and resource use efficiency.

CRISPR/dCas9 system also offers the added advantage of exploiting transcriptional activators/repressors for overexpression of genes of interest in a targeted manner. CRISPR activation (CRISPRa) and CRISPR interference (CRISPRi) are variants of CRISPR in which a dCas9 dependent transcription activation or interference is achieved. To improve guide RNA efficiency, specialized peptide epitopes are fused to the dCas9. The recently developed dCas9-SunTag system (Tanenbaum et al., 2014), was employed to engineer targeted gene activation and DNA methylation in Arabidopsis (Papikian et al., 2019).

Cis-engineering is observed to be a more promising way to regulate the gene expression because of lower detrimental pleiotropic effects in comparison to coding sequences (Pandiarajan and Grover, 2018; Wolter and Puchta, 2018). But very limited research has been carried out in plants where target gene expression was regulated by insertion of new cis regulatory element (CRE) or disruption/deletion of existing CREs. One such example of deletion was studied in rice where disease resistance was improved by regulatory fragment deletion spanning 149bp which included a transcription-activator-like effector (TALe)-Binding Element (EBE) in SWEET11's (SUGARS WILL EVENTUALLY BE EXPORTED TRANSPORTERS 11) promoter (Li et al., 2020). Insertion of transposable elements (TEs) can alter the expression of desirable genes. In tomato, the insertion of a transposable element in tomato homolog of *SEPALLATA4* and 564 bp in *FRUITFULL* homolog improved floral architecture along with higher fruit weight, number and yield (Soyk et al., 2017). Gain of function of alleles can also be obtained by cis-engineering involving CRISPR/Cas to alter CREs of introns and other downstream genes. One notable example is CArG element disruption downstream of SlWUS including two SNPs (Li et al., 2018b).

Promoter insertion and swapping can potentially be used to improve nutrient use efficiency which can be achieved by HDR. Under drought conditions maize yield was improved by moderate overexpression of *AGROS8,* an ethylene response negative regulator using site-specific insertion in 5′ UTR using Cas9 induced DSBs (Shi et al., 2017). Additionally, the epigenome can also be edited using CRISPR/Cas mediated cis-engineering. Although a very limited attempt has been made to describe epigenome editing involving alteration in histone acetylation (Roca Paixão et al., 2019) and DNA methylation (Gallego-Bartolomé et al., 2018; Papikian et al., 2019). Upstream open reading frames (uORFs) could be used to improve nutrient use efficiency. uORFs are protein-coding short elements localised in the 5' leader region which controls the quantity of downstream primary ORFs (pORFs) which are synthesised (von Arnim et al., 2014). In plants, the uORFs occupy 30–40% of the transcript proportion(von Arnim et al., 2014). Previously rice plant immunity was improved by the insertion of uORF upstream of pORF without yield reduction (Xu et al., 2017). To date, very limited research was conducted to improve nutrient use efficiency using CRISPR/Cas genome editing targeting positive regulators.

Transporters are one of the important positive regulators whose activity can be enhanced using CRISPR/Cas based genome editing. One attempt has been made to improve the activity of OsNRT1.1B using dCas9 (D10A) fused with rat's cytidine deaminase enzyme APOBEC1 to replace C to T resulting in Thr327Met mutation.

TABLE 1 | Compilation of genome editing case studies and targets for improving nutrient response and toxicity tolerance.

Trait	Genome editing targets	Functions	Effects of targeted editing or RNAi suppression	References
Nitrogen use efficiency	*LATERAL ORGAN BOUNDARY DOMAIN (LBD) gene family transcription factors (LBD37, LBD38, and LBD39)*	Repressors of anthocyanin biosynthesis and nitrate-responsive genes including NRT2.1, NRT2.2, NIA1, and NIA2		Rubin et al. (2009)
	SPL9 At upstream	Reduces the transcript levels of NRT1.1, NIA2, and NIR.	—	Krouk et al. (2010)
	Leaf width 5 (LW5)	LW5 is an allele of D1, encoding the rice G protein α subunit	The loss of LW5 leads to an increase in photosynthesis, N uptake, and chlorophyll content	Zhu et al. (2020)
Phosphorous use efficiency	*MYB1*	Negative regulator of Pi starvation signalling and Pi transporters (*PHT1;2, PHT1;9, PHT1;10, IPS, miR399j* and *PH O 2.1*)	Mutation of *MYB1* transcription factor using CRISPR Cas9 system showed an increase in Pi uptake and accumulation	Gu et al. (2017)
Salinity tolerance	*RECEPTOR FOR ACTIVATED C KINASE 1 (RACK1)*	Negative regulator of ABA responses in plants and	Constitutive RNAi-mediated down-regulation of RACK1 homolog in soybean has been found to increase drought and salt tolerance	Zhang et al. (2013), Chen et al. (2018), Zheng et al. (2019)
	STRESS-ASSOCIATED PROTEIN 1(SAP1)	—	Down-regulation of PagSAP1 in poplar enhances salinity tolerance by an increase in stress response genes	Yoon et al. (2018)
	OsRR22	Transcription factor involved in cytokinin signal transduction and metabolism	Loss of function and CRISPR/CAS mediated editing of OsRR22 resulted in salt tolerance in rice	Takagi et al. (2015), Zhang et al. (2019)
Metal toxicity	*OsAUX3*	Auxin influx transporter	The *Osaux3* mutants generated by CRISPR/Cas9 were insensitive to Al and accumulated less Al in tissues under Al toxicity	Wang M. et al., 2019
	OsARM1 (ARSENITE-RESPONSIVE MYB1)	—	CRISPR/Cas9 based genome editing of *OsARM1* resulted in reduced As uptake and As tolerance of rice	Wang et al. (2017)
	BET1 (BORON EXCESS TOLERANT 1)	—	Loss of function mutation improved B toxicity tolerance	Ochiai et al. (2011)
	OsCCX2/OsCTD1 (CATION/Ca EXCHANGER 2 or CADMIUM TOLERANCE 1)	Cadmium efflux channel protein expressed in stem nodes and flower of rice transports Cd to aerial parts	The knockout lines showed a lower Cd content	Hao et al. (2018)
	OsFR O 1 (FERRIC REDUCTASE OXIDASE 1)	Reduces ferric (Fe^{3+}) to ferrous (Fe^{2+}) iron	RNAi mediated knockdown of *OsFRO1* imparted enhanced tolerance to Fe toxicity in rice	Shou et al. (2019)
Biofortification	*OsNramp5*	Transport of Cadmium (Cd) to grains	CRISPR/Cas9 mediated knocking out of metal transporter *OsNramp5* it reduced Cd accumulation in rice grains	Tang et al. (2017)
	OsITPK6 (Inositol trisphosphate kinase 6)	ITPK enzymes catalyse the sequential phosphorylation of inositol phosphate to inositol hexakisphosphate/phytic acid, the primary storage form of phosphate in cereal grains	CRISPR/Cas9 mutant lines *Ositpk6_1* and *Ositpk6_2* had low phytic acid and higher levels of grain P	Jiang et al. (2019)
	GmFAD2		Targeted modification of the GmFAD2 by CRISPR/Cas9 increased the accumulation of oleic acids and reduced the linoleic and linolenic acids	Al Amin et al. (2019)

Though improvement in N use efficiency was not tested (Lu and Zhu 2017). Some key changes can enhance the activity of transporter viz., Tyr312Asp in PHT1; 1 for improved Pi transport (Fontenot et al., 2015), Phe130Se in HAK5 to increase the affinity to K^+ (Alemán et al., 2014). Nitrate transport can be increased by enhancing the activity of NRT1s and NRT2s. CRISPR activator led increase in binding of TFs NF-Y, NAC2, MADS57 and NLP4 may enhance the nitrate transport (He et al., 2015; Qu et al., 2015; Huang et al., 2019; Wu et al., 2020). The expression of *NRT1* can be augmented by increasing the CAT box and CArG motifs in the NRT1 promoter (Qu et al., 2015; Huang et al., 2019). Phosphate transport can be improved by amplifying the PHT1 activity which can be achieved by CRISPR activator led enhancement in the binding of PHR1 to P1BS, MYCS to P1BS (Chen et al., 2011). PHO2 negatively regulates the PHT1 activity involving ubiquitination (Liu et al., 2014). SPX1,2 of rice also acts as a negative regulator of PHT1 (Zhou et al., 2015; Gu

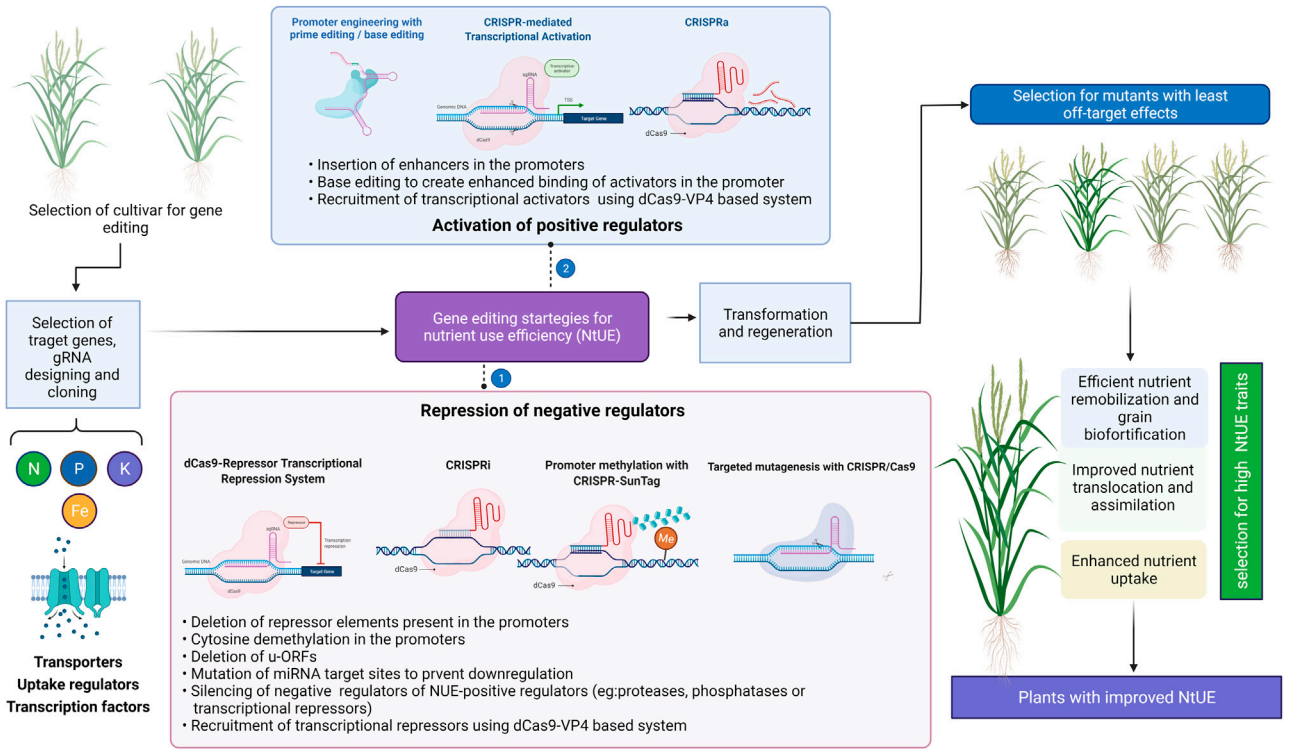


FIGURE 5 | Gene editing strategies for improving nutrient use efficiency (NtUE). The approaches to utilize the positive and negative regulators of the trait of interest are presented. Illustration was made with Biorender (https://biorender.com/).

et al., 2016). Repression of these negative regulators by dCas9 with specific gRNA may improve P transport. W-Box, P1BS, MYCS and other motifs can also be engineered to enhance Pi transport in plants (Ceasar et al., 2022). The expression of the *AMT1* plant gene family may be upregulated by enhancing the binding of IDD10 to DOF18 TFs through a CRISPR activator (Xuan et al., 2013; Wu et al., 2017). Further CRISPR activator mediated binding of RAP2.11, TFII_A, bHLH121, DDF2 and JLO TFs could amplify the expression of *HAK* transporter. GCC-box and AuxREs motifs can be engineered into the promoter of HAK to enhance its expression through CRISPR prime editing (PE) (Hong et al., 2013). bZIP19 and 23 binding to ZDRE can be complemented by CRISPR activator while CRISPR PE could be employed to increase ZDRE motifs elicitation of *ZIP* genes (Ceasar et al., 2022). Similarly other nutrient transcriptional factors and motifs could be engineered to improve the nutrient use efficiency.

SUMMARY AND CONCLUSION

CRISPR/Cas9 has evolved as cutting-edge technology for better understanding of gene functions, characterising molecular networks, and improving the yield, nutritional content, NtUE, and tolerance to biotic and abiotic stresses. A customized small sgRNA directs Cas9 nuclease at a specific genomic location. Cas9 makes a double-stranded break three to four bp upstream to the PAM site at the target site. Subsequently, the cell's natural repair mechanism of HDR or NHEJ repairs double-strand breakage (Miglani, 2017). As NHEJ repair is error-prone, it creates random insertions/deletions and indels, resulting in frameshift mutations and targeted gene knockouts (Jinek et al., 2012a; Feng et al., 2013). Compared to earlier genome editing methods, i.e. zinc-finger nucleases and transcription activator-like effector nucleases, the CRISPR/Cas9 offers the accurate and efficient targeted modification in the genome of any organism in a simple way. The specific genome editing targets for improving nutrient response and nutrient stress tolerance are compiled in **Table 1**. The major genome editing strategies for improving NtUE is presented in **Figure 5**. Further development in the CRISPR/Cas9 system allows for the editing of multiple genes at one time (Donohoue et al., 2018), and mutations can be targeted to the untranslated region of coding genes (Mao et al., 2018), promoter regions (Seth, 2016), microRNAs (Chang et al., 2016), non-coding RNAs (ncRNAs) (Wang et al., 2018c). CRISPRi and CRISPRa are among the development of technology in which transcription factors are fused with dCas9 to repress or enhance transcription by RNA polymerase and subsequently, upregulate or downregulate the expression of a gene/genes of interest. Researchers can effectively use the knowledge of genes regulating nutrient homeostasis in plants and advanced techniques in genome editing to design plants with desired traits. The development of resource use efficient, high NtUE plants through CRISPR-Cas technology will enhance the pace of genetic improvement for yield.

AUTHOR CONTRIBUTIONS

LS, BJ, DB, GK, SA, PP, and SN wrote the first draft. BJ and PP prepared the figures. LS revised the figures and compiled the first draft. LS, VC, SA, PP, BJ, ST, and SJ revised the manuscript. LS and VC conceived the idea and finalized the manuscript.

FUNDING

The work on genome editing at VC's lab is funded by NASF (ICAR), New Delhi, Grant Nos. F. No. NASF/CRISPR-Cas-7003/2017-18 and DBT, Govt of India, Grant Nos. BT/PR25820/GET/119/100/2017. LS acknowledges funding by DST-SERB F. No. ECR/2017/002982.

REFERENCES

Ahmed, S., Zhang, Y., Abdullah, M., Ma, Q., Wang, H., and Zhang, P. (2019). Current Status, Challenges, and Future Prospects of Plant Genome Editing in China. *Plant Biotechnol. Rep.* 13, 459–472. doi:10.1007/s11816-019-00577-6

Al Amin, N., Ahmad, N., Wu, N., Pu, X., Ma, T., Du, Y., et al. (2019). CRISPR-Cas9 Mediated Targeted Disruption of FAD2-2 Microsomal Omega-6 Desaturase in Soybean (Glycine max.L). *BMC Biotechnol.* 19, 9–10. doi:10.1186/s12896-019-0501-2

Alboresi, A., Gestin, C., Leydecker, M.-T., Bedu, M., Meyer, C., and Truong, H.-N. (2005). Nitrate, a Signal Relieving Seed Dormancy in Arabidopsis. *Plant Cell. Environ.* 28, 500–512. doi:10.1111/j.1365-3040.2005.01292.x

Alemãn, F., Caballero, F., Rã³denas, R., Rivero, R. M., MartÃ-nez, V., and Rubio, F. (2014). The F130S Point Mutation in the Arabidopsis High-Affinity K+ Transporter AtHAK5 Increases K+ over Na+ and Cs+ Selectivity and Confers Na+ and Cs+ Tolerance to Yeast under Heterologous Expression. *Front. Plant Sci.* 5. doi:10.3389/fpls.2014.00430

Alfatih, A., Wu, J., Zhang, Z.-S., Xia, J.-Q., Jan, S. U., Yu, L.-H., et al. (2020). Rice NIN-LIKE PROTEIN 1 Rapidly Responds to Nitrogen Deficiency and Improves Yield and Nitrogen Use Efficiency. *J. Exp. Bot.* 71, 6032–6042. doi:10.1093/jxb/eraa292

Alok, A., Sandhya, D., Jogam, P., Rodrigues, V., Bhati, K. K., Sharma, H., et al. (2020). The Rise of the CRISPR/Cpf1 System for Efficient Genome Editing in Plants. *Front. Plant Sci.* 11, 264. doi:10.3389/FPLS.2020.00264/BIBTEX

Alvarez-Fernández, A., Díaz-Benito, P., Abadía, A., López-Millán, A. F., and Abadía, J. (2014). Metal Species Involved in Long Distance Metal Transport in Plants. *Front. Plant Sci.* 5, 105. doi:10.3389/fpls.2014.00105

Anderson, S. A., Nizzi, C. P., Chang, Y.-I., Deck, K. M., Schmidt, P. J., Galy, B., et al. (2013). The IRP1-HIF-2α Axis Coordinates Iron and Oxygen Sensing with Erythropoiesis and Iron Absorption. *Cell. Metab.* 17, 282–290. doi:10.1016/j.cmet.2013.01.007

Aoyama, T., Kobayashi, T., Takahashi, M., Nagasaka, S., Usuda, K., Kakei, Y., et al. (2009). OsYSL18 Is a Rice Iron(III)-deoxymugineic Acid Transporter Specifically Expressed in Reproductive Organs and Phloem of Lamina Joints. *Plant Mol. Biol.* 70, 681–692. doi:10.1007/S11103-009-9500-3

Aranda-Sicilia, M. N., Aboukila, A., Armbruster, U., Cagnac, O., Schumann, T., Kunz, H.-H., et al. (2016). Envelope K+/H+ Antiporters AtKEA1 and AtKEA2 Function in Plastid Development. *Plant Physiol.* 172, 441–449. doi:10.1104/pp.16.00995

Armbruster, U., Carrillo, L. R., Venema, K., Pavlovic, L., Schmidtmann, E., Kornfeld, A., et al. (2014). Ion Antiport Accelerates Photosynthetic Acclimation in Fluctuating Light Environments. *Nat. Commun.* 5, 5439. doi:10.1038/ncomms6439

Balyan, H. S., Gahlaut, V., Kumar, A., Jaiswal, V., Dhariwal, R., Tyagi, S., et al. (2016). Nitrogen and Phosphorus Use Efficiencies in Wheat: Physiology, Phenotyping, Genetics, and Breeding. *Plant Breed. Rev.* 40, 167–234. doi:10.1002/9781119279723.CH4

Bari, R., Datt Pant, B., Stitt, M., and Scheible, W.-R. (2006). PHO2, microRNA399, and PHR1 Define a Phosphate-Signaling Pathway in Plants. *Plant Physiol.* 141, 988–999. doi:10.1104/PP.106.079707

Behera, S., Long, Y., Schmitz-Thom, I., Wang, X. P., Zhang, C., Li, H., et al. (2017). Two Spatially and Temporally Distinct Ca 2+ Signals Convey *Arabidopsis thaliana* Responses to K + Deficiency. *New Phytol.* 213, 739–750. doi:10.1111/nph.14145

Boonyaves, K., Gruissem, W., and Bhullar, N. K. (2016). NOD Promoter-Controlled AtIRT1 Expression Functions Synergistically with NAS and FERRITIN Genes to Increase Iron in Rice Grains. *Plant Mol. Biol.* 90, 207–215. doi:10.1007/S11103-015-0404-0

Brauer, E. K., Ahsan, N., Dale, R., Kato, N., Coluccio, A. E., Piñeros, M. A., et al. (2016). The Raf-like Kinase ILK1 and the High Affinity K+ Transporter HAK5 Are Required for Innate Immunity and Abiotic Stress Response. *Plant Physiol.* 171, 1470–1484. doi:10.1104/pp.16.00035

Buet, A., Galatro, A., Ramos-Artuso, F., and Simontacchi, M. (2019). Nitric Oxide and Plant Mineral Nutrition: Current Knowledge. *J. Exp. Bot.* 70, 4461–4476. doi:10.1093/jxb/erz129

Castro Marín, I., Loef, I., Bartetzko, L., Searle, I., Coupland, G., Stitt, M., et al. (2011). Nitrate Regulates Floral Induction in Arabidopsis, Acting Independently of Light, Gibberellin and Autonomous Pathways. *Planta* 233, 539–552. doi:10.1007/s00425-010-1316-5

Ceasar, S. A., Maharajan, T., Hillary, V. E., and Krishna, T. P. A. (2022). Insights to Improve the Plant Nutrient Transport by CRISPR/Cas System. *Biotechnol. Adv.* 59, 107963. doi:10.1016/j.biotechadv.2022.107963

Chandrasegaran, S., and Carroll, D. (2016). Origins of Programmable Nucleases for Genome Engineering. *J. Mol. Biol.* 428, 963–989. doi:10.1016/j.jmb.2015.10.014

Chang, H., Yi, B., Ma, R., Zhang, X., Zhao, H., and Xi, Y. (2016). CRISPR/cas9, a Novel Genomic Tool to Knock Down microRNA *In Vitro* and *In Vivo*. *Sci. Rep.* 6, 22312–12. doi:10.1038/srep22312

Chanroj, S., Padmanaban, S., Czerny, D. D., Jauh, G.-Y., and Sze, H. (2013). K+ Transporter AtCHX17 with its Hydrophilic C Tail Localizes to Membranes of the Secretory/endocytic System: Role in Reproduction and Seed Set. *Mol. Plant* 6, 1226–1246. doi:10.1093/mp/sst032

Che, J., Tsutsui, T., Yokosho, K., Yamaji, N., and Ma, J. F. (2018). Functional Characterization of an Aluminum (Al)-Inducible Transcription Factor, ART2, Revealed a Different Pathway for Al Tolerance in Rice. *New Phytol.* 220, 209–218. doi:10.1111/nph.15252

Chen, A., Gu, M., Sun, S., Zhu, L., Hong, S., and Xu, G. (2011). Identification of Two Conserved Cis -acting Elements, MYCS and P1BS, Involved in the Regulation of Mycorrhiza-activated Phosphate Transporters in Eudicot Species. *New Phytol.* 189, 1157–1169. doi:10.1111/j.1469-8137.2010.03556.x

Chen, C., Fenk, L. A., and de Bono, M. (2013). Efficient Genome Editing in *Caenorhabditis elegans* by CRISPR-Targeted Homologous Recombination. *Nucleic Acids Res.* 41, e193. doi:10.1093/nar/gkt805

Chen, G., Hu, Q., Luo, L., Yang, T., Zhang, S., Hu, Y., et al. (2015). Rice Potassium Transporter OsHAK1 Is Essential for Maintaining Potassium-Mediated Growth and Functions in Salt Tolerance over Low and High Potassium Concentration Ranges. *Plant Cell. Environ.* 38, 2747–2765. doi:10.1111/pce.12585

Chen, Y., Ma, J., Miller, A. J., Luo, B., Wang, M., Zhu, Z., et al. (2016). OsCHX14 Is Involved in the K+ Homeostasis in Rice (Oryza Sativa) Flowers. *Plant Cell. Physiol.* 57, 1530–1543. doi:10.1093/pcp/pcw088

Chen, Y., Han, Y.-H., Cao, Y., Zhu, Y.-G., Rathinasabapathi, B., and Ma, L. Q. (2017). Arsenic Transport in Rice and Biological Solutions to Reduce Arsenic Risk from Rice. *Front. Plant Sci.* 8, 268. doi:10.3389/fpls.2017.00268

Chen, Z. C., Yamaji, N., Fujii-Kashino, M., and Ma, J. F. (2016). A Cation-Chloride Cotransporter Gene Is Required for Cell Elongation and Osmoregulation in Rice. *Plant Physiol.* 171, 494–507. doi:10.1104/pp.16.00017

Chiou, T.-J., and Lin, S.-I. (2011). Signaling Network in Sensing Phosphate Availability in Plants. *Annu. Rev. Plant Biol.* 62, 185–206. doi:10.1146/annurev-arplant-042110-103849

Cuin, T., Dreyer, I., and Michard, E. (2018). The Role of Potassium Channels in *Arabidopsis thaliana* Long Distance Electrical Signalling: AKT2 Modulates Tissue Excitability while GORK Shapes Action Potentials. *Ijms* 19, 926. doi:10.3390/ijms19040926

Curie, C., Cassin, G., Couch, D., Divol, F., Higuchi, K., Le Jean, M., et al. (2009). Metal Movement within the Plant: Contribution of Nicotianamine and Yellow Stripe 1-like Transporters. *Ann. Bot.* 103, 1–11. doi:10.1093/AOB/MCN207

Curie, C., Panaviene, Z., Loulergue, C., Dellaporta, S. L., Briat, J.-F., and Walker, E. L. (2001). Maize Yellow Stripe1 Encodes a Membrane Protein Directly Involved in Fe(III) Uptake. *Nature* 409, 346–349. doi:10.1038/35053080

Demidchik, V. (2018). ROS-activated Ion Channels in Plants: Biophysical Characteristics, Physiological Functions and Molecular Nature. *Ijms* 19, 1263. doi:10.3390/ijms19041263

Devaiah, B. N., Madhuvanthi, R., Karthikeyan, A. S., and Raghothama, K. G. (2009). Phosphate Starvation Responses and Gibberellic Acid Biosynthesis Are Regulated by the MYB62 Transcription Factor in Arabidopsis. *Mol. Plant* 2, 43–58. doi:10.1093/mp/ssn081

DiDonato, R. J., Roberts, L. A., Sanderson, T., Eisley, R. B., and Walker, E. L. (2004). Arabidopsis Yellow Stripe-Like2 (YSL2): A Metal-Regulated Gene Encoding a Plasma Membrane Transporter of Nicotianamine-Metal Complexes. *Plant J.* 39, 403–414. doi:10.1111/J.1365-313X.2004.02128.X

Ding, F., Hu, Q., Wang, M., and Zhang, S. (2018). Knockout of SLSBPASE Suppresses Carbon Assimilation and Alters Nitrogen Metabolism in Tomato Plants. *Ijms* 19, 4046. doi:10.3390/ijms19124046

Ding, Y., Ye, Y., Jiang, Z., Wang, Y., and Zhu, C. (2016). MicroRNA390 Is Involved in Cadmium Tolerance and Accumulation in Rice. *Front. Plant Sci.* 7, 235. doi:10.3389/fpls.2016.00235

Dong, O. X., Yu, S., Jain, R., Zhang, N., Duong, P. Q., Butler, C., et al. (2020). Marker-free Carotenoid-Enriched Rice Generated through Targeted Gene Insertion Using CRISPR-Cas9. *Nat. Commun.* 11, 1178. doi:10.1038/s41467-020-14981-y

Donohoue, P. D., Barrangou, R., and May, A. P. (2018). Advances in Industrial Biotechnology Using CRISPR-Cas Systems. *Trends Biotechnol.* 36, 134–146. doi:10.1016/j.tibtech.2017.07.007

Drechsler, N., Zheng, Y., Bohner, A., Nobmann, B., von Wirén, N., Kunze, R., et al. (2015). Nitrate-dependent Control of Shoot K Homeostasis by NPF7.3/NRT1.5 and SKOR in Arabidopsis. *Plant Physiol.* 169, 2832–2847. pp.01152.2015. doi:10.1104/PP.15.01152

Engelsberger, W. R., and Schulze, W. X. (2012). Nitrate and Ammonium Lead to Distinct Global Dynamic Phosphorylation Patterns when Resupplied to Nitrogen-starved Arabidopsis Seedlings. *Plant J.* 69, 978–995. doi:10.1111/j.1365-313X.2011.04848.x

Feng, Y., Cui, R., Wang, S., He, M., Hua, Y., Shi, L., et al. (2020). Transcription Factor BnaA9.WRKY47 Contributes to the Adaptation of Brassica Napus to Low Boron Stress by Up-regulating the Boric Acid Channel Gene BnaA3.NIP5;1. *Plant Biotechnol. J.* 18, 1241–1254. doi:10.1111/pbi.13288

Feng, Z., Zhang, B., Ding, W., Liu, X., Yang, D.-L., Wei, P., et al. (2013). Efficient Genome Editing in Plants Using a CRISPR/Cas System. *Cell. Res.* 23, 1229–1232. doi:10.1038/cr.2013.114

Fontenot, E. B., DiTusa, S. F., Kato, N., Olivier, D. M., Dale, R., Lin, W.-Y., et al. (2015). Increased Phosphate Transport ofArabidopsis Thaliana Pht1;1 by Site-Directed Mutagenesis of Tyrosine 312 May Be Attributed to the Disruption of Homomeric Interactions. *Plant Cell. Environ.* 38, 2012–2022. doi:10.1111/pce.12522

Fourcroy, P., Sisó-Terraza, P., Sudre, D., Savirón, M., Reyt, G., Gaymard, F., et al. (2014). Involvement of the ABCG 37 Transporter in Secretion of Scopoletin and Derivatives by Arabidopsis Roots in Response to Iron Deficiency. *New Phytol.* 201, 155–167. doi:10.1111/NPH.12471

Gaj, T., Sirk, S. J., Shui, S.-l., and Liu, J. (2016). Genome-editing Technologies: Principles and Applications. *Cold Spring Harb. Perspect. Biol.* 8, a023754–20. doi:10.1101/cshperspect.a023754

Gallego-Bartolomé, J., Gardiner, J., Liu, W., Papikian, A., Ghoshal, B., Kuo, H. Y., et al. (2018). Targeted DNA Demethylation of the Arabidopsis Genome Using the Human TET1 Catalytic Domain. *Proc. Natl. Acad. Sci. U.S.A.* 115, E2125–E2134. doi:10.1073/pnas.1716945115

Ganie, S. A., Wani, S. H., Henry, R., and Hensel, G. (2021). Improving Rice Salt Tolerance by Precision Breeding in a New Era. *Curr. Opin. Plant Biol.* 60, 101996. doi:10.1016/J.PBI.2020.101996

Gao, Y., Xu, Z., Zhang, L., Li, S., Wang, S., Yang, H., et al. (2020). MYB61 Is Regulated by GRF4 and Promotes Nitrogen Utilization and Biomass Production in Rice. *Nat. Commun.* 11, 5219. doi:10.1038/s41467-020-19019-x

Gierth, M., Mäser, P., and Schroeder, J. I. (2005). The Potassium Transporter AtHAK5 Functions in K+ Deprivation-Induced High-Affinity K+ Uptake and AKT1 K+ Channel Contribution to K+ Uptake Kinetics in Arabidopsis Roots. *Plant Physiol.* 137, 1105–1114. doi:10.1104/pp.104.057216

Goritschnig, S., Weihmann, T., Zhang, Y., Fobert, P., McCourt, P., and Li, X. (2008). A Novel Role for Protein Farnesylation in Plant Innate Immunity. *Plant Physiol.* 148 (1), 348–357.

Grillet, L., and Schmidt, W. (2019). Iron Acquisition Strategies in Land Plants: Not So Different after All. *New Phytol.* 224, 11–18. doi:10.1111/NPH.16005

Gu, J., Li, W., Wang, S., Zhang, X., Coules, A., Ding, G., et al. (2019). Differential Alternative Splicing Genes in Response to Boron Deficiency in Brassica Napus. *Genes.* 10, 224. doi:10.3390/genes10030224

Gu, M., Chen, A., Sun, S., and Xu, G. (2016). Complex Regulation of Plant Phosphate Transporters and the Gap between Molecular Mechanisms and Practical Application: What Is Missing? *Mol. Plant* 9, 396–416. doi:10.1016/J.MOLP.2015.12.012

Gu, M., Zhang, J., Li, H., Meng, D., Li, R., Dai, X., et al. (2017). Maintenance of Phosphate Homeostasis and Root Development Are Coordinately Regulated by MYB1, an R2R3-type MYB Transcription Factor in Rice. *J. Exp. Bot.* 68, 3603–3615. doi:10.1093/jxb/erx174

Guo, F.-Q., Wang, R., Chen, M., and Crawford, N. M. (2001). The Arabidopsis Dual-Affinity Nitrate Transporter Gene AtNRT1.1 (CHL1) Is Activated and Functions in Nascent Organ Development during Vegetative and Reproductive Growth. *Plant Cell.* 13, 1761–1777. doi:10.1105/tpc.13.8.1761

Guo, J., Wang, J., Xi, L., Huang, W.-D., Liang, J., and Chen, J.-G. (2009). RACK1 Is a Negative Regulator of ABA Responses in Arabidopsis. *J. Exp. Bot.* 60, 3819–3833. doi:10.1093/jxb/erp221

Guo, M., Ruan, W., Li, C., Huang, F., Zeng, M., Liu, Y., et al. (2015). Integrative Comparison of the Role of the PHOSPHATE RESPONSE1 Subfamily in Phosphate Signaling and Homeostasis in Rice. *Plant Physiol.* 168, 1762–1776. doi:10.1104/pp.15.00736

Gupta, R. M., and Musunuru, K., (2014). Expanding the Genetic Editing Tool Kit: ZFNs, TALENs, and CRISPR-Cas9. *J. Clin. Invest.* 124, 4154–4161. doi:10.1172/jci72992

Halterman, D., Guenthner, J., Collinge, S., Butler, N., and Douches, D. (2016). Biotech Potatoes in the 21st Century: 20 Years since the First Biotech Potato. *Am. J. potato Res.* 93, 1–20. doi:10.1007/s12230-015-9485-1

Han, M., Wu, W., Wu, W.-H., and Wang, Y. (2016). Potassium Transporter KUP7 Is Involved in K + Acquisition and Translocation in Arabidopsis Root under K + -Limited Conditions. *Mol. Plant* 9, 437–446. doi:10.1016/j.molp.2016.01.012

Hao, X., Zeng, M., Wang, J., Zeng, Z., Dai, J., Xie, Z., et al. (2018). A Node-Expressed Transporter OsCCX2 Is Involved in Grain Cadmium Accumulation of Rice. *Front. Plant Sci.* 9, 476. doi:10.3389/fpls.2018.00476

Hasan, M. M., Hasan, M. M., Teixeira da Silva, J. A., and Li, X. (2016). Regulation of Phosphorus Uptake and Utilization: Transitioning from Current Knowledge to Practical Strategies. *Cell. Mol. Biol. Lett.* 21, 7–19. doi:10.1186/s11658-016-0008-y

He, X., Qu, B., Li, W., Zhao, X., Teng, W., Ma, W., et al. (2015). The Nitrate Inducible NAC Transcription Factor TaNAC2-5A Controls Nitrate Response and Increases Wheat Yield. *Plant Physiol.* 169, pp.00568.2015–2005. doi:10.1104/pp.15.00568

Hindt, M. N., and Guerinot, M. L. (2012). Getting a Sense for Signals: Regulation of the Plant Iron Deficiency Response. *Biochimica Biophysica Acta (BBA) - Mol. Cell. Res.* 1823, 1521–1530. doi:10.1016/j.bbamcr.2012.03.010

Hirel, B., Le Gouis, J., Ney, B., and Gallais, A. (2007). The Challenge of Improving Nitrogen Use Efficiency in Crop Plants: Towards a More Central Role for Genetic Variability and Quantitative Genetics within Integrated Approaches. *J. Exp. Bot.* 58, 2369–2387. doi:10.1093/jxb/erm097

Hisano, H., Abe, F., Hoffie, R. E., and Kumlehn, J. (2021). Targeted Genome Modifications in Cereal Crops. *Breed. Sci.* 71, 405–416. doi:10.1270/jsbbs.21019

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

Hong, J.-P., Takeshi, Y., Kondou, Y., Schachtman, D. P., Matsui, M., and Shin, R. (2013). Identification and Characterization of Transcription Factors Regulating Arabidopsis HAK5. *Plant Cell. Physiol.* 54, 1478–1490. doi:10.1093/pcp/pct094

Huang, S., Liang, Z., Chen, S., Sun, H., Fan, X., Wang, C., et al. (2019). A Transcription Factor, OsMADS57, Regulates Long-Distance Nitrate Transport and Root Elongation. *Plant Physiol.* 180, 882–895. doi:10.1104/PP.19.00142

Ibrahim, S., Saleem, B., Rehman, N., Zafar, S. A., Naeem, M. K., and Khan, M. R. (2022). CRISPR/Cas9 Mediated Disruption of Inositol Pentakisphosphate 2-Kinase 1 (TaIPK1) Reduces Phytic Acid and Improves Iron and Zinc

Accumulation in Wheat Grains. *J. Adv. Res.* 37, 33–41. doi:10.1016/J.JARE.2021.07.006

Inoue, H., Kobayashi, T., Nozoye, T., Takahashi, M., Kakei, Y., Suzuki, K., et al. (2009). Rice OsYSL15 Is an Iron-Regulated Iron(III)-deoxymugineic Acid Transporter Expressed in the Roots and Is Essential for Iron Uptake in Early Growth of the Seedlings. *J. Biol. Chem.* 284, 3470–3479. doi:10.1074/jbc.m806042200

Ishimaru, Y., Suzuki, M., Tsukamoto, T., Suzuki, K., Nakazono, M., Kobayashi, T., et al. (2006). Rice Plants Take up Iron as an Fe3+-Phytosiderophore and as Fe2+. *Plant J.* 45, 335–346. doi:10.1111/j.1365-313x.2005.02624.x

Ishino, Y., Shinagawa, H., Makino, K., Amemura, M., and Nakata, A. (1987). Nucleotide Sequence of the Iap Gene, Responsible for Alkaline Phosphatase Isozyme Conversion in *Escherichia coli*, and Identification of the Gene Product. *J. Bacteriol.* 169, 5429–5433. doi:10.1128/JB.169.12.5429-5433.1987

Jagadhesan, B., Sathee, L., Meena, H. S., Jha, S. K., Chinnusamy, V., Kumar, A., et al. (2020). Genome Wide Analysis of NLP Transcription Factors Reveals Their Role in Nitrogen Stress Tolerance of Rice. *Sci. Rep.* 10, 1–16. doi:10.1038/s41598-020-66338-6

Jan, S. U., Liaqat, A., Zhu, Y., Li, J., Zhang, H., Abdalla, M., et al. (2022). Arabidopsis NLP7 Improves Nitrogen Use Efficiency and Yield in Cotton. *J. Cott. Res.* 5, 1–10. doi:10.1186/S42397-021-00110-7/FIGURES/6

Jean, M. L., Schikora, A., Mari, S., Briat, J.-F., and Curie, C. (2005). A Loss-Of-Function Mutation in AtYSL1 Reveals its Role in Iron and Nicotianamine Seed Loading. *Plant J.* 44, 769–782. doi:10.1111/tpj.2005.44.issue-510.1111/j.1365-313x.2005.02569.x

Jia, H., Yin, Z., Xuan, D., Lian, W., Han, D., Zhu, Z., et al. (2022). Mutation of NtNRAMP3 Improves Cadmium Tolerance and its Accumulation in Tobacco Leaves by Regulating the Subcellular Distribution of Cadmium. *J. Hazard. Mater.* 432, 128701. doi:10.1016/J.JHAZMAT.2022.128701

Jiang, M., Liu, Y., Liu, Y., Tan, Y., Huang, J., and Shu, Q. (2019). Mutation of Inositol 1,3,4-trisphosphate 5/6-kinase6 Impairs Plant Growth and Phytic Acid Synthesis in Rice. *Plants* 8, 114. doi:10.3390/plants8050114

Jiang, Z., Hong, X., Zhang, S., Yao, R., and Xiao, Y. (2020). CRISPR Base Editing and Prime Editing: DSB and Template-free Editing Systems for Bacteria and Plants. *Synthetic Syst. Biotechnol.* 5 (4), 277–292. doi:10.1016/J.SYNBIO.2020.08.003

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012a). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816–821. doi:10.1126/science.1225829

Jinek, M., Chylinski, K., Fonfara, I., Hauer, M., Doudna, J. A., and Charpentier, E. (2012b). A Programmable Dual-RNA-Guided DNA Endonuclease in Adaptive Bacterial Immunity. *Science* 337, 816–821. doi:10.1126/SCIENCE.1225829

Joung, J. K., and Sander, J. D. (2013). TALENs: a Widely Applicable Technology for Targeted Genome Editing. *Nat. Rev. Mol. Cell. Biol.* 14, 49–55. doi:10.1038/nrm3486

Kanazawa, K., Higuchi, K., Nishizawa, N.-K., Fushiya, S., Chino, M., and Mori, S. (1994). Nicotianamine Aminotransferase Activities Are Correlated to the Phytosiderophore Secretions under Fe-Deficient Conditions in Gramineae. *J. Exp. Bot.* 45, 1903–1906. doi:10.1093/JXB/45.12.1903

Karunarathne, S. D., Han, Y., Zhang, X. Q., and Li, C. (2022). CRISPR/Cas9 Gene Editing and Natural Variation Analysis Demonstrate the Potential for HvARE1 in Improvement of Nitrogen Use Efficiency in Barley. *Integr. Plant Biol.* 64, 756–770. doi:10.1111/JIPB.13214/SUPPINFO

Kato-Inui, T., Takahashi, G., Hsu, S., and Miyaoka, Y. (2018). Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-associated Protein 9 with Improved Proof-Reading Enhances Homology-Directed Repair. *Nucleic Acids Res.* 46, 4677–4688. doi:10.1093/nar/gky264

Kaur, N., Alok, A., Shivani, N., Kaur, N., Pandey, P., Awasthi, P., et al. (2018). CRISPR/Cas9-mediated Efficient Editing in Phytoene Desaturase (PDS) Demonstrates Precise Manipulation in Banana Cv. Rasthali Genome. *Funct. Integr. Genomics* 18, 89–99. doi:10.1007/s10142-017-0577-5

Kim, J., and Kim, J.-S. (2016). Bypassing GMO Regulations with CRISPR Gene Editing. *Nat. Biotechnol.* 34, 1014–1015. doi:10.1038/nbt.3680

Kim, M. J., Ciani, S., and Schachtman, D. P. (2010). A Peroxidase Contributes to ROS Production during Arabidopsis Root Response to Potassium Deficiency. *Mol. Plant* 3, 420–427. doi:10.1093/mp/ssp121

Kobayashi, T., Nakanishi Itai, R., and Nishizawa, N. K. (2014). Iron Deficiency Responses in Rice Roots. *Rice* 7, 27. doi:10.1186/S12284-014-0027-0

Kobayashi, T., and Nishizawa, N. K. (2014). Iron Sensors and Signals in Response to Iron Deficiency. *Plant Sci.* 224, 36–43. doi:10.1016/J.PLANTSCI.2014.04.002

Kobayashi, T. (2019). Understanding the Complexity of Iron Sensing and Signaling Cascades in Plants. *Plant Cell. Physiol.* 60, 1440–1446. doi:10.1093/PCP/PCZ038

Krouk, G., Mirowski, P., LeCun, Y., Shasha, D. E., and Coruzzi, G. M. (2010). Predictive Network Modeling of the High-Resolution Dynamic Plant Transcriptome in Response to Nitrate. *Genome Biol.* 11, R123. doi:10.1186/gb-2010-11-12-r123

Kumar, A., Sharma, M., Kumar, S., Tyagi, P., Wani, S. H., Gajula, M. N. V. P., et al. (2018). Functional and Structural Insights into Candidate Genes Associated with Nitrogen and Phosphorus Nutrition in Wheat (*Triticum aestivum* L.). *Int. J. Biol. Macromol.* 118, 76–91. doi:10.1016/J.IJBIOMAC.2018.06.009

Kunz, H.-H., Gierth, M., Herdean, A., Satoh-Cruz, M., Kramer, D. M., Spetea, C., et al. (2014). Plastidial Transporters KEA1, -2, and -3 Are Essential for Chloroplast Osmoregulation, Integrity, and pH Regulation in Arabidopsis. *Proc. Natl. Acad. Sci. U.S.A.* 111, 7480–7485. doi:10.1073/pnas.1323899111

Lee, S., Chiecko, J. C., Kim, S. A., Walker, E. L., Lee, Y., Guerinot, M. L., et al. (2009). Disruption of OsYSL15 Leads to Iron Inefficiency in Rice Plants. *Plant Physiol.* 150, 786–800. doi:10.1104/PP.109.135418

Lei, G. J., Zhu, X. F., Wang, Z. W., Dong, F., Dong, N. Y., and Zheng, S. J. (2014). Abscisic Acid Alleviates Iron Deficiency by Promoting Root Iron Reutilization and Transport from Root to Shoot inArabidopsis. *Plant Cell. Environ.* 37, 852–863. doi:10.1111/PCE.12203

Lešková, A., Giehl, R. F. H., Hartmann, A., Fargašová, A., and von Wirén, N. (2017). Heavy Metals Induce Iron Deficiency Responses at Different Hierarchic and Regulatory Levels. *Plant Physiol.* 174, 1648–1668. doi:10.1104/pp.16.01916

Li, C., Li, W., Zhou, Z., Chen, H., Xie, C., and Lin, Y. (2020). A New Rice Breeding Method: CRISPR/Cas9 System Editing of the Xa13 Promoter to Cultivate Transgene-free Bacterial Blight-resistant Rice. *Plant Biotechnol. J.* 18, 313–315. doi:10.1111/pbi.13217

Li, J., Long, Y., Qi, G.-N., Li, J., Xu, Z.-J., Wu, W.-H., et al. (2014). The Os-AKT1 Channel Is Critical for K+ Uptake in Rice Roots and Is Modulated by the Rice CBL1-CIPK23 Complex. *Plant Cell.* 26, 3387–3402. doi:10.1105/tpc.114.123455

Li, J., Wu, W.-H., and Wang, Y. (2017). Potassium Channel AKT1 Is Involved in the Auxin-Mediated Root Growth Inhibition in Arabidopsis Response to Low K+ Stress. *J. Integr. Plant Biol.* 59, 895–909. doi:10.1111/jipb.12575

Li, L., Ye, L., Kong, Q., and Shou, H. (2019). A Vacuolar Membrane Ferric-Chelate Reductase, OsFRO1, Alleviates Fe Toxicity in Rice (Oryza Sativa l.). *Front. Plant Sci.* 10. doi:10.3389/FPLS.2019.00700

Li, S., Tian, Y., Wu, K., Ye, Y., Yu, J., Zhang, J., et al. (2018). Modulating Plant Growth-Metabolism Coordination for Sustainable Agriculture. *Nature* 560, 595–600. doi:10.1038/s41586-018-0415-5

Li, T., Yang, X., Yu, Y., Si, X., Zhai, X., Zhang, H., et al. (2018). Domestication of Wild Tomato Is Accelerated by Genome Editing. *Nat. Biotechnol.* 36, 1160–1163. doi:10.1038/nbt.4273

Liang, C., Piñeros, M. A., Tian, J., Yao, Z., Sun, L., Liu, J., et al. (2013). Low pH, Aluminum, and Phosphorus Coordinately Regulate Malate Exudation through GmALMT1 to Improve Soybean Adaptation to Acid Soils. *Plant Physiol.* 161, 1347–1361. doi:10.1104/pp.112.208934

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

Lightfoot, D. A. (2009). "Genes for use in improving nitrate use efficiency in crops," in *Genes for Plant Abiotic Stress*. Editors Jenks, J. A., and Wood, A. J. (Wiley-Blackwell), 167–182.

Lin, Q., Zong, Y., Xue, C., Wang, S., Jin, S., Zhu, Z., et al. (2020). Prime Genome Editing in Rice and Wheat. *Nat. Biotechnol.* 38, 582–585. doi:10.1038/s41587-020-0455-x

Lin, X. Y., Ye, Y. Q., Fan, S. K., Jin, C. W., and Zheng, S. J. (2016). Increased Sucrose Accumulation Regulates Iron-Deficiency Responses by Promoting Auxin Signaling in Arabidopsis Plants. *Plant Physiol.* 170, 907–920. doi:10.1104/pp.15.01598

Ling, H.-Q., Bauer, P., Bereczky, Z., Keller, B., and Ganal, M. (2002). The Tomato Fer Gene Encoding a bHLH Protein Controls Iron-Uptake Responses in Roots. *Proc. Natl. Acad. Sci. U.S.A.* 99, 13938–13943. doi:10.1073/pnas.212448699

Liu, H., Ding, Y., Zhou, Y., Jin, W., Xie, K., and Chen, L.-L. (2017). CRISPR-P 2.0: An Improved CRISPR-Cas9 Tool for Genome Editing in Plants. *Mol. Plant* 10, 530–532. doi:10.1016/j.molp.2017.01.010.1016/j.molp.2017.01.003

Liu, L., Zheng, C., Kuang, B., Wei, L., Yan, L., and Wang, T. (2016). Receptor-like Kinase RUPO Interacts with Potassium Transporters to Regulate Pollen Tube Growth and Integrity in Rice. *PLoS Genet.* 12, e1006085. doi:10.1371/journal.pgen.1006085

Liu, P., Chen, S., Song, A., Zhao, S., Fang, W., Guan, Z., et al. (2014). A Putative High Affinity Phosphate Transporter, CmPT1, Enhances Tolerance to Pi Deficiency of chrysanthemum. *BMC Plant Biol.* 14, 18. doi:10.1186/1471-2229-14-18

Long, T. A., Tsukagoshi, H., Busch, W., Lahner, B., Salt, D. E., and Benfey, P. N. (2010). The bHLH Transcription Factor POPEYE Regulates Response to Iron Deficiency inArabidopsisRoots. *Plant Cell.* 22, 2219–2236. doi:10.1105/TPC.110.074096

Lu, Y., and Zhu, J.-K. (2017). Precise Editing of a Target Base in the Rice Genome Using a Modified CRISPR/Cas9 System. *Mol. Plant* 10, 523–525. doi:10.1016/j.molp.2016.11.013

Lynch, J. P. (2007). Roots of the Second Green Revolution. *Aust. J. Bot.* 55, 493–512. doi:10.1071/bt06118

Mahajan, S., and Tuteja, N. (2005). Cold, Salinity and Drought Stresses: An Overview. *Archives Biochem. Biophysics* 444, 139–158. doi:10.1016/j.abb.2005.10.018

Manmathan, H., Shaner, D., Snelling, J., Tisserat, N., and Lapitan, N. (2013). Virus-Induced Gene Silencing of *Arabidopsis thaliana* Gene Homologues in Wheat Identifies Genes Conferring Improved Drought Tolerance. *J. Exp. Bot.* 64 (5), 1381–1392.

Mao, Y., Yang, X., Zhou, Y., Zhang, Z., Botella, J. R., and Zhu, J. K. (2018). Manipulating Plant RNA-Silencing Pathways to Improve the Gene Editing Efficiency of CRISPR/Cas9 Systems. *Genome Biol.* 19, 149. doi:10.1186/s13059-018-1529-7

Meng, S., Peng, J.-S., He, Y.-N., Zhang, G.-B., Yi, H.-Y., Fu, Y.-L., et al. (2016). Arabidopsis NRT1.5 Mediates the Suppression of Nitrate Starvation-Induced Leaf Senescence by Modulating Foliar Potassium Level. *Mol. Plant* 9, 461–470. doi:10.1016/j.molp.2015.12.015

Miglani, G. S. (2017). Genome Editing in Crop Improvement: Present Scenario and Future Prospects. *J. Crop Improv.* 31, 453–559. doi:10.1080/15427528.2017.1333192

Miller, J. C., Holmes, M. C., Wang, J., Guschin, D. Y., Lee, Y.-L., Rupniewski, I., et al. (2007). An Improved Zinc-Finger Nuclease Architecture for Highly Specific Genome Editing. *Nat. Biotechnol.* 25, 778–785. doi:10.1038/nbt1319

Munns, R., Passioura, J. B., Colmer, T. D., and Byrt, C. S. (2020). Osmotic Adjustment and Energy Limitations to Plant Growth in Saline Soil. *New Phytol.* 225, 1091–1096. doi:10.1111/nph.15862

Nam, Y.-J., Tran, L.-S. P., Kojima, M., Sakakibara, H., Nishiyama, R., and Shin, R. (2012). Regulatory Roles of Cytokinins and Cytokinin Signaling in Response to Potassium Deficiency in Arabidopsis. *PLoS One* 7, e47797. doi:10.1371/journal.pone.0047797

Nestler, J., Keyes, S. D., and Wissuwa, M. (2016). Root Hair Formation in Rice (Oryza sativaL.) Differs between Root Types and Is Altered in Artificial Growth Conditions. *Exbotj* 67, 3699–3708. doi:10.1093/jxb/erw115

Nguyen, G. N., Rothstein, S. J., Spangenberg, G., and Kant, S. (2015). Role of microRNAs Involved in Plant Response to Nitrogen and Phosphorous Limiting Conditions. *Front. Plant Sci.* 6, 629. doi:10.3389/fpls.2015.00629

Nieves-Cordones, M., Mohamed, S., Tanoi, K., Kobayashi, N. I., Takagi, K., Vernet, A., et al. (2017). Production of Low-Cs+ Rice Plants by Inactivation of the K+ Transporter OsHAK1 with the CRISPR-Cas System. *Plant J.* 92, 43–56. doi:10.1111/tpj.13632

Nieves-Cordones, M., Martínez, V., Benito, B., and Rubio, F. (2016). Comparison between Arabidopsis and Rice for Main Pathways of K+ and Na+ Uptake by Roots. *Front. Plant Sci.* 7, 992. doi:10.3389/fpls.2016.00992

Nishitani, C., Hirai, N., Komori, S., Wada, M., Okada, K., Osakabe, K., et al. (2016). Efficient Genome Editing in Apple Using a CRISPR/Cas9 System. *Sci. Rep.* 6, 31481–31488. doi:10.1038/srep31481

Niu, Y. F., Chai, R. S., Jin, G. L., Wang, H., Tang, C. X., and Zhang, Y. S. (2013). Responses of Root Architecture Development to Low Phosphorus Availability: a Review. *Ann. Bot.* 112, 391–408. doi:10.1093/aob/mcs285

Nozoye, T., Nagasaka, S., Kobayashi, T., Takahashi, M., Sato, Y., Sato, Y., et al. (2011). Phytosiderophore Efflux Transporters Are Crucial for Iron Acquisition in Graminaceous Plants. *J. Biol. Chem.* 286, 5446–5454. doi:10.1074/JBC.M110.180026

Ochiai, K., Shimizu, A., Okumoto, Y., Fujiwara, T., and Matoh, T. (2011). Suppression of a NAC-like Transcription Factor Gene Improves Boron-Toxicity Tolerance in Rice. *Plant Physiol.* 156, 1457–1463. doi:10.1104/pp.110.171470

Oelkers, E. H., and Valsami-Jones, E. (2008). Phosphate Mineral Reactivity and Global Sustainability. *Elements* 4, 83–87. doi:10.2113/gselements.4.2.83

Olsen, K. M., and Wendel, J. F. (2013). A Bountiful Harvest: Genomic Insights into Crop Domestication Phenotypes. *Annu. Rev. Plant Biol.* 64, 47–70. doi:10.1146/annurev-arplant-050312-120048

Pandiarajan, R., and Grover, A. (2018). *In Vivo* promoter Engineering in Plants: Are We Ready? *Plant Sci.* 277, 132–138. doi:10.1016/J.PLANTSCI.2018.10.011

Papikian, A., Liu, W., Gallego-Bartolomé, J., and Jacobsen, S. E. (2019). Site-specific Manipulation of Arabidopsis Loci Using CRISPR-Cas9 SunTag Systems. *Nat. Commun.* 10, 729. doi:10.1038/s41467-019-08736-7

Pathak, R. R., Ahmad, A., Lochab, S., and Raghuram, N. (2008). Molecular Physiology of Plant Nitrogen Use Efficiency and Biotechnological Options for its Enhancement. *Curr. Sci.* 94 (11), 1394–1403.

Poza-Carrión, C., and Paz-Ares, J. (2019). When Nitrate and Phosphate Sensors Meet. *Nat. Plants* 5, 339–340. doi:10.1038/s41477-019-0403-2

Prathap, V., Kumar, A., Maheshwari, C., and Tyagi, A. (2022). Phosphorus Homeostasis: Acquisition, Sensing, and Long-Distance Signaling in Plants. *Mol. Biol. Rep.* doi:10.1007/S11033-022-07354-9

Qu, B., He, X., Wang, J., Zhao, Y., Teng, W., Shao, A., et al. (2015). A Wheat CCAAT Box-Binding Transcription Factor Increases the Grain Yield of Wheat with Less Fertilizer Input. *Plant Physiol.* 167, 411–423. doi:10.1104/pp.114.246959

Ragel, P., Ródenas, R., García-Martín, E., Andrés, Z., Villalta, I., Nieves-Cordones, M., et al. (2015). The CBL-Interacting Protein Kinase CIPK23 Regulates HAK5-Mediated High-Affinity K+ Uptake in Arabidopsis Roots. *Plant Physiol.* 169, 2863–2873. pp.01401.2015. doi:10.1104/PP.15.01401

Rahayu, Y. S., Walch-Liu, P., Neumann, G., Römheld, V., Von Wirén, N., and Bangerth, F. (2005). Root-derived Cytokinins as Long-Distance Signals for NO3-Iinduced Stimulation of Leaf Growth. *J. Exp. Bot.* 56, 1143–1152. doi:10.1093/jxb/eri107

Reise, S. P., and Waller, N. G. (2009). Item Response Theory and Clinical Measurement. *Annu. Rev. Clin. Psychol.* 5, 27–48. doi:10.1146/annurev.clinpsy.032408.153553

Robinson, N. J., Procter, C. M., Connolly, E. L., and Guerinot, M. L. (1999). A Ferric-Chelate Reductase for Iron Uptake from Soils. *Nature* 397, 694–697. doi:10.1038/17800

Roca Paixão, J. F., Gillet, F.-X., Ribeiro, T. P., Bournaud, C., Lourenço-Tessutti, I. T., Noriega, D. D., et al. (2019). Improved Drought Stress Tolerance in Arabidopsis by CRISPR/dCas9 Fusion with a Histone AcetylTransferase. *Sci. Rep.* 9, 8080. doi:10.1038/s41598-019-44571-y

Rodríguez-Celma, J., Pan, I. C., Li, W., Lan, P., Buckhout, T. J., and Schmidt, W. (2013). The Transcriptional Response of Arabidopsis Leaves to Fe Deficiency. *Front. Plant Sci.* 4. doi:10.3389/FPLS.2013.00276

Römheld, V., and Marschner, H. (1986). Evidence for a Specific Uptake System for Iron Phytosiderophores in Roots of Grasses. *Plant Physiol.* 80, 175–180. doi:10.1104/PP.80.1.175

Ronzier, E., Corratgé-Faillie, C., Sanchez, F., Prado, K., Brière, C., Leonhardt, N., et al. (2014). CPK13, a Noncanonical Ca2+-dependent Protein Kinase, Specifically Inhibits KAT2 and KAT1 Shaker K+ Channels and Reduces Stomatal Opening. *Plant Physiol.* 166, 314–326. doi:10.1104/pp.114.240226

Rubin, G., Tohge, T., Matsuda, F., Saito, K., and Scheible, W.-R. (2009). Members of theLBDFamily of Transcription Factors Repress Anthocyanin Synthesis and Affect Additional Nitrogen Responses inArabidopsis. *Plant Cell.* 21, 3567–3584. doi:10.1105/tpc.109.067041

Saito, S., and Uozumi, N. (2019). Guard Cell Membrane Anion Transport Systems and Their Regulatory Components: an Elaborate Mechanism Controlling Stress-Induced Stomatal Closure. *Plants* 8, 9. doi:10.3390/plants8010009

Santi, S., and Schmidt, W. (2009). Dissecting Iron Deficiency-induced Proton Extrusion in Arabidopsis Roots. *New Phytol.* 183, 1072–1084. doi:10.1111/J.1469-8137.2009.02908.X

Schaaf, G., Schikora, A., Häberle, J., Vert, G., Ludewig, U., Briat, J.-F., et al. (2005). A Putative Function for the Arabidopsis Fe-Phytosiderophore Transporter Homolog AtYSL2 in Fe and Zn Homeostasis. *Plant Cell. Physiol.* 46, 762–774. doi:10.1093/PCP/PCI081

Schachtman, D. P. (2015). The Role of Ethylene in Plant Responses to K+ Deficiency. *Front. Plant Sci.* 6, 1153. doi:10.3389/fpls.2015.01153

Seth, K. (2016). Current Status of Potential Applications of Repurposed Cas9 for Structural and Functional Genomics of Plants. *Biochem. Biophysical Res. Commun.* 480 (4), 499–507. doi:10.1016/j.bbrc.2016.10.130

Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M., et al. (2017). ARGOS8 Variants Generated by CRISPR-Cas9 Improve Maize Grain Yield under Field Drought Stress Conditions. *Plant Biotechnol. J.* 15, 207–216. doi:10.1111/pbi.12603

Shin, L.-J., Lo, J.-C., Chen, G.-H., Callis, J., Fu, H., and Yeh, K.-C. (2013). IRT1 Degradation Factor1, a Ring E3 Ubiquitin Ligase, Regulates the Degradation of Iron-Regulated Transporter1 in Arabidopsis. *Plant Cell.* 25, 3039–3051. doi:10.1105/TPC.113.115212

Songmei, L., Jie, J., Yang, L., Jun, M., Shouling, X., Yuanyuan, T., et al. (2019). Characterization and Evaluation of OsLCT1 and OsNramp5 Mutants Generated through CRISPR/Cas9-mediated Mutagenesis for Breeding Low Cd Rice. *Rice Sci.* 26, 88–97. doi:10.1016/j.rsci.2019.01.002

Soyk, S., Lemmon, Z. H., Oved, M., Fisher, J., Liberatore, K. L., Park, S. J., et al. (2017). Bypassing Negative Epistasis on Yield in Tomato Imposed by a Domestication Gene. *Cell.* 169, 1142–1155. e12. doi:10.1016/j.cell.2017.04.032

Stigter, K., and Plaxton, W. (2015). Molecular Mechanisms of Phosphorus Metabolism and Transport during Leaf Senescence. *Plants* 4, 773–798. doi:10.3390/plants4040773

Straub, T., Ludewig, U., and Neuhäuser, B. (2017). The Kinase CIPK23 Inhibits Ammonium Transport in *Arabidopsis thaliana*. *Plant Cell.* 29, 409–422. doi:10.1105/tpc.16.00806

Sun, Y., Jiao, G., Liu, Z., Zhang, X., Li, J., Guo, X., et al. (2017). Generation of High-Amylose Rice through CRISPR/Cas9-mediated Targeted Mutagenesis of Starch Branching Enzymes. *Front. Plant Sci.* 8, 298. doi:10.3389/fpls.2017.00298

Szczepek, M., Brondani, V., Büchel, J., Serrano, L., Segal, D. J., and Cathomen, T. (2007). Structure-based Redesign of the Dimerization Interface Reduces the Toxicity of Zinc-Finger Nucleases. *Nat. Biotechnol.* 25, 786–793. doi:10.1038/nbt1317

Takahashi, R., Ishimaru, Y., Nakanishi, H., and Nishizawa, N. K. (2011). Role of the Iron Transporter OsNRAMP1 in Cadmium Uptake and Accumulation in Rice. *Plant Signal. Behav.* 6, 1813–1816. doi:10.4161/PSB.6.11.17587

Takasu, Y., Kobayashi, I., Beumer, K., Uchino, K., Sezutsu, H., Sajwan, S., et al. (2010). Targeted Mutagenesis in the Silkworm *Bombyx mori* Using Zinc Finger Nuclease mRNA Injection. *Insect Biochem. Mol. Biol.* 40, 759–765. doi:10.1016/j.ibmb.2010.07.012

Tanenbaum, M. E., Gilbert, L. A., Qi, L. S., Weissman, J. S., and Vale, R. D. (2014). A Protein-Tagging System for Signal Amplification in Gene Expression and Fluorescence Imaging. *Cell.* 159, 635–646. doi:10.1016/J.CELL.2014.09.039

Tang, T., Botwright Acuña, T., Spielmeyer, W., and Richards, R. A. (2020). Effect of Gibberellin-Sensitive Rht18 and Gibberellin-Insensitive Rht-D1b Dwarfing Genes on Vegetative and Reproductive Growth in Bread Wheat. *J. Exp. Bot.* 72, 445–458. doi:10.1093/jxb/eraa481

Tang, X., Lowder, L. G., Zhang, T., Malzahn, A. A., Zheng, X., Voytas, D. F., et al. (2017). A CRISPR-Cpf1 System for Efficient Genome Editing and Transcriptional Repression in Plants. *Nat. Plants* 3, 1–5. doi:10.1038/nplants.2017.18

Tian, S., Jiang, L., Gao, Q., Zhang, J., Zong, M., Zhang, H., et al. (2017). Efficient CRISPR/Cas9-based Gene Knockout in Watermelon. *Plant Cell. Rep.* 36, 399–406. doi:10.1007/s00299-016-2089-5

Tissot, N., Przybyla-Toscano, J., Reyt, G., Castel, B., Duc, C., Boucherez, J., et al. (2014). Iron Around the Clock. *Plant Sci.* 224, 112–119. doi:10.1016/J.PLANTSCI.2014.03.015

Ueda, Y., Kiba, T., and Yanagisawa, S. (2020). Nitrate-inducible NIGT1 Proteins Modulate Phosphate Uptake and Starvation Signalling via Transcriptional Regulation of SPX Genes. *Plant J.* 102, 448–466. doi:10.1111/tpj.14637

Véry, A.-A., Nieves-Cordones, M., Daly, M., Khan, I., Fizames, C., and Sentenac, H. (2014). Molecular Biology of K+ Transport across the Plant Cell Membrane: what Do We Learn from Comparison between Plant Species? *J. Plant Physiology* 171, 748–769. doi:10.1016/j.jplph.2014.01.011

Vlčko, T., and Ohnoutkova, L. (2020). Allelic Variants of CRISPR/Cas9 Induced Mutation in an Inositol Trisphosphate 5/6 Kinase Gene Manifest Different Phenotypes in Barley. *Plants* 9, 195. doi:10.3390/plants9020195

von Arnim, A. G., Jia, Q., and Vaughn, J. N. (2014). Regulation of Plant Translation by Upstream Open Reading Frames. *Plant Sci.* 214, 1–12. doi:10.1016/j.plantsci.2013.09.006

Waltz, E. (2018). With a Free Pass, CRISPR-Edited Plants Reach Market in Record Time. *Nat. Biotechnol.* 36, 6–7. doi:10.1038/nbt0118-6b

Wang, C., Chen, X., Yao, Q., Long, D., Fan, X., Kang, H., et al. (2019a). Overexpression of TtNRAMP6 Enhances the Accumulation of Cd in Arabidopsis. *Gene* 696, 225–232. doi:10.1016/j.gene.2019.02.008

Wang, F.-Z., Chen, M.-X., Yu, L.-J., Xie, L.-J., Yuan, L.-B., Qi, H., et al. (2017). OsARM1, an R2R3 MYB Transcription Factor, Is Involved in Regulation of the Response to Arsenic Stress in Rice. *Front. Plant Sci.* 8, 1868. doi:10.3389/fpls.2017.01868

Wang, F., Deng, M., Xu, J., Zhu, X., and Mao, C. (2018). "Molecular Mechanisms of Phosphate Transport and Signaling in Higher Plants," in *Seminars in Cell & Developmental Biology*, 74, 114–122. doi:10.1016/j.semcdb.2017.06.013

Wang, M., Qiao, J., Yu, C., Chen, H., Sun, C., Huang, L., et al. (2019). The Auxin Influx Carrier, OsAUX3, Regulates Rice Root Development and Responses to Aluminium Stress. *Plant Cell. Environ.* 42, 1125–1138. doi:10.1111/pce.13478

Wang, Q., Nian, J., Xie, X., Yu, H., Zhang, J., Bai, J., et al. (2018b). Genetic Variations in ARE1 Mediate Grain Yield by Modulating Nitrogen Utilization in Rice. *Nat. Commun.* 9, 1–10. doi:10.1038/s41467-017-02781-w

Wang, Q., Su, Q., Nian, J., Zhang, J., Guo, M., Dong, G., et al. (2021). The Ghd7 Transcription Factor Represses ARE1 Expression to Enhance Nitrogen Utilization and Grain Yield in Rice. *Mol. Plant* 14, 1012–1023. doi:10.1016/j.molp.2021.04.012

Wang, S., Chen, A., Xie, K., Yang, X., Luo, Z., Chen, J., et al. (2020). Functional Analysis of the OsNPF4.5 Nitrate Transporter Reveals a Conserved Mycorrhizal Pathway of Nitrogen Acquisition in Plants. *Proc. Natl. Acad. Sci. U.S.A.* 117, 16649–16659. doi:10.1073/pnas.2000926117

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, 43–49. doi:10.1038/s41586-018-0063-9

Wang, X.-P., Chen, L.-M., Liu, W.-X., Shen, L.-K., Wang, F.-L., Zhou, Y., et al. (2016). AtKC1 and CIPK23 Synergistically Modulate AKT1-Mediated Low-Potassium Stress Responses in Arabidopsis. *Plant Physiol.* 170, 2264–2277. doi:10.1104/pp.15.01493

Wang, Y., Beaith, M., Chalifoux, M., Ying, J., Uchacz, T., Sarvas, C., et al. (2009). Shoot-Specific Down-Regulation of Protein Farnesyltransferase (α-subunit) for Yield Protection Against Drought in Canola. *Mol. Plant* 2 (1), 191–200.

Wang, Y.-Y., Cheng, Y.-H., Chen, K.-E., and Tsay, Y.-F. (2018d). Nitrate Transport, Signaling, and Use Efficiency. *Annu. Rev. Plant Biol.* 69, 85–122. doi:10.1146/ANNUREV-ARPLANT-042817-040056

Wang, Y., and Wu, W.-H. (2017). Regulation of Potassium Transport and Signaling in Plants. *Curr. Opin. Plant Biol.* 39, 123–128. doi:10.1016/j.pbi.2017.06.006

Wang, Z.-Y., Qin, X.-H., Li, J.-H., Fan, L.-F., Zhou, Q., Wang, Y.-Q., et al. (2019c). Highly Reproducible Periodic Electrical Potential Changes Associated with Salt Tolerance in Wheat Plants. *Environ. Exp. Bot.* 160, 120–130. doi:10.1016/j.envexpbot.2019.01.014

Waters, B. M., Chu, H.-H., DiDonato, R. J., Roberts, L. A., Eisley, R. B., Lahner, B., et al. (2006). Mutations in Arabidopsis Yellow Stripe-Like1 and Yellow Stripe-Like3 Reveal Their Roles in Metal Ion Homeostasis and Loading of Metal Ions in Seeds. *Plant Physiol.* 141, 1446–1458. doi:10.1104/PP.106.082586

Wei, Y., Shi, A., Jia, X., Zhang, Z., Ma, X., Gu, M., et al. (2018). Nitrogen Supply and Leaf Age Affect the Expression of Tags1 or TaGS2 Driven by a Constitutive Promoter in Transgenic Tobacco. *Genes.* 9, 406. pii. doi:10.3390/genes9080406

Wolter, F., and Puchta, H. (2018). "Application of CRISPR/Cas to Understand Cis- and Trans-regulatory Elements in Plants," in *Plant Transcription Factors: Methods and Protocols*. Editor N. Yamaguchi (New York, NY: Springer New York), 23–40. doi:10.1007/978-1-4939-8657-6_2

Wu, J., Zhang, Z.-S., Xia, J.-Q., Alfatih, A., Song, Y., Huang, Y.-J., et al. (2020). Rice NIN-LIKE PROTEIN 4 Plays a Pivotal Role in Nitrogen Use Efficiency. *Plant Biotechnol. J.* 19 (3), 448–461. doi:10.1111/pbi.13475

Wu, P., and Wang, X. M. (2008). Role of OsPHR2 on Phosphorus Homoestasis and Root Hairs Development in Rice (Oryza sativaL.). *Plant Signal. Behav.* 3, 674–675. doi:10.4161/psb.3.9.5781

Wu, Y., Yang, W., Wei, J., Yoon, H., and An, G. (2017). Transcription Factor OsDOF18 Controls Ammonium Uptake by Inducing Ammonium

Transporters in Rice Roots. *Mol. Cells* 40, 178–185. doi:10.14348/molcells.2017.2261

Wu, Z., Ren, H., McGrath, S. P., Wu, P., and Zhao, F.-J. (2011). Investigating the Contribution of the Phosphate Transport Pathway to Arsenic Accumulation in Rice. *Plant Physiol.* 157, 498–508. doi:10.1104/pp.111.178921

Xia, X., Fan, X., Wei, J., Feng, H., Qu, H., Xie, D., et al. (2015). Rice Nitrate Transporter OsNPF2.4 Functions in Low-Affinity Acquisition and Long-Distance Transport. *J. Exp. Bot.* 66, 317–331. doi:10.1093/jxb/eru425

Xie, K., Minkenberg, B., and Yang, Y. (2015). Boosting CRISPR/Cas9 Multiplex Editing Capability with the Endogenous tRNA-Processing System. *Proc. Natl. Acad. Sci. U.S.A.* 112, 3570–3575. doi:10.1073/PNAS.1420294112

Xie, K., and Yang, Y. (2013). RNA-guided Genome Editing in Plants Using a CRISPR-Cas System. *Mol. Plant* 6, 1975–1983. doi:10.1093/mp/sst119

Xu, G., Yuan, M., Ai, C., Liu, L., Zhuang, E., Karapetyan, S., et al. (2017). uORF-Mediated Translation Allows Engineered Plant Disease Resistance without Fitness Costs. *Nature* 545, 491–494. doi:10.1038/nature22372

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, L., Zhao, H., Wan, R., Liu, Y., Xu, Z., Tian, W., et al. (2019). Identification of Vacuolar Phosphate Efflux Transporters in Land Plants. *Nat. Plants* 5, 84–94. doi:10.1038/s41477-018-0334-3

Xuan, Y. H., Priatama, R. A., Huang, J., Je, B. I., Liu, J. M., Park, S. J., et al. (2013). I Ndeterminate Domain 10 Regulates Ammonium-mediated Gene Expression in Rice Roots. *New Phytol.* 197, 791–804. doi:10.1111/nph.12075

Yang, A., Dai, X., and Zhang, W.-H. (2012). A R2R3-type MYB Gene, OsMYB2, Is Involved in Salt, Cold, and Dehydration Tolerance in Rice. *J. Exp. Bot.* 63, 2541–2556. doi:10.1093/jxb/err431

Yang, T., Zhang, S., Hu, Y., Wu, F., Hu, Q., Chen, G., et al. (2014). The Role of a Potassium Transporter OsHAK5 in Potassium Acquisition and Transport from Roots to Shoots in Rice at Low Potassium Supply Levels. *Plant Physiol.* 166, 945–959. doi:10.1104/pp.114.246520

Ye, J., Wang, X., Hu, T., Zhang, F., Wang, B., Li, C., et al. (2017). An InDel in the Promoter of Al-ACTIVATED Malate Transporter9 Selected during Tomato Domestication Determines Fruit Malate Contents and Aluminum Tolerance. *Plant Cell.* 29, 2249–2268. doi:10.1105/tpc.17.00211

Zeeshan, M., Lu, M., Sehar, S., Holford, P., and Wu, F. (2020). Comparison of Biochemical, Anatomical, Morphological, and Physiological Responses to Salinity Stress in Wheat and Barley Genotypes Deferring in Salinity Tolerance. *Agronomy* 10, 127. doi:10.3390/agronomy10010127

Zhang, D., Wang, Q., Zhu, T., Cao, J., Zhang, X., Wang, J., et al. (2013). RACK1 Promotes the Proliferation of THP1 Acute Myeloid Leukemia Cells. *Mol. Cell. Biochem.* 384, 197–202. doi:10.1007/s11010-013-1798-0

Zhang, F., Cheng, D., Wang, S., and Zhu, J. (2020). Crispr/Cas9-mediated Cleavages Facilitate Homologous Recombination during Genetic Engineering of a Large Chromosomal Region. *Biotechnol. Bioeng.* 117, 2816–2826. doi:10.1002/bit.27441

Zhang, H., Li, Y., Yao, X., Liang, G., and Yu, D. (2017). Positive Regulator of Iron Homeostasis1, OsPRI1, Facilitates Iron Homeostasis. *Plant Physiol.* 175, 543–554. doi:10.1104/pp.17.00794

Zhang, H., and Reynolds, M. (2019). Cadmium Exposure in Living Organisms: A Short Review. *Sci. Total Environ.* 678, 761–767. doi:10.1016/j.scitotenv.2019.04.395

Zhang, J., Zhang, H., Li, S., Li, J., Yan, L., and Xia, L. (2021a). Increasing Yield Potential through Manipulating of an ARE1 Ortholog Related to Nitrogen Use Efficiency in Wheat by CRISPR/Cas9. *J. Integr. Plant Biol.* 63, 1649–1663. doi:10.1111/JIPB.13151/SUPPINFO

Zhang, J., Zhou, Z., Bai, J., Tao, X., Wang, L., Zhang, H., et al. (2020). Disruption of MIR396e and MIR396f improves rice yield under nitrogen-deficient conditions. *Natl. Sci. Rev.* 7, 102–112. doi:10.1093/nsr/nwz142

Zhang, L., Sun, X., Chen, X., Huang, M., Xiao, T., Mo, Y., et al. (2020). A Potassium (K)-transporter, OsHAK3, Is Required for K-Homeostasis in Rice under Low-K and High-Salinity Conditions. *Authorea Prepr.* doi:10.22541/au.160269754.40987920/v1

Zhang, P., Wang, R., Ju, Q., Li, W., Tran, L.-S. P., and Xu, J. (2019). The R2R3-MYB Transcription Factor MYB49 Regulates Cadmium Accumulation. *Plant Physiol.* 180, 529–542. doi:10.1104/pp.18.01380

Zhang, T.-T., Kang, H., Fu, L.-L., Sun, W.-J., Gao, W.-S., You, C.-X., et al. (2021b). NIN-like Protein 7 Promotes Nitrate-Mediated Lateral Root Development by Activating Transcription of Tryptophan Aminotransferase Related 2. *Plant Sci.* 303, 110771. doi:10.1016/J.PLANTSCI.2020.110771

Zhang, Z. S., Xia, J. Q., Alfatih, A., Song, Y., Huang, Y. J., Sun, L. Q., et al. (2022). Rice Nin-Like Protein 3 Modulates Nitrogen Use Efficiency and Grain Yield under Nitrate-sufficient Conditions. *Plant Cell. & Environ.* 45, 1520–1536. doi:10.1111/PCE.14294

Zhao, L., Liu, F., Crawford, N., and Wang, Y. (2018). Molecular Regulation of Nitrate Responses in Plants. *Ijms* 19, 2039. doi:10.3390/ijms19072039

Zhao, P., You, Q., and Lei, M. (2019). A CRISPR/Cas9 Deletion into the Phosphate Transporter SlPHO1;1 Reveals its Role in Phosphate Nutrition of Tomato Seedlings. *Physiol. Plant.* 167, 556–563. doi:10.1111/ppl.12897

Zhao, S., Zhang, M.-L., Ma, T.-L., and Wang, Y. (2016). Phosphorylation of ARF2 Relieves its Repression of Transcription of the K+ Transporter Gene HAK5 in Response to Low Potassium Stress. *Plant Cell.* 28, 3005–3019. doi:10.1105/tpc.16.00684

Zhou, J., Deng, K., Cheng, Y., Zhong, Z., Tian, L., Tang, X., et al. (2017). CRISPR-cas9 Based Genome Editing Reveals New Insights into microRNA Function and Regulation in Rice. *Front. Plant Sci.* 8, 1598. doi:10.3389/FPLS.2017.01598/BIBTEX

Zhou, Y., Neuhäuser, B., Neumann, G., and Ludewig, U. (2020). LaALMT1 Mediates Malate Release from Phosphorus-deficient White Lupin Root Tips and Metal Root to Shoot Translocation. *Plant Cell. Environ.* 43, 1691–1706. doi:10.1111/pce.13762

Zhou, Z., Wang, Z., Lv, Q., Shi, J., Zhong, Y., Wu, P., et al. (2015). SPX Proteins Regulate Pi Homeostasis and Signaling in Different Subcellular Level. *Plant Signal. Behav.* 10, e1061163. doi:10.1080/15592324.2015.1061163

Zhu, J.-K. (2002). Salt and Drought Stress Signal Transduction in Plants. *Annu. Rev. Plant Biol.* 53, 247–273. doi:10.1146/annurev.arplant.53.091401.143329

Zhu, X. F., Wang, B., Song, W. F., Zheng, S. J., and Shen, R. F. (2015). Putrescine Alleviates Iron Deficiency via NO-dependent Reutilization of Root Cell-Wall Fe in Arabidopsis. *Plant Physiol.* 170, 558–567. doi:10.1104/PP.15.01617

Zhu, Y., Li, T., Xu, J., Wang, J., Wang, L., Zou, W., et al. (2020). Leaf Width Gene LW5/D1 Affects Plant Architecture and Yield in Rice by Regulating Nitrogen Utilization Efficiency. *Plant Physiology Biochem.* 157, 359–369. doi:10.1016/j.plaphy.2020.10.035

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.

Publisher's Note: All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 26 August 2022
DOI 10.3389/fgene.2022.978880

OPEN ACCESS

EDITED BY
Reyazul Rouf Mir,
Sher-e-Kashmir University of Agricultural Sciences and Technology, India

REVIEWED BY
Yuanfeng Hao,
Institute of Crop Sciences (CAAS), China
Sindhu Sareen,
Indian Institute of Wheat and Barley Research (ICAR), India
Shengjie Liu,
Northwest A&F University, China

*CORRESPONDENCE
Zhibin Xu,
xuzb@cib.ac.cn
Tao Wang,
wangtao@cib.ac.cn

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 26 June 2022
ACCEPTED 25 July 2022
PUBLISHED 26 August 2022

CITATION
Fan X, Liu X, Feng B, Zhou Q, Deng G, Long H, Cao J, Guo S, Ji G, Xu Z and Wang T (2022), Construction of a novel Wheat 55 K SNP array-derived genetic map and its utilization in QTL mapping for grain yield and quality related traits. *Front. Genet.* 13:978880. doi: 10.3389/fgene.2022.978880

Construction of a novel Wheat 55 K SNP array-derived genetic map and its utilization in QTL mapping for grain yield and quality related traits

Xiaoli Fan[1], Xiaofeng Liu[1,2], Bo Feng[1], Qiang Zhou[1], Guangbing Deng[1], Hai Long[1], Jun Cao[3], Shaodan Guo[1], Guangsi Ji[1,2], Zhibin Xu[1]* and Tao Wang[1,4]*

[1]Chengdu Institute of Biology, Chinese Academy of Sciences, Chengdu, China, [2]University of Chinese Academy of Sciences, Beijing, China, [3]Yibin University, Yibin, China, [4]Innovative Academy for Seed Design, Chinese Academy of Sciences, Beijing, China

Wheat is one of the most important staple crops for supplying nutrition and energy to people world. A new genetic map based on the Wheat 55 K SNP array was constructed using recombinant inbred lines derived from a cross between Zhongkemai138 and Kechengmai2 to explore the genetic foundation for wheat grain features. This new map covered 2,155.72 cM across the 21 wheat chromosomes with 11,455 markers. And 2,846 specific markers for this genetic map and 148 coincident markers among different maps were documented, which was helpful for improving and updating wheat genetic and genomic information. Using this map, a total of 68 additive QTLs and 82 pairs of epistatic QTLs were detected for grain features including yield, nutrient composition, and quality-related traits by QTLNetwork 2.1 and IciMapping 4.1 software. Fourteen additive QTLs and one pair of epistatic QTLs could be detected by both software programs and thus regarded as stable QTLs here, all of which explained higher phenotypic variance and thus could be utilized for wheat grain improvement. Additionally, thirteen additive QTLs were clustered into three genomic intervals (C4D.2, C5D, and C6D2), each of which had at least two stable QTLs. Among them, C4D.2 and C5D have been attributed to the famous dwarfing gene *Rht2* and the hardness locus *Pina*, respectively, while endowed with main effects on eight grain yield/quality related traits and epistatically interacted with each other to control moisture content, indicating that the correlation of involved traits was supported by the pleotropic of individual genes but also regulated by the gene interaction networks. Additionally, the stable additive effect of C6D2 (*QMc.cib-6D2* and *QTw.cib-6D2*) on moisture content was also highlighted, potentially affected by a novel locus, and validated by its flanking Kompetitive Allele-Specific PCR marker, and *TraesCS6D02G109500*, encoding aleurone layer morphogenesis protein, was deduced to be one of the candidate genes for this locus. This result observed at the QTL level the possible contribution of grain water content to the balances among yield, nutrients, and quality properties and reported a possible

new locus controlling grain moisture content as well as its linked molecular marker for further grain feature improvement.

KEYWORDS

grain yield, moisture content, wheat, genetic map, nutritional content

1 Introduction

Wheat (*Triticum aestivum* L.) as one of the most important staple foods, accounts for 20% of global calorie and protein intake (FAO, http://www.fao.org/faostat/). Starch and protein, as the two major compositions of wheat grain, account for 60%–75% and 8%–20% of the total dry mass of the mature grain (Šramková et al., 2009), respectively, greatly influencing flour processing and end-use attributes, such as gelatinization temperatures and tensile properties (Branlard et al., 2001; Zhu et al., 2009; Shevkani et al., 2017), as well as affecting nutrient concentrations, ultimately determining food nutrition supply amount. The famous starch and protein biosynthetic genes including *SUT* (encoding sucrose transporter), *Wx* (encoding granule-bound starch synthase), *SsI* (encoding soluble starch synthases), *GPC* (encoding NAC transcription factor *Grain Protein Content*) (Murai et al., 1999; Aoki et al., 2002; Hurkman et al., 2003; Uauy et al., 2006; Avni et al., 2014) have been well studied for wheat quality improvement.

Moisture content typically determines grain storability. Grain with a low moisture content (<12%) is more resistant to storage, but too low may result in poor seed viability (Al-Yahya, 2001; Karunakaran et al., 2001). Additionally, as one of the three components with the highest proportion in the seed (Al-Mahasneh and Rababah, 2007), moisture content has a balance with protein and starch content, also influencing nutritive value and several processing quality parameters (Fang and Campbell, 2003), especially grain hardness and test weight parameters, which are widely employed as key indicators in wheat grading and pricing (Schuler et al., 1995; Cabral et al., 2018), because both of them can predict the grain milling efficiency and flour production (Muhamad and Campbell, 2004) and are also involved with the level of starch and protein content (Anjum and Walker, 1991). The related controlling genes for hardness (*Pina*, *Pinb*) have been well studied (Bhave and Morris, 2008; Qamar et al., 2014). However, the genetic factors controlling moisture content or test weight are less well discovered.

These grain traits generally present a correlation with grain yield and thus are generally noticed in the genetic improvement process, especially the well-known trade-offs between protein content and grain yield (Michel et al., 2019; Sandhu et al., 2021). However, with no change in concentration, the absolute production of these grain inner compositions should positively correlate with the total grain yield, and thus the absolute output of these compositions also deserves attention and research. In particular, with the advancement of processing and purification technology, it is now more convenient to extract wheat grain protein, starch or other nutrient substances (Day et al., 2006), both of which can be used as nutrition for functional food production or as raw materials for producing gluten, resistant starch, monosodium glutamate, fermentation substrate, etc. (Mohamed and Rayas-Duarte, 2003; Day et al., 2006; Xie et al., 2008; Ortolan and Steel, 2017; Sardari et al., 2019). They have been widely used in the food and processing industries, and thus not only the proportions of inner grain compositions but also their absolute yield are also factors in determining wheat's commercial potential.

All these traits mentioned above are quantitative traits. Besides the known genes, previous studies have reported numerous QTLs providing genetic basis for these traits (Groos et al., 2003; Narasimhamoorthy et al., 2006; Zhang et al., 2011; Bonneau et al., 2013; Heo and Sherman, 2013; Kumar et al., 2019; Guo et al., 2020; Tu and Li, 2020; Colasuonno et al., 2021; Fradgley et al., 2022; Gudi et al., 2022). Although epistatic QTLs, which are involved in regulatory networks, are also crucial foundations for controlling complex traits, the majority of reported loci are additive QTLs (Xing et al., 2002; Malmberg and Mauricio, 2005). To date, several software programs have been developed to conduct both additive and epistatic QTL analyses, such as QTLNetwork (Yang et al., 2008) and IciMapping (Meng et al., 2015), both of which are helpful for determining the loci with main and interaction effects on the target traits.

With the development of the wheat genome, several arrays with high density SNP markers have been developed for wheat genetic research, such as 660, 90, and 55 K, etc. (Wang et al., 2014; Cui et al., 2017; Liu et al., 2018; Sun et al., 2020). These SNP arrays and their transformed KASP/CAPS markers provide great help for wheat genetic improvement. In this study, we constructed a new genetic map using a Wheat 55K array to complement wheat genomic information, and used this map to perform QTL analysis for grain features such as starch content, protein content, test weight and hardness of wheat grain, and to resolve their relationships with grain yield and grain nutrition yield at the QTL level.

2 Materials and methods

2.1 Plant materials and field trials

An $F_{6:7}$ recombinant inbred line (RIL) population with 152 lines derived from a cross between varieties Zhongkemai138 (ZKM138) and Kechengmai2 (indicated as ZK-RILs) was used in this study. ZKM138 and KCM2 are both widely adaptable varieties in southwestern China and released by Chengdu Biology of Institute, CAS, (CIBCAS) in recent years. ZK-RILs and their parents were evaluated at Chengdu (30°57′ N, 104°94′ E) and Deyang (31°13′ N, 104°16′ E) in 2016–2017 and 2017–2018. A total of four environments were designated, namely, 17CD, 17DY, 18CD and 18DY. For each environment, the materials were planted in two replicated blocks. Each block contained two rows that were 2 m long and 0.25 m apart, and 20 seeds were evenly planted in each row. All of the recommended agronomic practices were followed in each trial.

2.2 Phenotypic evaluation and statistical analysis

Five representative plants in the centre of the rows were randomly sampled at physiological maturity for phenotypic evaluation. The lines were naturally air dried at sunny days and then dried at 42°C for 5–7 days, until samples presented constant weight. Three grain nutrition concentration related traits, including protein content (PC), starch content (SC), and moisture content (MC), and three grain grading related traits, including test weight (TW), grain hardness (GH) and water absorption (ABS), were measured by near-infrared reflectance spectroscopy (NIRS) with a Perten DA-7250 instrument (Perten Instruments, Huddinge, Sweden) and expressed on a 14% moisture basis. The grain yield per plant (GY) was weighed and used to calculate three physical compositions of the yield, including grain protein weight per plant (GPW = GY×PC×100), grain starch weight per plant (GSW = GY×SC×100) and grain water weight per plant (GWW = GY×MC×100). The spike length and kernel length were evaluated at maturity in the 18CD environment. The spike length was measured from the base of the rachis to the tip of the terminal spikelet excluding awns. The kernel length was investigated by lining up 20 kernels length-wise along a ruler with a precision of 0.1 mm.

2.3 Map construction

Genomic DNA for the 152 RILs and their parental lines was extracted from tender leaves using the TiangenTM Plant Genomic DNA Kit, and quality was evaluated using an agarose gel. The concentration was detected using the a Thermo Scientific NanoDrop 2000. The qualified DNA was genotyped using the Affymetrix 55K SNP array by Compass Biotechnology Company (Beijing, China) for genetic map construction and QTL mapping. Chip genotyping was performed according to the Axiom R 2.0 Assay for 384 Samples User Manual (Cui et al., 2017). The reported primers and their corresponding PCR system were used to detect *Rht2* (Ellis et al., 2002; Zhang et al., 2006) and *Pina* (Li et al., 2006).

Biallelic pleomorphic SNPs with >10% missing data and $p < 0.01$ by chi-square test of segregation distortion (departure from the expected 1:1 segregation ratio) were removed, and the remaining high-quality SNPs were binned by their pattern of segregation using the BIN function of IciMapping 4.1. Each bin had several markers; the correlation coefficient between them was 1, and one marker with the lowest missing rate was chosen to represent this bin. If there were no missing data of the markers in one bin, one marker was chosen randomly. Markers were tested for significant segregation distortion using a chi-square test. SNPs were sorted into groups using the MAP function in IciMapping 4.1. A logarithm of the odds (LOD) score of 3.5 and a recombination fraction of 0.4 were used to sort the SNPs with the Kosambi mapping function. Groups were ordered with the Kosambi mapping function within JoinMap v. 4.0, using an LOD score ≥3 after preliminary analysis of SNPs with LOD scores ranging from 2 to 10. The polarity of each chromosome was identified from the IWGSC wheat survey sequence, and groups were oriented to have the short arm above the long arm.

The SNP flanking sequences mapped in the ZK-RILs map were used to perform BLAST searches against (E value cutoff of 10–5) the IWGSC wheat survey contig sequences and the wheat genome assembly of *T. aestivum* cv. Chinese spring (CS) (IWGSC RefSeq v1.1, ftp://ftp.ensemblgenomes.org/pub/plants/release-44/fasta/triticum_aestivum/) to get their physical locations. The flanking sequences of SNPs and their best matched contigs were further used to blast against CDS of IWGSC RefSeq v1.1 to identify the number of coding-region SNPs (cSNPs) and perigenic SNPs (pSNPs) and intergenic SNPs (iSNPs), respectively (Cui et al., 2017; Liu et al., 2018). In addition, contig sequences to which the SNPs were best hits were screened in a BLASTN search against the coding sequences (CDSs) of Brachypodiun, barley, rice and maize. All CDSs were downloaded from http://plants.ensembl.org/index.html. An expectation value (E) of 1E-10 was defined as the significance threshold. Synteny analyses with common wheat, Brachypodiun, barley, rice and maize genomes were performed based on the SNP orders in the ZK genetic map and on the corresponding CDSs in the genome sequences of Brachypodiun, barley, rice and maize genomes. The SNPs flanking sequences of

90 K (Wang et al., 2014) and 660 K (Cui et al., 2017) were used to blast against IWGSC RefSeq v1.1 assembly sequences to map their physical locations, and compared with the SNP locations of ZK-RILs map.

2.4 Statistical analysis and QTL mapping

Statistical analysis was conducted using the set of predicted genotype means (Best Linear Unbiased Predictors, BLUP) for six traits in four environments and GenStat 19th software (VSN International, Hemel Hempstead, United Kingdom). Pearson correlation was used to analyse the relationship among the traits. A total of five datasets, designated 17CD, 17DY, 18CD, 18DY and BLUP data (B), were used for further analysis.

Two softwares, QTLNetwork 2.1 (http://ibi.zju.edu.cn/index.html/BCL/software/qtlnetwork.html) based on the mixed-model composite interval mapping (MCIM) model and IciMapping 4.1 (http://www.isbreeding.net/) based on the inclusive composite interval mapping (ICIM) model, were both employed to detect additive and epistatic QTLs. By using QTLNetwork v2.1, one- and two-dimensional genome scans for QTLs were performed using a 10 cM testing window, a 0.1 cM walk speed and a 0.5 cM filtration window size. For QTL analysis with IciMapping 4.1, the threshold of LOD of 3.42 was used to declare the presence of a putative QTL based on the 1,000 permutation tests ($p = 0.05$) for each trait. The walking speed chosen for all QTLs was 1.0 cM, and the p-value inclusion threshold was 0.001.

In this study, a QTL with a phenotypic variance contribution (PVE) > 10% (on average) detected by either approach (ICMI or MCIM) was defined as a major QTL; a QTL repeatedly detected by both softwares was defined as a stable QTL.

2.5 Development of kompetitive allele-specific PCR markers

On the basis of the preliminary QTL mapping results, SNPs adjacent to the target intervals of the major QTLs were converted into KASP markers (Supplementary Table S1) following the previously described method (Li et al., 2021) and used to trace the targeted QTL.

3 Results

3.1 Linkage map construction

The ZK-RILs were genotyped using the Wheat55K SNP array, which yielded 13,651 polymorphic markers for the linkage analysis. After removing unlinked markers and SNPs with more than 10% missing data or a segregation distortion test $p = 0.01$, a unique high-density genomic map with 11,455 markers and a length of 2,155.72 cM was created (Supplementary Figure S1; Table 1; Supplementary Table S2). To represent each bin on the created map, 1,459 bin markers with the lowest miss rate in each bin were picked (Supplementary Figure S1). The A, B, and D genomes had map lengths of 652.15, 606.41 and 857.16 cM, respectively, and 4,396, 4,288 and 2,782 markers, respectively, resulting in a similar number of bins (495 for the A genome, 469 for the B genome and 495 for the D genome) (Table 1; Supplementary Table S2). The 11,455 markers were distributed unevenly over the 21 chromosomes, with 59 markers on 4D and 1,325 on 2A. On chromosomes 2A, 2B, 3B, 4A, 5D and 7A, six gaps (more than 20 cM but less than 30 cM) were observed (Supplementary Figure S1; Supplementary Table S2). Furthermore, chromosomes 1D (1D1, 1D2, and 1D3), 3A (3A1, 3A2, and 3A3), and 6D (6D1, 6D2 and 6D3) were classified into three groups based on their linkage (Supplementary Figure S1; Table 1; Supplementary Table S2). The number of markers in each bin ranged from 1 to 980, whereas the number of chromosomes in each bin ranged from 30 (4D) to 105 (5B) (Supplementary Table S2).

3.2 Comparative genomic analysis

Overall, 10,384 (90.66%) of the 11,455 mapped probes were best matches to 9818 CS contigs, with an average of 1.06 polymorphic markers per contig (Additional file 1). The SNP arrangement in the present genetic map matched that of the wheat genome assembly (Supplementary Figure S2). In addition, 90.86% of the 11,455 mapped SNPs exhibited consistency with their physical position, whereas 5.85% were mapped to the genetic map for their homologous chromosomes, 0.72 percent were localized to different chromosomes within the same subgenome, and 2.57% were disordered (Supplementary Tables S2, S3). Furthermore, 3.04% (348 SNPs), 74.36% (8,517 SNPs), and 13.39% (1,534 SNPs) of the SNP markers were classified as cSNPs, iSNPs, and pSNPs, respectively (Additional file 1).

According to contig information, 4,257 markers in the ZK-RILs map shared contigs with 3,832 markers and 766 markers in previously reported genetic maps based on the Wheat 660K (Cui et al., 2017) and 90K SNP arrays (Wang et al., 2014), respectively (Supplementary Figure S3). Only 343 markers on 19 chromosomes (excluding 1B and 3B) shared contigs with three different arrays (Supplementary Figure S3, Additional file 2). These markers and the physical information associated with them could be used to create integrated genetic maps as well as to compare locus results from different chips.

Furthermore, when compared to the other 55 K SNP array-derived genetic maps for hexaploid wheat (Liu et al., 2018; Liu et al., 2020; Li et al., 2021; Lin et al., 2021), a total of 8,970 common markers were identified in these four maps:

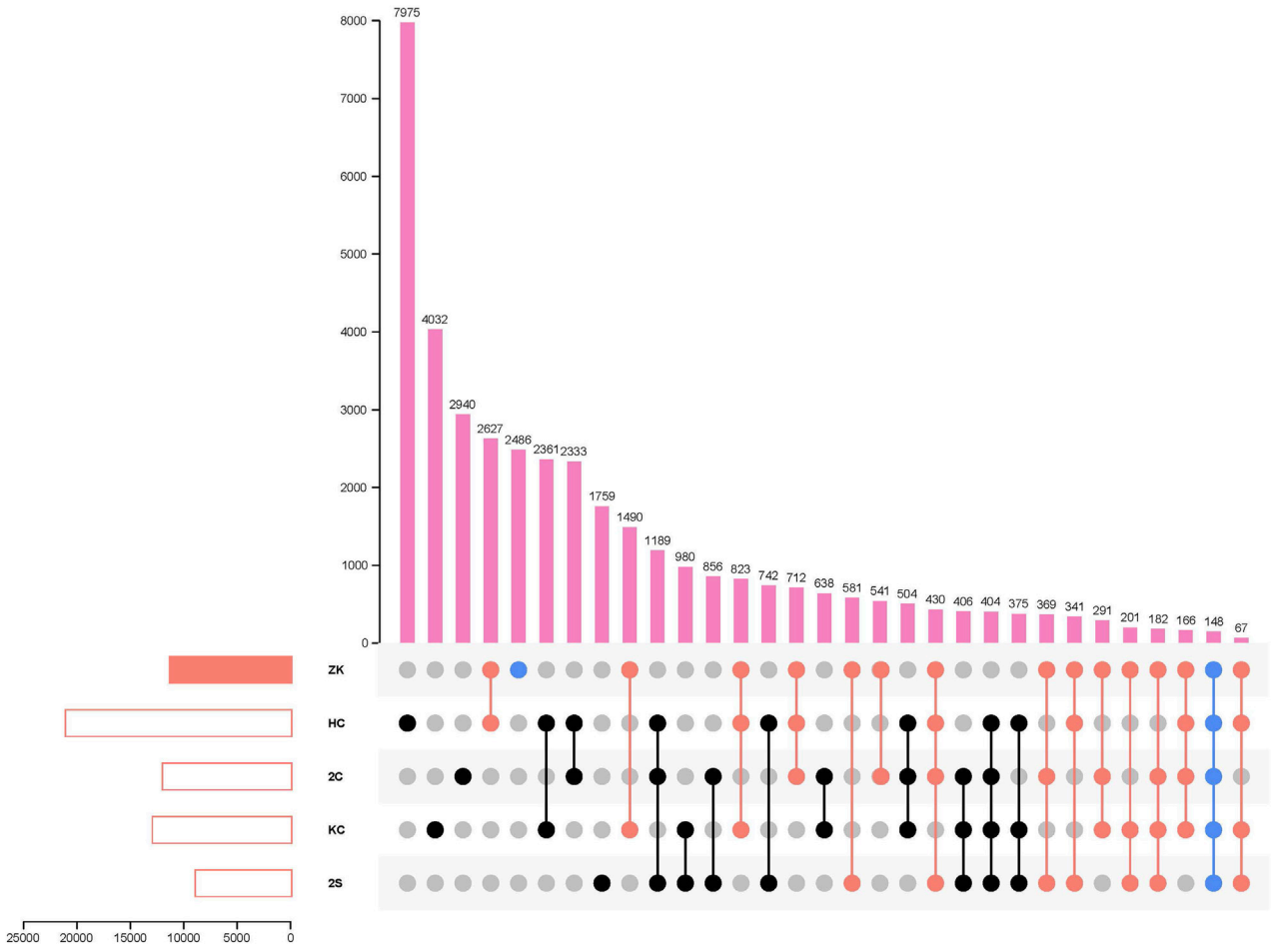


FIGURE 1
Venn diagram of the constructed Wheat 55K SNP array-derived genetic map based on ZK-RILs and the other four Wheat 55K SNP array-derived genetic maps for different populations.

2C-RILs map (Liu et al., 2018), 2S-RILs map (Liu et al., 2020), KC-DHs map (Li et al., 2021) and HC-RILs (Lin et al., 2021) (Figure 1; Supplementary Table S4). The most identical markers (2,627 SNPs) were discovered between ZK-RILs and HC-RILs maps (Lin et al., 2021), while only 541 common markers were discovered between ZK-RILs and 2C-RILs maps (Liu et al., 2018), which may be related to the number of markers in their respective genetic maps and the mapping parents' kinship. In particular, 2,486 special markers for ZK-RILs were likely less reported, showing that the development of new genetic populations and genetic maps could aid in the discovery of novel genes or alleles. Across five maps, just 148 common markers were detected. These shared markers facilitated the mapping of gene loci and pan-genomic studies. This small number of shared markers underlines the need for distinct genetic maps from different genetic backgrounds to expand and update wheat genetic and genomic information.

3.3 Phenotypic evaluation and correlation analysis

In all environments, ZKM138 clearly outperformed KCM2 in grain yield and its nutrient compositions (GPW, GSW and GWW), and ZKM138 had a relatively higher SC but a lower PC than KCM2 (Supplementary Table S5). In terms of GRRTs, ZKM138 only had significantly ($p < 0.05$) stronger GH and ABS but relatively lower TW than KCM2. The measured traits in the ZK-RILs showed transgressive segregation (Supplementary Table S5) and approximately continuous variance (Figure 2). Heritability ranged from 38.95% (ABS) to 89.46% (MC). The ANOVA results revealed that the genotype variance and the environmental effects of the investigated traits were both significant at $p < 0.001$ (Supplementary Table S5).

All correlation coefficients (r) (Figure 2) were considerably positive between GY and its three compositions ($r > 0.92$),

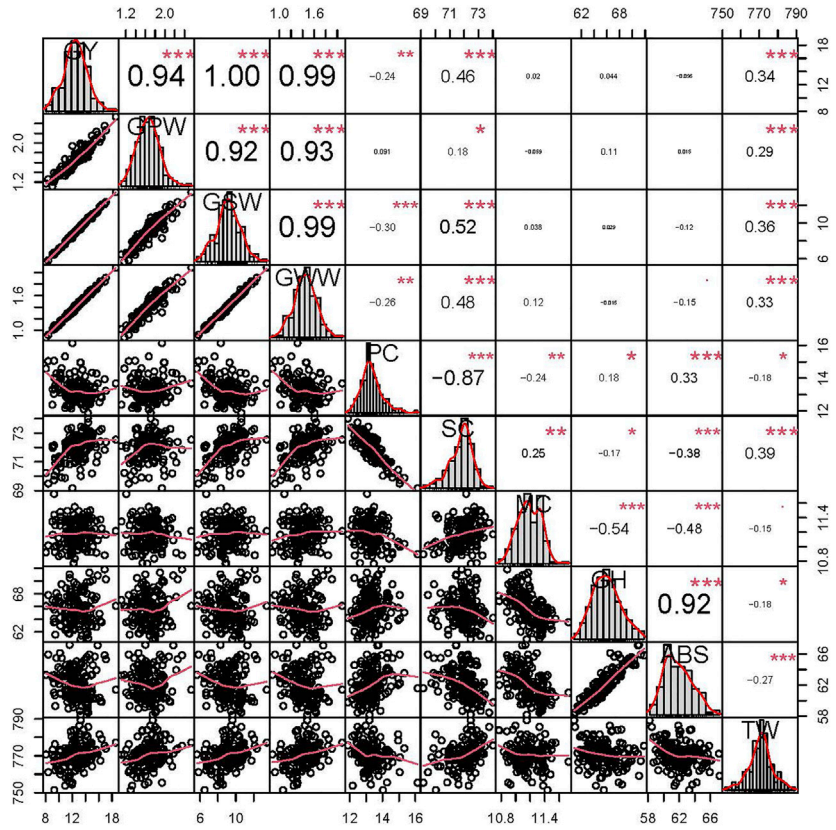


FIGURE 2
Phenotypic performance, distribution, and correlation coefficients of grain yield (GY), grain protein weight per plant (GPW), grain starch weight per plant (GSW), grain water weight per plant (GWW), protein content (PC), starch content (SC), moisture content (MC), grain hardness (GH), water absorption (ABS) and test weight (TW) in the ZK-RILs based on the BLUP data. "*", "**" and "***" represent significance at $p < 0.05$, $p < 0.01$, and $p < 0.001$, respectively.

including the correlation coefficient of GPW-GSW. However, PC was significantly negatively correlated with SC (r = −0.87), showing their well-known trade-offs. Except for the substantially strong correlation coefficient between ABS and GH (r = 0.92), significantly negative correlation coefficients (r = −0.18/−0.27) were observed for TW-GH and TW-ABS, respectively. In addition, for all measured traits, only SC and TW were shown to be significantly correlated with all of the other traits, particularly positively correlated with all four grain yield related traits, which were crucial determinants for wheat comprehensive commerciality.

3.4 Additive QTL mapping

3.4.1 Additive effect analysis in combined datasets based on the MCIM model

Using QTLNetwork 2.1, a total of 62 additive QTLs for ten measured traits were identified, distributed on 18 chromosomes except 1B, 1D, and 6B, and accounting for 0.25%–34.39% of the phenotypic variation (Supplementary Figure S4; Table 2; Supplementary Table S6). Seven and seventeen QTLs were classified as the major (PVE>10%) and moderate (5% < PVE<10%) QTLs, respectively, while the remaining loci were classified as the minor QTLs (PVE<5%).

In detail, for PC, SC, and MC, one major QTL for PC (*QPc.cib-4D*) was discovered on chromosome 4D, accounting for 10.49% of the phenotypic variation, and colocated with a moderate QTL for SC, *QSc.cib-4D.1*. The ZKM138-derived alleles decreased PC but increased SC at this locus, possibly providing the genetic foundation for the PC-SC trade-offs (Supplementary Figure S4; Table 2; Supplementary Table S6). Another major MC-related QTL, *QMc.cib-5D.1*, was discovered on chromosome 5D, with a ZKM138-derived allele decreasing MC.

For grain grading related traits, two major QTLs, affecting GH (*QGh.cib-5D.1*) and ABS (*QAbs.cib-5D.1*) explained 30.67%–34.39% of the phenotypic variance and were all mapped near *Pina* on chromosome 5D (Supplementary Figure S4; Table 2, Supplementary Table S6). At this location, a ZKM138-derived

allele may enhance both GH and ABS. Three moderate QTLs were discovered on chromosomes 2A (*QTw.cib-2A*), 6D2 (*QTw.cib-6D2*), and 7D (*QTw.cib-7D*), respectively. Among them, *QTw.cib-6D2* colocalized with another moderate QTL for MC (*QMc.cib-6D2*), with ZKM138-derived alleles decreasing TW but increasing MC.

Thirty-one QTLs were discovered for grain yield and three physical compositions (7 for GY, 9 for GWW, 8 for GPW, and 7 for GSW) (Supplementary Figure S4; Table 2; Supplementary Table S6). The interval *Rht2*–*AX-108905056* on the 4D chromosome was identified, which simultaneously clustered three major QTLs for GY, GPW, and GSW (*QGy.cib-4D*, *QGpw.cib-4D*, and *QGsw.cib-4D*) and one moderate QTL for GWW (*Gww.cib-4D*). The ZKM138 allele improved all four traits at this locus. Furthermore, the colocated interval on chromosome 2B grouped four QTLs (*QGy.cib-2B.1*, *QGsw.cib-2B.1*, *QGpw.cib-2B.1*, and *QGww.cib-2B*) with a moderate influence on all four traits and a positive additive effect from KCM2.

3.4.2 Additive effect analysis in single dataset based on the ICIM model

Using IciMapping 4.1, twenty QTLs with LOD thresholds greater than 3.42 were identified (Supplementary Figure S5; Table 2; Supplementary Table S7). These QTLs explained 5.46%–35.56% of the phenotypic variation, with an average LOD value ranging from 3.47 to 12.06. Among these loci, fourteen were consistently significant under both test conditions, eight of which clustered in three intervals on chromosomes 4D (*QGy.cib-4D*, *QGpw.cib-4D* and *QGww.cib-4D.1*), 5D (*QAbs.cib-5D.1*, *QGh.cib-5D.1* and *QMc.cib-5D.1*) and 6D (*QMc.cib-6D2* and *QTw.cib-6D2*), respectively. Noticeably, all three involved QTLs on chromosome 5D presented relatively higher PVE (35.56%, 32.98% and 24.46%, respectively) based on the ICIM model, which was consistent with the results detected by QTLNetwork 2.1, and they were all expressed across all datasets and thus were classified as major stable QTLs. Furthermore, because two loci on the 6D2 linkage (*QMc.cib-6D2* and *QTw.cib-6D2*) were persistently significant in all environments and repeatedly detected by two approaches, these two QTLs were modest but stable QTLs simultaneously controlling MC and TW. Only the GY composition-related QTLs were clustered for the interval on chromosome 4D.

3.5 Epistatic QTL mapping

For QTLnetwork 2.1, 47 pairs of epistatic QTLs were found, with PVE ranging from 0.91 to 11.77% (Supplementary Figure S5; Table 3; Supplementary Table S8). Two pairs that generated epistatic interactions for MC (*QMc.cib-4D*/*QMc.cib-5D.1* and *QMc.cib-5B*/*QMc.cib-6A*) also demonstrated an additive effect. Another QTL, *QAbs.cib-5D.1*, was shown to have an epistatic interaction with another epistatic QTL that had no additive effect (*QAbs.cib-3D*). There was just one pair of digenic QTLs (*QGy.cib-4A.1*/*QGy.cib-4A.2*) that showed a significant epistatic interaction effect (>10%) that explained 11.77% of the phenotypic variation. Except for this pair, the other 46 epistatic QTL pairs all interacted epistatically across different chromosomes.

ICIMapping 4.1 discovered thirty-five pairs of epistatic QTLs involving only three measured characteristics, namely ABS (8 pairs), GH (25 pairs), and GPW (1 pair) (Supplementary Figure S6, Supplementary Table S9). The other seven traits showed no epistasis over the LOD threshold (5.62). Only the epistatic interaction between chromosomes 1B (*AX-94383682–AX-108841881*) and 5A (*AX-110368018–AX-109731422*) demonstrated a higher PVE for GPW of 11.53%, which was also significant by QTLNetwork 2.1. Unlike the epistatic QTL mapping results from QTLNetwork 2.1, only three QTL pairs interacting across different chromosomes were detected by ICIMapping 4.1, but the epistatic effect of the remaining 32 pairs (91.43%) was generated by loci on the same chromosome.

3.6 QTL clusters

In this study, 13 intervals clustering two or more additive QTLs were observed (Supplementary Table S10). Only C4D.2, C5D, and C6D2 had at least two stable QTLs that could be detected repeatedly using both MCIM and ICIM methods (Figure 3; Supplementary Table S10). C4D.2, involving the famous semidwarfing gene *Rht2*, was clustered by four QTLs affecting grain yield and compositions with higher PVE values (moderate or major), one minor QTL for GH, and one moderate QTL for MC, showing that it mostly affects total yield and nutrient yield. C5D had two major QTLs for GH and ABS, probably associated with the famous hardness locus, *Pina*; two major or moderate QTLs for SC and MC; and one minor QTL for GWW, indicating its potential role in milling and processing. Furthermore, C6D2 has two moderate but consistent QTLs for TW and MC (*QTw.cib-6D2* and *QMc.cib-6D2*), implying a relation to wheat grading.

In addition to the additive QTLs, twenty-three epistatic QTLs controlling various attributes showed colocation in 10 intervals (Supplementary Table S11), indicating that epistasis could be pleotropic or linked. The epistatic interaction generated by the *AX-109412207–AX-108965184* interval on chromosome 3B and the *AX-109200636–AX-110438066* interval on chromosome 6D2 was shown to simultaneously influence GPW and GSW. Notably, *Pina* involved the epistatic interaction for GY, ABS, and MC with three different intervals on chromosomes 1B, 3D, and 4D (Figure 4), as well as the main additive effect on ABS, MC, GH, SC, and GWW. Among them, its interaction involving the

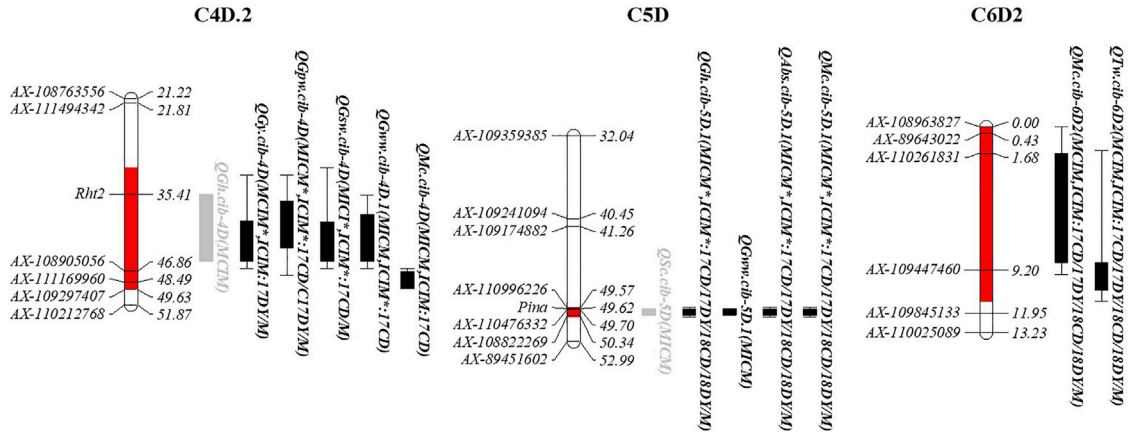


FIGURE 3
Genomic intervals harbouring at least two stable QTLs. The red interval on chromosomes indicates the confidence interval of this cluster. The rectangles to the right of the chromosome represent the QTLs. Black indicates significance under both detection approaches and is a stable QTL, while grey indicates it is only significant by one approach. The bar represents the confidence interval detected in one software, and the rectangle represents the overlapping confidence intervals. The brackets after the QTL name follow the detection method, and the asterisk means that both additive and epistatic effects are detected.

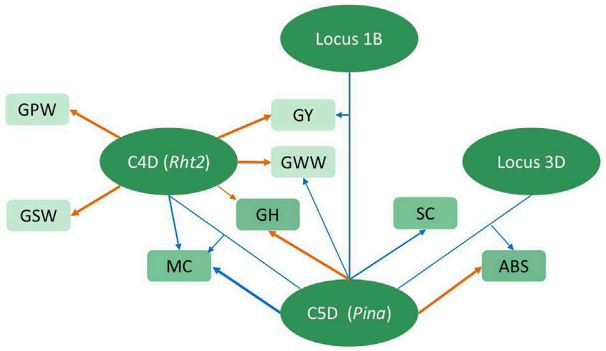


FIGURE 4
Schematic representation of the regulation of C4D and C5D on related traits. The ellipses represent loci, where C4D represents the *Rht2–AX-109297407* segment on chromosome 4D, possibly controlled by *Rht2*, C5D represents the *Pina–AX-110476332* segment on chromosome 5D, possibly controlled by *Pina*, Locus 1B represents the *AX-111234844–AX-110918909* segment on chromosome 1B, and Locus 3D represents the *AX-94528475–AX-110443023* segment on chromosome 3D. Rectangles represent the corresponding traits. Arrows emanating directly from the locus to the trait represent additive regulatory effects of the locus on the corresponding trait. Arrows linking two loci to the trait represent epistatic interactions between two loci regulating the corresponding trait. The different arrow thicknesses represent the magnitude of the PVE values. Orange arrows represent positive alleles derived from Zhongkemai138, and blue arrows represent positive alleles derived from KCM2.

MC was found to be involved with *Rht2*, which had the main effect on GY and compositions, indicating the complexity of the genetic effect on grain characteristics between these two genes, which might be explained by crucial structural genes being controlled by different regulatory factors to affect different traits.

3.7 Effects of *QTw/Mc.cib-6D2* on grain performance in the ZK-RILs

To validate the loci on chromosome 6D, a total of 15 KASP markers were developed and subsequently integrated into the genetic map of 6D2 linkage to trace *QTw.cib-6D2* and *QMc.cib-6D2* (Supplementary Table S1; Supplementary Table S12). Finally, a new genetic map for 6D2 linage with 51 markers were constructed to QTL mapping for the target phenotypes, i.e., TW and MC. Using the updated map, the confidence interval (CI) of *QTw.cib-6D2* and *QMc.cib-6D2* was relocated around a flanking KASP marker, *KASP14803* (Figure 5A) and its original flanking marker *AX-109447460*. Using the flanking marker *KASP14803*, ZK-RILs were genotyped to examine whether their accompanying phenotypes could be distinguished. Finally, TW and MC were remarkably different when the genotype was the AA (ZKM138 genotype) or BB (KCM2 genotype) (Figure 5B).

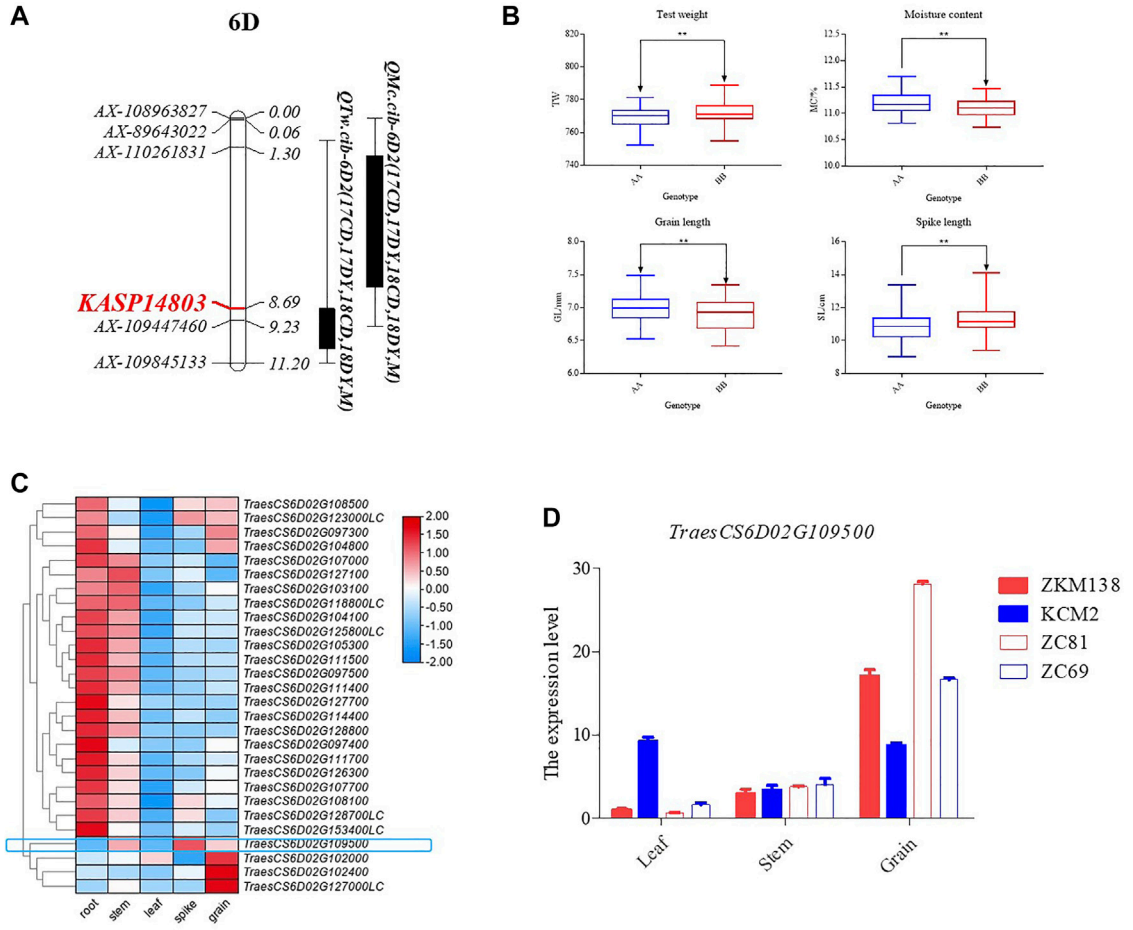


FIGURE 5
The candidate interval for *QTw/Mc-6D.2*. **(A)** The developed KASP marker for detecting *QTw/Mc-6D.2*s. The marker in red indicates the flanking KASP marker for *QTw/Mc-6D.2*. **(B)** The effect of *QTw/Mc-6D.2* on the corresponding traits validated by *KASP14803*. **(C)** Heatmap of the expression profiles for the candidate genes in this interval. **(D)** qPCR validation for *TraesCS6D02G109500*.

Aside from TW and MC, the group with ZKM138-derived alleles exhibited significantly shorter spike length (SL) but longer kernel length (KL) than the group with KCM2-derived alleles at this locus (Figure 5B). This result indicated that this locus not only controlled grain features but also affected spike formation and suggested that *KASP14803* was tightly linked to this target locus and could be convenient for the further molecular breeding.

4 Discussion

According to the physical location of these mapped SNPs in this study, these loci cover the majority of the wheat genome, demonstrating their applicability for further candidate gene excavation. However, several regions still failed to fill with markers on chromosomes 1D, 3A, and 6D, which were all composed of three individual linkage groups in ZK-RILs (Supplementary Figure S1; Supplementary Table S2), possibly explained by the low recombination events and frequency polymorphism in the corresponding intervals (Cui et al., 2017; Liu et al., 2018; Ren et al., 2018). Interval inversions were also found on multiple chromosomes, including 4D, 5D, and 7A (Supplementary Figure S1; Supplementary Table S2). These breakages and inversions are also frequently detected in other genetic maps (Cui et al., 2017; Liu et al., 2018; Ren et al., 2018; Li et al., 2021).

Furthermore, we discovered that in the other 55K SNP array derived genetic maps, some of their genetic markers do not match the physical location but share a similar position to the ZK-RILs genetic map (Supplementary Table S4), indicating that physical positions are relative and that different genetic backgrounds could induce alterations in genome arrangement information. The physical position of

TABLE 1 Details of the ZK-RILs genetic map.

Chromosome	Group	Bin	Marker number	Length (cM)	Bin resolution (cM)	Marker density (cM)
chr1A	1	56	767	45.02	0.80	0.06
chr2A	1	71	1,325	123.26	1.74	0.09
chr3A	1	14	72	10.68	0.76	0.15
	2	39	229	23.43	0.60	0.10
	3	16	74	25.92	1.62	0.35
chr4A	1	60	223	119.41	1.99	0.54
chr5A	1	99	484	117.19	1.18	0.24
chr6A	1	65	791	84.41	1.30	0.11
chr7A	1	75	431	102.83	1.37	0.24
chr1B	1	76	484	78.61	1.03	0.16
chr2B	1	68	439	108.79	1.60	0.25
chr3B	1	40	424	81.03	2.03	0.19
chr4B	1	63	774	103.24	1.64	0.13
chr5B	1	105	811	65.02	0.62	0.08
chr6B	1	52	888	58.41	1.12	0.07
chr7B	2	65	457	111.33	1.71	0.24
chr1D	1	50	326	87.66	1.75	0.27
	2	10	110	37.23	3.72	0.34
	3	6	6	8.80	1.47	1.47
chr2D	1	84	637	164.43	1.96	0.26
chr3D	1	55	556	83.41	1.52	0.15
chr4D	1	30	59	85.44	1.45	1.45
chr5D	1	85	252	130.59	1.55	0.52
chr6D	1	17	65	32.44	1.91	0.50
	2	37	202	77.41	2.09	0.38
	3	16	62	11.02	0.69	0.18
chr7D	3	105	507	178.73	1.70	0.35
A genome	12	495	4,396	652.15	1.32	0.15
B genome	8	469	4,277	606.41	1.29	0.14
D genome	19	495	2,782	897.16	1.81	0.32
Total	27	1,459	11,455	2,155.72	1.48	0.19

Note: Density was calculated by dividing their added genetic length by their added bin markers.

AX-111180568, for example, was on the 2D chromosome, but it was localized to chromosome 2A in this ZK-RILs map and the other four 55K array generated maps (Liu et al., 2018; Liu et al., 2020; Li et al., 2021; Lin et al., 2021) (Supplementary Table S4). Furthermore, the 6D genetic map is generally consistent with the physical map and other genetic maps, indicating that it might be well conserved. It should also be highlighted that some markers specific to the ZK-RILs genetic map were not consistent with physical positions and differed from other genetic map locations. For example, seventeen SNP markers mapped in the 31.06 cM-bin on chromosome 6D of the ZK-RILs genetic map were from chromosome 1A according to both physical location and other genetic maps (Cui et al., 2017; Lin et al., 2021) (Supplementary Table S2; Supplementary Table S4). This result demonstrated that maps from different genetic backgrounds could reflect their own genetic features inherited from different parents, laying the foundation for novel gene mining.

The dwarfing genes and hardness loci are commonly distributed in common wheat and affect wheat yield and quality (Ellis et al., 2002; Li et al., 2006; Qamar et al., 2014). To validate the usability of this map, we added molecular markers of the dwarfing gene *Rht2* (Ellis et al., 2002) and hardness locus *Pina* (Giroux and Morris, 1998; Li et al., 2006) for genotyping ZK-RILs and finally found that, as expected, they were integrated on chromosomes 4D (Börner et al., 1996) and 5D (Giroux and Morris 1998) and successfully located major QTLs controlling plant height (data not shown) and grain hardness (*QGh.cib-5D.1*), respectively. In this study, four major QTLs for grain yield and its physical composition

TABLE 2 Stable additive QTLs detected by both the MCIM model and ICIM model.

Trait	QTL[a]	Marker interval	Range[b]	MCIM		ICIM	
				Additive effect	H^2 (%)	Additive effect	PVE (%)
ABS	***QAbs.cib-5D.1****	*Pina–AX-110476332*	49.5–50.5	1.23	34.39	1.33	35.56
GH	***QGh.cib-5D.1***	*Pina–AX-110476332*	49.5–50.5	1.54	30.67	1.51	32.99
MC	*QMc.cib-4D**	*AX-108905056–AX-111169960*	46.5–49.5	−0.05	7.29	−0.05	7.19
	QMc.cib-5D.1	*Pina–AX-110476332*	49.5–50.5	−0.08	19.79	−0.10	24.46
	QMc.cib-6D2	*AX-110261831–AX-109447460*	0–9.5	0.06	9.82	0.06	8.13
TW	*QTw.cib-6D2*	*AX-110261831–AX-109447460*	1.5–11.2	−1.93	7.81	−2.18	8.21
	QTw.cib-7D	*AX-110033966–AX-108785845*	29.5–34.5	−2.04	8.7	−2.25	8.68
GY	***QGy.cib-4D***	*Rht2–AX-108905056*	32.5–46.5	0.74	13.83	0.85	11.97
SC	*QSc.cib-2D*	*AX-109525831–AX-108960866*	40.5–43.3	0.17	3.45	0.30	8.96
GPW	*QGpw.cib-3B*	*AX-109819016–AX-110392622*	39.5–40.5	0.07	5.9	0.12	9.21
	QGpw.cib-4B.1	*AX-109294476–AX-111176263*	15.5–20.6	−0.08	7.59	−0.11	6.58
	QGpw.cib-4D	*Rht2–AX-108905056*	32.5–47.5	0.11	14.94	0.13	10.94
GSW	***QGsw.cib-4D***	*Rht2–AX-108905056*	31.5–46.5	0.51	12.28	0.62	11.78
GWW	***QGww.cib-4D.1***	*Rht2–AX-108905056*	35.5–46.5	0.06	8.4	0.10	11.79

[a]The QTL, in bold are major additive QTLs; *Indicates the additive QTLs, with epistatic effects.
[b]The overlapping interval identified by QTLNetwork 2.1 based on the MCIM, model and IciMapping 2.1 based on the ICIM, model.

TABLE 3 Stable epistatic QTLs detected by both the MCIM model and ICIM model.

QTL1	Marker interval 1	Range 1	QTL2	Marker interval 2	Range 2	MCIM		ICIM	
						AA	H^2 (%)	AA	PVE (%)
QGpw.cib-1B.1	*AX-94383682–AX-108841881*	0.0–5.0	*QGpw.cib-5A*	*AX-109731422–AX-109342568*	94.9–97.0	0.03	1.51	0.09	11.53

were clustered at this Rht2 locus. Consistent with previous research indicating that *Rht2* has an effect on yield performance (Borrell et al., 1991), ZKM138-derived alleles at this locus could improve yield, possibly offering the genetic basis for ZKM138s excellent yield performance. Despite the fact that all three seed compositions, GWW, GSW, and GPW, exhibited a significant and positive correlation with grain yield (Figure 2), not all QTL influencing composition weight were able to influence the final grain yield. In contrast, only four of nine QTLs for GWW co-localized with GY-QTL (Supplementary Tables S6, S7). However, in the C4D.2 cluster associated with *Rht2*, not only GY- but also all GWW-, GSW-, and GPW-related QTLs were colocated in this region, indicating that *Rht2* may enhance the accumulation of all these physical substances to increase final grain yield, which might be related to the different competition capacity for carbohydrates or proteins between stem elongation and grain development (Gent and Kiyomoto, 1997; Triboi and Triboi-Blondel, 2002; Shearman et al., 2005). Although *Rht2* raised GY and its major compositions, there was no significant additive effect on the concentration parameters of protein and starch (PC and SC). Only an additive effect of *Rht2* controlling MC (*QMc.cib-4D*) was detected and showed an epistatic interaction for MC with *Pina* (Figure 4), indicating the existence of a complex regulatory network between these two genes and the potential relationship between plant height, yield composition, and grain hardness might be involved with the water accumulation and metabolism during grain development.

On the other hand, the C5D mapped additive QTL for SC, consistent with previous studies reporting that *Pina* regulating the grain starch granule synthesis (Capparelli et al., 2003; Wanjugi et al., 2007), indicating *Pina* is one of the dominant genes influencing flour quality. Its direct additive effect on GY has rarely been reported and also was not detected in this

study. However, the epistatic interaction on GY between *Pina* and the 589.95–668.67 Mb interval on chromosome 1B (*QGy.cib-1B*) was noticed. Therefore, the contribution of *Pina* to yield formation might also involve the accumulation of water and be governed by the epistasis of this *QGy.cib-1B*, taking into account the additive effect of GWW (*QGww.cib-5D*).

Previous research has shown that the Wheat 55K SNP array can satisfy primary QTL mapping in a genetic population of a similar size to that of this study (Liu et al., 2018; Ren et al., 2018). QTL mapping utilizing the primary mapping population to appropriately identify and find the genes responsible for particular agronomic traits is used for both major and moderate/minor QTLs (Cui et al., 2017). Previous research has shown that cloned genes are close to the positions detected in their primary mapping population (Cui et al., 2017). Three clusters on chromosomes 4D, 5D, and 6D were highlighted in this investigation because they all had at least two additive stable major/moderate QTLs that could be discovered repeatedly by both QTLNetwork 2.1 and IciMapping 4.1 (Table 2; Supplementary Table S10). However, C4D.2 and C5D have been discovered to be controlled by known genes. C6D2 for MC and TW might harbour novel loci, probably important in seeds storage and grading. According to the comparison of the resequencing results (data not shown) between the two parents, there were 174 identified genes with SNP or Indel variation (excluding intergenic regions) throughout the entire possible interval (60–95 Mb of the 6D chromosome). According to the public expression database (http://202.194.139.32/expression/wheat.html), only 28 of them had an overall expression level of greater than 5 in all tissues (Figure 5C). Only three genes were identified to be moderately highly expressed concurrently in grain and spike: *TraesCS6D02G123000LC*, *TraesCS6D02G108500*, and *TraesCS6D02G109500* (Figure 5C). Among them, *TraesCS6D02G109500*, which is located near 75 Mb of the 6D chromosome and encodes an aleurone layer morphogenesis protein, which might affect water adsorption and loss properties through controlling the development of the aleurone layer (constitutive of grain bran tissue) (Ying et al., 2020), finally regulating the water concentration and other related traits in seeds, was more likely the possible candidate gene, given its functional annotation for grain features and closeness to the overlapping candidate interval, but this conclusion is still preliminary.

The *TraesCS6D02G109500* expression pattern was preliminarily validated using qPCR, and the results were consistent with expectations, with grain expression being significantly higher than leaf and stem expression (Figure 5D). Furthermore, its expression was substantially higher in ZKM138 and line 81, both of which had ZKM138-derived genotypes validated by *KASP14803*, than in KCM2 and line 69, both of which had BB genotypes, indicating that the gene might have a role in grain features. This finding and the markers it yielded could be beneficial to grain performance genetic enhancement *via* MAS.

5 Conclusion

This study introduced a new genetic map using the Wheat 55K SNP array and presented a comparison with previously reported genetic and physical maps, which might provide information for wheat genetic and genomic studies. The additive and epistatic effects of QTLs were analysed using this map for six quality-related traits and four yield-related traits by two distinct QTL detection models. The major additive QTL affecting wheat grain yield and its compositions were localized around *Rht2*, suggesting that the dwarfing gene may affect yield by regulating the biomass accumulation of seed inner substances, while the major additive QTL influencing wheat seed hardness was localized around *Pina* and had epistatic interaction with *Rht2*. In addition to these two known genes, we also found a newly reported QTL for MC and TW on chromosome 6D that could be detected repeatedly in two different softwares, and finally located it around the flanking KASP marker (*KASP14803*). This marker could cleanly differentiate the MC and TW phenotypes separately and might be useful to the future molecular selection. Finally, a candidate gene encoding aleurone layer morphogenesis protein, *TraesCS6D02G109500*, was highlighted and requires further investigation and validation.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding authors.

Author contributions

XF undertook the field trials and subsequent analysis of all available data including the phenotyping and population

genotyping, and drafted this manuscript. ZX, HL, and GD assisted in field trials. XL, BF, QZ, JC, GJ, and SG participated in phenotyping. ZX and TW developed the population. XF, ZX, and TW discussed results. XF and TW designed the experiments, guided the entire study, participated in data analysis, discussed results and revised the manuscript.

Funding

This work is supported by the Sichuan Science and Technology Program, China (2022ZDZX0016) and the National Natural Science Foundation of China (31971934).

Acknowledgments

Triticeae Multi-omics Center (http://20219413932/) is also acknowledged for providing an integrated platform of genomic tools and data bringing great convenience to our work.

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.

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.978880/full#supplementary-material

References

Al-Mahasneh, M. A., and Rababah, T. M. (2007). Effect of moisture content on some physical properties of green wheat. *J. Food Eng.* 79, 1467–1473.

Al-Yahya, S. (2001). Effect of storage conditions on germination in wheat. *J. Agron. Crop Sci.* 186, 273–279.

Anjum, F. M., and Walker, C. (1991). Review on the significance of starch and protein to wheat kernel hardness. *J. Sci. Food Agric.* 56, 1–13. doi:10.1002/jsfa.2740560102

Aoki, N., Whitfeld, P., Hoeren, F., Scofield, G., Newell, K., Patrick, J., et al. (2002). Three sucrose transporter genes are expressed in the developing grain of hexaploid wheat. *Plant Mol. Biol.* 50, 453–462. doi:10.1023/a:1019846832163

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

Bhave, M., and Morris, C. F. (2008). Molecular genetics of puroindolines and related genes: Allelic diversity in wheat and other grasses. *Plant Mol. Biol.* 66, 205–219. doi:10.1007/s11103-007-9263-7

Bonneau, J., Taylor, J., Parent, B., Bennett, D., Reynolds, M., Feuillet, C., et al. (2013). Multi-environment analysis and improved mapping of a yield-related QTL on chromosome 3B of wheat. *Theor. Appl. Genet.* 126, 747–761. doi:10.1007/s00122-012-2015-3

Börner, A., Plaschke, J., Korzun, V., and Worland, A. (1996). The relationships between the dwarfing genes of wheat and rye. *Euphytica* 89, 69–75. doi:10.1007/bf00015721

Borrell, A. K., Incoll, L., and Dalling, M. J. (1991). The influence of the *Rht 1* and *Rht 2* alleles on the growth of wheat stems and ears. *Ann. Bot.* 67, 103–110. doi:10.1093/oxfordjournals.aob.a088108

Branlard, G., Dardevet, M., Saccomano, R., Lagoutte, F., and Gourdon, J. (2001). Genetic diversity of wheat storage proteins and bread wheat quality. *Euphytica* 119, 59–67. doi:10.1023/a:1017586220359

Cabral, A. L., Jordan, M. C., Larson, G., Somers, D. J., Humphreys, D. G., and McCartney, C. A. (2018). Relationship between QTL for grain shape, grain weight, test weight, milling yield, and plant height in the spring wheat cross RL4452/'AC Domain. *PLoS One* 13, e0190681. doi:10.1371/journal.pone.0190681

Capparelli, R., Borriello, G., Giroux, M., and Amoroso, M. (2003). Puroindoline A-gene expression is involved in association of puroindolines to starch. *Theor. Appl. Genet.* 107, 1463–1468. doi:10.1007/s00122-003-1392-z

Colasuonno, P., Marcotuli, I., Gadaleta, A., and Soriano, J. M. (2021). From genetic maps to QTL cloning: An overview for durum wheat. *Plants* 10, 315. doi:10.3390/plants10020315

Cui, F., Zhang, N., Fan, X. L., Zhang, W., Zhao, C. H., Yang, L. J., et al. (2017). Utilization of a Wheat660K SNP array-derived high-density genetic map for high-resolution mapping of a major QTL for kernel number. *Sci. Rep.* 7, 3788. doi:10.1038/s41598-017-04028-6

Day, L., Augustin, M., Batey, I., and Wrigley, C. (2006). Wheat-gluten uses and industry needs. *Trends Food Sci. Technol.* 17, 82–90. doi:10.1016/j.tifs.2005.10.003

Ellis, M., Spielmeyer, W., Gale, K., Rebetzke, G., and Richards, R. (2002). Perfect" markers for the *Rht-B1b* and *Rht-D1b* dwarfing genes in wheat. *Theor. Appl. Genet.* 105, 1038–1042. doi:10.1007/s00122-002-1048-4

Fang, C., and Campbell, G. M. (2003). On predicting roller milling performance V: Effect of moisture content on the particle size distribution from first break milling of wheat. *J. Cereal Sci.* 37, 31–41. doi:10.1006/jcrs.2002.0476

Fradgley, N. S., Gardner, K., Kerton, M., Swarbreck, S. M., and Bentley, A. R. (2022). Trade-offs in the genetic control of functional and nutritional quality traits in UK winter wheat. *Heredity* 128, 420–433. doi:10.1038/s41437-022-00503-7

Gent, M. P., and Kiyomoto, R. K. (1997). Physiological and agronomic consequences of *Rht* genes in wheat. *J. Crop Prod.* 1, 27–46. doi:10.1300/j144v01n01_02

Giroux, M. J., and Morris, C. F. (1998). Wheat grain hardness results from highly conserved mutations in the friabilin components puroindoline a and b. *Proc. Natl. Acad. Sci. U. S. A.* 95, 6262–6266. doi:10.1073/pnas.95.11.6262

Groos, C., Robert, N., Bervas, E., and Charmet, G. (2003). Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theor. Appl. Genet.* 106, 1032–1040. doi:10.1007/s00122-002-1111-1

Gudi, S., Saini, D. K., Singh, G., Halladakeri, P., Kumar, P., Shamshad, M., et al. (2022). Unravelling consensus genomic regions associated with quality traits in wheat using meta–Analysis of quantitative trait loci. *Planta* 255, 115–119. doi:10.1007/s00425-022-03904-4

Guo, Y., Zhang, G., Guo, B., Qu, C., Zhang, M., Kong, F., et al. (2020). QTL mapping for quality traits using a high-density genetic map of wheat. *PLoS One* 15, e0230601. doi:10.1371/journal.pone.0230601

Heo, H., and Sherman, J. (2013). Identification of QTL for grain protein content and grain hardness from winter wheat for genetic improvement of spring wheat. *Plant Breed. Biotechnol.* 1, 347–353. doi:10.9787/pbb.2013.1.4.347

Hurkman, W. J., McCue, K. F., Altenbach, S. B., Korn, A., Tanaka, C. K., Kothari, K. M., et al. (2003). Effect of temperature on expression of genes encoding enzymes for starch biosynthesis in developing wheat endosperm. *Plant Sci.* 164, 873–881. doi:10.1016/s0168-9452(03)00076-1

Karunakaran, C., Muir, W., Jayas, D., White, N., and Abramson, D. (2001). Safe storage time of high moisture wheat. *J. Stored Prod. Res.* 37, 303–312. doi:10.1016/s0022-474x(00)00033-3

Kumar, N., Orenday-Ortiz, J. M., Kiszonas, A. M., Boehm, J. D., Jr, and Morris, C. F. (2019). Genetic analysis of a unique 'super soft' kernel texture phenotype in soft white spring wheat. *J. Cereal Sci.* 85, 162–167. doi:10.1016/j.jcs.2018.12.003

Li, G., He, Z., Peña, R. J., Xia, X., Lillemo, M., and Sun, Q. (2006). Identification of novel secaloindoline–A and secaloindoline-b alleles in CIMMYT hexaploid triticale lines. *J. Cereal Sci.* 43, 378–386. doi:10.1016/j.jcs.2005.12.010

Li, T., Deng, G., Tang, Y., Su, Y., Wang, J., Cheng, J., et al. (2021). Identification and validation of a novel locus controlling spikelet number in bread wheat (*Triticum aestivum* L.). *Front. Plant Sci.* 12, 611106. doi:10.3389/fpls.2021.611106

Lin, Y., Jiang, X., Hu, H., Zhou, K., Wang, Q., Yu, S., et al. (2021). QTL mapping for grain number per spikelet in wheat using a high-density genetic map. *Crop J.* 9, 1108–1114. doi:10.1016/j.cj.2020.12.006

Liu, J., Luo, W., Qin, N., Ding, P., Zhang, H., Yang, C., et al. (2018). A 55 K SNP array-based genetic map and its utilization in QTL mapping for productive tiller number in common wheat. *Theor. Appl. Genet.* 131, 2439–2450. doi:10.1007/s00122-018-3164-9

Liu, J., Tang, H., Qu, X., Liu, H., Li, C., Tu, Y., et al. (2020). A novel, major, and validated QTL for the effective tiller number located on chromosome arm 1BL in bread wheat. *Plant Mol. Biol.* 104, 173–185. doi:10.1007/s11103-020-01035-6

Malmberg, R. L., and Mauricio, R. (2005). QTL-based evidence for the role of epistasis in evolution. *Genet. Res.* 86, 89–95. doi:10.1017/S0016672305007780

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

Michel, S., Löschenberger, F., Ametz, C., Pachler, B., Sparry, E., and Bürstmayr, H. (2019). Simultaneous selection for grain yield and protein content in genomics–Assisted wheat breeding. *Theor. Appl. Genet.* 132, 1745–1760. doi:10.1007/s00122-019-03312-5

Mohamed, A. A., and Rayas-Duarte, P. (2003). The effect of mixing and wheat protein/gluten on the gelatinization of wheat starch☆☆Names are necessary to report factually on available data; however, the USDA neither guarantee nor warrants the standard of the product, and the use of the name by the USDA implies no approval of the product to the exclusion of others that may also be suitable. *Food Chem.* 81, 533–545. doi:10.1016/s0308-8146(02)00487-9

MuhamadII, and Campbell, G. M. (2004). Effects of kernel hardness and moisture content on wheat breakage in the single kernel characterisation system. *Innovative Food Sci. Emerg. Technol.* 5, 119–125. doi:10.1016/j.ifset.2003.10.003

Murai, J., Taira, T., and Ohta, D. (1999). Isolation and characterization of the three *Waxy* genes encoding the granule-bound starch synthase in hexaploid wheat. *Gene* 234, 71–79. doi:10.1016/s0378-1119(99)00178-x

Narasimhamoorthy, B., Gill, B., Fritz, A., Nelson, J., and Brown-Guedira, G. (2006). Advanced backcross QTL analysis of a hard winter wheat x synthetic wheat population. *Theor. Appl. Genet.* 112, 787–796. doi:10.1007/s00122-005-0159-0

Ortolan, F., and Steel, C. J. (2017). Protein characteristics that affect the quality of vital wheat gluten to be used in baking: A review. *Compr. Rev. Food Sci. Food Saf.* 16, 369–381. doi:10.1111/1541-4337.12259

Qamar, Z. U., Bansal, U. K., Dong, C. M., Alfred, R. L., Bhave, M., and Bariana, H. S. (2014). Detection of puroindoline (*Pina-D1* and *Pinb-D1*) allelic variation in wheat landraces. *J. Cereal Sci.* 60, 610–616. doi:10.1016/j.jcs.2014.07.007

Ren, T., Hu, Y., Tang, Y., Li, C., Yan, B., Ren, Z., et al. (2018). Utilization of a Wheat55K SNP array for mapping of major QTL for temporal expression of the tiller number. *Front. Plant Sci.* 9, 333. doi:10.3389/fpls.2018.00333

Sandhu, K. S., Mihalyov, P. D., Lewien, M. J., Pumphrey, M. O., and Carter, A. H. (2021). Combining genomic and phenomic information for predicting grain protein content and grain yield in spring wheat. *Front. Plant Sci.* 12, 613300. doi:10.3389/fpls.2021.613300

Sardari, R. R., Sutiono, S., Azeem, H. A., Galbe, M., Larsson, M., Turner, C., et al. (2019). Evaluation of sequential processing for the extraction of starch, lipids, and proteins from wheat bran. *Front. Bioeng. Biotechnol.* 7, 413. doi:10.3389/fbioe.2019.00413

Schuler, S. F., Bacon, R. K., Finney, P. L., and Gbur, E. E. (1995). Relationship of test weight and kernel properties to milling and baking quality in soft red winter wheat. *Crop Sci.* 35, 949–953. doi:10.2135/cropsci1995.0011183x003500040001x

Shearman, V., Sylvester-Bradley, R., Scott, R., and Foulkes, M. (2005). Physiological processes associated with wheat yield progress in the UK. *Crop Sci.* 45, 175–185.

Shevkani, K., Singh, N., Bajaj, R., and Kaur, A. (2017). Wheat starch production, structure, functionality and applications—A review. *Int. J. Food Sci. Technol.* 52, 38–58. doi:10.1111/ijfs.13266

Šramková, Z., Gregová, E., and Šturdík, E. (2009). Chemical composition and nutritional quality of wheat grain. *Acta Chim. Slovaca* 2, 115–138.

Sun, C., Dong, Z., Zhao, L., Ren, Y., Zhang, N., and Chen, F. (2020). The Wheat 660K SNP array demonstrates great potential for marker-assisted selection in polyploid wheat. *Plant Biotechnol. J.* 18, 1354–1360. doi:10.1111/pbi.13361

Triboi, E., and Triboi-Blondel, A-M. (2002). Productivity and grain or seed composition: A new approach to an old problem—invited paper. *Eur. J. Agron.* 16, 163–186. doi:10.1016/s1161-0301(01)00146-0

Tu, M., and Li, Y. (2020). Toward the genetic basis and multiple QTLs of kernel hardness in wheat. *Plants* 9, 1631. doi:10.3390/plants9121631

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

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

Wanjugi, H., Hogg, A., Martin, J., and Giroux, M. (2007). The role of puroindoline A and B individually and in combination on grain hardness and starch association. *Crop Sci.* 47, 67–76. doi:10.2135/cropsci2006.05.0310

Xie, X. S., Cui, S. W., Li, W., and Tsao, R. (2008). Isolation and characterization of wheat bran starch. *Food Res. Int.* 41, 882–887. doi:10.1016/j.foodres.2008.07.016

Xing, Y., Tan, Y., Hua, J., Sun, X., Xu, C., and Zhang, Q. (2002). Characterization of the main effects, epistatic effects and their environmental interactions of QTLs on the genetic basis of yield traits in rice. *Theor. Appl. Genet.* 105, 248–257. doi:10.1007/s00122-002-0952-y

Yang, J., Hu, C., Hu, H., Yu, R., Xia, Z., Ye, X., et al. (2008). QTLNetwork: Mapping and visualizing genetic architecture of complex traits in experimental populations. *Bioinformatics* 24, 721–723. doi:10.1093/bioinformatics/btm494

Ying, R., Li, T., Wu, C., and Huang, M. (2020). Role of aleurone cell walls in water diffusion and distribution within cereal grains. *J. Cereal Sci.* 93, 102952. doi:10.1016/j.jcs.2020.102952

Zhang, X., Yang, S., Zhou, Y., He, Z., and Xia, X. (2006). Distribution of the *Rht-B1b, Rht-D1b* and *Rht8* reduced height genes in autumn-sown Chinese wheats detected by molecular markers. *Euphytica* 152, 109–116. doi:10.1007/s10681-006-9184-6

Zhang, Y., Tang, J., Zhang, Y., Yan, J., Xiao, Y., Zhang, Y., et al. (2011). QTL mapping for quantities of protein fractions in bread wheat (*Triticum aestivum* L). *Theor. Appl. Genet.* 122, 971–987. doi:10.1007/s00122-010-1503-6

Zhu, F., Cai, Y., Sun, M., and Corke, H. (2009). Effect of phytochemical extracts on the pasting, thermal, and gelling properties of wheat starch. *Food Chem.* 112, 919–923. doi:10.1016/j.foodchem.2008.06.079

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 05 September 2022
DOI 10.3389/fgene.2022.939255

OPEN ACCESS

EDITED BY
Ajay Kumar,
North Dakota State University,
United States

REVIEWED BY
Jun Cui,
Hunan Normal University, China
Qiang Wei,
Nanjing Forestry University, China

*CORRESPONDENCE
Liyun Wan,
wanliyun2019@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 09 May 2022
ACCEPTED 21 July 2022
PUBLISHED 05 September 2022

CITATION
Ren W, Zhang J, He J, Fang J and Wan L (2022), Identification, expression, and association analysis of calcineurin B-like protein–interacting protein kinase genes in peanut.
Front. Genet. 13:939255.
doi: 10.3389/fgene.2022.939255

Identification, expression, and association analysis of calcineurin B-like protein–interacting protein kinase genes in peanut

Weifang Ren[1†], Juncheng Zhang[2†], Jie He[1], Jiahai Fang[1] and Liyun Wan[1*]

[1]Key Laboratory of Crop Physiology, Ecology and Genetic Breeding, Ministry of Education, Jiangxi Agricultural University, Nanchang, China, [2]College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou, China

Plants usually respond to the external environment by initiating a series of signal transduction processes mediated by protein kinases, especially calcineurin B-like protein–interacting protein kinases (CIPKs). In this study, 54 *CIPKs* were identified in the peanut genome, of which 26 were from cultivated species (named *AhCIPKs*) and 28 from two diploid progenitors (*Arachis duranensis*—*AdCIPKs* and *Arachis ipaensis*—*AiCIPKs*). Evolution analysis revealed that the 54 *CIPKs* were composed of two different evolutionary branches. The *CIPK* members were unevenly distributed at different chromosomes. Synteny analysis strongly indicated that whole-genome duplication (allopolyploidization) contributed to the expansion of *CIPK*. Comparative genomics analysis showed that there was only one common collinear *CIPK* pairs among peanut, *Arabidopsis*, rice, grape, and soybean. The prediction results of *cis*-acting elements showed that *AhCIPKs*, *AdCIPKs*, and *AiCIPKs* contained different proportions of transcription factor binding motifs involved in regulating plant growth, abiotic stress, plant hormones, and light response elements. Spatial expression profiles revealed that almost all *AhCIPKs* had tissue-specific expression patterns. Furthermore, association analysis identified one polymorphic site in *AdCIPK12* (*AhCIPK11*), which was significantly associated with pod length, seed length, hundred seed weight, and shoot root ratio. Our results provide valuable information of *CIPKs* in peanut and facilitate better understanding of their biological functions.

KEYWORDS

peanut, expression, adversity response, association analysis, *CIPK* genes

Introduction

Calcium is an important second messenger in plants (Feng et al., 2017), and the signal transduction pathway mediated by calcium plays an important role in plant growth, development, and stress (Mao et al., 2016). Calcium signals formed during plant growth are further transmitted by calcium sensor proteins. Common calcium sensor proteins include calmodulins (CaMs), calmodulin-like proteins (CMLs), calcium-dependent protein kinases (CDPK), calcineurin B-like proteins (CBL), and CBL-interacting protein kinase (CIPK) (Zhao et al., 2009). CIPK, the interacting protein kinase of CBL, is a kind of protein that interacts with activated CBL as a downstream protein and is also a kind of Ca^{2+}-dependent serine/threonine protein kinase (Kim et al., 2000), which contains a conserved catalytic kinase domain at its N-terminal (Albrecht et al., 2001). The C-terminal is an NAF regulatory control domain composed of 24 amino acids, which is highly conserved and mediates the interaction between CBL and CIPK protein (Kolukisaoglu et al., 2004).

Up to now, 25, 30, 16, and 43 *CIPK* genes have been identified in *Arabidopsis thaliana* (Yu et al., 2007), *Oryza sativa* (Kolukisaoglu et al., 2004), *Vitis vinifera* (Lu et al., 2017), and *Zea mays* (Chen et al., 2011), respectively. *CIPK* genes were widely involved in stress response, growth, and development regulation of plants (Hu et al., 2015). Studies of *Arabidopsis* have shown that the protein encoded by *AtCIPK23* played an important role in potassium metabolism (Wang and Wu, 2009). The combination of *AtCBL1*, *AtCBL9*, and *AtCIPK23* activated the potassium transport channel (Cheong et al., 2007). AtCIPK15 interacted with PP2C phosphatase ABI2 to form a complex, which controlled the expression of ABA-related genes (Li et al., 2006). *AtCIPK3* responded to various abiotic stresses through ABA-dependent or ABA-independent pathway (Xu et al., 2006). It was found that overexpression of *OsCIPK3*, *OsCIPK12*, and *OsCIPK15* improved cold, drought, and salt stress tolerance of rice, respectively (Xiang et al., 2007). Overexpression of wheat *CIPK24* increased the contents of Na^+ and antioxidant protective enzymes (Hrabak et al., 2003) and further improved the tolerance of *A. thaliana* to high salt stress (Deng et al., 2013). Apple *CIPK6* interacted with *AtCBL4* protein after transferring *A. thaliana AtCBL4* gene into apple and improved the resistance of apple seedlings to low temperature, drought, and high salt stress (Wang and Liu, 2018). With the completion of large-scale plant genome sequencing, the research on *CIPK* genes in soybean (Feng et al., 2015), poplar (Sun et al., 2015), and many other plants received considerable attention, but research on *CIPK* genes in peanut has not been reported yet.

Peanut is one of the major oil and cash crops in China (Pandey et al., 2020). Calcium is the second largest nutrient element in peanut (Rui, 2015). In recent years, frequent occurrence of land drought, cold damage, and soil salinization has been a serious impact on the increase of peanut yield and the improvement of peanut quality (Lu et al., 2017). Under the pressure of reducing production cost and protecting the environment, screening peanut varieties with low calcium tolerance and exploring multistress response proteins and stress resistance mechanisms in an adverse environment have become the top priority in studying the stress resistance breeding of peanut (Sanders et al., 2002). To deal with abiotic stresses such as drought, cold, and salt damage and improve the yield and quality of peanut, it is of great significance to explore peanut *CIPK* genes and reveal their role in the calcium signaling pathway. In this study, the *CIPK* genes of cultivated peanut and its two diploid progenitors were comprehensively analyzed. At last, a total of 54 *CIPK* genes were explicitly identified. Their basic protein information, exon–intron structure, phylogeny, and *cis*-acting elements were systematically analyzed, which provided valuable theoretical basis and genetic resources for the high-yield breeding of peanut.

Materials and methods

Genome-wide identification of the *CIPK* genes in peanut

To identify *CIPK* genes in peanut, the protein sequences of three peanut genomes were downloaded from PeanutBase (http://www.peanutbase.org/). The conserved domains of all proteins encoded by peanut genome were analyzed using the HMMER 3.0 software, and genes including both the Pkinase (PF00069.24) and NAF (PF03822.13) domains were selected as peanut *CIPK* candidates. The PROSTIE and SMART software were used to verify the 54 CIPKs as calcineurin B-like protein–interacting protein kinases. Candidates with PROTEIN_KINASE_DOM (PS50011) and NAF (PS50816) in PROSTIE and the S_TKc (SM00220) domain in SMART were selected as CIPKs. The physicochemical data such as gene number, coding sequence (CDS) length, amino acid number, isoelectric point, molecular weight, and EF hand structure number were obtained from PeanutBase or analyzed using the ExPASy Proteomics Sever online tool (Gasteiger et al., 2005).

Evolution and structure analysis of *CIPK* genes in peanut

The gene structure diagram was drawn using the GSDS 2.0 mapping software based on the genome annotation information of *CIPK* genes with the GFF format, which was downloaded from PeanutBase (http://www.peanutbase.org/). The phylogenetic tree using 179 CIPK proteins from peanut, rice, grape, *Arabidopsis*, and soybean was constructed using the MEGA 5.2 software with the neighbor-joining method (Bootstrap value 1,000, Poisson model, uniform rates, pairwise deletion). The analyses of the composition of conserved motifs were conducted using MEME (http://meme-suite.org/tools/

10.3389/fgene.2022.939255

TABLE 1 Information on CIPK genes identified in peanuts.

Gene name	Gene locus	CDS length (bp)	AA[a]	MW[b] (kDa)	pI[c]	TMD[d]	Chr
AdCIPK1	*Aradu.9W61Z*	1,386	461	51.86	8.38	0	Aradu.A01
AdCIPK2	*Aradu.TL55R*	1,582	453	51.10	8.48	0	Aradu.A01
AdCIPK3	*Aradu.K8K3S*	2002	463	52.10	8.64	0	Aradu.A01
AdCIPK4	*Aradu.SX4F9*	1,670	452	50.61	8.35	0	Aradu.A01
AdCIPK5	*Aradu.MRA83*	1,645	434	49.26	8.44	0	Aradu.A02
AdCIPK6	*Aradu.T9ESX*	1,221	406	46.38	8.80	0	Aradu.A02
AdCIPK7	*Aradu.HS592*	2,199	457	51.51	9.22	0	Aradu.A03
AdCIPK8	*Aradu.8B2M9*	2,311	466	51.93	8.84	0	Aradu.A03
AdCIPK9	*Aradu.73JAV*	1982	441	50.46	6.42	0	Aradu.A07
AdCIPK10	*Aradu.7W2Z9*	2,110	456	50.40	8.30	0	Aradu.A07
AdCIPK11	*Aradu.V638G*	1,378	457	51.72	5.85	0	Aradu.A08
AdCIPK12	*Aradu.Z7XZ9*	1,681	461	52.41	8.75	0	Aradu.A09
AdCIPK13	*Aradu.Q5XDE*	1,653	550	61.66	8.17	0	Aradu.A10
AiCIPK1	*Araip.L2Z00*	1,582	455	51.24	8.06	0	Araip.B01
AiCIPK2	*Araip.X0WZQ*	1,386	461	51.93	8.38	0	Araip.B01
AiCIPK3	*Araip.I6C5W*	1,591	427	47.72	8.37	0	Araip.B01
AiCIPK4	*Araip.J6DER*	2,212	456	51.33	8.09	0	Araip.B01
AiCIPK5	*Araip.MS6UX*	1,146	381	43.51	8.81	0	Araip.B01
AiCIPK6	*Araip.A3V01*	1,239	412	46.65	8.52	0	Araip.B02
AiCIPK7	*Araip.CMC8E*	1,317	413	47.25	6.54	0	Araip.B02
AiCIPK8	*Araip.B16CX*	1,663	508	57.37	9.79	0	Araip.B03
AiCIPK9	*Araip.M3K7N*	2,359	465	51.84	8.84	0	Araip.B03
AiCIPK10	*Araip.7IS5A*	1991	441	50.46	6.42	0	Araip.B07
AiCIPK11	*Araip.WP1GX*	2021	456	50.40	8.02	0	Araip.B07
AiCIPK12	*Araip.Z7THM*	2,259	452	51.10	6.72	0	Araip.B07
AiCIPK13	*Araip.KS6V8*	1,677	461	52.56	8.49	0	Araip.B09
AiCIPK14	*Araip.W1LHP*	1,653	550	61.63	8.17	0	Araip.B10
AiCIPK15	*Araip.882H9*	1,387	436	48.93	9.00	0	Araip.B10
AhCIPK1	*Arahy.TE3LXI*	1,386	461	51.86	8.38	0	Arahy.01
AhCIPK2	*Arahy.N6YX8I*	2,735	453	51.10	8.48	0	Arahy.01
AhCIPK3	*Arahy.069RBA*	2014	446	50.44	8.79	0	Arahy.01
AhCIPK4	*Arahy.HIWA31*	3,017	570	64.19	9.22	0	Arahy.01
AhCIPK5	*Arahy.MA0DIS*	2,252	448	50.59	9.21	0	Arahy.02
AhCIPK6	*Arahy.T0XBGT*	1,516	398	45.07	9.31	0	Arahy.02
AhCIPK7	*Arahy.9ML3HZ*	2,762	497	55.49	8.90	0	Arahy.03
AhCIPK8	*Arahy.KQQ5DM*	1,218	405	46.26	6.26	0	Arahy.07
AhCIPK9	*Arahy.50QZS1*	2,813	456	50.40	8.30	0	Arahy.07
AhCIPK10	*Arahy.YEIN47*	2,295	549	61.97	7.56	0	Arahy.08
AhCIPK11	*Arahy.R30FAJ*	2,874	461	52.41	8.75	0	Arahy.09
AhCIPK12	*Arahy.L0AIUK*	1,653	550	61.66	8.17	0	Arahy.10
AhCIPK13	*Arahy.RMG3R2*	1,365	454	51.14	8.18	0	Arahy.11
AhCIPK14	*Arahy.M8C26Q*	1,386	461	51.90	8.38	0	Arahy.11
AhCIPK15	*Arahy.4 × 641H*	2,976	501	56.15	8.66	0	Arahy.11
AhCIPK16	*Arahy.QIVE9X*	2015	446	50.44	8.79	0	Arahy.11
AhCIPK17	*Arahy.I9I2G1*	1948	417	47.05	8.97	0	Arahy.12
AhCIPK18	*Arahy.LMT476*	1,318	427	48.48	9.13	0	Arahy.12
AhCIPK19	*Arahy.84D15R*	1,365	454	51.18	8.79	0	Arahy.13
AhCIPK20	*Arahy.JQ1SFF*	2,694	465	51.84	8.84	0	Arahy.13

(Continued on following page)

TABLE 1 (*Continued*) Information on CIPK genes identified in peanuts.

Gene name	Gene locus	CDS length (bp)	AA[a]	MW[b] (kDa)	pI[c]	TMD[d]	Chr
AhCIPK21	*Arahy.EY5MJ2*	1,218	405	46.26	6.26	0	Arahy.17
AhCIPK22	*Arahy.42B8G7*	2,818	456	50.40	8.02	0	Arahy.17
AhCIPK23	*Arahy.9NB62H*	3,177	492	55.53	6.95	0	Arahy.17
AhCIPK24	*Arahy.XP1WSF*	2,874	461	52.56	8.49	0	Arahy.19
AhCIPK25	*Arahy.VY9V4D*	1,653	550	61.63	8.17	0	Arahy.20
AhCIPK26	*Arahy.D01KFK*	1,361	436	48.93	9.00	0	Arahy.20

[a]Length of the amino acid sequence.
[b]Molecular weight of the amino acid sequence.
[c]Isoelectric point of the AhCIPKs.
[d]Number of transmembrane domains, as predicted by the TMHMM Server v2.0.
CDS, coding sequence.

meme) with the maximum number 20 (classic mode, zero or one occurrence per sequence).

Cis-acting elements analysis of peanut *CIPKs*

We defined the 2-kb upstream sequence of the initiation codon as the promoter of the peanut *CIPKs* and downloaded it from PeanutBase to search for *cis*-acting regulatory elements through PlantCARE (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Then, only the *cis*-acting elements related to adversity stress were screened out statistically.

Expression profiles of *AhCIPK* genes in different tissues and treatments

RNA-seq datasets of 22 peanut tissues were downloaded from PeanutBase (http://www.peanutbase.org/) and NCBI SRA (https://www.ncbi.nlm.nih.gov/sra/), and the expression levels of *AhCIPK* genes in different tissues were obtained (Clevenger et al., 2016) with all raw data deposited as BioSamples SAMN03944933–SAMN03944990. The expression data (FPKM value) of peanut *CIPK* genes were normalized and output using the TBtools software (Chen et al., 2020). *Ralstonia solanacearum* infection was carried out as described before according to (Zhang et al. 2017). Submergence treatment followed the method described by (Zeng et al. 2021).

Candidate gene association mapping

The genotype data of the *CIPK* genes used for association analyses were obtained from the transcriptome sequencing data of a peanut germplasm population with 146 accessions (unpublished data). The phenotypes were collected from five environments (Wuhan 2016, Wuhan 2017, Yangluo 2016, Yangluo 2017, Zhanjiang 2016). Three replicates were randomly planted in each environment, with 12 plants in each row.

Results

Genome-wide identification of the *CIPK* genes in cultivated peanut and its diploid progenitors

To systematically determine *CIPK* genes in peanut, genes containing both the conserved Pkinase (PF00069.24) and NAF (PF03822.13) domains were searched through the whole peanut genome. The SMART and PROSTIE software tools were used to verify the Pkinase domains. A total of 54 *CIPK* candidates were identified from the peanut genome of cultivated species *Arachis hypogaea* (26, namely *AhCIPK1–AhCIPK26*) and its two wild species *Arachis duranensis* (13, namely, *AdCIPK1–AdCIPK13*) and *Arachis ipaensis* (15, namely, *AiCIPK1–AiCIPK15*) (Table 1). Then, we determined their chromosome locations, mRNA length, number of amino acids (aa), MW, theoretical pI, and transmembrane domain (TMD) (Table 1). *AhCIPK* genes were distributed on chromosomes 1, 2, 3, 7, 8, 9, and 10 (A genome) and 11, 12, 13, 17, 19, and 20 (B genome). Two wild species–specific *CIPK* genes (*AdCIPK7* and *AiCIPK5*) were uncovered (Figure 1). The gene length of the peanut *CIPKs* ranged from 1,146 to 3,177 bps, of which the shortest was *AiCIPK4* with 1,146 bps and the longest length was *AhCIPK22* with 3,177 bps. The amino acid length of *CIPKs* varied from 381 to 570. The isoelectric point ranged from 5.8 (*AdCIPK11*) to 9.79 (*AiCIPK8*), and the molecular weight

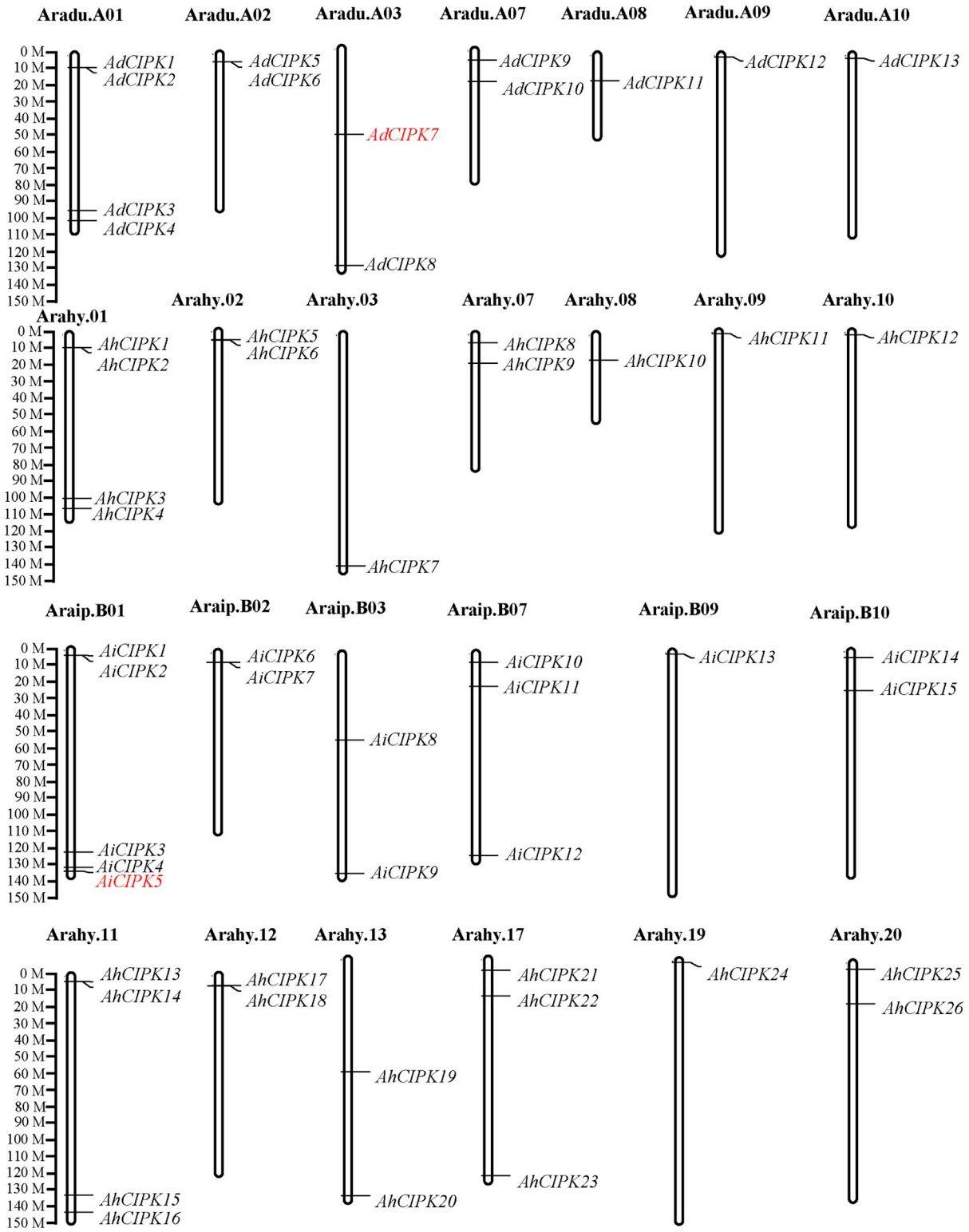


FIGURE 1
Chromosomal locations of peanut calcineurin B-like protein–interacting protein kinase (*CIPK*) genes. Chromosomal positions of the peanut *CIPK* genes were mapped based on GFF data downloaded from PeanutBase. The chromosome number is indicated above each chromosome. Genes in red mean wild species specific.

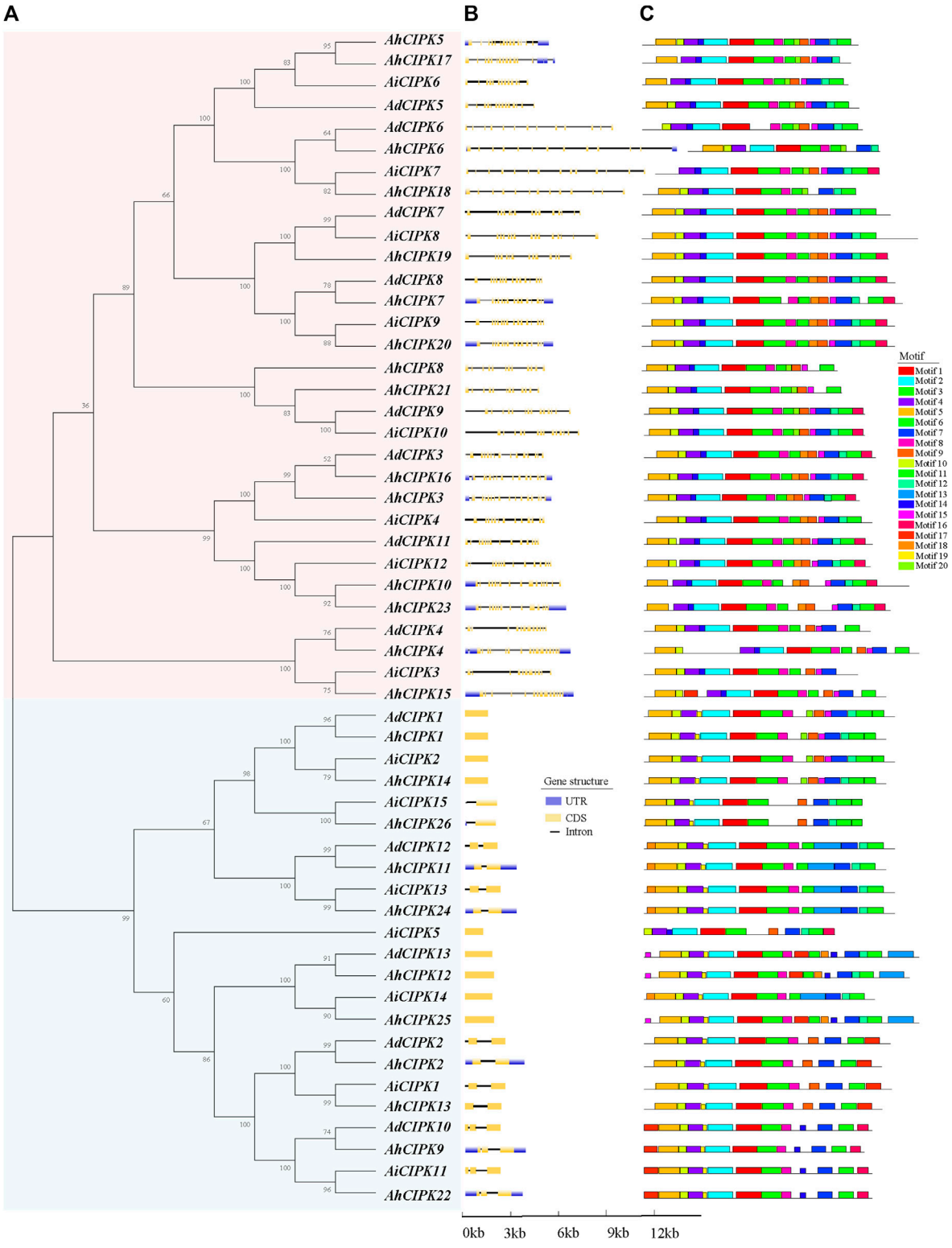


FIGURE 2
Comparison of the gene structure and motif of 54 *CIPK* genes in peanut. **(A)** unrooted phylogenetic tree with over 50% bootstrap value above the branch. Clades I and II are displayed in pink and blue colors, respectively. The names of the species are abbreviated to two letters, named as *Arachis duranensis* (Ad), *Arachis ipaensis* (Ai), and *Arachis hypogaea* L. *Tifrunner* (Ah). **(B)** exon/introns and untranslated regions (UTRs) of *CIPK*s. Blue boxes denote UTR (untranslated region); yellow boxes denote coding sequence (CDS); and black lines denote introns. The length of protein can be estimated using the scale at the bottom. **(C)** motif architectures of all *CIPK* genes. Each motif is illustrated with a specific color, and the distribution of identified motifs corresponds to their positions.

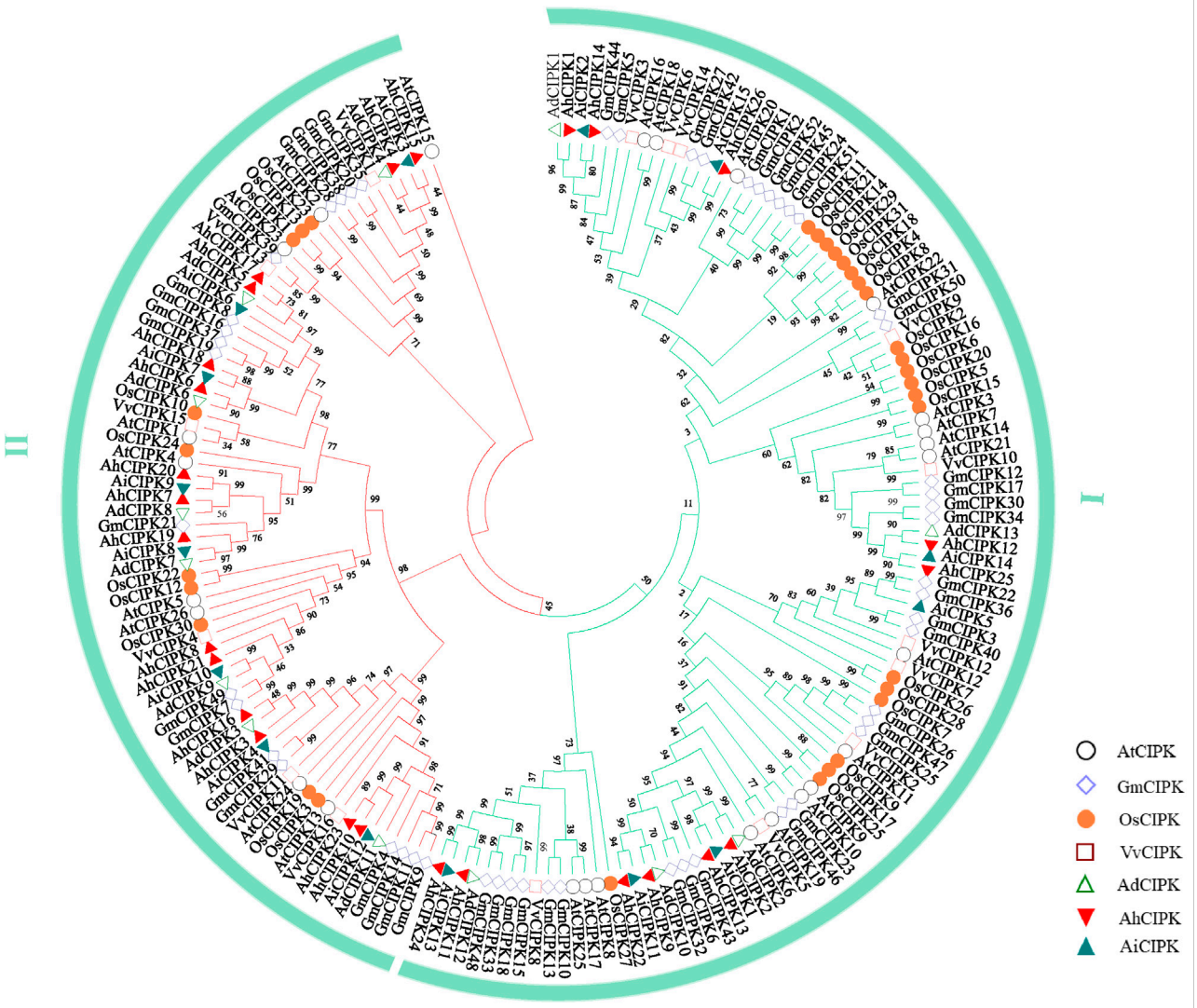


FIGURE 3
Phylogenetic tree of CIPKs in peanut and other plant species. The plants in the square frame indicate that the CIPK members outside of peanuts have the closest evolutionary relationship with soybean CIPKs.

ranged from 45,067.06 to 64,189.82 Da. All CIPKs do not contain a TMD.

Phylogenetic and gene structure analysis of *CIPKs* in peanut

To determine the evolutionary relationship of *CIPKs* among *A. hypogaea*, *A. duranensis*, and *A. ipaensis*, the phylogenetic tree of the 54 *CIPKs* was constructed. The results indicated that the *CIPKs* can be classified into two clades (I and II) (Figure 2A). Clades I and II consisted of 31 *CIPKs* (15 *AhCIPKs*, 8 *AdCIPKs*, and 8 *AiCIPKs*) and 23 *CIPKs* (11 *AhCIPKs*, 5 *AdCIPKs*, and 7 *AiCIPKs*), respectively. It is interesting that the results of the gene structure based on the genome annotations also showed that the *CIPK* genes can be divided into two groups, corresponding to the two phylogenetic families (the intron-rich group corresponded to phylogenetic clade I, and the intron-less group corresponded to phylogenetic clade II). The intron numbers of the intron-less group were less than 3 (0 to 2), while those of the intron-rich group were more than 10 (Figure 2B). Further, the phylogenetic relationship and classification of peanut CIPKs were supported by motif analysis. A total of 20 motifs were identified (Figure 2C); in general, peanut CIPKs had 11–17 motifs. Motif 1, Motif 2, and Motif 4 were the most common, present in all CIPK proteins. Otherwise, the vast majority of *CIPKs* included Motifs 3, 5, 6, 7, and 8, which covered more than 50 CIPK members. Motifs

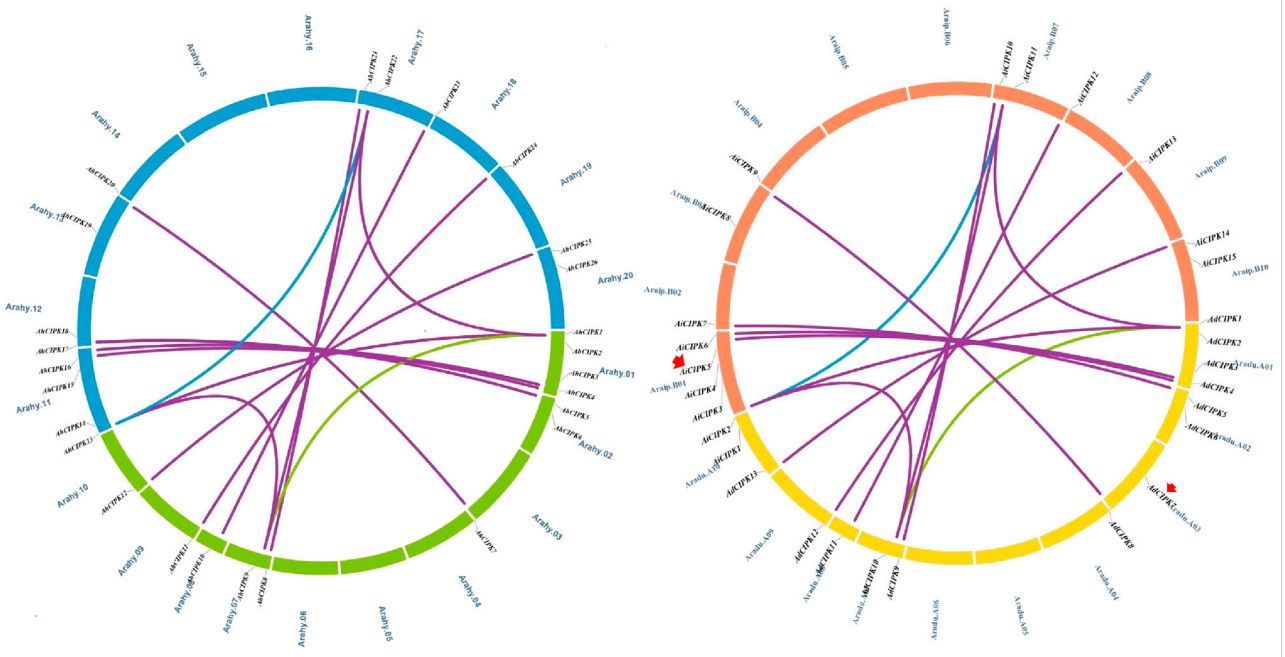


FIGURE 4
Chromosomal distribution and gene duplications of the *AdCIPK*, *AiCIPK*, and *AhCIPK* genes. The scales on the circle are in megabases. Each colored bar represents a chromosome as indicated. Gene IDs are labeled on the basis of their positions on the chromosomes. Red arrows indicate wild species–specific *CIPK* genes.

13 and 19 were clade-specific elements in clade II, and Motif 17 only existed in AhCIPK15 of clade I but 13 CIPKs in clade II.

Biological evolution analysis of CIPKs in peanut and other plant species

To further understand the relationship of CIPK members among different species, the phylogenetic tree of CIPK proteins of *Arabidopsis*, rice, grape, soybean [AtCIPK (26), OsCIPK (31), VvCIPK(16), and GmCIPK (52)], and peanut was constructed using maximum parsimony (Figure 3). The results showed that *CIPK* proteins of these species can be divided into two subfamilies (Ⅰ and Ⅱ). The analysis of the phylogenetic tree revealed that all the peanut CIPKs were clustered together (Figure 3). The relationships between the two wild species *A. duranensis* and *A. ipaensis* and cultivated species *A. hypogaea* were closer than that between the other four species. In addition, many CIPK members of peanut and soybean clustered together, suggesting that the two legume species were evolutionarily closer than others. TThe second evolutionary closest of peanut was *Arabidopsis*. There were more family members in peanut and soybean than in *Arabidopsis*, rice, and grape plants, suggesting a specific linear amplification of the gene family in legume plant. Whether these additional members of the genes have additional functions as well or whether they are produced only because functional redundancy requires further experimental verification.

Gene duplication and synteny analyses of peanut *CIPKs*

Chromosomal location analyses revealed that the 26 *AhCIPKs* distributed unevenly on 13 chromosomes (chromosomes 01, 02, 03, 07, 08, 09, 10, 11, 12, 13, 17, 19, and 20). The 13 *AdCIPK*s presented on chromosomes A01, A02, A03, A07, A08, A09, and A10, and the 15 *AiCIPK*s distributed on chromosomes B01, B04, B06, B09, and B10 (Figure 4). A total of 16 chromosomal fragment repeat gene pairs were identified without tandem repeats (Figure 4, Supplementary Table S1). Further, we calculated the Ks (synonymous) and Ka (nonsynonymous) values of the duplicated gene pairs and found that the Ka/Ks ratio for duplicated *AhCIPK* gene pairs ranged from 0.00 to 0.57 with an average of 0.17 (Supplementary Table S2). The ω values of all duplicated gene pairs were less than 1, showing that purifying selection occurred on these duplicated gene pairs. Synteny analysis with *Arabidopsis*, rice, grape, and soybean revealed one conserved *CIPK* gene (*AhCIPK14*) in these species (Figure 5, Supplementary Table S1). BLASTP methods were used to identify peanut *CIPK* gene orthologs between peanut and *Arabidopsis*. In total, we found 54 orthologous

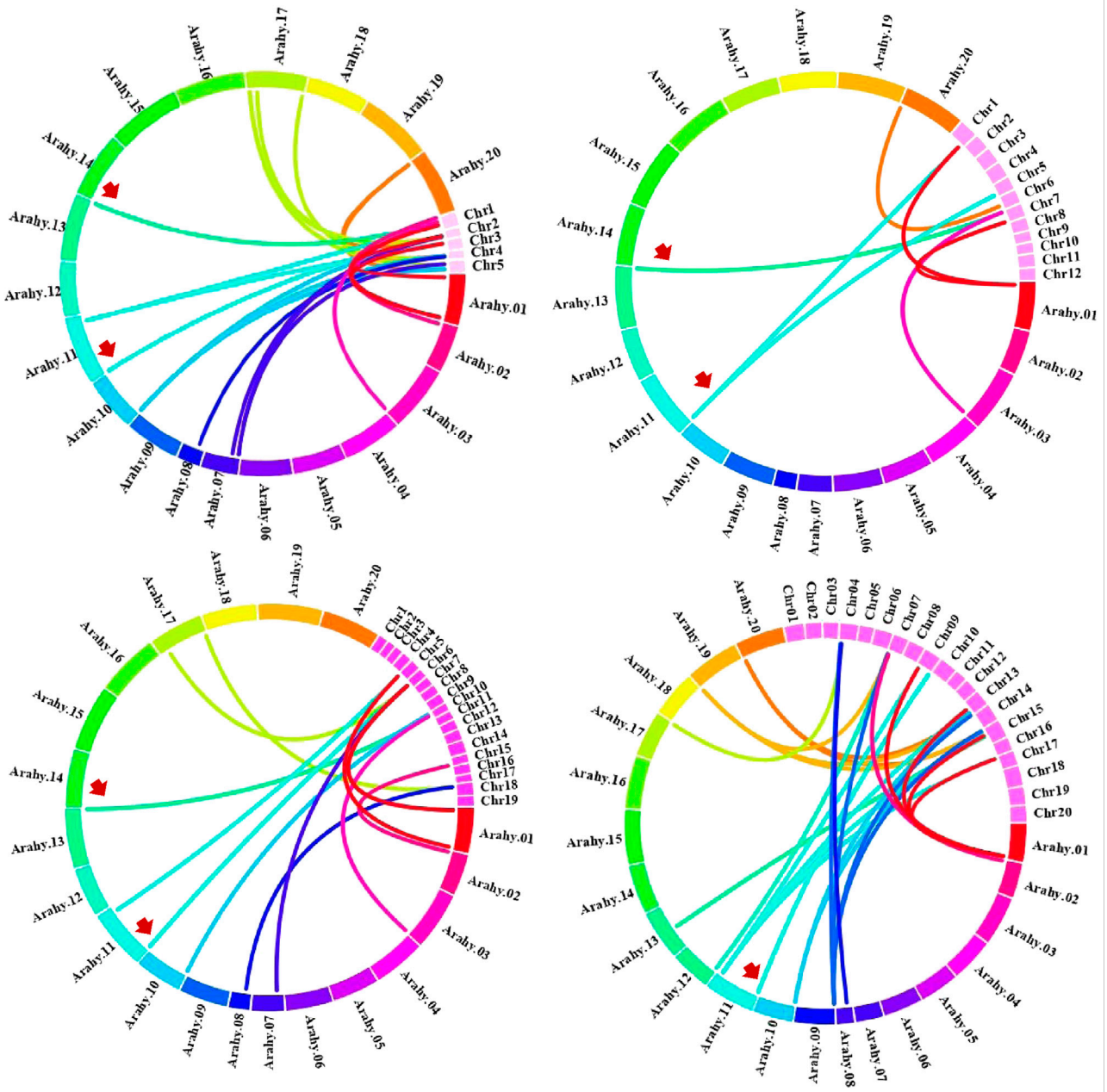


FIGURE 5
Comparative physical mapping showed orthologous relationships of AhCIPK genes with (**A**) *Arabidopsis,* (**B**) rice, (**C**) grape, and (**D**) soybean. Red arrows indicate common collinear *CIPK* orthologs.

gene pairs between peanut and *Arabidopsis* (Table 2). The orthologs in *Arabidopsis* included *AtCIPK12/AtWL4* and *AtCIPK5* participating in pollen germination and tube growth (Wang et al., 2008; Steinhorst et al., 2015), *AtCIPK24/AtSOS2* required for salt tolerance in *A. thaliana* (Halfter et al., 2000; Ishitani et al., 2000; Liu et al., 2000; Guo et al., 2001), and *AtCIPK1* and *AtCIPK3* relating to the ABA signal transduction (D'Angelo et al., 2006; Sanyal et al., 2017; Pandey et al., 2008; Kim et al., 2003). Therefore, we speculated that these *AhCIPK* homologous genes might play multiple roles not only in peanut growth and development but also in plant hormone and stress resistance.

TABLE 2 The function of AhCIPKs genes homologous to *Arabidopsis*.

Peanut	*Arabidopsis*	Function	References
AdCIPK4/AiCIPK3/AhCIPK4/AhCIPK15	*AtCIPK1* (*At3G17510*)	Controls abscisic acid-dependent and independent stress responses	D'Angelo et al. (2006)
AdCIPK5/AdCIPK6/AiCIPK6/AiCIPK7/AhCIPK5/AhCIPK6/AhCIPK17/AhCIPK18	*AtCIPK9* (*At1G01140*)	A calcium sensor-interacting protein kinase required for low-potassium tolerance	Pandey et al. (2007) Singh et al. (2018) Lara et al. (2020) Kanwar et al. (2022)
AdCIPK7/AdCIPK8/AiCIPK8/AiCIPK9/AhCIPK7/AhCIPK19/AhCIPK20	*AtCIPK23* (*At1G30270*)	Serves as a positive regulator of the potassium transporter AKT1 by directly phosphorylating AKT1	Sánchez-Barrena et al. (2020) Ragel et al. (2015) Tian et al. (2016) Wang et al. (2016) Xu et al. (2006) Cheong et al. (2007)
AhCIPK8/AhCIPK21/AdCIPK9/AiCIPK10	*AtCIPK3* (*At2G26980*)	Regulates Abscisic Acid and Cold Signal Transduction	Sanyal et al. (2017) Pandey et al. (2008) Kim et al. (2003)
AdCIPK2/AdCIPK10/AiCIPK1/AiCIPK11/AhCIPK2/AhCIPK9/AhCIPK13/AhCIPK22/	*AtCIPK11* (*At2G30360*)	A positive regulator in cadmium stress response	Zhou et al. (2015) Gu et al. (2021) Liu and Guo, (2011)
AdCIPK13/AiCIPK14/AhCIPK12/AhCIPK25	*AtCIPK12/AtWL4* (*At4G18700*)	Required for Polarized Pollen Tube Growth	Steinhorst et al. (2015)
AdCIPK11/AiCIPK12/AhCIPK10/AhCIPK23	*AtCIPK8* (*At4G24400*)	Regulates the low-affinity phase of the primary nitrate response	Hu et al. (2009) Gong et al. (2002)
AiCIPK5	*AtCIPK6* (*At4G30960*)	Required for development and salt tolerance	Sardar et al. (2017) Chen et al. (2013) Tsou et al. (2012) Held et al. (2011) Tripathi et al. (2009a) Tripathi et al. (2009b)
AdCIPK12/AiCIPK13/AhCIPK11/AhCIPK24	*AtCIPK5* (*At5G10930*)	Regulates potassium homeostasis under low oxygen	Schlücking et al. (2013) Tagliani et al. (2020)
AdCIPK3/AiCIPK4/AhCIPK3/AhCIPK16	*AtCIPK24/AtSOS2* (*At5G35410*)	SOS2 gene encodes a protein kinase that is required for salt tolerance	Liu et al. (2000) Halfter et al. (2000) Ishitani et al. (2000) Guo et al. (2001)
AdCIPK1/AiCIPK2/AiCIPK15/AhCIPK1/AhCIPK14/AhCIPK26	*AtCIPK5* (*At5G58380*)	Gene expression to accompany pollen germination and tube growth	Wang et al. (2008)

Cis-acting elements prediction of *CIPK* genes in peanut

Cis-acting elements in a promoter as the binding target of transcription factors are essential in the regulation of gene expression. In order to understand the regulation mechanisms of peanut *CIPK* genes, 2-kb upstream sequences of the peanut *CIPK* genes were analyzed via the PlantCARE database. In total, 54 *cis*-regulatory elements were detected. Four main categories were defined as the light responsiveness element, phytohormone responsiveness, abiotic stress responsiveness, and plant growth groups (Figures 6A–C, Supplementary Table S3). In the promoter region of the *AdCIPKs*, the largest subdivision was the light responsiveness group, containing 53.2% of the predicted *cis*-elements; phytohormone responsiveness elements ranked second (24.2%) (Figure 6A); abiotic stress response elements were 15.9%; and elements involved in plant growth accounted for 6.7% (Figure 6A). *AdCIPK2* had the greatest number of elements with 38 in total, which contained six abscisic acid responsiveness elements (ABREs) (Supplementary Table S3). For *AiCIPKs*, the percentage of light, phytohormone, abiotic stress, and plant growth responsiveness *cis*-elements was 57.7, 21.5, 16.6, and 4.2% (Figure 6B). *AiCIPK14* had the greatest number of elements at 40 in total, which also contained six ABREs. In

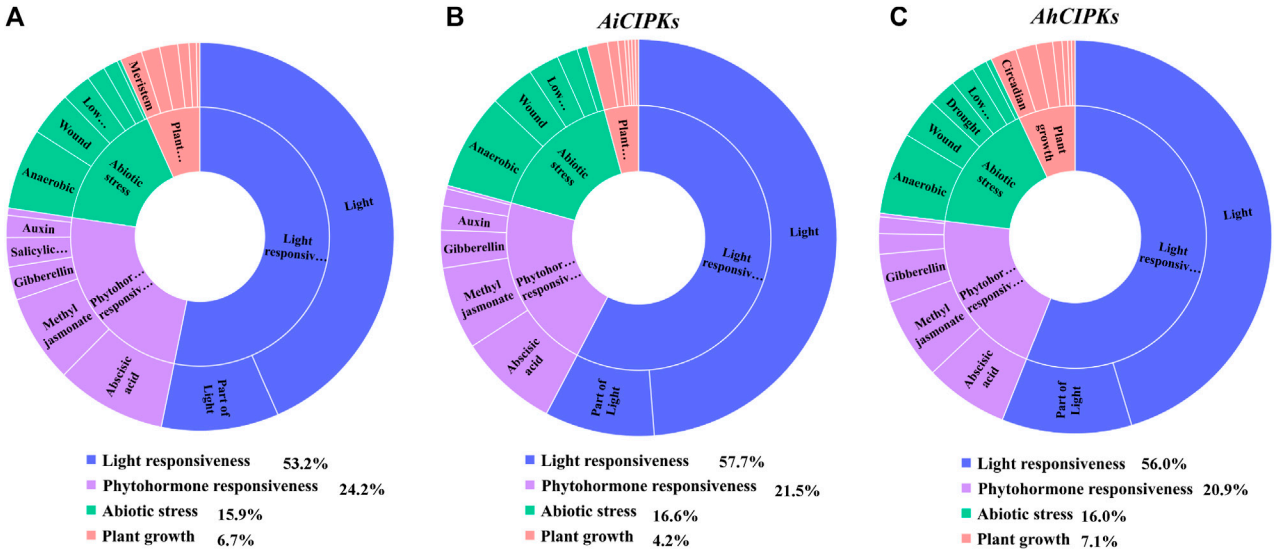


FIGURE 6
cis-acting elements in peanut *CIPK* genes. **(A)** assessment of *AdCIPK* different subclass and category proportions in sunburst chart. **(B)** assessment of *AiCIPK* different subclass and category proportions in sunburst chart. **(C)** assessment *AhCIPK* different subclass and category proportions in sunburst chart. The lengths of the petals are proportional to the number of elements in each subclass qualitatively. Blue, purple, green, and red represent light responsiveness, phytohormone responsiveness, abiotic stress responsiveness, and plant growth regulation, respectively.

AhCIPKs, the proportions were 56.0, 20.9, 16.0, and 7.1% (Figure 6C). In the light response category, Box 4 (light-responsive element) and GT1-motif (part of a module for light response) were the most dominant. Meanwhile, *cis*-acting elements responding to auxin, abscisic acid, gibberellin, flavonoids, methyl jasmonate, and salicylic acid were detected in the phytohormone responsiveness group.

Tissue expression profiles of *AhCIPKs*

To further study the expression pattern of peanut *CIPKs* in different tissues and explore its function in peanut growth and development, the tissue expression profiles of *CIPK* genes were analyzed by using the transcriptome data of 22 peanut tissues (Figure 7). The results showed that the 26 *AhCIPK* genes had distinct tissue-specific expression patterns across the 22 tissues (leaves, stem, roots, flower, pod, and seed). *AhCIPK9* and *AhCIPK25* showed a higher expression level in leaf. *AhCIPK7* and *AhCIPK20* were mostly expressed in reproductive shoot and pattee 1 stalk; interestingly, nearly half of *AhCIPK* genes were highly expressed in reproductive organs, among which *AhCIPK2*, *AhCIPK3*, *AhCIPK10*, *AhCIPK14*, *AhCIPK15*, *AhCIPK18*, and *AhCIPK22* had strong expression in perianth; *AhCIPK3*, *AhCIPK6*, *AhCIPK8*, *AhCIPK16*, *AhCIPK17*, and *AhCIPK21* were highly expressed in stamens; and *AhCIPK12* and *AhCIPK26* were obviously expressed in roots. *AhCIPK11*, *AhCIPK23*, and *AhCIPK24* were highly expressed in nodules. *AhCIPK4*, *AhCIPK12*, *AhCIPK19*, and *AhCIPK24* had strong expression during the relatively later pericarp developmental stage. In addition, *AhCIPK1* and *AhCIPK19* were enriched in the earlier seed developmental stage, while *AhCIPK4* and *AhCIPK13* were expressed highly in the later seed developmental stage (Figure 7).

Expression pattern of the *AhCIPKs* under submergence and *Ralstonia solanacearum* infection

Plants suffer from a wide variety of environmental stressors under natural conditions. To determine the abiotic and biotic stress responses, we detected the expression of *AhCIPK* genes responding to submergence and *R. solanacearum* infection. The results showed that except the two unexpressed members *AhCIPK6* and *20*, almost all other *AhCIPKs* respond to submergence stress (Figure 8A, Supplementary Table S4). *AhCIPK3*, *5*, *7*, 8, *9*, *17*, *19*, *20*, *21*, and *22* were upregulated rapidly after 6 h of the submergence treatment, while *AhCIPK10*, *11*, *23*, and *23* reached their highest expression at 24 h (Figure 8). By contrast, the expression levels of *AhCIPK1*, *2*, *4*, *14*, *15*, *16*, and *26* were inhibited under the whole submergence treatment process. We found it interesting that *AhCIPK12*, *13*, and *25* were first repressed under the earlier submergence treatment stages and were induced at the later stages. In addition, seven *AhCIPKs* (*AhCIPK1*, *5*, *7*, *11*, *14*, *19*, and *20*) were activated after 6 h *R. solanacearum* infection (Figure 8B), while six *AhCIPK* members (*AhCIPK2*, *9*, *11*, *12*, *22*, and *25*) were obviously upregulated and

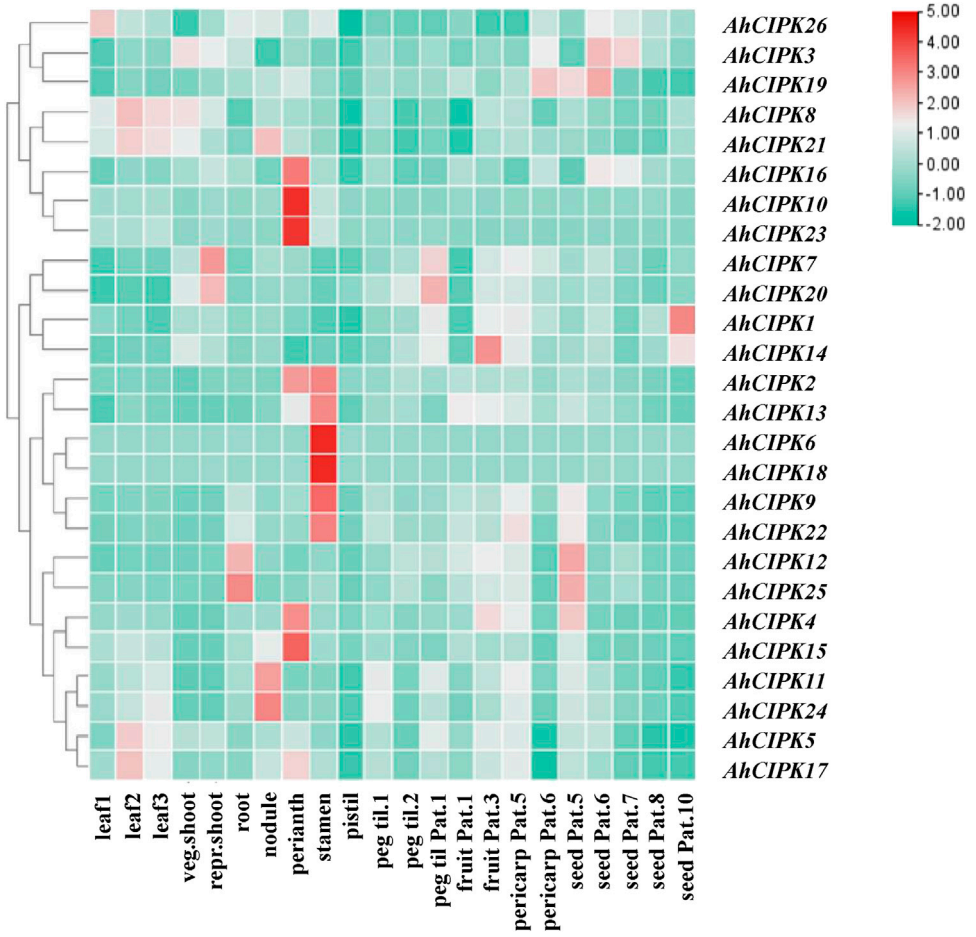


FIGURE 7
Expression profiles of *AhCIPK* genes. The heatmap of the *AhCIPK* gene expression levels was hierarchically clustered using TBtools with the data normalization method of Z-score standardization. The color scale bar from green to red represents low and high expressions, respectively. Abbreviations of the 22 tissues used in the expression profiles of *AhCIPK* genes were from Ren et al. (2022).

three (*AhCIPK3*, *16*, and *26*) were depressed after 48 h of *R. solanacearum* infection (Figure 8B, Supplementary Table S4); In total, *AhCIPK* genes might have functioned differentially in both abiotic and biotic environmental stress regulation.

Candidate gene association of peanut *CIPK*s polymorphisms with 104 traits

To further uncover the roles of *CIPK* genes in peanut development and stress response, we performed candidate gene association analysis using 22 single-nucleotide polymorphisms in *CIPKs* from transcriptome data of 146 peanut lines and 104 phenotypes related to peanut development and stress response collected in five environments. The results indicated that one polymorphic site [A09_903480$^{(G/K/T)}$] was significantly associated with pod length (PL), seed length (SL), hundred seed weight (HSW), and shoot root ratio (SR) traits (Figure 9A, Supplementary Tables S5–S7). Site B09_903480 mainly formed three haplotypes [B09_903480$^{(G/K/T)}$] (Figure 9B) in the population and was located in the predicted exon region of *AiCIPK10* (Figure 9C). Results showed that PL, SL, HSW, and SR in haplotype G were significantly higher than those in haplotype T (Figure 9D).

Discussion

Peanut *CIPK*s did not expand with genome duplication

CIPK genes are widely distributed widely in the biological world; however, their number varies greatly among different species. One, two, and seven *CIPK* genes were found in green algae, *Chlorella*, and *Physcomitrella patens*, respectively (Cheong et al., 2007; Weinl and

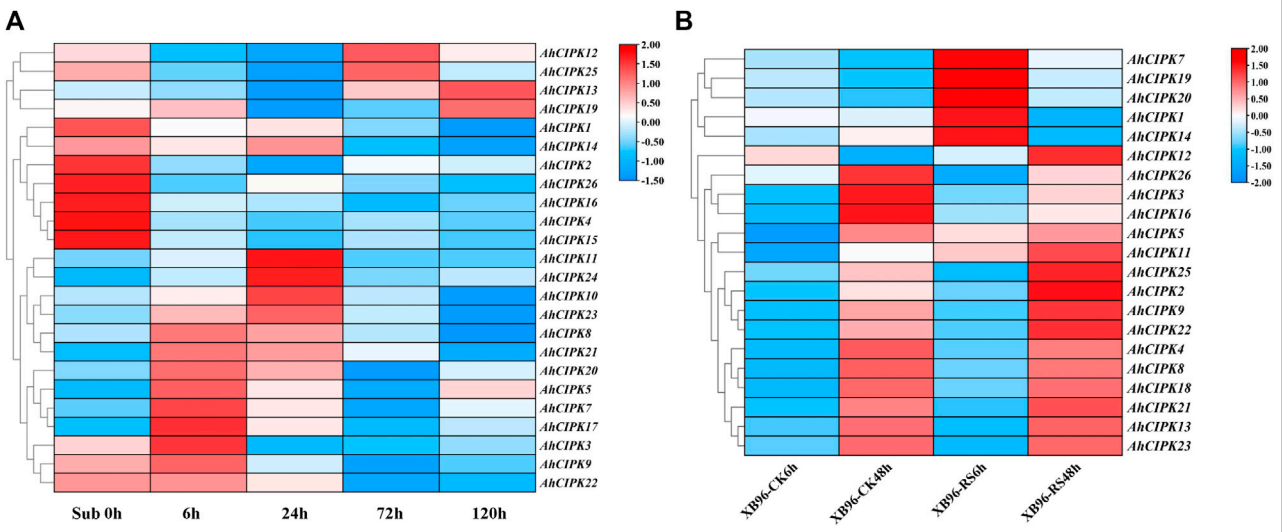


FIGURE 8
Relative expression of *AhCIPKs* under submergence and *Ralstonia solanacearum* infection. **(A)** expression characteristics of *AhCIPKs* in response to submergence at five time points (00, 06, 24, 72, and 120 h). **(B)** expression of *AhCIPKs* in response to *R. solanacearum* infection at two time points (6 and 48 h). The heatmap of the AhCIPK gene expression levels was hierarchically clustered using TBtools with the data normalization method of Z-score standardization. Legend at the right of the heatmap: red means upregulation and blue shows downregulated expression.

Kudla, 2009). According to previous studies, 25, 30, 27, 43, and 79 CIPK genes were identified in *A. thaliana* (125–155 Mb), rice (389 Mb), poplar (416 Mb), corn (2,400 Mb), and wheat (14,500 Mb) (Kolukisaoglu et al., 2004; Xiang et al., 2007; Chen et al., 2011; Sun et al., 2015; Zhu et al., 2021). Among *Solanaceae* plants, there were 21 and 22 *CIPK* members in tomato (Liu et al., 2017; Wang and Liu, 2018). Woody plants had 27 members (Zhang et al., 2008). This study uncovered 54 *CIPK* genes in three peanut genomes, among which 26 were from cultivated peanut (2,540 Mb). Our results support the hypothesis that the number of *CIPK* members in monocotyledonous plants is more than that in dicotyledonous plants. However, the number of peanut *CIPK* members did not expand with peanut genome size expansion.

The phylogenetic analysis revealed that *CIPK* members in *Arabidopsis*, rice, soybean, grape, and peanut could be coincidentally clustered into two distinct groups with different numbers of introns. This cluster pattern is the same as that in the previously reported *CIPK*-phylogenetic trees of *Arabidopsis* and rice (Kolukisaoglu et al., 2004). Our results further supported the hypothesis that the ancestor of *CIPKs* evolutionarily formed in the plant genome prior to the separation of the lineages of monocotyledons and dicotyledons (Kolukisaoglu et al., 2004; Yu et al., 2007).

Peanut *CIPK* genes functioned in stress response

The prediction of *cis*-acting elements can provide important clues for the study of gene expression regulation (Zhu et al., 2021). *Cis*-acting elements of biotic and abiotic stresses presented ubiquitously in the promoter region of peanut *CIPKs*, indicating these impact factors may interact to act on the *CIPK* regulatory mechanism (Figure 6). Compared to those in the *AhCIPKs*, the defense, stress, low-temperature response element TC-rich repeats and LTR were distributed concentratedly in several *AdCIPKs* or *AiCIPKs* in the two diploid progenitors; for example, four and three LTRs were identified in the promoter region of *AdCIPK4* and *AiCIPK11*, and four TC-rich repeats were found in the promoter region of *AdCIPK4*, indicating that the regulatory mechanism might be different between the cultivated peanut and the two diploid progenitors. Further the stress-induced expression data also elucidated the potential functions of peanut *CIPKs* in stress.

Among the orthologs between peanut and *Arabidopsis* (Table 2), the functions of the corresponding ortholog genes in *Arabidopsis* have been determined, and they functioned in influencing plant stress response (Table 2). *AdCIPK3*, *AiCIPK4*, *AhCIPK3*, and *AhCIPK16* were identified as the orthologous genes of the famous salt tolerance gene *AtCIPK24/AtSOS2* in *A. thaliana* (Halfter et al., 2000; Ishitani et al., 2000; Liu et al., 2000; Guo et al., 2001), and *AhCIPK8*, *AhCIPK21*, *AdCIPK9*, and *AiCIPK10* were found as the orthologs of *AtCIPK3* relating to the cold signal transduction (Kim et al., 2003; Pandey et al., 2008; Sanyal et al., 2017). Many other orthologs were also uncovered between peanut and the *Arabidopsis* abscisic acid-dependent and *Arabidopsis* abscisic acid-independent stress response gene *AtCIPK1* (D'Angelo et al., 2006), low-potassium tolerance

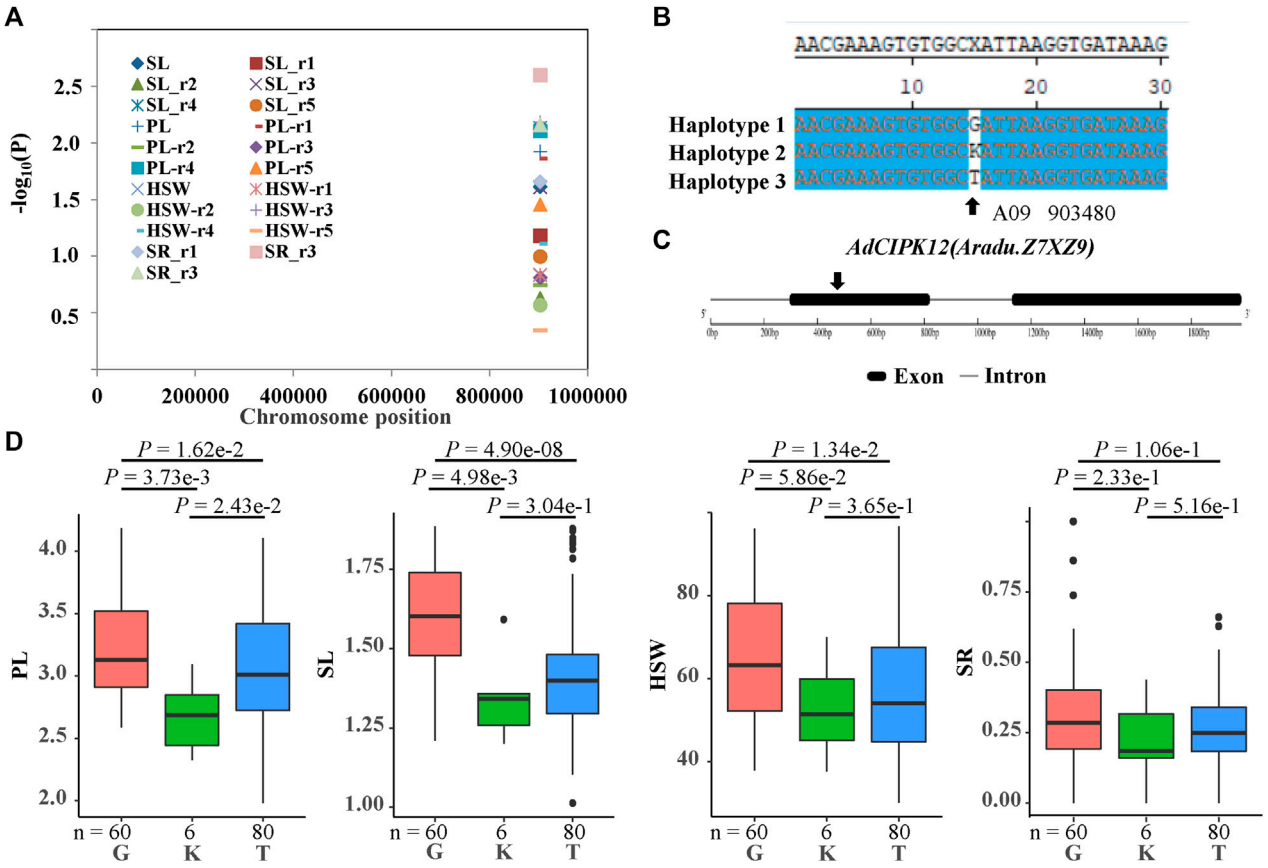


FIGURE 9
Association results and the phenotypes of the polymorphic sites of the peanut *CIPKs* associated with hundred seed weight (HSW)/pod length (PL)/ seed length (SL)/shoot root ratio (SR) variation. **(A)** association results between HSW/PL/SL/SR and the polymorphisms in peanut *CIPKs*. **(B)** sequences of the site significantly associated with HSW/PL/SL/SR variation. **(C)** gene structures of *AdCIPK12*(*Aradu.Z7XZ9*). **(D)** phenotypic comparison of haplotypes of the associated site with HSW/PL/SL/SR in five environments of the population.

gene *AtCIPK9* (Pandey et al., 2007; Singh et al., 2018; Lara et al., 2020; Kanwar et al., 2022), and cadmium stress response gene *AtCIPK11* (Liu and Guo, 2011; Zhou et al., 2015; Gu et al., 2021). Therefore, these peanut *CIPK* orthologous genes may also play multiple roles in peanut stress response, especially in salt response.

Peanut *CIPK* genes play important roles in growth and development

Many important genes were selectively expressed in specific tissues during various physiological and developmental processes (Wan et al., 2014). Our results showed that the 26 *AhCIPK* genes had distinct tissue-specific expression patterns, and several *AhCIPKs* showed higher expression level in the leaf, reproductive shoot, root, nodule, pod, and seed. We found it interesting that nearly half of *AhCIPK* genes are highly expressed in reproductive organs (Figure 7), indicating *AhCIPKs* play important roles in multiple tissue growth and development, especially the reproductive organs.

The single-nucleotide polymorphic sites in *AdCIPK12* (corresponding to *AhCIPK11*) were significantly associated with PL, SL, HSW, and SR variation. The polymorphic site in *AdCIPK12* [A09_ 903,480 $^{(G/K/T)}$], located in the predicted first exon region of the gene, A09_903480$^{(G/K/T)}$, led no transition (synonymous mutation) in the peanut population. Recent studies proved that synonymous mutations also have dramatic effects on protein output (Gillen et al., 2021). These results indicated that the B09_903480 $^{(G/K/T)}$ sequence polymorphisms might be the actual functional sites. Further, *AhCIPK11* was mainly expressed in the nodule, peg, pericarp, and seed, especially in the middle pericarp and seed development stages, which provided additional evidence for its function in peanut pod development (Figure 6). Further investigation was needed to confirm the roles of *AdCIPK12* (*AhCIPK11*) in the pod and seed development of peanut.

Conclusions

We definitely identified 54 *CIPK* members in cultivated and wild peanut for the first time and determined their chromosomal locations, gene structures, evolution, and expression patterns under biotic and abiotic conditions. We also focused on one gene, *AiCIPK10*/*AhCIPK21*, which was involved in pod and seed development. Our results provide valuable information for understanding the functions of the peanut *CIPK* gene family in regulating yield, quality, and stress responses in peanut.

Data availability statements

The datasets generated and analyzed during the current study are available from the corresponding author on reasonable request.

Author contributions

WR and JZ carried out all the experiments and data analyses. WR and JH prepared the figures and tables. LW, JF, and JZ made modifications to the article. All authors read and approved the final manuscript.

Funding

This research was funded by grants from the Jiangxi Agriculture Research System (No. JCARS-18) and Scientific Research Foundation for Scholars of HZNU (2022QDL034). Funders did not participate in the design of the study, analysis of the results and writing of the manuscript, but provided financial support for the manuscript.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.939255/full#supplementary-material

References

Albrecht, V., Ritz, O., Linder, S., Harter, K., and Kudla, J. (2001). The NAF domain defines a novel protein-protein interaction module conserved in Ca^{2+}regulated kinases. *Embo J.* 2 (5), 1051–1063. doi:10.1093/emboj/20.5.1051

Chen, C., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y., et al. (2020). TBtools: An integrative toolkit developed for interactive analyses of big biological data. *Mol. Plant* 13 (8), 1194–1202. doi:10.1016/j.molp.2020.06.009

Chen, L., Wang, Q. Q., Zhou, L., Ren, F., Li, D. D., and Li, X. B. (2013). *Arabidopsis* CBL-interacting protein kinase (CIPK6) is involved in plant response to salt/osmotic stress and ABA. *Mol. Biol. Rep.* 40 (8), 4759–4767. doi:10.1007/s11033-013-2572-9

Chen, X., Gu, Z., Xin, D., Hao, L., Liu, C., Huang, J., et al. (2011). Identification and characterization of putative *CIPK* genes in maize. *J. Genet. Genomics* 38 (2), 77–87. doi:10.1016/j.jcg.2011.01.005

Cheong, Y. H., Pandey, G. K., Grant, J. J., Batistic, O., Li, L., Kim, B. G., et al. (2007). Two calcineurin B-like calcium sensors, interacting with protein kinase CIPK23, regulate leaf transpiration and root potassium uptake in *Arabidopsis*. *Plant J.* 52 (2), 223–239. doi:10.1111/j.1365-313X.2007.03236.x

Clevenger, J., Chu, Y., Scheffler, B., and Ozias-Akins, P. (2016). A developmental transcriptome map for allotetraploid *Arachis hypogaea*. *Front. Plant Sci.* 7, 1446. doi:10.3389/fpls.2016.01446

D'Angelo, C., Weinl, S., Batistic, O., Pandey, G. K., Cheong, Y. H., Schültke, S., et al. (2006). Alternative complex formation of the Ca-regulated protein kinase CIPK1 controls abscisic acid-dependent and independent stress responses in *Arabidopsis*. *Plant J.* 48 (6), 857–872. doi:10.1111/j.1365-313X.2006.02921.x

Deng, X., Zhou, S., Hu, W., Feng, J., Zhang, F., Chen, L., et al. (2013). Ectopic expression of wheat *TaCIPK14*, encoding a calcineurin B-like protein-interacting protein kinase, confers salinity and cold tolerance in tobacco. *Physiol. Plant.* 149 (3), 367–377. doi:10.1111/ppl.12046

Feng, Z. J., Cui, X. Y., Cui, X. Y., Chen, M., Yang, G. X., Ma, Y. Z., et al. (2015). The soybean GmDi19-5 interacts with GmLEA3.1 and increases sensitivity of transgenic plants to abiotic stresses. *Front. Plant Sci.* 6, 179–186. doi:10.3389/fpls.2015.00179

Feng, Z. J., Xu, S. C., Liu, N., Zhang, G. W., Hu, Q. Z., and Gong, Y. M. (2017). Expression character of vegetable soybean CIPKs in response to abiotic stresses and hormones. *J. Plant Genet. Resour.* 18 (6), 1168–1178. doi:10.13430/j.cnki.jpgr.2017.06.019

Gasteiger, E., Hoogland, C., Gattiker, A., Duvaud, S., Wilkins, M. R., Appel, R. D., et al. (2005). "Protein identifification and analysis tools on the ExPASy server," in *The Proteomics protocols handbook* (New Jersey: Humana Press), 571–607. doi:10.1385/1-59259-890-0:571

Gillen, S. L., Waldron, J. A., and Bushell, M. (2021). Codon optimality in cancer. *Oncogene* 40 (45), 6309–6320. doi:10.1038/s41388-021-02022-x

Gong, D., Gong, Z., Guo, Y., Chen, X., and Zhu, J. K. (2002). Biochemical and functional characterization of PKS11, a novel *Arabidopsis* protein kinase. *J. Biol. Chem.* 277 (31), 28340–28350. doi:10.1074/jbc.M107719200

Gu, S., Wang, X., Bai, J., Wei, T., Sun, M., Zhu, L., et al. (2021). The kinase CIPK11 functions as a positive regulator in cadmium stress response in *Arabidopsis*. *Gene* 772, 145372. doi:10.1016/j.gene.2020.145372

Guo, Y., Halfter, U., Ishitani, M., and Zhu, J. K. (2001). Molecular characterization of functional domains in the protein kinase SOS2 that is required for plant salt tolerance. *Plant Cell* 13 (6), 1383–1400. doi:10.1105/tpc.13.6.1383

Halfter, U., Ishitani, M., and Zhu, J. K. (2000). The *Arabidopsis* SOS2 protein kinase physically interacts with and is activated by the calcium-binding protein SOS3. *Proc. Natl. Acad. Sci. U. S. A.* 97 (7), 3735–3740. doi:10.1073/pnas.040577697

Held, K., Pascaud, F., Eckert, C., Gajdanowicz, P., Hashimoto, K., Corratgé-Faillie, C., et al. (2011). Calcium-dependent modulation and plasma membrane targeting of the AKT2 potassium channel by the CBL4/CIPK6 calcium sensor/protein kinase complex. *Cell Res.* 21 (7), 1116–1130. doi:10.1038/cr.2011.50

Hrabak, E. M., Chan, C. W. M., Gribskov, M., Harper, J. F., Choi, J. H., Halford, N., et al. (2003). The *Arabidopsis* CDPK-SnRK super family of protein kinases. *Plant Physiol.* 132 (2), 666–680. doi:10.1104/pp.102.011999

Hu, H. C., Wang, Y. Y., and Tsay, Y. F. (2009). AtCIPK8, a CBL-interacting protein kinase, regulates the low-affinity phase of the primary nitrate response. *Plant J.* 57 (2), 264–278. doi:10.1111/j.1365-313X.2008.03685.x

Hu, W., Xia, Z., Yan, Y., Ding, Z., Tie, W., Wang, L., et al. (2015). Genome-wide gene phylogeny of CIPK family in cassava and expression analysis of partial drought-induced genes. *Front. Plant Sci.* 30 (6), 914–929. doi:10.3389/fpls.2015.00914

Ishitani, M., Liu, J., Halfter, U., Kim, C. S., Shi, W., and Zhu, J. K. (2000). SOS3 function in plant salt tolerance requires N-myristoylation and calcium binding. *Plant Cell* 12 (9), 1667–1678. doi:10.1105/tpc.12.9.1667

Kanwar, P., Sanyal, S. K., Mahiwal, S., Ravi, B., Kaur, K., Fernandes, J. L., et al. (2022). CIPK9 targets VDAC3 and modulates oxidative stress responses in *Arabidopsis*. *Plant J.* 109 (1), 241–260. doi:10.1111/tpj.15572

Kim, K. N., Cheong, Y. H., Grant, J. J., Pandey, G. K., and Luan, S. (2003). CIPK3, a calcium sensor-associated protein kinase that regulates abscisic acid and cold signal transduction in *Arabidopsis*. *Plant Cell* 15 (2), 411–423. doi:10.1105/tpc.006858

Kim, K. N., Cheong, Y. H., Gupta, R., and Luan, S. (2000). Interaction specificity of *Arabidopsis* calcineurin B-like calcium sensors and their target kinases. *Plant Physiol.* 124 (4), 1844–1853. doi:10.1104/pp.124.4.1844

Kolukisaoglu, Ü., Weinl, S., Blazevic, D., Batistic, O., and Kudla, J. (2004). Calcium sensors and their interacting protein kinases: genomics of the Arabidopsis and rice CBL-CIPK signaling networks.:genomics of the *Arabidopsis* and rice CBL-CIPK signaling networks. *Plant Physiol.* 134 (1), 43–58. doi:10.1104/pp.103.033068

Lara, A., Ródenas, R., Andrés, Z., Martínez, V., Quintero, F. J., Nieves-Cordones, M., et al. (2020). *Arabidopsis* K+ transporter HAK5-mediated high-affinity root K+ uptake is regulated by protein kinases CIPK1 and CIPK9. *J. Exp. Bot.* 71 (16), 5053–5060. doi:10.1093/jxb/eraa212

Li, L., Kim, B. G., Cheong, Y. H., Pandey, G. K., and Luan, S. (2006). A Ca^{2+} signaling pathway regulates a K^+ channel for low-K response in *Arabidopsis*. *Proc. Natl. Acad. Sci. U. S. A.* 103 (33), 12625–12630. doi:10.1073/pnas.0605129103

Liu, J., and Guo, Y. (2011). The alkaline tolerance in *Arabidopsis* requires stabilizing microfilament partially through inactivation of PKS5 kinase. *J. Genet. Genomics* 38 (7), 307–313. doi:10.1016/j.jgg.2011.05.006

Liu, J., Ishitani, M., Halfter, U., Kim, C. S., and Zhu, J. K. (2000). The *Arabidopsis thaliana* SOS2 gene encodes a protein kinase that is required for salt tolerance. *Proc. Natl. Acad. Sci. U. S. A.* 97 (7), 3730–3734. doi:10.1073/pnas.060034197

Liu, S. Y., Huang, H. F., Wang, X. D., Gong, C., Song, H. H., Chen, X. L., et al. (2017). Identification and characterization of *CBL* and *CIPK* gene families in pepper genome analyse. *Mol. Plant Breed.* 15 (8), 2977–2985. doi:10.1186/s12864-019-6125-z

Lu, Q., Li, S. X., Chen, X. P., Zhou, G. Y., Hong, Y. B., Li, H. F., et al. (2017). Current situation, problems and suggestions of peanut breeding in southern China. *Chin. J. Oil Crop Sci.* 39 (4), 556–566. doi:10.7505/j.issn.1007-9084.2017.04.019

Mao, J., Manik, S. M., Shi, S., Chao, J., Jin, Y., Wang, Q., et al. (2016). Mechanisms and physiological roles of the CBL-CIPK networking system in *Arabidopsis thaliana*. *Genes* 7 (9), 62–72. doi:10.3390/genes7090062

Pandey, G. K., Cheong, Y. H., Kim, B. G., Grant, J. J., Li, L., and Luan, S. (2007). CIPK9: a calcium sensor-interacting protein kinase required for low-potassium tolerance in *Arabidopsis*. *Cell Res.* 17 (5), 411–421. doi:10.1038/cr.2007.39

Pandey, G. K., Grant, J. J., Cheong, Y. H., Kim, B. G., Li, le.G., and Luan, S. (2008). Calcineurin-B-like protein CBL9 interacts with target kinase CIPK3 in the regulation of ABA response in seed germination. *Mol. Plant* 1 (2), 238–248. doi:10.1093/mp/ssn003

Pandey, M. K., Pandey, A. K., Kumar, R., Nwosu, C. V., Guo, B., Wright, G. C., et al. (2020). Translational genomics for achieving higher genetic gains in groundnut. *Theor. Appl. Genet.* 133 (5), 1679–1702. doi:10.1007/s00122-020-03592-2

Ragel, P., Ródenas, R., García-Martín, E., Andrés, Z., Villalta, I., Nieves-Cordones, M., et al. (2015). The CBL-interacting protein kinase CIPK23 regulates HAK5-mediated high-affinity K+ uptake in *Arabidopsis* roots. *Plant Physiol.* 169 (4), 2863–2873. doi:10.1104/pp.15.01401

Ren, W., Zeng, Z., Wang, S., Zhang, J., Fang, J., Wan, L., et al. (2022). Global Survey, Expressions and Association Analysis of CBLL Genes in Peanut. *Front. genet.* 13, 821163. doi:10.3389/fgene.2022.821163

Rui, Y. (2015). *Calcium availability to runner -type peanut (*Arachis hypogaea *L.) in the southeastern United States*. Auburn: Graduate Faculty of Auburn University.

Sánchez-Barrena, M. J., Chaves-Sanjuan, A., Raddatz, N., Mendoza, I., Cortés, Á., Gago, F., et al. (2020). Recognition and activation of the plant AKT1 potassium channel by the kinase CIPK23. *Plant Physiol.* 182 (4), 2143–2153. doi:10.1104/pp.19.01084

Sanders, D., Pelloux, J., Brownlee, C., and Harper, J. F. (2002). Calcium at the crossroads of signaling. *Plant Cell* 14, S401–S417. doi:10.1105/tpc.002899

Sanyal, S. K., Kanwar, P., Samtani, H., Kaur, K., Jha, S. K., and Pandey, G. K. (2017). Alternative splicing of CIPK3 results in distinct target selection to propagate ABA signaling in *Arabidopsis*. *Front. Plant Sci.* 8, 1924. doi:10.3389/fpls.2017.01924

Sardar, A., Nandi, A. K., and Chattopadhyay, D. (2017). CBL-interacting protein kinase 6 negatively regulates immune response to *Pseudomonas syringae* in *Arabidopsis*. *J. Exp. Bot.* 68 (13), 3573–3584. doi:10.1093/jxb/erx170

Schlücking, K., Edel, K. H., Köster, P., Drerup, M. M., Eckert, C., Steinhorst, L., et al. (2013). A new β-estradiol-inducible vector set that facilitates easy construction and efficient expression of transgenes reveals CBL3-dependent cytoplasm to tonoplast translocation of CIPK5. *Mol. Plant* 6 (6), 1814–1829. doi:10.1093/mp/sst065

Singh, A., Yadav, A. K., Kaur, K., Sanyal, S. K., Jha, S. K., Fernandes, J. L., et al. (2018). A protein phosphatase 2C, AP2C1, interacts with and negatively regulates the function of CIPK9 under potassium-deficient conditions in *Arabidopsis*. *J. Exp. Bot.* 69 (16), 4003–4015. doi:10.1093/jxb/ery182

Steinhorst, L., Mähs, A., Ischebeck, T., Zhang, C., Zhang, X., Arendt, S., et al. (2015). Vacuolar CBL-CIPK12 Ca(2+)-sensor-kinase complexes are required for polarized pollen tube growth. *Curr. Biol.* 25 (11), 1475–1482. doi:10.1016/j.cub.2015.03.053

Sun, T., Wang, Y., Wang, M., Li, T., Zhou, Y., Wang, X., et al. (2015). Identification and comprehensive analysis of the *CBL* and *CIPK* gene family in Wheat (*Triticum aestivum* L). *BMC Plant Biol.* 15, 269. doi:10.1186/s12870-015-0657-4

Tagliani, A., Tran, A. N., Novi, G., Di, Mambro.R., Pesenti, M., Sacchi, G. A., et al. (2020). The calcineurin β-like interacting protein kinase CIPK25 regulates potassium homeostasis under low oxygen in Arabidopsis. *J. Exp. Bot.* 71 (9), 2678–2689. doi:10.1093/jxb/eraa004

Tian, Q., Zhang, X., Yang, A., Wang, T., and Zhang, W. H. (2016). CIPK23 is involved in iron acquisition of *Arabidopsis* by affecting ferric chelate reductase activity. *Plant Sci.* 246, 70–79. doi:10.1016/j.plantsci.2016.01.010

Tripathi, V., Parasuraman, B., Laxmi, A., and Chattopadhyay, D. (2009a). CIPK6, a CBL-interacting protein kinase is required for development and salt tolerance in plants. *Plant J.* 58 (5), 778–790. doi:10.1111/j.1365-313X.2009.03812.x

Tripathi, V., Syed, N., Laxmi, A., and Chattopadhyay, D. (2009b). Role of CIPK6 in root growth and auxin transport. *Plant Signal. Behav.* 4 (7), 663–665. doi:10.1111/j.1365-313-X.2009.03812.x

Tsou, P. L., Lee, S. Y., Allen, N. S., Winter-Sederoff, H., and Robertson, D. (2012). An ER-targeted calcium-binding peptide confers salt and drought tolerance mediated by CIPK6 in *Arabidopsis*. *Planta* 235 (3), 539–552. doi:10.1007/s00425-011-1522-9

Wan, L., Wu, Y., Huang, J., Dai, X., Lei, Y., Yan, L., et al. (2014). Identification of ERF genes in peanuts and functional analysis of AhERF008 and AhERF019 in abiotic stress response. *Funct. Integr. Genomics* 14 (3), 467–477. doi:10.1007/s10142-014-0381-4

Wang, A. X., and Liu, S. Y. (2018). Identification and bioinformatics analysis on *CIPK* gene family in tomato. *J. Northeast Agric. Univ.* 49 (2), 31–38. doi:10.3969/j.issn.1005-9369.2018.02.004

Wang, X. P., Chen, L. M., Liu, W. X., Shen, L. K., Wang, F. L., Zhou, Y., et al. (2016). AtKC1 and CIPK23 synergistically modulate AKT1-mediated low-potassium stress responses in *Arabidopsis*. *Plant Physiol.* 170 (4), 2264–2277. doi:10.1104/pp.15.01493

Wang, Y., and Wu, W. H. (2004). Molecular genetic mechanism of high efficient potassium up take in plants. *Chin. Bull. Bot.* 44 (1), 27–36. doi:10.3969/j.issn.1674-3466.2009.01.003

Wang, Y., Zhang, W. Z., Song, L. F., Zou, J. J., Su, Z., and Wu, W. H. (2008). Transcriptome analyses show changes in gene expression to accompany pollen germination and tube growth in *Arabidopsis*. *Plant Physiol.* 148 (3), 1201–1211. doi:10.1104/pp.108.126375

Weinl, S., and Kudla, J. (2009). The CBL-CIPK Ca^{2+}-decoding signaling network: function and perspectives. *New Phytol.* 184, 517–528. doi:10.1111/j.1469-8137.2009.02938.x

Xiang, Y., Huang, Y., and Xiong, L. (2007). Characterization of stress-responsive *CIPK* genes in rice for stress tolerance improvement. *Plant Physiol.* 144 (3), 1416–1428. doi:10.1104/pp.107.101295

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 (7), 1347–1360. doi:10.1016/j.cell.2006.06.011

Yu, Y. H., Xia, X. L., Yin, W. L., and Zhang, H. C. (2007). Comparative genomic analysis of *CIPK* gene family in *Arabidopsis* and Populus. *Plant Growth Regul.* 52 (2), 101–110. doi:10.1007/s10725-007-9165-3

Zeng, R., Chen, T., Wang, X., Cao, J., Li, X., Xu, X., et al. (2021). Physiological and expressional regulation on photosynthesis, starch and sucrose metabolism response to waterlogging stress in peanut. *Front. Plant Sci.* 12, 601771. doi:10.3389/fpls.2021.601771

Zhang, C., Chen, H., Cai, T., Deng, Y., Zhuang, R., Zhang, N., et al. (2017). Overexpression of a novel peanut NBS-LRR gene *AhRRS5* enhances disease resistance to *Ralstonia solanacearum* in tobacco. *Plant Biotechnol. J.* 15 (1), 39–55. doi:10.1111/pbi.12589

Zhang, H., Yin, W., and Xia, X. (2008). Calcineurin B-like family in populus: comparative genome analysis and expression pattern under cold, drought and salt stress treatment. *Plant Growth Regul.* 56 (2), 129–140. doi:10.1007/s10725-008-9293-4

Zhao, J. F., Sun, Z. F., Zheng, J., Guo, X., Dong, Z., Huai, J., et al. (2009). Cloning and characterization of a novel CBL-interacting protein kinase from maize. *Plant Mol. Biol.* 69 (6), 661–674. doi:10.1007/s11103-008-9445-y

Zhou, X., Hao, H., Zhang, Y., Bai, Y., Zhu, W., Qin, Y., et al. (2015). SOS2-LIKE PROTEIN KINASE5, an SNF1-RELATED PROTEIN KINASE3-type protein kinase, is important for abscisic acid responses in *Arabidopsis* through phosphorylation of ABSCISIC ACID-INSENSITIVE5. *Plant Physiol.* 168 (2), 659–676. doi:10.1104/pp.114.255455

Zhu, T., Wang, L., Rimbert, H., Rodriguez, J. C., Deal, K. R., De, O. R., et al. (2021). Optical maps refine the bread wheat *Triticum aestivum* cv. Chinese Spring genome assembly. *Plant J.* 107 (1), 303–314. doi:10.1111/tpj.15289

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 27 September 2022
DOI 10.3389/fgene.2022.997901

OPEN ACCESS

EDITED BY
Reyazul Rouf Mir,
Sher-e-Kashmir University of Agricultural Sciences and Technology, India

REVIEWED BY
Kalpana Singh,
Indian Agricultural Statistics Research Institute (ICAR), India
Dalia Z. Alomari,
Aarhus University, Denmark
Pradeep Kumar Bhati,
The International Maize and Wheat Improvement Center (CIMMYT), India
Shamseldeen Shehabeldin Eltaher,
University of Sadat City, Egypt

*CORRESPONDENCE
Saba Akram,
sabaakram3463@yahoo.com
Mian Abdur Rehman Arif,
m.a.rehman.arif@gmail.com

SPECIALTY SECTION
This article was submitted to Plant Genomics,
a section of the journal
Frontiers in Genetics

RECEIVED 19 July 2022
ACCEPTED 15 August 2022
PUBLISHED 27 September 2022

CITATION
Akram S, Ghaffar M, Wadood A, Shokat S, Hameed A, Waheed MQ and Arif MAR (2022), A GBS-based genome-wide association study reveals the genetic basis of salinity tolerance at the seedling stage in bread wheat (Triticum aestivum L.).
Front. Genet. 13:997901.
doi: 10.3389/fgene.2022.997901

A GBS-based genome-wide association study reveals the genetic basis of salinity tolerance at the seedling stage in bread wheat (*Triticum aestivum* L.)

Saba Akram*, Maria Ghaffar, Ayesha Wadood, Sajid Shokat, Amjad Hameed, Muhammad Qandeel Waheed and Mian Abdur Rehman Arif*

Nuclear Institute for Agriculture and Biology College, Pakistan Institute of Engineering and Applied Sciences (NIAB-C, PIEAS), Faisalabad, Pakistan

High salinity levels affect 20% of the cultivated area and 9%–34% of the irrigated agricultural land worldwide, ultimately leading to yield losses of crops. The current study evaluated seven salt tolerance-related traits at the seedling stage in a set of 138 pre-breeding lines (PBLs) and identified 63 highly significant marker-trait associations (MTAs) linked to salt tolerance. Different candidate genes were identified in *in silico* analysis, many of which were involved in various stress conditions in plants, including glycine-rich cell wall structural protein 1-like, metacaspase-1, glyceraldehyde-3-phosphate dehydrogenase GAPA1, and plastidial GAPA1. Some of these genes coded for structural protein and participated in cell wall structure, some were linked to programmed cell death, and others were reported to show abiotic stress response roles in wheat and other plants. In addition, using the Multi-Trait Genotype-Ideotype Distance Index (MGIDI) protocol, the best-performing lines under salt stress were identified. The SNPs identified in this study and the genotypes with favorable alleles provide an excellent source to impart salt tolerance in wheat.

KEYWORDS

GBS, abiotic stress, salt tolerance, association mapping, *Triticum aestivum* L., candidate genes, MGIDI, SNP

Introduction

Salt stress is caused by an abundance of sodium chloride (NaCl) from irrigation with brackish water and crop evaporation (Flowers and Flowers, 2005). A saline soil can be characterized as one with electric conductivity (EC) >4 dS m^{-1} at 25°C and 15% exchangeable sodium. As much as 20% of the cultivated area and 9%–34% of the irrigated agricultural land worldwide is affected by high salinity (Ghassemi et al., 1995), which ultimately leads to yield losses in crops (Jamil et al., 2011). Salinity

affects >20% of modern agriculture (Li et al., 2020), making it a significant hurdle for sustainable agriculture production (Shokat and Großkinsky, 2020). Salinity affects plant growth by affecting two basic components of mineral transportation; initially, plants experience osmotic stress, which results in ion deposition and ensuing toxicity (Flowers and Flowers, 2005; Verslues et al., 2006; Munns and Tester, 2008).

While the wheat plant (*Triticum aestivum* L.) has moderate tolerance to salinity (Saddiq et al., 2019), significant yield losses still occur due to soil salinization (Oyiga et al., 2016). At 6–8 dS m^{-1} (Royo and Abió, 2003), wheat plants undergo metabolic changes that alter their life functions (Chen et al., 2016; Acosta-Motos et al., 2017). Furthermore, due to excess Na^+ ion accumulation and low water potential of soil, hyperosmotic and hyper-ionic stress occur, in addition to primary stresses (Huang et al., 2010). These results manifest as decreased germination percentage, reduced growth, reduced yield, and changes in reproductive behavior (Hasanuzzaman et al., 2017). Among important field crops, salinity causes more damage to wheat throughout its life cycle starting from germination to maturity. Moreover, the flowering to grain filling period is the most affected stage under salt stress, which eventually leads to low grain production. In Pakistan, average wheat yield losses of up to 65% are reported in moderately saline soils (Shafi et al., 2010). Thus, it is necessary to convene all available tools of conventional and modern plant breeding tools as well as agronomic practices to hasten the development of salt-tolerant cultivars that can meet this increasing demand (Ashraf and Harris, 2004; El Sabagh et al., 2021).

Different genes are involved in the regulation of salt stress and play roles in improving plant performance under salt stress by regulating diverse mechanisms including the antioxidant defense system, Na^+ exclusion, maintenance of Na^+/K^+ homeostasis, transpiration efficiency, and cytosolic K^+ retention (Shabala and Munns, 2012; Rahman et al., 2016). Several strategies are used to increase yield under salt stress using conventional breeding tools (Hasanuzzaman et al., 2017). Several characteristics are used as indicators for wheat salinity tolerance (Colmer et al., 2006), including Na^+ exclusion (the ability to minimize Na^+ concentrations entering the xylem) (Munns, 2005). Salt tolerance is a quantitative trait for which numerous quantitative trait loci (QTL) have been reported in wheat at the germination, seedling, and maturity stages, as well as plant survival (Zhou et al., 2012). Previous studies suggested that the shoot Na^+ exclusion phenotype is associated with two genes: *Nax1* (present at 2A) and *Nax2* (present at 5A) in durum wheat, a close relative of bread wheat (James et al., 2011). Many QTLs linked to salt stress tolerance in wheat have been studied previously; e.g., 65 QTLs linked to 13 different seedling traits of wheat have been identified (Masoudi et al., 2015). Similarly, QTLs associated with *NAX* were mapped to chromosome 2A and were responsible for a 10% increase in wheat biomass under salinity stress. Two QTLs—*qRNAX.7 A.3* and *qSNAX.7A.3*—mapped on chromosome 7A both showed 11% and 16% increases in salinity tolerance in wheat (Hussain et al., 2017). GWAS studies have also been conducted to identify QTLs and candidate genes linked to salinity stress in barley. Xue et al. (2009) identified 30 QTLs linked to ten different traits under salinity stress that accounted for 3%–30% the total phenotypic variation. Additionally, *HvNaX3* was mapped on the 7H chromosome of barley and was linked to salt stress tolerance (Sayed et al., 2021).

Due to the involvement of only two parents in the development of bi-parental populations, the QTL mapping approach fails to disclose the entire genetic architecture for salt tolerance (Shi et al., 2017). Therefore, genome-wide association studies (GWAS) were devised in which natural populations of hundreds of individuals with low genetic relationships are used to map desirable markers, known as marker-trait associations (MTAs) (Liu et al., 2017; Akram et al., 2021). The principle of GWAS is the linkage disequilibrium (LD), which is used to identify the relationship associations between a large number of DNA variants and traits in several genotypes from natural populations (Hu et al., 2011; Mwando et al., 2020). GWAS is a useful tool to genetically dissect biotic (Jighly et al., 2015; Arif et al., 2022; Dababat et al., 2021) and abiotic stress tolerance (Arif et al., 2012; Turki et al., 2015), physiological (Arif et al., 2021) adaptability traits (Akram et al., 2021), and nutrient uptake (Sharma et al., 2022) in wheat. While a plethora of indices has been devised to identify the best genotypes under a given environment/stress, many challenges still exist (Bizari et al., 2017). Owing to the limitations posed by previous indices (Céron-Rojas and Crossa, 2018), a new index was recently proposed based on genotype-ideotype distance and factorial analysis, which focused on the selection of superior genotypes based on multiple traits (Olivoto and Nardino, 2021).

The results of these investigations and identifications may allow the improvement of salt stress tolerance in wheat cultivars. The present study investigated a set of 138 wheat genotypes for salt-stress tolerance at the seedling stage. GWAS was applied to determine the extent of variation in response to salt stress, to identify molecular markers linked to salt tolerance, to search for candidate genes favorable for salt tolerance, and to identify salt-tolerant genotypes.

Materials and methods

The study was conducted on a set of 138 diverse wheat pre-breeding lines (PBLs) developed at CIMMYT (Supplementary Table S1). These 138 lines were selected from a larger set of 312 PBLs previously reported (Akram et al., 2021). This germplasm was the product of a large project, the "SeeDs of Discovery", which was implemented at CIMMYT, Mexico (Singh et al., 2018) where each line was the product of two elite

(best/approved cultivars) and one exotic line (GenBank accession). The current investigation used seeds obtained from the 2018–2019 harvest.

Experimental design and measured morphological traits

This investigation followed a completely randomized design (CRD). Initially, seeds from each line were treated with 10% NaOCl for 5 minutes followed by three washes with distilled water. In a growth chamber (Sanyo-Gallenkamp, United Kingdom) with controlled temperature of 28 ± 2°C and a 10-h photoperiod, 25 seeds of each genotype were grown on Whatman no. 1 filter paper moistened with a salt solution (Zafar et al., 2015). The experiment was conducted at three NaCl concentrations: 0, 150, and 250 mM NaCl corresponding to the control (S0), treatment 1 (S1), and treatment 2 (S2) groups, respectively.

On the eighth day of germination tests, the performance of the seedlings was assessed by recording the following morphological characteristics (Table 1). The mean value of each trait in each treatment was used for association analysis.

DNA extraction and genotyping

The genotyping used the flag leaves at the booting stage of $TC1F_5$ plants. DNA was extracted using the cetrylmethylammonium bromide (CTAB) method and quantified on a Nano-Drop instrument (http://www.diversityarrays.com/dart-application-dartseq), as described by Akram et al. (2021). A total of 58,378 high-quality SNP markers were generated, which were condensed to 6,887 SNPs by applying various stringent criteria including call rate (quality of genotyping) and reproducibility (marker consistency over replicated assays). the chromosomes, orders, and genetic distances of the mapped SNPs were obtained from the 100K-marker DArT-seq consensus map (http://www.diversityarrays.com/sequence-maps).

Statistical analysis

SPSS 16.0 was used to generate the descriptive statistics. All other analyses, including three way ANOVA, phenotypic histograms (using the "ggplot2" package) (Wickham, 2016), circular Manhattan plot ("CMplot") (Yin et al., 2021), and correlation ("qgraph") (Epskamp et al., 2012) were performed in RStudio version1.0.153. To assess the association of the

TABLE 1 Parameters measured to assess salinity tolerance.

Traits	Abbreviations	Description/formulas	References
Total germination percentage	TG	Total germination calculated by using the formula $Germination\ percentage = \frac{Number\ of\ seed\ germinated}{Total\ number\ of\ seed} \times 100$	Rajabi Dehnavi et al. (2020)
Numbers of roots	RN	Root numbers from each seedling/petro plate in control, S1 and S2	—
Coleoptile length	CL	Coleoptile length was calculated with the help of scale from each seedling/petri plate under control and all three replications of both treatments	—
Shoot length	SL	Shoot length was calculated with the help of scale from each seedling/petri plate under control and all three replications of both treatments	—
Roots length	RL	Roots length was calculated with the help of scale from each seedling/petri plate under control as well as S1 and S2	—
Root to shoot length ratio	R/S	Root to shoot length ratio was calculated with formula mentioned below $R/S = \frac{RL}{SL}$	—
Seedling vigor index	SVI	Seedling vigor index was calculated by following formula: $seedling\ vigor = (Average\ root\ length + Average\ shoot\ length) \times germination\ \%$	Kandil et al. (2012)
Relative total germination	RTG	Relative total germination was calculated by following formula: $RTG = \frac{TG\ of\ stressed\ plants}{TG\ of\ controlled\ plants} \times 100$	Kandil et al. (2012)
Relative numbers of roots	RRN	Relative numbers of roots was recorded by following formula: $RRN = \frac{RN\ of\ stressed\ plants}{RN\ of\ controlled\ plants} \times 100$	Fernandez (1993)
Relative coleoptile length	RCL	Relative coleoptile length was estimated by following formula: $RCL = \frac{CL\ of\ stressed\ plants}{CL\ of\ controlled\ plants} \times 100$	Fernandez (1993)
Relative shoot length	RSL	Relative shoot length was estimated by following formula: $RSL = \frac{Shoot\ length\ of\ stressed\ plants}{Shoot\ length\ of\ controlled\ plants} \times 100$	(Berger et al., 2012; Takahashi et al., 2015)
Relative root length	RRL	Relative root length was estimated by following formula: $RRL = \frac{RL\ of\ stressed\ plants}{RL\ of\ controlled\ plants} \times 100$	Zafar et al. (2015)
Relative root to shoot ratio	RR/S	Relative root to shoot ratio was calculated by following formula: $RR/S = \frac{R/S\ of\ stressed\ plants}{R/S\ of\ controlled\ plants} \times 100$	Berger et al. (2012)
Relative seedling vigor index	RSVI	Relative seedling vigor was estimated by following formula: $RSVI = \frac{SVI\ of\ stressed\ plants}{SVI\ of\ controlled\ plants} \times 100$	Fernandez (1993)

genotypes with the expressed phenotypes, principal component analysis (PCA) was performed using the "factoextra" package in R version 4.1.3 to reduce the dimensionality of the data (Kassambara and Mundt, 2017).

Genetic analysis

We used STRUCTURE version 2.3.4 to analyze population structure (Pritchard et al., 2003) where the K values ranged from 1 to 9 according to Akram et al. (2021). The online Structure Harvester software was used (Earl, 2012) to obtain the result files from STRUCTURE. The bar charts of population structure were plotted using STRUCTURE PLOT (Ramasamy et al., 2014).

TASSEL V5.2.43 software was used to perform marker-trait associations using the mean data for each treatment (Bradbury et al., 2007). The current study employed an MLM model that used population structure (Q-matrix generated by the structure) and kinship (K-matrix generated by TASSEL v 5.0) matrix as covariates to avoid false positives. Markers with p-values $<10^{-3}$ were defined as significant, whereas markers p-values less than the reciprocal of the number of markers ($<1.45 \times 10^{-4}$) were defined as highly significant associations (after Bonferroni correction) (Holm, 1979; Arif and Börner, 2020; Akram et al., 2021).

Identification of candidate genes

Sixty-nine base-pair-length sequences for each highly significant marker including 48 bp flanking regions of SNP marker were subjected to BLAST (Basic Local Alignment Search Tool) using the NCBI (National Center for Biotechnology Information) database. The BLAST search was conducted using the genome assembly IWGSC RefSeq v2.1 (Zhu et al., 2021). Hits with 100% identity and e-values $< 10^{-4}$ were selected.

Selection of tolerant wheat genotypes

The "metan" (Olivoto and Lúcio, 2020) package in R was used to differentiate the lines according to the MGIDI, where each trait (rX_{ij}) was standardized initially. This was followed by factor analysis to characterize the ideotype matrices. In the final step, an MGIDI index was computed by measuring the Euclidean distance between genotypes and ideotype scores using the following equation:

$$MGIDI = \sum_{j=1}^{f} [(\gamma ij - \gamma j)2]^{0.5}$$

where γ_{ij} represents the score represents of ith genotype (i = 1, 2,. . . ,t) in the jth factor (j = 1,2,. . . , f) and t and f are the number of genotypes and factors, respectively. The score of the ideal genotype was represented by γ_j. The lower the MGIDI value of a genotype, the closer it is to the ideal genotype (Olivoto and Nardino, 2021). A ~10% selection intensity (SI) was set to select the genotypes. Based on the ideotype concept, the traits were rescaled by assigning 0–100 values for all traits, in which 0 corresponded to the least valuable trait, and 100 to the most valuable/desired trait, to define the ideotype. In the present investigation, all traits were assigned with increasing values defining the quantitative morphological traits, which directly or indirectly affected the wheat response towards salt stress.

Results

Phenotypic variations

Salt stress significantly affected all traits. For example, TG decreased from 91.10 to 68.84 and 53.22 in S1 and S2, respectively, while the RTG after S1 and S2 were 75.06 and 57.71, respectively. In contrast, RN increased from 4.52 to 5.12 in S1 and 2.12 in S2, while the corresponding RRN was 114 in both treatments (S1 and S2). The CL in S1 was higher (2.94) than those in S0 and S2 (2.76 and 2.34, respectively). The RCL in S1 was also higher (106.78) than that in S2 (85.62). Among all traits, the highest decrease was observed in SL, which decreased from 9.38 (S0) to 6.08 (S1) and 3.25 (S2). The values for RSL_S1 and RSL_S2 were 67.30 and 33.02, respectively. A decreasing trend was also observed in RL, from 9.35 (S0) to 6.08 (S1) and 3.25 (S2). In contrast, the RRL in S1 was 65.67 and 35.36 in S2. The mean R/S in the control group was higher (1.14) than those for the S1 (0.98) and S2 (1.11) treatments. The relative R/S was higher in S2 (112.64) compared to that in S1 (99.50). The highest SVI was observed in the control group, with a mean value of 1708, followed by S1 (877, a decrease of 48% from the control) and S2 (353, a 91% decrease from the control). The mean RSVI in S1 (50.77) decreased to 20.38 in S2 (Figure 1; Supplementary Table S2).

Genotypes (G), treatments (T), and GxT showed significant differences in TG, RN, CL, SL, RL, R/S, SVI, RCL, RSL, RRL, RR/S, and RSVI. Among G and T, RTG showed highly significant differences, whereas no significant differences were observed for G×T. Only genotypes differed significantly in RRN, as compared to treatment plus G×T (Figure 1; Supplementary Table S2).

Correlations

Most traits were positively correlated in the untreated control, although SL_S0 was negatively correlated with TG_S0 and R/S_S0 (Figure 2; Supplementary Table S4).

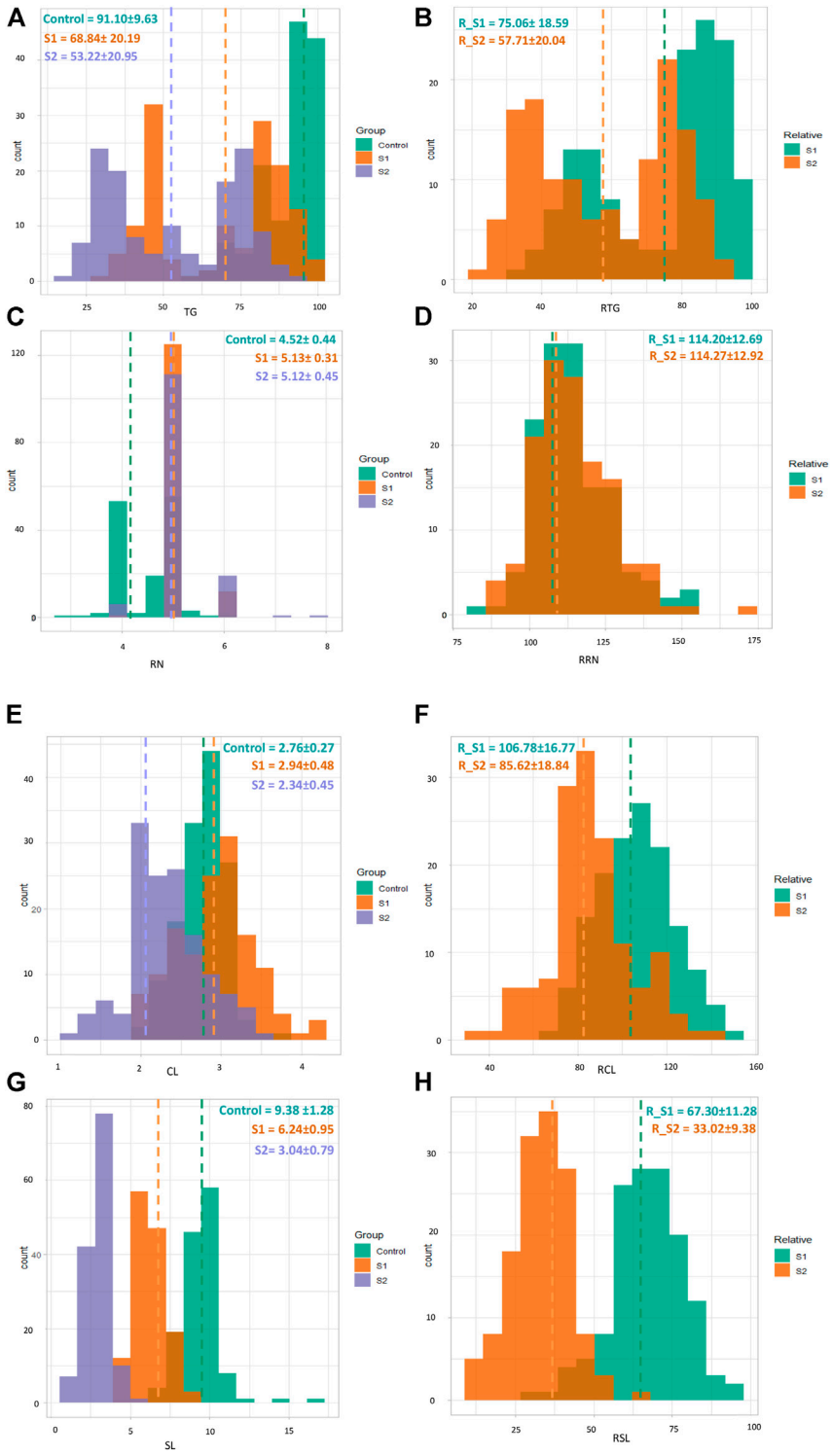


FIGURE 1
Overlaid histograms showing frequency distributions of TG **(A)**, RTG (relative) **(B)**, RN **(C)**, RRN (relative) **(D)**, CL **(E)** RCL (relative) **(F)**, SL **(G)**, RSL (relative) **(H)**, RL **(I)**, RRL (relative) **(J)**, R/S **(K)**, RR/S (relative) **(L)**, SVI **(M)**, and RSVI (relative) **(N)** across control (green), S1 (brown), and S2 (purple, for relative traits S1 (green) and S2 (brown). The vertical dotted lines indicate the mean values of each trait.

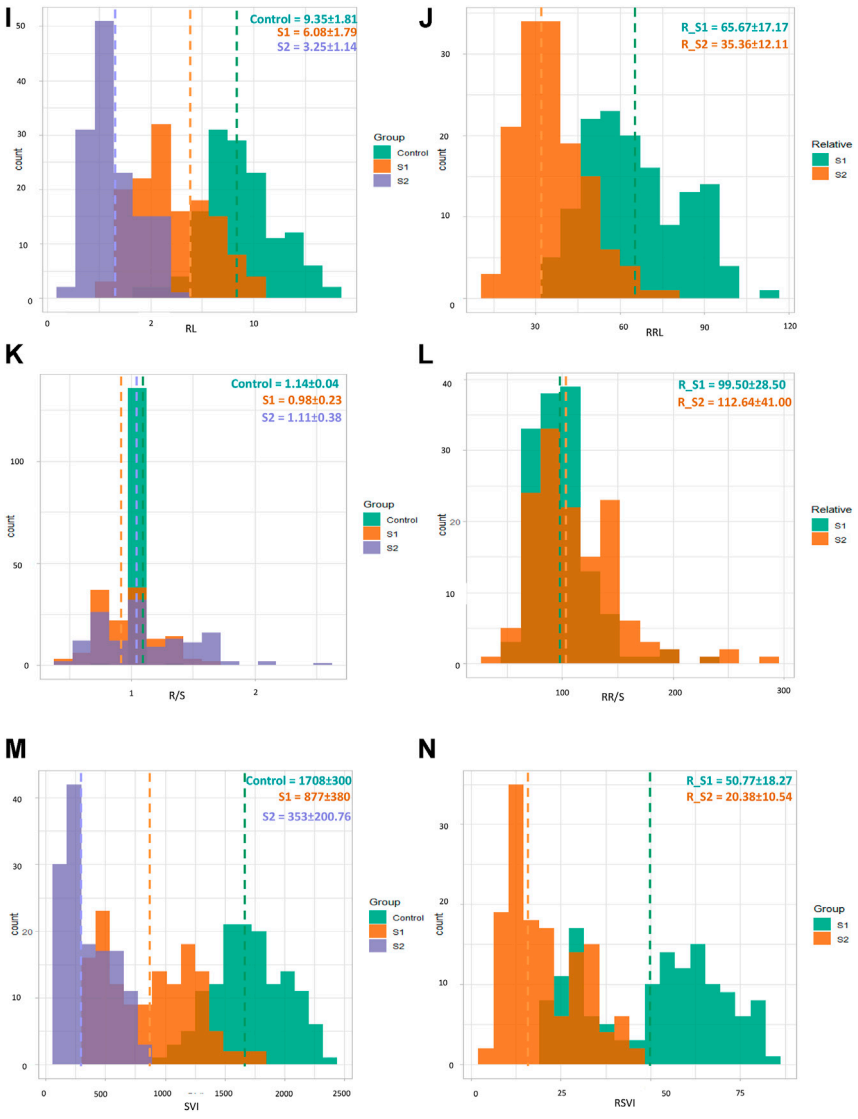


FIGURE 1 (Continued).

Likewise, R/S_S0 was also negatively correlated with SVI_S0. In contrast, all traits in S1 were positively correlated except for RN_S1 which did not show any correlation with any trait. Similar trends were observed in S2, except for RN_S2, which did not show any correlation with any other trait. In addition, SL_S2 was negatively correlated with R/S_S2. The relative traits were also positively correlated in most instances with their corresponding traits under salt stress.

Structure analysis

To correctly estimate the numbers of sub-groups in our germplasm, we plotted ΔK with a constant number of K sub-groups on the *x*-axis (Evanno et al., 2005), which showed a maximum ΔK value for K = 2 (Supplementary Figure S1). This value rose again at 5 and remained stable afterward. Therefore, we concluded that our germplasm carried five sub-populations (Supplementary Figure S2). The highest numbers of PBLs were observed in the second sub-group (Q2) followed by Q1, Q4, Q3, and Q5, with 27, 25, 23, and 18 PBLs, respectively. Supplementary Table S1 also provides information on the accessions regarding the pedigree and Q groups as identified from the STRUCTURE analysis, where the STRUCTURE results were consistent with the pedigrees of the collection. According to the pedigree, accessions with a last-crossed parent of Baj#1 were grouped in Q1 whereas most accessions with a last-crossed parent of KIRITATI were grouped in Q2.

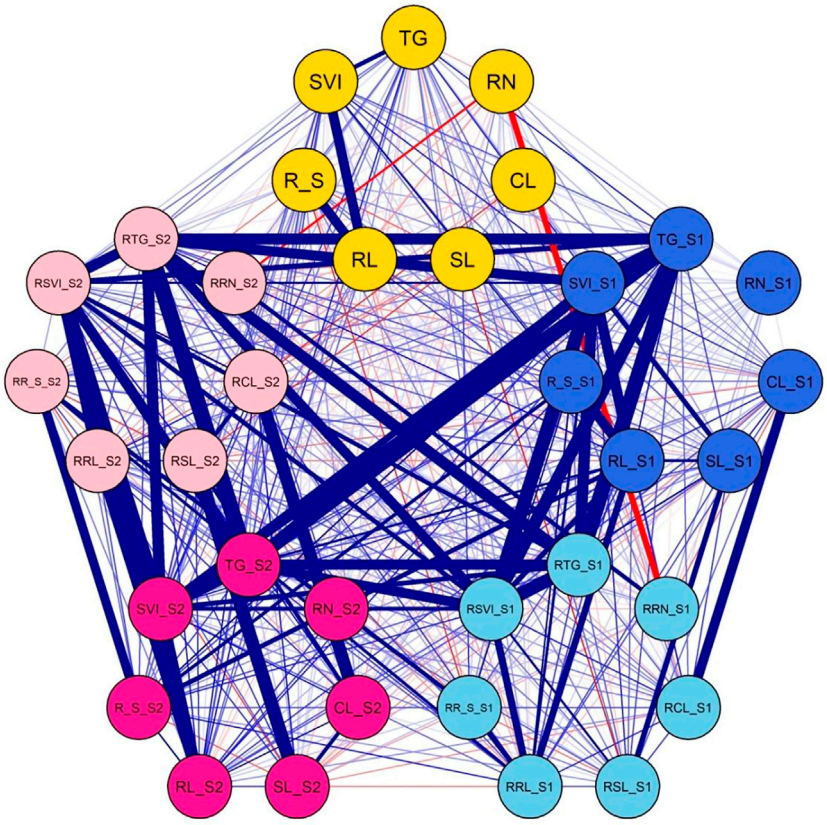


FIGURE 2
Correlations among matrixes of various traits under normal and salt-stress treatments. Yellow circle: traits under control. Blue circle: traits under S1. Sky blue circle: relative traits under S1. Bright pink circle: traits under S2. Light pink circle: relative traits under S2. Thick blue lines: highly significant correlations. Thin blue lines: significant correlations. Thick red lines: highly significant negative correlations. Thin red lines: significant negative correlations.

Accessions with last-crossed parents of Baj1, KACHU, and KIRITATI were grouped in Q3. The fourth subgroup included accessions in which the last-cross parent included VILLA JUARE2 F2009, while all other accessions with SUP152 as a parent in the last cross were grouped in Q5 (Supplementary Table S1).

Association mapping

A total of 195 MTAs showed an LOD of ≥3 ($p < 10^{-3}$) for the different traits observed in this study (Figure 3; Table 2). After Bonferroni correction, the number of associations decreased to 63 for all traits except for TG, CL, and SVI which did not show any association with any marker. Associations with $p < 1 \times 10^{-3}$ were considered significant, while those with $p < 1.452 \times 10^{-4}$ were considered highly significant.

The distributions of these MTAs in the wheat genome varied considerably. For example, chromosome 3B exhibited the highest (18 including nine highly significant) number, whereas chromosomes 3A (three including one highly significant), 6A (three including two highly significant) and 7B (three significant) exhibited the lowest numbers of MTAs (Table 2). Sixteen MTAs each were detected on chromosomes 2A (including six highly significant associations) and 1B (including four highly significant associations). Chromosome 4A carried 15 MTAs, including seven that were highly significant. There were 13 MTAs on each of chromosomes 5A (including five highly significantly) MTAs. Chromosome 2B also carried 13 MTAs, including two that were highly significant. Likewise, chromosomes 7A and 1A had 12 MTAs each, with six and one highly significant MTA, respectively. Similarly, 10 MTAs were detected on chromosomes 1D (including four highly significant associations) and 4B (including seven highly significant associations), respectively. Chromosome 2D showed eight significant MTAs, including one that was highly significant. Chromosomes 5B, 5D, 6D, and 7D each showed seven MTAs; each chromosome contained three highly significant MTAs except for 7D, which

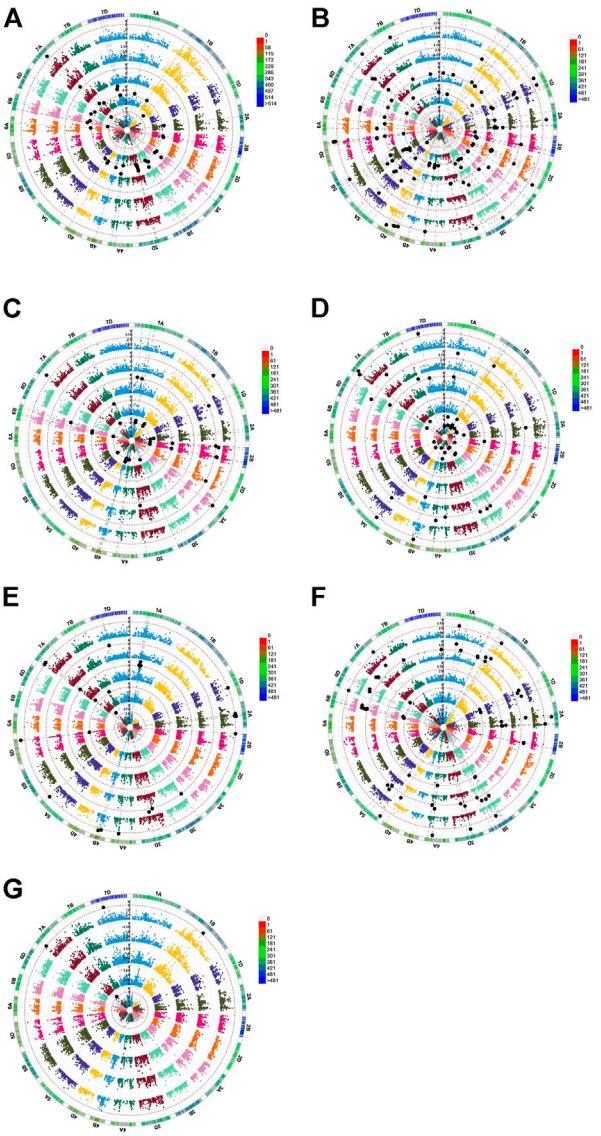


FIGURE 3
Genome-wide scan (GWAS analysis) of **(A)** TG, **(B)** RN, **(C)** CL, **(D)** SL, **(E)** RL, **(F)** R/S, and **(G)** SVI. S0 (inner circle), S1 (first outer circle), S2 (second outer circle), R_S1 (third outer circle), and R_S2 (fourth outer circle) are circular Manhattan plots in which the chromosomes are plotted at the outermost circle. The thin dotted red line indicates significance at $p < 0.001$ (–log10 = 3 or more) beyond which an association is counted as a true association (highlighted black dots). The scale between chromosomes 7D and 1A indicates the LOD threshold. The colored boxes outside on the top right side indicate the SNP density across the genome where green to red indicates less to more dense.

showed no highly significant MTAs. In addition, six MTAs were detected on chromosomes 3D (including two highly significant MTAs), while 6B showed five (including one highly significant) MTAs. Finally, four (including three highly significant MTAs) MTAs were detected on chromosome 4D (Figure 4).

TABLE 2 Chromosome-wide distributions of markers associated with various traits at $p < 1 \times 10^{-3}$ (normal text) and $p < 1.452 \times 10^{-4}$ (bold text).

Trait	Marker	Chr	Pos	F	*p*	R^2
CL_S2	M5289	1A	74.16	8.053285	5.25E-04	0.1227
RL_S2	M10801	1A	139.53	7.804446	6.61E-04	0.1289
SL_S2	M10801	1A	139.53	7.364397	9.75E-04	0.1176
RL_S2	M3085	1A	152.73	7.385107	9.39E-04	0.1129
RL_S2	**M406**	**1A**	**155.92**	**9.66784**	**1.30E-04**	**0.1529**
CL_S2	M5640	1A	159.59	7.605521	8.12E-04	0.1267
RL_S2	M11222	1A	170.55	7.735749	7.38E-04	0.1368
RR/S_S1	M8532	1A	224.72	8.297159	4.23E-04	0.1220
SL_S0	M6548	1A	235.18	9.168947	2.08E-04	0.1463
RR/S_S1	M8919	1A	480.78	8.353943	4.32E-04	0.1443
SL_S0	M7237	1A	490.67	9.146991	2.02E-04	0.1530
RR/S_S1	M7237	1A	490.67	7.462517	8.85E-04	0.1241
RR/S_S1	**M1426**	**1B**	**37.65**	**10.8378**	**4.98E-05**	**0.1626**
RRN_S1	M1426	1B	37.65	7.634138	7.80E-04	0.1225
SL_S0	**M11993**	**1B**	**61.58**	**11.5367**	**2.88E-05**	**0.1849**
RR/S_S1	M11993	1B	61.58	7.675678	7.64E-04	0.1367
SL_S0	M9230	1B	93.83	7.759434	6.77E-04	0.1218
SL_S0	M9015	1B	95	9.261137	1.83E-04	0.1447
SL_S0	M6072	1B	95	9.183687	1.96E-04	0.1498
RN_S2	M3534	1B	104	9.287996	1.76E-04	0.1384
RN_S2	M541	1B	194.87	7.783304	6.56E-04	0.1185
RN_S2	M5680	1B	205	7.901704	5.91E-04	0.1177
RSVI_S2	M7489	1B	285.98	8.025234	5.46E-04	0.1258
SL_S0	**M11428**	**1B**	**384.13**	**9.90536**	**1.07E-04**	**0.1584**
TG_S0	M11428	1B	384.13	9.222794	1.93E-04	0.1503
RN_S2	M11428	1B	384.13	8.169008	4.82E-04	0.1248
RCL_S2	M1718	1B	442.09	8.215706	4.73E-04	0.1363
RN_S1	**M9978**	**1B**	**492.15**	**11.2766**	**3.40E-05**	**0.1705**
RR/S_S1	**M10566**	**1D**	**93.71**	**10.6262**	**6.08E-05**	**0.1651**
RRN_S2	**M10810**	**1D**	**22.56**	**9.80467**	**1.29E-04**	**0.1718**
RRN_S1	M4912	1D	83.22	7.693458	7.52E-04	0.1276
SL_S0	**M10295**	**1D**	**90.1**	**9.63159**	**1.33E-04**	**0.1552**
RRN_S2	M8081	1D	100.36	9.542085	1.47E-04	0.1615
RCL_S1	M3676	1D	103.92	7.543196	8.35E-04	0.0709
RN_S2	M6138	1D	136.12	7.98738	5.52E-04	0.1191
RRL_S2	M10955	1D	151.27	7.340744	9.80E-04	0.1220
RN_S2	M10317	1D	164.21	9.49205	1.46E-04	0.1413
SL_S0	**M8113**	**1D**	**167.7**	**9.67561**	**1.36E-04**	**0.1614**
RN_S2	M7514	2A	76.95	8.021436	5.37E-04	0.1242
SL_S0	**M4431**	**2A**	**125.28**	**17.4073**	**3.07E-07**	**0.3133**
RR/S_S1	**M4431**	**2A**	**125.28**	**11.7778**	**2.48E-05**	**0.1837**
RSL_S1	M4431	2A	125.28	8.738744	3.12E-04	0.1495
RRN_S2	**M10796**	**2A**	**159.56**	**9.90685**	**1.09E-04**	**0.1538**
CL_S0	M1289	2A	168.8	7.575586	7.97E-04	0.1220
RN_S2	M5579	2A	214.32	7.451526	9.34E-04	0.1189
R/S_S2	**M765**	**2A**	**221.1**	**10.4296**	**7.22E-05**	**0.1892**
RN_S2	M765	2A	221.1	8.42717	3.97E-04	0.1543

(Continued on following page)

TABLE 2 (*Continued*) Chromosome-wide distributions of markers associated with various traits at $p < 1 \times 10^{-3}$ (normal text) and $p < 1.452 \times 10^{-4}$ (bold text).

Trait	Marker	Chr	Pos	F	*p*	R^2
RR/S_S2	M765	2A	221.1	7.845271	6.59E-04	0.1605
R/S_S2	M8894	2A	224.61	7.562287	8.18E-04	0.1106
RN_S2	**M9176**	**2A**	**231.79**	**16.8382**	**4.39E-07**	**0.3004**
RN_S2	**M662**	**2A**	**231.79**	**9.80564**	**1.22E-04**	**0.1573**
R/S_S2	M3109	2A	231.79	7.847723	6.22E-04	0.1085
RN_S2	M11432	2A	231.79	7.603874	7.75E-04	0.1158
R/S_S2	M3296	2A	231.79	7.386294	9.50E-04	0.1051
CL_S1	M3355	2B	7.26	7.579461	7.92E-04	0.0887
RCL_S1	M3355	2B	7.26	7.31986	9.98E-04	0.0663
RCL_S1	M2801	2B	24.73	8.93737	2.52E-04	0.0885
SL_S1	M592	2B	51.74	7.474094	8.70E-04	0.1109
RRN_S1	M2717	2B	147.53	8.397012	3.85E-04	0.1290
RR/S_S1	**M2717**	**2B**	**147.53**	**10.6548**	**5.47E-05**	**0.1545**
RCL_S1	M1480	2B	152.1	8.473657	3.74E-04	0.0868
RCL_S1	M8928	2B	155.58	8.837134	2.96E-04	0.0956
CL_S1	M8928	2B	155.58	7.526448	9.10E-04	0.1045
CL_S0	M6019	2B	156.75	7.508633	9.00E-04	0.1379
RR/S_S1	**M5489**	**2B**	**157.36**	**10.6002**	**6.12E-05**	**0.1611**
RN_S1	M5756	2B	162.47	8.266946	4.54E-04	0.1369
RN_S1	M1181	2B	198.95	9.488483	1.53E-04	0.1433
RN_S2	M7686	2D	99.87	8.163531	4.88E-04	0.1468
RN_S2	M8295	2D	113.22	9.128514	2.10E-04	0.1622
RN_S2	M10939	2D	134.09	7.609537	7.82E-04	0.1161
RN_S1	M6213	2D	153.45	8.50066	3.75E-04	0.1290
RRN_S1	M6643	2D	169.13	7.692638	7.19E-04	0.1132
SL_S0	M10941	2D	214.69	7.76266	6.78E-04	0.1241
RCL_S1	M754	2D	238.26	7.388606	9.42E-04	0.0697
RRN_S2	**M627**	**2D**	**289.68**	**10.8054**	**5.45E-05**	**0.1759**
RN_S2	**M1930**	**3A**	**75.97**	**15.3941**	**1.12E-06**	**0.2296**
RCL_S2	M1648	3A	88.71	7.598868	7.78E-04	0.1194
RN_S2	M11728	3A	229.22	8.641905	3.14E-04	0.1348
RN_S1	M4502	3B	58.82	8.172732	4.95E-04	0.1240
RRN_S2	**M5307**	**3B**	**68.2**	**11.2345**	**3.32E-05**	**0.1704**
SL_S0	**M7873**	**3B**	**103.59**	**17.5398**	**2.78E-07**	**0.3375**
RR/S_S1	**M7873**	**3B**	**103.59**	**11.097**	**4.32E-05**	**0.1783**
RSL_S1	M7873	3B	103.59	9.205768	2.10E-04	0.1616
R/S_S2	M1347	3B	116.53	7.685308	7.23E-04	0.1073
R/S_S2	M5248	3B	129.72	7.976771	5.93E-04	0.1179
RN_S1	M9174	3B	132.6	7.573937	8.01E-04	0.1143
RN_S1	M787	3B	159.94	7.836743	6.30E-04	0.1181
SL_S0	**M8515**	**3B**	**162.58**	**17.8368**	**1.55E-07**	**0.2757**
RR/S_S1	**M8515**	**3B**	**162.58**	**12.057**	**1.64E-05**	**0.1725**
RSL_S1	**M8515**	**3B**	**162.58**	**9.51185**	**1.43E-04**	**0.1382**
RR/S_S2	M8515	3B	162.58	8.461946	3.59E-04	0.1171
RR/S_S1	**M9138**	**3B**	**222.96**	**11.437**	**2.95E-05**	**0.1710**
RN_S2	**M11925**	**3B**	**227.41**	**16.3361**	**5.31E-07**	**0.2437**

(Continued in next column)

TABLE 2 (*Continued*) Chromosome-wide distributions of markers associated with various traits at $p < 1 \times 10^{-3}$ (normal text) and $p < 1.452 \times 10^{-4}$ (bold text).

Trait	Marker	Chr	Pos	F	*p*	R^2
CL_S0	M992	3B	294.83	8.731384	2.96E-04	0.1380
RN_S2	**M2025**	**3B**	**297.56**	**11.0286**	**4.09E-05**	**0.1748**
RN_S1	M4046	3B	297.56	9.507191	1.55E-04	0.1441
RRN_S2	M5581	3D	14.7	9.166748	2.15E-04	0.1532
RR/S_S1	**M1987**	**3D**	**107.75**	**10.5149**	**6.16E-05**	**0.1534**
RRL_S1	M7724	3D	108.86	7.690475	7.33E-04	0.1013
RR/S_S2	**M1019**	**3D**	**116.66**	**10.2482**	**8.18E-05**	**0.1494**
RN_S2	M2318	3D	184.67	7.958473	5.68E-04	0.1375
CL_S2	M4388	3D	265.96	8.999231	2.25E-04	0.1322
SL_S0	**M1398**	**4A**	**49.56**	**12.2238**	**1.43E-05**	**0.1889**
RSL_S1	M1398	4A	49.56	7.399761	9.21E-04	0.1075
RN_S1	M10965	4A	106.71	8.44769	4.01E-04	0.1404
SL_S0	**M11711**	**4A**	**180.05**	**20.4829**	**3.24E-08**	**0.3835**
RR/S_S1	**M11711**	**4A**	**180.05**	**10.8836**	**5.15E-05**	**0.1755**
RSL_S1	**M11711**	**4A**	**180.05**	**10.1833**	**9.21E-05**	**0.1798**
RN_S1	**M5589**	**4A**	**215.47**	**10.3647**	**7.31E-05**	**0.1705**
RN_S1	**M10678**	**4A**	**215.47**	**10.1663**	**8.65E-05**	**0.1679**
RN_S1	**M3948**	**4A**	**215.47**	**10.0451**	**1.02E-04**	**0.1527**
RN_S1	M1034	4A	215.47	8.555139	3.47E-04	0.1294
RN_S1	M2701	4A	215.47	7.851619	6.27E-04	0.1329
CL_S0	M8745	4A	215.47	7.818745	6.58E-04	0.1267
RR/S_S1	M10341	4A	219.65	7.883817	6.65E-04	0.1269
RR/S_S1	M2690	4A	221.42	8.133403	5.28E-04	0.1297
RRN_S2	M5221	4A	221.42	7.694232	7.42E-04	0.1237
CL_S0	M10528	4B	70.22	7.35443	9.94E-04	0.1170
RN_S1	M1225	4B	76.73	7.907482	5.99E-04	0.1193
RN_S1	M7413	4B	76.73	7.602701	8.52E-04	0.1753
CL_S0	M9224	4B	78.23	8.734255	3.03E-04	0.1489
RN_S1	M10770	4B	78.61	9.461426	1.62E-04	0.1567
CL_S0	M3043	4B	85.28	9.325857	1.71E-04	0.1503
RSL_S2	M10038	4B	90.17	8.880234	2.63E-04	0.1503
RRL_S2	M10038	4B	90.17	8.608923	3.33E-04	0.1499
SL_S0	M10038	4B	90.17	7.700859	7.35E-04	0.1255
SL_S2	M8833	4B	108.27	8.390578	4.21E-04	0.1387
RRN_S2	M4320	4D	1.14	9.312591	1.78E-04	0.1505
RRN_S2	**M4103**	**4D**	**34.78**	**9.93443**	**1.05E-04**	**0.1556**
RRN_S2	**M3343**	**4D**	**50.31**	**10.0131**	**9.62E-05**	**0.1525**
RR/S_S1	**M1974**	**4D**	**157.4**	**11.1122**	**4.19E-05**	**0.1728**
RN_S1	M11094	5A	85.73	9.569622	1.48E-04	0.1584
RSL_S2	M11935	5A	97	7.910871	6.14E-04	0.1332
RR/S_S1	**M8885**	**5A**	**113.15**	**11.209**	**3.87E-05**	**0.1720**
SL_S0	**M8885**	**5A**	**113.15**	**10.5227**	**6.84E-05**	**0.1954**
RSL_S1	M8885	5A	113.15	7.908211	6.32E-04	0.1305
RR/S_S1	M5275	5A	113.33	7.593138	8.15E-04	0.1151
RRN_S1	**M11486**	**5A**	**161.09**	**12.2524**	**1.43E-05**	**0.1818**
RN_S0	M10895	5A	161.09	7.412972	9.41E-04	0.1337

(Continued on following page)

TABLE 2 (*Continued*) Chromosome-wide distributions of markers associated with various traits at $p < 1 \times 10^{-3}$ (normal text) and $p < 1.452 \times 10^{-4}$ (bold text).

Trait	Marker	Chr	Pos	F	*p*	R^2
RR/S_S1	M56	5A	167.78	9.31266	1.78E-04	0.1370
RN_S2	M7106	5A	199.92	7.444092	9.25E-04	0.1140
RN_S2	**M4314**	**5A**	**209.83**	**13.8064**	**4.18E-06**	**0.2143**
RN_S2	**M3034**	**5A**	**229.39**	**16.1937**	**6.55E-07**	**0.2476**
CL_S0	M1688	5A	285.2	7.362705	9.64E-04	0.1131
RN_S2	**M4710**	**5B**	**56.97**	**16.4673**	**5.46E-07**	**0.2679**
RN_S2	**M2139**	**5B**	**56.97**	**9.73039**	**1.24E-04**	**0.1701**
RN_S1	M10042	5B	83.23	7.506657	8.63E-04	0.1135
RN_S1	**M1450**	**5B**	**125.96**	**10.6565**	**5.93E-05**	**0.1617**
RRN_S2	M5772	5B	145.97	7.384505	9.33E-04	0.1106
RN_S2	M8958	5B	169.96	8.360985	4.36E-04	0.1412
SL_S0	M5828	5B	256.53	8.074973	5.05E-04	0.1248
RN_S2	**M11102**	**5D**	**195.73**	**17.1946**	**2.72E-07**	**0.2852**
RN_S2	**M7775**	**5D**	**210.31**	**17.2919**	**2.72E-07**	**0.2875**
RN_S2	M8011	5D	223.19	7.825985	6.82E-04	0.1365
RN_S1	M7245	5D	228.94	8.147855	4.91E-04	0.1232
RL_S1	M4090	5D	228.94	7.589824	8.04E-04	0.1074
RRN_S2	**M337**	**5D**	**232.48**	**10.0153**	**9.55E-05**	**0.1504**
RRL_S2	M337	5D	232.48	9.507016	1.48E-04	0.1531
RN_S2	**M5347**	**6A**	**65.28**	**16.3246**	**5.36E-07**	**0.2435**
RN_S2	**M530**	**6A**	**73.17**	**16.3465**	**5.13E-07**	**0.2436**
RN_S2	M8998	6A	167.65	7.676664	7.57E-04	0.1279
CL_S1	M6385	6B	23.32	7.343408	9.77E-04	0.0848
RN_S2	M10569	6B	45.17	9.014767	2.31E-04	0.1536
CL_S0	M5384	6B	62.83	7.539925	8.44E-04	0.1175
RR/S_S1	**M4362**	**6B**	**69.05**	**10.4907**	**6.81E-05**	**0.1701**
CL_S2	M2278	6B	132.99	7.923197	5.93E-04	0.1185
RR/S_S2	M5749	6D	0	7.639943	7.99E-04	0.1119
RN_S2	M3987	6D	25.79	8.724641	3.24E-04	0.1565
RR/S_S1	**M11763**	**6D**	**90.8**	**10.7309**	**5.45E-05**	**0.1574**
SL_S0	**M1188**	**6D**	**121.22**	**17.9048**	**1.47E-07**	**0.2767**
RR/S_S1	**M1188**	**6D**	**121.22**	**11.4818**	**2.66E-05**	**0.1643**
RSL_S1	M1188	6D	121.22	9.290898	1.74E-04	0.1350
SL_S2	M10197	6D	181.69	8.790818	3.02E-04	0.1408
RL_S2	M10260	7A	38.12	7.674994	8.05E-04	0.1508
RR/S_S1	**M10555**	**7A**	**45.99**	**11.291**	**3.36E-05**	**0.1933**
RR/S_S1	**M10047**	**7A**	**45.99**	**10.9379**	**4.80E-05**	**0.1637**
RR/S_S1	**M9141**	**7A**	**45.99**	**10.6199**	**5.90E-05**	**0.1581**
RRN_S1	M9141	7A	45.99	7.803681	6.66E-04	0.1198
RL_S1	M8832	7A	55.34	9.453854	1.74E-04	0.1710
RTG_S2	M3854	7A	145.43	9.03094	2.34E-04	0.1356
TG_S2	M3854	7A	145.43	8.309385	4.36E-04	0.1223
RRN_S2	**M9550**	**7A**	**170.91**	**10.5026**	**6.22E-05**	**0.1581**
RL_S0	**M38**	**7A**	**214.7**	**10.7602**	**6.60E-05**	**0.2053**
SVI_S0	M38	7A	214.7	7.811965	7.53E-04	0.1426
RN_S1	**M9660**	**7A**	**276.04**	**11.7383**	**2.18E-05**	**0.1768**
RR/S_S1	**M9581**	**7B**	**114.48**	**7.671001**	**7.70E-04**	**0.1140**
RL_S1	**M9431**	**7B**	**117.72**	**7.838873**	**6.23E-04**	**0.1061**
SL_S2	**M7039**	**7B**	**234.97**	**9.060685**	**2.34E-04**	**0.1565**
RSL_S2	M3286	7D	113.89	8.872057	2.51E-04	0.1323
RN_S2	M7616	7D	205.95	8.168864	4.70E-04	0.1244
RSVI_S2	M3286	7D	113.89	7.843709	6.20E-04	0.1176
RRL_S1	M6322	7D	117.37	7.653348	7.58E-04	0.0988
SL_S0	M10389	7D	369.62	7.566639	8.88E-04	0.1370
RR/S_S1	M10389	7D	369.62	7.493323	9.46E-04	0.1260
SL_S2	M3286	7D	113.89	7.350426	9.62E-04	0.1068

(Continued in next column)

TABLE 2 (*Continued*) Chromosome-wide distributions of markers associated with various traits at $p < 1 \times 10^{-3}$ (normal text) and $p < 1.452 \times 10^{-4}$ (bold text).

Regarding traits, RN_S2 showed the highest number of MTAs (35 including thirteen that were highly significant), whereas RL_S0, RN_S0, RTG_S2, SL_S1, SVI_S0, TG_S2, and TG_S0 showed the lowest numbers of MTAs (one) for each trait. Twenty-eight MTAs (including eighteen that were highly significant) and 22 MTAs (including six that were highly significant) were detected for RR/S_S1 and RN_S1 respectively. We also detected 20 MTAs associated with SL_S0, thirteen with RRN_S2, and seven with RSL_S1 including eleven, eight, and two that were highly significantly MTAs respectively. Nine MTAs associated with SL_S0 (with no highly significant association), while six MTAs were detected for R/S_S2, RCL_S1 (including one significant MTA), and RRN_S1. Five MTAs were linked with RL_S2 (one significantly MTA) and SL_S0 (with no highly significant association). RR/S_S2 and CL_S2 were linked to four markers. RR/S_S2 was associated with one highly significant MTA. We also detected three MTAs for CL_S1, RL_S1, RRL_S2, and RSL_S2 and two MTAs for RCL_S2, RRL_S1, and RSVI_S2, with no highly significant association (Table 2).

Candidate genes

BLAST was performed for highly significant SNP markers. Hits with 100% identity and *e*-values < 10^{-4} were selected. Twelve candidate genes were identified through this analysis. These included putative disease resistance RPP13-like protein 1, disease resistance protein RGA2-like (involved in conferring disease resistance), glycine-rich cell wall structural protein 1-like (part of the cell wall that acts as a structural protein), two metacaspase-1-like proteins (play roles in programmed cell death), sphinganine

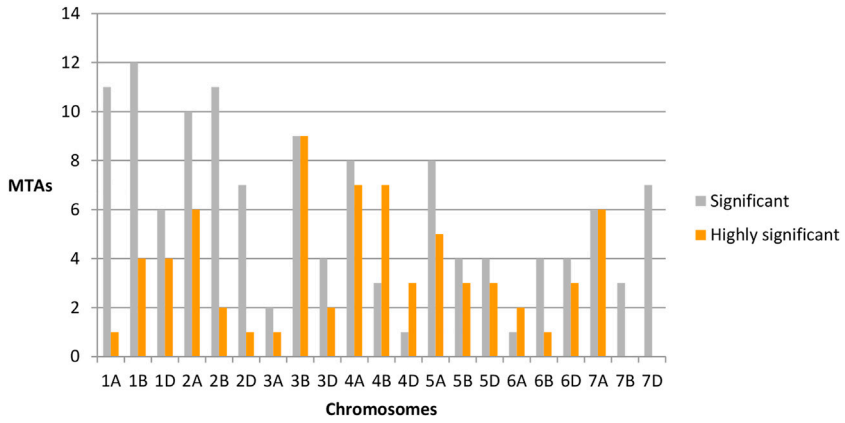


FIGURE 4
Chromosome-wide distribution of significant (gray) and highly significant (orange) marker-trait associations.

C4 monooxygenase 1-like (plays a role in sphingolipid biosynthesis), two 60S ribosomal protein L22-like proteins (ribosomal proteins), glyceraldehyde-3-phosphate dehydrogenase GAPA1, subtilisin-like protease SBT1.7, mRNA-decapping enzyme-like protein, and calmodulin-binding protein 60 D-like (involved in different stress responses, including biotic and abiotic responses). These candidate genes along with their physical locations and functions are shown in Table 3.

Multivariate analysis

The first two principal components explained significant variation (73.9% of the overall variation), with eigenvalues ≥1 (Figure 5), where PC1 accounted for 57.2% of the total variation and was positively correlated with all traits except RN. The second PC explained 16.7% of the total variation and was mainly influenced by R_S, TG, RL, and SVI. All traits were sorted into three groups (I–III) based on their distributions in the

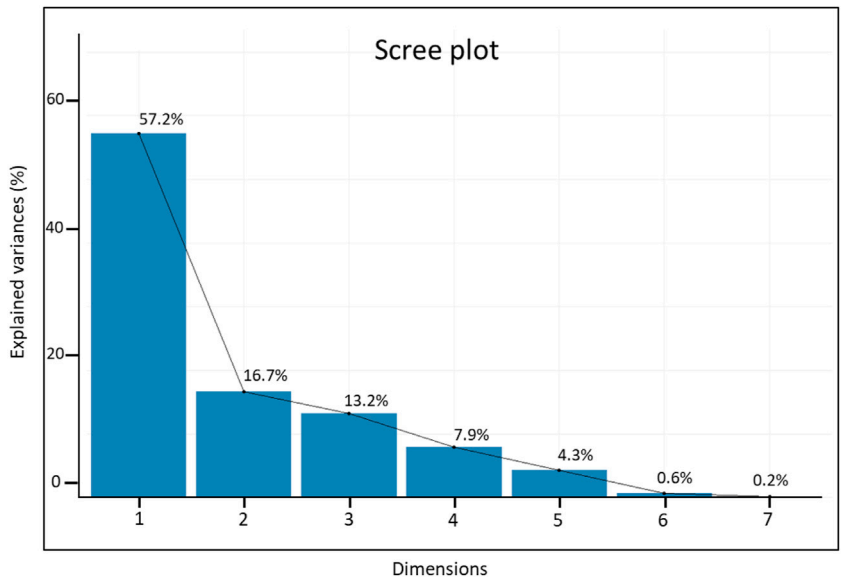


FIGURE 5
Scree plot showing the contributions of the first seven principal components.

TABLE 3 Identification of candidate genes involved in stress tolerance.

Sr. No.	Marker/trait/chr	Gene ID	Location	Length	Candidate genes
1	M11428/SL_S0/IB	LOC123145204	Chr1B: 648,595,158-648,602,293	4,385	*Triticum aestivum* putative disease resistance RPP13-like protein 1
2	M10810/RRN_S2/1D	LOC123179818	Chr1D: 6,607,035-6,641,306	3,071	*Triticum aestivum* disease resistance protein RGA2-like
3	M10796/RRN_S2/5D	LOC123121043	Chr5D: (507,153,271-507,153,732)	462	*Triticum aestivum* glycine-rich cell wall structural protein 1-like
4	M1930/RN_S2/3A	LOC123060422	Chr3A: 82,083,711-82,086,686	2,629	*Triticum aestivum* metacaspase-1-like (LOC123060422)
5	M5307/RRN_S2/3B	LOC123064730	Chr3B: 57,612,905-57,618,157	2,238	*Triticum aestivum* glyceraldehyde-3-phosphate dehydrogenase GAPA1, chloroplastic-like (LOC123064730)
6	M8515/RSL_S1.SL_S0.RR/S_S1/3B	LOC123072063	Chr3B: 734,929,010-734,930,868	1,478	*Triticum aestivum* metacaspase-1-like (LOC123072063)
7	M11711/RSL_S1.SL_S0.RR/S_S1/4A	LOC123082947	Chr4A: 659,640,114-659,640,987	738	*Triticum aestivum* sphinganine C4 monooxygenase 1-like (LOC123082947)
8	M3948/RN_S1/4A	LOC123088273	Chr4A: 709,761,323-709,762,376	1,054	*Triticum aestivum* 60S ribosomal protein L22-like (LOC123088273)
9	M10678/RN_S1/4A	LOC123088271	Chr4A: 709,405,204-709,406,325	1,122	*Triticum aestivum* 60S ribosomal protein L22-like (LOC123088271)
10	M4710/RN_S2/5A	LOC123124722	Chr5A: 301,269,821-301,272,263	2,443	*Triticum aestivum* subtilisin-like protease SBT1.7 (LOC123124722)
11	M7775/RN_S2/5D	LOC123123671	Chr5D: 542,615,453-542,621,186	2,031	*Triticum aestivum* mRNA-decapping enzyme-like protein (LOC123123671)
12	M11763/RR/S_S1/6D	LOC123144592	Chr6D: 260,335,200-260,341,291	2,089	*Triticum aestivum* calmodulin-binding protein 60 D-like (LOC123144592)

PCA biplot quadrants: group I included R_S, TG, RL, and SVI; group II consisted of SL and CL; and group III contained only RN (Figure 6). The bi-plot analysis also highlighted the correlation between the examined traits: The sharp angle between trait vectors indicated a positive correlation, while obtuse and right angles indicated negative and no correlations between the parameters, respectively.

Discussion

Wheat seedling development has three parts: germination, emergence, and early growth. All three stages are especially sensitive to salinity stress (Jamil et al., 2005). The most crucial stage of plant development is germination (Song et al., 2008). Salinity tolerance at the germination stage may provide the ultimate yield gains in terms of grains produced (Jajarmi, 2009). Salt stress also significantly affected the overall performance of the germplasm in the present study. TG decreased by 24% and 42% in S1 and S2 compared to the control. TG was also highly correlated with RN, CL, SL, RL, and SVI. Highly significant differences between genotypes and treatments were observed for all traits. The same trends were also observed for the relative values of all traits under all treatments. Salinity stress affected all genotypes during S1 and S2 treatments compared to the control. The most adverse effects of salt stress were observed at S2, in which the TG dropped to 68% (TG_S2) from 91% (TG_S0). Other important seedling traits like CL, SL, RL, and SVI also decreased in the presence of elevated salt stress (S2), consistent with the findings reported by El Hehdaey et al. (2011).

Owing to the significance of the germination stage in plant tolerance against salinity stress (Munns and James, 2003; Singh et al., 2012), germination tests are among the most suited approaches for the early-stage screening of germplasm collections for salt tolerance (Munns et al., 2006; Aflaki et al., 2017). In the present study, chromosomes 1B (at 384.13 cM) and 7A (at 145.43) carried MTAs associated with TG under control and salt stress, respectively. Batool et al. (2018) also reported one QTL associated with standard germination on chromosome 1B. A recent report also underscored the importance of loci on chromosome 2B with respect to germination under post-abiotic stress in wheat (Arif and Börner, 2020). Likewise, a QTL related to germination under abiotic stress (experimental aging) has also been reported on chromosome 7A (Arif et al., 2012).

The root is the first plant organ that experiences salt stress; thus, it plays a significant role in sensing the salinity level in the nearby environment (Galvan-Ampudia et al., 2013; Robbins et al., 2014) by signal transduction (Jiang et al., 2012; Choi et al., 2014). During salt stress, the cell cycle activity of the root meristem is reduced, resulting in reduced growth (West et al., 2004). Hence, RN and RL are important criteria in estimating the salinity tolerance in wheat. In our study, RN increased by 13% in both S1 and S2 compared to

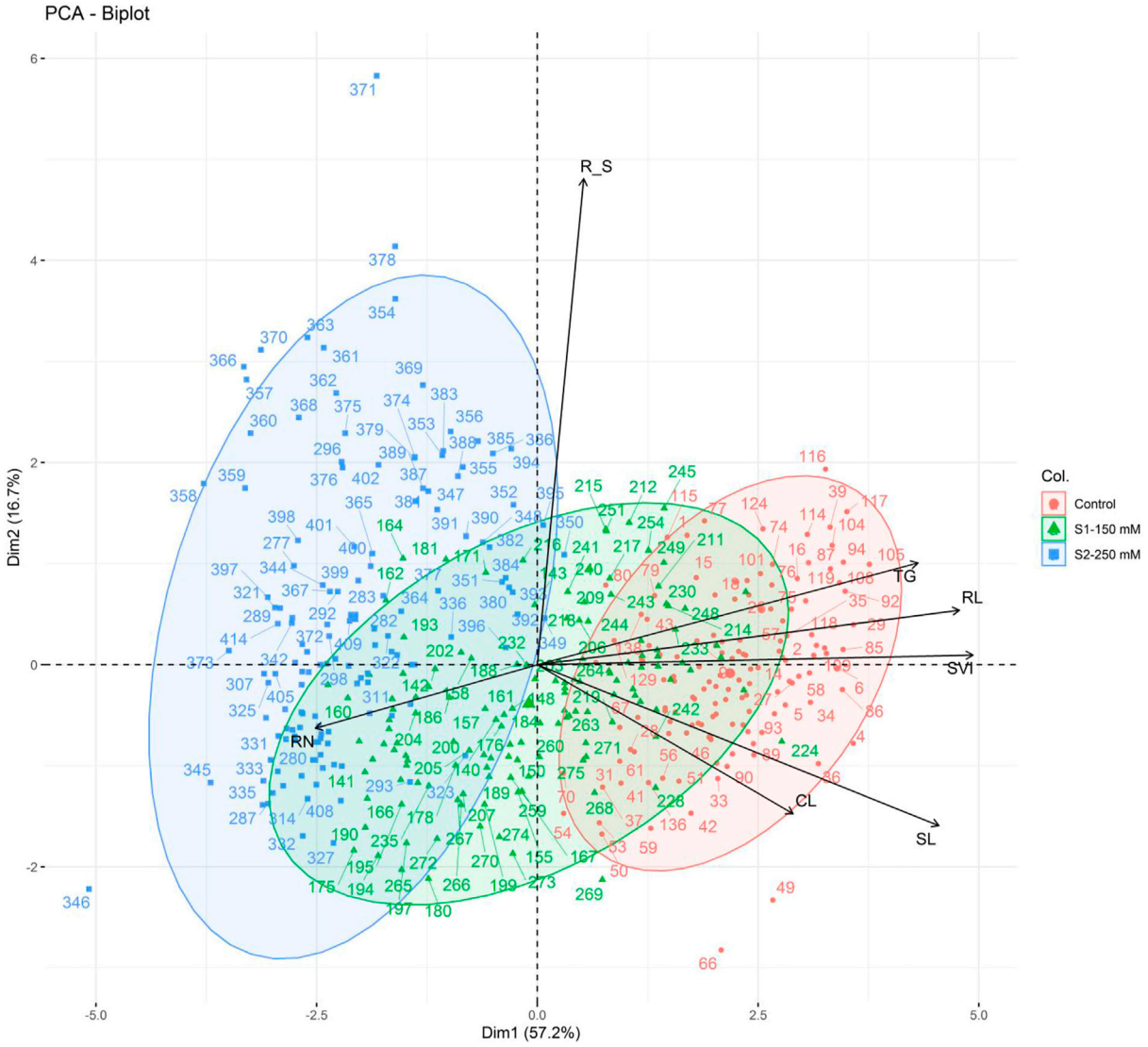


FIGURE 6
PCA of the morphological traits measured in 138 wheat genotypes in control (red) and salt stress [S1-150 Mm NaCl (blue) and S2-250 Mm NaCl)] conditions based on the first two components. TG, total germination percentage; RN, numbers of roots; CL, coleoptile length; SL, shoot length; RL, root length; R_S, root to shoot length ratio; SVI, seedling vigor index.

S0, with significant differences between G, T, and GXT. In contrast to RN, RL showed a significant decrease in S1 (35%) and S2 (65%) compared to the control. The same drop in RL was observed for relative traits, consistent with previous reports (Duan et al., 2013; Julkowska and Testerink, 2015). RL was highly correlated with SVI, whereas RN showed highly positive correlations with CL, SL, RL, and R/S. Previous studies have provided evidence of reduced root length due to elevated salt stress. A total of 58 MTAs linked to RN were detected on all chromosomes except for 1A, 4D, and 7B. Literature related to RN under salinity stress is scarce. However, our findings are consistent with those reported by Li et al. (2011), Salem and Mattar (2014), and Rufo et al. (2020). Likewise, chromosomes 1A, 1D, 3D, 4B, 5D, 7A, 7B, and 7D carried MTAs for RL and RRL under control and salt stress, similar to the results reported by Batool et al. (2018) and Salem and Mattar (2014). Marker M4090 present on 5D at 228.94 cM was linked to RL and RN, which showed a pleiotropic effect by controlling two characters on the same chromosome and location Markers controlling more than one trait are important for the improvement of salt stress tolerance in wheat (Batool et al., 2018).

The coleoptile protects the first leaf of the future wheat plant, which also functions as the driver to propel the leaf

outside the soil crust. Coleoptile strength and success is tantamount to the successful establishment and early plant vigor. Stress, may enhance the CL and shorten the SL (Zhang and Wang, 2012). In the present study, a 6% increase in coleoptile length was observed at S1 compared to S0. A 15% decrease was observed at S2 compared to S0. The CL also showed highly significant positive correlations with SL, RL, and R/S. Saboora et al. (2006) also observed the same trends in CL increases and decreases at moderate (75 mM) and higher levels of salt stress (150, 225, 300, and 375 mM). Moud and Maghsoudi (2008) reported that salt stress inhibited CL more than root growth. MTAs linked with CL under various conditions were detected on the chromosomes of group three and chromosomes 1A, 2A, 2B, 4A, and 4B. Li et al. (2011) reported QTLs associated with CL on chromosome 4B and 6B. In addition, two major QTLs of CLs were reported on chromosome 4B and 4D (Sidhu et al., 2019), corroborating our findings. Salem and Mattar (2014) reported QTLs related to CL under salinity stress at NaCl concentrations of 0 (on chromosome 1D), 150 (on 1D and 3BS), and 250 (on 1D, 4B, and 7D) mM; however, we detected no MTA for CL on chromosome 1D.

SL showed highly significant differences among treatments, with 33% and 68% decreases in S1 and S2, respectively. SL showed highly significant positive correlations with RL and R/S and a highly significant negative correlation with SVI. Bilkis et al. (2016) reported a 6%–36% decrease in shoot length under salt treatment. These findings were also similar to those of Datta et al. (2009) and Alom et al. (2016), who reported significantly reduced SL and RL at salinity levels <125 mM NaCl. SL is an important factor in the selection of genotypes against salt stress. The 20 MTAs of SL in various conditions were distributed on 14 different chromosomes, corroborating previous findings (Ghaedrahmati et al., 2014; Batool et al., 2018; Liu et al., 2018). The MTAs on chromosome 4B for SL_S0 (M8833 at 108.27 cM) and SL_S2 (M10038 at 90.17 cM) corresponded to the dwarfing gene *Rht-B1* on chromosome 4BS (Arif et al., 2021; Mo, 2018).

The R/S ratio is also disturbed under salt stress; however, this response is more tied to water stress than to salt stress (Hsiao and Xu, 2000). Increased RL as compared to SL may lead to the preservation of a large proportion of toxic ions in the roots and ameliorate their movement to the shoot, thus helping plant survival under salt stress (Cassaniti et al., 2009; Cassaniti et al., 2012). Çamlıca and Yaldız (2017) reported a decreased root/shoot length ratio with increasing salinity doses, with a greater reduction in root length than shoot length. In other words, the root length was more negatively affected than shoot length by increasing salinity doses. Landjeva et al. (2008) identified two QTLs on chromosome 3DL (*QRSRc.ipk-3D.1* and *QRSRc.ipk-3D.2*) and one QTL on chromosome 6DL (*QRSRp.ipk-6D*) that were associated with the R/S ratio under osmotic stress. We detected two highly significant MTAs (M1987 at 107.75 cM and M1019 at 116.66 cM) in S1 and S2 on chromosome 3D. Two MTAs, M11763 at 90.8 cM and M1188 at 121.22 in S1, were also detected on chromosome 6D, suggesting that both these loci were associated with R/S control under both salt and osmotic stresses.

Damaged plants show decreased viability, as represented by SVI (Copeland and McDonald, 2012). This is the most important trait for screening against salt stress. SVI is the product of many different factors and is related to genetics and environmental influences. The results of the current study showed highly significant differences between G, T, and G×T. SVI dropped by 48% and 91% at S1 and S2, respectively, from S0, implying that S2 was critical and damaging. A similar decrease was reported in *Brassica napus* (Batool et al., 2015) and *Hibiscus* species (Rashmi and Naik, 2014). In contrast (Batool et al., 2015), various QTLs associated with SVI at 150 mM NaCl on chromosomes 2A (*QSVI.2A.SG*), 4A (*QSVI.4A.SG*), 6D (*QSVI.6D.SG*) and 7B (*QSVI.7B.SG*) have been reported. We detected one MTA (M38 at 214.7 cM) on chromosome 7A that was associated with SVI_S0 on chromosome 7A. M38 was also associated with RL. Likewise, two MTAs (M7489 at 258.98 cM on chromosome 1B and M3286 at 113.89 on chromosome 7A) were also detected with RSVI_S2 on chromosome 7A, indicating that a wide variety of loci determine the SVI in wheat and are strongly dependent on RL and SL.

The BLAST search against the highly significant SNP markers identified in the present study revealed candidate genes involved in various stress conditions in plants. One such candidate gene (glycine-rich cell wall structural protein 1-like) was found on chromosome 5D. Glycine-rich proteins (GRP) are reportedly involved in stress responses including salinity, drought, etc., in many plants (Czolpinska and Rurek, 2018). Moreover, two candidate genes on chromosome 1B and 1D are involved in disease resistance. Additionally, multiple SNP markers matched candidate genes; i.e., metacaspase-1-like protein. Metacaspase-1 has a predominant role in the regulation of programmed cell death. The endoplasmic reticulum (ER) regulates protein synthesis. High salinity levels cause ER stress through the accumulation of misfolded proteins, which can lead to unfolded protein response (UPR) as a stress response mechanism. The UPR mechanism reverses misfolded proteins. UPR failure activates programmed cell death (Yusof et al., 2021). Metacaspase genes are key regulators of programmed cell death and might be the cardinal components of the saline stress pathway. Another candidate stress response gene (glyceraldehyde-3-phosphate dehydrogenase GAPA1, chloroplastic-like) located on chromosome 3B was also detected. Plastidial *GAPA1* has an abiotic stress response role in wheat and other plants (Chang et al., 2015; Li et al., 2019) and was associated with the M5307 marker related to RRN in S2 treatment at chromosome 3B. Munoz-Bertomeu et al. (2009) described the role of plastidial *GAPA1* in root development as this gene is involved in the biosynthesis of serine, which is essential for root development. The present study also showed the association of plastidial *GAPA1* in root development. Another gene, sphinganine C4-monooxygenase 1-like, involved in sphingolipid biosynthesis was also identified. Sphingolipids are ubiquitous and present in all types of plants.

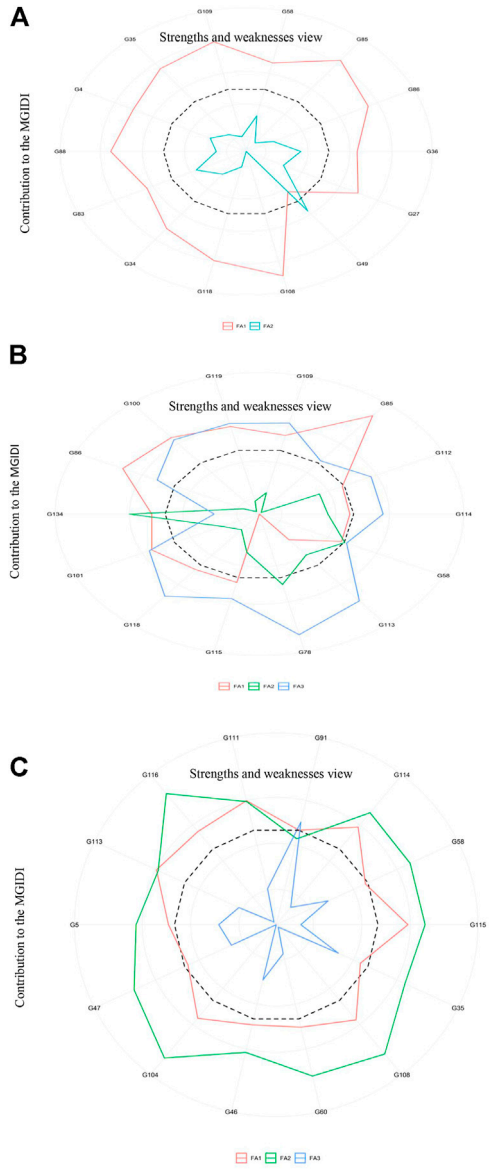


FIGURE 7
Strengths and weaknesses of the stable genotypes identified in the control **(A),** S1 (150 mM) **(B)**, and S2 (250 mM) **(C)** groups.

They comprise parts of plant cell membranes and endo-membranes. They also play roles in plant stress responses (Huby et al., 2020). Furthermore, another important candidate gene, calmodulin-binding protein 60 D-like, aligned to the M11763/RR/S_S1/6D marker. which corresponds to the shoot-to-root length ratio and is also involved in environmental stress responses in plants (Zeng et al., 2015). Calmodulin binding proteins play a significant role in plant growth, which corresponds to the results of the present study. Another stress response gene, subtilisin-like protease SBT1.7, was identified by the analysis in the present study. This gene plays a role in biotic stress response (Meyer et al., 2016). In addition, mRNA-decapping enzyme-like protein and 60S ribosomal protein L22-like were also identified as candidate genes. Kawa and Testerink (2017) and Liu et al. (2019) reported the role of both genes in salt stress response. Moreover, mRNA-decapping contributes to the regulation of ABA signaling (Munoz-Bertomeu et al., 2009). ABA mediates many developmental programs in plants, including seed dormancy or root growth (Finkelstein, 2013). The analysis of the M7775/RN_S2/5D marker in this study also suggested the involvement of the mRNA-decapping enzyme-like protein in root development.

PCA analysis disentangles a large data set into a small number of unrelated groups that can easily be plotted along independent linear axes. Closely linked variables in the same group may hint at latent relationships among them. Multiple traits often make it difficult to choose the best-performing genotypes. Therefore, several counter multivariate approaches such as cluster analysis, factor analysis, and PCA including other indices (Hazel, 1943; Williams, 1962) have been devised; however, each has limitations. We applied the MGIDI index for each treatment to identify superior genotypes. The analysis identified ten genotypes for further assessments; among these ten genotypes, only one was common to all three treatments. G58 showed the best ideotype, with positive genetic gains for all traits (Figure 7). MGIDI is a novel way to select genotypes. Other researchers have used this approach in different crops including strawberry, wheat, barley, guar, and soybean (Gabriel et al., 2019; Olivoto et al., 2021; Lima et al.; Farhad et al., 2022; Pour-Aboughadareh et al., 2021; Benakanahalli et al., 2021; Maranna et al., 2021).

According to the IMGIDI-based selection of genotypes, we identified the 10 best accessions (58, 85, 86, 108, 118, 35, 109, 113, 115, and 114) for all treatments (Figure 8). We observed the

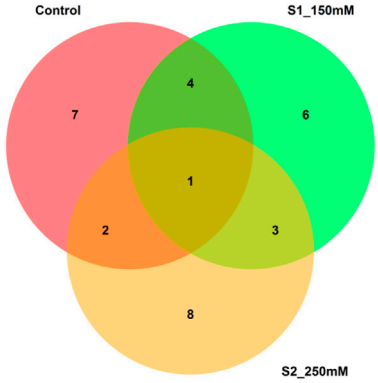


FIGURE 8
Venn diagram of the common genotypes shared across the treatments.

genotypic profile of these genotypes with respect to the highly significant MTAs discussed above. The identified genotypes carried from 37 to 44 positive alleles out of 48 possible positive alleles, with accessions 109 (GID: 7642809) and 115 (GID: 7642901) carrying the maximum numbers (44) of positive alleles. The phenotypic profiles of these accessions under both salinity levels (150 and 250 mM NaCl) indicated a percent increase in all traits except RN_S0, Sl_S0, RL_S0, R/S_S0, and SL_S2. After excluding these traits, all traits showed a mean increase of approximately 9.9% from the population mean (Supplementary Table S6). The use of accessions with more favorable alleles in wheat breeding can aid in improving salinity tolerance traits.

Conclusion

This study comprehensively dissected the performance of diverse bread wheat germplasm against different levels of salt stress. A total of 138 lines were screened at the seedling stage for seven traits at 0, 150, and 250 mM NaCl. We identified 195 significant SNPs/loci and 63 highly significant loci related to different traits. Most of the associations were present on the A genome, especially on chromosome 2A, and strengthened our findings regarding salinity tolerance. A total of 12 candidate genes were associated with highly significant SNP markers. The chromosomal localization of many of the important candidate genes such as Plastidial GAPA1, Metacaspase-1, etc., and their role in salt stress were also reported previously. These results of the extensive study of salt stress tolerance in *Triticum aestivum* L. could be a valuable reference for future studies. The best-performing lines with desirable allele combinations can be incorporated into wheat breeding programs.

Author's note

This experiment/research paper is a part of Saba Akram's PhD study.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, and further inquiries can be directed to the corresponding authors.

Author contributions

Basic idea, MA and AH; experimentation and theory, SA and MW; software, SA, MG, and AW; formal analysis, SA, MG, and AW; writing, SA, MG, AW, and SS; review and editing, MA and AH; supervision, MA and AH.

Acknowledgments

The author wants to acknowledge NIAB-C, PIEAS management for their help to complete the research work.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.997901/full#supplementary-material

References

Acosta-Motos, J. R., Ortuño, M. F., Bernal-Vicente, A., Diaz-Vivancos, P., Sanchez-Blanco, M. J., and Hernandez, J. A. (2017). Plant responses to salt stress: Adaptive mechanisms. *Agronomy* 7 (1), 18. doi:10.3390/agronomy7010018

Aflaki, F., Sedghi, M., Pazuki, A., and Pessarakli, M. (2017). Investigation of seed germination indices for early selection of salinity tolerant genotypes: A case study in wheat. *Emir. J. Food Agric.* 29, 222–226. doi:10.9755/ejfa.2016-12-1940

Akram, S., Arif, M. A. R., and Hameed, A. (2021). A GBS-based GWAS analysis of adaptability and yield traits in bread wheat (*Triticum aestivum* L.). *J. Appl. Genet.* 62 (1), 27–41. doi:10.1007/s13353-020-00593-1

Alom, R., Hasan, M. A., Islam, M. R., and Wang, Q.-F. (2016). Germination characters and early seedling growth of wheat (*Triticum aestivum* L.) genotypes under salt stress conditions. *J. Crop Sci. Biotechnol.* 19 (5), 383–392. doi:10.1007/s12892-016-0052-1

Arif, M. A. R., and Börner, A. (2020). An SNP based GWAS analysis of seed longevity in wheat. *Cereal Res. Commun.* 1–8. doi:10.1007/s42976-020-00025-0

Arif, M. A. R., Liaqat, M., Imran, M., Waheed, M. Q., Arif, A., Singh, S., et al. (2021). Genetic basis of some physiological traits and yield in early and late sowing conditions in bread wheat (*Triticum aestivum* L.). *J. Appl. Genet.* 62 (4), 601–605. doi:10.1007/s13353-021-00643-2

Arif, M. A. R., Waheed, M. Q., Lohwasser, U., Shokat, S., Alqudah, A. M., Volkmar, C., et al. (2022). Genetic insight into the insect resistance in bread wheat exploiting the untapped natural diversity. *Front. Genet.* 13, 828905. doi:10.3389/fgene.2022.828905

Arif, R., Nagel, M., Neumann, K., Kobiljski, B., Lohwasser, U., and Börner, A. (2012). Genetic studies of seed longevity in hexaploid wheat using segregation and association mapping approaches. *Euphytica* 186 (1), 1–13. doi:10.1007/s10681-011-0471-5

Ashraf, M., and Harris, P. (2004). Potential biochemical indicators of salinity tolerance in plants. *Plant Sci.* 166 (1), 3–16. doi:10.1016/j.plantsci.2003.10.024

Batool, N., Ilyas, N., Shahzad, A., Hauser, B. A., and Arshad, M. (2018). Quantitative trait loci (QTLs) mapping for salt stress tolerance in wheat at germination stage. *Pak. J. Agric. Sci.* 55 (1). doi:10.21162/pakjas/18.5426

Batool, N., Noor, T., Ilyas, N., and Shahzad, A. (2015). Salt stress impacts on seed germination and seedling growth of Brassica napus L. *Pab* 4 (3), 398–406. doi:10.19045/bspab.2015.43016

Benakanahalli, N. K., Sridhara, S., Ramesh, N., Olivoto, T., Sreekantappa, G., Tamam, N., et al. (2021). A framework for identification of stable genotypes basedon MTSI and MGDII indexes: An example in guar (*Cymopsis tetragonoloba* L.). *Agronomy* 11 (6), 1221. doi:10.3390/agronomy11061221

Berger, B., de Regt, B., and Tester, M. (2012). Trait dissection of salinity tolerance with plant phenomics. *Plant salt Toler.* 913, 399–413. doi:10.1007/978-1-61779-986-0_27

Bilkis, A., Islam, M., Hafiz, M., and Hasan, M. (2016). Effect of NaCl induced salinity on some physiological and agronomic traits of wheat. *Pak. J. Bot.* 48 (2), 455–460.

Bizari, E. H., Val, B. H. P., Pereira, E. d. M., Mauro, A. O. D., and Unêda-Trevisoli, S. H. (2017). Selection indices for agronomic traits in segregating populations of soybean1. *Rev. Cienc. Agron.* 48, 110–117. doi:10.5935/1806-6690.20170012

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 (19), 2633–2635. doi:10.1093/bioinformatics/btm308

Çamlıca, M., and Yaldız, G. (2017). Effect of salt stress on seed germination, shoot and root length in Basil (*Ocimum basilicum*). *Int. J. Second. Metab.* 4 (3), 69–76. doi:10.21448/ijsm.356250

Cassaniti, C., Leonardi, C., and Flowers, T. J. (2009). The effects of sodium chloride on ornamental shrubs. *Sci. Hortic.* 122 (4), 586–593. doi:10.1016/j.scienta.2009.06.032

Cassaniti, C., Romano, D., and Flowers, T. J. (2012). *The response of ornamental plants to saline irrigation water*, 131–158.*Irrigation Water Manag. Pollut. Altern. Strategies*

Céron-Rojas, J. J., and Crossa, J. (2018). *Linear selection indices in modern plant breeding*. Springer Nature.

Chang, L., Guo, A., Jin, X., Yang, Q., Wang, D., Sun, Y., et al. (2015). The beta subunit of glyceraldehyde 3-phosphate dehydrogenase is an important factor for maintaining photosynthesis and plant development under salt stress-Based on an integrative analysis of the structural, physiological and proteomic changes in chloroplasts in Thellungiella halophila. *Plant Sci.* 236, 223–238. doi:10.1016/j.plantsci.2015.04.010

Chen, W., Guo, C., Hussain, S., Zhu, B., Deng, F., Xue, Y., et al. (2016). Role of xylo-oligosaccharides in protection against salinity-induced adversities in Chinese cabbage. *Environ. Sci. Pollut. Res.* 23 (2), 1254–1264. doi:10.1007/s11356-015-5361-2

Choi, W.-G., Toyota, M., Kim, S.-H., Hilleary, R., and Gilroy, S. (2014). Salt stress-induced Ca 2+ waves are associated with rapid, long-distance root-to-shoot signaling in plants. *Proc. Natl. Acad. Sci. U.S.A.* 111 (17), 6497–6502. doi:10.1073/pnas.1319955111

Colmer, T., Munns, R., and Flowers, T. (2006). Improving salt tolerance of wheat and barley: Future prospects. *Aust. J. Exp. Agric.* 45 (11), 1425–1443. doi:10.1071/EA04162

Copeland, L. O., and McDonald, M. F. (2012). *Principles of seed science and technology*. Springer Science & Business Media.

Czolpinska, M., and Rurek, M. (2018). Plant glycine-rich proteins in stress response: An emerging, still prospective story. *Front. Plant Sci.* 9, 302. doi:10.3389/fpls.2018.00302

Datta, J., Nag, S., Banerjee, A., and Mondai, N. (2009). Impact of salt stress on five varieties of wheat (*Triticum aestivum* L.) cultivars under laboratory condition. *J. Appl. Sci. Environ. Manag.* 13 (3). doi:10.4314/jasem.v13i3.55372

Duan, L., Dietrich, D., Ng, C. H., Chan, P. M. Y., Bhalerao, R., Bennett, M. J., et al. (2013). Endodermal ABA signaling promotes lateral root quiescence during salt stress inArabidopsisSeedlings. *Plant Cell* 25 (1), 324–341. doi:10.1105/tpc.112.107227

Earl, D. A., and vonHoldt, B. M. (2012). Structure harvester: A website and program for visualizing STRUCTURE output and implementing the evanno method. *Conserv. Genet. Resour.* 4 (2), 359–361. doi:10.1007/s12686-011-9548-7

El Hendawy, S., Hu, Y., Sakagami, J.-I., and Schmidhalter, U. (2011). *Screening Egyptian wheat genotypes for salt tolerance at early growth stages*, 283–298.

El Sabagh, A., Islam, M. S., Skalicky, M., Ali Raza, M., Singh, K., Anwar Hossain, M., et al. (2021). Salinity stress in wheat (*Triticum aestivum* L.) in the changing climate: Adaptation and management strategies. *Front. Agron.* 3, 661932. doi:10.3389/fagro.2021.661932

Epskamp, S., Cramer, A. O., Waldorp, L. J., Schmittmann, V. D., and Borsboom, D. (2012). qgraph: Network visualizations of relationships in psychometric data. *J. Stat. Softw.* 48, 1–18. doi:10.18637/jss.v048.i04

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 (8), 2611–2620. doi:10.1111/j.1365-294x.2005.02553.x

Farhad, M., Tripathi, S. B., Singh, R. P., Joshi, A. K., Bhati, P. K., Vishwakarma, M. K., et al. (2022). Multi-trait selection of bread wheat ideotypes for adaptation to early sown condition. *Crop Sci.* 62 (1), 67–82. doi:10.1002/csc2.20628

Fernandez, G. C. (1993). *Adaptation of food crops to temperature and water stress*, 13–181992257270.Effective selection criteria for assessing plant stress tolerance

Finkelstein, R. (2013). Abscisic acid synthesis and response. *Arab. Book* 11, e0166. doi:10.1199/tab.0166

Flowers, T., and Flowers, S. (2005). Why does salinity pose such a difficult problem for plant breeders? *Agric. Water Manag.* 78 (1-2), 15–24. doi:10.1016/j.agwat.2005.04.015

Gabriel, A., Resende, J. T., Zeist, A. R., Resende, L. V., Resende, N. C., and Zeist, R. A. (2019). Phenotypic stability of strawberry cultivars based on physicochemical traits of fruits. *Hortic. Bras.* 37, 75–81. doi:10.1590/s0102-053620190112

Galvan-Ampudia, C. S., Julkowska, M. M., Darwish, E., Gandullo, J., Korver, R. A., Brunoud, G., et al. (2013). Halotropism is a response of plant roots to avoid a saline environment. *Curr. Biol.* 23 (20), 2044–2050. doi:10.1016/j.cub.2013.08.042

Ghaedrahmati, M., Mardi, M., Naghavi, M., Haravan, E. M., Nakhoda, B., Azadi, A., et al. (2014). Mapping QTLs associated with salt tolerance related traits in seedling stage of wheat (*Triticum aestivum* L.). *J. Agric. Sci. Tech.* 16, 1413–1428.

Ghassemi, F., Jakeman, A. J., and Nix, H. A. (1995). *Salinisation of land and water resources: Human causes, extent, management and case studies*. Wallingford, Oxfordshire: CAB international.

Großkinsky, D. K., Roitsch, T., and Liu, F. (2020). Higher activity of monodehydroascorbate reductase and lower activities of leaf and spike vacuolar invertase and glutathione S-transferase reveals higher number of grains per spike in spring wheat genotypes grown under well-watered and drought conditions. *BMC Plant Biol.*

Hasanuzzaman, M., Nahar, K., Rahman, A., Anee, T. I., Alam, M. U., Bhuiyan, T. F., et al. (2017). Approaches to enhance salt stress tolerance in wheat. *Wheat improvement, management and utilization*, 151–187.

Hazel, L. N. (1943). The genetic basis for constructing selection indexes. *Genetics* 28 (6), 476–490. doi:10.1093/genetics/28.6.476

Holm, S. (1979). A simple sequentially rejective multiple test procedure. *Scand. Stat. Theory Appl.* 65–70.

Hsiao, T. C., and Xu, L. K. (2000). Sensitivity of growth of roots versus leaves to water stress: Biophysical analysis and relation to water transport. *J. Exp. Bot.* 51 (350), 1595–1616. doi:10.1093/jexbot/51.350.1595

Hu, V. W., Addington, A., and Hyman, A. (2011). Novel autism subtype-dependent genetic variants are revealed by quantitative trait and subphenotype association analyses of published GWAS data. *PLoS One* 6 (4), e19067. doi:10.1371/journal.pone.0019067

Huang, Y., Bie, Z., He, S., Hua, B., Zhen, A., and Liu, Z. (2010). Improving cucumber tolerance to major nutrients induced salinity by grafting onto Cucurbita ficifolia. *Environ. Exp. Bot.* 69 (1), 32–38. doi:10.1016/j.envexpbot.2010.02.002

Huby, E., Napier, J. A., Baillieul, F., Michaelson, L. V., and Dhondt-Cordelier, S. (2020). Sphingolipids: Towards an integrated view of metabolism during the plant stress response. *New Phytol.* 225 (2), 659–670. doi:10.1111/nph.15997

Hussain, B., Lucas, S. J., Ozturk, L., and Budak, H. (2017). Mapping QTLs conferring salt tolerance and micronutrient concentrations at seedling stagein wheat. *Sci. Rep.* 7, 15662. doi:10.1038/s41598-017-15726-6

Jajarmi, V. (2009). Effect of water stress on germination indices in seven wheat cultivar. World Academy of Science. *Eng. Technol.* 49, 105–106.

James, R. A., Blake, C., Byrt, C. S., and Munns, R. (2011). Major genes for Na+ exclusion, Nax1 and Nax2 (wheat HKT1;4 and HKT1;5), decrease Na+ accumulation in bread wheat leaves under saline and waterlogged conditions. *J. Exp. Bot.* 62 (8), 2939–2947. doi:10.1093/jxb/err003

Jamil, A., Riaz, S., Ashraf, M., and Foolad, M. (2011). Gene expression profiling of plants under salt stress. *Crit. Rev. Plant Sci.* 30 (5), 435–458. doi:10.1080/07352689.2011.605739

Jamil, M., Lee, C. C., Rehman, S. U., Lee, D. B., Ashraf, M., and Rha, E. S. (2005). Salinity (NaCl) tolerance of Brassica species at germination and early seedling growth. Electronic Journal of Environmental. *J. Agric. Food Chem.* 4 (4), 970–976.

Jiang, C., Belfield, E. J., Mithani, A., Visscher, A., Ragoussis, J., Mott, R., et al. (2012). ROS-mediated vascular homeostatic control of root-to-shoot soil Na delivery inArabidopsis. *EMBO J.* 31 (22), 4359–4370. doi:10.1038/emboj.2012.273

Jighly, A., Oyiga, B. C., Makdis, F., Nazari, K., Youssef, O., Tadesse, W., et al. (2015). Genome-wide DArT and SNP scan for QTL associated with resistance to stripe rust (Puccinia striiformis f. sp. tritici) in elite ICARDA wheat (*Triticum aestivum* L.) germplasm. *Theor. Appl. Genet.* 128 (7), 1277–1295. doi:10.1007/s00122-015-2504-2

Julkowska, M. M., and Testerink, C. (2015). Tuning plant signaling and growth to survive salt. *Trends Plant Sci.* 20 (9), 586–594. doi:10.1016/j.tplants.2015.06.008

Kandil, A., Sharief, A., and Elokda, M. (2012). Germination and seedling characters of different wheat cultivars under salinity stress. *J. Basic Appl. Sci.* 8 (2), 585–596. doi:10.6000/1927-5129.2012.08.02.52

Kassambara, A., and Mundt, F. (2017). Package 'factoextra. *Extr. Vis. results Multivar. data analyses* 76 (2).

Kawa, D., and Testerink, C. (2017). Regulation of mRNA decay in plant responses to salt and osmotic stress. *Cell. Mol. Life Sci.* 74 (7), 1165–1176. doi:10.1007/s00018-016-2376-x

Landjeva, S., Neumann, K., Lohwasser, U., and Börner, A. (2008). Molecular mapping of genomic regions associated with wheat seedling growth under osmotic stress. *Biol. Plant.* 52 (2), 259–266. doi:10.1007/s10535-008-0056-x

Li, P., Chen, J., Wu, P., Zhang, J., Chu, C., See, D., et al. (2011). Quantitative trait loci analysis for the effect ofRht-B1Dwarfing gene on coleoptile length and seedling root length and number of bread wheat. *Crop Sci.* 51 (6), 2561–2568. doi:10.2135/cropsci2011.03.0116

Li, W., Zhang, H., Zeng, Y., Xiang, L., Lei, Z., Huang, Q., et al. (2020). A salt tolerance evaluation method for sunflower (Helianthus annuus L.) at the seed germination stage. *Sci. Rep.* 10 (1), 10626–10629. doi:10.1038/s41598-020-67210-3

Li, X., Wei, W., Li, F., Zhang, L., Deng, X., Liu, Y., et al. (2019). The plastidial glyceraldehyde-3-phosphate dehydrogenase is critical for abiotic stress response in wheat. *Ijms* 20 (5), 1104. doi:10.3390/ijms20051104

Lima, G. W., e Silva, C. M., Mezzomo, H. C., Casagrande, C. R., Olivoto, T., Borem, A., et al. (2022). Genetic diversity in tropical wheat germplasm and selection via multi-trait index. *J. Agron.* 114 (2), 887–899. doi:10.1002/agj2.20991

Liu, J., He, Z., Rasheed, A., Wen, W., Yan, J., Zhang, P., et al. (2017). Genome-wide association mapping of black point reaction in common wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 17, 220. doi:10.1186/s12870-017-1167-3

Liu, Y., Liu, Y., Zhang, Q., Fu, B., Cai, J., Wu, J., et al. (2018). Genome-wide association analysis of quantitative trait loci for salinity-tolerance related morphological indices in bread wheat. *Euphytica* 214 (10), 176. doi:10.1007/s10681-018-2265-5

Liu, Z., Zou, L., Chen, C., Zhao, H., Yan, Y., Wang, C., et al. (2019). iTRAQ-based quantitative proteomic analysis of salt stress in Spica Prunellae. *Sci. Rep.* 9, 9590. doi:10.1038/s41598-019-46043-9

Maranna, S., Nataraj, V., Kumawat, G., Chandra, S., Rajesh, V., Ramteke, R., et al. (2021). Breeding for higher yield, early maturity, wider adaptability and waterlogging tolerance in soybean (Glycine max L.): A case study. *Sci. Rep.* 11, 22853. doi:10.1038/s41598-021-02064-x

Masoudi, B., Mardi, M., Hervan, E. M., Bihamta, M. R., Naghavi, M. R., Nakhoda, B., et al. (2015). QTL mapping of salt tolerance traits with different effects at the seedling stage of bread wheat. *Plant Mol. Biol. Rep.* 33, 1790–1803. doi:10.1007/s11105-015-0874-x

Meyer, M., Huttenlocher, F., Cedzich, A., Procopio, S., Stroeder, J., Pau-Roblot, C., et al. (2016). The subtilisin-like protease SBT3 contributes to insect resistance in tomato. *Exbotj* 67 (14), 4325–4338. doi:10.1093/jxb/erw220

Mo, Y. (2018). *Exploring natural and induced genetic variation controlling plant height in wheat*. Davis: University of California.

Moud, A. M., and Maghsoudi, K. (2008). Salt stress effects on respiration and growth of germinated seeds of different wheat (*Triticum aestivum* L.) cultivars. *World J. Agric. Sci.* 4 (3), 351–358.

Munns, R. (2005). Genes and salt tolerance: Bringing them together. *New Phytol.* 167 (3), 645–663. doi:10.1111/j.1469-8137.2005.01487.x

Munns, R., James, R. A., and Läuchli, A. (2006). Approaches to increasing the salt tolerance of wheat and other cereals. *J. Exp. Bot.* 57 (5), 1025–1043. doi:10.1093/jxb/erj100

Munns, R., and James, R. A. (2003). Screening methods for salinity tolerance: A case study with tetraploid wheat. *Plant soil* 253 (1), 201–218. doi:10.1023/a:1024553303144

Munns, R., and Tester, M. (2008). Mechanisms of salinity tolerance. *Annu. Rev. Plant Biol.* 59, 651–681. doi:10.1146/annurev.arplant.59.032607.092911

Muñoz-Bertomeu, J., Cascales-Miñana, B., Mulet, J. M., Baroja-Fernández, E., Pozueta-Romero, J., Kuhn, J. M., et al. (2009). Plastidial glyceraldehyde-3-phosphate dehydrogenase deficiency leads to altered root development and affects the sugar and amino acid balance in Arabidopsis. *Plant Physiol.* 151 (2), 541–558. doi:10.1104/pp.109.143701

Mwando, E., Han, Y., Angessa, T. T., Zhou, G., Hill, C. B., Zhang, X.-Q., et al. (2020). Genome-wide association study of salinity tolerance during germination in barley (Hordeum vulgare L.). *Front. Plant Sci.* 11, 118. doi:10.3389/fpls.2020.00118

Olivoto, T., Diel, M. I., Schmidt, D., and Lúcio, A. D. C. (2021). Multivariate analysis of strawberry experiments: Where are we now and where can we go? *BioRxiv* 2020. doi:10.2012/2030.424876

Olivoto, T., and Lúcio, A. D. C. (2020). metan: An R package for multi-environment trial analysis. *Methods Ecol. Evol.* 11 (6), 783–789. doi:10.1111/2041-210x.13384

Olivoto, T., and Nardino, M. (2021). Mgidi: Toward an effective multivariate selection in biological experiments. *Bioinformatics* 37 (10), 1383–1389. doi:10.1093/bioinformatics/btaa981

Oyiga, B. C., Sharma, R., Shen, J., Baum, M., Ogbonnaya, F., Léon, J., et al. (2016). Identification and characterization of salt tolerance of wheat germplasm using a multivariable screening approach. *J. Agro Crop Sci.* 202 (6), 472–485. doi:10.1111/jac.12178

Pour-Aboughadareh, A., and Poczai, P. (2021). Dataset on the use of MGIDI index in screening drought-tolerant wild wheat accessions at the early growth stage. *Data Brief* 36, 107096. doi:10.1016/j.dib.2021.107096

Pour-Aboughadareh, A., Sanjani, S., Nikkhah-Chamanabad, H., Mehrvar, M. R., Asadi, A., and Amini, A. (2021). Identification of salt-tolerant barley genotypes using multiple-traits index and yield performance at the early growth and maturity stages. *Bull. Natl. Res. Cent.* 45 (1), 1–16. doi:10.1186/s42269-021-00576-0

Pritchard, J. K., Wen, W., and Falush, D. (2003). *Documentation for STRUCTURE software: Version 2*.

Rahman, A., Nahar, K., Hasanuzzaman, M., and Fujita, M. (2016). Calcium supplementation improves Na+/K+ ratio, antioxidant defense and glyoxalase systems in salt-stressed rice seedlings. *Front. Plant Sci.* 7, 609. doi:10.3389/fpls.2016.00609

Rajabi Dehnavi, A., Zahedi, M., Ludwiczak, A., Cardenas Perez, S., and Piernik, A. (2020). Effect of salinity on seed germination and seedling development of sorghum (Sorghum bicolor (L.) Moench) genotypes. *Agronomy* 10 (6), 859. doi:10.3390/agronomy10060859

Ramasamy, R. K., Ramasamy, S., Bindroo, B. B., and Naik, V. G. (2014). Structure plot: A program for drawing elegant STRUCTURE bar plots in user friendly interface. *SpringerPlus* 3 (1), 431–433. doi:10.1186/2193-1801-3-431

Rashmi, G., and Naik, G. (2014). A comparative study on the effect of salt stress on seed germination and early seedling growth of two Hibiscus species. *IOSR J. Agric. Vet. Sci.* 7, 90–96.

Robbins, N. E., Trontin, C., Duan, L., and Dinneny, J. R. (2014). Beyond the barrier: Communication in the root through the endodermis. *Plant Physiol.* 166 (2), 551–559. doi:10.1104/pp.114.244871

Royo, A., and Abió, D. (2003). Salt tolerance in durum wheat cultivars. *Span. J. Agric. Res.* 1 (3), 27–35. doi:10.5424/sjar/2003013-32

Rufo, R., Salvi, S., Royo, C., and Soriano, J. M. (2020). Exploring the genetic architecture of root-related traits in mediterranean bread wheat landraces by genome-wide association analysis. *Agronomy* 10 (5), 613. doi:10.3390/agronomy10050613

Saboora, A., Kiarostami, K., Behroozbayati, F., and Hajihashemi, S. (2006). Salinity (NaCl) tolerance of wheat genotypes at germination and early seedling growth. *Pak. J. Biol. Sci.* 9 (11), 2009–2021. doi:10.3923/pjbs.2006.2009.2021

Saddiq, M. S., Iqbal, S., Afzal, I., Ibrahim, A. M., Bakhtavar, M. A., Hafeez, M. B., et al. (2019). Mitigation of salinity stress in wheat (*Triticum aestivum* L.) seedlings through physiological seed enhancements. *J. Plant Nutr.* 42 (10), 1192–1204. doi:10.1080/01904167.2019.1609509

Salem, K. F., and Mattar, M. Z. (2014). Identification of microsatellite alleles for salt tolerance at seedling stage in wheat (*Triticum aestivum* L.). *Life Sci.* 11 (12).

Sayed, M. A., Nassar, S. M., Moustafa, E. S., Said, M. T., Börner, A., and Hamada, A. (2021). Genetic mapping reveals novel exotic and elite QTL alleles for salinity tolerance in barley. *Agronomy* 11 (9), 1774. doi:10.3390/agronomy11091774

Shabala, S., and Munns, R. (2012). Salinity stress: Physiological constraints and adaptive mechanisms. *Plant Stress Physiol.* 1, 59–93. doi:10.1079/9781845939953.0059

Shafi, M., Bakhat, J., Khan, M. J., Khan, M. A., and Anwar, S. (2010). Effect of salinity on yield and ion accumulation of wheat genotypes. *Pak. J. Bot.* 42 (6), 4113–4121.

Sharma, A., Arif, M. A. R., Shamshad, M., Rawale, K. S., Brar, A., Burgueño, D., et al. (2022). Preliminary dissection of grain yield and related traits at differential nitrogen levels in diverse pre-breeding wheat germplasm through association mapping. *Mol. Biotechnol.* doi:10.1007/s12033-022-00535-8

Shi, Y., Gao, L., Wu, Z., Zhang, X., Wang, M., Zhang, C., et al. (2017). Genome-wide association study of salt tolerance at the seed germination stage in rice. *BMC Plant Biol.* 17 (1), 92. doi:10.1186/s12870-017-1044-0

Shokat, S., and Großkinsky, D. K. (2019). Tackling salinity in sustainable agriculture—What developing countries may learn from approaches of the developed world. *Sustainability* 11 (17), 4558.

Sidhu, J. S., Singh, D., Gill, H. S., Brar, N. K., Qiu, Y., Halder, J., et al. (2019). Genome-wide association study uncovers novel genomic regions associated with coleoptile length in hard winter wheat. *Front. Genet.* 10, 1345. doi:10.3389/fgene.2019.01345

Singh, J., Sastry, E. D., and Singh, V. (2012). Effect of salinity on tomato (*Lycopersicon esculentum* Mill.) during seed germination stage. *Physiol. Mol. Biol. Plants.* 18 (1), 45–50. doi:10.1007/s12298-011-0097-z

Singh, S., Vikram, P., Sehgal, D., Burgueño, J., Sharma, A., Singh, S. K., et al. (2018). Harnessing genetic potential of wheat germplasm banks through impact-oriented-prebreeding for future food and nutritional security. *Sci. Rep.* 8, 12527. doi:10.1038/s41598-018-30667-4

Song, J., Fan, H., Zhao, Y., Jia, Y., Du, X., and Wang, B. (2008). Effect of salinity on germination, seedling emergence, seedling growth and ion accumulation of a euhalophyte Suaeda salsa in an intertidal zone and on saline inland. *Aquat. Bot.* 88 (4), 331–337. doi:10.1016/j.aquabot.2007.11.004

Takahashi, F., Tilbrook, J., Trittermann, C., Berger, B., Roy, S. J., Seki, M., et al. (2015). Comparison of leaf sheath transcriptome profiles with physiological traits of bread wheat cultivars under salinity stress. *PLoS One* 10 (8), e0133322. doi:10.1371/journal.pone.0133322

Turki, N., Shehzad, T., Harrabi, M., and Okuno, K. (2015). Detection of QTLs associated with salinity tolerance in durum wheat based on association analysis. *Euphytica* 201 (1), 29–41. doi:10.1007/s10681-014-1164-7

Verslues, P. E., Agarwal, M., Katiyar-Agarwal, S., Zhu, J., and Zhu, J. K. (2006). Methods and concepts in quantifying resistance to drought, salt and freezing, abiotic stresses that affect plant water status. *Plant J.* 45 (4), 523–539. doi:10.1111/j.1365-313x.2005.02593.x

West, G., Inze´, D., and Beemster, G. T. (2004). Cell cycle modulation in the response of the primary root of Arabidopsis to salt stress. *Plant Physiol.* 135 (2), 1050–1058. doi:10.1104/pp.104.040022

Wickham, H. (2016). "Data analysis," in *ggplot2* (Springer), 189–201. doi:10.1007/978-3-319-24277-4_9

Williams, J. (1962). The evaluation of a selection index. *Biometrics* 18 (3), 375–393. doi:10.2307/2527479

Xue, D., Huang, Y., Zhang, X., Wei, K., Westcott, S., Li, C., et al. (2009). Identification of QTLs associated with salinity tolerance at late growth stage in barley. *Euphytica* 169 (2), 187–196. doi:10.1007/s10681-009-9919-2

Yin, L., Zhang, H., Tang, Z., Xu, J., Yin, D., Zhang, Z., et al. (2021). rMVP: a memory-efficient, visualization-enhanced, and parallel-accelerated tool for genome-wide association study. *Genomics, Proteomics Bioinforma.* 19 (4), 619–628. doi:10.1016/j.gpb.2020.10.007

Yusof, N. F. M., Saparin, N. F., Seman, Z. A., Ab Rahman, Z., Sew, Y. S., Shaharuddin, N. A., et al. (2021). *Overexpression of rice metacaspase, OsMC4, increases endoplasmic reticulum stress tolerance in transgenic rice calli*.

Zafar, S., Ashraf, M. Y., Niaz, M., Kausar, A., and Hussain, J. (2015). Evaluation of wheat genotypes for salinity tolerance using physiological indices as screening tool. *Pak. J. Bot.* 47 (2), 397–405.

Zeng, H., Xu, L., Singh, A., Wang, H., Du, L., and Poovaiah, B. (2015). Involvement of calmodulin and calmodulin-like proteins in plant responses to abiotic stresses. *Front. Plant Sci.* 6, 600. doi:10.3389/fpls.2015.00600

Zhang, H., and Wang, H. (2012). Evaluation of drought tolerance from a wheat recombination inbred line population at the early seedling growth stage. *Afr. J. Agric. Res.* 7 (46), 6167–6172. doi:10.5897/AJAR12.2004

Zhou, G., Johnson, P., Ryan, P. R., Delhaize, E., and Zhou, M. (2012). Quantitative trait loci for salinity tolerance in barley (Hordeum vulgare L.). *Mol. Breed.* 29 (2), 427–436. doi:10.1007/s11032-011-9559-9

Zhu, T., Wang, L., Rimbert, H., Rodriguez, J. C., Deal, K. R., De Oliveira, R., et al. (2021). Optical maps refine the bread wheat *Triticum aestivum* cv. Chinese Spring genome assembly. *Plant J.* 107 (1), 303–314. doi:10.1111/tpj.15289

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 13 October 2022
DOI 10.3389/fgene.2022.1007896

OPEN ACCESS

EDITED BY
Manish Kumar Pandey,
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

REVIEWED BY
Chaolei Liu,
China National Rice Research Institute (CAAS), China
Jinpeng Wang,
Institute of Botany (CAS), China

*CORRESPONDENCE
Xianjin Qiu,
xjqiu216@yangtzeu.edu.cn
Jianlong Xu,
xujlcaas@126.com

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 31 July 2022
ACCEPTED 03 October 2022
PUBLISHED 13 October 2022

CITATION
Wang W, Zhang F, Liu D, Chen K, Du B, Qiu X, Xu J and Xing D (2022), Distribution characteristics of selenium, cadmium and arsenic in rice grains and their genetic dissection by genome-wide association study.
Front. Genet. 13:1007896.
doi: 10.3389/fgene.2022.1007896

Distribution characteristics of selenium, cadmium and arsenic in rice grains and their genetic dissection by genome-wide association study

Wenxi Wang[1,2], Fan Zhang[3], Dapu Liu[3], Kai Chen[4], Bin Du[2], Xianjin Qiu[2]*, Jianlong Xu[3,4,5]* and Danying Xing[2]

[1]College of Economy and Management, Hubei University of Technology, Wuhan, China, [2]MARA Key Laboratory of Sustainable Crop Production in the Middle Reaches of the Yangtze River (Co-construction by Ministry and Province), College of Agriculture, Yangtze University, Jingzhou, China, [3]Institute of Crop Science, Chinese Academy of Agricultural Sciences, Beijing, China, [4]Shenzhen Branch, Guangdong Laboratory for Lingnan Modern Agriculture, Agricultural Genomics Institute at Shenzhen, Chinese Academy of Agricultural Sciences, Shenzhen, China, [5]National Nanfan Research Institute (Sanya), Chinese Academy of Agricultural Sciences, Sanya, China

High selenium (Se) and low cadmium (Cd) and arsenic (As) contents in rice grains were good for human health. The genetic basis and relationship of Se, Cd and As concentrations in rice grains are still largely unknown. In the present study, large variations were observed in Se, Cd and As concentrations in brown and milled rice in normal and Se treatment conditions in 307 rice accessions from 3K Rice Genomes Project. Se fertilizer treatment greatly increased Se concentrations but had no obvious changes in concentrations of Cd and As both in brown and milled rice. Total of 237 QTL were identified for Se, Cd and As concentrations in brown and milled rice in normal and Se treatment conditions as well as ratio of concentrations under Se treatment to normal conditions. Only 19 QTL (13.4%) were mapped for concentrations of Se and Cd, Se and As, and Se, Cd and As in the same or adjacent regions, indicating that most Se concentration QTL are independent of Cd and As concentration QTL. Forty-three favorable alleles were identified for 40 candidate genes by gene-based association study and haplotype analysis in 14 important QTL regions. Se-enriched rice variety will be developed by pyramiding favorable alleles at different Se QTL and excluding undesirable alleles at Cd and As QTL, or combining favorable alleles at Se QTL with the alleles at Se-sensitive QTL by marker-assisted selection.

KEYWORDS

rice, selenium, cadmium, arsenic, genome-wide association study, quantitative trait locus/loci (QTL), candidate gene

Introduction

Selenium (Se) is an essential element for humans, which is important component for many enzymes in our body, and has immune function (Meplan and Hughes, 2020). Se deficiency would damage our health, such as Keshan disease, Kashin-Beck disease and cancer (Tapiero et al., 2003; Cox and Bastiaans, 2007). It is mainly taken up by food, and the recommended daily intake of Se is 50–200 μg day^{-1} (Gailer, 2009). In some area of China, Se concentration in soil is very low, resulting in the Se deficiency in human diet (Tapiero et al., 2003), and the average daily intake of Se in this area is 26–32 μg (Chen et al., 2002). On the other hand, with the intensive anthropogenic activities during last century such as metal mining, smelting, industrial activities and use of excess fertilizers (Bolan et al., 2013), two heavy metals, cadmium (Cd) and arsenic (As), have been large problems and become a growing concern (Feng et al., 2013). They are mainly taken up by food and water, and exceeded intake of them would cause diseases such as renal dysfunction, osteoporosis and cancer (Rikans and Yamano, 2000; Liu et al., 2007; Akesson et al., 2014). Because of soil and water contamination, the concentrations of these two elements exceeded the Chinese maximum permissible concentrations of them in rice grains in some areas in southern China (Shang et al., 2012; Du et al., 2013; Zhu et al., 2016; Wang et al., 2019).

Rice is one of the most important staple crops, and provides food for two-thirds of world's population, especially for Chinese people (Zhang J. et al., 2018), thus it is also one of the main resources of above three elements. Biofortification is an effective way to improve Se concentration in rice grains. However, many variables are involved in Se biofortification strategies, such as the Se administration mode (soil fertilization, foliar spray, or hydroponics), Se dose, species and fertilizer form, crop species, and variety and growth stage. Foliar application is a valid alternative for Se enrichment of agricultural products. Besides, it was reported that Se can inhibit the uptake of heavy metals in plant including Cd and As (Feng et al., 2013) and reduced harmful effect of Cd on plants (Pedrero et al., 2008). Compared to Se fertilization to the soil, foliar application by-passes any interference due to soil chemistry and microbiology issues, ensuring a higher efficacy even with low volumes of Se solution (D'Amato et al., 2020). Eventually, it is important and fundamental to solve the problem by developing new variety with rich Se concentration but minimum Cd and As accumulations in rice grains, both in brown rice and milled rice.

The concentrations of Se, Cd and As in brown rice range from 0.025 to 7.500 mg kg^{-1}, 0.002–0.475 mg kg^{-1}, and 0.010–0.750 mg kg^{-1} in rice germplasms, respectively (Zhang et al., 2010; Wang et al., 2012; Kuramata et al., 2013; Zhang G. M. et al., 2018; Liu et al., 2019; Norton et al., 2019). All of them varied among different rice varieties (Zou et al., 2018), and belong to quantitative traits controlled by multiple genes (Zhang G. M. et al., 2018). Using molecular marker technology, some QTL controlling concentrations of Se, Cd and As in brown rice have been identified. For Se concentration, two QTL were identified on chromosome 5 using an F_2 population derived from Xinfenghongmi and Minghui 100 (Zhang et al., 2010). For Cd, more than 35 QTL were detected on all 12 chromosomes using different populations (Li et al., 2020; Pan et al., 2021). For As concentration, three QTL were identified on chromosomes 6 and 8 using an F_2 population for methylated arsenic (Kuramata et al., 2013), and two QTL for As accumulation were identified using an doubled haploid population (Zhang et al., 2008). Till now, only six genes including *OsHMA2*, *OsHMA3*, *OsLCT1*, *CAL1*, *OsCd1* and *OsNAMP5* have been cloned for Cd concentration in brown rice (Chen et al., 2019; Yan et al., 2019; Chang et al., 2020).

In recent years, a new approach named genome-wide association study (GWAS) was used for identifying QTL for rice complex traits. It is based on rice germplasm and linkage disequilibrium (LD). Using this strategy, some QTL for Se, Cd and As have been detected, and some candidate genes were identified. Liu *et al.* (2019) identified 17 and 22 QTL for Cd and As concentrations in rice grains, respectively in 276 accessions using 416K SNP genotypes. Norton et al. (2019) detected 74 QTL for As concentration in rice grains in 266 landraces using 2M SNP genotypes. Among them, six QTL were stably expressed across different environments. Zhang G. M. et al. (2018) identified 3 and 62 QTL for Se and Cd concentrations in rice grains, respectively, in 698 *xian* and *geng* germplasms using 27K SNP genotypes. Further, 198 candidate genes for Cd concentration were detected using 2.9M SNP genotypes. Yan et al. (2019) detected 12 QTL for Cd accumulation in rice grains using 3M SNP genotypes in 127 rice cultivars, and then successfully cloned *OsCd1* for Cd accumulation. Qiu T. C. et al. (2021) identified 9 QTL for Se concentration in milled rice in 207 accessions using 50K SNPs.

In the present study, Se, Cd and As concentrations in brown and milled rice were measured under normal and Se fertilizer treatment conditions, in a diverse panel of 307 accessions from 3K RGP (Wang et al., 2018). Based on analysis of distribution and relationship of the three elements in brown and milled rice under the two conditions, QTL for Se, Cd and As concentrations in brown and milled rice were identified by GWAS using high-quality SNPs from 3K GRP. Then, gene-based association studies were performed in some important QTL regions using all available SNPs from RFGB v2.0 database (Wang et al., 2020). Finally, haplotype analysis was used to identify the candidate genes and novel alleles. The aim in this study is to explore effects of Se fertilizer application on concentrations of Se, Cd and As and the relationships among them both in brown and milled rice at phenotypic level and genetic mechanisms underlying the

concentrations of them. Our results will help us better understand the genetic basis of Se, Cd and As concentrations in brown and milled rice and provide candidate genes and favorable alleles for rice breeding of high Se, low Cd and As concentration by marker-assisted selection (MAS).

Materials and methods

Association mapping panel

In our previous study, a total of 1,016 rice accessions from 3K RAP (Wang et al., 2018) were identified to flower normally in Jinghzou of Hubei province, China (Qiu et al., 2020). To avoid the negative effects of lodging and large difference of heading date on the concentrations of Se, Cd and As in grains, 307 rice accessions with plant height below 140 cm and heading date within 10 days difference were selected as materials in the present study. They were from 33 countries worldwide (Supplementary Table S1).

Field experiment and evaluation of Se, Cd and As concentrations in grains

All accessions were grown in the summer seasons of 2019 and 2020 at the Experimental Farm of College of Agriculture, Yangtze University, Jingzhou City (30.2°N, 112.7°E), Hubei province, China. The Se, Cd and As concentrations in the soil were averagely 0.33, 0.27 and 7.15 mg kg^{-1} in the 2 years. Field experiment was carried out using a randomized complete design. Each accession was planted in three rows and ten plants per row with the density of 17 cm between plants and 20 cm between rows for two replications in two conditions, i.e., normal condition where field management was followed as local farmers' practices, and Se treatment condition where Se fertilizer in diluent with 3.30 g Na_2SeO_3 dissolved in 40 kg water per 666.7 m^2 was sprayed to rice plants at booting stage and other field management was completely same as the normal condition. At mature stage, seeds were bulk-harvested from each plot, and air-dried and stored for more than 3 months at room temperature before assessing of concentrations of selenium (Se), cadmium (Cd) and arsenic (As) in brown and milled rice.

Seeds of each plot were de-hulled into brown rice, and then polished into milled rice according to the method described previously (Qiu et al., 2015). Brown and milled rice were ground into flour for assessing the three element concentrations by atomic fluorescence spectrometry (AFS-230E, Beijing Jinsuokun Technology Developing Co. Ltd, China) after microwave digestion (ECH-2, Shanghai Sineo Microwave Chemistry Technology Co. Ltd, China) following the methods described as Zhang G. M. et al. (2018) and Liu et al. (2019). The evaluated traits included Se concentration in brown rice and milled rice in normal condition (SBN, SMN), Cd concentration in brown rice and milled rice in normal condition (CBN, CMN), As concentration in brown rice and milled rice in normal condition (ABN, AMN), Se concentration in brown rice and milled rice in treatment condition (SBT, SMT), Cd concentration in brown rice and milled rice in treatment condition (CBT, CMT), and As concentration in brown rice and milled rice in treatment condition (ABT, AMT). To reflect effects of Se fertilizer addressing on concentrations of Se, Cd and As in rice grains, ratio of Se concentration in brown rice (RSB) and milled rice (RSM) under Se treatment to normal conditions, ratio of Cd concentration in brown rice (RCB) and milled rice (RCM) under Se treatment to normal conditions, and ratio of As concentration in brown rice (RAB) and milled rice (RAM) under Se treatment to normal conditions were calculated. All concentrations were independently tested three times per sample, and the mean value was used as the trait value for data analysis. Then, Statistica 5.5 was used for analysis of statistical parameters and correlations among different traits (Qiu et al., 2017).

SNP data

The 4.8 M SNP dataset of 3K RGP was download from the Rice SNP-Seek Database (http://snp-seek.irri.org) (Alexandrov et al., 2015). For structure analysis, PLINK 1.9 was used to obtain 73,162 independent SNPs (Purcell et al., 2007). The parameter was set as MAF >0.05 and missing data ratio <0.10. ADMIXTURE program and GCTA software were used to identify genetic structure and principal component (Alexander et al., 2009). For GWAS, PLINK 1.9 was also used to identify high quality SNPs with similar parameter setting as MAF >0.05 and missing data ratio <0.20.

QTL detection for concentrations of three elements by GWAS

GWAS was conducted to identify QTL for Se, Cd and As concentrations, and ratio of Se, Cd and As concentrations of treatment to normal conditions in brown and milled rice using 405,150 high quality SNPs and trait values of the 307 accessions. The SVS software package v8.4.0 was used for association analysis (Qiu X. et al., 2021). The EMMAX (Efficient Mix-Model Association eXpedited) implementation of the single-locus mixed linear model was fitted to the marker dataset (Vilhjálmsson and Nordborg, 2013). The threshold was set as $p < 1.0E-05$. Since the LD decay of 3K RGP was about 200 kb (Wang et al., 2018), QTL located in the region of 100 kb each side at the peak SNPs and containing more than ten SNP with p value above threshold were considered as one QTL.

TABLE 1 Concentrations of Se, Cd and As in brown and milled rice in normal and treatment conditions and their ratios of treatment to normal conditions in 307 accessions.

Trait[a]	2019			2020		
	Mean ± SD	Range	CV (%)	Mean ± SD	Range	CV (%)
SBN (mg kg^{-1})	0.041±0.045	0.002–0.452	109.87	0.091±0.232	0.013–1.994	254.54
SMN (mg kg^{-1})	0.020±0.021	0–0.209	107.24	0.054±0.196	0.006–1.775	360.53
CBN (mg kg^{-1})	0.007±0.031	0.001–0.430	417.94	0.012±0.091	0.004–0.707	74.59
CMN (mg kg^{-1})	0.003±0.002	0–0.025	76.74	0.009±0.003	0.003–0.018	35.10
ABN (mg kg^{-1})	0.419±0.125	0.024–1.037	29.83	0.173±0.042	0.082–0.335	24.43
AMN (mg kg^{-1})	0.240±0.087	0.106–0.680	36.11	0.125±0.030	0.059–0.199	24.14
SBT (mg kg^{-1})	0.161±0.150	0.029–1.014	92.90	0.265±0.191	0.038–1.113	72.08
SMT (mg kg^{-1})	0.100±0.982	0–0.769	99.11	0.180±0.138	0–0.533	76.55
CBT (mg kg^{-1})	0.006±0.007	0.001–0.069	122.66	0.010±0.004	0.003–0.019	35.90
CMT (mg kg^{-1})	0.003±0.002	0.001–0.009	50.68	0.008±0.003	0–0.016	39.39
ABT (mg kg^{-1})	0.374±0.252	0.141–3.392	57.31	0.179±0.039	0.076–0.300	21.63
AMT (mg kg^{-1})	0.242±0.084	0.022–0.512	34.90	0.133±0.032	0–0.239	24.12
RSB	5.587±4.920	0.304–31.57	88.08	7.878±7.792	0.222–32.298	98.91
RSM	6.991±6.901	0.034–43.016	98.71	9.866±11.947	0.879–59.369	121.09
RCB	1.541±2.395	0.209–19.222	155.38	1.185±0.750	0.206–3.374	63.27
RCM	1.557±1.425	0.310–10.468	91.51	1.026±0.556	0.404–3.147	54.22
RAB	0.961±0.718	0.247–8.438	74.71	1.114±0.393	0.353–3.085	35.31
RAM	1.153±0.519	0.170–3.411	45.01	1.163±0.453	0.330–3.202	38.94

[a]SBN, se concentration in brown rice in normal condition; SMN, se concentration in milled rice in normal condition; CBN, cd concentration in brown rice in normal condition; CMN, cd concentration in milled rice in normal condition; ABN, as concentration in brown rice in normal condition; AMN, as concentration in milled rice in normal condition; SBT, se concentration in brown rice in treatment condition; SMT, se concentration in milled rice in treatment condition; CBT, cd concentration in brown rice in treatment condition; CMT, cd concentration in milled rice in treatment condition; ABT, as concentration in brown rice in treatment condition; AMT, as concentration in milled rice in treatment condition; RSB, ratio of Se concentration in brown rice; RSM, ratio of Se concentration in milled rice; RCB, ratio of Cd concentration in brown rice; RCM, ratio of Cd concentration in milled rice; RAB, ratio of As concentration in brown rice; RAM, ratio of As concentration in milled rice.

Gene-based association mapping and haplotypes analysis

Gene-based association mapping and haplotype analysis were conducted to detect candidate genes and favorable alleles in the important QTL regions. QTL regions identified for the same trait in both years and/or for more than two different traits were considered as important QTL. QTL with peak SNP having lowest p value in the regions were further analyzed. Firstly, all genes located within 100 kb up and down the peak SNP were retrieved from the Rice Genome Annotation Project (http://rice.plantbiology.msu.edu/). Secondly, all SNPs located on these genes were retrieved from 32 M SNPs data generated from 3K RGP in the RFGB v2.0 database (Wang et al., 2020), and they were used for gene-based association analysis. The threshold was also set as–$\log_{10}$(P) > 5.0 (p < 1.0E-05). Thirdly, for each QTL region, all significant SNPs located in promoter region, 5′ UTR region, non-synonymous SNPs in the exon regions, splice region in intron and 3' UTR region of each candidate gene were used for identifying haplotypes. Candidate genes were assigned by testing significant difference for relevant traits among major haplotypes (samples more than 8) for each important QTL by ANOVA with $p < 0.01$. When QTL containing cloned genes for element-related traits, the cloned genes were considered as only candidate gene (s) of the target QTL.

Results

Phenotypic variations in the natural population

A wide range of concentrations of Se, Cd and As in both brown and milled rice under normal, treatment and ratio of treatment to normal conditions reflected the substantial genotypic variations associated with concentrations of three elements in the panel of accessions (Table 1). Phenotypic variations had a similar trend in 2 years. Take the data in 2019 as example, SMN, CMN and AMN amounted to 48.8%, 42.9% and 57.3% of SBN, CBN and ABN, respectively, indicating that concentrations of the three elements in milled rice (endosperm) accounted for approximately half of total

concentrations in brown rice. Averagely, SBT, SMT, CBT, CMT, ABT and AMT were 3.9, 5.0, 0.9, 1.0, 0.9, and 1.0 times as much as SBN, SMN, CBN, CMN, ABN and AMN, respectively, indicating that Se fertilizer treatment substantially improved Se concentrations but no obvious changes in concentrations of Cd and As both in brown and milled rice. The ratios of treatment to normal conditions for the three element concentrations in milled rice were higher than those in brown rice, suggesting that Se-treatment accelerated accumulation of the three elements in endosperm. This was also reflected by the fact that SMT, CMT and AMT amounted to 62.1%, 50.0% and 64.7% of SBT, CBT and ABT, respectively, meaning that Se-treatment much improved the proportions of the three elements in milled rice as compared with those in normal condition.

Among the 307 accessions, three accessions (IRIS_313–8208, IRIS_313–8856 and IRIS_313–11968) were consistently identified to have high Se concentration, and low Cd and As concentrations in milled rice in normal condition in 2 years (Supplementary Table S2). Specifically, IRIS_313–8208 averagely had 0.067 mg kg^{-1} SMN, 0.005 mg kg^{-1} CMN and 0.161 mg kg^{-1} AMN; IRIS_313–8856 averagely had 0.108 mg kg^{-1} SMN, 0.004 mg kg^{-1} CMN and 0.189 mg kg^{-1} AMN; and IRIS_313–11968 averagely had 0.044 mg kg^{-1} SMN, 0.004 mg kg^{-1} CMN and 0.130 mg kg^{-1} AMN. IRIS_313–10083 had high Se concentration, and low Cd and As concentrations in milled rice in Se treatment condition in 2 years, with average 0.333 mg kg^{-1} SMT, 0.003 mg kg^{-1} CMT and 0.139 mg kg^{-1} AMT. Another four accessions (CX115, IRIS_313–11039, IRIS_313–11197 and IRIS_313–11943) were consistently identified as high ratio of Se concentration and low ratios of Cd and As concentrations in milled rice (Supplementary Table S2).

Correlation of concentrations in the panel of accessions

We analyzed correlations of three element concentrations in different parts under different treatments (Supplementary Table S3). For the same element between brown and milled rice under normal and treatment conditions, all three elements showed consistently highly significant positive correlations between brown and milled rice both in 2 years, suggesting accessions with high concentrations of three elements in brown rice would had high concentrations in milled rice. Se and Cd concentrations had no significant correlation both in brown and milled rice in the 2 years under normal and treatment conditions except for a highly positive correlation in milled rice in 2020 under treatment condition, suggesting that Se was almost independent of Cd in milled and brown rice under normal condition. Se and As had highly significant negative correlations in brown and milled rice only in 2019 under normal condition, and highly significant negative correlation in milled rice under treatment condition in 2 years, indicating that Se was antagonistic to As in milled rice. There were no significant correlations between Cd and As in brown and milled rice under normal and treatment conditions in 2 years except that in milled rice in 2020 under treatment, indicating that Cd and As are mostly independent. Correlations of the same elements Se, Cd and As in the same parts were all no significant between normal and treatment conditions in brown rice in 2 years. In milled rice, however, correlation of Se and As was significant positive between normal and treatment conditions in 2019, indicating that Se and As were had consistent responses to Se treatment in milled rice than brown rice among different accessions.

Detection of QTL by GWAS

Three main subpopulations were identified by admixture analysis and principal component analysis (Supplementary Figure S1), including *xian* (*indica*, 138), *geng* (*japonica*, 132), *aus* (32). In addition, five accessions belong to *admix*. Using 405,150 high quality of 307 accessions, a total of 237 QTL were identified for 16 traits, ranging from four for AMT to 23 for RCB and RAB. Among them, three QTL were identified in both years (Supplementary Table S4; Figure 1; Supplementary Table S2). None QTL was detected for SMT and CMT.

In normal condition, 87 QTL were detected for six traits (Supplementary Table S4; Figure 1; Supplementary Table S2). For Se concentrations, 21 QTL for SBN and 18 for SMN were detected on all 12 chromosomes in 2 years, and explained phenotypic variations of 1.00%–18.69%. Among them, *qSMN11.1* was stably detected in both years. For Cd concentrations, 17 QTL for CBN and 13 QTL for CMN were detected on all 12 chromosomes with average phenotypic variation of 7.38% and a range from 1.01% to 41.35% in 2 years. Of them, two QTL (*qCBN5* and *qCBN9*) were stably detected in both years. For As concentration, nine QTL for each of ABN and AMN were identified on 12 chromosomes except chromosome 10 in 2019, accounting for phenotypic variation ranging from 1.01% to 23.84%.

A total of 55 QTL were identified for concentrations of three elements in treatment condition (Supplementary Table S4; Figure 1; Supplementary Table S2). Ten SBT QTL were detected on chromosomes 1, 2, 5, 6, 8–10, 12, with an average phenotypic variation of 11.12%. Nineteen CBT QTL were detected on all 12 chromosomes in 2019 with 1.45%–12.76% of phenotypic variations. For As concentration, 22 and 4 QTL for ABT and AMT were identified in 2019 and 2020, respectively. They explained phenotypic variation from 1.21% to 13.08%.

For ratio of concentrations of three elements under Se treatment to normal conditions, a total of 95 QTL were

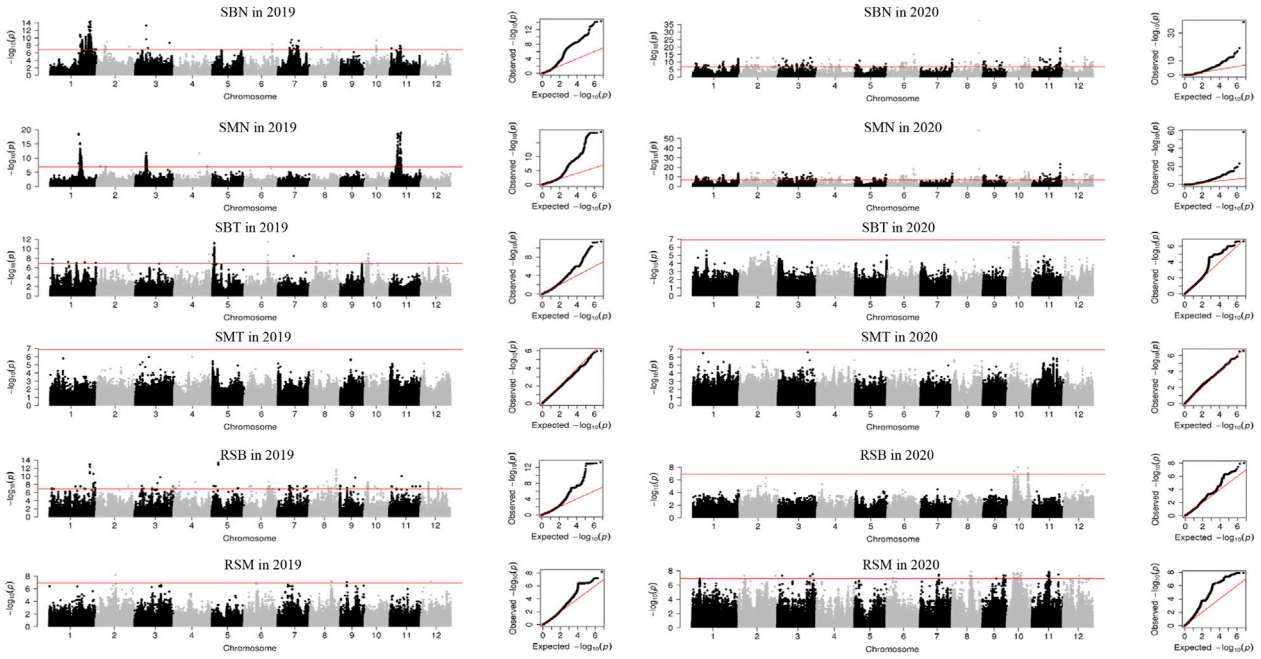


FIGURE 1
Manhattan plots and QQ plots of GWAS for Se concentration in 2019 and 2020 in the 307 rice accessions. SBN, Se concentration in brown rice in normal condition; SMN, Se concentration in milled rice in normal condition; SBT, Se concentration in brown rice in treatment condition; SMT, Se concentration in milled rice in treatment condition; RSB, ratio of the Se concentration in brown rice; RSM, ratio of the Se concentration in milled rice.

identified for RSB, RSM, RCB, RCM, RAB and RAM (Supplementary Table S4; Figure 1; Supplementary Table S2). Nineteen and 13 QTL for RSB and RSM were detected on all 12 chromosomes, respectively, explaining phenotypic variation of 1.79%–40.14%. Nineteen and 22 QTL were identified for RCB and RCM, respectively, with mean phenotypic variation of 4.14%, ranging from 1.21% to 12.76%. Twenty-three and five QTL were found for RAB and RAM on all chromosomes with phenotypic variation of 0.21%–18.37%.

Genetic overlap among QTL underlying different elements in grains

Comparison of QTL located in the same or adjacent regions can reveal genetic relationships of the concentration between brown and milled rice and among three elements. Total of 46 regions harboring QTL for different traits were detected (Supplementary Table S5). Among them, 11 regions were identified for concentrations of the same element between brown rice and milled rice, including 10 harboring *qSBN2.2* and *qSMN2.1*, *qSBN2.3* and *qSMN2.2*, *qSBN3.1* and *qSMN3.1*, *qSBN4.2* and *qSMN4.2*, *qSBN5.2* and *qSMN5.2*, *qSBN6* and *qSMN6*, *qSBN7.2* and *qSMN7*, *qSBN8.2* and *qSMN8*, *qSBN10* and *qSMN10*, and *qSBN11.2* and *qSMN11.2* for Se concentration in brown and milled rice under normal conditions, and one harboring *qAMN2* and *qABN2* for As concentration in brown and milled rice under normal condition. It was indicated that pleiotropy is most probably responsible for association of Se and As concentrations between brown and milled rice. Seventeen regions were identified for concentrations of two or three elements, including three harboring *qSMN1.2* and *qCBN1.2*, *qSBN8.2*, *qSMN8* and *qCBN8.2*, and *qSBT12* and *qCBT12* affecting Se and Cd concentrations in normal or treatment conditions, 10 harboring *qCBN2*, *qABN2* and *qAMN2*, *qCBT2.1* and *qABT2.1*, *qCBN3*, *qCBT3.1* and *qABT3.2*, *qCBT3.2* and *qABT3.3*, *qCBT4.1*, *qABT4.1* and *qAMT4*, *qCBT7* and *qABT7*, *qCBT8.2* and *qABT8.3*, *qCBT9.1* and *qABT9.1*, *qCBT10* and *qABT10.1*, and *qCBN11* and *qAMN11* for Cd and As concentrations under normal or treatment condition, and four harboring *qSBT1.2*, *qCBT1.2* and *qABT1.2*, *qSBT5*, *qCBT5* and *qABT5*, *qSBT8*, *qCMN8* and *qABT8.1*, and *qSBT9*, *qCBT9.2* and *qABT9.2* affecting Se, Cd and As concentrations in brown rice under treatment or normal condition. Twenty regions were identified for QTL overlap between concentrations in normal or treatment conditions and concentration ratio of the three elements. It was indicated that partial genetic overlaps exist due to pleiotropy or tightly linkage for the three elements, especially Cd and As under treatment condition.

TABLE 2 Information of candidate genes in 14 important QTL regions.

QTL	Chr	Confidence interval (Mb)	MSU ID	Annotation	Trait affected
qRSB1.1	1	1.95–2.15	*LOC_Os01g04580*	Ser/Thr protein kinase, putative, expressed	RSB/RCM
			LOC_Os01g04590	expressed protein	RCM/SBT/RSB
qABT1.2	1	24.33–24.53	*LOC_Os01g42870*	transferase family protein, putative, expressed	ABT/SBT/CBT
			LOC_Os01g42909	hypothetical protein	ABT/CBT
			LOC_Os01g42950	protein kinase family protein, putative, expressed	ABT/CBT
qRSB1.4	1	42.31–42.51	*LOC_Os01g73040*	CBS domain-containing protein, putative, expressed	RSB/RCM/SBT
			LOC_Os01g73080	expressed protein	RSB/RCM/SBT
			LOC_Os01g73110	expressed protein	RSB/RCM/SBT
			LOC_Os01g73120	expressed protein	RSB/RCM/SBT
			LOC_Os01g73130	vacuolar ATP synthase, putative, expressed	RSB/RCM/SBT
			LOC_Os01g73140	ubiquitin-fold modifier 1 precursor, putative, expressed	RSB/RCM/SBT
			LOC_Os01g73150	expressed protein	RSB/RCM/SBT
			LOC_Os01g73170	peroxidase precursor, putative, expressed	RSB/RCM/SBT
			LOC_Os01g73250	abscisic stress-ripening, putative, expressed	RCM
qRAB3.1	3	13.20–13.40	*LOC_Os03g22840*	retrotransposon protein, putative, unclassified, expressed	RAB/ABT/RCB
qRSB4.1	4	5.22–5.42	*LOC_Os04g09880*	expressed protein	RSB/RCM
qABT5	5	1.69–1.89	*LOC_Os05g03934*	expressed protein	ABT/CBT
			LOC_Os05g03972	plant protein of unknown function domain containing protein, expressed	ABT/CBT/ RCB/RAB
			LOC_Os05g04000	expressed protein	ABT/CBT/ RCB/RAB
			LOC_Os05g04020	plant protein of unknown function domain containing protein, expressed	ABT/CBT/RAB
qRAB5.2	5	4.11–4.31	*LOC_Os05g07810*	universal stress protein domain containing protein, putative, expressed	RAB
qCBN5	5	8.95–9.15	*LOC_Os05g15960*	retrotransposon protein, putative, unclassified, expressed	CBN
qCBT7	7	21.02–21.22	*LOC_Os07g35290*	TKL_IRAK_DUF26-lc.10 - DUF26 kinases have homology to DUF26 containing loci, expressed	CBT/ABT/RAB
			LOC_Os07g35310	TKL_IRAK_DUF26-lc.12 - DUF26 kinases have homology to DUF26 containing loci, expressed	CBT/ABT/ RCB/RAB
qSMN8	8	24.64–24.84	*LOC_Os08g39120*	expressed protein	SMN/SBN
			LOC_Os08g39174	retrotransposon protein, putative, Ty1-copia subclass, expressed	SMN/SBN
			LOC_Os08g39180	OsWAK73 - OsWAK receptor-like protein kinase, expressed	SMN/SBN
qABT8.3	8	25.50–25.70	*LOC_Os08g40510*	KID-containing protein, putative, expressed	ABT/CBT
			LOC_Os08g40570	pyridoxamine 5%27-phosphate oxidase family protein, putative, expressed	ABT/CBT/RCB
qRAB9.2	9	20.08–20.28	*LOC_Os09g34200*	OsFBX338 - F-box domain containing protein, expressed	RAB/CBT/ ABT/RCB
			LOC_Os09g34214	UDP-glucoronosyl and UDP-glucosyl transferase domain containing protein, expressed	RAB/CBT/ ABT/RCB
			LOC_Os09g34230	UDP-glucoronosyl/UDP-glucosyl transferase, putative, expressed	RAB/CBT/ ABT/RCB
			LOC_Os09g34250	UDP-glucoronosyl and UDP-glucosyl transferase domain containing protein, expressed	RAB/CBT/ ABT/RCB
			LOC_Os09g34270	UDP-glucoronosyl and UDP-glucosyl transferase domain containing protein, expressed	RAB/CBT/ ABT/RCB
qSMN10	10	15.95–16.15	*LOC_Os10g30790*	inorganic phosphate transporter, putative, expressed	SMN/SBN/RAB
			LOC_Os10g30800	expressed protein	SMN/SBN/RAB
qSMN11.2	12	25.94–26.14	*LOC_Os11g43140*	expressed protein	SMN/SBN
			LOC_Os11g43150	transposon protein, putative, CACTA, En/Spm sub-class, expressed	SMN/SBN
			LOC_Os11g43160	ransposon protein, putative, CACTA, En/Spm sub-class	SMN/SBN
			LOC_Os11g43170	hypothetical protein	SMN/SBN

SBN, se concentration in brown rice in normal condition; SMN, se concentration in milled rice in normal condition; CBN, cd concentration in brown rice in normal condition; SBT, se concentration in brown rice in treatment condition; CBT, cd concentration in brown rice in treatment condition; ABT, as concentration in brown rice in treatment condition; RSB, ratio of Se concentration in brown rice; RCB, ratio of Cd concentration in brown rice; RCM, ratio of Cd concentration in milled rice; RAB, ratio of As concentration in brown rice. The QTL listed in Supplementary Table S4.

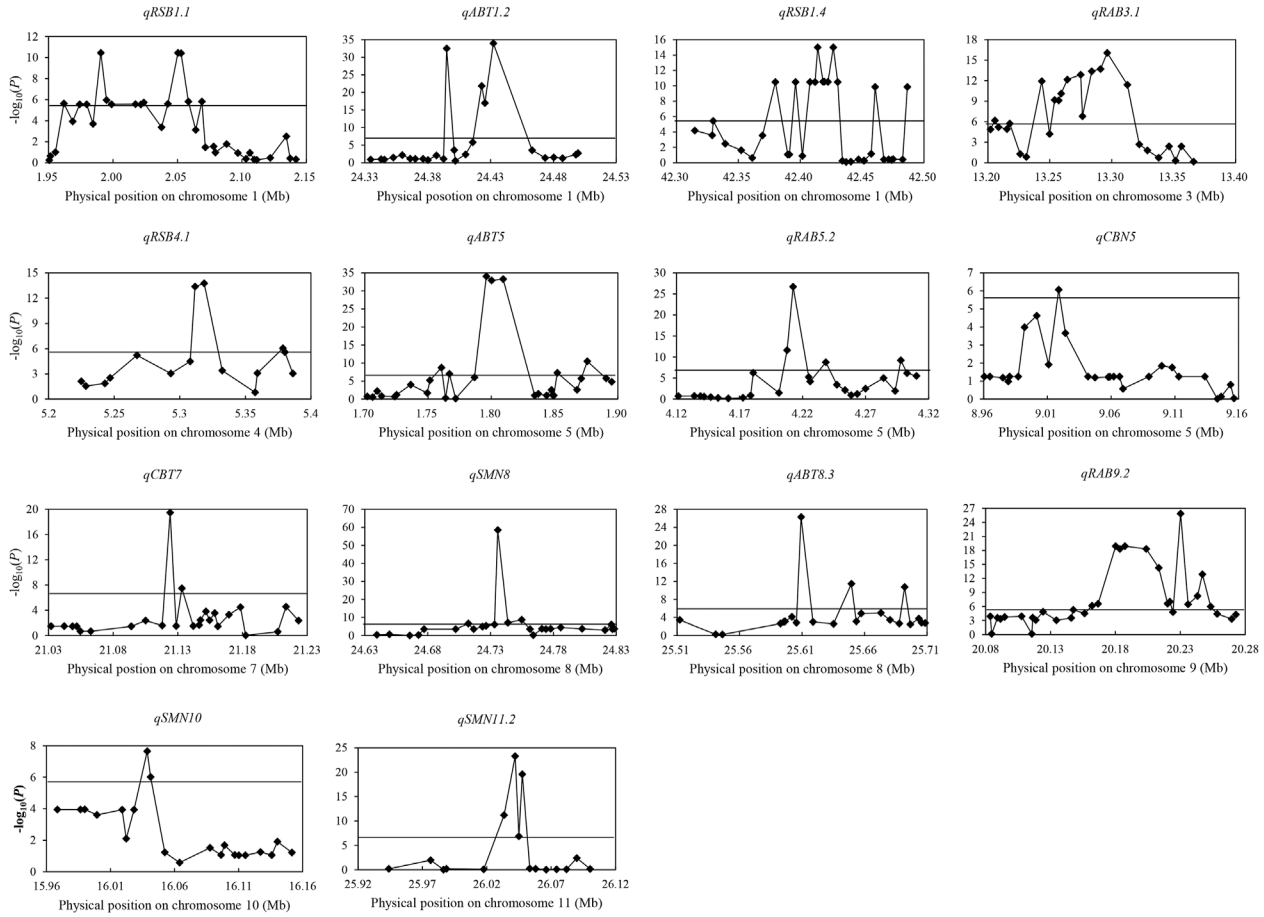


FIGURE 2
Manhattan plots of gene-based association analysis in 14 important QTL regions. The rhomb-shaped point represents annotation genes in the candidate region; the horizontal line represents the threshold.

Candidate gene analysis for 14 important QTL

Among 237 QTL for concentrations of three elements, 14 important QTL were identified in both years and/or had effect on more than two different traits. A total of 40 candidate genes and 43 favorable alleles were identified by gene-based association analysis and haplotype analysis (Table 2; Supplementary Table S6; Figure 2).

For *qRSB1.1* in 1.95–2.15 Mb on chromosome 1, 1,598 SNPs were found in 34 genes. Among them, 14 genes were identified. Significant differences were identified in RSB and RCM in 2019 between different haplotypes for two candidate genes (*LOC_Os01g04580* and *LOC_Os01g04590*). Also, there was significant difference in SBT in 2019 between two haplotypes in *LOC_Os01g04590*. No haplotype was found for the rest 12 genes.

For *qABT1.2* in the region of 24.33–24.53 Mb on chromosome 1, 449 SNPs were found in 25 genes, and five candidate genes were identified. Significant differences were detected in ABT and CBT in 2019 between different haplotypes in candidate genes *LOC_Os01g42870*, *LOC_Os01g42909* and *LOC_Os01g42950*. Besides, there was significant difference in SBT in 2019 between two haplotypes in *LOC_Os01g42870*. No haplotype was found for the other two genes (*LOC_Os01g42920* and *LOC_Os01g42934*).

For *qRSB1.4* in 42.31–42.51 Mb on chromosome 1, 1,137 SNPs were found in 34 genes, and 13 candidate genes were identified. Significant differences in RSB, SBT and RCM in 2019 were detected between two haplotypes in *LOC_Os01g73040*, *LOC_Os01g73080*, *LOC_Os01g73110*, *LOC_Os01g73120*, *LOC_Os01g73130*, *LOC_Os01g73140*, *LOC_Os01g73150* and *LOC_Os01g73170*. Besides, there was significant difference in RCM in 2019 between two haplotypes in *LOC_Os01g73250*. No

haplotype was found for *LOC_Os01g72980*, *LOC_Os01g73100*, *LOC_Os01g73160* and *LOC_Os01g73310*.

For *qRAB3.1* in the region of 13.20–13.40 Mb on chromosome 3, 305 SNPs were found in 26 genes, and 14 candidate genes were identified. Significant differences were detected in RAB, ABT and RCB in 2019 between different haplotypes in *LOC_Os03g22840*. There was no significant difference in any trait between different haplotypes or no haplotype available in the rest 13 genes.

For *qRSB4.1* in the region of 5.22–5.42 Mb on chromosome 4, 818 SNPs were found in 15 genes, and five candidate genes were identified. Significant differences were detected in RSB and RCM in 2019 between two haplotypes in *LOC_Os04g09880*. There was no significant difference in any trait between different haplotypes or no haplotype identified in the rest four genes.

For *qABT5* in the region of 1.69–1.89 Mb on chromosome 5, 1,086 SNPs were found in 29 genes, and 11 candidate genes were identified. Significant differences were detected in ABT and CBT in 2019 between different haplotypes in *LOC_Os05g03934*, *LOC_Os05g03972*, *LOC_Os05g04000* and *LOC_Os05g04020*. Besides, significant differences in RCB and RAB in 2019 were found between two haplotypes of each of *LOC_Os05g03972* and *LOC_Os05g04000*, and there was also significant difference in RAB in 2019 between two haplotypes in *LOC_Os05g04020*. There was no significant difference in any trait between different haplotypes or no haplotype available in the rest seven genes.

For *qRAB5.2* in the region of 4.11–4.31 Mb on chromosome 5, 1,508 SNPs were found in 26 genes, and nine candidate genes were identified. Significant difference was detected in RAB in 2019 between two haplotypes in *LOC_Os05g07810*. No haplotype was found in the rest eight genes.

For *qCBN5* in the region of 8.95–9.15 Mb on chromosome 5, 596 SNPs were found in 28 genes, and only one candidate gene (*LOC_Os05g15960*) was identified. Significant difference was detected in CBN in 2019 between the two haplotypes.

For *qCBT7* in the region of 21.02–21.22 Mb on chromosome 7, 1,245 SNPs were found in 26 genes, and two candidate genes (*LOC_Os07g35290* and *LOC_Os07g35310*) were identified. Significant differences were detected in CBT, ABT and RAB in 2019 between two haplotypes in both the two genes. Besides, there was significant difference in RCB in 2019 between two haplotypes in *LOC_Os07g35310*.

For *qSMN8* in the region of 24.64–24.84 Mb on chromosome 8, 629 SNPs were found in 25 genes, and seven candidate genes were identified. Significant differences were detected in SMN and SBN in 2020 between two haplotypes in *LOC_Os08g39120*, *LOC_Os08g39174* and *LOC_Os08g39180*. No haplotype was found in *LOC_Os08g39150*, *LOC_Os08g39160*, *LOC_Os08g39170* and *LOC_Os08g39280*.

For *qABT8.3* in the region of 25.50–25.70 Mb on chromosome 8, 620 SNPs were found in 22 genes, and three candidate genes (*LOC_Os08g40470*, *LOC_Os08g40510* and *LOC_Os08g40570*) were identified. Significant differences were detected in ABT and CBT in 2019 between two haplotypes in *LOC_Os08g40510* and *LOC_Os08g40570*. Besides, significant difference was also found in RCB in 2019 between two haplotypes in *LOC_Os08g40570*. No haplotype was found in *LOC_Os08g40470*.

For *qRAB9.2* in the region of 20.08–20.28 Mb on chromosome 9, 1,162 SNPs were found in 32 genes, and 15 candidate genes were identified. Significant differences were detected in RAB, ABT, CBT and RCB in 2019 between different haplotypes in *LOC_Os09g34200*, *LOC_Os09g34214*, *LOC_Os09g34230*, *LOC_Os09g34250* and *LOC_Os09g34270*. There was no significant difference in any trait between different haplotypes or no haplotype available in the rest 10 genes.

For *qSMN10* in the region of 15.95–16.15 Mb on chromosome 10, 807 SNPs were found in 21 genes, and two candidate genes (*LOC_Os10g30790* and *LOC_Os10g30800*) were identified. Significant differences were detected in SBN and SMN in 2020, and RAB in 2019 between two haplotypes in both the two genes.

For *qSMN11.2* in the region of 25.94–26.14 Mb on chromosome 11, 376 SNPs were found in 16 genes, and four candidate genes (*LOC_Os11g43140, LOC_Os11g43150*, *LOC_Os11g43160* and *LOC_Os11g43170*) were identified. Significant differences were detected in SMN and SBN in 2020 between two haplotypes in all four candidate genes.

Discussion

Distribution characteristics of three elements in rice grains and Se-treatment effect on their concentrations

Most Chinese people are likely to consume white rice or polished rice which is produced from brown rice. Although rice bran only accounts for 6–10% of the total weight of rice (Wang et al., 2017), it contains 64% of important nutrients or toxic elements (Yang et al., 2004). Our research showed that the levels of toxic elements such as As and Cd, and nutritional element Se have a good linear relationship between rice bran and polished rice, and most of them are enriched in rice bran as indicated that SBN, CBN and ABN accounted for almost more than half of those in the whole grain. Relatively speaking, the proportion of Se concentration in polished rice to the whole grain is higher than that of Cd. The differences of element distribution in rice grains provide an effective way to remove most of toxic elements enriched in bran by polishing brown rice, thus better ensuring food security for people who consume the polished rice.

It was reported that the average Se concentration in rice grains is 0.025 mg kg^{-1} in different regions of China, which

results in the Se deficiency of inhabitants in China because the rice Se is the main source of Se in the Chinese diet (Chen et al., 2002). Biofortification of rice using Se fertilizers not only increases grain Se concentration, but also reduces grain Cd concentration (Tan et al., 2014). This was supported by our finding that Se fertilizer treatment substantially improved Se concentrations but had no obvious changes in concentrations of Cd and As both in brown and milled rice in this study (Table 1). It is the first report in the present study that the reactions of Cd and As in rice grain to Se fertilizer treatment varied largely among rice accessions. Take the concentrations in milled rice as example, 26.51% and 20.11% accessions significantly decreased in Cd and As concentrations (RCM and RAM below 0.80) after Se treatment, 27.71% and 15.64% accessions significantly increased in Cd and As concentrations (RCM and RAM above 1.5), and 45.78% and 64.25% accessions were insensitive to Se treatment in Cd and As concentrations (RCM and RAM between 0.8 and 1.5), respectively. Thus, screening varieties whose Cd and As concentrations are insensitive to Se treatment or even decreased in milled rice after Se treatment would provide an effective way of biofortification for rice to enhance Se concentration and avoid enrichment of toxic elements through spraying Se fertilizer on leaf.

Comparisons of important QTL regions with ones previously reported and their candidate gene inference

In the present study, a total of 14 important QTL regions was identified in both years and/or had effect on more than two different traits (Table 2). Most of them were identified in the same or adjacent regions with ones previously reported (Table 2; Supplementary Table S5). For example, *qRSB1.1*, *qSBT1.1* and *qRCM1.1* at the peak position of 2.05 Mb, *qRSB1.4*, *qSBT1.3* and *qRCM1.2* in the region of 42.25–42.60 Mb on chromosome 1 were identified in the same regions of *qCd1-2* for grain Cd concentration (Zhang G. M. et al., 2018) and *qGCD8* for grain Cd concentration (Liu et al., 2019), and *OsLCD* for Cd tolerance and accumulation (Chen et al., 2019). *qRAB3.1*, *qABT3.1* and *qRCB3.1* in the region of 13.16–13.30 Mb on chromosome 3 were detected together with *qGAS15* for grain As accumulation (Liu et al., 2019). *qRCM4.1*, and *qRSB4.1* in the 5.15–5.32 region on chromosome 4 were found near *qCd4-3* for grain Cd concentration (Zhang G. M. et al., 2018). *qABT5*, *qCBT5*, *qRCB5.1* and *qRAB5.1* in the region of 1.79–1.80 Mb, and *qRAB5.2* in the region of 4.21–4.33 Mb on chromosome 5 were identified in the same regions of *qSe5-1* for grain Se concentration (Zhang et al., 2010), *qCd5-1* for grain Cd concentration (Zhang G. M. et al., 2018), *OsMTP1* for Cd translocation, and *OsZIP6* for Cd transport (Chen et al., 2019). *qCBT7*, *qABT7*, *qRCB7.2* and *qRAB7* in the peak position of 21.12 Mb on chromosome 7 were detected near *qGAS13* for grain As accumulation (Liu et al., 2019) and *qCd7-2* for grain Cd accumulation (Zhang G. M. et al., 2018). *qABT8.3*, *qCBT8.2*, *qRCB8.3* and *qRAB8.3* at the peak position of 25.60 Mb on chromosome 8 were mapped together with *qCd8-3* for grain Cd concentration (Zhang G. M. et al., 2018). *qRAB9.2*, *qABT9.2*, *qCBT9.2* and *qRCB9.2* in the region of 20.18–20.19 Mb on chromosome 9 were detected near *qRCd-5* for grain Cd accumulation (Pan et al., 2021). Whether the QTL identified in this study are allelic to the previously reported QTL or genes need to be validated by fine-mapping and transgenic strategy.

The candidate genes were further inferred by bioinformatics and gene expression at filling or milk stage of grain. Heavy metal accumulation in rice grains may occur at grain filling or milk stage. To identify the candidate genes for QTL involved in this process, we analyzed the expression pattern of candidate genes for each QTL using the RNA-seq database from MBKBASE. For three candidate genes of *qABT1.2*, *LOC_Os01g42909*, encoding hypothetical protein, was specifically expressed at milk stage of embryo and endosperm (Supplementary Figure S3). For *qRSB1.4*, *LOC_Os01g73040* and *LOC_Os01g73130* encode CBS domain-containing protein and vacuolar ATP synthase, respectively. A rice gene (*OsCBSX4*) with CBS domain involved in heavy metal tolerance (Singh et al., 2012). Besides, some heavy metal ATPase, OsHMA2 and OsHMA3, had function on Cd concentrations in grains (Yan et al., 2019). *LOC_Os01g73140* encodes precursor of UFM1 (ubiquitin-fold modifier 1) and is highly expressed in the milk stage of grain and filling stage of endosperm (Supplementary Figure S4). UFM1 protein functions as new post-translational UBLs (ubiquitin-like proteins) with similar structure and regulatory mechanism with ubiquitin. Although the function of UFM1 has not been reported in plants, the human UFM1 participates in the ER (endoplasmic reticulum) stress response that may involve in the folding and transport of protein (Kang et al., 2007; Li et al., 2018), suggesting that *LOC_Os01g73140* may involve in the regulation of element content. So, *LOC_Os01g73040*, *LOC_Os01g73130* and *LOC_Os01g73140* are likely candidate genes for *qRSB1.4*. Of three candidate genes for *qSMN8*, *LOC_Os08g39120* encodes an unknown expressed protein, with higher expression level in the milk stage of grain and filling stage of endosperm compared with the other two genes (Supplementary Figure S5), suggesting that *LOC_Os08g39120* is the most likely candidate gene for *qSMN8*. For *qRAB9.2*, five candidate genes were identified, and *LOC_Os09g34200* encodes F-box domain proteins. In previous study, many F-box proteins were highly expressed at Cd stress (Zhou et al., 2012; Chen et al., 2014), implying that they may involve in Cd transport in plants and affect Cd accumulation. Another candidate gene, *LOC_Os09g34230*, encoding UDP-glucoronosyl/UDP-glucosyl transferase protein, is highly expressed at milk stage of grain and filling stage of endosperm, as well as mature stage of grain (Supplementary Figure S6), indicating that *LOC_Os09g34230* plays important

roles in the rice grain ripening process. UDP-glucosyl transferase has been found to be involved in biologic and abiotic stress responses in wheat, *Arabidopsis* and *Rhazya stricta* (Hajrah et al., 2017; Sharma et al., 2018; Rao et al., 2019). The candidate gene *LOC_Os10g30790* for *qSMN10*, encoding inorganic phosphate transporter, is expressed at milk stage of grain, filling stage of endosperm, mature stage of grain and especially highly expressed at mature stage of aleurone that exists in brown rice (Supplementary Figure S7). Importantly, the expression of phosphate transporter could increase arsenic tolerance in *Arabidopsis* (Feng et al., 2021), indicating that *LOC_Os10g30790* is the most likely candidate gene for *qSMN10*, which mediates the response to arsenic stress. Above most likely candidate genes will be validated using CRISPR-Cas9 and transgenic technology in future.

Application in rice breeding for Se enrichment with decreased Cd and As

In China, the general population in 72% of the total land area is facing a Se deficiency problem (Feng et al., 2013). Therefore, it is of great significance to develop rice variety with Se enrichment but minimum contents of heavy metals such as Cd and As in milled rice. Liu et al. (2020) reported that new elite varieties with enriched Se content in grains can be realized by pyramiding different main-effect QTL that considerably facilitated high Zn/Se enrichment while low Cd accumulation in grains. In the present study, among 142 QTL for concentrations of three elements detected in normal or treatment conditions (Supplementary Table S4), only 19 QTL (13.4%) were commonly detected for concentrations of Se and Cd, Se and As, and Se, Cd and As in the same or adjacent regions (Supplementary Table S5), indicating that most genetic loci underlying Se concentration are independent of those controlling Cd and As concentrations. Breeders could pyramid favorable alleles at different Se QTL and excluding undesirable alleles at Cd and As QTL, thus, simultaneously increase Se concentration and minimize Cd and As concentrations. Based on haplotype analysis of candidate genes in 14 important QTL regions identified in this study, Hap2 of candidate gene *LOC_Os08g39120* at *qSMN8*, *LOC_Os10g30790* at *qSMN10*, and Hap1 of *LOC_Os11g43140* or *LOC_Os11g43150*, or *LOC_Os11g43170* at *qSMN11.2* could increase Se concentration in both brown and milled rice. Three accessions, IRIS_313–8208, IRIS_313–8856 and IRIS_313–11968 identified with high Se and low Cd and As concentrations in milled rice carried 2, 3 and 1 favorable alleles of the candidate genes at the three loci underlying Se concentration, respectively (Supplementary Table S7). Meanwhile, the three accessions all carried favorable Hap 1 of *LOC_Os05g15960* at *qCBN5* which could decrease CBN. So, the three accessions could be used as donor parents to introgress then pyramid different favorable alleles at the three QTL for Se concentration and one QTL for Cd concentration by marker-assisted selection (MAS).

In most cases, Se deficiency can be corrected by the application of Se fertilizer into the soils or on rice leaf, which is termed as agronomic biofortification (Lyons et al., 2004). As indicated in this study, Se fertilizer treatment substantially improved Se concentrations but had no obvious changes in concentrations of Cd and As both in brown and milled rice, meaning as compared with Se, Cd and As in grains were insensitive to Se fertilizer treatment. Three important QTL regions (1.95–2.15 Mb and 42.31–42.51 Mb on chromosome 1, and 5.22–5.42 Mb on chromosome 4) were identified for both RSB and RCM in this study, and Hap1 of *LOC_Os01g04580* at *qRSB1.1*/*qRCM1.1* and *LOC_Os04g09880* at *qRSB4.1*/*qRCM4.1*, and Hap2 of *LOC_Os01g04590* at *qRSB1.1*/*qRCM1.1*, *LOC_Os01g73040*, *LOC_Os01g73130* and *LOC_Os01g73140* at *qRSB1.4*/*qRCM1.2* could enhance about 5.1–31.6 times of the Se concentration in brown rice and 5.4–10.5 times of Cd concentrations in milled rice after Se treatment. So, introgressing and pyramiding favorable genes for above three QTL could much improve Se concentration in grains after Se fertilizer treatment although Cd concentration could be increased to some extent due to synergistic effect of Se at these RSB QTL. Four accessions (CX115, IRIS_313–11039, IRIS_313–11197 and IRIS_313–11943) with high RSM and low RCM and RAM carried all favorable alleles at above QTL (Supplementary Table S7). Therefore, to develop higher Se-rich variety with more efficient response to Se treatment, above four accessions could be used as donor parent for introgressing and pyramiding of favorable alleles at *qRSB1.1*/*qRCM1.1*, *qRSB1.4*/*qRCM1.2* and *qRSB4.1*/*qRCM4.1* by MAS. It is worth noting that *qRSB3.2*/*qRSM3.1* for both RSB and RSM is only one locus which is independent of QTL for ratio of Cd and As concentrations. So, after mining favorable alleles at this locus using more diverse germplasms, the favorable alleles can be used for further enhancing Se concentrations in both brown and milled rice after Se treatment without interference of Cd and As in rice breeding for Se biofortification.

Finally, another efficient strategy to enhance Se concentration in rice grains is probably to pyramid favorable alleles of different kinds of QTL, i.e., QTL for Se enrichment identified in normal condition and QTL for Se sensitivity identified under Se fertilizer treatment by MAS. Thus, new variety pyramiding above two kinds of QTL will show high Se concentration in normal condition and more higher Se concentration after Se fertilizer application in some Se-deficient areas.

Conclusion

Large variations in concentrations of Se, Cd and As in grains existed in the panel of 307 rice accessions in normal and

Se treatment conditions. Se fertilizer treatment greatly improved Se concentrations but had no obvious changes in concentrations of Cd and As both in brown and milled rice. A total of 237 QTL were identified for Se, Cd and As concentrations in normal, Se treatment and ratio of treatment to normal conditions by GWAS. Most QTL for Se concentration is independent of those for Cd and As concentrations in view of only 13.4% QTL commonly detected in the same or adjacent regions. Forty-three favorable alleles were identified for 40 candidate genes in 14 important QTL regions. Pyramiding of favorable alleles at Se QTL and excluding undesirable alleles at Cd and As QTL, or combining favorable alleles at Se QTL detected in normal condition with the alleles at Se-sensitive QTL detected under Se treatment by MAS will facilitate development of rice variety with Se enrichment and minimum concentrations of Cd and As.

Data availability statement

Publicly available datasets were analyzed in this study. This data can be found here: http://snp-seek.irri.org.

Author contributions

JX and DX contributed to the conception of the study; WW, KC and BD collected the phenotypic data; FZ and DL analyzed the data; XQ wrote the original manuscript; JX revised the manuscript. The authors read and approved the final manuscript.

Funding

This work was funded by National Public Industry (Agriculture) Research Special Project (201303106), and Shenzhen Technology Research & Development (JCYJ20200109150713553).

Conflict of interest

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

The reviewer JW declared a past co-authorship with the authors FZ and JX to the handling editor.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1007896/full#supplementary-material

References

Akesson, A., Barregard, L., Bergdahl, I. A., Nordberg, G. F., Nordberg, M., and Skerfving, S. (2014). Non-renal effects and the risk assessment of environmental cadmium exposure. *Environ. Health Perspect.* 122, 431–438. doi:10.1289/ehp.1307110

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

Alexandrov, N., Tai, S., Wang, W., Mansueto, L., Palis, K., Fuentes, R. R., et al. (2015). SNP-Seek database of SNPs derived from 3000 rice genomes. *Nucleic Acids Res.* 43, D1023–D1027. doi:10.1093/nar/gku1039

Bolan, N. S., Makino, T., Kunhikrishnan, T., Kim, P. J., Ishikawa, S., Murakami, M., et al. (2013). Cadmium contamination and its risk management in rice ecosystems. *Adv. Agron.* 119, 183–273. doi:10.1016/B978-0-12-407247-3.00004-4

Chang, J. D., Huang, S., Konishi, N., Wang, P., Chen, J., Huang, X. Y., et al. (2020). Overexpression of the manganese/cadmium transporter OsNRAMP5 reduces cadmium accumulation in rice grain. *J. Exp. Bot.* 71, 5705–5715. doi:10.1093/jxb/eraa287

Chen, J., Zou, W., Meng, L., Fan, X., Xu, G., and Ye, G. (2019). Advances in the uptake and transport mechanisms and QTLs mapping of cadmium in rice. *Int. J. Mol. Sci.* 20, E3417. doi:10.3390/ijms20143417

Chen, L. C., Yang, F. M., Xu, J., Hu, Y., Hu, Q. H., Zhang, Y. L., et al. (2002). Determination of selenium concentration of rice in China and effect of fertilization of selenite and selenate on selenium content of rice. *J. Agric. Food Chem.* 50, 5128–5130. doi:10.1021/jf0201374

Chen, R., Guo, W., Yin, Y., and Gong, Z. H. (2014). A novel F-box protein CaF-box is involved in responses to plant hormones and abiotic stress in pepper (Capsicum annuum L.) *Int. J. Mol. Sci.* 15, 2413–2430. doi:10.3390/ijms15022413

Cox, D. N., and Bastiaans, K. (2007). Understanding Australian consumers' perceptions of selenium and motivations to consume selenium enriched foods. *Food Qual. prefer.* 18, 66–76. doi:10.1016/j.foodqual.2005.07.015

D'Amato, R., Regni, L., Falcinelli, B., Mattioli, S., Benincasa, P., Dal Bosco, A., et al. (2020). Current knowledge on selenium biofortification to improve the nutraceutical profile of food: A comprehensive review. *J. Agric. Food Chem.* 68, 4075–4097. doi:10.1021/acs.jafc.0c00172

Du, Y., Hu, X. F., Wu, X. H., Shu, Y., Jiang, Y., and Yan, X. J. (2013). Affects of mining activities on Cd pollution to the paddy soils and rice grain in Hunan province, Central South China. *Environ. Monit. Assess.* 185, 9843–9856. doi:10.1007/s10661-013-3296-y

Feng, H. Y., Li, X. Y., Sun, D., Chen, Y. S., Xu, G. H., Cao, Y., et al. (2021). Expressing. phosphate transporter PvPht2;1 enhances P transport to the chloroplasts and increases arsenic tolerance in *Arabidopsis thaliana*. *Environ. Sci. Technol.* 55, 2276–2284. doi:10.1021/acs.est.0c03316

Feng, R. W., Wei, C. Y., and Tu, S. X. (2013). The roles of selenium in protecting plants against abiotic stresses. *Environ. Exp. Bot.* 87, 58–68. doi:10.1016/j.envexpbot.2012.09.002

Gailer, J. (2009). Chronic toxicity of as(III) in mammals: The role of (GS)(2) AsSe(-). *Biochimie* 91, 1268–1272. doi:10.1016/j.biochi.2009.06.004

Hajrah, N. H., Obaid, A. Y., Atef, A., Ramadan, A. M., Arasappan, D., Nelson, C. A., et al. (2017). Transcriptomic analysis of salt stress responsive genes in Rhazya stricta. *PLoS One* 12, e0177589. doi:10.1371/journal.pone.0177589

Kang, S. H., Kim, G. R., Seong, M., Baek, S. H., Seol, J. H., Bang, O. S., et al. (2007). Two novel ubiquitin-fold modifier 1 (Ufm1)-specific proteases, UfSP1 and UfSP2. *J. Biol. Chem.* 282, 5256–5262. doi:10.1074/jbc.M610590200

Kuramata, M., Abe, T., Kawasaki, A., Ebana, K., Shibaya, T., Yano, M., et al. (2013). Genetic diversity of arsenic accumulation in rice and QTL analysis of methylated arsenic in rice grains. *Rice* 6, 3. doi:10.1186/1939-8433-6-3

Li, J., Yue, G., Ma, W., Zhang, A., Zou, J., Cai, Y., et al. (2018). Ufm1-Specific ligase Ufl1 regulates endoplasmic reticulum homeostasis and protects against heart failure. *Circ. Heart Fail.* 11doi, e004917. doi:10.1161/circheartfailure.118.004917

Li, W. X., Ouyang, L. J., Wang, W., Chen, X. R., Peng, X. S., He, H. H., et al. (2020). Advances in research on Cadmium accumulation characteristics and genetic regulation in rice. *J. Chin. Cereal Oil Assoc.* 35, 182–189. doi:10.3969/j.issn.1003-0174.2020.06.027

Liu, C., Ding, S., Zhang, A., Hong, K., Jiang, H., Yang, S., et al. (2020). Development of nutritious rice with high zinc/selenium and low cadmium in grains through QTL pyramiding. *J. Integr. Plant Biol.* 62, 349–359. doi:10.1111/jipb.12909

Liu, X., Chen, S., Chen, M., Zheng, G., Peng, Y., Shi, X., et al. (2019). Association study reveals genetic loci responsible for Arsenic, Cadmium and lead accumulation in rice grain in contaminated farmlands. *Front. Plant Sci.* 10, 61. doi:10.3389/fpls.2019.00061

Liu, Z. Y., Yang, J. X., Chen, G. Z., and Diao, J. M. (2007). Influence of arsenic in soil on the growth of different rice (*Oryza sativa* L.) varieties. *Ecol. Environ.* 16, 1700–1704. doi:10.16258/j.cnki.1674-5906.2007.06.043

Lyons, G. H., Lewis, J., Lorimer, M. F., Holloway, R. E., Brace, D. M., Stangoulis, J. C. R., et al. (2004). High-selenium wheat: Agronomic biofortification strategies to improve human nutrition. *Food Agric. Environ.* 2, 171–178.

Meplan, C., and Hughes, D. J. (2020). The role of selenium in health and disease: Emerging and recurring trends. *Nutrients* 12, 1049. doi:10.3390/nu12041049

Norton, G. J., Travis, A. J., Talukdar, P., Hossain, M., Islam, M. R., Douglas, A., et al. (2019). Genetic loci regulating arsenic content in rice grains when grown flooded or under alternative wetting and drying irrigation. *Rice* 12, 54. doi:10.1186/s12284-019-0307-9

Pan, C. Y., Ye, H. F., Zhou, W. Y., Wang, S., Li, M. J., Lu, M., et al. (2021). QTL mapping of candidate genes involved in Cd accumulation in rice grain. *Chin. Bull. Bot.* 56, 25–32. doi:10.11983/cbb20148

Pedrero, Z., Madrid, Y., Hartikainen, H., and Cámara, C. (2008). Protective effect of selenium in Broccoli (*Brassica oleracea*) plants subjected to cadmium exposure. *J. Agric. Food Chem.* 56, 266–271. doi:10.1021/jf072266w

Purcell, S., Neale, B., Todd-Brown, K., Thomas, L., Ferreira, M. A., Bender, D., et al. (2007). Plink: A tool set for whole-genome association and population-based linkage analyses. *Am. J. Hum. Genet.* 81, 559–575. doi:10.1086/519795

Qiu, T. C., Li, J., Cong, X., Cheng, H., Yu, S. B., Cheng, S. Y., et al. (2021a). Genome-wide association study of total Selenium content in polished rice. *Food Sci. Technol.* 46, 21–25. doi:10.13684/j.cnki.spkj.2021.12.004

Qiu, X., Chen, K., Lv, W., Ou, X., Zhu, Y., Xing, D., et al. (2017). Examining two sets of introgression lines reveals background-independent and stably expressed QTL that improve grain appearance quality in rice (*Oryza sativa* L.) *Theor. Appl. Genet.* 130, 951–967. doi:10.1007/s00122-017-2862-z

Qiu, X. J., Zhu, S. B., Hu, H., Wang, C. C., Lv, W. K., He, L. P., et al. (2020). Genome-wide association mapping for grain shape in rice accession. *Int. J. Agric. Biol.* 23, 512–618. doi:10.17957/IJAB/15.1327

Qiu, X., Pang, Y., Yuan, Z., Xing, D., Xu, J., Dingkuhn, M., et al. (2015). Genome-wide association study of grain appearance and milling quality in a worldwide collection of Indica rice germplasm. *PLoS One* 10, e0145577. doi:10.1371/journal.pone.0145577

Qiu, X., Yang, J., Zhang, F., Niu, Y., Zhao, X., Shen, C., et al. (2021b). Genetic dissection of rice appearance quality and cooked rice elongation by genome-wide association study. *Crop J.* 9, 1470–1480. doi:10.1016/j.cj.2020.12.010

Rao, M. J., Xu, Y., Huang, Y., Tang, X., Deng, X., and Xu, Q. (2019). Ectopic expression of citrus *UDP-GLUCOSYL TRANSFERASE* gene enhances anthocyanin and proanthocyanidins contents and confers high light tolerance in Arabidopsis. *BMC Plant Biol.* 19, 603. doi:10.1186/s12870-019-2212-1

Rikans, L. E., and Yamano, T. (2000). Mechanisms of cadmium-mediated acute hepatotoxicity. *J. Biochem. Mol. Toxicol.* 14, 110–117. doi:10.1002/(sici)1099-0461(2000)14:2<110::aid-jbt7>3.0.co;2-j

Shang, D. R., Zhao, Y. F., Guo, Y. Y., Zhai, Y. X., Ning, J. S., Sheng, X. F., et al. (2012). Safety evaluation of arsenic and arsenic coumpounds in food. *Chin. Fish. Qual. Stand.* 2, 21–32.

Sharma, P., Gangola, M. P., Huang, C., Kutcher, H. R., Ganeshan, S., and Chibbar, R. N. (2018). Single nucleotide polymorphisms in B-genome specific UDP-Glucosyl Transferases associated with fusarium head blight resistance and reduced deoxynivalenol accumulation in wheat grain. *Phytopathology* 108, 124–132. doi:10.1094/phyto-04-17-0159-r

Singh, A. K., Kumar, R., Pareek, A., Sopory, S. K., and Singla-Pareek, S. L. (2012). Overexpression of rice CBS domain containing protein improves salinity, oxidative, and heavy metal tolerance in transgenic tobacco. *Mol. Biotechnol.* 52, 205–216. doi:10.1007/s12033-011-9487-2

Tan, X. S., Zeng, Y. H., Li, Z. M., Guan, F., and Guan, E. X. (2014). Effects of Selenium fertilizer on the content of heavy metal Cadmium in eice. *Chin. Rice* 20, 76–77. doi:10.3969/j.issn.1006-8082.2014.01.018

Tapiero, H., Townsend, D. M., and Tew, K. D. (2003). The antioxidant role of selenium and seleno-compounds. *Biomed. Pharmacother.* 57, 134–144. doi:10.1016/s0753-3322(03)00035-0

Vilhjálmsson, B. J., and Nordborg, M. (2013). The nature of confounding in genome-wide association studies. *Nat. Rev. Genet.* 14, 1–2. doi:10.1038/nrg3382

Wang, C. C., Yu, H., Huang, J., Wang, W. S., Faruquee, M., Zhang, F., et al. (2020). Towards a deeper haplotype mining of complex traits in rice with RFGB v2.0. *Plant Biotechnol. J.* 18, 14–16. doi:10.1111/pbi.13215

Wang, L. Y., Zhu, C. N., Wang, J. J., Zhang, L. X., Jin, Q. S., and Shi, C. H. (2012). Screen for rice (*Oryza sativa*) genotypes with lower Cd, Pb and as contents. *Acta. Agric. Zhejiang* 24, 133–138. doi:10.3969/j.issn.1004-1524.2012.01.025

Wang, P., Chen, H. P., Kopittke, P. M., and Zhao, F. J. (2019). Cadmium contamination in agricultural soils of China and the impact on food safety. *Environ. Pollut.* 249, 1038–1048. doi:10.1016/j.envpol.2019.03.063

Wang, T., Khir, R., Pan, Z., and Yuan, Q. (2017). Simultaneous rough rice drying and rice bran stabilization using infrared radiation heating. *LWT-Food Sci. Technol.* 78, 281–288. doi:10.1016/j.lwt.2016.12.041

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, 43–49. doi:10.1038/s41586-018-0063-9

Yan, H., Xu, W., Xie, J., Gao, Y., Wu, L., Sun, L., et al. (2019). Variation of a major facilitator superfamily gene contributes to differential cadmium accumulation between rice subspecies. *Nat. Commun.* 10, 2562. doi:10.1038/s41467-019-10544-y

Yang, H. P., Wang, S. Y., Song, W., Jin, J., and Zhu, G. F. (2004). Study on extracting rice bran oil from rice bran by aqueous enzymatic method. *Food Sci.* 25, 106–109. doi:10.1007/BF02911033

Zhang, G. M., Zheng, T. Q., Chen, Z., Wang, Y. L., Wang, Y., Shi, Y. M., et al. (2018a). 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, J., Ou, X. X., Hu, H., Du, B., Lv, W. K., Yang, L. W., et al. (2018b). Identification of QTLs for yield-related traits using two sets of introgression lines with a common donor parent in rice. *Int. J. Agric. Biol.* 20, 15–24. doi:10.17957/ijab/15.0321

Zhang, J., Zhu, Y. G., Zeng, D. L., Cheng, W. D., Qian, Q., and Duan, G. L. (2008). Mapping quantitative trait loci associated with arsenic accumulation in rice (Oryza sativa). *New Phytol.* 177, 350–355. doi:10.1111/j.1469-8137.2007.02267.x

Zhang, X. W., Yang, L., Zhang, T., Zheng, J. K., Jin, L., Zheng, J. M., et al. (2010). Inheritance of Selenium contents and QTL detection in rice grains. *J. Plant. Genet. Res.* 11, 445–450.

Zhou, Z. S., Song, J. B., and Yang, Z. M. (2012). Genome-wide identification of Brassica napus microRNAs and their targets in response to cadmium. *J. Exp. Bot.* 63, 4597–4613. doi:10.1093/jxb/ers136

Zhu, H. H., Chen, C., Xu, C., Zhu, Q. H., and Huang, D. Y. (2016). Effects of soil acidification and liming on the phytoavailability of cadmium in paddy soils of central subtropical China. *Environ. Pollut.* 219, 99–106. doi:10.1016/j.envpol.2016.10.043

Zou, W. L., Leng, Y. J., Li, J. L., Meng, L. J., He, H. H., Chen, J. G., et al. (2018). Uptake and translocation mechanism of Cadmium accumulation in rice and QTL mapping related to Cadmium stress. *Mol. Plant Breed.* 16, 8128–8141. doi:10.13271/j.mpb.016.008128

frontiers | Frontiers in Genetics

TYPE Review
PUBLISHED 28 October 2022
DOI 10.3389/fgene.2022.1001001

OPEN ACCESS

EDITED BY
Reyazul Rouf Mir,
Sher-e-Kashmir University of Agricultural Sciences and Technology, India

REVIEWED BY
Jiuran Zhao,
Beijing Academy of Agricultural and Forestry Sciences, China
Tim L. Setter,
Cornell University, United States

*CORRESPONDENCE
Walter P. Suza,
wpsuza@iastate.edu

†Present address: Michael S. McMillen, Syngenta, Greensboro, NC, United States

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 22 July 2022
ACCEPTED 14 October 2022
PUBLISHED 28 October 2022

CITATION
McMillen MS, Mahama AA, Sibiya J, Lübberstedt T and Suza WP (2022), Improving drought tolerance in maize: Tools and techniques.
Front. Genet. 13:1001001.
doi: 10.3389/fgene.2022.1001001

Improving drought tolerance in maize: Tools and techniques

Michael S. McMillen[1†], Anthony A. Mahama[1], Julia Sibiya[2], Thomas Lübberstedt[1] and Walter P. Suza[1*]

[1]Department of Agronomy, Iowa State University, Ames, IA, United States, [2]School of Agricultural, Earth and Environmental Sciences, University of KwaZulu-Natal, Pietermaritzburg, South Africa

Drought is an important constraint to agricultural productivity worldwide and is expected to worsen with climate change. To assist farmers, especially in sub-Saharan Africa (SSA), to adapt to climate change, continuous generation of stress-tolerant and farmer-preferred crop varieties, and their adoption by farmers, is critical to curb food insecurity. Maize is the most widely grown staple crop in SSA and plays a significant role in food security. The aim of this review is to present an overview of a broad range of tools and techniques used to improve drought tolerance in maize. We also present a summary of progress in breeding for maize drought tolerance, while incorporating research findings from disciplines such as physiology, molecular biology, and systems modeling. The review is expected to complement existing knowledge about breeding maize for climate resilience. Collaborative maize drought tolerance breeding projects in SSA emphasize the value of public-private partnerships in increasing access to genomic techniques and useful transgenes. To sustain the impact of maize drought tolerance projects in SSA, there must be complementary efforts to train the next generation of plant breeders and crop scientists.

KEYWORDS

drought tolerance, food security, maize breeding, genomics assisted selection, genome mapping, model-assisted approaches, plant breeding education

Introduction

Maize is the primary staple food for more than 900 million people globally and the third most important source of calories after rice and wheat (Shiferaw et al., 2011; Adebayo and Menkir, 2015). With a decrease in rice production in China and India, and an increase in the demand for dairy and meat, global demand for maize is projected to double by 2050 (Rosegrant et al., 2009). Maize is also the most widely grown staple crop in SSA and serves an important role in food security, but it is also highly susceptible to drought, with 15%–20% of its yield lost to drought each year (FAOSTAT, 2010; Bankole et al., 2017; Lunduka et al., 2017; FAO, 2021). Because of maize yield loss to drought, from 2005 to 2915, developing countries experienced a revenue loss of up to USD 29 billion (FAOSTAT, 2010; Liu and Qin, 2021). Changes in rainfall magnitude, distribution, and timing, as well as an increase in temperature, all interact to destabilize maize production further (IPCC, 2012; Meseka et al., 2018). Data from more than 20,000 trials in Africa revealed a 1% reduction in maize yield for each "degree day" above 30°C (Lobell et al.,

2011; Bankole et al., 2017; Lunduka et al., 2017). When other factors are not limiting, the combination of heat and drought stress causes a 1.7% yield reduction for each excess degree day (Lobell et al., 2011).

An assessment of the impact of physical drought on maize revealed an alarming vulnerability in SSA (Kamali et al., 2018). On average, 5 to 10 drought events were experienced between 1970 and 2004 in most parts of SSA (Fisher et al., 2015). This makes drought a key constraint to maize production in the region (Heisey and Edmeades, 1999). Further, climate change is expected to exacerbate the impact of drought in the region and cause a reduction in maize production by almost 22% in 2050 (Schlenker and Lobell, 2010; Challinor et al., 2016; Tesfaye et al., 2018; Barbosa et al., 2021). Therefore, meeting the increased demand while stabilizing production requires a strategy that includes the genetic improvement of maize for drought tolerance (Messina et al., 2021). A recent review by Sheoran et al. (2022) discussed new breeding technologies and approaches to improve maize drought tolerance. This review presents an overview of a broad range of tools and techniques, the integration of plant physiology, molecular biology, and systems modeling, and examples of private-public partnerships in developing drought tolerant maize for Africa. The review also discusses the importance of plant breeding education to address the shortage of plant breeders in SSA.

Breeding for drought tolerance in maize

Drought is a key abiotic stress that causes low-income countries to lose billions of dollars (FAO, 2021). Levitt (1972) provided functional definitions for water deficit and drought stress which can be used to develop breeding targets. Water deficit occurs when plant transpiration cannot fully meet the atmospheric demand due to a lack of water in the environment. This deficit causes damage and induces a stress response proportional to the rate of deficit (Blum, 2014).

The impact of drought stress depends on the interaction of plant, environmental, and management factors: the crop development stage, rate of water deficit development, peak intensity of the deficit, and planting density. Regarding the development stage, maize is most sensitive to drought stress during flowering (Bolanos and Edmeades, 1993). Severe water deficits during the period of a few days before silking to roughly 25 days can eliminate yield entirely (Claassen and Shaw, 1970). Drought stress causes a delay in ear growth and silking, increasing the anthesis-silking interval (ASI) to the point where it inhibits fertilization. The result is a barren ear or one with few kernels (Sah et al., 2020). Even with successful pollination, kernel abortion beginning as early as 2–3 days after pollination can reduce kernel number (Westgate and Boyer, 1986). Drought stress at the start of grain filling can also significantly lower or eliminate yield (Barker et al., 2005). Since the crop reaches full size prior to flowering, water use is at a maximum. Drought stress induces premature leaf senescence and reduces ear growth, with severe stress causing complete desiccation. Consequently, kernel weight is significantly reduced due to lowered photosynthate accumulation.

According to Ribaut et al. (2009), maize responds to and mitigates the impact of water deficit using three primary strategies: drought escape, drought avoidance, and drought tolerance. Drought escape is a strategy to prevent the coincidence of water deficit with key developmental stages and is primarily achieved by early flowering and maturation. Drought avoidance, on the other hand, is the capacity to avoid or reduce plant water deficit but maintaining turgor through an increase in water uptake (using a deeper and/or a larger root system for example) and/or a reduction in water use (for instance, decreased stomatal conductance). Drought tolerance is the ability to maintain plant function during water deficit, which can be achieved by alleviating oxidative stress, for instance.

The goal of a drought tolerance breeding program is two-fold: 1) reduce the gap between yields in optimal and stress conditions without sacrificing yield potential, and 2) improve yield stability for a range of stress conditions. According to Lunduka et al. (2017), a drought tolerant maize variety is one that gives at least a yield of 30% of its potential under water-stress, especially during the flowering and grain-filling stages. In addition, Messina et al. (2021) found that although root systems architecture and yield have changed because of breeding for maize drought tolerance, the uptake of water has not changed.

While drought escape does create an advantage under drought conditions, using earlier maturing maize varieties which generally yield less than full season varieties (White et al., 1922) in non-drought years creates a lower yield potential (Ke and Ma, 2021). To improve both yield potential and stability, the use of hybrids with a maturity suited to the wettest part of the year in the target environment might be an option. Together, the capacity of a plant to avoid or reduce water deficit (dehydration avoidance), sustain function under water deficit (dehydration tolerance), improve crop productivity and provide avenues for plant improvement.

Even without drought tolerance as an explicit breeding goal, selecting for high yield potential under well-watered conditions has consistently led to increased yield in both water deficit and non-deficient environments (Castleberry et al., 1984). Selecting for high yield potential extends tolerance to other abiotic stresses as well, such as heat, cold, and low soil fertility. The improvement in drought tolerance and the rate of genetic gain has been evaluated through experiments testing historic cultivars under drought conditions applied at various growth stages, demonstrating a gain of 124 kg ha^{-1} yr^{-1} for flowering stress and 91 kg ha^{-1} yr^{-1} for mid grain fill stress (Barker et al., 2005). Nevertheless, the ability of newer hybrids to tolerate drought stress is primarily due to the

adaptation of parent lines to higher planting densities (Tollenaar and Lee, 2006).

Although selecting for high yield potential in favorable environments has led to increases in drought tolerance, the correlation of yield for hybrids grown under well-watered and water-stressed conditions is reduced with the degree of stress. In addition, the degree of genotype by environment interaction necessitates screening materials under both well-watered and water stressed conditions. Therefore, because drought is unpredictable in the target environment, sites in rain-free environments, which use irrigation and planting date to control the timing and severity of drought, are needed (Campos et al., 2004).

Choosing testing environments

The rate of achievable genetic gain is highly dependent on the choice of testing environments, especially how closely the selection environments mirror the target environments (Cooper et al., 2006). Since the distribution and amount of rainfall, temperature, soil water holding capacity, and the developmental stage of the crop all interact to create distinct drought scenarios, an important preliminary step in developing a breeding program for drought tolerance is capturing this information to identify the target population of environments (TPE) for which the crop will be adapted (Cooper et al., 1997).

Environmental characterization and TPE delineation are crucial for several reasons. Information on the TPE can be used to explain and predict a considerable portion of genotype by environment (G x E) interactions. This is necessary since the effect of a particular allele can be different depending on environmental conditions (Chenu et al., 2009). In one drought scenario a trait can be advantageous, while in others it can be detrimental. For instance, breeding for water use efficiency can improve yield in very dry environments, while reducing the potential yield in mild drought conditions (Tardieu, 2012).

Characterizing the TPE is also crucial in identifying appropriate managed stress environments for multi-environment trials (METs) since the expected performance gain is dependent on the similarity between environments represented in METs and the TPE (Ribaut et al., 2009). Furthermore, by weighting phenotype data from METs based on how representative individual trials of the TPE are, selection gain over generations can be improved. This improvement in gain was demonstrated by Podlich and Cooper (1998) using a genetic simulation model. Comparing the drought scenarios of the TPE with other drought-prone regions facilitates methodological and adapted germplasm exchange to other parts of the world (Chenu et al., 2011). Increases in temperature and changes in weather patterns brought about by climate change highlight the need to characterize the target environment in advance, since the environmental conditions could be significantly different by the time a variety is developed and disseminated, resulting in lower-than-expected yields (Challinor et al., 2016). These seasonal variability and resource constraints often lead to multi-environment breeding trials offering biased representation of the TPE. One way to avoid this is use of weighted analysis based on representative trials which can help breeders select for germplasm better adapted to the TPE. Managed-environment trials are another way to evaluate performance in representative environments or for stresses, allowing detailed assessment of germplasm, traits, or genes of interest (Chenu, 2015).

As mentioned above, crop simulation models have been used to characterize TPEs, and to evaluate how well MET locations fit the TPE. Studies using these models help to improve breeding efficiency for multiple crops in regions around the world. For instance, the "*Cerrado*" environments of central Brazil have been characterized for rice and maize production (Heinemann et al., 2008). More recently, maize growing environments in Eastern and Southern Africa were characterized using the APSIM model, leading to the distinction of four environmental types to which breeding objectives can be catered (Seyoum et al., 2017).

In addition to similarity to the target environment, site homogeneity and the ability to manage water inputs determine the success of testing (Blum, 2011b). Differences in soil texture, effective rooting depth, micronutrient concentration, salinity, pH, and the presence of pathogens all increase residual variability, thus minimizing the precision in estimating genotypic means. Regarding the ability to manage water inputs, use of desert and off-season environments, and rainout shelters provide options to prevent the effect of precipitation (Blum, 2011b). One issue with the use of desert environments is the effect of temperature extremes. Estimating drought tolerance is confounded by the occurrence of extreme temperatures during ear growth (Otegui and Andrade, 2000). Recent improvements in modeling techniques have increased their effectiveness. Liu et al. (2021) concluded that DSSAT CERES-Maize can adequately simulate regional maize yields using the CERES-Maize module calibrated to regional soil and daily weather databases. Adnan et al. (2020) successfully use the CERES-Maize model to generate data for GEI and stability studies of maize genotype in the absence of observed field data, and Ramirez-Villegas et al. (2020) pointed out the varying important roles of crop modeling in breeding efforts, including assessing genotypic adaptability and stability, characterizing and identifying target breeding environments, identifying tradeoffs among traits for such environments, and making predictions of the likely breeding value of the genotypes. Recognizing the successes from simulation modeling, Hajjapoor et al. (2022) pointed out the problems that still exist in identifying MET environments that fit TPEs, to deal with Genotype-by-Environment-by-Management (G x E x M) interactions and proposed a simple step-by-step approach to bring the capacity of process-based

models to better define target population of environment, within which the clustering of subunits will allow for the reduction of G x E × M interactions.

Incorporating secondary traits

Although maximizing grain yield is the primary objective of breeding for improved drought tolerance, reduced genotypic variance and high G × E interaction contribute to reduced heritability of yield when testing under drought conditions. Incorporating secondary traits into the selection process, that is, traits which more directly reflect the physiological effect of drought stress, can increase selection efficiency, and improve gain (Ribaut et al., 2009). Edmeades et al. (1996a) developed criteria for ideal secondary traits for drought screening. In addition to high heritability, high genetic variability, genetic correlation with yield, and no association with yield loss under non-limiting conditions, ideal traits should be simple, cheap, non-destructive, and fast to assay. Using these criteria, only ASI, ears per plant (EPP), barrenness, kernels per ear, and stay green have been found to be suitable secondary traits. Of those traits, ASI and EPP have been identified as the best performing traits, with ASI being the most widely used. While tassel growth is not as affected by drought stress, silk emergence and thus ASI can be used as an indicator for ear and plant growth rates during flowering. When testing under conditions that reduce yield by more than 50%, the incorporation of these secondary traits into a selection index has led to a selection efficiency on a par with testing under optimal conditions (Chapman et al., 1997). Using yield alone for selection resulted in significantly lower gains but the use of secondary traits for selection has resulted in improved genetic gains in other instances as well (Campos et al., 2004).

Application of genomic mapping tools

While the use of secondary traits provides more heritable selection targets, the advancing field of genomics has been used to further improve selection efficiency by identifying and selecting the genomic regions responsible for improved secondary traits and tolerant phenotypes. By developing a better understanding of the genetic and physiological basis of drought tolerance, and using this information during selection, the value of genomics in improving drought tolerance has been demonstrated in maize (Tuberosa et al., 2007; Tsonev et al., 2009), in addition to several other crops (Tuberosa and Salvi, 2006; Mir et al., 2012). With marker-assisted selection, the initial step is to identify markers and genes associated with drought tolerance. These associations have been identified using a variety of approaches. In addition, the use of parental haplotype sharing can help increase the power, precision, and accuracy in Quantitative Trait Loci (QTL) mapping (Jansen, Jannink, and Beavis, 2003).

Most understanding of the genetic basis of drought tolerance was gained from traditional QTL mapping, using a relatively limited number of markers in bi-parental populations (Cattivelli et al., 2008). While a large number of associations have been found, there are limitations with traditional QTL mapping: 1) the resulting QTL are comparatively large due to limited genetic resolution; 2) extra time is required to develop a mapping population; 3) by using a mapping population, only a small proportion of the total allelic diversity expected in the germplasm pool is sampled; and 4) QTL for the same trait can segregate differently in other mapping populations.

Using linkage mapping, numerous QTL relating to morphological traits such as flowering and tassel size, as well as physiological parameters such as ABA and carbohydrate metabolism have been identified. Several of these QTL studies have been summarized by Ribaut (2006). To identify genome regions and candidate genes consistent across populations, which convey drought tolerance, QTL from various studies have been analyzed and compiled into consensus maps (Tuberosa et al., 2002; Sawkins et al., 2004; Hao et al., 2009; Semagn et al., 2013; Zhao et al., 2018).

Linkage disequilibrium (LD)-based association mapping has been used to overcome some of the constraints of linkage mapping. The advantages of LD association mapping are summarized by Mir et al. (2012) and include: less time and resources are required since a natural germplasm collection can be used, and it provides a higher mapping resolution. LD studies also have the advantage of being able to simultaneously evaluate the varying effects of multiple alleles in multiple backgrounds at one time (Buckler and Thornsberry, 2002). Since diverse germplasm is used for association mapping, identified markers are more likely to convey drought tolerance in multiple backgrounds which is particularly valuable to breeders.

Although linkage and LD-mapping both have unique advantages, the approaches are complementary (Myles et al., 2009) and have been combined to better identify QTL associated with drought tolerance. Using the combined techniques, SNP markers were identified, which better explained the phenotypic variation regarding ASI compared to either technique alone (Lu et al., 2010). However, association mapping is not without limitations. One drawback is the difficulty in detecting associations with traits underpinned by many rare variants with a large effect size, or by many common variants that have a small effect (Korte and Farlow, 2013). Since the effect size of the allelic variants as well as their frequency in the sampling population determine the phenotypic variance, rare variants, and a combination of alleles with small effect sizes are difficult to associate (Korte and Farlow, 2013).

As sequencing technologies have advanced and costs have dropped, genome-wide association studies (GWAS) have become a common approach to uncover the genetic basis of drought

tolerance (Yamada and Dwiyanti, 2013). GWAS has been successful in identifying genome regions associated with drought tolerance. Using 60,000 SNPs on a hybrid testcross background allowed dominant alleles to be detected. Farfan et al. (2015) identified 10 quantitative trait variants (QTVs) for flowering time, plant and ear height, and yield. Under both well-watered and stressed conditions, three of these QTVs explained 5–10% of yield variation. Many of the QTVs also co-located with QTL from other studies, which confirmed their association with drought tolerance. Rather than using a high-density coverage of the maize genome, other association studies have taken a more targeted approach and instead have used a comparatively limited number of SNPs (1536) selected from candidate genes associated with drought response (Setter et al., 2011).

While Setter et al. (2011) identified several significant loci and candidate genes, it is possible that many were missed due to a lack of genome coverage. To enlarge the panel of markers for the collection without additional genotyping, Zhang et al. (2016) used imputation based on 556,809 SNPs from 368 diverse inbred lines that were previously genotyped using RNA sequencing. This method resulted in the identification of 26 new loci associated with metabolic and physiological traits in leaf tissue, and only one of the six loci significantly associated with drought-related metabolites from the Setter study (Setter et al., 2011) was still significant. Positional cloning of QTL in conjunction with association mapping helped to identify genes and noncoding sequences associated with flowering time, a trait strongly associated with drought response (Salvi et al., 2007).

Nested Association Mapping (NAM), which relies on a specifically designed population, also combines the advantages of linkage and association mapping while minimizing the disadvantages of each (Yu et al., 2008). By crossing 25 inbred lines to the B73 inbred and selfing the F_2 populations to the F_6 generation, 200 recombinant inbred lines (RILs) were created for each of the 25 populations. These RILs were genotyped using the same 1106 markers to identify recombination blocks and the parents were sequenced, resulting in 5000 RILs (known as the US-NAM population) that were genotyped at high density and could be compared and analyzed together (Yu et al., 2008). Researchers from the Institute of Crop Science at the Chinese Academy of Agricultural Sciences used both the US-NAM and China NAM populations to identify 52 candidate genes with differential expression under two water treatments, allowing them to make genomic predictions of drought-related traits with a mean accuracy of 0.57 (Li et al., 2016).

Candidate genes have also been identified by incorporating functional genomics technologies, such as transcriptome and metabolome analysis. These approaches help reveal how biochemical, physiological, and regulatory networks change in response to stress, and expose differences in drought response among tissues and genotypes (Mir et al., 2012). By studying gene expression in pre-fertilization ears using both cDNA microarray and genome-survey technology, Zinselmeier et al. (2002) demonstrated differential gene expression among tissues, and identified several genes not previously associated with drought stress response. Microarray technology also facilitated the identification of 22 differentially expressed genes which co-located on the genetic map with QTL previously associated with drought tolerance (Marino et al., 2008).

Like gene expression analysis, examining changes in the plant metabolome can connect the agronomic phenotype with the underlying genetics, and can be useful in identifying genes that are not as affected by environmental factors (Riedelsheimer et al., 2012). A distinct advantage of profiling the metabolome (and proteome) over the use of transcriptomics is that effects from posttranscriptional and posttranslational regulation can be accounted for (Tuberosa et al., 2007). By examining metabolic changes in maize leaves and ears due to water deficit and dissecting the genetic basis of those traits using GWAS, Zhang et al. (2016) identified 23 metabolite-associated loci and validated 10 as responsive to drought stress. Using the same technique on leaves alone, Riedelsheimer et al. (2012) identified 26 SNPs strongly associated with changes in distinct metabolites, which explained up to 32% of the observed genetic variance.

More exciting research results have been achieved in recent years on the genetics of drought tolerance/resistance. The identification of naturally occurring loci or genes associated with drought tolerance can serve as direct targets for both engineering and selecting improved maize for drought regions. Increased expression of the NAC gene (*ZmNAC111*) with the MITE (significantly associated with natural variation in maize drought tolerance) insertion in its promoter enhanced drought tolerance in maize seedlings (Mao et al., 2015). Wang et al. (2016) reported that transgenic maize with enhanced expression *ZmVPP1* (also with natural variation) exhibited improved seedling drought tolerance. In a study of a NAC-encoding gene of *ZmNAC080308*, a functional marker developed for use in predicting drought stress tolerance in a US maize inbred line panel showed that lines carrying Hap2 produced greater than 10% grain yield than those carrying Hap1 under drought stress condition (Wang et al., 2021). Wang et al. (2022) identified and reported the overexpression of a transcription factor, *ZmERF21*, is tightly associated with drought tolerance in maize seedlings, expressed by significantly increased chlorophyll content and activities of antioxidant enzymes under drought conditions. Another GWAS study (Khan et al., 2022) identified candidate genes and their key variations that will contribute to an understanding of the genetic basis of drought tolerance, especially for the female inflorescence, and will be important in facilitating drought-tolerant maize breeding.

Bridging the gene-to-phenotype gap

Advances in molecular biology have provided tools and strategies to associate genomic regions with traits that convey

improved yield in a drought scenario. With genomic selection, these associations can be used for phenotypic prediction and breeding decisions. However, predicting the effects of genes across scales of biological organization is made difficult by the complex interaction of genes and environmental factors (Hammer et al., 2006). Much of the time the effect of a particular allele on a complex trait such as drought tolerance is confounded by gene interactions such as epistasis and pleiotropy (Hammer et al., 2006). In addition to gene-gene interactions, the drought scenario strongly influences whether an allele will convey a significant advantage. In many cases, the effect of a trait on yield is a trade-off which depends on environmental conditions (Tardieu and Tuberosa, 2010). For instance, reduced transpiration and biomass accumulation protect against drought stress, yet reduce yield potential (Blum, 2009).

Modeling can be incorporated to address the challenges presented by gene-gene and gene-environment interactions in the physiological and genetic dissection of traits which convey drought tolerance, and ultimately in developing improved cultivars (Tardieu and Tuberosa, 2010). In addition to improving efficiency by dissecting complex traits into more measurable targets, which helps in the development of phenotyping strategies, using crop growth and development models to evaluate traits and predict phenotypes in the TPE is useful for assessing breeding strategies and allocating resources (Hammer et al., 2006; Messina et al., 2011). With the addition of environment modeling, the range of drought scenarios that comprise a TPE over time can be accounted for, helping to efficiently identify associations between traits and the set of environmental conditions which maximize yield.

Not only has modeling been used to enhance the physiological and genetic dissection of drought tolerance, it also has been used to augment a maize breeding program. By extending the concept of fitness landscapes to the characterization of yield-trait performance landscapes and extending the E (NK) model of trait genetic architecture to incorporate biophysical, physiological, and statistical components, Messina et al. (2011) developed a graphical representation of the associated yield-trait performance landscape that could be used in selection decisions. Doubled haploid (DH) lines were selected based on their relative position in the performance landscape, their predicted performance, and their potential to contribute to further yield improvement. This approach showed that an understanding of yield-trait performance landscapes can be used to improve genomic selection and phenotyping strategies.

Studies which use modeling have improved the predictive power of the combined effects of major QTL. In one study, a model based on the combined effects of the major QTLs was able to predict 74% of the variability for maize leaf elongation rate under a range of temperature and water deficit conditions (Reymond et al., 2003). Modeling has also been combined with phenotypic analysis to identify QTL for drought traits and to characterize genotypes (Tardieu, 2006). The development of response curves, which more quantitatively define the relationship between the phenotypic trait and environmental conditions, are unique to each genotype and can be compared to select optimal genotypes for a particular range of conditions. A similar approach has been used to identify common QTL for leaf growth and ASI under drought conditions, suggesting that the genetic determination of leaf growth and silk elongation rate is at least partially shared (Welcker et al., 2006). By extending the APSIM crop model to include genotype-specific parameters, Chenu et al. (2009) simulated the effects of QTL on leaf and silk elongation, and ultimately yield. The study demonstrated the high level of QTL × environment interaction, opening the possibility of exploiting these interactions for drought tolerance (Chenu et al., 2009).

Marker assisted selection in breeding

While many drought tolerance candidate genes and QTL have been documented, few have been validated to produce a clear genetic gain when selected for in diverse germplasm under field conditions (Ribaut et al., 2009). This is in part due to many studies being limited to putative association based on colocalization of candidate genes and QTL along genetic maps. Another reason is that QTL or genes affecting drought tolerance are not distinguished based on how they are expressed under contrasting drought treatments. For breeding, QTL which are constitutively expressed and affect the yield consistently should be given priority, since they show limited interaction with the environment (Vargas et al., 2006). However, relatively simple heritable constitutive plant morphological and developmental traits are often ignored when evaluated by functional genomics, even though they have a considerable effect on performance under drought stress (Blum, 2011a). Other factors which limit the practical use of previous genetic investigations are the germplasm used for the mapping population, and the drought treatment used. Non-commercially viable lines are often used for association studies and identified QTL may have a null or negative effect when introgressed into elite materials (Monneveux and Ribaut, 2006). Regarding the drought treatment, many of the expressed genes identified in association studies are due to imposing rapid stress in a laboratory setting, but the slow progress of drought stress in field conditions results in minimal expression of these stress-responsive genes (Barker et al., 2005).

Although most information gathered from the genetic dissection of drought tolerance has not directly led to the development of improved cultivars, marker assisted selection has been used successfully. By introgressing favorable alleles at five genomic regions, MABC-derived hybrids were selected, which yielded 50% more than control hybrids under severe

water stress conditions. Under mild stress and well-watered conditions, the hybrids performed as well as the control (Ribaut and Ragot, 2006). While the experiment did improve the recipient line through the introgression of drought QTL, further improvement is limited after the recurrent parent is fixed for the new QTL (Mir et al., 2012). To address this limitation, MARS has been used to take advantage of desirable alleles in multiple lines. The value of MARS in breeding for improved maize grain yield in SSA was demonstrated by Beyene et al. (2016) who estimated the rate of genetic gain under drought and non-drought environments. Since the value of a set of QTL alleles depends on the germplasm used in breeding, another MAS strategy was proposed by Podlich et al. (2004) to attempt to account for the presence of epistasis and G × E interaction. The "Mapping As You Go" (MAYG) strategy works by re-estimating the value of each QTL allele upon creation of a new set of germplasm, and its effectiveness has been estimated by simulation (Podlich et al., 2004).

Rather than selection based on known markers with significant associations to drought traits, the use of genomic selection (GS) or genome-wide selection (GWS) to develop drought tolerance promises improved gains (Bernardo and Yu, 2007). The application of GS in yield trials of tropical maize lines across multiple locations in SSA produced selection candidates at lower cost than phenotypic selection (Beyene et al., 2019). The cost aspect is important because in situations where doubled haploids (DH) lines are used, the capacity to phenotype testcross materials across multiple sites can limit progress (Beyene et al., 2021).

Genomic Selection described as "test-half-predict half" approach uses random markers to genotype a phenotyped training population. The marker and phenotyping data are used to develop breeding values of alternative alleles, which are fitted as random effects in a linear model. Selection in each recurrent generation is based on the sum of those breeding values, known as the genomic estimated breeding value (GEBV) (Meuwissen et al., 2001). This approach has considerable advantages, such as high selection accuracy when selecting on markers alone, and no prior knowledge of QTL positions is needed (Resende et al., 2014). The efficiency of the GS approach has been compared experimentally to selection based on yield alone along with selection incorporating secondary traits (Ziyomo and Bernardo, 2013). Compared to selection based on yield, secondary trait assisted selection was slightly higher or lower depending on the trait, while GS was significantly more efficient.

Sources of genetic diversity

There are many sources of genetic diversity which potentially hold alleles promoting drought resistance (Barbosa et al., 2021). While conventional germplasm bases contain enormous levels of allelic polymorphism (Guo et al., 2004), the development of mutagenic and transgenic lines create a virtually endless supply of genetic diversity. Most sources of natural variation have remained relatively untapped. Globally, only 5% of the available maize germplasm is used commercially (Brown, 1975), and exotic germplasm constitutes only 1% of the US germplasm base (Goodman, 1983). This section aims to present resources that have been used or with potential in breeding for or dissecting drought tolerance.

An ideal germplasm base 1) segregates for the trait of interest, 2) has a high probability of containing desired alleles, and 3) those alleles should be relatively easy to introgress into cultivars (Blum, 2011b). The existing drought tolerance, genetic diversity, and ease of introgression make the agronomic germplasm pool an ideal first choice. Agronomic germplasm has been used to develop the Pioneer AQUAmax® product line of drought-tolerant maize hybrids, improving upon the drought tolerance of commercial hybrids that already possess a high degree of tolerance (Cooper et al., 2014).

Alleles from landraces have been introgressed into elite varieties, leading to improved drought tolerance. Researchers at the International Institute of Tropical Agriculture (IITA) in Ibadan, Nigeria, crossed six landraces pre-screened for drought tolerance with an elite maize variety (AK9443-DMRSR), resulting in some of the BC_1F_2 populations expressing improved yield potential under random-drought conditions (Meseka et al., 2013). In dry, marginal growing regions of Kenya, farmers use local landraces rather than hybrids because they are believed to produce better under low or no input use (Sammons, 1987), and possess resistance to biotic and abiotic stress.

While wild germplasm and related species, such as *Zea maxicana* or *Tripsacum floridanum*, are excellent sources of novel genes for drought tolerance improvement (Singh, 2010), there are some difficulties with their use. First, the technical difficulty in making a wide cross between an agronomic genotype and the donor. Second, the need to eliminate the introgression of negative traits carried by the donor, and third, the value of genes or alleles conveying drought resistance once introgressed (Blum, 2011b). Since adaptation to drought stress in wild type and related species enhances survival rather than optimizing yield, the genes which impart that resistance may not be beneficial when introgressed into breeding germplasm.

A molecular approach particularly suited to addressing this complication and identifying beneficial alleles from wild sources is advanced backcross QTL (AB-QTL) analysis (Tanksley and Nelson, 1996). After developing backcross families and eliminating lines displaying yield-reducing characteristics in the BC_1 and BC_2 generations, QTL analysis is performed. This approach has the benefit of identifying valuable QTL and developing superior genotypes simultaneously and has been used to analyze more conventional crosses as well (Ho et al., 2002).

Transgenic and gene-editing tools

The application of mutagenesis techniques to improve drought tolerance in maize is limited (Blum, 2011b; Gao et al., 2014; Ruswandi et al., 2014). Therefore, transgenic approaches are useful in functional analysis of genes affecting stress response and adaptation (Yang et al., 2010), and can act as a bridge to move valuable genes into breeding germplasm (Blum, 2011b). Nelson et al. (2007) transformed plants to overexpress transcription factor *ZmNFYB2* resulting in increased drought tolerance and yield. Plants engineered to overexpress the *ZmAsr1* gene, a putative transcription factor, resulted in lines with improved water use efficiency and dry weight accumulation (Jeanneau et al., 2002). In another study, Nuccio et al. (2015) engineered plants to overexpress a gene encoding rice trehalose-6-phosphate phosphatase (*TPP*), which increased sucrose concentration in ear spikelets, kernel set, and harvest index, leading to improved yield in mild and severe drought. Guo et al. (2014) generated transgenic maize overexpressing the *ARGOS1* (*ZAR1*) gene resulting in enhanced maize organ growth, grain yield, and drought-stress tolerance.

While the transgenic approach has helped to identify genes and mechanisms that improve drought tolerance, there are obstacles to their use in released cultivars. Regarding transgenic events tested in laboratory conditions, the effect of the transgene might not yield an advantage in the TPE because of environmental interaction (Bänziger and Araus, 2007). When placed in agronomic germplasm, the expression of a single gene may not significantly alter the final phenotype due to dampening and compensation from other processes (Sinclair and Purcell, 2005). The mechanism of tolerance can also be dependent on the developmental stage (Flowers, 2004).

Despite these obstacles, there has been some success with the use of transgenic cultivars to mitigate the impact of drought stress on crops. One released cultivar developed using transgenics is the Droughtgard™ hybrids from Monsanto. By isolating and transferring a cold-shock protein gene (*cspB*) from the soil bacteria "*Bacillus subtilis*", yield improvements from 11%–21% were achieved under drought conditions, with no negative effects under normal conditions. The protein cspB acts as a chaperone for other proteins and is believed to help disentangle RNA which has folded abnormally due to drought. Yield improvement due to the expression of *cspB* is primarily due to an increase in the number of kernels per plant (Castiglioni et al., 2008).

Novel genome editing tools, such as CRISPR-Cas9, have also been used to create genetic diversity resulting in improved drought response. Using CRIPR-Cas9, the native maize *GOS2* promoter was used by Shi et al. (2017) to both replace and supplement the native *ARGOS8* promoter. This created variants with altered expression of *ARGOS8*, a negative regulator of ethylene responses. Some variants achieved a yield gain of five bushels per acre under flowering stress conditions, with no yield penalty under well-watered conditions, demonstrating the viability of the editing tool (Shi et al., 2017). Gene knockout study of *ahb2* in maize *via* CRISPR/Cas9 resulted in quicker closure of stomata in response to water deficit stress, and three independent homozygous lines for the *i1*, *d2* and *d35* alleles that are tolerant to drought stress have been obtained (Liu et al., 2020). Gene editing approach will most likely be more attractive since crop varieties derived from this method could be considered non-genetically modified and thus be more acceptable.

Breeding for maize drought tolerance in SSA

Decades of partnerships with public and private sector institutions by the International Maize and Wheat Improvement Center (CIMMYT) have resulted in successful breeding and deployment of elite stress-tolerant maize cultivars across SSA (Edmeades et al., 1996b; Prasanna et al., 2021). CIMMYT leveraged molecular tools such as QTL and MARS to improve maize tolerance to drought, as well as other important stresses such as nitrogen use efficiency, maize streak virus, and maize lethal necrosis (Semagn et al., 2015). Advances in molecular and systems biology provided new opportunities to accelerate the maize improvement progress in SSA (Wossen et al., 2017). As a result, the adoption of improved maize varieties tolerant to drought and other stresses has increased across the region (Chivasa et al., 2022).

Two projects that stand out in SSA are public-private funded collaborations: the Drought Tolerant Maize for Africa (DTMA) and Water Efficient Maize for Africa (WEMA) (Oikeh et al., 2014; Nasser et al., 2020). DTMA is implemented jointly by CIMMYT and the International Institute for Tropical Agriculture, in collaboration with national agricultural research systems in participating nations. WEMA is a partnership involving the Bill & Melinda Gates Foundation, USAID, the Howard G. Buffett Foundation, CIMMYT, Monsanto, the National Agricultural Research Systems (NARS) from the participating nations (Kenya, Uganda, Tanzania, South Africa, Ethiopia and Mozambique), and led by the African Agricultural Technology Foundation (AATF) a nonprofit organization. The projects used conventional breeding, double haploid technology (DHT), transgenic technology, and marker assisted breeding. The objectives of WEMA and DTMA projects were to develop and distribute drought tolerant white maize hybrids that would yield more under drought stress than commercially used varieties (Oikeh et al., 2014; Nasser et al., 2020).

In the DTMA project, researchers used MARS to improve locally adapted germplasm and identify genomic regions associated with drought tolerance. The DTMA project resulted in the registration of 160 drought tolerant maize hybrids for release across 13 SSA countries between 2007 and 2013, with an

adoption rate of between 9%–61% in six countries (Fisher et al., 2015). In 2020, 27 new, multiple-stress-tolerant maize hybrids and open-pollinated varieties were released by small and medium sized enterprises and national agricultural research system (NARS) partners for commercialization in SSA.

In the WEMA project, lines used to create the hybrids were also developed using DHT and selection achieved through MARS (Edge et al., 2017). About 106 non-transgenic drought tolerant maize hybrids were released across five participating countries (Edge et al., 2017), and sold under the name, DroughtTEGO, as well as several genetically modified (GM) varieties marketed under the brand name TELA, which are both drought tolerant and insect resistant from the Bt gene (Oikeh et al., 2014; Edge et al., 2017).

Plant breeding education

The shortage of plant breeders in SSA presents a serious challenge because it limits the development of improved varieties of crops for regional food security (Suza et al., 2016). To sustain the success of the maize drought tolerance projects, a pipeline of plant breeders and crop scientists must be in place in SSA. Support for human capacity enhancement can be modeled using successful partnerships such as Improved Master of Science in Cultivar Development for Africa and Plant Breeding E-Learning in Africa (Suza et al., 2016). In addition, centers in Africa, such as the African Centre for Crop Improvement (ACCI) in South Africa and West Africa Centre for Crop Improvement (WACCI) in Ghana, with funding largely from the Alliance for a Green Revolution in Africa (AGRA) have contributed plant breeding capacity in Africa. Several of the breeders from ACCI and WACCI have identified germplasm that contain useful alleles for drought tolerance improvement of maize inbred lines and hybrids under drought stress (Derera et al., 2008; Adebayo and Menkir, 2015; Meseka et al., 2018; Nasser et al., 2020). Evaluation of these germplasm has been achieved mainly through the utilization of conventional breeding *via* indirect selection using secondary traits drought selection indices. Other studies involving plant breeders from ACCI and WACCI also utilized marker-assisted recurrent selection and reported genetic gain under drought stress ranging from 22.7 kg ha^{-1} yr^{-1}–118 kg ha^{-1} yr^{-1} (Beyene et al., 2016; Bankole et al., 2017; Masuka et al., 2017).

Conclusion

Since the development of hybrid maize, progress in breeding for yield potential has consistently improved drought tolerance. However, new strategies are needed to address the increase in demand and challenges brought about by climate change. While the goal is improved yield under drought conditions and yield stability, selecting on yield alone is inefficient due to the low heritability under stress conditions. Physiological dissection of yield into more heritable and selectable secondary traits can lead to better phenotyping strategies and improved gains. Advances in molecular biology provide new tools to understand the genetic basis of drought tolerance. Such tools have helped to identify associations between QTL and drought tolerance, yet gene interactions such as epistasis and pleiotropy, and gene by environment interactions complicate the use of genomic information for breeding. Integrating molecular genetics with physiology will help untangle the complex network of gene and environmental interactions, promoting the identification of the most promising loci conveying drought tolerance. Ecophysiological modeling can be used to better understand how genetic variability translates into the final phenotype by incorporating genomic information with known principles of crop growth and development. By integrating the vast amount of environmental, genetic, and physiological knowledge and applying it to a well-chosen germplasm base, plant breeders will be able to meet the challenge of drought stress and continue to deliver improved varieties in the future. The DTMA and WEMA projects in SSA emphasize the importance of multidisciplinary approaches that incorporate multiple breeding tools and approaches In addition, financial investment in human and institutional infrastructure capacity strengthening are needed to sustain the application and adoption innovations such as DTMA and WEMA in Africa.

Author contributions

MM and WS designed the manuscript. All authors contributed to writing the manuscript.

Funding

This work was supported by Iowa State University (ISU).

Acknowledgments

Siddique Aboobucker for reviewing the manuscript. We apologize to those colleagues whose important contributions may have been omitted.

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.

Publisher's note

References

Adebayo, M. A., and Menkir, A. (2015). Combining ability of adapted and exotic drought-tolerant maize inbred lines under full irrigation and rainfed conditions in Nigeria. *J. Crop Improv.* 29, 117–130. doi:10.1080/15427528.2014.980484

Adnan, A. A., Diels, J., Jibrin, J. M., Kamara, A. Y., Shaibu, A. S., Craufurd, P., et al. (2020). CERES-Maize model for simulating genotype-by-environment interaction of maize and its stability in the dry and wet savannas of Nigeria. *Field Crops Res.* 253, 107826. doi:10.1016/j.fcr.2020.107826

Bankole, F., Menkir, A., Olaoye, G., Crossa, J., Hearne, S., Unachukwu, N., et al. (2017). Genetic gains in yield and yield related traits under drought stress and favorable environments in a maize population improved using marker assisted recurrent selection. *Front. Plant Sci.* 8, 808. doi:10.3389/fpls.2017.00808

Bänziger, M., and Araus, J.-L. (2007). "Recent advances in breeding maize for drought and salinity stress tolerance," in *Advances in molecular breeding toward drought and salt tolerant crops*. Editors M. A. Jenks, P. M. Hasegawa, and S. M. Jain (Netherlands: Springer), 587–601.

Barbosa, P. A. M., Fritsche-Neto, R., Andrade, M. C., Petroli, C. D., Burgueño, J., Galli, G., et al. (2021). Introgression of maize diversity for drought tolerance: Subtropical maize landraces as source of new positive variants. *Front. Plant Sci.* 12, 691211. doi:10.3389/fpls.2021.691211

Barker, T., Campos, H., Cooper, M., Dolan, D., Edmeades, G., Habben, J., et al. (2005). in *Improving drought tolerance in maizePlant Breeding Reviews*. Editor J. Janick (Oxford, UK: John Wiley & Sons), 173–253.

Bernardo, R., and Yu, J. (2007). Prospects for genomewide selection for quantitative traits in maize. *Crop Sci.* 47 (3), 1082–1090. doi:10.2135/cropsci2006.11.0690

Beyene, Y., Gowda, M., Olsen, M., Robbins, K. R., Pérez-Rodríguez, P., Alvarado, G., et al. (2019). Empirical comparison of tropical maize hybrids selected through genomic and phenotypic selections. *Front. Plant Sci.* 10, 1502. doi:10.3389/fpls.2019.01502

Beyene, Y., Gowda, M., Pérez-Rodríguez, P., Olsen, M., Robbins, K. R., Burgueño, J., et al. (2021). Empirical comparison of tropical maize hybrids selected through genomic and phenotypic selections. *Front. Plant Sci.* 12, 1502. doi:10.3389/fpls.2019.01502

Beyene, Y., Semagn, K., Crossa, J., Mugo, S., Atlin, G. N., Tarekegne, A., et al. (2016). Improving maize grain yield under drought stress and non-stress environments in sub-saharan Africa using marker-assisted recurrent selection. *Crop Sci.* 56 (1), 344–353. doi:10.2135/cropsci2015.02.0135

Blum, A. (2011a). Drought resistance - is it really a complex trait? *Funct. Plant Biol.* 38 (10), 753–757. doi:10.1071/FP11101

Blum, A. (2009). Effective use of water (EUW) and not water-use efficiency (WUE) is the target of crop yield improvement under drought stress. *Field Crops Res.* 112 (2), 119–123. doi:10.1016/j.fcr.2009.03.009

Blum, A. (2014). Genomics for drought resistance – getting down to Earth. *Funct. Plant Biol.* 41 (11), 1191–1198. doi:10.1071/FP14018

Blum, A. (2011b). *Plant breeding for water-limited environments*. New York: Springer.

Bolanos, J., and Edmeades, G. O. (1993). Eight cycles of selection for drought tolerance in lowland tropical maize. I. Responses in grain yield, biomass, and radiation utilization. *Field Crops Res.* 31 (3-4), 233–252. doi:10.1016/0378-4290(93)90064-t

Brown, W. L. (1975). "A broader germplasm base in corn and sorghum", in: Proceedings of the 30th Annual Corn and Sorghum Research Conference.

Buckler, E. S., and Thornsberry, J. M. (2002). Plant molecular diversity and applications to genomics. *Curr. Opin. Plant Biol.* 5 (2), 107–111. doi:10.1016/S1369-5266(02)00238-8

Campos, H., Cooper, M., Habben, J. E., Edmeades, G. O., and Schussler, J. R. (2004). Improving drought tolerance in maize: A view from industry. *Field Crops Res.* 90 (1), 19–34. doi:10.1016/j.fcr.2004.07.003

Castiglioni, P., Warner, D., Navarro, S., Back, S., Fernandes, M., Targolli, J., et al. (2008). Bacterial RNA chaperones confer abiotic stress tolerance in plants and improved grain yield in maize under water-limited conditions. *Plant Physiol.* 147 (2), 446–455. doi:10.1104/pp.108.118828

Castleberry, R. M., Crum, C. W., and Krull, C. F. (1984). Genetic yield improvement of U.S. maize cultivars under varying fertility and climatic environments. *Crop Sci.* 24 (1), 33–36. doi:10.2135/cropsci1984.0011183X002400010008x

Cattivelli, L., Rizza, F., Badeck, F.-W., Mazzucotelli, E., Mastrangelo, A. M., Francia, E., et al. (2008). Drought tolerance improvement in crop plants: An integrated view from breeding to genomics. *Field Crops Res.* 105 (1), 1–14. doi:10.1016/j.fcr.2007.07.004

Challinor, A. J., Koehler, A. K., Ramirez-Villegas, J., Whitfield, S., and Das, B. (2016). Current warming will reduce yields unless maize breeding and seed systems adapt immediately. *Nat. Clim. Chang.* 6 (10), 954–958. doi:10.1038/NCLIMATE3061

Chapman, S. C., Crossa, J., and Edmeades, G. O. (1997). Genotype by environment effects and selection for drought tolerance in tropical maize. I. Two mode pattern analysis of yield. *Euphytica* 95 (1), 01–09. doi:10.1023/A:1002918008679

Chenu, K., Chapman, S. C., Tardieu, F., McLean, G., Welcker, C., and Hammer, G. L. (2009). Simulating the yield impacts of organ-level quantitative trait loci associated with drought response in maize: A "gene-to-phenotype" modeling approach. *Genetics* 183 (4), 1507–1523. doi:10.1534/genetics.109.105429

Chenu, K. (2015). "Chapter 13 - characterizing the crop environment – nature, significance and applications," in *Crop physiology*. Editors Victor O. Sadras and Daniel F. Calderini. Second Edition (Academic Press), 321–348.

Chenu, K., Cooper, M., Hammer, G. L., Mathews, K. L., Dreccer, M. F., and Chapman, S. C. (2011). Environment characterization as an aid to wheat improvement: Interpreting genotype–environment interactions by modelling water-deficit patterns in north-eastern Australia. *J. Exp. Bot.* 62 (6), 1743–1755. doi:10.1093/jxb/erq459

Chivasa, W., Worku, M., Teklewold, A., Setimela, P., Gethi, J., Magorokosho, C., et al. (2022). Maize varietal replacement in Eastern and Southern Africa: Bottlenecks, drivers and strategies for improvement. *Glob. Food Sec.* 32, 100589. doi:10.1016/j.gfs.2021.100589

Claassen, M. M., and Shaw, R. H. (1970). Water deficit effects on corn. II. Grain components. *Agron. J.* 62 (5), 652–655. doi:10.2134/agronj1970.00021962006200050032x

Cooper, M., Gho, C., Leafgren, R., Tang, T., and Messina, C. (2014). Breeding drought-tolerant maize hybrids for the US corn-belt: Discovery to product. *J. Exp. Bot.* 65 (21), 6191–6204. doi:10.1093/jxb/eru064

Cooper, M., Stucker, R. E., DeLacy, I. H., and Harch, B. D. (1997). Wheat breeding nurseries, target environments, and indirect selection for grain yield. *Crop Sci.* 37 (4), 1168–1176. doi:10.2135/cropsci1997.0011183X003700040024x

Cooper, M., van Eeuwijk, F. A., Chapman, S., Podlich, D. W., and Löffler, C. (2006). "Genotype-by-environment interactions under water-limited conditions," in *Drought adaptation in cereals*. Editor J.-M. Ribaut (New York: The Haworth Press), 51–95.

Derera, J., Tongoona, P., Vivek, B. S., and Laing, M. D. (2008). Gene action controlling grain yield and secondary traits in southern African maize hybrids under drought and non-drought environments. *Euphytica* 162, 411–422. doi:10.1007/s10681-007-9582-4

Edge, M., Oikeh, S. O., Kyetere, D., Mugo, S., and Mashingaidze, K. (2017). "Water efficient maize for Africa: A public-private partnership in technology transfer to smallholder farmers in sub-saharan Africa," in *From agroscience to agribusiness*. Editors N. Kalaitzandonakes, E. Carayannis, E. Grigoroudis, and S. Rozakis (Cham: Springer International Publishing), 391–412.

Edmeades, G. O., Bänziger, M., Beck, D. L., Bolaños, J. A., and Ortega, C. A. (1996). "Development and per se performance of CIMMYT maize populations asdrought-tolerant sources," in *Developing drought and low N-tolerant Maize Proceedings of a symposium* El Batán, Mexico, March 25–9, 1996.

G. O. Edmeades, M. Bänziger, H. R. Mickelson, and C. B. Peña-Valdivia (Mexico: CIMMYT.

Edmeades, G. O., Bolanos, J., and Chapman, S. C. (1996). "Value of secondary traits in selecting for drought tolerance in tropical maize," in *Developing drought- and low N-tolerant maize. Proceedings of a symposium, March 25-29, 1996, CIMMYT, El Batán, Mexico. Mexico, D.F.* Editors G. O. Edmeades, M. Bänziger, H. R. Mickelson, and C. B. Peña-Valdivia (Mexico: CIMMYT).

FAO (2021). *The impact of disasters and crises on agriculture and food security*. Rome: Food and agriculture organization of the United Nations.

Faostat (2010). *Statistical database of the food and agriculture organization of the united nations*. Rome: FAO.

Farfan, I. D. B., De La Fuente, G. N., Murray, S. C., Isakeit, T., Huang, P.-C., Warburton, M., et al. (2015). Genome wide association study for drought, aflatoxin resistance, and important agronomic traits of maize hybrids in the sub-tropics. *PLoS One* 10 (2), e0117737. doi:10.1371/journal.pone.0117737

IPCC (2012). "Changes in climate extremes and their impacts on the natural physical environment," in *Managing the risks of extreme events and disasters to advance climate change adaptation*. Editors C. B. B. V. Field, T. F. Stocker, D. Qin, D. J. Dokken, K. L. Ebi, et al. (Cambridge, England: Cambridge University Press).

Fisher, M., Abate, T., Lunduka, R. W., Asnake, W., Alemayehu, Y., and Madulu, R. B. (2015). Drought tolerant maize for farmer adaptation to drought in sub-Saharan Africa: Determinants of adoption in eastern and southern Africa. *Clim. Change* 133 (2), 283–299. doi:10.1007/s10584-015-1459-2

Flowers, T. J. (2004). Improving crop salt tolerance. *J. Exp. Bot.* 55 (396), 307–319. doi:10.1093/jxb/erh003

Gao, Z., Liu, H., Wang, H., Li, N., Wang, D., Song, Y., et al. (2014). Generation of the genetic mutant population for the screening and characterization of the mutants in response to drought in maize. *Chin. Sci. Bull.* 59 (8), 766–775. doi:10.1007/s11434-013-0031-6

Guo, M., Rupe, M. A., Wei, J., Winkler, C., Goncalves-Butruille, M., Weers, B. P., et al. (2014). Maize ARGOS1 (ZAR1) transgenic alleles increase hybrid maize yield. *J. Exp. Bot.* 65 (1), 249–260. doi:10.1093/jxb/ert370

Guo, M., Rupe, M. A., Zinselmeier, C., Habben, J., Bowen, B. A., and Smith, O. S. (2004). Allelic variation of gene expression in maize hybrids. *Plant Cell* 16 (7), 1707–1716. doi:10.1105/tpc.022087

Hajjapoor, A., Nelson, W. C. D., and Vadez, V. (2022). How process-based modeling can help plant breeding deal with G x E x M interactions. *Field Crops Res.* 283, 108554. doi:10.1016/j.fcr.2022.108554

Hammer, G., Cooper, M., Tardieu, F., Welch, S., Walsh, B., van Eeuwijk, F., et al. (2006). Models for navigating biological complexity in breeding improved crop plants. *Trends Plant Sci.* 11 (12), 587–593. doi:10.1016/j.tplants.2006.10.006

Hao, Z., Li, X., Liu, X., Xie, C., Li, M., Zhang, D., et al. (2009). Meta-analysis of constitutive and adaptive QTL for drought tolerance in maize. *Euphytica* 174 (2), 165–177. doi:10.1007/s10681-009-0091-5

Heinemann, A. B., Dingkuhn, M., Luquet, D., Combres, J. C., and Chapman, S. (2008). Characterization of drought stress environments for upland rice and maize in central Brazil. *Euphytica* 162 (3), 395–410. doi:10.1007/s10681-007-9579-z

Heisey, P. W., and Edmeades, G. O. (1999). *Maize production in drought-stressed environments: Technical options and research resource allocation*. Mexico D.F.: CIMMYT.Maize production in drought-stressed environments: Technical options and research resource allocation. Part 1 of CIMMYT 1997/98 world maize facts and trends

Ho, J. C., McCouch, S. R., and Smith, M. E. (2002). Improvement of hybrid yield by advanced backcross QTL analysis in elite maize. *Theor. Appl. Genet.* 105 (2-3), 440–448. doi:10.1007/s00122-002-0945-x

Jansen, R. C., Jannink, J-L., and Beavis, W. D. (2003). Mapping quantitative trait loci in plant breeding populations: Use of parental haplotype sharing. *Crop Sci.* 44 (3), 829–834. doi:10.2135/cropsci2003.8290

Jeanneau, M., Gerentes, D., Foueillassar, X., Zivy, M., Vidal, J., Toppan, A., et al. (2002). Improvement of drought tolerance in maize: Towards the functional validation of the zm-asr1 gene and increase of water use efficiency by over-expressing C4–PEPC. *Biochimie* 84 (11), 1127–1135. doi:10.1016/S0300-9084(02)00024-X

Kamali, B., Abbaspou, K. C., Lehmann, A., Wehrli, B., and Yang, H. (2018). Spatial assessment of maize physical drought vulnerability in sub-Saharan Africa: Linking drought exposure with crop failure. *Environ. Res. Lett.* 12, 074010. doi:10.1088/1748-9326/aacb37

Ke, F., and Ma, X. (2021). Responses of maize hybrids with contrasting maturity to planting date in Northeast China. *Sci. Rep.* 11 (1), 15776. doi:10.1038/s41598-021-95328-5

Khan, S. U., Zheng, Y., Chachar, Z., Zhang, X., Zhou, G., Zong, N., et al. (2022). Dissection of maize drought tolerance at the flowering stage using genome-wide association studies. *Genes* 13, 564. doi:10.3390/genes13040564

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* 9 (1), 29. doi:10.1186/1746-4811-9-29

Levitt, J. (1972). *Responses of plants to environmental stresses*. New York: Academic Press.

Li, C., Sun, B., Li, Y., Liu, C., Wu, X., Zhang, D., et al. (2016). Numerous genetic loci identified for drought tolerance in the maize nested association mapping populations. *BMC genomics* 17 (1), 894. doi:10.1186/s12864-016-3170-8

Liu, S., Li, C., Wang, H., Wang, S., Yang, S., Liu, X., et al. (2020). Mapping regulatory variants controlling gene expression in drought response and tolerance in maize. *Genome Biol.* 21 (1), 163. doi:10.1186/s13059-020-02069-1

Liu, S., and Qin, F. (2021). Genetic dissection of maize drought tolerance for trait improvement. *Mol. Breed.* 41, 8. doi:10.1007/s11032-020-01194-w

Liu, S., Yang, J., Yang, X., Drury, C. F., Jiang, R., and Reynolds, W. D. (2021). Simulating maize yield at county scale in southern Ontario using the decision support system for agrotechnology transfer model. *Can. J. Soil Sci.* 101 (4), 734–748. doi:10.1139/cjss-2020-0116

Lobell, D. B., Schlenker, W., and Costa-Roberts, J. (2011). Climate trends and global crop production since 1980. *Science* 333 (6042), 616–620. doi:10.1126/science.1204531

Lu, Y., Zhang, S., Shah, T., Xie, C., Hao, Z., Li, X., et al. (2010). Joint linkage–linkage disequilibrium mapping is a powerful approach to detecting quantitative trait loci underlying drought tolerance in maize. *Proc. Natl. Acad. Sci. U. S. A.* 107 (45), 19585–19590. doi:10.1073/pnas.1006105107

Lunduka, R. W., Mateva, K. I., Magorokosho, C., and Manjeru, P. (2017). Impact of adoption of drought-tolerant maize varieties on total maize production in south Eastern Zimbabwe. *Clim. Dev.* 11, 35–46. doi:10.1080/17565529.2017.1372269

Mao, H., Wang, H., Liu, S., Li, Z., Yang, X., Yan, J., et al. (2015). A transposable element in a NAC gene is associated with drought tolerance in maize seedlings. *Nat. Commun.* 6, 8326. doi:10.1038/ncomms9326

Marino, R., Ponnaiah, M., Krajewski, P., Frova, C., Gianfranceschi, L., Pè, M. E., et al. (2008). Addressing drought tolerance in maize by transcriptional profiling and mapping. *Mol. Genet. Genomics* 281 (2), 163–179. doi:10.1007/s00438-008-0401-y

Masuka, B., Atlin, G. N., Olsen, M., Magorokosho, C., Labuschagne, M., Crossa, J., et al. (2017). Gains in maize genetic improvement in eastern and southern Africa: I. CIMMYT hybrid breeding pipeline. *Crop Sci.* 57 (1), 168–179. doi:10.2135/cropsci2016.05.0343

Meseka, S., Fakorede, M., Ajala, S., Badu-Apraku, B., and Menkir, A. (2013). Introgression of alleles from maize landraces to improve drought tolerance in an adapted germplasm. *J. Crop Improv.* 27 (1), 96–112. doi:10.1080/15427528.2012.729259

Meseka, S., Menkir, A., Bossey, B., and Mengesha, W. (2018). Performance assessment of drought tolerant maize hybrids under combined drought and heat stress. *Agronomy* 8, 274. doi:10.3390/agronomy8120274

Messina, C. D., Podlich, D., Zhanshan, D., Samples, M., and Cooper, M. (2011). Yield–trait performance landscapes: From theory to application in breeding maize for drought tolerance. *J. Exp. Bot.* 62 (3), 855–868. doi:10.1093/jxb/erq329

Messina, C., McDonald, D., Poffenberger, H., Clark, R., Salinas, A., Fang, Y., et al. (2021). Reproductive resilience but not root architecture underpins yield improvement under drought in maize. *J. Exp. Bot.* 72, 5235–5245. doi:10.1093/jxb/erab231

Meuwissen, T. H. E., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157 (4), 1819–1829. doi:10.1093/genetics/157.4.1819

Mir, R. R., Zaman-Allah, M., Sreenivasulu, N., Trethowan, R., and Varshney, R. K. (2012). Integrated genomics, physiology and breeding approaches for improving drought tolerance in crops. *Theor. Appl. Genet.* 125 (4), 625–645. doi:10.1007/s00122-012-1904-9

Monneveux, P., and Ribaut, J.-M. (2006). "Secondary traits for drought tolerance improvement in cereals," in *Drought adaptation in cereals*. Editor J.-M. Ribaut (Binghamton, NY: The Haworth Press, Inc;), 97–143.

Myles, S., Peiffer, J., Brown, P. J., Ersoz, E. S., Zhang, Z., Costich, D. E., et al. (2009). Association mapping: Critical considerations shift from genotyping to experimental design. *Plant Cell* 21 (8), 2194–2202. doi:10.1105/tpc.109.068437

Nasser, L. M., Badu-Apraku, B., Gracen, V. E., and Mafouasson, N. A. (2020). Combining ability of early-maturing Yellow maize inbreds under combined drought and heat stress and well-watered environments. *Agronomy* 10 (10), 1585. doi:10.3390/agronomy10101585

Nelson, D. E., Repetti, P. P., Adams, T. R., Creelman, R. A., Wu, J., Warner, D. C., et al. (2007). Plant nuclear factor Y (NF-Y) B subunits confer drought tolerance and lead to improved corn yields on water-limited acres. *Proc. Natl. Acad. Sci. U. S. A.* 104 (42), 16450–16455. doi:10.1073/pnas.0707193104

Nuccio, M. L., Wu, J., Mowers, R., Zhou, H.-P., Meghji, M., Primavesi, L. F., et al. (2015). Expression of trehalose-6-phosphate phosphatase in maize ears improves yield in well-watered and drought conditions. *Nat. Biotechnol.* 33 (8), 862–869. doi:10.1038/nbt.3277

Oikeh, S., Ngonyamo-Majee, D., Mugo, S. I. N., Mashingaidze, K., Cook, V., and Stephens, M. (2014). "The water efficient maize for Africa project as an example of a public–private partnership," in *Convergence of food security, Energy security and sustainable agriculture. Biotechnology in agriculture and Forestry*. Editors D. Songstad, J. Hatfield, and D. Tomes (Berlin, Heidelberg: Springer). doi:10.1007/978-3-642-55262-5_13

Otegui, M. E., and Andrade, F. H. (2000). "New relationships between light interception, ear growth, and kernel set in maize," in *Physiology and modeling kernel set in maize*. Editors M. Westgate and K. Boote (Baltimore, Maryland: CSSA Special Publication Number 29), 89–102.

Podlich, D. W., and Cooper, M. (1998). QU-GENE: A simulation platform for quantitative analysis of genetic models. *Bioinformatics* 14 (7), 632–653. doi:10.1093/bioinformatics/14.7.632

Podlich, D. W., Winkler, C. R., and Cooper, M. (2004). Mapping as you go: An effective approach for marker-assisted selection of complex traits. *Crop Sci.* 44 (5), 1560–1571. doi:10.2135/cropsci2004.1560

Ramirez-Villegas, J., Milan, A. M., Alexandrov, N., Asseng, S., Challinor, A. J., Crossa, J., et al. (2020). CGIAR modeling approaches for resource-constrained scenarios: I. Accelerating crop breeding for a changing climate. *Crop Sci.* 60, 547–567. doi:10.1002/csc2.20048

Resende, M. D. V. d., Silva, F. F. e., Resende, M. F. R., and Azevedo, C. F. (2014). "Chapter 5 - genome-wide selection (GWS)," in *Biotechnology and plant breeding*. Editors A. Borem and R. Fritsche-Neto (San Diego: Academic Press), 105–133.

Reymond, M., Muller, B., Leonardi, A., Charcosset, A., and Tardieu, F. (2003). Combining quantitative trait loci analysis and an ecophysiological model to analyze the genetic variability of the responses of maize leaf growth to temperature and water deficit. *Plant Physiol.* 131 (2), 664–675. doi:10.1104/pp.013839

Ribaut, J.-M., Betran, J., Monneveux, P., and Setter, T. (2009). "Drought tolerance in maize," in *Handbook of maize: Its biology*. Editors J. L. Bennetzen and S. C. Hake (New York: Springer), 311–344.

Ribaut, J.-M. (2006). *Drought adaptation in cereals*. Binghamton, NY: The Haworth Press.

Ribaut, J.-M., and Ragot, M. (2006). Marker-assisted selection to improve drought adaptation in maize: The backcross approach, perspectives, limitations, and alternatives. *J. Exp. Bot.* 58 (2), 351–360. doi:10.1093/jxb/erl214

Riedelsheimer, C., Lisec, J., Czedik-Eysenberg, A., Sulpice, R., Flis, A., Grieder, C., et al. (2012). Genome-wide association mapping of leaf metabolic profiles for dissecting complex traits in maize. *Proc. Natl. Acad. Sci. U. S. A.* 109 (23), 8872–8877. doi:10.1073/pnas.1120813109

Rosegrant, M. R., Ringler, C., Sulser, T. B., Ewing, M., Palazzo, A., Zhu, T., et al. (2009). *Agriculture and food security under global change: Prospects for 2025/2050*. Washington, DC: International Food Policy Research Institute.

Ruswandi, D., AgustianAnggia, E. P., Canama, A., Marta, H., Ruswandi, S., et al. (2014). Mutation breeding of maize for anticipating global climate change in Indonesia. *Asian J. Agric. Res.* 8, 234–247. doi:10.3923/ajar.2014.234.247

Sah, R. P., Chakraborty, M., Prasad, K., Pandit, M., Tudu, V. K., Chakravarty, M. K., et al. (2020). Impact of water deficit stress in maize: Phenology and yield components. *Sci. Rep.* 10, 2944–2945. doi:10.1038/s41598-020-59689-7

Salvi, S., Sponza, G., Morgante, M., Tomes, D., Niu, X., Fengler, K. A., et al. (2007). Conserved noncoding genomic sequences associated with a flowering-time quantitative trait locus in maize. *Proc. Natl. Acad. Sci. U. S. A.* 104 (27), 11376–11381. doi:10.1073/pnas.0704145104

Sammons, D. J. (1987). *Origin and early history of maize in East Africa with special reference to Kenya*. Njoro, Kenya: Egerton University Research Paper Series, 2–11.

Sawkins, M. C., Farmer, A. D., Hoisington, D., Sullivan, J., Tolopko, A., Jiang, Z., et al. (2004). Comparative map and trait viewer (CMTV): An integrated bioinformatic tool to construct consensus maps and compare QTL and functional genomics data across genomes and experiments. *Plant Mol. Biol.* 56 (3), 465–480. doi:10.1007/s11103-004-4950-0

Schlenker, W., and Lobell, D. B. (2010). Robust negative impacts of climate change on African agriculture. *Environ. Res. Lett.* 5 (1), 014010. doi:10.1088/1748-9326/5/1/014010

Semagn, K., Beyene, Y., Babu, R., Nair, S., Gowda, M., Das, B., et al. (2015). Quantitative trait loci mapping and molecular breeding for developing stress resilient maize for sub-saharan Africa. *Crop Sci.* 55 (4), 1449–1459. doi:10.2135/cropsci2014.09.0646

Semagn, K., Beyene, Y., Warburton, M. L., Tarekegne, A., Mugo, S., Meisel, B., et al. (2013). Meta-analyses of QTL for grain yield and anthesis silking interval in 18 maize populations evaluated under water-stressed and well-watered environments. *BMC genomics* 14 (1), 313. doi:10.1186/1471-2164-14-313

Setter, T. L., Yan, J., Warburton, M., Ribaut, J.-M., Xu, Y., Sawkins, M., et al. (2011). Genetic association mapping identifies single nucleotide polymorphisms in genes that affect abscisic acid levels in maize floral tissues during drought. *J. Exp. Bot.* 62 (2), 701–716. doi:10.1093/jxb/erq308

Seyoum, S., Chauhan, Y., Rachaputi, R., Fekybelu, S., and Boddupalli, P. (2017). Characterising production environments for maize in eastern and southern Africa using the APSIM Model. *Agric. For. Meteorology* 247, 445–453. doi:10.1016/j.agrformet.2017.08.023

Shi, J., Gao, H., Wang, H., Lafitte, H. R., Archibald, R. L., Yang, M., et al. (2017). ARGOS8 variants generated by CRISPR-Cas9 improve maize grain yield under field drought stress conditions. *Plant Biotechnol. J.* 15 (2), 207–216. doi:10.1111/pbi.12603

Shiferaw, B., Prasanna, B. M., Hellin, J., and Bänziger, M. (2011). Crops that feed the world 6. Past successes and future challenges to the role played by maize in global food security. *Food Secur.* 3 (3), 307–327. doi:10.1007/s12571-011-0140-5

Sinclair, T. R., and Purcell, L. C. (2005). Is a physiological perspective relevant in a 'genocentric' age? *J. Exp. Bot.* 56 (421), 2777–2782. doi:10.1093/jxb/eri297

Singh, B. D. (2010). *Plant breeding: Principles and methods*. New Delhi - Ludhiana: Kalyani Publishers.

Suza, W. P., Gibson, P., Edema, R., Akromah, R., Sibiya, J., Madakadze, R., et al. (2016). Plant breeding capacity building in Africa. *Nat. Clim. Chang.* 6, 976. doi:10.1038/nclimate3139

Tanksley, S. D., and Nelson, J. C. (1996). Advanced backcross QTL analysis: A method for the simultaneous discovery and transfer of valuable QTLs from unadapted germplasm into elite breeding lines. *Theor. Appl. Genet.* 92 (2), 191–203. doi:10.1007/BF00223376

Tardieu, F. (2012). Any trait or trait-related allele can confer drought tolerance: Just design the right drought scenario. *J. Exp. Bot.* 63 (1), 25–31. doi:10.1093/jxb/err269

Tardieu, F. (2006). "Leaf growth under water-limited conditions," in *Drought adaptation in cereals*. Editor J.-M. Ribaut (Binghamton, NY: The Haworth Press).

Tardieu, F., and Tuberosa, R. (2010). Dissection and modelling of abiotic stress tolerance in plants. *Curr. Opin. Plant Biol.* 13 (2), 206–212. doi:10.1016/j.pbi.2009.12.012

Tesfaye, K., Kruseman, G., Cairns, J. E., Zaman-Allah, M., Wegary, D., Zaidi, P. H., et al. (2018). Potential benefits of drought and heat tolerance for adapting maize to climate change in tropical environments. *Clim. Risk Manag.* 19, 106–119. doi:10.1016/j.crm.2017.10.001

Tollenaar, M., and Lee, E. (2006). Physiological dissection of grain yield in maize by examining genetic improvement and heterosis. *Maydica* 51 (2), 399. doi:10.2135/cropsci2007.04.0010IPBS

Tsonev, S., Todorovska, E., Avramova, V., Kolev, S. E., Abu-Mhadi, N., and Christov, N. K. (2009). Genomics assisted improvement of drought tolerance in maize: QTL approaches. *Biotechnol. Biotechnol. Equip.* 23, 1410–1413. doi:10.2478/V10133-009-0004-8

Tuberosa, R., and Salvi, S. (2006). Genomics-based approaches to improve drought tolerance of crops. *Trends Plant Sci.* 11 (8), 405–412. doi:10.1016/j.tplants.2006.06.003

Tuberosa, R., Salvi, S., Giuliani, S., Sanguineti, M. C., Bellotti, M., Conti, S., et al. (2007). Genome-wide approaches to investigate and improve maize response to drought. *Crop Sci.* 47, S-120–S-141. doi:10.2135/cropsci2007.04.0001IPBS

Tuberosa, R., Salvi, S., Sanguineti, M. C., Landi, P., Maccaferri, M., and Conti, S. (2002). Mapping QTLs regulating morpho-physiological traits and yield: Case studies, shortcomings and perspectives in drought-stressed maize. *Ann. Bot.* 89 (7), 941–963. doi:10.1093/aob/mcf134

Vargas, M., van Eeuwijk, F. A., Crossa, J., and Ribaut, J.-M. (2006). Mapping QTLs and QTL x environment interaction for CIMMYT maize drought stress program using factorial regression and partial least squares methods. *Theor. Appl. Genet.* 112 (6), 1009–1023. doi:10.1007/s00122-005-0204-z

Wang, N., Cheng, M., Chen, Y., Liu, B., Wang, X., Li, G., et al. (2021). Natural variations in the non-coding region of ZmNAC080308 contributes maintaining grain yield under drought stress in maize. *BMC Plant Biol.* 21, 305. doi:10.1186/s12870-021-03072-9

Wang, X., Wang, H., Liu, S., Ferjani, A., Li, J., Yan, J., et al. (2016). Genetic variation in ZmVPP1 contributes to drought tolerance in maize seedlings. *Nat. Genet.* 48, 1233–1241. doi:10.1038/ng.3636

Wang, Z., Zhao, X., Ren, Z., Abou-Elwafa, S. F., Pu, X., Zhu, Y., et al. (2022). ZmERF21 directly regulates hormone signaling and stress-responsive gene expression to influence drought tolerance in maize seedlings. *Plant Cell Environ.* 45, 312–328. doi:10.1111/pce.14243

Welcker, C., Boussuge, B., Bencivenni, C., Ribaut, J. M., and Tardieu, F. (2006). Are source and sink strengths genetically linked in maize plants subjected to water deficit? A QTL study of the responses of leaf growth and of anthesis-silking interval to water deficit. *J. Exp. Bot.* 58 (2), 339–349. doi:10.1093/jxb/erl227

Westgate, M. E., and Boyer, J. S. (1986). Reproduction at low and pollen water potentials in maize. *Crop Sci.* 26 (5), 951–956. doi:10.2135/cropsci1986.0011183X002600050023x

White, G. C., Chapman, L. M., Slate, W. L., Jr., and Brown, B. A. (1922). A comparison of early, medium and late maturing varieties of silage corn for milk production: A report of progress. *J. Dairy Sci.* 5 (4), 333–347. doi:10.3168/jds.s0022-0302(22)94161-x

Wossen, T., Abdoulaye, T., Alene, A., Feleke, S., Menkir, A., and Manyong, V. (2017). Measuring the impacts of adaptation strategies to drought stress: The case of drought tolerant maize varieties. *J. Environ. Manage.* 203 (1), 106–113. doi:10.1016/j.jenvman.2017.06.058

Yamada, T., and Dwiyanti, M. (2013). "Molecular mapping and breeding for genes/QTLS related to climate change," in *Genomics and breeding for climate -resilient crops*. Editor C. Kole (Berlin, Heidelberg: Springer Berlin Heidelberg), 179–212.

Yang, S., Vanderbeld, B., Wan, J., and Huang, Y. (2010). Narrowing down the targets: Towards successful genetic engineering of drought-tolerant crops. *Mol. Plant* 3 (3), 469–490. doi:10.1093/mp/ssq016

Yu, J., Holland, J. B., McMullen, M. D., and Buckler, E. S. (2008). Genetic design and statistical power of nested association mapping in maize. *Genetics* 178 (1), 539–551. doi:10.1534/genetics.107.074245

Zhang, X., Warburton, M. L., Setter, T., Liu, H., Xue, Y., Yang, N., et al. (2016). Genome-wide association studies of drought-related metabolic changes in maize using an enlarged SNP panel. *Theor. Appl. Genet.* 129 (8), 1449–1463. doi:10.1007/s00122-016-2716-0

Zhao, X., Peng, Y., Zhang, J., Fang, P., Wu, B., and Wang, Y. (2018). Shenmai injection protects mitochondria from oxidative injury in myocardial cells and its mechanism. *Crop Sci.* 58, 507–513. doi:10.3785/j.issn.1008-9292.2018.10.10

Zinselmeier, C., Sun, Y., Helentjaris, T., Beatty, M., Yang, S., Smith, H., et al. (2002). The use of gene expression profiling to dissect the stress sensitivity of reproductive development in maize. *Field Crops Res.* 75 (2), 111–121. doi:10.1016/S0378-4290(02)00021-7

Ziyomo, C., and Bernardo, R. (2013). Drought tolerance in maize: Indirect selection through secondary traits versus genome wide selection. *Crop Sci.* 53 (4), 1269–1275. doi:10.2135/cropsci2012.11.0651

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fgene.2022.1089389

OPEN ACCESS

EDITED BY
Mahendar Thudi,
Dr. Rajendra Prasad Central Agricultural University, India

REVIEWED BY
Lixian Qiao,
Qingdao Agricultural University, China
Dongmei Bai,
Shanxi Academy of Agricultural Sciences, China

*CORRESPONDENCE
Xinyou Zhang,
haasz@126.com
Suoyi Han,
suoyi_han@126.com

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 04 November 2022
ACCEPTED 28 November 2022
PUBLISHED 04 January 2023

CITATION
Guo J, Qi F, Qin L, Zhang M, Sun Z, Li H, Cui M, Zhang M, Li C, Li X, Zhao Q, Luo D, Tian M, Liu H, Xu J, Miao L, Huang B, Dong W, Han S and Zhang X (2023), Mapping of a QTL associated with sucrose content in peanut kernels using BSA-seq.
Front. Genet. 13:1089389.
doi: 10.3389/fgene.2022.1089389

Mapping of a QTL associated with sucrose content in peanut kernels using BSA-seq

Junjia Guo, Feiyan Qi, Li Qin, Maoning Zhang, Ziqi Sun, Hongyan Li, Mengjie Cui, Mengyuan Zhang, Chenyu Li, Xiaona Li, Qi Zhao, Dandan Luo, Mengdi Tian, Hua Liu, Jing Xu, Lijuan Miao, Bingyan Huang, Wenzhao Dong, Suoyi Han* and Xinyou Zhang*

Henan Academy of Crops Molecular Breeding/Postgraduate T&R Base of Zhengzhou University/The Shennong Laboratory/Key Laboratory of Oil Crops in Huang-Huai-Hai Plains, Ministry of Agriculture/Henan Provincial Key Laboratory for Oil Crop Improvement/National Centre for Plant Breeding, Xinxiang, Henan, China

As an important factor affecting the edible quality of peanut kernels, sucrose content is a complex quantitative trait regulated by multiple factors. In this study, an F_2 segregating population and a recombinant inbred line (RIL) population, derived from a cross between the high sucrose content variety Jihuatian 1 and the low sucrose content line PI478819, were used as materials to map a quantitative trait locus (QTL) associated with sucrose content in peanut kernels. Four QTLs were initially located on chromosomes A03 and A06 based on BSA-seq technology, and multiple kompetitive allele-specific PCR markers were developed based on single-nucleotide polymorphisms (SNPs) in the intervals. The markers were genotyped in the RIL population and finely mapped to a stable QTL, *qSUCA06*, located on chromosome A06 within a 0.29-Mb physical genomic interval (112367085–112662675 bp), which accounted for 31.95%–41.05% of the phenotypic variance explained. SNP and insertion/deletion annotations were performed on genes in the candidate interval, and having screened out those genes with mutations in exons, candidate genes were verified by qRT-PCR. The results revealed that *Arahy.Y2LWD9* may be the main gene regulating sucrose content. The QTL identified in this study will not only contribute to marker-assisted breeding for improvement of peanut sucrose content but also paves the way for identifying gene function.

KEYWORDS

peanut, sucrose content, BSA-seq, QTL, KASP

Abbreviations: AFLP, amplified fragment length polymorphism; ANOVA, analysis of variance; ACBP, acyl-CoA-binding protein; BSA, bulk segregation analysis; ED, Euclidean distance; GATK, The genome analysis toolkit; KASP, Kompetitive allele-specific PCR; LG, linkage group; LOD, log of odds (genetic linkage score); NIR, near infrared; PVE, phenotypic variation explained; QTL, quantitative trait locus; RIL, recombinant inbred line; qRT-PCR, quantitative real-time PCR; RFLP, restriction fragment length polymorphism.

Introduction

Peanuts (*Arachis hypogaea* L.), an important oil and cash crop, are rich in vegetable oil and protein, and widely cultivated worldwide (Liu et al., 2020a). In recent years, during which there have been increases in the production and consumption of edible peanuts, increasing attention has focused on the edible quality of peanut kernels, an important index of which is sweetness. Indeed, some studies have reported correlation values of as high as 0.88 between sweetness and peanut kernel taste quality (Pattee et al., 1998). The most direct factor affecting the sweetness of peanut is the content of soluble sugars in kernels. These sugars consist primarily of sucrose, fructose, and glucose, among which, sucrose accounts for the largest proportion, and makes the largest contribution to the sweetness of peanuts (McDaniel et al., 2012). Given that peanuts with kernel sucrose contents exceeding 6% are considered to have a better taste (P), determining the main genetic loci controlling sucrose content in kernels would make a valuable contribution to enhancing the sucrose content and edible quality of peanuts.

In plants, sucrose, the main product of leaf photosynthesis, is exported to different non-photosynthetic organs according to demands for the synthesis of carbon and storage materials required for growth (Davis and Dean, 2016). Sucrose transported to the developing seeds is synthesized to yield lipid (oil) or protein storage substances under the action of a series of enzymes such as invertase. Genetic studies have shown that there are multiple factors affecting the sucrose content in kernels, including the influences of environmental factors and plant genotype (Pattee et al., 1981), maturity (Sanders and Bett, 1995), and genotype–environment interactions (Pattee et al., 2000). Furthermore, it has been established that there is a significant difference in the kernel sucrose contents of plants derived from direct and reciprocal crosses, which tends to indicate that this trait is matrilineally determined (Isleib et al., 2004). Collectively, the aforementioned findings provide evidence to indicate that the sucrose content of peanut kernels is a complex quantitative character influenced by multiple factors.

Bulked-segregant analysis (BSA) can be applied to rapidly and efficiently mine causal genes without the necessity of constructing a genetic map. The technique, which uses amplified fragment length polymorphic (AFLP) and restriction fragment length polymorphic (RFLP) markers, was initially used in lettuce and tomato (Giovannoni et al., 1991; Michelmore et al., 1991). The principle of the BSA-seq method is based on the selection of individuals in a population with bipolar characteristics to construct mixed pools, the whole genomes of which are sequenced to identify causal genes associated with traits of interest (Liu et al., 2020b). With the emergence and development of high-throughput sequencing technology, given its high efficiency, the BSA-seq method has been widely used in the analysis of important agronomic characters of soybean (Xie et al., 2021), rice (Yang et al., 2021), sesame (Sheng et al., 2021), and other crops. In peanut, this method has been used for quantitative trait gene mining for traits such as fresh seed dormancy (Kumar et al., 2020), seed coat color (Chen et al., 2021), and late leaf spot resistance (Clevenger et al., 2018; Han et al., 2022).

In this study, using peanut genotypes Jihuatian one and PI478819 as parental plants, we used a combination of BSA-seq and kompetitive allele-specific PCR (KASP) markers to determine the quantitative trait locus (QTL) controlling the sucrose content of peanut kernels and to predict candidate genes. Our findings will provide a theoretical basis for further elucidation of the control of sucrose content in peanuts, and thereby contribute to breeding for enhanced edible quality.

Materials and methods

Plant materials and phenotypic evaluation

In the present study, we used the peanut genotypes Jihuatian 1 and PI478819. The high-sucrose variety Jihuatian 1 is a Spanish-type cultivar developed by the Hebei Academy of Agriculture and Forestry Sciences, China, whereas the low-sucrose line PI478819 is a Virginia-type variety introduced from the United States. These germplasms were used as the female and male parents, respectively, which were crossed to obtain an F_1 population. KASP molecular marker technique was used to identify true and false hybrids of F_1 seeds. The KASP markers with obvious differences between parents were designed, and the genomic DNA of parents and F_1 seeds were extracted and detected by KASP molecular markers. The homozygous type with the same genotype as the parent was false hybrid, and the heterozygous genotype was expressed as true hybrid (Qin et al., 2020). A subsequent F_2 segregating population consisting of 831 lines was obtained by selfing. In addition, a population of recombinant inbred lines (RILs) was obtained based on single-seed descent. F_2 population and parental individuals were planted on the experimental farm of Henan Academy of Agricultural Sciences in Xinxiang (Henan province) in May 2017, and a total of 251 lines of the RIL population were planted in Xinxiang (Henan), Kaifeng (Henan), and Zhumadian (Henan) in May 2021. Seeds of both the F_2 and RIL populations were planted individually in holes. For the RIL population, RILs were planted in a randomized complete block design with two replications. Each RIL comprising 10 plants in one replicate was planted in a single row with inter-plant spacing of 0.2 m and inter-row spacing of 0.5 m in each of the testing environments. Crop management was conducted following regular agricultural practices (Zhang et al., 2022).

Mature pods harvested from the experimental plants were naturally sun-dried, and the sucrose contents of peanut kernels were measured using a near infra-red (NIR) spectrometer

(DA7200; Perten). For measurement, we selected three replicates of approximately 20 uniform kernels, which were evenly packed into a sample cup, and NIR spectral information was collected in the wavelength range 950–1,650 nm (Qin et al., 2016).

Mixed pool construction and whole-genome resequencing

Young leaves were collected from all F_2 population lines and the parent plants, from which genomic DNAs were extracted using a plant genomic DNA extraction kit (DP305-03; TianGen), followed by determination of DNA quality. On the basis of the determined sucrose contents of F_2 individuals, we selected 20 plants with extremely high sucrose content and 20 with low sucrose content, the respective DNAs of which were mixed in equal quantities to give two extreme phenotypic mixed pools. DNAs from these two mixed pools and both parents were then subjected to whole-genome resequencing using the Illumina HiSeq/DNBSEQ platform in conjunction with a double-terminal 150-bp sequencing strategy. The sequencing depth was 20×, and the reference genome used was the Tifrunner_V20190521 version of cultivated peanut (https://www.peanutbase.org/).

Data analysis and filtering

For the detection of SNP and insertion/deletion (InDel) variants, we used GATK software (McKenna et al., 2010), and SnpEff software (Cingolani et al., 2012) was used to perform variant annotation and predict variant impact. In order to obtain high-quality SNPs for association analysis, the SNPs were initially filtered by removing SNP loci with multiple genotypes, and then those loci with read support values of less than 4. Parent SNP information was then used filter out those sites with different phenotypes derived from the same parent. The SNPs that remained were deemed credible.

BSA-seq analysis

SNP-index is a marker association analysis method based on differences in genotype frequencies of mixed pools (Fekih et al., 2013). The main purpose of this method is to detect significant differences in the genotype frequencies of mixed pools using Δ (SNP-index) statistics (Fekih et al., 2013). The stronger the correlation between marker SNPs and traits, the closer Δ (SNP-index) is to 1. The Euclidean distance (ED) algorithm is a method where by sequencing data is used to detect significant differences between the markers of mixed pools and to evaluate the intervals associated with traits (Hill et al., 2013). In the context of the present study, with the exception of differences in sucrose content-related sites, other sites of the two mixed pools constructed using the BSA should be relatively consistent, and consequently, the ED values of non-target sites should be approximately 0. In contrast, the higher the ED value, the greater is the difference of the marker between the two mixed pools. In this study, we used *G* statistics for the purpose of gene detection. The *G* value of each SNP is calculated according to the allele sequencing depth, and is weighted according to the physical distance of the adjacent SNP (Magwene et al., 2011). In addition, given that *G* values are close to the lognormal distribution, the non-parametric estimation of the zero distribution of *G* values can be used to estimate the *p*-value of each SNP (Mansfeld and Grumet, 2018). When the values of *P* and the false discovery rate are both less than 0.01, it is considered that this interval may be the main effect area affecting the trait of interest (i.e., sucrose content in the present study).

Associations among the Δ (SNP-index), ED, *G* statistics, and *p* values of the SNP loci were analyzed using the website https://github.com/xiekunwhy/bsa. The four methods used are all based on a 2-Mb sliding window with a step size of 10 kb, which is applied to calculate the average and smooth the map. A 99% confidence level was selected as the threshold for screening, and the window above the confidence level was defined as the area associated with sucrose content. The intervals obtained using the four correlation analysis methods were compared, and the overlapping interval was regarded as the QTL interval associated sucrose content. The genes and polymorphic sites in the candidate interval were annotated using the website https://www.peanutbase.org/.

Development of KASP markers and verification of the initial positioning results

Young leaves were collected from RIL population plants, and genomic DNAs were extracted using a plant genomic DNA extraction kit (DP305-03, TianGen). On the basis of the differential SNP information obtained for the two parents Jihuatian 1 and PI478819 in the initial mapping interval of the QTL, we designed 23 pairs of KASP primers using Primer Premier 5.0. FAM or HEX fluorescent splice sequences were attached to the 5′ ends of the primers and synthesized by the LGC Genomics company. The PCR reaction mixtures used contained the following: 1 μl of template DNA at a concentration of 50–100 ng/μl and 1 μl of a mixture of 1× Master Mix and Primer Mix. We performed LGC water bath PCR amplification, using the following amplification program: pre-denaturation at 94°C for 15 min; 10 cycles of denaturation at 94 °C and extension at 55°C–61°C for 1 min; 26 cycles of denaturation at 94°C and extension at 55°C for 1 min; and preservation at 10 °C. After the reactions were completed, the genotypes of each site were determined using the SNPline genotyping platform (Majeed et al., 2018).

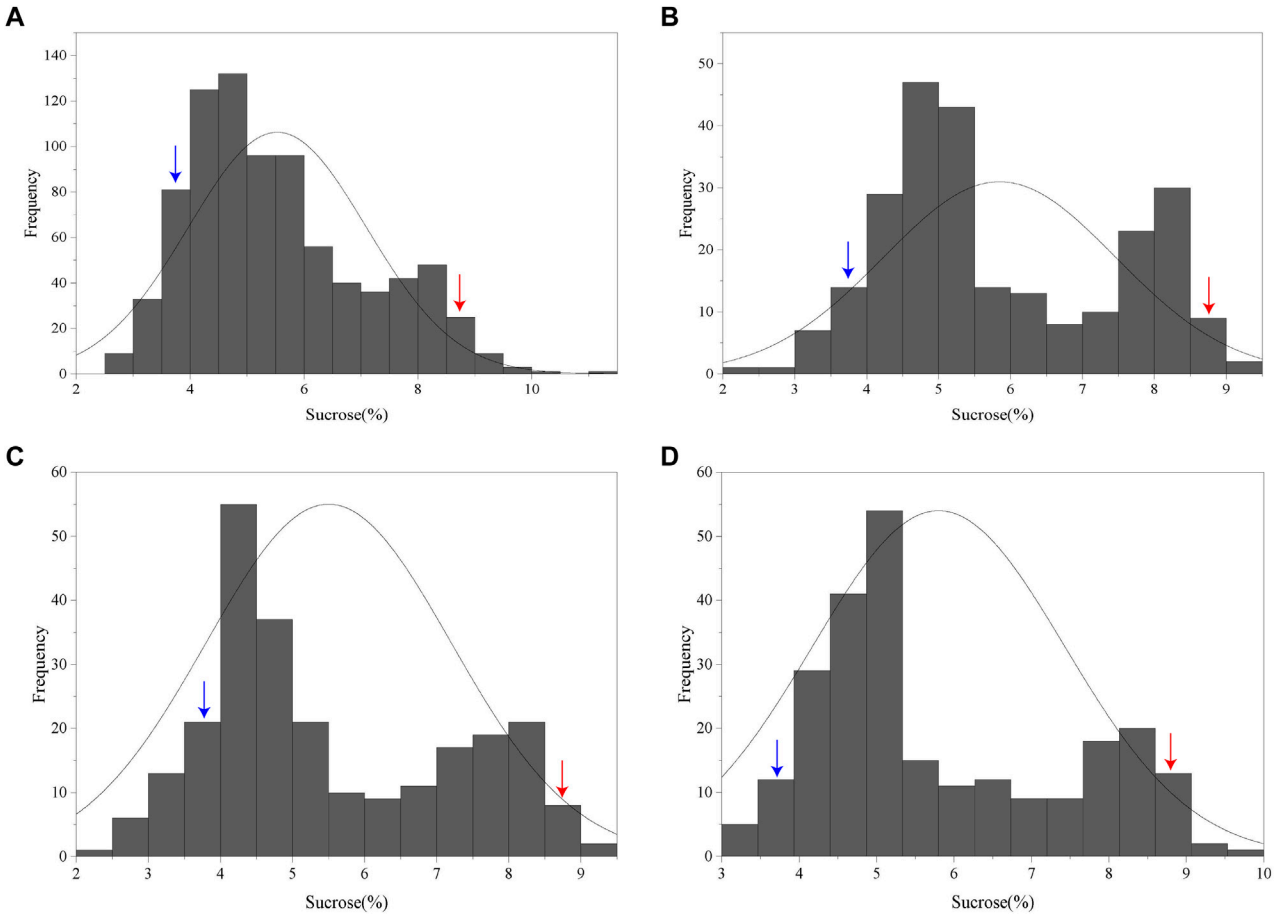


FIGURE 1
Frequency distribution of sucrose content in F_2 and RIL populations of peanut (*Arachis hypogaea*). **(A)** Frequency distribution of sucrose contents in the F_2 population. **(B–D)** Frequency distribution of sucrose content in the RIL population planted in Xinxiang, Kaifeng, and Zhumadian (Henan Province). The blue and red arrows indicate the sucrose contents of the male parent PI478819 and female parent Jihuatian 1, respectively.

TABLE 1 Variation of sucrose content in different populations.

Population	Mean	SD	CV(%)	Min	Max	Kurt	Skew
F_2	5.53	1.56	0.28	2.52	11.03	−0.29	0.68
F_9-XX	5.89	1.68	0.29	2.26	9.34	−1.11	0.34
F_9-KF	5.64	1.79	0.32	2.27	9.67	−0.92	0.45
F_9-ZMD	5.91	1.69	0.29	2.78	9.92	−1.00	0.50

Note: F_2: F_2 segregation population; F_9-XX: F_9 RIL, population planted in Xinxiang; F_9-KF: F_9 RIL, population planted in Kaifeng; F_9-ZMD: F_9 RIL, population planted in Zhumadian. same as below.

QTLs for the sucrose contents in plants cultivated in each environment and at different stages of growth were detected based on the replication mean using QTL IciMapping (Li et al., 2007; Meng et al., 2015), setting the mapping step size as 1 cM and the logarithm of odds (LOD) threshold as 3.0. The QTL region of LG06 was drawn using MapChart 2.3 (Voorrips, 2002). QTLs were designated as follows: q+ the abbreviated trait name + linkage group number, or named as q+ the abbreviated trait name + linkage group number + a number designating one of multiple QTLs in a single linkage group, following the International Rules of Genetic Nomenclature (Liu et al., 2008).

TABLE 2 Analysis of variance for sucrose content in RIL population.

Sucrose	df	SS	MS	F-value	*p*-value
Genotype	250	3823.557	15.294	37.403	<0.01
Environment	2	34.744	17.372	42.484	<0.01
Genotype×Environment	500	262.067	0.524	1.282	<0.01
Error	753	307.905	0.409		

Candidate gene analysis

On the basis of BSA-seq analysis and fine mapping combined with gene annotation information, we performed a preliminary determination of candidate genes. Following a previously described procedure (Pattee et al., 1974), kernel tissue were collected from both parents at 20, 35, 50, and 60 days after flowering (stages S1–S4), with three biological replicates for each period. S1 is the early development stage, S2 and S3 are the developing stages, and S4 is the seed maturity stage. Total RNA was extracted from the collected tissues using a RNAprep Pure Plant Plus Kit (DP441, TIANGEN) and the concentration and purity of the extracted RNA were examined. High-quality RNA samples were selected based on the obtained purity values and concentration values were used to determine the amount of RNA template. The isolated RNA was subsequently reversed transcribed to cDNA using a FastKing RT Kit (With gDNase) (KR116, TIANGEN), and the cDNA thus obtained was diluted with sterile double-distilled water. qPCR reaction systems were prepared according to the requirements of a PowerUp SYBR Green Master Mix kit. A Quant Studio 5 real-time quantitative PCR instrument was used to run the reactions, and the $2^{-\Delta\Delta CT}$ method was used to determine gene expression levels (Livak and Schmittgen, 2001). For each sample, we assessed three biological replicates, for each of which, we also analyzed three technical replicates. The relative expression of candidate genes at the different developmental stages of Jihuatian 1 and PI4788 was determined based on normalization analysis of the gene expression data, using the *ADH3* gene as an internal reference gene (Brand and Hovav, 2010). The cDNA sequences of candidate genes and *ADH3* were downloaded from the Peanutbase website (https://www.peanutbase.org/), and corresponding primers were designed using Primer Premier 5.0.

Results

Phenotypic identification of F_2 and RIL populations

NIR spectrometric analysis indicated that the sucrose contents of the female parent Jihuatian 1 and male parent PI478819 were 8.96% and 3.70%, respectively. For the sucrose content of the F_2 population, we obtained maximum and minimum values of 11.03% and 2.52%, respectively, with a coefficient of variation of 0.28%, (Supplementary Table S1). The sucrose content per plant in the F_2 population showed continuous variation and an approximate normal distribution, which is typical of a quantitative character (Figure 1A). The sucrose content of the RIL population was measured in three environments, with mean values of 5.89%, 5.64%, and 5.91% and coefficient of variation ranging from 0.29% to 032% being obtained (Table 1). In each of the three growth environments, we detected a continuous frequency distribution of sucrose content in the RIL population, indicating that the population may contain multiple major genes or QTLs associated with the control of sucrose content (Figures 1B–D). ANOVA revealed that sucrose content is influenced by genotype, the environment, and genotype–environment interactions (Table 2).

Identification of candidate SNPs associated with sucrose content using BSA

Using the measured phenotype data, individuals with extreme phenotypes were used to form two mixed pools (Supplementary Table S1). The original data obtained from whole-genome re-sequencing of the two mixed pools and two

TABLE 3 Sequencing data evaluation and comparison with reference genome statistics.

Sample_ID	Clean_reads	GC_rate (%)	Q20 (%)	Q30 (%)	Mapped (%)	Coverage_rate (%)	Mean_depth
JHT 1	439154736	38.26	92.55	84.09	95.54	98.53	25.76×
PI478819	604226914	37.47	95.69	89.88	96.81	98.87	35.45×
HSP	600610150	36.91	97.38	93.52	98.13	98.73	35.23×
LSP	600223092	36.66	97.11	92.93	98.07	98.74	35.21×

Note: Sample_ID, sample number; Clean_bases, number of bases filtered; Clean_reads, number of Clean reads filtered; GC (%), sample GC, content, that is, percentage of G and C type bases in total bases; Q20 (%), percentage of bases with mass value greater than or equal to 20 in total bases; Q30 (%), percentage of bases with mass value greater than or equal to 30 in total bases. Mapped (%), percentage of Clean Reads to reference genome in total Clean Reads; Coverage_ratio (%), percentage of overlay sites in genome; Mean—depth, average sequencing depth.

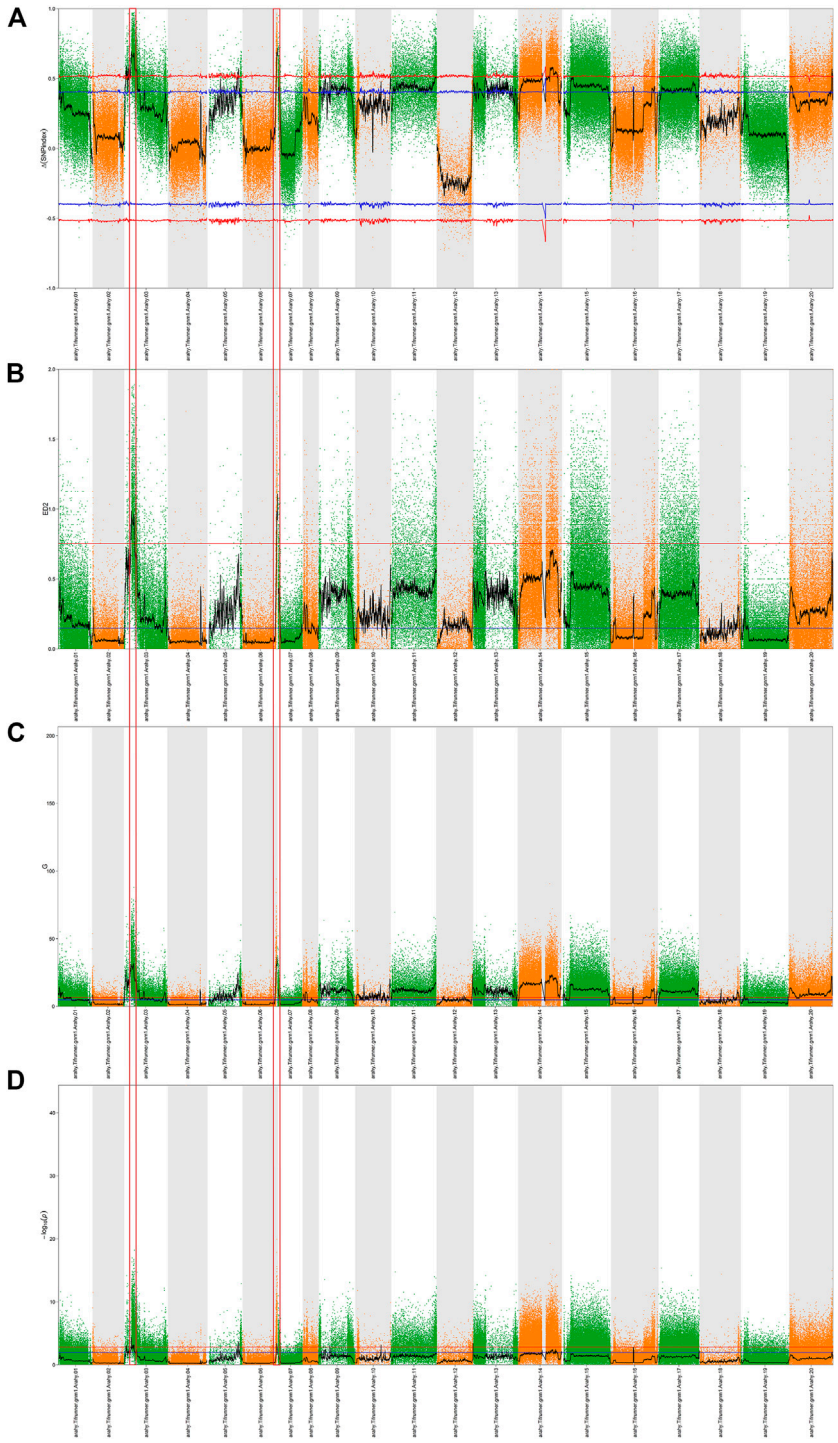


FIGURE 2
QTL analysis of sucrose based on BSA-seq. A Manhattan plot showing the distribution of Δ (SNP-index) **(A)**, the square of the Euclidean distance **(B)**, the distribution of *G*-values **(C)**, and the distribution of -log 10 (*p*-values) based on Fisher's exact tests **(D)** on chromosomes. The blue and red lines represent 95% and 99% confidence intervals, respectively. The black lines are average values of the four algorithms and were drawn based on sliding window analysis. The numbers on the horizontal coordinates represent the chromosome numbers. The part circled by the red box is the target region.

TABLE 4 Initial positioning QTL interval information.

Chromsome	Start(bp)	End(bp)	Size(Mb)
Arahy.03	5540001	8340000	2.80
Arahy.03	20800001	34180000	13.38
Arahy.06	109810001	114840000	5.03

parents were filtered to obtain a total of 336.63 Gbp clean reads, with a Q30 value ≥84.09%, GC content ranging from 36.66% to 38.26%, and the distribution of insert sizes showing a unimodal normal distribution. The average comparison efficiency between samples and the reference genome was 97.14%, the average sequencing depth was 32.92×, and we obtained 98.72% genome coverage (Table 3). The values of

TABLE 5 QTL fine mapping of sucrose content in peanut kernels.

Environment	Chromsome	Position	LeftMarker	RightMarker	LOD	PVE(%)	Add
F_9-XX	Arahy.06	15.70	A06.112437412	A06.112662675	44.84	41.05	−1.0196
F_9-KF	Arahy.06	14.80	A06.112367085	A06.112437412	28.70	31.95	−0.9461
F_9-ZMD	Arahy.06	15.20	A06.112437412	A06.112662675	39.61	37.64	−0.9842
F_9-XX	Arahy.03	54.00	A03.31223142	A03.33012101	3.23	6.03	−0.3867
F_9-KF	Arahy.03	13.00	A03.5557721	A03.6982931	2.58	4.68	−0.3863
F_9-ZMD	Arahy.03	54.00	A03.31223142	A03.33012101	2.90	5.75	−0.3721

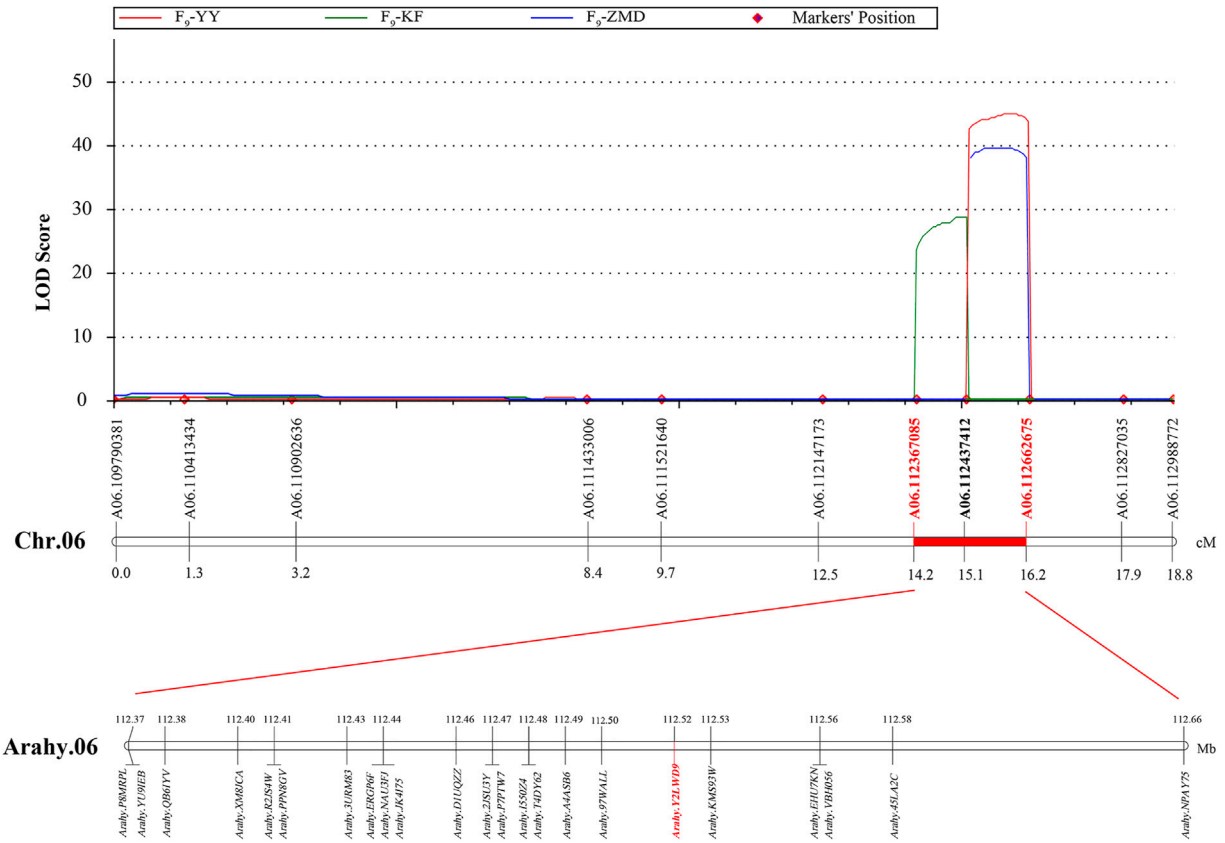


FIGURE 3
QTL fine mapping of sucrose content on peanut chromosome A06.

TABLE 6 The information of primers for quantitation real-time PCR.

Gene ID	Forward primer	Reverse primer	Product Length(bp)
Arahy.Y2LWD9	ATGAACCTCAACCAATGCCTCT	CAGGAACAGCAAACCCAGAA	204
ADH3	GACGCTTGGCGAGATCAACA	AACCGGACAACCACCACATG	140

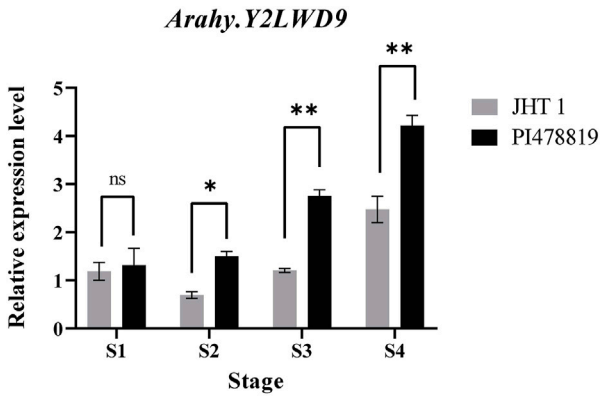


FIGURE 4
Expression analysis of a candidate gene in two parents at different growth stages (S1 to S4). Asterisks indicate a significant difference between the Jihuatian 1 (JHT 1) and PI478819 parents at each of the four growth stages, as determined using Student's *t*-test ($^{ns}P > 0.05$; $^{*}p < 0.05$; $^{**}p < 0.01$).

these parameters indicated the sufficiently good quality of sequencing and a high percentage matches with the peanut reference genome, thereby indicating that the obtained sequences could be used for subsequent variant detection and analysis.

Prior to bulked segregant analysis, we filtered out low-quality SNPs, and thereby finally obtained 318,057 high-quality credible SNPs. These high-quality SNPs were subjected to Δ (SNP-index) (Figure 2A), ED (Figure 2B), *G*-value (Figure 2C), and Fisher's exact test (Figure 2D) association analyses, and plotted according to the chromosomal distribution of each parameter. Using a 99% confidence level as the screening threshold for associated chromosomal intervals, all four methods identified multiple candidate intervals on multiple chromosomes (Supplementary Table S2). The three overlapping intervals obtained using these four methods were identified as candidate intervals associated with peanut sucrose content (Table 4).

Verification and narrowing of the positioning range

According to the different SNP information of Jihuatian 1 and PI478819 in three overlapping QTL intervals, the KASP primers were designed (Supplementary Table S3). The markers were genotyped in 251 RIL lines grown in three environments and subjected to genetic linkage analysis. The results revealed that the QTL detected in the initial mapping interval on chromosome A03 was not identified in the RIL population, indicating that this locus might be a false positive locus (Table 5). For all three assessed environments, we detected a candidate interval on chromosome A06 with phenotypic variance explained (PVE) and LOD values of 31.95%–41.05% and 28.70–44.84, respectively, which was considered to be a major QTL, which we designated *qSUCA06* (Figure 3). The genetic distance of the *qSUCA06* interval was 2.01 cM and the physical distance was 0.29 Mb (112367085–112662675 bp) (Table 5).

Candidate gene annotation and expression analysis

The total of 23 genes were identified in the *qSUCA06* interval (Figure 3, Supplementary Table S4). And then we detected eight genes changed in exon regions by further amplification and identification in this interval. Among which, six and three genes characterized by SNP and InDel differences, respectively, between Jihuatian 1 and PI478819 (Supplementary Table S5). We have made in-depth functional annotation on several databases and identified two genes related to protein synthesis and metabolism, three genes related to signal transduction, two genes related to cell cycle, and one gene encoding transcription factor through the analysis of the biological process of gene expression products.

The candidate gene designated *Arahy.Y2LWD9*, which encodes acyl-CoA-binding domain 3 (ACBD), is a domain of acyl-CoA-binding proteins, a class of lipid transporter family proteins, which may be associated with sucrose. Given the detected correlation between *Arahy.Y2LWD9* and sucrose accumulation, we analyzed the levels of *Arahy.Y2LWD9* expression in the two parents. On the basis of the cDNA sequences of *Arahy.Y2LWD9*, we designed primers (Table 6) and performed qRT-PCR analyses of candidate genes, using *ADH3* as the internal reference control (Brand and Hovav, 2010). The results showed that whereas there were no significant difference between two parents at the S1 stage of seed development with respect to the relative expression of *Arahy.Y2LWD9*, we detected significant differences in expression at stages S2, S3, and S4 (Figure 4). Overall, the expression of *Arahy.Y2LWD9* in the two parents showed an upward trend, which was opposite to the observed accumulation of sucrose,

thereby tending to indicate this gene may play a negative regulatory role in the accumulation of sucrose in peanut (Li et al., 2021).

Discussion

Given their high nutritional value, peanuts are probably the most widely consumed type of nut. For consumers, it is desirable that peanut kernels are of high quality with a good taste, an important contributory factor of which is sucrose content, which imparts a sweet taste. In this study, using BSA-seq technology, we investigated the potential genetic mechanisms underlying the control of sucrose content of peanut. For the purposes of QTL mapping analysis, we used an F_2 segregating population of 831 plants and an RIL population comprising 251 lines. Our ANOVA results revealed that the growth environment has a significant influence on the sucrose content of peanut kernels. To minimize the effect of environment on the mapping results, we cultivated the RIL population in three different locations, which also contributed to a more accurate and reliable identification of QTLs.

To date, there have been a few studies that have examined the QTLs or genes associated with sucrose content in peanut kernels. In one of these studies, the transcriptomes of two peanut cultivars with different sucrose contents were comparatively analyzed based on weighted gene correlation network analysis and qRT-PCR across multiple developmental stages, and six genes with high expression levels were finally identified in the derived RILs (Li et al., 2021). However, whereas none of the six genes reported were detected within the confidence intervals of the QTL, the QTL *qSUCA06* identified in present study was found to have a negative additive effect with a PVE of 31.95%–41.05%, which accordingly tended to indicate the high probability that novel genes regulating sucrose content are located in this region.

The candidate gene *Arahy.Y2LWD9* identified within the *qSUCA06* QTL, which encodes acyl-CoA-binding domain 3. might be one of such gene modulating sucrose content. Previous studies have shown that ACBD is a domain of acyl-CoA-binding proteins, which play key roles in plant fat metabolism (Ye and Chye, 2016). In eukaryotic cells, these proteins are involved in the transport of acyl-CoA esters and the formation and maintenance of the cytosolic acyl-CoA pool, thereby contributing to the regulation of lipid metabolism (Guo et al., 2017). On the basis of principal component analysis, Yu et al. (2020) found that the sucrose content in peanut kernels was negatively correlated with fat content. In the present study, we detected significant differences between the two parents with respect to the expression of *Arahy.Y2LWD9* during different stages of development, and that overall, there was an upward trend in expression with growth progression, which was opposite to the accumulation of sucrose. Acyl-CoA-binding proteins contain a class of highly conserved acyl coenzyme A that has been identified from rice (Meng et al., 2011), *Arabidopsis thaliana* (Xiao and Chye, 2009), *Agave americana* (Guerrero et al., 2006), and *Brassica napus* (Hills et al., 1994). Studies have shown that overexpression of *OsACBP2* in rice can promote a significant increase in the contents of triglycerides and long-chain fatty acids in seeds (Guo et al., 2019). In view of the strong activity of *Arabidopsis thaliana AtACBP6 pro::GUS* in the cotyledons of developmental embryos and the accumulation of oleyl and linoleyl CoA esters in ACBP6 seedlings, it is speculated that *AtACBP6*, together with *AtACBP4* and *AtACBP5*, may play a role in seed oil synthesis (Hsiao et al., 2014). Consequently, it is plausible that *Arahy.Y2LWD9* indirectly regulates sucrose content by regulating lipid metabolism in peanut kernels; however, this specific function needs to be further verified based on either overexpression or loss-of-function analyses.

Conclusion

4In this study, we identified a major stable QTL, *qSUCA06*, for peanut sucrose content based on BSA-seq analysis and fine mapping. Within the QTL interval, we detected a candidate gene, *Arahy.Y2LWD9*, which was verified by qRT-PCR to be negatively corrected with peanut sucrose content. The findings of this study provide a theoretical basis for further analysis of the genetic regulation of sucrose content in peanut, and will contribute to breeding for both oil and sucrose contents, taking into consideration the requirements of industry and consumers.

Data availability statement

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: https://bigd.big.ac.cn/gsa/browse/CRA009024.

Author contributions

JG, XZ and SH conceived the study. JG, FQ, LQ, CL, XL, HL, DL, MT, HL, JX, LM, BH and WD collected plant materials and performed the experiments. MZ, ZS, MC, MZ and QZ participated in handling figures and tables. JG drafted the manuscript. JG, SH and XZ revised the manuscript. All authors read and approved the final manuscript.

Funding

This work was supported by National key R&D plan (2022YFD1200402), the China Agriculture Research System (CARS-13), the Major Technology Research and Development of Henan Province, China (201300111000, 221100110300) and the Henan Provincial Agriculture Research System, China (S2012-5).

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2022.1089389/full#supplementary-material

References

Brand, Y., and Hovav, R. (2010). Identification of suitable internal control genes for quantitative real-Time PCR expression analyses in peanut (*Arachis hypogaea*). *Peanut Sci.* 37, 12–19. doi:10.3146/PS09-014.1

Chen, H., Chen, X., Xu, R., Liu, W., Liu, N., Huang, L., et al. (2021). Fine-mapping and gene candidate analysis for *AhRt1*, a major dominant locus responsible for testa color in cultivated peanut. *Theor. Appl. Genet.* 134, 3721–3730. doi:10.1007/s00122-021-03924-w

Cingolani, P., Platts, A., Wang, L. L., Coon, M., Nguyen, T., Wang, L., et al. (2012). A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of *Drosophila melanogaster* strain w1118; iso-2; iso-3. *Fly* 6, 80–92. doi:10.4161/fly.19695

Clevenger, J., Chu, Y., Chavarro, C., Botton, S., Culbreath, A., Isleib, T. G., et al. (2018). Mapping late leaf spot resistance in peanut (*Arachis hypogaea*) using QTL-seq reveals markers for marker-Assisted selection. *Front. Plant Sci.* 9, 83. doi:10.3389/fpls.2018.00083

Davis, J. P., and Dean, L. L. (2016). Peanut composition, flavor and nutrition. *Peanuts*, 289–345. doi:10.1016/B978-1-63067-038-2.00011-3

Fekih, R., Takagi, H., Tamiru, M., Abe, A., Natsume, S., Yaegashi, H., et al. (2013). MutMap+: Genetic mapping and mutant identification without crossing in rice. *PLoS One* 8, e68529. doi:10.1371/journal.pone.0068529

Giovannoni, J. J., Wing, R. A., Ganal, M. W., and Tanksley, S. D. (1991). Isolation of molecular markers from specific chromosomal intervals using DNA pools from existing mapping populations. *Nucleic Acids Res.* 19, 6553–6558. doi:10.1093/nar/19.23.6553

Guerrero, C., Martín-Rufián, M., Reina, J. J., and Heredia, A. (2006). Isolation and characterization of a cDNA encoding a membrane bound acyl-CoA binding protein from *Agave americana* L. epidermis. *Plant Physiol. biochem.* 44, 85–90. doi:10.1016/j.plaphy.2006.01.002

Guo, Z. H., Chan, W. H. Y., Kong, G. K. W., Hao, Q., and Chye, M. L. (2017). The first plant acyl-CoA-binding protein structures: The close homologues *OsACBP1* and *OsACBP2* from rice. *Acta Crystallogr. D. Struct. Biol.* 73, 438–448. doi:10.1107/S2059798317004193

Guo, Z. H., Haslam, R. P., Michaelson, L. V., Yeung, E. C., Lung, S. C., Napier, J. A., et al. (2019). The overexpression of rice ACYL-CoA-BINDING PROTEIN2 increases grain size and bran oil content in transgenic rice. *Plant J.* 100, 1132–1147. doi:10.1111/tpj.14503

Han, S., Zhou, X., Shi, L., Zhang, H., Geng, Y., Fang, Y., et al. (2022). *AhNPR3* regulates the expression of WRKY and PR genes, and mediates the immune response of the peanut (*Arachis hypogaea* L.). *Plant J.* 110, 735–747. doi:10.1111/tpj.15700

Hills, M. J., Dann, R., Lydiate, D., and Sharpe, A. (1994). Molecular cloning of a cDNA from *Brassica napus* L. for a homologue of acyl-CoA-binding protein. *Plant Mol. Biol.* 25, 917–920. doi:10.1007/BF00028886

Hill, J. T., Demarest, B. L., Bisgrove, B. W., Gorsi, B., Su, Y. C., and Yost, H. J. (2013). Mmappr: Mutation mapping analysis pipeline for pooled RNA-seq. *Genome Res.* 23, 687–697. doi:10.1101/gr.146936.112

Hsiao, A. S., Haslam, R. P., Michaelson, L. V., Liao, P., Chen, Q. F., Sooriyaarachchi, S., et al. (2014). *Arabidopsis* cytosolic acyl-CoA-binding proteins *ACBP4, ACBP5* and *ACBP6* have overlapping but distinct roles in seed development. *Biosci. Rep.* 34, e00165. doi:10.1042/BSR20140139

Isleib, T. G., Pattee, H. E., and Giesbrecht, F. G. (2004). Oil, sugar, and starch characteristics in peanut breeding lines selected for low and high oil content and their combining ability. *J. Agric. Food Chem.* 52, 3165–3168. doi:10.1021/jf035465y

Kumar, R., Janila, P., Vishwakarma, M. K., Khan, A. W., Manohar, S. S., Gangurde, S. S., et al. (2020). Whole-genome resequencing-based QTL-seq identified candidate genes and molecular markers for fresh seed dormancy in groundnut. *Plant Biotechnol. J.* 18, 992–1003. doi:10.1111/pbi.13266

Li, H. H., Ye, G. Y., and Wang, J. K. (2007). A modified algorithm for the improvement of composite interval mapping. *Genetics* 175, 361–374. doi:10.1534/genetics.106.066811

Li, W., Huang, L., Liu, N., Pandey, M. K., Chen, Y., Cheng, L., et al. (2021). Key regulators of sucrose metabolism identified through comprehensive comparative transcriptome analysis in peanuts. *Int. J. Mol. Sci.* 22, 7266. doi:10.3390/ijms22147266

Liu, S. B., Cai, S. B., t Graybosch, R., Chen, C. X., and Bai, G. H. (2008). Quantitative trait loci for resistance to pre-harvest sprouting in US hard white winter wheat Rio Blanco. *Theor. Appl. Genet.* 117, 691–699. doi:10.1007/s00122-008-0810-7

Liu, H., Sun, Z., Zhang, X., Qin, L., Dong, W., Wang, Z., et al. (2020a). QTL mapping of web blotch resistance in peanut by high-throughput genome-wide sequencing. *BMC Plant Biol.* 20, 249. doi:10.1186/s12870-020-02455-8

Liu, H., Zhou, F., Zhou, T., Yang, Y., and Zhao, Y. (2020b). Fine mapping of a novel male-sterile mutant showing wrinkled-leaf in sesame by BSA-Seq technology. *Ind. Crops Prod.* 156, 112862. doi:10.1016/j.indcrop.2020.112862

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408. doi:10.1006/meth.2001.1262

Magwene, P. M., Willis, J. H., and Kelly, J. K. (2011). The statistics of bulk segregant analysis using next generation sequencing. *PLoS Comput. Biol.* 7, e1002255. doi:10.1371/journal.pcbi.1002255

Majeed, U., Darwish, E., Rehman, S. U., and Zhang, X. (2018). Kompetitive allele specific PCR (KASP): A singleplex genotyping platform and its application. *J. Agric. Sci. (Tor).* 11, 11. doi:10.5539/jas.v11n1p11

Mansfeld, B. N., and Grumet, R. (2018). QTLseqr: An R package for bulk segregant analysis with next-generation sequencing. *Plant Genome* 11, 180006. doi:10.3835/plantgenome2018.01.0006

McDaniel, K. A., White, B. L., Dean, L. L., Sanders, T. H., and Davis, J. P. (2012). Compositional and mechanical properties of peanuts roasted to equivalent colors using different time/temperature combinations. *J. Food Sci.* 77, C1293–C1299. doi:10.1111/j.1750-3841.2012.02979.x

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, 1297–1303. doi:10.1101/gr.107524.110

Meng, W., Su, Y. C. F., Saunders, R. M. K., and Chye, M. L. (2011). The rice acyl-CoA-binding protein gene family: Phylogeny, expression and functional analysis. *New Phytol.* 189, 1170–1184. doi:10.1111/j.1469-8137.2010.03546.x

Meng, L., Li, H. H., Zhang, L. Y., and Wang, J. K. (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

Michelmore, R. W., Paran, I., and Kesseli, R. V. (1991). Identification of markers linked to disease-resistance genes by bulked segregant analysis: A rapid method to detect markers in specific genomic regions by using segregating populations. *Proc. Natl. Acad. Sci. U. S. A.* 88, 9828–9832. doi:10.1073/pnas.88.21.9828

Pattee, H. E., Johns, E. B., Singleton, J. A., and Sanders, T. H. (1974). Composition changes of peanut fruit parts during maturation1. *Peanut Sci.* 1, 57–62. doi:10.3146/i0095-3679-1-2-6

Pattee, H. E., Young, C. T., and Giesbrecht, F. G. (1981). Seed size and storage effects on carbohydrates of peanuts. *J. Agric. Food Chem.* 29, 800–802. doi:10.1021/jf00106a028

Pattee, H. E., Isleib, T. G., and Giesbrecht, F. G. (1998). Variation in intensity of sweet and bitter sensory attributes across peanut genotypes1. *Peanut Sci.* 25, 63–69. doi:10.3146/i0095-3679-25-2-2

Pattee, H. E., Isleib, T. G., Giesbrecht, F. G., and McFeeters, R. F. (2000). Investigations into genotypic variations of peanut carbohydrates. *J. Agric. Food Chem.* 48, 750–756. doi:10.1021/jf9910739

Qin, L., Liu, H., Du, P., Dong, W. Z., Huang, B. Y., Han, S. Y., et al. (2016). Determination of sucrose content in peanut seed kernel based on infrared spectroscopy. *Chin. J. Oil Crop Sci.* 38, 666–671. doi:10.7505/j.issn.1007-9084.2016.05.018

Qin, L., Liu, H., Zhang, X, Y., Du, P., Dai, X, D., Sun, Z, Q., et al. (2020). Genetic analysis of sugar content in peanut kernel via mixed major gene plus polygene inheritance model in multi-generation combined population. *Chin. J. Oil Crop Sci.* 43, 590–599. doi:10.19802/j.issn.1007-9084.2020185

Sanders, T. H., and Bett, K. L. (1995). Effect of harvest date on maturity, maturity distribution, and flavor of florunner peanuts. *Peanut Sci.* 22, 124–129. doi:10.3146/i0095-3679-22-2-10

Sheng, C., Song, S., Zhou, R., Li, D., Gao, Y., Cui, X., et al. (2021). QTL-Seq and transcriptome analysis disclose major QTL and candidate genes controlling leaf size in sesame (*Sesamum indicum* L.). *Front. Plant Sci.* 12, 580846. doi:10.3389/fpls.2021.580846

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

Xiao, S., and Chye, M. L. (2009). An Arabidopsis family of six acyl-CoA-binding proteins has three cytosolic members. *Plant Physiol. biochem.* 47, 479–484. doi:10.1016/j.plaphy.2008.12.002

Xie, J., Wang, Q., Zhang, Z., Xiong, X., Yang, M., Qi, Z., et al. (2021). QTL-seq identified QTL and candidate genes for two-seed pod length and width in soybean (*Glycine max*). *Plant Breed.* 140, 453–463. doi:10.1111/pbr.12920

Yang, L., Wang, J., Han, Z., Lei, L., Liu, H. L., Zheng, H., et al. (2021). Combining QTL-seq and linkage mapping to fine map a candidate gene in *qCTS6* for cold tolerance at the seedling stage in rice. *BMC Plant Biol.* 21, 278. doi:10.1186/s12870-021-03076-5

Ye, Z. W., and Chye, M. L. (2016). Plant cytosolic Acyl-CoA-Binding proteins. *Lipids* 51, 1–13. doi:10.1007/s11745-015-4103-z

Yu, H., Liu, H., Erasmus, S. W., Zhao, S., Wang, Q., and van Ruth, S. M. (2020). Rapid high-throughput determination of major components and amino acids in a single peanut kernel based on portable near-infrared spectroscopy combined with chemometrics. *Ind. Crops Prod.* 158, 112956. doi:10.1016/j.indcrop.2020.112956

Zhang, M. N., Zeng, Q., Liu, H., Qi, F. Y., Sun, Z. Q., Miao, L. J., et al. (2022). Identification of a stable major QTL for fresh-seed germination on chromosome Arahy.04 in cultivated peanut (*Arachis hypogaea* L.). *Crop J.*, 2214–5141. doi:10.1016/j.cj.2022.03.012

frontiers | Frontiers in Genetics

TYPE Review
PUBLISHED 05 January 2023
DOI 10.3389/fgene.2022.1032601

OPEN ACCESS

EDITED BY
Manish Kumar Pandey,
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India

REVIEWED BY
Waseem Hussain,
International Rice Research Institute (IRRI), Philippines
Meriem Aoun,
Oklahoma State University, United States

*CORRESPONDENCE
Mohamed Mergoum,
mmergoum@uga.edu

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 31 August 2022
ACCEPTED 20 December 2022
PUBLISHED 05 January 2023

CITATION
Subedi M, Ghimire B, Bagwell JW, Buck JW and Mergoum M (2023), Wheat end-use quality: State of art, genetics, genomics-assisted improvement, future challenges, and opportunities.
Front. Genet. 13:1032601.
doi: 10.3389/fgene.2022.1032601

Wheat end-use quality: State of art, genetics, genomics-assisted improvement, future challenges, and opportunities

Madhav Subedi [1], Bikash Ghimire [2], John White Bagwell [1], James W. Buck [2] and Mohamed Mergoum[3]*

[1]Institute of Plant Breeding, Genetics and Genomics, University of Georgia, Griffin Campus, Griffin, GA, United States, [2]Department of Plant Pathology, University of Georgia, Griffin Campus, Griffin, GA, United States, [3]Department of Crop and Soil Sciences, University of Georgia, Griffin Campus, Griffin, GA, United States

Wheat is the most important source of food, feed, and nutrition for humans and livestock around the world. The expanding population has increasing demands for various wheat products with different quality attributes requiring the development of wheat cultivars that fulfills specific demands of end-users including millers and bakers in the international market. Therefore, wheat breeding programs continually strive to meet these quality standards by screening their improved breeding lines every year. However, the direct measurement of various end-use quality traits such as milling and baking qualities requires a large quantity of grain, traits-specific expensive instruments, time, and an expert workforce which limits the screening process. With the advancement of sequencing technologies, the study of the entire plant genome is possible, and genetic mapping techniques such as quantitative trait locus mapping and genome-wide association studies have enabled researchers to identify loci/genes associated with various end-use quality traits in wheat. Modern breeding techniques such as marker-assisted selection and genomic selection allow the utilization of these genomic resources for the prediction of quality attributes with high accuracy and efficiency which speeds up crop improvement and cultivar development endeavors. In addition, the candidate gene approach through functional as well as comparative genomics has facilitated the translation of the genomic information from several crop species including wild relatives to wheat. This review discusses the various end-use quality traits of wheat, their genetic control mechanisms, the use of genetics and genomics approaches for their improvement, and future challenges and opportunities for wheat breeding.

KEYWORDS

end-use quality, genomics-assisted breeding, quantitative trait locus mapping, marker-assisted selection, genomic selection, candidate gene approach, translational genomics, genome-wide association study

1 Introduction

Wheat is a major cereal crop due to its production and utilization throughout the world. Wheat has been the major source of energy (carbohydrate), protein, and dietary fiber for humankind (Ma et al., 2013). The wheat grain consists of 8%–20% protein and the carbohydrates make up 85% (w/w) of grain, most of which is starch (Anjum et al., 2007; Shewry and Hey, 2015; Nigro et al., 2019). Thus, the nutrition and food security of the world relies on the quantity and quality of wheat being produced worldwide. However, the demand for

wheat consumption still surpasses its productivity and there is an urgent need to increase food production by nearly 60% to feed the ever-increasing human population that is predicted to reach 9 billion by 2050 (Zhang et al., 2014; Varshney et al., 2015). Nearly 821 million people, which accounts for one in nine people in the world, suffered from hunger in 2018 (WHO, 2018) further emphasizing the importance of wheat in alleviating hunger. Increased food production is also challenged by decreasing cropland area, finite resources, and the impacts of climate change on overall crop performance.

Wheat breeding programs throughout the world face a common challenge to maintain or improve agronomic performance while simultaneously improving quality traits to fulfill the needs of the diverse international market and end-users including growers, millers, bakers, and consumers (Gale, 2005; Huang et al., 2006; Echeverry-Solarte et al., 2015). To increase the grain yield, researchers have focused on various yield components (grain weight, grain length, spike length, kernel number/spike, spike number/unit area etc.), its correlated trait (plant height, chlorophyll content etc.) and resistance to biotic and abiotic stresses including heat tolerance, and resistance to rusts and *Fusarium* head blight (Zhang et al., 2009; Hanif, 2016; Sun et al., 2017; Kuzay, 2019; Pinto et al., 2019; Sapkota et al., 2019; Yang et al., 2019; Kumar et al., 2007; Ghimire et al., 2020; Liu et al., 2020; Pradhan et al., 2020; Zhang et al., 2020; Ghimire et al., 2022). Consequently, significant progress has been achieved in the past 50 years as wheat yield worldwide increased from 1.4 MT/ha in 1970 to 3.5 MT/ha in 2019 (FAO, 2021).

Overall improvement in the end-use quality of wheat is an inherently complex breeding objective since it is determined by the combination of many component quality traits that are underpinned by diverse metabolic pathways (Mann et al., 2009). Many of these are also correlated with each other which further adds the complexity (Prasad et al., 2003; Mann et al., 2009; Li J et al., 2012). Besides this, the selection of optimum quality trait(s) targeting one user group often comes as a tradeoff for others (Echeverry-Solarte et al., 2015). For example, in North America and Western Europe, bread is made from wheat varieties that produce strongly elastic dough with some extensibility. However, these varieties may not be suitable for making other wheat products such as cookies (biscuits) which are made from highly extensible dough (Payne, 1983). Similarly, chapatti and noodles consumed in South Asia are made from wheat varieties with intermediate properties between the two extremes and pizza and bagels are made from wheat with high gluten strength (Lindgren and Simsek, 2016). Therefore, wheat breeders have to make targets of developing wheat varieties with quality parameters that meet particular demands from millers and consumers (Prasad et al., 2003; Mann et al., 2009).

Through the 19th century, wheat quality essentially meant bread quality to people (Kiszonas and Morris, 2017). In the past few decades, people are more conscious of their dietary habits and have developed a preference for different wheat products with specific quality attributes (Peña, 2007). The development of such a wide range of products was possible due to the studies on the functional and molecular genetic basis of wheat quality particularly wheat protein and its subunits during the 20th century as summarized in several reviews (Payne et al., 1982; Payne, 1987; Weegels et al., 1996; Shewry et al., 1997). The wide range of wheat products creates a broad spectrum of performance specifications that are determined by different end-use quality traits, their genetic and environmental factors and their complex interactions. Improvement in end-use quality of wheat depends will depend on our understanding of all these components, therefore, further studies on the end-use quality of wheat are critical.

Since the turn of the 21st century, research has been carried out in diverse fields including QTL mapping, association mapping and marker-assisted selection (Patil et al., 2009; Reif et al., 2011). There is also diversity in the traits being studied such as starch, grain hardness, flour color, milling and baking quality that has helped to discover many attributes that define the quality of wheat (Zanetti et al., 2001; Sourdille et al., 2003; Kuchel et al., 2006; McCartney et al., 2006). Based on the results of these studies, a few reviews were published that summarized the QTLs/genes and their diagnostic markers to be used for marker-assisted breeding of wheat end-use quality (Gale, 2005; Liu et al., 2012). More recently, studies on genomic selection and translational genomics for end-use quality traits are being carried out which still warrants exhaustive efforts across wheat breeding programs to achieve satisfactory results as is the case for grain yield or disease resistance (Kristensen et al., 2019; Nigro et al., 2019; Aoun et al., 2021). There are limited reviews that capture these modern genetic and genomic studies on diverse end-use quality traits. This review summarizes the key genetic and genomic findings made using different genomic tools on various important end-use quality traits and also points out the current challenges and future opportunities for such studies in wheat.

2 Wheat classes and their end-use products

Wheat grown throughout the US are classified into five different classes based on the grain color (red and white wheat), texture (hard and soft wheat), and growth habit (winter and spring wheat) (Clark and Bayles, 1935; https://www.uswheat.org/) thus, categorizing them as hard red winter wheat (HRWW), hard red spring wheat (HRSW), soft red winter wheat (SRWW), soft white wheat (SWW), and hard white wheat (HWW). These classes of wheat also differ based on the products that can be made from them (Clark and Bayles, 1935). For instance, HRWW is suitable for making flat bread, hard rolls, hearth bread, croissants, and all-purpose flour because of its excellent milling and baking characteristics (Chang and Chambers Iv, 1992; https://www.uswheat.org/). The HRSW with high protein content is also referred to as "aristocrat of wheat" for designing wheat products like rolls, bagels, croissants, hearth bread, and pizza crust. SRWW with low protein content is used for making cookies, cakes, crackers, pretzels, and pastries (Faridi et al., 1994). Similarly, SWW also has low protein content and is used for making the best quality pastry, cakes, and other confectionary products. HWW is used for making Asian noodles, pan, and flatbreads (Chang and Chambers Iv, 1992). Durum wheat is the hardest of all kinds of wheat with high protein content and is used for making pasta and couscous (Dick and Youngs, 1988).

3 Major end-use quality traits

There are multiple complex traits to consider for defining the quality of wheat importance to wheat producers, end-users, and breeders (McCartney et al., 2006). These include traits related to grain characteristics (protein content, color, weight, grain hardness/texture), milling properties (flour yield, protein content, moisture

content, ash content), flour and dough properties (starch content, falling number, gluten characteristics, dough rheology) and baking qualities (loaf height, volume and texture, elongation, mixing time, cookie diameter, baking score) (Mohler et al., 2014; Guzmán and Alvarez, 2016; Hayes et al., 2017; Naraghi et al., 2019). Above all, grain protein content (GPC) has been a focus for both plant breeders and end-users since it directly affects nutritional value, and dough rheological and baking properties (Joppa et al., 1997; Alamri et al., 2009; Brevis et al., 2010). GPC is a critical marketing characteristic and it influences the quality performance of wheat end products in general, including pasta and bread (Mesfin et al., 1999; Distelfeld et al., 2006; Brevis et al., 2010; Naraghi et al., 2019) and it is used as one of the parameters for wheat classification (Tu and Li, 2020). The price of wheat is also determined based on the GPC where wheat with higher GPC is valued more than lower GPC (Farquharson, 2006). Wheat GPC should usually be above 12.5% to be used in bread making (Turner et al., 2004).

Wheat endosperm harbors the majority of the grain content (Osborne, 1907). Gluten protein is a rubbery mass left behind when the starch granules and water-soluble constituents of wheat dough are removed (Wieser, 2007). The quality and quantity of gluten protein is an important trait to be considered for wheat breeding since it determines the baking quality of wheat dough by conferring its viscosity, cohesivity, elasticity, and water absorption capacity (Wieser, 2007). As the protein quality increases, the dough strength, firmness, and stability and cooked weight also increase (Payne, 1987; Dick and Youngs, 1988). Wheat gluten has high glutamine and proline amino acid content (Wieser, 2007). Based on the molecular size in dissociating solvents, wheat endosperm gluten can be categorized into two storage proteins, gliadin and glutenin (Huebner, 1970). The gliadins are small with no disulfide-bonded subunit structure and are soluble in aqueous alcohol, whereas, the glutenins are large, heterogeneous molecules connected by disulfide bonds that are insoluble in aqueous alcohol (Osborne, 1907; Wall, 1979; D'Ovidio and Masci 2004; Kumar et al., 2013). Based on gel electrophoresis, the glutenin subunits can be further divided into predominant low-molecular-weight glutenin subunits (LMW-GS) and high-molecular-weight glutenin subunits (HMW-GS), whereas, gliadin can be separated into four groups (α, β, γ, and ω) (Wall, 1979; D'Ovidio and Masci, 2004). Hydrated glutenin is cohesive and elastic which provides strength and elasticity to dough, whereas, hydrated gliadin is less elastic and cohesive than glutenin and provides dough viscosity and extensibility (Payne et al., 1984; Wieser, 2007; Kumar et al., 2013). Wieser (2007) defines gliadins as a "plasticizer or solvent for glutenins." Gluten strength can be measured by the Sodium dodecyl sulfate (SDS)-microsedimentation test or sedimentation volume (Kumar et al., 2013).

Grain hardness or texture is a fundamental basis for differentiating wheat market class and trade worldwide (Guzmán and Alvarez 2016). The durum wheat grains are classified as very hard and used for making pasta and couscous whereas the common wheat grains are classified as hard and soft and used for making bread, cakes, noodles, and cookies (Giroux and Morris 1997; Bushuk, 1998). Grain hardness also has a profound effect on the milling, baking, and end-use qualities of wheat (Giroux and Morris 1997). The most common methods for grain texture measurement are Particle Size Index (PSI), Near-Infrared Reflectance (NIR), and the Single Kernel Characterization System (SKCS) (Morris, 2002). The hardness or texture of wheat grains is molecularly determined by two puroindoline proteins, Pina and Pinb, where the grain texture is considered soft when both proteins are functional. However, when one is absent or mutated, the texture is hard, and when both proteins are absent as in the case of durum wheat, the texture is very hard (Morris, 2002).

Starch content and its pasting property also significantly influence wheat end-use products. Starch comprises about 70% of the endosperm dry weight, and it affects grain weight and quality as well as the capacity of plant sink tissues to accept and convert photoassimilates (Dale and Housley, 1986; Kumar et al., 2018; Zi et al., 2018). The starch pasting property has been found to affect the texture and quality of end-use products (Blazek and Copeland 2008). The reserved starch in the plant is comprised mainly of two macromolecules, amylose (22%–35%) and amylopectin (68%–75%) (Nakamura et al., 1995). Amylose in wheat grain is synthesized by granule-bound starch synthase (GBSSI), or waxy protein (Guzmán et al., 2012a). Wheat with reduced amylose is referred to as "partial waxy," and wheat with no amylose is referred to as "waxy" (Graybosch, 1998). Absence of amylose, i.e., 100% amylopectin, in waxy wheat grains is supposed to help reduce the staling of flour products, especially bread, and keep baked goods fresh for longer periods (Zi et al., 2018). However, waxy wheat lines have reduced grain yield and increased amylose content has also been associated with nutritional and physiological effects (Nakamura et al., 1995; Blazek and Copeland, 2008; Zi et al., 2018). Partial waxy wheat is considered useful for producing high quality thick white noodles such as udon noodle used in Japanese cuisine (Zhang et al., 2022). Non-waxy and waxy starches can be differentiated by staining with iodine, where non-waxy starch stains blue-black and waxy starch stains red-brown (Nakamura et al., 1995). The relative amounts of amylose and amylopectin determine the physical and chemical properties of starch, such as pasting, gelation, and gelatinization, which determine the quality of end-product (Fredriksson et al., 1998; Guzmán and Alvarez 2016). There are various methods for measuring starch, amylose, and amylopectin, including the dual wavelength iodine binding method, differential scanning calorimetry (DSC), and high-performance size-exclusion chromatography (HPSEC) (Zhu et al., 2008).

The end-use quality of wheat is also determined by dough rheological properties and falling numbers. The rheological properties of wheat are estimated by water absorption, dough development time, dough stability, maximum dough resistance, dough extensibility, and flour paste viscosity (Guo J. et al., 2020). Dough rheology can be measured by using farinograph, extensograph, and alveograph instruments (AbuHammad et al., 2012). These dough rheological properties along with other traits such as flour protein content (FPC), particle size, loaf volume, and crumb score can be used for estimating the baking quality of wheat (Kuchel et al., 2006). During germination, starch in the wheat grain needs to be converted to simple sugars to feed the embryo. Alpha-amylase is one of the primary enzymes responsible for the starch degradation causing sprouting of the grain (Newberry et al., 2018). Such alpha-amylase activity is only desirable if the grain has been planted (Thomason et al., 2019). However, some wheat genotypes are characterized by an excessive level of alpha-amylase from the grain development stage to harvest causing pre-harvest sprouting (PHS). PHS results in lower yield and affects the end-use quality of wheat such as dough softening, sticky bread crumb, and problems while slicing bread (Mohler et al., 2014; Newberry et al., 2018). The falling number determines the effect of alpha-amylase activity on damaging starch by examining the starch pasting property (Mohler et al., 2014). Increased level of the alpha-amylase reduces the value of falling number (Thomason et al., 2019).

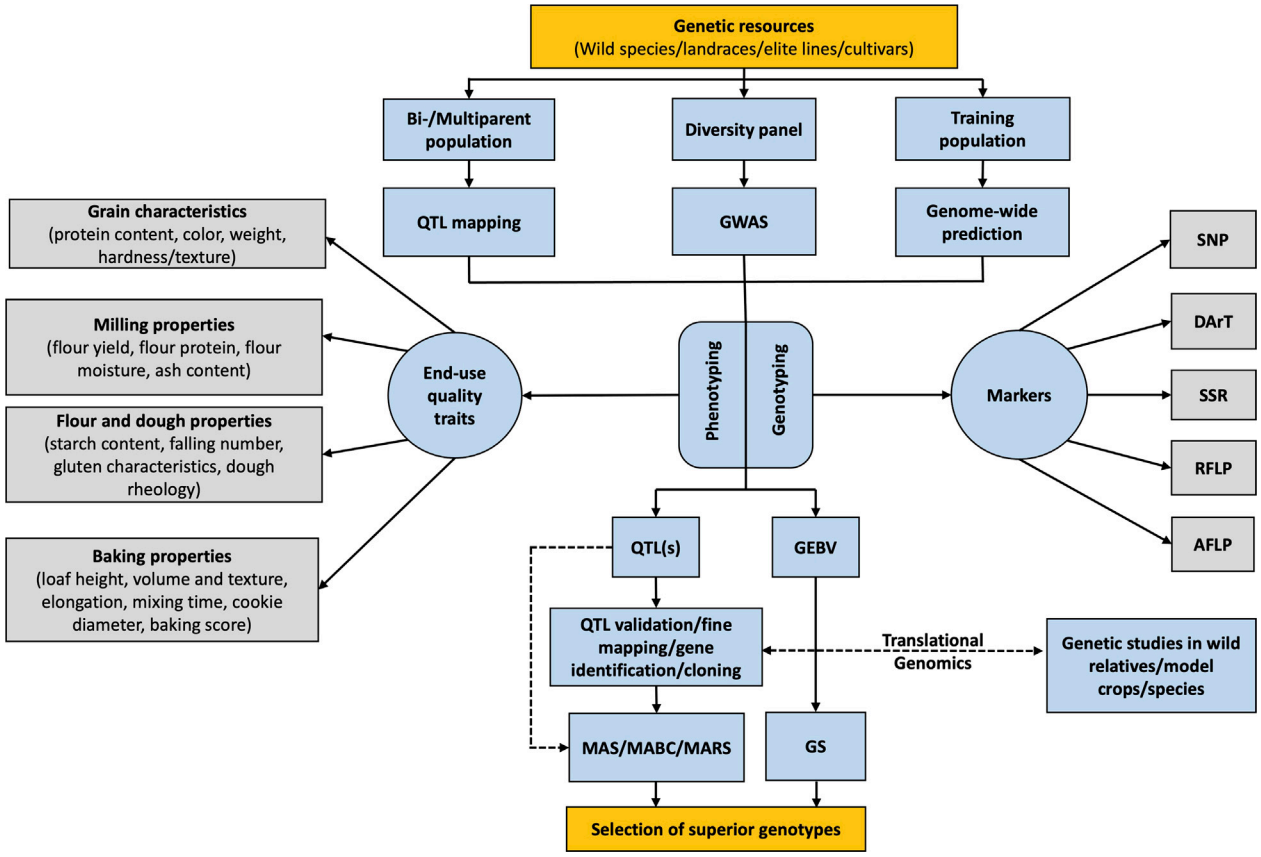


FIGURE 1
Schematic flow diagram showing the use of genomics-assisted breeding and translational genomics for improvement of end-use quality traits in wheat breeding program.

4 Breeding methods for end-use quality traits of wheat

4.1 Direct phenotypic selection

To select individual plants with superior quality traits, wheat breeders used to grow multiple individuals/lines in different environments, harvest grain, process it, and evaluate it for different milling and baking quality parameters. All these wheat qualities had to be evaluated precisely using appropriate instruments under lab conditions which makes this procedure challenging because it is expensive, time-consuming, labor-intensive, and typically requires large seed samples (Blanco et al., 1996; Prasad et al., 2003; Naraghi et al., 2019; Yang et al., 2020). Analyses requires time, which often leads to increase in 1 year for the breeding program or passing wheat lines with undesirable quality alleles into the next growing season (Sandhu et al., 2021). Moreover, if these evaluations are carried out later in the breeding program, the developed wheat lines could end up with poor end-use quality and be discarded leading to waste of resources as the primary focus was centered on improving other traits such as grain yield and disease resistance (Naraghi et al., 2019). Therefore the direct selection method for end-use qualities of wheat has been more complicated for wheat breeders (Yang et al., 2020). Traits like GPC and FPC are also influenced by the genotype by environment (G×E) interaction leading to low heritability (Blanco et al., 1996; Groos et al., 2003; Prasad et al., 2003). As a result, accurate assessment of these traits in the breeding programs is quite challenging which makes the selection process even more complicated.

4.2 Modern approaches: Genomics-assisted breeding (GAB) and translational genomics

Genome-based technologies are an important means of breeding for improved wheat quality (Kiszonas and Morris, 2017). The advancement in sequencing technologies and their utilization for genomic research brought increased precision and efficiency to crop breeding (Varshney et al., 2005; Varshney et al., 2015). Various kinds of molecular markers such as restriction fragment length polymorphism (RFLP), simple sequence repeats (SSR), and single nucleotide polymorphism (SNP) are being used for genetic mapping to identify quantitative trait loci (QTL) and functional markers related to the genes of interest (Figure 1; Table 1) (Groos et al., 2003; Prasad et al., 2003; Rimbert et al., 2018). These markers can be exploited for marker-assisted selection (MAS) and genomic selection (GS) which allow screening for superior end-use quality traits earlier in the breeding program (Figure 1; Tables 2, 3) (Naraghi et al., 2019). These approaches are convenient and faster than the traditional selection method. This allowed the transition from solely phenotype-based traditional selection to a genotype-linked

TABLE 1 Summary of representative QTL identification studies for end-use quality traits in wheat. Phenotypic variance is indicated collectively for all the QTLs or individually for each QTL.

Model species	Population type[a]	Mapping population	End-use quality traits[b]	Marker type[c]	Number of QTLs identified	Genome location	Phenotypic variance explained	References
Durum wheat	RIL	Messapia (cultivar of *T. turgidum* L. subsp. Durum) × MG4343 (wild accession of *T. turgidum* L var *dicoccoides*)	GPC	RFLP	6	4BS, 5AL, 6AS, 6BS, and 7BS	6.0%–23.5% individually	Blanco et al. (1996)
Durum wheat	RICL	LDN (DIC-6B) × LDN durum	GPC	RLFP	1	6BS	66%	Joppa et al. (1997)
Durum wheat	RIL	Messapia (cultivar of *T. turgidum* L. subsp. Durum) × MG4343 (wild accession of *T. turgidum* L var *dicoccoides*)	SV	RFLP	7	1AL, 1B, 3AS, 3BL, 5AL, 6AL, 7BS	36%–64% collectively	Blanco et al. (1998)
Common wheat	RIL	Glupro × Keene, Glupro × Bergen, ND683 × Bergen	GPC	RFLP	1	6BS, 6BL	12.4%–34.6% across three populations	Mesfin et al. (1999)
Common wheat and spelt	RIL	*T. aestivum* L variety "Forno" × *T. spelta* L variety 'Oberkulmer	FPC, SV, dough properties, GH, TKW	RFLP	FPC: 9; SV: 9; dough properties: 10; GH: 10; TKW:8	FPC: 3B, 4A, 5B, 6B, 7A, 7B, 7D; SV: 1B, 1D, 2A, 3A, 5A, 5B, 5D; dough properties: 1A, 1B, 2A, 2B, 2D, 3A, 3B, 3DL, 4A, 4D, 5A, 5D,7A, 7B, 7D; GH: 2A, 3B, 4A, 4D, 5A, 6B, 7A,7B, 7D; TKW: 1B, 2B, 3B, 5A, 5B, 7B	Protein: 51%; SV: 47%; GH: 54%; TKW: 54%; and dough properties: 39% collectively	Zanetti et al. (2001)
Common wheat	RIL	Renan (variety) × Recital (variety)	GPC, TKW	SSR, RFLP, AFLP	GPC: 10; TKW: 9	GPC:1A, 2A, 3A, 3B,4A, 4D, 5B, 6A, 7A, 7D; TKW: 1D, 2B, 2D, 3A, 5B, 6A, 6D, 7A, 7D	GPC: 4.2%–10.4%; TKW: 4.7%–19.7% individually	Groos et al. (2003)
Common wheat	DH	Courtot (cultivar) × Chinese spring (cultivar)	Protein content GH, Dough strength	SSR, AFLP	Protein content: 2; GH: 3; dough strength: 3	Protein content: 1BL, 6AS; GH: 1AL, 5DS, 6AL; dough strength: 1AL, 3BL, 5DS	Protein content: 6.5%–17.1%; GH: 3.1%–66.9%; dough strength: 9.4%–19.5% individually	Sourdille et al. (2003)
Common wheat	RIL, NIL	RIL: WL711 × PH132 NIL: WL711 × PH132, WL711 × PH133, HD2329 × PH132	GPC	SSR	13	2A, 2B, 2D, 3D, 4A, 6B, 7A, 7D	2.95%–32.44% individually	Prasad et al. (2003)
Common wheat	DH	Trident (cultivar) × Molineux (cultivar)	FPC, Flour color, Flour brightness, dough properties, baking quality	SSR, STS	FPC: 5; flour brightness: 2; flour color: 3; dough properties: 18; baking quality: 4	FPC: 1B, 6A, 6D, 7A, 7D; flour brightness: 1A, 7B; flour color: 1A, 7B; dough properties: 1A, 1B, 2A, 2B, 2D, 3D, 7A, 7D; baking quality: 2A, 3A	FPC: 6%–13%; flour brightness: 10%–22%; flour color: 6%–77%, dough properties:5%–20%, baking quality: 5%–13% individually	Kuchel et al. (2006)
Common wheat	DH	AC karma (variety) × 87E03-S2B1 (breeding line)	GPC, FPC, MG, SV, TKW	SSR	GPC: 2; FPC: 3; MG: 18; SV: 3; TKW: 6	GPC: 4D, 7B; FPC: 2D,4D,7B; MG: 1B,1D,3B,4D,5D; SV: 1B,2D,5D; TKW: 2B, 2D, 3B, 4B, 4D, 6A	GPC: 12.6%–32.7%; FPC: 6.6%–28.6%; MG: 6%–55.9%; SV: 8.8%–14.9%; TKW: 3.7%–26.3% individually	Huang et al. (2006)
Common wheat	DH	RL4452 × AC Domain	47 traits including Milling (GPC, FPC, PSI, FY), MG, FG, baking, starch properties, noodle color, and others	SSR	Milling: 12; MG: 24; FG: 9; baking: 14; starch properties: 19; noodle color: 11; others: 10	1A, 1B, 1D,2A, 2B, 3A, 3B, 3D, 4A, 4B, 4D, 5B, 5D, 6A, 6B, 7A, 7B, 7D for all traits combined	GPC: 6.2–29.8; FPC: 6.1%–28.7%; FY: 7.9%–11.9%; PSI:28.2%; MG: 4.4%–42%; FG: 4.6%–35.4%; baking: 7.9%–24.8%; starch properties: 4.1%–54.4%;	McCartney et al. (2006)

(Continued on following page)

TABLE 1 (*Continued*) Summary of representative QTL identification studies for end-use quality traits in wheat. Phenotypic variance is indicated collectively for all the QTLs or individually for each QTL.

Model species	Population type[a]	Mapping population	End-use quality traits[b]	Marker type[c]	Number of QTLs identified	Genome location	Phenotypic variance explained	References
							noodle color: 7.7%–36% individually	
Durum wheat	RIL	PDW 233 × Bhalegaon 4 (Landrace)	SV, GPC, MG	SSR, ISSR, SCAR, TRAP	26 main effect QTLs in total from 3 environments	SV: 1B; GPC: 7B; MG: 1A, 1B, 2B, 4B, 7A, 7B	SV: 6.7%–40.66%; GPC: 9.64%; MG: 6.75%–21.32% individually	Patil et al. (2009)
Common wheat	DH	Kurki (cultivar) × Janz (cultivar)	GPC, sponge and dough baking performance, GH	SSR	-	GPC: 1B, 3A, 5A, 5B, 7A; sponge and dough making: 1B,1D,3A, 4D, 5B, 5D, 7A, 7B; GH: 1A, 4D, 5D	-	Mann et al. (2009)
Durum wheat	DH	Rugby (cultivar) × Maier (cultivar)	Gluten strength	DArT, STS, EST-SSR	3	major QTL in 1BS	90% by single major QTL	Kumar et al. (2013)
Common wheat	RIL	WCB414 (elite line) × WCB617(exotic line)	TKW, KVW, GPC, FE, MG	DArT	TKW: 11; KVW: 10; GPC: 11; FE: 6, MG: 31	TKW: 2A, 2B, 2D, 3A, 3B, 4A, 5A, 6B; KVW: 1A, 1B, 2A, 4B, 5B, 6A, 6B, 7B; FE: 1A, 1B, 2B, 3D, 4A, 6A; GPC: 1A, 1B, 2B, 2D, 3D, 4B, 5B, 6B, 7B; MG: 2B, 3A, 6A, 6B, 7A, 7B, 7D	TKW: 6.2%–17.1%; KVW: 6.7%–22.5%; FE: 4.9–19; GPC: 4.7–16.9; MG: 5.3–19.9 individually	Echeverry-Solarte et al. (2015)
Common wheat	DH	Yumechikara (HRWW variety) × Kitahonami (SRWW variety)	GPC, FPC	SSR	1	2B	GPC: 32%; FPC: 16.5%	Terasawa et al. (2016)
Common wheat	RIL	Shannong01-35 (variety) × Gaocheng9411 (variety)	Pasting property	SNP, DArT, SSR	43 QTLs from 3 environments	1A, 1B, 2B, 2D, 3A, 4A, 4B, 5B, 6A, 6B, 7A	.11%–37.68% individually	Wang et al. (2017)
Common wheat	RIL	Glenn × Traverse (both hard red spring wheat cultivars)	GPC, FE and MG, Baking properties	SNP	GPC: additive QTL-11, digenic epistatic QTL-18; FE and MG: additive-32, digenic epistatic-51; baking properties: additive QTL-31, digenic epistatic-15	GPC (additive QTL): 1A, 1B, 2A, 2B, 3A, 3B, 4B, 5B, 7A; FE and MG (additive QTL): 1D, 2B, 3D, 5A, 6A, 6D; baking properties (additive QTL): all chromosomes except 1A, 2B, 3D, 6A	GPC: 6.5%–20%; FE and MG: 1%–24%, and baking properties: 2%–28% individually for additive QTLs	Naraghi et al. (2019)
Durum wheat	Diversity panel	7 *T. turgidum* subspecies: *durum, durum* var *ethiopicum, turanicum, polonicum, turgidum, carthlicum, dicoccum,* and *dicoccoides*	GPC	SNP	11 stable QTLs in at least 3 environments out of 7 tested environments	2BS, 3AL, 3BL,4AS, 4BL, 5AS, 5BL, 6BL, 7AS, 7BL	5.1%–8.7% individually across all environments	Nigro et al. (2019)
Common wheat	RIL	ND 705 (elite line) × PI 414566 (exotic line)	GPC, KH, FE	SNP	GPC: 14; KH: 7; FE: 11	GPC: 1A, 1B, 2A, 4B, 4D, 5A, 5B, 6D, 7A, 7B, 7D; KH: 1A, 1B, 4B, 5A, 7A; FE: 1A, 2A, 3A, 3B, 4D, 5A, 5B, 5D, 7D	GPC: 5.6%–24.9%; KH: 7.4%–21.3%; FE: 6.3%–12.6% individually	Kumar A et al. (2019)
Common wheat	Diversity panel	SRWW	FY, FPC, GH, FPC, SRC	SN	FY: 3; FPC: 5; SE: 2; SRC: 8	FY: 1B, 2A, 2B; FPC: 5A, 6A, 7A; GH: 4A, 4B; SRC: 1A, 1B, 3A, 4B, 5B, 7D	3.4%–6.0% individually	Gaire et al. (2019)
Common wheat	Diversity panel	Winter wheat	Grain quality traits (GPC, test weight, WGC, SV, FY, TSC), dough rheological	SNP	Grain quality traits: 246 QTNs; Dough rheological properties: 86	40 stable QTNs found in 1A, AD, 1D, 2A, 2B, 2D, 3A, 3B, 3D, 4B, 4D, 5A, 5D,	40 stable QTNs explained 4.38%–18.6% individually	Yang et al. (2020)

(Continued on following page)

TABLE 1 (*Continued*) Summary of representative QTL identification studies for end-use quality traits in wheat. Phenotypic variance is indicated collectively for all the QTLs or individually for each QTL.

Model species	Population type[a]	Mapping population	End-use quality traits[b]	Marker type[c]	Number of QTLs identified	Genome location	Phenotypic variance explained	References
			properties (DDT, DS and DWA)			6A, 6B, 6D, 7A, 7B, 7D		
Common wheat	DH	Two DH populations, Yecora Rojo × Ksu106 and Klasic × Ksu105	FG, MG traits	SNP	Total 176 additive QTLs for FG and MG traits in two populations combined	QTLs found in almost all 21 chromosomes in both populations	103 QTLs with PV ≥ 10% individually in all traits combining two populations	Barakat et al. (2020a)
Common wheat	DH	Two DH populations, Yecora Rojo × Ksu106 (YK) and Klasic × Ksu105(KK)	8 traits including FPC, gluten index, FN, AC, test weight, FM	SNP	Total 127 additive QTLs for all traits in two populations combined	QTLs found in almost all 21 chromosomes in both populations	72 QTLs with PV ≥ 10% individually in all traits combining two populations	Barakat et al. (2020b)
Common wheat	RIL	Tainong 18 (cultivar)× Linmai 6 (elite line)	GPC, SV, FG, FN, Starch pasting properties	SNP, DArT, SSR, EST-SSR	GPC: 10; SV: 11; FG: 17, FN: 4; starch pasting: 64	Protein: 1B,3B, 4A, 4B, 4D, 6A; SV: 1A, 1D, 4B, 5B, 5D, 6A, 6D	Protein: 7%–15.5%; SV: 5.3%–18.73%; FG: 5.6%–35%; FN: 8.4–10.7: starch pasting: 6.3%–17.9% individually	Guo et al. (2020a)
Common wheat	Diversity panel	SWW breeding lines and DH	14 traits categorized as grain characteristics, milling traits, flour characteristics and baking quality	SNP	Total 178 MTA including 12 large effect MTA	Large effect QTLs for Grain characteristics: 1B, 2B, 4B, 5A, and 6B; Milling traits: 1B, 1D, 5A, and 6B; Flour characteristics: 1B, 1D and 4A	-	Aoun et al. (2021)

[a]RIL, recombinant inbred line; RICL, recombinant inbred chromosome line; NIL, near-isogenic lines; DH, doubled haploid.
[b]GPC, grain protein content; FPC, flour protein content; SV, sedimentation volume; MG, mixograph; FG, farinograph; FN, falling number; GH, grain hardness; PSI, particle size index; FY, flour yield; FE, flour extraction; AC, ash content; FM, flour moisture; TKW, thousand kernel weight; KVW, kernel volume weight; SRC, solvent retention capacity; WGC, wet gluten content; TSC, total starch content; DDT, dough development time; DS, dough stability; DWA, dough water absorption.
[c]RFLP, restriction fragment length polymorphism; SSR, simple sequence repeat; AFLP, amplified fragment length polymorphism; STS, sequence-tagged site; ISSR, inter simple sequence repeat; SCAR, sequence characterized amplified region; TRAP, target region amplified polymorphism; DArT, diversity array technology; SNP, single nucleotide polymorphism; EST-SSR, expressed sequence tag-derived simple sequence repeat.

phenotype-assisted selection of superior lines for different plant species (D'hoop et al., 2014; dos Santos et al., 2016; Proietti et al., 2018; Bhattarai and Subudhi, 2019). However, given the complex inheritence nature of these quantitative traits and the significant effects of the environment and the G×E interaction for these traits, phenotypic selection is still required to confirm the efficacity of modern methods including MAS and GS.

The availability of plant genome sequence helps in the identification of candidate genes that are linked to available markers. However, initial sequencing was expensive and conducted by fewer labs which resulted in sequenced genomes of only a few model plant species including *Arabidopsis* (The Arabidopsis Genome Initiative, 2000), rice (International Rice Genome Sequencing Project and Sasaki, 2005), and maize (Schnable et al., 2009). Therefore, a lot of earlier genomic studies were conducted in these model plants and their findings were eventually transferred to other crop species which is now referred as the translational genomics apporach (Varshney et al., 2015). Rice was the first plant species to be used as a model cereal crop due to its small genome size and the availability of a well-mapped and characterized genome compared to other cereals like maize and wheat (Eckardt, 2000). With the revolutionization in next-generation sequencing (NGS) technologies such as Illumina and Nanopore sequencing, genome sequencing is much more accessible and affordable than ever before. Currrently, it is possible to genotype 10,000 lines at the same cost associated to phenotyping 1,000 wheat lines for end-use quality traits (Battenfield et al., 2016). Various molecular breeding tools including, but not limited to, biparental linkage mapping, genome-wide association studies (GWAS), MAS, and GS have been developed and are being extensively utilized for identifying functional loci/genes and for predicting the genomic/breeding value of individuals (Figure 1). This genomic information is then utilized for crop improvement and is referred to as genomics-assisted breeding (GAB). The molecular breeding tools used in GAB for wheat quality are discussed later in this chapter.

GAB facilitates the generation of an integrated database of genetic and genomic resources of many crops that can be used for translational genomics (Varshney et al., 2015). Translational genomics utilizes the scope of comparative, functional, and evolutionary genomics for finding relevant information from model species (Kang et al., 2016). Candidate gene approach (CGA) is one of the important tools of translational genomics. CGA assumes that genes within the model and target species could govern similar functions or influence the same trait (Salentijn et al., 2007). These genes could be either functional candidate genes with an identified or predicted function or positional candidate genes that are co-localized with a trait locus (Pflieger et al., 2001). As an example, the flowering time trait of many angiosperm species is functionally related to *Flowering Locus T (FT)* gene homologs (Pin and Nilsson, 2012; Subedi et al., 2021) and can be utilized for the study of the trait in any other flowering species. Positional candidate genes used for CGA

TABLE 2 An overview of markers used for marker-assisted selection (MAS) of end-use quality traits in wheat.

End-use quality trait[a]	Genes	Allele(s)/subunits	Chromosome	Marker type(s)[b]	References
GPC	*Gpc-B1*	Gene specific	6BS	SSR (*Xucw108*)	Uauy et al. (2006), Vishwakarma et al. (2014), Gautam et al. (2020)
		Linked	6BS	SSR (*Xuhw89*)	Distelfeld et al. (2006), Vishwakarma et al. (2014), Gupta et al. (2022)
HMW-GS	*Glu-A1*	Ax1, AxNull, Ax2*	1AL	AS-PCR, KASP	de Bustos et al. (2001), Ma et al. (2003), Ravel et al. (2020)
	Glu-B1	Bx7, Bx8, Bx9, Bx13, Bx14, Bx15, Bx16, Bx17, Bx20, Bx23, $Bx7^{OE}$	1BL	AS-PCR, STS, KASP	Ma et al. (2003), Butow et al. (2004), Lei et al. (2006), Ravel et al. (2020)
	Glu-D1	Dx2, Dx5, Dx3, Dy10, Dy12, gene specific	1DL	AS-PCR, KASP	Smith et al. (1994), de Bustos et al. (2001), Ma et al. (2003), Ravel et al. (2020), D'Ovidio and Anderson (1994)
LMW-GS	*Glu-A3*	*a, b, c, d, e, f, g*	1AS	STS, KASP	Wang et al. (2010), Dreisigacker et al. (2020)
	Glu-B3	*a, b, c, d, e, f, fg, g, h, i*	1BS	AS-PCR, STS, KASP	Shan et al. (2007), Wang et al. (2009), Dreisigacker et al. (2020)
	Glu-D3	*a, b, c, d, e, g, h, i, j, k*	1DS	STS	Zhao et al. (2007), Appelbee et al. (2009)
Grain hardness	*Pina-D1*	*Pina-D1a, Pina-D1b*	5DS	AS-PCR, STS	Gautier et al. (1994), Huang and Brûlé-Babel (2011), Chen et al. (2013), Rai et al. (2019) Qamar et al. (2014)
	Pinb-D1	*Pinb-D1a, Pinb-D1b, Pinb-D1c, Pinb-D1d, Pinb-D1e, Pinb-D1p*	5DS	STS, CAPS	Gautier et al. (1994), Tranquilli et al. (1999), Lillemo and Morris, (2000), Huang and Brûlé-Babel (2011), Qamar et al. (2014)
Starch	*Wx-A1*	*Wx-A1a, Wx-A1b, WxA1-c, Wx-A1d, Wx-A1e, Wx-A1f, Wx-A1g, Wx-A1h, Wx-A1i, Wx-A1* null mutant	7AS	SSR, AS-PCR, STS, RFLP, KASP	Vrinten et al. (1999), McLauchlan et al. (2001), Nakamura et al. (2002), Saito et al. (2004), Monari et al. (2005), Shan et al. (2007), Vanzetti et al. (2010), Yamamori and Guzmán (2013), Zhang et al. (2022)
	Wx-B1	Null *Wx-B1, Wx-B1a, Wx-B1b, Wx-B1e, Wx-B1*	4AL	SSR, AS-PCR, STS, RFLP	Briney et al. (1998), Vrinten et al. (1999), McLauchlan et al. (2001), Saito et al. (2004), Monari et al. (2005), Shan et al. (2007), Saito et al. (2009), Divashuk et al. (2011)
	Wx-D1	*WxD1a, Wx-D1b, Wx-D1*	7DS	AS-PCR, SSR, STS, RFLP	Vrinten et al. (1999), McLauchlan et al. (2001), Nakamura et al. (2002), Saito et al. (2004), Monari et al. (2005)

[a]GPC, grain protein content; HMW-GS, high molecular weight-glutenin subunit; LMW-GS, low molecular weight-glutenin subunit.
[b]STS, sequence-tagged site; CAPS, cleaved amplified polymorphic sequences; AS-PCR, allele specific-polymerase chain reaction; RFLP, restriction fragment length polymorphism; SSR, simple sequence repeat; KASP, kompetitive allele specific PCR.

can be genes identified through QTL mapping, marker-trait association, or syntenic regions between different genomes (Salentijn et al., 2007). Once the candidate gene has been identified, it can be translated to the target crop and validated through reverse genetic approaches (Figure 1).

There are several ways for validating orthologous genes in the target crop including the Targeting induced Local Lesion in genomes (TILLING) (Chen et al., 2012) and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) (Liang et al., 2017). TILLING by sequencing (TbyS) was successfully carried out for translation of genes in Mung bean from medicago, cowpea, and soybean (Tsai et al., 2011). CRISPR gene editing techniques have become extremely popular and are widely used in various crop species. There are many reviews available on the use of CRISPR gene editing technique for crop improvement (e.g., Bortesi and Fischer, 2015; Song et al., 2016; Arora and Narula, 2017). One of the concerns related to CRISPR technology is the presence of transgene after editing and the possibility of off-target mutations which can face regulatory restrictions (Liang et al., 2017). However, methods have been developed and deployed in wheat to create mutants with no detectable transgenes (Liang et al., 2017). This eliminates the need for laborious and time-consuming backcrossing steps to segregate away the CRISPR/Cas9 cassette. Therefore, researchers can select the validation methods according to their interest and resource availability. To complement the validation step of translational genomics, fine-mapping can be done in a population segregating for the trait and associating polymorphism within the candidate gene to the phenotypic variation (Salentijn et al., 2007). Since plant genome sequencing has become fairly routine, the goal now is to sequence every crop species and apply GAB techniques for crop improvement. In this context, there are a plethora of genomic resources from many different crops that can be utilized for translational genomics.

5 Genetic and genomic studies on wheat end-use quality traits: An overview

Genetic studies of quality traits are complex due to the polygenic nature of the traits and high G×E interaction (Zanetti et al., 2001;

TABLE 3 Summary of few representative genome-based prediction studies for end-use quality traits in wheat.

End-use quality traits[a]	Germplasm[b]	Marker[c]	Model(s) used[d]	Accuracy	References
FY, softness, gluten strength, water absorption, FPC, starch properties	Two DH biparental population of SWW of 209 and 174 individuals	DArT, AFLP, RFLP, SSR, STS	Multiple linear regression, Ridge regression, Bayes-C∏	.42–.66	Heffner et al. (2011)
GPC, FPC, test weight, TKW, grain hardness, FY, dough rheology, loaf volume, gluten strength	5,520 advanced lines of common wheat	SNP	PLSR, elastic net, Random Forest	.32–.62	Battenfield et al. (2016)
19 traits total including grain characteristics, milling traits, dough rheology, noodle traits, baking traits	2076 common wheat accessions	SNP	NA	>.5 for many traits	Hayes et al. (2017)
Protein content, dough rheology, baking quality	840 winter wheat lines including inbreds and DHs	DArT	RR-BLUP, W-BLUP	.38–.63	Michel et al. (2018)
Protein content	Two durum wheat panels of 189 and 159 genotypes consisting of varieties and breeding lines	DArT, SNP	RR-BLUP	.40–.46	Rapp et al. (2018)
GPC, alveograph measurements	Durum wheat panel consisting of 170 varieties and advanced breeding lines, and 154 DHs	SNP	Single trait models: RR-BLUP, G-BLUP, Bayes A, Bayes B, Bayesian LASSO, RKHS, Multi-trait models: MT-Bayes A, MT-Matrix, MT-SI	.5–.8 for single trait models	Haile et al. (2018)
FY, alveograph measurements	635 winter wheat breeding lines	SNP	G-BLUP, Bayesian Power Lasso	.50–.79	Kristensen et al. (2019)
14 traits including grain characteristics, milling traits, flour parameters, and baking traits	SWW panel of 666 lines consisting of inbreds, advanced breeding lines, and DHs	SNP	RR-BLUP, Bayes A, Bayes B, Bayes Lasso, Bayes C, Random Forest, SVM, MLP, CNN	average .58–.63 for all traits and models	Sandhu et al. (2021)
GPC, FPC, FY, dough rheology, bake absorption, bake mixing time, loaf volume, baking score	462 advanced breeding lines of winter wheat	SNP	G-BLUP	.25–.55	Zhang-Biehn et al. (2021)
14 traits including grain characteristics, milling traits, flour parameters, and baking traits	SWW panel of 672 lines consisting of breeding lines, cultivars, and DHs	SNP	G-BLUP	~.35–.65	Aoun et al. (2021)

[a]GPC, grain protein content; FPC, flour protein content; FY, flour yield; FPC, flour protein content; TKW, thousand kernel weight.
[b]DHs, double haploids; SWW, soft white wheat.
[c]SNP, single nucleotide polymorphism; STS, sequence-tagged site; DArT, diversity array technology; RFLP, restriction fragment length polymorphism; AFLP, amplified fragment length polymorphism.
[d]PLSR, partial least squares regression; SVM, support vector machine; MLP, multilayer perceptron; CNN, convolutional neural network; G-BLUP, genomic best linear unbiased prediction model; W-BLUP, weighted genomic best linear unbiased prediction; RR-BLUP, ridge regression best linear unbiased prediction.

Prasad et al., 2003; Huang et al., 2006). In addition, due to the polyploidy nature of common wheat, recessive mutation phenotypic effects are masked by the effective homoeoloci of other genomes creating an extra challenge in intrachromosomal mapping (Kuspira and Unrau, 1957; Chao et al., 1989). As a result, researchers have focused on increasing marker density to correctly identify loci associated with a certain phenotype. The use of molecular markers for the assessment of quality traits has been of interest since the late 1990s (Galande et al., 2001). Researchers are also selecting genotyping methods that allow better genome coverage and can identify a higher number of genetic variations.

Initial research for end-use quality traits was often carried out using RFLP markers that are unlimited in number, generally codominant, and able to recognize individual loci, thus being effective for the hexaploid genome of common wheat (Chao et al., 1989; Blanco et al., 1996; Joppa et al., 1997). However, RFLP markers were found to show a low level of polymorphism in wheat, especially when studying progenies of closely related genotypes (Devos et al., 1992). Later, microsatellites also called SSR, became more popular since they show a higher level of polymorphism than RFLP markers (Huang et al., 2006). SSR markers were used for linkage mapping and QTL identification for end-use quality traits (Table 1) (Prasad et al., 2003; Sourdille et al., 2003; Kuchel et al., 2006). A combination of RFLP, SSR, and Amplified Fragment Length Polymorphism (AFLP) was also used to increase the marker density in maps (Table 1) (Sourdille et al., 2003; Mohler et al., 2014). With the advancement in sequencing technologies, the use of SNP has been prioritized since they are abundant and uniformly distributed in a genome providing high genomic resolution (Gupta et al., 2008; Kang et al., 2016; Wang et al., 2017; Naraghi et al., 2019; Guo Y. et al., 2020). The availability of high throughput sequencing platform for SNPs makes it a preferred genotyping method in general. Genotyping-by-sequencing (GBS) has been a popular genotyping method that allows capturing SNPs through a reduced representation of the genome (Poland et al., 2012). A few of the representative QTL mapping studies utilizing various genotyping methods for end-use quality traits in wheat are summarized in Table 1.

QTLs are identified either by QTL mapping in a biparental mapping population or by association mapping (Beló et al., 2008). The majority of prior research work in wheat quality traits was concentrated on QTL mapping using different mapping populations like F2 generation, recombinant inbred lines (RILs), doubled-haploids (DHs), and near-isogenic lines (NILs) (Table 1). Recently, GWAS has also been a popular tool for the genetic study of

complex polygenic traits in species such as wheat, rice, maize, and barley because of many advantages over QTL mapping (Yang et al., 2020). Since GWAS is carried out in a diversity panel, it has higher allelic diversity and higher resolution as a result of genetic recombination events (Korte and Farlow, 2013). GWAS also allows the detection of quantitative trait nucleotides (QTNs) with small effects associated with complex traits (Yang et al., 2020). However, GWAS has less power and requires a larger population to detect an association compared to QTL mapping. On the positive side, GWAS can be used for the selection of parents for subsequent QTL analysis thus complementing each other (Korte and Farlow, 2013). A few representative GWAS studies for end-use quality traits are summarized in Table 1. With the assembly being available, research can be concentrated on the identification of candidate genes using QTL mapping and association studies (Li and Yang, 2017).

The generation of wheat genome assembly has been a great milestone for improving wheat end-use quality traits. A reference genome has allowed us to physically map the genes/loci related to end-use quality traits in wheat (Guo J. et al., 2020). The QTLs identified from the mapping studies are first located in the physical map of the chromosome and the candidate genes located in the QTL are noted using the gene annotation available from the assembly (Li Y et al., 2012). This has allowed the identification of novel QTLs/genes related to end-use quality traits in wheat (Li Y et al., 2012). Based on the gene function annotation and existing knowledge of quality development metabolic pathways, the function of these candidate genes are classified (Yang et al., 2020). Fine mapping has also been done to narrow down the list of candidate genes or to even identify the causal gene for traits such as GPC, grain weight and size in wheat (Olmos et al., 2003; Distelfeld et al., 2006; Uauy et al., 2006; Röder et al., 2008; Zhang et al., 2018). Markers related to the gene can then be used for marker-assisted selection of individuals having the desired genotype of interest. Wheat genome assembly has also allowed study of the homologs of the genes that could be present in all three subgenomes of wheat i.e., A, B and D genome. Moreover, it also allows assessments of intraspecies genomic variation and the availability of multiple assemblies from different lines can be further used for comparative mapping (Walkowiak et al., 2020).

All these genetic and genomic studies in wheat have generated valuable resources for wheat including reference genomes, whole-genome sequencing data, shotgun sequencing data, genome-wide genetic marker data, gene expression atlas, and diagnostic markers which are available across various platforms. Genome reference sequence browsers are available for different cultivars/species of wheat including the Chinese spring wheat (Appels et al., 2018), HRWW varieties such as Jagger and Mace, spelt, durum wheat, and wild relatives of wheat such as wild emmer and *Ae. tauschii* (https://wheat.pw.usda.gov/GG3/genome_browser#). PlanGDB (http://www.plantgdb.org/), Gramene (http://www.gramene.org/), and Phytozome (https://phytozome.jgi.doe.gov/pz/portal.html#) are additional sources for genomic information on many crops including wheat. In 2019, an international collaboration project "10 + genome project" released a wheat pan-genome containing a reference sequence of 10 global panels of wheat varieties (http://www.10wheatgenomes.com). As of November 2022, sequencing data from more than 61,500 bio-samples related to common wheat have been deposited in NCBI and are publicly available (https://www.ncbi.nlm.nih.gov/sra). These resources are being used for advancing studies of end-use quality traits in wheat and will continue to be useful in the future.

6 QTL/genes related to end-use quality traits in wheat

The genetic, genomic, and physiological information for various end-quality traits in wheat continues to grow. This includes the identification of major QTLs/genes controlling various traits of interest. These results will be instrumental for the continued success of GAB in wheat. The following are some of the major findings for end-use quality traits in wheat.

6.1 Grain protein content (GPC)

Joppa et al. (1997) identified a significant QTL related to GPC on chromosome 6BS contributing up to 66% phenotypic variance (PV) from a population of RIL of T. turgidium L. var dicoccoides (Table 1). This QTL was later mapped as a single Mendelian locus, Gpc-6B1, flanked by Xcdo365 and xUCW67 markers at 1.5 cM and 1.2 cM, respectively (Olmos et al., 2003). Distelfeld et al. (2006) carried out microcolinearity between the rice reference genome and wheat expressed sequence tags (ESTs) which narrowed down the position of the Gpc-B1 gene to .3 cM flanked by PCR markers Xucw79 and Xucw71. The same authors also identified a more tightly linked (.1 cM) high throughput codominant marker Xuhw89 that was suggested for the initial screening of the Gpc-B1 gene. So far, many QTLs have been identified for GPC, however improvement in GPC of wheat has been limited to the introgression of this major gene, *Gpc-B1* (Sengar and Singh, 2018). Another important finding for GPC was done by Terasawa et al. (2016) who studied DH lines of HRWW and identified a single major QTL, *QGpc.2B-yume*, on chromosome 2BS that explained 32% of PV for GPC and 16.5% of PV for FPC. The authors also recommended the use of flanking SSR marker *Xgpw4382* for MAS of the identified QTL. Similarly, Kumar A et al. (2019) studied GPC in RIL of common wheat where they identified two major and stable QTLs, *QGPC. ndsu.7A.2* and *QGPC. ndsu.7B*, in chromosome 7AL and 7B that explained 14.6% and 24.9% of PV, respectively. Additional studies on GPC have been summarized in Table 1.

6.2 Glutenin and gliadin

In hexaploid wheat, HMW-GS is encoded by three loci, Glu-A1, Glu-B1, and Glu-D1, found on the distal half of the long arm of chromosomes 1A, 1B, and 1D, respectively, whereas the LMW-GS loci, *Glu-A2*, *Glu-B3*, *Glu-D3* are found on short arms of those same group 1 chromosomes (Payne, 1987; Singh and Shepherd 1988). HMW-GS are extensively studied in wheat as they play a major role in dough elasticity (Payne, 1987). Wheat varieties can contain tightly linked genes with multiple alleles within these loci that code for "x" and "y" type glutenin subunits (Payne et al., 1981; Payne, 1987; Ravel et al., 2006; Mann et al., 2009). Ravel et al. (2020) identified 8, 22 and 9 different alleles at *Glu-A1*, *Glu-B1* and *Glu-D1* locus, respectively using the sodium dodecyl sulfate–polyacrylamide gel-electrophoresis (SDS–PAGE) method. The subunit "Dx5+Dy10"

coded by Glu-D1d, "Ax2" coded by Glu-A1b, and "Bx7+Bx8" coded by Glu-B1b have been found to have a positive effect on dough properties resulting in good bread making quality (Payne and Lawrence 1983; Shewry 2003; Pirozi et al., 2008; Ravel et al., 2020). Ravel et al. (2006) carried out an association study in bread wheat and identified Glu-B1-1 as the candidate gene for determining the quantity of HMW glutenin. Payne et al. (1984) identified the association of gluten strength and LMW glutenin subunits (LMW-GS), and the gliadins were found to be tightly linked to the LMW-GS. The α and β gliadin genes are found on the short arms of chromosomes 6A, 6B, and 6D, whereas the γ and ω gliadins occur in the same locations as the LMW glutenin subunits (Payne et al., 1984; Payne, 1987).

6.3 Starch

There are three homoeologous waxy genes in common wheat, Wx-A1, Wx-B1, and Wx-D1, located on chromosomes 7AS, 4AL, and 7DS, respectively (Chao et al., 1989), that produce three distinct Wx proteins (Nakamura et al., 1993). Chao et al. (1989) indicated that the Waxy gene originally present on chromosome 7B was translocated to chromosome 4AL. Based on the presence/absence of Wx proteins, wheat plants were categorized into eight groups: Type 1 for wild type (having all Wx-A1, Wx-B1, and Wx-D1 proteins), Type 8 for waxy wheat (no Wx protein), and Type 2–7 for partial waxy lines (with missing one or two Wx proteins) (Nakamura et al., 2002). Nakamura et al. (1993) developed the first waxy tetraploid (amylose free) by crossing partially waxy plants, Type7 common wheat and Type IV durum. Such waxy wheat has been used to products such as Asian wet noodles (Graybosch, 1998). Guzmán and Alvarez (2016) reported that there are 19 different waxy protein variants, and so far, molecular studies have identified 19, 15, and 7 allelic variants for Wx-A1, Wx-B1, and Wx-D1 gene, respectively. The *Wx-A1a*, *Wx-B1a* and *Wx-D1*a are the wild type alleles for these genes (Yamamori et al., 1994; Nakamura et al., 1995). Apart from normal functional alleles, alleles with loss of function (i.e., null allele) have also been identified for all three genes with null alleles for *Wx-A1* and *Wx-B1* gene being more common than *Wx-D1 gene* (Guzmán et al., 2012b). The use of null alleles allows researchers to create partial waxy and waxy starch (Zhang et al., 2022), that has a beneficial effect on quality of products as discussed above. Therefore, identification and discovery of the alleles has been of major focus for grain starch. The *Wx* genes consist of 12 exons and 11 introns similar to that in barley and molecular markers have been successfully used to identify null mutants for these genes (Guzmán and Alvarez, 2016).

6.4 Grain hardness

Early studies (Symes, 1965) speculated that the grain hardness that distinguishes between hard and soft wheat is caused by a single major gene. This Hardness (Ha) locus is present in the short arm of chromosome 5D and contains Pina-D1, Pinb-D1, and Gsp-1 genes which encode for puroindolines a (Pina), puroindolines b (Pinb), and Grain Softness protein-1, respectively (Morris, 2002; Bhave and Morris, 2008; Guzmán and Alvarez, 2016). So far, nine Pina alleles and 17 Pinb alleles have been identified, out of which Pina-D1b, Pinb-D1b, Pinb-D1c, and Pinb-D1d are the major alleles identified in hard wheat cultivars (Bhave and Morris, 2008; Huang and Brûlé-Babel, 2011).

The durum wheat (genome AABB) has no D genome (Morris et al., 2011) therefore, the *Ha* locus is completely absent and it expresses extremely hard texture, harder than "hard" common wheat (Morris et al., 2011). A novel soft grain texture referred as "super soft" has also been identified among common as well as durum wheat and is found to be associated with higher break flour yield (Morris et al., 2011; Wang et al., 2012; Ibba et al., 2019; Kumar N et al., 2019). Kumar N et al. (2019) identified four major QTLs on chromosome 1BS, 4BS, 5AL and 7AS explaining 15%–19% of phenotypic variance for soft kernel texture in spring wheat. Wang et al. (2012) and Ibba et al. (2019) also identified QTLs associated to grain texture on chromosome 1BS, 4BS, 5BS, 2DS, 4DS and 5DL. These results provide evidence that the grain hardness is also controlled by genomic regions other than the *Ha* locus. The identification of super soft grain texture is also very useful as it could possibly provide bakers a new type of flour (Kumar A et al., 2019). These findings will further help to improve the understanding of kernel texture in common and durum wheats.

6.5 Sedimentation volume (SV)

Several studies have identified many loci associated with SV on chromosomes 3A, 3B, 4A, 4B, 5A, 5B, 6A, 7A, and 7B (Blanco et al., 1998; Elouafi et al., 2000; Patil et al., 2009). Patil et al. (2009) identified three main effect QTLs for SV in proximity to Glu-B1, Glu-B2, and Glu-B3 glutenin coding loci. In the same study, the QTL QSv.macs-1B.1, flanked by marker interval Xgwm550 and Glu-B3, explained 9.18%–40.6% of PV across five environments. The same authors also found that 22 main effect QTLs for various quality traits such as SV, GPC, and mixograph parameters formed five different clusters on chromosomes 1B, 4B, 7A, and 7B.

To summarize, genetic mapping studies for end-use quality traits in wheat have provided valuable information regarding putative causal/candidate genes or loci related to these traits. Besides this, many studies have also reported an important physiological correlation between various quality traits in wheat including the negative correlations between GPC and grain yield (Cox et al., 1985; Simmonds, 1995). Yang et al. (2020) discovered that the GPC, wet gluten content, and starch content in wheat to be highly positively correlated. The authors also suggested the use of such pleiotropic QTL for the selection of improved quality traits. A significant positive correlation has also been found between dough rheological properties such as dough development time and dough stability time (Li J et al., 2012; Yang et al., 2020). However, a negative correlation was reported between GPC and dough stability time and grain hardiness, wet gluten content and dough stability time and wet gluten content and dough development time (Li J et al., 2012). This information will be useful for improvement of end-use quality traits in wheat.

7 Marker-assisted breeding (MAB) and translational genomics for end-use quality traits

The use of molecular markers has significantly impacted the area of plant breeding and genetics. The linkage between major genes and quantitative trait effects was first reported by Sax (1923) and later

Thoday (1961) reported the use of gene markers to locate the QTL. Molecular markers improve the efficiency of breeding programs by allowing early generation screening of individuals for targeted trait(s) which saves/reduces resources, energy, and time (Gale, 2005; Xu and Crouch, 2008). Molecular markers have been utilized for the selection of individuals in a breeding program through MAS and GS. Since quality traits are evaluated on harvested grains, i.e., at the end of the crop season, they are ideal targets for MAS and GS (Hayes et al., 2017).

7.1 Marker-assisted selection (MAS)

MAS is an important tool in plant breeding. Molecular markers can either be linked to the gene/locus or are diagnostic for a trait of interest (Gale, 2005; Xu and Crouch 2008). QTL mapping and GWAS studies in wheat have been able to identify markers that are tightly linked to the genes/locus influencing the trait and are therefore co-inherited with the trait. On the other hand, diagnostic/functional markers are those that are developed from actual gene sequence influencing the trait and they do not need independent validation for each parental line in a breeding program (Gale, 2005; Lau et al., 2015). MAS is used in many conditions including: 1) breeding for traits where conventional phenotypic selection is difficult, costly or time-consuming; 2) breeding for traits with high environmental influence or for traits whose selection depends on a specific environment and/or developmental stage; 3) speed breeding backcrossing or maintenance of recessive alleles during backcrossing also referred as marker-assisted backcrossing (MABC); and 4) pyramiding multiple favorable alleles within a single population also referred as marker-assisted recurrent selection (MARS) (Xu and Crouch, 2008).

The use of MAS method for end-use quality traits is a much more convenient, efficient, and faster alternative than the conventional method of phenotypic measurement and selection of individuals. In the past two decades, several studies have identified markers related to quality traits that were recommended for use in MAS. For wheat, there are more than 97 functional markers that are being used in breeding programs for various traits of interest including end-use quality traits (Nadeem et al., 2018). Table 2 provides an overview of markers being used for MAS of five end-use quality traits in wheat.

The PCR markers, Xucw71 and Xuhw89 (codominant), developed by Distelfeld et al. (2006) have been used for selecting genotypes having the Gpc-B1 gene related to high GPC content (https://maswheat.ucdavis.edu/protocols/HGPC) (Table 2). Bokore et al. (2019) also validated the use of Xucw71 and Xuhw89 markers for MAS of genotypes with or without the Gpc-B1 gene from four common hexaploid wheat populations. The authors reported an increment in the GPC content of lines with the Gpc-B1 positive allele. Similarly, Ravel et al. (2006) claimed that the Glu-B1 linked markers can be used by breeders for the indirect selection of high protein genotypes.

MAS has also been used for traits other than GPC in wheat. Huang and Brûlé-Babel (2011) developed simple and co-dominant PCR markers to select grain hardness targeting the *Pina* and *Pinb* alleles (Table 2). Patil et al. (2009) recommended the use of markers *Xgwm550* and *Glu-B3* flanking the QTL *QSv.macs-1B.1* for MAB of SV. Likewise, Ibba et al. (2018) developed haplotype specific molecular markers for identification of specific haplotypes of *Glu-A3*, *Glu-B3* and *Glu-D3* locus associated with gluten strength.

The negative relationship between grain yield and GPC, and the significant interactions between these traits and the environment make it challenging for GPC improvement through conventional breeding (Cox et al., 1985; Simmonds, 1995; Rapp et al., 2018) and MAS has been suggested for improvement in both traits. The negative relationship between these traits could be due to plants requiring more energy to produce the same amount of protein than carbohydrate (De Vries et al., 1974; Blanco et al., 1996). One of the strategy for ameliorating such negative effect is to do the MAS of GPC simultaneously with the phenotypic selection for other yield and quality traits (Tabbita et al., 2017). As an example, Kumar et al. (2011) carried out MAS of *Gpc-B1* and reported that MAS-derived progenies did not show proportional decline of grain yield. Therefore, the authors were able to select for lines with high GPC and no penalty for grain yield. The other alternative would be to identify genes with no negative effect on the grain yield and use it in breeding program. Terasawa et al. (2016) identified a major QTL, *QGpc.2B-yume*, on chromosome 2B which had no significant negative effect on grain yield and other yield components traits. The effect of *QGpc.2B-yume* should be tested in other populations and environments as well. These findings are promising for the development of wheat varieties with improved yield and end-use quality.

Vishwakarma et al. (2014) carried out MABC for introgression of *Gpc-B1* from a high GPC variety, Glu269, to an elite variety HUW468. Elite lines were developed within a relatively short period of two and half years (five crop cycles), and improved lines were selected with significantly higher GPC and consisted of 88.4%–92.3% of the recurrent parent plant genome. The authors utilized the SSR marker *Xucw108* developed by Uauy et al. (2006) for foreground selection, and 86 other polymorphic SSR markers were used for background selection on recovery of the recurrent parent plant genome. Similarly; Rai et al. (2019) carried out MABC for the introgression of the grain softness gene, *PinaD1a*, into a hard-grained variety to develop soft grain wheat lines. In this study, the authors carried out a foreground selection for the *PinaD1a* allele using *PinA* marker (Gautier et al., 1994), background selection using 173 SSR markers covering all 21 chromosomes, and a negative selection for *Pina-D1b* allele using Pina-N1 marker (Chen et al., 2013). MARS has also been used in wheat breeding for grain yield and end-use quality traits. Maich et al. (2020) reported results from 12 cycles (24 years) of recurrent selection for grain yield among 83 F1 hybrids obtained by crossing 16 commercial varieties. The authors obtained a yield increment of 1.3% per year and at the end of 12 cycles and the improvement in yield did not affect the baking quality of wheat.

Although MAS seems a promising step forward in the field of MAB, it does have some limitations. MAS is based on the prediction accuracy of previously identified significant markers linked to major traits and these markers are few in number (Heffner et al., 2011). This limits the use of MAS for complex quantitative traits as accurate prediction of these traits will require markers for all of the associated genes (Heffner et al., 2011). These genes/QTLs also interact with each other at different expression levels thus changing the phenotype of an individual. MAS cannot account for these interaction effects because most of the markers are initially developed from a mapping population segregating for a single or few QTL which usually leads to overestimation of the QTL effect. Therefore, MAS is mostly constrained to simply inherited or monogenic traits (Michel et al., 2018). In addition, MAS also does not factor the environmental

influence on the trait leading to lower phenotype prediction accuracy (Xu and Crouch, 2008; Heffner et al., 2009). Breeding programs now have integrated GS in their MAB as it considers all of these limitations of MAS, and these two tools complement each other.

7.2 Genomic selection (GS)

In GS, individuals are selected based on whole-genome marker profiling data that provides an overall performance evaluation of the plant (Varshney et al., 2015). The use of a whole-genome marker increases the chance that all QTLs are in linkage disequilibrium with at least one marker (Varshney et al., 2015). Due to this reason, unlike MAS, the GS does not require prior knowledge about large effect QTLs (Guo Y. et al., 2020). Also, for quantitative traits, GS usually has higher prediction accuracy than conventional MAS. In GS, phenotypic and genotypic marker data from the training population are fitted into the statistical models to get an estimation of all marker effects which are then used on unobserved genotypes of a testing population to calculate the genomic estimated breeding value (GEBV) (Heffner et al., 2009). Individuals are then selected based on the GEBV. The training population must be representative of the testing population to maximize the GEBV accuracy (Heffner et al., 2009). These training populations are advanced breeding materials that have been well characterized phenotypically and the GEBV of the testing population genotypes is predicted based on their genetic relation with the training population (Guzman et al., 2016; Michel et al., 2017; Michel et al., 2018). To assure proper selection, well-known check varieties, such as varieties with both good and poor bread-making quality, are included in the GS process that will help in validation of the selection as well as provide a reference for quality profiles (Guzman et al., 2016). However, extensive and good phenotypic data of the training population is required at the forefront in the wheat breeding program to generate accurate GEBV before GS can be used to select quantitative traits.

GS allows a 2–3 years earlier selection of many end-use quality traits in wheat than traditional selections and it is done on a much broader population allowing the selection of lines with good end-use quality and higher yield (Michel et al., 2018). Table 3 summarizes some of the GS studies carried out over the last decade for end-use quality traits in wheat. CIMMYT has been conducting GS for end-use quality traits in wheat since 2012 (Battenfield et al., 2016). In the CIMMYT breeding program, ~10,000 first-year yield trials lines are genotyped and GEBV values are estimated for grain yield and end-use quality traits (Guzman et al., 2016). Heffner et al. (2011) compared MAS, GS, and phenotypic prediction accuracy of genetic value for nine different grain quality traits in biparental populations of soft winter wheat and found that the GS to phenotypic selection accuracy was as high as .66 and the GS prediction accuracy was superior to conventional MAS (Table 3). Similarly, Hayes et al. (2017) conducted a study on genomic prediction of 19 end-use quality traits and observed prediction accuracy greater than .5 for many of the traits (Table 3). Besides this, Michel et al. (2018) also carried out GS of baking quality in wheat and reported that an acceptable prediction accuracy of .38–.63 can be obtained in all dough rheological traits.

GS is still emerging, and advanced computation models are being developed to increase the accuracy in the prediction of GEBV value. Several GS models have been developed which also account for G×E interaction thus increasing prediction accuracy across environments (Heslot et al., 2014; Jarquín et al., 2014). Known candidate genes can be used to correlate the genotype-level and gene-level G×E interactions which can be used to predict the influence of G×E interaction (Li et al., 2018). Models are being developed for various G×E conditions including: 1) predicting tested genotypes in untested environments; 2) predicting untested genotypes in tested environments, and 3) predicting untested genotypes in untested environments (Li et al., 2018). Multi-trait GS models have also been developed that can account for the genetic correlation among the traits and improve the prediction accuracy of the primary trait when data for secondary correlated traits are available (Gill et al., 2021). Sandhu et al. (2022) carried out genomic prediction for seven end-use quality traits in winter wheat and found that the multi-trait models performed 5.5% and 7.9% superior to uni-trait GS models for within-environment and across location predictions, respectively, and multi-trait-multi-environment models performed 10.5% superior to the uni-trait models. Similar results have been found in other GS studies in wheat as well (Hayes et al., 2017; Guo J. et al., 2020; Gill et al., 2021). These results will certainly increase breeder's confidence in integrating GS into their breeding program.

The UGA and SUNGRAINS (Southern Universities Grains, http://www.sungrains.lsu.edu/data.shtml), a cooperative research program between seven universities has been utilizing GAB for the development of superior germplasm of SRWW for the southeast United States. In collaboration with the USDA-ARS genotyping center (lead by Dr. Gina Brown-Guedira, https://wheat.pw.usda.gov/GenotypingLabs/?q=about), UGA and the SUNGRAINS breeding programs are using GS for important economic traits since 2015. MAS is also an important part of the program. Diagnostic markers are used for the selection of desirable genotypes for traits such as grain texture/softness using *Pinb-D1a* and *Pinb-D1a* markers, GPC using *Gpc-B1* marker, and HMW-GS using *Glu-A1* (Ax2, Ax1, and Axnull subunits), *Glu-B1* ($Bx7^{OE}$), and *Glu-D1* (Dx5+Dx10, Dx2+Dx12) markers etc.

7.3 Candidate gene approach (CGA)

CGA is one of the translational genomics tools that has been practiced in wheat for different traits including thousand kernel weight (TKW), kernel size, GPC, yield, plant height, and disease resistance (Faris et al., 1999; Zhang et al., 2014; Nigro et al., 2019; Sehgal et al., 2019; Zhang et al., 2020). Zhang et al. (2014) carried out a translational genomics study of grain size regulation gene in wheat where they utilized the rice candidate gene, OsGS3. This gene is well characterized, negative regulator of grain size and explained 80%–90% of phenotypic variation for grain weight and grain length in rice. The orthologous TaGS gene in wheat, *TaGS-D1* on chromosome 7DS was identified based on sequence identity and predicted protein similarity (Zhang et al., 2014). The authors were also able to identify QTL associated with *TaGS-D1* using SSR markers for chromosome 7D. Similarly, a recent study by Nigro et al. (2019) utilized 14 candidate genes for GPC and/or yield for gene-based association mapping by identification of SNP for candidate genes across a diversity panel. The authors found that genes, *AlaAT4A*, *ASN1-5A*, *NR-6A*, and *GS2-2B* were significantly associated with GPC across seven environments. They further identified 11 stable QTL for GPC using GWAS and suggested that the utilization of the CGA and GWAS in parallel can

increase the power and precision of QTL detection by reducing Type I and Type II error rates. Garg et al. (2009) identified a novel pair of HMW glutenin subunits, *Glu-S*^{s}*1* from a study on Ae. searsii, an S genome wild ancestor of wheat. They indicated that these would be good candidate subunits that can improve the bread-making quality of wheat.

8 Challenges and future opportunities for improving wheat end-use qualities

Phenotyping for end-use quality traits is challenging and requires a great deal of resources and time, owing to the fact that the quality assessment for direct screening of elite breeding lines is only feasible later in the advanced generations when other important traits including grain yield, disease and pest resistance have been selected for. Therefore, it is imperative to optimize phenotyping and selection of end-use quality traits. An effective strategy is to study the G×E interaction for end-use quality traits. Traits highly influenced by environment, such as GPC and FPC, should be tested in multiple environments to allow breeders to get a better estimation of the value of a particular genotype (Aoun et al., 2021). Traits not as affected by environment can be tested in fewer environments which saves resources to test additional lines in the breeding program (Aoun et al., 2021). In addition, the G×E interaction provides information on genotype stability across environment and identifies the target environment for maximizing genetic gain (Dias et al., 2018). Moreover, future advancements in phenotypic analysis equipment and rapid assay techniques, including automated phenotyping and phenomics, could further reduce time to analyze multiple samples and reduce resources.

Not all research facilities and breeding programs have cutting-edge laboratories, equipment, and expertise to operate and assess all quality attributes. In fact, only a few U.S. states and federal agencies (e.g., USDA-ARS) have wheat end-use quality assessment labs available. With regards to such traditional assessment methods, early generation MAS of desirable genotype(s) for the end-use quality traits is much more convenient and efficient. However, available markers are limited to only some of the genes identified so far and are already being used for selecting associated genes in the breeding lines. Therefore, there is a continual need for introgression of other major genes associated with these traits. One of the significant bottlenecks to QTL mapping and identification is the lack of validation studies. Hence, the plethora of QTLs identified to date for several end-use quality traits have few practical implications. Kiszonas and Morris, (2017) stated: "wheat breeding community have collected numerous markers and QTLs, but very few of them are being used for the improvement of wheat …. collecting countless QTLs is rather meaningless unless we take the next step and develop and use markers to improve the quality and consistency of wheat." Therefore, the wheat breeding community should focus fine mapping on identified, important major QTLs, functionally validate and clone the candidate gene(s), and develop markers associated with the genes for MAS of end-use quality traits. However, it is also important to consider that most QTLs identified so far are not major and/or they explain small portion of PV and are population and/or environment specific. Minor QTLs can be useful for their additive effect, but it is obvious that the major QTLs such as *GPC-B1* are the ones that add the most value in a breeding program. Therefore, the quest for the major/novel QTL(s)/gene(s) should continue in balance with validation studies. Moreover, the identified major genes can also be integrated as a fixed effect in the genomic prediction models for further improving the accuracy of GS (Bernardo, 2014).

Conventional genetic map-based QTL/gene discovery of traits requires screening of numerous breeding lines to identify the individuals segregating for the traits of interest. However, individuals segregating for traits of interest may not be present in the germplasm managed by the breeding program or the identified loci for segregating population could only explain a small amount of phenotypic variance. Besides this, the development of crosses could also be troublesome in the plants due to flower morphology or unsynchronized flowering between the lines. Therefore, researchers instead can look for studies done on the same traits across crops or even across related plant species to find genomic information that can be translated to the target crop, wheat. For example, wild emmer, *T. turgidum* L ssp *dicoccoides*, contains higher GPC than most hexaploid/tetraploid wheat cultivars (Avivi, 1978; Nevo et al., 2002) and could potentially be used to identify positive alleles of genes governing the GPC (Distelfeld et al., 2006). Therefore, studies done on GPC in wild emmer can provide useful genomic information to improve the GPC of durum and hexaploid wheat using the methods discussed above. In addition, other crop species within the Gramineae family such as rice, maize, barley, and oat have a greater chance of synteny or gene co-linearity with wheat and can influence similar phenotypes. Therefore, genetic findings on end-use quality traits in these crops should be further explored for translation to wheat.

Translational genomics through comparative genomics and functional genomics allows for a more integrated improvement of crop species. Non-etheless, translational genomics can be challenging due to the limitations of the molecular breeding approaches and tools used by the breeding programs. The use of MAS could be hindered if there are limited significant markers for the traits of interest. The GS can be limited based on the models used for the prediction of breeding values and the density of the markers. The CGA approach is solely based on the identification of functionally characterized genes or positional candidate genes which can be limited in number. However, these challenges can be overcome as additional research and findings are being put forth requiring continued and improved access to data. Therefore, it is necessary for the development of a single database that allows access to all these resources as proposed by (Kang et al., 2016). Infrastructure and management system of these data and the competence to handle large/sophisticated data are critical (Varshney et al., 2015). Additionally, the future of translational genomics also depends on how accessible these NGS technologies are to breeding/research programs. Genomic information and breakthroughs are the foundations for translational genomics for crop improvement in the future. This would also be beneficial to neglected and underutilized, yet nutritionally important crops that are being left behind from genetic improvement such as quinoa, finger millet etc. (Subedi, 2020; Gautam et al., 2021).

Consideration should also be given for the establishment of expanded private-public research collaborations which have been in limited number and scale in wheat breeding. These two institutions worked independently with a common goal of developing improved varieties of wheat. As a result, these two institutions were in competing space with each other. Therefore, there is a need for establishment of a collaboration where both benefit without compromising the

institutional integrity and success. Such collaboration could be made by sharing resources such as epidemiological information, technologies such as sequencing platforms, labs, equipment such as planters and harvesters or research fields for testing varieties in diverse environmental conditions. One such example is the International Wheat Yield Partnership (IWYP) (https://iwyp.org/) where both private and public institutions are working together to achieve a common goal of increasing the genetic yield potential of wheat by 50% by 2035 (https://iwyp.org/global-challenge/). The partnership combines the discovery, and research expertise of the public institutes with the private industries excellence in taking validated discoveries into breeding programs, developing putative varieties, large-scale testing, multiplying seeds of the varieties and delivery of improved varieties to farmers across the globe. There is a need of such bilateral relationship between universities and industries and the joint findings can be further disseminated in collaboration with the extension service within each of the institutions.

To conclude, research towards improved end-use quality of wheat utilizing recently developed modern tools/methodologies and technologies such as MAS, GS, and CGA clearly indicate that the adoption of GAB is increasing and will likely continue to expand in the future. This approach facilitates the plant breeding objectives of identifying and selecting individuals with desirable and economic end-use quality traits in a more effective and efficient way. It speeds up the process of crop improvement and the development of adapted cultivars. Utilization of GAB approaches, combined with targeted phenotypic assessments to validate plant selections, provides wheat breeders with ever-powerful tools to maximize wheat improvement. It is also vital to innovate ways that allow both private and public sectors to work hand-in-hand to address future challenges, including increasing productivity and end-use quality of major crops. Overall, understanding and modifying the crop genomic architecture will advance wheat germplasm with promising end-use quality traits in addition to increased grain yield that can ensure food and nutritional security for the escalating global population.

Author contributions

MS and MM conceptualized the review. MS designed the overall outline and prepared the first draft. BG, JWB, JB, and MM provided guidance, critical suggestions, and feedback throughout the manuscript preparation. All authors contributed to the review and approved the submitted version.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

AbuHammad, W. A., Elias, E. M., Manthey, F. A., Alamri, M. S., and Mergoum, M. (2012). A comparison of methods for assessing dough and gluten strength of durum wheat and their relationship to pasta cooking quality. *quality. Int. J. Food Sci. Technol.* 47(12), 2561–2573. doi:10.1111/j.1365-2621.2012.03135.x

Alamri, M., Manthey, F., Mergoum, M., Elias, E., and Khan, K. (2009). Assessing spring wheat quality using the glutograph instrument. *Cereal Foods World* 54, 124–131. doi:10.1094/cfw-54-3-0124

Anjum, F. M., Khan, M. R., Din, A., Saeed, M., Pasha, I., and Arshad, M. U. (2007). Wheat gluten: High molecular weight glutenin subunits—structure, genetics, and relation to dough elasticity. *J. food Sci.* 72 (3), R56–R63. doi:10.1111/j.1750-3841.2007.00292.x

Aoun, M., Carter, A., Thompson, Y. A., Ward, B., Morris, C. F., and Elias, E. M. (2021). Novel stripe rust all-stage resistance loci identified in a worldwide collection of durum wheat using genome-wide association mapping. *Plant Genome* 14 (3), e20136. doi:10.1002/tpg2.20128

Appelbee, M.-J., Mekuria, G. T., Nagasandra, V., Bonneau, J. P., Eagles, H. A., Eastwood, R. F., et al. (2009). Novel allelic variants encoded at the Glu-D3 locus in bread wheat. *J. Cereal Sci.* 49 (2), 254–261. doi:10.1016/j.jcs.2008.10.011

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 (6403), eaar7191. doi:10.1126/science.aar7191

Arora, L., and Narula, A. (2017). Gene editing and crop improvement using CRISPR-cas9 system. *Front. Plant Sci.* 8, 1932. doi:10.3389/fpls.2017.01932

Avivi, L. (1978). "High grain protein content in wild tetraploid wheat Triticum dicoccoides Korn," in Proceeding of the 5th International Wheat genetics Symposium (IEEE), 372–380.

Barakat, M., Al-Doss, A., Moustafa, K., Motawei, M., Alamri, M., Mergoum, M., et al. (2020a). QTL analysis of farinograph and mixograph related traits in spring wheat under heat stress conditions. *Mol. Biol. Rep.* 47 (7), 5477–5486. doi:10.1007/s11033-020-05638-6

Barakat, M. N., Al-Doss, A. A., Moustafa, K. A., Motawei, M. I., Alamri, M. S., Mergoum, M., et al. (2020b). Molecular detection of QTLs for flour quality traits in two doubled haploid populations in spring wheat under heat stress. *Cereal Res. Commun.* 48 (4), 525–532. doi:10.1007/s42976-020-00084-3

Battenfield, S. D., Guzmán, C., Gaynor, R. C., Singh, R. P., Peña, R. J., Dreisigacker, S., et al. (2016). Genomic selection for processing and end-use quality traits in the CIMMYT spring bread wheat breeding program. *Plant Genome* 9 (2). doi:10.3835/plantgenome2016.01.0005

Beló, A., Zheng, P., Luck, S., Shen, B., Meyer, D. J., Li, B., et al. (2008). Whole genome scan detects an allelic variant of fad2 associated with increased oleic acid levels in maize. *Mol. Genet. Genomics* 279, 1–10. doi:10.1007/s00438-007-0289-y

Bernardo, R. (2014). Genomewide selection when major genes are known. *Crop Sci.* 54 (1), 68–75. doi:10.2135/cropsci2013.05.0315

Bhattarai, U., and Subudhi, P. K. (2019). Genetic diversity, population structure, and marker-trait association for drought tolerance in US rice germplasm. *Plants* 8 (12), 530. doi:10.3390/plants8120530

Bhave, M., and Morris, C. F. (2008). Molecular genetics of puroindolines and related genes: Allelic diversity in wheat and other grasses. *Plant Mol. Biol.* 66 (3), 205–219. doi:10.1007/s11103-007-9263-7

Blanco, A., Giovanni, C. D., Laddomada, B., Sciancalepore, A., Simeone, R., Devos, K. M., et al. (1996). Quantitative trait loci influencing grain protein content in tetraploid wheats. *Plant Breed.* 115 (5), 310–316. doi:10.1111/j.1439-0523.1996.tb00925.x

Blanco, A., Bellomo, M. P., Lotti, C., Maniglio, T., Pasqualone, A., Simeone, R., et al. (1998). Genetic mapping of sedimentation volume across environments using recombinant inbred lines of durum wheat. *Plant Breed.* 117 (5), 413–417. doi:10.1111/j.1439-0523.1998.tb01965.x

Blazek, J., and Copeland, L. (2008). Pasting and swelling properties of wheat flour and starch in relation to amylose content. *Carbohydr. Polym.* 71 (3), 380–387. doi:10.1016/j.carbpol.2007.06.010

Bokore, F. E., Knox, R. E., DePauw, R. M., Cuthbert, R. D., Valerio, I. P., Clarke, F. R., et al. (2019). Validation of the effects of the Gpc-B1 high grain protein concentration locus from Lillian hard red spring wheat (*Triticum aestivum* L.) using locus specific markers. *Euphytica* 215 (1), 2. doi:10.1007/s10681-018-2322-0

Bortesi, L., and Fischer, R. (2015). The CRISPR/Cas9 system for plant genome editing and beyond. *Biotechnol. Adv.* 33 (1), 41–52. doi:10.1016/j.biotechadv.2014.12.006

Brevis, J. C., Morris, C. F., Manthey, F., and Dubcovsky, J. (2010). Effect of the grain protein content locus Gpc-B1 on bread and pasta quality. *J. Cereal Sci.* 51 (3), 357–365. doi:10.1016/j.jcs.2010.02.004

Briney, A., Wilson, R., Potter, R. H., Barclay, I., Crosbie, G., Jones, M. G. K., et al. (1998). A PCR-based marker for selection of starch and potential noodle quality in wheat. *Mol. Breed.* 4, 427–433. doi:10.1023/a:1009664917998

Bushuk, W. (1998). Wheat breeding for end-product use. *Euphytica* 100, 137–145. doi:10.1023/a:1018368316547

Butow, B. J., Gale, K. R., Ikea, J., Juhasz, A., Bedo, Z., Tamas, L., et al. (2004). Dissemination of the highly expressed Bx7 glutenin subunit (Glu-B1al allele) in wheat as revealed by novel PCR markers and RP-HPLC. *Theor. Appl. Genet.* 109 (7), 1525–1535. doi:10.1007/s00122-004-1776-8

Chang, C.-Y., and Chambers Iv, E. (1992). *Flavor characterization of breads made from hard red winter wheat and hard white winter wheat*'. American association of cereal chemists, 4.

Chao, S., Sharp, P. J., Worland, A. J., Warham, E. J., Koebner, R. M. D., and Gale, M. D. (1989). RFLP-based genetic maps of wheat homoeologous group 7 chromosomes. *Theor. Appl. Genet.* 78 (4), 495–504. doi:10.1007/BF00290833

Chen, L., Huang, L., Min, D., Phillips, A., Wang, S., Madgwick, P. J., et al. (2012). "Development and characterization of a new TILLING population of common bread wheat (*Triticum aestivum* L.)," in *PLoS ONE*. Editor M. Grebe (Sweden: Umeå Plant Science Centre), 7, e41570. doi:10.1371/journal.pone.0041570

Chen, F., Li, H., and Cui, D. (2013). Discovery, distribution and diversity of Puroindoline-D1 genes in bread wheat from five countries (*Triticum aestivum* L.). *BMC Plant Biol.* 13 (1), 125. doi:10.1186/1471-2229-13-125

Clark, J. A., and Bayles, B. B. (1935). *Classification of wheat varieties grown in the United States*. Washington, DC: United states department of agriculture, 214.

Cox, M. C., Qualset, C. O., and Rains, D. W. (1985). Genetic variation for nitrogen assimilation and translocation in wheat. I. Dry matter and nitrogen accumulation [1]. *Crop Sci.* 25 (3), 430–435. doi:10.2135/cropsci1985.0011183X002500030002x

Dale, E. M., and Housley, T. L. (1986). Sucrose synthase activity in developing wheat endosperms differing in maximum weight. *Plant Physiol.* 82 (1), 7–10. doi:10.1104/pp.82.1.7

de Bustos, A., Rubio, P., Soler, C., García, P., and Jouve, N. (2001). "Marker assisted selection to improve HMW-glutenins in wheat," in Wheat in a global environment. *Developments in plant breeding*. Editors Z. Bedö and L. Láng (Dordrecht: Springer Netherlands), 171–176. doi:10.1007/978-94-017-3674-9_19

De Vries, F. W. T. P., Brunsting, A. H. M., and Van Laar, H. H. (1974). Products, requirements and efficiency of biosynthesis a quantitative approach. *J. Theor. Biol.* 45 (2), 339–377. doi:10.1016/0022-5193(74)90119-2

Devos, K. M., Atkinson, M. D., Chinoy, C. N., Liu, C. J., and Gale, M. D. (1992). RFLP-based genetic map of the homoeologous group 3 chromosomes of wheat and rye. *Theor. Appl. Genet.* 83 (8), 931–939. doi:10.1007/BF00232953

D'hoop, B. B., Keizer, P. L. C., Paulo, M. J., Visser, R. G. F., van Eeuwijk, F. A., and van Eck, H. J. (2014). Identification of agronomically important QTL in tetraploid potato cultivars using a marker–trait association analysis. *Theor. Appl. Genet.* 127 (3), 731–748. doi:10.1007/s00122-013-2254-y

Dias, K. O. D. G., Gezan, S. A., Guimarães, C. T., Parentoni, S. N., Guimarães, P. E. de O., Carneiro, N. P., et al. (2018). Estimating genotype × environment interaction for and genetic correlations among drought tolerance traits in maize via factor analytic multiplicative mixed models. *Crop Sci.* 58 (1), 72–83. doi:10.2135/cropsci2016.07.0566

Dick, J. W., and Youngs, V. L. (1988). Evaluation of durum wheat, semolina, and pasta in the United States. Available at: https://agris.fao.org/agris-search/search.do?recordID=US19890103346.

Distelfeld, A., Uauy, C., Fahima, T., and Dubcovsky, J. (2006). Physical map of the wheat high-grain protein content gene Gpc-B1 and development of a high-throughput molecular marker. *New Phytol.* 169 (4), 753–763. doi:10.1111/j.1469-8137.2005.01627.x

Divashuk, M. G., Klimushina, M. V., and Karlov, G. I. (2011). Molecular genetic characteristics of the Wx-B1e allele from common wheat and applicability of the DNA markers for its identification. *Russ. J. Genet.* 47 (12), 1428–1432. doi:10.1134/S1022795411120039

dos Santos, J. P. R., Pires, L. P. M., de Castro Vasconcellos, R. C., Pereira, G. S., Von Pinho, R. G., and Balestre, M. (2016). Genomic selection to resistance to Stenocarpella maydis in maize lines using DArTseq markers. *BMC Genet.* 17 (1), 86. doi:10.1186/s12863-016-0392-3

D'Ovidio, R., and Anderson, O. D. (1994). PCR analysis to distinguish between alleles of a member of a multigene family correlated with wheat bread-making quality. *Theor. Appl. Genet.* 88 (6–7), 759–763. doi:10.1007/BF01253982

D'Ovidio, R., and Masci, S. (2004). The low-molecular-weight glutenin subunits of wheat gluten. *J. Cereal Sci.* 39 (3), 321–339. doi:10.1016/j.jcs.2003.12.002

Dreisigacker, S., Xiao, Y., Sehgal, D., Guzman, C., He, Z., Xia, X., et al. (2020). SNP markers for low molecular glutenin subunits (LMW-GSs) at the Glu-A3 and Glu-B3 loci in bread wheat. *PLoS ONE* 15, e0233056. doi:10.1371/journal.pone.0233056

Echeverry-Solarte, M., Kumar, A., Kianian, S., Simsek, S., Alamri, M. S., Mantovani, E. E., et al. (2015). New QTL alleles for quality-related traits in spring wheat revealed by RIL population derived from supernumerary × non-supernumerary spikelet genotypes. *Theor. Appl. Genet.* 128 (5), 893–912. doi:10.1007/s00122-015-2478-0

Eckardt, N. A. (2000). Sequencing the rice genome. *Plant Cell.* 12 (11), 2011–2017. doi:10.1105/tpc.12.11.2011

Elouafi, I., Nachit, M. M., Elsaleh, A., Asbati, A., and Mather, D. E. (2000). *QTL-mapping of genomic regions controlling gluten strength in durum (Triticum turgidum L. var. durum)*. Zaragoza, Spain: CIHEAM, 6.

FAO (2021). FAOSTAT statistical database. Available at: https://search.library.wisc.edu/catalog/999890171702121.

Faridi, H., Gaines, C., and Finney, P. (1994). "Soft wheat quality in production of cookies and crackers," in *Wheat: Production, properties and quality*. Editors W. Bushuk and V. F. Rasper (Boston, MA: Springer US), 154–168. doi:10.1007/978-1-4615-2672-8_11

Faris, J. D., Li, W. L., Liu, D. J., Chen, P. D., and Gill, B. S. (1999). Candidate gene analysis of quantitative disease resistance in wheat. *Theor. Appl. Genet.* 98 (2), 219–225. doi:10.1007/s001220051061

Farquharson, R. J. (2006). Production response and input demand in decision making: Nitrogen fertilizer and wheat growers. *Aust. Agribus. Rev.* 15, 5. doi:10.22004/ag.econ.126105

Fredriksson, H., Silverio, J., Andersson, R., Eliasson, A.-C., and Åman, P. (1998). The influence of amylose and amylopectin characteristics on gelatinization and retrogradation properties of different starches. *Carbohydr. Polym.* 35 (3), 119–134. doi:10.1016/S0144-8617(97)00247-6

Galande, A. A., Tiwari, R., Ammiraju, J. S. S., Santra, D. K., Lagu, M. D., Rao, V. S., et al. (2001). Genetic analysis of kernel hardness in bread wheat using PCR-based markers. *Theor. Appl. Genet.* 103, 601–606. doi:10.1007/pl00002915

Gale, K. R. (2005). Diagnostic DNA markers for quality traits in wheat. *J. Cereal Sci.* 41 (2), 181–192. doi:10.1016/j.jcs.2004.09.002

Garg, M., Tanaka, H., Ishikawa, N., Takata, K., Yanaka, M., and Tsujimoto, H. (2009). A novel pair of HMW glutenin subunits from *Aegilops searsii* improves quality of hexaploid wheat. *Cereal Chem. J.* 86 (1), 26–32. doi:10.1094/CCHEM-86-1-0026

Gautam, T., Dhillon, G. S., Saripalli, G., RakhiSingh, V. P., Prasad, P., Kaur, S., et al. (2020). Marker-assisted pyramiding of genes/QTL for grain quality and rust resistance in wheat (*Triticum aestivum* L.). *Mol. Breed.* 40 (5), 49. doi:10.1007/s11032-020-01125-9

Gautam, N., Garg, S., and Yadav, S. (2021). Underutilized finger millet crop for starch extraction, characterization, and utilization in the development of flexible thin film. *J. Food Sci. Technol.* 58 (11), 4411–4419. doi:10.1007/s13197-020-04926-0

Gautier, M.-F., Aleman, M.-E., Guirao, A., Marion, D., and Joudrier, P. (1994). *Triticum aestivum* puroindolines, two basic cystine-rich seed proteins: cDNA sequence analysis and developmental gene expression. *Plant Mol. Biol.* 25 (1), 43–57. doi:10.1007/BF00024197

Gaire, R., Huang, M., Sneller, C., Griffey, C., and Brown-Guedira, G. (2019). Association Analysis of Baking and Milling Quality Traits in an Elite Soft Red Winter Wheat Population. *Crop Sci.* 59 (3), 1085–1094. doi:10.2135/cropsci2018.12.0751

Ghimire, B., Sapkota, S., Bahri, B. A., Martinez-Espinoza, A. D., Buck, J. W., and Mergoum, M. (2020). Fusarium head blight and rust diseases in soft red winter wheat in the southeast United States: State of the art, challenges and future perspective for breeding. *Front. Plant Sci.* 11, 1080. doi:10.3389/fpls.2020.01080

Ghimire, B., Mergoum, M., Martinez-Espinoza, A. D., Sapkota, S., Pradhan, S., Babar, M. A., et al. (2022). Genetics of Fusarium head blight resistance in soft red winter wheat using a genome-wide association study. *Plant Genome* 15, e20222. doi:10.1002/tpg2.20222

Gill, H. S., Halder, J., Zhang, J., Brar, N. K., Rai, T. S., Hall, C., et al. (2021). Multi-trait multi-environment genomic prediction of agronomic traits in advanced breeding lines of winter wheat. *Front. Plant Sci.* 12, 709545. doi:10.3389/fpls.2021.709545

Giroux, M. J., and Morris, C. F. (1997). A glycine to serine change in puroindoline b is associated with wheat grain hardness and low levels of starch-surface friabilin. *Theor. Appl. Genet.* 95 (5–6), 857–864. doi:10.1007/s001220050636

Graybosch, R. A. (1998). Waxy wheats: Origin, properties, and prospects 11Joint contribution of the United States Department of Agriculture, Agricultural Research Service and the Department of Agronomy, University of Nebraska-Lincoln as Journal Series Paper No 12125. Mention of firm names or trade products does not imply that they are endorsed or recommended by the USDA or the University of Nebraska over other firms or products not mentioned. *Trends food Sci. Technol.* 9, 135–142. doi:10.1016/s0924-2244(98)00034-x

Groos, C., Robert, N., Bervas, E., and Charmet, G. (2003). Genetic analysis of grain protein-content, grain yield and thousand-kernel weight in bread wheat. *Theor. Appl. Genet.* 106 (6), 1032–1040. doi:10.1007/s00122-002-1111-1

Gupta, P. K., Rustgi, S., and Mir, R. R. (2008). Array-based high-throughput DNA markers for crop improvement. *Heredity* 101 (1), 5–18. doi:10.1038/hdy.2008.35

Guo, J., Khan, J., Pradhan, S., Shahi, D., Khan, N., Avci, M., et al. (2020a). Multi-trait genomic prediction of yield-related traits in US soft wheat under variable water regimes. *Genes* 11 (11), 1270. doi:10.3390/genes11111270

Guo, Y., Zhang, G., Guo, B., Qu, C., Zhang, M., Kong, F., et al. (2020b). "QTL mapping for quality traits using a high-density genetic map of wheat," in *Plos one*. Editor A. Kumar, 15, e0230601. doi:10.1371/journal.pone.0230601

Gupta, P. K., Balyan, H. S., Chhuneja, P., Jaiswal, J. P., Tamhankar, S., Mishra, V. K., et al. (2022). Pyramiding of genes for grain protein content, grain quality, and rust resistance in eleven Indian bread wheat cultivars: A multi-institutional effort. *Mol. Breed.* 42 (4), 21. doi:10.1007/s11032-022-01277-w

Guzmán, C., and Alvarez, J. B. (2016). Wheat waxy proteins: Polymorphism, molecular characterization and effects on starch properties. *Theor. Appl. Genet.* 129 (1), 1–16. doi:10.1007/s00122-015-2595-9

Guzmán, C., Caballero, L., Martín, L. M., and Alvarez, J. B. (2012a). Waxy genes from spelt wheat: New alleles for modern wheat breeding and new phylogenetic inferences about the origin of this species. *Ann. Bot.* 110 (6), 1161–1171. doi:10.1093/aob/mcs201

Guzmán, C., Caballero, L., Yamamori, M., and Alvarez, J. B. (2012b). Molecular characterization of a new waxy allele with partial expression in spelt wheat. *Planta* 235 (6), 1331–1339. doi:10.1007/s00425-011-1577-7

Guzman, C., Peña, R. J., Singh, R., Autrique, E., Dreisigacker, S., Crossa, J., et al. (2016). Wheat quality improvement at CIMMYT and the use of genomic selection on it. *Appl. Transl. Genomics* 11, 3–8. doi:10.1016/j.atg.2016.10.004

Haile, J. K., N'Diaye, A., Clarke, F., Clarke, J., Knox, R., Rutkoski, J., et al. (2018). Genomic selection for grain yield and quality traits in durum wheat. *Mol. Breed.* 38 (6), 75. doi:10.1007/s11032-018-0818-x

Hanif, M., Gao, F., Liu, J., Wen, W., Zhang, Y., Rasheed, A., et al. (2016). TaTGW6-A1, an ortholog of rice TGW6, is associated with grain weight and yield in bread wheat. *Mol. Breed.* 36, 1. doi:10.1007/s11032-015-0425-z

Hayes, B. J., Panozzo, J., Walker, C. K., Choy, A. L., Kant, S., Wong, D., et al. (2017). Accelerating wheat breeding for end-use quality with multi-trait genomic predictions incorporating near infrared and nuclear magnetic resonance-derived phenotypes. *Theor. Appl. Genet.* 130 (12), 2505–2519. doi:10.1007/s00122-017-2972-7

Heffner, E. L., Jannink, J.-L., Iwata, H., Souza, E., and Sorrells, M. E. (2011). Genomic selection accuracy for grain quality traits in biparental wheat populations. *Crop Sci.* 51 (6), 2597–2606. doi:10.2135/cropsci2011.05.0253

Heffner, E. L., Sorrells, M. E., and Jannink, J.-L. (2009). Genomic selection for crop improvement. *Crop Sci.* 49 (1), 1–12. doi:10.2135/cropsci2008.08.0512

Heslot, N., Akdemir, D., Sorrells, M. E., and Jannink, J.-L. (2014). Integrating environmental covariates and crop modeling into the genomic selection framework to predict genotype by environment interactions. *Theor. Appl. Genet.* 127 (2), 463–480. doi:10.1007/s00122-013-2231-5

Huang, X.-Q., and Brûlé-Babel, A. (2011). Development of simple and co-dominant PCR markers to genotype puroindoline a and b alleles for grain hardness in bread wheat (*Triticum aestivum* L.). *J. Cereal Sci.* 53 (3), 277–284. doi:10.1016/j.jcs.2011.01.008

Huang, X. Q., Cloutier, S., Lycar, L., Radovanovic, N., Humphreys, D. G., Noll, J. S., et al. (2006). Molecular detection of QTLs for agronomic and quality traits in a doubled haploid population derived from two Canadian wheats (*Triticum aestivum* L.). *Theor. Appl. Genet.* 113 (4), 753–766. doi:10.1007/s00122-006-0346-7

Huebner, F. R. (1970). Comparative studies on glutenins from different classes of wheat. *J. Agric. Food Chem.* 18 (2), 256–259. doi:10.1021/jf60168a006

Ibba, M. I., Kiszonas, A. M., and Morris, C. F. (2018). Development of haplotype-specific molecular markers for the low-molecular-weight glutenin subunits. *Mol. Breed.* 38 (5), 68–12. doi:10.1007/s11032-018-0827-9

Ibba, M. I., Kiszonas, A. M., See, D. R., Skinner, D. Z., and Morris, C. F. (2019). Mapping kernel texture in a soft durum (Triticum turgidum subsp. durum) wheat population. *J. Cereal Sci.* 85, 20–26. doi:10.1016/j.jcs.2018.10.006

Jarquín, D., Crossa, J., Lacaze, X., Du Cheyron, P., Daucourt, J., Lorgeou, J., et al. (2014). A reaction norm model for genomic selection using high-dimensional genomic and environmental data. *Theor. Appl. Genet.* 127 (3), 595–607. doi:10.1007/s00122-013-2243-1

Joppa, L. R., Du, C., Hart, G. E., and Hareland, G. A. (1997). Mapping gene(s) for grain protein in tetraploid wheat (*Triticum turgidum* L.) using a population of recombinant inbred chromosome lines. *Crop Sci.* 37 (5), 1586–1589. doi:10.2135/cropsci1997.0011183X003700050030x

Kang, Y. J., Lee, T., Lee, J., Shim, S., Jeong, H., Satyawan, D., et al. (2016). Translational genomics for plant breeding with the genome sequence explosion. *Plant Biotechnol. J.* 14 (4), 1057–1069. doi:10.1111/pbi.12449

Kiszonas, A. M., and Morris, C. F. (2017). Wheat breeding for quality: A historical review. *Cereal Chem. J.* 95 17–34. CCHEM-05-17-0103-FI. doi:10.1094/CCHEM-05-17-0103-FI

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: A review. *Plant Methods* 9 (1), 29. doi:10.1186/1746-4811-9-29

Kristensen, P. S., Jensen, J., Andersen, J. R., Guzmán, C., Orabi, J., and Jahoor, A. (2019). Genomic prediction and genome-wide association studies of flour yield and alveograph quality traits using advanced winter wheat breeding material. *Genes* 10 (9), 669. doi:10.3390/genes10090669

Kuchel, H., Langridge, P., Mosionek, L., Williams, K., and Jefferies, S. P. (2006). The genetic control of milling yield, dough rheology and baking quality of wheat. *Theor. Appl. Genet.* 112 (8), 1487–1495. doi:10.1007/s00122-006-0252-z

Kumar, A., Elias, E. M., Ghavami, F., Xu, X., Jain, S., Manthey, F. A., et al. (2013). A major QTL for gluten strength in durum wheat (Triticum turgidum L. var. durum). *J. Cereal Sci.* 57 (1), 21–29. doi:10.1016/j.jcs.2012.09.006

Kumar, A., Mantovani, E. E., Simsek, S., Jain, S., Elias, E. M., and Mergoum, M. (2019). "Genome wide genetic dissection of wheat quality and yield related traits and their relationship with grain shape and size traits in an elite × non-adapted bread wheat cross," in *Plos one*. Editor A. Zhang, 14, e0221826. doi:10.1371/journal.pone.0221826

Kumar, N., Orenday-Ortiz, J. M., Kiszonas, A. M., Boehm, J. D., and Morris, C. F. (2019). Genetic analysis of a unique 'super soft' kernel texture phenotype in soft white spring wheat. *J. Cereal Sci.* 85, 162–167. doi:10.1016/j.jcs.2018.12.003

Kumar, N., Kulwal, P. L., Balyan, H. S., and Gupta, P. K. (2007). QTL mapping for yield and yield contributing traits in two mapping populations of bread wheat. *Mol. Breed* 19 (2), 163–177. doi:10.1007/s11032-006-9056-8

Kumar, J., Jaiswal, V., Kumar, A., Kumar, N., and Mir, R. R. (2011). Introgression of a major gene for high grain protein content in some Indian bread wheat cultivars. *Field Crops Res.* 123 (3), 226–233. doi:10.1016/j.fcr.2011.05.013

Kumar, R., Mukherjee, S., and Ayele, B. T. (2018). Molecular aspects of sucrose transport and its metabolism to starch during seed development in wheat: A comprehensive review. *Biotechnol. Adv.* 36(4), 954–967. doi:10.1016/j.biotechadv.2018.02.015

Kuspira, J., and Unrau, J. (1957). Genetic analysis of certain characters in common wheat using chromsome substitution lines. *Can. J. Plant Sci.* 37, 300–326. doi:10.4141/cjps57-037

Kuzay, S. (2019). Identification of a candidate gene for a QTL for spikelet number per spike on wheat chromosome arm 7AL by high-resolution genetic mapping. *Theor. Appl. Genet.* 132, 17. doi:10.1007/s00122-019-03382-5

Lau, W. C. P., Rafii, M. Y., Ismail, M. R., Puteh, A., Latif, M. A., and Ramli, A. (2015). Review of functional markers for improving cooking, eating, and the nutritional qualities of rice. *Front. Plant Sci.* 6, 832. doi:10.3389/fpls.2015.00832

Lei, Z. S., Gale, K. R., He, Z. H., Gianibelli, C., Larroque, O., Xia, X. C., et al. (2006). Y-type gene specific markers for enhanced discrimination of high-molecular weight glutenin alleles at the Glu-B1 locus in hexaploid wheat. *J. Cereal Sci.* 43 (1), 94–101. doi:10.1016/j.jcs.2005.08.003

Li, J., Cui, F., Ding, A., Zhao, C., Wang, X., Wang, L., et al. (2012). QTL detection of seven quality traits in wheat using two related recombinant inbred line populations. *Euphytica* 183 (2), 207–226. doi:10.1007/s10681-011-0448-4

Li, Y., Zhou, R., Wang, J., Liao, X., Branlard, G., and Jia, J. (2012). Novel and favorable QTL allele clusters for end-use quality revealed by introgression lines derived from synthetic wheat. *Mol. Breed.* 29 (3), 627–643. doi:10.1007/s11032-011-9578-6

Li, X., Guo, T., Mu, Q., Li, X., and Yu, J. (2018). Genomic and environmental determinants and their interplay underlying phenotypic plasticity. *Proc. Natl. Acad. Sci.* 115 (26), 6679–6684. doi:10.1073/pnas.1718326115

Li, W., and Yang, B. (2017). Translational genomics of grain size regulation in wheat. *Theor. Appl. Genet.* 130 (9), 1765–1771. doi:10.1007/s00122-017-2953-x

Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q., et al. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat. Commun.* 8 (1), 14261. doi:10.1038/ncomms14261

Lillemo, M., and Morris, C. F. (2000). A leucine to proline mutation in puroindoline b is frequently present in hard wheats from Northern Europe. *Theor. Appl. Genet.* 100 (7), 1100–1107. doi:10.1007/s001220051392

Lindgren, A., and Simsek, S. (2016). Evaluation of hard red spring wheat mill stream fractions using solvent retention capacity test: Solvent retention capacity of mill fractions. *J. Food Process. Preserv.* 40 (2), 131–139. doi:10.1111/jfpp.12590

Liu, Y., He, Z., Appels, R., and Xia, X. (2012). Functional markers in wheat: Current status and future prospects. *Theor. Appl. Genet.* 125 (1), 1–10. doi:10.1007/s00122-012-1829-3

Liu, H., Zhang, X., Xu, Y., Ma, F., Zhang, J., Cao, Y., et al. (2020). Identification and validation of quantitative trait loci for kernel traits in common wheat (*Triticum aestivum* L.). *BMC Plant Biol.* 20 (1), 529. doi:10.1186/s12870-020-02661-4

Ma, W., Zhang, W., and Gale, K. R. (2003). Multiplex-PCR typing of high molecular weight glutenin alleles in wheat. *Euphytica* 134, 51–60. doi:10.1023/a:1026191918704

Ma, C., Yang, Y., Li, X., Ge, P., Guo, G., Subburaj, S., et al. (2013). Molecular cloning and characterization of six novel HMW-GS genes from *Aegilops speltoides* and *Aegilops kotschyi*. *McIntosh Ed. Plant Breed.* 132 (3), 284–289. doi:10.1111/pbr.12046

Maich, R. H., Manera, G. A., Londero, W. H., Vanzetti, L. S., Helguera, M., and Di Rienzo, J. A. (2020). Changes in yield determinants and quality traits in wheat after 12 cycles of recurrent selection. *Cereal Res. Commun.* 48 (4), 533–538. doi:10.1007/s42976-020-00053-w

Mann, G., Diffey, S., Cullis, B., Azanza, F., Martin, D., Kelly, A., et al. (2009). Genetic control of wheat quality: Interactions between chromosomal regions determining protein content and composition, dough rheology, and sponge and dough baking properties. *Theor. Appl. Genet.* 118 (8), 1519–1537. doi:10.1007/s00122-009-1000-y

McCartney, C. A., Somers, D. J., Lukow, O., Ames, N., Noll, J., Cloutier, S., et al. (2006). QTL analysis of quality traits in the spring wheat cross RL4452 x 'AC Domain'. *Plant Breed.* 125 (6), 565–575. doi:10.1111/j.1439-0523.2006.01256.x

McLauchlan, A., Ogbonnaya, F. C., Hollingsworth, B., Carter, M., Gale, K. R., Henry, R. J., et al. (2001). Development of robust PCR-based DNA markers for each homoeo-allele of granule-bound starch synthase and their application in wheat breeding programs. *Aust. J. Agric. Res.* 52 (12), 1409. doi:10.1071/AR01036

Mesfin, A., Frohberg, R. C., and Anderson, J. A. (1999). RFLP markers associated with high grain protein from *Triticum turgidum* L. Var. *dicoccoides* introgressed into hard red spring wheat. *Crop Sci.* 39 (2), 508–513. doi:10.2135/cropsci1999.0011183X003900020035x

Michel, S., Ametz, C., Gungor, H., Akgöl, B., Epure, D., Grausgruber, H., et al. (2017). Genomic assisted selection for enhancing line breeding: Merging genomic and phenotypic selection in winter wheat breeding programs with preliminary yield trials. *Theor. Appl. Genet.* 130 (2), 363–376. doi:10.1007/s00122-016-2818-8

Michel, S., Kummer, C., Gallee, M., Hellinger, J., Ametz, C., Akgöl, B., et al. (2018). Improving the baking quality of bread wheat by genomic selection in early generations. *Theor. Appl. Genet.* 131 (2), 477–493. doi:10.1007/s00122-017-2998-x

Mohler, V., Albrecht, T., Mrva, K., Schweizer, G., and Hartl, L. (2014). "Genetic analysis of falling number in three bi-parental common winter wheat populations," in *Plant breeding*. Editor K. Pillen, 133, 448–453. doi:10.1111/pbr.12177

Monari, A. M., Simeone, M. C., Urbano, M., Margiotta, B., and Lafiandra, D. (2005). Molecular characterization of new waxy mutants identified in bread and durum wheat. *Theor. Appl. Genet.* 110 (8), 1481–1489. doi:10.1007/s00122-005-1983-y

Morris, C. F. (2002). Puroindolines: The molecular genetic basis of wheat grain hardness. *Plant Mol. Biol.* 48, 633–647. doi:10.1023/a:1014837431178

Morris, C. F., Simeone, M. C., King, G. E., and Lafiandra, D. (2011). Transfer of soft kernel texture from *Triticum aestivum* to durum wheat, *Triticum turgidum* ssp. durum. *Crop Sci.* 51 (1), 114–122. doi:10.2135/cropsci2010.05.0306

Nadeem, M. A., Nawaz, M. A., Shahid, M. Q., Doğan, Y., Comertpay, G., Yıldız, M., et al. (2018). DNA molecular markers in plant breeding: Current status and recent advancements in genomic selection and genome editing. *Biotechnol. Biotechnol. Equip.* 32 (2), 261–285. doi:10.1080/13102818.2017.1400401

Nakamura, T., Yamamori, M., Hirano, H., and Hidaka, S. (1993). Identification of three Wx proteins in wheat (*Triticum aestivum* L.). *Biochem. Genet.* 31 (1–2), 75–86. doi:10.1007/BF02399821

Nakamura, T., Yamamori, M., Hirano, H., Hidaka, S., and Nagamine, T. (1995). Production of waxy (amylose-free) wheats. *Mol. General Genet. MGG* 248 (3), 253–259. doi:10.1007/BF02191591

Nakamura, T., Vrinten, P., Saito, M., and Konda, M. (2002). Rapid classification of partial waxy wheats using PCR-based markers. *Genome* 45, 1150–1156. doi:10.1139/g02-090

Naraghi, S. M., Simsek, S., Kumar, A., Al Rabbi, S. M. H., Alamri, M. S., Elias, E. M., et al. (2019). Deciphering the genetics of major end-use quality traits in wheat. *G3 Genes|Genomes|Genetics* 9 (5), 1405–1427. doi:10.1534/g3.119.400050

Nevo, E., Korol, A. B., Beiles, A., and Tzion, F. (2002). *Evolution of wild emmer and wheat improvement: Population genetics, genetic resources, and genome organization of wheat's progenitor, Triticum dicoccoides*. Berlin Heidelberg: Springer-Verlag.

Newberry, M., Zwart, A. B., Whan, A., Mieog, J. C., Sun, M., Leyne, E., et al. (2018). Does late maturity alpha-amylase impact wheat baking quality? *Front. Plant Sci.* 9, 1356. doi:10.3389/fpls.2018.01356

Nigro, D., Gadaleta, A., Mangini, G., Colasuonno, P., Marcotuli, I., Giancaspro, A., et al. (2019). Candidate genes and genome-wide association study of grain protein content and protein deviation in durum wheat. *Planta* 249 (4), 1157–1175. doi:10.1007/s00425-018-03075-1

Olmos, S., Distelfeld, A., Chicaiza, O., Schlatter, A. R., Fahima, T., Echenique, V., et al. (2003). Precise mapping of a locus affecting grain protein content in durum wheat. *Theor. Appl. Genet.* 107 (7), 1243–1251. doi:10.1007/s00122-003-1377-y

Osborne, T. B. (1907). *Proteins of wheat kernel*. Washigton: Carneie Institution of Washington. doi:10.5962/bhl.title.26152

Patil, R. M., Oak, M. D., Tamhankar, S. A., and Rao, V. S. (2009). Molecular mapping of QTLs for gluten strength as measured by sedimentation volume and mixograph in durum wheat (Triticum turgidum L. ssp durum). *J. Cereal Sci.* 49 (3), 378–386. doi:10.1016/j.jcs.2009.01.001

Payne, P. I. (1983). "Breeding for protein quantity and protein quality in seed crops," in *Seed proteins. Annual proceedings of the phytochemical society of Europe, held at versailles, september 1981.*, 223–253.

Payne, P. I. (1987). Genetics of wheat storage proteins and the effect of allelic variation on bread-making quality. *Plant Physiol.* 38, 13. doi:10.1146/annurev.pp.38.060187.001041

Payne, P. I., Holt, L. M., and Law, C. N. (1981). Structural and genetical studies on the high-molecular-weight subunits of wheat glutenin: Part 1: Allelic variation in subunits amongst varieties of wheat (*Triticum aestivum*). *TAG. Theor. Appl. Genet. Theor. und angewandte Genet.* 60 (4), 229–236. doi:10.1007/BF02342544

Payne, P. I., Jackson, E. A., and Holt, L. M. (1984). The association between γ-gliadin 45 and gluten strength in durum wheat varieties: A direct causal effect or the result of genetic linkage? *J. Cereal Sci.* 2 (2), 73–81. doi:10.1016/S0733-5210(84)80020-X

Payne, P. I., Holt, L. M., Lawrence, G. J., and Law, C. N. (1982). The genetics of gliadin and glutenin, the major storage proteins of the wheat endosperm. *Qual. Plant Foods Hum. Nutr.* 31 (3), 229–241. doi:10.1007/BF01108632

Payne, P., and Lawrence, G. (1983). Catalogue of alleles for the complex gene loci, Glu-A1, Glu-B1, and Glu-D1 which code for high-molecular-weight subunits of glutenin in hexaploid wheat. *Cereal Res. Commun.* 11, 8.

Peña, R. J. (2007). "Current and future trends of wheat quality needs," in *Wheat production in stressed environments. Developments in plant breeding*. Editors H. T. Buck, J. E. Nisi, and N. Salomón (Dordrecht: Springer Netherlands), 411–424.

Pflieger, S., Lefebvre, V., and Causse, M. (2001). The candidate gene approach in plant genetics: A review. *Mol. Breed.* 7, 275–291. doi:10.1023/a:1011605013259

Pinto, J. G. C. P., Munaro, L. B., Jaenisch, B. R., Nagaoka, A. K., and Lollato, R. P. (2019). Wheat variety response to seed cleaning and treatment after Fusarium head blight infection. *Agrosystems, Geosciences Environ.* 2 (1), 1–8. doi:10.2134/age2019.05.0034

Pin, P. A., and Nilsson, O. (1953). The multifaceted roles of FLOWERING LOCUS T in plant development: FT, a multifunctional protein. *Plant Cell Environ.* 35 (10), 1742–1755. doi:10.1111/j.1365-3040.2012.02558.x

Pirozi, M. R., Margiotta, B., Lafiandra, D., and MacRitchie, F. (2008). Composition of polymeric proteins and bread-making quality of wheat lines with allelic HMW-GS differing in number of cysteines. *J. Cereal Sci.* 48 (1), 117–122. doi:10.1016/j.jcs.2007.08.011

Poland, J. A., Brown, P. J., Sorrells, M. E., and Jannink, J.-L. (2012). "Development of high-density genetic maps for barley and wheat using a novel two-enzyme genotyping-by-sequencing approach," in *PLoS ONE*. Editor T. Yin, 7, e32253. doi:10.1371/journal.pone.0032253

Pradhan, S., Babar, M. A., Bai, G., Khan, J., Shahi, D., Avci, M., et al. (2020). Genetic dissection of heat-responsive physiological traits to improve adaptation and increase yield potential in soft winter wheat. *BMC Genomics* 21 (1), 315. doi:10.1186/s12864-020-6717-7

Prasad, M., Kumar, N., Kulwal, P., Röder, M., Balyan, H., Dhaliwal, H., et al. (2003). QTL analysis for grain protein content using SSR markers and validation studies using NILs in bread wheat. *Theor. Appl. Genet.* 106 (4), 659–667. doi:10.1007/s00122-002-1114-y

Proietti, S., Caarls, L., Coolen, S., Van Pelt, J. A., Van Wees, S. C. M., and Pieterse, C. M. J. (2018). Genome-wide association study reveals novel players in defense hormone crosstalk in *Arabidopsis*: Genome-wide association mapping of hormone cross-talk genes. *Plant Cell. Environ.* 41 (10), 2342–2356. doi:10.1111/pce.13357

Qamar, Z. U., Bansal, U. K., Dong, C. M., Alfred, R. L., Bhave, M., and Bariana, H. S. (2014). Detection of puroindoline (Pina-D1 and Pinb-D1) allelic variation in wheat landraces. *J. Cereal Sci.* 60 (3), 610–616. doi:10.1016/j.jcs.2014.07.007

Rai, A., Mahendru-Singh, A., Raghunandan, K., Kumar, T. P. J., Sharma, P., Ahlawat, A. K., et al. (2019). Marker-assisted transfer of PinaD1a gene to develop soft grain wheat cultivars. *3 Biotech.* 9 (5), 183. doi:10.1007/s13205-019-1717-5

Rapp, M., Lein, V., Lacoudre, F., Lafferty, J., Müller, E., Vida, G., et al. (2018). Simultaneous improvement of grain yield and protein content in durum wheat by different phenotypic indices and genomic selection. *Theor. Appl. Genet.* 131 (6), 1315–1329. doi:10.1007/s00122-018-3080-z

Ravel, C., Praud, S., Murigneux, A., Linossier, L., Dardevet, M., Balfourier, F., et al. (2006). Identification of Glu-B1-1 as a candidate gene for the quantity of high-molecular-weight glutenin in bread wheat (*Triticum aestivum* L.) by means of an association study. *Theor. Appl. Genet.* 112 (4), 738–743. doi:10.1007/s00122-005-0178-x

Ravel, C., Faye, A., Ben-Sadoun, S., Ranoux, M., Dardevet, M., Dupuits, C., et al. (2020). SNP markers for early identification of high molecular weight glutenin subunits (HMW-GSs) in bread wheat. *Theor. Appl. Genet.* 133 (3), 751–770. doi:10.1007/s00122-019-03505-y

Reif, J. C., Gowda, M., Maurer, H. P., Longin, C. F. H., Korzun, V., Ebmeyer, E., et al. (2011). Association mapping for quality traits in soft winter wheat. *Theor. Appl. Genet.* 122 (5), 961–970. doi:10.1007/s00122-010-1502-7

Röder, M. S., Huang, X. -Q., and Börner, A. (2008). Fine mapping of the region on wheat chromosome 7D controlling grain weight. *Funct. Integr. Genomics* 8 (1), 79–86. doi:10.1007/s10142-007-0053-8

Rimbert, H., Darrier, B., Navarro, J., Kitt, J., Choulet, F., Leveugle, M., et al. (2018). "High throughput SNP discovery and genotyping in hexaploid wheat," in *Plos one*. Editor A. Zhang, 13, e0186329. doi:10.1371/journal.pone.0186329

Saito, M., Konda, M., Vrinten, P., Nakamura, K., and Nakamura, T. (2004). Molecular comparison of waxy null alleles in common wheat and identification of a unique null allele. *Theor. Appl. Genet.* 108 (7), 1205–1211. doi:10.1007/s00122-003-1549-9

Saito, M., Vrinten, P., Ishikawa, G., Graybosch, R., and Nakamura, T. (2009). A novel codominant marker for selection of the null Wx-B1 allele in wheat breeding programs. *Mol. Breed.* 23 (2), 209–217. doi:10.1007/s11032-008-9226-y

Salentijn, E. M. J., Pereira, A., Angenent, G. C., van der Linden, C. G., Krens, F., Smulders, M. J. M., et al. (2007). Plant translational genomics: From model species to crops. *Mol. Breed.* 20 (1), 1–13. doi:10.1007/s11032-006-9069-3

Sandhu, K., Aoun, M., Morris, C., and Carter, A. (2021). Genomic selection for end-use quality and processing traits in soft white winter wheat breeding program with machine and deep learning models. *Biology* 10 (7), 689. doi:10.3390/biology10070689

Sandhu, K. S., Patil, S. S., Aoun, M., and Carter, A. H. (2022). Multi-trait multi-environment genomic prediction for end-use quality traits in winter wheat. *Front. Genet.* 13, 831020. doi:10.3389/fgene.2022.831020

Sapkota, S., Hao, Y., Johnson, J., Lopez, B., Bland, D., Chen, Z., et al. (2019). Genetic mapping of a major gene for leaf rust resistance in soft red winter wheat cultivar AGS 2000. *Mol. Breed.* 39 (1), 8. doi:10.1007/s11032-018-0909-8

International Rice Genome Sequencing Project and Sasaki, 2005 International Rice Genome Sequencing Project, and Sasaki, T. (2005). The map-based sequence of the rice genome. *Nature* 436 (7052), 793–800. doi:10.1038/nature03895

Sax, K. (1923). The association of size differences with seed-coat pattern and pigmentation in phaseolus vulgaris. *Genetics* 8 (6), 552–560. doi:10.1093/genetics/8.6.552

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 (5956), 1112–1115. doi:10.1126/science.1178534

Sehgal, D., Mondal, S., Guzman, C., Garcia Barrios, G., Franco, C., Singh, R., et al. (2019). Validation of candidate gene-based markers and identification of novel loci for thousand-grain weight in spring bread wheat. *Front. Plant Sci.* 10, 1189. doi:10.3389/fpls.2019.01189

R. S. Sengar and A. Singh (Editors) (2018). *Eco-friendly agro-biological techniques for enhancing crop productivity* (Singapore: Springer Singapore).

Shan, X., Clayshulte, S. R., Haley, S. D., and Byrne, P. F. (2007). Variation for glutenin and waxy alleles in the US hard winter wheat germplasm. *J. Cereal Sci.* 45 (2), 199–208. doi:10.1016/j.jcs.2006.09.007

Shewry, Peter.R., and Hey, S. J. (2015). The contribution of wheat to human diet and health. *Food Energy Secur.* 4 (3), 178–202. doi:10.1002/fes3.64

Shewry, P. R. (2003). "Improving wheat quality: The role of biotechnology," in *Bread making: Improving quality* (Cambridge: Woodhead Publishing), 168–186.

Shewry, P. R., Tatham, A. S., and Lazzeri, P. (1997). Biotechnology of Wheat Quality. *J. Sci. Food Agric.* 73 (4), 397–406. doi:10.1002/(SICI)1097-0010(199704)73:4<397::AID-JSFA758>3.0.CO;2-Q

Simmonds, N. W. (1995). The relation between yield and protein in cereal grain. *J. Sci. Food Agric.* 67 (3), 309–315. doi:10.1002/jsfa.2740670306

Singh, N. K., and Shepherd, K. W. (1988). Linkage mapping of genes controlling endosperm storage proteins in wheat. *Theor. Appl. Genet.* 75, 14. doi:10.1007/BF00289132

Smith, R. L., Schweder, M. E., and Barnett, R. D. (1994). Identification of glutenin alleles in wheat and triticale using PCR-generated DNA markers. *Crop Sci.* 34 (5), 1373–1378. doi:10.2135/cropsci1994.0011183X003400050042x

Song, G., Jia, M., Chen, K., Kong, X., Khattak, B., Xie, C., et al. (2016). CRISPR/Cas9: A powerful tool for crop genome editing. *Crop J.* 4 (2), 75–82. doi:10.1016/j.cj.2015.12.002

Sourdille, P., Cadalen, T., Guyomarc'h, H., Snape, J., Perretant, M., Charmet, G., et al. (2003). An update of the Courtot x Chinese Spring intervarietal molecular marker linkage map for the QTL detection of agronomic traits in wheat. *Theor. Appl. Genet.* 106 (3), 530–538. doi:10.1007/s00122-002-1044-8

Subedi, M. (2020). *Developing Chenopodium ficifolium as a diploid model system relevant to genetic characterization and improvement of allotetraploid C. quinoa*. Master's thesis. New Hampshire: University of New Hampshire.

Subedi, M., Neff, E., and Davis, T. M. (2021). Developing Chenopodium ficifolium as a potential B genome diploid model system for genetic characterization and improvement of allotetraploid quinoa (Chenopodium quinoa). *BMC Plant Biol.* 21 (1), 490. doi:10.1186/s12870-021-03270-5

Sun, C., Zhang, F., Yan, X., Zhang, X., Dong, Z., Cui, D., et al. (2017). Genome-wide association study for 13 agronomic traits reveals distribution of superior alleles in bread wheat from the Yellow and Huai Valley of China. *Plant Biotechnol. J.* 15 (8), 953–969. doi:10.1111/pbi.12690

Symes, J. K. (1965). The inheritance of grain hardness in wheat as measured by the particle size index. *Aust. J. Agric. Res.* 16 (2), 113. doi:10.1071/ar9650113

Tabbita, F., Pearce, S., and Barneix, A. J (2017). Breeding for increased grain protein and micronutrient content in wheat: Ten years of the GPC-B1 gene. *J. Cereal Sci.* 73, 183–191. doi:10.1016/j.jcs.2017.01.003

Terasawa, Y., Ito, M., Tabiki, T., Nagasawa, K., Hatta, K., and Nishio, Z. (2016). Mapping of a major QTL associated with protein content on chromosome 2B in hard red winter wheat (*Triticum aestivum* L.). *Breed. Sci.* 66 (4), 471–480. doi:10.1270/jsbbs.16026

The Arabidopsis Genome Initiative (2000). Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature* 408 (6814), 796–815. doi:10.1038/35048692

Thoday, J. M. (1961). Location of polygenes. *Nature* 191 (4786), 368–370. doi:10.1038/191368a0

Thomason, W., Griffey, C. A., Parrish, D. J., and Barbeau, W. E. (2019). *Understanding pre-harvest sprouting of wheat*. Blacksburg, Virginia: Virginia cooperative extension.

Tranquilli, G., Lijavetzky, D., Muzzi, G., and Dubcovsky, J. (1999). Genetic and physical characterization of grain texture-related loci in diploid wheat. *Mol. General Genet. MGG* 262 (4–5), 846–850. doi:10.1007/s004380051149

Tsai, H., Howell, T., Nitcher, R., Missirian, V., and Watson, B. (2011). Discovery of Rare Mutations in Populations: TILLING by Sequencing. *Plant Physiol.* 156 (3), 1257–1268. doi:10.1104/pp.110.169748

Tu, M., and Li, Y. (2020). Toward the genetic basis and multiple QTLs of kernel hardness in wheat. *Plants* 9 (12), 1631. doi:10.3390/plants9121631

Turner, A. S., Bradburne, R. P., Fish, L., and Snape, J. W. (2004). New quantitative trait loci influencing grain texture and protein content in bread wheat. *J. Cereal Sci.* 40 (1), 51–60. doi:10.1016/j.jcs.2004.03.001

Uauy, C., Distelfeld, A., Fahima, T., Blechl, A., and Dubcovsky, J. (2006). A NAC gene regulating senescence improves grain protein, zinc, and iron content in wheat. *Science* 314 (5803), 1298–1301. doi:10.1126/science.1133649

Vanzetti, L. S., Pflüger, L., Bainotti, C. T., Jensen, C., and Helguera, M. (2010). Identification of a null allele at the wx-A1 locus in durum wheat (Triticum turgidum L. ssp. durum desf.): A novel wx-A1 null allele in durum wheat. *Plant Breed.* 129 (6), 718–720. doi:10.1111/j.1439-0523.2009.01741.x

Varshney, R., Graner, A., and Sorrells, M. (2005). Genomics-assisted breeding for crop improvement. *Trends Plant Sci.* 10 (12), 621–630. doi:10.1016/j.tplants.2005.10.004

Varshney, R. K., Kudapa, H., Pazhamala, L., Chitikineni, A., Thudi, M., Bohra, A., et al. (2015). Translational genomics in agriculture: Some examples in grain legumes. *Crit. Rev. Plant Sci.* 34 (1–3), 169–194. doi:10.1080/07352689.2014.897909

Vishwakarma, M. K., Mishra, V. K., Gupta, P. K., Yadav, P. S., Kumar, H., and Joshi, A. K. (2014). Introgression of the high grain protein gene Gpc-B1 in an elite wheat variety of Indo-Gangetic Plains through marker assisted backcross breeding. *Curr. Plant Biol.* 1, 60–67. doi:10.1016/j.cpb.2014.09.003

Vrinten, P., Nakamura, T., and Yamamori, M. (1999). Molecular characterization of waxy mutations in wheat. *Mol. General Genet. MGG* 261 (3), 463–471. doi:10.1007/s004380050989

Walkowiak, S., Gao, L., Monat, C., Haberer, G., Kassa, M. T., Brinton, J., et al. (2020). Multiple wheat genomes reveal global variation in modern breeding. *Nature* 588 (7837), 277–283. doi:10.1038/s41586-020-2961-x

Wall, J. S. (1979). The role of wheat proteins in determining baking quality. *Recent Adv. Biochem. Cereals*275–311. Available at: https://ci.nii.ac.jp/naid/10020269251/.

Wang, L. H., Zhao, X. L., He, Z. H., Ma, W., Appels, R., Peña, R. J., et al. (2009). Characterization of low-molecular-weight glutenin subunit Glu-B3 genes and development of STS markers in common wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 118 (3), 525–539. doi:10.1007/s00122-008-0918-9

Wang, L., Li, G., Peña, R. J., Xia, X., and He, Z. (2010). Development of STS markers and establishment of multiplex PCR for Glu-A3 alleles in common wheat (*Triticum aestivum* L.). *J. Cereal Sci.* 51 (3), 305–312. doi:10.1016/j.jcs.2010.01.005

Wang, G., Leonard, J. M., Ross, A. S., Peterson, C. J., Zemetra, R. S., Garland Campbell, K., et al. (2012). Identification of genetic factors controlling kernel hardness and related traits in a recombinant inbred population derived from a soft × 'extra-soft' wheat (*Triticum aestivum* L.) cross. *Theor. Appl. Genet.* 124 (1), 207–221. doi:10.1007/s00122-011-1699-0

Wang, F., Liu, T., Li, Q., An, Y., Xie, C., Sun, X., et al. (2017). QTL mapping of the pasting properties of wheat flour treated by papain digestion: QTL mapping of the pasting properties. *Starch - Stärke* 69 (3–4), 1600077. doi:10.1002/star.201600077

Weegels, P. L., Hamer, R. M., and Schofield, J. D. (1996). Functional Properties of Wheat Glutenin. *J. Cereal Sci.* 23 (1), 1–17. doi:10.1006/jcrs.1996.0001

WHO (2018). *Accelerating the end of hunger and malnutrition global event*. Bangkok, Thailand: WHO. Available at: https://www.who.int/news-room/events/detail/2018/11/29/default-calendar/accelerating-the-end-of-hunger-and-malnutrition-global-event.

Wieser, H. (2007). Chemistry of gluten proteins. *Food Microbiol.* 24 (2), 115–119. doi:10.1016/j.fm.2006.07.004

Xu, Y., and Crouch, J. H. (2008). Marker-assisted selection in plant breeding: From publications to practice. *Crop Sci.* 48 (2), 391–407. doi:10.2135/cropsci2007.04.0191

Yamamori, M., and Guzmán, C. (2013). SNPs and an insertion sequence in five Wx-A1 alleles as factors for variant Wx-A1 protein in wheat. *Euphytica* 192 (3), 325–338. doi:10.1007/s10681-012-0850-6

Yamamori, M., Nakamura, T., Endo, T. R., and Nagamine, T. (1994). Waxy protein deficiency and chromosomal location of coding genes in common wheat. *Theor. Appl. Genet.* 89 (2), 179–184. doi:10.1007/BF00225138

Yang, J., Zhou, Y., Wu, Q., Chen, Y., Zhang, P., Zhang, Y., et al. (2019). Molecular characterization of a novel TaGL3-5A allele and its association with grain length in wheat (*Triticum aestivum* L.). *Theor. Appl. Genet.* 132 (6), 1799–1814. doi:10.1007/s00122-019-03316-1

Yang, Y., Chai, Y., Zhang, X., Lu, S., Zhao, Z., Wei, D., et al. (2020). Multi-locus GWAS of quality traits in bread wheat: Mining more candidate genes and possible regulatory Network. *Front. Plant Sci.* 11, 1091. doi:10.3389/fpls.2020.01091

Zanetti, S., Winzeler, M., Feuillet, C., Keller, B., and Messmer, M. (2001). Genetic analysis of bread-making quality in wheat and spelt. *Plant Breed.* 120 (1), 13–19. doi:10.1046/j.1439-0523.2001.00552.x

Zhang, K., Zhang, Y., Chen, G., and Tian, J. (2009). Genetic analysis of grain yield and leaf chlorophyll content in common wheat. *Cereal Res. Commun.* 37 (4), 499–511. doi:10.1556/CRC.37.2009.4.3

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

Zhang, L., Liu, P., Wu, J., Qiao, L., Zhao, G., Jia, J., et al. (2020). Identification of a novel ERF gene, TaERF8, associated with plant height and yield in wheat. *BMC Plant Biol.* 20 (1), 263. doi:10.1186/s12870-020-02473-6

Zhang, X., Karim, H., Feng, X., Lan, J., Tang, H., Guzmán, C., et al. (2022). A single base change at exon of Wx-A1 caused gene inactivation and starch properties modified in a wheat EMS mutant line. *J. Sci. Food Agric.* 102 (5), 2012–2022. doi:10.1002/jsfa.11540

Zhang-Biehn, S., Fritz, A. K., Zhang, G., Evers, B., Regan, R., and Poland, J. (2021). Accelerating wheat breeding for end-use quality through association mapping and multivariate genomic prediction. *Plant Genome* 14 (3), e20164. doi:10.1002/tpg2.20164

Zhang, X., Liu, G., Zhang, L., Xia, C., and Zhao, T. (2018). Fine Mapping of a Novel Heading Date Gene, TaHdm605, in Hexaploid Wheat. *Front. Plant Sci.* 9 , 1059. doi:10.3389/fpls.2018.01059

Zhao, X. L., Xia, X. C., He, Z. H., Lei, Z. S., Appels, R., Yang, Y., et al. (2007). Novel DNA variations to characterize low molecular weight glutenin Glu-D3 genes and develop STS markers in common wheat. *Theor. Appl. Genet.* 114 (3), 451–460. doi:10.1007/s00122-006-0445-5

Zhu, T., Jackson, D. S., Wehling, R. L., and Geera, B. (2008). Comparison of amylose determination methods and the development of a dual wavelength iodine binding technique. *Cereal Chem.* 85 (1), 51–58. doi:10.1094/CCHEM-85-1-0051

Zi, Y., Ding, J., Song, J., Humphreys, G., Peng, Y., Li, C., et al. (2018). Grain yield, starch content and activities of key enzymes of waxy and non-waxy wheat (*Triticum aestivum* L.). *Sci. Rep.* 8 (1), 4548. doi:10.1038/s41598-018-22587-0

frontiers | Frontiers in Genetics

TYPE Review
PUBLISHED 09 January 2023
DOI 10.3389/fgene.2022.1007552

OPEN ACCESS

EDITED BY
Reyazul Rouf Mir,
Sher-e-Kashmir University of Agricultural Sciences and Technology, India

REVIEWED BY
Manjusha Verma,
National Bureau of Plant Genetic Resources (ICAR), India
Paul Kiprotich Kimurto,
Egerton University, Kenya
Vanisri Satturu,
Professor Jayashankar Telangana State Agricultural University, India

*CORRESPONDENCE
J. Lydia Pramitha,
lydiapramitha@gmail.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal *Frontiers in Genetics*

RECEIVED 30 July 2022
ACCEPTED 07 December 2022
PUBLISHED 09 January 2023

CITATION
Lydia Pramitha J, Ganesan J, Francis N, Rajasekharan R and Thinakaran J (2023), Revitalization of small millets for nutritional and food security by advanced genetics and genomics approaches.
Front. Genet. 13:1007552.
doi: 10.3389/fgene.2022.1007552

Revitalization of small millets for nutritional and food security by advanced genetics and genomics approaches

J. Lydia Pramitha[1*†], Jeeva Ganesan[2†], Neethu Francis[1†], Ravikesavan Rajasekharan[2†] and Jenita Thinakaran[1†]

[1]Karunya Institute of Technology and Sciences, Coimbatore, India, [2]Tamil Nadu Agricultural University, Coimbatore, India

Small millets, also known as nutri-cereals, are smart foods that are expected to dominate food industries and diets to achieve nutritional security. Nutri-cereals are climate resilient and nutritious. Small millet-based foods are becoming popular in markets and are preferred for patients with celiac and diabetes. These crops once ruled as food and fodder but were pushed out of mainstream cultivation with shifts in dietary habits to staple crops during the green revolution. Nevertheless, small millets are rich in micronutrients and essential amino acids for regulatory activities. Hence, international and national organizations have recently aimed to restore these lost crops for their desirable traits. The major goal in reviving these crops is to boost the immune system of the upcoming generations to tackle emerging pandemics and disease infestations in crops. Earlier periods of civilization consumed these crops, which had a greater significance in ethnobotanical values. Along with nutrition, these crops also possess therapeutic traits and have shown vast medicinal use in tribal communities for the treatment of diseases like cancer, cardiovascular disease, and gastrointestinal issues. This review highlights the significance of small millets, their values in cultural heritage, and their prospects. Furthermore, this review dissects the nutritional and therapeutic traits of small millets for developing sustainable diets in near future.

KEYWORDS

small millets, nutrition, therapeutic traits, ethnobotany, multi-omics

1 Introduction

Recent changes in food habits related to multi-grains have generated enormous interest in food and nutritional security. People have now started to focus on the nutritional uptake of small millets and growing children are provided processed foods that meet the daily nutritional requirement (Muthamilarasan and Prasad, 2021). Small millets, otherwise known as "wonder cereals", possess numerous health benefits and thrive in harsh conditions (Barretto et al., 2021). Increasing demand for nutrition has paved the way to revive these crops, which once ruled nations. The 11 small millet

species—finger millet, foxtail millet, barnyard millet, little millet, proso millet, Kodo millet, fonio millet, teff, brown top millet, Job's tears, and guinea millet—are commonly known as the lost crops of the world. They are so-called 'small' owing to their smaller seed size. Small millets are rich in micronutrients, essential amino acids, and vitamin B complex, which are very rare in our staple diets. Phytochemical studies in small millets have demonstrated their higher antioxidant contents and lower glycemic indexes compared to other food crops (Vetriventhan et al., 2020).

Several programs for diet schedules are now popularizing the use and consumption of small millets in various forms. This has resulted in the food processing industries producing value-added products like flakes, noodles, biscuits, cookies, batter, flour, bread, and rice analogs. Rice analogs are one of the most popular developments that increase the palatable value of small millets by heat extrusion (Zhang et al., 2020). Hence, the demand for small millets is increasing in food industries. In addition, small millets are free of gluten, and it is highly suitable for patients with celiac disease (Deshpande et al., 2015). Despite nutrition, traditional practices have emphasized their medicinal values. These crops were previously not only foods but were also used to treat diseases including cancer and snakebites. Hence, analyzing the therapeutic traits of small millets paves a new way for smarter food in near future. Their uptake in the regular diet would provide their caloric, nutritional, medicinal, and fiber rich-properties for a sustainable life (Banerjee and Maitra, 2020).

In response to emerging pandemics in the changing climatic scenario, crop scientists have focused on boosting the human immune system with natural supplements through food diversification. These efforts have necessitated the acceleration of small millet breeding programs (Upadhyaya and Vetriventhan, 2018). Small millet genomics has recently started analyzing agronomic traits. Dissecting the genes underlying nutritional and therapeutic traits by applying advanced omics will reveal novel metabolic pathways in cereals for biofortification (Kumar, and Kumar, 2015). Next-generation and third-generation crop sequencing and breeding have demonstrated the potential of genetic databases and tools to analyze the genomes of small millets (Aggarwal et al., 2022). Hence, the future relies on developing smart crops with climate resilience, higher nutrition, and therapeutic traits. Moreover, conserving the traditional landraces of small millets, their cultural heritage, and ethnobotanical values by national and international organization programs to protect the rights of tribal farmers is needed to preserve novel alleles for the future (Saha et al., 2016). The present review focuses on the significance of small millets, breeding techniques, and advanced approaches for improving their productivity by enhancing their desirable features in a sustainable cropping system.

2 Ethnobotanical significance and cultural heritage of small millets

Small millets have a profound significance in our cultural heritage and until now have played significant roles in temple festivals in tribal regions. These activities are preserved as traditional knowledge in the regulations put forth by PPVFRA, 2001 (Satyarthi et al., 2018). These traditions demonstrate that these grains were recognized by our ancestors for their nutraceutical and therapeutic values (Figure 1). One common practice in small grains was their presentation as a wedding gift to the bridegroom. The amount of grains gifted was treated as a prestige; the grains were also cooked, especially during puberty and childbirth celebrations (Rawat et al., 2021). Small millets are rich in folic acid; therefore, they were treated as a special entity for women to overcome anemic disorders. In Africa, fonio millet, commonly known as "hungry rice," has similar importance. These grains are predominantly used to prepare couscous known as wusu-wusu. Fonio is best used in the preparation of beverages called *Tchapulo*, which is rich in minerals. Finger millet has a similar value in the processing of beer, and its malted products are often used in African tribal communities (Hitu et al., 1997). *Arake*, a distilled liquor, is prepared in Ethiopia with finger millet flour. Furthermore, people residing in Sudan predominantly consume a hot porridge of finger millet with banana or sugar juice, which is a staple dish in tribal zones. A sour bread known as *injera* is made from teff and is used in spicy stews by well-off tribal individuals. Teff has a unique role in Africa after fonio millet. Due to its cold tolerance in higher altitudes, it is popularly known as love grass. The novel features of these lost crops are also being conserved. The major morphotypes in fonio millet are Yoro, Ipordapia, Ipordawoun, Ipoagoa, and Iporni are conserved by communities including the *Hausa* and *pagans* in west African regions (Dansi et al., 2010; National Research Council, 1996).

Regarding the conservation of cultural heritage in India, the Malayali in Eastern Ghats continue to cultivate and conserve small millet landraces (Newmaster et al., 2013). The landraces of little, foxtail, and proso millets are conserved by the Kolli Hill tribes; the characteristics of these millets suggest the presence of novel alleles for future breeding programs (Venkatesan et al., 2015; Ragupathy et al., 2016). Tracking the records of the utilization of small millets in Chhattisgarh, India has revealed the use of Kodo millet straw in mud to make termite-resistant walls. The farmers of this region also used Kodo millet straw as fertilizers in onion fields to increase productivity. Another traditional practice was the application of finger millet husk in green gram and pigeon pea products to protect these grains from pest infestation during storage. The pot makers in Northern India also are using Kodo millet straw when baking pots. Moreover, the leaves of Kodo millet possess lecithin and are used for the treatment of snakebites, stomach disorders, and joint pain. In cattle, Kodo millet straw had a significant impact on

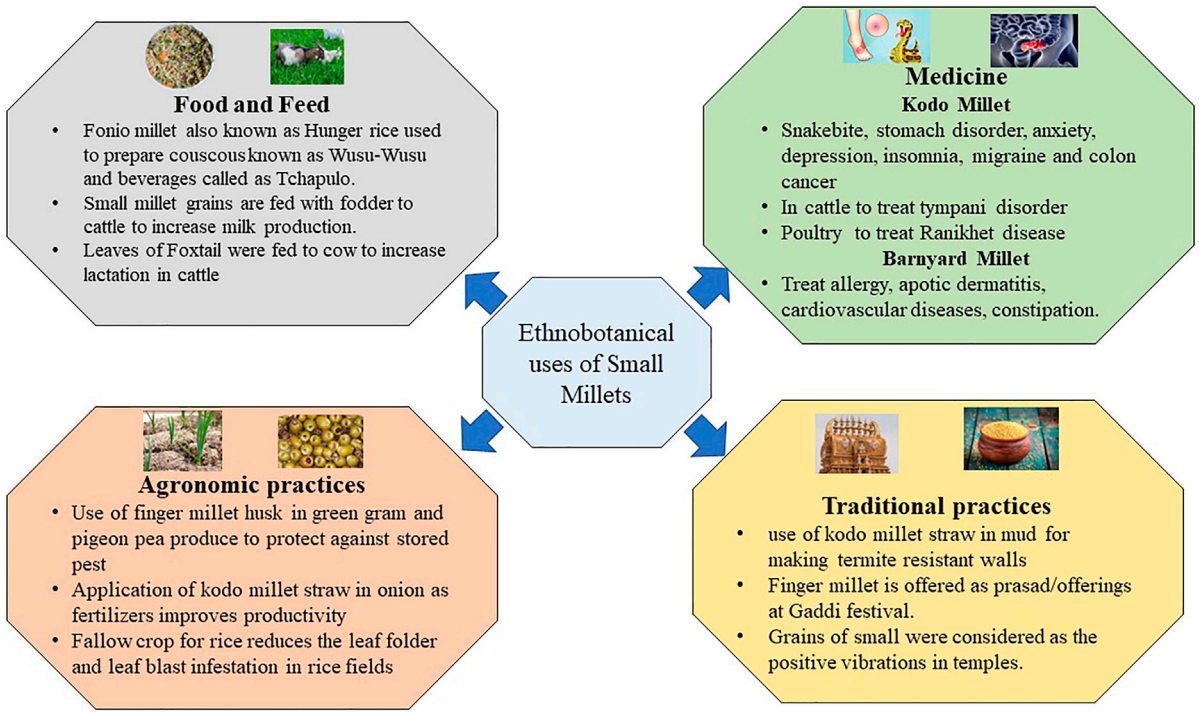


FIGURE 1
Ethnobotanical values of small millets.

treating tympani disorder. Additionally, the older grains of Kodo millet (3–4 years) were used to cure Ranikhet disease in poultry. In Africa and India, Kodo millet was a fallow crop after rice; in other rice fields, Kodo millet straw is usually spread in the fields to protect against leaf folder and blast (Rawat et al., 2021).

The small millet grains were also previously mixed with fodder to increase milk production in cows. Several recent agro-start-ups in India for cattle feed also practice this technique to enhance milk production in rural dairy farms (Bhat et al., 2018). In traditional practice, finger millet is often a prasad/offering in the Gaddi festival. This is believed to enhance the fruiting of non-flowering mango and tamarind trees. Furthermore, thick pastes of finger millet flour are used to treat fire burns, and these grains are considered to offer positive energy in temples (Sahu and Sharma, 2013).

In addition, barnyard millet is used to treat allergies, atopic dermatitis, cardiovascular diseases, constipation, and blood-related disorders. Moreover, Kodo millet is preferred for overcoming anxiety, depression, insomnia, migraine, and colon cancer. Foxtail millet is used to treat chicken pox, heart attack, fever, cholera, and gastric problems. The leaves of foxtail millet are also fed to cattle to increase lactation. These practices underscore the nutritional and therapeutic value of small millets in our heritage. The genetics underlying these traits could be explored and used to develop sustainable diets (Satyarthi et al., 2018).

3 Nutritional importance and quality parameters of small millets

Small millets are being considered for their paramount importance in nutritional aspects not provided by other staple crops. These species have profound nutritional benefits, especially their micronutrient and protein profiles (Pramitha Lydia et al., 2021). The 11 small millets species each represent a unique beneficial feature for dietary bowls (Figure 2). Considering their acceptability for consumption, they are highly preferred for patients with diabetes and celiac disease owing to their gluten-free and higher fiber content (Table 1). Our ancestors were involved in laborious work and these foods predominantly helped them to sustain healthy lives. Due to the comfortability provided by major staples in later generations, the importance of these crops was almost forgotten, and they lost their commercial value. However, some traditional farmers still preserve this cultural heritage through festivals and traditional practices. Arising pandemics and emerging unprecedented weather have caused us to reassess

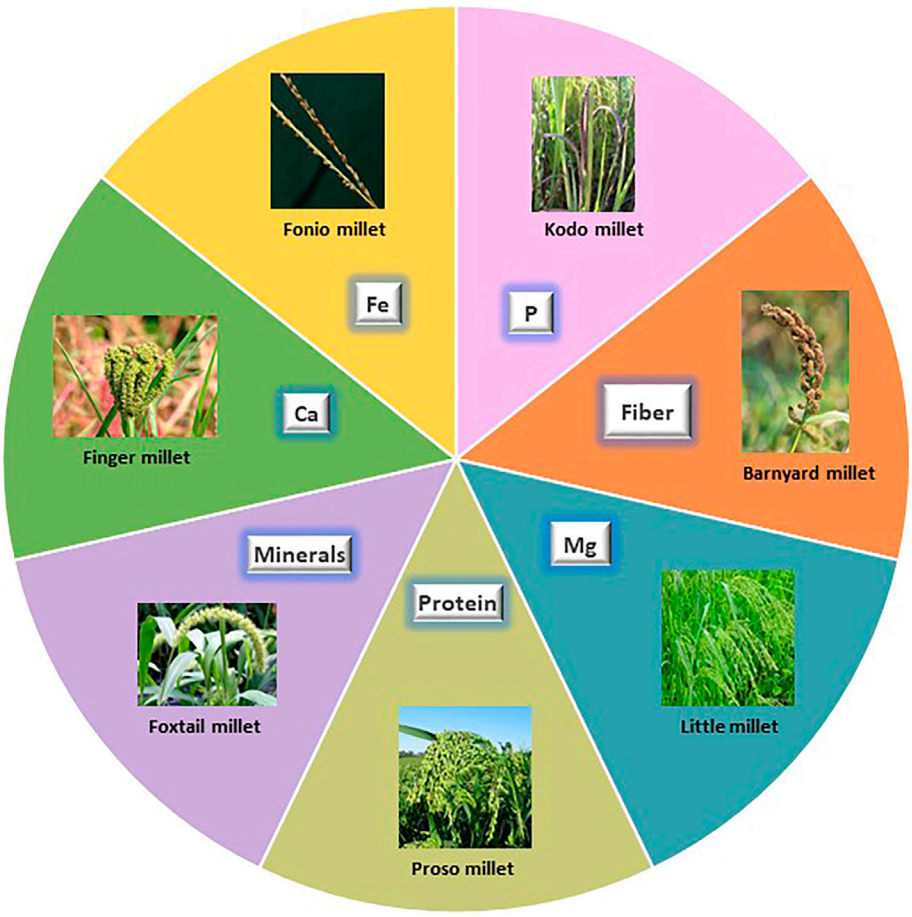


FIGURE 2
Nutritional uniqueness of small millets. Each sphere depicts the specialty of small millets in food.

TABLE 1 Nutritional composition of small millets.

Crop	CHO (g)	Protein (g)	Fat (g)	Crude fiber (g)	Ash (g)	Ca (mg)	P (mg)	Fe (mg)	Zn (mg)	Mg (mg)
Finger millet	72.60	7.70	1.50	3.60	2.70	344.00	250.00	6.30	2.30	130.00
Foxtail millet	60.90	12.30	4.30	8.00	3.30	31.00	290.00	2.80	2.40	81.00
Proso millet	60.90	12.50	1.10	5.20	1.90	14.00	206.00	0.80	1.40	81.00
Barnyard millet	65.50	6.20	4.40	13.60	2.20	20.00	280.00	8.00	3.00	137.00
Little millet	65.60	10.40	1.30	7.60	1.30	16.10	220.00	1.30	3.70	133.00
Kodo millet	66.20	8.90	2.60	5.20	1.70	15.30	188.00	2.30	0.70	147.00
Rice	78.20	7.90	0.50	1.00	11.19	7.50	160.00	0.70	1.30	64.00
Wheat	64.00	10.60	1.50	2.00	0.94	41.00	306.00	5.30	2.70	138.00
Maize	18.70	3.27	1.35	2.00	4.83	10.00	89.00	0.52	0.46	37.00

Sources: Dey et al., 2022; Gopalan et al., 2009; Geervani and Eggum, 1989; Longvah et al., 2017.

the nutritional benefits of these crops (Muthamilarasan and Prasad, 2021).

Among the eleven species of small millets, finger millet and foxtail millet have larger areas of cultivation. Finger millet is known for its calcium, phosphorous, iron, and zinc content. It also has a higher caloric value due to its carbohydrate content. Foxtail millet is mostly preferred for its protein content, which is higher than wheat. It is also rich in minerals like phosphorous, iron, and zinc. Barnyard millet is known for its low glycemic index and high phosphorous and magnesium content. Among small millets, barnyard millet has the highest dietary fiber and niacin content (Pramitha Lydia et al., 2021). Kodo millet has the highest phosphorous content and radical scavenging activity owing to its high phenol content. Consuming Kodo millet reduces the risk of cardiovascular dysfunction (Rao Dayakar et al., 2017; Deepak Taggelli and Thakur, 2018). Little millet is known for its magnesium, phosphorous, and protein content. The unique feature of little millet is that it is rich in PUFA and flavonoids (Indirani and Devasena, 2021). Proso millet has a high protein profile following foxtail millet and also has higher magnesium and niacin content. Proso millet is also rich in essential amino acids like lysine, leucine, isoleucine, and methionine, unlike other major cereals (Pramitha Lydia et al., 2021). Fonio millet is preferably found in African regions and has high iron, dietary fiber, crude protein, flavonoid, GABA, and riboflavin content. It is highly nutritious owing to its high methionine and cysteine content. Teff is yet another wonder millet in Africa, which features the smallest grain and highest calcium and iron content. Teff is especially rich in lysine, which is deficient in all grains, and has comparatively higher protein and starch content. The other minor millets include brown top millet, Job's tears, and guinea millet. These are also rich in phosphorous, iron, zinc, and vitamin B (Durairaj et al., 2019).

The demand for small millets is increasing in upcoming markets due to value-added products. Thus, their quality aspects should be considered for commercial production. Small millets are small-grained cereals and usually have a lower milling recovery than other grains. Therefore, uniformly sized, unbroken seeds are of concern in small millet production (Aggarwal et al., 2022). The colors vary from black to light yellow-colored grains, with consumers predominantly preferring light yellow- to brown-colored seeds for flour and cooking purposes (Durgad et al., 2021). Yellow-colored grains are also reportedly more aromatic compared to black grains (Aggarwal et al., 2022). Small millets are also a great source of resistant starch. The highest resistant starch values have been reported in Kodo millet, little millet, barnyard millet, and foxtail millet. Thus, these grains have nutraceutical value. As the grain yields of cereals deteriorate owing to harsh climates, small millets can be used to manufacture and process resistant starch in various industries (Kaimal et al., 2021). Rice analogs from small millet flours are also currently produced by heat extrusion and gelatinization. These rice granules are comparatively more desirable than normal rice varieties (Zhang et al., 2020).

Other quality parameters in small millet breeding include their fodder and forage value. Small millets are also used as fodder crops due to their higher biomass. Kodo millet, little millet, and proso millet are the most preferred animal feeds due to their higher palatability and crude protein content. Despite these factors, rancidity affects the storage of these flours in homes (Vetriventhan et al., 2020; Sruthi and Rao, 2021). Although the grains of small millets have a longer storage value, their flours are subject to rancidity; thus, efforts in advanced breeding techniques are needed to enhance the shelf life of small millet flours (He et al., 2015; Hariprasanna et al., 2014).

4 Breeding objectives and prospects in overcoming constraints

Small millets are among the less demanding crops in cultivation. They require minimum input from irrigation to fertilizers and pesticides. They can also thrive in harsh conditions, making them a reliable smart crop in the future (Durairaj et al., 2019). The major breeding objectives in small millets concern cultivation and post-harvest techniques. The initial objective predominantly focuses on higher yields. Yield is a complex trait that depends on numerous variables. The major factors causing yield loss in small millets are lodging and shattering. The increase of tillers in small millets with higher biomass at maturity results in lodging. Also, some crops are prone to shattering at harvest, which can be overcome with a stronger culm diameter while breeding for higher yield. Second, small millets have a non-preference in ideotype due to the spined hairy shoots and leaves. This causes challenges in managing field activities such as weeding and pesticide spraying. Hence, breeding for reduced bristles, spines in shoots, and leaves helps in proper crop management (Ganapathy and Patil, 2017).

Small millets are also less prone to diseases. Some of the major diseases infecting them are shown in Table 2. These infestations are observed only in endemic regions and conducive conditions caused by poor cultivation practices (Nagaraja and Das, 2016). Therefore, the crop protection schedule and its inputs in these crops are predominantly minimal (Ravikesavan et al., 2022). Finally, as a major objective in the post-harvest technique, smaller grain size causes poor milling recovery; thus, breeding for larger seeds would help minimize post-harvest losses (Table 3). Future efforts in small millet breeding should focus on improving the color, nutritional profile, fodder yield, flour quality, and reduced antinutritional traits (Vetriventhan and Upadhyaya, 2019).

Although small millets have significant potential, the number of research programs remains low compared to other crops. This is due to the constraints involved in the smaller inflorescence and spikelets. These features restrict the possibility of attaining desirable recombinants. Recent approaches in hot water

TABLE 2 Major diseases affecting small millets.

S. No.	Millet	Disease	Causal agent
1	Finger millet	Blast	*Pyricularia grisea*
		Rust	*Puccinia substriata*
		Smut	*Melanopsichium eleusinis*
		Downy mildew	*Sclerophthora macrospora*
		Seedling and leaf blight	*Drechlsera nodulosum*
		Cercospora leaf spot	*Cercospora eleusinis*
		Banded Blight	*Rhizoctonia solanii*
		Wilt or foot rot	*Sclerotium rolfsii*
		Bacterial leaf spot	*Xanthomonas eleusinae*
		Bacterial leaf blight	*Xanthomonas axonopodis pv. corocane*
		Bacterial leaf stripe	*Pseudomonas eleusinae*
		Ragi severe mosaic	*Sugarcane mosaic virus*
		Ragi mottle streak	*Ragi mottle streak virus*
		Ragi streak	*Maize streak virus*
	Foxtail millet	Blast	*Pyricularia setaria*
		Rust	*Uromyces-setaria italicae*
		Smut	*Ustilago crameri*
		Downy mildew	*Sclerospora graminicola*
		Udbatta	*Ephelis* sp.
		Bacterial leaf blight	*Pseudomonas avenae*
	Kodo millet	Head smut	*Sorosporium paspali*
		Rust	*Puccinia substriata*
		Udbatta	*Ephelis* sp.
		Kodua poisoning	*Aspergillus flavus/tamarii*
	Barnyard millet	Head Smut	*Ustilago crus-galli*
		Kernel smut	*Ustilago panici-frumentacei*
		Bacterial leaf blight	*Pseudomonas avenae*
	Proso millet	Head smut	*Sphacelotheca destruens*
		Bacterial leaf blight	*Pseudomonas avenae*
	Little millet	Rust	*Uromyces linearis*
	Teff	Rust	*Uromyces eragrostidis*
		Damping off	*Helminthospoirum poae*

Source: Nagaraja and Das, 2016.

emasculation have resulted in the successful development of recombinants. Advances in mutation breeding with MutMap+ and genotyping facilities have allowed the dissection of novel alleles in small millets. Hence, there is a need for increased focus on small millet genetics and genomics using advanced omics approaches (Muthamilarasan and Prasad, 2015).

TABLE 3 Specific breeding objectives for small millet resources.

S. No.	Small millet	Progenitor	Origin	Race	Breeding objective
1	Finger Millet (*Eleusine corocana* L.)	*Eleusine indica* x *Eleusine floccifolia/E. tristachya*	West Africa	1. Elongata a. Laxa	i) Higher yield, productivity, and non-lodging efficiency ii) Reduced tannins iii) Higher calcium and micronutrients iii) Bold seeded with higher finger length iv) Tolerant to drought, salinity, and diseases-like blast
				b. Reclusa	
				c. Sparsa	
				2. Plana a. Seriata	
				b. Confundere	
				c. Grandigluma	
				3. Compacta	
2	Barnyard millet				
	1. Japanese barnyard millet *(Echinochloa esculenta)*	*Echinochloa crus-galli*	Japan	1. Utilis	i) Higher yield and productivity ii) Bold seeded type iii) Higher quality and micronutrient profile
				2. Intermedia	
	2. Indian barnyard millet (*Echinochloa frumentacea*)	*Echinochloa colona*	India and Africa	1. Laxa	
				2. Robusta	
				3. Intermedia	
				4. Stolonifera	
3	Foxtail millet (*Setaria italica L.*)	*Setaria viridis*	China	1. Indica	i) Breeding for higher yield and productivity
				2. Maxima	ii) Higher protein, micronutrient status and palatability for fodder
				3. Moharia	iii) Breeding cultivars to resistance to pest and diseases
					iv) Varieties suitable for intercropping and early maturing
4	Little millet (*Panicum sumatrense* Roth. ex. Roem. and Schultz)	*Panicum psilopidium*	India and South East China	1. Nana	i) Bold seeded and higher productivity
				2. Robusta	ii) Higher protein and crude protein for enhanced fodder quality
					iii) Non-shattering and non-lodging type
					iv) early maturing, high fiber, and tolerant to shoot fly
5	Proso millet (*Panicum miliaceaum*)	*Panicum capillare* and *Panicum repens*	China	1. Miliaceaum	i) Suitable for intercropping and early maturing
				2. Patentissimum	ii) Higher productivity for grain and fodder
				3. Contactum	iii) Non-shattering and tolerant to biotic and abiotic stresses

(Continued on following page)

TABLE 3 (*Continued*) Specific breeding objectives for small millet resources.

S. No.	Small millet	Progenitor	Origin	Race	Breeding objective
				4. Compactum	
				5. Ovatum	
6	Kodo millet (*Paspalum scrobiculatum* L.)	*Paspalum sanguinale*	Africa	1. Regularis	i) Early maturing, non-lodging, and photosynthetic efficiency ii) Higher yield, productivity, bold seeded, and higher fodder quality iii) Cultivars for sole and mixed cropping iv) Tolerance to smut and other diseases
				2. Irregularis	
				3. Variabilis	

(Ravikesavan et al., 2022; Vetriventhan and Upadhyaya, 2018).

5 Progress of crop improvement in small millets

Initiatives on small millet improvement began in the early 1950s, when conventional breeding including pureline selection and pedigree breeding were ruling the varietal releases in all crops (Paroda and Mal, 1989). All small millets were highly self-pollinated and local collections by breeders were evaluated for line development. Varieties in small millets including CO 6 and CO (7) thenai in foxtail; CO (PV) 5 in proso millet; CO (samai) 4 in little millet; and CO 9, CO 13, CO 14, and Paiyur 2 in finger millet are examples of successful lines developed from recombination breeding. These varieties were released by Tamil Nadu Agricultural University by using standardized protocols for emasculation like hot water treatment and approaches (Ravikesavan et al., 2022). The millets were dominant in earlier civilizations; however, due to the palatability of major staples like rice and wheat, the small millets lost their presence among the population. Additionally, constraints like shattering, low yield, poor milling recovery, flour rancidity, and cooking time in small millets were major factors that drove these crops out of the staple source in the pipeline. Hence, breeder efforts were lost in the middle of the century after the green revolution (Vetriventhan et al., 2020).

Although these crops lost their economic value, few farmers relied entirely on this cropping pattern due to their adaptability in arid and harsh environments. Thus, these crops were always preserved as a cultural heritage among tribal populations and traditional farmers across generations. Overcoming the difficulties in crossing, mutation breeding in small millets later began to arise for varietal development. Mutation breeding in small millets started in the 1970s based on EMS and gamma rays. The frequency spectrum of macromutations with EMS and gamma in proso millet was studied by Ganapathy et al. (2008) and Bhave et al. (2016). Eswari et al. (2014) and Ambavane et al. (2015) also evaluated the effect of mutagenic frequencies with EMS and gamma rays in the previously released cultivars of finger millet. Foxtail mutation breeding experimented with different doses of EMS, DES, and gamma which also tended to affix the LD_{50} for isolating desirable mutants (Anittha and Mullainathan, 2018). Ramesh et al. (2019)assess the efficiencies of EMS and gamma rays in the CO (Kv) 2 strain of barnyard millet. Sood et al. (2020) evaluated the efficacies of different concentrations of sodium azide and gamma-ray treatments in barnyard millets. Thus, the LD_{50} values in small millets from earlier studies provided an understanding of the frequency of mutations induced in these genomes. Dosages of 500–600 Gy for finger millet and barnyard millet; 0.3%–0.45% EMS for finger millet, barnyard millet, and Kodo millet; 0.2% EMS for teff; 0.1 M nitrous acid in fonio millet for 4 h; and 0.03% nitrosoguanidine for finger millet induced desirable variations (Jency et al., 2016; Vetriventhan et al., 2020; Francis et al., 2022). Simultaneously, the genetic resources in small millet germplasm collections were also conserved as core collections in international research institutes. Due to efforts by Upadhyaya et al. (2016), the core collections from ICRISAT for little millet (460 accessions), foxtail millet (155 accessions), barnyard millet (89 accessions), and Kodo millet (75 accessions) were preserved and utilized. Other genetic reservoirs for small millets include NBPGR, the Ethiopian Biodiversity Institute, ICARDA, USDA, the N.I. Vavilov Russian Scientific Research Institute, the Ustymivika Experimental Research Station, and the Kenya Agricultural Livestock Research Station (Upadhyaya et al., 2008; Joshi et al., 2021).

The breeding of small millets has only recently accelerated due to increases in patients with diabetes and children with nutritional absorption issues. Moreover, the publication of the genome sequence of foxtail millet brought attention to the importance of these millets. Hence, relatively fewer molecular studies are published in small millets compared to the major staples. Still, almost all kinds of markers, from RFLP to AFLP, RAPD, EST-SSR, SSR, and SNPs have been utilized for marker-assisted selection. The first linkage map in small millets was developed with RFLP in foxtail millet; later, SSR and SNPs were used to map the QTLs in a high-density linkage map (Wang et al., 1998; Wang et al., 2017). Linkage maps with SNP markers were also later developed in proso millet by Rajput et al. (2016). A high-density linkage map with SNPs for finger millet was recently developed by Pendergast et al. (2022). Subsequently, molecular

TABLE 4 Major QTLs detected in small millets.

S. No.	Crop	Trait	QTL	Chromosome	References
1	Foxtail millet	Tiller number	*Tdll*	5	Doust et al., 2004
		Axillary branching	*SQUAX 1*	6	
		Number of spikelets	*SPK*	9	
		Bristle number/primary branch	*BR*	8	
		Panicle length	*qPL 6.1*	6	Fang et al. (2016)
		Straw weight/plant	*qSWP 1.1*	1	
		Node number of the main stem	*qNMS 1.1*	1	
		Tiller number	*qTN 5*	5	Zhang et al. (2019)
		Plant height	*4115*	5	
		First main internode diameter	*qFMID 9.1*	9	Wang et al. (2017)
		Second main internode diameter	*qSM1D9.1*	9	
		Heading date	*qDTH2*	2	Yoshitsu et al. (2017)
			qDTH7	7	
		Plant height	*qPII5-2*	5	Wang et al. (2017)
		Panicle diameter	*qPD5-2*	5	
		Panicle weight	*qPW5-1*	5	
		Pericarp color	*qPC7-2*	7	
		Grain weight per plant	*qGWP3.3*	3	Liu et al. (2020)
		Straw weight per plant	*qSWP7.4*	7	
		Straw weight per plant	*qSWP9.1*	9	
		Plant height	*qPH1.1, qPH1.2*	1	He et al. (2021)
			qPH3.2	3	
			qPH5.1, qPH5.2	5	
			qPH6.3	6	
			qPH8	8	
			qPH9.1, qPH9.2, qPH9.4, qPH9.5	9	
		Blast	*QLB-czas1*	1	Tian et al. (2021)
			QLB-czas2	2	
			QLB-czas8	8	
		Panicle length	*qPL9.5*	9	Zhi et al. (2021)
		Panicle diameter	*qPD9.2*	9	
2	Finger millet	Blast	*UGEP24*	3B	Babu (2014)
			UGEP81	6B	
		Neck blast	*UGEP18*	1B	
		Flowering date	-	1B	Pendergast IV et al. (2022)
		Plant height	-	3B	

(Continued on following page)

TABLE 4 (*Continued*) Major QTLs detected in small millets.

S. No.	Crop	Trait	QTL	Chromosome	References
		Panicle number	-	3B	
		Leaf blast severity	-	1B	
		Leaf sheath color	-	4B	Fukunaga et al. (2022a)
		Stb	-	2	
		DTH	-	2	
3	Proso millet	Lodging	*QLh.unac-lg5*	LG-5	Rajput et al. (2016)
		Peduncle length	*QPdl.unac-lg4*	LG-45	
		Grain shattering	*QGs.unac-lg5*	LG-5	
		100-grain weight	*QGw.unac-lg40*	LG-40	
		Grains per panicle	*QGpp.unac-lg4*	LG-1	
		Plant height	PH1.1	1	Liu C et al. (2022)
4	Teff	Days to heading	-	LG 17	Chanyalew et al. (2005)
		Days to maturity	-	LG 4	
		Grain yield	-	LG 6	Zeid et al. (2011)
		Panicle weight	-	LG 7	
		Panicle seed weight	-	LG 7	
		Panicle length	-	LG 7	
		Lodging index	-	LG 9	

markers such as AFLP, RAPD, CAP, miRNA, EST, ISSR, SRAP, DEG, and SNP in proso millet were implied for genotyping the diversity, while EST, RAPD, and AFLP were used to analyze the calcium dynamics in finger millet (Habiyaremye et al., 2017). Fukunago et al. (2002b) and Desai et al. (2021) studied chloroplast and mitochondrial diversity in foxtail millet and barnyard millet, respectively. Among all molecular markers, SSRs were predominantly utilized in small millets and were also used in comparative genomics to analyze their lineages. Hence, a separate marker database for SSRs in foxtail was developed by Bonthala et al. (2014).

QTL mapping and trait mapping in small millets are in the initial stages; until now, QTL mapping for agronomic traits was performed (Table 4). The QTLs for traits like tiller number, branching, number of spikelets, bristles, panicle length and weight, plant height, grain weight/plant, and pericarp color have been identified in foxtail millet (Aggarwal et al., 2022). In finger millet, Babu (2014) and Ramakrishnan et al. (2016) identified QTLs for blast, neck blast, and leaf blast. Furthermore, in finger millet, the QTLs for P efficiency (Ramakrishnan et al., 2017), agronomic traits (Sharma et al., 2018), and biochemical traits like APX, CAT, GR, and SOD and POD activities (David et al., 2021) have also been studied. The populations mostly used for mapping in small millets included F2–F6 and RILs. The use of other mapping populations like NILs and double haploids remains to be explored (Vetriventhan et al., 2020; Zhang et al., 2021). Association mapping in foxtail millet and proso millet were predominantly studied with SSR and SNPs to determine the LD values for the diverse collections (Aggarwal et al., 2022; Boukail et al., 2021).

Among small millets, foxtail millet is the most exploited crop in genomics analysis (Table 5). Thus, foxtail millet is a model crop to understand genes in other crops, owing to its amenable genome and crop duration. Hence, advanced omics, genome editing, and double haploid techniques have been standardized and utilized from foxtail millet to analyze the genomics in other small millets (Aggarwal et al., 2022). The first sequence of foxtail millet was completed by Bennetzen et al. (2012) and Zhang et al. (2012) with two different cultivars. An updated sequence of this millet was later released by Ni et al. (2017). Successively, the finger millet whole genome sequence was published by Hittlamani et al. (2017) and Hatakeyama et al. (2018). The genome of sequences of barnyard millet and proso millet were also completed and published in China by Guo et al. (2017) and Zou et al. (2019), respectively. The genome sequences of teff (VanBuren et al., 2020) and fonio millet were recently completed (Abrouk et al., 2020). In addition, the protocols for genetic transformation with agrobacterium and biolistic approaches in

TABLE 5 Major omics approaches conducted in small millets.

S. No.	Crop	Omics approach	Key finding	Reference
1	Foxtail millet	Genomics	Reference Genome sequence of "Yugu1" genotype. Genome size was predicted to be 510 Mb	Bennetzen et al. (2012)
		Genomics	Draft genome sequence of "Zhang gu" genotype. Estimated genome size was 423 Mb	Zhang et al. (2012)
		Transcriptomics and metabolomics	Phenylpropanoid, flavonoid, and lignin biosynthesis pathways, and lysophospholipids plays an important role in salinity tolerance	Pan et al. (2020)
		Metabolomics	Metabolite profiling of seeds. Region-specific differential expression of 20 metabolites	Yang et al., 2021
		Transcriptomics and metabolomics	Identified vital genes involved in carotenoid metabolism and regulation	Li et al. (2022)
2	Finger millet	Genomics	Draft genome sequence of "ML-365" genotype	Hittalmani et al. (2017)
		Genomics	Draft genome sequence of "PR202" genotype. Estimated genome size was 1.5 GB	
				Hatakeyama et al. (2018)
		Integrated transcriptomics and proteomics	Differentially expressed genes (DEGs) associated with corresponding differentially expressed proteins (DEPs) involved in drought tolerance were identified. They were enhanced in hydrolase activity, glycosyl bond formation, oxidoreductase activity, carbohydrate binding and biosynthesis of unsaturated fatty acids	Li W et al. (2021)
		Multiomics	Fifteen putative genes involved in Fe and Zn homeostasis pathways were identified	
			Function annotation of the genes identified high similarity with rice, wheat, maize, barley, and foxtail millet.	Chandra et al. (2020)
		Transcriptomics and metabolomics	Supplementation of silica to osmotic-stressed plants reprograms fatty acid biosynthesis to impart tolerance	Mundada et al. (2021)
3	Proso millet	Genomics	Draft genome sequence of 'Longmi4' genotype. Estimated genome size was 887.8 Mb	Shi et al. (2019)
		Genomics	Draft genome sequence of a landrace from Northern China (accession number 00000390). Estimated genome size was 923 Mb	Zou et al. (2019)
		Transcriptomics	Low N-tolerant genotype had higher efficiency of N uptake and utilization and photosynthesis of leaves	Liu T et al. (2022)
		Transcriptomics	Faster recovery of the photosynthetic genes, ROS scavenging system transcriptional responses, and regulation of jasmonic acid signal transduction pathway played a critical role in the drought-tolerant genotype 'Neimi 5'	Zhang et al. (2019)
		Metabolomics	Grain metabolite and phenolic profiling in Korean varieties	Kim et al. (2013)
		Metabolomics	No significant differences were observed between metabolites of millets grown conventionally and organically	Liang et al. (2018)
		Metabolomics	Hundred metabolites were differentially expressed between colored grain and white grain type. Identified metabolites involved responsible for antioxidant and quality characters in the seed	Li J et al. (2021)
4	Barnyard millet	Genomics	Draft genome sequence of 'STB08'	Guo L et al. (2017)
		Transcriptomics	Differentially expressed genes and regulatory mechanisms involved in drought tolerance and Fe and Zn accumulation	Jayakodi et al. (2019)
		Metabolomics	Metabolomic profiling of contrasting genotypes of Fe content identified differentially expressed metabolites at different stages of spike development	Padhiyar et al. (2022)
5	Teff	Genomics	Draft genome sequence of "Tsedey (DZ-Cr-37)"	Cannarozzi et al. (2014)
		Transcriptomics	Annotation of 3800 transcripts	Cannarozzi et al. (2014)

(Continued on following page)

TABLE 5 (*Continued*) Major omics approaches conducted in small millets.

S. No.	Crop	Omics approach	Key finding	Reference
		Comparative genomics	Synteny with sorghum	Cannarozzi et al. (2014)
6	Little millet	Metabolomics	Twenty-five metabolic pathways were impacted by drought stress	Dhawale (2022)
		Transcriptomics	SSR marker identification and functional annotation of unigenes	Desai et al. (2021)
		Transcriptomics	Molecular mechanism and DEGs under drought and saline stress	Das et al. (2020)
7	Kodo millet	Transcriptomics	Identification of differentially expressed genes and molecular pathways involved in dehydration stress	Suresh et al. (2022)

finger millet, agrobacterium and biolistic methods in foxtail millet, biolistic methods in barnyard millet, agrobacterium-mediated approaches in teff and fonio, and agrobacterium-mediated methods in Kodo millet are now available (Vetriventhan et al., 2020; Bhatt et al., 2021). These advances are expected to enhance the prospects of understanding the underlying genomics of small millets. However, there remain research gaps in small millets regarding key nutritional and medicinal traits that could be used to revitalize our diets.

6 Genomics of climate resilience in small millets

The average global temperature is predicted to rise by 4–5°C by the end of the twenty-first century, which will adversely affect the growth of cereals crops like wheat and rice. Small millet crops, apart from being nutritionally superior to major cereal crops, are inherently tolerant to abiotic stresses like drought, high temperature, cold, poor soil fertility, and salinity (Goron and Raizada, 2015; Singh et al., 2021). They are physiologically sustainable under adverse environmental conditions owing to their excellent and efficient water and nitrogen use (Baltensperger, 2002; Saha et al., 2016). These characteristics highlight small millets as ideal smart crops for cultivation in the context of climate change. They can also aid in creating climate resilience in major cereals like wheat and rice (Bandyopadhyay et al., 2017). Hence, understanding the genetic and molecular mechanisms controlling stress tolerance in millets is critical. Modern and conventional breeding should work together to expedite the dissection and utilization of these complex mechanisms.

The advent of next-generation sequencing technologies (NGS) has revolutionized the field of genomics and created significant molecular information. Although the benefits of this technology were delayed in small millets due to limited funding and research, the genome sequences of foxtail millet, finger millet, proso millet, teff, Japanese barnyard millet, and white fonio are now available (Zou et al., 2019; Vetriventhan et al., 2020; Wang et al., 2021). Whole-genome sequencing and annotation of the ML-365 finger millet genome revealed TFs (transcription factors) and genes related to drought tolerance and the C_4 photosynthetic pathway. A total of 2,866 drought-responsive genes were associated with WRKY, MYB, MYC, ZFHD, NAC, ABF, AREB, GRF, and NF-Y transcription factors (Hittalmani et al., 2017). In proso millet, 180 *NAC* TFs were identified and differentially expressed under various drought treatments. The expression levels of 31% of *PmNAC* (Proso millet *NAC*) genes were upregulated in the roots, indicating the critical role of root characteristics in drought tolerance (Shan et al., 2020). Numerous enzymes, including signal recognition particle receptor, farnesyl pyrophosphate synthase, calcineurin B-like interacting protein kinase 31, serine–threonine protein phosphatase 2A, and others were activated by drought stress in finger millet. Several housekeeping and basal regulatory genes were also activated by drought. Novel drought-associated genes reported in the crop included pentatricopeptide repeat proteins and tetratricopeptide repeat proteins (Parvathi et al., 2019).

Salinity-tolerant finger millet genotypes show upregulation of many genes governing cell growth and differentiation. In the salinity-tolerant strain Trichy 1, genes involved in flavonoid biosynthesis were selectively down-regulated (Rahman et al., 2014). Transcriptomic and metabolomic analyses of contrasting genotypes for salinity tolerance showed that lysophospholipids, phenylpropanoid, flavonoid, and lignin biosynthesis pathways were crucial in the salinity tolerance of foxtail millet. The tolerant Yugu2 strain showed increased antioxidant enzyme activity and non-enzymatic antioxidant content (Pan et al., 2020; Puranik et al., 2011). The genome-wide gene expression profile in foxtail millet showed that most of the crop's drought-responsive genes were associated with photosynthesis, signal transduction, and TFs (Qin et al., 2020). Apart from the well-known stress tolerance-associated genes like peroxidases and glutamine synthetase, leaf tissue-specific expression of ricin-B lectin-like was reported in little millet under salt and drought stress. Alcohol dehydrogenase was also upregulated in both root and leaf tissues of the crop under both stresses (Das et al., 2020). Photosynthesis was a key factor in the ability of Indian barnyard millet to adapt to dry conditions, as

shown in a comparative transcriptome analysis between the plant and its wild relative, barnyard grass (*Echinochloa crus-galli*) (Jayakodi et al., 2019). Thus, the unique expression for climate resilience in small millets suggests novel alleles for use in stress breeding programs (Saha et al., 2016).

7 Innovative breeding methods to improve nutrition and stress tolerance

Innovative breeding methods can be used to speed breeding progress and reveal genes responsible for stress tolerance and nutritional quality improvement in small millets. Additionally, this information can be used to enhance these traits in susceptible major cereal crops. Whole-genome sequencing has revolutionized the breeding and biotechnological approaches in crop improvement. The genome sequences of small millets have generated genomic resources that can be used to identify the gene position, selection, and introgression of desirable traits into other varieties, species, or crops. By considering the genome sequence of a single individual as a reference genome, we may miss many key variations present in crop genetic diversity. This reference bias can be overcome by the concept of a pangenome, wherein several individuals of the species are sequenced to better represent the crop's diversity. This method can distinguish the diversity as core (shared among different genotypes of the species) and variable (may be absent in some individuals) genome (Tettelin et al., 2005; Morgante et al., 2007). In small millets, where genome sequence information is limited, the sequences of panicoid species can be combined using pangenome and comparative genome approaches to improve breeding efficiency (Bennetzen and Freeling, 1997; Sharma et al., 2018; Tay Fernandez et al., 2022). Moreover, trait-based pangenomes can also be devised for stress tolerance and nutritional quality (Tay Fernandez et al., 2022).

Comparative genomics between C_4 and C_3 grasses provides a better understanding of their evolutionary paths to accelerate the research goal of improving the photosynthetic efficiency and drought tolerance in common staple crops (Chapman et al., 2022). High-throughput sequencing platforms have made the identification of single-nucleotide polymorphisms (SNPs) easier. Currently, SNPs are the main genetic variations or markers used in marker-assisted selection (Lata and Prasad, 2013). Genome-wide association studies (GWAS) and nested association mapping (NAM) can be used to establish marker–trait associations and detect powerful QTLs. NAM uses multiple mapping populations, has higher recombination events, and can detect tightly associated QTLs compared to traditional QTL mapping. Though this approach is widely used to identify candidate genes in major cereal crops, it has not yet been applied in underutilized crops like small millets (McMullen et al., 2009; Bouchet et al., 2017).

Mutation breeding has evolved from conventional methods to MutMap populations and targeted mutagenesis. MutMap and MutMap + are genome sequence-based mutation breeding approaches that facilitate rapid gene identification and isolation. The MutMap method has been used to map the dwarfing gene *D3* on chromosome 8 of the foxtail millet genome. The dwarf mutant also shows reduced drought tolerance, indicating critical trait associations (Fan et al., 2017). Mutmap-based cloning was used to establish the regulatory role of the WRKY transcription factor in panicle and seed development. MutMap + does not require crossing between mutant and wild type and is a desirable approach in millet crops in which crossing is tedious (Fekih et al., 2013). New plant breeding techniques like CRISPR-Cas9, TALENs, ZFNs, mega nucleases, oligonucleotide-directed mutagenesis (ODM), cis-genesis, transgenesis, and RNA-dependent methylation (RDM) can also be explored for small millet crop improvement (Schaart et al., 2016). These techniques can be used to decrease anti-nutritional content like phytates, polyphenols, and tannins. The higher levels of proteases and amylose inhibitors that impede millet digestibility can also be reduced using site-targeted modifications (Vinoth and Ravindhran, 2017). In foxtail millet, bioinformatic database information was used to identify target genes for waterlogging tolerance in designing guide RNA for clustered regularly interspaced short palindromic repeat (CRISPR)-aided activation of tolerance-associated transcripts (Abdulla et al., 2021).

Various approaches exist for the transgenic biofortification of finger millet and other crops. Increasing the calcium (Ca^{2+}) storage capacity in edible parts by modifying the expression of transporter proteins, including the overexpression of channel proteins and Ca^{2+}-binding proteins to increase calcium accumulation and mutagenesis-induced alterations of calcium content (Wyatt et al., 2002; Morris et al., 2008; Conn and Gilliham, 2010; Sharma et al., 2018). Other modern techniques like double haploid production and reverse breeding in millet crops may be useful for fixing genetically desirable cross combinations (Gis et al., 2019). Creating such platforms in the model crop foxtail millet may provide opportunities to dissect complex C4 pathways in millets and exploit their role in stress tolerance (Jacquier et al., 2020). Cheng et al. (2021) reported haploid embryo induction through the CRISPR-Cas9-mediated knockout of *SiMTL* in foxtail, which is an ortholog of MATRILINEAL/NOT-LIKE-DAD/PHOSPHOLIPASE A (*MTL/NLD/ZmPLA*) in maize used for haploid induction (Gilles et al., 2017; Kleter et al., 2019).

Transcriptomics is yet another method that has generated enormous genomic information even in crops in which genome sequence data are not available. Transcriptome sequencing (RNA-Seq) reveals genome-wide information on functional genes, differential expression of these genes, and their regulatory mechanisms. This method is widely used as it is

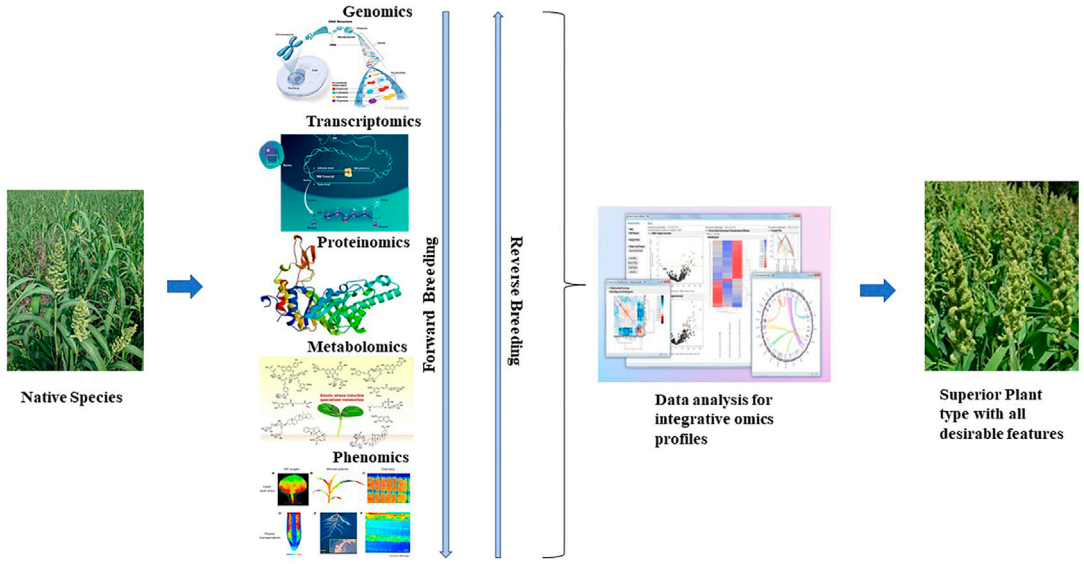


FIGURE 3
Flow of multi-omics in rapid breeding programs.

cheaper than building a genome assembly and reveals the transcriptional activity based on the time and location of the observation. It can also provide insights into various metabolic pathways of the crop (Guo et al., 2021). Metabolomics, another promising omics method, was used to profile metabolites associated with drought stress response in little millet (Dhawale. 2022). The integrative use of transcriptomics and metabolomics was applied to identify key pathways involved in salinity tolerance in foxtail millet (Pan et al., 2020). Metabolomics analysis revealed the environmental and geographical influences on the bioactive nutrient profile of foxtail millet and the key biochemical pathways (Yang et al., 2021). Thus, combining advanced techniques to understand the genetic regulation of crops provides new gateways for the development of desirable plant types.

8 Advanced omics approaches in small millets for trait-specific breeding

Recent advances in next-generation sequencing have applied multiple omics. Integrated omics is a rapid platform to hasten the selection process in plant breeding programs (Figure 3). Genomics plays roles from locating QTLs to developing trait-associated markers. Structural genomics has been utilized to locate the trait of interest in genetic and physical maps. Furthermore, positional cloning of QTLs by advanced QTL-seq approach has also allowed the identification of the exact location and sequence of candidate genes, paving the way for the design of allele-specific markers (Fan et al., 2017). Inclusively, genotype by sequencing has more precisely revealed the population structure of diverse collections, which has enabled the detection of novel alleles. Allele-specific markers are now used in plant breeding programs and are more reliable than the conventional markers used previously (Scossa et al., 2021).

Functional genomics also plays a major role in identifying the functions of genes relative to the trait of interest. Mutagenomics, epigenomics, and pangenome analysis are involved in depicting gene functions. Metabolomics, proteomics, trancriptomics, ionomics, and phenomics are also used in plants to dissect functional roles in gene expression. However, to understand the complete biological process of an organism, multi-omics is essential and also plays a role in the rapid breeding of crops (Kumar et al., 2021). Previously, metabolomics and transcriptomics were associated with plant phenotypes of major traits (Table 5). Recently, owing to advances in omics technology, an integrated approach has been implemented in crops. GWAS combined with metabolite profiling has been used to detect biochemical and genetic processes in major staples (Luo, 2015; Matsuda et al., 2015). Combined HRPF and GWAS was used to inform biomass and yield in rice (Yang et al., 2015); photosynthesis and growth rates in maize were analyzed by the combination of metabolomics and ionomics (Guo R et al., 2017); metabolomics, ionomics, and genomics involving QTL mapping were applied to evaluate iron and zinc in rice (Pinson et al., 2015); integrated mutagenomics and phenomics were used to understand increased lycopene in tomato (Li et al., 2018); mQTL analysis and mGWAS have been applied in maize and rice (Li K. et al., 2019; Tiozon, 2021), and combined

transcriptome and metabolome study have been used to understand defense response in wheat *Trichoderma harzianum* strain T22 (Coppola et al., 2019) to detect key metabolites and candidate genes involved in the pathways for grain yield, secondary metabolites, and disease resistance (Yang et al., 2021).

Finally, functional genomics plays a major role in hybrid performance. The prediction of genotype performance by multi-omics is more precise than the conventional techniques (Yang et al., 2021). Future goals of integrated omics include the development of a proteome and metabolome atlas depicting the major genes in metabolic pathways, which also ensures the development of advanced databases for integrated big data in molecular breeding (Peng et al., 2020). These future innovations could frame a new way to design crops for favorable traits by *de novo* domestication and allow the rapid development of crop ideotypes (Choi, 2019).

9 Therapeutic value and its incorporation through genomics-assisted breeding techniques

Small millets are known for their nutritional value; however, understanding their significance in ethnobotany reveals their therapeutic value. The wonder grains of small millets are remembered as medicines for cures; thus, studying these values offers a greater scope for feeding people with a healthier diet. Each small millet has a potential medicinal value; for example, barnyard millet is known for its fiber content and thus is highly recommended to patients with heart disease and diabetes mellitus. Barnyard millet also has a higher lysine and iso-leucine for blood formation and lipid metabolism in organisms. The components of barnyard millet have the highest linoleic acid, which is a major explanation for the plant's antioxidant and immunological activities. The phytochemicals in barnyard millet include alkaloids, tannins, terpenoids, and flavonoids (Renganathan et al., 2020). Moreover, finger millet is rich in tryptophan; therefore, it is prescribed for decreasing appetite in weight loss programs. Finger millet also includes threonine, which obstructs the formation of fat in the liver and regulates cholesterol levels. Due to its higher calcium, potassium, and iron levels, this millet is also suggested for use in growing children, lactating women, and anemic people (Verma et al., 2018). Ragi is rich in phytates, tannins, and trypsin inhibitors. Proso millet is predominantly known for its highest protein content and is prescribed as a potential therapeutic agent for diabetes mellitus II. This is combined with folic acid, vitamin B6, phenolic acid, ferulic acid, chlorogenic acid, syringic acid, and caffeic acid. Proso millet has shown effectiveness in curing heart disease and preventing breast cancer (Goron and Raizada, 2015). However, foxtail millet has higher linoleic acid, tocopherol, phenol, and flavonoid levels, which enhance its therapeutic value. Additionally, it supplies copper for regulated metabolic reactions. Moreover, small millets contain flavonoids and folic acid and are often used to enhance immune function. Finally, fonio millet is a good source of GABA and folic acid and has the highest iron content (Satyarthi et al., 2018).

These small millet components have been associated with health. Research on their therapeutic value and their stability in the grains is needed and could be a foundation for improving the medicinal values of major staples. The biofortification of grains with micronutrients is reaching new heights and incorporating flavonoids, folates, and GABA in mainstream foods to strengthen our immune response to overcome new emerging diseases. Although rice landraces are used as sources of folate, GABA, and flavonoids, their expression in processed grains is a concern. Whole brown rice and pigmented rice are highly nutritious; however, low percentages of people have access to these landraces. Hence, small millets could be an efficient source for enhancing the medicinal value of food, and its genomics may reveal the expression of these traits in processed foods (Tiozon 2021).

Thus, studies are needed for the identification of QTLs for these traits from mapping populations; association studies in small millets to identify factors related to micronutrients, therapeutic traits, and yield; metabolite genome-wide association studies; functional characterization of key genes; dissection of stable donors; and candidate gene analysis (Xia et al., 2021). Correlation and association studies on medicinal and yield values would help improve multiple traits in small millets (Tiozon 2021). Linkage maps for agronomic traits are currently available in foxtail millet and must be further improved to include micronutrient and medicinal traits. Since small millets are climate-resilient, achieving overall improvements in yields, micronutrients, and therapeutic profiles will be rewarding for smart foods (Aggarwal et al., 2022).

10 Road map for enhancing nutritional security with small millets

The current status in crop production sheds light on enhancing small millets productivity, which is influenced by many factors. The priority is to increase the areas under small millet cultivation. Until now, the cultivation of small millets was restricted to marginal and rainfed farmers who use these as cover crops to conserve arable lands in harsh conditions (Merga, 2018). These farmers also have poor access to high-quality seeds for sowing and the availability of pure small millet varietal seeds must be meticulously formulated from seed production plots. Second, increasing the cultivation area requires accelerating the market demand. Market consumers must be familiarized with diversifying their food habits for a sustainable life (Ravikesavan et al., 2022). Food diversification results in crop diversification and the potential to reintroduce lost and underutilized crops to

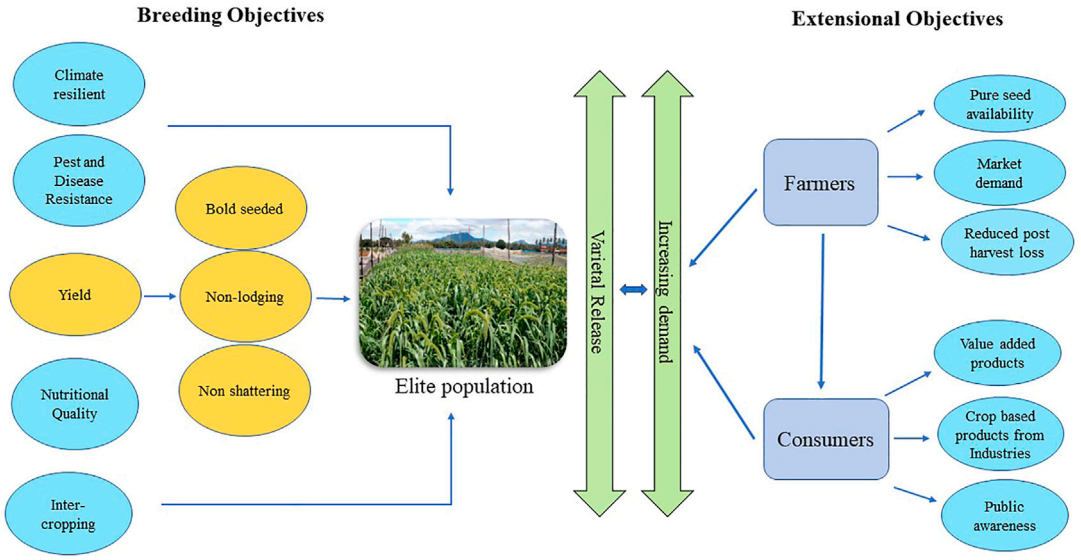


FIGURE 4
Road map for revitalizing small millets.

mainstream cultivation. Recent cropping systems have also been altered to include small millets and they are now incorporated as an intercrop with legumes for a higher profit (Maitra, 2020). This positively impacts our ecosystem for a balanced chain and may also minimize genetic erosion (Deshpande et al., 2015).

Moreover, public communities have recently focused on physical fitness, with trainers in urban regions counseling their trainees on food habits. These efforts involve nutritional supplements and popular advertisements for artificial health supplements have appeared in diet schedules. These schedules could soon be replaced with small millet-based supplements. Small millets are rich in folic acid, flavonoids, terpenoids, resistant starch, and other phytochemicals that regulate cellular metabolic activities. Hence, in-depth research on the therapeutic value of small millets and their enrichment could replace artificial supplements in fitness programs (Durairaj et al., 2019; Lloyd and Kossmann, 2021).

Another major stream in small millets that requires consideration is the incorporation of breeding strategies to yield desirable traits. Weedy features in small millets need improvement through selection. The characteristic features like shattering, lodging, small seeds, spined shoots, bristles, and awns must also be considered in the development of crop ideotypes in small millet breeding (Banerjee and Maitra, 2020). Further breeding objectives for analyzing their nutritional stability, bioavailability in processed foods, multigrain products, phytochemical expression, uniform maturity, fertilizer responsiveness, and biotic stress tolerance must be emphasized to provide higher yield and productivity (Figure 4). Thus, small millets could be an alternative source of nutritional and therapeutic traits in our regular diet (Adekunle et al., 2018; Upadhyaya and Vetriventhan, 2018). Food processing industries are also now moving toward small millets for the extraction of resistant starch, which opens a new market for small millet demand. In response, countries have increased their focus on small millets (Kaimal et al., 2021).

Small millets are arising as major cereals in countries like Ethiopia where other crops strive hard to feed the populations. The staple crop of Ethiopia is teff and Europe and the US are increasing their production of this grain owing to its nutritional value (Woldeyohannes et al., 2022). Proso millet is yet another "poor man's food". The cultivation of proso millet is increasing in the northern US and China. From a catch crop, it is now evolving as an alternative to the major food basket across countries (Bhat et al., 2019). Foxtail millet is well renowned for its grain and fodder. Recent ventures in foxtail yield and productivity are welcomed by farmers in Telangana and Andhra Pradesh in India. Internationally, foxtail millet is also favored for its various savories and recipes in China, Europe, and the US. Millet is also being explored as a biofuel, which has generated large industrial interest (Petti et al., 2013; Aggarwal et al., 2022). In addition, Western Africa is also now considering the fonio millet value chain in its market as a future for crop and food diversification based on small millets (Kanlindogbe et al., 2020; Ibrahim Bio Yerima and Yerima, 2021). The new-age policies from the government for introducing finger millet in Nepal and enhancing its cultivation in Ethiopia (Gebreyohannes et al., 2021) provide a new gateway for future diets. Recently, a global strategy for the intake and conservation of small millets proposed by Bramel et al. (2022) also enforces the need to

enhance the cultivation of small millets to meet the continuously increasing demand. As discussed in the previous works on small millets, these crops can thrive to address future food and nutritional hunger (Muthamilarasan and Prasad, 2021; Renganathan et al., 2020).

11 Conclusion

This review highlights the overall value of small millets in our daily lives. Small millets were important food crops in ancient civilizations. Due to shifts in human dietary habits, these crops are now underutilized as food. By analyzing recent progress toward the use of small millets, we can realize support from government-aided projects and international and national collaborations to revitalize the genetic resources in small millets. Insights into the novel traits in small millets have demonstrated their paramount importance in nutritional and climate resilience. Therefore, this states the reasons behind the reverence of these crops in the cultural heritage of our ancestors. From these ethnobotanical records, we can conclude that breeding programs for improved varieties will restore the importance of on-farm conservation of wild and traditional landraces of small millets.

Small millets are climate-resilient crops that can meet the need for food and fodder and act as nutritional supplements. Hence, they are well considered as nutri-cereals and smart crops. Our detailed survey of studies in small millets revealed their great scope for applying advanced omics techniques to tap the genetic reasons for climate resilience and nutrition. Among small millets, the prospects of research in foxtail are higher and it has been utilized as a model crop system. Thus, small millets could act as a model crop for novel genes and mechanisms for manipulation by comparative genomics in mainstream cereals for increased stress tolerance and therapeutic traits. In this context, recent trends in food habits among the public have also begun to shift toward small millets, as evidenced by multigrain value-added products, processed foods, and new recipes in the markets. Therefore, small millets could be a major crop in future and a component in diversifying our food habits for a healthier life.

Author contributions

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

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

References

Aggarwal, P.i, Lydia, P., Choudhary, P., Singh, R., Shukla, P., Prasad, M., et al. (2022). Multi-omics intervention in Setaria to dissect climate-resilient traits: Progress and prospects. *Front. Plant Sci.* 13, 892736. doi:10.3389/fpls.2022.892736

Abdullah, S. N. A., Mayes, S., and Moradpour, M. (2021). Target gene identification and sgRNA design for waterlogging tolerance in foxtail millet via CRISPR-based transcriptional activation. *Curr. Chin. Sci.* 1 (5), 523–533. doi:10.2174/2210298101666210709104258

Abrouk, M., Ahmed, H. I., Cubry, P., Šimoníková, D., Cauet, S., Pailles, Y., et al. (2020). Fonio millet genome unlocks African orphan crop diversity for agriculture in a changing climate. *Nat. Commun.* 11 (1), 1–13. doi:10.1038/s41467-020-18329-4

Adekunle, A., Darwin, L., Valérie, O., and Raghavan, V. (2018). Helping agribusinesses—small millets value chain—to grow in India. *Agriculture* 8 (3), 44. doi:10.3390/agriculture8030044

Ambavane, A. R., Sawardekar, S. V., Sawantdesai, S. A., and Gokhale, N. (2015). Studies on mutagenic effectiveness and efficiency of gamma rays and its effect on quantitative traits in finger millet (Eleusine coracana L. Gaertn). *J. Radiat. Res. Appl. Sci.* 8, 120–125. doi:10.1016/j.jrras.2014.12.004

Anittha, I., and Mullainathan, L. (2018). Mutagenic effect of EMS and DES on Tenai (Setaria italica) in M1 generation. *J. Phytology*, 06–08. doi:10.25081/jp.2018.v10.3406

Babu, B. K. (2014). Molecular analysis of world collection of finger millet accessions for blast disease resistance using functional SSR markers. *SABRAO J. Breed. Genet.* 46 (2), 202–216.

Baltensperger, D. D. (2002). "Progress with proso, pearl and other millets," in *Trends in new crops and new uses*, 100–103.

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

Banerjee, P., and Maitra, S. (2020). The role of small millets as functional food to combat malnutrition in developing countries. *Indian J. Nat. Sci.* 10 (60), 20412. doi:10.1073/pnas.0402892101

Barretto, R., Buenavista, R. M., Rivera, J. L., Wang, S., Prasad, P. V., and Siliveru, K. (2021). Teff (eragrostis tef) processing, utilization and future opportunities: A review. *Int. J. Food Sci. Technol.* 56 (7), 3125–3137. doi:10.1111/ijfs.14872

Bennetzen, J. L., and Freeling, M. (1997). The unified grass genome: Synergy in synteny. *Genome Res.* 7, 301–306. doi:10.1101/gr.7.4.301

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 (6), 555–561. doi:10.1038/nbt.2196

Bhat, S., Nandini, C., Srinathareddy, S., Jayarame, G., and Prabhakar (2019). Proso millet (*Panicum miliaceum* l.)-a climate resilient crop for food and nutritional security: A review. *Environ. Conserv. J.* 20 (3), 113–124. doi:10.36953/ECJ.2019.20315

Bhat, S., Nandini, C., and Tippeswamy, V. (2018). Significance of small millets in nutrition and health-A review. *J. Dairy Foods Home Sci.* 37 (1), 35–40. doi:10.18805/ajdfr.DR-1306

Bhatt, R., Asopa, P. P., Jain, R., Kothari-Chajer, A., Kothari, S. L., and Kachhwaha, S. (2021). Optimization of agrobacterium mediated genetic transformation in paspalum scrobiculatum L. (Kodo millet). *Agronomy* 11 (6), 1104. doi:10.3390/agronomy11061104

Bhave, K. G., Dalvi, V. V., Thaware, B. L., Mahadik, S. G., Kasture, M. C., and Desai, S. S. (2016). Mutagenesis in proso millet (Panicum miliaceum L.). *Int. J. Sci. Res.* 5, 1635–1638.

Bonthala, V. S., Muthamilarasan, M., Roy, R., and Prasad, M. (2014). FmTFDb: A foxtail millet transcription factors database for expediting functional genomics in millets. *Mol. Biol. Rep.* 41 (10), 6343–6348. doi:10.1007/s11033-014-3574-y

Bouchet, S., Marcus, O., Olatoye, S. R., Marla, S. R., Ramasamy, P., Tesfaye, T., et al. (2017). Increased power to dissect adaptive traits in global sorghum diversity using a nested association mapping population. *Genetics* 206 (2), 573–585. doi:10.1534/genetics.116.198499

Boukail, S., Macharia, M., Miculan, M., Masoni, A., Calamai, A., Palchetti, E., et al. (2021). Genome wide association study of agronomic and seed traits in a world collection of proso millet (Panicum miliaceum L.). *BMC Plant Biol.* 21 (1), 1–12. doi:10.1186/s12870-021-03111-5

Bramel, P., Giovannini, P., and Eshan Dulloo, M. (2022). *Global strategy for the conservation and use of genetic resources of selected millets*. Bonn, Germany: Global Crop Diversity Trust.

Cannarozzi, G., Plaza-wüthrich, S., Esfeld, K., Larti, S., Wilson, Y. S., Girma, D., et al. (2014). Genome and transcriptome sequencing identifies breeding targets in the orphan crop tef (*Eragrostis tef*). *BMC Genomics* 15, 581. doi:10.1186/1471-2164-15-581

Chandra, A. K., Pandey, D., Tiwari, A., Sharma, D., Agarwal, A., Sood, S., et al. (2020). An omics study of iron and zinc homeostasis in finger millet: Biofortified foods for micronutrient deficiency in an era of climate change? *OMICS A J. Integr. Biol.* 24 (12), 688–705. doi:10.1089/omi.2020.0095

Chanyalew, S., Singh, H., Tefera, H., and Sorrels, M. (2005). Molecular genetic map and QTL analysis of agronomic traits based on a Eragrostis tef x E. pilosa recombinant inbred population. *J. Genet. Breed.* 59 (1), 53.

Chapman, M. A., He, Y., and Zhou, M. (2022). Beyond a reference genome: Pangenomes and population genomics of underutilized and orphan crops for future food and nutrition security. *New Phytol.* 234 (5), 1583–1597. doi:10.1111/nph.18021

Cheng, Z., Sun, Y., Yang, S., Zhi, H., Yin, T., Ma, X., and Sui, Y. (2021). Establishing in planta haploid inducer line by edited SiMTL in foxtail millet (Setaria italica). *Plant Biotechnology Journal*, 19(6), 1089. doi:10.1111/pbi.13584

Choi, H-K (2019). Translational genomics and multi-omics integrated approaches as a useful strategy for crop breeding. *Genes. Genomics* 41 (2), 133–146. doi:10.1007/s13258-018-0751-8

Conn, S., and Gilliham, M. (2010). Comparative physiology of elemental distributions in plants. *Ann. Bot.* 105, 1081–1102. doi:10.1093/aob/mcq027

Coppola, M., Diretto, G., Cristina Digilio, M., Woo, S., Giuliano, G., Molisso, D., et al. (2019). Transcriptome and metabolome reprogramming in tomato plants by Trichoderma harzianum strain T22 primes and enhances defense responses against aphids. *Front. Physiol.* 10, 745. doi:10.3389/fphys.2019.00745

Dansi, A., Adoukonou-Sagbadja, H., and Vodouhe, R. (2010). Diversity, conservation and related wild species of Fonio millet (Digitaria spp.) in the northwest of Benin. *Genet. Resour. Crop Evol.* 57 (6), 827–839. doi:10.1007/s10722-009-9522-3

Das, R. R., Pradhan, S., and Parida, A. (2020). De-novo transcriptome analysis unveils differentially expressed genes regulating drought and salt stress response in Panicum sumatrense. *Sci. Rep.* 10 (1), 1–14. doi:10.1038/s41598-020-78118-3

David, R. H. A., Ramakrishnan, M., Maharajan, T., BarathiKannan, K., Babu, G. A., Daniel, M. A., et al. (2021). Mining QTL and genes for root traits and biochemical parameters under vegetative drought in South Indian genotypes of finger millet (*Eleusine coracana* (L.) Gaertn) by association mapping and *in silico* comparative genomics. *Biocatal. Agric. Biotechnol.* 32, 101935. doi:10.1016/j.bcab.2021.101935

Deepak Taggelli, R. G., and Thakur, V. (2018). Minor millets-their potential health benefits and medicinal properties: A review. *Int. J. Pure Appl. Sci.* 6 (1), 1677.

Desai, H., Hamid, R., Ghorbanzadeh, Z., Bhut, N., Padhiyar, S. M., Kheni, J., et al. (2021). Genic microsatellite marker characterization and development in little millet (Panicum sumatrense) using transcriptome sequencing. *Sci. Rep.* 11 (1), 1–14. doi:10.1038/s41598-021-00100-4

Deshpande, S. S., Mohapatra, D., Tripathi, M. K., and Sadvatha, R. H. (2015). Kodo millet-nutritional value and utilization in Indian foods. *J. grain Process. storage* 2 (2), 16.

Dey, S., Saxena, A., Kumar, Y., Maity, T., and Tarafdar, A. (2022). Understanding the antinutritional factors and bioactive compounds of kodo millet (paspalum scrobiculatum) and little millet (Panicum sumatrense). *J. Food Qual.* 2022, 1–19. doi:10.1155/2022/1578448

Dhawale, R. N. (2022). Metabolomic profiling of drought-tolerant little millet (Panicum sumatrense L.) genotype in response to drought stress. 11(4), 1754.

Doust, A. N., Devos, K. M., Gadberry, M. D., Gale, M. D., and Kellogg, E. A. (2004). Genetic control of branching in foxtail millet. *Proc. Natl. Acad. Sci. U.S.A.* 101 (24), 9045–9050.

Durairaj, M., Gurumurthy, G., Nachimuthu, V., Karthikeyan, M., and Balasubramanian, S. (2019). Dehulled small millets: The promising nutricereals for improving the nutrition of children. *Matern. Child. Nutr.* 15, e12791. doi:10.1111/mcn.12791

Durgad, A. G., AmruthaJoshi, T., and Hiremath, G. M. (2021). Consumer preference for foxtail and little millets in north eastern region of Karnataka. *Econ. Aff.* 66 (1), 101–108. doi:10.46852/0424-2513.1.2021.13

Eswari, K., Gogulan, G., Hari Prathab, K. A., and Sakila, M. (2014). Development of early maturing mutants in finger millet. *Res. J. Agric. For. Sci. , ISSN* 2320, 6063.

Fan, X., Tang, S., Zhi, H., He, M., Ma, W., Jia, Y., et al. (2017). Identification and fine mapping of SiDWARF3 (D3), a pleiotropic locus controlling environment-independent dwarfism in foxtail millet. *Crop Sci.* 57 (5), 2431–2442. doi:10.2135/cropsci2016.11.0952

Fang, X., Dong, K., Wang, X., Liu, T., Jihong, H., 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 (1), 1–12. doi:10.1186/s12864-016-2628-z

Fekih, R., Takagi, H., Tamiru, M., Abe, A., Natsume, S., Yaegashi, H., et al. (2013). MutMap+: Genetic mapping and mutant identification without crossing in rice. *PloS one* 8 (7), pe68529. doi:10.1371/journal.pone.0068529

Francis, N., Rajasekaran, R., Krishnamoorthy, I., Muthurajan, R., Thiyagarajan, C., and Alagarswamy, S. (2022). Gamma irradiation to induce beneficial mutants in proso millet (Panicum miliaceum L.): An underutilized food crop. *Int. J. Radiat. Biol.* 98, 1277–1288. doi:10.1080/09553002.2022.2024292

Fukunaga, K., Abe, A., Mukainari, Y., Komori, K., Tanaka, K., Fujihara, A., et al. (2022a). Recombinant inbred lines and next-generation sequencing enable rapid identification of candidate genes involved in morphological and agronomic traits in foxtail millet. *Sci. Rep.* 12 (1), 1–12. doi:10.1038/s41598-021-04012-1

Fukunaga, K., Wang, Z., Kato, K., and Kawase, M. (2002b). Geographical variation of nuclear genome RFLPs and genetic differentiation in foxtail millet, Setaria italica (L.) P. Beauv. *Genet. Resour. Crop Evol.* 49 (1), 95–101. doi:10.1023/a:1013852007770

Ganapathy, K. N., and Patil, J. V. (2017). Improvement in finger millet: Status and future prospects. *Millets sorghum Biol. Genet. Improv.*, 87–111. doi:10.1002/9781119130765.ch3

Ganapathy, S. A., Nirmalakumari, N., Senthil, J., Souframanien, and Raveendran, T. S. (2008). Isolation of macromutations and mutagenic effectiveness and efficiency in little millet varieties. *World J. Agric. Sci.*, 4, 483. ISSN 1817- 3047.

Gebreyohannes, A., Shimelis, H., Laing, M., Mathew, I., Odeny, D. A., and Ojulong, H. (2021). Finger millet production in Ethiopia: Opportunities, problem diagnosis, key challenges and recommendations for breeding. *Sustainability* 13 (23), 13463. doi:10.3390/su132313463

Geervani, P., and Eggum, O. (1989). Nutrient composition and protein quality of minor millets. *Plant Foods Hum. Nutr.* 39, 201–208. doi:10.1007/BF01091900

Gilles, L. M., Khaled, A., Laffaire, J. B., Chaignon, S., Gendrot, G., Laplaige, J., et al. (2017). Loss of pollen-specific phospholipase NOT LIKE DAD triggers gynogenesis in maize. *Embo J.* 36, 707–717. doi:10.15252/embj.201796603

Gopalan, C., Rama Sastri, B. V., and Balasubramanian, S. C. (2009). *Nutritive value of Indian foods*. Hyderabad, India: National Institute of Nutrition, Indian Council of Medical Research.

Goron, T. L., and Raizada, M. (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, J., Huang, Z., Sun, J., Cui, X., and Liu, Y. (2021). Research progress and future development trends in medicinal plant transcriptomics. *Front. Plant Sci.* 12, 691838. doi:10.3389/fpls.2021.691838

Guo, L, L., Qiu, J., Ye, C., Jin, G., Mao, L., Zhang, H., et al. (2017). Echinochloa crus-galli genome analysis provides insight into its adaptation and invasiveness as a weed. *Nat. Commun.* 8 (1), 1–10. doi:10.1038/s41467-017-01067-5

Guo, R, R., Shi, L., Yan, C., Zhong, X., Gu, F., Liu, Q., et al. (2017). Ionomic and metabolic responses to neutral salt or alkaline salt stresses in maize (Zea mays L.) seedlings. *BMC Plant Biol.* 17, 41. doi:10.1186/s12870-017-0994-6

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, US: A review. *Front. Plant Sci.* 7, 1961. doi:10.3389/fpls.2016.01961

Hariprasanna, K., Agte, V., Elangovan, M., and Patil, J. V. (2014). Genetic variability for grain iron and zinc content in cultivars, breeding lines and selected germplasm accessions of sorghum [*Sorghum bicolor* (L.) Moench]. *Ind. Jrnl. Gen. Plnt. Bree.* 74, 42–49. doi:10.5958/j.0975-6906.74.1.006

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 (1), 39–47. doi:10.1093/dnares/dsx036

He, L., Zhang, B., Wang, X., Li, H., and Han, Y. (2015). Foxtail millet: Nutritional and eating quality, and prospects for genetic improvement. *Front. Agric. Sci. Eng.* 2, 124–133. doi:10.15302/J-FASE-2015054

He, Q., Zhi, H., Tang, S., Xing, L., Wang, S., Wang, H., et al. (2021). QTL mapping for foxtail millet plant height in multi-environment using an ultra-high density bin map. *Theor. Appl. Genet.* 134 (2), 557–572. doi:10.1007/s00122-020-03714-w

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 (1), 1–16. doi:10.1186/s12864-017-3850-z

Hitu, K. W., M'ribu, K., Liang, H., and Mandelbaum, C. (1997). Fonio millets: Ethnobotany, genetic diversity and evolution. *South Afr. J. Bot.* 63 (4), 185–190. doi:10.1016/S0254-6299(15)30742-0

Ibrahim Bio Yerima, A. R., and Achigan-Dako, E. G. (2021). A review of the orphan small grain cereals improvement with a comprehensive plan for genomics-assisted breeding of fonio millet in West Africa. *Plant Breed.* 140 (4), 561–574. doi:10.1111/pbr.12930

Indirani, K., and Devasena, M. (2021). Review on nutritional profiles and health benefits of little millets –India. *Int. J. Res. Eng. Sci. (IJRES)* 9, 7.

Jacquier, N. M. A., Gilles, L. M., Pyott, D. E., Martinant, J. P., Rogowsky, P. M., and Widiez, T. (2020). Puzzling out plant reproduction by haploid induction for innovations in plant breeding. *Nat. Plants* 6, 610–619. doi:10.1038/s41477-020-0664-9

Jayakodi, M., Madheswaran, M., Adhimoolam, K., Perumal, S., Manickam, D., 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 (5), 66–11. doi:10.1007/s11738-019-2855-4

Jency, J. P., Ravikesavan, R., Sumathi, P., and Raveendran, M. (2016). Determination of lethal dose and effect of physical mutagen on germination percentage and seedling parameters in kodo millet variety CO 3. *Electron. Journ. Plan. Breed.* 7 (4), 1122–1126. doi:10.5958/0975-928x.2016.00155.1

Joshi, D. C., Meena, R. P., and Chandora, R. (2021). "Genetic resources: Collection, characterization, conservation, and documentation," in *Millets and Pseudo Cereals* (India: Woodhead Publishing), 19–31. doi:10.1016/B978-0-12-820089-6.00003-3

Kaimal, A. M., Mujumdar, A. S., and Thorat, B. N. (2021). Resistant starch from millets: Recent developments and applications in food industries. *Trends Food Sci. Technol.* 111, 563–580. doi:10.1016/j.tifs.2021.02.074

Kanlindogbe, C., Sekloka, E., and Kwon-Ndung, E. H. (2020). Genetic resources and varietal environment of grown fonio millets in West Africa: Challenges and perspectives. *Plant Breed. Biotech.* 8 (2), 77–88. doi:10.9787/PBB.2020.8.2.77

Kim, J. K., Park, S. Y., Yeo, Y., Cho, H. S., Kim, Y. B., Bae, H., et al. (2013). Metabolic profiling of millet (*Panicum miliaceum*) using gas chromatography-time-of-flight mass spectrometry (GC-TOFMS) for quality assessment. *Plant Omics* 6 (1), 73–78.

Kleter, G. A., Kuiper, H. A., and Kok, E. J. (2019). Gene-edited crops: Towards a harmonized safety assessment. *Trends Biotechnol.* 37 (5), 443–447. doi:10.1016/j.tibtech.2018.11.014

Kumar, A., Kumar, S., and Kumar, S. (2015). Promoting minor millets for food and nutritional security. *Acta Hortic.* 1241, 571–576. doi:10.17660/actahortic.2019.1241.84

Kumar, R., Sharma, V., Srinivas, S., Ramrao, D., Akash, V., Kumar, S., et al. (2021). Understanding omics driven plant improvement and de novo crop domestication: Some examples. *Front. Genet.* 415, 637141. doi:10.3389/fgene.2021.637141

Lata, C., and Prasad, M. (2013). Validation of an allele-specific marker associated with dehydration stress tolerance in a core set of foxtail millet accessions. *Plant Breed.* 132, 496–499.

Li, H., Han, S., Huo, Y., Ma, G., Sun, Z., Li, H., et al. (2022). Comparative metabolomic and transcriptomic analysis reveals a coexpression network of the carotenoid metabolism pathway in the panicle of Setaria italica. *BMC Plant Biol.* 22 (1), 105–116. doi:10.1186/s12870-022-03467-2

Li, K., Wen, W., Alseekh, S., Yang, X., Guo, H., Li, W., et al. (2019). Largescale metabolite quantitative trait locus analysis provides new insights for high-quality maize improvement. *Plant J.* 99, 216–230. doi:10.1111/tpj.14317

Li, T., Yang, X., Yu, Y., Si, X., Zhai, X., Zhang, H., et al. (2018). Domestication of wild tomato is accelerated by genome editing. *Nat. Biotechnol.* 36, 1160–1163. doi:10.1038/nbt.4273

Li, J, J., Wang, Y., Wang, L., Zhu, J., Deng, J., Tang, R., et al. (2021). Integration of transcriptomic and proteomic analyses for finger millet [Eleusine coracana (L.) Gaertn.] in response to drought stress. *Plos one* 16 (2), e0247181. doi:10.1371/journal.pone.0247181

Li, W, W., Wen, L., Chen, Z., Zhang, Z., Pang, X., Deng, Z., et al. (2021). Study on metabolic variation in whole grains of four proso millet varieties reveals metabolites important for antioxidant properties and quality traits. *Food Chem.* 357, 129791. doi:10.1016/j.foodchem.2021.129791

Liang, K., Liang, S., Lu, L., Zhu, D., Zhu, H., Liu, P., et al. (2018). Metabolic variation and cooking qualities of millet cultivars grown both organically and conventionally. *Food Res. Int.* 106, 825–833. doi:10.1016/j.foodres.2018.01.023

Liu, T., He, J., Dong, K., Wang, X., Wang, W., Yang, P., et al. (2020). QTL mapping of yield component traits on bin map generated from resequencing a RIL population of foxtail millet (Setaria italica). *BMC genomics* 21 (1), 1–13. doi:10.1186/s12864-020-6553-9

Liu, C, C., Yuan, Y., Liu, J., Wang, H., Ma, Q., Zhou, Y., et al. (2022). Comparative transcriptome and physiological analysis unravel proso millet (Panicum miliaceum L.) source leaf adaptation to nitrogen deficiency with high nitrogen use efficiency. *Environ. Exp. Bot.* 199, 104891. doi:10.1016/j.envexpbot.2022.104891

Liu, T, T., Liu, X., He, J., Dong, K., Pan, W., Zhang, L., et al. (2022). Identification and fine-mapping of a major QTL (PH1. 1) conferring plant height in broomcorn millet (Panicum miliaceum). *Front. Plant Sci.* 13, 1010057. doi:10.3389/fpls.2022.1010057

Lloyd, J. R., and Kossmann, J. (2021). Improving crops for a changing world. *Front. Plant Sci.* 12, 728328. doi:10.3389/fpls.2021.728328

Longvah, T., Ananthan, R., Bhaskarachary, K., and Venkaiah, K. (2017). *Indian food composition tables*. Hyderabad, India: National Institute of Nutrition.

Luo, J. (2015). Metabolite-based genome-wide association studies in plants. *Curr. Opin. Plant Biol.* 24, 31–38. doi:10.1016/j.pbi.2015.01.006

Maitra, S. (2020). *Intercropping of small millets for agricultural sustainability in drylands: A review*. Odisha, India: Crop Research, 55.

Matsuda, F., Nakabayashi, R., Yang, Z., Okazaki, Y., Yonemaru, J., Ebana, K., et al. (2015). Metabolome-genome-wide association study dissects genetic architecture for generating natural variation in rice secondary metabolism. *Plant J.* 81, 13–23. doi:10.1111/tpj.12681

McMullen, M. D., Stephen, K., Sanchez Villeda., H., Peter, B., Huihui, L., Qi, S., et al. (2009). Genetic properties of the maize nested association mapping population. *Science* 325, 737–740. doi:10.1126/science.1174320

Merga, M. (2018). Progress, achievements and challenges of tef breeding in Ethiopia. *J. Agric. Res.* 9 (1), 1–8.

Morgante, M., De Paoli, E., and Radovic, S. (2007). Transposable elements and the plant pan-genomes. *Curr. Opin. Plant Biol.* 10, 149–155. doi:10.1016/j.pbi.2007.02.001

Morris, J., Hawthorne, K. M., Hotze, T., Abrams, S. A., and Hirschi, K. D. (2008). Nutritional impact of elevated calcium transport activity in carrots. *Proc. Natl. Acad. Sci. U. S. A.* 105, 1431–1435. doi:10.1073/pnas.0709005105

Mundada, P. S., Barvkar, V. T., Umdale, S. D., Kumar, S. A., Nikam, T. D., and Ahire, M. L. (2021). An insight into the role of silicon on retaliation to osmotic stress in finger millet (Eleusine coracana (L.) Gaertn). *J. Hazard. Mat.* 403, 124078. doi:10.1016/j.jhazmat.2020.124078

Muthamilarasan, M., and Prasad, M. (2015). Advances in Setaria genomics for genetic improvement of cereals and bioenergy grasses. *Theor. Appl. Genet.* 128 (1), 1–14. doi:10.1007/s00122-014-2399-3

Muthamilarasan, M., and Prasad, M. (2021). Small millets for enduring food security amidst pandemics. *Trends Plant Sci.* 26 (1), 33–40. doi:10.1016/j.tplants.2020.08.008

Nagaraja, A., and Das, I. K. (2016). "Disease resistance in pearl millet and small millets," in *Biotic stress resistance in millets* (Academic Press).

National Research Council (1996). *Lost crops of Africa: Volume I: Grains*. Washington, DC: National Academies Press. doi:10.17226/2305

Newmaster, S. G., Ragupathy, S., Dhivya, S., Jijo, C. J., Sathishkumar, R., and Patel, K. (2013). Chitilappilly Joseph Jijo., Ramalingam Sathishkumar., and Kirit PatelGenomic valorization of the fine scale classification of small millet landraces in southern India. *Genome* 56 (2), 123–127. doi:10.1139/gen-2012-0183

Ni, X., Xia, Q., Zhang, H., Cheng, S., Li, H., Fan, G., et al. (2017). Updated foxtail millet genome assembly and gene mapping of nine key agronomic traits by resequencing a RIL population. *GigaScience* 6 (2), 1–8. doi:10.1093/gigascience/giw005

Padhiyar, S. M., Kheni, J., and Singh, H. V. D. R. (2022). Metabolomic characterization of barnyard millet (*Echinochloa frumentacea* L.) involved in different stages of spike development. *Pharma innov.* 11 (6), 604–611.

Pan, J., Li, Z., Dai, S., Ding, H., Wang, Q., Li, X., et al. (2020). Integrative analyses of transcriptomics and metabolomics upon seed germination of foxtail millet in response to salinity. *Sci. Rep.* 10 (1), 1–16. Available from. doi:10.1038/s41598-020-70520-1

Paroda, R. S., and Mal, B. (1989). *New plant sources for food and industry in*. New crops food industry.

Parvathi, M. S., Nataraja, K. N., Reddy, Y. A. N., Naika, M. B. N., and Gowda, M. V. C. (2019). Transcriptome analysis of finger millet (Eleusine coracana (L.) Gaertn.) reveals unique drought responsive genes. *J. Genet.* 98 (2), 46–12. doi:10.1007/s12041-019-1087-0

Pendergast IV, T. H., Qi, P., Odeny, D. A., Dida, M. M., and Devos, K. M. (2022). A high-density linkage map of finger millet provides QTL for blast resistance and other agronomic traits. *Plant Genome* 15 (1), e20175. doi:10.1002/tpg2.20175

Peng, H., Wang, K., Chen, Z., Cao, Y., Gao, Q., Yan, L., et al. (2020). MBKbase for rice: An integrated omics knowledgebase for molecular breeding in rice. *Nucleic Acids Res.* 48 (1), D1085-D1092–D1092. doi:10.1093/nar/gkz921

Petti, C., Shearer, A., Tateno, M., Ruwaya, M., Nokes, S., Brutnell, T., et al. (2013). Comparative feedstock analysis in Setaria viridis L. as a model for C4 bioenergy grasses and Panicoid crop species. *Front. Plant Sci.* 4, 181. doi:10.3389/fpls.2013.00181

Pinson, S. R. M., Tarpley, L., Yan, W., Yeater, K., Lahner, B., Yakubova, E., et al. (2015). Worldwide genetic diversity for mineral element concentrations in rice grain. *Crop Sci.* 55, 294–311. doi:10.2135/cropsci2013.10.0656

Pramitha Lydia, J., Subhalakshmi, V. K. I., Jeeva, G., and Surendar, A. (2021). *Small millets exploring the hidden potential*. North West Delhi, India: Scripown Publications.

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 (3), 280–287. doi:10.1016/j.jplph.2010.07.005

Qin, L., Chen, E., Li, F., Yu, X., Liu, Z., Yang, Y., et al. (2020). Genome-wide gene expression profiles analysis reveal novel insights into drought stress in foxtail millet (Setaria italica L.). *Int. J. Mol. Sci.* 21 (22), 8520–8521. doi:10.3390/ijms21228520

Ragupathy, S., Dhivya, S., Patel., Kirit, Abiran, S., Sambandan, K., Gartaula, H., et al. (2016). DNA record of some traditional small millet landraces in India and Nepal. *3 Biotech.* 6 (2), 1–19. doi:10.1007/s13205-016-0450-6

Rahman, H., Jagadeesh, s. 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 (4), 485–503. doi:10.1007/s11103-014-0199-4

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 (4), 37–18. doi:10.1007/s11032-016-0460-4

Ramakrishnan, M., Antony Ceasar, S., Duraipandiyan, V., Vinod, K. K., Kalpana, K., Al-Dhabi, N. A., et al. (2016). Tracing QTLs for leaf blast resistance and agronomic performance of finger millet (Eleusine coracana (L.) Gaertn.) genotypes through association mapping and *in silico* comparative genomics analyses. *PLoS One* 11 (7), pe0159264. doi:10.1371/journal.pone.0159264

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 (8), e0183261. doi:10.1371/journal.pone.0183261

Ramesh, M., Vanniarajan, C., Ravikesavan, R., Aiyan, K. E. A., and Mahendran, P. P. (2019). Determination of lethal dose and effect of EMS and gamma ray on germination percentage and seedling parameters in barnyard millet variety Co Kv 2. *Electron. Journ. Plan. Breed.* 10 (2), 957–962. doi:10.5958/0975-928x.2019.00123.6

Rao Dayakar, B., Bhaskarachary, K., Arlene Christina, G. D., Sudha Devi, G., Vilas, A. T., and Tonapi, A. (2017). *Nutritional and health benefits of millets*. Rajendranagar, Hyderabad: ICAR, Indian Institute of Millets Research IIMR, 112.

Ravikesavan, R., Sivamurugan, A. P., Iyanar, K., Pramitha, J. L., and Nirmalakumari, A. (2022). *Millet cultivation: An overview. Handbook of millets-processing, quality, and nutrition status*, 23.

Rawat, J. M., Pandey, S., Debbarma, P., and Rawat, B. (2021). Preparation of alcoholic beverages by tribal communities in the Indian himalayan region: A review on traditional and ethnic consideration. *Front. Sustain. Food Syst.* 5, 672411. doi:10.3389/fsufs.2021.672411

Renganathan, V. G., 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

Saha, M. V., Gowda, C., Arya, L., Verma, M., and Bansal, K; C. (2016). Genetic and genomic resources of small millets. *Crit. Rev. Plant Sci.* 35 (1), 56–79. doi:10.1080/07352689.2016.1147907

Sahu, K. R., and Sharma, M. L. (2013). Medicinal and other uses of small millets by the tribal farmers of the Bastar Plateau Zone of Chhattisgarh. *Agric. Update* 4 (8), 596–599.

Satyarthi, K. S., Sharma, Y. P., Sharma, P., Vaidiya, P., and Randhawa, S. S. (2018). *Traditional food grain crops of Himachal Pradesh*. Shimla, Himachal Pradesh: Himachal Pradesh State Biodiversity Board. Himachal Pradesh Council for Science, Technology and Environment, Bemole.

Schaart, J. G., van de Wiel, C. C., Lotz, L. A., and Smulders, M. J. (2016). Opportunities for products of new plant breeding techniques. *Trends in Plant Science*, 21(5), 438–449. doi:10.1016/j.tplants.2015.11.006

Scossa, F., Saleh, A., and Fernie, A. R. (2021). Integrating multi-omics data for crop improvement. *J. Plant Physiol.* 257, 153352. doi:10.1016/j.jplph.2020.153352

Shan, Z., Jiang, Y., Li, H., Guo, J., Dong, M., Zhang, J., et al. (2020). Genome-wide analysis of the NAC transcription factor family in broomcorn millet (Panicum miliaceum L.) and expression analysis under drought stress. *BMC Genomics* 21 (1), 96–13. doi:10.1186/s12864-020-6479-2

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 (8), e0199444. doi:10.1371/journal.pone.0199444

Shi, J., Ma, X., Zhang, J., Zhou, Y., Liu, M., Huang, L., et al. (2019). Chromosome conformation capture resolved near complete genome assembly of broomcorn millet. *Nat. Commun.* 10 (1), 464–469. doi:10.1038/s41467-018-07876-6

Singh, R. K., Muthamilarasan, M., and Prasad, M. (2021). Biotechnological approaches to dissect climate-resilient traits in millets and their application in crop improvement. *J. Biotechnol.* 327, 64–73. doi:10.1016/j.jbiotec.2021.01.002

Sood, S., Joshi, D. C., and Pattanayak, A. (2020). Breeding advancements in barnyard millet. *Accel. Plant Breed.* 1, 391. doi:10.1007/978-3-030-41866-3_15

Sruthi, N. U., and Rao, P. S. (2021). Effect of processing on storage stability of millet flour: A review. *Trends Food Sci. Technol.* 112, 58–74. doi:10.1016/j.tifs.2021.03.043

Suresh, B. V., Choudhary, P., Aggarwal, P. R., Rana, S., Singh, R. K., Ravikesavan, R., et al. (2022). De novo transcriptome analysis identifies key genes involved in dehydration stress response in kodo millet (Paspalum scrobiculatum L.). *Genomics* 114 (3), 110347. doi:10.1016/j.ygeno.2022.110347

Tay Fernandez, C. G., Nestor, B. J., Danilevicz, M. F., Gill, M., Petereit, J., Bayer, P. E., et al. (2022). Pangenomes as a resource to accelerate breeding of under-utilised crop species. *Int. J. Mol. Sci.* 23 (5), 2671. doi:10.3390/ijms23052671

Tettelin, H., Masignani, V., Cieslewicz, M. J., Donati, C., Medini, D., Ward, N. L., et al. (2005). Genome analysis of multiple pathogenic isolates of Streptococcus agalactiae: Implications for the microbial "pan- genome". *Proc. Natl. Acad. Sci. U. S. A.* 102, 13950–13955. doi:10.1073/pnas.0506758102

Tian, B., Zhang, L., Liu, Y., Wu, P., Wang, W., Zhang, Y., et al. (2021). Identification of QTL for resistance to leaf blast in foxtail millet by genome re-sequencing analysis. (2), 743–754. doi:10.1007/s00122-020-03730-w

Tiozon, R., Sartagoda, K. J. D., Fernie, A. R., and Sreenivasulu, N. (2021). Kristel june D., sartagoda, alisdair R. Fernie., and nese SreenivasuluThe nutritional profile and human health benefit of pigmented rice and the impact of post-harvest processes and product development on the nutritional components: A review. *Crit. Rev. Food Sci. Nutr.*, 1–28. doi:10.1080/10408398.2021.1995697

Upadhyaya, Hari D., Gowda, C. L. L., Gopal Reddy, V., and Singh, S. (2008). *Diversity of small millets germplasm in gene bank at ICRISAT*, 173.

Upadhyaya, H. D., Vetriventhan, M., Dwivedi, S. L., Pattanashetti, S. K., and Singh, S. K. (2016). "Proso, barnyard, little, and kodo millets," in *Genetic and genomic resources for grain cereals improvement* (Academic Press), 321.

Upadhyaya, H. D., and Vetriventhan, M. (2018). *Underutilized climate-smart nutrient rich small millets for food and nutritional security*. Patancheru, India: International Crops Research Institute for the Semi-Arid Tropics (ICRISAT).

VanBuren, R., Man Wai, C., Wang, X., Pardo, J., Yocca, A. E., Wang, H., et al. (2020). Exceptional sub genome stability and functional divergence in the allotetraploid Ethiopian cereal teff. *Nat. Commun.* 11 (1), 884. doi:10.1038/s41467-020-14724-z

Venkatesan, P., Sundaramari, M., and Venkattakumar, R. (2015). *Ethno taxonomical classification of little millet (Panicum sumatrance) by the tribal people in Tamilnadu*. India.

Verma, V. C., Verma., V. C., Singh., A., and Agrawal, S. (2018). Ethnobotanical study of small millets from India: Prodigious grain for nutritional and industrial aspects. *Int. J. Chem. Stud.* 6 (4), 2155.

Vetriventhan, M., and Upadhyaya, H. D. (2019). Variability for productivity and nutritional traits in germplasm of kodo millet, an underutilized nutrient-

rich climate smart crop. *Crop Sci.* 59 (3), 1095–1106. doi:10.2135/cropsci2018.07.0450

Vetriventhan, M., Azevedo, V. C., Upadhyaya, H. D., Nirmalakumari, A., Kane-Potaka, J., Anitha, S., et al. (2020). Genetic and genomic resources, and breeding for accelerating improvement of small millets: Current status and future interventions. *Nucleus* 63 (3), 217–239. doi:10.1007/s13237-020-00322-3

Vetriventhan, Mani., and Upadhyaya, Hari D. (2018). Diversity and trait-specific sources for productivity and nutritional traits in the global proso millet (Panicum miliaceum L.) germplasm collection. *Crop J.* 6 (5), 451–463. doi:10.1016/j.cj.2018.04.002

Vinoth, A., and Ravindhran, R. (2017). Biofortification in millets: A sustainable approach for nutritional security. *Front. Plant Sci.* 8, 29. doi:10.3389/fpls.2017.00029

Wang, J., Wang, Z. L., Du, X. F., Yang, H. Q., Han, F., Han, Y. H., 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, X., Chen, S., Ma, X., Yssel, A. E., Chaluvadi, S. R., Johnson, M. S., et al. (2021). Genome sequence and genetic diversity analysis of an under-domesticated orphan crop, white fonio (Digitaria exilis). *Gigascience* 10 (3), giab013. doi:10.1093/gigascience/giab013

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 (1), 31–36. doi:10.1007/s001220050705

Woldeyohannes, A. B., Desta, E. A., Fadda, C., Pè, M. E., and Dell'Acqua, M. (2022). Value of teff (eragrostis tef) genetic resources to support breeding for conventional and smallholder farming: A review. *CABI Agric. Biosci.* 3 (1), 27–16. doi:10.1186/s43170-022-00076-9

Wyatt, S. E., Tsou, P. L., and Robertson, D. (2002). Expression of the high-capacity calcium binding domain of calreticulin increases bioavailable calcium stores in plants. *Transgenic Res.* 11, 1–10. doi:10.1023/A:1013917701701

Xia, D., Zhou, H., Wang, Y., Li, P., Fu, P., Wu, B., et al. (2021). How rice organs are colored: The genetic basis of anthocyanin biosynthesis in rice. *Crop J.* 9 (3), 598–608. doi:10.1016/j.cj.2021.03.013

Yang, L., Li, R., Cui, Y., Qin, X., and Li, Z. (2021). Comparison of nutritional compositions of foxtail millet from the different cultivation regions by UPLC-Q-Orbitrap HRMS based metabolomics approach. *J. Food Biochem.* 45 (10), e13940. doi:10.1111/jfbc.13940

Yang, W., Guo, Z., Huang, C., Wang, K., Jiang, N., Fen, H., et al. (2015). Genome wide association study of rice (Oryza sativa L.) leaf traits with a high-throughput leaf scorer. *J. Exp. Bot.* 66, 5605–5615. doi:10.1093/jxb/erv100

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 (5), 518–527. doi:10.1270/jsbbs.17061

Zeid, M., Belay, G., Mulkey, S., Poland, J., and Sorrells, M. E. (2011). QTL mapping for yield and lodging resistance in an enhanced SSR-based map for tef. *Theor. Appl. Genet.* 122 (1), 77–93. doi:10.1007/s00122-010-1424-4

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 (6), 549–554. doi:10.1038/nbt.2195

Zhang, H., Tang, S., Schnable, J. C., He, Q., Gao, Y., Luo, M., et al. (2021). SiSTL2 is required for cell cycle, leaf organ development, chloroplast biogenesis, and has effects on C4 photosynthesis in setaria italica (L.) P. Beauv. *Front. Plant Sci.* 12, 1103. doi:10.3389/fpls.2018.01103

Zhang, K., Jia, X., Zhu, Z., and Xue, W. (2020). Physicochemical properties of rice analogs based on multi-level: Influence of the interaction of extrusion parameters. *Int. J. Food Prop.* 23 (1), 2033–2049. doi:10.1080/10942912.2020.1840389

Zhang, Y., Gao, X., Li, J., Gong, X., Yang, P., Gao, J., et al. (2019). Comparative analysis of proso millet (Panicum miliaceum L.) leaf transcriptomes for insight into drought tolerance mechanisms. *BMC Plant Biol.* 19 (1), 1–17. doi:10.1186/s12870-019-2001-x

Zhi, H., He, Q., Tang, S., Yang, J., Zhang, W., Liu, H., et al. (2021). Genetic control and phenotypic characterization of panicle architecture and grain yield-related traits in foxtail millet (Setaria italica). *Theor. Appl. Genet.* 134 (9), 3023–3036. doi:10.1007/s00122-021-03875-2

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

frontiers | Frontiers in Genetics

TYPE Review
PUBLISHED 25 January 2023
DOI 10.3389/fgene.2023.1053810

OPEN ACCESS

EDITED BY
Reyazul Rouf Mir,
Sher-e-Kashmir University of Agricultural Sciences and Technology, India

REVIEWED BY
Rutwik Barmukh,
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India
Sandhya Tyagi,
United Arab Emirates University, United Arab Emirates

*CORRESPONDENCE
Astha Gupta,
asthanbri@gmail.com
Rajendra Kumar,
rajendrak64@yahoo.co.in

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Plant Genomics, a section of the journal Frontiers in Genetics

RECEIVED 26 September 2022
ACCEPTED 05 January 2023
PUBLISHED 25 January 2023

CITATION
Gupta A, Sharma T, Singh SP, Bhardwaj A, Srivastava D and Kumar R (2023), Prospects of microgreens as budding living functional food: Breeding and biofortification through OMICS and other approaches for nutritional security. *Front. Genet.* 14:1053810. doi: 10.3389/fgene.2023.1053810

Prospects of microgreens as budding living functional food: Breeding and biofortification through OMICS and other approaches for nutritional security

Astha Gupta[1*†], Tripti Sharma[1†], Surendra Pratap Singh[2], Archana Bhardwaj[3], Deepti Srivastava[4] and Rajendra Kumar[5*]

[1]Sharda School of Agricultural Sciences, Sharda University, Greater Noida, India, [2]Plant Molecular Biology Laboratory, Department of Botany, Dayanand Anglo-Vedic (PG) College, Chhatrapati Shahu Ji Maharaj University,, Kanpur, India, [3]Molecular Biology and Biotechnology Division, CSIR-National Botanical Research Institute, Rana Pratap Marg, Lucknow, India, [4]Department of Agriculture, Integral Institute of Agricultural Science and Technology, Integral University, Lucknow, Uttar Pradesh, India, [5]Division of Genetics, ICAR-Indian Agricultural Research Institute, New Delhi, India

Nutrient deficiency has resulted in impaired growth and development of the population globally. Microgreens are considered immature greens (required light for photosynthesis and growing medium) and developed from the seeds of vegetables, legumes, herbs, and cereals. These are considered "living superfood/ functional food" due to the presence of chlorophyll, beta carotene, lutein, and minerals like magnesium (Mg), Potassium (K), Phosphorus (P), and Calcium (Ca). Microgreens are rich at the nutritional level and contain several phytoactive compounds (carotenoids, phenols, glucosinolates, polysterols) that are helpful for human health on Earth and in space due to their anti-microbial, anti-inflammatory, antioxidant, and anti-carcinogenic properties. Microgreens can be used as plant-based nutritive vegetarian foods that will be fruitful as a nourishing constituent in the food industryfor garnish purposes, complement flavor, texture, and color to salads, soups, flat-breads, pizzas, and sandwiches (substitute to lettuce in tacos, sandwich, burger). Good handling practices may enhance microgreens'stability, storage, and shelf-life under appropriate conditions, including light, temperature, nutrients, humidity, and substrate. Moreover, the substrate may be a nutritive liquid solution (hydroponic system) or solid medium (coco peat, coconut fiber, coir dust and husks, sand, vermicompost, sugarcane filter cake, *etc.*) based on a variety of microgreens. However integrated multiomics approaches alongwith nutriomics and foodomics may be explored and utilized to identify and breed most potential microgreen genotypes, biofortify including increasing the nutritional content (macro-elements:K, Ca and Mg; oligo-elements: Fe and Zn and antioxidant activity) and microgreens related other traits viz., fast growth, good nutritional values, high germination percentage, and appropriate shelf-life through the implementation of integrated approaches includes genomics, transcriptomics, sequencing-based approaches, molecular breeding, machine learning, nanoparticles, and seed priming strategiesetc.

KEYWORDS

biofortification, functional food, microgreen, nutrition, OMICS, shelf life

1 Introduction

In the past few decades, interest in organic and nutritional vegetables has gained momentum among people. That has increased the demand for sprouts and microgreens. Sprouts are considered germinated seeds that can be harvested before the growth of true leaves and consumed whole with the seed (Di Gioia et al., 2017). On the other hand, microgreens are defined as tender, immature greens that need light for photosynthesis, growing medium (soil or nutrient solution medium), and represent a 7–28 days growth cycle. Microgreens can be developed from the seeds of cereals, vegetables, legumes, and herbs, comprising two completely expanded cotyledon leaves with or without the appearance of a rudimentary pair of first true leaves (Xiao et al., 2012). The history of microgreen production can be traced back to the 1980s when it first seemed on the menu of chefs in San Francisco, California (United States Department of Agriculture, 2014). Following its popularity further, its cultivation began in the southern part of California in the 1990s, such that now microgreens are regarded as "functional foods" or "Superfoods."

Some cereals (rice, corn, oats, wheat, and barley) and legumes (chickpeas, beans, and lentils) can also be exploited for microgreens cultivation and production due to their nutritional values. Microgreens may be used for sweet and savory dishes as a garnish. They can complement the flavor, texture, and color of salads, soups, flatbreads, pizzas, and sandwiches (alternative to lettuce in tacos/burgers/sandwiches) due to a good proportion of nutritional values and some specific metabolites. People can also supplement microgreens to prepare smoothies, juices, and health drinks. Several value-added products can be synthesized by using fresh and dry microgreens in food science laboratories as one ingredient, for example, cookies, noodles, snacks, and chips are developed. Microgreens products would be an up-and-coming resource for vegetarians as a healthy lifestyle change and generation of employment for the population engaged in agriculture and food processing sectors.

Very few reports are available on the cultivation of legumes microgreens, for example, chickpea micro-greens (Sreenivasan, 2020). The small-seeded legumes promise unmarked resources of potential ingredients for the fortification of traditional or staple foods with bioactive compounds, nutrients, and minerals (Butkutė et al., 2018). Erosion of preserved seed nutrients was observed in chickpea during the process to stimulate seedling growth that improved nutritional value with digestible protein levels ranging from 18.96% to 28.69% in 6-day-old sprouts identifying the genotypes BG-1092, ICC-11378, JG-74 as potential resources for accumulating proteins and other nutritional components at the seedling stage (Kumar et al., 2022a). Therefore, the cultivation of microgreens can be easily practiced, and cheaper sources of budding superfoods in terms of nutrition and antioxidant properties with minimum food wastage (only root) during consumption may be obtained. Thus, in the present report, we have discussed the prospects of nutraceutical aspects, health benefits, growth, and cultivation practices for microgreen production and mentioned related strategies for the first time to overcome the limitations through the utilization/integration of various OMICS approaches.

2 Nutraceutical and health benefits of microgreens: The superfoods/functional foods

Nutrients deficiency may cause serious diseases and health related issues and appropriate proportions of micronutrients, macronutrients, flavonoids and polyphenols regulate immunity and prevent from several health threats viz; osteoporosis, pharmacotherapy, COVID-19 etc (Batiha et al., 2022; Martiniakova et al., 2022; Bansal et al., 2022a; Bansal et al., 2022b). Microgreens are partially mature greens that elicit their intense flavour, aroma, texture, and nutrient properties with sensory attributes and acceptance are discussed in this review. Desirable sensory qualities and intense flavours of these superfoods have gained acclamation to be consumed as salads, garnishes, *etc.* The presence of minerals, vitamins, and their precursors - ascorbic acid and several other bioactive compounds tocopherols, carotenoids, betaine, phenols, glucosinolates, phytosterolsetc in microgreens add to their health and nutrition-related functional aspects (Xiao et al., 2012; Kyriacou et al., 2019a)as presented in Table 1 and Table 2. The sensory qualities of microgreens are influenced by their chemical composition (Xiao et al., 2015). This study highlights the correlation between the concentration of total phenols and with overall acceptability of sensory attributes and acceptance in terms of sweetness, sourness, bitterness, and astringency. The sensory and nutritional qualities of microgreens also vary with the growing methods. The sensory attributes and nutritional content of microgreens grown hydroponically and in soil procured from a commercial and local farm were compared (Tan et al., 2020). Nutritionally microgreens grown by either method were comparable. However, a significant difference in vitamin C content was reported. Variations in both micro-minerals (iron, zinc, copper, and manganese) and macro-mineral content (calcium, phosphorus, sodium, magnesium, chloride, potassium, and sulfur),phytochemical profile, and antioxidant capacities have been found to vary with genotype rather than growth stage as studied in microgreens of four Brassicaceae genotypes-Komatsuna, Mizuna, Pak Choi and Mibuna (Kyriacou et al., 2021a). In addition, delaying harvest from the arrival of the first to second true leaf does not seem relevant for improving bioactive compounds in microgreens (Kyriacou et al., 2021a). Regulated feeding of nutrient solution (NS) to microgreens and fertigation treatment can also influence the composition of phytochemicals and antioxidant activity apart from growth and yield (Petropoulos et al., 2021). Moreover, spinach microgreens were assessed for the effect of nutrient deprivation (0, 5, 10, and 20 days) and fertigation treatment before harvesting (Petropoulos et al., 2021). This study reported that NS feeding for a longer duration of 20 days resulted in enhanced fresh yield and content of photosynthetic pigments, including chlorophyll, beta carotene, and lutein. In contrast, the concentration of minerals like Calcium (Ca), Potassium(K), magnesium (Mg), and Phosphorus (P) were found to be lowest after 20 days in contrast to being maximum in control and 5 days of NS feeding (Petropoulos et al., 2021). Feeding the spinach microgreens for 10 days with NS yielded the best combination of yield, minerals (high), and nitrate (low) while maintaining the concentrations of bioactive compounds (Petropoulos et al., 2021). Simultaneous to the accumulation of secondary metabolites, microgreens are also known to accumulate anti-nutritive compounds like nitrate. Therefore, nutrient deprivation before harvest (DBH) was employed to reduce nitrate levels by substituting NS with osmotic water for 6 and 12 days in a garden rocket, lettuce, and mustard microgreens grown on a peat-based substrate (Kyriacou et al., 2021b). Nutrient deprivation proved a good strategy for lowering nitrate content with effective treatment duration varying from species to species. Even nitrate hyper-accumulating species like garden rockets showed an abrupt decline in nitrate concentration (Kyriacou et al., 2021b). Further, abundant secondary metabolites like flavanol glycosides, quercetin,

TABLE 1 Assessment of microgreens studies in different plants.

S.No.	Crop species	Objective of study	Key findings	Conclusion/ Recommendation	References
Vegetables					
1	Cabbage, Kale, Argula and Mustard	To study yield and appearance quality in response to variation in blue light	Fresh and dry weight remained unaffected, however hypocotyl length and cotyledon area decreased	Blue light 15% and 5% were best for cabbage other three microgreens respectively	Ying et al.(2020a)
2	Kale (*Brassica napus* L. 'Red Russian'), mustard (*Brassica juncea* L. 'Ruby Streaks'), cabbage (*Brassica oleracea* L.), and arugula (*Eruca sativa* L.).	To study effect of Single Source (SS) LED on growth, yield and quality. Further, to develop mathematical models to understand these relationships	Increased fresh and dry weight with increase in light intensity while hypocotyl length and hue angle decreased linearly. Phenotypic plasticity exhibited by arugula and mustard were greater compared to kale and cabbage	Optimum Sole Source-LED light intensity for these four microgreens depending on genotype, production system and goal	Jones-Baumgardt et al.(2019)
3	Arugula, cabbage, mustard and kale	To study the effects of photoperiod shortening on elongation growth	Blue light promoted elongation was evident from length of petiole and rate of stem extension	Blue light promotes elongation growth for 16–24 h photoperiod and is beneficial for indoor production methods	Kong et al.(2019)
4	Mustard and arugula	To study the effect of treatment of Blue light and combined effect of Blue and Far-Red light during night on yield, plant quality and elongation	Under Blue light plant height increased by 34% and 18% in mustard and arugula respectively. Combination of B and FR light also gave similar results without compromising with yield and quality of microgreens under either treatment conditions	Treatment with blue light alone at night can promote elongation in microgreens while maintaining their yield and quality	Ying et al. (2020b)
5	Broccoli	To study effect of application of $CaCl_2$ pre-harvest and UV-B post-harvest on levels of Glucosinolates (GLS) and glucoerucin (GLE) i.e. storage quality of microgreens	Content of total aliphatic glucosinolates in microgreens was four times as compared to mature counterparts. Treatment with 10 mM $CaCl_2$ and UV-B also enhanced GLS levels	Spraying microgreens with $CaCl_2$ prior to harvest not only enhances levels of GLS it also improves visual appearance and storage/shelf life, alongwith UV-B exposure post-harvest	Lu et al.(2018)
Cereals					
1	*Triticum aestivum* (wheat) and *Hordeum vulgare* (barley)	Profiling content of chlorophyll and carotenoid during 7 and 16 days on dry basis and evaluate pigment accumulation rate	Content of chlorophylls and carotenoids had strong correlation with number of growth days and progressively increased up to day 16. Accumulation of pigments was maximum between day 7–10 in wheat and day 10–13 in barley	Cereal microgreens can be considered for *in vivo* studies for potential use in nutraceutical and pharmaceutical industry	Niroula et al.(2019)
Herbs—Medicinal/Culinary					
1	*Daucas carota* L. var. *New Kuroda, Foeniculum vulgare* Mill., *Trigonella foenum –graecum* L., *Ocimum basilicum* L., *Alium cepa* L. var. *Light Red Gavran, Hibiscus sabdariffa* L. (white var.), *Raphanus sativus* L. var. *Imp. Chetki, Spinacia oleraceae* L. var. *All Green, Helianthus annus* L., *Brassica juncea* L	To study the bioactive phytochemicals (Overall phytochemical composite index -OPCI)and overall antioxidant activity measured as Antioxidant potential composite index APCI using appropriate parameters), of ten culinary microgreens. Further, the phytochemicals contributing in antioxidant potential were also identified	Antioxidant potential and Phytochemical profile was reported highest in *Foeniculum vulgare* Mill. And *Hibiscus sabdariffa* L. microgreens. Total phenols and total flavonoids contributed maximum to OPCI, APCI and radical scavenging activity	Wholesome nutritional status of ten culinary microgreens was reported	Ghoora and Srividya, (2020)
Legumes					
1	Lentil and Mung bean	Analysis of diversity in phytochemical profile, antioxidant capacity and content of micro as well as macro nutrients in 20 genotypes of lentil and mung bean grown in plain and high altitude regions	L830 and MH810 genotypes of lentil and mung bean respectively, were identified as superior based on antioxidant activity, total flavonoids, ascorbic acid, carotenoids, and phenol content	Nutritional profiles of same genotypes showed variation when grown in two different altitude regions of Delhi and Leh –Ladakh	Mishra et al. (2021)

(Continued on following page)

TABLE 1 (*Continued*) Assessment of microgreens studies in different plants.

S.No.	Crop species	Objective of study	Key findings	Conclusion/ Recommendation	References
2	Black Gram, Chickpea and Mung bean	To analyse and compare the nutritional profile of these legumes cultivated in water, soil and cocopeat supplemented with nutrient solution	Mineral content, phenol proximate and amino acid composition, antioxidant activity showed variation in different species and substrate used for cultivation	Nutritional qualities of the crops varied with substrate used and crop species	Kaur et al. (2022)
3	*Trifolium pratense* *T. medium, Medicago sativa, M. lupulina, Onobrychis viciifolia, Astragalus glycyphyllos* and *A. cicer*	To determine the nutritional potential, phytochemical and mineral profile of forage legumes by analysing the seeds, sprouts and microgreens	All species exhibited high nutritional potential, phytochemical and mineral values. Mineral and protein levels were high in all three forms studied. Increase in quantity of protein and phytochemicals was observed from seed to microgreen stage, however trends were opposite for total carbohydrates	Small seeded legumes, (especially as microgreens) have the potential of being used for fortification to enhance nutrients as well as bioactive compounds in staple food	Butkutė *et al.* (2018)
4	*Vigna radiata* L. (Mungbean), *Lens culinaris* subsp. *culinaris* (lentil), and *Brassica juncea* L. (Indian mustard)	To optimize and evaluate yield, shelf-life, sensory parameters and microbial load in microgreens of these crop species	Optimum seed density (to obtain maximum yield), time of harvest was reported for three species. High correlation between seed size and yield for both legume species. No pathogenic bacteria were found in microbial load	Proper cultivation and storage of microgreens can aid in their safe human	Sangwan et al. (2022)
5	Alfalfa, Fenugreek, Lentil, and Daikon Radish	To determine polyamine content in seed, sprouts and microgreens of these crops and find the stage with superior quantity of polyamines. Also to determine the enzymatic capacity of sprouts to degrade unwanted biogenic amines	Polyamines of nutritional importance (spermine, spermidine and agmatine) were in abundance in microgreens. Cadaverine was highest in sprouts of legumes. On the other hand, nutritionally important polyamines were higher in sprouts of daikon radish than their microgreen counterparts	Microgreens are a rich source of nutritionally beneficial polyamines	KraljCigic et al. (2020)

and kaempferol glycosides were detected in Brassicaceae, and caffeoyl quinic acid in lettuce microgreens. However, total phenols increased in lettuce, reduced in the garden rocket, and unaffected in mustard microgreens in response to nutrient deprivation (Kyriacou et al., 2021b).

The presence of bioactive compounds renders microgreens as a health beneficiary, antioxidant,anti-microbial, anti-inflammatory, and anti-carcinogenic. The health-related properties of bioactive compounds in food and herbs depend not only on their content and the amount consumed but also on their bioavailability (de la Fuente et al., 2019). Quantity and bio-accessibility of bioactive antioxidant compounds (total anthocyanins, total soluble polyphenols, ascorbic acid, total isothiocyanates), antioxidant capacity (Trolox Equivalent Antioxidant Capacity, and Oxygen Radical Absorbance capacity), macro-elements (K, Ca and Mg) and oligo elements (Fe and Zn) have been evaluated in four hydroponic Brassicaceae microgreens-broccoli, radish, kale and mustard (de la Fuente et al., 2019). The optimum amount of nutrients (macro and oligo-elements) is required for proper growth metabolism, and in contrast, deficiency may lead to life-threatening diseases in extreme circumstances. The essential nutrients are widely distributed in foods, and most people can obtain sufficient amounts by consuming a varied diet. Moreover,macro-elements (K, Ca, and Mg) and oligo-elements (Fe,Se, Cu, and Zn) remarkably stimulate the function of the immune system and are helpful in cardiopulmonary bypass (Al-Bader et al., 1998). Microgreens of soybean, green pea, garden rocket, radish, and red Rambo radish were cultivated under fluorescent and LED light conditions. The variation in anti-proliferative/pro-oxidant efficiencies of these microgreens was studied using Ewing sarcoma lines RD-ES and A673 (Truzzi et al., 2021). It was observed that all microgreen extracts could reduce cell proliferation in 2-dimensional cell cultures, while extracts from pea microgreens grown under LED light showed anti-proliferative and pro-apoptotic activity on 3-dimensional A673 and RD-ES spheroids without showing cytotoxicity on healthy L929 fibroblasts (Truzzi et al., 2021). LED and fluorescent light illuminated Red Rambo radish also exhibited anti-tumor effects on RD-ES spheroids. Further, the effects of UV-A, B, and C on inducing polyphenol content and anti-tumor activities of UV-illuminated microgreens can be an area of exploration in the future. In a study, the effect of salinity in combination with different wavelengths of light in *Brassica carinata* extracts from microgreens grown under different treatments of salinity and light were checked for their ability to stimulate antioxidant enzymes, including catalase (CAT), superoxide dismutase (SOD) and expressions of Nrf2 (nuclear transcription factor-erythroid 2 related factor) and HO-1 proteins (heme-oxygenase -1)on human colorectal carcinoma cells-HCT116 (Maina et al., 2021). Activation of antioxidant enzymes (SOD and CAT) and stimulation of HO-1 and Nrf2 make them preferable for the prevention and treatment of oxidative stress and inflammatory disorders (Maina et al., 2021). Microgreens can be a powerful source of nutrients owing to higher concentrations of several phytochemicals than their matured counterparts. These can be used as supplements to overcome deficiencies of several nutrients. Being a rich source of bioactive

TABLE 2 Comparative analysis of nutritional components in microgreeens, mature part and different stages of plants.

Table-2				
Plant	Immature/Microgreen (7 days after seed sowing)	Seedling (15 days after seed sowing)	Mature (adult stage-30 days after seed sowing	References
Brassica rapa subsp. Chinensis var. Parachinensis (Choy sum)	**Essential amino acids**- 15.8	**Essential amino acids**- 7.8%	**Essential amino acids**-10.4%	Zou et al. (2021)
	Total concentration of sugars-0.01g/100 g FW	**Total concentration of sugars**-0.2 g/100 g FW	**Total concentration of sugars**- 0.2g/100 g FW	
	vitamin B9 (folate)	**vitamin B9** (folate)	**vitamin B9** (folate)	
	CHO-folate- 46 ± 5 μg//100 g FW	CHO-folate- 46 ± 4 μg/100 g FW	CHO-folate-31 ± 7 μg/100 g FW	
	Folic Acid- 1.6 ± 0.22 μg//100 g FW	Folic Acid-3.0 ± 1.26 μg//100 g FW	Folic Acid- 2.6 ± 1.57 μg//100 g FW	
	Tetrahydrofolate- 36 ± 6 μg// 100 g FW	Tetrahydrofolate- 34 ± 5 μg// 100 g FW	Tetrahydrofolate- 33 ± 9 μg// 100 g FW	
	5-methyltetrahydrofolate -23 ± 2 μg// 100 g FW	5-methyltetrahydrofolate- 32 ± 2 μg// 100 g FW	5-methyltetrahydrofolate-20 ± 3 μg// 100 g FW	
	Lipid soluble Vitamin A	**Lipid soluble Vitamin A**	**Lipid soluble Vitamin A**	
	β-Cryptoxanthin- 97 ± 19 μg// 100 g FW	β-Cryptoxanthin- 101 ± 18 μg// 100 g FW	β-Cryptoxanthin- 35 ± 48 μg// 100 g FW	
	Neoxanthin- 2,105 ± 279 μg// 100 g FW	Neoxanthin - 2,909 ± 490 μg// 100 g FW	Neoxanthin -2,243 ± 486 μg//100 g FW, Violaxanthin - 2,201 ± 602 μg// 100 g FW	
	Violaxanthin- 2040 ± 296 μg// 100 g FW	Violaxanthin - 4,336 ± 731 μg// 100 g FW		
	α-**Tocopherol & γ-Tocopherol:** no significant change at 3 stages			
	Vitamin K1- 377 ± 29 μg/100 g FW	**Vitamin K1**- 433 ± 33 μg/100 g FW	**Vitamin K1**- 363 ± 27 μg/100 g FW	
	Glucosinolates: Gluconapoleiferin-517 ± 125 μg/100 g FW, Gluconapin - 5,576 ± 1,431 μg/100 g FW	**Glucosinolates:** Gluconapoleiferin - 214 ± 40 μg/100 g FW, Gluconapin - 952 ± 257 μg/100 g FW	**Glucosinolates:** Gluconapoleiferin-85 ± 16	
			μg/100 g FW, Gluconapin - 3,488 ± 181 μg/100 g FW	
	Minerals (μg/100 g FW)	**Minerals** (μg/100 g FW)	**Minerals** (μg/100 g FW)	
	Copper (Cu)- 27 ± 4 Iron (Fe) - 504 ± 77, Magnesium (Mg)- 30,411 ± 5,705, Potassium (K)- 411,908 ± 46,579, Zinc (Zn)- 472 ± 80	Copper (Cu)- 21 ± 8, Iron (Fe)- 358 ± 77, Magnesium (Mg)- 22,504 ± 4,685, Potassium (K)- 229,953 ± 40,028, Zinc (Zn)- 412 ± 93	Copper (Cu) - 20 ± 5, Iron (Fe)-325 ± 48, Magnesium (Mg)- 26,838 ± 5,994, Potassium (K)- 222,630 ± 69,219, Zinc (Zn)- 280 ± 55	
Broccoli microgreens	Copper and selenium were 13.7 fold higher in microgreens as compared to the mature stage	-	-	Johnson et al. (2021)
Red cabbage microgreens	2.1 fold higher phosphorus, 2.4 fold higher iron, 3.8 fold higher zinc and 9.1 fold higher copper contents were explored in microgreens in comparison to mature stage	-	-	Johnson et al. (2021); Podse dek et al. (2006)
	Vitamin E (0.06 mg/100 g FW) was forty times higher in comparison to the mature red cabbage			
Red beet microgreens	10.2 fold higher selenium, 3.1 fold and 2. 5 fold higher chromium	-	-	Johnson et al. (2021)
Red amaranth microgreens	2.2 times higher concentration of copper than the fully developed plants	-	-	Johnson et al. (2021)
Pea microgreens	12.2 and 16.8 times higher contents of molybdenum and selenium were recorded in comparison to 38 days old pea plants	-	-	Johnson et al. (2021)

(Continued on following page)

TABLE 2 (*Continued*) **Comparative analysis of nutritional components in microgreeens, mature part and different stages of plants.**

Table-2				
Plant	Immature/Microgreen (7 days after seed sowing)	Seedling (15 days after seed sowing)	Mature (adult stage-30 days after seed sowing	References
Golden pea	Quantity of α-tocopherol and γ-tocopherol was significantly higher (4.9 mg/100 g FW and 3.0 mg/100 g FW respectively) than mature spinach leaves (α-tocopherol-2.0 mg/100 g FW; γ-tocopherol -0.2 mg/100 g FW)	-	-	Xiao et al. (2012)
Cilantro microgreens	5 times and 2.8 times more violaxanthin content was estimated as observed in the mature leaves of cilantro (1.4 mg/100 g FW) and spinach leaves (2.7 mg/100 g FW) respectively	-	-	Bunea et al. (2008); Kobori and Amaya, (2008)

compounds like carotenoids, phenols, glucosinolates, polysterols, and many others, microgreens possess health benefits due to anti-microbial, antioxidant, anti-inflammatory, and anti-carcinogenic properties (Choe et al., 2018; Le et al., 2020).

Most of the microgreens viz., broccoli, red beet, red amaranth, red cabbage and pea microgreens exhibit high proportion of minerals and nutritionally rich components (copper, selenium, phosphorus, iron, zinc, molybdenum and chromium) as compared to mature and other counterparts (Johnson et al., 2021) as presented in Table 2 and are being narrated as given further.

3 Comparison for nutrients and phytochemicals in microgreens and different stages of crop plants

The comparative nutrient profiling was explored in *Brassica rapa* subsp. Chinensis var. Parachinensis (Choy sum) at three different growth stages *viz.* Microgreen, seedling and adult/mature stages (7, 15 and 30 days after seed sowing) respectively (Zou et al., 2021). The content of essential amino acid was high at microgreen stage (15.8%) as compared to seedling (7.8%) and mature (10.4%) stages. Moreover, the contents of metabolites and minerals were estimated along with their changes at maturity in six crops microgreens belonging to three distinct families Brassicaceae (*Eruca sativa* L.) Cav. - Arugula, *Brassica oleracea* L. Italica Group-Broccoli and *Brassica oleracea* L. Capitata Group - Red cabbage), Fabaceae (*Pisum sativum* - Pea) and Amaranthaceae—(*Amaranthus tricolor* L. -Red amaranth and *Beta vulgaris* L. Crassa Group- Red beet) (Johnson et al., 2021). Moreover, metabolites showed significant difference ($p < 0.05$) and were ≥2- times greater in microgreens as compared to their mature counterparts (Johnson et al., 2021). There are 95 such metabolites in broccoli, 110 in red cabbage, 87 in arugula, 80 in red beet, 93 in pea and 101 in red amaranth microgreens, respectively. These metabolites were majorly classified as peptides, saccharides, nucleotides, amines, phenolics, lipids, organo-sulfurs, alkaloids and vitamins with prevalence of lipids and phenolics (Johnson et al., 2021). In another study, good quantities of α-tocopherol and γ-tocopherol were estimated in golden pea tendrils as 4.9 mg/100 g FW and 3.0 mg/100 g FW respectively (Xiao et al., 2012). Interestingly, cilantro microgreens and red cabbage microgreens have been reported to have more violaxanthin and vitamin E contents as compared to their mature counterparts, respectively (Podse dek et al., 2006; Bunea et al., 2008; Kobori and Amaya, 2008).

Further, now it is a proven fact that the red cabbage microgreen regulates the levels of lipids, cholesterols in human blood vessels and protects against cardiovascular diseases (Huang et al., 2016). The mustard and coriander microgreens have been reported to be very rich sources for their antioxidant, antimicrobial, anti-cancerous, anti-obesity, anti-inflammatory and antidiabetic activities (de la Fuente et al., 2020; Le et al., 2020; Saengha et al., 2021; Truzzi al., 2021; Dhakshayani and Alias 2022). Hence, appropriate daily consumption of microgreens would be a potential protective nutritional strategy to manage health, nutrition and chronic degenerative diseases (de la Fuente et al., 2020).

4 Microgreens: A novel, live, super functional food

Microgreens' demand and preference are attributed to their aroma, tender texture, vivid colour, flavour, sensory attributes, and quick production or cultivation. Several microgreen species have peculiar colours; for example, microgreens of broccoli, spinach, and celery have been reported to be green in colour, microgreens of radish, red basil, and red cabbage are crimson, while multicolour microgreens have been developed in mustard and beet (Di Gioia et al., 2015). Some of the previous studies related to microgreens have been summarized in Table 1. Similarly, microgreens are also known to possess a distinct flavour viz; spinach and rapini taste-neutral, arugula, radish, and watercress spicy and *Cucurbitaceae* microgreens as bitter have been reported (Di Gioia et al., 2017).

Most of the species and varieties used in current microgreen production come from the Brassicaceae and Amaranthaceae families (Xiao et al., 2015; Kyriacou et al., 2016). In the Amaranthaceae family, some of the more popular species, subspecies, and varieties include beet, chard, and amaranth; in the *Brassicaceae* family, radish, broccoli, kale, cabbage, tatsoi, pakchoi, mizuna, arugula, and mustard. Microgreens of grain crops such as buckwheat, wheat, and rye have also been grown. Various medicinal and culinary herbs have also been used for microgreen production, including borage (or starflower), parsley, basil, and fenugreek, among many others (Verlinden, 2020). Microgreens' taste can fluctuate

significantly based on the variety. The widely used varieties have been explored for microgreens production from seeds of diverse plant families as mentioned further;Brassicaceae (Cauliflower, broccoli, cabbage, watercress, radish, and arugula), Apiaceae (Dill, carrot, fennel, and celery), Asteraceae (Lettuce, endive, chicory, and radicchio), Amaranthaceae (Amaranth, quinoa swiss chard, beet, and spinach), Amaryllidaceae (Garlic, onion, and leek) and Cucurbitaceae (Melon, cucumber, and squash).

Increasing interest in the production and consumption of microgreens is also due to their high nutritional content, high yield, rapid production, aroma, and other qualities. Their high nutritional qualities are mainly due to the presence of phytochemicals and other bioactive compounds, along with their antioxidant capacities. They are considered highly nutritious food because of the presence of nutrients that include proteins, minerals, vitamins, carotenoids, phenols, and glucosinolates (Ebert, 2013; Di Gioia et al., 2017). The concentrations of bioactive compounds found in microgreens and even sprouts are reported to be much higher than their mature counterparts (Kyriacou et al., 2016). For example, Broccoli microgreens grown hydroponically and in compost were found to have more nutrient content (Mn, Cu, P, K, Na, Mg, and Fe) than mature broccoli vegetables (Weber, 2017). Contrary to this, hydroponically grown fenugreek (*Trigonella foenum-graecum* L.), broccoli, and garden rocket (*Eruca vesicaria subsp. Sativa*) microgreens were reported to have lower mineral contents than their mature plants to be eaten as vegetables. Among the three, only fenugreek microgreens efficiently uptake iron in caco-2 cells (Khoja et al., 2020). Chickpea microgreens also contains a good amount of zinc, calcium, iron, antioxidants, vitamins, carbohydrates, fiber,fat, and high protein content. Moreover, enriched nutrient contents containing plants may produce microgreens with high biomass in a limited time in a cost-effective manner that would improve the cultivation and yield of microgreens.

5 Growth and cultivation practices for the production of microgreens

The basic requirements for microgreen cultivation are the availability of substances and the effect of light as narrated below.

5.1 Availability of substrate for cultivation of microgreens

Various substrates have been used to grow microgreens, and their influence on yield and nutritional quality has been studied. In a study, three different substrates - vermiculite, cotton, and jute fiber were used to grow microgreens of green basil—*Ocimum basilicum* L., Red basil—*Ocimum basilicum* var. Purpurecsens and garden rocket *Eruca sativa* Mill in a Micro Experimental Growing System (MEG) fitted with LED lamps for light supply. In addition to that, several other substrates are also available to use further as primary medium or in combinations, for example, coco peat, coconut fiber, coconut coir dust, coconut husks, sand, jute fiber, vermicompost, sugarcane filter cake, peat and white sphagnum peat substrates presented in Table 3. A high yield of 2–3 kg/m2 was obtained. The three microgreens varied in nutritional quality, with red basil accounting for high antioxidant compounds on vermiculite and jute fiber media. At the same time, the qualitative parameters were found to be species-dependent (Bulgari et al., 2021).

5.2 Effect of light on nutritional quality and growth of microgreens

Essential growth factors like light (wavelength, intensity, and photoperiod) also influence microgreens' biosynthesis and accumulation of phytochemicals (Delian et al., 2015). Recently, many studies have been carried out on the effect of artificial light sources like halogen lamps, high-pressure sodium lamps, fluorescent lamps, and LED lights on the growth, yield, nutrient quality, and phytochemical content of microgreens. Many studies have focused on the benefit of illumination with LED light on plant growth, quality, and accumulation of phytonutrients. Some of these studies have been tabulated (Table 4). The effect of different wavelengths of light in combination with salinity was assessed on quality in terms of antioxidant capacity, the content of phenolics and glucosinolates, as well as yield of microgreens of *Brassica carinata* (Maina et al., 2021). Stable cultivation of microgreens was achieved under fluorescent and blue plus red (B1R1) light conditions, which resulted in a high accumulation of biomass and glucobrassicin. Under saline conditions, blue LEDs and fluorescent light promoted antioxidant activity together with the accumulation of phenols, sinigrin and glucosinolate (Maina et al., 2021).

6 Limitations of microgreen production

The consumption and storage of microgreens pose several limitations. Poor shelf life and postharvest management of microgreens are major challenges for the concerning researchers. It is important to know if a correlation exists between produced spoilage and contamination by the human pathogen. Studies have reported seed-to-harvest pathogen infection in several microgreens like basil, lettuce, parsley, melons, and spinach (Alegbeleye et al., 2018). Growing conditions that promote microorganism growth or transfer, processing practices that expose the commodity to contaminants from animals or humans, and physiological characteristics of the plant that allow contact and binding with microorganisms, all of these factors put crops at risk (Maluin et al., 2021). As microgreens are grown in a controlled environment, they are unexposed to external agents like pests and insects. Also, there is minimal or no contamination due to almost no external application of fertilizers or manures. Strategies concerning postharvest management and enhancing the shelf life of microgreens need to be particularly developed.

7 Strategies to overcome limitations of microgreen production

There are various strategies to encounter the limitations related to microgreen cultivation and production as presented through Figure 2 and discussed below.

TABLE 3 Different growth media/substrate used in different plants for microgreens cultivation.

Crop/Plant	Substrate used	Remarks/Findings	Reference
Ocimum basilicum L.- Basil, *Eruca vesicaria* (L.) Cav. subsp. *Sativa* (Mill.) Thell.- Rocket	Hydroponics (Soil-less medium)	High concentrations of some minerals	Bulgari et al. (2017)
Eruca sativa Mill. - Rocket, *Ocimum basilicum* L. - Green Basil, *Ocimum basilicum* var. Purpurecsens - Red basil	Vermiculite, coconut fiber, jute	Substrate significantly regulates nitrate concentration, yield and dry matter percentage	Bulgari et al. (2021)
Hairy basil (*Ocimum basilicum* L.f. var. *citratum* Back),	Sand, vermicompost, coconut coir dust, sugarcane filter cake, peat	Local organic biomaterials were identified as suitable substitutes to costly peat-based media for cultivating microgreens.	Muchjajib et al. (2015)
Sweet basil (*Ocimum basilicum* Linn.),			
Holy basil (*Ocimum sanctum* Linn.),			
Huanmoo (*Dregea volubilis* Stapf),			
Sano (*Sesbania javanica* Mig.),			
Vine spinach (*Basella alba* Linn.),			
Rat- tailed radish (*Raphanus sativus* var. caudatus Linn),			
Leaf mustard (*Brassica juncea* Czern. & Coss.),			
Kangkong (*Ipomoea aquatica* Forsk.)			
Krathin (*Leucaena leucocephala* de Wit.),			
Red radish (*Raphanussativus*) var "Sango"	White sphagnum peat substrate, Coco coir dust	Microgreens grown on these substrates had permissible levels of nitrate content and microbial growth.	Thuong and Minh (2020)

7.1 Selection and validation of potential crops/genotypes for microgreen traits

Identifying a diverse collection of genotypes is very important to explore the most promising genotypes containing microgreens-related traits viz., high nutrition, shelf life, sensory attributes, acceptable taste, and yield. However, microgreens' phytoactive compounds, antioxidant capacity, shelf life, and nutrient content depend on genotypes' genetic makeup and environmental conditions. For instance, twenty diverse genotypes of lentil and mungbean were grown as microgreens in plain and high altitude two regions (Delhi and Leh Ladakh). The investigation for profiling of phytochemical, macro and micronutrients content along with antioxidant capacity was accomplished (Mishra et al., 2021). Based on phytochemical profiles, lentil genotype L830 and mungbean genotype MH810 were identified as superior to other genotypes for contents of total flavonoids, carotenoids, ascorbic acid, antioxidant parameters, and phenols. The difference in nutritional profiles of identical genotypes under different environmental conditions was observed, probably due to variable gene expression. However, the genes and pathways governing the variation in response to different environmental conditions are yet to be elucidated. This study has provided new insights into the potential application of microgreens in harbouring the nutritional security of inhabitants of harsh environments such as the high altitudes of Leh-Ladakh (Mishra et al., 2021). Some other legume microgreens viz., sainfoin, red clover, alfalfa, chickpea, lentils, maize, cowpea and mung bean were tested for phenolic, antioxidant activity, flavonoid, carotenoid, ascorbic acid, total chlorophyll, chlorophyll a, and chlorophyll b concentrations The highest total antioxidant activity (TAA: 4,789.373 mg TE g-1), total phenolic contents (TPC: 791.770 mg GAE 100–1 g-1) in red clover and highest total flavonoid content (672.177 mg QE 100 g-1) in maize were estimated (Altuner et al., 2022). Thus, it was concluded that total phenolic contents (TPC) was considerable good in red clover (46%) and maize (73%) as compared to cowpea and other studied legumes (Altuner et al., 2022). Although, biochemical parameters have not been correlated in legumes but pigment parameters were positively associated in legumes and cereals (Altuner, 2021). Further, positive correlation was found for total antioxidant activity, total phenolic content and total ascorbic acid in cereals (Altuner, 2021).

Variation in phytochemical and antioxidant profile and macronutrient composition of red and green butterhead lettuce cultivars (*Lactuca sativa* L. var. *Capitata*–green and red Salanova) has been reported with varying stages of development at harvest (El-Nakhel et al., 2020). Though microgreens of both cultivars were rich in calcium and magnesium, the red Salanova microgreens were concluded to be highly nutrient enriched. The nutritional profiles of Chicory and lettuce microgreens of the Asteraceae family and two genotypes of *Brassica* (broccoli) were compared. However, the *Brassica* microgreens were the richest source of phenols and vitamin E; the Asteraceae microgreens were rich in carotenoids and alpha-tocopherol (Paradiso et al., 2018).

These aforesaid studies suggest that microgreens with desirable nutritional contents can be obtained by exploring and manipulating the available genetic diversity.

7.2 Good agricultural practices (GAP) and optimized storage conditions for pre/post-harvest management to prolong the shelf life/quality of microgreens

Microgreens are preferred to be consumed afresh, either wholesome or as garnishes or seasonings and its cultivation at

TABLE 4 Influence of light on growth and phytochemical quality of microgreens.

S.No	Microgreen	Light treatment	Trait/Aspect studied	Key findings	References
Vegetables					
1	*Brassica juncea* 'Red Lace' (Mustard) and Brassica napus'Red Russian' (kale)	Blue:red light (LED)	Mineral nutrient content	Increase in blue light decreased elongation and enhanced accumulation of micro as well as macro nutrients	Brazaitytė et al.(2021)
2	Broccoli microgreens	Red:blue:green (1:1:1) LEDs	Growth and phytochemical content	Increased fresh weight, dry weight and moisture content, further, elevated chlorophyll and reduced carotenoid content with increasing light intensity. Contents of some other phytochemicals like vitamin C, soluble proteins and sugar, flavonoid, free amino acid, and glucosinolates except progoitrin also increased	Gao et al.(2021)
3	*Amaranthus tricolor* L. and *Brassica rapa* L. subsp. *oleifera*	White LED, Blue LED and Red LED	Yield and nutritional quality	Blue light is most effective in promoting growth and nutritional qualities. Red light had pronounced effects on accumulation of fresh biomass as well as growth of hypocotyl	Toscano et al.(2021)
4	Chinese Kale	Red, white, Blue LED and sunlight (control)	Growth and antioxidant system	Low intensity red light increased fresh weight ang hypocotyl growth. White LED promoted accumulation of phenolic compounds, glucosinolates and ascorbic acid	Tantharapornrerk et al.(2021)
5	*Eruca sativa* L.- Arugula, *Brassica oleracea* L. var. Capitata f. rubra - red cabbage, Brassica napus L. subsp. napus var. Pabularia - 'Red Russian' kale, Brassica juncea L. - 'Mizuna' mustard	LEDs supplying blue (5%–30%) and red light (70%–95%)	Phytochemical profiles	20% blue light enhanced ascorbate levels (both reduced and total) in arugula, mustard as well as kale microgreens. 30% blue light stimulated accumulation of phenols in Kale and mustard. Total anthocyanin content showed proportional increase with the % of blue light supplied up to 30 percent in all microgreens, except mustard	Ying et al.(2021)
6	Leafy vegetable amaranth and Red amaranth	LEDs	Growth and nutritional value	Red + blue in ratio 70R:30B (PPFD-280 μmol/m^2/s^1; Photoperiod- 16 h) improved fresh yield, vitamin C, content of photosynthetic pigments (carotenoids and chlorophylls), anthocyanins, and levels in both red amaranth and leafy vegetable amaranth microgreens. Further, total antioxidant capacity was also increased	Meas et al.(2020)
7	Amaranth, cress (edible herb), mizuna, purslane	Red, Blue, Blue-Red	Differences in productivity levels, polyphenolic and antioxidant profiles together with content of mineral–carotenoid	Higher nitrate accumulation, Increased concentrations of Na and K, while decreased calcium and magnesium concentration. Enhanced lipophilic antioxidant activity, β-carotene and lutein. Decreased polyphenolic content	Kyriacou et al.(2019b)
8	*Brassica oleracea* var. *gongylodes* (kohlrabi), *Brassica rapa* var. Japonica (mizuna) and *Brassica oleracea* (broccoli) *Brassica oleracea* var. gongylodes—Kohlrabi, *Brassica rapa* var. *Japonica*—Mizuna and *Brassica oleracea* - Broccoli	1^0 light spectrum comprising of red light (638 and 665 nm), far-red light (731 nm) and blue light (447 nm), or supplemented by yellow (595 nm), green light (520 nm), or orange (622 nm), LED source	Nutrient levels	Metabolic changes resulted in increase in essential nutrients like Iron, Magnesium, Calcium, beta carotene, soluble carbohydrates, ascorbic acid	Samuolienė et al.(2019)

(Continued on following page)

TABLE 4 (*Continued*) Influence of light on growth and phytochemical quality of microgreens.

S.No	Microgreen	Light treatment	Trait/Aspect studied	Key findings	References
9	Radish	White, Blue, UV-A and dark; light conditions combined with Hydrogen rich water (HRW)	Anthocyanin accumulation	Blue light and UV-A combined with HRW resulted in higher phenolic content. Increased content of anthocyanin compounds	Zhang et al.(2019)
10	Mustard, Beet and Parsley	Blue light treatment	Carotenoid and tocopherol content	Quantity of chlorophylls, carotenoids, alpha-carotenes and beta-carotenes, zeaxanthin, violaxanthin, and lutein, increased 1.2 to 4.3 folds	Samuolienė et al.(2017)
11	*Brassica oleracea* var. *gongylodes*—Kohlrabi, *Brassica juncea* 'Garnet Giant´- mustard, and *Brassica rapa* var. *Japonica* - mizuna	SS- LEDs; % Ratios as follows: Red_{87}:$Blue_{13}$, Red_{84}:Far Red_7: $Blue_9$,orRed_{74}:$Green_{18}$:$Blue_8$	Phytochemical synthesis	Total carotenoids in mizuna and mustard microgreens lowered, as light intensities increased. Higher values of total integrated chlorophyll were observed in kohlrabi at Red_{87}: $Blue_{13}$compared to mustard microgreens at Red_{84}:Far Red_7: $Blue_9$ and Red_{74}:$Green_{18}$: $Blue_8$.Total concentration of anthocyanins also increased as light intensity increased	Craver et al.(2017)
12	*Brassica rapa* var. *Chinensis*—Red pakchoi, *Brassica juncea* L.—Mustard and *Brassica rapa* var. *rosularis*—Tatsoi	Pulsed LED	Phytochemical levels	Total phenolic content decreased in response to pulsed LED while total anthocyanin content increased	Vasõtakaitė et al. (2017)
Herbs—Medicinal/Culinary					
1	Chia (*Salvia hispanica* L.) (Dark grown)	Constant light (100 µmol photons/m^2/s) for 24h and 48 h	Antioxidant activity and metabolic profile	Significant increase in antioxidant activity, chlorophyll and carotenoid synthesis, total soluble phenols and ascorbic acid content	Mlinarić et al.(2020)
2	*Ocimumbasilicum* L. (Acyanic and cyanic basil)	Light Emitting Diodes - Red andBlue	Microgreen morphometric parametres and bioactive compounds	Blue light illumination affected growth parameters resulting in increased cotyledon surface area, fresh weight, anthocyanin concentration and chlorophyll levels. Red light triggered synthesis of phenols and capacity to scavenge free radicals in green cultivar while in red cultivar blue light was found effective for same	Lobiuc et al.(2017)
Legumes					
1	Soybean	LED light spectra	Growth, antioxidant capacity, phenolic compounds profile	Decreased seedling heightas well as yield while increase in phenols. UV-A and Blue light significantly increased antioxidant capacity, total phenols and total flavonoid content	Zhang et al. (2019)

home can be an excellent practical approach concerning price and sustainability that also provides fresh life and functional food on the table for growing kids and families. Maintaining food safety standards with minimal or no microbial contamination that can cause potential health hazards without compromising the sensory qualities is of utmost importance.

Microbial contamination can cause spoilage of food and products, rendering them unfit for sale and consumption. This requires adopting hygienic and healthy practices throughout the cycle, beginning from cultivation until reaching the end user. Traditional methods like chilling, freezing, pasteurization and antimicrobial compounds (chemical or biological) compromise the sensory attributes (Tropea et al., 2021). Research is now focused on improving the quality and safety of food while maintaining its nutritional and organoleptic properties. One such technology is nisin-containing nano-carriers that can be applied safely as antimicrobial agents on food products (Bahrami et al., 2019).

Microgreens have a high respiration rate at the time of harvest that affects their shelf-life and storage (Chandra et al., 2012). Hence, application of 10 mM calcium chloride to microgreens prior to harvest is effective in delaying senescence, enhancing the visual appearance, and diminishing the growth of microorganisms during storage in broccoli microgreens (Kou et al., 2014). Further, Lu et al. (2018) studied the effect of applications of $CaCl_2$ as pre-harvest and

UV-B as post-harvest on levels of Glucosinolates (GLS) and glucoerucin (GLE) for assessing and enhancing the storage quality of microgreens. It was found that the treatments with 10 mM $CaCl_2$ followed by UV-B enhanced GLS levels andcontent of total aliphatic glucosinolates in microgreens was four times as compared to mature counterparts. These microgreens had increased biomass, calcium content, and activities of antioxidant enzymes superoxide dismutase and peroxidase. Overall shelf life, productivity, and post-harvest of microgreens were improved.

The shelf life of microgreens is a very important concern that varies from 10–15 days after harvesting, depending on the category of microgreens. Microgreens are potential plant-based food/diet full of nutrition, fibers, and antioxidants, which diminish the risk of cardiovascular disease and numerous types of cancer. The microgreens may be packed in polypropylene bags and stored at 5°C in a climate chamber or incubator for 10 days with controlled temperature and humidity. However, the 1°C storage temperature was optimum due to no chilling injury (Xiao et al., 2014). A combination of pre-harvest and post-harvest treatments, different packaging materials, and modified atmosphere packaging (MAP) regulates the shelf-life of fresh-cut microgreens and diverse sensorial characteristics. Moreover, macro-perforated packaging, including PET clamshell and LDPE self-seal bags, was also assessed for longer shelf life in radish and roselle microgreens (Ghoora and Srividya, 2020).

After harvesting, packets of microgreens should be kept at a 4–5°C and consumed within 8–10 days. Storage conditions and maintenance of shelf-life are very important to preserve the microgreens in good quality with stable nutrition. Several factors, viz., storage temperature, atmospheric composition, post-harvest light exposure, and packaging technologies, are associated with conserving fresh-cut microgreens. Further, during value-added product development, processing avenues (freezing, drying, waving, microwaving, frying, toasting, and boiling)are equally required to maintain the bio-availability of bioactive and phytochemical components of microgreens.

Different technologies and methods have been explored to maintain microgreens' shelf life and postharvest quality for preparing ready-to-eat products through wash steps and foliage spray. Aloe vera gel-based pre-harvest spray treatment and postharvest dip coatings were tested in radish and roselle microgreens for extended shelf life due to regulation of stomata closure (Ghoora and Srividya, 2020). These procedures regulate lower physiological weight loss, respiration rate, electrolyte leakage, microbial counts, and good overall acceptability. Further, researchers concluded that aloe vera gel-based-coated microgreens exhibit minimum deteriorative postharvest changes and higher ascorbic acid content than the uncoated control. Preharvest 10 mmolL−1 CaCl2 spray without postharvest dip displayed good yield, visual quality, and extended storage life (Kou et al., 2015). The optimum quality and highest shelf life of buckwheat microgreens can be maintained and stored at 5°C through moderately high O_2 (14.0–16.5 kPa) and low CO_2 (1.0–1.5 kPa) content with the treatment of chlorinated wash to reduce microbial counts (Kou et al., 2013). A very interesting plant regulator is 1-methyl cyclopropane (1-MCP) which binds competitively to ethylene receptors and delays senescence resulting in active treatment to prolong the shelf life of fruits, vegetables, and edible flowers (Turner et al., 2020). However, washing treatment is equally important in maintaining prolonged shelf life and the least microbial load. Comparative analysis of treated (chlorine wash) and controlled (unwashed) Ruby radish microgreens determines that 100 ppm chlorine wash enhances the visual quality and reduces electrolyte leakage (Turner et al., 2020). In addition, 0.25%–0.50% citric acid wash followed by 50% ethanol spray and 0.25% ascorbic acid is also effective in augmenting quality score. Further, to enhance the shelf life, a potential application of "nano packaging" technology concerning microgreens can also be explored for effective postharvest management. Thus, in coherence with the farm-to-fork tradition, good agricultural practices and handling practices are crucial.

7.3 Fortification through agronomic approaches, nano-technology and seed priming for enhancing preferred qualities of microgreens

Microgreens are considered potential nutrient sources that can help overcome the deficiency of many nutrients which are not met up with the seeds or mature parts of the plant. Effective fortification strategies for producing microgreens with desired nutritional traits and shelf life can be effective tools. Insufficient availability of iron and zinc in the human diet has posed a risk of malnutrition in young children and women. To address the deficiency of iron and zinc micronutrients, microgreen produce fortification can serve as an effective but short-term approach. Fortification can be done through several approaches such as agronomic practices, application of nanotechnology, *etc.*

Agronomic practices are cheap and simple but non-heritable and must be done with great care due to the application method, kind and environmental considerations. This strategy emphasizes improved nutrient accessibility to plants, efficient usage of nutrients, plant mobility, and increased microbial activity. Microbes like *Bacillus*, *Rhizobium*, *Azotobacte*r, *Actinomycete,* and some fungal strains, i.e., *Pseudomonas indica*, are used to increase nutrient availability and their uptake. Mineral nutrients show great potential for fortification when applied to the soil and the leaves. The most popular fertilizer is based on nitrogen, phosphorus, and potassium (NPK), which is vital for the health of both plants and mankind. Crops also require other micro-minerals such as iodine, zinc, copper, iron, nickel, molybdenum, manganese, *etc.*

In a recent study, the fortification of Brassicaceae microgreens was attempted for iron and zinc enrichment (Di Gioia et al., 2019). It involved growing red cabbage, red mustard, and arugula microgreens in nutrient solution supplemented with sulfate salts of iron and zinc at 0, 10, 20, and 40 mg L^{-1} and 0, 5, 10, and 20 mg L^{-1} concentrations, respectively. Further, investigations on the growth, yield, and mineral composition of these microgreens grown in these media composition exhibited accumulation of both iron and zinc minerals in microgreens of all three Brassicaceae members in a genotype-specific manner. Thus, this study also indicated that soil-less cultivation systems could be exploited for the production of fortification of microgreens by altering the composition of the nutrient medium. Pannico et al. (2020) have fortified selenium in green basil, purple basil, coriander, and tatsoi microgreens through a modified quarter-strength Hoagland nutrient solution with sodium selenite compound.

Nanofortification is the approach to fortify the plants or microgreens through nanoparticle application of some essential nutrients (Cu, Se, Fe, and Zn) in the form of liquid treatment as foliar and nano-fertilizers in soil or water medium (El-Ramady et al., 2021). Nanoparticles (NPs) are small materials ranging from 1 to 100 nm in size or dimension (Laurent et al., 2008), with a large surface area that allows its application in diverse fields, including fortification in plant systems. Their size and large surface area: volume ratio also

contributes to their physical and chemical properties. Due to their size, the optical properties of these particles impart unique characteristic colours. Their size, property, and shape are categorized into different groups: metal NPs, ceramic NPs, polymeric NPs, and fullerenes. They find applications in several fields, including environmental engineering, biotechnology, textiles, food processing/packaging, cosmetics, plant sciences, and agriculture. Several methods are employed for the synthesis and detection of nanoparticles. The techniques used for synthesis include chemical synthesis, thermal decomposition, photo-reduction, and green synthesis. Further, characterization of synthesized NPs may be performed using UV-Vis spectroscopy, X-ray diffraction assay, Scanning electron microscopy (SEM), transmission electron microscopy (TEM), and energy dispersive analysis. The conventional chemical methods of NP synthesis are costly, and toxic chemicals used for synthesis also pose a hazard to the environment. These pave the way for the need to synthesize NPs from biological methods using plants, microorganisms, and enzymes. These methods are not only cost-effective and rapid but are also environment-friendly and safe. Several metal NPs have been synthesized using the green synthesis approach, like silver NPs, gold NPs, copper, zinc oxide, and iron oxide (Bibi et al., 2022; Eltaweil et al., 2022; Nguyen et al., 2022). These may be utilized for seed priming to improve germination, seedling growth, and nutrition. Several plant parts like roots, leaves, flowers, fruits, and stems have been used to synthesize NPs *via* green methods. For example, zinc oxide nanoparticles are stable oxides of metal, eco-friendly, and have no harmful effects on humans and animals. These are most interesting to researchers due to their magnetic, optical, thermal, and chemical properties. ZnO NPs also exhibit adsorption ability which increases the catalytic efficiency. Nanoparticles have been applied extensively in agriculture research and may be used as nano-fertilizers to promote plant growth.

Seed priming is an innovative and user-friendly approach to fortify seeds by treating with an appropriate number of desirable nanoparticles. Sundaria et al. (2019) investigated the effect of iron oxide nanoparticles (25–600 ppm) on wheat genotypes (WL711 and IITR2). They observed increased germination percentage, shoot length, growth parameters, and accumulation of grain iron in WL711 and IITR2 at 200ppm and 400ppm, respectively. In addition, cold plasma (CP) treatment is a pollution-free way to improve seed germination, water use efficiency, nutrient uptake, photo- and thermo-dormancy, and plant yield (Mahanta et al., 2022). Moreover, CP treatment would be a potential approach to improve microgreens performance as it plays a critical role in numerous physiological, biological, and developmental processes in plants that improve seed performance, bacterial load on seeds, altering seed coat structures, enhance seedlings growth and its association with machine learning is a sustainable approach for seed priming (Shelar et al., 2022).

7.4 OMICS and breeding approaches for microgreen biofortification

Microgreens contain various favourable attributes like a pleasing palette of colours, quality, textures, and flavours (Aroma volatiles associated with flavour) but limit their commercial use due to short shelf life. NASA scientists have also explored microgreens in space due to dynamic properties like the availability of oxygen generation, nitrogen, essential nutrients, and photoactive compounds to enhance the morale of astronauts during stretched stays away from Earth (Kyriacou et al., 2017). Several breeding approaches and multiOmics have been explored to augment the shelf life, developmental rate and nutrient content of vegetables and fruits (Kalia and Singh, 2018; Mathiazhagan et al., 2021; Reda et al., 2021; Valdes et al., 2021; AieseCigliano et al., 2022; Chakravorty et al., 2022; Dutta et al., 2022; Kumari et al., 2022; Parmar et al., 2022; Sahoo et al., 2022; Sharma et al., 2022). However, it has been observed in tomatoes that the dominance component was higher than the additive component for shelf life (Pavan and Gangaprasad, 2022). To maintain the shelf life with natural colour and flavour, an antisense gene was introduced in tomatoes by a Californian company Calgene in the 1980s and developed improved shelf-life tomatoes, i.e., popularly known as FlavrSavr tomatoes (Bruening and Lyons, 2000). In addition to that natural variant of cucumber fruit (DC-48:high shelf life) was also explored through qRT-PCR for fresh green colour and shelf life and found that Expansin (EXP), Polygalacturunase (PG), and xyloglucan endotransglucosylase linked with cell wall degradation process and regulates to maintain the fruit firmness (Pradeepkumara et al., 2022). Several such efforts are underway in several vegetable and fruit crops.

The term "biofortification" refers to the process of enhancing the nutritional value of a plant's edible parts. It provides a long-term and sustainable alternative for supplying people with micronutrient-rich crops (Garg et al., 2018). Biofortification offers an effective, economically feasible, and sustainable means of enhancing nutritional content in crops that contribute to staple diets. This strategy typically involves interventions for improving the nutritional content, including vitamins, essential amino acids, minerals, and fatty acids, while simultaneously reducing anti-nutritive factors that hinder the bioavailability of nutrients in crop plants (Garcia-Casal et al., 2017). It has the potential to overcome malnutrition prevalent as 'hidden hunger. In the current climate change scenario, where crop yield and nutritional quality are adversely affected, biofortification can be a successful and game-changing strategy for overcoming nutrient deficiencies, especially in developing nations. Biofortification can be done by understanding the genes, pathways and regulatory networks responsible for absorbing and transporting nutrients and adopting genetic means that involve using the natural germplasm during conventional breeding, genetic engineering (through genetic engineering or production of transgenics) and other OMICS approaches. The significance of multiomics, nutriomics and foodomics have been explained for development of desirable genotypes with potential microgreen related traits (shelf life and nutrient content) those will be also helpful for improvement of breeding cycles (Kalia and Singh, 2018; Mathiazhagan et al., 2021; Reda et al., 2021; Valdes et al., 2021; AieseCigliano et al., 2022; Bansal et al., 2022a; Chakravorty et al., 2022; Dutta et al., 2022; Kumari et al., 2022; Parmar et al., 2022; Sahoo et al., 2022; Sharma et al., 2022).

Conventional breeding depends on the genetic diversity of the gene pool for the trait of interest (TOI). The desired genes are pyramided using traditional crossing approaches, and then segregation populations are thoroughly screened. Biofortification *via* genetic means offers a cost-effective and relatively efficient strategy with the pre-requisite of the availability of inbred lines with high nutrient content in several crops. For millions of under privileged rural residents, sorghum is one of the most essential basic foods. It can flourish in challenging conditions. Moreover, HT12 protein increased the lysine content in sorghum (Zhao et al.,

2003; Lipkie et al., 2013). The fact that sorghum is less easily digested than other main staple crops is one of the problems with eating it. Its kafirin seed storage protein is immune to protease digestion. The RNA silencing of kafirin increases the digestibility index of sorghum in combination with the suppression of kafirin-1, kafirin-2, and kafirin A1 genes (Grootboom et al., 2014; Elkonin et al., 2016). Accordingly, in soybean the globally preferred crop due to its vegetable oil and high-quality protein, the expression of the bacterial PSY gene (beta-carotene) increases the level of provitamin-A, oleic acid, and other protein contents of seed (Schmidt et al., 2015). The fruit color and freshness in strawberry were regulated by specific anthocyanin, anthocyanin related transcription factors and biosynthesis-associated gene expression (Lee et al., 2022). Thus, utilization of such nutrition-enriched developed varieties may be explored for microgreens production.

New breeding approaches, including transgenic breeding, RNA interference (RNAi), and genome editing etc. are crucial for the biofortification of crops because they provide new opportunities for developing unique genetic varieties and are being discussed under following sub heads.

7.5 Genomics and transcriptomics for microgreen traits

A number of DNA-based molecular markers viz; SSR (Simple sequence repeats),microRNA-based SSR, AFLP (Amplified fragment length polymorphism), SNP (Single Nucleotide Polymorphisms), etc. are available for diversity analysis and identification of most potential genotypes for desirable traits in different plants (Gupta et al., 2013; Maurya et al., 2015; Liu et al., 2019; Gupta et al., 2020a; Pradeepkumara et al., 2022) that would be useful to detect most potential genotypes based on phylogeny and genetics studies for particular genotype-specific microgreens. Nutritional profiling of the most diverse genotypes (range: 300–1,000 genotypes) for microgreen-related traits may be performed to detect major phytonutrient components. Evaluation of the most promising genotypes of potential plants (legumes, cereals, herbs, and vegetables) would be helpful for microgreen production based on performance for microgreen-related traits (aroma, tender texture, vivid colour, flavour, and rapid production)at different locations (high/low altitude). Genome-wide association mapping and quantitative trait loci (QTL) mapping will be helpful in detecting potential genotypes and for candidate gene identification related to microgreen related desirable traitsformolecular breeding programs.

Specific QTLs may be identified for desirable microgreens related traits (shelf life and nutrients content: Fe, Zn) in particular crop on the basis of contrasting parents and corresponding data of Genomics. The Quantitative Trait Loci have been identified for phytoactive compounds, iron, zinc and shelf-life related microgreen traits incabbage (Wu et al., 2008), broccoli (Gardner et al., 2016), wheat (Krishnappa et al., 2022), lettuce (Hayashi et al., 2012), melon (Dai et al., 2022)and chickpea (Mahto et al., 2022).

Accordingly, transcriptomics based specific mRNA expression quantitative trait loci (eQTLs)and splicing quantitative trait loci (sQTL) may also be identified for desirable microgreens related traits (shelf life, Fe, Zn) in particular crop on the basis of contrasting parents, standard population size and corresponding nutrition data (Agarwal et al., 2014; Zhu et al., 2022) as identified for several complex traits (Khokhar et al., 2019; Qi et al., 2022). Based on the desirable trait evaluation, a set of candidate genotypes may be selected for transcriptomics study to identify the differentially expressed genes (DEGs) for nutrition and related traits (Figure 2). The candidate genotypes should differ for traits, including bioactive compounds, phytochemicals, antioxidant capacities, mineral composition, yield, and biomass-related traits. The transcriptomic analysis will help to identify the differentially expressed transcripts, biological processes, and molecular pathways for all the contrasting traits, including nutritional and shelf-life-related traits. Once identified, the differential transcripts/genes will be converted into user-friendly markersto be utilized through breeding approaches for future microgreens production.

7.6 Proteomics and metabolomics for microgreen traits

Proteomics and metabolomics have been explored in several microgreens viz; broccoli (Sun et al., 2015), brassica (Castellaneta et al., 2022) and other leafy vegetables also (Sahoo et al., 2022). Novel datasets for microgreens may be generated through research activities, and prospects of encroachment in OMICS approaches. The improvements will consequence through the combined relationship of proteomics and metabolomics for nutritionally rich microgreen development. Protein specific quantitative trait loci (pQTLs), metabolic quantitative trait loci (mQTLs) and micronutrient quantitative trait loci (nutriQTL) play dynamic role in Physiological processes and molecular pathways that may be further identified for desirable microgreens related traits (shelf life, Fe, Zn) as investigated earlier (Engelken et al., 2016; Zhou et al., 2021). Furthermore, proteomics and metabolomics approaches will be a major advancement for microgreen improvement in terms of proteins, metabolites, and bioactive compounds. They will be helpful from plant breeding to OMICS-assisted plant molecular breeding (Langridge and Fleury, 2011).

Variations in the quality and quantity of microgreen proteins and metabolites can be investigated through analysis of proteome composition and changes to developmental stages, including stress-response mechanisms for the enhancement of proteome coverage data and further improvement of protein quality and shelf-life. The establishment of novel approaches related to proteomic and metabolites pipelines would be useful for data analysis associated with different kinds of growth and stress conditions for microgreen-related traits as already explored in several crops through proteome mapping, comparative analysis of proteomics, post-translational modifications, and protein-protein interaction networks, 2D gel electrophoresis coupled with MALDI-TOF (Vanderschuren et al., 2013; Katam et al., 2015). These approaches would be helpful for the purposeful annotation of desirable proteins that participate in metabolism (nitrogen, amino acid, carbon and energy, and Reactive Oxygen Species), stress response, secondary metabolism, and signal transduction (Gupta et al., 2019) that will be further regulated and improve nutrition quality and extended shelf-life.

Good accuracy, speed improvements, sensitivity perfections in mass spectrometry (MS) applications, and software tool improvements have all benefitted high-throughput protein quantification that will be further useful for comparative analysis of proteomics profile in association with differential expression analysis

as explored related to stress responses in legume crops (Pandey et al., 2008; Abdallah et al., 2012; Hu et al., 2015).

Metabolite profiling provides the appropriate data and depth information on metabolic networks responsible for a diverse range of desirable phenotypic traits and undesirable traits that can be regulated through plant metabolic engineering (Fernie and Schauer, 2009). The literature highlights two important nuclear magnetic resonance (NMR) and mass spectrometry (MS)-based metabolomics profiling techniques. It was usually necessary to combine several analytical methods to extract a wider variety of multiple plant metabolites from a single MS (Arbona et al., 2013). Additional methods include Fourier Transform Infrared spectroscopy and MS (FIA/MS), and flow injection-based analysis.

Integrating metabolomics, transcriptomics, bioinformatics platforms, and phenomics to evaluate genetically diverse individuals and improve gene identification accuracy enables the detection of unique metabolic QTLs and candidate genes for the targeted trait that will be cooperative for microgreens improvements. Moreover, a combination of metabolomics screening and genomic-assisted selection strategy has been identified to increase yields, reducing the time spent discovering novel traits and allelic mutations (Fernie and Schauer, 2009).

7.7 Pan genomics for microgreen traits

A species' pan-genome refers to all of its genes collectively. Due to the variety in genomic sequences, it has been determined that a single organism cannot have all of a species' genes. Completeness (i.e., the presence of all functioning genes), stability (i.e., the presence of distinctive catechistic properties), comprehensibility (i.e., the presence of all genomic data for all species or individuals), and effectiveness are the desired characteristics of an ideal pan-genome (i.e., organized data structure). Recently, a 592.58 Mb chickpea pangenome with 29,870 genes was created (Varshney et al., 2021). In order to create the pan-genome, 3,366 accessions totalling 3,171 farmed and 195 wild ones were used in whole genome sequencing. This comprehensive genome analysis provided important details on the genomic regions frequently chosen during domestication, the best haplotypes, and the locations of harmful allele targets. The newly discovered genes that encode reactions to oxidative stress, stimuli, heat shock proteins, cellular (acidic pH) and cold responses may help to modify microgreen cultivation.

7.8 Transgenic approaches for microgreen traits

In order to introduce tolerance or resistance to diverse abiotic and biotic problems, genetic manipulation has been extensively used to identify and transfer resistant gene(s) from a variety of resources to desirable plants for a targeted trait. Today, different genes are used in plants, and transgenic plants have been created through *Agrobacterium*-mediated transformation (Sharma et al., 2006), electroporation of intact axillary buds (Chowrira et al., 1996), particle gun bombardment (Indurker et al., 2007). *Agrobacterium*-mediated explant transformation is the most frequently employed technique to create transgenic pulse crops by inserting transgenes from diverse sources to produce transgenic plants.

Various transgenic plants have been developed for several desired traits. Further, several genes have also been identified for insect pest (protease inhibitor genes, α-amylase inhibitor genes, lectin genes, Cry genes from *Bacillus thuringiensis*, chitinase gene) and disease resistance for example fungal (antifungal protein genes, stilbene synthase gene), viral (coat protein genes of viruses) and bacterial (T4 lysozyme gene) (Dita et al., 2006; Eapen, 2008). The impact of endogenous genes could be regulated by modifying biological processes and metabolic pathways to boost carotenoids and flavanoids using various abiotic stimuli such as drought, salinity, mineral toxins, cold, temperature and RNA interference technologies (Eapen, 2008). Interestingly, desirable multi-trait transgenic plant can be developed through *in vitro* gene stacking system: GuanNan Stacking (Qin et al., 2022). However, transgenic rice have been developed by construction of binary vector and insertion of five desirable foreign genes that would be helpful for regulation of metabolic engineering and trait improvements through breeding and multiomics approaches in future (Qin et al., 2022). Moreover, improvement of chickpea and pigeonpea have been explored through transgenic and molecular approaches (Arya and Mishra 2022).

7.9 Genome editing for microgreen traits

In plant genome editing, sequence-specific nucleases modify specific genes in the selected crop to construct transgene free plants. Moreover, different sequence-specific nucleases including ZFNs, TALENs, and the CRISPR-Cas9 systems have been explored to alter the genome of the targeted plant, fruits and vegetables (Upadhyaya et al., 2009; Mathiazhagan et al., 2021; Chakravorty et al., 2022; Kumari et al., 2022). CRISPR genome editing uses RNA-guided DNA endonucleases (Cas9/13), but these complexes form at the specific target site to execute targeted gene editing (Ducreux et al., 2004; Dancs et al., 2008). Enhancing shelf life and plant storage features by genome editing would be a valuable venture (Kuzmina, 2020). Although, biofortification for cytokinin using gene editing for improving nutrition in chickpeas has recently been projected (Mahto et al., 2022).

7.10 Sequencing-based approaches for microgreen traits

With the advances in the NGS based technologies, trait mapping has become an easy job to do. Not only are these technologies time saving but also reduces the cost at basal levels. The genetic mapping is based on recombination (the exchange of DNA sequence between sister chromatids during meiosis) and the centimorgan (cM) distance measured between the markers by representing approximately 1% of the recombination frequency, while the physical map is based on the alignment of the DNA sequences with distance between markers measured in base pairs. However, the high-resolution physical maps serve as the scaffold for genome sequence assembly to identify the most accurate distance between the markers and the genes linked in addition to exploring of the potential candidate gene(s) linked to desired traits. The trait mapping through sequencing approaches may be categorized into two classes i) Sequencing of complete populations

for trait mapping and ii) Sequencing of pooled samples for trait mapping (Singh et al., 2022). Researchers have great interest in its genomic properties, which provides a valuable marker for crop improvement. Wu et al. (2020) identified multiple high-quality SNPs that would serve as an important resource for the mungbean's nutritional improvement and cultivation. Due to the advancement in sequencing techniques, Dasgupta et al. (2021) conducted the bulk-RNASeq-based gene expression analysis across mungbean genotypes to identify disease-resistance genes. Bhardwaj et al. (2015) did an RNAseq-based analysis to identify drought stress-regulated genes in *Brassica juncea*. Guo et al. (2021) emphasized the genomic-based structural variant that indicated the diversification of different morphotypes of *Brassica oleracea*. These approaches are equally applicable and may be explored for improving microgreen traits (Mishra et al., 2021).

7.12 Epigenomics for microgreen traits

The increasing world population and changing climate increase the demand for greater crop productivity. The selection of appropriate genetic techniques and desirable heritable DNA sequences have led to notable genetic advancements in many crop species. Specific methylation quantitative trait loci (meQTLs) may be identified for desirable microgreens related traits (shelf life, Fe, Zn) (Kumar et al., 2016; Kumar et al., 2022b). Further, correlation between transcriptome (eQTL) and methylome (meQTLs) have been established for genetic regulation of complex traits (Oliva et al., 2022). In addition, a better comprehension of and capacity to choose advantageous epigenomic modifications is suggested to incorporate a more effective and comprehensive approach to crop improvement (AieseCigliano et al., 2022; Chandana et al., 2022). This is because many plant stress responses are governed by epigenomic processes, notably through cell-autonomous epigenetic switching. This makes it possible to register and remember random genetic signals. According to a report, the memory-directed alteration may result in an increased ability to resist stress in the future (Berr et al., 2011).

The mechanisms governing plant-stress interactions and conditions are revealed by studying the roles of epigenetics causing stressors, such as histone changes and DNA methylations (Chinnusamy and Zhu, 2009). Numerous heritable modifications originating from mitotic and meiotic divisions (variations in the heredity of epigenetic markers) were observed during gene expression studies and are steadily transmitted from one generation to the next that were not encoded in the DNA sequence itself (Tsaftaris and Polidors, 2000; Berger et al., 2009; Chen et al., 2010).

Stout epigenetic alterations are mitotically transmitted through genomic imprinting, but transient epigenetic alterations are not heritable (Spillane et al., 2001). Until they are lost or removed, epigenetic alterations produced during meiosis are always transmissible from one generation to the next without the need for initial stimulation. The loss could be from a genetic mutation, unintentional (for unknown reasons), or result from environmental factors. These are distinct from the ones that brought about the original epigenetic changes. In plants, heterosis is exhibited in hybrids for high biomass (Gupta et al., 2020b), a straightforward epigenetic assumption.

7.13 Genomics to artificial intelligence for microgreen traits

Machine learning (ML) and Deep Learning (DL) is the component of artificial intelligence (AI) that includes mathematical models of data to improve plant performance based on statistics, predictive modeling, and data analysis. It may be considered artificial intelligence based on plant breeding or crop improvement. Precision agriculture and crop improvement through artificial intelligence are new for plants and are untouched areas for microgreens. Prediction of shelf-life is also possible through exact measurement of plant characters (accurate interpretation of high-throughput phenotypic data) by good quality imaging techniques (greater than 7,000 high-resolution images of 300 GBytes) and proficient analysis of refined extracted data using artificial intelligence.

Microgreens-specific genomics and machine learning-based innovative approaches may be employed in other crop plants that have been proposed worldwide. Genomics approaches help to identify the "SNP" and "SSR " markers and annotate the genes. Furthermore, a detailed analysis of these markers could help identify the nearest genes (QTLs), as shown in Figure 1. Another important genomic layer of information is "Bulk-RNASeq," which can be implemented for accurate data analysis and prediction. ML aims at providing innovative approaches for prediction-based model development; Girma (2019) discussed the ML approach in mungbean to classify the raw quality of samples by analyzing digital imaging data; Jung et al. (2021) applied the deep learning algorithm to correctly identify the *Brassica napa* varieties, followed by a cross-validation approach. Another important microgreen vegetable crop is Broccoli (*Brassica oleracea* L. var. *italica*). The importance lies in the broccoli head portion, which helps to assess plant quality and different biotic and abiotic stress. Zhou et al. (2020) used the "Improved ResNet" to extract the broccoli pixels from the background data. In another study based on broccoli head estimation, Kusumam et al. (2017) applied the deep learning approach for 3D-vision-based detection. DL based semantic segmentation models have been applied in another microgreen plant cabbage for crop estimation (Jo et al., 2021). One of the studies by Wolanin et al. (2020) used the DL method for the time series data to estimate the wheat yield in the Indian wheat belt. AI has been applied to the barley seeds too. Various studies have been published where machine learning has been used for the identification of lentil-based rust disease identification (Singh et al., 2019). Applying machine learning models and techniques predicted the shelf life of Okra (Iorliam et al., 2021) and muskmelons (Albert-Weiss and Osman, 2022). Thus, active learning would be helpful to predict the shelf life of different types of microgreens through Support Vector Machine, Logistic Regression K-Nearest Neighbour algorithms,Naïve Bayes and Decision Tree.

7.14 Bioinformatics and molecular databases for microgreen traits

The plant research group requires efficient bioinformatics pipelines and a system to support efforts to analyse microgreen-related targeted plant genomes through functional genomics due to the rapid progress of publicly available databases from different kinds of tissues, development, environments, and stress treatment. The comprehensive model plant genomics, transcriptomics, and proteomics databases can be used to identify appropriate microgreen genotypes. The genome sequences of several plants (*Medicago truncatula*, *Glycine max*, and *Lotus japonicus*) and

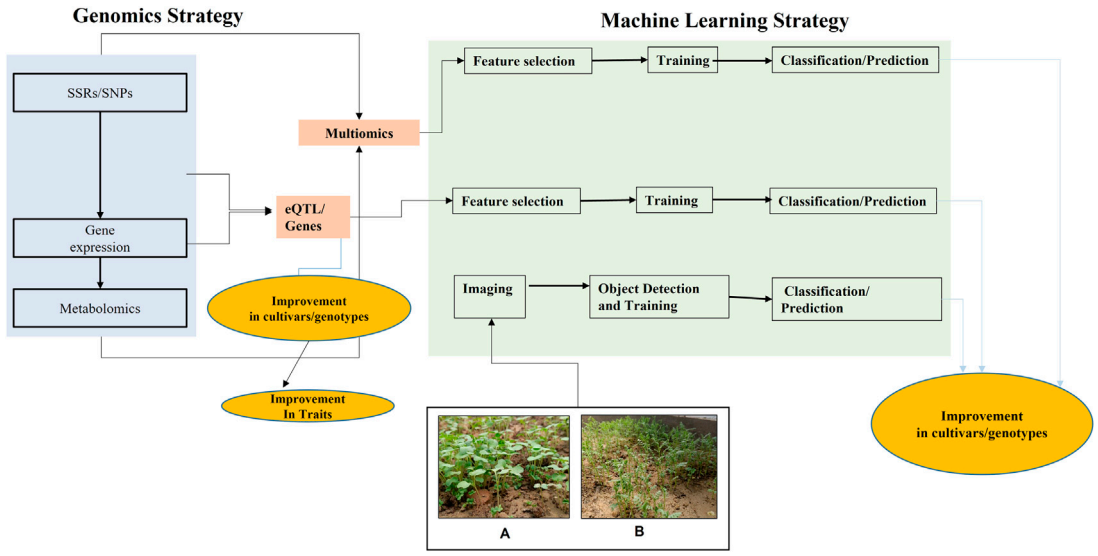


FIGURE 1
Systematic workflow of OMICS analysis and its association with machine learning for microgreens improvement and shelf-life prediction: Genomic data could be used to explore the markers in the form of SSRs, SNPs. Bulk-RNASeq and metabolomics data provides another layer of information that would enable Multi-OMICS analysis for cultivation improvement. Machine learning strategies that make the use of both "numeric" and "imaging" data would enable the development of prediction and classification models.

reference plant species (*Arabidopsis thaliana* and *Populus trichocarpa*) are already available to investigate gene function, biological processes, metabolic pathways, and genome evolution (Li et al., 2012). The available data bases viz., The Legume Information System (LIS; https://legumeinfo.org) and KnowPulse (https://knowpulse.usask.ca) are very informative and useful computational genomics platforms to evaluate molecular markers, diversity analysis, comparative genomics, gene annotation, novel transcription factors, sequence variants, phenotypic traits informationand to map SNPs, QTLs, long non-coding RNAs and to identify candidate genesfor selection of microgreen oriented suitable chickpea, faba bean, common bean, lentil, andfield pea germplasm (Doddamani et al., 2015; Verma et al., 2015; Dash et al., 2016; Gayaliet al., 2016; Sanderson et al., 2019; Lee et al., 2022).

8 Integrating various OMICS approaches for microgreen traits

Prospective OMICS approaches have been investigated in many plants to improve the desirable traits and elucidated earlier (AieseCigliano et al., 2022; Chakravorty et al., 2022; Dutta et al., 2022; Gupta, 2022; Gupta and Tyagi, 2022; Parmar et al., 2022; Sharma and Gupta, 2022; Singh et al., 2022). Re-sequencing activities of whole-genome employing genetic diversity, domestication patterns, evolutionary analysis, population structure, and linkage disequilibrium for chickpea improvement as a result of the technical advancements that upgraded chickpea (an orphan crop) to a potential geneticcrop (Varshney et al., 2019).

Recent genomics methods offer the potential to accelerate gene discovery, marker creation, molecular breeding, trait mapping, and productivity advances in microgreens, among other processes (Figure 2). Integration of precise phenotypic variation, low-frequency variants, and sequence information approach would be helpful for the selection of the most appropriate accessions with desirable key traits like biomass components, biotic and abiotic stress tolerance, and nutritional traits (Roorkiwal et al., 2020). A broad range of molecular markers (SSR, SNP, and DArT) have been discovered in chickpea that has been facilitated by NGS technology (Whole-genome re-sequencing, genotyping by sequencing, skim sequencing, RAD-Seq, and lower-depth sequencing). They can also be used for the development of chickpea microgreens (Kale et al., 2015; Varshney, 2016; Varshney et al., 2018).

There is a huge gap from genome to phenome in agricultural plants to identify the particular phenotype based on their DNA sequence information and genetics. Thus, it is crucial to integrate multi-OMICS information in one place from several branches of OMICs platforms, including phenomics, genomics transcriptomics, proteomics, epigenomics, and metabolomics. Using all the OMICS technology, the genotype-phenotype divide in any microgreens can be closed with precision phenotyping.

To learn new things about the potential genes and biological processes involved, analysis at genomics, transcriptomics, proteomics, epigenomics, and metabolomics levels can be done depending on the study's goal. Moreover, it has been reported by Mannur et al. (2019) that *Fusarium* wilt resistance loci (foc 4) from WR 315 Annigeri 1 has been made available as "Super Annigeri one" for commercial production in India using a genomics strategy. Thus, it is evident that various other studies have also proposed OMICS-based integration methods (Argelaguet et al., 2018; Bhardwaj and Steen, 2020; Mahto et al., 2022). Cai et al., 2021 discussed applicability of OMICS data (transcriptome and proteomics) for barley and identified the connection between sugar metabolism and wild barley. Li et al. (2019) used a single omics layer of information, i.e., transcriptome, and revealed the gene expression patterns of sulforaphane metabolism

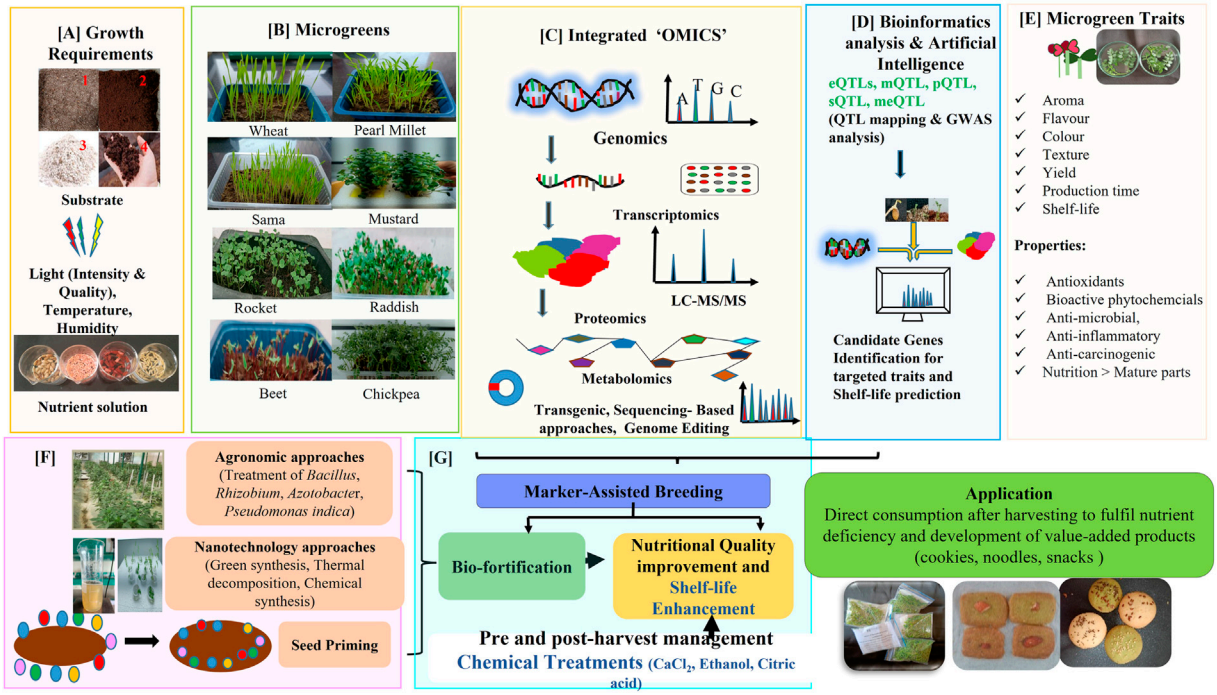


FIGURE 2
Summarizes prospects of microgreens as budding live functional food. **(A)** Growth conditions required for microgreens cultivation which includes a variety of substrates like vermiculite (1), cocopeat (2), perlite (3) and vermicompost (4), light (quality, intensity and duration), temperature, humidity and nutrient solution. **(B)** Microgreens of different plant species: Wheat, Pearl millet, Sama, Mustard, Rocket, Raddish, Beet and Chickpea. **(C)** Numerous "OMICS' approaches such as Genomics, Transcriptomics, Proteomics, Metabolomics, Epigenomics, along with Transgenics, Gene editing and Sequencing based approaches can be integrated with bioinformatics tools and artificial intelligence **(D)** to tag Quantitative Trait Loci (eQTLs: mRNA expression Quantitative Trait Loci; meQTLs: Methylation Quantitative Trait Loci; PQTL: Protein Quantitative Trait Loci; sQTL: splicing Quantitative Trait Loci; mQTLs: Metabolic Quantitative Trait Loci) and candidate genes identification for microgreen related desirable traits **(E)** like nutrients, flavour, colour, early germination, yield and shelf life. The data generated from Integrated "OMICS' approaches can be further utilized in molecular breeding to produce nutritionally rich varieties with improved shelf life through Marker-Assisted Breeding and producing biofortified microgreens with targeted micro/macro nutrients (like iron, zinc, magnesium, calcium) enrichments. **(F)** Biofortification of target microgreen is also possible by agronomic approaches (incubation with microorganisms like *Bacillus, Rhizobium, Azotobacter, Pseudomonas indica*), nanotechnology (nano-biofortification) and seed priming. Bioavailability of nutrients and minerals of microgreens can be stabilized through pre and post-harvest management strategy **(G)**. Improved microgreens with desired nutrients can be either consumed fresh as garnishes in soups, sandwiches, salads or processed to develop value-added products (like noodles, breads, drinks, cookies *etc.*) to overcome nutrients deficiency.

in Broccoli florets. For microgreens, this field is still lagging (Figure 1). There is a need to exploit the information from different resources to identify various biological factors that would help microgreen plants' cultivation and breeding processes.

Candidate genes have been identified and incorporated in desirable plant through Genomics (QTL mapping), transcriptomics and transgenic approaches for microgreens related traits (shelf life and nutrient content: Fe, Zn) in different plant systems viz cabbage (Wu et al., 2008), broccoli (Gardner et al., 2016), wheat (Krishnappa et al., 2022), lettuce (Hayashi et al., 2012), melon (Dai et al., 2022) and chickpea (Mahto et al., 2022). Metabolomics and proteomics have been explored in several microgreens viz., broccoli (Sun et al., 2015), brassica (Castellaneta et al., 2022) and other leafy vegetables also (Sahoo et al., 2022). Thus, it is evident that we may identify and predict candidate genes associated with microgreens related traits (shelf life, desirable nutrients content, developmental rate and phytochemicals) may be incorporated in targeted crop varieties of fruits and vegetables utilizing the crop specific data bases of genomics, transcriptomics and metabolomics, marker-assisted selection, GWAS, bioinformatics, AI approaches utilizing the databases of specific crop transgenic CRISPR/Cas9 and gene editing approaches (Mathiazhagan et al., 2021; Valdes et al., 2021; Chakravorty et al., 2022; Parmar et al., 2022; Sahoo et al., 2022). Thus combination of multiomics, foodomics in association with nutriomics would be fruitful for regulation of nutritional balance, health management and treatment of diseases (Bansal et al., 2022a).

9 Conclusion and future perspective

Optimization of light, substrate, and temperature would be helpful for good quality microgreen cultivation containing desirable aroma traits, tender texture, vivid colour, flavour, sensory attributes, and rapid production. Further evaluation of genotypes would be helpful for the selection of the most potential genotypes for the mass production of microgreens. Genomics and transcriptomics approaches may be explored for candidate gene identification for microgreens and important nutritional traits. Explored genes or associated SNPs may be developed and explored for user-friendly markers for marker-assisted selection and metabolic pathways with the integration of multi-OMICS approaches. Characterization of microgreens can be performed by combining the most standardized condition for the growth of microgreens with enhanced nutrient quality and bioavailability. Profiling of

phytochemicals, nutrients, and minerals should be studied for good quality microgreens and biofortify further using traditional and novel biofortification approaches. For commercialization and popularization of the microgreen's cultivation and harvest, shelf-life should be focused. The development of post-harvest technology for enhanced storability of microgreens is very important for synthesizing value-added, tasty, and nutritional products.

Due to the nutritional content inclusion of microgreens in the diet have several health benefits, as evident from the literature. However, microgreens' growth, yield, and nutritional content can vary with the growing method (soil, compost, or hydroponic), intensity and quality of illumination, and composition of plant nutrient solution. Establishing hydroponics and vertical farming systems will enhance the cultivation of nutritionally rich microgreens with an easy harvest. The development of technologies to preserve microgreens is needed for a more extended period with minimal changes in their phytochemicals and nutrients. Novel value-added products (for example, drink, juices, cookies, noodles, *etc.*) may be developed with the integration of microgreens as one of the ingredients having wider acceptability and enhanced nutrition, especially for elderly persons, infants, young growing children, and sick persons due to excellent digestibility. Microgreens cultivation is a very easy and promising strategy to initiate at home or as a start-up for beginners/poor farmers in their respective localities, including people with malnutrition in India. The persons engaged in microgreens cultivation, production, and utilization may explore the food processing and packaging industry to enhance the market regarding agriculture and nutritionally rich products. Thus, microgreens will occupy a central place in the future food industry.\

Author contributions

AG and RK conceptualized and supervised the manuscript writing. TS, DS, AB and SS collected the related literature and contributed to the original writing. RK extended their help in inference, review and AG did editing of the manuscript. All authors went through the final manuscript draft and approved it.

Funding

The study was supported by School of Agricultural Sciences, Sharda University, and ICAR-Indian Agricultural Research Institute, New Delhi.

Acknowledgments

The authors gratefully acknowledge the Sharda School of Agricultural Sciences, Sharda University, Greater Noida, Uttar Pradesh, India; ICAR-Indian Agricultural Research Institute, New Delhi, India; Integral Institute of Agricultural Sharda School of Agricultural Sciences, Sharda University, Greater Noida Uttar Pradesh, India; Institute of Clinical Molecular Biology (IKMB), Kiel, Germany; Dayanand Anglo-Vedic (PG) College, Chhatrapati Shahu Ji Maharaj University, Kanpur, Uttar Pradesh, India. All authors are grateful to Dolly Wattal Dhar, Dean, Sharda School of Agricultural Sciences, Sharda University, Greater Noida, Uttar Pradesh, India for professional editing. The lead author RK gratefully acknowledges the DST-Science and Engineering Research Board for providing the financial support to carry out the manuscript preparation.

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.

Publisher's note

References

Abdallah, C., Dumas-Gaudot, E., Renaut, J., and Sergeant, K. (2012). Gel-based and gel-free quantitative proteomics approaches at a glance. *Int. J. Plant Genomics* 2012, 1–17. doi:10.1155/2012/494572

Agarwal, S., Vgn, T. V., Kotla, A., Mangrauthia, S. K., and Neelamraju, S. (2014). Expression patterns of QTL based and other candidate genes in Madhukar× Swarna RILs with contrasting levels of iron and zinc in unpolished rice grains. *Gene* 546 (2), 430–436. doi:10.1016/j.gene.2014.05.069

AieseCigliano, R., Aversano, R., Di Matteo, A., Palombieri, S., Termolino, P., Angelini, C., et al. (2022). Multi-omics data integration provides insights into the post-harvest biology of a long shelf-life tomato landrace. *Hortic. Res.* 9 uhab042, doi:10.1093/hr/uhab042

Al-Bader, A., Christenson, J. T., Simonet, F., Abul, H., Dashti, H., and Schmuziger, M. (1998). Inflammatory response and oligo-element alterations following cardiopulmonary bypass in patients undergoing coronary artery bypass grafting. *Cardiovasc. Surg.* 6 (4), 406–414. doi:10.1016/s0967-2109(98)00022-2

Albert-Weiss, D., and Osman, A. (2022). Interactive deep learning for shelf life prediction of muskmelons based on an active learning approach. *Sensors* 22 (2), 414. doi:10.3390/s22020414

Alegbeleye, O. O., Singleton, I., and Sant'Ana, A. S. (2018). Sources and contamination routes of microbial pathogens to fresh produce during field cultivation: A review. *Food Microbiol.* 73, 177–208. doi:10.1016/j.fm.2018.01.003

Altuner, F. (2021). Determination of biochemical composition and pigment content in legume and cereal microgreens. *Legume Res-An Int. J.* 1, 8. doi:10.18805/LR-635

Altuner, F., Tunçtürk, R., Oral, E., and Tunçtürk, M. (2022). Determınatıon of the content of antıoxıdants and bıochemıcal composition of legume mıcrogreens. *J. Elem.* 1 (2022), 165–180. doi:10.5601/jelem.2022.27.1.2178

Arbona, V., Manzi, M., de Ollas, C., and Gómez-Cadenas, A. (2013). Metabolomics as a tool to investigate abiotic stress tolerance in plants. *Int. J. Mol. Sci.* 14 (3), 4885–4911. doi:10.3390/ijms14034885

Argelaguet, R., Velten, B., Arnol, D., Dietrich, S., Zenz, T., Marioni, J. C., et al. (2018). Multi-omics factor analysis—A framework for unsupervised integration of multi-omics data sets. *Mol. Sys. Biol.* 14 (6), 8124. doi:10.15252/msb.20178124

Arya, M., and Mishra, S. B. (2022). "Transgenic and molecular approaches for pigeonpea and chick pea improvement," in *Technologies in plant biotechnology and breeding of field crops* (Singapore: Springer), 239–272.

Bahrami, A., Delshadi, R., Jafari, S. M., and Williams, L. (2019). Nano encapsulated nisin: An engineered natural antimicrobial system for the food industry. *Trends Food Sci. Technol.* 94, 20–31. doi:10.1016/j.tifs.2019.10.002

Bansal, P., Bansal, R., and Arora, M. (2022a). "Traditional nutritional approaches and nutriomics evidence: Nutrition from inception to evidence," in *Nutriomics* (United States: CRC Press), 23–47.

Bansal, T., Kawatra, A., and Sangwan, V. (2022b). Sensorial, nutritional and shelf life evaluation of bio-fortified millet based cookies supplemented with carrot powder and sesame. *Asian JDairy Food Res.* 41 (1), 83–88. doi:10.18805/ajdfr.DR-1696

Batiha, G. E. S., Al-Gareeb, A. I., Qusti, S., Alshammari, E. M., Kaushik, D., Verma, R., et al. (2022). Deciphering the immunoboosting potential of macro and micronutrients in COVID support therapy. *Environ Sci Pollut. Res* 1, 43516–43531. doi:10.1007/s11356-022-20075-7

Berger, S. L., Kouzarides, T., Shiekhattar, R., and Shilatifard, A. (2009). An operational definition of epigenetics. *Genes & Dev.* 23 (7), 781–783. doi:10.1101/gad.1787609

Berr, A., Shafiq, S., and Shen, W. H. (2011). Histone modifications in transcriptional activation during plant development. *Biochim. Biophys. Acta.* 1809 (10), 567–576. doi:10.1016/j.bbagrm.2011.07.001

Bhardwaj, A. R., Joshi, G., Kukreja, B., Malik, V., Arora, P., Pandey, R., et al. (2015). Global insights into high temperature and drought stress regulated genes by RNA-Seq in economically important oilseed crop *Brassica juncea*. *BMC Plant Biol.* 15 (1), 9–15. doi:10.1186/s12870-014-0405-1

Bhardwaj, A., and Steen, K. V. (2020). "Multi-omics data and analytics integration in ovarian cancer," in IFIP International Conference on Artificial Intelligence Applications and Innovations, Neos Marmaras, Greece, 5-7 June, 347–435.

Bibi, I., Ghulam, T., Kamal, S., Jilani, K., Alwadai, N., and Iqbal, M. (2022). Green synthesis of iron nanoparticles and photocatalytic activity evaluation for the degradation of methylene blue dye. *Z. Phys. Chem.* 236, 1191–1201. doi:10.1515/zpch-2021-3128

Brazaitytė, A., Miliauskienė, J., Vaštakaitė-Kairienė, V., Sutulienė, R., Laužikė, K., Duchovskis, P., et al. (2021). Effect of different ratios of blue and red LED light on *Brassicaceae* microgreens under a controlled environment. *Plants* 10 (4), 801. doi:10.3390/plants10040801

Bruening, G., and Lyons, J. (2000). The case of the FLAVR SAVR tomato. *Calif. Agri* 54 (4), 6–7. doi:10.3733/ca.v054n04p6

Bulgari, R., Baldi, A., Ferrante, A., and Lenzi, A. (2017). Yield and quality of basil, Swiss chard, and rocket microgreens grown in a hydroponic system. *N. Z. J. Crop Hortic. Sci.* 45 (2), 119–129. doi:10.1080/01140671.2016.1259642

Bulgari, R., Negri, M., Santoro, P., and Ferrante, A. (2021). Quality evaluation of indoor-grown microgreens cultivated on three different substrates. *Horticulturae* 7 (5), 96. doi:10.3390/horticulturae7050096

Bunea, A., Andjelkovic, M., Socaciu, C., Bobis, O., Neacsu, M., Verhé, R., et al. (2008). Total and individual carotenoids and phenolic acids content in fresh, refrigerated and processed spinach (*Spinacia oleracea* L.). *Food Chem.* 108 (2), 649–656. doi:10.1016/j.foodchem.2007.11.056

Butkutė, B., Taujenis, L., and Norkevičienė, E. (2018). Small-seeded legumes as a novel food source. Variation of nutritional, mineral and phytochemical profiles in the chain: Raw seeds-sprouted seeds-microgreens. *Molecules* 24 (1), 133. doi:10.3390/molecules24010133

Cai, S., Shen, Q., Huang, Y., Han, Z., Wu, D., Chen, Z. H., et al. (2021). Multi-omics analysis reveals the mechanism underlying the edaphic adaptation in wild barley at evolution slope (tabigha). *Adv. Sci.* 8 (20), 2101374. doi:10.1002/advs.202101374

Castellaneta, A., Losito, I., Cisternino, G., Leoni, B., Santamaria, P., Calvano, C. D., et al. (2022). All ion fragmentation analysis enhances the untargeted profiling of glucosinolates in Brassica microgreens by liquid chromatography and high-resolution mass spectrometry. *J. AmSoc Mass Spectrom.* 33 (11), 2108–2119. doi:10.1021/jasms.2c00208

Chakravorty, M., Nanda, M., Arora, N., Singh, S., Kumar, V., and Deshwal, S. (2022). CRISPR-Cas9 mediated genome tailoring to improve nutritional quality and shelf life in crops: A review. *Plant* 31, 100369. doi:10.1016/j.plgene.2022.100369

Chandana, B. S., Mahto, R. K., Singh, R. K., Ford, R., Vaghefi, N., Gupta, S. K., et al. (2022). Epigenomics as potential tools for enhancing magnitude of breeding approaches for developing climate resilient chickpea. *Front. Genet.* 13, 900253. doi:10.3389/fgene.2022.900253

Chandra, D., Kim, J. G., and Kim, Y. P. (2012). Changes in microbial population and quality of microgreens treated with different sanitizers and packaging films. *Hortic. Environ. Biotech.* 53 (1), 32–40. doi:10.1007/s13580-012-00756

Chen, H. M., Chen, L. T., Patel, K., Li, Y. H., Baulcombe, D. C., and Wu, S. H. (2010). 22-Nucleotide RNAs trigger secondary siRNA biogenesis in plants. *Proc. Natl. Acad. Sci. U. S. A.* 107 (34), 15269–15274. doi:10.1073/pnas.1001738107

Chinnusamy, V., and Zhu, J. K. (2009). Epigenetic regulation of stress responses in plants. *Curr.Opin. Plant Biol.* 12 (2), 133–139. doi:10.1016/j.pbi.2008.12.006

Choe, U., Yu, L. L., and Wang, T. T. (2018). The science behind microgreens as an exciting new food for the 21st century. *J. Agric. Food Chem.* 66 (44), 11519–11530. doi:10.1021/acs.jafc.8b03096

Chowrira, G. M., Akella, V., Fuerst, P. E., and Lurquin, P. F. (1996). Transgenic grain legumes obtained by in planta electroporation-mediated gene transfer. *Mol. Biotech.* 5 (2), 85–96. doi:10.1007/BF02789058

Craver, J. K., Gerovac, J. R., Lopez, R. G., and Kopsell, D. A. (2017). Light intensity and light quality from sole-source light-emitting diodes impact phytochemical concentrations within *Brassica* microgreens. *J. Am. Soc. Hortic. Sci.* 142 (1), 3–12. doi:10.21273/JASHS03830-16

Dai, D., Zeng, S., Wang, L., Li, J., Ji, P., Liu, H., et al. (2022). Identification of fruit firmness QTL ff2. 1 by SLAF-BSA and QTL mapping in melon. *Euphytica* 218 (5), 1–15. doi:10.21203/rs.3.rs-1006787/v1

Dancs, G., Kondrak, M., and Banfalvi, Z. (2008). The effects of enhanced methionine synthesis on amino acid and anthocyanin content of potato tubers. *BMC Plant Biol.* 8 (1), 65. doi:10.1186/1471-2229-8-65

Dasgupta, U., Mishra, G. P., Dikshit, H. K., Mishra, D. C., Bosamia, T., Roy, A., et al. (2021). Comparative RNA-Seq analysis unfolds a complex regulatory network imparting yellow mosaic disease resistance in mungbean [*Vigna radiata* (L.) R. Wilczek]. *Plos one* 16 (1), 0244593. doi:10.1371/journal.pone.0244593

Dash, S., Campbell, J. D., Cannon, E. K., Cleary, A. M., Huang, W., Kalberer, S. R., et al. (2016). Legume information system (legume info. Org): A key component of a set of federated data resources for the legume family. *Nucl. Acids Res.* 44 (D1), D1181–D1188. doi:10.1093/nar/gkv1159

De la Fuente, B., López-García, G., Máñez, V., Alegría, A., Barberá, R., and Cilla, A. (2020). Antiproliferative effect of bioaccessible fractions of four Brassicaceae microgreens on human colon cancer cells linked to their phytochemical composition. *Antioxidants* 9 (5), 368–383. doi:10.3390/antiox9050368

de la Fuente, B., López-García, G., Máñez, V., Alegría, A., Barberá, R., and Cilla, A. (2019). Evaluation of the bioaccessibility of antioxidant bioactive compounds and minerals of four genotypes of *Brassicaceae* microgreens. *Foods* 8 (7), 250. doi:10.3390/foods8070250

Delian, E., Chira, A., Bădulescu, L., and Chira, L. (2015). Insights into microgreens physiology. *Sci. Pap. Ser. B Hortic.* 59, 447–454.

Dhakshayani, G. M., and Alias, P. S. J. (2022). A comparative study of phytochemical, antioxidant, anticarcinogenic, and antidiabetic potential of coriander (*Coriandrum sativum* L.): Microgreen and mature plant. *Foods Raw Materi* 10 (2), 283–294. doi:10.21603/2308-4057-2022-2-539

Di Gioia, F., Leoni, B., and Santamaria, P. (2015). "The selection of the species to grow," in *Eco-logicaeditore*. Microgreens(Bari) (Bari: Eco-logica).

Di Gioia, F., Petropoulos, S. A., Ozores-Hampton, M., Morgan, K., and Rosskopf, E. N. (2019). Zinc and iron agronomic biofortification of *Brassicaceae* microgreens. *Agronomy* 9 (11), 677. doi:10.3390/agronomy9110677

Di Gioia, F., Renna, M., and Santamaria, P. (2017). "*Sprouts, microgreens and "baby leaf" vegetables*"," in *Minimally processed refrigerated fruits and vegetables* (Boston: Springer), 403–432.

Dita, M. A., Rispail, N., Prats, E., Rubiales, D., and Singh, K. B. (2006). Biotechnology approaches to overcome biotic and abiotic stress constraints in legumes. *Euphytica* 147 (1), 1–24. doi:10.1007/s10681-006-6156-9

Doddamani, D., Khan, A. W., Katta, M. A. V. S. K., Agarwal, G., Thudi, M., Ruperao, P., et al. (2015). CicArVarDB: SNP and InDel database for advancing genetics research and breeding applications in chickpea. *Database* 19, bav078–7. doi:10.1093/database/bav078

Ducreux, L. J. M., Morris, W. L., Hedley, P. E., Shepherd, T., Davies, H. V., Millam, S., et al. (2004). Metabolic engineering of high carotenoid potato tubers containing enhanced levels of β-carotene and lutein. *J. Exp. Bot.* 56 (409), 81–89. doi:10.1093/jxb/eri016

Dutta, B., Lahiri, D., Nag, M., Abukhader, R., Sarkar, T., Pati, S., et al. (2022). Multi-omics approach in amelioration of food products. *FrontMicrobiol* 13, 955683. doi:10.3389/fmicb.2022.955683

Eapen, S. (2008). Advances in development of transgenic pulse crops. *Biotechnol. Adv.* 26 (2), 162–168. doi:10.1016/j.biotechadv.2007.11.001

Ebert, A. W. (2013). "Sprouts, microgreens, and edible flowers: The potential for high value specialty produce in asia," in SEAVEG 2012: High Value Vegetables in Southeast Asia: Production, Supply and Demand, Chiang Mai, Thailand, 24-26 January 2012, 216–227.

El-Nakhel, C., Pannico, A., Graziani, G., Kyriacou, M. C., Giordano, M., Ritieni, A., et al. (2020). Variation in macronutrient content, phytochemical constitution and *in vitro* antioxidant capacity of green and red butterhead lettuce dictated by different developmental stages of harvest maturity. *Antioxidants* 9 (4), 300. doi:10.3390/antiox9040300

El-Ramady, H., Abdalla, N., Elbasiouny, H., Elbehiry, F., Elsakhawy, T., Omara, A. E. D., et al. (2021). Nano-biofortification of different crops to immune against COVID-19: A review. *Ecotoxicol. Environ. Saf.* 222, 112500. doi:10.1016/j.ecoenv.2021.112500

Elkonin, L. A., ItalianskayaI, J. V., Domanina, V. N., Selivanov, N. Y., Rakitin, A. L., and Ravi, N. V. (2016). Transgenic Sorghum with improved digestibility of storage proteins obtained by *Agrobacterium*-mediated transformation. *Russ. J. Plant Physiol.* 63 (5), 678–689. doi:10.1134/S1021443716050046

Eltaweil, A. S., Fawzy, M., Hosny, M., Abd El-Monaem, E. M., Tamer, T. M., and Omer, A. M. (2022). Green synthesis of platinum nanoparticles using *Atriplex halimus* leaves for potential antimicrobial, antioxidant, and catalytic applications. *Arab. J. Chem.* 15, 103517. doi:10.1016/j.arabjc.2021.103517

Engelken, J., Espadas, G., Mancuso, F. M., Bonet, N., Scherr, A. L., Jímenez-Álvarez, V., et al. (2016). Signatures of evolutionary adaptation in quantitative trait loci influencing trace element homeostasis in liver. *Mol. BiolEvol* 33 (3), 738–754. doi:10.1093/molbev/msv267

Fernie, A. R., and Schauer, N. (2009). Metabolomics-assisted breeding: A viable option for crop improvement? *Trends Genet.* 25 (1), 39–48. doi:10.1016/j.tig.2008.10.010

Gao, M., He, R., Shi, R., Zhang, Y., Song, S., Su, W., et al. (2021). Differential effects of low light intensity on broccoli microgreens growth and phytochemicals. *Agronomy* 11 (3), 537. doi:10.3390/agronomy11030537

Garcia-Casal, M. N., Peña-Rosas, J. P., and Giyose, B.Consultation Working Groups (2017). Staple crops biofortified with increased vitamins and minerals: Considerations for a public health strategy. *Ann. N. Y. Acad. Sci.* 1390 (1), 3–13. doi:10.1111/nyas.13293

Gardner, A. M., Brown, A. F., and Juvik, J. A. (2016). QTL analysis for the identification of candidate genes controlling phenolic compound accumulation in broccoli (*Brassica oleracea* L. var. italica). *Mol. Breed.* 36 (6), 81–12. doi:10.1007/s11032-016-0497-4

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

Gayali, S., Acharya, S., Lande, N. V., Pandey, A., Chakraborty, S., and Chakraborty, N. (2016). CicerTransDB 1.0: A resource for expression and functional study of chickpea transcription factors. *BMC Plant Biol.* 16 (1), 169. doi:10.1186/s12870-016-0860-y

Ghoora, M. D., and Srividya, N. (2020). Effect of packaging and coating technique on postharvest quality and shelf life of *Raphanus sativus* L. and *Hibiscus sabdariffa* L. microgreens. *Foods* 9 (5), 653. doi:10.3390/foods9050653

Girma, T. (2019). *Feature extraction and classification of green mung bean using machine learning techniques*. Ethiopia: Diss. ADDIS ABABA Science and Technology University.

Grootboom, A. W., Mkhonza, N. L., Mbambo, Z., O'Kennedy, M. M., Da Silva, L. S., Taylor, J., et al. (2014). Co-suppression of synthesis of major α-kafirin sub-class together with γ-kafirin-1 and γ-kafirin-2 required for substantially improved protein digestibility in transgenic Sorghum. *Plant Cell Rep.* 33 (3), 521–537. doi:10.1007/s00299-013-1556-5

Guo, N., Wang, S., Gao, L., Liu, Y., Wang, X., Lai, E., et al. (2021). Genome sequencing sheds light on the contribution of structural variants to *Brassica oleracea* diversification. *BMC Biol.* 19 (1), 93–15. doi:10.1186/s12915-021-01031-2

Gupta, A., Bhardwaj, A., Sawant, S. V., and Yadav, H. K. (2019). Utilization and characterization of genome-wide SNP markers for assessment of ecotypic differentiation in *Arabidopsis thaliana*. *Int. J. Curr. Microbiol. App. Sci.* 8 (6), 158–173. doi:10.20546/ijcmas.2019.806.020

Gupta, A., Jaiswal, V., Sawant, S. V., and Yadav, H. K. (2020a). Mapping QTLs for 15 morphometric traits in *Arabidopsis thaliana* using Col-0 × Don-0 population. *Physiol. Mol. Biol. Plants* 26 (5), 1021–1034. doi:10.1007/s12298-020-00800-7

Gupta, A. (2022). "Mainstreaming of underutilized oilseed safflower crop through biotechnological approaches for improving economic and environmental sustainability," in *Biotechnological innovations for environmental bioremediation* (Singapore: Springer), 397–418.

Gupta, A., Maurya, R., Roy, R. K., Sawant, S. V., and Yadav, H. K. (2013). AFLP based genetic relationship and population structure analysis of *Canna* - an ornamental plant. *Sci. Hortic.* 154, 1–7. doi:10.1016/j.scienta.2013.02.005

Gupta, A., Sawant, S. V., and Yadav, H. K. (2020b). Assessment of phenotypic developmental traits and hybrid vigour in *Arabidopsis thaliana*. *Indian J. Biotechnol.* 19 (2), 102–112.

Gupta, A., and Tyagi, T. (2022). "Phytoremediation and therapeutic potential of neglected plants: An invasive aquatic weeds and ornamental plant," in *Biotechnological innovations for environmental bioremediation* (Singapore: Springer), 259–290.

Hayashi, E., You, Y., Lewis, R., Calderon, M. C., Wan, G., and Still, D. W. (2012). Mapping QTL, epistasis and genotype× environment interaction of antioxidant activity, chlorophyll content and head formation in domesticated lettuce (*Lactuca sativa*). *TheorApplGenet* 124 (8), 1487–1502. doi:10.1007/s00122-012-1803-0

Hu, J., Rampitsch, C., and Bykova, N. V. (2015). Advances in plant proteomics toward improvement of crop productivity and stress resistancex. *Front. Plant Sci.* 6, 209. doi:10.3389/fpls.2015.00209

Huang, H., Jiang, X., Xiao, Z., Yu, L., Pham, Q., Sun, J., et al. (2016). Red cabbage microgreens lower circulating low-density lipoprotein (LDL), liver cholesterol, and inflammatory cytokines in mice fed a high-fat diet. *J. Agric. Food Chem.* 64 (48), 9161–9171. doi:10.1021/acs.jafc.6b03805

Indurker, S., Misra, H. S., and Eapen, S. (2007). Genetic transformation of chickpea (*Cicerarietinum* L.) with insecticidal crystal protein gene using particle gun bombardment. *Plant Cell Rep.* 26 (6), 755–763. doi:10.1007/s00299-006-0283-6

Iorliam, I. B., Ikyo, B. A., Iorliam, A., Okube, E. O., Kwaghtyo, K. D., and Shehu, Y. I. (2021). Application of machine learning techniques for Okra shelf-life prediction. *J. Data Anal. Info Process.* 9 (3), 136–150. doi:10.4236/jdaip.2021.93009

Jo, Y., Lee, S., Lee, Y., Kahng, H., Park, S., Bae, S., et al. (2021). Semantic segmentation of cabbage in the South Korea highlands with images by unmanned aerial vehicles. *Appl. Sci.* 11 (10), 4493. doi:10.3390/app11104493

Johnson, S. A., Prenni, J. E., Heuberger, A. L., Isweiri, H., Chaparro, J. M., Newman, S. E., et al. (2021). Comprehensive evaluation of metabolites and minerals in 6 microgreen species and the influence of maturity. *Curr. Dev. Nutr.* 5 (2), nzaa180–12. doi:10.1093/cdn/nzaa180

Jones-Baumgardt, C., Llewellyn, D., Ying, Q., and Zheng, Y. (2019). Intensity of sole-source light-emitting diodes affects growth, yield, and quality of *Brassicaceae* microgreens. *Hort. Sci.* 54 (7), 1168–1174. doi:10.21273/HORTSCI13788-18

Jung, M., Song, J. S., Hong, S., Kim, S., Go, S., Lim, Y. P., et al. (2021). Deep learning algorithms correctly classify *Brassica rapa* varieties using digital images. *Front. Plant Sci.* 12, 738685. doi:10.3389/fpls.2021.738685

Kale, S. M., Jaganathan, D., Ruperao, P., Chen, C., Punna, R., Kudapa, H., et al. (2015). Prioritization of candidate genes in "QTL-hotspot" region for drought tolerance in chickpea (*Cicer arietinum* L.). *Sci. Rep.* 5 (1), 15296–15314. doi:10.1038/srep15296

Kalia, P., and Singh, S. (2018). "Breeding for improving nutritional qualities and shelf life in vegetable crops," in *Advances in postharvest technologies of vegetable crops* (Canada: Apple Academic Press), 127–169.

Katam, K., Jones, K. A., and Sakata, K. (2015). Advances in proteomics and bioinformatics in agriculture research and crop improvement. *J. Proteomics Bioinform* 8 (3), 39. doi:10.4172/2Fjpb.1000372

Kaur, N., Singh, B., Kaur, A., and Yadav, M. P. (2022). Impact of growing conditions on proximate, mineral, phenolic composition, amino acid profile, and antioxidant properties of black gram, mung bean, and chickpea microgreens. *J. Food Process Preserv* 46, e16655. doi:10.1111/jfpp.16655

Khoja, K., Buckley, A., Aslam, F. M., Sharp, A. P., and Latunde-Dada, G. O. (2020). *In vitro* bioaccessibility and bioavailability of iron from mature and microgreen Fenugreek, Rocket and Broccoli. *Nutrients* 12 (4), 1057. doi:10.3390/nu12041057

Khokhar, W., Hassan, M. A., Reddy, A. S., Chaudhary, S., Jabre, I., Byrne, L. J., et al. (2019). Genome-wide identification of splicing quantitative trait loci (sQTLs) in diverse ecotypes of *Arabidopsis thaliana*. *Front plant sci* 10, 1160. doi:10.3389/fpls.2019.01160

Kobori, C. N., and Amaya, D. B. R. (2008). Uncultivated Brazilian green leaves are richer sources of carotenoids than are commercially produced leafy vegetables. *Food Nutr. Bull.* 29 (4), 320–328. doi:10.1177/156482650802900408

Kong, Y., Kamath, D., and Zheng, Y. (2019). Blue versus red light can promote elongation growth independent of photoperiod: A study in four *Brassica* microgreens species. *HortSci* 54 (11), 1955–1961. doi:10.21273/HORTSCI14286-19

Kou, L., Luo, Y., Yang, T., Xiao, Z., Turner, E. R., Lester, G. E., et al. (2013). Postharvest biology, quality and shelf life of buckwheat microgreens. *LWT-Food Sci. Tech.* 51 (1), 73–78. doi:10.1016/j.lwt.2012.11.017

Kou, L., Yang, T., Liu, X., and Luo, Y. (2015). Effects of pre-and postharvest calcium treatments on shelf life and postharvest quality of broccoli microgreens. *HortSci* 50 (12), 1801–1808. doi:10.21273/HORTSCI.50.12.1801

Kou, L., Yang, T., Luo, Y., Liu, X., Huang, L., and Codling, E. (2014). Pre-harvest calcium application increases biomass and delays senescence of broccoli microgreens. *Postharvest Biol. Technol.* 87, 70–78. doi:10.1016/j.postharvbio.2013.08.004

KraljCigić, I., Rupnik, S., Rijavec, T., PoklarUlrih, N., and Cigić, B. (2020). Accumulation of agmatine, spermidine, and spermine in sprouts and microgreens of alfalfa, fenugreek, lentil, and daikon radish. *Foods* 9 (5), 547–567. doi:10.3390/foods9050547

Krishnappa, G., Khan, H., Krishna, H., Kumar, S., Mishra, C. N., Parkash, O., et al. (2022). Genetic dissection of grain iron and zinc, and thousand kernel weight in wheat (*Triticum aestivum* L.) using genome-wide association study. *Sci. Rep.* 12 (1), 12444–12514. doi:10.1038/s41598-022-15992-z

Kumar, S., Hash, C. T., Thirunavukkarasu, N., Singh, G., Rajaram, V., Rathore, A., et al. (2016). Mapping quantitative trait loci controlling high iron and zinc content in self and open pollinated grains of pearl millet [*Pennisetum glaucum* (L.) R. Br.]. *Front Plant Sci* 7, 1636. doi:10.3389/fpls.2016.01636

Kumar, R., Singh, R. K., Misra, J., Yadav, A., Kumar, A., Yadav, R., et al. (2022a). Dissecting proteomic estimates for enhanced bioavailable nutrition during varied stages of germination and identification of potential genotypes in chickpea. *Legume Res.* 45 (9), 1082–1087 doi:10.18805/LR-4531

Kumar, S., Seem, K., Kumar, S., Vinod, K. K., Chinnusamy, V., and Mohapatra, T. (2022b). Pup1 QTL regulates gene expression through epigenetic modification of DNA under phosphate starvation stress in rice. *Front. Plant Sci.* 13, 871890. doi:10.3389/fpls.2022.871890

Kumari, C., Sharma, M., Kumar, V., Sharma, R., Kumar, V., Sharma, P., et al. (2022). Genome editing technology for genetic amelioration of fruits and vegetables for alleviating post-harvest loss. *Bioengineering* 9 (4), 176. doi:10.3390/bioengineering9040176

Kusumam, K., Krajník, T., Pearson, S., Duckett, T., and Cielniak, G. (2017). 3D-vision based detection, localization, and sizing of broccoli heads in the field. *J. Field Robot.* 34 (8), 1505–1518. doi:10.1002/rob.21726

Kuzmina, Y. (2020). Methods of genome editing for increasing the shelf life of tomato fruit. *Plant Biotechnol. Breed.* 3, 31–39. doi:10.30901/2658-6266-2020-1-o6

Kyriacou, M. C., De Pascale, S., Kyratzis, A., and Rouphael, Y. (2017). Microgreens as a component of space life support systems: A cornucopia of functional food. *Front. Plant Sci.* 1587, 1587. doi:10.3389/fpls.2017.01587

Kyriacou, M. C., El-Nakhel, C., Graziani, G., Pannico, A., Soteriou, G. A., Giordano, M., et al. (2019a). Functional quality in novel food sources: Genotypic variation in the nutritive and phytochemical composition of thirteen microgreens species. *Food Chem.* 277, 107–118. doi:10.1016/j.foodchem.2018.10.098

Kyriacou, M. C., El-Nakhel, C., Pannico, A., Graziani, G., Soteriou, G. A., Giordano, M., et al. (2019b). Genotype-specific modulatory effects of select spectral bandwidths on the nutritive and phytochemical composition of microgreens. *Front. Plant Sci.* 10, 1501. doi:10.3389/fpls.2019.01501

Kyriacou, M. C., El-Nakhel, C., Pannico, A., Graziani, G., Zarrelli, A., Soteriou, G. A., et al. (2021a). Ontogenetic variation in the mineral, phytochemical and yield attributes of brassicaceous microgreens. *Foods* 10 (5), 1032. doi:10.3390/foods10051032

Kyriacou, M. C., El-Nakhel, C., Soteriou, G. A., Graziani, G., Kyratzis, A., Antoniou, C., et al. (2021b). Pre harvest nutrient deprivation reconfigures nitrate, mineral, and phytochemical content of microgreens. *Foods* 10 (6), 1333. doi:10.3390/foods10061333

Kyriacou, M. C., Rouphael, Y., Di Gioia, F., Kyratzis, A., Serio, F., Renna, M., et al. (2016). Micro-scale vegetable production and the rise of microgreens. *Trends Food Sci. Technol.* 57, 103–115. doi:10.1016/j.tifs.2016.09.005

Langridge, P., and Fleury, D. (2011). Making the most of 'omics' for crop breeding. *Trends Biotech.* 29 (1), 33–40. doi:10.1016/j.tibtech.2010.09.006

Laurent, S., Forge, D., Port, M., Roch, A., Robic, C., Vander Elst, L., et al. (2008). Magnetic iron oxide nanoparticles: Synthesis, stabilization, vectorization, physicochemical characterizations, and biological applications. *Chem. Rev.* 108 (6), 2064–2110. doi:10.1021/cr068445e

Le, T. N., Chiu, C. H., and Hsieh, P. C. (2020). Bioactive compounds and bioactivities of *Brassicaoleracea* L. var. italica sprouts and microgreens: An updated overview from a nutraceutical perspective. *Plants* 9 (8), 946–969. doi:10.3390/plants9080946

Lee, C., Lee, J., and Lee, J. (2022). Relationship of fruit color and anthocyanin content with related gene expression differ in strawberry cultivars during shelf life. *SciHortic* 301, 111109. doi:10.1016/j.scienta.2022.111109

Li, J., Dai, X., Liu, T., and Zhao, P. X. (2012). Legume IP: An integrative database for comparative genomics and transcriptomics of model legumes. *Nucl. Acids Res.* 40 (D1), D1221–D1229. doi:10.1093/nar/gkr939

Li, Z., Liu, Y., Li, L., Fang, Z., Yang, L., Zhuang, M., et al. (2019). Transcriptome reveals the gene expression patterns of sulforaphane metabolism in broccoli florets. *PLoS One* 14 (3), 0213902. doi:10.1371/journal.pone.0213902

Lipkie, T. E., De Moura, F. F., Zhao, Z-Y., Albertsen, M. C., Che, P., Glassman, K., et al. (2013). Bioaccessibility of carotenoids from transgenic provitamin A biofortified Sorghum. *J. Agric. Food Chem.* 61 (24), 5764–5771. doi:10.1021/jf305361s

Liu, W., Min, X., and Wang, Y. (2019). Genome-wide development of microRNA-based SSR markers in *Medicago truncatula* with their transferability analysis and utilization in related legume species. *Model Legume Medicago truncatula* 2019, 936–945. doi:10.1002/9781119409144.ch121

Lobiuc, A., Vasilache, V., Oroian, M., Stoleru, T., Burducea, M., Pintilie, O., et al. (2017). Blue and red LED illumination improves growth and bioactive compounds contents in acyanic and cyanic *Ocimumbasilicum* L. microgreens. *Molecules* 22 (12), 2111. doi:10.3390/molecules22122111

Lu, Y., Dong, W., Alcazar, J., Yang, T., Luo, Y., Wang, Q., et al. (2018). Effect of pre harvest $CaCl_2$ spray and postharvest UV-B radiation on storage quality of Broccoli microgreens, a richer source of glucosinolates. *J. Food Compost Anal.* 67, 55–62. doi:10.1016/j.jfca.2017.12.035

Mahanta, S., Habib, M. R., and Moore, J. M. (2022). Effect of high-voltage atmospheric cold plasma treatment on germination and heavy metal uptake by soybeans (*Glycine max*). *Int. J. Mol. Sci.* 23 (3), 1611. doi:10.3390/ijms23031611

Mahto, R. K., Singh, C., Chandana, B. S., Singh, R. K., Verma, S., Gahlaut, V., et al. (2022). Chickpea biofortification for cytokinin dehydrogenase via genome editing to enhance abiotic-biotic stress tolerance and food security. *Front. Genet.* 13, 900324. doi:10.3389/fgene.2022.900324

Maina, S., Ryu, D. H., Cho, J. Y., Jung, D. S., Park, J. E., Nho, C. W., et al. (2021). Exposure to salinity and light spectra regulates glucosinolates, phenolics, and antioxidant capacity of *Brassica carinata* L. *Microgreens. Antioxidants.* 10 (8), 1183. doi:10.3390/antiox10081183

Maluin, F. N., Hussein, M. Z., Nik Ibrahim, N. N. L., Wayayok, A., and Hashim, N. (2021). Some emerging opportunities of nanotechnology development for soilless and microgreen farming. *Agronomy* 11 (6), 1213. doi:10.3390/agronomy11061213

Mannur, D. M., Babbar, A., Thudi, M., Sabbavarapu, M. M., Roorkiwal, M., Yeri, S. B., et al. (2019). Super Annigeri 1 and improved JG 74: Two *Fusarium* wilt-resistant introgression lines developed using marker-assisted backcrossing approach in chickpea (*Cicer arietinum* L.). *Mol. Breed.* 39 (1), 2–13. doi:10.1007/s11032-018-0908-9

Martiniakova, M., Babikova, M., Mondockova, V., Blahova, J., Kovacova, V., and Omelka, R. (2022). The role of macronutrients, micronutrients and flavonoid polyphenols in the prevention and treatment of osteoporosis. *Nutrients* 14 (3), 523. doi:10.3390/nu14030523

Mathiazhagan, M., Chidambara, B., Hunashikatti, L. R., and Ravishankar, K. V. (2021). Genomic approaches for improvement of tropical fruits: Fruit quality, shelf life and nutrient content. *Genes* 12 (12), 1881. doi:10.3390/genes12121881

Maurya, R., Gupta, A., Singh, S. K., Rai, K. M., Katiyar, R., Sawant, S. V., et al. (2015). Genomic-derived microsatellite markers for diversity analysis in *Jatropha curcas*. *Trees* 29 (3), 849–858. doi:10.1007/s00468-015-1166-7

Meas, S., Luengwilai, K., and Thongket, T. (2020). Enhancing growth and phytochemicals of two amaranth microgreens by LEDs light irradiation. *Sci. Hortic.* 265, 109204. doi:10.1016/j.scienta.2020.109204

Mishra, G. P., Dikshit, H. K., Tontang, M. T., Stobdan, T., Sangwan, S., Aski, M. S., et al. (2021). Diversity in phytochemical composition, antioxidant capacities, and nutrient contents among Mungbean and Lentil microgreens when grown at plain-altitude region (Delhi) and high-altitude region (Leh-Ladakh), India. *Front. Plant Sci.* 12, 710812. doi:10.3389/fpls.2021.710812

Mlinarić, S., Gvozdić, V., Vuković, A., Varga, M., Vlašiček, I., Cesar, V., et al. (2020). The effect of light on antioxidant properties and metabolic profile of chia Microgreens. *Appl. Sci.* 10 (17), 5731. doi:10.3390/app10175731

Muchjajib, U., Muchjajib, S., Suknikom, S., and Butsai, J. (2015). Evaluation of organic media alternatives for the production of microgreens in Thailand. *Acta Hortic.* 1102, 157–162. doi:10.17660/ActaHortic.2015.1102.19

Nguyen, T. H. A., Nguyen, V. C., Phan, T. N. H., Vasseghian, Y., Trubitsyn, M. A., Nguyen, A. T., et al. (2022). Novel biogenic silver and gold nanoparticles for multifunctional applications: Green synthesis, catalytic and antibacterial activity, and colorimetric detection of Fe (III) ions. *Chemosphere* 287, 132271. doi:10.1016/j.chemosphere.2021.132271

Niroula, A., Khatri, S., Timilsina, R., Khadka, D., Khadka, A., and Ojha, P. (2019). Profile of chlorophylls and carotenoids of wheat (*Triticum aestivum* L.) and barley (*Hordeum vulgare* L.) microgreens. *J. Food Sci. Tech.* 56 (5), 2758–2763. doi:10.1007/s13197-019-03768-9

Oliva, M., Demanelis, K., Lu, Y., Chernoff, M., Jasmine, F., Ahsan, H., et al. (2022). DNA methylation QTL mapping across diverse human tissues provides molecular links between genetic variation and complex traits. *Nat. Genet.* 2022, 1–11. doi:10.1038/s41588-022-01248-z

Pandey, A., Chakraborty, S., Datta, A., and Chakraborty, N. (2008). Proteomics approach to identify dehydration responsive nuclear proteins from chickpea (*Cicer arietinum* L.). *Mol. Cell Proteom* 7 (1), 88–107. doi:10.1074/mcp.M700314-MCP200

Pannico, A., El-Nakhel, C., Graziani, G., Kyriacou, M. C., Giordano, M., Soteriou, G. A., et al. (2020). Selenium biofortification impacts the nutritive value, polyphenolic content, and bioactive constitution of variable microgreens genotypes. *Antioxidants* 9 (4), 272. doi:10.3390/antiox9040272

Paradiso, V. M., Castellino, M., Renna, M., Gattullo, C. E., Calasso, M., Terzano, R., et al. (2018). Nutritional characterization and shelf-life of packaged microgreens. *Food Funct.* 9 (11), 5629–5640. doi:10.1039/C8FO01182F

Parmar, S., Sharma, V., Bomireddy, D., Soni, P., Joshi, P., Gangurde, S. S., et al. (2022). Recent advances in genetics, genomics, and breeding for nutritional quality in groundnut, *Accel. Plant Breed.* 4, 111–137. doi:10.1007/978-3-030-81107-5_4

Pavan, M. P., and Gangaprasad, S. (2022). Studies on mode of gene action for fruit quality characteristics governing shelf life in tomato (*Solanum lycopersicum* L.). *Sci. Hortic.* 293, 110687. doi:10.1016/j.scienta.2021.110687

Petropoulos, S. A., El-Nakhel, C., Graziani, G., Kyriacou, M. C., and Rouphael, Y. (2021). The effects of nutrient solution feeding regime on yield, mineral profile, and phytochemical composition of spinach microgreens. *Horticulturae* 7 (7), 162. doi:10.3390/horticulturae7070162

Podse dek, A., Sosnowska, D., Redzynia, M., and Anders, B. (2006). Antioxidant capacity and content of *Brassica oleracea* dietary antioxidants. *Int. J. Food Sci. Technol.* 41 (1), 49–58. doi:10.1111/j.1365-2621.2006.01260.x

Pradeepkumara, N., Sharma, P. K., Munshi, A. D., Behera, T. K., Bhatia, R., Kumari, K., et al. (2022). Fruit transcriptional profiling of the contrasting genotypes for shelf life reveals the key candidate genes and molecular pathways regulating post-harvest biology in cucumber. *Genomics* 114 (2), 110273. doi:10.1016/j.ygeno.2022.110273

Qi, T., Wu, Y., Fang, H., Zhang, F., Liu, S., Zeng, J., et al. (2022). Genetic control of RNA splicing and its distinct role in complex trait variation. *Nat. Genet.* 54 (9), 1355–1363. doi:10.1038/s41588-022-01154-4

Qin, G., Wu, S., Zhang, L., Li, Y., Liu, C., Yu, J., et al. (2022). An efficient modular gateway recombinase-based gene stacking system for generating multi-trait transgenic plants. *Plants* 11 (4), 488. doi:10.3390/plants11040488

Reda, T., Thavarajah, P., Polomski, R., Bridges, W., Shipe, E., and Thavarajah, D. (2021). Reaching the highest shelf: A review of organic production, nutritional quality, and shelf life of kale (*Brassica oleracea* var. acephala). *Plants, People, Planet* 3 (4), 308–318. doi:10.1002/ppp3.10183

Roorkiwal, M., Bharadwaj, C., Barmukh, R., Dixit, G. P., Thudi, M., Gaur, P. M., et al. (2020). Integrating genomics for chickpea improvement: Achievements and opportunities. *Theor. Appl. Genet.* 133 (5), 1703–1720. doi:10.1007/s00122-020-03584-2

Saengha, W., Karirat, T., Buranrat, B., Matra, K., Deeseenthum, S., Katisart, T., et al. (2021). Cold plasma treatment on mustard green seeds and its effect on growth, isothiocyanates, antioxidant activity and anticancer activity of microgreens. *Int. J. Agric. Biol.* 25 (33), 667–676. doi:10.17957/IJAB/15.1715

Sahoo, M. R., Naresh, P., Kumari, M., and Acharya, G. C. (2022). "Omics in leafy vegetables: Genomics, transcriptomics, proteomics, metabolomics, and multiomics approaches," in *Omics in horticultural crops* (United States: Academic Press), 281–302.

Samuolienė, G., Brazaitytė, A., Viršilė, A., Miliauskienė, J., Vaštakaitė-Kairienė, V., and Duchovskis, P. (2019). Nutrient levels in *Brassicaceae* microgreens increase under tailored light-emitting diode spectra. *Front. Plant Sci.* 10, 1475. doi:10.3389/fpls.2019.01475

Samuolienė, G., Viršilė, A., Brazaitytė, A., Jankauskienė, J., Sakalauskienė, S., Vaštakaitė, V., et al. (2017). Blue light dosage affects carotenoids and tocopherols in microgreens. *Food Chem.* 228, 50–56. doi:10.1016/j.foodchem.2017.01.144

Sanderson, L. A., Caron, C. T., Tan, R., Shen, Y., Liu, R., and Bett, K. E. (2019). KnowPulse: A web-resource focused on diversity data for pulse crop improvement. *Front. Plant Sci.* 10, 965. doi:10.3389/fpls.2019.00965

Sangwan, S., Kukreja, B., Mishra, G. P., Dikshit, H. K., Singh, A., Aski, M., et al. (2022). Yield optimization, microbial load analysis, and sensory evaluation of mungbean (*Vigna radiata* L.), lentil (*Lens culinaris* subsp. *culinaris*), and Indian mustard (*Brassica juncea* L.) microgreens grown under greenhouse conditions. *Plos one* 17 (5), 0268085. doi:10.1371/journal.pone.0268085

Schmidt, M. A., Parrott, W. A., Hildebrand, D. F., Berg, R. H., Cooksey, A., Pendarvis, K., et al. (2015). Transgenic soyabean seeds accumulating β-carotene exhibit the collateral enhancements of oleate and protein content traits. *Plant Biotech. J.* 13 (4), 590–600. doi:10.1111/pbi.12286

Sharma, K. K., Lavanya, M., and Anjaiah, V. (2006). *Agrobacterium*-mediated production of transgenic pigeon pea (*Cajanus cajan* L. Mill sp.) expressing the synthetic Bt cry1Ab gene. *Vitro Cell Dev. Biol. Plant* 42 (2), 165–173. doi:10.1079/IVP2005730

Sharma, S., Shree, B., Sharma, D., Kumar, S., Kumar, V., Sharma, R., et al. (2022). Vegetable microgreens: The gleam of next generation super foods, their genetic enhancement, health benefits and processing approaches. *Food Res. Int.* 155, 111038. doi:10.1016/j.foodres.2022.111038

Sharma, T., and Gupta, A. (2022). "Integrated OMICS approaches to ameliorate the abiotic stress in," in *Plant stress: Challenges and management in the new decade* (Cham: Springer), 361–373.

Shelar, A., Singh, A. V., Dietrich, P., Maharjan, R. S., Thissen, A., Didwal, P. N., et al. (2022). Emerging cold plasma treatment and machine learning prospects for seed priming: A step towards sustainable food production. *RSC Adv.* 12 (17), 10467–10488. doi:10.1039/D2RA00809B

Singh, K., Kumar, S., and Kaur, P. (2019). Automatic detection of rust disease of Lentil by machine learning system using microscopic images. *Int. J. Elect. Com. Eng.* 9 (1), 660–666. doi:10.11591/ijece.v9i1.pp660-666

Singh, R. K., Singh, C., Chandana, B. S., Mahto, R. K., Patial, R., Gupta, A., et al. (2022). Exploring chickpea germplasm diversity for broadening the genetic base utilizing genomic resourses. *Front. Genet.* 1836, 905771. doi:10.3389/fgene.2022.905771

Spillane, C., Steimer, A., Grossniklaus, U., and GrossniklaUs, U. (2001). Apomixis in agriculture: The quest for clonal seeds. *Sex. Plant Reprod.* 14 (4), 179–187. doi:10.1007/s00497-001-0117-1

Sreenivasan, E. (2020). Preliminary report on multiple harvests of microgreens from Chickpea (*Cicer arietinum*) seeds. *Intl J. Trend Sci. Res. Dev.* 4, 306–308. IJTSRD33337.

Sun, J., Kou, L., Geng, P., Huang, H., Yang, T., Luo, Y., et al. (2015). Metabolomic assessment reveals an elevated level of glucosinolate content in $CaCl_2$ treated broccoli microgreens. *J. Agric. Food Chem.* 63 (6), 1863–1868. doi:10.1021/jf504710r

Sundaria, N., Singh, M., Upreti, P., Chauhan, R. P., Jaiswal, J. P., and Kumar, A. (2019). Seed priming with iron oxide nanoparticles triggers iron acquisition and biofortification in wheat (*Triticum aestivum* L.) grains. *J. Plant Growth Reg.* 38 (1), 122–131. doi:10.1007/s00344-018-9818-7

Tan, L., Nuffer, H., Feng, J., Kwan, S. H., Chen, H., Tong, X., et al. (2020). Antioxidant properties and sensory evaluation of microgreens from commercial and local farms. *Food Sci. Hum. Wellness.* 9 (1), 45–51. doi:10.1016/j.fshw.2019.12.002

Tantharapornrerk, N., Vichitsoonthonkul, T., Techavuthiporn, C., and Photchanachai, S. (2021). Growth and antioxidant system of Chinese kale microgreens in response to different illumination of light sources. *N. Z. J. Crop Hort. Sci.* 2021, 1–15. doi:10.1080/01140671.2021.1958876

Thuong, V. T., and Minh, H. G. (2020). Effects of growing substrates and seed density on yield and quality of radish (*Raphanus sativus*) microgreens. *Res. Crops* 21 (3), 579–586. doi:10.31830/2348-7542.2020.091

Toscano, S., Cavallaro, V., Ferrante, A., Romano, D., and Patané, C. (2021). Effects of different light spectra on final biomass production and nutritional quality of two microgreens. *Plants* 10 (8), 1584. doi:10.3390/plants10081584

Tropea, A., Potortì, A. G., Lo Turco, V., Russo, E., Vadalà, R., Rando, R., et al. (2021). Aqua feed production from fermented fish waste and lemon peel. *Fermentation* 7 (4), 272. doi:10.3390/fermentation7040272

Truzzi, F., Whittaker, A., Roncuzzi, C., Saltari, A., Levesque, M. P., and Dinelli, G. (2021). Microgreens: Functional food with anti proliferative cancer properties influenced by light. *Foods* 10 (8), 1690. doi:10.3390/foods10081690

Tsaftaris, A. S., and Polidoros, A. N. (2000). DNA methylation and plant breeding. *Plant Breed. Rev.* 18, 87–176. doi:10.1002/9780470650158.ch3

Turner, E. R., Luo, Y., and Buchanan, R. L. (2020). Microgreen nutrition, food safety, and shelf life: A review. *J. Food Sci.* 85 (4), 870–882. doi:10.1111/1750-3841.15049

Upadhyaya, C. P., Young, K. E., Akula, N., Soon Kim, H., Heung, J. J., Oh, O., et al. (2009). Over-expression of strawberry d-galacturonic acid reductase in potato leads to accumulation of vitamin C with enhanced abiotic stress tolerance. *Plant Sci.* 177 (6), 659–667. doi:10.1016/j.plantsci.2009.08.004

United States Department of Agriculture (2014). Specialty greens pack a nutritionalpunch, *AgResearch Magazine*. Available at: http://agresearchmag.ars.usda.gov/2014/jan/greens.

Valdes, A., Álvarez-Rivera, G., Socas-Rodríguez, B., Herrero, M., Ibáñez, E., and Cifuentes, A. (2021). Foodomics: Analytical opportunities and challenges. *Anal. Chem.* 94 (1), 366–381. doi:10.1021/acs.analchem.1c04678

Vanderschuren, H., Lentz, E., Zainuddin, I., and Gruissem, W. (2013). Proteomics of model and crop plant species: Status, current limitations and strategic advances for crop improvement. *J. Proteomics* 93, 5–19. doi:10.1016/j.jprot.2013.05.036

Varshney, R. K. (2016). Exciting journey of 10 years from genomes to fields and markets: Some success stories of genomics-assisted breeding in chickpea, pigeonpea and groundnut. *Plant Sci.* 242, 98–107. doi:10.1016/j.plantsci.2015.09.009

Varshney, R. K., Pandey, M. K., Bohra, A., Singh, V. K., Thudi, M., and Saxena, R. K. (2019). Toward the sequence-based breeding in legumes in the post-genome sequencing era. *TheorAppl Genet.* 132 (3), 797–816. doi:10.1007/s00122-018-3252-x

Varshney, R. K., Roorkiwal, M., Sun, S., Bajaj, P., Chitikineni, A., Thudi, M., et al. (2021). A chickpea genetic variation map based on the sequencing of 3,366 genomes. *Nature* 599 (7886), 622–627. doi:10.1038/s41586-021-04066-1

Varshney, R. K., Thudi, M., Pandey, M. K., Tardieu, F., Ojiewo, C., Vadez, V., et al. (2018). Accelerating genetic gains in legumes for the development of prosperous smallholder agriculture: Integrating genomics, phenotyping, systems modelling and agronomy. *J. Exp. Bot.* 69 (13), 3293–3312. doi:10.1093/jxb/ery088

Vasõtakaitė, V., Virsõilė, A., Brazaitytė, A., Samuolienė, G., Jankauskienė, J., Novicõkovas, A., et al. (2017). Pulsed light-emitting diodes for a higher phytochemical level in microgreens. *J. Agri Food Chem.* 65 (31), 6529–6534. doi:10.1021/acs.jafc.7b01214

Verlinden, S. (2020). Microgreens: Definitions, product types, and production practices. *Hortic. Rev.* 47, 85–124. doi:10.1002/9781119625407.ch3

Verma, M., Kumar, V., Patel, R. K., Garg, R., and Jain, M. (2015). CTDB: An integrated chickpea transcriptome database for functional and applied genomics. *PLoS ONE* 10 (8), e0136880. doi:10.1371/journal.pone.0136880

Weber, C. F. (2017). Broccoli microgreens: A mineral-rich crop that can diversify food systems. *Front. Nutr.* 4, 7. doi:10.3389/fnut.2017.00007

Wolanin, A., Mateo-García, G., Camps-Valls, G., Gómez-Chova, L., Meroni, M., Duveiller, G., et al. (2020). Estimating and understanding crop yields with explainable deep learning in the Indian Wheat Belt. *Environ. Res. Lett.* 15 (2), 024019. doi:10.1088/1748-9326/ab68ac

Wu, J., Yuan, Y. X., Zhang, X. W., Zhao, J., Song, X., Li, Y., et al. (2008). Mapping QTLs for mineral accumulation and shoot dry biomass under different Zn nutritional conditions in Chinese cabbage (*Brassica rapa* L. ssp. pekinensis). *Plant Soil* 310 (1), 25–40. doi:10.1007/s11104-008-9625-1

Wu, X., Islam, A. F., Limpot, N., Mackasmiel, L., Mierzwa, J., Cortés, A. J., et al. (2020). Genome-wide Snp identification and association mapping for seed mineral concentration in mung bean (*Vigna Radiata* L.). *Front. Genet.* 11, 656. doi:10.3389/fgene.2020.00656

Xiao, Z., Lester, G. E., Luo, Y., and Wang, Q. (2012). Assessment of vitamin and carotenoid concentrations of emerging food products: Edible microgreens. *J. Agri Food Chem.* 60 (31), 7644–7651. doi:10.1021/jf300459b

Xiao, Z., Lester, G. E., Park, E., Saftner, R. A., Luo, Y., and Wang, Q. (2015). Evaluation and correlation of sensory attributes and chemical compositions of emerging fresh produce: Microgreens. *Postharvest Biol. Tech.* 110, 140–148. doi:10.1016/j.postharvbio.2015.07.021

Xiao, Z., Luo, Y., Lester, G. E., Kou, L., Yang, T., and Wang, Q. (2014). Postharvest quality and shelf life of radish microgreens as impacted by storage temperature, packaging film, and chlorine wash treatment. *LWT-Food Sci. Tech.* 55 (2), 551–558. doi:10.1016/j.lwt.2013.09.009

Ying, Q., Jones-Baumgardt, C., Zheng, Y., and Bozzo, G. (2021). The Proportion of blue light from light-emitting diodes alters microgreen phytochemical profiles in a species-specific manner. *HortSci* 56 (1), 13–20. doi:10.21273/HORTSCI15371-20

Ying, Q., Kong, Y., Jones-Baumgardt, C., and Zheng, Y. (2020a). Responses of yield and appearance quality of four *Brassicaceae* microgreens to varied blue light proportion in red and blue light-emitting diodes lighting. *Sci. Hortic.* 259, 108857. doi:10.1016/j.scienta.2019.108857

Ying, Q., Kong, Y., and Zheng, Y. (2020b). Applying blue light alone, or in combination with far-red light, during nighttime increases elongation without compromising yield and quality of indoor-grown microgreens. *HortSci* 55 (6), 876–881. doi:10.21273/hortsci14899-20

Zhang, X., Bian, Z., Li, S., Chen, X., and Lu, C. (2019). Comparative analysis of phenolic compound profiles, antioxidant capacities, and expressions of phenolic biosynthesis-related genes in soybean microgreens grown under different light spectra. *J Agri Food Chem* 67 (49), 13577–13588. doi:10.1021/acs.jafc.9b05594

Zhang, X., Wei, J., Tian, J., Li, N., Jia, L., Shen, W., et al. (2019). Enhanced anthocyanin accumulation of immature radish microgreens by hydrogen-rich water under short wavelength light. *Sci. Horti* 247, 75–85. doi:10.1016/j.scienta.2018.11.060

Zhao, Z. Y., Glassman, K., Sewalt, V., Wang, N., Miller, M., Chang, S., et al. (2003). "Nutritionally improved transgenic sorghum"," in *Plant Biotechnology2002 and beyond* (Netherlands: Springer), 413–416.

Zhou, C., Hu, J., Xu, Z., Yue, J., Ye, H., and Yang, G. (2020). A monitoring system for the segmentation and grading of broccoli head based on deep learning and neural networks. *Front. Plant Sci.* 11, 402. doi:10.3389/fpls.2020.00402

Zhou, Q., Fu, Z., Liu, H., Wang, J., Guo, Z., Zhang, X., et al. (2021). Mining novel kernel size-related genes by pQTL mapping and multi-omics integrative analysis in developing maize kernels. *Plant Biotechnol. J.* 19 (8), 1489–1491. doi:10.1111/pbi.13634

Zhu, F., Jadhav, S. S., Tohge, T., Salem, M. A., Lee, J. M., Giovannoni, J. J., et al. (2022). A comparative transcriptomics and eQTL approach identifies SlWD40 as a tomato fruit ripening regulator. *Plant Physiol.* 190 (1), 250–266. doi:10.1093/plphys/kiac200

Zou, L., Tan, W. K., Du, Y., Lee, H. W., Liang, X., Lei, J., et al. (2021). Nutritional metabolites in *Brassica rapa* subsp. chinensis var. parachinensis (choy sum) at three different growth stages: Microgreen, seedling and adult plant. *Food Chem.* 357, 129535. doi:10.1016/j.foodchem.2021.129535

frontiers | Frontiers in Genetics

TYPE Original Research
PUBLISHED 16 March 2023
DOI 10.3389/fgene.2023.1128182

OPEN ACCESS

EDITED BY
Milind B. Ratnaparkhe,
ICAR Indian Institute of Soybean Research, India

REVIEWED BY
Radhakrishnan T.,
Directorate of Groundnut Research (ICAR-DGR), India
Sushil Satish Chhapekar,
University of Missouri, United States

*CORRESPONDENCE
Manish K. Pandey,
Manish.Pandey@icrisat.org

SPECIALTY SECTION
This article was submitted to Plant Genomics,a section of the journal Frontiers in Genetics

RECEIVED 20 December 2022
ACCEPTED 20 February 2023
PUBLISHED 16 March 2023

CITATION
Gangurde SS, Pasupuleti J, Parmar S, Variath MT, Bomireddy D, Manohar SS, Varshney RK, Singam P, Guo B and Pandey MK (2023), Genetic mapping identifies genomic regions and candidate genes for seed weight and shelling percentage in groundnut.
Front. Genet. 14:1128182.
doi: 10.3389/fgene.2023.1128182

Genetic mapping identifies genomic regions and candidate genes for seed weight and shelling percentage in groundnut

Sunil S. Gangurde[1,2,3], Janila Pasupuleti[1], Sejal Parmar[1,2], Murali T. Variath[1], Deekshitha Bomireddy[1], Surendra S. Manohar[1], Rajeev K. Varshney[1,4], Prashant Singam[2], Baozhu Guo[3] and Manish K. Pandey[1,2]*

[1]Center of Excellence in Genomics & Systems Biology (CEGSB), International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Hyderabad, India, [2]Department of Genetics, Osmania University, Hyderabad, India, [3]USDA-ARS, Crops Genetics and Breeding Research Unit, Tifton, GA, United States, [4]State Agricultural Biotechnology Centre, Centre for Crop & Food Innovation, Food Futures Institute, Murdoch University, Murdoch, WA, Australia

Seed size is not only a yield-related trait but also an important measure to determine the commercial value of groundnut in the international market. For instance, small size is preferred in oil production, whereas large-sized seeds are preferred in confectioneries. In order to identify the genomic regions associated with 100-seed weight (HSW) and shelling percentage (SHP), the recombinant inbred line (RIL) population (Chico × ICGV 02251) of 352 individuals was phenotyped for three seasons and genotyped with an Axiom_*Arachis* array containing 58K SNPs. A genetic map with 4199 SNP loci was constructed, spanning a map distance of 2708.36 cM. QTL analysis identified six QTLs for SHP, with three consistent QTLs on chromosomes A05, A08, and B10. Similarly, for HSW, seven QTLs located on chromosomes A01, A02, A04, A10, B05, B06, and B09 were identified. *BIG SEED* locus and *spermidine synthase* candidate genes associated with seed weight were identified in the QTL region on chromosome B09. Laccase, fibre protein, lipid transfer protein, senescence-associated protein, and disease-resistant NBS-LRR proteins were identified in the QTL regions associated with shelling percentage. The associated markers for major-effect QTLs for both traits successfully distinguished between the small- and large-seeded RILs. QTLs identified for HSW and SHP can be used for developing potential selectable markers to improve the cultivars with desired seed size and shelling percentage to meet the demands of confectionery industries.

KEYWORDS

candidate gene discovery, peanut, diagnostic markers, high-density genotyping, Axiom_*Arachis* array

1 Introduction

Groundnut or peanut is a self-pollinated, allotetraploid (AABB) (2n = 4X = 40), and leguminous oilseed crop with ~2.7 GB genome size (Bertioli et al., 2019). Presently, groundnut is cultivated globally on 36.18 million hectares of area, yielding 71.68 million tonnes of pods in the year 2020 (FAOSTAT, 2020). Groundnut seeds contain the most nutritious components; 100 g of groundnuts contains proteins (16 g), oil or fat (49 g),

carbohydrate (26 g), and dietary fibres (9 g) (Parmar et al., 2022). The estimated demand of edible oil based on current population projections and *per capita* consumption is likely to be 240 metric tonnes by 2050, which is nearly twice the current requirement (Corley, 2009). Improvement of yield and quality traits are the major objectives of many groundnut breeding programs (Gangurde et al., 2019; Pandey et al., 2020a). Groundnut yield is influenced by hundred-seed weight (HSW), shelling percentage (SHP), and number of seeds per pod. Among them, HSW and seed number per pod are important allied traits that are positively correlated with groundnut yield. HSW is an important yield attributing trait that is positively correlated with yield per plant. Being a quantitative trait, seed weight is controlled by multiple genes and also influenced by the environment. Therefore, understanding the regulation of seed size has always been an important area of research for groundnut improvement.

In the pre-genomic era, efforts on genetic mapping for seed weight produced large QTL intervals, making identification of the key candidate genes very difficult (Varshney et al., 2009; Mondal & Badigannavar, 2019). QTLs for seed size and pod size in cultivated and wild relatives of groundnut were discovered using an advanced backcross population (Fonceka et al., 2012). Consistent QTLs for seed weight on chromosome A07 and B06 and for shelling percentage on A10 and B06 were identified using an SSR-based genetic map (Chen et al., 2017). During the last few years, since the availability of groundnut diploid genomes (Bertioli et al., 2016; Chen et al., 2016) and tetraploid genomes (Bertioli et al., 2019; Chen et al., 2019; Zhuang et al., 2019), several sequencing-based trait mapping efforts have been reported to fine map the genomic regions for key traits in groundnut. The discovered candidate genes include those associated with leaf rust and late leaf spot resistance (Pandey et al., 2017a), stem rot resistance (Dodia et al., 2019), fresh seed dormancy (Kumar et al., 2020), shelling percentage (Luo et al., 2019a), bacterial wilt resistance (Luo et al., 2019b), early leaf spot and late leaf spot resistance (Agarwal et al., 2018), tomato spotted wilt virus resistance (Agarwal et al., 2019), and yield-related traits (Jadhav et al., 2021). In the post-genomic era, a significant amount of genomic resources at the genome and transcriptome level have been developed in groundnut (Pandey et al., 2020a). The transcriptome map for subsp. *hypogaea* was developed to understand the differential expression of genes at various growth stages (Clevenger et al., 2016). Moreover, another gene expression atlas for subsp. *fastigiata* was developed for 20 tissues at various developmental stages, ranging from the seedling stage to the maturity stage (Sinha et al., 2020).

The development of high-density SNP chips in groundnut allowed the construction of high-density genetic maps that helped to saturate the large QTL intervals (Pandey et al., 2017b). SNP arrays were successfully used to dissect the yield-related traits (Pandey et al., 2020b), root-knot nematode resistance (Ballén-Taborda et al., 2019), stem rot resistance (Luo et al., 2020), late leaf spot resistance (Han et al., 2018; Chu et al., 2019; Zhang et al., 2020), fresh seed dormancy (Wang et al., 2022), leaf chlorophyll content (Zou et al., 2022), salinity tolerance (Zou et al., 2020a), background genome recovery during a marker-assisted backcross selection in groundnut (Shasidhar et al., 2020), and germplasm diversity analysis (Nabi et al., 2021). Specific locus amplified fragment sequencing (SLAF-seq)-based high-density genetic map and phenotyping under multiple environments identified a total of 27 QTLs for seed weight, seed length, and width (Zhang et al., 2019). Recently, US-based nested association mapping (NAM) populations genotyped with a 58K SNP array discovered the genomic regions associated with seed and pod weights in groundnut (Gangurde et al., 2020). The SSR and SNP array-based genetic map identified the major genomic region on chromosome B06 and homologous region on chromosome A07/B07 (Chavarro et al., 2020). It has also been reported that in the US mini-core collection, the QTL on chromosome A05 is conserved with major effects on groundnut seed size (Chu et al., 2020). Similarly, in Chinese germplasm, the QTL on chromosome A05 was identified for the seed number per pod (Chen et al., 2019). Recently, the 58K SNP array was used to identify genomic regions associated with the seed aspect ratio (length width ratio) using GWAS on the US mini-core and Korean germplasm (Zou et al., 2020b). Although a large number of studies have been conducted globally to identify the QTLs for groundnut seed weight, there are limited reports on candidate genes or diagnostic markers for genomic-assisted breeding to improve seed weight.

Therefore, in order to identify the genomic regions and candidate genes associated with HSW and SHP, we developed a recombinant inbred line (RIL) population (Chico × ICGV 02251). The parental genotypes included a large-seeded cultivar "ICGV 02251" and a small-seeded germplasm line "Chico." A 58K high-density SNP array was used to genotype the RIL population along with the two parents to construct a dense genetic map. The genetic map along with genotyping data and multiple seasons phenotyping data was used to identify the genomic regions associated with HSW and SHP in groundnut.

2 Materials and methods

2.1 Plant material and phenotyping

A RIL population (Chico × ICGV 02251) comprising 352 RILs was developed by crossing Chico and ICGV 02251 and advanced using the single-seed decent (SSD) method. Both parents were allotetraploid (AABB), *Arachis hypogaea* and subsp. *fastigiata*, and the male parent, ICGV 02251, a late-maturing, Virginia bunch, has a significantly higher HSW and larger pod size than the female parent. Chico is an early-maturing Spanish bunch; a selection from PI 268661 was released in 1973 by the United States Department of Agriculture (USDA) in Georgia, Virginia, and Oklahoma (Bailey and Hammons, 1975). The RIL population was phenotyped for three seasons at ICRISAT, Patancheru, Hyderabad (India) during post-rainy 2013–14 (S1), rainy 2014 (S2), and rainy 2019 (S3). During each season, the RILs and parental lines were planted in three replications with a spacing of 30 × 10 cm in two rows of 2 m, with standard agronomic practices. The weather data for monthly average high and low temperatures and rainfall (mm) for the years 2013, 2014, and 2019 are shown in Supplementary Figure S1. The weight of 100 mature groundnut seeds from each RIL was measured as hundred-seed weight (gm), while for shelling percentage, 100 gm pods were shelled and the seed weight from these pods (gm) was

divided by pod's weight and multiplied by 100 as a measure of the shelling percentage (%). The multi-season phenotypic data for SHP and HSW on 352 RILs and both parents were used for the identification of genomic regions associated with HSW and SHP.

2.2 DNA extraction and genotyping with an "Axiom_*Arachis*" array

DNA from 352 RILs and both parents was extracted using the NucleoSpin Plant II kit (Macharey-Nigel, Duren, Germany). The DNA quality was analysed on 0.8% agarose gel, and concentration was measured using a NanoDrop 8000 spectrophotometer (Thermo Scientific). An Affymetrix GeneTitan® platform was used to genotype the RIL population with the 58K SNPs "Axiom_*Arachis*" array (Pandey et al., 2017b). Initially, the target probes for 352 samples were used in at least 20 µL DNA, with a concentration of 10 ng/µL. The samples were then amplified, fragmented, and hybridized on the array chip, followed by single-base extension through DNA ligation and signal amplification, according to the procedure explained in the Affymetrix Axiom 2.0 Assay manual (axiom_2_assay_auto_workflow_user_guide.pdf) (Pandey et al., 2020a). Axiome_*Arachis* is an SNP array developed for genotyping genetic populations in groundnut for trait mapping and association mapping (Pandey et al., 2017b).

2.3 SNP allele calling and quality analysis

We used the "*Best Practices*" workflow to perform quality control (QC) analysis of samples to select only those that pass the QC test for further downstream analysis. The "*Sample QC*" workflow was used to produce genotype calls for the samples that passed the QC test. The "*Genotyping*" workflow was used to perform genotyping on the imported CEL files regardless of the sample QC matrix. Before making the genotyping calls, samples that did not pass the QC were removed as their inclusion may reduce the quality of the analysed results. Finally, the "*Summary Only*" workflow was used to produce a summary containing details on the intensities for the probe sets for use in copy number analysis tools. It also allows exporting the SNP data after the analysis is completed for downstream analysis. The genotyping data with a towere filtered for monomorphismtal of 58,233 SNPs for 352 RILs was extracted from Axiom analysis suit as explained in Pandey et al., 2017b (Supplementary Table S1).

2.4 Construction of a genetic map using RIL population (Chico × ICGV 02251)

The 58,233 SNPs were filtered for monomorphism and highly missing (>30%), and only the selected polymorphic 10,236 SNPs between parental genotypes, ICGV 02251, and Chico were retained. The selected SNPs were subjected to the chi-square (χ^2) test to determine the goodness-of-fit to the expected 1: 1 segregation ratio; highly distorted markers were filtered out and not considered for the linkage map construction. Finally, after stringent filtration, informative SNPs were used for the construction of a genetic map. The alleles of ICGV 02251 were coded as "AA," Chico as "BB," and heterozygotes as "H." JoinMap (v4.0) software was used for the construction of a genetic map. Kosambi's mapping function was used to estimate the genetic distance to convert the recombination frequencies into map distances in centimorgans (cMs) (Kosambi, 1944). A total of 20 linkage groups were constructed individually by applying the LOD score (logarithm of the odds) with an LOD threshold of 3.0, and the recombination frequency (*rf*) threshold (∂) was set to 50%. MapChart software was used to finalize the marker position with the final genetic map (Voorrips, 2002).

2.5 Identification of QTLs for seed weight and shelling percentage

Multi-season phenotyping data for seed weight and shelling percentage generated during S1, S2, and S3 was used with genetic map information along with genotyping data for QTL analysis. The inclusive composite interval mapping-additive (ICIM-ADD) algorithm implemented in inclusive composite interval mapping (ICIM) software was used for identification of the main-effect QTLs (Meng et al., 2015). Epistatic QTLs for seed weight and shelling percentage were identified to understand the combined effect of any two genomic regions on seed weight and shelling percentage. QTLs with >10% phenotypic variance explained (PVE) were considered major QTLs; the remaining were considered minor QTLs. Realizing that the groundnut seed weight is a complex trait and also affected by the environment, we carried out epistatic QTL (Q × Q) and environment effect QTL (E-QTL) analysis in ICIM. The ICIM-EPI algorithm from ICIM software was used for epistatic QTL analysis. The environmental-effect QTLs were identified using multi-environment trails (METs) in ICIM. The LOD threshold score of 3.0 was used as the minimum significance level for the main effect QTLs, epistatic QTLs, and environmental-effect QTLs. The major QTL regions were validated using extreme RILs for both HSW and SHP from RIL population. A total of 15 extremely small-seeded and 15 extremely large-seeded lines were used for the validation of QTL regions. In order to validate the major QTLs, the RILs distinguished by the alleles of the flanking markers of QTLs were compared with the mean values of phenotypes.

2.6 Identification and expression analysis of candidate genes in QTL regions

Major QTL regions identified with >10% PVE identified for both seed weight and shelling percentage were targeted for candidate gene discovery. Candidate genes associated were mined in the QTL interval between the positions of flanking markers on the physical map of groundnut genome (https://peanutbase.org/). The expression data for the candidate genes were accessed from the *Arachis hypogaea* gene expression atlas for subsp. *fastigiata* (Sinha et al., 2020). The heatmaps for the expression data were generated using the R-software package "Pheatmap" (Kolde, 2019).

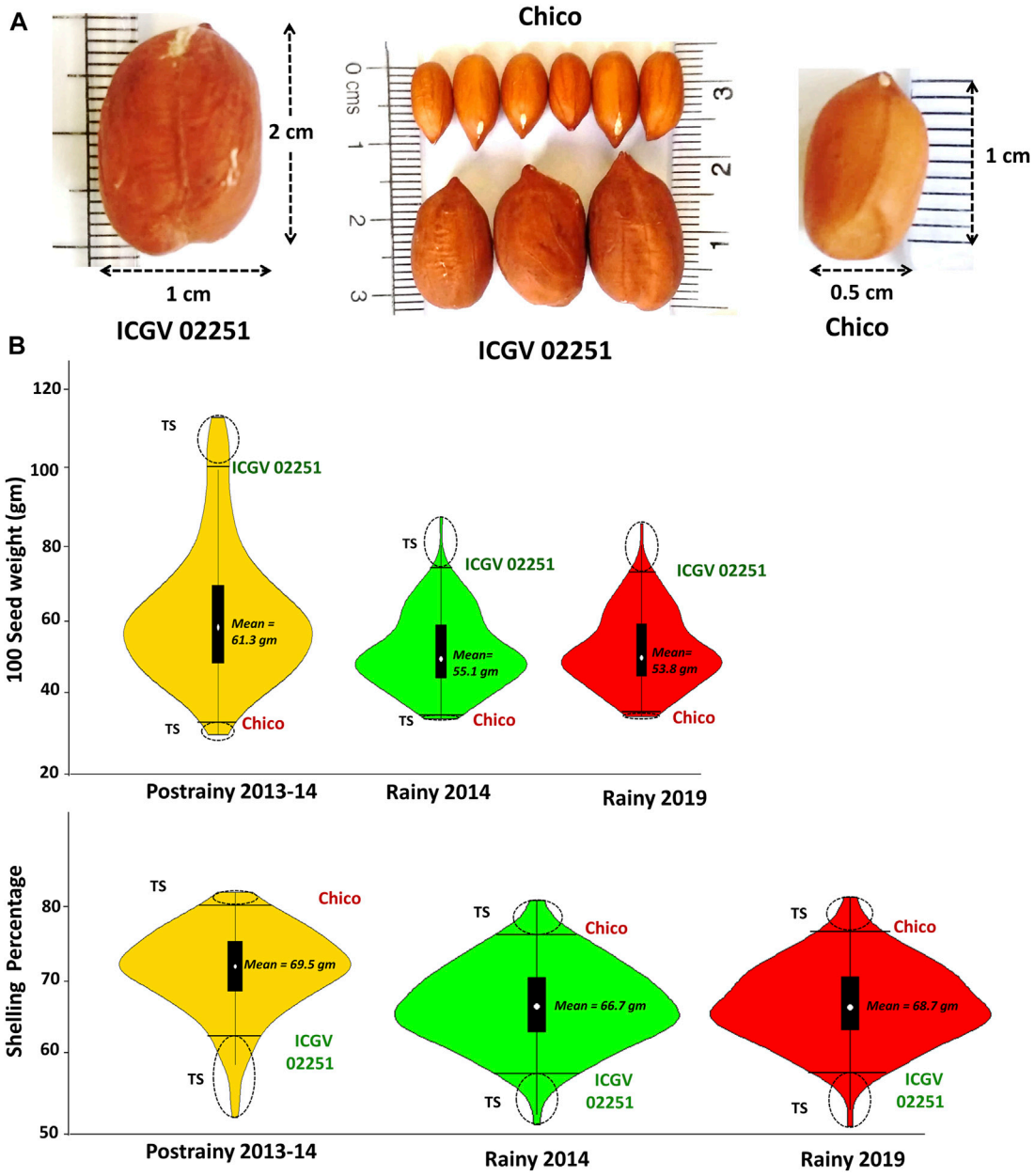


FIGURE 1
Variation in the phenotypic data generated for **(A)** hundred seed weight and **(B)** shelling percentage in the RIL population (Chico × ICGV 02251) during three seasons (Postrainy 2013-14, Rainy 2014 and Rainy 2019).

3 Results

3.1 Phenotypic variation for seed weight and shelling percentage in RIL population

Multi-season phenotyping data was generated during three seasons on RIL population (Chico × ICGV 02251). ICGV 02251 was used as the male parent with HSW of 101.8 gm, while Chico (small-seeded cultivar) was used as a female parent with HSW of 35.0 gm. Furthermore, Chico has a high SHP with 80.5%, whereas ICGV 02251 (large-seeded cultivar) had a medium SHP of 69.5%. The average HSW in RIL population was 61.5 gm during S1, 43.5 gm during S2, and 45.5 gm during S3, whereas the average SHP was 69.40% during S1, 66.7% during S2, and 68.75% during S3. The seed weight and shelling percentage were higher in season S1 than those in S2 and S3. The phenotypic data generated in three seasons for both HSW and SHP showed normal distribution on violin plots (Figures 1A, B).

3.2 Important features of SNP array-based genetic map

The filtered 6235 SNPs were used for genetic map construction. A total of 4199 loci were mapped on A and B

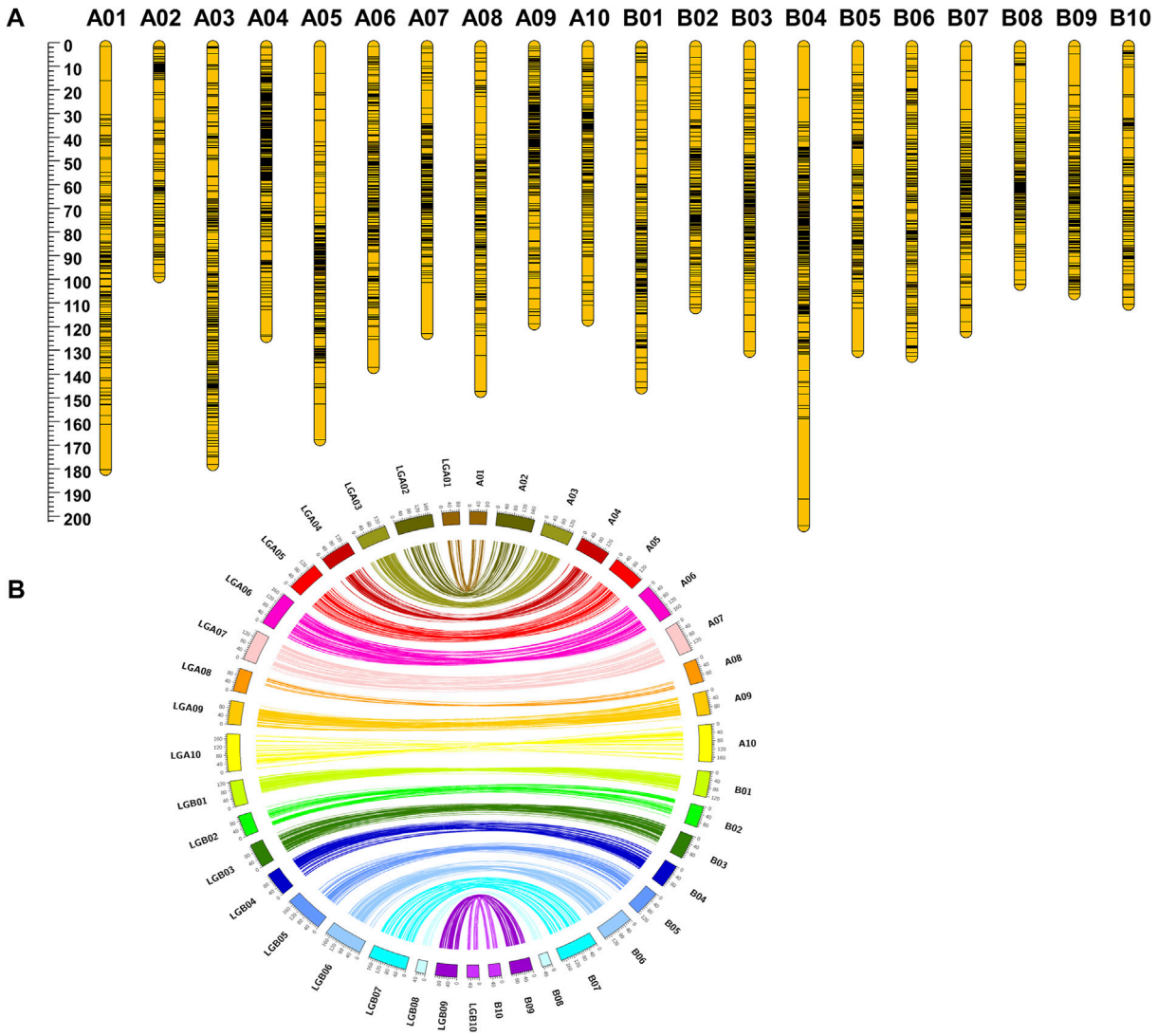


FIGURE 2
Highly collinear genetic map constructed using RIL population Chico × ICGV 02251. **(A)** SNP loci density on the 20 linkage groups. **(B)** Collinearity of the genetic map with the physical map.

subgenomes with a total distance of 2708.36 cM. A total of 2036 SNPs did not show any linkage with the SNP markers in the generated genetic map. No attempt was made to map the unlinked SNPs in the final genetic map to avoid noise during QTL analysis. Among the 4199 SNP loci, 2343 loci were mapped on A subgenome, whereas 1856 loci were mapped for B subgenome with a distance of 1456.96 cM and 1251.4 cM, respectively. The A and B subgenomes reached an average inter-marker distance of 0.66 and 0.67 cM/loci, respectively. A maximum number of loci mapped in a specific linkage group ranged from 128 (B10) to 447 (A04). The average inter-marker distance for each linkage group ranged from 97.34 cM (A02) to 202.5 cM (B04). The average inter-marker distance was maximum, 1.17 cM/loci, for the linkage group A01 and minimum, 0.47 (cM/loci), for the linkage group A09 (Table 1; Figure 2).

3.3 Identification of the major-effect QTLs associated with seed weight and shelling percentage

A total of 13 major QTLs were identified for HSW and SHP; six QTLs were identified for SHP with 5.3%–15.8% PVE and LOD score ranging from 2.51 to 7.16 during three seasons (S1, S2, and S3); and seven QTLs were identified for HSW with 6.96%–21.29% PVE and LOD score ranging from 3.9 to 11.7 during three seasons (Table 2).

3.4 QTLs identified for shelling percentage

Of the six QTLs identified for SHP, a QTL identified on chromosome A04 (*qShP-A04.1*) showed a minor effect of 5.32% PVE. A QTL was consistently detected on A05 (*qShP-A05.1*) with

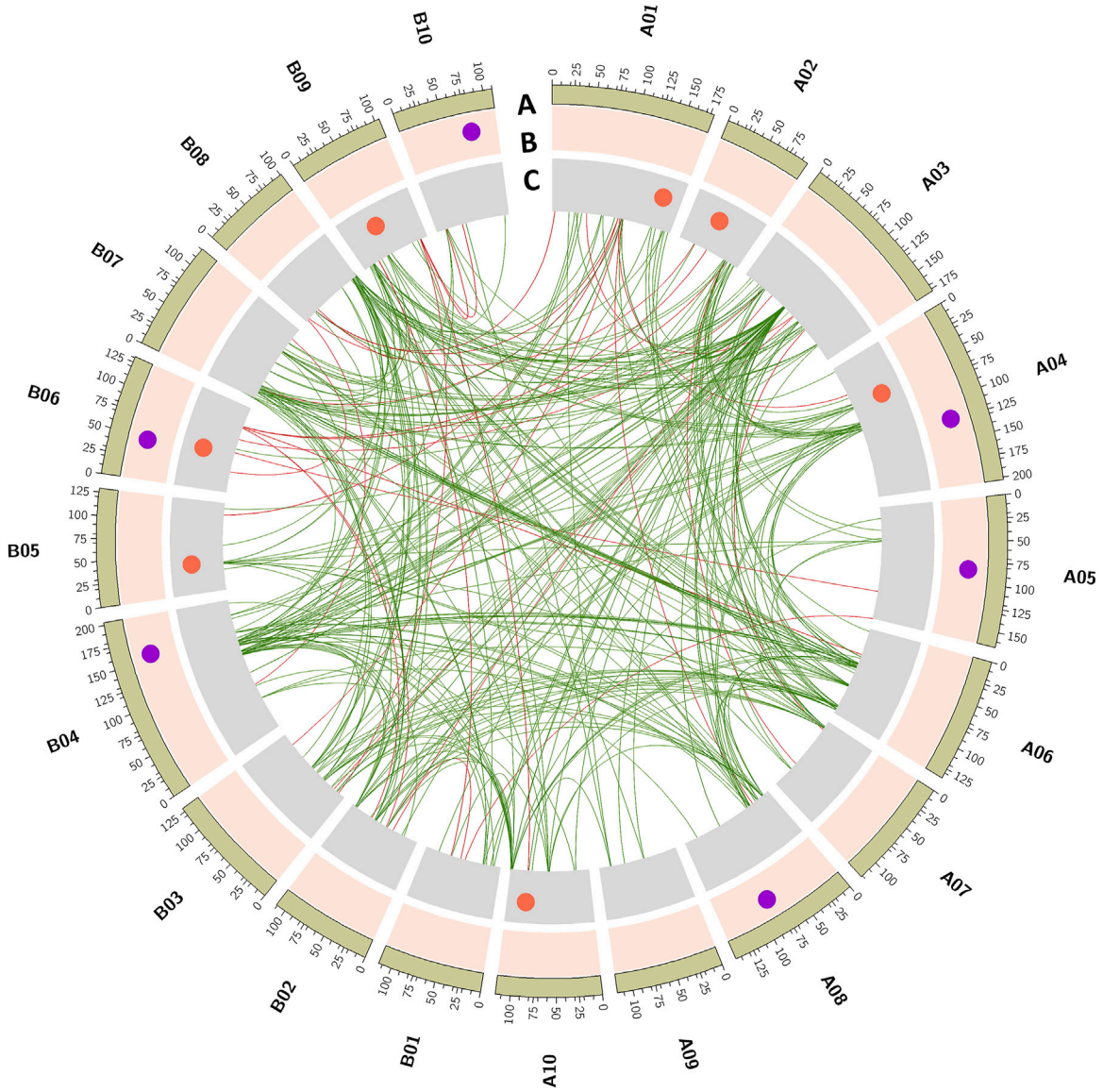


FIGURE 3
Main-effect QTLs identified for hundred-seed weight (HSW) and shelling percentage (SHP). The tracks outside to inside illustrates, **(A)** 20 chromosomes of cultivated groundnut labeled as A01 to A10 and B01 to B10, **(B)** QTLs identified for shelling percentage (SHP), **(C)** QTLs identified for hundred seed weight. Inner links represent epistatic (Q × Q) interactions. Green color links represent the epistatic interaction for shelling percentage, and red color links are epistatic interactions for seed weight.

LOD score of 3.5 and 11.07% PVE in seasons S2 and S3, respectively. A QTL *qShP-A08.1* on A08 with LOD score of 3.81 and 9.41% PVE was also consistently identified in seasons S2 and S3. A QTL (*qShP-B04.1*) identified on B04 with 9.11% PVE and *qShP-B06.1* identified on B06 with 7.16 LOD and 15.80% PVE. A QTL *qShP-B10.1* identified on B10 with LOD score of 4.64 and 12.8% PVE was consistently identified in seasons S2 and S3 (Table 2; Figure 3).

3.5 QTLs for hundred-seed weight

A total of seven QTLs were identified for HSW with a PVE range of 6.9%–21.29% and LOD score ranging from 3.9 to 11.7. A single QTL (*qHSW-A01.1*) was identified on A01 with 20.65% PVE and 8.82 LOD. On A02, a minor QTL (*qHSW-A02.1*) was identified with 4.2 LOD and 6.96% PVE. A major effect QTL *qHSW-A04.1* was identified on A04 with ~17.12% PVE in both seasons S1 and S3. Two QTLs were identified in S2, *qHSW-A10.1* on A10 with 3.98 LOD and 10.68% PVE and *qHSW-B05.1* on B05 with 7.66 LOD and 21.29% PVE. A major-effect QTL (*qHSW-B09.1*) was detected on B09 with 6.85 LOD and 12.0% PVE. A QTL *qHSW-B06.1* on B06 explained 13.65% PVE (Table 2; Figure 3).

3.6 Identification of epistatic interactions for seed weight and shelling percentage

A total of 375 E-QTLs were identified for shelling percentage and seed weight-related traits. A total of 42 E-QTLs were detected

TABLE 1 Summary of a genetic map constructed based on RIL population Chico × ICGV 02251.

Linkage group	Total loci	Mapped loci	Length of LG (cM)	Average inter-marker distance (cM/loci)	Linkage group	Total loci	Mapped loci	Length of LG (cM)	Average inter-marker distance (cM/loci)
A subgenome					B subgenome				
A01	4714	153	178.79	1.17	B01	2405	176	114.3	0.65
A02	3167	159	97.34	0.61	B02	3112	184	110.5	0.60
A03	3478	267	176.65	0.66	B03	3443	212	128.8	0.61
A04	2693	447	202.5	0.64	B04	2588	315	202.5	0.64
A05	2624	228	166.17	0.73	B05	2576	162	128.8	0.80
A06	2764	270	135.57	0.50	B06	2793	198	131	0.66
A07	2303	234	121.34	0.52	B07	2638	162	120.5	0.74
A08	2921	168	145.5	0.87	B08	2671	145	100.6	0.69
A09	2790	249	117.3	0.47	B09	3152	174	104.9	0.60
A10	2529	168	115.8	0.69	B10	2872	128	109.5	0.86
Total	29983	2343	1456.96	0.66	Total	28250	1856	1251.4	0.67
Grand total	58233	4199	2708.36	0.65					

for SHP and 332 E-QTLs for HSW. The phenotypic variation explained by the E-QTLs identified for seed weight and shelling percentage ranged from 10.0%–11.2% to 10.58%–27.09%, respectively.

A total of 332 epistatic QTLs were detected for HSW, of which 14 epistatic QTLs were identified with 3.60–7.32 LOD score and 10.02%–16.28% PVE during S2. On the other hand, 37 epistatic QTLs were identified with 3.60–6.91 LOD score and 10.02%–16.28% PVE during S1. A total of 182 epistatic QTLs were identified with 3.61–6.28 LOD, 10.20%–11.12% PVE during S3, of which 29 major epistatic QTLs were identified (LOD > 5.0) for HSW.

A total of 42 epistatic QTLs were detected for shelling percentage, and among them, eight had 3.05–3.55 LOD score and 10.65%–27.09% PVE during S1. The rest of the QTLs included 18 epistatic QTLs with 3.03–5.47 LOD score and 10.58%–26.31% PVE during S2 and 18 epistatic QTLs with 3.01–5.26 LOD score and 10.86%–25.14% PVE during S3. Major epistatic QTL for shelling percentage was identified with PVE of 15.46% and an LOD score of 5.4 that showed the interaction between genomic regions of chromosomes A02 and B06 (Figure 3; Supplementary Table S2).

3.7 Identification of environmental effect QTLs (Q × E) for seed weight and shelling percentage

A total of 15 environmental-effect QTLs were identified for seed weight and shelling percentage with LOD score value > 3.0. A total of seven environmental-effect QTLs were identified for seed weight with 4.11–9.56 LOD scores and 4.81%–27.18% PVE. A major environmental-effect QTL (*EqtlSW.B05.1*) on B05 was identified for seed weight with PVE of 10.03%. The same QTL region (*qHSW-B05.1*) was also identified as a main-effect QTL for HSW (Figure 4A). A total of three such E-QTLs for HSW showed high PVE on chromosome A01 (27.1%), B05 (10.0%), and B09 (12.3%). Here, we concluded that the QTL region on chromosome B05 explains the higher phenotypic variance due to the environmental effect with partial effects from a background genome. A total of eight environmental-effect QTLs were identified for SHP with 3.68–9.32 LOD score and 3.96%–11.06% PVE. A total of two major E-QTLs for SHP each on chromosome B06 and B10 were identified with PVE of 11.06% on B10 (Figure 4B). We plotted QTL additive effects against additive by environmental effects to find the QTLs which are highly influenced by the environment. We observed that both SHP and HSW showed higher (>10%) phenotypic variance explained by environmental effects (Supplementary Table S3).

3.8 Candidate genes for hundred-seed weight and shelling percentage

QTLs for both HSW and SHP with PVE >10.0 were targeted for candidate gene discovery. With this criterion, a total of three QTLs for HSW, namely, *qHSW_B05.1*, *qHSW_B06.1*, and *qHSW_B09.1* were investigated for candidate gene discovery. In the genomic region of *qHSW_B05.1* (0.9 Mb), a total of 16 candidate genes including nodulin MtN21 (*Araip.IJC5B*), acyl-CoA acyltransferase (*Araip.FGM9R*), yellow stripe-like (YSL) proteins (*Araip.Y0XXQ*), and serine–threonine protein phosphatase (*Araip.QEE5K*) were identified. A total of 47 genes were identified in the QTL region

TABLE 2 Main effect of QTLs for hundred-seed weight and shelling percentage.

QTL name	Season	Chr	Pos	Left marker	Right marker	LOD	PVE (%)	ADD	Contributing parent
Shelling percentage (SHP)									
qShP-A04.1	S1	A04	122	A04_59857421	A04_22131549	2.51	5.32	−2.10	Chico
qShP-A05.1	S2	A05	82	A05_22902580	A05_97942457	3.54	11.00	−2.42	Chico
	S3	A05	82	A05_22902580	A05_97942457	3.62	11.07	−2.41	Chico
qShP-A08.1	S3	A08	83	A08_43239201	A08_34847702	3.81	9.41	−2.04	Chico
	S2	A08	83	A08_43239201	A08_34847702	3.81	9.41	−2.04	Chico
qShP-B04.1	S1	B04	156	B04_130804224	B04_112728451	3.49	9.11	−1.98	Chico
qShP-B06.1	S1	B06	46	B06_125401834	B06_122810153	7.16	15.80	−2.68	Chico
qShP-B10.1	S3	B10	78	B10_135814931	B10_132665342	4.73	12.97	2.40	ICGV 02251
	S2	B10	78	B10_135814931	B10_132665342	4.64	12.80	2.39	ICGV 02251
Hundred-seed weight (HSW)									
qHSW-A01.1	S1	A01	152	A01_95668523	A01_104923102	8.82	20.65	10.25	ICGV 02251
qHSW-A02.1	S1	A02	40	A02_1992802	A02_84637409	4.20	6.96	4.79	ICGV 02251
qHSW-A04.1	S1	A04	61	A04_118815935	A04_33608076	8.01	14.49	3.53	ICGV 02251
	S3	A04	61	A04_118815935	A04_33608076	11.67	17.12	4.19	ICGV 02251
qHSW-A10.1	S2	A10	89	A10_109094772	A10_104674897	3.98	10.68	2.22	ICGV 02251
qHSW-B05.1	S2	B05	39	B05_25495898	B05_26588237	7.66	21.29	3.39	ICGV 02251
qHSW-B06.1	S3	B06	54	B06_112472327	B06_110468072	7.13	13.65	3.17	ICGV 02251
qHSW-B09.1	S1	B09	44	B09_145481859	B09_146042834	6.85	12.00	5.78	ICGV 02251

S1, Post-rainy 2013–14; **S2**, rainy 2014; **S3**, rainy 2019; **Chr**, chromosome; **Pos**, position; **LOD**, logarithm of odds; **PVE**, phenotypic variance explained; **ADD**, additive effect.

qHSW_B06.1 (2.0 Mb) on chromosome B06. The key genes in this region included zinc-finger proteins (*Araip.DYC48*), MYB transcription factors (*Araip.6N0ZN*), receptor-like serine/threonine kinases (*Araip.QR111*), E3 ubiquitin–protein ligase (*Araip.R3W9S*), methyltransferase-like protein (*Araip.YM7ML*), and glutamate dehydrogenase (*Araip.Q3F5T*). A total of 51 candidate genes were identified in the QTL region *qHSW_B09.1* (0.5 Mb) on chromosome B09 (Figure 5A). In this region, an important candidate gene from TIFY family proteins called *BIG SEED* locus was identified (*Araip.YK09Y*) along with *spermidine synthase* (*Araip.8Z3VM*), seed linoleate lipoxygenase (*Araip.D6PZJ*), epidermal patterning factor (*Araip.9SE5V*), cytochrome P450 superfamily protein (*Araip.A4X5Z*), and sugar transporters (*Araip.4HV2H*) (Figure 5B; Supplementary Table S4).

Similarly, for shelling percentage, we targeted three main major effect QTLs, namely, *qShP_A05.1*, *qShP_A08.1*, and *qShP_B10.1* (Figure 5C). The QTL on chromosome A05 (*qShP_A05.1*) harboured 47 candidate genes including *spermidine synthase* (*Aradu.FH71M*), receptor-like protein kinases (*Aradu.CU9HL*), disease-resistance protein (*Aradu.G0IJA*), Fe superoxide dismutase (*Aradu.3HY3W*), Ubiquitin–protein ligase, FAR1-related sequence 5-like (*Aradu.EW4Y3*), and tetratricopeptide repeat protein (*Aradu.1L0UP*). A total of 120 genes were identified in the QTL region (*qShP_A08.1*) on chromosome A08. The genes in this region included C_2H_2 zinc-finger protein (*Aradu.8HA7W*), cellulose synthase family protein (*Aradu.8HA7W*), chilling-induced protein (*Aradu.F7DPM*), defensin (*Aradu.CK6H7*), disease-resistance protein (TIR-NBS-LRR) (*Aradu.HLR71*), SUMO proteins (*Aradu.3IM1W*), fibre protein (*Aradu.3IM1W*), MYB transcription factor, NAC domain, senescence-associated proteins (*Aradu.GR4IB*), and Ulp proteases (*Aradu.G32CD*). A total of 84 genes were identified in the QTL region (*qShP_B08.1*) on chromosome B08. The candidate genes in this region included calcium-dependent lipid-binding protein (*Araip.B3EI3*), cellulose synthase (*Araip.U1RD3*), sugar transporters (SWEET) (*Araip.3H2LW*), Rubisco methyltransferase family protein (*Araip.N03N5*), disease-resistance protein (*Araip.P42EJ*), and glycine-rich abscisic acid inducible gene (*Araip.MZ5SZ*) (Table 3; Supplementary Table S4).

3.9 Expression of potential candidate genes at seed and pod developmental stages

The two parents used for developing the RIL population were subsp. *fastigiata*; hence, we used the *fastigiata* gene expression atlas (Sinha et al., 2020) to study the tissue-specific expression of candidate genes identified for HSW and SHP. In the QTL region

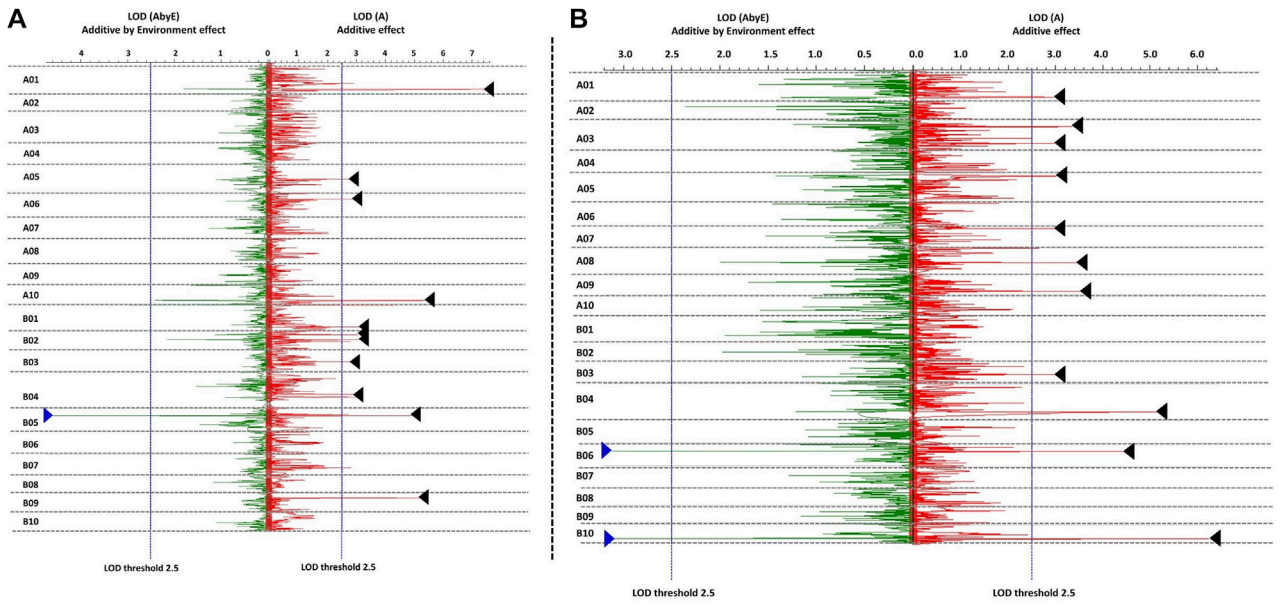


FIGURE 4
Main effect QTLs with major environment effect identified for **(A)** hundred seed weight, and **(B)** shelling percentage (%). Blue triangles represent additive by environment effect and black triangles represent the effect due genetic background.

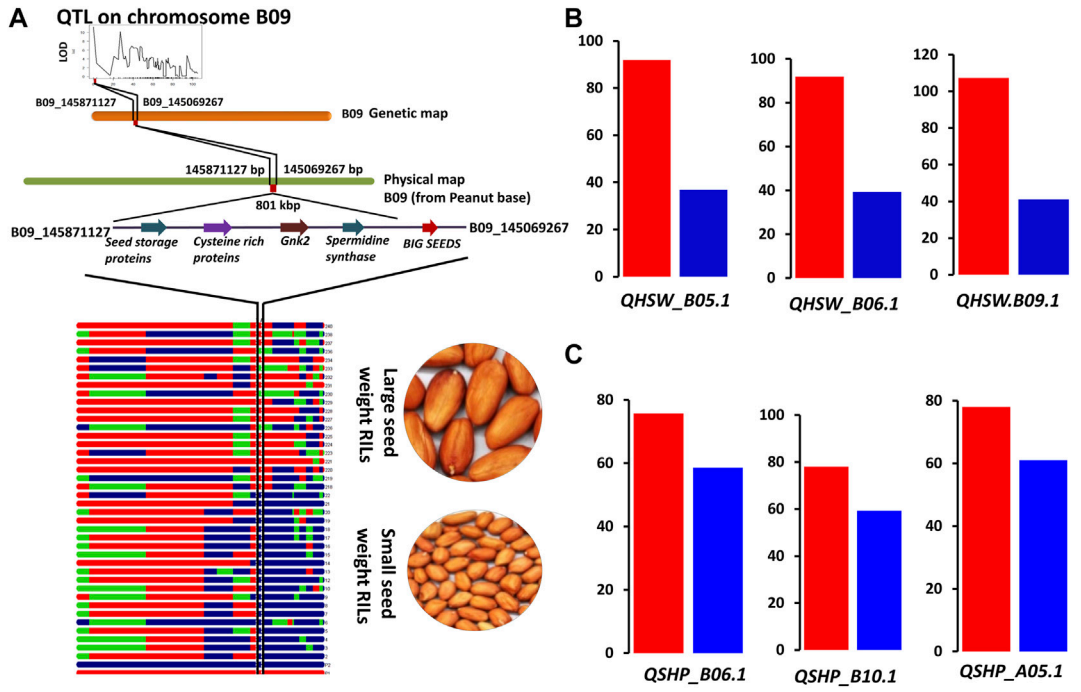


FIGURE 5
Validation of QTL regions using RILs for seed weight and shelling percentage. **(A)** QTL region and candidate genes on chromosome B09 associated with groundnut seed weight, **(B)** validation of main-effect QTLs for hundred-seed weight (HSW). **(C)** Validation of QTLs for shelling percentage (SHP). Red bars represent average shelling percentage of high-shelling percentage RILs, and blue bars represent average shelling percentage of low shelling percentage RILs.

qHSW_B05.1, YSL-like protein (*Araip.W1XHN*) was highly expressed in embryo and seed tissues. We identified two isoforms of YSL-like protein (*Araip.W1XHN* and *Araip.Y0XXQ*). Acyl-CoA acyltransferase (*Araip.FGM9R*) showed high expression in pod walls as compared to seed developmental stages. From the QTL region (*qHSW_B06.1*), two isoforms of *spermidine synthase*

TABLE 3 Potential candidate genes identified for hundred-seed weight and shelling percentage in major QTL regions.

QTL	Chromosome	Gene ID	Start	End	Annotation
QShP_A05.1	Aradu.A05	*Aradu.CU9HL*	22986869	22990293	Receptor-like protein kinase
	Aradu.A05	*Aradu.B4D3A*	24879878	24881965	U-box domain-containing protein
	Aradu.A05	*Aradu.MYH5B*	24792286	24793231	Adaptor protein complex
	Aradu.A05	*Aradu.FH71M*	24920996	24925015	Spermidine synthase
	Aradu.A05	*Aradu.G0IJA*	24901015	24904320	Disease resistance protein
	Aradu.A05	*Aradu.3HY3W*	24892006	24900621	Fe superoxide dismutase
	Aradu.A05	*Aradu.H04NN*	23876875	23880528	Protein kinase superfamily protein
	Aradu.A05	*Aradu.48LC5*	23381620	23385794	ATP-citrate lyase
	Aradu.A05	*Aradu.D08CT*	23649272	23653365	E3 ubiquitin-protein ligase
	Aradu.A05	*Aradu.Y354W*	23590437	23591175	Serine/threonine-protein phosphatase
	Aradu.A05	*Aradu.AP3A5*	23740398	23740895	GRF zinc finger protein
QShP_A08.1	Aradu.A08	*Aradu.6G4CM*	37286958	37292277	Calcium-binding EF-hand family protein
	Aradu.A08	*Aradu.440M4*	37377596	37378358	Defensin related
	Aradu.A08	*Aradu.HLR71*	43189988	43192287	Disease resistance protein (TIR-NBS-LRR class)
	Aradu.A08	*Aradu.7I425*	36999018	37001500	Fiber protein
	Aradu.A08	*Aradu.QL5QA*	41446652	41447645	Laccase
	Aradu.A08	*Aradu.V4L4B*	37275421	37281934	Leucine-rich repeat receptor-like protein
	Aradu.A08	*Aradu.FXW2H*	38389997	38390929	Lipid transfer protein
	Aradu.A08	*Aradu.AK182*	37321258	37328555	MLO-like protein 11-like
	Aradu.A08	*Aradu.888CN*	38843334	38844343	MYB transcription factor MYB82
	Aradu.A08	*Aradu.USH95*	38011876	38013744	NAC domain protein
	Aradu.A08	*Aradu.GR4IB*	42739667	42742321	Senescence-associated protein
	Aradu.A08	*Aradu.G32CD*	42492569	42494401	Ulp1 protease family C-terminal catalytic domain containing protein n
	Aradu.A08	*Aradu.2X4SQ*	39936927	39948439	Zinc finger CCCH domain-containing protein 37-like
qHSW_B05.1	Araip.B05	*Araip.W1XHN*	26367457	26370442	Protein YLS7-like
	Araip.B05	*Araip.QEE5K*	26463556	26466288	Serine/threonine-protein phosphatase
	Araip.B05	*Araip.8GK8W*	26024832	26025395	Protein FAR1-RELATED SEQUENCE 6-like isoform
	Araip.B05	*Araip.IJC5B*	26188466	26198447	Nodulin MtN21 /EamA-like transporter family protein
	Araip.B05	*Araip.FGM9R*	26261340	26263632	Acyl-CoA N-acyltransferases (NAT) superfamily protein
qHSW_B06.1	Araip.B06	*Araip.JGK52*	110857554	110860089	Spermidine hydroxycinnamoyl transferase-like
	Araip.B06	*Araip.Z5DXR*	111296004	111297328	Zinc-binding alcohol dehydrogenase family protein
	Araip.B06	*Araip.6N0ZN*	111333418	111334852	MYB transcription factor
	Araip.B06	*Araip.MK1X6*	111913983	111914486	Serine/threonine-protein phosphatase
	Araip.B06	*Araip.R3W9S*	112048940	112052397	E3 ubiquitin-protein ligase n
	Araip.B06	*Araip.86RJH*	112430911	112431479	CLP-similar protein
qHSW_B09.1	Araip.B09	*Araip.M2BYP*	145798830	145800398	Lipid transfer protein
	Araip.B09	*Araip.4HV2H*	145518107	145520042	Sugar transporter

(Continued on following page)

TABLE 3 (*Continued*) Potential candidate genes identified for hundred-seed weight and shelling percentage in major QTL regions.

QTL	Chromosome	Gene ID	Start	End	Annotation
	Araip.B09	*Araip.T9QAQ*	145492177	145493838	E3 ubiquitin-protein ligase
	Araip.B09	*Araip.U0WFW*	145740761	145742266	Seed maturation protein
	Araip.B09	*Araip.A4X5Z*	145725368	145727373	Cytochrome P450 superfamily protein
	Araip.B09	*Araip.9SE5V*	145953611	145955081	EPIDERMAL PATTERNING FACTOR-like protein 4-like
	Araip.B09	*Araip.D6PZJ*	145587000	145594367	Seed linoleate 9S-lipoxygenase
	Araip.B09	*Araip.Q00X2*	145593792	145596313	Spermidine synthase
	Araip.B09	*Araip.YK09Y*	145971775	145975639	Protein TIFY 4B-like isoform (*BIG SEED locus*)
QShP_B10.1	Araip.B10	*Araip.B3EI3*	133700271	133702677	Calcium-dependent lipid-binding (CaLB domain) family protein
	Araip.B10	*Araip.U1RD3*	133778488	133782909	Cellulose synthase
	Araip.B10	*Araip.RHZ53*	135758161	135760450	Cold acclimation protein *WCOR413* family
	Araip.B10	*Araip.P42EJ*	134195386	134198229	Disease resistance protein
	Araip.B10	*Araip.XJC3R*	133165446	133172336	E3 ubiquitin-protein ligase *RGLG2*-like isoform X4
	Araip.B10	*Araip.EJ1RQ*	135156885	135158036	F-box family protein
	Araip.B10	*Araip.I3ZKC*	135136304	135136707	LRR and NB-ARC domain disease resistance protein
	Araip.B10	*Araip.UA0W9*	133594505	133595933	NAC domain protein
	Araip.B10	*Araip.JW7JM*	135500237	135501639	Receptor-like protein kinase-like
	Araip.B10	*Araip.I4J8B*	134451076	134452602	Serine/threonine-protein phosphatase long form homolog
	Araip.B10	*Araip.3H2LW*	134153729	134154083	Sugar transporter SWEET

(*Araip.JGK52* and *Araip.J2AQK*) showed contrasting expression. The expression of *Araip.J2AQK* was higher in flower embryo and seed developmental stages. In the QTL region (*qHSW_B09.1*), sugar transporters (*Araip.4HV2H*) were highly expressed in flowers. Seed maturation protein (*Araip.U0WFW*) was expressed at the time of maturity in seeds. The TIFY family protein (*BIG SEED* locus) (*Araip.YK09Y*) was highly expressed in embryo and seed developmental stages. The epidermal patterning factor (*Araip.9SE5V*) showed high expression in flowers, seeds, and shells. Seed linoleate 9S-lipoxygenase (*Araip.9SE5V*) is shown in Figure 6A (Supplementary Table S5).

In the case of shelling percentage, several disease-resistance genes were highly expressed, such as MYB, NBS-LRR, and NAC domain proteins. In the QTL region *qShP_A05.1*, the *spermidine synthase* (*Araip.8Z3VM*) and ATP citrate lyase (*Aradu.48LC5*) showed higher expression in all the seed and pod tissues. In the QTL region *qShP_A08.1*, the genes involved in the synthesis of cellulose were highly expressed in the SHP QTLs. For instance, cellulose synthase (*Aradu.GKF95*) and fibre proteins (*Aradu.7I425*) showed high expression in all seed and pod developmental stages. Senescence-associated protein (*Aradu.GR4IB*) was highly expressed in pod walls at the time of maturity. In the QTL region *qShP_B10.1*, cold acclimation protein (*Araip.RHZ53*) was highly expressed in seed and pods. The lipases (*Araip.4V9G7*) and NAC domains (*Araip.UA0W9*) were highly expressed in flowers and mature shells. The expression analysis of genes showed that there are key genes involved in the cellulose biosynthesis pathway. However, a group of disease-resistant genes were highly expressed in all SHP QTL regions (Figure 6B; Supplementary Table S6).

3.10 Development of KASP markers for hundred-seed weight in groundnut

We used multiple approaches for genetic dissection of groundnut seed weight. In a previous study, we used NAM population to map the genomic regions associated with seed weight and pod weight in groundnut (Gangurde et al., 2020). A KASP marker (snpAH00173) on chromosome A05 at 101618480 bp was developed and validated on small- and large-seeded groundnut germplasms. A sequencing-based trait mapping approach "QTL-seq" was also used to identify the genomic regions for HSW of groundnut. An overlapping genomic region was identified on chromosome B09 in the present genetic mapping study and QTL-seq analysis for groundnut seed weight. A total of four KASP markers (snpAH0031, snpAH0033, snpAH0037, and snpAH0038) were recently developed from the same population using the QTL-seq approach (Gangurde et al., 2021). Because of high polymorphism, the seed weight KASP's markers were included in the quality control panel for their use in confirmation of F_1s and hybrid purity testing. Moreover, the KASP markers can also be used in the marker-assisted selection breeding programs to improve the seed-weight trait in important groundnut cultivars. In this study, we discovered a novel genomic region on chromosome B09, containing

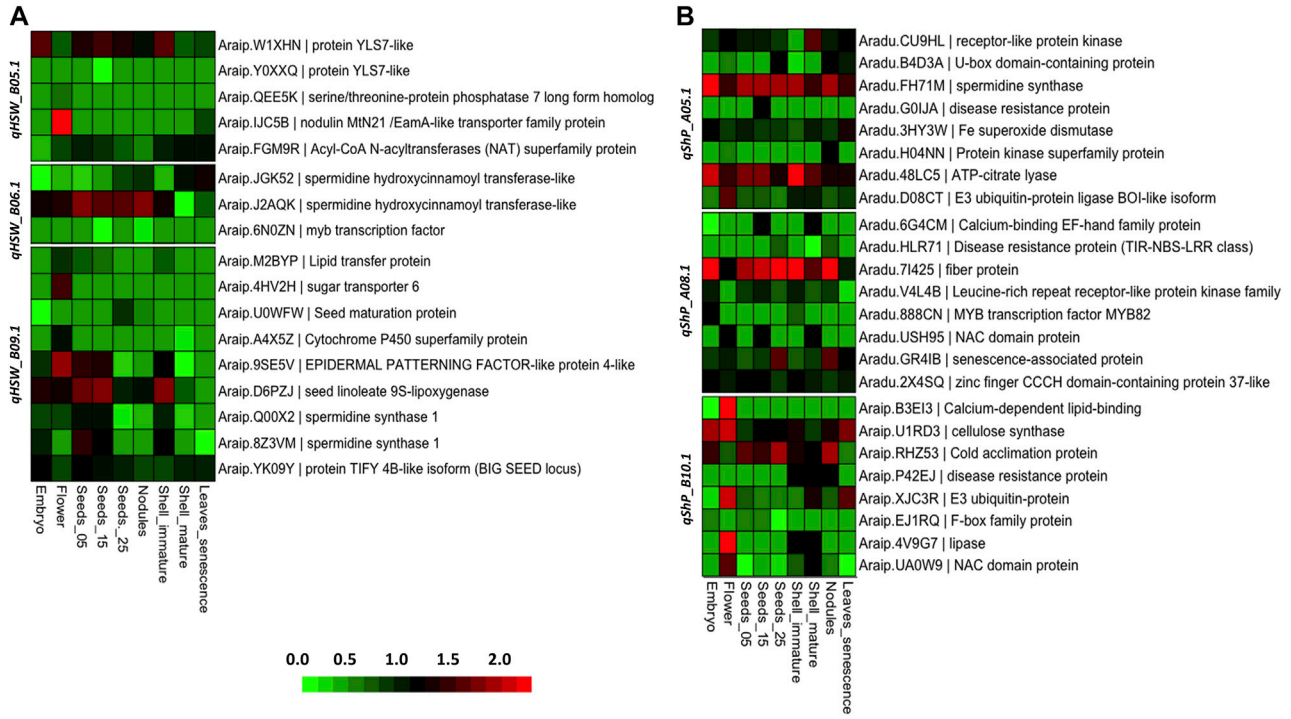


FIGURE 6
Heatmap illustrating the expression of candidate genes in various seed and pod development stages. **(A)** Heatmap for differential gene expression of genes identified in HSW QTL regions. **(B)** Heatmap for differential gene expression of genes identified in SHP QTL regions.

important genes, such as *BIG SEED* locus and *spermidine synthase* (*spds*), associated with seed development.

4 Discussion

In the present study, a RIL population (Chico × ICGV 02251) was used for mapping the genomic regions associated with HSW and SHP in groundnut. The three seasons of phenotyping data and SNP array-based genotyping data were generated to identify the QTLs linked with HSW and SHP. We observed that HSW was comparatively higher in post-rainy seasons than in rainy seasons as confirmed by repeated planting in two consecutive rainy seasons. This might be due to high disease pressure in rainy seasons that affects seed size in groundnut. A high-density 58K SNP array was used to construct a dense genetic map comprising 4199 SNP loci in a map distance of 2708.36 cM with an average inter-marker distance of 0.65 cM. Only 7.2% of SNP loci (4199 SNPs out of 58,233 SNPs on the array) were mapped on 20 linkage groups of groundnut. Genetic diversity analysis in groundnut reported that it has a very narrow genetic base; therefore, the construction of very high-density genetic map in groundnut is very challenging (Pandey et al., 2012). The density of this genetic map was the highest when compared to the previous genetic maps constructed using the SNP array (Pandey M. K. et al., 2020) and genotyping by sequencing (GBS) (Dodia et al., 2019; Jadhav et al., 2021). Moreover, a genetic map with 3630 markers grouped in 2636 bins was used to identify the QTLs for groundnut seed weight (Wang et al., 2018). Genome-wide association analysis with DArT markers identified nine marker–trait associations for seed length and five for HSW, but due to the unavailability of annotated reference genome for groundnut, the researchers could not reach to candidate genes associated with HSW (Pandey et al., 2014). Earlier genotyping by SSR markers was laborious and time consuming. Now, allelic SNP markers along with high-quality reference genomes allows for genetic dissection of complex traits and the identification of candidate genes (Bertioli et al., 2019; Chen et al., 2019; Zhuang et al., 2019). In our previous study, we successfully used the SNP array for mapping the genomic regions associated with seed and pod weight in groundnut using NAM population (Gangurde et al., 2020). The seed size QTLs on chromosomes A05 and A07 were reported from two RIL populations (Luo et al., 2017; Luo et al., 2018). Interestingly, in this study, we identified a consistent major-effect QTL for shelling percentage on chromosome A05 with >10% PVE, and a major QTL on chromosome B05 was identified with 21% PVE with 7.7 LOD. In addition, a major-effect QTL was identified on chromosome B09 with 13.0% and 11% PVE. Therefore, the genomic region on chromosome B09 was targeted for the identification of candidate genes using PeanutBase (www.peanutbase.org).

We also demonstrated that the major effect of a QTL is not just because of the genetic background; sometimes, it might be due to the environmental effect or a combination of these two. A major QTL for HSW on chromosome B05 and two major QTLs for shelling

percentage on chromosome B06 and B10 exhibited major additive by environmental effect (Q × E). Similar findings have been reported on the background effect and QTL × Environment effect for yield traits in rice (Wang et al., 2014).

Genes, such as *BIG SEED* locus and *spermidine synthase*, located in the QTL region on chromosome B09 negatively regulates the seed weight. Gene cloning for *BIG SEED* locus has been reported in *Medicago truncatula*, and the *BIG SEED* gene was isolated from soybean and overexpressed in the model legume crop, *Medicago truncatula*, which resulted in small seeds in transgenic lines (Ge et al., 2016). *Spermidine synthase* is also reported as the negative regulator of seed weight and size in rice (Tao et al., 2018). Therefore, we concluded that the *BIG SEED* locus and *spermidine synthase* genes can be targeted for the genome editing to enhance the groundnut seed weight. In the QTL region on chromosome B09, genes such as lipid transfer protein, sugar transporter, seed maturation protein, epidermal patterning factor-like proteins, and seed linoleate 9S-lipoxygenase associated with seed growth and development were identified. Seed linoleate 9S-lipoxygenase is a fat-metabolising gene linked with seed oil content (Wang et al., 2016). The epidermal patterning factor-like proteins are associated with plant epidermal cell growth factors and widely reported as a regulator for plant growth and development (Endo and Torii, 2019). A gene, serine threonine protein phosphatase identified in almost all QTL regions of seed weight and shelling percentage is associated with reactive oxygen species metabolism (ROS), plant's cold tolerance, and abscisic acid signalling (Hou et al., 2016). We identified a copy of *spermidine synthase* in the QTL region (*QShP_A05*), identified for the shelling percentage on chromosome A05, along with receptor-like kinases (RLKs) which play a major role in plant growth and stress response (Cui et al., 2021).

In the present study, we identified several disease-resistance genes, such as disease-resistance protein (*Aradu.G0IJA*) and Leucine-rich repeat receptor-like protein (*Aradu.V4L4B*), that are from NBS-LRR class in the QTL regions of shelling percentage. In addition, MLO-like proteins, laccase, senescence-associated proteins, cold acclimation proteins, and F-box family proteins were identified in the QTL regions of shelling percentage. The groundnut shell is made up of cellulose and fibre; cellulose synthase, fibre proteins, and laccase were identified in the QTL regions for shelling percentage. SWEET genes encoding sugar transporters play a major role for plant growth and development (Gupta, 2020).

In this study, both parental genotypes used in developing the RIL population were from subsp. *fastigiata*. Therefore, we used a gene expression atlas developed from subsp. *fastigiata* (Sinha et al., 2020). From gene expression patterns, we observed that the isoforms of YSL7-like protein showed multiple gene-expression patterns. For instance, isoform *Araip.W1XHN* was highly expressed, and *Araip.Y0XXQ* was not expressed in seed and pod tissues. We identified multiple isoforms of *spermidine synthases* (*Araip.Q00X2* and *Araip.8Z3VM*) in the major QTLs of seed weight. The *BIG SEED* locus encoded by protein TIFY 4B-like isoform showed high expression in all seed developmental stages. In addition, we observed that the seed linoleate 9s-lipoxygenase (*Araip.D6PZJ*) and epidermal patterning factor-like proteins (*Araip.9SE5V*) are the most expressed genes in the seed tissues (Figure 6A). In the case of shelling percentage, the disease-resistance genes in QTL regions of shelling percentage were confirmed with gene expression atlas. Almost all disease-resistance genes in the QTL regions of shelling percentage were highly expressed in pod walls and seed tissues. Interestingly, the *spermidine synthase* genes were identified in the QTL regions of both seed weight and shelling percentage. At the time of maturity, the senescence-associated protein (*Araip.GR4IB*) was highly expressed in pod walls. Cellulose synthase and calcium-binding proteins NAC domains were differentially expressed in seed and pod tissues.

The identified genes, particularly, *spermidine synthase*, *BIG SEED* locus, and seed linoleate 9s lipoxygenase genes, can be targeted for functional validation and can be used in improving the seed weight and shelling percentage in groundnut. Furthermore, the QTLs will be validated on diverse seed weight groundnut genotypes for their use in genomics-assisted breeding for improving HSW and SHP.

Data availability statement

The original contributions presented in the study are included in the article/Supplementary Material, further inquiries can be directed to the corresponding author.

Author contributions

MP conceived the idea, conceptualized this research, and supervised the entire study. JP developed RIL population. SSG, SM, MV, and JP generated the phenotyping data. SSG generated the genotyping data and analysed the data. SSG and MP interpreted the results and wrote the manuscript. SP and DB carried out formal analysis and writing. MP, PS, RV, and BG revised the final manuscript.

Funding

The funding was received from Indian Council of Agricultural Research (ICAR) through ICAR-ICRISAT project, Department of Biotechnology (DBT), Government of India, MARS-Wrigley and Bill and Melinda Gates Foundation (BMGF), United States.

Acknowledgments

SSG acknowledges the Department of Science and Technology, Government of India for DST-INSPIRE Fellowship and SP acknowledges Council of Scientific and Industrial Research (CSIR), and Govt. of India for the award of CSIR Junior Research Fellowship (JRF). The authors are thankful to the Indian Council of Agricultural Research (ICAR) and Department of Biotechnology (DBT) of Government of India, MARS-Wrigley and Bill and Melinda Gates Foundation (BMGF), United States for partial financial assistance.

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.

Publisher's note

All claims expressed in this article are solely those of the authors and do not necessarily represent those of their affiliated organizations, or those of the publisher, the editors, and the reviewers. Any product that may be evaluated in this article, or claim that may be made by its manufacturer, is not guaranteed or endorsed by the publisher.

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fgene.2023.1128182/full#supplementary-material

References

Agarwal, G., Clevenger, J., Kale, S. M., Wang, H., Pandey, M. K., Choudhary, D., et al. (2019). A recombination bin-map identified a major QTL for resistance to Tomato Spotted Wilt Virus in peanut (*Arachis hypogaea*). *Sci. Rep.* 9 (1), 18246–18313. doi:10.1038/s41598-019-54747-1

Agarwal, G., Clevenger, J., Pandey, M. K., Wang, H., Shasidhar, Y., Chu, Y., et al. (2018). High-density genetic map using whole-genome resequencing for fine mapping and candidate gene discovery for disease resistance in peanut. *Plant Biotechnol. J.* 16 (11), 1954–1967. doi:10.1111/pbi.12930

Bailey, W. K., and Hammons, R. O. (1975). Registration of Chico peanut germplasm 1 (reg. No. GP 2). *Crop Sci.* 15 (1), 105. doi:10.2135/cropsci1975.0011183x001500010050x

Ballén-Taborda, C., Chu, Y., Ozias-Akins, P., Timper, P., Holbrook, C. C., Jackson, S. A., et al. (2019). A new source of root-knot nematode resistance from Arachis stenosperma incorporated into allotetraploid peanut (*Arachis hypogaea*). *Sci. Rep.* 9, 17702. doi:10.1038/s41598-019-54183-1

Bertioli, D. J., Cannon, S. B., Froenicke, L., Huang, G., Farmer, A. D., Cannon, E. K. S., et al. (2016). The genome sequences of *Arachis duranensis* and *Arachis ipaensis*, the diploid ancestors of cultivated peanut. *Nat. Genet.* 48 (4), 438–446. doi:10.1038/ng.3517

Bertioli, D. J., Jenkins, J., Clevenger, J., Dudchenko, O., Gao, D., Seijo, G., et al. (2019). The genome sequence of segmental allotetraploid peanut *Arachis hypogaea*. Arachis hypogaea. *Nat. Genet.* 51 (5), 877–884. doi:10.1038/s41588-019-0405-z

Chavarro, C., Chu, Y., Holbrook, C., Isleib, T., Bertioli, D., Hovav, R., et al. (2020). Pod and seed trait QTL identification to assist breeding for peanut market preferences. *G3 Genes, Genomes, Genet.* 10 (7), 2297–2315. doi:10.1534/g3.120.401147

Chen, W., Jiao, Y., Cheng, L., Huang, L., Liao, B., Tang, M., et al. (2016). Quantitative trait locus analysis for pod-and kernel-related traits in the cultivated peanut (*Arachis hypogaea* L.) *BMC Genet.* 17 (1), 25–29. doi:10.1186/s12863-016-0337-x

Chen, X., Lu, Q., Liu, H., Zhang, J., Hong, Y., Lan, H., et al. (2019). Sequencing of cultivated peanut, *Arachis hypogaea*, yields insights into genome evolution and oil improvement. *Mol. Plant* 12 (7), 920–934. doi:10.1016/j.molp.2019.03.005

Chen, Y., Ren, X., Zheng, Y., Zhou, X., Huang, L., Yan, L., et al. (2017). Genetic mapping of yield traits using RIL population derived from Fuchuan Dahuasheng and ICG6375 of peanut (*Arachis hypogaea* L.) *Mol. Breed.* 37 (2), 17–14. doi:10.1007/s11032-016-0587-3

Chu, Y., Chee, P., Culbreath, A., Isleib, T. G., Holbrook, C. C., and Ozias-Akins, P. (2019). Major QTLs for resistance to early and late leaf spot diseases are identified on chromosomes 3 and 5 in peanut (*Arachis hypogaea* L.) *Front. Plant Sci. 10* 10, 883. doi:10.3389/fpls.2019.00883p.883

Chu, Y., Chee, P., Isleib, T. G., Holbrook, C. C., and Ozias-Akins, P. (2020). Major seed size QTL on chromosome A05 of peanut *(Arachis hypogaea)* is conserved in the US mini core germplasm collection. *Mol. Breed.* 40 (1), 6–16. doi:10.1007/s11032-019-1082-4

Clevenger, J., Chu, Y., Scheffler, B., and Ozias-Akins, P. (2016). A developmental transcriptome map for allotetraploid *Arachis hypogaea*. Arachis hypogaea. *Front. Plant Sci.* 7, 1446. doi:10.3389/fpls.2016.01446

Corley, R. H. (2009). How much palm oil do we need? *Environ. Sci. Policy* 12 (2), 134–139. doi:10.1016/j.envsci.2008.10.011

Cui, Y., Lu, X., and Gou, X. (2021). Receptor-like protein kinases in plant reproduction: Current understanding and future perspectives. *Plant Commun. 100273* 3, 100273. doi:10.1016/j.xplc.2021.100273

Dodia, S. M., Joshi, B., Gangurde, S. S., Thirumalaisamy, P. P., Mishra, G. P., Narandrakumar, D., et al. (2019). Genotyping-by-sequencing based genetic mapping reveals large number of epistatic interactions for stem rot resistance in groundnut. *Theor. Appl. Genet.* 132 (4), 1001–1016. doi:10.1007/s00122-018-3255-7

Endo, H., and Torii, K. U. (2019). Stomatal development and perspectives toward agricultural improvement. *Cold Spring Harb. Perspect. Biol.* 11 (5), a034660. doi:10.1101/cshperspect.a034660

Fonceka, D., Tossim, H. A., Rivallan, R., Vignes, H., Faye, I., Ndoye, O., et al. (2012). Fostered and left behind alleles in peanut: Interspecific QTL mapping reveals footprints of domestication and useful natural variation for breeding. *BMC Plant Biol.* 12 (1), 26–16. doi:10.1186/1471-2229-12-26

Food and Agriculture Organization of the United Nations Database (FAOSTAT) (2018). Online database at: https://www.fao.org/faostat/en/#data March Accessed on 23, 2022).

Gangurde, S. S., Kumar, R., Pandey, A. K., Burow, M., Laza, H. E., Nayak, S. N., et al. (2019). Climate-smart groundnuts for achieving high productivity and improved quality: Current status, challenges, and opportunities. *Genomic Des. climate-smart oilseed crops*, 133–172. doi:10.1007/978-3-319-93536-2_3

Gangurde, S. S., Nayak, S. N., Joshi, P., Purohit, S., Sudini, H. K., Chitikineni, A., et al. (2021). Comparative transcriptome analysis identified candidate genes for late leaf spot resistance and cause of defoliation in groundnut.*Int*. J. Mol. Sci.*22(9)* 4491.

Gangurde, S. S., Wang, H., Yaduru, S., Pandey, M. K., Fountain, J. C., Chu, Y., et al. (2020). Nested-association mapping (NAM)-based genetic dissection uncovers candidate genes for seed and pod weights in peanut (*Arachis hypogaea*). *Plant Biotechnol. J.* 18 (6), 1457–1471. doi:10.1111/pbi.13311

Ge, L., Yu, J., Wang, H., Luth, D., Bai, G., Wang, K., et al. (2016). Increasing seed size and quality by manipulating *BIG SEEDS1* in legume species. *Proc. Natl. Acad. Sci.* 113 (44), 12414–12419. doi:10.1073/pnas.1611763113

Gupta, P. K. (2020). SWEET genes for disease resistance in plants. *Trends Genet.* 36 (12), 901–904. doi:10.1016/j.tig.2020.08.007

Han, S., Yuan, M., Clevenger, J. P., Li, C., Hagan, A., Zhang, X., et al. (2018). A SNP-based linkage map revealed QTLs for resistance to early and late leaf spot diseases in peanut (*Arachis hypogaea* L.) *Front. Plant Sci.* 9, 10.3389/fpls.2018.01012.

Hou, Y. J., Zhu, Y., Wang, P., Zhao, Y., Xie, S., Batelli, G., et al. (2016). Type one protein phosphatase 1 and its regulatory protein inhibitor 2 negatively regulate ABA signaling. *PLoS Genet.* 12 (3), e1005835. doi:10.1371/journal.pgen.1005835

Jadhav, M. P., Gangurde, S. S., Hake, A. A., Yadawad, A., Mahadevaiah, S. S., Pattanashetti, S. K., et al. (2021). Genotyping-by-sequencing based genetic mapping identified major and consistent genomic regions for productivity and quality traits in peanut. *Front. Plant Sci.* 12, 668020. doi:10.3389/fpls.2021.6680202034

Kolde, R. (2019). Pheatmap: Pretty heatmaps version 1.0.12. Available at: https://cran.r-project.org/web/packages/pheatmap/index.html.

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

Kumar, R., Janila, P., Vishwakarma, M. K., Khan, A. W., Manohar, S. S., Gangurde, S. S., et al. (2020). Whole-genome resequencing-based QTL-seq identified candidate genes and molecular markers for fresh seed dormancy in groundnut. *Plant Biotechnol. J.* 18 (4), 992–1003. doi:10.1111/pbi.13266

Luo, H., Guo, J., Ren, X., Chen, W., Huang, L., Zhou, X., et al. (2018). Chromosomes A07 and A05 associated with stable and major QTLs for pod weight and size in cultivated peanut (*Arachis hypogaea* L.) *Theor. Appl. Genet.* 131 (2), 267–282. doi:10.1007/s00122-017-3000-7

Luo, H., Pandey, M. K., Khan, A. W., Guo, J., Wu, B., Cai, Y., et al. (2019a). Discovery of genomic regions and candidate genes controlling shelling percentage using QTL-seq approach in cultivated peanut (*Arachis hypogaea* L.) *Plant Biotechnol. J.* 17 (7), 1248–1260. doi:10.1111/pbi.13050

Luo, H., Pandey, M. K., Khan, A. W., Wu, B., Guo, J., Ren, X., et al. (2019b). Next-generation sequencing identified genomic region and diagnostic markers for resistance to bacterial wilt on chromosome B02 in peanut (*Arachis hypogaea* L.) *Plant Biotechnol. J.* 17 (12), 2356–2369. doi:10.1111/pbi.13153

Luo, H., Ren, X., Li, Z., Xu, Z., Li, X., Huang, L., et al. (2017). Co-localization of major quantitative trait loci for pod size and weight to a 3.7 cM interval on chromosome

A05 in cultivated peanut (*Arachis hypogaea* L.) *BMC Genom.* 18 (1), 58–12. doi:10.1186/s12864-016-3456-x

Luo, Z., Cui, R., Chavarro, C., Tseng, Y. C., Zhou, H., Peng, Z., et al. (2020). Mapping quantitative trait loci (QTLs) and estimating the epistasis controlling stem rot resistance in cultivated peanut (*Arachis hypogaea*). *Theor. Appl. Genet.* 133 (4), 1201–1212. doi:10.1007/s00122-020-03542-y

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

Mondal, S., and Badigannavar, A. M. (2019). Identification of major consensus QTLs for seed size and minor QTLs for pod traits in cultivated groundnut (*Arachis hypogaea* L.) *3 Biotech.* 9 (9), 347–349. doi:10.1007/s13205-019-1881-7

Nabi, R. B. S., Cho, K. S., Tayade, R., Oh, K. W., Lee, M. H., Kim, J. I., et al. (2021). Genetic diversity analysis of Korean peanut germplasm using 48 K SNPs 'Axiom_Arachis' Array and its application for cultivar differentiation. *Sci. Rep.* 11 (1), 16630. doi:10.1038/s41598-021-96074-4

Pandey, M. K., Khan, A. W., Singh, V. K., Vishwakarma, M. K., Shasidhar, Y., Kumar, V., et al. (2017a). QTL-seq approach identified genomic regions and diagnostic markers for rust and late leaf spot resistance in groundnut (*Arachis hypogaea* L.) *Plant. Biotechnol. J.* 15 (8), 927–941.

Pandey, M. K., Agarwal, G., Kale, S. M., Clevenger, J., Nayak, S. N., Sriswathi, M., et al. (2017b). Development and evaluation of a high-density genotyping 'Axiom_Arachis' array with 58 K SNPs for accelerating genetics and breeding in groundnut. *Sci. Rep.* 7 (1), 40577–40610. doi:10.1038/srep40577

Pandey, M. K., Pandey, A. K., Kumar, R., Nwosu, C. V., Guo, B., Wright, G. C., et al. (2020a). Translational genomics for achieving higher genetic gains in groundnut. *Theor. Appl. Genet.* 133 (5), 1679–1702. doi:10.1007/s00122-020-03592-2

Pandey, M. K., Gangurde, S. S., Sharma, V., Pattanashetti, S. K., Naidu, G. K., Faye, I., et al. (2020b). Improved genetic map identified major QTLs for drought tolerance-and iron deficiency tolerance-related traits in groundnut. *Genes* 12 (1), 37. doi:10.3390/genes12010037

Pandey, M. K., Monyo, E., Ozias-Akins, P., Liang, X., Guimarães, P., Nigam, S. N., et al. (2012). Advances in Arachis genomics for peanut improvement. *Biotechnol. Adv.* 30 (3), 639–651. doi:10.1016/j.biotechadv.2011.11.001

Pandey, M. K., Upadhyaya, H. D., Rathore, A., Vadez, V., Sheshshayee, M. S., Sriswathi, M., et al. (2014). Genome-wide association studies for 50 agronomic traits in peanut using the 'reference set' comprising 300 genotypes from 48 countries of the semi-arid tropics of the world. *PLoS one* 9 (8), e105228. doi:10.1371/journal.pone.0105228

Parmar, S., Sharma, V., Bomireddy, D., Soni, P., Joshi, P., Gangurde, S. S., et al. (2022). Recent advances in genetics, genomics, and breeding for nutritional quality in groundnut. *Accel. Plant Breed.* 4, 111–137.

Shasidhar, Y., Variath, M. T., Vishwakarma, M. K., Manohar, S. S., Gangurde, S. S., Sriswathi, M., et al. (2020). Improvement of three popular Indian groundnut varieties for foliar disease resistance and high oleic acid using SSR markers and SNP array in marker-assisted backcrossing. *Crop J.* 8 (1), 1–15. doi:10.1016/j.cj.2019.07.001

Sinha, P., Bajaj, P., Pazhamala, L. T., Nayak, S. N., Pandey, M. K., Chitikineni, A., et al. (2020). *Arachis hypogaea* gene expression atlas for *fastigiata* subspecies of cultivated groundnut to accelerate functional and translational genomics applications. *Plant Biotechnol. J.* 18 (11), 2187–2200. doi:10.1111/pbi.13374

Tao, Y., Wang, J., Miao, J., Chen, J., Wu, S., Zhu, J., et al. (2018). The *spermine synthase OsSPMS1* regulates seed germination, grain size, and yield. *Plant Physiol.* 178 (4), 1522–1536. doi:10.1104/pp.18.00877

Varshney, R. K., Bertioli, D. J., Moretzsohn, M., de, C., Vadez, V., Krishnamurthy, L., et al. (2009). The first SSR-based genetic linkage map for cultivated groundnut (*Arachis hypogaea* L.) *Theor. Appl. Genet.* 118 (4), 729–739. doi:10.1007/s00122-008-0933-x

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

Wang, M. L., Wang, H., Zhao, C., Tonnis, B., Tallury, S., Wang, X., et al. (2022). Identification of QTLs for seed dormancy in cultivated peanut using a recombinant inbred line mapping population. *Plant Mol. Biol. Rep.* 40, 208–217. doi:10.1007/s11105-021-01315-5

Wang, X., Pang, Y., Zhang, J., Zhang, Q., Tao, Y., Feng, B., et al. (2014). Genetic background effects on QTL and QTL× environment interaction for yield and its component traits as revealed by reciprocal introgression lines in rice. *Crop J.* 2 (6), 345–357. doi:10.1016/j.cj.2014.06.004

Wang, Y., Ma, X., Zhang, X., He, X., Li, H., Cui, D., et al. (2016). ITRAQ-based proteomic analysis of the metabolic mechanisms behind lipid accumulation and degradation during peanut seed development and post-germination. *J. Proteome Res.* 15 (12), 4277–4289. doi:10.1021/acs.jproteome.6b00345

Wang, Z., Huai, D., Zhang, Z., Cheng, K., Kang, Y., Wan, L., et al. (2018). Development of a high-density genetic map based on specific length amplified fragment sequencing and its application in quantitative trait loci analysis for yield-related traits in cultivated peanut. *Front. Plant Sci.* 9, 827. doi:10.3389/fpls.2018.00827

Zhang, H., Chu, Y., Dang, P., Tang, Y., Jiang, T., Clevenger, J. P., et al. (2020). Identification of QTLs for resistance to leaf spots in cultivated peanut (*Arachis hypogaea* L.) through GWAS analysis. *Theor. Appl. Genet.* 133 (7), 2051–2061. doi:10.1007/s00122-020-03576-2

Zhang, S., Hu, X., Miao, H., Chu, Y., Cui, F., Yang, W., et al. (2019). QTL identification for seed weight and size based on a high-density SLAF-seq genetic map in peanut (*Arachis hypogaea* L.) *BMC Plant Biol.* 19 (1), 1–15. doi:10.1186/s12870-019-2164-5

Zhuang, W., Chen, H., Yang, M., Wang, J., Pandey, M. K., Zhang, C., et al. (2019). The genome of cultivated peanut provides insight into legume karyotypes, polyploid evolution and crop domestication. *Nat. Genet.* 51 (5), 865–876. doi:10.1038/s41588-019-0402-2

Zou, K., Kang, D., Kim, K. S., and Jun, T. H. (2020b). Screening of salinity tolerance and genome-wide association study in 249 peanut accessions (*Arachis hypogaea* L.) *Plant Breed. Biotech.* 8 (4), 434–438. doi:10.9787/pbb.2020.8.4.434

Zou, K., Kim, K. S., Kang, D., Kim, M. C., Ha, J., Moon, J. K., et al. (2022). Genome-wide association study of leaf chlorophyll content using high-density SNP array in peanuts (*Arachis hypogaea* L.) *Agronomy* 12 (1), 152. doi:10.3390/agronomy12010152

Zou, K., Kim, K. S., Kim, K., Kang, D., Park, Y. H., Sun, H., et al. (2020a). Genetic diversity and genome-wide association study of seed aspect ratio using a high-density SNP array in peanut (*Arachis hypogaea* L.) *Genes* 12 (1), 2. doi:10.3390/genes12010002

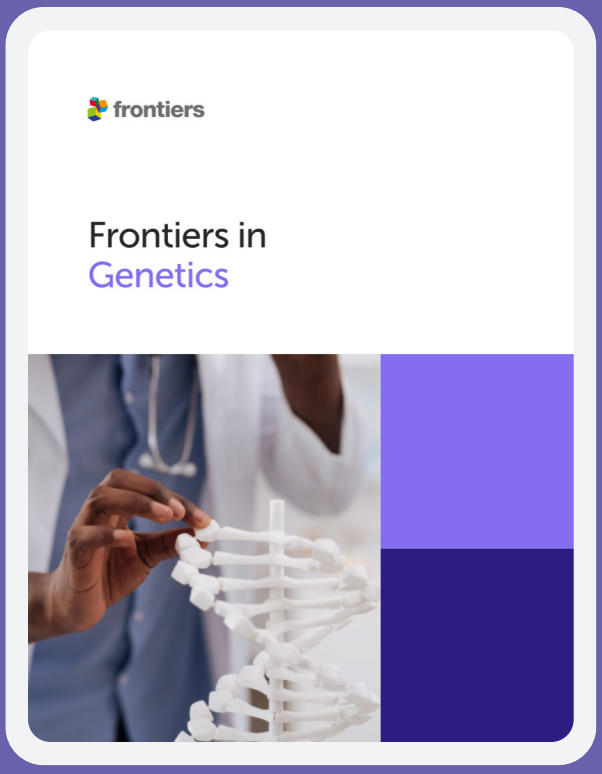
Frontiers in Genetics

Highlights genetic and genomic inquiry relating to all domains of life

The most cited genetics and heredity journal, which advances our understanding of genes from humans to plants and other model organisms. It highlights developments in the function and variability of the genome, and the use of genomic tools.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact