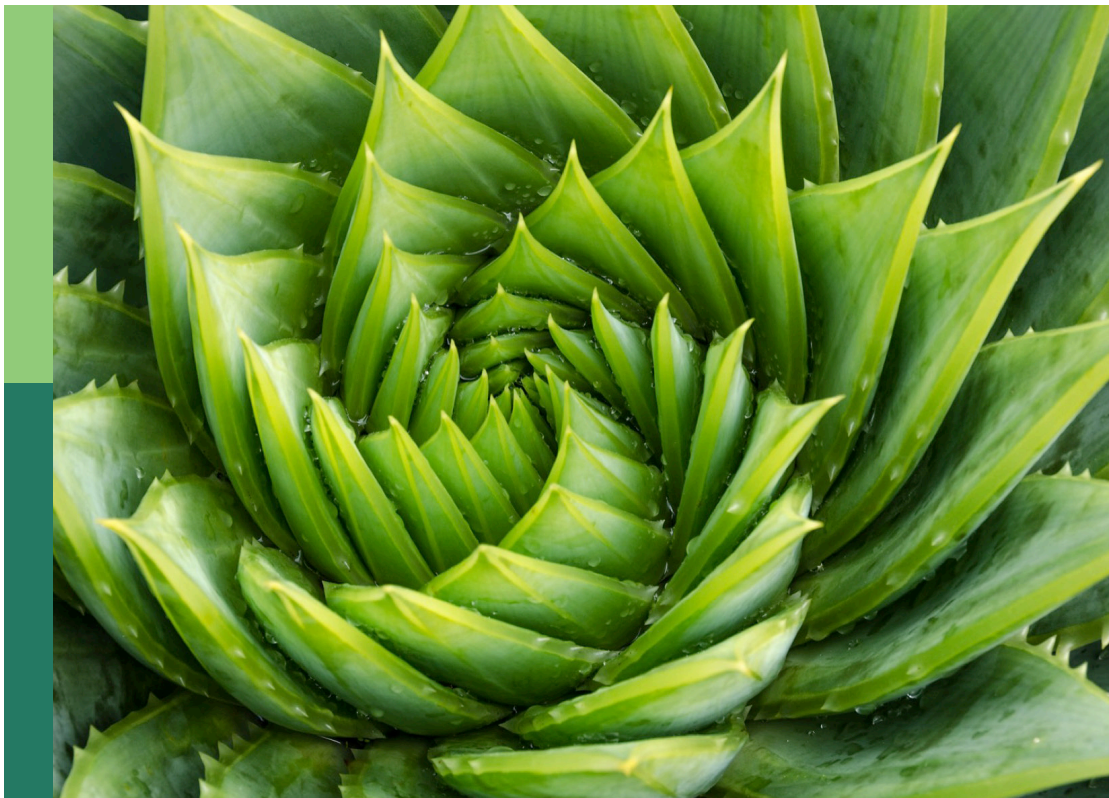
frontiers | Research Topics

Neglected and underutilized crop species for sustainable food and nutritional security: Prospects and hidden potential

Edited by
Omena Bernard Ojuederie, David Okeh Igwe,
Ndomelele Ndiko Ludidi and Beckley Ikhajiagbe

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-4332-0
DOI 10.3389/978-2-8325-4332-0

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

Neglected and underutilized crop species for sustainable food and nutritional security: Prospects and hidden potential

Topic editors

Omena Bernard Ojuederie — Kings University, Nigeria
David Okeh Igwe — Cornell University, United States
Ndomelele Ndiko Ludidi — University of Mpumalanga, South Africa
Beckley Ikhajiagbe — University of Benin, Nigeria

Citation

Ojuederie, O. B., Igwe, D. O., Ludidi, N. N., Ikhajiagbe, B., eds. (2024). *Neglected and underutilized crop species for sustainable food and nutritional security: Prospects and hidden potential*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-4332-0

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 09 January 2024
DOI 10.3389/fpls.2023.1358220

OPEN ACCESS

EDITED AND REVIEWED BY
Zerihun Tadele,
University of Bern, Switzerland

*CORRESPONDENCE
Omena Bernard Ojuederie
ob.ojuederie@kingsuniversity.edu.ng

RECEIVED 19 December 2023
ACCEPTED 29 December 2023
PUBLISHED 09 January 2024

CITATION
Ojuederie OB, Igwe DO, Ludidi NN and Ikhajiagbe B (2024) Editorial: Neglected and underutilized crop species for sustainable food and nutritional security: prospects and hidden potential.
Front. Plant Sci. 14:1358220.
doi: 10.3389/fpls.2023.1358220

Editorial: Neglected and underutilized crop species for sustainable food and nutritional security: prospects and hidden potential

Omena Bernard Ojuederie[1,2]*, David Okeh Igwe[3], Ndiko Ndomele Ludidi[4] and Beckley Ikhajiagbe[5]

[1]Department of Biological Sciences, Biotechnology Unit, Faculty of Science, Kings University, Ode-Omu, Osun, Nigeria, [2]Food Security and Safety Focus, Faculty of Natural and Agricultural Sciences, North-West University, Mmabatho, South Africa, [3]Plant Pathology and Plant-Microbe Biology Section, School of Integrative Plant Science, Cornell University, Ithaca, NY, United States, [4]Department of Biotechnology, University of the Western Cape, Bellville, South Africa, [5]Department of Plant Biology and Biotechnology, Faculty of Science, University of Benin, Benin City, Edo, Nigeria

KEYWORDS

food security, genetic improvement, orphan crops, sustainable agriculture, zero hunger

Editorial on the Research Topic
Neglected and underutilized crop species for sustainable food and nutritional security: prospects and hidden potential

The exploration of neglected and underutilized crop species (NUCs) is indeed crucial in tackling global food insecurity. These nutrient-rich, climate-resilient crops, often disregarded for limited commercial value, hold the key to combating malnutrition and boosting food security, especially in vulnerable regions. These crops, which have not been previously categorized as major crops and are mainly confined to smallholder farming areas, are nutrient-dense, climate-resilient, and locally adaptable (Li and Siddique, 2020; Mudau et al., 2022). The erosion of these crops can hinder the nutritional status and food security of the poor, and their greater use can augment nutrition and fight hidden hunger (Dansi et al., 2012; Ojuederie et al., 2015; Joy and Siddhuraju, 2017). It is crucial that we recognize the hidden potential of these crops and harness them to achieve a more sustainable future.

This editorial spotlight promising research showcasing the hidden potential of NUCs and exploring their utilization through modern advancements. The range of research showcased in this editorial on the Research topic *'Neglected and Underutilized Crop Species for Sustainable Food and Nutritional Security: Prospects and Hidden Potential'* is impressive and covers various aspects of these crops, from their genetic improvement to their potential applications in diverse fields. The Research Topic consists of 9 publications: 6 original research articles and 3 reviews focusing on the genetic improvement, conservation and utilization of some NUCs in addressing global food and nutrition challenges.

Citrus grandis (L.) Osbeck, commonly called pomelo, is an underutilized citrus fruit, whose potential as a source of flavonoids, phenols, and antioxidants, have been overlooked.

The investigation into drying methods for *Citrus grandis* fruits by Kumar et al. sheds light on optimizing processes to retain their bioactive compounds, potentially opening avenues for functional food development and nutraceuticals. In the parched lands where water is a precious commodity, foxtail millet stands tall, a beacon of hope for farmers and food security.

The research by Singh et al. investigated the potential for improving foxtail millet (*Setaria italica* L. beauv) yields in rainfed conditions through germplasm assessment and analysis. From this research, five superior genotypes emerged through genetic analysis: Kangni-7, Kangni-1, Kangni-6, Kangni-5, and Kangni-4 offering a promising approach to revolutionize its production under harsh conditions and bolster food security and enhance livelihoods.

The effectiveness of using mock genomes for conducting genomic studies in orphan crops, which lack a reference genome was investigated by Machado et al. The genotyping-by-sequencing-Mock approach provided comparable results to standard methods in the areas of genetic diversity studies, dividing heterotic groups, selecting testers, and predicting the performance of single crosses. The study offers a promising approach for leveraging genomic technologies to improve the development of orphan crops, which can have significant benefits for global food security and nutrition. Duckweed, with rapid growth and high protein content, shows potential for biofuel production, animal feed, and wastewater treatment.

The assessment of genetic variability among duckweed clones using DNA barcoding markers has been a subject of interest (Borisjuk et al., 2014; Al-Dakhil et al., 2021). Exploring DNA barcoding techniques and assessing the biomass accumulation rates of duckweed species, particularly of native Iranian types, offer promising opportunities for biotechnological applications.

Taghipour et al. identified four duckweed Iranian species using two standard chloroplast markers and assessed their growth rates, emphasizing their potential for sustainable biotechnological solutions. Growth rates of selected duckweed clones were found to be higher than common crops, demonstrating their potential for large-scale biomass production. As a result of the high protein content of the Native Iranian duckweed, it can be utilized as a sustainable source of protein for food and feed applications and in the development of valuable recombinant proteins.

The genetic dissection of agronomic traits in Andean lupin (*L. mutabilis*) and Bambara groundnut, offer invaluable insights for breeding improved varieties and promoting these crops' consumption. We highlight the potential impact of the research of Gulisano et al. on Andean lupin improvement for sustainable agriculture. Their research identifies genetic factors for breeding adapted Andean lupin varieties, laying the groundwork for its successful cultivation in Europe and establish a foundation for further research on plant development and phenology in *L. mutabilis*, paving the way for improved breeding programs and a better understanding of plant development in this species.

Understanding the genetic basis of drought tolerance in Bambara groundnut is crucial for breeding programs. Odesola et al. considered both genetic diversity and environmental interactions for effective breeding programs. The study paves the way for developing improved Bambara varieties resilient to climate change through targeted breeding approaches.

Underutilized legumes (NULs) in Africa are valuable sources of nutrition and bioactive compounds offering health-promoting benefits. Despite their nutritional value, they remain neglected and underutilized. Legume seeds possessing bioactive compounds with antioxidant activity could lessen the negative impact of oxidative stress, and enhance the well-being of man. Popoola et al. reviewed selected NULs and their nutritional, functional, and bioactive properties. their potential and challenges, proposing strategies for increased utilization and their role in developing sustainable and healthy food systems and strategies proposed for their increased exploitation.

Moth bean, a highly adaptable and nutritious crop, is highlighted by Kanishka et al. as a climate-smart crop for food security and income generation. This underscores the importance of further research and development efforts. The authors advocates for increased research and development to unlock its full potential, leveraging its unique strengths and untapped potential to address agricultural challenges in a changing climate.

The review of Kudapa et al. focusing on biofortification of millets through genetic and genomic interventions showcases the potential of cutting-edge technologies like CRISPR-Cas9 in addressing micronutrient malnutrition, presenting a viable solution for sustainable food and nutritional security. Combining genetic and genomic interventions with harnessing the key characteristics of millets, offer promising solutions to combat micronutrient malnutrition and contribute to sustainable food and nutritional security. Overall, this comprehensive range of research efforts highlights the multifaceted approach needed to unlock the hidden potential of neglected and underutilized crops, paving the way for a more sustainable and food-secure future.

Author contributions

OO: Conceptualization, Writing – original draft, Writing – review & editing. DI: Writing – original draft, Writing – review & editing. NL: Writing – original draft, Writing – review & editing. BI: Writing – original draft, Writing – review & editing.

Acknowledgments

We thank all contributing authors for their participation in the Research Topic.

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

Al-Dakhil, M., Alghamdi, S., Migdadi, H., Afzal, M., and Ali, A. (2021). Morphological characterization and DNA barcoding of duckweed species in Saudi Arabia. *Plants* 10 (11), 2438. doi: 10.3390/plants10112438

Borisjuk, N., Chu, P., Gutierrez, R., Zhang, H., Acosta, K., Friesen, N., et al. (2015). Assessment, validation and deployment strategy of a two-barcode protocol for facile genotyping of duckweed species. *Plant Biol.* 17 (s1), 42–49. doi: 10.1111/plb.12229

Dansi, A., Vodouhè, R., Azokpota, P., Yédomonhan, H., Assogba, P., Adjatin, A., et al. (2012). Diversity of the neglected and underutilized crop species of importance in Benin. *Sci. World J.* 2012, 1–19. doi: 10.1100/2012/932947

Joy, J., and Siddhuraju, P. (2017). Studies on nutritional profile and antioxidant potential of different *Dioscorea sp* with *Plectranthus rotundifolius*. *Int. J. Curr. Pharma Res.* 9 (4), 65. doi: 10.22159/ijcpr.2017v9i4.20959

Li, X., and Siddique, K. (2020). Future smart food: harnessing the potential of neglected and underutilized species for zero hunger. *Maternal Child Nutr.* 16 (S3), e13008. doi: 10.1111/mcn.13008

Mudau, F., Chimonyo, V., Modi, A., and Mabhaudhi, T. (2022). Neglected and underutilised crops: a systematic review of their potential as food and herbal medicinal crops in south Africa. *Front. Pharmacol.* 12. doi: 10.3389/fphar.2021.809866

Ojuederie, O. B., Balogun, M. O., Akande, S. R., Korie, S., and Omodele, T. (2015). Intraspecific variability in agro-morphological traits of African yam bean (*Sphenostylis stenocarpa* (Hochst ex. A. Rich) Harms. *J. Crop Sci. Biotechnol.* 18 (2), 53–62. doi: 10.1007/s12892-014-0109-y

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 29 November 2022
DOI 10.3389/fpls.2022.1034238

OPEN ACCESS

EDITED BY
Beckley Ikhajiagbe,
University of Benin, Nigeria

REVIEWED BY
Sezai Ercisli,
Atatürk University, Turkey
David Dewez,
Université du Québec à Montréal, Canada

*CORRESPONDENCE
Ali Hatef Salmanian
salman@nigeb.ac.ir

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 01 September 2022
ACCEPTED 01 November 2022
PUBLISHED 29 November 2022

CITATION
Taghipour E, Bog M, Frootan F, Shojaei S, Rad N, Arezoumandi M, Jafari M and Salmanian AH (2022) DNA barcoding and biomass accumulation rates of native Iranian duckweed species for biotechnological applications. *Front. Plant Sci.* 13:1034238. doi: 10.3389/fpls.2022.1034238

DNA barcoding and biomass accumulation rates of native Iranian duckweed species for biotechnological applications

Elham Taghipour[1], Manuela Bog[2], Fateme Frootan[1], Sadegh Shojaei[1], Nima Rad[1], Mahdi Arezoumandi[1], Mahyat Jafari[1] and Ali Hatef Salmanian[1]*

[1]National Institute of Genetic Engineering and Biotechnology (NIGEB), Department of Agricultural Biotechnology, Tehran, Iran, [2]University of Greifswald, Institute of Botany and Landscape Ecology, Greifswald, Germany

The Lemnaceae family (duckweed) consists of at least three recognized genera with six reported species in Iran that are distributed in wetlands. Duckweeds are the simplest and smallest flowering aquatic monocots with free-floating fronds that can reproduce asexually every 2–3 days. Duckweed could be a major source of balanced amino acids and high protein content, which is increasingly promising for biotechnological applications. For molecular classification and species identification of the collected samples, DNA barcoding was performed using two standard chloroplast markers, the spacer region between the ATP synthase subunits F and H (*atpF-atpH*) and the intron region of the ribosomal protein S16 (*rps16*). The results confirm the presence of four species belonging to the two genera *Lemna* and *Spirodela*. In addition, *L. turionifera* was detected for the first time in Iran. Due to the high growth rates of duckweed, measurement of biomass accumulation and doubling time are important factors in determining growth potential, especially for native species. The relative growth rates (RGR), doubling times (DT), biomass accumulation, and relative weekly yields (RY) of 40 distinct duckweed clones were determined under standard cultivation conditions. The dry weight–based RGR ranged from 0.149 to more than 0.600 per day, DT from 1.12 to 9 days, and RY from 7 to 108.9 per week. All values are comparable with previous studies. RGR and RY of selected clones are higher than the growth potential for a wide range of wild plants and common crops. These data support that native duckweed has high productivity value and should be further investigated as a potentially rich protein source for alternative human food, livestock feed, and recombinant protein production.

KEYWORDS

duckweed, *Lemna*, RGR, DNA barcoding, biomass accumulation, doubling time, biotechnology, food security

Introduction

The cosmopolitan monocotyledonous family Lemnaceae (duckweed), which includes the smallest and fastest growing angiosperms known, comprises 36 species in five genera: *Spirodela* Schleid., *Landoltia* Les & Crawford, *Lemna* L., *Wolffia* Horkel ex Schleid., and *Wolffiella* Hegelm. (Bog et al., 2020). However, some authors claim that duckweeds should be reclassified as subfamily Lemnoideae in the Araceae family based on a close phylogenetic relationship. However, the inclusion of duckweed in the Araceae would remove a useful and well-defined taxonomic category of an angiosperm family that has been used by duckweed biologists for many years (Tippery et al., 2021). These tiny aquatic plants grow on or below the surface of slow-flowing, nutrient-enriched water bodies. Their morphology is highly reduced to simple leaf-like structures also known as fronds that appear genus-specific with or without roots. Duckweeds can create genetically uniform populations by their rapid vegetative propagation. These properties, their small size, rapid and high yield growth, and additionally relatively small genome sizes make duckweeds an ideal experimental material and a powerful platform for various biotechnological applications (Appenroth et al., 2013; Heenatigala et al., 2020; Tippery et al., 2021). Under optimized growth conditions, duckweed contains a high protein content of up to 45% with high-quality and easily digestible amino acids close to the recommendations of the World Health Organization (WHO), which is an important nutritional index (Mes et al., 2022a; Pagliuso et al., 2022). Therefore, there is increased interest in the use of duckweed species from the genera *Wolffia* and *Lemna* as a good protein source, particularly for use in human food and animal nutrition (Edelman and Colt, 2016; Appenroth et al., 2017; Appenroth et al., 2018). Moreover, the high fiber content (~25% of dry weight) and polyunsaturated fatty acids (more than 60% of total fat) are shown to be a unique nutrient composition of duckweed (Appenroth et al., 2018). While major crops such as rice, maize, and wheat often have an imbalance in nutrient composition, the amino acids, vitamins, mineral profiles, and fatty acid fractions of duckweed have a high-quality composition (Edelman and Colt, 2016). The increase in world population, climate change, and decrease in food supply have increased pressure on food systems. In addition, excessive land use and agricultural activities have led to soil erosion, resulting in a 0.4% per year decline in global crop yields (Pennock, 2019). Therefore, there are increasing demands for the development of a sustainable food and feed safety system (Pagliuso et al., 2022). For example, the European Food and Safety Authority considers all genera of duckweed as novel foods (Mes et al., 2022b). Most notably, the ease of cultivation of these tiny aquatic plants in multilayered vertical farming systems and their high tolerance to a wide range of environmental conditions around the world may reduce competition with terrestrial crops (Coughlan et al., 2022; Pagliuso et al., 2022).

In addition, duckweed can be used in phytoremediation, water quality measurement, and wastewater treatment. Duckweeds have the ability to accumulate macronutrients and micronutrients hundreds of times compared with the mineral concentration of the water in which they proliferate (Chakrabarti et al., 2018; Walsh et al., 2021). The high biomass accumulation with a typically high protein content between 24% and 45% makes duckweed a suitable supplement for animal feed (Iatrou et al., 2015; Pagliuso et al., 2022). Further, duckweed can be used for bioethanol production due to its high starch accumulation under stress conditions (Sree and Appenroth, 2014; Liu et al., 2019).

The potential application of duckweed in plant bioreactors has attracted increasing attention due to its rapid doubling time (DT) and proliferation of uniform clones with nearly exponential growth (Edelman et al., 2020). Their growth rate is nearly 28 times faster than conventional crops used for human nutrition (Pagliuso et al., 2022; Coughlan et al., 2022). Undoubtedly, this higher reproductive rate will significantly shorten the production cycle of duckweed in bioreactors, leading to a maximum biomass accumulation of up to 100 tons of dry matter per hectare per year (Cao et al., 2018). Therefore, the study of biomass production and growth factors in duckweed under different cultivation conditions could be interesting. There is a number of pilot studies on duckweed biomass production under environmental conditions where wastewater or enriched medium is used to grow duckweed. The reports describe high-yield biomass production of 8 t dw/ha/y for *Wolffia arrhiza* (L.) Horkel ex Wimm. (Fujita et al., 1999) and 36 t dw/ha/y for *Spirodela polyrhiza* (L.) Schleid. (Xu et al., 2012) and up to 104 t/ha/year for *Lemna minor* L. (Frederic et al., 2006), which is comparable to the average yields of major land crops reported by the Food and Agriculture Organization of the United Nations (FAO, 2013) and the U.S. Department of Agriculture (USDA) (Ziegler et al., 2015).

The quality value of duckweed as a food source depends on the content and composition of constituents, particularly amino acid profiles, protein content, and high potential for rapid growth. Therefore, cultivation conditions and especially the genetic background of ecotypes are assumed to play a crucial role (Appenroth et al., 2018; Chakrabarti et al., 2018). Thus, assuming that different geographic isolates (ecotypes) of duckweeds have genetically differentiated due to adaptation to specific environmental conditions, the resulting clones may exhibit different physiological and growth behavior (Ziegler et al., 2015; Chakrabarti et al., 2018; Walsh et al., 2021). Studying a wide range of these ecotypes around the world may lead to the identification of superior clones in terms of growth characteristics that can be eligible for other studies in different

fields, such as biochemical analysis to introduce an alternative food supply.

To date, only a limited number of extensive studies have been conducted to investigate growth factors and biomass production in duckweed ecotypes from different parts of the world. Bergmann et al. (2000) studied 41 geographic isolates of 12 species from Landolt's worldwide stock collection at ETH Zurich (now hosted by the Istituto di Biologia e Biotecnologia Agraria in Milano, Italy) under *in vitro* conditions in a synthetic medium. To assess growth, they reported only wet weight gain and the percentage dry weight during the 11-day growth period for selection of superior geographic isolates (Bergmann et al., 2000). Among others, the relative growth rate (RGR) is an important growth factor that reflects the growth potential, especially in duckweeds. A comprehensive study of the RGR of duckweed was already carried out by Landolt (1957), who investigated 71 clones of 13 species. Ziegler et al. (2015) presented two other growth factors in addition to RGR to provide comparable data with other reports, e.g., on terrestrial crops. RGR, DT, and relative weekly yield (RY) of 39 ecotypes from 13 duckweed species were determined under standard cultivation conditions using a modified Schenk–Hildebrand medium for 7 days. Here, the mean RGR was 0.304 per day for *Spirodela* and 0.396 per day for *Lemna*. In general, RGR ranged from 0.153 to 0.519 per day, DT from 1.34 to 4.54 days, and RY from 2.9 to 37.8 per week for the duckweed species studied (Ziegler et al., 2015). Sree et al. (2015) investigated the RGR of 25 clones representing all 11 species of the genus *Wolffia*, the genus commonly used for human nutrition. They present a clone of *Wolffia microscopica* (Griff.) Kurz with a doubling time of 29.3 h, which is the fastest growing flowering plant (Sree et al., 2015). Other reports determining the RGR of duckweed species have reported an RGR of 0.31 per day for *Lemna minor* and 0.30 and 0.42 per day for *Lemna gibba* (Lasfar et al., 2007).

One of the first important steps is the precise identification of plant species. Several methods, such as morphological, biochemical, and molecular comparisons, can be used for correct identification. Among the mentioned marker types, comparison of molecular data with references is the most reliable (Dogan et al., 2014; Hasanbegovic et al., 2021; Saran et al., 2021). Specifically for duckweed, identification using only morphological characteristics is nearly impossible even for experts because the morphological structure of duckweed is greatly reduced. Using molecular methods, such as DNA barcoding, which is based on DNA markers, it is possible to reproducibly and reliably identify most duckweed species (Borisjuk et al., 2015; Bog et al., 2019). To the best of our knowledge, the native duckweeds of Iran have not been studied at the molecular level or in terms of growth rate. This is the first report on DNA barcoding of the only duckweed collection in Iran.

In the present study, RGR, DT, and RY for duckweed species native to Iran are investigated for the first time. Most of the clones studied are from the north of Iran, where most of the duckweed habitats are located. The investigation of 40 Iranian clones, which can be assigned to four species within the two genera *Lemna* and *Spirodela*, aims to determine accumulation of biomass yields and to identify the superior ecotypes with high growth potential under laboratory conditions for future biotechnological applications and food safety research.

Material and methods

Plant material

Duckweed samples were collected from different natural ponds in the north of Iran (Mazandaran and Gilan provinces) and a region in the west of Iran (Kermanshah province). The geographical distribution map of sampling can be found in Supplementary Figure 1. Fronds were rinsed in clean tap water and sterilized using 2.5% sodium hypochlorite solution for 1 min and subsequently washed three times with sterile distilled water. Sterilized fronds of all *Lemna* species were then cultivated in modified Hoagland medium (Khvatkov et al., 2019) except *L. gibba*, which was cultivated in NF medium (Muranaka et al., 2015). Schenk–Hildebrand medium (Schenk and Hildebrandt, 1972) was used for the *Spirodela* species. All media were supplemented with 1% sucrose. All clones were cultivated under standard cultivation conditions (ISO 20079, 2005) with a 16/8 h light/dark photoperiod with 80 µmol/m^2/s light intensity from fluorescent light tubes (40W) (Pars shahab, Tehran, Iran) at 25°C.

Morphological identification and molecular analysis

Duckweed samples were primarily identified morphologically using the key from Riedl (1976) and the updated key from Bog et al. (2020) based on frond shape, frond size, and number of roots and veins. The identity of the clones that were chosen for the biomass accumulation test was confirmed by DNA barcoding. For molecular analysis, total DNA was extracted by a modified CTAB protocol (Murray and Thompson, 1980). The chloroplast marker from the noncoding spacer *atpF-atpH* was amplified using the primers *atpF-atpH* forward (5' ACTCGCACACACTCCCTTTCC 3') and *atpF-atpH* reverse (5' GCTTTTATGGAAGCTTTAACAAT 3') as described previously (Wang et al., 2010). The PCR conditions were predenaturation at 94°C for 2 min, followed by 35 cycles of 94°C, 15 s; 51°C, 15 s; 72°C, 40 s; and a final extension at 72°

C for 5 min. The primer set of the second marker, *rps16*, was used to amplify the chloroplast ribosomal protein S16 gene intron with the degenerate primers *rps16* F (5' AAACGATGTGGTARAAAGCAAC 3') and *rps16* R (5' AACATCWATTGCAASGATTCGATA 3') as described previously (Shaw et al., 2005). The PCR conditions were predenaturation at 94°C for 5 min, followed by 45 cycles at 94°C, 30 s; 61°C, 50 s; 72°C, 80s; and a final extension at 72°C for 7 min. The PCR fragments were purified and further processed for sequencing by the Beijing Genomic Institute (BGI, Shenzhen, China). Sequences were deposited in GenBank (https://www.ncbi.nlm.nih.gov/). Accession numbers of the *rps16* and *atpF-atpH* sequences are listed in Table 1.

TABLE 1 Duckweed species investigated for this study with their NCBI accession numbers.

Row	Species	Strain	Origin	Accession number	
				atpF-atpH	*rps 16*
1	*Lemna minor*	1C	Tonekabon pond, Iran	MT891062	MW308182
2	*Lemna minor*	2C	Tonekabon pond, Iran	MT891063	MZ422535
3	*Lemna minor*	3a	Tonekabon pond, Iran	MT891064	MW308183
4	*Lemna minor*	4BM	Mansoori pond, Iran	MT891065	MZ422534
5	*Lemna minor*	5C	Tonekabon pond, Iran	MT891066	MW308184
6	*Lemna minor*	6a	Tonekabon pond, Iran	MT891067	MW308185
7	*Lemna minor*	7W	Tonekabon pond, Iran	MT891068	MW308186
8	*Lemna minor*	8C	Tonekabon pond, Iran	MT891069	MW308187
9	*Lemna minor*	9a	Tonekabon pond, Iran	MT891070	MW308188
10	*Lemna minor*	10C	Tonekabon pond, Iran	MT891071	MW308189
11	*Lemna minor*	11C	Tonekabon pond, Iran	MT891072	MW308190
12	*Lemna minor*	12W	Tonekabon pond, Iran	MT891073	MZ422533
13	*Lemna minor*	13C	Tonekabon pond, Iran	MT891074	MW308191
14	*Lemna minor*	14BM	Mansoori pond, Iran	MT891075	MZ422532
15	*Lemna minor*	15C	Tonekabon pond, Iran	MT891076	MW308192
16	*Lemna minor*	16C	Tonekabon pond, Iran	MT891077	MW308193
17	*Lemna turionifera*	17AM	Amir kelaye international lagoon, Iran	MT891086	MW308200
18	*Lemna minor*	18C	Tonekabon pond, Iran	MT891078	MW308194
19	*Lemna minor*	19C	Tonekabon pond, Iran	MT891079	MW308195
20	*Lemna turionifera*	20AM	Amir kelaye international lagoon, Iran	MT891087	MW308201
21	*Lemna minor*	21BM	Mansoori pond, Iran	MT891080	MW308196
22	*Lemna minor*	22a	Tonekabon pond, Iran	MT891081	MW308197
23	*Lemna minor*	23C	Tonekabon pond, Iran	MT891082	MW308198
24	*Lemna minor*	24W	Tonekabon pond, Iran	MT891083	MZ422531
25	*Lemna minor*	25C	Tonekabon pond, Iran	MT891084	MZ422530
26	*Lemna minor*	26W	Tonekabon pond, Iran	MT891085	MW308199
27	*Lemna gibba*	1LA	Langarud paddy,Iran	MT891088	MW308202
28	*Lemna gibba*	2S	Soostan Lagoon, Iran	MT891089	MW308203
29	*Lemna gibba*	3g	Govaver, Iran	MT891092	MW308206
30	*Lemna gibba*	4LA	Langarud paddy, Iran	MT891093	MW308207
31	*Lemna gibba*	5S	Soostan Lagoon, Iran	MT891091	MW308205
32	*Lemna gibba*	6LA	Langarud paddy, Iran	MT891090	MW308204
33	*Lemna gibba*	7S	Soostan Lagoon, Iran	MT891094	MZ422528
34	*Lemna gibba*	8S	Soostan Lagoon, Iran	MT891095	MZ422529
35	*Spirodela polyrhiza*	1AM	Amir kelaye international lagoon, Iran	MT891096	MW308208
36	*Spirodela polyrhiza*	2AM	Amir kelaye international lagoon, Iran	MT891097	MW308209
37	*Spirodela polyrhiza*	3LA	Langarud lagoon, Iran	MT891098	MW308210
38	*Spirodela polyrhiza*	4AM	Amir kelaye international lagoon, Iran	MT891099	MW308211
39	*Spirodela polyrhiza*	5BM	Mansoori pond, Iran	MT891100	MW308212
40	*Spirodela polyrhiza*	6AM	Amir kelaye international lagoon, Iran	MT891101	MZ422536

DNA barcoding analysis

To analyze the genetic diversity, *atpF-atpH* and *rps16* sequences were checked using Chromas Lite 2.6.2 (Technelysium Pty Ltd, South Brisbane, Australia) and aligned using the BioEdit Sequence Alignment Editor 7.1.3.0 (Hall, 1999). Reference sequences of *atpF-atpH* and *rps16* markers for DNA barcoding were prepared from already identified clones from the duckweed stock collection of the University Greifswald (Germany) or taken from GenBank (Table 2). They were chosen to represent a wide geographical distribution of the species. SeqState 1.4.1 (Müller, 2005) was used to recode insertion and deletion (indel) positions using the implemented Simmons and Ochoterena simple coding algorithm, leading to a final alignment length of 798 sites including 14 indel coded sites for *rps16* and a final alignment length of 661 sites including 12 indel coded sites for *atpF-atpH*. The indel coded alignments can be found as Supplementary Material 1 and 2. Subsequently, TCS 1.23 (Clement et al., 2000) was used with default settings to build haplotype networks for each chloroplast marker. Based on the haplotype results, the alignments were collapsed to unique haplotypes for which a maximum-likelihood tree was built using iqTREE 2.1.3 (Minh et al., 2020) with 1000 bootstrap replicates. The implemented ModelFinder (Kalyaanamoorthy et al., 2017) chose F81+F (*atpF-atpH*) and K3Pu+F (*rps16*) as the best-fit models according to the Bayesian information criterion. Finally, DnaSP 6.12.03 (Rozas et al., 2017) was run to count polymorphic and parsimony informative sites and to estimate nucleotide and haplotype diversity.

Culture conditions for growth factor analysis

Forty geographic isolates representing four species from two genera were used for biomass accumulation and DT analysis as described in Table 1. According to the ISO 20079 protocol (ISO 20079, 2005), the sterilized duckweed samples were precultivated for 1 month to acclimatize the clones to the cultivation conditions. Nutrient media were replenished every week. A single clone with the same frond number from each species was used for initial inoculation of 50 ml nutrient medium in glass jars covered with plastic caps. All glasses were kept under axenic conditions at 25°C in a standard growth chamber. The investigated duckweed species showed optimum growth in different media (unpublished results). For this reason, we used a specific nutrient medium for each species instead of the Steinberg medium specified in the ISO 20079 protocol as mentioned. The growth factors were determined starting with a four-frond colony for each species, and the initial weight was determined. The main cultivation phase lasted 7 days, taking care that the fronds never completely covered the surface of the medium, which may limit growth.

Calculation of growth parameters

All growth parameters were determined at the onset of the experiment (t_0) and 7 days later (t7). The number of fronds (FN_0 and FN_7), fresh weight (FW_0 and FW_7), and dry weight (DW_0 and DW_7) were measured. At the initiation of the experiment, the frond numbers were recorded. Then, an equal frond number and size was surface-dried by filter paper and weighed (FW_0). These reference samples were dried at 37°C for 72 h to determine the dry weight of the preliminary inoculum (DW_0). Frond number, fresh weight, and dry weight of fronds at t_7 were determined as for the reference samples from t_0. Three independent experiments were conducted with each of the clones.

RGR was calculated using Equation (1) (Naumann et al., 2007; Ziegler et al., 2015). This equation was simplified to

TABLE 2 Reference sequences of *atpF-atpH* and *rps16* markers for DNA barcoding.

Row	Reference sequences	Strain	Origin	Accession number	
				atpF-atpH	*rps 16*
1	*Lemna minor*	7123	Canada, Saskachewan, Saskatoon	MG000397	*
2	*Lemna minor*	8292	Iran, Mazanda, Ramsar, Ghassem Abbath	*	*
3	*Lemna minor*	9441	Germany, Marburg (clone St)	*	*
4	*Lemna turionifera*	6573	USA, Montana, Lincoln Co.	MG775403	*
5	*Lemna turionifera*	7683	Korea, Kyonggi, Sosa	MG775404	*
6	*Lemna turionifera*	9434	Russia, Lake Baikal	MG775405	*
7	*Lemna gibba*	7589	USA, California, Los Angeles Co., Covina	GU454219	*
8	*Lemna gibba*	7741	Italy, Sicilia, Siracusa (clone G3)	KX212887	*
9	*Lemna gibba*	8703	Japan, Honshu Aichi	GU454222	*
10	*Spirodela polyrhiza*	7373	Egypt, Mahallet, El Rahabein	HG938145	HG938250
11	*Spirodela polyrhiza*	7498	USA, North Carolina, Durham Co., Durham	GU454204	HG938251
12	*Spirodela polyrhiza*	9500	Germany, Jena, Porstendorf 1967 (clone SJ)	GU454208	HG938257

*sequenced but no Genbank number yet.

Equation (2) for better interpretation of growth potentials. The values of measured parameters x (Frond number or fresh and dry weight) in two time points (t_0 and t_7) were placed in Equation (2).

$$X_t = x_{t0} \times e^{RGR \times t} \quad (1)$$

$$RGR = (lnx_{t7} - lnx_{t0}) / (t7 - t0) \quad (2)$$

The RGR unit is based on time (per day). DT (days) or biomass accumulation (per day), was calculated by Equation (3), when RGR is measured with frond number values or fresh weight and dry weight of fronds at the two time points, respectively.

$$(BA)\ DT = ln2 / RGR \quad (3)$$

The yield obtained from the initial inoculum of one frond (or 1 mg duckweed biomass) after 7 days of cultivation is known as RY. It was calculated using Equation (4):

$$RY = \ln x_{t7} = \ln x_{to} + RGR \times (t7 - t0) \quad (4)$$

RY is equal to $\ln x_{t7}$, and x is one of the growth parameters measured in the experiment, such as FN, FW, and DW at t_0 ($\ln x_{t0}$) and at t_7 ($\ln x_{t7}$). The RY of one frond or 1 mg duckweed initial inoculum after 7 days has the unit per week.

Data analysis

Statistical analysis was carried out in SPSS 16.0 (IBM, USA). The normality of the data was confirmed by the Kolmogorov–Smirnov test. Therefore, parametric methods were used to compare the means. The variation of means among groups was compared with one-way ANOVA using the Student–Newman–Keuls test (SNK), a *post hoc* test for analysis of the differences in means, at the level of $P \leq .05$.

Results

DNA barcoding of duckweed ecotypes based on *rps16* and *atpF-atpH* sequences

A total of 40 duckweed clones were collected from the north of Iran (lakes of Tonekabon, Lahijan, and Langarud) and the Kermanshah Govaver River. The collected duckweed accessions were morphologically determined and validated by molecular methods, i.e., DNA barcoding (Figure 1).

In summary, a total of 24 ecotypes of *L. minor*, eight ecotypes of *L. gibba*, six ecotypes of *S. polyrhiza*, and two ecotypes of *L. turionifera*, a rare species for Iran, were identified with the chloroplast fragments *rps16* and *atpF-atpH*. All identified clones were successfully propagated to produce pure clones.

The species could be very well-distinguished by both chloroplast markers (*rps16* and *atpF-atpH*) as represented in the maximum-likelihood phylogenetic trees (Supplementary Figures 2A, B). Comparison of the two sequence alignments for both markers separately shows that *rps16* has a higher haplotype and nucleotide diversity than *atpF-atpH* although differences between different haplotypes within one species are most often caused by indels (Table 3, Figures 2, 3). For *rps16*, both clones identified as *L. turionifera* from Iran showed the same haplotype (LT1) as the reference sequence of clone 9434 from Lake Baikal, Russia. For *L. minor* the Iranian clones were identical to the haplotype (LM1) of 8292, a reference clone from Iran, too. Two further clones (25C – haplotype LM3 and 12W – haplotype LM4) showed one or two additional bases but are more similar to the main haplotype LM1 found for Iran than to the other two reference clones from Canada and Germany (haplotype LM2) (Figures 2A, B). For the marker *atpF-atpH* only *L. gibba* revealed different haplotypes, in which the Iranian clones differed by an additional stretch of three A's (haplotype lg1) from the three reference clones, which showed the same haplotype (lg2) (Figures 3A, B).

Biomass accumulation and doubling time

The growth potential measured as RGR and RY based on fresh and dry weight and DT based on frond numbers of 40 geographic duckweed isolates under axenic cultivation conditions are shown in Table 4. Overcrowding of populations was not observed during the experiment or at the end after 7 days. This ensured that the growth was not inhibited by intraspecific competition. In addition, axenic cultivation prevented inhibitory effects of undesirable microorganisms (Ziegler et al., 2015). During the 7 days of the experiment, the increase in frond number, fresh weight, and dry weight never deviated from an exponential progression.

The mean RGR for all 40 investigated ecotypes was 0.301 per day for fresh weight, ranging from 0.094 to 0.472 per day for individual clones belonging to *L. minor* 26W and *S. polyrhiza* 2AM, respectively. Additionally, the mean RGR for dry weight was 0.435 per day with a range from 0.149 to 0.784 per day for *Spirodela* clones (Table 4).

As shown in Figure 4A, the mean RGR based on fresh weight for four species belonging to two duckweed genera ranged from 0.284 per day for *S. polyrhiza* to 0.332 per day for *L. gibba*. The mean RGR for *L. minor* is 0.274 per day. The RGR for *L. gibba* was significantly higher than for any other species ($P<.05$). This was due to a higher weight gain and more fronds after 7 days of cultivation. There were no significant differences between the mean RGR_{FW} of *L. minor* and *S. polyrhiza* due to a wide range of RGR values among *L. minor* clones.

The dry weight analysis gave similar results to fresh weight. Among the 24 *L. minor* clones, the highest RGR_{DW} values were

10.3389/fpls.2022.1034238

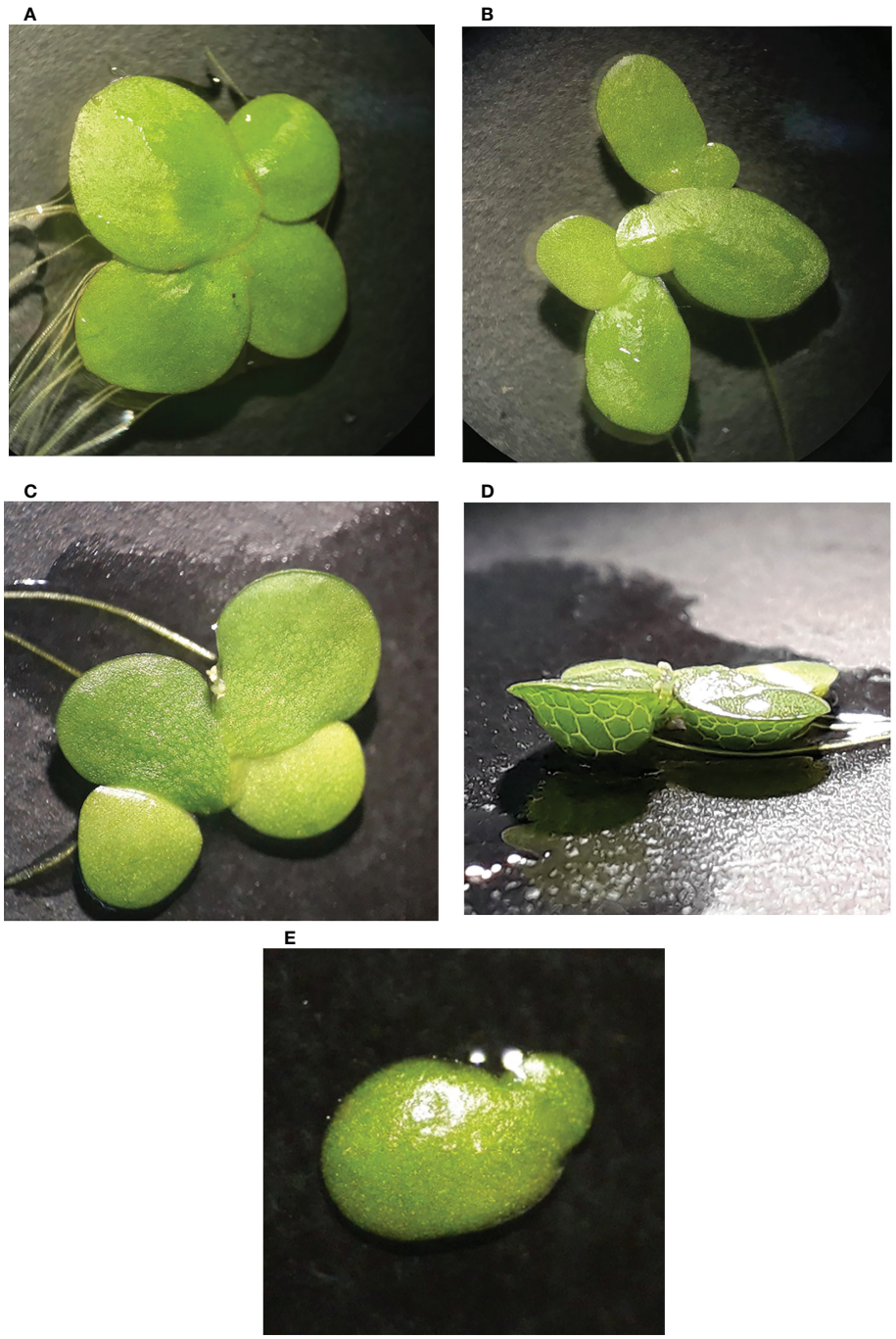


FIGURE 1
Four investigated duckweed species native to Iran; **(A)** *Spirodela polyrhiza*. **(B)** *Lemna minor*. **(C)** *L. gibba* in dorsal view. **(D)** *L. gibba* in ventral view. **(E)** *L. turionifera*.

obtained by clones 11C, 23C, and 24W with 0.563, 0.612, and 0.654 per day, respectively. These values were not significantly different from each other (P<.05). The clone with the significantly highest RGR_{DW} among the eight *L. gibba* clones was clone 7S with 0.420 per day. Among the six clones of *S. polyrhiza*, 4AM showed the highest value with RGR_{DW} = 0.784 per day. Fresh and dry weight RGR within species are significantly different (P<.05), especially in *L. minor* and *S.*

10.3389/fpls.2022.1034238

TABLE 3 Alignment characteristics for the two investigated chloroplast markers.

species	numberclones	*rps16*							*atpF-atpH*						
		alignmentlength (bp) *	indel codedsites	PS	PI	H_{num}	H_d ± SD	π ± SD	alignmentlength (bp) *	indel codedsites	PS	PI	H_{num}	H_d ± SD	π ± SD
S. polyrhiza	9	798	14	0	0	1	0.000± 0.000	0.0000± 0.0000	661	12	0	0	1	0.000± 0.000	0.0000± 0.0000
L. gibba	11			0	0	1	0.000± 0.000	0.0000± 0.0000			1	1	2	0.436± 0.133	0.0008± 0.0002
L. turionifera	5			2	2	2	0.600± 0.175	0.0016± 0.0005			0	0	1	0.000± 0.000	0.0000± 0.0000
L. minor	27			3	2	4	0.276± 0.109	0.0005± 0.0002			0	0	1	0.000± 0.000	0.0000± 0.0000

PS, polymorphic sites; PI, parsimony informative sites; H_{num}, number of haplotypes; H_d, haplotype diversity; π, nucleotide diversity; SD, standard deviation;* including indel coded sites.

polyrhiza. Some of the ecotypes of *L. minor* and *S. polyrhiza* studied as well as one of the clones of *L. turionifera* had an RGR higher than 0.600 per day (Table 4). This is a significantly higher RGR value compared with other ecotypes.

DT, which is based on the number of fronds, and BA, based on the fresh or dry weight, reflect the RGR value but numerically in the opposite way (see Equation (3) above). It indicates how much time is needed to double the number of fronds or biomass. DT based on frond number was investigated for all ecotypes. The lowest DT (rapid growth) within species was measured for *L. minor* 11C (2.96 days), *L. gibba* 6LA (2.14 days), and *S. polyrhiza* 4AM (2.33 days), which doubled their frond number every 51 to 71 h (Table 4). A comparison of the mean values for DT of the investigated clones of the four species is shown in Figure 4B. The mean value of frond doubling time for *L. gibba* with 2.26 days is significantly lower than that of the other species (P<.05).

The mean biomass accumulation based on fresh weight (BA_{FW}) of eight *L. gibba* clones was 2.12 days. Thus, this species had significantly higher productivity than the other species (P<.05; Figure 4C). On the other hand, the mean biomass accumulation based on dry weight (BA_{DW}) of *L. gibba* is consistent with BA_{FW}. However, this is not significantly different between species (Figure 4E). The significantly fastest BA_{FW} within *L. minor* was observed for strain 11C with 1.87 days (P<.05; Table 4). The two clones of the rare *L. turionifera* had BA rates of 1.60 and 3.78 days. To increase the accuracy of growth parameters in *L. turionifera*, it is necessary to continue the work with more ecotypes.

The RY of biomass accumulation based on fresh weight (RY_{FW}) of all investigated ecotypes ranged from 108.92 (*L. gibba* 2S) to 7.2 per week (*L. turionifera* 17AM). The highest mean RY_{FW} among the species belonged to *L. gibba* (82.34 per week) and the lowest mean RY_{FW} to *Spirodela* (35.76 per week) and *L. minor* (24.61 per week). The differences among the four species were statistically significant (P<.05; Figure 4D). Analysis of the RY value showed that the results were consistent with BA and DT. In addition, intraspecies data for *L. minor* showed that the highest relative yield of 37.09 per week was obtained by *L. minor* 11C due to its better growth potential. Despite the results for mean fresh weight (Figures 4D, F), where *L. gibba* showed the highest RY_{FW}, *S. polyrhiza* had the highest value in RY_{DW} in the species comparison with 23.78 per week (Figure 4F) while *L. gibba* (5.72 per week) and *L. minor* (3.70 per week) were both in the lower range.

Discussion

In the present study, biomass production screening was performed based on growth potential analysis on native duckweed species from Iran coupled with DNA barcoding

10.3389/fpls.2022.1034238

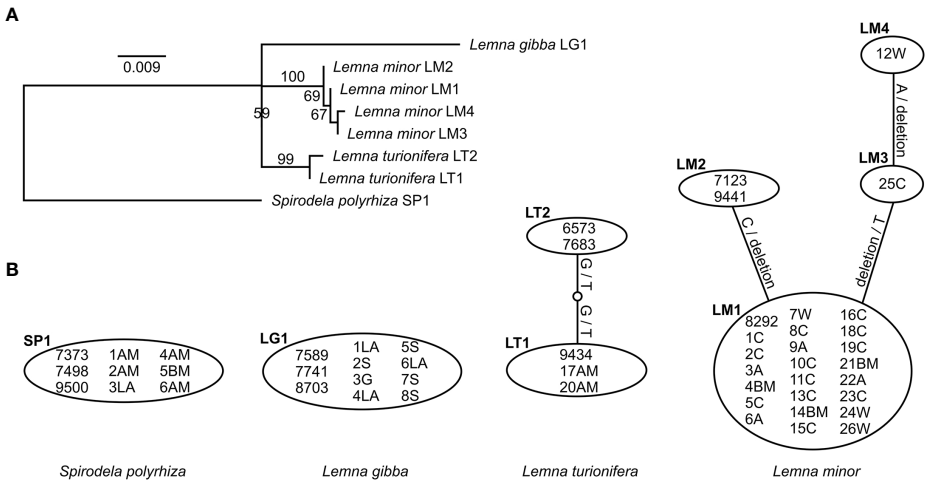


FIGURE 2
Molecular genetic results of the cp marker *rps16*. **(A)** Maximum-likelihood tree of unique haplotypes found for all investigated clones. Bootstrap values based on 1,000 replicates. *Spirodela polyrhiza* was set as outgroup. Scale indicates number of substitutions per site. **(B)** Identification of haplotypes and their differences for all investigated clones. Substitutions are given on the lines. - denotes deletion.

based on two standard chloroplast markers. Growth parameters were used to identify the most productive ecotypes of each of the four native duckweed species (*Spirodela polyrhiza*, *Lemna minor*, *L. gibba*, and *L. turionifera*) among the 40 clones studied. *Lemna minor* 11C, *L. gibba* 2S, and *S. polyrhiza* 4AM showed growth rates and a relative yield even higher than any terrestrial plants reported in previous studies (Ziegler et al., 2015; Koca and Erekul, 2016). These data demonstrate the importance of comprehensive studies on duckweed for biomass accumulation in biological production systems and food security. In addition, for the first time, the species *L. turionifera* was detected for Iran based on DNA barcoding analysis, and the two clones were included in the biomass screening.

Primary identification of duckweed based on morphological characters identified three species belonging to two genera: *Spirodela polyrhiza*, *Lemna minor*, and *L. gibba*. However,

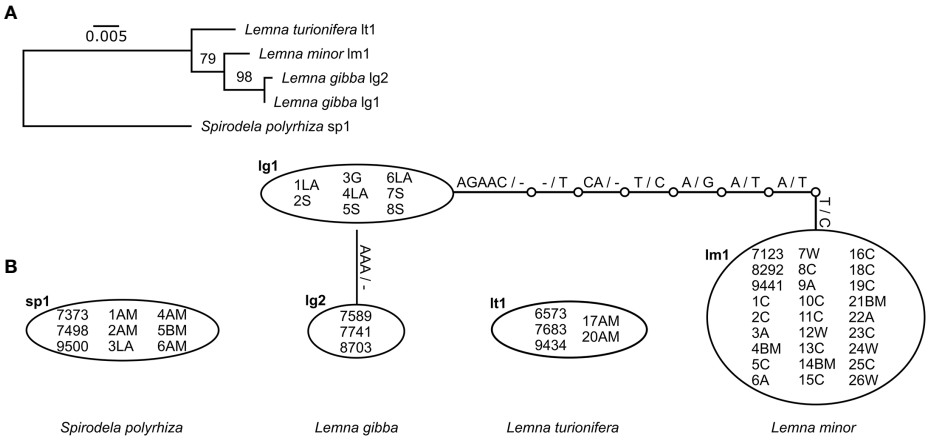


FIGURE 3
Molecular genetic results of the cp marker *atpF-atpH*. **(A)** Maximum-likelihood tree of unique haplotypes found for all investigated clones. Bootstrap values based on 1,000 replicates. *Spirodela polyrhiza* was set as outgroup. Scale indicates number of substitutions per site. **(B)** Identification of haplotypes and their differences for all investigated clones. Substitutions are given on the lines. - denotes deletion.

10.3389/fpls.2022.1034238

TABLE 4 Fresh and dry weight biomass accumulation and doubling time of 40 geographical isolates of native Iranian duckweed representing four species.

Row	Species	strain	DT_{FN}	RGR_{FW}	RGR_{DW}	BA_{FW}	BA_{DW}	RY_{FW}	RY_{DW}
1	*Lemna minor*	1C	3.77 ± 0.11	0.337 ± 0.029	0.478 ± 0.01	2.10 ± 0.18	1.51 ± 0.02	28.94 ± 5.55	3.71 ± 0.02
2	*Lemna minor*	2C	3.44 ± 0.12	0.279 ± 0.005	0.255 ± 0.01	2.49 ± 0.05	2.72 ± 0.06	30.68 ± 1.14	2.14 ± 0.09
3	*Lemna minor*	3a	5.97 ± 0.42	0.215 ± 0.001	0.307 ± 0.01	3.10 ± 0.06	2.39 ± 0.02	14.44 ± 0.25	2.05 ± 0.03
4	*Lemna minor*	4BM	3.73 ± 0.001	0.338 ± 0.006	0.277 ± 0.01	2.05 ± 0.03	2.50 ± 0.06	29.69 ± 1.14	1.75 ± 0.09
5	*Lemna minor*	5C	3.67 ± 0.15	0.343 ± 0.002	0.256 ± 0.01	2.02 ± 0.01	2.72 ± 0.11	34.87 ± 0.43	2.05 ± 0.14
6	*Lemna minor*	6a	7.05 ± 0.02 (-)	0.301 ± 0.025	0.287 ± 0.02	2.35 ± 0.20	2.44 ± 0.14	28.21 ± 4.84	1.75 ± 0.2
7	*Lemna minor*	7W	4.04 ± 0.01	0.215 ± 0.001	0.332 ± 0.02	3.23 ± 0.02	2.18 ± 0.03	28.06 ± 0.20	2.94 ± 0.09
8	*Lemna minor*	8C	5.34 ± 0.16	0.202 ± 0.005	0.156 ± 0.01 (-)	3.44 ± 0.08	4.52 ± 0.37 (-)	12.65 ± 0.43	1.30 ± 0.12
9	*Lemna minor*	9a	3.74 ± 0.01	0.277 ± 0.007	0.294 ± 0.003	2.50 ± 0.07	2.36 ± 0.06	22.78 ± 1.14	1.10 ± 0.06
10	*Lemna minor*	10C	3.40 ± 0.04	0.260 ± 0.002	0.273 ± 0.01	2.67 ± 0.01	2.68 ± 0.01	15.86 ± 0.01	1.40 ± 0.01
11	*Lemna minor*	11C *	2.96 ± 0.02 (+)	0.373 ± 0.017 (+)	0.563 ± 0.01	1.87 ± 0.08 (+)	1.23 ± 0.03	37.09 ± 4.27 (+)	2.59 ± 0.23
12	*Lemna minor*	12W	4.97 ± 0.01	0.293 ± 0.014	0.302 ± 0.02	2.38 ± 0.12	2.31 ± 0.12	20.31± 2	1.35 ± 0.14
13	*Lemna minor*	13C	4.43 ± 0.01	0.306 ± 0.007	0.191 ± 0.01	2.27 ± 0.05	3.63 ± 0.14	23.77 ± 1.14	2.25 ± 0.11
14	*Lemna minor*	14BM	4.04 ± 0.01	0.301 ± 0.003	0.258 ± 0.01	2.23 ± 0.03	2.71 ± 0.12	29.44 ± 0.71	3.24 ± 0.26
15	*Lemna minor*	15C	4.97 ± 0.02	0.350 ± 0.005	0.369 ± 0.02	1.98 ± 0.03	1.90 ± 0.12	25.25 ± 0.85	2.74 ± 0.43
16	*Lemna minor*	16C	4.43 ± 0.01	0.326 ± 0.002	0.489 ± 0.01	2.12 ± 0.01	1.42 ± 0.04	26.24 ± 0.28	1.55 ± 0.14
18	*Lemna minor*	18C	4.68 ± 0.11	0.287 ± 0.029	0.326 ± 0.03	2.40 ± 0.30	2.18 ± 0.19	19.40 ± 0.24	1.35 ± 0.26
19	*Lemna minor*	19C	4.61 ± 0.45	0.234 ± 0.015	0.216 ± 0.02	3 ± 0.19	3.26 ± 0.24	21.79 ± 2.28	1.85 ± 0.20
21	*Lemna minor*	21BM	5.07 ± 0.27	0.256 ± 0.006	0.327 ± 0.02	2.72 ± 0.07	2.14 ± 0.10	26.73 ± 1.14	1.10 ± 0.12 (-)
22	*Lemna minor*	22a	3.78 ± 0.26	0.317 ± 0.019	0.512 ± 0.01	2.21 ± 0.13	1.36 ± 0.04	35.60 ± 4.55	3.64 ± 0.37
23	*Lemna minor*	23C	3.96 ± 0.19	0.260 ± 0.009	0.612 ± 0.01	2.68 ± 0.09	1.15 ± 0.01	33.14 ± 1.99	8.21 ± 0.85 (+)
24	*Lemna minor*	24W	5.07 ± 0.27	0.229 ± 0.007	0.654 ± 0.03 (+)	3.03 ± 0.09	1.06 ± 0.01 (+)	17.35 ± 0.86	1.95 ± 0.03
25	*Lemna minor*	25C	2.97 ± 0.02	0.175 ± 0.002	0.406 ± 0.07	3.95 ± 0.05	1.90 ± 0.35	20.31 ± 0.29	4.59 ± 0.34
26	*Lemna minor*	26W	5.07 ± 0.27	0.094 ± 0.010 (-)	0.229 ± 0.01	7.62 ± 0.84 (-)	3.04 ± 0.13	7.95 ± 0.57 (-)	2.14 ± 0.14
	Lemna minor	mean	4.38 ± 0.05	0.274 ± 0.002	0.349 ± 0.02	2.77 ± 0.04	2.31 ± 0.03	24.61 ± 0.42	3.70 ± 0.05
17	*Lemna turionifera*	17AM	9.57 ± 0.03	0.196 ± 0.033	0.199 ± 0.03	3.36 ± 0.36	3.78 ± 0.61	7.20 ± 1.57	0.95 ± 0.20
20	*Lemna toriunifera*	20AM	4.22 ± 0.30	0.433 ± 0.003	0.628 ± 0.02	1.60 ± 0.01	1.11 ± 0.04	14.38 ± 0.29	0.85 ± 0.14
	Lemna turionifera	mean	6.90 ± 0.13	0.314 ± 0.018	0.413 ± 0.02	2.48 ± 0.19	2.45 ± 0.28	10.79 ± 0.93	0.90 ± 0.03
27	*Lemna gibba*	1LA	2.50 ± 0.01 (-)	0.274 ± 0.002 (-)	0.329 ± 0.01 (-)	2.53 ± 0.02 (-)	2.11 ± 0.02	60.49 ± 0.99 (-)	3.98 ± 0.01 (-)
28	*Lemna gibba*	2S *	2.16 ± 0.04	0.393 ± 0.011 (+)	0.376 ± 0.01	1.77 ± 0.05 (+)	1.85 ± 0.06	108.9 ± 7.94 (+)	6.96 ± 0.57 (+)
29	*Lemna gibba*	3g	2.16 ± 0.07	0.379 ± 0.011	0.391 ± 0.03	1.83 ± 0.05	1.76 ± 0.06	69.10 ± 5.11	4.68 ± 0.40
30	*Lemna gibba*	4LA	2.17 ± 0.03	0.355 ± 0.002	0.338 ± 0.01	1.95 ± 0.01	2.05 ± 0.02	97.13 ± 1.42	6.36 ± 0.11
31	*Lemna gibba*	5S	2.35 ± 0.02	0.306 ± 0.004	0.403 ± 0.01	2.26 ± 0.03	1.72 ± 0.02 (+)	75.98 ± 2.27	5.02 ± 0.20
32	*Lemna gibba*	6LA	2.14 ± 0.01 (+)	0.321 ± 0.002	0.333 ± 0.03	2.16 ± 0.01	2.15 ± 0.01 (-)	93.20 ± 1.42	6.66 ± 0.06
33	*Lemna gibba*	7S	2.48 ± 0.05	0.298 ± 0.007	0.420 ± 0.02 (+)	2.33 ± 0.06	1.80 ± 0.01	60.73 ± 3.13	5.92 ± 0.03
34	*Lemna gibba*	8S	2.16 ± 0.01	0.333 ± 0.003	0.412 ± 0.01	2.08 ± 0.01	1.77 ± 0.02	93.10 ± 0.01	6.17 ± 0.06
	Lemna gibba	mean	2.26 ± 0.01	0.332 ± 0.003	0.375 ± 0.01	2.12 ± 0.01	1.90 ± 0.02	82.34 ± 1.65	5.72 ± 0.13
35	*Spirodela polyrhiza*	1AM	7.22 ± 0.93 (-)	0.323 ± 0.011	0.425 ± 0.04	2.16 ± 0.08	1.63 ± 0.02	28.70 ± 2.28	2.74 ± 0.09 (-)
36	*Spirodela polyrhiza*	2AM *	2.84 ± 0.02	0.472 ± 0.003 (+)	0.498 ± 0.05	1.47 ± 0.01 (+)	1.43 ± 0.14	59.25 ± 1.14 (+)	10.85 ± 0.03
37	*Spirodela polyrhiza*	3LA	3.86 ± 0.23	0.379 ± 0.011	0.382 ± 0.01	1.83 ± 0.05	1.82 ± 0.06	52.36 ± 3.98	4.08 ± 0.34
38	*Spirodela polyrhiza*	4AM *	2.33 ± 0.05 (+)	0.200 ± 0.009	0.784 ± 0.02 (+)	3.48 ± 0.16	0.88 ± 0.02 (+)	40.04 ± 2.56	50.88 ± 4.83 (+)
39	*Spirodela polyrhiza*	5BM	2.93 ± 0.15	0.214 ± 0.018	0.650 ± 0.02	3.30 ± 0.27	1.07 ± 0.04	16.36 ± 2 (-)	40.07 ± 0.04
40	*Spirodela polyrhiza*	6AM	3.9 ± 0.21	0.117 ± 0.014 (-)	0.149 ± 0.07 (-)	6.16 ± 0.73 (-)	13.55 ± 6.35 (-)	17.84 ± 1.71	34.01 ± 0.01
	Spirodela polyrhiza	mean	3.85 ± 0.21	0.284 ± 0.002	0.481 ± 0.01	3.07 ± 0.08	3.40 ± 1.05	35.76 ± 0.47	23.78 ± 0.76

DT, doubling time (days); BA, biomass accumulation (per day); RGR, relative growth rate (per day); RY, relative yield (per week). Mean values for each species are presented in high light rows. Values are mean ± SE.
(+) indicates the highest growth rate values and (-) indicates the lowest growth rate values.
*Strains with the highest growth rate.

10.3389/fpls.2022.1034238

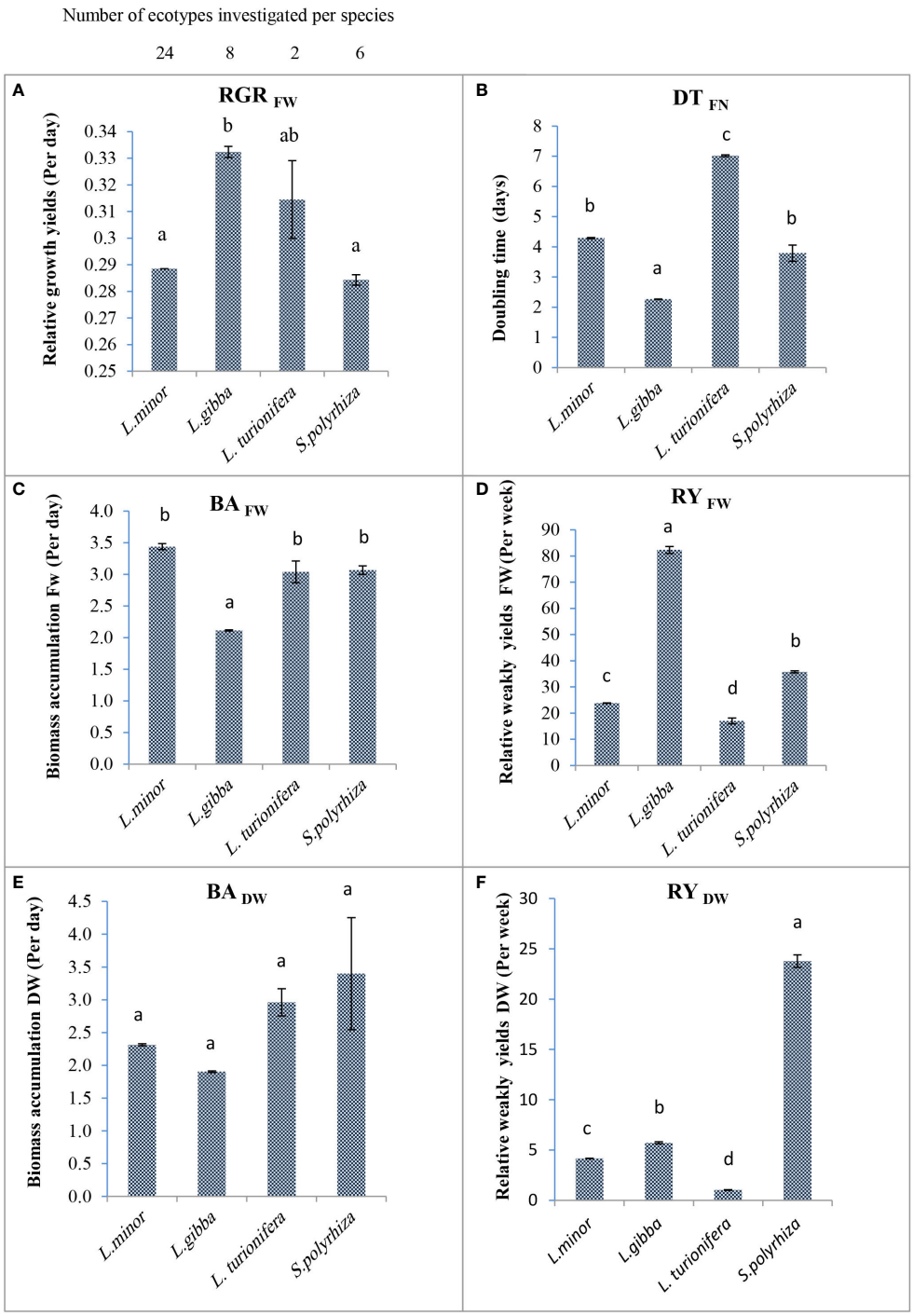


FIGURE 4
The mean fresh and dry weight–based relative growth rate (biomass accumulation, relative weekly yield, and doubling time based on frond number) of the four investigated duckweed species. **(A)** The mean RGR for the species represented by several clones. The number of clones per species is given above. **(B)** The mean value of DT_{FN} between species is significantly different. *Lemna gibba* by doubling its fronds every 2.26 days is faster than other species. **(C)** The mean BA_{FW} of four species of two genera is shown. **(D)** The maximum rate of mean RY_{FW} of four species belongs to *L. gibba* (82.3 per week). **(E)** The mean of BA_{DW} is not significantly different among species. **(F)** The mean dry weight (mg) produced after 1 week (RY) from primary inoculum among species is significantly different ($P<.05$). In all figures: The column height shows the mean growth parameters of the four investigated species. Error bar was indicated in the figures. Letters above the columns indicate significance according to ANOVA: means of columns marked with the same letter (either 'a' or 'b') do not differ to a statistically significant extent; differences are statistically significant when the columns are marked with single, different letters. The means of the columns marked with 'ab' do not differ significantly from means of columns marked with either 'a' or 'b'.

subsequent DNA barcoding analysis revealed a fourth species: *L. turionifera*, which is not easily distinguishable from *L. minor* due to the strong reduction in their morphology. After an appropriate literature research, *L. turionifera* was not listed for Iran in Landolt's monograph (Landolt, 1986), nor in the online Flora of Iran (2022) (http://flora-iran.com/central-herbarium-of-tehran-university/plant-list/), nor in the two online platforms "Plants of the World Online (2022)" (https://powo.science.kew.org/) and "Global Biodiversity Information Facility – GBIF (2022)" (https://www.gbif.org/).

The variation of the studied sequences is rather low, as is known for several duckweed species (Borisjuk et al., 2015; Chen et al., 2022), but also to be expected for a DNA barcode as they should show lower genetic variation within than between species (Dasmahapatra and Mallet, 2006). Interestingly, with the exception of two accessions, all other accessions identified as *L. minor* show the same haplotype as the reference sequence of *rps16* for clone 8292 from Iran, which was already mentioned in Landolt and Urbanska-Worytkiewicz (1980) and, thus, has been kept in culture for more than 40 years, which again could indicate a relatively constant haplotype pool. However, further studies are needed to reach similar conclusions as for *S. polyrhiza*, namely, that the mutation rate in this species is very low (Ho et al., 2019; Xu et al., 2019). Because the molecular markers used do not allow us to draw conclusions about hybridization events, the possibility that the *L. minor* clones identified here may be hybrids between *L. minor* and *L. turionifera* cannot be ruled out. Both species occur in the area, and hybridization between these two species has already been demonstrated by Braglia et al. (2021).

RGR is an important factor to show physiological responses of plants to light, temperature, CO_2, and nutrients, but its interpretation is less intuitive (Buxbaum et al., 2022). Therefore, it is often mathematically transformed and reported as BA or RY, which are common parameters for large-scale screening of plant growth potential (Ziegler et al., 2015). Fast-growing species show higher RGR under standard cultivation conditions. Higher RGR is due to efficient nutrient and CO_2 uptake in fronds (Naumann et al., 2007; Ziegler et al., 2015; Ghanem et al., 2019). High RGR causes duckweed to rapidly double its frond number and biomass, which increases its photosynthetically active surface area per unit area. One of the most important variations is the nutrient medium. Many studies demonstrate that duckweed exhibits optimal growth potential and accumulates more biomass in species-specific nutrient media (Kittiwongwattana and Vuttipongchaikij, 2013; Muranaka et al., 2015; Ghanem et al., 2019). This study was the first to use a nutrient medium optimized for growth for each species (data in preparation).

In this study, a relative linear relationship between RGR, DT or BA, and RY was investigated (Figure 4). In addition, growth factors based on dry weight, especially RY_{DW}, are more reliable parameters to study the stability of biomass gain in the duckweed family. Approximately 92%–94% of the fresh weight of duckweed consist of water, which is lost after drying (Pagliuso et al., 2022). This is confirmed by the relatively low yield (RY_{DW}) results for *L. gibba* shown in Figure 4F. Despite the highest mean RY_{FW} value (Figure 4D, F) for *L. gibba*, almost 90% of the fresh weight was lost to drying. In contrast, for *Spirodela*, only 35% of the fresh weight was lost to drying, and on average, 65% of the weight was retained as dried biomass (Table 4 and Figure 4F).

Based on the analysis of growth potential of 40 ecotypes, ecotypes with better growth potential under standard growing conditions were identified. *Spirodela polyrhiza* and *L. gibba* showed higher average growth factors among the studied species in this experiment. On the other hand, *L. minor*, the most widespread and easily manipulated species, with ecotype 11C provided one of the most productive duckweed clones in this experiment for future research. Considering that growth potential is shown to have a wide range of values within genera and species and, thus, significant differences among clones and ecotypes, it is concluded that species or genus specificity is not a reliable method for screening growth potential. In other words, growth parameters determined for one species or genus cannot be generalized to all clones of a species. This is consistent with Bergmann et al. (2000), who suggest focusing on the geographic isolate (ecotypes) rather than the species level. Ziegler et al. (2015) also confirm that screening on the ecotype level is the most reliable method for screening growth factors in duckweed. We found three ecotypes with high growth potential such as *S. polyrhiza* 2AM with an RGR_{FW} of 0.472 per day, which is higher than the value reported by Ziegler et al. (2015), ranging from 0.168 to 0.386 per day for seven *S. polyrhiza* clones. Other dependent parameters, such as BA_{FW} (1.47 per day) and RY_{FW} (59.25 per week), are consistent with those reported in the literature (Ziegler et al., 2015). In addition, *L. gibba* 2S was found to be the fastest ecotype in biomass accumulation with 1.77 per day. Similarly, the relative yield after 7 days of inoculation is the highest value (108.92 per week) for *L. gibba* 2S, which is consistent with those from previous reports (Ziegler et al., 2015). It is noteworthy that *L. gibba* already has a good potential to gain biomass in a short time. This is due to the rapid proliferation rate and gibbous fronds in this species. As mentioned earlier, under unsuitable culture conditions, *L. gibba* loses its gibbosity and looks like *L. minor*. NF is the best culture medium for *L. gibba* to form gibbous fronds and increase the rate of biomass accumulation (Muranaka et al., 2015). In addition, the widely distributed duckweed species *L. minor* with ecotype 11C has an RGR_{FW} value of 0.373 per day. This is comparable to the value of RGR_{FW} = 0.422 per day reported by Chakrabarti et al. (2018), where *L. minor* was cultivated with high-efficiency organic manure. The RGR_{DW} of *L. minor* 11C was 0.563 per day, which was among the highest values in the

study, while Petersen et al. (2021) obtained a maximum RGR_{DW} of 0.23 per day for *L. minor* under nonsterile conditions and with high-yielding agricultural fertilizer treatment. Despite the different cultivation conditions, the highly concentrated nitrate-N medium and light intensity (270 µmol/m^2/s higher than in the present study) were effective factors for optimal growth.

Duckweed is described as one of the fastest angiosperms due to its ability to double its biomass in a short time period. Demmig-Adams et al. (2022) suggest that the rapid growth of duckweed, even under limited light conditions, may be related to relatively thin photosynthetic organs (fronds) without complex structures on the water surface, allowing all chloroplasts to be involved in sugar production. The free-floating fronds with high availability of nutrient resources have higher photosynthetic yields. Terrestrial plants, on the other hand, use a significant amount of sugar to build the complex structures of their stems, leaves, and roots, which is a time-consuming process that involves the production of some organs that are unusable for human nutrition. The presented results shown in Table 4 confirm the superiority of growth rate of duckweeds by comparing their RGR with the RGR of crops. Of course, plant species differ greatly in their relative growth rate even when compared under similar environmental conditions (Tomlinson et al., 2014). However, duckweed is shown to have a higher RGR_{DW} compared with many crops despite its reduced and leaf-like structure that makes it one of the lightweights among crops. Poorter (1989) reports RGR_{DW} of eight herbaceous wild species under optimized growth conditions with the highest RGR value of 0.268 per day for *Urtica dioica*. Potter and Jones (1977) report RGR_{DW} after 28 days for nine species of important crops with a value ranging from 0.202 per day for soybean to 0.391 per day for sorghum. In contrast, an RGR value of 0.255 per day was determined for maize under the best experimental conditions. The data presented in Table 4 and Figure 4 show that most of the duckweed clones that were studied had an RGR based on fresh and dry weight (and corresponding RY) that was higher than the 0.255 per day described for maize. Overall, the superior growth characteristics and short harvest time of duckweed compared with crops may lead to higher biomass and economic production.

As reported by the Food and Agriculture Organization of the United Nations (FAO), food security is one of the most important challenges in the world (https://www.fao.org/state-of-food-security-nutrition/2021). Due to the increase in world population, reduction in food resources, depletion of nutrients in soils, and global climate change leading to a reduction in crop production, it is imperative to pay more attention to alternatives with low water and soil requirements that are cost-efficient, have short harvesting times, and produce more biomass. Duckweed as an aquatic crop has several advantages over terrestrial crops: it absorbs nutrients directly from water, is easy to grow and harvest, has low water requirements (due to its short growing season), and does not compete with crops in agricultural land use, allowing for higher biomass production per hectare even in dry areas (Toyama et al., 2017; Tursi, 2019). To date, interest in large-scale cultivation of duckweed in greenhouses and under protective structures has grown through commercial companies, such as LENTEIN™ (https://www.parabel.com/) and Rubisco Foods (2022) (https://rubiscofoods.com/).

Conclusion

The results of this study show that three selected ecotypes of *Lemna* and *Spirodela* species can provide high yields of fresh and dry biomass under optimal growth conditions. The data confirm that most ecotypes of duckweed can grow faster than traditional crop plants. However, this high RGR obtained under the optimized culture conditions for the selected ecotypes may be different under real environmental conditions. However, the fact that duckweed as an emergent crop has a high relative yield and accumulates more biomass in a short time was also confirmed under agricultural cultivation conditions. The data in this study illustrate the numerous potentials of the selected duckweed ecotypes for commercial biomass production and biotechnological application. However, several strategies are needed to optimize duckweed growth with low cost, simplicity, and scalability.

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 in the article/Supplementary Material.

Author contributions

AS is the executor of plan, ET is PhD student and project manager, MB contributed to DNA barcoding analysis, FF, SS, NR and MA cooperated in biomass measurements, MJ organized the database. All authors contributed to manuscript revision, read, and approved the submitted version.

Funding

This research was supported partially by grants from Iran National Science Foundation (INSF) (97024978), Center for International Scientific Studies and Collaboration (CISSC) (990160) and National Institute of Genetic Engineering and Biotechnology (NIGEB) (980301-II-714).

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/fpls.2022.1034238/full#supplementary-material

References

Appenroth, K. J., Borisjuk, N., and Lam, E. (2013). Telling duckweed apart: genotyping technologies for the lemnaceae. *Chin. J. Appl. Environ. Biol.* 19, 1–10. doi: 10.3724/SP.J.1145.2013.00001

Appenroth, K. J., Sree, K. S., Bog, M., Ecker, J., Seeliger, C., Böhm, V., et al. (2018). Nutritional value of the duckweed species of the genus wolffia (Lemnaceae) as human food. *Front. Chem.* 6, 483. doi: 10.3389/fchem.2018.00483

Appenroth, K.-J., Sree, K. S., Böhm, V., Hammann, S., Vetter, W., Leiterer, M., et al. (2017). Nutritional value of duckweeds (Lemnaceae) as human food. *Food Chem.* 217, 266–273. doi: 10.1016/j.foodchem.2016.08.116

Bergmann, B. A., Cheng, J., Classen, J., and Stomp, A. M. (2000). *In vitro* selection of duckweed geographical isolates for potential use in swine lagoon effluent renovation. *Bioresour. Technol.* 73 (1), 13–20. doi: 10.1016/S0960-8524(99)00137-6

Bog, M., Appenroth, K. J., and Sree, K. S. (2019). Duckweed (Lemnaceae): its molecular taxonomy. *Front. Sustain. Food Syst.* 3, 117. doi: 10.3389/fsufs.2019.00117

Bog, M., Appenroth, K. J., and Sree, K. S. (2020). Key to the determination of taxa of Lemnaceae: an update. *Nord. J. Bot.* 38 (8), e02658. doi: 10.1111/njb.02658

Borisjuk, N., Chu, P., Gutierrez, R., Zhang, H., Acosta, K., Friesen, N., et al. (2015). Assessment, validation and deployment strategy of a two-barcode protocol for facile genotyping of duckweed species. *Plant Biol.* 17, 42–49. doi: 10.1111/plb.12229

Braglia, L., Lauria, M., Appenroth, K. J., Bog, M., Breviario, D., Grasso, A., et al. (2021). Duckweed species genotyping and interspecific hybrid discovery by tubulin-based polymorphism fingerprinting. *Front. Plant Sci.* 12, 625670. doi: 10.3389/fpls.2021.625670

Buxbaum, N., Lieth, J. H., and Earles, M. (2022). Non-destructive plant biomass monitoring with high spatio-temporal resolution *via* proximal RGB-d imagery and end-to-End deep learning. *Front. Plant Sci.* 13, 758818. doi: 10.3389/fpls.2022.758818

Cao, H. X., Fourounjian, P., and Wang, W. (2018). "The importance and potential of duckweeds as a model and crop plant for biomass-based applications and beyond," in *Handbook of environmental materials management* (Cham: Springer International Publishing), 1–16.

Chakrabarti, R., Clark, W. D., Sharma, J. G., Goswami, R. K., Shrivastav, A. K., and Tocher, D. R. (2018). Mass production of *Lemna minor* and its amino acid and fatty acid profiles. *Front. Chem.* 6, 479. doi: 10.3389/fchem.2018.00479

Chen, G., Stepanenko, A., Lakhneko, O., Zhou, Y., Kishchenko, O., Peterson, A., et al. (2022). Biodiversity of duckweed (Lemnaceae) in water reservoirs of Ukraine and China assessed by chloroplast DNA barcoding. *Plants* 11 (11), 1468. doi: 10.3390/plants11111468

Clement, M., Posada, D. C. K. A., and Crandall, K. A. (2000). TCS: a computer program to estimate gene genealogies. *Mol. Ecol.* 9 (10), 1657–1659. doi: 10.1046/j.1365-294x.2000.01020.x

Coughlan, N. E., Walsh, É., Bolger, P., Burnell, G., O'Leary, N., O'Mahoney, M., et al. (2022). Duckweed bioreactors: Challenges and opportunities for large-scale indoor cultivation of lemnaceae. *J. Clean. Prod.* 336, 130285. doi: 10.1016/j.jclepro.2021.130285

Dasmahapatra, K. K., and Mallet, J. (2006). DNA Barcodes: recent successes and future prospects. *Heredity* 97 (4), 254–255. doi: 10.1038/sj.hdy.6800858

Demmig-Adams, B., López-Pozo, M., Polutchko, S. K., Fourounjian, P., Stewart, J. J., Zenir, M. C., et al. (2022). Growth and nutritional quality of lemnaceae viewed comparatively in an ecological and evolutionary context. *Plants* 11 (2), 145. doi: 10.3390/plants11020145

Dogan, H., Ercisli, S., Temim, E., Hadziabulic, A., Tosun, M., Yilmaz, S. O., et al. (2014). Diversity of chemical content and biological activity in flower buds of a wide number of wild grown caper (*Capparis ovate* desf.) genotypes from Turkey. *C R Acad. Bulg Sci.* 67 (11), 1593–1600.

Edelman, M., Appenroth, K. J., and Sree, K. S. (2020). Duckweed: Biological chemistry and applications. *Front. Sustain. Food Syst.* 4, 615135. doi: 10.3389/fsufs.2020.615135

Edelman, M., and Colt, M. (2016). Nutrient value of leaf vs. seed. front. *Chem.* 4, 32. doi: 10.3389/fchem.2016.00032

FAO. (2013). (*Food and Agriculture of the United Nations*). Available at: https://www.faostat3.fao.org/faostat-gateway (accessed 30 May 2013).

Flora of Iran (2022) *Plant-list*. Available at: http://flora-iran.com/central-herbarium-of-tehran-university/plant-list/ (Accessed 30/05/2022).

Food and Agriculture Organization of the United Nations (2021) *Food-security*. Available at: https://www.fao.org/state-of-food-security-nutrition (Accessed 23/07/2022).

Frederic, M., Samir, L., Louise, M., and Abdelkrim, A. (2006). Comprehensive modeling of mat density effect on duckweed (*Lemna minor*) growth under controlled eutrophication. *Water Res.* 40, 2901–2910. doi: 10.1016/j.watres.2006.05.026

Fujita, M., Mori, K., and Kodera, T. (1999). Nutrient removal and starch production through cultivation of *Wolffia arrhiza*. *J. Biosci. Bioeng.* 87, 194–198. doi: 10.1016/S1389-1723(99)89012-4

Ghanem, H., Haddad, A., Baydoun, S., Abou Hamdan, H., Korfali, S., and Chalak, L. (2019). *In vitro* proliferation of Lebanese *Lemna minor* and *Lemna gibba* on different nutrient media. *J. Taibah Univ. Sci.* 13 (1), 497–503. doi: 10.1080/16583655.2019.1597450

Global Biodiversity Information Facility – GBIF (2022). Available at: https://www.gbif.org/ (Accessed 30/05/2022).

Hall, T. A. (1999). "BioEdit: a user-friendly biological sequence alignment editor and analysis program for windows 95/98/NT," in *Nucleic acids symp. ser*, vol. Vol. 41. (London: Information Retrieval Ltd), 95–98.

Hasanbegovic, J., Hadziabulic, S., Kurtovic, M., Gasi, F., Lazovic, B., Dorbic, B., et al. (2021). Genetic characterization of almond (Prunus amygdalus l) using microsatellite markers in the area of Adriatic Sea. *Turk. J. Agric. For.* 45, 797–806. doi: 10.3906/tar-2103-82

Heenatigala, P. P. M., Sun, Z., Yang, J., Zhao, X., and Hou, H. (2020). Expression of LamB vaccine antigen in *Wolffia globosa* (duck weed) against fish vibriosis. *Front. Immunol.* 11, 1857. doi: 10.3389/fimmu.2020.01857

Ho, E. K., Bartkowska, M., Wright, S. I., and Agrawal, A. F. (2019). Population genomics of the facultatively asexual duckweed *Spirodela polyrhiza*. *New Phytol.* 224 (3), 1361–1371. doi: 10.1111/nph.16056

Iatrou, E. I., Stasinakis, A. S., and Aloupi, M. (2015). Cultivating duckweed *Lemna minor* in urine and treated domestic wastewater for simultaneous biomass production and removal of nutrients and antimicrobials. *Ecol. Eng.* 84, 632–639. doi: 10.1016/j.ecoleng.2015.09.071

ISO 20079 (2005). Water quality – determination of the toxic effect of water constituents and waste water on duckweed (*Lemna minor*) – duckweed growth inhibition test. ISO TC 147/SC 5/WG 5.

Kalyaanamoorthy, S., Minh, B. Q., Wong, T. K., Von Haeseler, A., and Jermiin, L. S. (2017). ModelFinder: fast model selection for accurate phylogenetic estimates. *Nat. Methods* 14 (6), 587–589. doi: 10.1038/nmeth.4285

Khvatkov, P., Chernobrovkina, M., Okuneva, A., and Dolgov, S. (2019). Creation of culture media for efficient duckweeds micropropagation (*Wolffia arrhiza* and *Lemna minor*) using artificial mathematical optimization models. *PCTOC* 136 (1), 85–100. doi: 10.1007/s11240-018-1494-6

Kittiwongwattana, C., and Vuttipongchaikij, S. (2013). Effects of nutrient media on vegetative growth of *Lemna minor* and *Landoltia punctata* during *in vitro* and ex vitro cultivation. *Maejo Int. J. Sci. Technol.* 7 (1), 60–69. doi: 10.14456/mijst.2013.5

Koca, Y. O., and Erekul, O. (2016). Changes of dry matter, biomass and relative growth rate with different phenological stages of corn. *Agric. Agric. Sci. Procedia.* 10, 67–75. doi: 10.1016/j.aaspro.2016.09.015

Landolt, E. (1986). *The Family of Lemnaceae - A Monographic Study. Vol. 1. Biosystematic Investigations in the Family of Duckweeds (Lemnaceae)*. Veröffentlichungen des Geobotanischen Institutes der Eidg. Technische Hochschule, Stiftung Rübel, Zurich.

Landolt, E. (1957). "Physiologische und Ökologische untersuchungen an lemnaceen," in *Berichte der schweizerischen botanischen gesellschaft*, (Bern, Switzerland: Druck und Kommissionsverlag von Büchler & Co.) 67, 271–410.

Landolt, E., and Urbanska-Worytkiewicz, K. (1980). "List of the studied lemnaceae samples: origin and chromosome numbers," in *Biosystematische untersuchungen in der familie der wasserlinsen (Lemnaceae)*. (Zürich, Switzerland: Veröff. Geobot. Inst. Rübel) 1, 205–247.

Lasfar, S., Monette, F., Millette, L., and Azzouz, A. (2007). Intrinsic growth rate: a new approach to evaluate the effects of temperature, photoperiod and phosphorus–nitrogen concentrations on duckweed growth under controlled eutrophication. *Water Res.* 41 (11), 2333–2340. doi: 10.1016/j.watres.2007.01.059

Lemnature AquaFarms - A Sustainable Aquaculture Company (2022). Available at: https://www.parabel.com/ (Accessed 06/05/2022).

Liu, Y., Chen, X., Wang, X., Fang, Y., Zhang, Y., Huang, M., et al. (2019). The influence of different plant hormones on biomass and starch accumulation of duckweed: A renewable feedstock for bioethanol production. *Renew. Energy.* 138, 659–665. doi: 10.1016/j.renene.2019.01.128

Mes, J. J., Esser, D., Oosterink, E., van den Dool, R. T., Engel, J., de Jong, G. A., et al. (2022a). A controlled human intervention trial to study protein quality by amino acid uptake kinetics with the novel *Lemna* protein concentrate as case study. *Int. J. Food. Sci. Nutr.* 73 (2), 251–262. doi: 10.1080/09637486.2021.1960958

Mes, J. J., Esser, D., Somhorst, D., Oosterink, E., van der Haar, S., Ummels, M., et al. (2022b). Daily intake of *Lemna minor* or spinach as vegetable does not show significant difference on health parameters and taste preference. *Plant Foods Hum. Nutr.* 77 (1), 121–127. doi: 10.1007/s11130-022-00952-9

Minh, B. Q., Schmidt, H. A., Chernomor, O., Schrempf, D., Woodhams, M. D., Von Haeseler, A., et al. (2020). IQ-TREE 2: new models and efficient methods for phylogenetic inference in the genomic era. *Mol. Biol. Evol.* 37 (5), 1530–1534. doi: 10.1093/molbev/msaa015

Müller, K. (2005). SeqState - primer design and sequence statistics for phylogenetic DNA data sets. *Appl. Bioinf.* 4 (1), 65–69. doi: 10.2165/00822942-200504010-00008

Muranaka, T., Okada, M., Yomo, J., Kubota, S., and Oyama, T. (2015). Characterization of circadian rhythms of various duckweeds. *Plant Biol.* 17, 66–74. doi: 10.1111/plb.12202

Murray, M. G., and Thompson, W. (1980). Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res.* 8 (19), 4321–4326. doi: 10.1093/nar/8.19.4321

Naumann, B., Eberius, M., and Appenroth, K. J. (2007). Growth rate based dose–response relationships and EC-values of ten heavy metals using the duckweed growth inhibition test (ISO 20079) with *Lemna minor* l. clone St. *J. Plant Physiol.* 164 (12), 1656–1664. doi: 10.1016/j.jplph.2006.10.011

Pagliuso, D., Grandis, A., Fortirer, J. S., Camargo, P., Floh, E. I. S., and Buckeridge, M. S. (2022). Duckweeds as promising food feedstocks globally. *Agron.* 12 (4), 796. doi: 10.3390/agronomy12040796

Pennock, D. (2019). *Soil erosion: The greatest challenge for sustainable soil management* (Food and Agriculture Organization of the United Nations (FAO).

Petersen, F., Demann, J., Restemeyer, D., Ulbrich, A., Olfs, H. W., Westendarp, H., et al. (2021). Influence of the nitrate-n to ammonium-n ratio on relative growth rate and crude protein content in the duckweeds *Lemna minor* and *Wolffiella hyalina*. *Plants* 10 (8), 1741. doi: 10.3390/plants10081741

Plants of the World Online (2022). Available at: https://powo.science.kew.org/ (Accessed 30/05/2022).

Poorter, H. (1989). Plant growth analysis: towards a synthesis of the classical and the functional approach. *Physiol. Plant* 75, 237–244. doi: 10.1111/j.1399-3054.1989.tb06175.x

Potter, J. R., and Jones, J. W. (1977). Leaf area partitioning as an important factor in growth. *Plant Physiol.* 59 (1), 10–14. doi: 10.1104/pp.59.1.10

Riedl, H. (1976). *Lemnaceae in K. H. Rechinger, Flora Iranica, no. 119. -Graz.*

Rozas, J., Ferrer-Mata, A., Sánchez-DelBarrio, J. C., Guirao-Rico, S., Librado, P., Ramos-Onsins, S. E., et al. (2017). DnaSP 6: DNA sequence polymorphism analysis of large data sets. *Mol. Biol. Evol.* 34 (12), 3299–3302. doi: 10.1093/molbev/msx248

Rubisco Foods (2022) *Next generation ingredients*. Available at: https://rubiscofoods.com/ (Accessed 06/05/2022).

Saran, P. L., Singh, S., Solanki, V., Choudhary, R., and Manivel, P. (2021). Evaluation of asparagus adscendens accessions for root yield and shatavarin IV content in India. *Turk J. Agric. For.* 45 (4), 475–483. doi: 10.3906/tar-2006-42

Schenk, R. U., and Hildebrandt, A. C. (1972). Medium and techniques for induction and growth of monocotyledonous and dicotyledonous plant cell cultures. *Can. J. Bot.* 50, 199–204. doi: 10.1139/b72-026

Shaw, J., Lickey, E. B., Beck, J. T., Farmer, S. B., Liu, W., Miller, J., et al. (2005). The tortoise and the hare II: relative utility of 21 noncoding chloroplast DNA sequences for phylogenetic analysis. *Am. J. Bot.* 92 (1), 142–166. doi: 10.3732/ajb.92.1.142

Sree, K. S., and Appenroth, K. J. (2014). Increase of starch accumulation in the duckweed *Lemna minor* under abiotic stress. *Albanian J. Agric. Sci.*, 13, 11–14.

Sree, K. S., Sudakaran, S., and Appenroth, K. J. (2015). How fast can angiosperms grow? species and clonal diversity of growth rates in the genus *Wolffia* (Lemnaceae). *Acta Physiol. Plant* 10 (37), 1–7. doi: 10.1007/s11738-015-1951-3

Tippery, N. P., Les, D. H., Appenroth, K. J., Sree, K. S., Crawford, D. J., and Bog, M. (2021). Lemnaceae and orontiaceae are phylogenetically and morphologically distinct from araceae. *Plants.* 10 (12), 2639. doi: 10.3390/plants10122639

Tomlinson, K. W., Poorter, L., Bongers, F., Borghetti, F., Jacobs, L., and van Langevelde, F. (2014). Relative growth rate variation of evergreen and deciduous savanna tree species is driven by different traits. *Ann. Bot.* 114 (2), 315–324. doi: 10.1093/aob/mcu107

Toyama, T., Kuroda, M., Ogata, Y., Hachiya, Y., Quach, A., Tokura, K., et al. (2017). Enhanced biomass production of duckweeds by inoculating a plant growth-promoting bacterium, acinetobacter calcoaceticus P23, in sterile medium and non-sterile environmental waters. *Water Sci. Technol.* 76 (6), 1418–1428. doi: 10.2166/wst.2017.296

Tursi, A. (2019). A review on biomass: importance, chemistry, classification, and conversion. *Biofuel Res. J.* 6 (2), 962–979. doi: 10.18331/BRJ2019.6.2.3

Walsh, É., Kuehnhold, H., O'Brien, S., Coughlan, N. E., and Jansen, M. A. (2021). Light intensity alters the phytoremediation potential of *Lemna minor*. environ. *Sci. pollut. Res.* 28 (13), 16394–16407. doi: 10.1007/s11356-020-11792-y

Wang, W., Wu, Y., Yan, Y., Ermakova, M., Kerstetter, R., and Messing, J. (2010). DNA Barcoding of the lemnaceae, a family of aquatic monocots. *BMC Plant Biol.* 10 (1), 1–11. doi: 10.1186/1471-2229-10-205

Xu, J., Cheng, J. J., and Stomp, A. M. (2012). Growing *Spirodela polyrrhiza* in swine wastewater for the production of animal feed and fuel ethanol: a pilot study. *Clean (Weinh).* 40 (7), 760–765. doi: 10.1002/clen.201100108

Xu, S., Stapley, J., Gablenz, S., Boyer, J., Appenroth, K. J., Sree, K. S., et al. (2019). Low genetic variation is associated with low mutation rate in the giant duckweed. *Nat. Commun.* 10 (1), 1–6. doi: 10.1038/s41467-019-09235-5

Ziegler, P., Adelmann, K., Zimmer, S., Schmidt, C., and Appenroth, K. J. (2015). Relative *in vitro* growth rates of duckweeds (Lemnaceae)–the most rapidly growing higher plants. *Plant Biol.* 17 (s1), 33–41. doi: 10.1111/plb.12184

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 04 January 2023
DOI 10.3389/fpls.2022.1099293

OPEN ACCESS

EDITED BY
Beckley Ikhajiagbe,
University of Benin, Nigeria

REVIEWED BY
Hengyou Zhang,
Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences (CAS), China
Wricha Tyagi,
Central Agricultural University, India

*CORRESPONDENCE
Luisa M. Trindade
luisa.trindade@wur.nl

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 15 November 2022
ACCEPTED 12 December 2022
PUBLISHED 04 January 2023

CITATION
Gulisano A, Lippolis A, van Loo EN, Paulo M-J and Trindade LM (2023) A genome wide association study to dissect the genetic architecture of agronomic traits in Andean lupin (*Lupinus mutabilis*).
Front. Plant Sci. 13:1099293.
doi: 10.3389/fpls.2022.1099293

A genome wide association study to dissect the genetic architecture of agronomic traits in Andean lupin (*Lupinus mutabilis*)

Agata Gulisano[1], Antonio Lippolis[1], Eibertus N. van Loo[1], Maria-João Paulo[2] and Luisa M. Trindade[1]*

[1]Wageningen University and Research Plant Breeding, Wageningen University, Wageningen, Netherlands, [2]Wageningen University and Research Biometris, Wageningen Research, Wageningen, Netherlands

Establishing *Lupinus mutabilis* as a protein and oil crop requires improved varieties adapted to EU climates. The genetic regulation of strategic breeding traits, including plant architecture, growing cycle length and yield, is unknown. This study aimed to identify associations between 16 669 single nucleotide polymorphisms (SNPs) and 9 agronomic traits on a panel of 223 *L. mutabilis* accessions, grown in four environments, by applying a genome wide association study (GWAS). Seven environment-specific QTLs linked to vegetative yield, plant height, pods number and flowering time, were identified as major effect QTLs, being able to capture 6 to 20% of the phenotypic variation observed in these traits. Furthermore, two QTLs across environments were identified for flowering time on chromosome 8. The genes *FAF*, *GAMYB* and *LNK*, regulating major pathways involved in flowering and growth habit, as well as *GA30X1*, *BIM1*, *Dr1*, *HDA15*, *HAT3*, interacting with these pathways in response to hormonal and environmental cues, were prosed as candidate genes. These results are pivotal to accelerate the development of *L. mutabilis* varieties adapted to European cropping conditions by using marker-assisted selection (MAS), as well as to provide a framework for further functional studies on plant development and phenology in this species.

KEYWORDS

***Lupinus mutabilis*, molecular markers, SNP, flowering time, association mapping, plant architecture**

1 Introduction

Lupinus mutabilis, also known as "Andean lupin", is an endemic legume of the Andean region of South America. Firstly domesticated in the north of Peru (Atchison et al., 2016), *L. mutabilis* has been traditionally cultivated in Ecuador, Peru and Bolivia for soil enrichment and as a food crop (Gross and von Baer, 1981). Similarly to other Andean grains, in the last 500 years its cultivation has been marginalized and neglected due to the introduction of western pulses characterized by higher productivity. The scarce dissemination of its nutraceutical properties, and the presence of alkaloids providing a bitter taste to the seeds, have also partly contributed to this negligence (Chirinos-Arias et al., 2015). Only in the past decades, the demand for alternative plant protein sources and the need to maximize agricultural land, by making use of marginal lands, has sparked renewed interest in this crop, not only in South America but also in Europe. *Lupinus mutabilis* seeds are characterized by a high content of protein and oil (44% dm and 18% dm, on average respectively), which exceeds that of any other lupin species and is comparable to soybean. On top of that, *L. mutabilis* is adapted to low input farming in temperate climates and can effectively contribute to the improvement of poor soils by fixing nitrogen and mobilizing soil phosphates. In Europe and other temperate climatic regions, the combination of these features makes *L. mutabilis* a potentially superior alternative to current plant-based sources of protein and oil. Despite this, some challenges need to be addressed to expand its cultivation on large scale, as *L. mutabilis* remains to date an under studied crop. Pivotal to achieving this aim are breeding programs focused on guaranteeing economic viability and consumer acceptance of the crop.

In the past decades, numerous studies have investigated the nutritional profile and potential applications of *L. mutabilis* seeds in a wide range of end-use products, from protein, oil and food additives to cosmetics, medicines and bio-pesticides. In contrast, few studies have addressed the agronomic aspect of its cultivation. Research on the adaptation of *L. mutabilis* to European soil and climate conditions started only 30 years ago, when the first European project focusing on 16 selected lines was initiated. More recently, a second European project (LIBBIO) has expanded this investigation to 225 *L. mutabilis* accessions, by evaluating a wide panel of Andean germplasm both as a winter crop in Mediterranean conditions and as a summer crop in North-Central European conditions. Both projects have pointed out the need of breeding for a better plant architecture, early maturity and yield stability (Caligari et al., 2000; Gulisano et al., 2022a). These traits are highly interconnected. Plant architecture has been identified as the main factor limiting yield in European environments (Caligari et al., 2000). This is to a great extent due to the indeterminate growth habit characterizing *L. mutabilis*, which leads to an overlap of vegetative and reproductive phases, hindering uniform maturation and delaying reproductive growth. The ability to prolong indefinitely vegetative growth after the onset of flowering, results in flowers and pods abortion in temperate climatic conditions during the long maturation period (Galek et al., 2007). Conversely, in Mediterranean environments, dry conditions at the end of the cropping season can drastically affect biomass yield and pods set, decreasing considerably seed yield (Hardy et al., 1997). Early flowering genotypes can contribute to shorten the growing season and to escape terminal drought and heat stress (Gulisano et al., 2022a), but only the combination of determinacy and earliness can ultimately lead to higher and more stable seed yield.

Similar challenges have characterized the domestication of other legumes, including "old world" lupin species. Indeterminate growth habit is typical of many wild relatives of grain legumes (e.g. pea, soybean, common bean) and its switch to determinate growth can be considered as one of the most important traits of their domestication (Krylova et al., 2020). The identification of genes responsible for growth habit has been pivotal for the development of determinate varieties, and has highlighted the interconnection of growth habit with stem length, flowering duration, yield, resistance to lodging and suitability to mechanized cultivation (Krylova et al., 2020). Mutation of *TERMINAL FLOWER 1* (*TFL1*)-like gene controlling transition to flowering has led to determinate growth habit in many legumes, including pea (Foucher et al., 2003), faba bean (Avila et al., 2006), common bean (Kwak et al., 2008) and soybean (Tian et al., 2010). *TFL1* belongs to the small gene family *CENTRORADIALIS/TERMINAL FLOWER1/SELF–PRUNING (CETS)*, controlling the developmental transition from indeterminate to determinate growth habit (Wickland and Hanzawa, 2015). Besides regulating flowering, genes of the *CETS* family are also involved in other processes, including stomatal opening or gibberellic and abscisic acid signaling pathways (Xi et al., 2010; Ando et al., 2013). Within the Lupin genus, determinate cultivars have been obtained in *L. albus, L. luteus*, and *L. angustifolius* through selection of spontaneous or induced mutants. The first determinate types in *L. mutabilis* have been obtained through induced mutation in Poland, distinguished by medium-tall stems without lateral branches, resistance to lodgings and early generative growth (Galek et al., 2007). Nevertheless, the molecular mechanisms underlying plant-architecture and other yield related traits remains still unknown in *L. mutabilis*, due to the lack of genetic and molecular studies on this crop.

Given the importance of these agronomic traits, a large number of genomic regions (Quantitative Trait Loci, QTLs) associated to plant-architecture related traits, flowering time and seed yield has been identified in the past decades for many grain legumes (Dargahi et al., 2014; Yao et al., 2015; Ávila et al., 2017; Klein et al., 2020). These studies have contributed to significantly increase knowledge on the genetic basis of these traits and to accelerate breeding of these crops. However, the majority of

them has focused on identifying QTLs in biparental populations, hence limiting genetic variation and mapping resolution (Gupta et al., 2005). With the current development of sequencing and genotyping technologies at affordable cost, genome wide association studies (GWAS) have rapidly become a more common and powerful tool to investigate natural variation and to identify genomic regions underlying important agronomic traits. Moreover, GWAS approach allows to exploit higher phenotypic diversity than biparental mapping populations derived from targeted crosses, as well as a direct application of the results from research to breeding.

In this study, we evaluated a panel of 223 diverse accessions of *L. mutabilis* in the native Andean region, and over two cropping conditions in Europe, in order to identify Single Nucleotide polymorphisms (SNPs) underlying the variation in plant-architecture, flowering and yield related traits in this species. A GWAS approach was used to capture the natural diversity present in the panel, with the objective to develop genetic markers and highlight genomic regions (QTLs) harboring causal candidate genes, which are critical for assisting and speeding up the breeding of *L. mutabilis*.

2 Material and methods

2.1 Plant materials and field trials

The GWAS panel used in this study comprised 223 *L. mutabilis* accessions, 201 provided by the Instituto Nacional de Investigaciones Agropecuarias of Quito (INIAP, Ecuador), comprising landraces, varieties and wild material collected across the Andean region, and 22 *L. mutabilis* lines developed in European breeding programs. The panel was evaluated in a total of four field trials in Ecuador and Europe during the growing seasons 2019 and 2020. The locations were chosen to represent an example of cultivation in the native environment as well as in Mediterranean (winter cycle) and North-Central European (summer cycle) climates and photoperiod regimes. The trial in Ecuador (EC, Cotopaxi 0° 55′ 35′′ S, 78° 40′ 07.4′′ W) was sown in December 2019 and harvested in June 2020. The field layout was an alpha-lattice design with 3 replicates. Plants were arranged in plots with 5 rows (80 cm between rows), containing 40 plants spaced 20 cm. In Europe, the field trials were set up during winter (Nov 2019 – May 2020) for the Portuguese site (PT, Lisbon 38° 42′ 33.5′′ N, 9° 11′ 0.5′′W), and during two consecutive growing seasons from April to October in two locations in The Netherlands (NL-Sc 2019, Scheemda 53° 09′ 60′′ N, 6° 57′ 59.9′′E; NL-Wi 2020, Winschoten 53° 10′ 11.′′ N, 7° 2′ 56.09′′E). A randomized complete block design (RCBD) with three replicates was adopted for the three EU trials. Out of 20 plants present in each plot, at a distance of 30x30 cm, phenotyping was conducted on the 6 central plants. In all the locations, plants were cultivated under rain-fed conditions and without the aid of any fertilization, following local cultivation practices.

2.2 Phenotyping of the GWAS panel

The phenotypic evaluation of the *L. mutabilis* GWAS panel across the four environments included the scoring of quantitative traits related to plant morphology, phenology and agronomic performance (Gulisano et al., 2022a). The phenotypes of interest were: flowering time, plant height, number of branching orders, vegetative yield, number of pods and seeds produced on the main stem, total number of pods and seeds produced on the overall plants, and 100 seeds weight. The traits were scored as described in our previous study (Gulisano et al., 2022a). Briefly, flowering time was scored as number of days from sowing until 50% of the plants in a plot had started flowering. Due to unforeseen circumstances, longer intervals in scoring of flowering in NL-Sc resulted in suboptimal phenotyping, as also shown by the low heritability estimate for flowering in this trial (H^2 = 0.24; Gulisano et al., 2022a), hence these data were excluded from this study. Instead, in Ecuador, due to the impossibility of collecting data during Covid restrictions, scoring of flowering time and vegetative yield was not possible. Phenotyping was conducted on the six central plant of the plots (five in Ecuador). At harvest, height of the main stem (cm) and number of branching orders were scored (0 = main stem only, 1 = main stem and first branching order, etc.). Number of pods on the main stem (Pods MS) and total number of pods (Pods T) was also recorded. After harvesting, seeds were air-dried and counted separately on the main stem (Seeds MS) and on the total plant (Seeds T). Vegetative yield (dw, g/plant) was estimated as the difference between the total amount of biomass harvested (dw) and the seed yield per plant.

2.3 Statistical analysis

Phenotypic data were analyzed using the SpATS mixed model approach, implemented in *statgenSTA* R package (v1.0.8), to correct for spatial gradients in the field by adopting a 2-dimensional smoothing with P-splines (Rodriguez-Alvarez et al., 2018). The analysis was conducted in a two-stages approach, where the final adjusted mean across trial was calculated after performing a single trial analysis. The best linear unbiased estimations (BLUEs) of genotypic means were obtained from the models and then used for the rest of the analyses. In addition, a random effects model was fitted using *lme4* R package to calculate variance components of genotype (G), environment (E) and genotype by environment interaction (GEI) and estimates of broad sense heritability across trials. Broad sense heritability was calculated across the three environments as in (Renaud et al., 2014):

$$H^2 = V_G / \left(V_G + \frac{V_{GEI}}{nE} + \frac{V_\epsilon}{nE * nBlock} \right)$$

where *VG*, *VGEI*, V_ϵ represent respectively the estimated genetic, GEI and error variance components, while *nE* represents the number of environments and *nBlock* the number of blocks in each environment. Pearson's correlations between BLUEs genotypic values in each trial were estimated and plotted using the R package *corrplot* (v0.92).

2.4 Genotyping and SNP development

Reduced representation sequencing and single nucleotide polymorphism (SNP) typing was performed as previously described in Gulisano et al. (2022a). Briefly, genomic DNA was isolated from young grinded *L. mutabilis* leaves (~20–400 mg, freeze dried material) using acetyl trimethyl ammonium bromide (CTAB) method (Doyle and Doyle, 1987) as described in Petit et al. (2020). To cover some degree of genetic heterogeneity expected in the accessions, DNA was extracted from a pool of 10 individuals/accession. Restriction site-associated DNA sequencing (RAD-seq) was used to identify SNPs distributed over the genome, by digesting 1 ug of high-quality genomic DNA (at a concentration ≥ 25 ng/µl) using the restriction enzyme EcoRI. RAD library preparation and sequencing were performed by Beijing Genomics Institute (BGI, Hong Kong). The Burrows–Wheeler Alignment Tool based on BWA- MEM algorithm were used to map the clean sequence reads to the *L. angustifolius* 'Tanjil' (LupAngTanjil_v1.0 refSeq GCF_001865875.1) genome reference (Hane et al., 2017). The average mapping rate was 76.9%, and the properly paired average 63.7%. BCFtools (v1.9) was used to call SNPs in each sample based on genotype likelihoods. At each locus, SNPs were called as percentage of the reference allele present on the total number of reads generated. After SNPs quality control and removal of SNPs not assigned to any chromosome, a total set of 16,781 biallelic SNPs was selected for the genetic analysis.

2.5 Genome-wide association studies

Single trait GWAS analysis was conducted separately for each environment, using 16,669 polymorphic markers after filtration to remove SNPs with minor allele frequencies (MAF) < 0.02. The GWAS model was based on a linear mixed model for association mapping as implemented in StatgenGWAS package v1.0.5 (van Rossum et al., 2020). Single trait GWAS in statgenGWAS follows the approach of Kang et al. (2008), by performing a two steps procedure. Firstly, an 'empty' model without any SNP effect is fitted in order to obtain REML-estimates of the genetic and residual variance components, computed using the Efficient Mixed Model Association (EMMA) algorithm (Kang et al., 2008). Secondly, the single SNP-effect of interest is tested by using generalized least-squares (GLS) and F-test, obtaining the effect-size and P-values for all SNPs. Population structure and individuals' relatedness were taken into account by fitting a Van Raden kinship matrix and adding origin of the accessions as corrections. The adequateness of genetic relatedness correction was assessed by evaluating genomic inflation factors. The Bonferroni correction was used to correct for multiple testing, thus obtaining a threshold of 5.52 for $-\log_{10}$ (p) that was used to state statistically significant SNPs. Manhattan plots were visualized using *qqman* package in R (Turner, 2018). Linkage disequilibrium for this panel was already estimated in Gulisano et al. (2022a), following the approach of (Vos et al., 2017). LD was estimated to decay around 80 kbp of distance.

2.6 Candidate genes identification

The size of the genomic regions investigated to identify putative candidate genes, controlling the traits under study, was defined by the extend of the average Linkage Disequilibrium across the genome. Starting from the position of the detected significant SNPs, candidate genes were proposed whether harbored in a maximum physical distance, upstream and downstream, of 80 kbp. All candidate genes were selected based on the information contained in the NCBI *Lupinus angustifolius* Annotation Release 100 for the genome assembly GCF_001865875.1 of LupAngTanjil_v1.0 (https://ftp.ncbi.nlm.nih.gov/genomes/all/annotation_releases/3871/100/). Special attention was given to genes with predicted functions related to the flowering time and regulation of plant growth and development, based on gene description and relevant research paper.

3 Results

3.1 Phenotypic analysis

Nine traits relevant for breeding of *L. mutabilis* accessions adapted to different cropping conditions in Europe and Ecuador were investigated using a panel of 223 accessions. Extensive phenotypic variations were observed for all traits, namely flowering time, plant height, number of branching orders, vegetative yield, number of pods and seeds on the main stem, total number of pods and seeds and seed weight (depicted in Figure 1). Remarkable variation in flowering time was observed between the two European cropping conditions. In particular, flowering occurred at an earlier time in the Netherland than Portugal, with an average of 81 and 110 days for NL-Wi and PT respectively. Relevant phenotypic diversity in time to flowering was observed within the panel, as highlighted by coefficients of

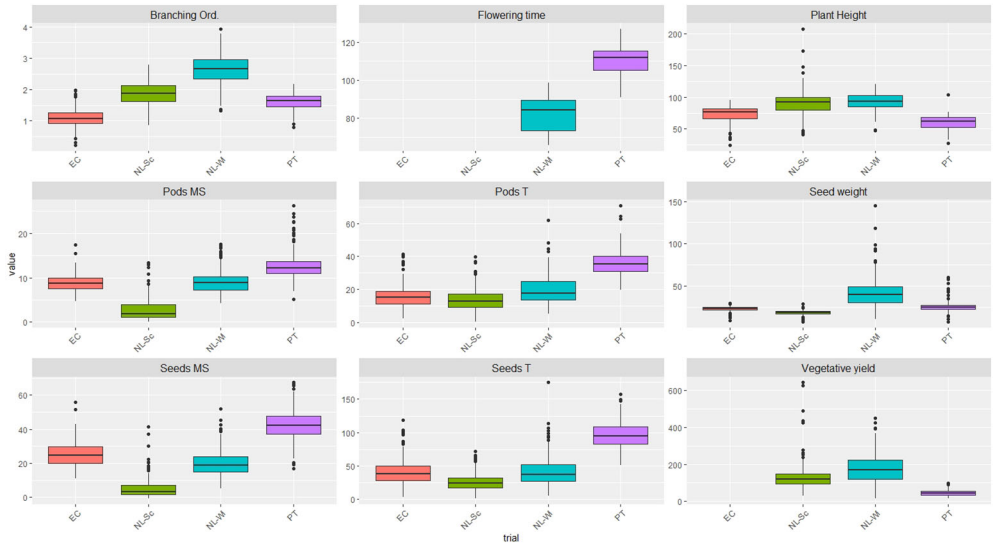


FIGURE 1
Boxplots showing the variation of the plant architecture and yield related traits (BLUES data) under study in a collection of 223 *L. mutabilis* accessions. The upper and lower end of the boxes indicate respectively the 75th and 25th percentiles, while the line in the middle represents the 50th percentile (median). The whiskers represent the highest and lowest values, while circles represent outliers outside the 5th-95th percentile interval.

variation (CV %) of 6% in Portugal and 8% in the Netherlands 2020. The overall plant height across different locations ranged from 24 cm to and 248 cm, indicating that *L. mutabilis* small genotypes can be 10 times smaller than tall genotypes as combination of environmental and genetic effects. As general trend, plants were shorter on average in PT (60 cm) and EC (73 cm) compared to NL-Sc (90 cm) and NL-Wi (92 cm) and characterized on average by a first branching order only. Contrarily, in NL-Sc and NL-Wi the main stem reached an average height of 90 and 92 cm respectively, with two/three branching orders produced. Plant height had high phenotypic variability within the 223 accessions, as pointed by the CV values of 17%, 18%, 20%, 13% in Ecuador, Portugal, NL-Sc and NL-Wi respectively. Vegetative yield reflected the development of *L. mutabilis* in different environments. In Portugal, the average vegetative yield recorded (45.8 g/plant) was between three and four times lower than the average yield of NL-Sc (133.6 g/plant) and NL-Wi (174.9 g/plant). Concerning grain yield components, the average number of pods and seeds produced by the total plant, as well as by the main stem only, was higher in winter-Mediterranean conditions, followed by the native environment (EC) and summer-North European environments. The number of seeds harvested per plant was on average 42 in EC (CV 47%), 95 in PT (CV 19%), 26 in NL-Sc (CV 49%) and 42 (CV 53%) in NL-Wi. The number of pods and seeds produced on the main stem were respectively 9 (CV 20%) and 25 (CV 28%) in EC, 13 (CV 23%) and 42 (CV 21%) in PT, 3 (CV 89%) and 14 (CV 47%) in NL-Sc and 9 (CV 27%) and 20 (CV 37%) in NL-Wi. Conversely, the yield of biomass agricultural residues was lowest in PT (45.8 g/plant) and highest in NL-Wi (174.9 g/plant). An exhaustive analysis of the phenotypic traits under study was conducted in a previous study of diversity and agronomic adaptation of this collection. There, the mean values, range of variation and Pearson's correlation between plant architecture and yield-related traits are reported (Gulisano et al., 2022a).

3.2 Heritability estimates and identification of relevant breeding traits

In the four separate locations, estimates of heritability were high for flowering time (0.87-0.93) and plant height (0.5-0.8). Moderate to high heritability was also estimated for number of pods on the main stem (0.4-0.8), number of seeds on the main stem (0.4-0.8), total number of pods (0.2-0.6), total number of seeds (0.2- 0.7), 100 seeds weight (0.1-0.67), and moderate to low value for vegetative yield (0.1-0.5) and branching order (0.1-0.4). When looking at the data across locations (averaged data), the analysis of variance indicated that genotype (G) and environment (E) had a significant effect on all the traits analyzed ($P< 0.001$). The effect of genotype by environment interaction (GEI) was also highly significant ($P< 0.001$) for the majority of the traits but had no significant effect on vegetative yield. As shown in Table 1, heritability values across locations were higher for plant height (0.82) and flowering time (0.69), two traits well known for being under the influence of a high quantitative genetic component in many crops, as well as for the

TABLE 1 Estimates of heritability (H_2) of breeding traits in single locations (from Gulisano et al., 2022a) and across trials.

	EC	PT	NL-Sc	NL-Wi	Across trials					
	H_2	H_2	H_2	H_2	G	E	Block	GEI	ε	H_2
Plant height	0.72	0.86	0.56	0.57	16.4	45.1	1.9	3.8	32.9	0.82
Flowering time	–	0.93		0.87	10.8	86.6	0.2	1.4	0.9	0.88
Branching orders	0.33	0.44	0.46	0.07	0.4	44.5	2.7	0.4	42.9	0.06
Vegetative yield	–	0.54	0.17	0.21	3.5	27.1	5.0	1.5	62.9	0.39
Pods on main stem	0.48	0.42	0.83	0.46	6.2	57.0	0.8	6.6	29.5	0.60
Seeds on main stem	0.46	0.57	0.83	0.53	4.7	70.2	0.3	4.7	20.4	0.63
Pods Total	0.23	0.41	0.69	0.58	1.4	50.0	1.1	9.6	38.0	0.20
Seeds Total	0.21	0.36	0.70	0.64	2.3	58.2	0.9	6.1	32.5	0.35
100 Seed weight	0.67	0.17	0.63	0.24	2.7	39.5	0.2	3.5	54.1	0.33

For the analysis of variance across trials, the percentage of variance explained by the different components of variance is reported. Total Variance was decomposed in: genotype (G), environment (E), block effects within trial (Block), interaction of genotype by environment (GEI) and Residuals (ε) effects, reported as percentage of the total variance. Traits that were not measured are indicated with a dash. Crossed out traits were discarded for this study.

number of pods and seeds produced on the main stem (0.60-0.63). Contrarily, lower heritability values were estimated for the total number of pods and seeds produced (0.20-0.35), and almost null for the number of branching orders (0.06). Overall, variance component estimates showed a preponderant effect of the environment (Table 1). For this reason, the genetic analysis of these traits should be conducted on the singular specific environments separately.

3.3 Genome wide association studies

The GWAS analysis was performed using the R package StatGenGWAS. The Kinship relationship matrix among samples, calculated with the Van Raden method, was included together with the country of origin of the accessions as covariate, to adjust for population structure. The quantile distribution of the observed p-values versus the expected p-values (QQ plot) confirmed adequacy of the population structure correction.

Due to the large effect of genotype by environment interaction on the phenotypic variation, GWAS was performed analyzing each trial separately, as well as across all the locations. By applying the Bonferroni threshold ($-\log_{10}(p) > 5.52$), a total of 5 significant SNPs were identified as associated with flowering time, 1 SNP was associated with plant height, 1 SNP with vegetative yield and 4 SNPs with the number of pods on the main stem (Pods MS) (reported in Figure 2 and Table 2). Contrarily, no significant SNPs were found in association with the remaining phenotypic traits. Single location GWAS for flowering time revealed the association of the SNP M11043, located on chromosome 13, with flowering time in the Portuguese environment. M11034 explained ~6% of the phenotypic variance observed. In NL-Wi, SNP M7399 on chromosome 8 was found to explain almost 14% of variation in time to flowering. When the two trials were analyzed together, both M11034 and M7399 remained significantly associated to flowering time explaining respectively 5.4 and 13.2% of variation across trials. Additionally, M7412, M7413 and M7670 on chromosome 8 were also detected as significant across trials, explaining between 7 and 14% of variation in flowering time. Notably, it was possible to observe a similar piling up of SNPs on chromosome 8 also in PT, harboring some of the same markers detected across trials in association with flowering time (M7670 and M7399), but below the stringent Bonferroni detection threshold (Figure 3). Instead, the SNP M11034 was also associated with the number of Pods MS in the trial located in the Netherlands in 2019 (NL-Sc). For Pods MS, one more SNP was detected on chromosome 8 across trials (M7272), while two more SNPs were detected on single trials respectively on chromosome 18 (M14832) in Ecuador and chromosome 17 (M14675) in NL-Sc. About 10% of the plant height variation in Ecuador was ascribable to the SNPs M396 on chromosome 1, while no other SNPs associated with this trait passed the Bonferroni threshold in the other locations, despite the high heritability and genetic variation recorded. However, the visual inspection of Manhattan plots revealed the presence of different regions likely associated with plant height in the other trials (Figure 4). The importance of these SNPs needs to be further investigated considering that Bonferroni is a quite conservative threshold. As an example, an interesting piling up of SNPs right below the threshold was observed on chromosome 16 for PT, while another piling up in association with this trait was observed on chromosome 12 for both NL-Sc and NL-Wi. Finally, 1 SNP associated to vegetative yield was detected on

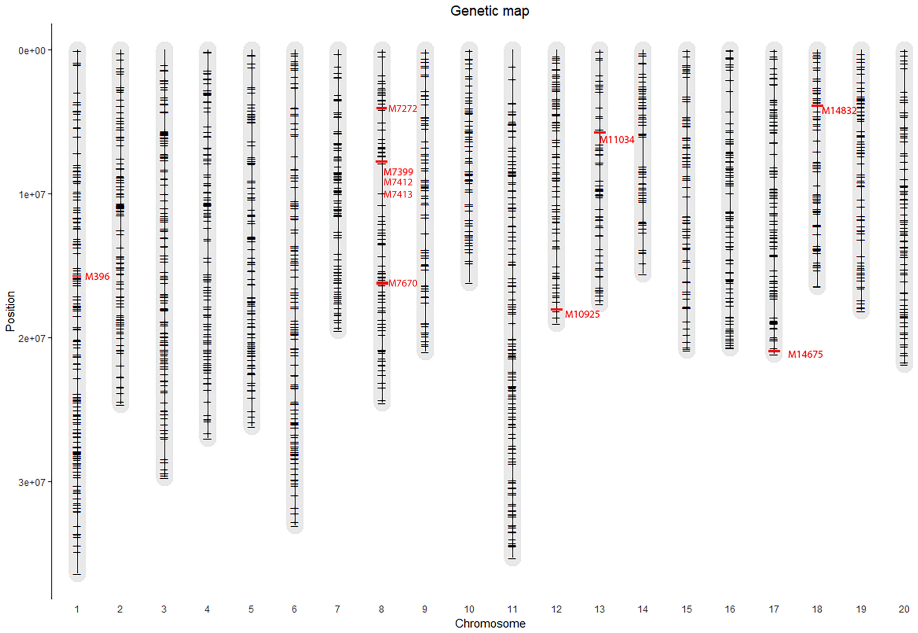


FIGURE 2
Markers distribution on the 20 chromosomes of *L. angustifolius* genome. In red, significant markers detected in association with agronomic traits using Bonferroni threshold (LOD 5.52, experiment *wise error* $\alpha = 0.05$).

TABLE 2 SNPs found in significant association with plant-architecture and yield related traits for 223 accessions of *L. mutabilis* grown in four different environments.

Traits	Markers	Chr	Position	Environment	LOD	% Var	Effect	N. of genes
Flowering time (days)	**M11034**	LG13	5,615,305	PT	6.39	6.35	15.49	14
	M7399	LG08	7,668,768	NL-Wi	6.69	13.9	9.44	20
	M11034	LG13	5,615,305	All Trials	6.65	5.48	16.12	14
	M7399	LG08	7,668,768	All Trials	7.67	13.2	7.65	20
	M7412	LG08	7,670,152	All Trials	5.74	7.25	14.29	20
	M7413	LG08	7,680,541	All Trials	6.14	8.16	7.72	18
	M7670	LG08	16,132,804	All Trials	5.95	14.33	-21.86	1
Plant height (cm)	M396	LG01	15,690,000	EC	5.99	10.3	14.24	2
Vegetative yield (g)	M10925	LG12	18,150,269	NL-Sc	5.95	19.3	-168.78	29
Pods on the main stem (number)	**M11034**	LG13	5,615,305	NL-Sc	5.99	5.93	-5.67	14
	M14832	LG18	3,860,231	EC	5.62	10.7	9.02	5
	M14675	LG17	20,841,361	NL-Wi	6.03	7.00	4.03	16
	M7272	LG08	4,062,540	All Trials	5.65	4.75	12.41	5

SNPs that are detected in more than one environment, or in association with more than one trait, are reported in bold. For each marker we report: the position on *L. angustifolius* chromosome (Chr), the trial where the association was detected as significant (Environment), the LOD value of association, the phenotypic variance explained (% Var), the allelic effect on the phenotypic mean of the trait (Effect) and the number of genes found in a window of ± 80 kbp from the marker.

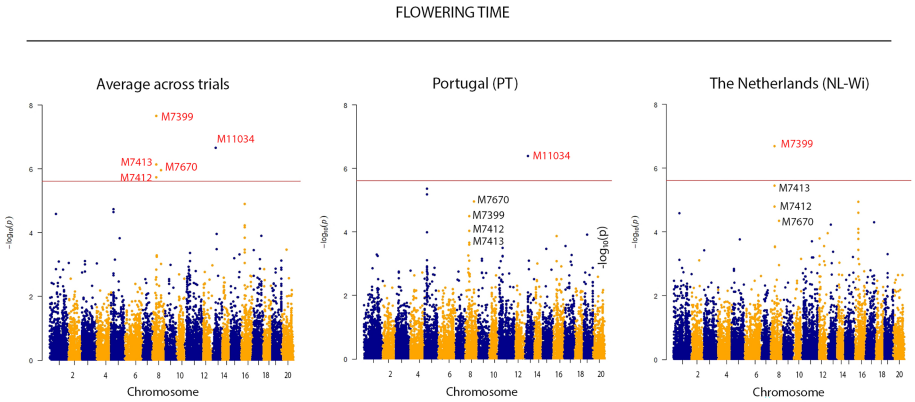


FIGURE 3
Manhattan plots displaying SNP markers-trait associations identified for Flowering time in GWAS using 223 accessions of *L. mutabilis*. The red line indicates the Bonferroni threshold (LOD = 5.52). Common SNPs showed significant association with variation in flowering time across locations, even if not all of them were detected as significant.

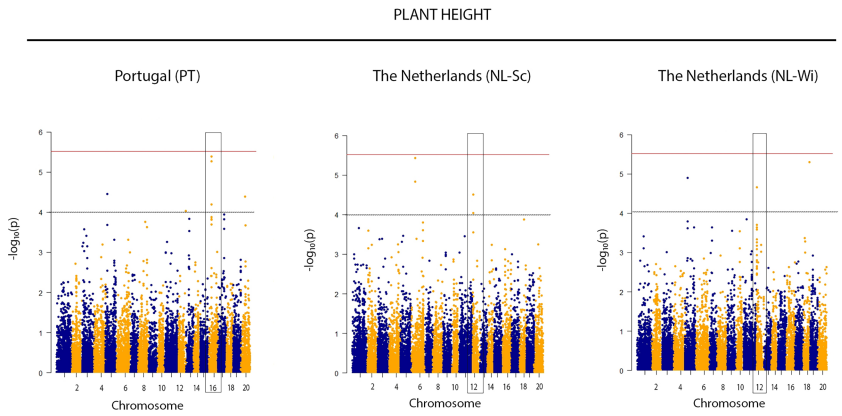


FIGURE 4
Manhattan plots suggestive of a SNP markers-trait association for plant height on chromosome LG16 and LG12, respectively for *L. mutabilis* growing in Mediterranean (Portugal) and North-Central European cropping conditions *(NL-Sc and NL-Wi)*. The red line indicates the Bonferroni threshold (LOD = 5.52). The black dotted line indicates a suggestive threshold of LOD= 4. The rectangles highlight interesting piling up of SNPs below the threshold of significance, harboring interesting candidate genes related to plant height.

chromosome 12 (M10925), but only in NL-Sc. The SNPs detected as significant are reported in Table 2.

3.4 Putative candidate genes

We searched for putative candidate genes in regions close to the peak of the eight QTLs detected. An interval of ± 80 kbp around the peak of these QTLs was taken as search window for candidate genes as this is about the size of the linkage disequilibrium blocks, hence genes outside this window are not likely to be causative for the QTL effect. All search regions together harbored a total of 91 unique genes on the *L. angustifolius* genome. Moreover, the investigation of other interesting regions previously mentioned as likely associated with plant height, which did not pass the stringent Bonferroni correction, led to the identification of 37 additional genes, mostly related to plant development. The investigation of these genes, based on functional annotation of their orthologs on *Arabidopsis* genome and relevant literature, resulted in a subset of 10 genes as possibly involved in the control of traits of interest. These genes, listed in Table 3, are known to play a role in regulation of

TABLE 3 Functional annotation of ortholog genes proposed as potential candidate genes for plant architecture and yield-related traits in *L. mutabilis*.

Trait	Chr	Candidate genes	Function annotation/common name
Plant height (PT)	LG16*	LOC109329023	transcription factor RF2b-like
Plant height (NL)	LG12*	LOC109362850	homeobox-leucine zipper protein HAT3
Vegetative yield (NL-Sc)	LG12	LOC109361654	gibberellin 3-beta-dioxygenase 1/GA3OX1
	LG12	LOC109362161	histone deacetylase 15/HDA15
	LG12	LOC109362682	transcription factor BIM1
	LG12	LOC109362691	protein Dr1 homolog
Flowering	LG08	LOC109354086	protein FANTASTIC FOUR 3
Flowering (PT)/ Pods main stem (NL-Sc)	LG13	LOC109325309	protein LNK3
	LG13	LOC109363186	transcription factor GAMYB
Pods main stem (NL-Wi)	LG17	LOC109330763	E3 ubiquitin-protein ligase RING1

*Putative candidate genes found in linkage disequilibrium with SNPs detected below the Bonferroni threshold (LOD ≥ 4).

flowering and are involved in different plant growth and development processes. For instance, the identified genes encoding the protein FANTASTIC FOUR 3 FAF3, and the transcription factors LNK3 and GAMYB, are known to contribute to the genetic control of flowering time in Arabidopsis and other species (Wahl et al., 2010). *FAF3* is located on chromosome 8 at a distance of 77 kbp from the SNP M7399 associated with flowering, while *LNK3* and *GAMYB* are located on chromosome 13 (respectively, 62 kbp upstream and 44 kbp downstream of SNP M11034). A gene encoding an E3-ubiquitin protein ligase RING1 was detected on chromosome 17 in linkage with the SNPs M14675, that showed significant association with Pods MS. Members of the E3 ubiquitin ligases are well known to play a role in plant growth and development. Additionally, important genes involved in plant growth were identified on chromosome 12 and 16. Genetic polymorphism in these genes may be associated with plant height and vegetative yield and captured by blocks of closely located markers: M12928 to M12945 on chromosome 16 and M10488, M10515 and M10527 on chromosome 12 (Figure 4). These included genes encoding transcription factors involved in vascular development (*RF2b*), shade-induced plant growth (*HAT3*), transcriptional regulators of flowering time (*Dr1, BIM1*) and genes involved in the gibberellin signaling pathway (*GA3OX1*).

4 Discussion

Large phenotypic variation for plant height, flowering time, biomass and yield related traits was observed among the 223 *L. mutabilis* accessions studied, indicating that the used panel is suitable for a GWAS study that requires genetic variation of the traits. The large phenotypic variation observed in our population is beneficial to select breeding candidates in order to initiate a genetic improvement program aimed at developing superior varieties adapted to Europe and high performing. One of the key factors to promote the cultivation of high yielding *L. mutabilis* in Europe is breeding towards earliness and semi-determinate growth habit (Caligari et al., 2000; Gulisano et al., 2022a). In our extensive phenotypic evaluation of this collection, we highlighted the presence of strong correlations between plant height, flowering time and seed yield (Gulisano et al., 2022a). These findings are in line with previous observations carried out in other legumes (González et al., 2016; Annicchiarico et al., 2021). Breeding for plant architecture should be tailored to specific growing conditions, considering that environmental conditions highly impact the morphology and yield of *L. mutabilis* (Gulisano et al., 2022a). Generally, restricted plant architecture, caused by a limited number of branching orders, is associated with higher seed yield in *L. mutabilis*. In fact, indeterminate growth and extensive branching have a detrimental effect on seed yield, since Andean lupin tends to produce the largest share of seeds on the main stem and on the first branching order (Sinjushin, 2015; Gulisano et al., 2022a). However, from our study emerges that total seed production positively correlates with late flowering and higher vegetative biomass when Andean Lupin is cultivated in Mediterranean cropping conditions (Portugal). This observation supports the hypothesis that plants with indeterminate growth, thus larger branching orders, outperform determinate varieties in adverse climates. One possible explanation is that, under high temperature or drought, the indeterminate growth habit may better compensates yield loss, which is more evident in determinate types, where pods set is stronger negatively affected by the environmental factors (Martins et al., 2016; Gulisano et al., 2022a). However, it is important to report that

extensive branching in Mediterranean conditions is already constrained by environmental factors such as water scarcity, hence even plants characterized by an indeterminate growth habit will not produce more than two branching orders, maintaining overall a semi-determinate growth. Yet, the effect that a larger contribution of higher branching orders to seed yield can have on seed quality and composition remains to be further investigated.

Given the importance of breeding for earliness and growth habit in this crop, the collection under study showed enough phenotypic and genetic diversity in flowering time (about one month difference between early and late flowering genotypes), and vegetative yield, to potentially breed for novel early and high yielding lines. The highlighted variation is therefore suitable to support breeding aimed at the introduction of *L. mutabilis* as a protein and oil crop in Europe. For genetic improvement purposes, understanding the genetic architecture of earliness, plants architecture and seed-yield related traits is pivotal. By evaluating strategic breeding parameters including heritability and genotype-by-environment interactions (GxE), we demonstrated that some phenotypic traits of interest are highly influenced by their genetic components. We report high heritability values for flowering (H^2 = 0.88), plant height (H^2 = 0.82) and production of pods and seeds on the main stem (H^2 = 0.6) across the different trials, thus across environments and growing conditions (Table 1). Our results are in line with the heritability values reported for plant height, productivity on the main stem and flowering time by previous authors (Hardy et al., 1998; Guilengue et al., 2020). These findings encourage breeding towards early varieties with restricted plant architecture and improved yield in Europe, since high heritability is always advantageous to increase the expected genetic gain in crop selection.

To the best of our knowledge, this is the first genome-wide association study (GWAS) investigating the genetic architecture of agronomic traits in a collection of *L. mutabilis*. Since genotype by environment interaction can alter the QTLs effects, an initial GWAS analysis carried out on single trial data, followed by a genome-wide scan across all locations, was considered the best approach to analyze our data. We have already shown in previous studies that in the absence of a reference genome for *L. mutabilis*, the use of *L. angustifolius* pseudochromosomes assembly can be suitable for genome-wide associations in this species (Gulisano et al., 2022b). In addition, our GWAS approach also shows that, in the absence of pure highly stable homozygous lines, a mapping experiment can be conducted using allele frequencies as marker score, estimated at each marker locus in pooled DNA samples. We estimated allele frequencies by counting the number of reads carrying the reference alleles on the total number of generated reads at a given variant position. In order to ensure a more accurate estimation of allele frequencies, as well as accounting for the presence of alleles of different individuals in the same DNA sample, we have discarded SNPs with a reads depth<45x from the final markers set. This approach of treating markers as a variable with a continuous distribution was already adopted by Petit et al. (2020), and is a good option for mapping studies in crops where a certain degree of cross-pollination is expected within the accessions. The Efficient Mixed Model Association (EMMA) approach showed robust performance on the traits and in the population studied. In GWAS, false discoveries are a major concern and spurious associations between phenotypes and testing markers can arise as a consequence of population structure and differences in relatedness among individuals in the tested population (Sul et al., 2018). The association model we used incorporated kinship and the origin of accessions as covariates to control for spurious association and false discoveries. The QQ-plots generated after the genome scan confirmed the high quality of the GWAS models. The difference between observed and expected –log10(p) values, in fact, showed no inflation and values of inflation factors close to the ideal value of 1 for all the traits.

GWAS analysis identified a total of 5 significative SNPs associated with flowering time (M11034, chr 13, position 5,615,305; M7399, chr 8, position 7,668,768; M7412 chr 8, position 7,670,152; M7413, chr 8, position 7,680,541; M7670, chr 8, position 16,132,804), 1 SNP associated with plant height (M369, chr 1, position 15,690,000), 1 SNP associated with vegetative yield (M10925, chr 12, position 18,150,269), and 3 SNPs associated with number of pods on the main stem (M1134, chr 13, position 5,615,305; M14832 on chr 18, position 3,860,231; M14675, chr 17, position 20,841,361). In agreement with the idea that the genetic control of agronomic traits might differ in response to different environmental conditions, our study points out that the association of genomic regions with phenotypic traits can be location specific, thus different QTLs can be detected in different environments. Indeed, most of the SNPs reported were identified as single location markers. Taking into account that the Bonferroni correction method is a very conservative statistics to set a threshold *P*-value (Kaler and Purcell, 2019), we have also investigated some genomic regions where SNPs were piling up in the QQ plots, even if they were not passing the stringent Bonferroni threshold. We therefore report on likely association on chromosome 12 and chromosome 16 for plant height in the Netherlands (NL-Sc and NL-Wi) and in Portugal (PT), respectively.

Flowering time is a quantitative traits and its genetic control has been investigated in several crops, including *L. albus*, where it has been described as quantitative and under the control of major QTLs and numerous regulatory genes, including *FLOWERING LOCUS T (Fta1), CONSTANS (CO), FY, MOTHER OF FT AND TERMINAL FLOWERING LOCUS 1 (TFL1)*, but also of genes related to the response to photoperiod and vernalization, as well as regions involved in hormone signaling pathways (i.e, gibberellin) (Rychel et al., 2019). In our panel, we did not detect these major candidate genes conserved in several crops. However, the

variation in flowering time in our panel resulted associated with two large QTL on chromosome 8 (one identified by the SNPs M7399, M7412 and M7413 and one indicated by the SNP M7670) with a QTL on chromosome 13 (M11043). In addition, single trials analysis revealed that the effect of the QTL on chromosome 8 is more relevant in NL, while the effect of the QTL on chromosome 13 is more relevant in the Mediterranean environment. Given that M7399 and M7670 respectively accounted for ~11% and 16% of the phenotypic variation observed across trials, it is possible to hypothesize the presence of major QTLs in their surrounding genomic regions. Our investigation highlighted the presence of the gene *FANTASTIC FOUR 3* (*FAF3*) at a distance between 65 kb and 77 kb from M7413, M7412 and M7399, thus located in a linkage disequilibrium block captured by these three proximal markers. *FAF3* is a member of the FANTASTIC FOUR (FAF) genes family, which contains four members in *Arabidopsis* (*FAF1-FAF4*), dynamically expressed throughout development (Wahl et al., 2010), and mainly involved in the modulation of the shoot meristem size. Wahl et al., 2010 reported *FAF* genes as redundant in function, indicating that all the four proteins may perform a very similar role. Interestingly, the *FAF* genes are expressed in flowering buds and inflorescence, with an increased expression during the transition to flowering. Moreover, a regulatory interaction has been reported between *FAF2* and *FAF4* with the WUSCHEL (WU)-CLAVATA (CLV3) signaling pathway that plays a central role in regulating stem cell proliferation and differentiation in crop plants, supporting organogenesis in the floral meristem through a complex molecular pathway (Wahl et al., 2010). Notably, studies on *Arabidopsis thaliana* report *FAF1/FAF2* and *FAF3/FAF4* to be recently duplicated paralogs (Blanc and Wolfe, 2004) that could therefore be characterized by similar structure and functions. Unfortunately, it was not possible to detect any annotated and characterized gene in the genomic region surrounding M7670, hence further studies are needed to elucidate which gene is responsible for the variation in flowering time associated with this SNP. In addition to these markers detected on chromosome 8, we report the finding of SNP M11043 on chromosome 13 as associated to flowering time across location and in Mediterranean growing conditions specifically. M11043 is located in a genomic area that harbors two genes potentially involved in the regulation of flowering. The gene *GAMYB* is located at a distance of 44.2 kbp from the detected marker, while the gene *NIGHTLIGHT–INDUCIBLE AND CLOCK-REGULATED 3 (LNK3)* at a distance of 62.3 kbp. *GAMYB* is described as a regulator of flower induction, *via* the transcriptional activation of the LEAFY (*LFY*) gene (Gocal et al., 1999; Zhang et al., 2020), that results in early flowering both in dicots and monocot (Weigel and Nilsson, 1995; He et al., 2000). Instead, members of the family of *NIGHTLIGHT–INDUCIBLE AND CLOCK-REGULATED* (*LNK1* and *LNK2*) genes are known to control photomorphogenic and photoperiodic responses, as well as circadian rhythms (Rugnone et al., 2013). Recently, Li et al., 2021 obtained an early flowering line in soybean by targeted mutagenesis of the four *LNK2* genes, that molecularly interact with major flowering genes, showing that targeted breeding on these genes can contribute to soybean expansion to high latitudes in Europe where shorter growing cycles are needed (Li et al., 2021). Despite we report on *LNK3*, that was not previously described as involved in flowering, it will be beneficial to further investigate its role in *L. mutabilis*, based on the strong significant association detected, and the fact that the polymorphism detected explains ~6% of the flowering variation recorded.

The genetic analysis carried out in this study showed that plant height and vegetative yield are likely affected by QTL-by-environment interaction, since the detected SNPs and the respective candidate QTLs in linkage with them appeared to be environment-specific. Hence, different transcriptional regulators can similarly affect the same traits in response to different environmental stimuli. In particular, plant height showed association with the SNP M396 (chr 1, position 15,690,000) in Ecuador, but a piling up of SNP (not significant based on Bonferroni) was instead observed on chromosome 16 in Portugal, and on chromosome 12 in the Netherlands. On chromosome 1 we did not report any relevant candidate genes linked to M396 and plant height. Contrarily, we report the finding of a transcription factor *RF2b-like* and a homeobox-leucine zipper protein *HAT3*, both harbored in the genomic regions where we visualize an interesting piling up of SNPs, that likely resulted as false negative when applying Bonferroni threshold to detect significant associations. These genes are respectively involved in vascular development and shoot tissue organization (*RF2b)* and in shade-induced plant growth (*HAT3)*, and can both lead, through different mechanisms, to the development of semi-dwarf varieties (Dai et al., 2004; Bou-Torrent et al., 2012).

The association between genetic variants and vegetative yield was only significant in North-Central European summer conditions, where the SNP M10925 (chromosome 12, position 18,150,269) was detected. M10925 showed a strong association, and it is able to explain a large share (~20%) of the phenotypic variation. This suggest the presence of a major QTL, considering that vegetative yield is a quantitative and complex trait. Harbored in this area, we identified 4 different candidate genes: *GIBERELLIN 3-BETA-DIOXYGENASE1 (GA30X1)*, *HISTONE-DEACETYLASE15 (HDA15)*, transcription factor *BIM1* and *D*r1 homolog, which are gene involved in the gibberellin pathway (*GA30X1*) and regulation of flowering time (*BIM, Dr1*). On one hand, *GA30X1* is involved in the gibberellin pathways as the enzyme catalyzing the final step of the synthesis of bioactive gibberellin during vegetative growth,

and its negative regulation might directly lead to substantial decreases in biomass production and to the development of semidwarf types (Hu et al., 2008). On the other hand, transcriptional regulators of flowering time can promote flowering in response to hormonal shifts or environmental stimuli and, as a consequence of the strong interplay between flowering time and growth habit in legumes, indirectly act on biomass production and vegetative yield. For example, the transcription factor *BIM1*, part of the bHLH family of transcription factors, can lead to the inhibition of floral transition in Arabidopsis by transduction of hormonal brassinosteroid signals to the activation of *FLOWERING LOCUS C* (*FLC*) and consequent floral repression, promoting in turn vegetative growth. Conversely, defects in *BIM1* can lead to early floral transition (Li et al., 2018). Instead, *Dr1* and *HDA15* play important roles in determining plant response to different environmental stimuli, such as drought and elevated ambient temperature (Zotova et al., 2019; Shen et al., 2019). For example, it is suggested that Dr1 protein can bind and deactivate genes sensing environmental cues for drought, releasing in turn the activity of vernalization and flowering regulators (*Vrn1* ad *FT1*). This mechanism has been described as a successful drought escape strategy in wheat (Zotova et al., 2019), and could be linked in *L. mutabilis* to other successful escape strategies like accelerated seed maturation and pod filling, which rely on a preferential partitioning of nutrients towards reproductive growth, at the expense of vegetative growth.

Finally, four QTLs were also detected in association with number of pods produced on the main stem (Pods MS). One SNP was detected across trials (M7272, chr 8, position 4,062,540), while the remaining three SNPs indicated QTLs location-specific. M11034 (chromosome 13, position 5,615,305) was detected in the NL-Sc, M1483 (chromosome 18, position 3,860,231) in EC and M14675 (chromosome 17, position 20,841,361) in NL-Wi. On chromosome 17, we highlighted the presence of the E3 ubiquitin ligases *RING1* gene in linkage disequilibrium with the respective significant marker. E3 ubiquitin ligases have been documented to play an important role in the regulation of plant growth and development, such as seed dormancy and germination, root growth and flowering time control, as well as regulation of several abiotic stress responses (Zhang et al., 2015; Shu and Wenyu, 2017). Increasing numbers of studies have documented the key role of the different types of E3s, including E3-RING as involved in seed biology, root elongation, flowering time control, light response, ABA signaling transduction and response to drought and salinity stresses.

L. mutabilis breeding in Europe can beneficiate from the findings of this study as alleles for the QTLs can be typed and potentially used to assist the genetic selection. Furthermore, the highlighted association of genomic regions with the reported phenotypic traits represents a molecular framework for further deeper investigations of the intricate networks regulating plant development and phenology in this species.

Data availability statement

The genotypic data used for the genome wide association study and generated by genotyping by restriction site-associated DNA sequencing (RAD-seq) are openly available at the 4TU. Research Data repository (http://doi.org/10.4121/21334146.v1).

Author contributions

AG and LT conceived and designed the experiment. AG performed all the analysis, interpreted the data and wrote the manuscript. AL contributed to the interpretation of the data and the writing of the manuscript. LT wrote the proposal and revised the manuscript. M-JP and EL provided guidance in the statistical analysis and genome-wide association study and revised the manuscript. All authors read and approved the manuscript.

Funding

This project has received funding from the BioBased Industries Joint Undertaking under the European Union's Horizon 2020 research and innovation program under grant agreement No. 720726 (LIBBIO).

Acknowledgments

The authors would like to acknowledge the help of all the people that collaborated in scoring the agronomic traits. In particular, Prof. João Neves Martins, Sofia Alves and Gerard Bremer for coordinating the field trials and Annick Hoogendam and Eline Disselhorst for joining the harvest.

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

Ando, E., Ohnishi, M., Wang, Y., Matsushita, T., Watanabe, A., Hayashi, Y., et al. (2013). TWIN SISTER OF FT, GIGANTEA, and CONSTANS have a positive but indirect effect on blue light-induced stomatal opening in arabidopsis. *Plant Physiol.* 162 (3), 1529–1538. doi: 10.1104/pp.113.217984

Annicchiarico, P., Nazzicari, N., Notario, T., Martin, C. M., Romani, M., Ferrari, B., et al. (2021). Pea breeding for intercropping with cereals: variation for competitive ability and associated traits, and assessment of phenotypic and genomic selection strategies. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.731949

Atchison, G. W., Nevado, B., Eastwood, R. J., Contreras-Ortiz, N., Reynel, C., Madriñán, S., et al. (2016). Lost crops of the incas: Origins of domestication of the Andean pulse crop tarwi, lupinus mutabilis. *Am. J. Bot.* 103 (9), 1592–1606. doi: 10.3732/ajb.1600171

Avila, C. M., Nadal, S., Teresa Moreno, M., and Torres, A. M. (2006). Development of a simple PCR-based marker for the determination of growth habit in vicia faba l. using a candidate gene approach. *Mol. Breed.* 17 (3), 185–190. doi: 10.1007/s11032-005-4075-4

Ávila, C. M., Ruiz-Rodríguez, M. D., Cruz-Izquierdo, S., Atienza, S. G., Cubero, J. I., and Torres, A. M. (2017). Identification of plant architecture and yield-related QTL in vicia faba l. *Mol. Breed.* 37 (7), 1–13. doi: 10.1007/s11032-017-0688-7

Blanc, G., and Wolfe, K. H. (2004). Functional divergence of duplicated genes formed by polyploidy during arabidopsis evolution. *Plant Cell* 16 (7), 1679–1691. doi: 10.1105/tpc.021410

Bou-Torrent, J., Salla-Martret, M., Brandt, R., Musielak, T., Palauqui, J.-C., Martínez-García, J. F., et al. (2012). ATHB4 and HAT3, two class II HD-ZIP transcription factors, control leaf development in arabidopsis. *Plant Signaling Behav.* 7 (11), 1382–1387. doi: 10.4161/psb.21824

Caligari, P. D. S., Römer, P., Rahim, M. A., Huyghe, C., Neves-Martins, J., and Sawicka-Sienkiewicz, E. J. (2000). "The potential of lupinus mutabilis as a crop," in *Linking research and marketing opportunities for pulses in the 21st century: Proceedings of the third international food legumes research conference*. Ed. R. Knight (Dordrecht: Springer Netherlands), 569–573. doi: 10.1007/978-94-011-4385-1_54

Chirinos-Arias, M., Jiménez, J. E., Lizbeth, S., and Vilca, -.-M. (2015). Analysis of genetic variability among thirty accessions of Andean lupin (Lupinus mutabilis sweet) using ISSR molecular markers. *SCI. AGRON.* 6, 17–30. doi: 10.17268/sci.agropecu.2015.01.02

Dai, S., Zhang, Z., Chen, S., and Beachy, R. N. (2004). RF2b, a rice bZIP transcription activator, interacts with RF2a and is involved in symptom development of rice tungro disease. *Proceedings of the National Academy of Sciences* 101 (2), 687–692.

Dargahi, H., Tanya, P., Somta, P., Abe, J., and Srinives, P. (2014). Mapping quantitative trait loci for yield-related traits in soybean (Glycine max l.). *Breed. Sci.* 64 (4), 282–290. doi: 10.1270/jsbbs.64.282

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* 19, 11–15.

Foucher, F., Morin, J., Courtiade, J., Cadioux, S., Ellis, N., Banfield, M. J., et al. (2003). DETERMINATE and LATE FLOWERING are two TERMINAL FLOWER1/CENTRORADIALIS homologs that control two distinct phases of flowering initiation and development in pea. *Plant Cell* 15 (11), 2742–2754. doi: 10.1105/tpc.015701

Galek, R., Sawicka-Sienkiewicz, E., and Zalewski, D. (2007). Evaluation of interspecific hybrids of andean lupin and their parental forms with regard to some morphological and quantitative characters. *Fragment. Agron. (Poland)* 24 (94), 81–87.

Gocal, G. F. W., Poole, A. T., Gubler, F., Watts, R. J., Blundell, C., and King, R. W. (1999). Long-day up-regulation of a GAMYB gene during lolium temulentum inflorescence formation. *Plant Physiol.* 119 (4), 1271–1278. doi: 10.1104/pp.119.4.1271

González, A. M., Yuste-Lisbona, F. J., Saburido, S., Bretones, S., De Ron, A. M., Lozano, R., et al. (2016). Major contribution of flowering time and vegetative growth to plant production in common bean as deduced from a comparative genetic mapping. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01940

Gross, R., and von Baer, E. (1981). Lupine, eine neue kulturpflanze in den anden. II. umwelt-und genotypisch bedingte einflusse auf olgehalt und-qualitat im samen von tarwi (Lupinus mutabilis sweet). *Z. fur Acker-und Pflanzenbau. J. Agron. Crop Sci.* 150 (1), 27–38.

Guilengue, N., Alves, S., Talhinhas, P., and Neves-Martins, J. (2020). Genetic and genomic diversity in a tarwi (Lupinus mutabilis sweet) germplasm collection and adaptability to Mediterranean climate conditions. *Agronomy* 10 (1), 215. doi: 10.3390/agronomy10010021

Gulisano, A., Alves, S., Rodriguez, D., Murillo, A., Dinter, B.-J. V., Torres, A. F., et al. (2022a). Diversity and agronomic performance of lupinus mutabilis germplasm in European and Andean environments. *Front. Plant Sci.* 13, 903661. doi: 10.3389/fpls.2022.903661

Gulisano, A., Dechesne, A., Paulo, M.-J., and Trindade, L. (2022b). Investigating the potential of Andean lupin as a lignocellulosic feedstock for Europe: first genome-wide association study on l. mutabilis biomass quality. *GCB Bioenergy* 15, 38-57. doi: 10.1111/gcbb.13006

Gupta, P. K., Rustgi, S., and Kulwal, P. L. (2005). Linkage disequilibrium and association studies in higher plants: Present status and future prospects. *Plant Mol. Biol.* 57, 461–485. doi: 10.1007/s11103-005-0257-z

Hane, J. K., Ming, Y., Kamphuis, L. G., Nelson, M. N., Garg, G., Atkins, C. A., et al. (2017). A comprehensive draft genome sequence for lupin (Lupinus angustifolius), an emerging health food: insights into p klant–microbe interactions and legume evolution.". *Plant Biotechnol. J.* 15 (3), 318–330. doi: 10.1111/pbi.12615

Hardy, A., Huyghe, C., and Papineau, J. (1997). Dry matter accumulation and partitioning, and seed yield in indeterminate Andean lupin (Lupinus mutabilis sweet). *Aust. J. Agric. Res.* 48 (1), 91–102.

Hardy, A., Huyghe, C., Rahim, M. A., Roemer, P., Neves-Martins, J. M., Sawicka-Sienkiewicz, E., et al. (1998). Effects of genotype and environment on architecture and flowering time of indeterminate Andean lupins (*Lupinus mutabilis sweet*). *Aust. J. Agric. Res.* 49 (8), 1241–1252. doi: 10.1071/A98060

He, Z., Zhu, Q., Dabi, T., Li, D., Weigel, D., and Lamb, C. (2000). Transformation of rice with the arabidopsis floral regulator LEAFY causes early heading. *Transgenic Res.* 9 (3), 223–227. doi: 10.1023/A:1008992719010

Hu, J., Mitchum, M. G., Barnaby, N., Ayele, B. T., Ogawa, M., Nam, E., et al (2008). Potential sites of bioactive gibberellin production during reproductive growth in Arabidopsis. *The Plant cell* 20 (2), 320–336. doi: 10.1105/tpc.107.057752

Kaler, A. S., and Purcell, L. C. (2019). Estimation of a significance threshold for genome-wide association studies. *BMC Genomics* 20 (1), 1–85. doi: 10.1186/s12864-019-5992-7

Kang, H. M., Zaitlen, N. A., Wade, C. M., Kirby, A., Heckerman, D., Daly, M. J., et al. (2008). Efficient control of population structure in model organism association mapping. *Genetics* 178 (3), 1709–1723. doi: 10.1534/genetics.107.080101

Klein, A., Houtin, H., Rond-Coissieux, C., Naudet-Huart, M., Touratier, M., Marget, P., et al. (2020). Meta-analysis of QTL reveals the genetic control of yield-related traits and seed protein content in pea. *Sci. Rep.* 10 (1), 1–11. doi: 10.1038/s41598-020-72548-9

Krylova, E. A., Khlestkina, E. K., Burlyaeva, M. O., and Vishnyakova, M. A. (2020). Determinate growth habit of grain legumes: role in domestication and selection, genetic control. *Ecol. Genet.* 18 (1), 43–58. doi: 10.17816/ecogen16141

Kwak, M., Velasco, D., and Gepts, P. (2008). Mapping homologous sequences for determinacy and photoperiod sensitivity in common bean (Phaseolus vulgaris).". *J. Hered.* 99 (3), 283–291. doi: 10.1093/jhered/esn005

Li, Z., Cheng, Q., Gan, Z., Hou, Z., Zhang, Y., Li, Y., et al. (2021). Multiplex CRISPR/Cas9-mediated knockout of soybean LNK2 advances flowering time. *Crop J.* 9 (4), 767–776. doi: 10.1016/j.cj.2020.09.005

Li, Z., Ou, Y., Zhang, Z., Li, J., and He, Y. (2018). Brassinosteroid signaling recruits histone 3 lysine-27 demethylation activity to FLOWERING LOCUS c chromatin to inhibit the floral transition in arabidopsis. *Mol. Plant* 11 (9), 1135–1146. doi: 10.1016/j.molp.2018.06.007

Martins, J. M. N., Talhinhas, P., and de Sousa, R. B. (2016). Yield and seed chemical composition of lupinus mutabilis in Portugal. *Rev. Ciências Agrárias* 39, 518–525. doi: 10.19084/RCA16079

Petit, J., Salentijn, E. M. J., Paulo, M.-J., Denneboom, C., Van Loo, E. N., and Trindade, L. M. (2020). Elucidating the genetic architecture of fiber quality in hemp (Cannabis sativa l.) using a genome-wide association study. *Front. Genet.* 11, 566314. doi: 10.3389/fgene.2020.566314

Renaud, E. N. C., van Bueren, E. T.L., Myers, J. R., Paulo, M. J., Eeuwijk, F. A.V., Zhu, N., et al. (2014). Variation in broccoli cultivar phytochemical content under organic and conventional management systems: Implications in breeding for nutrition. *PloS One* 9 (7), e95683. doi: 10.1371/journal.pone.0095683

Rodriguez-Alvarez, M. X., Boer, M. P., van Eeuwijk, F. A., and Eilers, P. H. C. (2018). Correcting for spatial heterogeneity in plant breeding experiments with p-splines. *Spatial Stat* 23, 52–71. doi: 10.1016/j.spasta.2017.10.003

Rugnone, M. L., Soverna, A. F., Sanchez, S. E., Schlaen, R. G., Hernando, C. E., Seymour, D. K., et al. (2013). LNK genes integrate light and clock signaling

networks at the core of the arabidopsis oscillator. *Proc. Natl. Acad. Sci.* 110 (29), 12120–12125. doi: 10.1073/pnas.1302170110

Rychel, S., Książkiewicz, M., Tomaszewska, M., Bielski, W., and Wolko, B. (2019). FLOWERING LOCUS T, GIGANTEA, SEPALLATA, and FRIGIDA homologs are candidate genes involved in white lupin (Lupinus albus l.) early flowering. *Mol. Breed.* 39 (3), 1–17. doi: 10.1007/s11032-019-0952-0

Shen, Y., Lei, T., Cui, X., Liu, X., Zhou, S., Zheng, Y., et al. (2019). Arabidopsis histone deacetylase HDA 15 directly represses plant response to elevated ambient temperature. *Plant J.* 100 (5), 991–1006. doi: 10.1111/tpj.14492

Shu, K., and Wenyu, Y. (2017). E3 ubiquitin ligases: Ubiquitous actors in plant development and abiotic stress responses. *Plant Cell Physiol.* 58 (9), 1461–1476. doi: 10.1093/pcp/pcx071

Sinjushin, A. A. (2015). Mutations of determinate growth and their application in legume breeding. *Legume Perspect.* 6, 14–15.

Sul, J. H., Martin, L. S., and Eskin, E. (2018). Population structure in genetic studies: Confounding factors and mixed models.". *PloS Genet.* 14 (12), e1007309. doi: 10.1371/journal.pgen.1007309

Tian, Z., Wang, X., Lee, R., Li, Y., Specht, J. E., Nelson, R. L., et al. (2010). Artificial selection for determinate growth habit in soybean. *Proc. Natl. Acad. Sci.* 107 (19), 8563–8568. doi: 10.1073/pnas.1000088107

Turner, S. D. (2018). Qqman: an r package for visualizing GWAS results using QQ and manhattan plots. *J. Open Source Software* 3, 731. doi: 10.21105/joss.00731

van Rossum, B.-J., Kruijer, W., Eeuwijk, F. v., Boer, M., Malosetti, M., Bustos-Korts, D., et al. (2020). *Package 'statgenGWAS'* Vol. 7 (R package version 1.0). Available at https://CRAN.R-project. org/package=statgenGWAS

Vos, P. G., Paulo, M. J., Voorrips, R. E., Visser, R. G. F., Eck, H. J.v., and Eeuwijk, F. A.v. (2017). Evaluation of LD decay and various LD-decay estimators in simulated and SNP-array data of tetraploid potato. *Theor. Appl. Genet.* 130 (1), 123–135. doi: 10.1007/s00122-016-2798-8

Wahl, V., Brand, L. H., Guo, Y.-L., and Schmid, M. (2010). The FANTASTIC FOUR proteins influence shoot meristem size in arabidopsis thaliana. *BMC Plant Biol.* 10 (1), 1–12. doi: 10.1186/1471-2229-10-285

Weigel, D., and Nilsson, O. (1995). A developmental switch sufficient for flower initiation in diverse plants. *Nature* 377 (6549), 495–500. doi: 10.1038/377495a0

Wickland, D. P., and Hanzawa, Y. (2015). The FLOWERING LOCUS T/TERMINAL FLOWER 1 gene family: functional evolution and molecular mechanisms. *Mol. Plant* 8 (7), 983–997. doi: 10.1016/j.molp.2015.01.007

Xi, W., Liu, C., Hou, X., and Yu, H. (2010). MOTHER OF FT AND TFL1 regulates seed germination through a negative feedback loop modulating ABA signaling in arabidopsis. *Plant Cell* 22 (6), 1733–1748. doi: 10.1105/tpc.109.073072

Yao, D., Liu, Z. Z., Zhang, J., Liu, S. Y., Qu, J., Guan, S. Y., et al. (2015). Analysis of quantitative trait loci for main plant traits in soybean. *Genet. Mol. Res.* 14 (2), 6101–6109. doi: 10.4238/2015.June.8.8

Zhang, H., Cui, F., Wu, Y., Lou, L., Liu, L., Tian, M., et al. (2015). The RING finger ubiquitin E3 ligase SDIR1 targets SDIR1-INTERACTING PROTEIN1 for degradation to modulate the salt stress response and ABA signaling in arabidopsis. *Plant Cell* 27 (1), 214–227. doi: 10.1105/tpc.114.134163

Zhang, Y., Zhang, B., Yang, T., Zhang, J., Liu, B., Zhan, X., et al. (2020). The GAMYB-like gene SlMYB33 mediates flowering and pollen development in tomato. *Horticult. Res.* 7 (2020), 133. doi: 10.1038/s41438-020-00366-1

Zotova, L., Kurishbayev, A., Jatayev, S., Goncharov, N. P., Shamambayeva, N., Kashapov, A., et al. (2019). The general transcription repressor TaDr1 is co-expressed with TaVrn1 and TaFT1 in bread wheat under drought. *Front. Genet.* 10, 63. doi: 10.3389/fgene.2019.00063

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 14 February 2023
DOI 10.3389/fpls.2023.1104417

OPEN ACCESS

EDITED BY
Ndiko Ndomelele Ludidi,
University of the Western Cape,
South Africa

REVIEWED BY
Sezai Ercisli,
Atatürk University, Türkiye
Uday Chand Jha,
Indian Institute of Pulses Research
(ICAR), India

*CORRESPONDENCE
Kafilat Abiodun Odesola
kaodesola@bellsuniversity.edu.ng

SPECIALTY SECTION
This article was submitted to
Crop and Product Physiology,
a section of the journal
Frontiers in Plant Science

RECEIVED 21 November 2022
ACCEPTED 05 January 2023
PUBLISHED 14 February 2023

CITATION
Odesola KA, Olawuyi OJ, Paliwal R,
Oyatomi OA and Abberton MT (2023)
Genome-Wide association analysis of
phenotypic traits in Bambara groundnut
under drought-stressed and non-stressed
conditions based on DArTseq SNP.
Front. Plant Sci. 14:1104417.
doi: 10.3389/fpls.2023.1104417

Genome-Wide association analysis of phenotypic traits in Bambara groundnut under drought-stressed and non-stressed conditions based on DArTseq SNP

Kafilat Abiodun Odesola[1,2,3]*, Odunayo Joseph Olawuyi[3], Rajneesh Paliwal[2], Olaniyi Ajewole Oyatomi[2] and Michael T. Abberton[2]

[1]Department of Biological Sciences, Bells University of Technology, Sango Otta, Ogun State, Nigeria, [2]Genetic Resources Centre, International Institute of Tropical Agriculture, Ibadan, Oyo State, Nigeria, [3]Department of Botany, University of Ibadan, Ibadan, Oyo State, Nigeria

Introduction: Bambara groundnut (BG) (Vigna subterranea [L.] Verdc) is an indigenous, resilient, but underutilized leguminous crop that occurs mostly as genetically heterogeneous landraces with limited information on the drought tolerant attributes. This study elucidates the associations between sequencing-based diversity array technology (DArTseq) and phenotypic character as well as differing indices related to drought tolerance in one hundred accessions of Bambara groundnut.

Methods: The field experiments were conducted at IITA research stations in Kano and Ibadan between 2016 and 2018 planting seasons. The experiments were arranged in randomised complete block design with three replications, under the different water regimes. The phenotypic traits evaluated was further to construct the dendrogram. Genome-wide association mapping was conducted based on 5927 DArTs loci with < 20% missing data.

Results and Discussions: The genome wide association study predicted drought tolerance in Bambara accessions for geometric mean productivity (GMP) and stress tolerance index (STI). TVSu-423 had the highest GMP and STI values (28.50, 2.40), while TVSu-2017 had the lowest at GMP (1.74) and STI (0.01) respectively. The relative water content (%) was significantly higher for accessions; TVSu-266 (60.35, 61.49), TVSu-2 (58.29, 53.94), and TVSu-411 (55.17, 58.92) in 2016/2017 and 2017/2018, respectively. The phenotypic characters studied delineated the accessions into two major clusters and five distinct sub-clusters, indicating variations across all the geographical locations. The 5,927 DArTseq genomic markers in association with STI further grouped the 100 accessions into two main clusters. TVSu-1897 from Botswana (Southern Africa) was in the first cluster, while the remaining 99 accessions from Western, Central, and Eastern Africa made up the second cluster. The eight significant Quantitative Trait Loci (QTLs) (24346377|F|0-22:A>G-22:A>G, 24384105|F|0-56:A>G33 :A> G, 24385643|F|0-53:G>C-53:G>C, 24385696|F|0-43:A>G-43:A>G, 4177257|F|0-44:A>T-44:A>T, 4182070|F|0-66:G>A-66:G>A, 4183483|F|0-24:

G>A-24:G>A, 4183904|F|0-11:C>T-11:C>T) identified with Bonferroni threshold was in association with STI, indicative of variations under the drought-stressed condition. The observation of consistent SNPs in the 2016 and 2017 planting seasons, as well as in combination with the 2016 and 2017 planting seasons, led to the designation of these QTLs as significant. The drought selected accessions could form basis for hybridization breeding. The identified quantitative trait loci could be useful in marker-assisted selection in drought molecular breeding programs.

KEYWORDS

accessions, underutilized, chromosome, drought, snps, DArTseq

Introduction

Given its unique resilience attributes, Bambara groundnut (BG) (*Vigna subterranea*) belongs to the group of underutilized crops presently being promoted as climate-smart crops (Feldman et al., 2019). Underutilized plants provide several impressive health benefits. They are also better adapted to marginal lands and biotic and abiotic stress conditions. They can contribute significantly to the diversification and resilience of agroecosystems (Dogan et al., 2014; Saran et al., 2021; Bozhuyuk, 2022). Throughout much of sub-Saharan Africa, it is a widely cultivated grain legume (Directorate Plant Production, 2011; Ahmad, 2013). BG is a highly self-fertilizing crop with a rather variable yield owing to its inconsistency in its response to different environmental circumstances, among which is drought stress (Bamshaiye et al., 2011). Drought tolerance is a quantitative trait with a lot of complex genetic and phenotypic control and gradual improvement (Sinclair, 2011). Prior research work on quantitative traits was based on the identification of single gene effects on phenotype (Orgogozo et al., 2015), until the recent advances in genomics and molecular biology improved the identification of candidate genes and quantitative trait loci (QTLs) through the simultaneous dissection of these complex traits using agronomic traits, gene region and association analysis (Jha et al., 2022; Wang et al., 2022).

The genome-wide association study (GWAS) is the commonest approach to understanding the association between the complexities controlling most agronomic traits of interest to their genetic bases. It relies on the association of several representative markers and a huge genetically diverse population of organisms. It finds the linkage disequilibrium between the phenotypic trait and the genetic markers (Korte and Farlow, 2013) without any prior knowledge of the level of the kinship of the organism. This mapping approach makes use of the occurrence of rare alternative forms of a gene on a particular chromosome at different gene locations to identify a common site of interest for any genetic relationship (Flint-Garcia et al., 2003; Mackay and Powell, 2007). Another benefit of association mapping for the study of a quantitative trait is that it is highly useful for researching species that are hard to cross or clone or that take a long time to reproduce (Nordborg and Weigel, 2008). Overall GWAS helps find markers, genes, or QTLs associated with phenotypic traits that can conveniently be used for gene introgression, gene discovery, or marker-assisted breeding.

The basic DNA-based marker systems used in mapping the most complex traits in different crop species include simple sequence repeat (SSR) and Amplified fragment length polymorphism (AFLP), which are mostly gel-based markers and are limited in their ability to perform rapid analysis on a large number of marker loci. Single Nucleotide Polymorphism (SNP) and the Diversity Arrays Technology (DArT) marker system are the most recent and better mapping tools, and they become more robust when used together (Nadeem et al., 2018). Diversity Array Technology sequencing (DarTseq) is one novel platform under the next-generation sequencing (NGS) platform developed to perform a simultaneous Single Nucleotide (SNP) discovery of unique nucleotides among a population, particularly for non-model germplasm sets (Kilian et al., 2012; Raman et al., 2014). This method for studying genetic diversity is based on the reduction of complexity through the application of restriction enzymes that target gene-rich areas. It, thus, thoroughly covers the genome using DarTseqTM technology to provide a high-density genetic map that increases the likelihood of finding QTL (Thudi et al., 2011). The combination of SNPs and the NGS reported for other crops of interest had been identified to bring uniqueness to the identification of genetic variants (SNPs), phylogenetics, germplasm assessment, and population structure (Grzebelus et al., 2014). A few of the grain crops on which DarT markers have previously been used include pigeon pea (Yang et al., 2011), soybean (Vu et al., 2015), and common bean, (Valdisser et al., 2017). Given the abundance of SNPs, the uniqueness of the NGS platform, and the complexity of drought trait, this study, therefore, aims to identify marker-trait association of some phenotypic traits of Bambara groundnut using the representative DarTseq to discover relevant QTLs for possible future breeding work.

The specific objectives of the present study are:

(ii) To evaluate the performance of DarTseq method-derived markers in Bambara groundnut,

(ii) Map QTLs/for drought tolerance in Bambara groundnut.

Materials and methods

Plant material and experimental design

For the study, 100 accessions of Bambara groundnut seeds from 12 African countries were collected from the International

Institute of Tropical Agriculture (IITA), Ibadan, Nigeria. The 100 accessions (Table 1) were subjected to intense phenotyping of key traits related to drought. Six of these traits were finally selected after a thorough field experiment involving two planting seasons (2016/2017 and 2017/2018), two water regimes (well-watered (WW) and water-stressed (WS) conditions), and two different locations (Ibadan and Kano stations research fields of IITA). The trials were set up using a Randomized Complete Block Design (RCBD) with three replications. Each plot measured 1 m by 2.5 m and the distances between rows were 1.00 meters and 0.25 meters, respectively. The population was phenotyped for the number of days to 50% flowering (FLW), Visual scoring (VSL) at three weeks, at six weeks and at the termination of the experiment after the imposition of drought stress, leaf chlorophyll content (CHL), number of seeds per plot (SPL), weight of seeds per plot (WTPL), and grain yield per plot (GYLD). The data were further subjected to yield trait analysis using the eleven tolerant indices: Stress Susceptibility Index (SSI), Relative Drought Index (RDI), Stress Tolerance Index (STI), Geometric Mean Production (GMP), Tolerance Index (TOL), Mean Production (MP), Yield Index (YI), Drought Resistance Index (DI), Yield Stability Index (YSI), Stress Susceptibility Percentage Index (SSPI), and Modified Stress Tolerance (MSTI).

DNA extraction and quantification

The 100 accessions of Bambara groundnut, plus two as control, were raised in controlled environments at the IITA research field and leaf tissues were sampled for DNA extraction at 4 weeks after planting. Young leaf tissue weighing one gram was harvested, promptly frozen in liquid nitrogen, and then kept at 80°C. Using a typical cetyltrimethylammonium bromide (CTAB), chloroform, and isoamyl alcohol technique with a little modification by the addition of 20% Sodium Dodecyl Sulphate (SDS), genomic DNA was recovered from the frozen leaves (Doyle and Doyle, 1987). The quality of the DNA was performed with 0.8% agarose gel while purity checks at 260/230 and 260/280 nm absorbance ratios were done with Nanodrop spectrophotometer (ND2000 V3.5, NanoDrop Technologies, Inc.). Still using the Nanodrop spectrophotometer, a final adjustment of the DNA was performed to 100 ng/µl for subsequent DarT and SNP genotyping.

Genotyping of individual samples using DarTseq technology

The 102 accessions of Bambara groundnut were forwarded to Diversity Arrays Technologies commercial service Ltd., Australia (www.diversityarrays.com) for individual genotyping with the HiSeq 2000 (Illumina) next-generation sequencer. After rigorous quality control and filtering with a call rate of 80%, marker reproducibility of 95%, and missing data of 20%, 5,927 SNPs were found to be polymorphic out of a total of 11,821 DarTseq markers generated for further analysis. Phenotypic data were subjected to analysis of variance (ANOVA) using the Statistical Analysis System (SAS 9.0) while the treatment means were all compared using the Duncan Multiple range test at a probability level of 0.05. To ascertain the association significance using the established threshold $P<1.68\times 10^{-4}$ (Li et al., 2016), the field data were subsequently subjected to the Best Linear Unbiased Prediction (BLUP) and correlated with the genetic sequence using the TASSEL molecular software version 5.2.23 for the Genome-Wide Association Study.

Results

Phenotypic traits evaluation

The result of mean square variance showed significant differences ($P< 0.001$) in the performances of the 100 Bambara groundnut accessions with regards to location, year water regime, and their interaction effects for all the studied traits. High and positive correlations were observed for all traits, however, no significant difference was recorded for the year effect on the 50% flowering rate while the second-order interactive effect of water and location on chlorophyll was also not significant. Analysis of variance further indicated significant differences in the chlorophyll content index among the accessions with respect to location, water treatment condition, and all other studied interactions except for water treatment/location combination (Table 2).

SNPs polymorphism, heterozygosity, minor allele frequency

A total number of 11,821 DarTseq markers were discovered and from these, 5,927 SNPs showed polymorphism and were subsequently retained for further analysis after thorough quality control and filtering were ascertained. Quality control was ensured at a consistent call rate ≥ 80%, while marker reproducibility was at ≥ 95%, and missing data ≤ 20%. The most frequently occurring polymorphism type was a cytosine to thymine combination (907; 15.3%), followed by guanine to adenine (871; 14.7%), with the least being observed with guanine to thymine at 4% and a proportion of 255 out of the whole 5927 SNPs (Table 3). The sequenced DNAs generated more transition (Ts) substitutions than transversion (Tv) substitutions at a percentage ratio of 58% to 39%, leading to an overall value of 1.52 for Ts/Tv (Table 4). Transitions C↔T and A↔G were represented with 29.8% and 28.9% of the total substitutions, respectively, while the four transversion classes occurrences were as displayed in Figure 1. The higher occurrence of transition showed that the SNPs occurrence is more in the coding region than in the non-coding region. Furthermore, the fraction of homozygous (G:G, C:C, A:A, and T:T) observations for individual SNPs were all in the percentage ratio of 0.5% to 0.8%. A range of 1% to 23% heterozygosity was observed among the 100 accessions, with an overall average of 0.06%. Of the accessions, 21 had 1% heterozygosity, with 67 of the accessions falling within the < 5% of recommended heterozygous rate for a self-pollinating plant. The remaining 12 accessions had heterozygous tendencies between 6% to 23%. The identified individual accession level of outcrossing evidenced the opportunity that could accrue from exploiting some of these accessions for further breeding purposes (Figure 2A). The set criteria for the minor allele frequency (MAF) were at <0.01 to show

10.3389/fpls.2023.1104417

TABLE 1 List of the accessions with their countries of origin.

S/N	Acc Name	Country of origin	S/N	Acc Name	Country of origin
1	TVSu-188	Benin	51	TVSu-250	Gambia
2	TVSu-189	Benin	52	TVSu-251	Gambia
3	TVSu-193	Benin	53	TVSu-252	Gambia
4	TVSu-194	Benin	54	TVSu-1	Nigeria
5	TVSu-203	Benin	55	TVSu-2	Nigeria
6	TVSu-288	Benin	56	TVSu-4	Nigeria
7	TVSu-1092	Benin	57	TVSu-5	Nigeria
8	TVSu-1897	Botswana	58	TVSu-9	Nigeria
9	TVSu-85	Burkina Faso	59	TVSu-10	Nigeria
10	TVSu-86	Burkina Faso	60	TVSu-22	Nigeria
11	TVSu-292	Burkina Faso	61	TVSu-174	Nigeria
12	TVSu-1157	Burkina Faso	62	TVSu-182	Nigeria
13	TVSu-1161	Burkina Faso	63	TVSu-258	Nigeria
14	TVSu-1162	Burkina Faso	64	TVSu-265	Nigeria
15	TVSu-1164	Burkina Faso	65	TVSu-266	Nigeria
16	TVSu-1166	Burkina Faso	66	TVSu-282	Nigeria
17	TVSu-1175	Burkina Faso	67	TVSu-286	Nigeria
18	TVSu-2017	Burundi	68	TVSu-326	Nigeria
19	TVSu-2018	Burundi	69	TVSu-338	Nigeria
20	TVSu-395	Cameroon	70	TVSu-354	Nigeria
21	TVSu-399	Cameroon	71	TVSu-362	Nigeria
22	TVSu-409	Cameroon	72	TVSu-593	Nigeria
23	TVSu-411	Cameroon	73	TVSu-595	Nigeria
24	TVSu-416	Cameroon	74	TVSu-682	Zambia
25	TVSu-418	Cameroon	75	TVSu-687	Zambia
26	TVSu-421	Cameroon	76	TVSu-689	Zambia
27	TVSu-423	Cameroon	77	TVSu-690	Zambia
28	TVSu-434	Cameroon	78	TVSu-691	Zambia
29	TVSu-442	Cameroon	79	TVSu-692	Zambia
30	TVSu-445	Cameroon	80	TVSu-693	Zambia
31	TVSu-447	Cameroon	81	TVSu-699	Zambia
32	TVSu-448	Cameroon	82	TVSu-702	Zambia
33	TVSu-449	Cameroon	83	TVSu-713	Zambia
34	TVSu-459	Cameroon	84	TVSu-716	Zambia
35	TVSu-504	Cameroon	85	TVSu-719	Zambia
36	TVSu-1277	Central African Republic	86	TVSu-725	Zambia
37	TVSu-1278	Central African Republic	87	TVSu-731	Zambia
38	TVSu-1284	Central African Republic	88	TVSu-736	Zambia
39	TVSu-1285	Central African Republic	89	TVSu-745	Zambia
40	TVSu-1289	Central African Republic	90	TVSu-978	Zimbabwe

(Continued)

TABLE 1 Continued

S/N	Acc Name	Country of origin	S/N	Acc Name	Country of origin
41	TVSu-1290	Central African Republic	91	TVSu-989	Zimbabwe
42	TVSu-1296	Central African Republic	92	TVSu-1011	Zimbabwe
43	TVSu-1309	Central African Republic	93	TVSu-1014	Zimbabwe
44	TVSu-1320	Central African Republic	94	TVSu-1015	Zimbabwe
45	TVSu-1373	Central African Republic	95	TVSu-1023	Zimbabwe
46	TVSu-1991	Congo	96	TVSu-1026	Zimbabwe
47	TVSu-115	Ivory Coast	97	TVSu-1034	Zimbabwe
48	TVSu-116	Ivory Coast	98	TVSu-1051	Zimbabwe
49	TVSu-118	Ivory Coast	99	TVSu-1056	Zimbabwe
50	TVSu-247	Gambia	100	TVSu-1078	Zimbabwe

TVSu, Tropical Vigna Subterranea.

that the rare allele frequency was well distributed and that the 5,927 SNP markers generated by DArTseq came from distinct sequences. The majority of MAFs were well distributed (Figure 2B).

Population distribution

Neighbor-Joining Phylogeny was generated with the DArT markers, and these produced several sub-clusters of related accessions. Clustering was mostly based on agro-climatic areas and a possible similar genetic background. This was evident with an observed delineation into two main clusters, with the first cluster harboring just one accession sampled from South Africa. The remaining 99 accessions were within three sub-clusters, having accessions from West, Central, and East Africa all clustering together. Within the 44 West Africa accessions, there were approximately seven Central Africa accessions clustered within. The uniqueness of these seven accessions was in the fact that they all originated from Cameroon. While 28 selections were from Central Africa, the remaining 27 accessions were from East Africa. Accession TVSu-362 from West Africa was found among the bulk of the Central Africa accessions, while four different accessions from East Africa and one accession from West Africa were observed to have clustered among the 27 Central Africa accessions (Figure 3).

Principal coordinate analysis (PCA)

The principal coordinate analysis (PCA) was exploited to measure the observed variations based on the DArTseq markers. The PCA is explained by fifteen components with wide-ranging variations, as the first two principal coordinates explain the largest proportion of the genetic diversity. Of the total variations, 34 (34%) were accounted for

TABLE 2 Mean squares for the traits for 100 accessions of Bambara groundnut planted in Ibadan and Kano in the 2015/16 and 2016/17 planting seasons.

Source	DF	FLW	CHL	VSL	RWC	SPL	WTPL
ACCNS	99	708.34***	1038.03***	3.95***	3066.66***	4771.58***	810.47***
LOCATN	1	1950.49***	19084.83***	98.91***	879801.16***	725903.82***	158246.62***
YEAR	1	6876.45NS	41384.07***	725.49***	8825.63***	85697.04***	26723.03***
WTRT	1	369.72***	1604.99***	43.97***	3142.25**	434093.37***	66457.48***
WTRT*LOCATN	2	405.13**	10.89NS	20531.97***	114.98***	143303.4****	800.24***
LOCATN*YEAR	1	192.71***	7181.83***	1021.87***	34969.76***	114.98***	1586.89***
ACCNS*WTRT	99	18.24***	95.45NS	2.02**	609.23**	2554.66***	626.7***
ACCNS*LOCATN	99	126.76***	297.34***	3.14***	1448.56***	3922.05***	413.09***
ACCNS*YEAR	99	273.05***	432.67***	3.28***	916.01***	2631.76***	442.45***
ACCN*WTRT*LOCAT*YEAR	398	53.74NS	192.54***	1.86***	596.15**	0.95***	350.18***
POOLED ERROR	801	179.86***	412.22***	6.66***	2225.64***	4439.87***	807.97***

DF-Degree of freedom; FLW-Days to flowering; VSL-Visual Scoring; RWC-Relative water content; SPL-Number of seeds/plotWTPL-Weight of seeds/Plot; Levels of significance - (**P<-0.01%, ***P<0.001%.)

ACCNS-Accessions; WTRT-Water treatment; LOCATN-location; YEAR-Year.

ACCNS*WTRT–Accessions/water treatment interaction; ACCNS*LOCATN-Accession/location interaction.

ACCNS*WTRT-accession/water treatment interaction; ACCNS*YEAR-accession/year interaction.

WTRT*LOCATN–Water treatment by location interaction; LOCATN*YEAR-location/year interaction.

ACCN*WTRT*LOCAT*YEAR-Accession/water treatment/location/year interaction.

10.3389/fpls.2023.1104417

TABLE 3 Summary of genomic sequencing.

Alleles	Number	Proportion	Frequency
C	144338	0.23875	0.25002
G	139777	0.23121	0.24212
A	138614	0.22928	0.24011
T	138098	0.22843	0.23921
N	27250	0.04507	0.0472
Y	5637	0.00932	0.00976
R	4896	0.0081	0.00848
W	1852	0.00306	0.00321
M	1479	0.00245	0.00256
S	1418	0.00235	0.00246
K	1195	0.00198	0.00207
C:T	907	0.15303s	
G:A	871	0.14695	
T:C	860	0.1451	
A:G	839	0.14156	
T:A	329	0.05551	
A:T	305	0.05146	
G:C	305	0.05146	
A:C	289	0.04876	
C:G	285	0.04809	
C:A	266	0.04488	
T:G	259	0.0437	
G:T	255	0.04302	
G:G	48	0.0081	
C:C	44	0.00742	
A:A	33	0.00557	
T:T	32	0.0054	
Total SNPs	5927		

C-Cytosine, G-Guanine, A-Adenine, T-Thymine, Y-pyrimidine (C or T), R-Purine (A or G), S-Strong (G or C), N-Any nucleotide, W-(A or T), M-amino (A or C), K- Keto (G or T).

TABLE 4 Degree of transition and transversion identified using DarTseq.

Polymorphism Type	Allele	Number of allelic sites	Percentage of allelic sites	Total percentage
Transition	A↔G	1710	28.9%	
	C↔T	1767	29.8%	**3,477(58.7%)**
	A↔C	555	9.40%	
	A↔T	634	10.70%	
	G↔C	590	10.00%	
	G↔T	514	8.70%	**2,293(38.7%)**

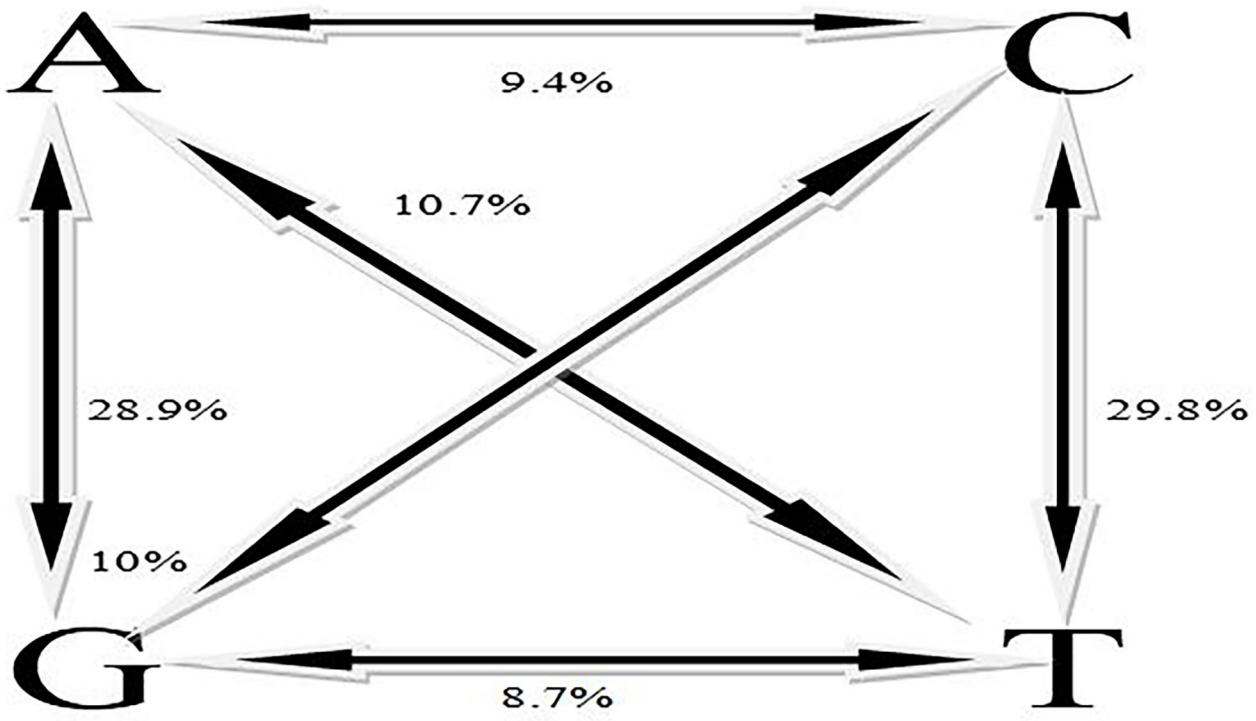


FIGURE 1
Percentage distribution of transitions and transversions among the SNPs.

by the PC1 while the remaining variations observed by the rest of the components were accounted for at just 2.82%. (Figure 4A). Furthermore, the PCA approach, which is mainly based on the genetic distance matrix and other data standardization, showed a consistent pattern with the phylogeny and grouped the 100 accessions of Bambara into three major clusters, similar to what was observed with the dendrogram. On the left side of the quadrant and colored red, are the accessions tagged to have originated from West Africa, with the few accessions colored green being known to have been from Central Africa. Likewise, on the right are accessions from East Africa with few accessions from Central Africa clustering together. In addition to the cluster is the accession originating from South Africa. (Figure 4B). In the dendrogram, the Southern Africa accessions occupied a main cluster delineating it from other accessions. The middle of the quadrant is occupied by accessions coming from Central Africa with some accessions coming from East Africa.

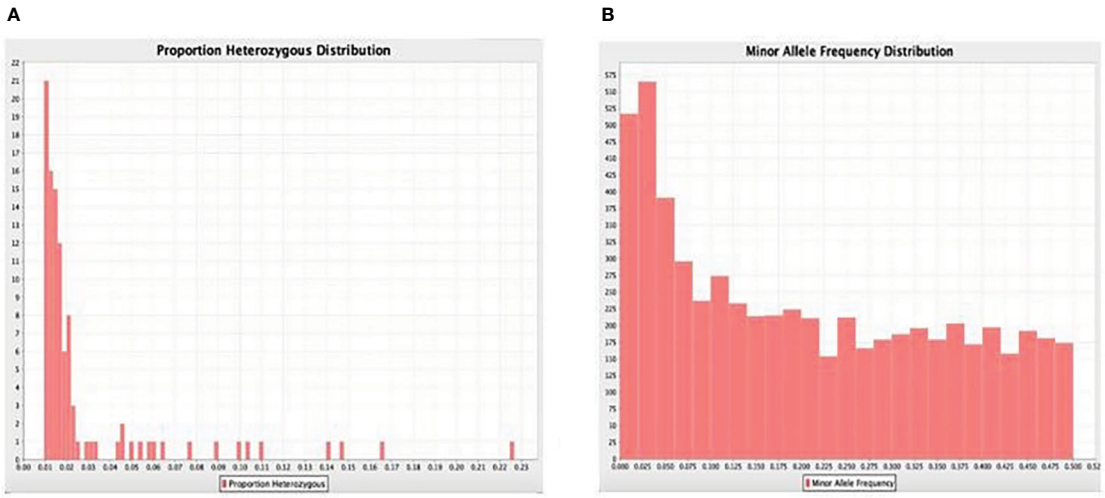


FIGURE 2
(A) Gene summary showing the frequency proportions of the heterozygosity frequency and **(B)** minor (rare) alleles.

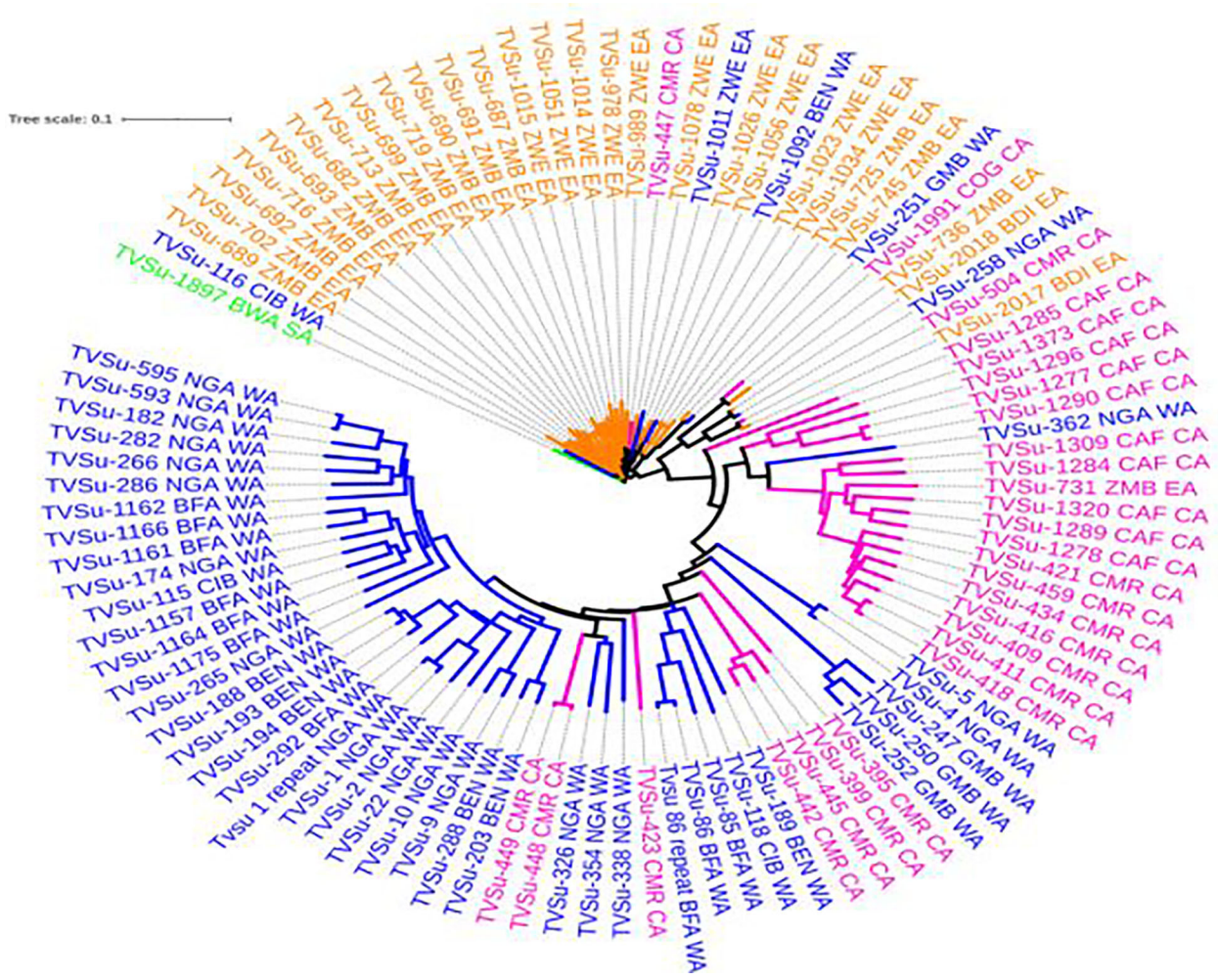


FIGURE 3
Phylogenetic relationship among Bambara groundnut accessions from 12 countries. Interactive tree of Life (ITOL) Letunic and Bork (2016) *Nucleic Acids Res.*

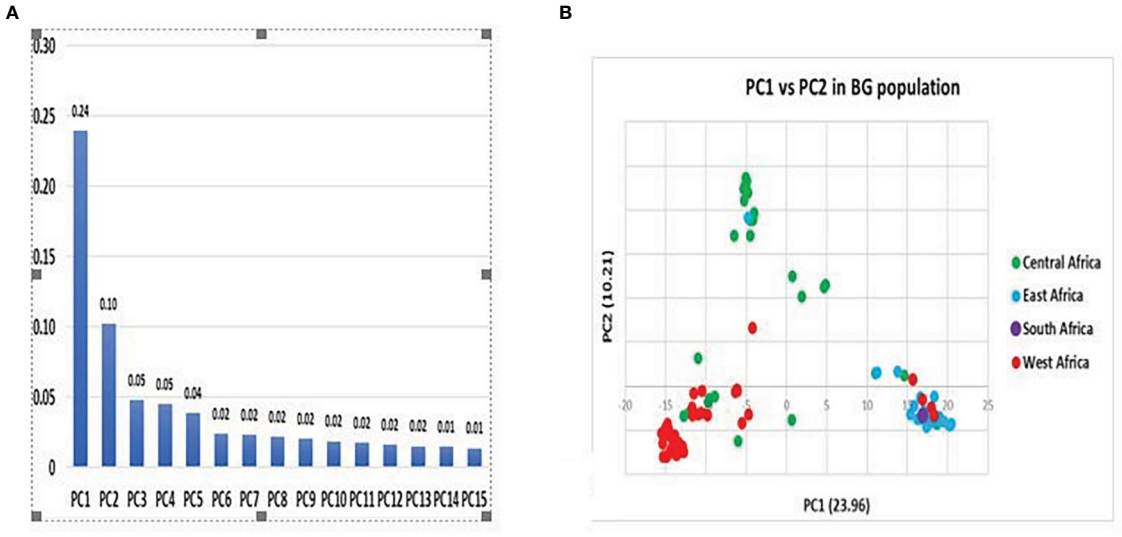


FIGURE 4
(A) Principal component contribution and **(B)** Principal coordinate analysis of total variation of Bambara accessions population based on DArTseq.

10.3389/fpls.2023.1104417

Marker-trait associations

A total of 11,083 "raw" SNPs with potential InDels were found; following careful filtering of SNPs, which excluded markers with more than 20% incomplete data and, thus, a read depth of >80% marker's call rate, and minor allele frequency (MAF > 0.01) on the 102 lines, there was a yield of 5927 SNPs. Of the 5927 SNPs, 4095 sequences could not be definitively linked to any of the referred mung bean genomes, while the remaining 1595 SNPs were evenly scattered across the whole "Mung bean chromosome". Chromosomes 5 and 7 contained the most SNPs (203 SNPs), followed by Chromosome 8 (199 SNPs), Chromosome 1 (158 SNPs), Chromosome 6 (148 SNPs), Chromosome 11 (134 SNPs), Chromosome 10 (132 SNPs), Chromosome 3 (120 SNPs), Chromosome 2 (119 SNPs), and Chromosome 9 (98 SNPs). Approximately 5% of markers were mapped to scaffolds that are still unrelated to a chromosome, making up a tiny portion of the total. The (-log p> 5.5, a= 1) threshold was taken as the cutoff point for significant detection of the marker-trait relationship, as shown in the first Quantile-Quantile figure (QQ plot) (Figure 5A). The different colors (blue, red, and pink) were observed to be aligned over this threshold, and traits associated with these colors were identified to be STI Kano 16, STI Kano 16 & 17, and STI Ibadan 16. Delineating the QQ plot on a location basis gave a clearer picture of correlation at the same threshold with the distinct red color observed for the Ibadan location and the colors red, green, and blue above the threshold relating the significance of associated trait for the individual years and

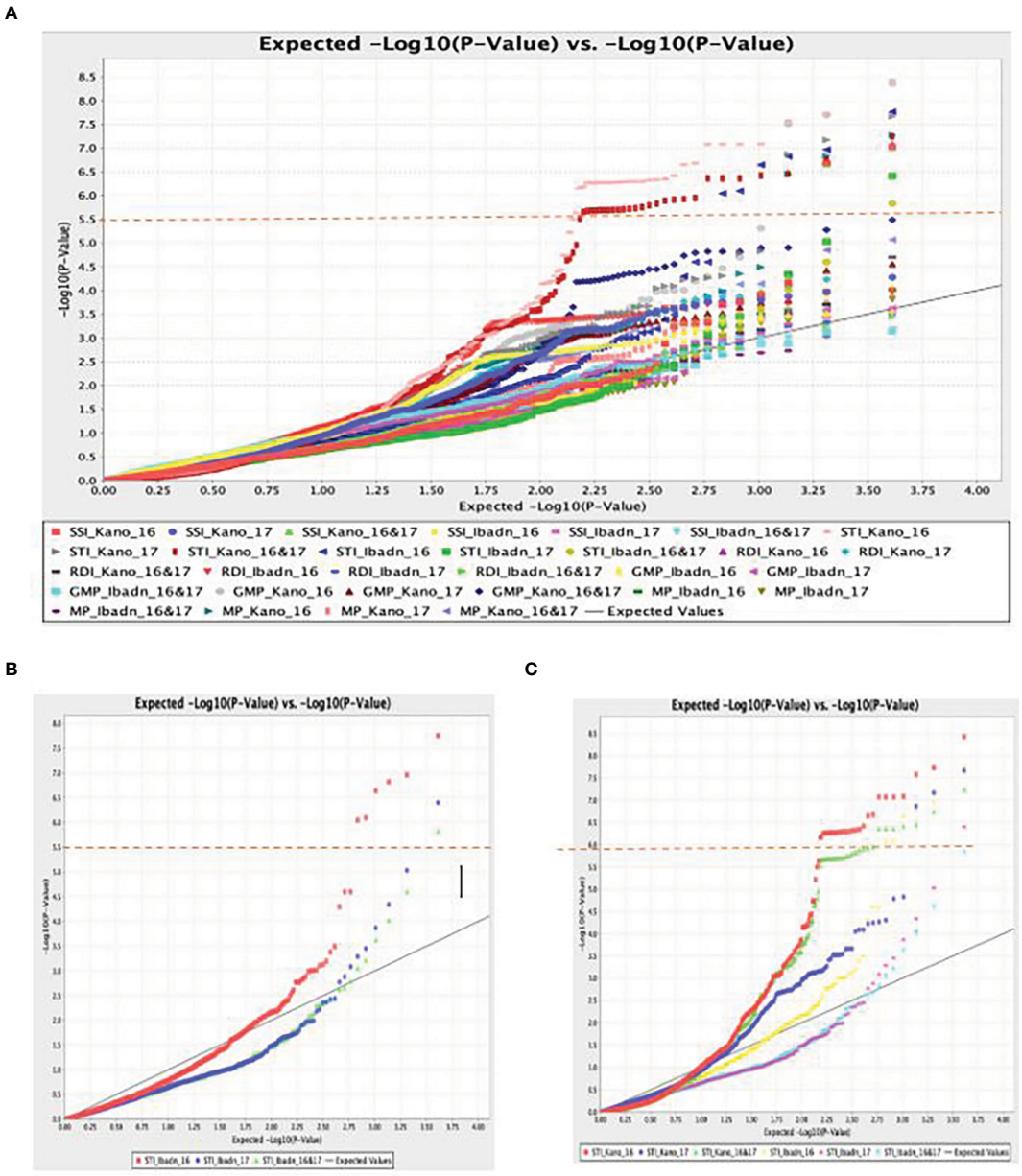


FIGURE 5
(A) Quantile–quantile (Q-Q) plot of a mixed linear model for Stress Tolerant Indices (STI) for drought studies in a panel of accessions of Bambara groundnut. **(B)** The p values of the SNPs and Quantile–quantile (Q-Q) plot of p values for stress-tolerant indices under the two water regimes in Ibadan. **(C)** The p values of the SNPs and Quantile–quantile (Q-Q) plot of p values for stress-tolerant indices under the two water regimes in Kano for the two years.

the combination of the years for the two locations of Ibadan and Kano (Figures 5B, C). The total 5,927 SNPs identified thus fulfilled these conditions, and as such, this number of SNPs was used to perform the association analysis with the stress tolerance indices. This significant marker was generated using the mixed linear model (MLM) association analysis and it was only detected for Stress Tolerance Index (STI), Geometric mean productivity (GMP), Relative Drought Index, and Mean Productivity (MP). No significant markers were detected for the other drought indices, namely, Tolerance index (TOL), Stress Susceptibility Index, Drought Index (DI), Yield index, Yield stability index (YSI), Modified stress tolerance index (k1STI and k2STI), and Susceptibility percentage index (SSPI) under the different water regimes for the two years of planting and the two locations generated data. The significant association between the phenotypic traits and the SNPs marker is further displayed with the Manhattan plots (Figure 6). The Manhattan plot is showing scattered outliers that are situated on the upper part of the plot at a surpassed threshold of 0.5. The genome-wide significance threshold is indicated by the red line and the SNPs associated with the trait of interest are highlighted in purple dots above the threshold. The SNPs were significantly associated with the stress-tolerant indices for 2016, 2017, and a combination of the 2016 and 2017 years of planting of the accession in the Kano state. Eight significant SNPs were identified based on the mixed linear model (MLM) association analysis. Worthy of note was the fact that the identified SNPs were unique for the Kano location in association with the Stress Tolerance Indices. The markers range for the expression of the phenotypic variation (R2) was between 0.19% and 0.31%, which is within the typical P value for an SNP to be significant. The eight SNPs, thus, were significant given that all p values increased exponentially beyond the standard adjustment threshold of $p < 5 \times 10-8$ (Table 5).

Discussion

Identifying the key genetic underpinnings of complex features is very critical in organisms as it gives an accurate measure of its diversity with respect to any trait of interest. This study, thus, explored the association of the genomic region and the yield-related traits of a diverse population of Bambara groundnut for drought tolerance using the New Generation Sequencing (NGS) platform. The application of the NGS platform has been described to give in-depth knowledge of an organism's genome compared to the regular markers (AFLP, RAPD, ISSR, SRAP, and SSR) application. A wide range of performances within and across environments for most of the studied drought-related traits observed with high estimates of values as well as significant and positive correlations among the 100 accessions of Bambara groundnut was obtained. For an effective association between the phenotype and genotype traits of interest through the GWAS approach, the population size, trait of interest, marker density, and phenotypic evaluation were all considered to be some of the key factors to variations. The high population size gave allowance for the production of a higher number of markers and, thus, a strong verdict of the power of the QTLs. This study used 100 accessions and this agrees with what was observed on GWAS by Upadhyaya et al. (2013), Dang et al. (2016), and Zhou et al. (2017), who all documented a population range of 100-300 for an effective GWAS analysis. Although these large numbers of accessions can sometimes be hindered by factors, such as the cost of

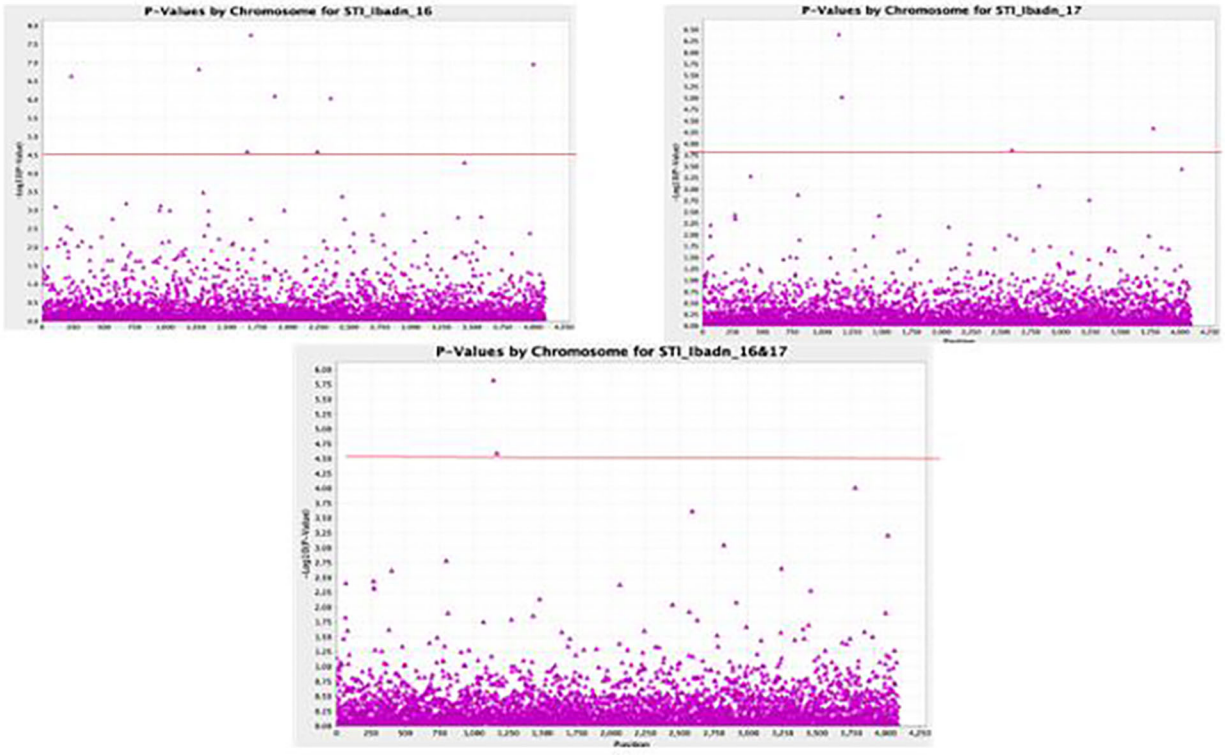


FIGURE 6
GWAS Manhattan Plot identifying the QTLs for Stress Tolerance Indices in Kano for two years.

analysis, availability of the platform for the analysis, as well as the plant species itself, and the specific traits of interest. Research has been conducted on the effective utilization of DarTseq for screening, which includes drought tolerance in wheat (Mwadzingeni et al., 2017), barley (Wójcik-Jagła et al., 2018), and rust resistance in *Pisum fulvum* (Barilli et al., 2018), as well as the comparative genome analysis in fish (Shams et al., 2019) amongst others.

To determine the most appropriate stress tolerance criteria, the overall grain yield was subjected to a correlation coefficient between yield in irrigated (Yp) and yield in stress (Ys) using the 11 quantitative indices of stress tolerance. This research identified STI and GMP as significant among the screened indices to be used in screening Bambara groundnut, as a positive response was observed with both indices. In addition to other indices, such as GMP and MP that were collectively interpreted as being suggestive of greater drought stress tolerance, Fernandez (1992) had suggested using STI as a measure to distinguish between genotypes with high yield and stress tolerance potentials. Similar studies on 14 soybeans genotype identified GMP and STI as correlating with both stress and non-stress yields, (Kargar et al. 2014). Furthermore, reports by Khodarahmpour et al. (2011) indicated the use of SSI, STI, and GMP indices to be more accurate criteria for selecting high yield and heat tolerance in some genotypes of maize. Research on maize, common bean, barley, and pea by Kristin et al. (1997); Khalili et al. (2004); Karami et al. (2006), and Souri et al. (2005) all proved to identify both STI and GMP were applicable and efficient to separate the different genotypes.

Given that Bambara is extremely cleistogamic, the examined accessions demonstrated significantly greater observed heterozygosity (Ho = 0.5) than what was expected (Ho = 0.01). This high heterogeneity observed among a few of the accessions (up to 20%) may serve as a breeding source for future purposes. Contrary to this, Molosiwa et al. (2015) observed a practically total deficiency of heterozygosity among the individual plants of Bambara groundnut landraces studied, while Odongo et al. (2015) observed an average value of 0.345, which were all lower than what was observed in this study. Blair et al. (2009) also identified a value of 0.64% among the 604 genotypes of common bean analyzed using a set of 36 SSR markers. Likewise, the review by Aliyu et al. (2016) affirmed less than 5% heterozygosity is usually predicted for the self-breeding plant. It is, thus, projected that the population size, possible markers number, and the species type could be key factors in explaining the huge differences observed in the rate of heterozygosity of the studied accessions and can, thus, be explored for subsequent breeding research.

TABLE 5 Summary of the eight diagnostic SNPs.

ID No	Trait	Marker	P-value	MarkerR2
8	STI_Kano_16		5.41E-07	0.34844
9	**STI_Kano_16&17**	**24346377|F|0-22:A>G-22:A>G**	**2.06E-06**	**0.30767**
10	STI_Kano_17		9.63E-04	0.15484
31	STI_Kano_16		7.11E-05	0.21889
32	**STI_Kano_16&17**	**24384105|F|0-56:A>G-56:A>G**	**2.57E-04**	**0.18418**
33	STI_Kano_17		3.59E-04	0.17897
63	STI_Kano_16		1.82E-05	0.25324
64	**STI_Kano_16&17**	**24385643|F|0-53:G>C-53:G>C**	**1.24E-04**	**0.20176**
65	STI_Kano_17		8.07E-05	0.21605
67	STI_Kano_16		1.84E-05	0.25299
68	**STI_Kano_16&17**	**24385696|F|0-43:A>G-43:A>G**	**1.25E-04**	**0.20175**
69	STI_Kano_17		8.91E-05	0.21353
136	STI_Kano_16		4.74E-07	0.35216
137	**STI_Kano_16&17**	**4177257|F|0-44:A>T-44:A>T**	**2.00E-06**	**0.30857**
138	STI_Kano_17		7.08E-04	0.16227
182	STI_Kano_16		4.70E-07	0.35242
183	**STI_Kano_16&17**	**4182070|F|0-66:G>A-66:G>A**	**1.77E-06**	**0.31191**
184	STI_Kano_17		7.44E-04	0.16107
201	STI_Kano_16		3.64E-05	0.23528
202	**STI_Kano_16&17**	**4183483|F|0-24:G>A-24:G>A**	**1.07E-04**	**0.2055**
203	STI_Kano_17		6.81E-04	0.1632
209	STI_Kano_16		5.35E-07	0.34985
210	**STI_Kano_16&17**	**4183904|F|0-11:C>T-11:C>T**	**2.22E-06**	**0.30628**
211	STI_Kano_17		7.06E-04	0.16268

The cluster analysis revealed a distinct pattern of relationships among the 100 accessions included in this study. This was brought about by the mutually exclusive grouping of similar descriptions into the same cluster to authenticate the similar pedigree based on common generated data. The SNPs presented thousands of markers that differentiated the accessions on the basis of geographical origin as it had usually been known with Bambara accessions. The South African accessions stood out distinctively from other accessions from other regions while the West African accessions were found clustering with Central African accessions as earlier evidenced by Olukolu et al. (2012), whose research observed that East Africa accessions did not form a unique cluster but rather clustered with accessions from West Africa. Further supporting the idea that there are multiple centers of diversity and/or domestication for various Bambara species, Aliyu et al. (2016) presented significant evidence for the existence of two primary subpopulations for Bambara groundnut that have developed as separate groups. Previous reports on diversity studies on Bambara had always observed that clustering of genotypes or landraces to be on a known geographical location similar to what was observed in this study, (Amadou et al., 2001; Massawe et al., 2002; Singrün and Schenkel, 2003; Ntundu et al., 2004). This, thus, reaffirmed the possible justification of West Africa being the center of origin of this crop.

The statistical method was another key determinant for the success of the GWAS (Zhang et al., 2015). The MLM is preferred to the GLM model as, by calculating the population structure and the unnecessary relations among the tested individuals, it allows for a massive reduction of false associations. According to the GWAS results of these 100 Bambara accessions, eight SNPs were generated and identified to be significantly associated only with stress-tolerant index (STI) using the MLM model and this is explained by over 80% of the phenotypic variation of the Bambara accessions. The minimum p value of the significant SNPs identified was with markers 24346377, 4177257, 4182070, and 4183904 at ($P \leq 1E\text{-}6$), while the remaining four markers were at an exponential p value of ($P \leq 1E\text{-}4$). Similar to what was observed in this research work, Martínez-Ballesta et al. (2015) explained their findings on SNPs that are associated with SSI and STI in relation to the percentage spikelet sterility and yield per plant in rice, which explains a range of 6 – 21% for the phenotypic variation. Edae et al. (2014) also reported one DArT marker on chromosome 4A in wheat and this explained 4% of the phenotypic variation for the Grain Yield-Stress Susceptible Index. Particularly for drought tolerance using DArT markers, Zhang et al. (2017) identified 16 SNP loci and eight candidate genes that were significantly associated while performing a GWAS in a panel of 66 canola accessions, while Schulz et al. (2016) used the GWAS platform to focus on anthocyanin and carotenoid content of petals from 96 diverse rose genotypes. Furthermore, previous studies on Bambara groundnut identified the nearest flanking markers from the pre-selected common marker set for internode length QTL in Bambara groundnut, Ho et al. (2017). All these loci, thus, present a region that could be considered for the development of markers for the subsequent selection of drought tolerance in Bambara groundnut.

Genome-wide association study (GWAS) methodologies applied in this study were able to identify relationships between the SNP markers and stress tolerance index as the phenotypes of interest for the association. The complexity of the trait of interest, known to be controlled by multiple genes, and the critical nature of GWAS facilitated effective and repeated field experiments for different years and different locations. Genome-wide association studies (GWAS) are, thus, valuable tools needed for identifying candidate genes, as well as the genetic loci that are responsible for the variations observed in a targeted quantitative trait.

Conclusions

In order to feed the continent's ever-growing population, underutilized crops, such as Bambara groundnut, need an urgent revolution to increase their productivity by utilizing modern technologies. Marker Trait Association is, therefore, a profound key in locating genomic areas linked to these phenotypic traits of breeding importance and, by doing so, can alleviate most of the production constraints. The genome-wide association study on Bambara groundnut was able to establish a strong meeting point between the phenotypic and genotypic traits. With this, it will make it easier to take more steps toward the future exploitation of the identified quantitative loci for marker-assisted selection, marker-assisted backcrossing, and other genomic and breeding initiatives. Future prospects can be improved by conducting an omics study on the identified drought-tolerant accessions to gain a thorough understanding of the genes that can be up- or down-regulated, providing clues as to potential genes that could be introduced for advantageous plant breeding processes for drought stress tolerance.

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/s.

Author contributions

KO, OAO, and MA conceived and designed the experiments; KO and RP performed the statistical analysis; KO wrote the manuscript and prepared the references; KO, OAO, MA, RP, and OJO, revised the manuscript. All authors contributed to the article and approved the submitted version.

Funding

The study was supported by the Genetic Resources center (GRC) of the Institute of Tropical Agriculture (IITA).

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

Ahmad, N. S. (2013). *Genetic analysis of plant morphology in bambara groundnut (Vigna subterranea (L.) verdc.)* (Doctoral dissertation, University of Nottingham).

Aliyu, S., Festo, M., and Sean, M. (2016). Genetic diversity and population structure of bambara groundnut (*Vigna subterranea*(L.) verdc.): Synopsis of the past two decades of analysis and implications for crop improvement programmes. *Genet. Resour. Crop Evol.* 63 (6), 925–943. doi: 10.1007/s10722-016-0406-z

Aliyu, S., Massawe, F., and Mayes, S. (2016). Genetic diversity and population structure of Bambara groundnut (Vigna subterranea (L.) Verdc.): synopsis of the past two decades of analysis and implications for crop improvement programmes. *Gen. Res. Crop Evol.*, 63 (6), 925–943.

Amadou, H. I., Bebeli, P. J., and Kaltsikes, P. J. (2001). Genetic diversity in bambara groundnut (*Vigna subterranea*L.) germplasm revealed by RAPD markers. *Genome* 44.6, 995–999. doi: 10.1139/g01-096

Barilli, E., Cobos, M. J., Carrillo, E., Kilian, A., Carling, J., and Rubiales, D. (2018). A high-density integrated DArTseq SNP-based genetic map of pisum fulvum and identification of QTLs controlling rust resistance. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00167

Blair, W. M., Diaz, M. L., Buendia, F. H., and Duque, M. C. (2009). Genetic diversity, seed size associations and population structure of common beans (*Phaseolus vulgaris* l.). *Theor. Appl. Genet.* 119, 955–972. doi: 10.1007/s00122-009-1064-8

Bamshaiye, O. M. Adegbola, J. A. Bamishaiye,, E. I. (2011). Bambara groundnut: an underutilized nut in Africa. *Advances in Agricultural Biotechnology* 1.1: 60–72

Bozhuyuk, M. R. (2022). Morphological and biochemical characterization of wild sour cherry (Prunus cerasus l.) germplasm. *Erwerbs-Obstbau* 64, 357–363. doi: 10.1007/s10341-022-00656-z

Dang, X., Liu, E., Liang, Y., Liu, Q., Breria, C. M., and Hong, D. (2016). QTL detection and elite alleles mining for stigma traits in oryza sativa by association mapping. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01188

Directorate Plant Production (2011). *Production guideline for bambara groundnuts* (Pretoria, South Africa: Department of Agriculture. Forestry and fisheries), p1–10.

Dogan, H., Ercisli, S., Temim, E., Hadziabulic, A., Tosun, M., Yilmaz, S. O., et al. (2014). Diversity of chemical content and biological activity in flower buds of a wide number of wild grown caper (Capparis ovate desf.) genotypes from Turkey. *C. R. Acad. Bulg. Sci.* 67, 1593–1600.

Doyle, J. J., and Doyle, J. L. (1987). A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem. Bull.*

Edae, E. A., Byrne, P. F., Haley, S. D., Lopes, M. S., and Reynolds, M. P. (2014). Genome-wide association mapping of yield and yield components of spring wheat under contrasting moisture regimes. *Theor. Appl. Genet.* 127 (4), 791–807. doi: 10.1007/s00122-013-2257-8

Feldman, A., Ho, W. K., Massawe, F., and Mayes, S. (2019). "Bambara groundnut is a climate-resilient crop: how could a drought-tolerant and nutritious legume improve community resilience in the face of climate change?," in *Sustainable solutions for food security*. Eds. A. Sarkar, S. Sensarma and G. vanLoon (Cham, Switzerland: Springer), pp 151–pp 167. doi: 10.1007/978-3-319-77878-5_8

Fernandez, G. C. J. (1992). "Effective selection criteria for assessing plant stress tolerance," in *Proceedings of symposium*. (Shanhua, Taiwan), 257–270.

Flint-Garcia, S., Thornsberry, J., and Buckler, E. (2003). Structure of linkage disequilibrium in plants. *Annu. Rev. Plant Biol.* 54, 357–374. doi: 10.1146/annurev.arplant.54.031902.134907

Grzebelus, D., Iorizzo, M., Senalik, D., Ellison, S., Cavagnaro, P., Macko-Podgorni, A., et al. (2014). Diversity, genetic mapping, and signatures of domestication in the carrot (*Daucus carota* l.) genome, as revealed by diversity arrays technology (DArT) markers. *Mol. Breed.* 33, 625–637. doi: 10.1007/s11032-013-9979-9

Ho, W. K., Chai, H. H., Kendabie, P., Ahmad, N. S., Jani, J., Massawe, F., et al. (2017). Integrating genetic maps in bambara groundnut [*Vigna subterranea*(L) verdc.] and their syntenic relationships among closely related legumes. *BMC Genomics* 18 (1), 192. doi: 10.1186/s12864-016-3393-8

Jha, U. C., Nayyar, H., Parida, S. K., Bakır, M., von Wettberg, E. J. B., and Siddique, K. H. M. (2022). Progress of genomics-driven approaches for sustaining underutilized legume crops in the post-genomic era. *Front. Genet.* 13. doi: 10.3389/fgene.2022.831656

Karami, A. A., Ghanadha, M. R., Naghavi, M. R., and Mardi, M. (2006). Identification of drought tolerance varieties in barley (Hordeum vulgare l.). Iran. *J. Crop Sci.* 37 (2), 371–379.

Kargar, S. M. A. Mostafaie, A. Hervan, E. M. Pourdad, S. S. (2014). Evaluation of soybean genotypes using drought stress tolerant indices. *International J. Agronomy and Agricultural Research* 5(2), 103–113.

Khalili, M., Kazemi, M., Moghaddam, A., and Shakiba, M. (2004). "Evaluation of drought tolerance indices at different growth stages of late-maturing corn genotypes," in *Proceedings 8th Iranian congress of crop science and breeding*. (Rasht, Iran) 25–27.

Khodarahmpour, Z., Choukan, R., Bihamta, M. R., and Majidi Harvan, E. (2011). Determination of the best heat stress tolerance indices in maize (Zea mays l.) inbred lines and hybrids under khuzestan province conditions. *J. Agric. Sci. Tech.* 13, 111–121.

Kilian, A., Wenzl, P., Huttner, E., Carling, J., Xia, L., Blois, H., et al. (2012). "Diversity arrays technology: A generic genome profiling technology on open platforms. *Methods Mol. Biol.* , 888, 67–89. doi: 10.1007/978-1-61779-870-2_5

Korte, A., and Farlow, A. (2013). The advantages and limitations of trait analysis with GWAS: a review. *Plant Methods* 9, 29. doi: 10.1186/1746-4811-9-29

Kristin, A. S., Serna, R. R., Perez, F. I., Enriquez, B. C., Gallegos, J. A. A., Vallejo, P. R., et al. (1997). Improving common bean performance under drought stress. *Crop Sci.* 37, 51–60. doi: 10.2135/cropsci1997.0011183X003700010007x

Letunic, and Bork,. (2019). Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Res.* 44 (W1), W242–W245.

Li, F., Chen, B., Xu, K., Gao, G., Yan, G., Qiao, J., et al (2016) A genome-wide association study of plant height and primary branch number in rapeseed (Brassi canapus). *Plant Sci* 242:169–177.

Mackay, I., and Powell, W. (2007). Methods for linkage disequilibrium mapping in crops. *Trends Plant Sci.* 12, 57–63. doi: 10.1016/j.tplants.2006.12.001

Martínez-Ballesta, M. C., López-Pérez, L., Muries, B., Muñoz-Azcarate, O., and Carvajal, M. (2015). "Climate change and plant water balance: the role of aquaporins–a review," in *Climate change, intercropping, pest control and beneficial microorganisms* (Dordrecht: Springer), 71–89. doi: 10.1007/978-90-481-2716-0_5

Massawe, F. J., Dickinson, M., Roberts, J. A., and Azam-Ali, S. N. (2002). Genetic diversity in bambara groundnut (*Vigna subterranea*(L.) verdc.) landraces revealed by AFLP markers. *Genome* 45, 1175–1180. doi: 10.1139/g02-093

Molosiwa, O. O., Aliyu, S., Stadler, F., Mayes, K., Massawe, F., Kilian, A., et al. (2015). SSR marker development, genetic diversity and population structure analysis of bambara groundnut [*Vigna subterranea*(L.) verdc.] landraces. *Genet. Resour. Crop Evol.* 62, 1225–1243. doi: 10.1007/s10722-015-0226-6

Mwadzingeni, L., Shimelis, H., Rees, D. J. G., and Tsilo, T. J. (2017). Genome-wide association analysis of agronomic traits in wheat under drought-stressed and non-stressed conditions. *PloS One* 12 (2), e0171692. doi: 10.1371/journal.pone.0171692

Nadeem, M. A., Habyarimana, E., Çiftçi, V., Nawaz, M. A., Karaköy, T., Comertpay, G., et al (2018). Characterization of genetic diversity in Turkish common bean gene pool using phenotypic and whole-genome DArTseq-generated silicoDArT marker information. *PLoS ONE* 13(10): e0205363.

Nordborg, M., and Weigel, D. (2008). Next-generation genetics in plants. *Nature* 456, 720–723. doi: 10.1371/journal.pone.0171692.s001

Ntundu, W. H., Bach, I. C., Christiansen, J. L., and Andersen, S. B. (2004). Analysis of genetic diversity in bambara groundnut [*Vigna subterranea*(L.) verdc] landraces using amplified fragment length polymorphism (AFLP) markers. *Afr. J. Biotechnol.* 3 (4), 220–225. eISSN: 1684-5315.

Odongo, F. O., Oyoo, M. E., Wasike, V., Owuoche, J. O., Karanja, L., and Korir, P. (2015). Genetic diversity of bambara groundnut (*Vigna subterranea*(L) verdc.) landraces in Kenya using microsatellite markers. *Afr. J. Biotechnol.* 14 (4), 283:291. doi: 10.5897/AJB2014.14082

Olukolu, B. A., Mayes, S., Stadler, F., Ng, N. Q., Fawole, I., Dominique, D., et al. (2012). Genetic diversity in bambara groundnut (*Vigna subterranea*(L.) verdc.) as revealed by phenotypic descriptors and DArT marker analysis. *Genet. Resour. Crop Evol.* 59, 347–358. doi: 10.1007/s10722-011-9686-5

Orgogozo, V., Morizot, B., and Martin, A. (2015). The differential view of genotype–phenotype relationships. *Front. Genet.* 6, 179. doi: 10.3389/fgene.2015.00179

Raman, H., Raman, R., Kilian, A., Detering, F., Carling, J., Coombes, N., et al. (2014). Genome-wide delineation of natural variation for pod shatter resistance in brassica napus. doi: 10.1371/journal.pone.0101673

Saran, P. L., Singh, S., Solanki, V., Choudhary, R., and Manivel, P. (2021). Evaluation of asparagus adscendens accessions for root yield and shatavarin IV content in India. *Turkish. J. Agric. Forestry.* 45 (4), 475–483. doi: 10.3906/tar-2006-42

Schulz, D. F., Schott, R. T., Voorrips, R. E., Smulders, M. J., Linde, M., and Debener, T. (2016). Genome-wide association analysis of the anthocyanin and carotenoid contents of rose petals. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.01798

Shams, F., Dyer, F., Thompson, R., Duncan, R. P., Thiem, J. D., Kilian, A., et al. (2019). Application of DArT seq derived SNP tags for comparative genome analysis in fishes; an alternative pipeline using sequence data from a non-traditional model species, macquaria ambigua. *PloS One* 14, 12. doi: 10.1371/journal.pone.0226365

Sinclair, T. R. (2011). Challenges in breeding for yield increase for drought. *Trends Plant Sci.* 16 (6), 289–293. doi: 10.1016/j.tplants.2011.02.008

Singrün, C., and Schenkel, W. (2003). "Fingerprinting of bambara groundnut germplasm with molecular markers," in *Proceedings of the international bambara groundnut symposium* (Botswana college of Agriculture, Botswana), 8–12.

Souri, J., Dehghani, H., and Sabaghpour, S. H. (2005). Study pea (Pisum sativum l.) genotypes in water stress condition. *Iran. J. Agric. Sci.* 36 (6), 1517–1527.

Thudi, M., Bohra, A., Nayak, S. N., Varghese, N., Shah, T. M., Penmetsa, R. V., et al. (2011). Novel SSR markers from BAC-end sequences, DArT arrays and a comprehensive genetic map with 1,291 marker loci for chickpea (*Cicer arietinum l.*). *PloS One* 6 (11), e27275. doi: 10.1371/journal.pone.0027275

Upadhyaya, H. D., Wang, Y. H., Gowda, C. L. L., and Sharma, S. (2013). Association mapping of maturity and plant height using SNP markers with the sorghum mini core collection. *Theor. Appl. Genet.* 126, 2003–2015. doi: 10.1007/s00122-013-2113-x

Valdisser, P. A. M. R., Pereira, W. J., Almeida Filho, J. E., Müller, B. S. F., Coelho, G. R. C., de Menezes, I. P. P., et al. (2017). In-depth genome characterization of a Brazilian common bean core collection using DArTseq high-density SNP genotyping. *BMC Genomics* 18 (1), 1–19. doi: 10.1186/s12864-017-3805-4

Vu, H. T. T., Kilian, A., James, A. T., Bielig, L. M., and Lawn, R. J. (2015). Use of DArT molecular markers for QTL analysis of drought-stress responses in soybean. II. marker identification and QTL analyses. *Crop Past. Sci.* 66, 817–830. doi: 10.1071/CP14304

Wang, H., Jia, J., Cai, Z., Duan, M., Jiang, Z., Xia, Q., et al. (2022). Identification of quantitative trait loci (QTLs) and candidate genes of seed iron and zinc content in soybean [Glycine max (L.) merr.]. *BMC Genomics* 23 (1), 1–14. doi: 10.1186/s12864-022-08313-1

Wójcik-Jagła, M., Fiust, A., Kościelniak, J., and Rapacz, M. (2018). Association mapping of drought tolerance-related traits in barley to complement a traditional biparental QTL mapping study. *Theor. Appl. Genet.* 131.1, 167–181. doi: 10.1007/s00122-017-2994-1

Yang, S. Y., Saxena, R. K., Kulwal, P. L., Ash, G. J., Dubey, A., Harper, J. D., et al. (2011). The first genetic map of pigeon pea based on diversity arrays technology (DArT) markers. *J. Genet.* 90, 103–109. doi: 10.1007/s12041-011-0050-5

Zhang, J., Mason, A. S., Wu, J., Liu, S., Zhang, X., Luo, T., et al. (2015). Identification of putative candidate genes for water stress tolerance in canola (Brassica napus). *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.01058

Zhang, J., Zhao, J., Xu, Y., Liang, J., Chang, P., Yan, F., et al. (2017). Genome-wide association mapping for tomato volatiles positively contributing to tomato flavor. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.01042

Zhou, H., Zeng, W., and Yan, H. (2017). *In vitro* induction of tetraploids in cassava variety 'Xinxuan 048' using colchicine. *Plant Cell Tiss. Organ Cult.* 128, 723–729. doi: 10.1007/s11240-016-1141-z

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 06 March 2023
DOI 10.3389/fpls.2023.1123655

OPEN ACCESS

EDITED BY
Omena Bernard Ojuederie,
Kings University, Nigeria

REVIEWED BY
Ravish Choudhary,
Indian Agricultural Research Institute (ICAR), India
Jacob Popoola,
Bowen University, Nigeria

*CORRESPONDENCE
Himabindu Kudapa
k.himabindu@cgiar.org
Mahalingam Govindaraj
m.govindaraj@cgiar.org

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

RECEIVED 14 December 2022
ACCEPTED 20 February 2023
PUBLISHED 06 March 2023

CITATION
Kudapa H, Barmukh R, Vemuri H, Gorthy S, Pinnamaneni R, Vetriventhan M, Srivastava RK, Joshi P, Habyarimana E, Gupta SK and Govindaraj M (2023) Genetic and genomic interventions in crop biofortification: Examples in millets. *Front. Plant Sci.* 14:1123655. doi: 10.3389/fpls.2023.1123655

Genetic and genomic interventions in crop biofortification: Examples in millets

Himabindu Kudapa[1]*, Rutwik Barmukh[1], Hindu Vemuri[2], Sunita Gorthy[1], Rajasekhar Pinnamaneni[3], Mani Vetriventhan[1], Rakesh K. Srivastava[1], Priyanka Joshi[1], Ephrem Habyarimana[1], S. K. Gupta[1] and Mahalingam Govindaraj[1,4]*

[1]International Crops Research Institute for the Semi-Arid Tropics, Patancheru, Telangana, India, [2]International Maize and Wheat Improvement Center (CIMMYT), Patancheru, Telangana, India, [3]Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India, [4]HarvestPlus Program, Alliance of Bioversity International and the International Center for Tropical Agriculture (CIAT), Cali, Colombia

Micronutrient malnutrition is a serious threat to the developing world's human population, which largely relies on a cereal-based diet that lacks diversity and micronutrients. Besides major cereals, millets represent the key sources of energy, protein, vitamins, and minerals for people residing in the dryland tropics and drought-prone areas of South Asia and sub-Saharan Africa. Millets serve as multi-purpose crops with several salient traits including tolerance to abiotic stresses, adaptation to diverse agro-ecologies, higher productivity in nutrient-poor soils, and rich nutritional characteristics. Considering the potential of millets in empowering smallholder farmers, adapting to changing climate, and transforming agrifood systems, the year 2023 has been declared by the United Nations as the International Year of Millets. In this review, we highlight recent genetic and genomic innovations that can be explored to enhance grain micronutrient density in millets. We summarize the advances made in high-throughput phenotyping to accurately measure grain micronutrient content in cereals. We shed light on genetic diversity in millet germplasm collections existing globally that can be exploited for developing nutrient-dense and high-yielding varieties to address food and nutritional security. Furthermore, we describe the progress made in the fields of genomics, proteomics, metabolomics, and phenomics with an emphasis on enhancing the grain nutritional content for designing competitive biofortified varieties for the future. Considering the close genetic-relatedness within cereals, upcoming research should focus on identifying the genetic and genomic basis of nutritional traits in millets and introgressing them into major cereals through integrated omics approaches. Recent breakthroughs in the genome editing toolbox would be crucial for mainstreaming biofortification in millets.

KEYWORDS

micronutrients, genetic resources, genomics-assisted breeding, omics, precision phenotyping, genome editing

Introduction

The current world population of 7.9 billion, increasing at an alarming rate of 1.05% annually, is anticipated to reach 10 billion by 2057 (https://www.worldometers.info/world-population/#pastfuture). To fill the hungry stomachs of the burgeoning human population, approximately 8–9 billion tons of food are produced worldwide (FAO, 2021). Despite this, global hunger affected more than 828 million individuals in 2021 and was associated predominantly with increased conflict, violence, and climate variability (FAO, 2021). Access to healthy and nutritious food will be challenging in the future, with malnutrition already influencing about one out of three people globally. The minerals that are essential for human health are classified into two major categories: (i) macronutrients that include sodium (Na), potassium (K), calcium (Ca), magnesium (Mg), sulfur (S), and phosphorus (P) and (ii) micronutrients including iron (Fe), copper (Cu), zinc (Zn), manganese (Mn), fluorine (F), iodine (I), and molybdenum (Mo), among others (Bouis and Welch, 2010). Notably, the distressingly enhancing population growth has led to more than three billion people worldwide suffering from "hidden hunger" or "micronutrient deficiencies," wherein they fail to obtain enough nutrients or micronutrients from the foods they eat to lead healthy and productive lives. This is particularly severe in the case of children, as they fail to develop to their full mental and physical potential (HarvestPlus, 2019).

The increasing prevalence of micronutrient malnutrition is likely to have major repercussions among the poor in developing countries (particularly South Asia and sub-Saharan Africa) who rely heavily on staple crops such as maize, wheat, and rice, and eat few micronutrient-rich foods such as fruits, vegetables, and animal and fish products. For instance, the population in developing nations have an average intake of Ca less than half of that by the population that resides in developed countries (Nordin, 2000). Furthermore, Zn deficiency affects 17% of the world's population, with the highest risk occurring in South Asia and sub-Saharan Africa, while Fe deficiency affects 32.9% of people worldwide (Kassebaum et al., 2014). In addition to vitamin A, I, Zn, and Fe deficiencies that are regarded as major health concerns globally according to The World Health Report (2000), the deficiencies of other nutrients such as folate are also placing human health at risk (Mayer et al., 2008).

To improve the nutritional status of the global human population, many intervention strategies were proposed to combat the ill effects of micronutrient malnourishment such as dietary diversification, mineral supplementation, food fortification, and biofortification (Cominelli et al., 2020; Ofori et al., 2022). For instance, dietary diversification provides diverse staple food to contribute macro- and micronutrients at the recommended dietary allowance (RDA) in a sustainable way (Lowe, 2021). Mineral supplementation pertains to the intake of extra nutrients *via* capsules, tablets, or syrups to increase the nutrient levels obtained through food consumption to contribute to meeting the RDA requirements in a short term (Tam et al., 2020). Food fortification involves adding particular nutrients to the food for increasing their nutritional content in order to aid consumers achieve the RDA for such nutrients with recurrent investments (Das et al., 2019). The aforementioned interventions have been less successful because of political, socio-economic, infrastructure-related, and technical constraints that are apparent in many developing countries. As a result, biofortification is a cost-effective, sustainable, and consumer friendly solution for meeting target levels of micronutrients such as Fe and Zn in human populations with one-time investment (Govindaraj et al., 2019; Kumar et al., 2019; Kiran et al., 2022). Biofortification is a crop-breeding process of enhancing the essential nutritional value of staple food crops as they grow, as opposed to adding nutrients while processing the edible parts into food products. This process is facilitated by conventional breeding and genetic improvement approaches that are based on natural genetic variation, modern selection methods, and the detection of novel genes and gene combinations associated with grain micronutrient content. For instance, HarvestPlus aims to develop nutritious crop varieties through biofortification to provide higher amounts of Zn, Fe, and Provitamin-A to vulnerable populations (children and women) *via* the staple food that they eat (www.harvestplus.org). To date, 290 biofortified varieties of 12 distinct staple crops (rice, wheat, maize, cassava, pearl millet, bean, sweet potato, lentil, cowpea, banana/plantain, sorghum, and Irish potato) have been released in more than 60 countries across the world (HarvestPlus, 2018). One example of successful biofortification is the vitamin A-rich orange-fleshed sweet potato that is grown in several African countries.

Taking into consideration the present scenario of worldwide nutritional insecurity, in this review, we attempt to highlight the role of cereals (particularly millets) in increasing nutritional security of the global population. Here, we refer "small-seeded grasses" including sorghum, pearl millet, finger millet, and other small millet crops by the term "millets." We provide an update on the emerging trends, high-throughput phenotyping approaches, and the potential of genetics- and genomics-based strategies to address hidden hunger within a limited time frame using nutrient-rich millets as model crops.

Relevance of millets in the human diet

The major cereal crops and millets that are grown on hundreds of millions of hectares worldwide (www.fao.org/faostat; Figure 1) are vital to diets, cultures, and economies around the world, especially in populous developing regions. The global demand for all cereals and millets is rising and offers critical entry points for improving nutrition (Shiferaw et al., 2011; Shewry and Hey, 2015). They contain enhanced levels of proteins, vitamins, minerals, and antioxidants, which offer nutritional superiority over other grain crops. These crops also possess high levels of low glycemic index non-starchy polysaccharides and dietary fibers, apart from their enhanced micro- and macronutrient content. In this section, we describe the composition of key nutrients and minerals in millets and their relevance in human diets.

An extraneous group of forage grasses that produce small-sized grains are referred to as "millets." Millets provide a rich source of

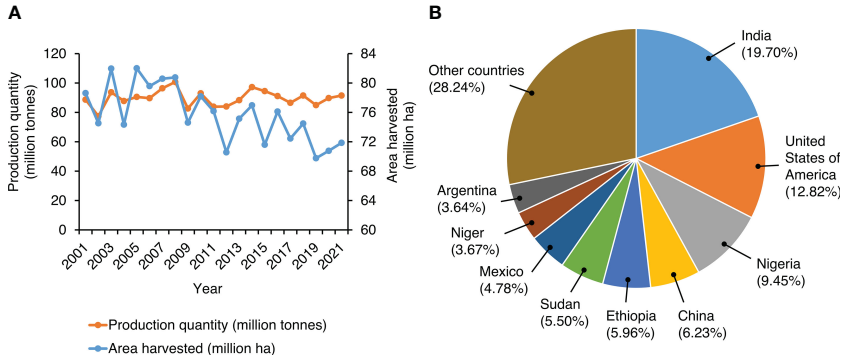


FIGURE 1
Global trends in millets production and cultivation. **(A)** Worldwide millet production in the last two decades. The graph represents production quantity (million tons) and the area under cultivation (million ha) from 2001 to 2021. Data represent the sum of values reported for millets and sorghum by FAOSTAT (2021). **(B)** Relative importance of the leading millet producing countries in the world. Data represent the sum of values reported for millets and sorghum production in 2021. *Source:* FAOSTAT (2021).

nutrition, show minimum vulnerability to pathogens, and are tolerant to abiotic factors including drought and salinity (Muthamilarasan et al., 2016; Govindaraj et al., 2020; Vetriventhan et al., 2020; Babele et al., 2022). As a result, millets serve as ideal staple crops for the semi-arid tropics of Asia and Africa. They are highly nutritious and superior to rice and wheat because they are rich in proteins, fibers, minerals, and vitamins (Saleh et al., 2013). Approximately 80% of millets production is utilized for human consumption, whereas the remaining is used for livestock feed and beer production (Saleh et al., 2013; Shivran, 2016). Millet grains are highly endorsed for the well-being of children, adolescent girls, lactating mothers, elderly, and convalescents.

Sorghum is grown on over 40.93 million hectares of land globally, producing 61.36 million tons, while in India, it is cultivated on approximately 4.24 million hectares and produces 4.81 million tons (FAOSTAT, 2021). Sorghum is one of the cheapest sources of nutrition, provides a high amount of energy, protein, Fe, and Zn, and contributes to >50% of the Fe and Zn requirement (Ashok Kumar et al., 2011; Ashok Kumar et al., 2013). In general, sorghum grain contains 79%–83% starch, 7%–14% protein, and 1%–7% fat (Rhodes et al., 2017). Sorghum is mostly used for food purposes (55%), consumed in the form of flat breads and porridges (thick or thin), and the vegetative part of the sorghum plant is used as stover for livestock, especially in drylands. The nutritional value and chemical composition of sorghum are similar to major cereals including rice, wheat, and maize. The energy value (per 100 g) of sorghum grain lies between 296 and 356 kcal and constitutes mainly polysaccharides such as starch and non-starch, proteins, and lipids (Table 1). Mkandawire et al. (2013) demonstrated that sorghum contains the lowest starch digestibility value when compared to other carbohydrate-rich crops. For instance, most of the starch particles of sorghum are gradually digestible (30.0%–66.2%), whereas others are quickly digestible (15.3%–26.6%) or resilient (16.7%–43.2%). The higher proportion of slowly digestible and resistant starch in sorghum has an added advantage, as it offers a low glycemic index and less threat of chronic diseases like type II diabetes and obesity (Simnadis et al., 2016). Non-starch polysaccharides in sorghum consist of 75%–90% insoluble fibers (mainly arabinoxylans) and 10%–25% soluble fibers (Martino et al., 2012). The lipid content of sorghum is low (1.24–3.07 g/100g) and contains approximately 83%–88% unsaturated fatty acids. Furthermore, the levels of mono-unsaturated fatty acids in several sorghum varieties are lower than those of polyunsaturated fatty acids, and linoleic (45.6%–51.1%), oleic (32.2%–42.0%), and palmitic (12.4%–16.0%) acids are some of the major fatty acids detected in sorghum grains (Afify et al., 2012). As a result, the natural variation available for protein, fat, and starch content in sorghum can be exploited to improve nutritional quality through crop improvement programs.

Other millets possess 10 genera and at least 14 species, of which pearl millet (*Pennisetum glaucum*) belongs to Paniceae tribe of the family Poaceae (Vetriventhan et al., 2020) and occupies 95% of the production (Nedumaran et al., 2014). Pearl millet contains high levels of micronutrients including Fe, Zn, and lysine (17–65 mg/g of protein) relative to other millet crops (McDonough et al., 2000; Haldimani et al., 2001). It is a good source of carbohydrates, resistant starch, proteins, dietary fibers, α-amylase activity, minerals, vitamins (A and B), and antioxidants, among others (Goswami et al., 2020) (Table 1). Increased levels of unsaturated fatty acids (75%) and phytic acid in pearl millet grains serve as a valuable resource for lowering cholesterol and phytate levels in individuals, which in turn decreases cancer risk. It contains high levels of antioxidants such as polyphenols, anthocyanins, tannins, phytates, and pinacosanols, which play a crucial role in regulating the aging process. Pearl millet is free of gluten and is suitable for consumption by people suffering from celiac disease who are usually allergic to the gluten from wheat and other cereals. Due to its nutritional superiority, pearl millet is beneficial for individuals suffering from diseases like diabetes, obesity, heart disorders, and atherosclerosis (Satyavathi et al., 2021). Furthermore, one of the key millet crops, finger millet (*Eleusine coracana* (L.) Gaertn.), serves as an important source of nutrition for people living in the developing world (Vetriventhan et al., 2016; Hittalmani et al., 2017). Finger millet is an important source of key nutrients including 18% dietary fiber, 6%–13% protein, 2.5%–3.5% minerals, 0.3%–3% phenolic compounds, and 0.34% calcium

TABLE 1 Nutritional status of millet grains compared with other cereal crops (per 100 g).

Nutrients	Sorghum	Millet	Rice	Wheat	Maize
Water	12.4 g	8.67 g	12.9 g	10.9 g	10.8 g
Energy	329 kcal	378 kcal	360 kcal	339 kcal	364 kcal
Protein	10.6 g	11 g	6.61 g	13.7 g	8.75 g
Total lipid (fat)	3.46 g	4.22 g	0.58 g	2.47 g	5.09 g
Ash	1.43 g	3.25 g	0.58 g	1.78 g	1.44 g
Carbohydrates	72.1 g	72.8 g	79.3 g	71.1 g	73.9 g
Dietary fiber	6.7 g	8.5 g	1.4 g	–	8.4 g
Sugars	2.53 g	–	–	–	–
Minerals					
Calcium, Ca	13 mg	8 mg	9 mg	34 mg	5 mg
Iron, Fe	3.36 mg	3.01 mg	4.36 mg	3.52 mg	1.74 mg
Magnesium, Mg	165 mg	114 mg	35 mg	144 mg	110 mg
Phosphorous, P	289 mg	285 mg	108 mg	508 mg	263 mg
Potassium, K	363 mg	195 mg	86 mg	431 mg	381 mg
Sodium, Na	2 mg	5 mg	1 mg	2 mg	5 mg
Zinc, Zn	1.67 mg	1.68 mg	1.16 mg	4.16 mg	2.24 mg
Copper, Cu	0.284 mg	0.75 mg	0.11 mg	0.553 mg	0.154 mg
Manganese, Mn	1.6 mg	1.63 mg	1.1 mg	3.01 mg	0.54 mg
Vitamins					
Thiamin (B1)	0.332 mg	0.421 mg	0.578 mg	0.419 mg	0.16 mg
Riboflavin (B2)	0.096 mg	0.29 mg	0.048 mg	0.121 mg	0.23 mg
Niacin (B3)	3.69 mg	4.72 mg	5.09 mg	6.74 mg	2.6 mg
Pantothenic acid (B5)	0.367 mg	0.848 mg	1.34 mg	0.935 mg	0.55 mg
Vitamin B6	0.443 mg	0.384 mg	0.145 mg	0.419 mg	0.47 mg
Folate (B9)	20 µg	85 µg	231 µg	43 µg	–
Vitamin B12	0 µg	0 µg	0 µg	0 µg	–
Vitamin C	0 mg	0 mg	0 mg	0 mg	–
Vitamin E	0.5 mg	0.05 mg	0.11 mg	–	–
Vitamin K	–	0.9 µg	–	–	–

Source: USDA FoodData Central Database (https://fdc.nal.usda.gov/fdc-app.html#/).

(Chandra et al., 2016). The crop is also valued for its health benefits like anti-diabetic, anti-tumorigenic, antioxidant, and antimicrobial properties (Nakarani et al., 2020; Kumar et al., 2021). Millets serve as a rich source of vitamins and trace elements that are essential for normal physiological functions in human. Among different millet crops, foxtail millet has maximum thiamine content (0.59 mg/100g), whereas proso millet is rich in riboflavin (0.28 mg/100g), which is very high compared to the riboflavin content of rice (0.04 mg/100g) and wheat (0.1 mg/100g). Furthermore, kodo millet contains high levels of iron (4.0 mg/100g), which is followed by finger millet (3.4 mg/100g), and foxtail millet (2.7 mg/100g) (Chandel et al., 2014). In addition, zinc levels were also found to be higher in foxtail millet (4 mg/100g), followed by barnyard millet (3.6 mg/100g) and finger millet (2 mg/100g). Taken together, these studies indicate that the micronutrient composition of millets are several fold higher as compared to the average micronutrient content in key non-millet cereals, thereby offering an inexpensive and sustainable solution to malnutrition.

Precision phenotyping for grain micronutrient content

Precision phenotyping for grain micronutrient content is the key for the success of millet biofortification programs. Phenotyping for micronutrient traits such as Fe and Zn is challenging, as they are accumulated in minute quantities in the grains. Various phenotyping methods are being used to measure Fe and Zn concentrations, which include simple staining procedures to complex analytical protocols. The use of Perl's Prussian blue for Fe and diphenyl thiocarbazone-based dithizone (DTZ) for Zn is a simple technique that gives a rough qualitative estimation of Fe and Zn in grain (Elango et al., 2021). Furthermore, elemental analysis techniques such as atomic absorption spectrometer (AAS), inductively coupled plasma–optical emission spectrometer (ICP-OES), micro-X-ray fluorescence spectroscopy (μ-XRF), secondary ion mass spectrometry (NanoSIMS) synchrotron X-ray, near-infrared reflectance spectrophotometer (NIRS), and fluorescence spectroscopy have been used to precisely measure grain micronutrient content (Rai et al., 2012; Shobana et al., 2013; Govindaraj et al., 2022). For instance, Govindaraj et al. (2016) employed X-ray fluorescence spectrometry (XRF) and ICP-OES methods for estimating Fe and Zn densities in diverse pearl millet genotypes and identified highly significant and positive correlations between these two advanced techniques for Fe (r = up to 0.97, $p<0.01$) and Zn (r = up to 0.98, $p<0.01$) contents. Furthermore, the application of high-resolution nanoscale secondary ion mass spectrometry (NanoSIMS) allows to qualitatively and quantitatively map the distribution of these micronutrients across the sorghum kernel components (Gaddameedi et al., 2022). Importantly, this knowledge will help reduce the micronutrient loss during seed and food processing.

In millets, the AAS, ICP-OES, and XRF methods are widely used for assessing grain Fe and Zn levels (Ashok Kumar et al., 2015; Pujar et al., 2020). In the case of finger millet, mineral and protein content were analytically determined using ICP-OES (Puranik et al., 2020). In addition, seed nitrogen content was determined using combustion, which was then followed by thermal conductivity using the Leco FP-528 Nitrogen/Protein Determinator (Puranik et al., 2020). Among all, XRF is a low-cost and high-throughput method for assessing grain Fe and Zn, and there is good correspondence between ICP-OES and XRF methods for assessing the grain Fe and Zn, but ICP-OES is more accurate. Hence, one can use XRF to discard the lines, segregate populations with low Fe and Zn, and validate all high Fe and Zn lines with ICP-OES method. Recently, Vetriventhan et al. (2021) reported the use of ICP-OES method to estimate Fe, Zn, and Ca content in 200 diverse little millet landraces. Of the total number of accessions evaluated, approximately 80% of accessions revealed consistent protein and Zn content, whereas less consistency was observed for Fe (64%) and Ca (30%) content. Notably, significant positive correlation (R^2 = 0.69–0.74, $p \leq 0.001$) was observed for trait-specific accessions possessing higher grain weight (10 accessions), grain yield (15), biomass (15), and having consistently higher grain nutrient levels (30) over a 2-year period. In addition, five promising accessions possessing higher grain yield and Ca content were also identified (Vetriventhan et al., 2021). After multi-location field evaluation, promising accessions with superior agronomic performance and higher grain nutrient levels can be released for commercial cultivation or used in millet improvement programs.

Genetic resources for nutritional traits discovery and utilization

Genetic variability acts as a raw material, and its utilization by plant breeders is a key step in biofortification programs. Genetically diverse accessions for micronutrient availability that are conserved in genebanks serve as a rich source of genetic resources for designing nutrient-dense and high-yielding crops for food and nutritional security. Understanding the extent of genetic variability for micronutrients in plant genetic resources along with underlying genetics of accumulation mechanisms is critical for the development of nutrient-rich varieties.

Sorghum is a highly diverse species with approximately 256,000 germplasm accessions conserved globally, and the genebank at ICRISAT conserves the largest collection of over 42,000 accessions originating from 92 countries. These accessions are well characterized for various morpho-physiological and agronomic traits (Ashok Kumar et al., 2013). The core collection (2,247 accessions) and mini core collection (242 accessions) representing the global sorghum collection conserved at the ICRISAT genebank have been developed (Grenier et al., 2001; Upadhyaya et al., 2009). On accessing the Fe and Zn concentrations of all the accessions belonging to the core collection, large variability was found for Fe and Zn content ranging from 26–60 mg kg^{-1} and 21–57 mg kg^{-1}, respectively (Ashok Kumar et al., 2012). In addition, when compared to the germplasm accessions, cultivars and breeding lines were found to possess less Fe and Zn levels (Reddy et al., 2005). The quantitative inheritance nature of these traits could be the possible reason for the observed variability in Fe and Zn concentration. On the other hand, the lower concentration of these two minerals in cultivars and breeding lines relative to landraces can be explained by the lack of breeding interest in these traits until the end of the second millennium. Based on the consumption level of sorghum (200 g) and estimated bioavailability of Fe (10%) and Zn (25%) in sorghum, it is necessary to enhance the Fe concentration by 30 mg kg^{-1} and zinc concentration by12 mg kg^{-1} over the base levels (30 mg kg^{-1} of Fe and 20 mg kg^{-1} of Zn) to meet the RDA requirements. In addition to the variability observed in grain Fe and Zn content, a huge variability was detected in grain phytate content, but not much variability in β-carotene content. On studying diverse sorghum cultivars (yellow endosperm lines, germplasm lines, high-protein digestible lines, high-lysine lines, and waxy lines), significant genetic differences were observed for Fe, Zn, and phytate concentrations and agronomic and grain quality traits (Reddy et al., 2005). Large genetic variability for grain Fe and Zn concentrations was also noted in sorghum hybrid parents (>500 B-lines and 100 R-lines) and 67 commercial hybrids (Ashok Kumar et al., 2009; Ashok

Kumar et al., 2012). Ng'uni et al. (2012) reported that grain Fe concentration ranged from 2.8 to 6.3 mg/100 g, and grain Zn content varied from 2.3 to 5.5 mg/100 g in South African sorghum cultivars. A large genotype × environment interaction was observed for Fe and Zn content in sorghum. That said, top-ranking genotypes excelled in most environments and years (Ashok Kumar et al., 2013). Hence, multi-location and multi-season evaluation is critical for the effective phenotyping for grain Fe and Zn concentrations to identify stable lines.

Many pearl millet germplasm collections are available worldwide, which includes the Pearl Millet inbred Germplasm Association Panel (PMiGAP) (Varshney et al., 2017). The PMiGAP has been developed at ICRISAT, India based on a core collection of 1,000 accessions, including landraces and cultivars originating from across three continents and representing genetic diversity within 27 countries (Sehgal et al., 2015; Srivastava et al., 2020). The PMiGAP has been completely re-sequenced with approximately 29 million genome-wide single nucleotide polymorphism (SNPs) (Srivastava et al., 2020). Besides this, the Iniari germplasm collection developed by ICRISAT includes landraces originating from West Africa, which possess high grain filling capability under terminal water deficit, big seeds, long panicles, and wide leaves (Ito et al., 1999). Taking advantage of cross-pollination driven by the protogynous flowers in pearl millet, breeders are in a continuous search for identifying nutritionally elite varieties suited for local environments. Recent studies have demonstrated high variability for grain Fe and Zn content in different breeding materials (Rai et al., 2014a; Rai et al., 2016; Upadhyaya et al., 2016). For instance, Rai et al. (2014a) screened seeds of parental progenies and restorer parent progenies for evaluating Fe and Zn variability using X-ray fluorescence spectroscopy. The mean Fe density of these progenies was found to be 5%–66% higher compared to the control cultivars. Furthermore, Rai et al. (2016) and Upadhyaya et al. (2016) used inductively coupled plasma atomic emission spectroscopy to scrutinize diverse germplasm accessions and landraces and observed substantial variability for Fe (51–121 mg kg^{-1}) and Zn (46–87 mg kg^{-1}). In finger millet, grain nutrients assessment of the core germplasm collection (622 accessions) revealed substantial variability for Fe (21.71–65.23 mg kg^{-1}), Zn (16.58–25.33 mg kg^{-1}), Ca (1.84–4.89 g kg^{-1}), and protein (6.00–11.09%) (Upadhyaya et al., 2011a). In another study, Puranik et al. (2020) also reported substantial variability in finger millet (190 accessions) for calcium (223.63–422.56 mg 100 g^{-1}), potassium (266.60–668.83 mg 100 g^{-1}), magnesium (106.36–179.99 mg 100 g^{-1}), and protein content (3.86–11.27% w/w). Similarly, grain nutrients assessment of a diverse set of germplasm including core collection conserved at the ICRISAT genebank revealed a significant variability for grain nutrients in foxtail millet [Fe (24.1–68.0 mg kg^{-1}), Zn (33.6–74.2 mg kg^{-1}), Ca (90–288 mg kg^{-1}), and protein (10.7–18.5%)] (Upadhyaya et al., 2011b), proso millet [Fe (41–73 mg kg^{-1}), Zn (26–47 mg kg^{-1}), Ca (91–241 mg kg^{-1}), and protein (11–19%)] (Vetriventhan and Upadhyaya, 2018), kodo millet [Fe (14.4–56.4 mg kg^{-1}), Zn (17.0–31.5 mg kg^{-1}), Ca (121–321 mg kg^{-1}), and protein (5.6%–11.3%)] (Vetriventhan and Upadhyaya, 2019), and little millet [Fe (17.6–58.0 mg kg^{-1}), Zn (19.4–39.5 mg kg^{-1}), Ca (92–390 mg kg^{-1}), and protein (6–15.6%)] (Vetriventhan et al., 2021). In the case of foxtail millet, a large variability was detected in the diversity panel (93 accessions) for potassium (1,477.9–3,365.5 ppm), magnesium (764.7–1,795.3 ppm), phosphorus (1,756.0–4,135.0 ppm), and sulfur content (703.0–1,908.8 ppm) (Jaiswal et al., 2019). Taken together, these studies reveal that continuation of breeding program with planned crosses will pave the way towards detecting nutrient-rich parental genotypes suitable for fluctuating environmental conditions.

Breeding approaches for enhancing micronutrient content in millets

Conventional breeding

The conventional breeding approach involves the selection and crossing of two parental genotypes possessing high micronutrient concentrations (e.g., Fe and Zn) to develop a hybrid expressing the desired trait(s). Natural genetic variation for target traits existing within the crop gene pool holds the key to successful crop genetic improvement through conventional breeding approaches.

Biofortified (Zn and Fe rich) sorghum lines (e.g., ICSR 14001 also known as "*Parbhani Shakti*") and hybrids (ICSH 14002, ICSA 661 × ICSR 196, ICSA 318 × ICSR 94, and ICSA 336 × IS 3760) have been jointly developed by ICRISAT-Vasantrao Naik Marathwada Krishi Vidyapeeeth (VNMKV) and released for commercial cultivation in India. Furthermore, two sorghum varieties, improved Deko (12KNICSV-188) and improved Zabuwa (12KNICSV-22), possessing high Fe and Zn content have been released in Nigeria (Okwuonu et al., 2021). The improved Deko variety has Fe content (126 ppm), which is three times higher than that obtained from local sorghum varieties (40 ppm) (ICRISAT, 2016). These improved varieties have the potential to boost the nutritional levels of malnourished populations, especially children, in Nigeria. These new varieties involved the crossing of local Nigerian germplasm with improved lines from ICRISAT (Mali). Similarly, in India, a biofortified (Fe and Zn rich) pearl millet variety (named *Dhanashakti*) and a hybrid ICMH 1201 (named Shakti-1201) were released by ICRISAT and HarvestPlus in 2014. In addition, two varieties, namely, ICMH 1202 (Nirmal-7) and ICMH 1301, are being evaluated under advanced varietal trials. Several well-adapted commercial pearl millet varieties and their progenies and hybrids with high grain Fe and Zn content have also been previously reported (Velu et al., 2007; Rai et al., 2012). However, breeding for higher levels of grain micronutrient content also depends on the existing environmental conditions, including soil mineral composition, which further complicates the process (Feil et al., 2005). As a result, the stability of phenotyping data should be verified through multi-environmental trials while breeding biofortified crops for future climates.

Integrated *omics* approaches

The conventional breeding approaches have been successful in the past decades in developing biofortified millet varieties (Rai et al., 2014b; Satyavathi et al., 2021). However, these approaches alone are not sufficient to keep pace with the future food and nutritional demands of the burgeoning population. The realization of biofortified sorghum and millet varieties for future climates can be significantly accelerated if assisted by recent innovations in omics approaches (Varshney et al., 2021a). Furthermore, a huge genetic variation for micronutrient concentration prevails among the sorghum and millet germplasm; however, it is yet to be characterized in detail. Evaluating the available germplasm accessions for micronutrient concentrations and detection of high-fidelity molecular markers through genome-wide association studies (GWAS) will facilitate the identification of quantitative trait loci (QTLs)/candidate genes/alleles regulating nutritional traits of interest. The identified QTLs/genes can then be introgressed into elite cultivars *via* genomics-assisted breeding or transgene-based methods to enhance the grain micronutrient content (Varshney et al., 2021b). The genetic relatedness of millets with other cereals, such as rice, wheat, and barley, further enables introgression of these nutritional traits into major cereal crops. The integration of different omics approaches such as genomics, transcriptomics, proteomics, metabolomics, and ionomics is crucial in designing biofortified crops for the future (Figure 2).

Genomics: QTL mapping and genome-wide association studies

Radical developments in genomics techniques have led to the development of various marker systems including mapped microsatellite or simple sequence repeat (SSR) markers, SNPs markers, insertion–deletion (InDel) markers, and more recently haplotype-based SNP markers. Linkage mapping and association mapping are the two most used tools for dissecting the genetics of complex nutritional traits in plants (Roorkiwal et al., 2022). Traditional linkage mapping/QTL mapping explores the recombination events and marker–trait associations (MTAs) in bi-parental segregating populations, such as F_2, doubled haploid (DH), and recombinant inbred lines (RILs). This method is very powerful in capturing major genes with larger effects and rare alleles (Kumar et al., 2016). On the other hand, GWAS explores functional variations within genetically diverse panels through linkage disequilibrium (LD) analysis, which is very efficient and effective for identifying new genomic regions (Yu and Buckler, 2006; Roorkiwal et al., 2022). Here, we highlight some key genomics efforts undertaken in millets to identify genomic regions associated with grain nutrient content (Table 2).

One of the major concerns in sorghum is the limited bioavailability of micronutrients in the grain due to complexity of polyphenols and tannins. For the identification of QTLs for grain Fe and Zn content, Phuke et al. (2017) developed a RIL population from the cross 296B × PVK 801, with parents contrasting for grain Fe and Zn content. This study showed a wide variation in grain Zn (10.2–58.7 mg kg^{-1}) and Fe (10.8–76.4 mg kg^{-1}) content existing within the RIL population. The grain Fe and Zn content was found to be negatively correlated with yield and positively correlated with 100-seed weight, suggesting a selection pressure of bolder seeds in biofortification programs. A significant G×E interaction was observed for both the micronutrients, with Fe being highly influenced by the environment compared to Zn (Phuke et al., 2017). Furthermore, a genetic map was constructed for 296B × PVK 801 RIL population with 2,088 markers (1148 DArTs, 927 DArTSeqs, and 13 SSRs), covering a distance of 1,355.52 cM with an average marker interval of 0.6 cM (Kotla et al., 2019). Here, 11 QTLs (individual environments) and 3 QTLs (across environments) for Fe and Zn were identified. A common genomic region was identified for ICRISAT environment, which explained a

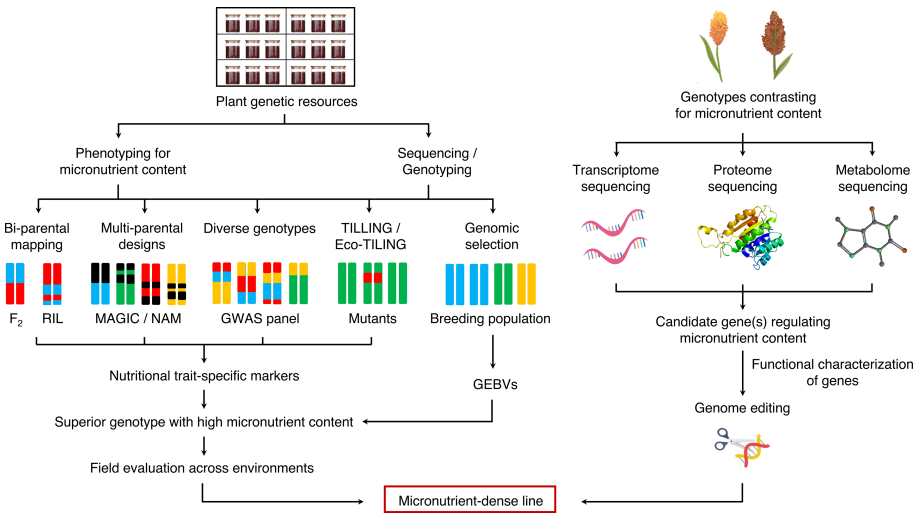


FIGURE 2
Genomic breeding strategies for enhancing nutritional content in millets. RIL, recombinant inbred line; MAGIC, multi-parent advanced generation inter-cross; NAM, nested association mapping; GWAS, genome-wide association study; TILLING, targeting induced local lesions in genomes; GEBV, genomic estimated breeding value.

TABLE 2 A list of some key QTLs identified for grain nutrient content in millets.

Crop	Trait	Mapping population	Population type	QTL(s)/MTAs/nearest marker	PVE (%)	Reference
Sorghum	Fe content	296B × PVK 801	RIL	*qfe1.1, qfe6.1, qfe7.1, qfe7.2*	5.09-6.80	Kotla et al., 2019
	Zn content	296B × PVK 801	RIL	*qzn7.1, qzn7.2, qzn7.3, qzn7.4*	5.70-9.42	Kotla et al., 2019
	Starch content	Rio × BTx623	RIL	7 QTLs	6.00-26.00	Murray et al., 2008
	Fat content	Rio × BTx623	RIL	3 QTLs	8.00-14.00	Murray et al., 2008
	Crude protein content	Rio × BTx623	RIL	7 QTLs	8.00-24.00	Murray et al., 2008
	Lutein content	KS115 × Macia	RIL	Lu-1.1, Lu-1.2, Lu-1.3, Lu-2.1, Lu-3.1, Lu-10b.1	7.13-28.50	Salas Fernandez et al., 2008
	Zeaxanthin content	KS115 × Macia	RIL	Ze-6.1, Ze-6.2	18.18-25.40	Salas Fernandez et al., 2008
	β-carotene content	KS115 × Macia	RIL	Bc-1.1, Bc-2.1, Bc-2.2, Bc-2.3, Bc-10b.1	8.00-15.15	Salas Fernandez et al., 2008
	Total carotenoids	KS115 × Macia	RIL	Tc-2.1, Tc-2.2	7.72-20.12	Salas Fernandez et al., 2008
	β-carotene content	–	Association mapping panel	14 MTAs	–	Cruet-Burgos et al., 2020
	Zeaxanthin content	–	Association mapping panel	38 MTAs	–	Cruet-Burgos et al., 2020
	Protein content	–	Association mapping panel	4 MTAs	34.00-36.00	Rhodes et al., 2017
	Fat content	–	Association mapping panel	41 MTAs	29.00-35.00	Rhodes et al., 2017
	Fe content	–	Association mapping panel	Sobic.001G213400	–	Shakoor et al., 2016
	Zn content	–	Association mapping panel	Sobic.007G064900	–	Shakoor et al., 2016
	Mn content	–	Association mapping panel	Sobic.003G349200	–	Shakoor et al., 2016
	Mg content	–	Association mapping panel	Sobic.001G443900	–	Shakoor et al., 2016
	Cd content	–	Association mapping panel	Sobic.002G083000	–	Shakoor et al., 2016
Pearl millet	Fe content	PPMI 683 × PPMI 627	RIL	14 QTLs	2.85-19.66	Singhal et al., 2021
	Zn content	PPMI 683 × PPMI 627	RIL	8 QTLs	2.93-25.95	Singhal et al., 2021
	Fe content	ICMB 841-P3 × 863B-P2	RIL	Xpsmp2214-Xipes142, Xpsmp322-Xipes181, pgpb11029-pgpb8456	18.10-19.40	Kumar et al., 2016
	Zn content	ICMB 841-P3 × 863B-P2	RIL	Xpsmp2214-Xipes142, Xpsmp2214-Xipes142, Xpsmp2040-pgpb10727	19.70-50.10	Kumar et al., 2016
	Fe content	ICMS 8511-S1-17-2-1-1-B-P03 × AIMP 92901-S1-183-2-2-B-08	RIL	11 QTLs	9.00-31.90	Kumar et al., 2018
	Zn content	ICMS 8511-S1-17-2-1-1-B-P03 × AIMP 92901-S1-183-2-2-B-08	RIL	8 QTLs	9.40-30.40	Kumar et al., 2018

(Continued)

TABLE 2 Continued

Crop	Trait	Mapping population	Population type	QTL(s)/MTAs/nearest marker	PVE (%)	Reference
	Fe content	–	Advanced inbred lines	18 MTAs	5.07-8.23	Pujar et al., 2020
	Zn content	–	Advanced inbred lines	43 MTAs	5.10-8.00	Pujar et al., 2020
	Protein content	–	Advanced inbred lines	17 MTAs	5.11-5.86	Pujar et al., 2020
	Fe content	–	Association mapping panel	21 MTAs	7.00-18.10	Anuradha et al., 2017
	Zn content	–	Association mapping panel	24 MTAs	8.20-16.50	Anuradha et al., 2017
Finger millet	Ca content	–	Association mapping panel	44 MTAs	4.80-17.79	Sharma et al., 2022
	Fe content[a]	–	Association mapping panel	895 MTAs	5.66-32.31	Puranik et al., 2020
	K content[a]	–	Association mapping panel	454 MTAs	5.78-19.89	Puranik et al., 2020
	Mg content[a]	–	Association mapping panel	5 MTAs	12.39-18.84	Puranik et al., 2020
	Na content[a]	–	Association mapping panel	643 MTAs	5.73-26.46	Puranik et al., 2020
	Zn content[a]	–	Association mapping panel	96 MTAs	7.55-17.07	Puranik et al., 2020
	Ca content	–	Association mapping panel	9 MTAs	7.90-41.00	Kumar et al., 2015
Foxtail millet	Fe content	–	Association mapping panel	1 MTA	–	Jaiswal et al., 2019
	Zn content	–	Association mapping panel	5 MTAs	–	Jaiswal et al., 2019
	Mg content	–	Association mapping panel	3 MTAs	–	Jaiswal et al., 2019
	B content	–	Association mapping panel	1 MTA	–	Jaiswal et al., 2019

[a]For finger millet, the number of MTAs indicate significant associations [$-\log_{10}(p) \geq 3.00$; $p \leq .001$ (for GLM) and $-\log_{10}(p) \geq 2.00$; $p \leq .01$ (for MLM), and an FDR < 0.1] reported by Puranik et al. (2020).

PVE, phenotypic variance explained; QTL, quantitative trait loci; MTA, marker trait association; Fe, iron; Zn, zinc; Mn, manganese; Mg, magnesium; Cd, cadmium; K, potassium; Na, sodium; Ca, calcium; B, boron.

phenotypic variance of 9.42% and 5.82% for Zn and Fe, respectively. After validation, the linked markers identified in this study can be deployed for developing high-grain Fe and Zn cultivars in sorghum improvement programs globally (Kotla et al., 2019). To unravel the genetic determinants of natural variation observed for seed element concentration, Shakoor et al. (2016) performed a GWA mapping of alleles regulating 20 traits influencing sorghum seed ionome, i.e., the mineral nutrient and trace element composition representing the inorganic component of cellular and organismal systems (Salt et al., 2008). This study also identified putative genes regulating Zn, Mn, Ni, Ca, and Cd accumulation in sorghum seeds.

Genetic variation observed among adapted pearl millet inbreds and hybrids suggests the possibility of improving grain micronutrient concentrations by selective breeding. In a previous study, Kumar et al. (2016) performed QTL mapping with 305 (96 SSRs; 208 DArT) markers distributed across seven linkage groups, covering a distance of 1,749 cM. Based on the phenotypic data collected across two different environments, QTLs for Fe and Zn content were found to co-localize on linkage group (LG) 3. QTL analysis using a pearl millet RIL population derived from PPMI 683 × PPMI 627 cross resulted in the identification of 14 QTLs for Fe and 8 QTLs for Zn with phenotypic variation ranging up to 19.66% and 25.95%, respectively (Singhal et al., 2021). These QTLs encompassed genes encoding ferritin, and Al^{3+}, K^{+}, Zn^{2+}, and Mg^{2+} transporters. Furthermore, Anuradha et al. (2017) performed GWAS using an association mapping panel comprising of 130 diverse lines of pearl millet, showing a wide range of grain micronutrient content. This study identified a total of 16 genomic regions for grain Fe and Zn content. Some regions were consistent across locations and years, while others were specific to a

particular location or year. Furthermore, a GWAS performed using 3 million SNPs generated using GBS resulted in the identification of several hundred significant MTAs for grain Fe and Zn content. This study also revealed six candidate genes linked with Fe/Zn uptake. The most significant candidate was found to be the YUCCA-11 gene, which is known to drive Zn efficiency by auxin biosynthesis (Manwaring, 2018). Furthermore, Pujar et al. (2020) performed GWAS analysis using a diverse panel of 281 advanced inbred lines to identify MTAs for grain Fe, Zn, and protein content. This study revealed 78 MTAs (including 18 MTAs for Fe, 43 MTAs for Zn, and 17 MTAs for protein content), and some promising candidate genes associated with grain Fe, Zn, and protein content in pearl millet. For finger millet, Puranik et al. (2020) identified 418 common MTAs associated with diverse mineral content using general linear model and mixed model approaches. Among these, 34 MTAs were above the Bonferroni threshold. From these 34 MTAs, 18 revealed homology with candidate genes involved in binding, remobilization, or metal ion transport (Puranik et al., 2020). After functional validation of these MTAs, these markers can be deployed in breeding efforts using genomic breeding approaches to develop high-grain nutritional quality in finger millet.

Transcriptomics

Recent advances in next-generation sequencing technologies have greatly revolutionized transcriptome sequencing. Being cost effective and with high coverage, transcriptomics has been performed in many crop species to identify candidate genes associated with nutrient biosynthesis and accumulation (Mishra et al., 2019). Additionally, RNA-sequencing (RNA-seq) provides information about the relative abundance of transcripts (at a given stage/condition) and enables molecular marker development in a high-throughput fashion. For instance, RNA-seq in grains of three sorghum cultivars differing in grain color led to the identification of >3,000 differentially expressed genes, which were mainly enriched in carbohydrates, amino acid, and flavonoid metabolism that may influence the grain nutritional content (Zhou et al., 2020). Furthermore, comparing the expression patterns of the genes (from an RNA-seq dataset) underlying a GWAS QTL led to the identification of a putative alpha-amylase 3 gene as a strong candidate associated with the variation in protein and fat content in sorghum grains (Rhodes et al., 2017). In a study done by Mahendrakar et al. (2020), in a set of contrasting mapping population parents for grain Fe and Zn content, diverse growth stages revealed tissue- and stage-specific expression patterns for a total of 29 Fe and Zn metabolism genes. Gene families including *PglZIP*, *PglNRAMP*, and *PglFER* were found to be candidates for grain Fe and Zn content, with ferritin-like gene, *PglFER1*, as the potential candidate gene for grain Fe content.

In a recent study, transcriptome profiling of stage-specific spikes of pearl millet genotypes contrasting for grain Fe and Zn content was performed to identify candidate genes expressed in developing spikes and those associated with the variation in Fe and Zn levels (Satyavathi et al., 2022). Here, 155 and 251 transcripts were found to be up- and downregulated, respectively, in the genotypes showing high Fe and Zn content, whereas 349 and 378 transcripts were differentially expressed during the flowering and milking stages of development, respectively. Gene Ontology analysis revealed that the genes involved in metabolic activities were primarily associated with uptake and transport of Zn and Fe in pearl millet (Satyavathi et al., 2022). Since finger millet grains contain higher levels of calcium content, RNA-seq of spike tissues of two finger millet genotypes contrasting for calcium content (GP-45, high calcium content; GP-1, low calcium content) was performed (Singh et al., 2014). A comparison of the relative abundance of transcripts revealed high expression of 24 calcium sensor genes (e.g., CaM, CaMLs, CBLs, CIPKs, and CRKs) in high calcium genotype. Collectively, these studies demonstrate the application of transcriptomics and encourage RNA-seq to be carried out in millets to comprehensively identify and functionally characterize the function of genes regulating nutritional content in these crops.

Proteomics and metabolomics

Proteomics and metabolomics represent some of the major players among omics approaches, as they are critical for characterizing the biomolecules and trace elements that have a high nutritional value. These approaches hold potential in identifying the protein function and in enhancing the production of crucial metabolic compounds in millets by offering key insights into the biological pathways. For example, in a recent study, non-targeted metabolomics analysis using a set of 61 diverse sorghum accessions enabled the differentiation of temperate and tropical sorghums based on the accumulation of phenolic acids, phytosterols, flavonoids, carotenoids, amino acids, sugars, and fatty acids (Ramalingam et al., 2021). This study offered new opportunities for generating biofortified sorghum varieties with enhanced nutraceutical and therapeutic characteristics. Although no major attempts have been made to analyze the nutrients at a protein and metabolite scale, these studies will pave the way for the identification of the proteins and metabolites in sorghum and/or millet seeds in response to several factors. Importantly, these approaches will facilitate the development of biomarkers for key nutritional traits in these crops. Taking into account the intricate and interrelatedness of physiological and metabolic pathways in crops, proteomics and metabolomics strategies will enable detection of putative genes having nutritive properties *via* a bottom–up approach. While the grains of millets are the major source of nutrition for people living in the developing countries, proteome and metabolome at diverse stages of seed development should be analyzed to detect candidate genes and their role in nutrient biosynthesis and accumulation.

Ionomics

With continuous technological innovations in genotyping and sequencing platforms, high-throughput phenotyping assumes higher importance for enhancing grain micronutrient content in different crops. To this end, ionomics has emerged as a high-throughput "elemental profiling" approach for accurately measuring mineral nutritional content of a living organism (Baxter, 2009; Huang and Salt, 2016). Precise and accurate

estimation of grain micronutrients is crucial for expediting the identification of genotypes possessing high micronutrient content (Swamy et al., 2016). A large number of elemental analysis techniques such as AAS, ICP-OES, μ-XRF, and NIRS have been used to correctly measure grain micronutrient levels (Trijatmiko et al., 2016; Manickavelu et al., 2017; Khokhar et al., 2018). In this context, community-oriented databases such as ionomic HUB or iHUB (http://www.ionomicshub.org/home/PiiMS) have been established to allow researchers to freely access ionomic resources of different plants including *Arabidopsis*, rice, and soybean (see Baxter et al., 2007). From the nutrition point of view, ionomics will serve as an effective approach to identify mineral transport mechanisms in millets by detecting the transporter genes and characterizing their molecular functions (Kumar et al., 2014). Although ionomic studies in crops are still in their preliminary stage, the role of ionome can be extrapolated to identify nutrient levels in grain crops. This information will in turn facilitate the development of biofortified crops and the required experimental work plan to enable efficient bioavailability of nutrients.

Transgene-based approaches

Transgenic approaches can be utilized to simultaneously integrate genes involved in improving micronutrient concentration and their bioavailability and reducing anti-nutrient concentration that limit the bioavailability of nutrients in plants. In addition, genetic engineering can also be used to redistribute micronutrients among tissues, enhance micronutrient levels in edible parts, and modulate biochemical pathways to increase grain micronutrient concentration in commercial crops (Sharma and Agarwal, 2005). Although the development of biofortified crops using genetic engineering is a time-consuming, labor-intensive, and costly effort during the research and development stage, eventually, it serves as a cost-effective and sustainable strategy, unlike conventional biofortification programs (White and Broadley, 2005). Transgenic crops with enhanced micronutrient contents hold the key to address hidden hunger especially among the smallholder farming households and vulnerable populations (women and young children) in the developing countries (Hirschi, 2009).

To date, there is no major study available for improvement of Fe and Zn by transgenic methods in millets. That said, sorghum has been targeted to improve provitamin A (beta-carotene) concentration by expressing Homo188-A (Lipkie et al., 2013). A biofortified sorghum with enhanced and stabilized pro-vitamin A providing 20%–90% of the estimated average requirement (EAR) for children under age 3 was developed recently using genetic transformation (Zhao et al., 2019). Moreover, it was demonstrated that provitamin A can be stabilized in sorghum by the co-expression of vitamin E through ectopic expression of homogentisate geranylgeranyltransferase (HGGT) and that vitamin E can enhance the stability of provitamin A *in planta*. These findings have the potential to impact directly the lives of millions of people who suffer from vitamin A deficiency, and they can be applied to enhance provitamin A stability in many food crops (Che et al., 2016). Similarly, the content of essential amino acid lysine was improved in sorghum by the introduction of a genetically engineered high lysine protein in the living sorghum cells (Zhao et al., 2003).

Genome editing

A recent development in the genome editing toolbox, Clustered Regulatory Interspaced Short Palindromic Repeats (CRISPR)/CRISPR-associated nuclease protein (Cas) systems for precise modification within the genome, gives researchers a possibility for accurately targeting the genes or genomic regions of interest (Zhu et al., 2020). CRISPR/Cas-mediated genome editing has been successfully used in major cereal crops; however, most millets have largely remained eluded from this success (Pixley et al., 2022). This is mainly because of the availability of limited genomic resources and lack of precise transformation systems developed in millets. A completely annotated whole-genome sequence is a major prerequisite for predicting the target genomic regions and designing guide RNA for the CRISPR/Cas system. Among millet crops, since a complete and annotated genome sequence is available only for foxtail millet to date (Vetriventhan et al., 2020), the off-target effects of genome-edited plants could not be studied precisely in other millets. Therefore, an effective application of genome editing systems in the future will rely mainly on the complete and annotated genomes of millets.

Notably, genome editing can aid in unravelling the mechanisms underlying nutrient fortification and transferring these key traits to major cereals. The development of nutrient-rich millet grains can also benefit from this method by tweaking the expression of genes involved in homeostasis and/or by editing the regulatory element of homeostasis genes. For instance, CRISPR/Cas9 approach is being used to target the k1C genes to create variants with reduced kafirin levels and improved protein quality and digestibility and improved lysine content in sorghum (Li et al., 2018). In addition, remarkably higher levels of mineral nutrients such as Ca (finger millet), Fe (barnyard millet), and vitamin B in millet seeds can be explored using genome editing tools to discover the transporters and signaling genes controlling seed biofortification. These traits can be transferred from millets to mainstream cereals to strengthen food sand nutritional security. In millets, biofortification is mainly limited due to the presence of antinutritional factors such as phytic acid, tannins, and polyphenols present in grains (Boncompagni et al., 2018). To this end, precision genome editing approaches like base editing can be used to reduce the quantity of antinutrients and enhance the bioavailability of macro- and micronutrients in millet grains. Taken together, the availability of high-quality reference genomes and efficient transformation systems could galvanize biofortification in millets using advanced genome editing tools, for sustainable food and nutritional security in the dryland tropics of South Asia and sub-Saharan Africa.

Conclusion and future perspectives

Millets have the nutritional potential for feeding vulnerable and poor populations in the semi-arid tropics of South Asia and sub-Saharan Africa. However, research efforts targeted towards exploring and utilizing nutrient-dense millet crops to address micronutrient malnutrition are still inadequate. Although millet grains possess significantly higher levels of essential amino acids, vitamins, and minerals, the bioavailability of nutrients in grains requires further research for improvement. It is now well known that biofortification serves as a promising and cost-effective approach to increase micronutrient content in food crops for enhancing the nutritional status of target populations (especially young children, adolescent girls, and women) across the world. Globally, committed efforts by HarvestPlus, ICRISAT, CGIAR, and other international and national initiatives are serving as pillars to achieve these goals. However, biofortification in millets is a challenging endeavor to make an impact in non-traditional areas and urban markets. Precision phenotyping for grain micronutrient content and characterizing diverse millet germplasm at the genotypic and phenotypic levels (for nutritional traits in combination with adaptive traits) would be useful for discovering novel sources of variation in nutritional traits. Recent innovations in next-generation sequencing technologies will enable the development of high-quality reference genomes for millets in a faster and more precise way. As a result, the identification of closely linked markers for nutritional traits will accelerate mainstreaming efforts to develop biofortified varieties and will also uncover the candidate genes controlling these traits. Integration of different omics approaches/genome editing with conventional biofortification programs to a larger extent can reflect in the rapid delivery of nutrient-dense millet varieties to address future food and nutritional security. Furthermore, since millets display cross-genera transferability, introgression of genomic regions/candidate genes associated with key nutritional traits from millets into mainstream cereals will be facilitated using genomics-assisted breeding or genome editing approaches. Considering their potential in addressing micronutrient malnutrition and hidden hunger among poor populations from developing countries, biofortified millets developed through genetic and genomic interventions hold a bright future.

Author contributions

HK and MG conceptualized the idea. HK, RB, HV, SG, and RP wrote the manuscript. HK, RB, VM, RS, PJ, EH, SG, SKG, and MG revised and finalized the manuscript. All authors contributed to the article and approved the submitted version.

Acknowledgments

HK acknowledges the Science and Engineering Research Board (Grant ID no. CRG/2019/005966), Department of Science and Technology (SERB-DST), Government of India, New Delhi, India, for the support.

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

Afify, A.E.-M.M.R., El-Beltagi, H. S., El-Salam, S.M.A.B.D., and Omran, A. A. (2012). Oil and fatty acid contents of white sorghum varieties under soaking, cooking, germination and fermentation processing for improving cereal quality. *Not. Bot. Horti Agrobot. Cluj-Napoca* 40, 86–92. doi: 10.15835/nbha4017585

Anuradha, N., Satyavathi, C. T., Bharadwaj, C., Nepolean, T., Sankar, S. M., Singh, S. P., et al. (2017). Deciphering genomic regions for high grain iron and zinc content using association mapping in pearl millet. *Front. Plant Sci.* 8, 412. doi: 10.3389/fpls.2017.00412

Ashok Kumar, A., Anuradha, K., Ramaiah, B., Grando, S., Frederick, H., Rattunde, W., et al. (2015). Recent advances in sorghum biofortification research. *Plant Breed. Rev.* 39, 89–124. doi: 10.1002/9781119107743.CH03

Ashok Kumar, A., Reddy, B. V. S., Ramaiah, B., Sahrawat, K. L., and Pfeiffer, W. H. (2012). Genetic variability and character association for grain iron and zinc contents in sorghum germplasm accessions and commercial cultivars. *Eur. J. Plant Sci. Biotechnol.* 6, 66–70.

Ashok Kumar, A., Reddy, B. V. S., Ramaiah, B., Sahrawat, K. L., Wolfgang, H., and Pfeiffer, W. H. (2013). Gene effects and heterosis for grain iron and zinc concentration in sorghum [*Sorghum bicolor* (L.) moench]. *Field Crops Res.* 146, 86–95. doi: 10.1016/j.fcr.2013.03.001

Ashok Kumar, A., Reddy, B. V. S., Ramaiah, B., Sanjana Reddy, P., Sahrawat, K. L., and Upadhyaya, H. D. (2009). Genetic variability and plant character association of grain Fe and Zn in selected core collection accessions of sorghum germplasm and breeding lines. *J. SAT Agric. Res.* 7, 1–4.

Ashok Kumar, A., Reddy, B. V. S., Sharma, H. C., Hash, C. T., Srinivasa Rao, P., Ramaiah, B., et al. (2011). Recent advances in sorghum genetic enhancement research at ICRISAT. *Am. J. Plant Sci.* 2, 589–600. doi: 10.4236/ajps.2011.24070

Babele, P. K., Kudapa, H., Singh, Y., Varshney, R. K., and Kumar, A. (2022). Mainstreaming orphan millets for advancing climate smart agriculture to secure nutrition and health. *Front. Plant Sci.* 13, 902536. doi: 10.3389/fpls.2022.902536

Baxter, I., Ouzzani, M., Orcun, S., Kennedy, B, Jandhyala, S.S., and Salt, D.E. (2007). Purdue ionomics information management system. An integrated functional genomics platform. *Plant Physiol.* 143, 600-611. doi: 10.1104/pp.106.092528

Baxter, I. (2009). Ionomics: studying the social network of mineral nutrients. *Curr. Opin. Plant Biol.* 12, 381–386. doi: 10.1016/j.pbi.2009.05.002

Boncompagni, E., Orozco-Arroyo, G., Cominelli, E., Gangashetty, P. I., Grando, S., Kwaku Zu, T. T., et al. (2018). Antinutritional factors in pearl millet grains: Phytate and goitrogens content variability and molecular characterization of genes involved in their pathways. *PloS One* 13, e0198394. doi: 10.1371/journal.pone.0198394

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

Chandel, G., Meena, R. K., Dubey, M., and Kumar, M. (2014). Nutritional properties of minor millets: neglected cereals with potentials to combat malnutrition. *Curr. Sci.* 107, 1109–1111. doi: 10.18520/CS/V107/I7/1109-1111

Chandra, D., Chandra, S., Pallavi,, and Sharma, A. K. (2016). Review of finger millet (*Eleusine coracana* (L.) gaertn): A power house of health benefiting nutrients. *Food Sci. Hum. Wellness* 5, 149–155. doi: 10.1016/j.fshw.2016.05.004

Che, P., Zhao, Z. Y., Glassman, K., Dolde, D., Hu, T. X., Jones, T. J., et al. (2016). Elevated vitamin E content improves all-trans β-carotene accumulation and stability in biofortified sorghum. *Proc. Natl. Acad. Sci. U.S.A.* 113, 11040–11045. doi: 10.1073/pnas.1605689113

Cominelli, E., Pilu, R., and Sparvoli, F. (2020). Phytic acid and mineral biofortification strategies: from plant science to breeding and biotechnological approaches. *Plants* 9, 553. doi: 10.3390/plants9050553

Cruet-Burgos, C., Cox, S., Ioerger, B. P., Perumal, R., Hu, Z., Herald, T. J., et al. (2020). Advancing provitamin a biofortification in sorghum: genome-wide association studies of grain carotenoids in global germplasm. *Plant Genome* 13, e20013. doi: 10.1002/tpg2.20013

Das, J., Salam, R., Mahmood, S., Moin, A., Kumar, R., Mukhtar, K., et al. (2019). Food fortification with multiple micronutrients: impact on health outcomes in general population. *Cochrane Database Syst. Rev.* 12, CD011400. doi: 10.1002/14651858.CD011400.pub2

Elango, D., Kanatti, A., Wang, W., Devi, A. R., Ramachandran, M., and Jabeen, A. (2021). Analytical methods for iron and zinc quantification in plant samples. *Commun. Plant Sci. Soil Anal.* 52, 1069–1075. doi: 10.1080/00103624.2021.1872608

FAO (2021). *World food and agriculture - statistical yearbook 2021* (Rome: Food and Agriculture Organization of the United Nations). doi: 10.4060/cb4477en

FAOSTAT (2021) *Crops and livestock products*. Available at: http://www.fao.org/faostat/en/#data/QCL (Accessed 01 February, 2022).

Feil, B., Moser, S., Jampatong, S., and Stamp, P. (2005). Mineral composition of the grains of tropical maize varieties as affected by pre-anthesis drought and rate of nitrogen fertilization. *Crop Sci.* 45, 516–523. doi: 10.2135/cropsci2005.0516

Gaddameedi, A., Sheraz, S., Kumar, A., Li, K., Pellny, T., Gupta, R., et al. (2022). The location of iron and zinc in grain of conventional and biofortified lines of sorghum. *J. Cereal Sci.* 107, 103531. doi: 10.1016/j.jcs.2022.103531

Goswami, S., Asrani, P., Ansheef Ali, T. P., Kumar, R. D., Vinutha, T., Veda, K., et al. (2020). Rancidity matrix: development of biochemical indicators for analysing the keeping quality of pearl millet flour. *Food Anal. Meth.* 13, 2147–2164. doi: 10.1007/s12161-020-01831-2

Govindaraj, M., Kanatti, A., Rai, K.N., Pfeiffer, W.H., and Shivade, H. (2022). Association of grain iron and zinc content with other nutrients in pearl millet germplasm, breeding lines, and hybrids. *Front. Nutr.* 8, 1357. doi: 10.3389/fnut.2021.746625

Govindaraj, M., Rai, K. N., Kanatti, A., Rao, A. S., and Shivade, H. (2019). Nutritional security in drylands: Fast-track intra-population genetic improvement for grain iron and zinc densities in pearl millet. *Front. Nutr.* 6, 74. doi: 10.3389/fnut.2019.00074

Govindaraj, M., Rai, K. N., Kanatti, A., Upadhyaya, H. D., Shivade, H., and Rao, A. S. (2020). Exploring the genetic variability and diversity of pearl millet core collection germplasm for grain nutritional traits improvement. *Sci. Rep.* 10, 21177. doi: 10.1038/s41598-020-77818-0

Govindaraj, M., Rai, K. N., Pfeiffer, W. H., Kanatti, A., and Shivade, H. (2016). Energy-dispersive X-ray fluorescence spectrometry for cost-effective and rapid screening of pearl millet germplasm and breeding lines for grain iron and zinc density. *Commun. Soil Sci. Plant Anal.* 47, 2126–2134. doi: 10.1080/00103624.2016.1228938

Grenier, C., Hamon, P., and Bramel-Cox, P. J. (2001). Core collection of sorghum: II. comparison of three random sampling strategies. *Crop Sci.* 41, 241–246. doi: 10.2135/cropsci2001.411241x

Haldimani, N. A., Muralikrishna, G., Tharanathan, R. N., and Malleshi, N. G. (2001). Nature of carbohydrates and proteins in three pearl millet varieties varying in processing characteristics and kernel texture. *J. Cereal Sci.* 33, 17–25. doi: 10.1006/jcrs.2000.0342

HarvestPlus (2018). *Catalyzing biofortified food systems: 2018 annual report* (Washington, DC: International Food Policy Research Institute). Available at: https://www.harvestplus.org/wp-content/uploads/2019/05/153613_HarvestPlus_AR_2018_Proof-1.pdf.

HarvestPlus (2019). *Getting biofortified food on everyone's plate: 2019 annual report* (Washington, DC: International Food Policy Research Institute). Available at: https://www.harvestplus.org/wp-content/uploads/2022/01/HarvestPlus-2019-Annual-Report.pdf.

Hittalmani, S., Mahesh, H., Shirke, M. D., Biradar, H., Uday, G., Aruna, Y., et al. (2017). Genome and transcriptome sequence of finger millet (*Eleusine coracana* (L.) gaertn.) provides insights into drought tolerance and nutraceutical properties. *BMC Genomics* 18, 465. doi: 10.1186/s12864-017-3850-z

Hirschi, K.D. (2009). Nutrient biofortification of food crops. *Annu. Rev. Nutr.* 29, 401-421. doi: 10.1146/annurev-nutr-080508-141143

Huang, X. Y., and Salt, D. E. (2016). Plant ionomics: From elemental profiling to environmental adaptation. *Mol. Plant* 9, 787–797. doi: 10.1016/j.molp.2016.05.003

ICRISAT (2016). Available at: https://www.icrisat.org/new-sorghum-varieties-will-fight-malnutrition-and-climate-change-in-sudan-and-sahel-region-of-nigeria/.

Ito, O., O'Toole, J. C., and Hardy, B. (1999). "Genetic improvement of rice for water-limited environments," in *Proceedings of the workshop on genetic improvement of rice for water-limited environments* (Los Baños: International Rice Research Institute), 353.

Jaiswal, V., Bandopadhyay, T., Gahlaut, V., Gupta, S., Dhaka, A., Ramchiary, N., et al. (2019). Genome-wide association study (GWAS) delineates genomic loci for ten nutritional elements in foxtail millet (*Setaria italica* l.). *J. Cereal Sci.* 85, 48–55. doi: 10.1016/j.jcs.2018.11.006

Kassebaum, N. J., Jasrasaria, R., Naghavi, M., Wulf, S. K., Johns, N., Lozano, R., et al. (2014). A systematic analysis of global anemia burden from 1990 to 2010. *Blood* 123, 615–624. doi: 10.1182/blood-2013-06-508325

Khokhar, J. S., Sareen, S., Tyagi, B. S., Singh, G., Wilson, L., King, I. P., et al. (2018). Variation in grain zn concentration, and the grain ionome, in field-grown Indian wheat. *PloS One* 13, e0192026. doi: 10.1371/journal.pone.0192026

Kiran, A., Wakeel, A., Mahmood, K., Mubaraka, R., Hafsa,, and Haefele, S. M. (2022). Biofortification of staple crops to alleviate human malnutrition: contributions and potential in developing countries. *Agronomy* 12, 452. doi: 10.3390/agronomy12020452

Kotla, A., Phuke, R., Hariprasanna, K., Mehtre, S., Rathore, A., Gorthy, S., et al. (2019). Identification of QTLs and candidate genes for high grain fe and zn concentration in sorghum [*Sorghum bicolor* (L.)Moench]. *J. Cereal Sci.* 90, 102850. doi: 10.1016/j.jcs.2019.102850

Kumar, A., Gaur, V. S., Goel, A., and Gupta, A. K. (2014). *De novo* assembly and characterization of developing spikes transcriptome of finger millet (*Eleusine coracana*): a minor crop having nutraceutical properties. *Plant Mol. Biol. Rep.* 33, 905–922. doi: 10.1007/s11105-014-0802-5

Kumar, S., Hash, C. T., Nepolean, T., Mahendrakar, M. D., Tara Satyavathi, C., Singh,, G., et al. (2018). Mapping grain iron and zinc content quantitative trait loci in an Iniadi-derived immortal population of pearl millet. *Genes (Basel).* 9, 248. doi: 10.3390/genes9050248

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, S., Palve, A., Joshi, C., Srivastava, R. K., and Rukhsar, (2019). Crop biofortification for iron (Fe), zinc (Zn) and vitamin a with transgenic approaches. *Heliyon* 5, e01914. doi: 10.1016/j.heliyon.2019.e01914

Kumar, A., Rani, M., Mani, S., Shah, P., Singh, D. B., Kudapa, H., et al. (2021). Nutritional significance and antioxidant-mediated antiaging effects of finger millet: Molecular insights and prospects. *Front. Sustain. Food Syst.* 5, 684318. doi: 10.3389/fsufs.2021.684318

Kumar, A., Yadav, S., Panwar, P., Gaur, V. S., and Sood, S. (2015). Identification of anchored simple sequence repeat markers associated with calcium content in finger millet (*Eleusine coracana*). *Proc. Natl. Acad. Sci. India B Biol. Sci.* 85, 311–317. doi: 10.1007/s40011-013-0296-1

Li, A., Jia, S., Yobi, A., Ge, Z., Sato, S. J., Zhang, C., et al. (2018). Editing of an alpha-kafirin gene family increases, digestibility and protein quality in sorghum. *Plant Physiol.* 177, 1425–1438. doi: 10.1104/pp.18.00200

Lipkie, T. E., De Moura, F. F., Zhao, Z.-U., Albertsen, M. C., Che, P., Glassman, K., et al. (2013). Bioaccessibility of carotenoids from transgenic provitamin a sorghum. *J. Agric. Food Chem.* 61, 5764–5771. doi: 10.1021/jf305361s

Lowe, N. M. (2021). The global challenge of hidden hunger: perspectives from the field. *Proc. Nutr. Soc* 80, 283–289. doi: 10.1017/S0029665121000902

Mahendrakar, M. D., Parveda, M., Kishor, P. B. K., and Srivastava, R. K. (2020). Discovery and validation of candidate genes for grain iron and zinc metabolism in pearl millet [*Pennisetum glaucum* (L.) r. br.]. *Sci. Rep.* 10, 16562. doi: 10.1038/s41598-020-73241-7

Manickavelu, A., Hattori, T., Yamaoka, S., Yoshimura, K., Kondou, Y., Onogi, A., et al. (2017). Genetic nature of elemental contents in wheat grains and its genomic prediction: Toward the effective use of wheat landraces from Afghanistan. *PloS One* 12, e0169416. doi: 10.1371/journal.pone.0169416

Manwaring, H. R. (2018). *Accessing and dissecting useful genetic variations for iron and zinc content in pearl millet (Pennisetum glaucum) using genome wide association studies* (Aberystwyth University, Aberystwyth, United Kingdom). Available at: https://pure.aber.ac.uk/portal/files/29240846/Manwaring_Hanna.pdf.

Martino, H. S. D., Tomaz, P. A., Moraes, E. A., Conceicao, L. L. D., Oliveira, D. D. S., Queiroz, V. A. V., et al. (2012). Chemical characterization and size distribution of sorghum genotypes for human consumption. *Rev. do Inst. Adolfo Lutz.* 71, 337–344.

Mayer, K. H., Bradford, J. B., Makadon, H. J., Stall, R., Goldhammer, H., and Landers, S. (2008). Sexual and gender minority health: what we know and what needs to be done. *Am. J. Public Health* 98, 989–995. doi: 10.2105/AJPH.2007.127811

McDonough, C. M., Rooney, L. W., and Serna-Saldivar, S. O. (2000). "The millets," in *Handbook of cereal science and technology (Food science and technology)*. Eds. K. Kulp and J. G. Ponte (New York, NY: Marcel Dekker Inc), 177–202.

Mishra, V. K., Gupta, S., Chand, R., Yadav, P. S., Singh, S. K., Joshi, A. K., et al. (2019). Comparative transcriptomic profiling of high- and low- grain zinc and iron containing Indian wheat genotypes. *Curr. Plant Biol.* 18, 100105. doi: 10.1016/j.cpb.2019.100105

Mkandawire, N. L., Kaufman, R. C., Bean, S. R., Weller, C. L., Jackson, D. S., and Rose, D. J. (2013). Effects of sorghum (*Sorghum bicolor* (L.) moench) tannins on α-amylase activity and *in vitro* digestibility of starch in raw and processed flours. *J. Agric. Food Chem.* 61, 4448–4454. doi: 10.1021/jf400464j

Murray, S. C., Sharma, A., Rooney, W. L., Klein, P. E., Mullet, J. E., Mitchell, S. E., et al. (2008). Genetic improvement of sorghum as a biofuel feedstock I: QTL for stem sugar and grain nonstructural carbohydrates. *Crop Sci.* 48, 2165–2179. doi: 10.2135/cropsci2008.01.0016

Muthamilarasan, M., Dhaka, A., Yadav, R., and Prasad, M. (2016). Exploration of millet models for developing nutrient rich graminaceous crops. *Plant Sci.* 242, 89–97. doi: 10.1016/j.plantsci.2015.08.023

Nakarani, U. M., Singh, D., Suthar, K. P., Karmakar, N., Faldu, P., and Patil, H. E. (2020). Nutritional and phytochemical profiling of nutracereal finger millet (*Eleusine coracana* l.) genotypes. *Food Chem.* 341, 128271. doi: 10.1016/j.foodchem.2020.128271

Nedumaran, S., Bantilan, M. C. S., Gupta, S. K., Irshad, A., and Davis, J. S. (2014). *Potential welfare benefit of millets improvement research at ICRISAT: Multi country-economic surplus model approach, socioeconomics discussion paper series number 15* (Hyderabad, India: ICRISAT).

Ng'uni, D., Geleta, M., Hofvander, P., Fatih, M., and Bryngelsson, T. (2012). Comparative genetic diversity and nutritional quality variation among some important southern African sorghum accessions [*Sorghum bicolor* (L.) moench]. *Aust. J. Crop Sci.* 6, 56–64.

Nordin, B. E. C. (2000). Calcium requirement is a sliding scale. *Am. J. Clin. Nutr.* 71, 1381–1383. doi: 10.1093/ajcn/71.6.1381

Ofori, K. F., Antoniello, S., English, M. M., and Aryee, A. N. A. (2022). Improving nutrition through biofortification–a systematic review. *Front. Nutr.* 9, 1043655. doi: 10.3389/fnut.2022.1043655

Okwuonu, I. C., Narayanan, N. N., Egesi, C. N., and Taylor, N. J. (2021). Opportunities and challenges for biofortification of cassava to address iron and zinc deficiency in Nigeria. *Glob. Food Sec.* 28, 100478. doi: 10.1016/j.gfs.2020.100478

Phuke, R. M., Anuradha, K., Radhika, K., Jabeen, F., Anuradha, G., Ramesh, T., et al. (2017). Genetic variability, genotype × environment interaction, correlation, and GGE biplot analysis for grain iron and zinc concentration and other agronomic traits in RIL population of sorghum (*Sorghum bicolor* l. moench). *Front. Plant Sci.* 8, 712. doi: 10.3389/fpls.2017.00712

Pixley, K. V., Falck-Zepeda, J. B., Paarlberg, R. L., Phillips, P. W. B., Slamet-Loedin, I. H., Dhugga, K. S., et al. (2022). Genome-edited crops for improved food security of smallholder farmers. *Nat. Genet.* 54, 364–367. doi: 10.1038/s41588-022-01046-7

Pujar, M., Govindaraj, M., Gangaprasad, S., Kanatti, A., and Shivade, H. (2020). Genetic variation and diversity for grain iron, zinc, protein and agronomic traits in advanced breeding lines of pearl millet [*Pennisetum glaucum* (L.) r. br.] for biofortification breeding. *Genet. Resour. Crop Evol.* 67, 2009–2022. doi: 10.1007/s10722-020-00956-x

Puranik, S., Sahu, P. P., Beynon, S., Srivastava, R. K., Sehgal, D., Ojulong, H., et al. (2020). Genome-wide association mapping and comparative genomics identifies genomic regions governing grain nutritional traits in finger millet (*Eleusine coracana* l. gaertn.). *Plants People Planet* 2, 649–662. doi: 10.1002/ppp3.10120

Rai, K. N., Govindaraj, M., and Rao, A. S. (2012). Genetic enhancement of grain iron and zinc content in pearl millet. *Qual. Assur. Saf. Crops Foods.* 4, 119–125. doi: 10.1111/j.1757-837X.2012.00135.x

Rai, K. N., Gupta, S. K., Sharma, R., Govindaraj, M., Rao, A. S., Shivade, H., et al. (2014a). Pearl millet breeding lines developed at ICRISAT: a reservoir of variability and useful source of non-target traits. *SAT eJournal* 1, 1–13.

Rai, K. N., Patil, H. T., Yadav, O. P., Govindaraj, M., Khairwal, I. S., Cherian, B., et al. (2014b). Dhanashakti: a high-iron pearl millet variety. *Ind. Farm.* 64, 32–34.

Rai, K. N., Yadav, O. P., Govindaraj, M., Pfeiffer, W. H., Yadav, H. P., Rajpurohit, B. S., et al. (2016). Grain iron and zinc densities in released and commercial cultivars of pearl millet (*Pennisetum glaucum*). *Indian J. Agric. Sci.* 86, 291–296.

Ramalingam, A. P., Mohanavel, W., Premnath, A., Muthurajan, R., Prasad, P. V. V., and Perumal, R. (2021). Large-Scale non-targeted metabolomics reveals antioxidant, nutraceutical and therapeutic potentials of sorghum. *Antioxidants* 10, 1511. doi: 10.3390/antiox10101511

Reddy, B. V. S., Ramesh, S., and Longvah, T. (2005). Prospects of breeding for micronutrients and b-carotene-dense sorghums. *Int. Sorghum Millets News.l.* 46, 10–14.

Rhodes, D. H., Hoffmann, L.Jr., Rooney, W. L., Herald, T. J., Bean, S., Boyles, R., et al. (2017). Genetic architecture of kernel composition in global sorghum germplasm. *BMC Genomics* 18, 15. doi: 10.1186/s12864-016-3403-x

Roorkiwal, M., Bhandari, A., Barmukh, R., Bajaj, P., Valluri, V. K., Chitikineni, A., et al. (2022). Genome-wide association mapping of nutritional traits for designing superior chickpea varieties. *Front. Plant Sci.* 13, 843911. doi: 10.3389/fpls.2022.843911

Salas Fernandez, M. G., Hamblin, M. T., Li, L., Rooney, W. L., Tuinstra, M. R., and Kresovich, S. (2008). Quantitative trait loci analysis of endosperm color and carotenoid content in sorghum grain. *Crop Sci.* 48, 1732–1743. doi: 10.2135/cropsci2007.12.0684

Saleh, A. S. M., Zhang, Q., Chen, J., and Shen, Q. (2013). Millet grains: Nutritional quality, processing, and potential health benefits. *Compr. Rev. Food Sci. Food Saf.* 12, 281–295. doi: 10.1111/1541-4337.12012

Salt, D. E., Baxter, I., and Lahner, B. (2008). Ionomics and the study of the plant ionome. *Annu. Rev. Plant Biol.* 59, 709–733. doi: 10.1146/annurev.arplant.59.032607.092942

Satyavathi, C. T., Ambawat, S., Khandelwal, V., and Srivastava, R. K. (2021). Pearl millet: A climate-resilient nutricereal for mitigating hidden hunger and provide nutritional security. *Front. Plant Sci.* 12, 659938. doi: 10.3389/fpls.2021.659938

Satyavathi, C. T., Tomar, R. S., Ambawat, S., Kheni, J., Padhiyar, S. M., Desai, H., et al. (2022). Stage specific comparative transcriptomic analysis to reveal gene networks regulating iron and zinc content in pearl millet [*Pennisetum glaucum* (L.) r. br.]. *Sci. Rep.* 12, 276. doi: 10.1038/s41598-021-04388-0

Sehgal, D., Skot, L., Singh, R., Srivastava, R. K., Das, S. P., Taunk, J., et al. (2015). Exploring potential of pearl millet germplasm association panel for association mapping of drought tolerance traits. *PloS One* 10, e0122165. doi: 10.1371/journal.pone.0122165

Shakoor, N., Ziegler, G., Dilkes, B. P., Brenton, Z., Boyles, R., Connolly, E. L., et al. (2016). Integration of experiments across diverse environments identifies the genetic determinants of variation in *Sorghum bicolor* seed element composition. *Plant Physiol.* 170, 1989–1998. doi: 10.1104/pp.15.01971

Sharma, R. K., and Agarwal, M. (2005). Biological effects of heavy metals: An overview. *J. Environ. Biol.* 26, 301–313.

Sharma, D., Tiwari, A., Sood, S., Meher, P. K., and Kumar, A. (2022). Identification and validation of candidate genes for high calcium content in finger millet [*Eleusine coracana* (L.) gaertn.] through genome-wide association study. *J. Cereal Sci.* 107, 103517. doi: 10.1016/j.jcs.2022.103517

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

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 Sec.* 3, 307. doi: 10.1007/s12571-011-0140-5

Shivran, A. C. (2016). "Biofortification for nutrient-rich millets," in *Biofortification of food crops*. Eds. U. Singh, C. S. Praharaj, S. S. Singh and N. P. Singh (New Delhi: Springer), 409–420. doi: 10.1007/978-81-322-2716-8_30

Shobana, S., Krishnaswamy, K., Sudha, V., Malleshi, N. G., Anjana, R. M., Palaniappan, L., et al. (2013). Finger millet (ragi, *Eleusine coracana* l.): A review of its nutritional properties, processing, and plausible health benefits. *Adv. Food Nutr. Res.* 69, 1–39. doi: 10.1016/B978-0-12-410540-9.00001-6

Simnadis, T. G., Tapsell, L. C., and Beck, E. J. (2016). Effect of sorghum consumption on health outcomes: a systematic review. *Nutr. Rev.* 74, 690–707. doi: 10.1093/nutrit/nuw036

Singh, U. M., Chandra, M., Shankhdhar, S. C., and Kumar, A. (2014). Transcriptome wide identification and validation of calcium sensor gene family in the developing spikes of finger millet genotypes for elucidating its role in grain calcium accumulation. *PloS One* 9, e103963. doi: 10.1371/journal.pone.0103963

Singhal, T., Satyavathi, C. T., Singh, S. P., Kumar, A., Sankar, S. M., Bhardwaj, C., et al. (2021). Multi-environment quantitative trait loci mapping for grain iron and zinc content using bi-parental recombinant inbred line mapping population in pearl millet. *Front. Plant Sci.* 12, 659789. doi: 10.3389/fpls.2021.659789

Srivastava, R. K., Singh, R. B., Pujarula, V. L., Bollam, S., Pusuluri, M., Chellapilla, T. S., et al. (2020). Genome-wide association studies and genomic selection in pearl millet: Advances and prospects. *Front. Genet.* 10, 1389. doi: 10.3389/fgene.2019.01389

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

Tam, E., Keats, E. C., Rind, F., Das, J. K., and Bhutta, Z. A. (2020). Micronutrient supplementation and fortification interventions on health and development outcomes among children under-five in low-and middle-income countries: a systematic review and meta-analysis. *Nutrients* 12, 289. doi: 10.3390/nu12020289

The World Health Report (2000). *Health systems: improving performance* (World Health Organisation, Geneva, Switzerland). Available at: https://www.who.int/publications/i/item/924156198X.

Trijatmiko, K., Dueñas, C., Tsakirpaloglou, N., Torrizo, L., Arines, F. M., Adeva, C., et al. (2016). Biofortified indica rice attains iron and zinc nutrition dietary targets in the field. *Sci. Rep.* 6, 19792. doi: 10.1038/srep19792

Upadhyaya, H. D., Pundir, R. P. S., Dwivedi, S. L., Gowda, C. L. L., Reddy, V. G., and Singh, S. (2009). Developing a mini core collection of sorghum for diversified utilization of germplasm. *Crop Sci.* 49, 1769–1780. doi: 10.2135/cropsci2009.01.0014

Upadhyaya, H. D., Ramesh, S., Sharma, S., Singh, S. K., Varshney, S. K., Sarma, N. D. R. K., et al. (2011a). Genetic diversity for grain nutrients contents in a core collection of finger millet (*Eleusine coracana* (L.) gaertn.) germplasm. *Field Crops Res.* 121, 42–52. doi: 10.1016/j.fcr.2010.11.017

Upadhyaya, H. D., Ravishankar, C. R., Narasimhudu, Y., Sarma, N. D. R. K., Singh, S. K., Varshney, S. K., et al. (2011b). Identification of trait-specific germplasm and developing a mini core collection for efficient use of foxtail millet genetic resources in crop improvement. *Field Crops Res.* 124, 459–467. doi: 10.1016/j.fcr.2011.08.004

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*. Eds. M. Singh and H. D. Upadhyaya (Academic Press: Elsevier, USA), 321–343.

Varshney, R. K., Bohra, A., Roorkiwal, M., Barmukh, R., Cowling, W. A., Chitikineni, A., et al. (2021a). Fast-forward breeding for a food-secure world. *Trends Genet.* 37, 1124–1136. doi: 10.1016/j.tig.2021.08.002

Varshney, R. K., Bohra, A., Yu, J., Graner, A., Zhang, Q., and Sorrells, M. E. (2021b). Designing future crops: Genomics-assisted breeding comes of age. *Trends Plant Sci.* 26, 631–649. doi: 10.1016/j.tplants.2021.03.010

Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35, 969–976. doi: 10.1038/nbt.3943

Velu, G., Rai, K.N., Muralidharan, V., Kulkarni, V.N., Longvah, T., and Raveendran, T.S.. (2007). Prospects of breeding biofortified pearl millet with high grain iron and zinc content. *Plant Breed.* 126, 182–185. doi: 10.1111/j.1439-0523.2007.01322.x

Vetriventhan, M., Azevedo, V. C. R., 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. *Nucl.* 63, 217–239. doi: 10.1007/s13237-020-00322-3

Vetriventhan, M., and Upadhyaya, H. 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, 451–463. doi: 10.1016/j.cj.2018.04.002

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, 1095–1106. doi: 10.2135/cropsci2018.07.0450

Vetriventhan, M., Upadhyaya, H. D., Azevedo, V. C., Allan, V., and Anitha, S. (2021). Variability and trait-specific accessions for grain yield and nutritional traits in germplasm of little millet (*Panicum sumatrense* roth. ex. roem. & schult.). *Crop Sci.* 61, 2658–2679. doi: 10.1002/csc2.20527

Vetriventhan, M., Upadhyaya, H. D., Dwivedi, S. L., Pattanashetti, S. K., and Singh, S. K. (2016). "Finger and foxtail millets," in *Genetic and genomic resources for grain cereals improvement*. Eds. M. Singh and H. D. Upadhyaya (Academic Press: Elsevier, USA), 291–319.

White, P., and Broadley, M. (2005). Biofortifying crops with essential mineral elements. *Trends Plant Sci.* 10, 586–593. doi: 10.1016/j.tplants.2005.10.001

Yu, J., and Buckler, E. S. (2006). Genetic association mapping and genome organization of maize. *Curr. Opin. Biotechnol.* 17, 155–160. doi: 10.1016/j.copbio.2006.02.003

Zhao, Z. Y., Che, P., Glassman, K., and Albertsen, M. (2019). Nutritionally enhanced sorghum for the arid and semiarid tropical areas of Africa. *Methods Mol. Biol.* 1931, 197–207. doi: 10.1007/978-1-4939-9039-9_14

Zhao, Z. Y., Glassman, K., Sewalt, V., Wang, N., Miller, M., Chang, S., et al. (2003). "Nutritionally improved transgenic sorghum," in *Plant biotechnology 2002 and beyond*. Ed. I. K. Vasil (Dordrecht: Kluwer Academic Publishers), 413–416.

Zhou, W., Zhou, Y., Wang, Z., Li, Y., Liu, H., Cui, F., et al. (2020). Comparative transcriptome analysis in three sorghum (*Sorghum bicolor*) cultivars reveal genomic basis of differential seed quality. *Plant Biosyst.* 156, 232–241. doi: 10.1080/11263504.2020.1851790

Zhu, H., Li, C., and Gao, C. (2020). Applications of CRISPR–cas in agriculture and plant biotechnology. *Nat. Rev. Mol. Cell Biol.* 21, 661–677. doi: 10.1038/s41580-020-00288-9

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 March 2023
DOI 10.3389/fpls.2023.1017652

OPEN ACCESS

EDITED BY
Ndiko Ndomelele Ludidi,
University of the Western Cape,
South Africa

REVIEWED BY
Rasha Ezzat Mahdy,
Assiut University, Egypt
Guanglong Hu,
Beijing Academy of Agricultural and
Forestry Sciences, China

*CORRESPONDENCE
Rohini Karunakaran
rohini@aimst.edu.my

SPECIALTY SECTION
This article was submitted to
Crop and Product Physiology,
a section of the journal
Frontiers in Plant Science

RECEIVED 12 August 2022
ACCEPTED 24 January 2023
PUBLISHED 08 March 2023

CITATION
Singh D, Lawrence K, Marker S,
Bhattacharjee I, Lawrence R, Choudhary R,
Ercisli S and Karunakaran R (2023) Rainfed
assessment of foxtail millet (*Setaria italica* L.
beauv) germplasms through genotyping
and principal component analysis.
Front. Plant Sci. 14:1017652.
doi: 10.3389/fpls.2023.1017652

Rainfed assessment of foxtail millet (*Setaria italica* L. beauv) germplasms through genotyping and principal component analysis

Divya Singh[1], Kapil Lawrence[1], Shailesh Marker[2], Indranil Bhattacharjee[2], Reena Lawrence[3], Ravish Choudhary[4], Sezai Ercisli[5] and Rohini Karunakaran[6,7,8]*

[1]Department of Biochemistry and Biochemical Engineering, Jacob Institute of Biotechnology and Bioengineering, Sam Higginbottom University of Agriculture, Technology and Sciences, Prayagraj, Uttar Pradesh, India, [2]Department of Genetics and Plant Breeding, Sam Higginbottom University of Agriculture, Technology and Sciences, Prayagraj, Uttar Pradesh, India, [3]Department of Chemistry, Sam Higginbottom University of Agriculture, Technology and Sciences, Prayagraj, Uttar Pradesh, India, [4]Division of Seed Science and Technology, ICAR-Indian Agricultural Research Institute, New Delhi, India, [5]Department of Horticulture, Faculty of Agriculture, Ataturk University, Erzurum, Türkiye, [6]Unit of Biochemistry, Faculty of Medicine, Semeling, Kedah, Malaysia, [7]Department of Computational Biology, Institute of Bioinformatics, Saveetha School of Engineering (SSE), SIMATS, Thandalam, Chennai, Tamil Nadu, India, [8]Centre of Excellence for Biomaterials Science, AIMST University, Semeling, Bedong, Malaysia

Introduction: Foxtail millet (*Setaria italica* L. beauv) is an important crop in underdeveloped countries; however, yield levels are low. The use of varied germplasm in a breeding approach is critical for increasing productivity. Foxtail millet can be cultivated effectively in a wide range of environmental circumstances but it is best suited to hot and dry climates.

Methods: In the current study, multivariant traits were used to define 50 genotypes in the first year and 10 genotypes in the second year. The phenotypic correlations among all traits in the entire germplasm were assessed, and the data acquired for all quantitative characters were subjected to analysis of variance for augmented block design. Furthermore, WINDOWS STAT statistical software was used to carry out a principal component analysis (PCA). The presence of substantial variations in most symptoms was shown by analysis of variance.

Results: Genotypic coefficient of variation (GCV) projections for grain yields were the highest, followed by panicle lengths and biological yields. Plant height and leaf length had the highest PCV estimates, followed by leaf width. Low GCV and phenotypic coefficient of variation (PCV) were measured as leaf length and 50% flowering in days. According to the PCV study, direct selection based on characters, panicle weight, test weight, and straw weight had a high and positive effect on grain yield per plant in both the rainy and summer seasons, indicating the true relationship between these characters and grain yield per plant, which aids indirect selection for these traits and thus improves grain yield per plant. Variability in foxtail millet germplasm enables plant breeders to effectively select appropriate donor lines for foxtail millet genetic improvement.

10.3389/fpls.2023.1017652

Discussion: Based on the average performance of genotypes considered superior in terms of grain yield components under Prayagraj agroclimatic conditions, the best five genotypes were: Kangni-7 (GS62), Kangni-1 (G5-14), Kangni-6 (GS-55), Kangni-5 (GS-389), and Kangni-4 (GS-368).

KEYWORDS

genotypic coefficient of variation (GCV), phenotypic coefficient of variation (PCV), genotypic path, phenotypic path, principal component analysis

1 Introduction

Plant genetic resources (PGR) are the backbone of the agricultural system, playing a positive and distinguishing role in the development of new cultivars from the past to the present, including the restructuring of existing ones (Sapkota et al., 2016). Genes for such traits are typically available in wild animals and landraces, allowing for the development of genotypes that can endure biotic and abiotic pressures. The current study concentrated on the genetic diversity of wild crop relatives. Genetic diversity, endangered plant species, species diversity, ecosystem stability, global floristic diversity in food plants, genetic resources in India, wild collections of major crops, plant genetic resources vis-à-vis crop breeding emphasis, and conservation of plant genetic resources are among the information needed to develop a breeding plan for sustainable agriculture: foxtail millet is a C4 crop that is diploid (2n = 18) (Mohammadi and Prasanna, 2003; Aghaee-Sarbarzeh and Amini, 2012). Foxtail millet cultivation is currently restricted to a few pockets, and in some locations it has been replaced by other crops that require irrigation. Its high nutritional value, combined with its low water requirement, makes it a climate-resilient crop appropriate for production in dryland agricultural systems. It has a tiny genome, and its use as a model crop for bioenergy has generated much interest, with more groups working on it than ever before. This troop's floral morphology and flowering behavior make it challenging to establish crosses between the desired parents. As a result, we have seen several published studies on creating strategies for crossing in foxtail millet to date. The experiment addresses floral biology, crossing procedures, and the generation of cytoplasmic male sterile (CMS) lines (Bhat et al., 2018).

The main component of foxtail millet grain is starch. Aside from grain, protein and fats are found in significant proportions. There are also some free sugar and non-starchy carbohydrates (CSE, 2007). Starch is widely used as a raw material in a variety of sectors, including textile, food, pharmaceutical, and paper. Native starch has relatively few industrial applications. Physical, chemical, or enzymatic processes can be used to create modified starches with specified qualities for a variety of uses (Kim et al., 2010). Owing to the rapid expansion of foxtail millet improvement in recent decades, as in other crops, foxtail millet landraces have been replaced by current cultivars, resulting in a significant loss of genetic diversity. As a result, established techniques of maintaining and multiplying foxtail millet landraces must be reconsidered. This could give germplasm conservationists and breeders some insight into the domestication, evolution, selection, and preservation of the world's oldest cereal crop (Ahmed et al., 2013). Foxtail millet is a promising source of micronutrients and protein compared with other cereals. Foxtail millet grain is (per 100 g) rich in protein (12.3%), iron (2.8 mg), and calcium (31 mg) compared with rice (7.9% protein and 1.8 mg iron) according to the Millet Network of India (MINI). Additionally, they contain a high quantity of beta-carotene and have a higher proportion of non-starchy polysaccharides and dietary fiber. Foxtail millet releases sugars very slowly and thus has a low glycemic index (GI) and hence can be used in a therapeutic diet but its potential role as low GI food has remained unrealized and unexploited. The low glycemic index diet has been shown to reduce blood glucose levels (Ahmed et al., 2013).

For selecting a new variety, GPB hybridization is one of the most efficient methods at present, with the ultimate goal of selecting a new variety (Islam, 2004). Appropriate parental line selection is the most important aspect of a dry lab experiment to improve the genetic recombination of potential breeds (Verma et al., 2018; Singh et al., 2019). Additionally, a vast number of morphologically documented germplasm studies are needed to determine the differences between all germplasm populations and their breeding potential. Breeders assessed a huge number of germplasm varieties, some of which may or may not have enough discriminatory power for germplasm selection, characterization, assessment, and management (Maji and Shaibu, 2012). If this is the case, then principal component analysis (PCA) can be used to determine parentage and reduce duplication in experimental data sets in which morphological and physiological variation occurs on a regular basis in GPB sciences (Singh and Verma, 2020). PCA is a multivariate statistical methodology that seeks to simplify and analyze the relationships between a large number of variants in terms of a relatively small number of variables or components while retaining all crucial information from the original genotype data set. (Amy and Pritts, 1991; Adams, 1995). However, these genotypes of foxtail millets have not been systemically determined so far; therefore, the current investigation provides a detailed overview of the rainfed assessment of *S. italica* genotypes through genotyping and principal component analysis.

2 Materials and methods

The experimental material consisted of accessions from 2018 and 2019, including 50 germplasm accessions of foxtail millet from 2018

10.3389/fpls.2023.1017652

and 10 from 2019. These 50 germplasm accessions were collected from ICRISAT and NBPGR, New Delhi during Kharif 2018. For evaluation and characterization, these 50 germplasm accessions and three check varieties were grown in a randomized complete block design (RBD) at the Field Experimental Centre, Department of Genetics and Plant Biotechnology, SHUATS, Prayagraj, India. The selection of the 10 best genotypes from 2019 was based on the yield of 50 foxtail millet germplasm accessions. The characterization site, Naini Prayagraj, is located at 13° 05' N latitude and 77° 34' E longitude. The Centre is 924 m above mean sea level. The annual rainfall ranges from 528 to 1374.4 mm with a mean of 915.8 mm. The germplasm accession was divided into three blocks, each consisting of 46 accessions and four check varieties (ISE375, ISE1468, ISE132, and ISE376). Each accession was grown in a single row 3 m in length and spaced 30 cm apart, and plant-to-plant spacing within the row was 10 cm. After 15 days, the crop was supplied with the recommended dose of fertilizer (10 kg N and 20 kg P-05 ha as a basal dose and 10 kg N at the time of earthing up). Irrigation was not provided and crops only received rainwater and were protected from weeds, pests, and diseases. For all characters, except days to emergence and days to maturity, observations were made for five randomly selected plants in each entry of each replication. The phenotypic correlation coefficients were obtained using the formula proposed by Vinsonias (2018).

The phenotypic correlations of all traits in the complete germplasm were estimated, and numerous significant correlations were found. Data for all quantitative characters were collected and subjected to analysis of variance for augmented block design using the method described by Kempthorne (1957). PCA was calculated for 15 quantitative traits to examine the relative value of various traits in capturing variation across the entire germplasm. The PCA was performed using WINDOWS STAT statistical software, as recommended by Fujita et al. (1996).

3 Results

Accessions showed variability among the quantitative and qualitative characters studied. The genetic parameters of 10 genotypes for 15 characters of foxtail millet were observed. Genotypic variance was high for plant height and low for leaf width. Phenotypic variance was at maximum for plant height and at minimum for leaf width. GCV was at maximum for economic yield and at minimum for conductivity temperature and depth (CTD). PCV was at maximum for economic yield and at minimum for CID. Heritability was at maximum for days of 50% flowering and at minimum for economic yield. Genetic advance (GA) was at maximum for plant height and at minimum for harvest index (Figures 1, 2). Ten accessions were used for calculations of genotypic and phenotypic coefficient of variation for different parameters in *S. Italica* (Singh et al., 2021). GCV and PCV ratios were highest for economic yield, while GCV and PCV ratios were very low for CTD. GCV and PCV ratios for all characters were as follows: days to 50% flowering, 6.50 and 7.03; days to 70% flowering, 5.68 and 5.88; plant height, 20.27 cm and 22.38 cm; leaf width, 10.10 cm and 13.06 cm; leaf length, 10.69 cm and 13.08 cm; pedicle length, 19.07 cm and 20.43 cm; panicle length, 23.97 cm and 25.25 cm; panicle weight, 32.77 g and 37.36 g; leaf area index, 25.23 and 28.86; stem girth, 8.42 cm and 10.31 cm; soil plant analysis development (SPAD), 7.01 and 10.38; CTD, 3.08 and 3.22; harvest index, 31.70% and 40.77%; biological yield, 32.55 g and 38.37 g; and grain yield, 47.24 g and 82.29g.

The heritability and GA ratio was high for the leaf area index and very low for economic yield and the harvest index. The heritability and GA ratios for all characters were as follows: days to 50% flowering, 96.26 and 7.29; days to 70% flowering, 93.31 and 7.29; plant height, 82.00 cm and 42.46 cm; leaf width, 59.86 cm and 0.22 cm; leaf length,

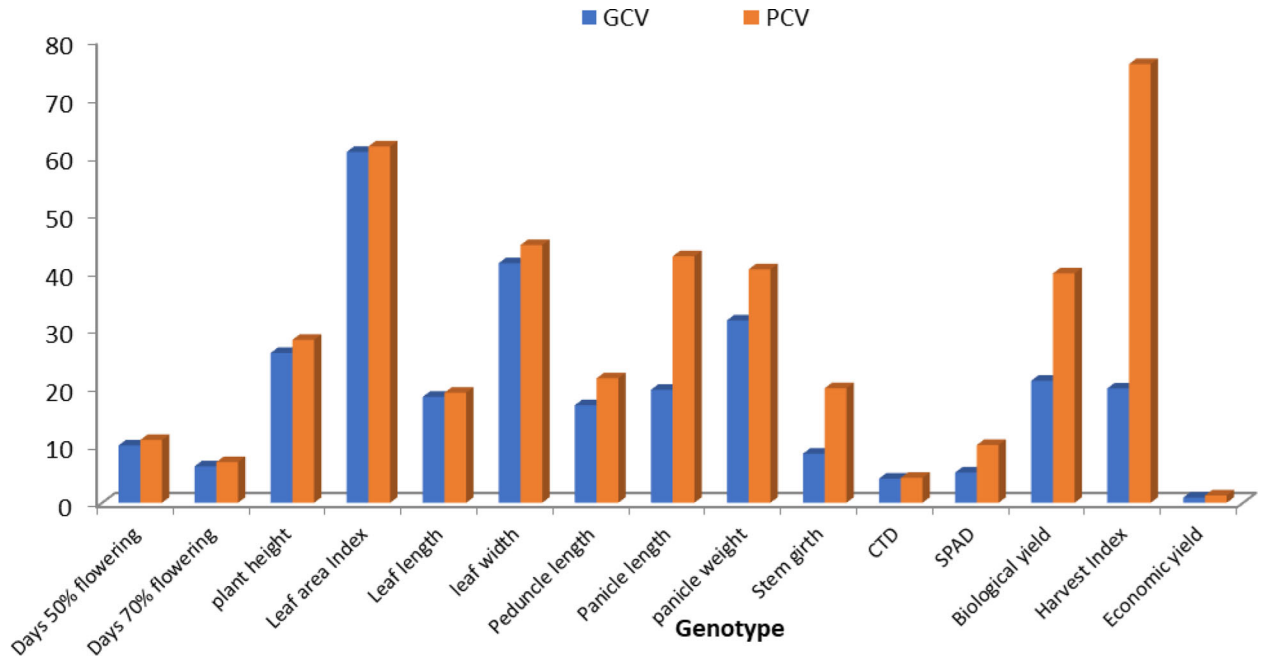


FIGURE 1
The genotypic and phenotypic coefficient of variation in 50 accessions for different parameters in *S. italica*.

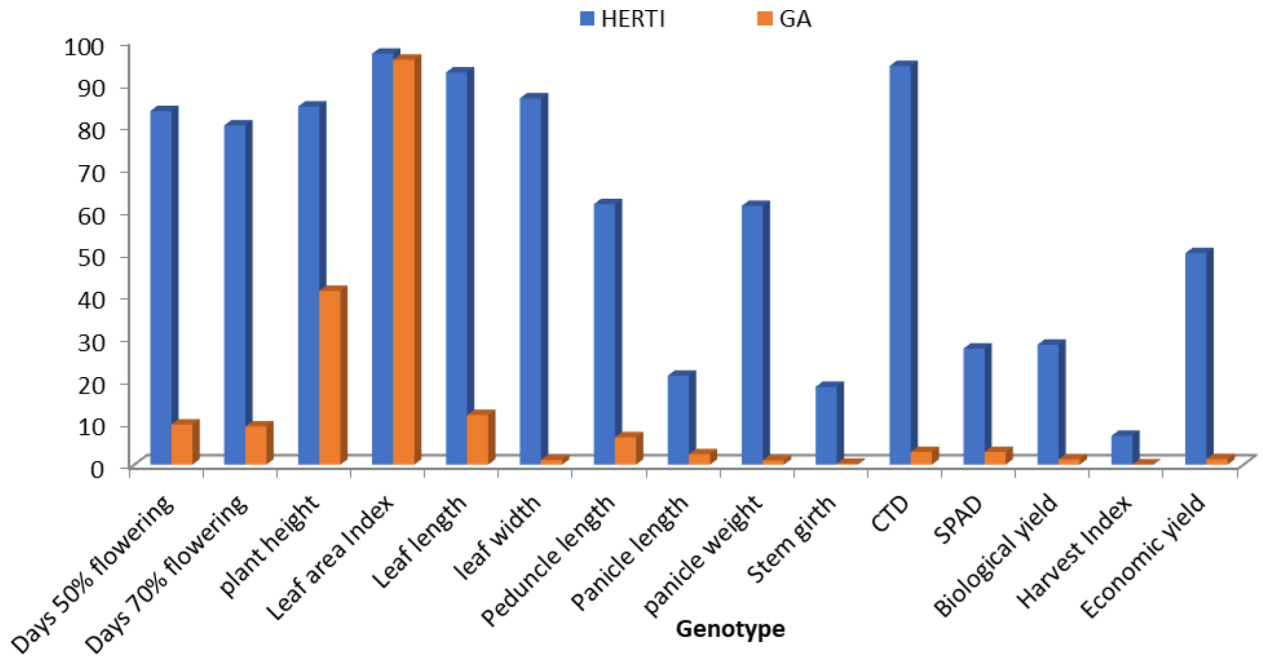


FIGURE 2
Histogram depicting heritability and genetic advance coefficient of variation for different parameters in *S. italica*.

66.80 cm and 6.40 cm; peduncle length, 87.17 cm and 10.09 cm; panicle length, 90.13 cm and 8.10 cm; panicle weight, 76.96 gm and 1.59 gm; leaf area index, 76.43 and 37.96; stem girth, 66.63 cm and 0.39 cm; SPAD, 45.61 nm and 5.26 nm; CTD, 91.43°C and 2.12°C; harvest index, 60.47 gm and 0.21 gm; biological yield, 71.99 gm and 4.26 gm; and economic yield, 32.96 gm and 1.70 gm. Brunda et al. (2014) reported that different crops have contributed to the overall parameters which involve the GCV and PCV traits.

3.1 Correlation analysis

The genotypic and phenotypic correlation between yield and yield components and the interrelationships among them are estimated and presented in Tables 1, 2. The qualitative and quantitative characters of the 50 genotypes from 2018 were analyzed to help identify the 10 best genotypes. The same methodology was used to select the five best genotypes from the 10 from 2019.

3.1.1 Genotypic correlation

For the 50 genotypes from 2018, the following genotypic correlations were obtained: days to 70% flowering had a 1% significant genotypic correlation with days to 50% flowering (0.466*); grain yield had a 50% significant genotypic correlation with leaf width (0.184 cm); grain yield had a 1% significant genotypic correlation with biological yield (0.554**); grain yield had a 1% significant genotypic correlation with harvest index (1.059%); and grain yield showed a negative genotypic correlation with SPAD (−0.403), CTD (−0,037), stem girth (−0.326 cm), and panicle length (−0.048 cm).

For the 10 genotypes from 2019, the genotypic correlations are shown in Table 2 and were as follows: grain yield with plant height (0.915), peduncle length (0.568 cm*), panicle length (0.551 cm), panicle weight (1.028 g), SPAD (0.609), harvest index (1.062%*), and biological yield (1.197 g) showed a 1% significant genotypic correlation; grain yield with leaf width (−0.034 cm), stem girth (−0.275 cm), and CTD (−0.053) showed a negative genotypic correlation; and grain yield with days to 50% flowering (0.109), days to 70% flowering (0.102), plant height (0.015 cm), leaf length (0.249 cm), peduncle length (0.568 cm), panicle length (0.551 cm), leaf area index (1.028), SPAD (0.098), harvest index (0.609%), biological yield (1.062 g), and grain yield (1.197 g) showed a positive genotypic correlation.

3.1.2 Phenotypic correlation

Phenotypic correlations were calculated among the 50 genotypes from 2018. There was 1% significance with days to 70% flowering, plant height, leaf area index, leaf length, leaf width, peduncle length, harvest index, biological yield, grain yield, peduncle weight, and stem girth. Grain yield with days to 70% flowering (−0.015), panicle length (−0.036 cm), stem girth (−0.023 cm), CTD (0.043), and SPAD (0.150) showed a negative phenotypic correlation. Grain yield showed a significantly increased positive correlation with days to 50% flowering (0.071), plant height (0.053 cm), leaf area index (0.099), leaf length (0.117 cm), leaf width (0.084 cm), peduncle length (0.024 cm), panicle weight (0.207 g), biological yield (0.297 g), harvest index (0.506%), and grain yield (1.000 g). Grain yield showed an increased positive phenotypic correlation in 10 genotypes.

The relationship pattern of grain yield with panicle weight, panicle length, leaf width, leaf length, leaf area index, and plant height was comparable at genotypic and phenotypic levels for the 50 genotypes from 2018. A profoundly huge positive affiliation was observed for grain yield per plant with panicle length, leaf width, leaf length, and leaf area index at both the genotypic and phenotypic

TABLE 1 Estimation of 50 accessions for the genotypic correlation coefficient 15 yield component in foxtail millet from 2018.

S. No		Days of 50% flowering	Days of 70% flowering	Leaf length (cm)	Leaf width (cm)	Panicle length (cm)	Panicle weight (gm)	Peduncle length (cm)	Plant height (cm)	SPAD	Stem girth (cm)	Biological yield (gm)	CTD	Harvest index (%)	Leaf area index	Economic yield (gm)
1	Days of 50% flowering	1.000	0.474**	0.354**	0.261*	0.302**	0.117	0.413**	0.385**	0.318**	0.016	0.067	0.137	0.080	0.361**	0.085
2	Days of 70% flowering	0.474**	1.000	0.081	0.137	0.232*	0.212*	0.232*	0.277**	0.126	-0.046	-0.091	0.164*	0.170*	0.252*	-0.007
3	Leaf length (cm)	0.354**	0.081	1.000	0.330**	0.343**	0.223*	0.287**	0.447**	0.097	0.224*	0.030	-0.279**	0.125	0.386**	0.140
4	Leaf width (cm)	0.261*	0.137	0.330**	1.000	0.368**	0.250*	0.210*	0.331**	-0.118	0.002	0.311**	-0.012	-0.089	0.875**	0.151
5	Panicle length (cm)	0.302**	0.232*	0.343**	0.368**	1.000	0.026	0.245*	0.494**	-0.101	0.160	0.029	0.074	-0.140	0.295**	-0.032
6	Panicle weight (gm)	0.117	0.212*	0.223*	0.250*	0.026	1.000	0.234*	0.386**	-0.216*	-0.004	0.205*	0.010	0.027	0.298**	0.150
7	Peduncle length (cm)	0.413**	0.232*	0.287**	0.210*	0.245*	0.234*	1.000	0.709**	-0.027	0.645**	0.096	0.171*	-0.140	0.107	-0.022
8	Plant height (cm)	0.385**	0.277**	0.447**	0.331**	0.494**	0.386**	0.709**	1.000	0.052	0.445**	0.234*	0.058	-0.145	0.298**	0.123
9	SPAD	0.318**	0.126	0.097	-0.118	-0.101	-0.216*	-0.027	0.052	1.000	0.032	-0.085	-0.216*	-0.376**	-0.076	-0.369**
10	Stem girth (cm)	0.016	-0.046	0.224*	0.002	0.160	-0.004	0.645**	0.445**	0.032	1.000	0.005	-0.273**	-0.299**	-0.193*	-0.150
11	Biological yield (gm)	0.067	-0.091	0.030	0.311**	0.029	0.205*	0.096	0.234*	-0.085	0.005	1.000	0.252*	-0.312**	0.241*	0.422**
12	CTD	0.137	0.164*	-0.279**	-0.012	0.074	0.010	0.171*	0.058	-0.216*	-0.273**	0.252*	1.000	-0.080	-0.083	-0.021
13	Harvest index (%)	0.080	0.170*	0.125	-0.089	-0.140	0.027	-0.140	-0.145	-0.376**	-0.299**	-0.312**	-0.080	1.000	-0.011	0.557**
14	Leaf area index	0.361**	0.252*	0.386**	0.875**	0.295**	0.298**	0.107	0.298**	-0.076	-0.193*	0.241*	-0.083	-0.011	1.000	0.123
15	Economic yield (gm)	0.085	-0.007	0.140	0.151	-0.032	0.150	-0.022	0.123	-0.369**	-0.150	0.422**	-0.021	0.557**	0.123	1.000

*Significant at the 5% level; ** Significant at 1%.

10.3389/fpls.2023.1017652

TABLE 2 Estimation of 10 accessions for the genotypic correlation coefficient 15 yield component in foxtail millet from 2019.

S. No.		Days to 50% flowering	Days to 70% flowering	Peduncle length (cm)	Panicle length (cm)	Panicle weight (gm)	Plant height (cm)	SPAD	Stem girth (cm)	Harvest index (%)	Leaf area index	Leaf length (cm)	Leaf width (cm)	Biological yield (gm)	CTD	Economic yield (cm)
1	Days to 50% flowering	1.000	1.467	-0.225	-0.480*	-0.100	0.030	0.059	-0.656**	0.477*	-0.555*	-0.631**	-0.636**	-0.241	0.692**	0.111
2	Days to 70% flowering	1.467	1.000	-0.475*	-0.703**	-0.072	0.131	-0.077	-0.255	0.680**	-0.637**	-0.695**	-0.424*	-0.329	0.761**	0.145
3	Peduncle length (cm)	-0.225	-0.475*	1.000	0.935**	0.414*	-0.147	-0.226	-0.545*	0.024	0.455*	0.421*	0.378*	0.567*	-0.485*	0.443*
4	Panicle length (cm)	-0.480*	-0.703**	0.935**	1.000	0.503*	0.027	-0.062	-0.399*	-0.029	0.570*	0.624**	0.525*	0.610**	-0.765**	0.436*
5	Panicle weight (gm)	-0.100	-0.072	0.414*	0.503*	1.000	0.635**	0.241	-0.165	0.515*	0.599**	0.652**	0.565*	0.878**	-0.326	0.811**
6	Plant height (cm)	0.030	0.131	-0.147	0.027	0.635**	1.000	0.546*	-0.076	0.615**	0.130	0.391*	0.010	0.609**	-0.363*	0.685**
7	SPAD	0.059	-0.077	-0.226	-0.062	0.241	0.546*	1.000	0.152	0.004	0.062	0.297	-0.135	0.360	0.392*	0.226
8	Stem girth (cm)	-0.656**	-0.255	-0.545*	-0.399*	-0.165	-0.076	0.152	1.000	0.094	-0.180	-0.096	0.007	-0.352	-0.049	-0.192
9	Harvest index (%)	0.477*	0.680**	0.024	-0.029	0.515*	0.615**	0.004	0.094	1.000	-0.227	-0.177	-0.228	0.362*	0.061	0.789**
10	Leaf area index	-0.555*	-0.637**	0.455*	0.570*	0.599**	0.130	0.062	-0.180	-0.227	1.000	1.045	1.082	0.499*	-0.837**	0.097
11	Leaf length (cm)	-0.631**	-0.695**	0.421*	0.624**	0.652**	0.391*	0.297	-0.096	-0.177	1.045	1.000	0.894**	0.590**	-1.048**	0.184
12	Leaf width (cm)	-0.636**	-0.424*	0.378*	0.525*	0.565*	0.010	-0.135	0.007	-0.228	1.082	0.894**	1.000	0.288	-0.959**	-0.044
13	Biological yield (gm)	-0.241	-0.329	0.567*	0.610**	0.878**	0.609**	0.360	-0.352	0.362*	0.499*	0.590**	0.288	1.000	-0.284	0.842**
14	CTD	0.692**	0.761**	-0.485*	-0.765**	-0.326	-0.363*	0.392*	-0.049	0.061	-0.837**	-1.048**	-0.959**	-0.284	1.000	-0.102
15	Economic yield (cm)	0.111	0.145	0.443*	0.436*	0.811**	0.685**	0.226	-0.192	0.789**	0.097	0.184	-0.044	0.842**	-0.102	1.000

**Significant at 1%; *significant at the 5% level.

10.3389/fpls.2023.1017652

levels. Comparative outcomes showed that grain yield per plant had a positively huge relationship at the two levels in terms of days to development, panicle length, panicle weight, plant stature, and test weight. Connection examinations likewise give provided data about the relationship between other plant attributes. Plant height had an exceptionally critical positive relationship with number of tillers and panicle length, which was in line with Nirmlakumari and Vetriventhan (2010), and furthermore with panicle width and panicle weight. Leaf area index and panicle weight showed an exceptionally critical positive relationship among themselves. Thus, the determination of both of these qualities increases the chances of improving the other characteristics; therefore, both attributes further improve grain yield.

It is fascinating to note that stem girth shows a positively huge relationship with plant height and peduncle length. Phenotypic correlations in the 10 genotypes collected in 2019 are shown in Table 2. Economic yield showed a 5% significant phenotypic correlation with plant height (0.457*), a 1% significant phenotypic correlation with panicle weight (0.615*), harvest index (0.553), and biological yield (0.521*), a negative phenotypic correlation with leaf width (−0.046), stem girth (−0.064), SPAD (−0.113), and CTD (−0.088), and a positive correlation with days to 50% flowering (0.104), days to 70% flowering (0.071), plant height (0.116), leaf length (0.091), peduncle length (0.304), panicle length (0.331), biological yield (0.521), and economic yield (1.000).

3.2 Path coefficient analysis

The direct and indirect effects of different yield components in grain yield were calculated through path coefficient analysis at genotypic and phenotypic levels and are shown in Tables 3, 4. The phenotypic and genotypic correlations reveal the extent and direction of association between different characters. These are in agreement with the results obtained by Brunda et al. (2015) in foxtail millet and suggest that selection for these traits indirectly improves grain yield.

3.2.1 Genotypic path correlation

Genotypic path correlation revealed a highly positive direct effect of panicle length. Days to 50% flowering, plant height, leaf width, biological yield, economic yield, and harvest index showed a negative genotypic path coefficient analysis with plant height (−0.1375), leaf width (−0.1275), peduncle length (−0.0190), panicle length (−0.1401), panicle weight (−0.051), stem girth (−0.6868), CTD (−0.1372), and SPAD (0.3700). Harvest index showed a significant increase and positive correlation coefficient analysis with days to 50% flowering (0.0979), days to 70% flowering (0.2048), leaf area index (0.0218), leaf length (0.2456), biological yield (0.0109), and economic yield (0.7547).

3.2.2 Phenotypic path correlation

The phenotypic path in the 50 genotypes collected in 2018 is shown in Table 5 and revealed a highly positive direct effect of plant height, leaf length, leaf width, panicle length, and biological yield. Harvest index showed a negative phenotypic path correlation with plant height (−0.0630), leaf area index (−0.0032), leaf width (−0.0431), peduncle length (−0.0561), panicle length (−0.0988), stem girth (−0.0201), CTD (−0.281), SPAD (−0.1502), and biological yield (−0.2820), and a positive phenotypic path coefficient analysis with days to 50% flowering (0.0419), days to 70% flowering (0.0582), leaf length (0,0740), and panicle weight (0.0359).

For the 10 genotypes collected in 2019, genotypic path coefficient analysis is shown in Table 6, which reveals the highly positive direct effect of harvest index, panicle weight, and plant height. Genotypic path coefficient analysis showed a positive genotypic path for days to 50% flowering (2.364), leaf length (0.427), panicle length (0.741), and leaf area index (2.747). Biological yield showed a positive genotypic path coefficient analysis with days to 50% flowering (0.234), days to 70% flowering (0.296), stem girth (0.366), and CTD (0.236). The immediate and roundabout impacts of various yield segments on grain yield were determined through weight examination at phenotypic and genotypic levels. This examination uncovered the high and immediate impact of plant height, peduncle length, and leaf width on grain yield per plant in 50 germplasm assortments. This demonstrates a genuine connection between these characters with grain yield per plant and the direct determination of these attributes helps to improve grain yield per plant. Comparable investigations of grain yield were carried out at the genotypic and phenotypic level in terms of panicle weight, test weight, and straw weight. Weight examination showed that plant height and peduncle length significantly and immediately affected grain yield. This positive direct impact of plant stature and peduncle length on grain yield suggests that the biomass of a plant should be built up to increase yield. The weight investigation revealed that the immediate impact of plant height, leaf length, leaf width, and panicle length on grain yield was positive. For this characteristic to produce the desired results, it would seem that determination must be focused in a particular direction.

The immediate impact of days to 70% flowering on grain yield was low and negative in both genotype and phenotype. The immediate impact of stem girth on grain yield was positive and low, which corroborates the findings of Nirmlakumari and Vetriventhan (2010). This positive direct impact of plant height on grain yield is attractive as it offers a way to increase straw and grain yield. The immediate impact of test weight on grain yield per plant was positive and high in both seasons, which demonstrates the genuine relationship between these attributes and a straightforward method for increasing grain yield. There is a tendency to believe that the determination of panicle length in foxtail millet will lead to plant height, stem girth, and leaf length being targeted to expand grain yield per plant. In light of the consequences of the weight examination, there is a tendency to infer that increasing characters such as panicle length, plant height, and stem girth, which had a positive connection with and direct impact on yield, will improve yield. Henceforth, lavish plants with enormous panicles, increased grain weight, and high panicle weight might bring about a better return in genotypes of foxtail millet.

The phenotypic path in the 10 genotypes collected in 2019 is shown in Table 7 and revealed a highly positive direct effect of plant height, harvest index, and biological yield. Biological yield showed a positive phenotypic path with plant height (0.037), leaf width (0.012), leaf length (0.031), pedicle length (-0.034), panicle length (0.036), panicle weight (0.052), and SPAD (0.027) and a negative phenotypic path with days to 70% flowering (−0.013), days to 50% flowering

TABLE 3 Estimation of 50 accessions for the phenotypic correlation coefficient 15 yield component in foxtail millet from 2018.

S. No.		Days of 50% flowering	Days of 70% flowering	Leaf length (cm)	Leaf width (cm)	Panicle length (cm)	Panicle weight (gm)	Peduncle length (cm)	Plant height (cm)	SPAD	Stem girth (cm)	Biological yield (gm)	CTD	Harvest index (%)	Leaf area index	Economic yield (gm)
1	Days of 50% flowering	1.000	0.454**	0.296**	0.215*	0.268**	0.114	0.347**	0.322**	0.154	0.013	0.063	0.033	0.076	0.309**	0.086
2	Days of 70% flowering	0.454**	1.000	0.053	0.098	0.179*	0.182*	0.152	0.195*	0.037	-0.066	-0.073	0.023	0.151	0.204*	0.002
3	Leaf length (cm)	0.296**	0.053	1.000	0.311**	0.309**	0.201*	0.235*	0.394**	0.054	0.177*	0.028	-0.140	0.115	0.369**	0.128
4	Leaf width (cm)	0.215*	0.098	0.311**	1.000	0.331**	0.225*	0.168*	0.304**	-0.050	-0.007	0.282**	-0.024	-0.082	0.833**	0.131
5	Panicle length (cm)	0.268**	0.179*	0.309**	0.331**	1.000	0.020	0.182*	0.414**	-0.044	0.104	0.060	0.026	-0.141	0.257*	-0.021
6	Panicle weight (gm)	0.114	0.182*	0.201*	0.225*	0.020	1.000	0.184*	0.339**	-0.118	-0.009	0.199*	0.011	0.048	0.263*	0.157
7	Peduncle length (cm)	0.347**	0.152	0.235*	0.168*	0.182*	0.184*	1.000	0.605**	0.049	0.357**	0.068	0.103	-0.125	0.098	-0.028
8	Plant height (cm)	0.322**	0.195*	0.394**	0.304**	0.414**	0.339**	0.605**	1.000	0.043	0.283**	0.198*	0.068	-0.148	0.271**	0.086
9	SPAD	0.154	0.037	0.054	-0.050	-0.044	-0.118	0.049	0.043	1.000	0.023	-0.054	-0.039	-0.210*	-0.053	-0.193*
10	Stem girth (cm)	0.013	-0.066	0.177*	-0.007	0.104	-0.009	0.357**	0.283**	0.023	1.000	0.016	-0.089	-0.198*	-0.130	-0.085
11	Biological yield (gm)	0.063	-0.073	0.028	0.282**	0.060	0.199*	0.068	0.198*	-0.054	0.016	1.000	0.095	-0.310**	0.214*	0.399**
12	CTD	0.033	0.023	-0.140	-0.024	0.026	0.011	0.103	0.068	-0.039	-0.089	0.095	1.000	-0.029	-0.065	-0.015
13	Harvest index (%)	0.076	0.151	0.115	-0.082	-0.141	0.048	-0.125	-0.148	-0.210*	-0.198*	-0.310**	-0.029	1.000	-0.011	0.539**
14	Leaf area index	0.309**	0.204*	0.369**	0.833**	0.257*	0.263*	0.098	0.271**	-0.053	-0.130	0.214*	-0.065	-0.011	1.000	0.116
15	Economic yield (gm)	0.086	0.002	0.128	0.131	-0.021	0.157	-0.028	0.086	-0.193*	-0.085	0.399**	-0.015	0.539**	0.116	1.000

**Significant at 1%; *significant at the 5% level.

10.3389/fpls.2023.1017652

TABLE 4 Estimation of 10 accessions for the phenotypic correlation coefficient 15 yield component in foxtail millet from 2019.

	Days to 50% flowering	Days to 70% flowering	Peduncle length (cm)	Panicle length (cm)	Panicle weight (gm)	Plant height (cm)	SPAD	Stem girth (cm)	Harvest index (%)	Leaf area index	Leaf length (cm)	Leaf width (cm)	Biological yield (gm)	CTD	Economic yield (gm)
Days to 50% flowering	1.000	0.556*	-0.217	-0.323	-0.021	0.041	-0.071	-0.251	0.327	-0.314	-0.405*	-0.230	-0.152	0.395*	0.098
Days to 70% flowering	0.556*	1.000	-0.213	-0.390*	-0.067	0.044	-0.120	-0.272	0.389*	-0.280	-0.376*	-0.370*	-0.186	0.276	0.068
Peduncle length (cm)	-0.217	-0.213	1.000	0.879**	0.362*	-0.138	-0.188	-0.356	0.031	0.413*	0.321	0.262	0.538*	-0.217	0.401*
Panicle length (cm)	-0.323	-0.390*	0.879**	1.000	0.485*	0.025	-0.042	-0.277	-0.020	0.540*	0.471*	0.358	0.600**	-0.408*	0.425*
Panicle weight (gm)	-0.021	-0.067	0.362*	0.485*	1.000	0.611**	0.181	-0.087	0.504*	0.578**	0.488*	0.411*	0.857**	-0.229	0.793**
Plant height (cm)	0.041	0.044	-0.138	0.025	0.611**	1.000	0.335	-0.025	0.576**	0.156	0.272	-0.078	0.569*	-0.105	0.650**
SPAD	-0.071	-0.120	-0.188	-0.042	0.181	0.335	1.000	0.135	0.039	0.001	0.042	-0.216	0.234	0.032	0.175
Stem girth (cm)	-0.251	-0.272	-0.356	-0.277	-0.087	-0.025	0.135	1.000	0.060	-0.132	-0.198	0.049	-0.203	0.112	-0.125
Harvest index (%)	0.327	0.389*	0.031	-0.020	0.504*	0.576**	0.039	0.060	1.000	-0.225	-0.182	-0.199	0.351	0.013	0.772**
Leaf area index	-0.314	-0.280	0.413*	0.540*	0.578**	0.156	0.001	-0.132	-0.225	1.000	0.821**	0.746**	0.478*	-0.405*	0.096
Leaf length (cm)	-0.405*	-0.376*	0.321	0.471*	0.488*	0.272	0.042	-0.198	-0.182	0.821**	1.000	0.760**	0.492*	-0.436*	0.136
Leaf width (cm)	-0.230	-0.370*	0.262	0.358	0.411*	-0.078	-0.216	0.049	-0.199	0.746**	0.760**	1.000	0.253	-0.357	-0.037
Biological yield (gm)	-0.152	-0.186	0.538*	0.600**	0.857**	0.569*	0.234	-0.203	0.351	0.478*	0.492*	0.253	1.000	-0.101	0.831**
CTD	0.395*	0.276	-0.217	-0.408*	-0.229	-0.105	0.032	0.112	0.013	-0.405*	-0.436*	-0.357	-0.101	1.000	-0.032
Economic yield (gm)	0.098	0.068	0.401*	0.425*	0.793**	0.650**	0.175	-0.125	0.772**	0.096	0.136	-0.037	0.831**	-0.032	1.000

**Significant at 1%; *significant at the 5% level.

10.3389/fpls.2023.1017652

TABLE 5 Genotypic path of 15 yield component traits in 50 foxtail millet accessions from 2018.

	Days of 50% flowering	Days of 70% flowering	Leaf length (cm)	Leaf width (cm)	Panicle length (cm)	Panicle weight (gm)	Peduncle length (cm)	Plant height (cm)	SPAD	Stem girth (cm)	Biological yield (cm)	CTD	Harvest Index (%)	Leaf area index	Economic yield (gm)
Days of 50% flowering	0.123	0.059	0.044	0.032	0.037	0.015	0.051	0.048	0.039	0.002	0.008	0.017	0.010	0.045	0.085
Days of 70% flowering	-0.012	-0.025	-0.002	-0.003	-0.006	-0.005	-0.006	-0.007	-0.003	0.001	0.002	-0.004	-0.004	-0.006	-0.007
Leaf length (cm)	-0.008	-0.002	-0.023	-0.008	-0.008	-0.005	-0.007	-0.011	-0.002	-0.005	-0.001	0.007	-0.003	-0.009	0.140
Leaf width (cm)	0.073	0.038	0.092	0.279	0.103	0.070	0.058	0.092	-0.033	0.000	0.087	-0.003	-0.025	0.244	0.151
Panicle length (cm)	-0.014	-0.011	-0.016	-0.017	-0.047	-0.001	-0.012	-0.023	0.005	-0.008	-0.001	-0.004	0.007	-0.014	-0.032
Panicle weight (gm)	-0.010	-0.017	-0.018	-0.020	-0.002	-0.081	-0.019	-0.031	0.018	0.000	-0.017	-0.001	-0.002	-0.024	0.150
Peduncle length (cm)	-0.010	-0.005	-0.007	-0.005	-0.006	-0.005	-0.023	-0.017	0.001	-0.015	-0.002	-0.004	0.003	-0.003	-0.022
Plant height (cm)	0.096	0.069	0.112	0.083	0.124	0.097	0.178	0.250	0.013	0.111	0.059	0.015	-0.036	0.075	0.123
SPAD	-0.061	-0.024	-0.019	0.023	0.019	0.042	0.005	-0.010	-0.192	-0.006	0.016	0.042	0.072	0.015	-0.369**
Stem girth (cm)	-0.003	0.009	-0.044	0.000	-0.031	0.001	-0.126	-0.087	-0.006	-0.196	-0.001	0.053	0.059	0.038	-0.150
Biological yield (cm)	0.043	-0.059	0.019	0.200	0.019	0.132	0.062	0.151	-0.055	0.003	0.644	0.162	-0.201	0.155	0.422**
CTD	-0.039	-0.046	0.079	0.003	-0.021	-0.003	-0.048	-0.016	0.061	0.077	-0.071	-0.282	0.023	0.024	-0.021
Harvest index (gm)	0.052	0.111	0.081	-0.058	-0.091	0.018	-0.091	-0.095	-0.245	-0.195	-0.203	-0.052	0.651	-0.007	0.557**
Leaf area index	-0.147	-0.103	-0.158	-0.357	-0.121	-0.122	-0.044	-0.122	0.031	0.079	-0.098	0.034	0.005	-0.408	0.123
Economic yield (gm)	0.085	-0.007	0.140	0.151	-0.032	0.150	-0.022	0.123	-0.369**	-0.150	0.422**	-0.021	0.557**	0.123	1.000
Partial R^2	0.010	0.000	-0.003	0.042	0.002	-0.012	0.001	0.031	0.071	0.029	0.272	0.006	0.363	-0.050	

**Significant at 1%.

10.3389/fpls.2023.1017652

TABLE 6 Genotypic path of 15 yield component traits in 10 foxtail millet accessions from 2019.

	Days of 50% flowering	Days of 70% flowering	Leaf length (cm)	Leaf width (cm)	Panicle length (cm)	Panicle weight (gm)	Peduncle length (cm)	Plant height (cm)	SPAD	Stem girth (cm)	Biological yield (cm)	CTD	Harvest index (%)	Leaf area index	Economic yield (gm)
Days of 50% flowering	0.007	0.010	-0.002	-0.003	-0.001	0.000	0.000	-0.005	0.003	-0.004	-0.004	-0.004	-0.002	0.005	0.111
Days of 70% flowering	-0.040	-0.027	0.013	0.019	0.002	-0.004	0.002	0.007	-0.018	0.017	0.019	0.012	0.009	-0.021	0.145
Leaf length (cm)	0.014	0.030	-0.063	-0.058	-0.026	0.009	0.014	0.034	-0.002	-0.029	-0.026	-0.024	-0.035	0.030	0.443
Leaf width (cm)	-0.114	-0.167	0.222	0.237	0.119	0.007	-0.015	-0.095	-0.007	0.135	0.148	0.124	0.145	-0.181	0.436
Panicle length (cm)	-0.017	-0.012	0.070	0.085	0.170	0.108	0.041	-0.028	0.087	0.102	0.111	0.096	0.149	-0.055	0.811
Panicle weight (gm)	-0.001	-0.005	0.005	-0.001	-0.023	-0.036	-0.020	0.003	-0.022	-0.005	-0.014	0.000	-0.022	0.013	0.685
Peduncle length (cm)	-0.001	0.001	0.003	0.001	-0.003	-0.007	-0.013	-0.002	0.000	-0.001	-0.004	0.002	-0.005	-0.005	0.226
Plant height (cm)	-0.003	-0.001	-0.002	-0.002	-0.001	0.000	0.001	0.004	0.000	-0.001	0.000	0.000	-0.002	0.000	-0.192
SPAD	0.235	0.336	0.012	-0.014	0.254	0.304	0.002	0.047	0.494	-0.112	-0.087	-0.113	0.179	0.030	0.789
Stem girth (cm)	0.111	0.127	-0.091	-0.114	-0.120	-0.026	-0.012	0.036	0.045	-0.200	-0.209	-0.216	-0.100	0.167	0.097
Biological yield (cm)	-0.024	-0.027	0.016	0.024	0.025	0.015	0.011	-0.004	-0.007	0.040	0.039	0.034	0.023	-0.040	0.184
CTD	0.054	0.036	-0.032	-0.045	-0.048	-0.001	0.011	-0.001	0.019	-0.092	-0.076	-0.085	-0.025	0.081	-0.044
Harvest index (gm)	-0.129	-0.176	0.303	0.326	0.469	0.325	0.192	-0.188	0.193	0.267	0.315	0.154	0.534	-0.152	0.842
Leaf area index	0.018	0.019	-0.012	-0.019	-0.008	-0.009	0.010	-0.001	0.002	-0.021	-0.026	-0.024	-0.007	0.025	-0.102
Economic yield (gm)	0.111	0.145	0.443	0.436	0.811	0.685	0.226	-0.192	0.789	0.097	0.184	-0.044	0.842	-0.102	1.000
Partial R^2	0.001	-0.004	-0.028	0.103	0.138	-0.025	-0.003	-0.001	0.390	-0.020	0.007	0.004	0.450	-0.003	

10.3389/fpls.2023.1017652

TABLE 7 Phenotypic path of 15 yield component traits in 50 foxtail millet accessions from 2018.

	Days of 50% flowering	Days of 70% flowering	Leaf length (cm)	Leaf width (cm)	Panicle length (cm)	Panicle weight (gm)	Peduncle length (cm)	Plant height (cm)	SPAD	Stem girth (cm)	Biological yield (cm)	CTD	Harvest index (%)	Leaf area index	Economic yield (gm)
Days of 50% flowering	0.025	0.014	-0.005	-0.008	-0.001	0.001	-0.002	-0.006	0.008	-0.008	-0.010	-0.006	-0.004	0.010	0.098
Days of 70% flowering	-0.022	-0.040	0.008	0.016	0.003	-0.002	0.005	0.011	-0.015	0.011	0.015	0.015	0.007	-0.011	0.068
Leaf length (cm)	0.017	0.017	-0.078	-0.069	-0.028	0.011	0.015	0.028	-0.002	-0.032	-0.025	-0.020	-0.042	0.017	0.401
Leaf width (cm)	-0.064	-0.078	0.175	0.199	0.096	0.005	-0.008	-0.055	-0.004	0.107	0.094	0.071	0.119	-0.081	0.425
Panicle length (cm)	-0.003	-0.008	0.044	0.058	0.120	0.073	0.022	-0.010	0.061	0.070	0.059	0.049	0.103	-0.028	0.793
Panicle weight (gm)	-0.002	-0.002	0.005	-0.001	-0.023	-0.037	-0.012	0.001	-0.021	-0.006	-0.010	0.003	-0.021	0.004	0.650
Peduncle length (cm)	0.002	0.003	0.004	0.001	-0.004	-0.007	-0.021	-0.003	-0.001	0.000	-0.001	0.005	-0.005	-0.001	0.175
Plant height (cm)	0.004	0.004	0.005	0.004	0.001	0.000	-0.002	-0.014	-0.001	0.002	0.003	-0.001	0.003	-0.002	-0.125
SPAD	0.161	0.191	0.015	-0.010	0.248	0.283	0.019	0.030	0.492	-0.111	-0.090	-0.098	0.173	0.006	0.772
Stem girth (cm)	0.050	0.045	-0.066	-0.086	-0.092	-0.025	0.000	0.021	0.036	-0.159	-0.131	-0.119	-0.076	0.065	0.096
Biological yield (cm)	-0.012	-0.011	0.009	0.014	0.014	0.008	0.001	-0.006	-0.005	0.024	0.029	0.022	0.014	-0.013	0.136
CTD	0.023	0.036	-0.026	-0.035	-0.040	0.008	0.021	-0.005	0.019	-0.073	-0.074	-0.098	-0.025	0.035	-0.037
Harvest index (gm)	-0.089	-0.109	0.316	0.352	0.503	0.334	0.137	-0.119	0.206	0.281	0.289	0.149	0.587	-0.059	0.831
Leaf area index	0.010	0.007	-0.006	-0.010	-0.006	-0.003	0.001	0.003	0.000	-0.010	-0.011	-0.009	-0.003	0.025	-0.032
Economic yield (gm)	0.098	0.068	0.401	0.425	0.793	0.650	0.175	-0.125	0.772	0.096	0.136	-0.037	0.831	-0.032	1.000
Partial R^2	0.002	-0.003	-0.031	0.084	0.095	-0.024	-0.004	0.002	0.380	-0.015	0.004	0.004	0.488	-0.001	

10.3389/fpls.2023.1017652

(−0.016), stem girth (−0.016), CTD (−0.011), harvest index (0.013), and biological yield. (0.016). Path analysis indicated that plant height and harvest index had a highly positive direct effect on grain yield (Table 8). This positive direct effect of plant height and biological yield on economic yield provides scope to increase the biomass of plants with increased yield.

3.3 Principal component analysis

PCA reduces a very large series of data into a smaller number of components by looking for groups with very strong intercorrelations with a set of variables, and each component is explained as a percentage of variation to the overall variability. The first main component explains most of the overall population variation, followed by subsequent components for huge data, PCA was used to reduce the multivariate data to determine the importance and contribution of each component to the total variance. From the data shown in Figure 3, total variation could be 100% explained by 15 principal components (PCs). The PC eigenvalues of 50 genotypes were calculated and represented F1 to F15. F1 was the largest contributing PC, followed by F2, F3, F4, F5, F6, F7, F8, F9, F10, F11, F12, F13, F14, and F15. PC F1 was the most important contributing character, with an eigen value of 3.590, variability of 23.931%, and a cumulative variance of 23.931%. For the entire germplasm, PC F1 mainly separated accessions with the following 14 traits: days of 50% flowering (0.324), days of 70% flowering (0.231), leaf length (0.318), leaf width (0.362), panicle length (0.291), panicle weight (0.242), peduncle length (0.327), plant height (0.416), SPAD (0.004), stem girth (0.128), biological yield (0.167), CTD (0.039), economic yield (0.103), and leaf area index (0.360), which had the highest loadings in PC F1, indicating the significant importance of this component. These traits explained the largest portion of the variability.

The results of the PCA are shown in Figures 3–6. PC F2 had an eigenvalue of 2.069, a variability of 13.794%, and a cumulative variability of 37.725%. PC F2 mainly separates accessions with eight traits: leaf length (0.005), leaf width (0.194), panicle weight (0.155), biological yield (0.190), CTD (0.044), economic yield (0.499), harvest index (0.416), and leaf area index (0.253), indicating their significant importance for these components. The remaining characters contributed negatively to the first component. The main traits for PC F3 were days of 50% flowering (0.230), days of 70% flowering (0.261), leaf length (0.225), panicle length (0.018), peduncle length (0.064), plant height (0.017), SPAD (0.073), stem girth (0.108), economic yield (0.165), and harvest Index (0.557). PC F4 had an eigenvalue of 1.512, a variability of 10.080%, and a cumulative variability (CV) of 58.221%. PC F5 had an eigenvalue of 1.235, a variability of 8.236%, and a CV of 66.457%. PC F6 had an eigenvalue of 1.050, a variability of 7.002%, and a CV of 73.458%. PC F7 had an eigenvalue of 0.902, a variability of 6.014%, and a CV of 79.473%. PC F8 had an eigenvalue of 0.798, a variability of 5.323%, and a CV of 84.796%. PC F9 had an eigenvalue of 0.671, a variability of 4.471%, and a CV of 89.267%. PC F10 had an eigenvalue of 0.460, a variability of 3.065%, and a CV of 92.332%. PC F11 had an eigenvalue of 0.419, a variability of 2.795%, and a CV of 95.128%. PC F12 had an eigenvalue of 0.323, a variability of 2.156%, and a CV of 97.284%. PC F13 had an eigen value of 0.204, a variability of 1.360%, and a CV of 98.643%. PC F14 had an eigenvalue of 0.109, a variability of 0.729%, and a CV of 99.373%. For the PCA for the entire germplasm, six traits (days to 50% flowering, plant height, peduncle length, panicle weight, leaf length, and economic yield) explained most of the variance in the first five principal components, indicating their importance for the characterization of foxtail millet germplasm accessions.

The principal component analysis featured the eigenvalue, variability (%), and cumulative variability (%) with respect to principal components F1 – F9. Component F1 was the largest contributing principal component followed by F2, F3, F4, F5, F6, F7, F8, and F9. PC F1 had the highest eigenvalue (5.839), with a variability of 38.926% and cumulative variability of 38.926%. PC F2 had the second highest eigenvalue (3.613), with a variability of 24.087%, and a cumulative variability of 63.013%. PC F3 had the third highest values, followed by F4 and F5. PCs F9 and F8 had the lowest eigenvalues. PC F9 had an eigenvalue of 0.031, a variability of 0.204%, and a cumulative variability of 100.00%. PC F8 had an eigenvalue of 0.091, a variability of 0.607%, and a cumulative variability of 99.796%. PC F6 and F7 had moderate ECV values (Table 4). In the scree plot, the red line represents cumulative variability (%) with respect to PCs F1 to F9. In the biplot graph, the PCA in general confirmed the groupings, which were obtained through cluster analysis. The results of PCA are shown in Figure 7. The first two PCs with an eigenvalue of >1 accounted for 63.013% of the total variance. Accessions GS-14 and GS-62 had more PCA value than the other genotype principal components.

The breeding of high-yielding varieties is dependent on the yield-contributing morphological features, and we chose a small number of key traits with a favorable association. F lag leaf area, plant height, peduncle length, and tiller count per plant are major morphological yield contributing factors that are positively connected with yield per plant (Eberhart and Russel, 1966). This experiment suggested that high yielding foxtail millet accessions can be selected through indirect selection of panicle length, panicle weight, stem girth, and economic yield. The accessions GS-14 and GS-62 demonstrated the best performance for the majority of yield-related parameters, and hence can be relevant for further investigation in other regions of Uttar Pradesh similar to the Naini regions.

4 Discussion

The short-term strategy for identifying foxtail millet genotypes rich in grain nutrients to fulfil the urgent requirement of target micronutrient and protein-deficient populations is to analyze, detect, and explore existing genetic diversity. Significant heterogeneity in all grain nutrients was identified in the foxtail millet core collection, implying that there is plenty of room for selecting nutrient-rich accessions for use in breeding approaches. Field trials were carried out at a variety of sites and across two seasons (2018 and 2019). Estimates of variability, heritability, genetic advance, genotypic correlation coefficient, phenotypic correlation coefficient, genotypic route, and phenotypic path were obtained from the data. Significant differences were observed in Kharif Season-2018 to Kharif Season-2019 among the genotypes for all the characters studied. The results showed that analysis of variance revealed significant

TABLE 8 Phenotypic path of 15 yield component traits in 10 foxtail millet accessions from 2019.

	Days of 50% flowering	Days of 70% flowering	Leaf length (cm)	Leaf width (cm)	Panicle length (cm)	Panicle weight (gm)	Peduncle length (cm)	Plant height (cm)	SPAD	Stem girth (cm)	Biological yield (cm)	CTD	Harvest index (%)	Leaf area index	Economic yield (gm)
Days of 50% flowering	0.013	0.006	0.004	0.003	0.003	0.001	0.004	0.004	0.002	0.000	0.001	0.000	0.001	0.004	0.086
Days of 70% flowering	-0.038	-0.083	-0.004	-0.008	-0.015	-0.015	-0.013	-0.016	-0.003	0.006	0.006	-0.002	-0.013	-0.017	0.002
Leaf length (cm)	-0.007	-0.001	-0.024	-0.007	-0.007	-0.005	-0.006	-0.009	-0.001	-0.004	-0.001	0.003	-0.003	-0.009	0.128
Leaf width (cm)	0.008	0.004	0.012	0.038	0.013	0.009	0.006	0.012	-0.002	0.000	0.011	-0.001	-0.003	0.032	0.131
Panicle length (cm)	0.008	0.005	0.009	0.010	0.029	0.001	0.005	0.012	-0.001	0.003	0.002	0.001	-0.004	0.008	-0.021
Panicle weight(gm)	-0.001	-0.002	-0.002	-0.002	0.000	-0.011	-0.002	-0.004	0.001	0.000	-0.002	0.000	-0.001	-0.003	0.157
Peduncle length (cm)	-0.012	-0.005	-0.008	-0.006	-0.006	-0.006	-0.035	-0.021	-0.002	-0.012	-0.002	-0.004	0.004	-0.003	-0.028
Plant height (cm)	0.036	0.022	0.044	0.034	0.046	0.038	0.068	0.112	0.005	0.032	0.022	0.008	-0.017	0.030	0.086
SPAD	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	-0.003	0.000	0.000	0.000	0.001	0.000	-0.193*
Stem girth (cm)	0.000	-0.002	0.004	0.000	0.002	0.000	0.008	0.006	0.001	0.022	0.000	-0.002	-0.004	-0.003	-0.085
Biological yield (cm)	0.039	-0.045	0.017	0.173	0.037	0.122	0.042	0.122	-0.034	0.010	0.616	0.059	-0.191	0.132	0.399**
CTD	-0.002	-0.001	0.008	0.001	-0.002	-0.001	-0.006	-0.004	0.002	0.005	-0.006	-0.058	0.002	0.004	-0.015
Harvest index (gm)	0.059	0.116	0.088	-0.063	-0.108	0.037	-0.096	-0.113	-0.161	-0.152	-0.237	-0.022	0.765	-0.008	0.539**
Leaf area index	-0.015	-0.010	-0.018	-0.041	-0.013	-0.013	-0.005	-0.013	0.003	0.006	-0.011	0.003	0.001	-0.050	0.116
Economic yield (gm)	0.086	0.002	0.128	0.131	-0.021	0.157	-0.028	0.086	-0.193*	-0.085	0.399**	-0.015	0.539**	0.116	1.000
Partial R^2	0.001	0.000	-0.003	0.005	-0.001	-0.002	0.001	0.010	0.001	-0.002	0.246	0.001	0.412	-0.006	

**Significant at 1%; *significant at the 5% level.

10.3389/fpls.2023.1017652

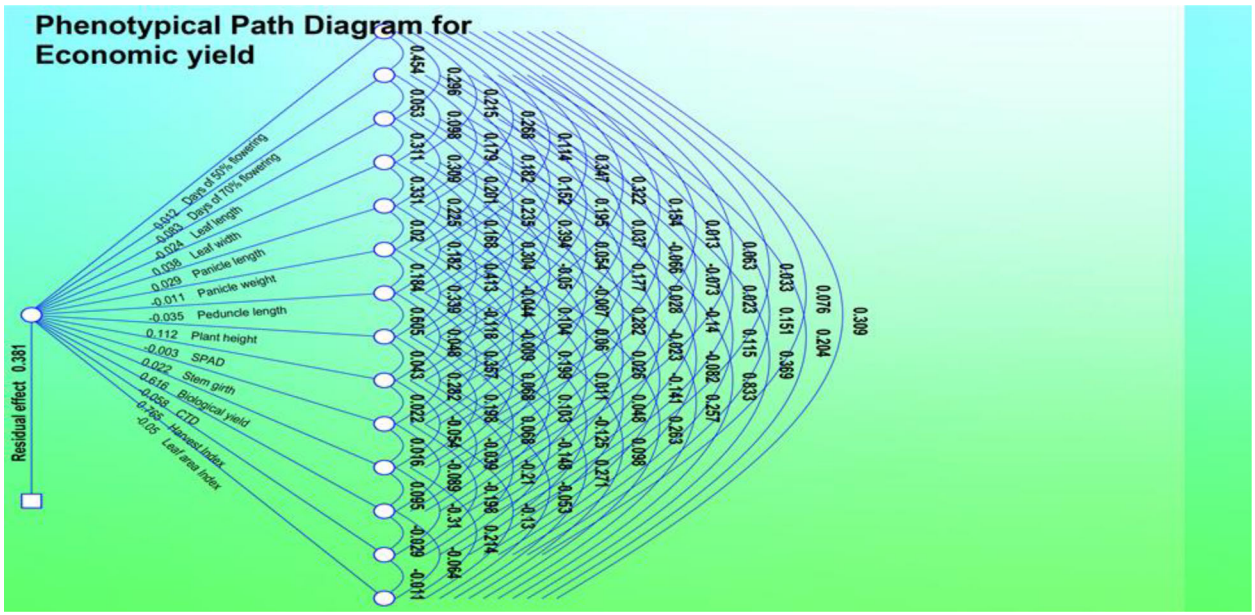


FIGURE 3
Phenotypic path diagram for grain yield of 15 yield component in 50 foxtail millet accessions collected in 2018.

differences for most of the traits, including days to 50% flowering, days to 75% maturity, plant height, leaf length, leaf width, leaf area index, panicle length, panicle weight, biological yield, economic yield, harvest index, and test weight, indicating that all genotypes were genetically diverse for most of the traits. GCV estimations for grain yield were the highest, followed by panicle length and biological yield. Leaf length had the highest PCV estimates, followed by plant height and leaf width. Leaf length had the highest heritability, followed by

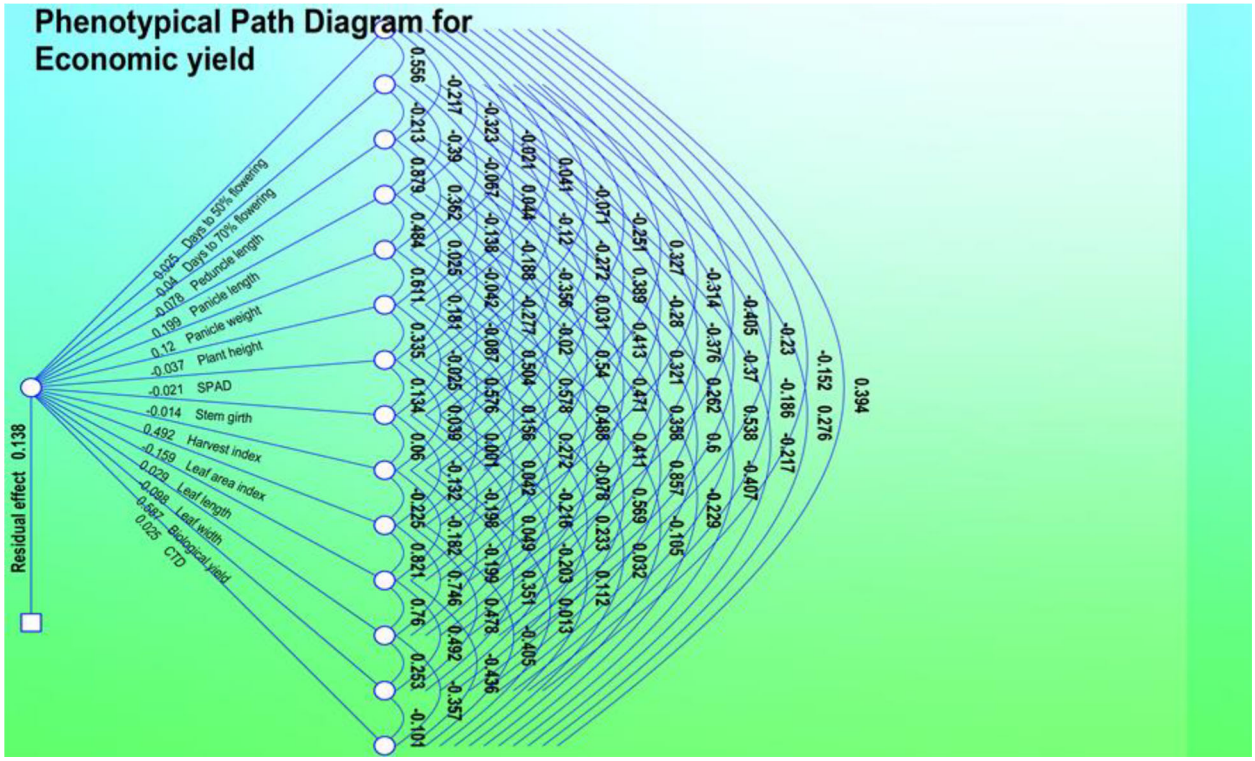


FIGURE 4
Phenotypic path diagram for grain yield of 15 yield component in 10 foxtail millet accessions collected in 2019.

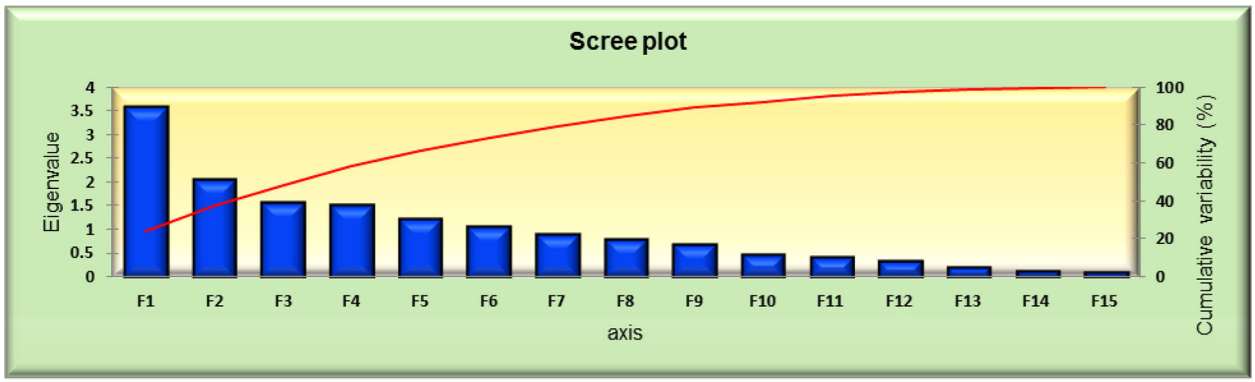


FIGURE 5
PCA scree plot series 1 and 2 for 50 foxtail millet accessions collected in 2018.

pedicle length. Biological yield, economic yield, and agricultural yield all showed significant genetic progress. H arvest index, leaf breadth, panicle weight, panicle length, and leaf area index were all moderately recorded. L ow GCV and PCV were recorded in leaf length and days to 50% flowering. In conclusion, the genotypes Kangni-1, Kangni-7, Kangni-6, Kangni-5, and Kangni-4 showed the best mean performance in the agroclimatic conditions of Allahabad. The direct influence of biological yield on economic yield per plant was positive and high in both years, which indicates that this feature has a true link and that direct selection using this attribute will be effective.

In foxtail millet, the direct selection of biological yield resulted in the simultaneous indirect selection of several panicles, panicle length, pedicle length, panicle weight, number of productive tillers, and biological yield for higher economic production per plant. Seed yield per plant was found to be positively and significantly linked with biological yield, panicle weight, harvest index, leaf length, leaf area index, leaf breadth, plant height, days to flowering, and days to maturity. This suggests that these traits are mostly driven by additive gene action, and thus direct selection for these traits will result in increased grain yield. Similar results were reported by Vinsonias (2018) for plant height and panicle length for plant height for 1,000 g grain weight and flag leaf blade length (Brunda et al., 2015).

Characters such as leaf length, days to 50% flowering, and days to 75% maturity demonstrated high heritability combined with moderate genetic advance, indicating that there is a greater chance of inheritance from progeny to offspring, and thus these characters should be prioritized for effective selection. Earlier studies have also reported a significantly positive association of biological yield per plant with productive panicle and peduncle length (Brunda et al., 2014; Kumar et al., 2015; Kavya et al., 2017). The positive correlation of yield with other characters indicated that all these characters could be simultaneously improved and that an increase

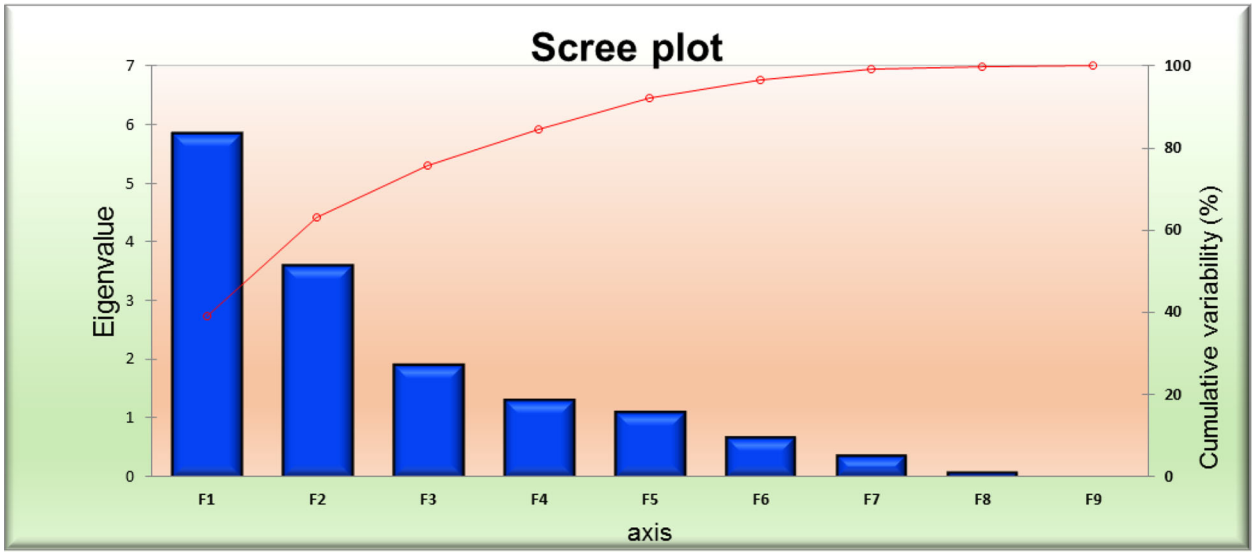


FIGURE 6
PCA scree plot series 1 and 2 for 10 foxtail millet accessions collected in 2019.

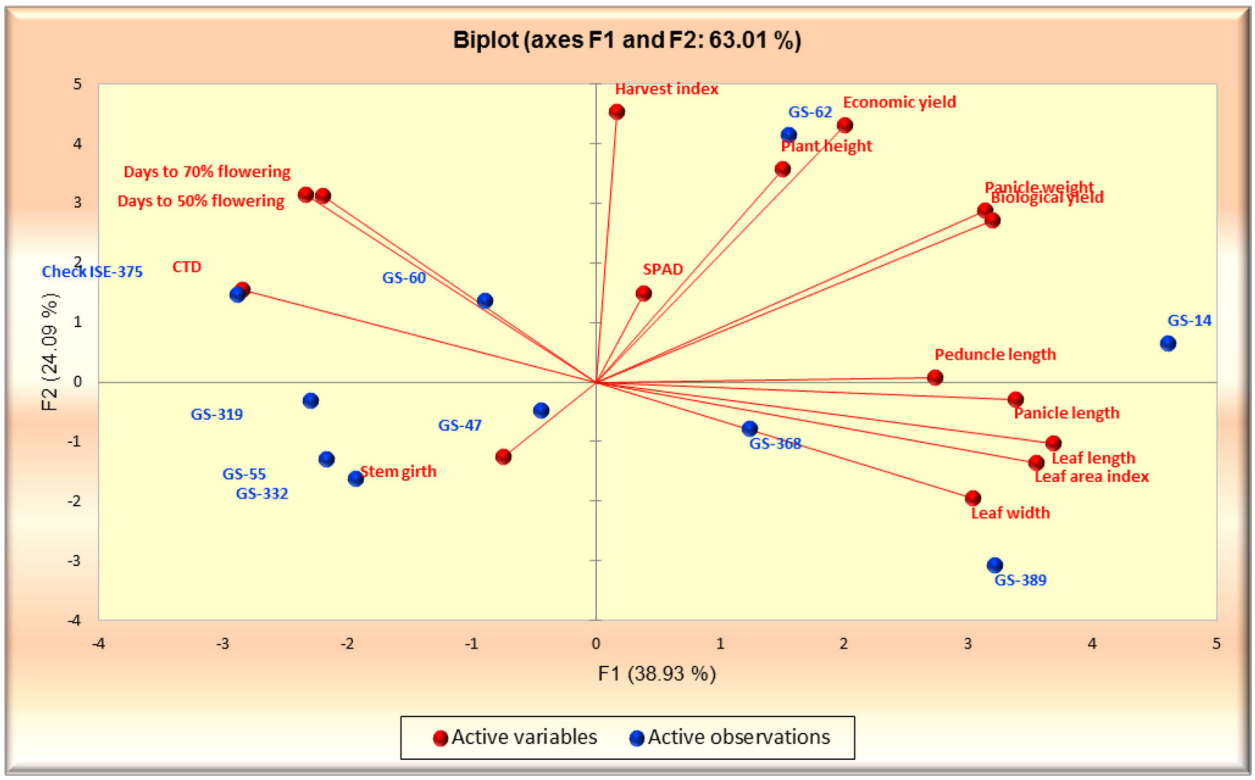


FIGURE 7
PCA biplot series 1 and 2 for 10 foxtail millet accessions collected in 2019.

in any one of them would lead to an improvement of other characters. Selection criteria should consider all these characters for to improve biological yield in foxtail millet. The PCA data reduction technique extracts the most important information from the data table (Islam, 2004), compresses the size of the data set by keeping only the important information (Maji and Shaibu, 2012), simplifies the description of the data set (Adams, 1995), and analyzes the structure of the observations and the variables (Amy and Pritts, 1991). Often, only the important information needs to be extracted from a data matrix, and the number of components that are needed should be considered. This problem can be overcome by using some guidelines. The first procedure is to plot the eigenvalues according to their size and to see whether there is a point in the graph (elbow) such that the slope of the graph goes from steep to flat and keep only the components that occur before the elbow. This procedure is called the scree or elbow test (Cattell, 1966; Jolliffe, 2002). Germplasm evaluation and characterization for plant breeders and multivariate statistical analysis estimate the genotypic and phenotypic parameters. The characteristics described in the list of pre-harvest and post-harvest observations were used for selecting the five best genotypes. PCV values were higher than GCV values, which indicates the effect of the environment on the expression of characters. These results are based on data for 2 years. The genotypes Kangni-1 (GS-14), Kangni-7 (GPF-7), Kangni-6 (GS-55), Kangni-5 (GS-389), and Kangni-4 (GS-368) cannot be found anywhere except SHUATS. That is why these genotypes are named by SHUATS. These five best genotypes will be further analyzed through biochemical trait analysis (Singh et al., 2022).

5 Conclusion

The present study found substantial diversity in the 50 genotypes of foxtail millet investigated for several agro-morphological variables that might be exploited efficiently in crop improvement approaches for diverse traits. According to the findings of this study, plant height and leaf length had the highest PCV estimates, followed by leaf width. Leaf length and 50% flowering in days determines the low GCV and PCV. Furthermore, direct selection based on panicle weight, test weight, and straw weight had a high and positive effect on grain yield per plant in both the rainy and summer seasons, indicating the true relationship between these characters and grain yield per plant, which aids indirect selection for these traits and thus improves grain yield per plant. The top five genotypes were therefore chosen using the pre-harvest and post-harvest attribute observation list. Based on the average performance of the best genotypes in terms of grain yield components in the agroclimatic conditions of Prayagraj, the best five genotypes were Kangni-1 (GS-14), Kangni-7 (GPF-7), Kangni-6 (GS-55), Kangni-5 (GS-389), and Kangni-4 (GS-368). As these findings are based on 2 years of data, biochemical testing of these genotypes validated their consistency.

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.

Ethics statement

All the genotypes were obtained from the NBPGR and ICRICIT for this project by the Directorate of Research at SHUATS.

Author contributions

Conceptualization and methodology, DS and KL; Data curation, DS, KL, and SM; Investigation, DS; Writing—original draft, DS and KL; Writing—review and editing, DS, KL, SM, IB, RL, SE, RC, and RK. All authors have read and agreed the submitted manuscript.

Acknowledgments

The authors wish to thank UGC and the corresponding author for funding the research, the Advisor for their support, and the Directorate of Research at SHUATS Prayagraj.

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

Adams, M. W. (1995). An estimate of homogeneity in crop plants with special reference to genetic vulnerability in dry season. phseolus vulgaris. *Ephytica.* 26, 665–679.

Aghaee-Sarbarzeh, M., and Amini, A. (2012). Genetic variability for agronomy traits in bread wheat genotype collection of Iran. *Seed. Plant Improv. J.* 1, 581–599.

Ahmed, S. M., Qing, Z., Jing, C., and Qun, S. (2013). Millet grains: Nutritional quality, processing, and potential health benefits. *Compr. Rev. Food Sci. Food Saf.* 12, 281–296. doi: 10.1111/1541-4337.12012

Amy, E. L., and Pritts, M. P. (1991). Application of principal component analysis to horticultural research. *Hort. Sci.* 26 (4), 334–338.

Bhat, B. V., Rao, B. D., and Tonapi, V. A. (2018). *ICAR -Indian institute of millets research, Hyderabad karnataka state department of agriculture, bengaluru, India with ICAR-Indian institute of millets research, Hyderabad, India inputs from: Prabhakar, b. boraiah and prabhu c. ganiger (All Indian coordinated research project on small millets* (Bengaluru, India: University of Agricultural Sciences).

Brunda, S. M., Kamatar, M. Y., Hundekar, R., and Kumar, K. L. N. (2015). Studies on correlation and path analysis in foxtail millet genotypes [*Setaria italica* (l.) p. b.]. *Green Farming.* 6 (5), 966–969. doi: 10.17577/IJERTV4IS040075

Brunda, S. M., Kamatar, M. Y., Naveen, K. L., and Hundekar, R. (2014). Study of genetic variability, heritability, and genetic advance in foxtail millet in both rainy and post rainy season. *J. Agri. Vet. Sci.* 7, 34–37. doi: 10.9790/2380-071133437

Cattell, R. B. (1966). The scree test for the number of factors. *Multivariate. Behav. Res.* 1, 245–276.

CSE (2009). *Rainfed areas of India-center for science and environment*. Available at: http://www.cseindia.org/programme/nrml/rainfed_specials.htm.

Eberhart S. A. and Russel, W. L. (1966). Stability parameters for comparing varieties. *Crop Sci.* 6, 34–40.

Fujita, S., Sugimoto, Y., Yamashita, Y., and Fuwa, H. (1996). Physicochemical studies of starch from foxtail millet (*Setaria italica* beauv). *Food Chem.* 3, 209–213.

Islam, M. R. (2004). Genetic diversity in irrigated rice. *Pakistan J. Biol. Sci.* 2, 226–229.

Jolliffe, I. T. (2002). *Principal component analysis* (New York: Springer).

Kavya, P., Sujatha, M., Pandravada, S. R., and Hymavathi, T. V. (2017). Variability studies in foxtail, millet *(Setaria italica (L.) p. beauv)* department of genetics and plant breeding, college of agriculture, 2National bureau of plant genetic resources, regional station, 3Department of food and nutrition PJTSAU, rajendra nagar, Hyderabad-500030, telangana, India original research article. *Int. J. Curr. Microbiol. App. Sci.* 6 (9), 955–960. doi: 10.20546/ijcmas.2017.609.115

Kempthorne, O. (1957). *An introduction to genetic statistics* (New York: John wiley and sons, Inc.).

Kim, J. S., Hyun, T. K., and Kim, M. J. (2010). Anti-oxidative activities of sorghum, foxtail millet and proso millet extracts. *Afr. J. Biotechnol.* 9, 2683–2690.

Kumar, G. V., Vanaja, M., Lakshmi, N. J., and Maheswari, M. (2015). Studies on variability, heritability, and genetic advance for quantitative traits in black gram [*Vigna mungo* (L.) hepper]. *Agric. Res. J.* 52 (4), 28–31.

Maji, A. T., and Shaibu, A. A. (2012). Application of principal component analysis for rice germplasm characterization and evaluation. *J. Plant Breed. Crop Sci.* 4 (6), 87–93.

Mohammadi, S. A., and Prasanna, B. M. (2003). Analysis of genetic diversity in crop plants, salient statistical tools and considerations. *Crop Sci.* 43, 1235–1248.

Nirmlakumari, A., and Vetriventhan, M. (2010). Characterization of foxtail millet germplasm collections for yield contributing traits. *Electronic. J. Plant Breed.* 1 (2), 140–147.

Sapkota, M., Pandey, M., and Thapa, D. (2016). Agromorphological characterisation of foxtail millet (*Setaria italica* L.) *at. Rampur. Chitwan. Nepal. Int. J. Appl. Sci. Biotechnol.* 4, 298–307. doi: 10.3126/ijasbt.v4i3.15148

Singh, D., Hembrom, S., and Raj, A. (2019). Neuroprotective effect of flavonoids: A systematic review, Ph.D. scholar of department of BCBE, JSBB SHUATS, prayagraj, gust faculty of department of microbiology, dr. shyama prasad mukherjee university, Ranchi, jharkhand, journal of pharmacognosy and Phytochemistry8(1): 699-707.doi: 10.13140/RG.2.2.18566.04169.

Singh, D., Lawrence, K., Singh, S., Ercisli, S., and Choudhary, R. (2022). *In-vivo* hyperglycemic, antioxidant, histopathological changes, and simultaneous measurement of kaempferol verified by high-performance thin layer chromatography of *Setaria italica* in streptozotocin -induced diabetic rats. *Saudi. J. Biol. Sci.* 29 (5), 3772–3790. doi: 10.1016/j.sjbs.2022.03.009

Singh, D., Marker, S., Suresh, B. G., and Lawrence, K. (2021). Under rainfed effect of foxtail millet [*Setaria italica* (L.) beauv] germplasms evaluation on genotypic variance, correlation and path analysis. *Indian J. Agri. Res.* 56 (2), 141–146. doi: 10.18805/IJARe.A-5862

Singh, D., and Verma, S. B. (2020). Pre-sowing treatments of maize (Zea mays l.) effect on seed treatments in growth and seed quality on different parameters department ofBCBE, SHUATS, prayagraj, U.P. India;AIOA, amity university noida. *J. Xidian. Univ.* 14. doi: 10.37896/jxu14.6/335

Verma, S. B., Singh, D., and Chauhan, R. N. (2018). Effect of neem (Azadirachta indica l.), mustard (Brassica juncea) de-oiled seed cake and bio-fertilizer on the growth and yield of wheat (Triticum aestivumL.), AIOA amity university noida, uttar pradesh, department of BCBE, SHUATS, prayagraj 211008 India. *J. Pharmacog. Phytochem.* 7 (5), 2416–2427. doi: 10.13140/RG.2.2.17108.24968

Vinsonias, (2018). Available at: http://vinsonias.com/DynImg/4d57c5b8-86ee-491b88336a5191b78be7.pdf.

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 14 April 2023
DOI 10.3389/fpls.2023.1105364

OPEN ACCESS

EDITED BY
María Serrano,
Miguel Hernández University of Elche, Spain

REVIEWED BY
Kuldeep Tripathi,
National Bureau of Plant Genetic Resources, Indian Council of Agricultural Research (ICAR), India
Sergio Medina Godoy,
National Polytechnic Institute (IPN), Mexico

*CORRESPONDENCE
Jacob Olagbenro Popoola
jacob.popoola@bowen.edu.ng
Omena B. Ojuederie
ob.ojuederie@kingsuniversity.edu.ng

SPECIALTY SECTION
This article was submitted to Crop and Product Physiology, a section of the journal Frontiers in Plant Science

RECEIVED 22 November 2022
ACCEPTED 16 March 2023
PUBLISHED 14 April 2023

CITATION
Popoola JO, Ojuederie OB, Aworunse OS, Adelekan A, Oyelakin AS, Oyesola OL, Akinduti PA, Dahunsi SO, Adegboyega TT, Oranusi SU, Ayilara MS and Omonhinmin CA (2023) Nutritional, functional, and bioactive properties of african underutilized legumes. *Front. Plant Sci.* 14:1105364. doi: 10.3389/fpls.2023.1105364

Nutritional, functional, and bioactive properties of african underutilized legumes

Jacob Olagbenro Popoola [1,2]*, Omena B. Ojuederie [3,4]*, Oluwadurotimi Samuel Aworunse [2], Aminat Adelekan [5], Abiodun S. Oyelakin [6], Olusola Luke Oyesola [2], Paul A. Akinduti [2], Samuel Olatunde Dahunsi [7,8], Taofeek T. Adegboyega [4,9], Solomon U. Oranusi [2], Modupe S. Ayilara [3,4] and Conrad A. Omonhinmin [2]

[1]Pure and Applied Biology Programme, College of Agriculture, Engineering and Science, Bowen University, Iwo, Osun, Nigeria, [2]Department of Biological Sciences/Biotechnology Cluster, Covenant University, Ota, Ogun, Nigeria, [3]Department of Biological Sciences, Kings University, Ode-Omu, Osun, Nigeria, [4]Food Security and Safety Focus, Faculty of Natural and Agricultural Sciences, North-West University, Mmabatho, South Africa, [5]Department of Chemical and Food Sciences, College of Natural and Applied Sciences, Bells University of Technology, Ota, Ogun, Nigeria, [6]Department of Pure and Applied Botany, College of Biosciences, Federal University of Agriculture, Abeokuta, Nigeria, [7]Microbiology Programme, College of Agriculture, Engineering and Science, Bowen University, Iwo, Osun, Nigeria, [8]The Radcliffe Institute for Advanced Study, Harvard University, Cambridge, MA, United States, [9]Biology Unit, Faculty of Science, Air Force Institute of Technology, Kaduna, Nigeria

Globally, legumes are vital constituents of diet and perform critical roles in maintaining well-being owing to the dense nutritional contents and functional properties of their seeds. While much emphasis has been placed on the major grain legumes over the years, the neglected and underutilized legumes (NULs) are gaining significant recognition as probable crops to alleviate malnutrition and give a boost to food security in Africa. Consumption of these underutilized legumes has been associated with several health-promoting benefits and can be utilized as functional foods due to their rich dietary fibers, vitamins, polyunsaturated fatty acids (PUFAs), proteins/essential amino acids, micro-nutrients, and bioactive compounds. Despite the plethora of nutritional benefits, the underutilized legumes have not received much research attention compared to common mainstream grain legumes, thus hindering their adoption and utilization. Consequently, research efforts geared toward improvement, utilization, and incorporation into mainstream agriculture in Africa are more convincing than ever. This work reviews some selected NULs of Africa (Adzuki beans (*Vigna angularis*), African yam bean (*Sphenostylis stenocarpa*), Bambara groundnut (*Vigna subterranea*), Jack bean (*Canavalia ensiformis*), Kidney bean (*Phaseolus vulgaris*), Lima bean (*Phaseolus lunatus*), Marama bean (*Tylosema esculentum*), Mung bean, (*Vigna radiata*), Rice bean (*Vigna Umbellata*), and Winged bean (*Psophocarpus tetragonolobus*)), and their nutritional, and functional properties. Furthermore, we highlight the prospects and current

challenges associated with the utilization of the NULs and discusses the strategies to facilitate their exploitation as not only sources of vital nutrients, but also their integration for the development of cheap and accessible functional foods.

KEYWORDS

antioxidants, bioactive compounds, functional food products, under-exploited legumes, sustainable development

1 Introduction

Legumes are a group of flowering plants and are classified under the Fabaceae family. This family is the third-largest in terms of angiosperm groups, consisting of over 800 different types and around 20,000 species. Within the Fabaceae family, there are three subfamilies known as Papilionoideae, Caesalpinioideae, and Mimosoideae. Of these, the edible legumes are grouped in the subfamily Papilionoideae. Globally, legumes are regarded as a valuable and inexpensive alternative protein sources and rank second after cereals as the most important food crop (Maphosa and Jideani, 2017). Apart from the rich protein and amino acid content, legume seeds provide a substantial amount of carbohydrates, minerals, and vitamins (Vadivel et al., 2012; Barman et al., 2018). In addition to having no cholesterol and gluten, legumes possess low fat and glycemic index and are rich in dietary fiber and antioxidants. These legumes possess bioactive compounds which possess antidiabetic, antimicrobial, anti-atherogenic, anti-thrombogenic, anti-hypertensive, and anticancer properties amongst others. Legumes also serve as fodder for livestock and fix atmospheric nitrogen in soils, thereby enhancing soil fertility and invariably promoting agricultural sustainability. They are also adapted to diverse agro-ecological zones and unfavorable environmental conditions, possessing structures for augmenting the sustainability of dry subtropical and tropical agricultural systems (Khoury, 2015).

It is interesting to note that some legumes also produce underground tubers in addition to edible seeds. However, only a few of these legumes are incorporated into the human diet. Such dual food legumes fall into the category of neglected and underutilized legumes (NULs) simply because they have not received much research focus and are still cultivated at the subsistence level by resource-poor farmers who hold the genetic resources of these plants. Tuberous underutilized legumes are gradually gaining recognition. These include the African yam bean (AYB) (*Sphenostylis stenocarpa*) cultivated in West Africa for the seeds and in East and Central Africa for the tubers (Adewale and Nnamani, 2022); winged bean (*Psophocarpus tetragonolobus*), grown and cultivated in Papua New Guinea Highland, northern Ghana, and Burma; the Marama bean (*Tylosema esculentum*) cultivated in the Southern Africa regions of Botswana, Namibia, Mozambique, Zambia, and in northern South Africa (Abberton et al., 2020a; Abberton et al., 2020b; Ojuederie et al., 2021; Sriwichai et al., 2021); Mexican yam bean (*Pachyrhizus erosus*); Zombi pea (*Vigna vexillata*) an underutilized legume with a pantropical distribution; hyacinth bean (*Lablab purpureus*) grown in North Africa; as well as Tala (*Neoapaloxylon tuberosum*) cultivated in Madagascar (Von Wettberg et al., 2021). Different tuber shapes and sizes of some tuberous underutilized legumes are presented in Figure 1.

The Bambara groundnut (*Vigna subterranea*) is a crop that is extensively grown for its seeds in certain regions of West and Southern Africa. Nigeria has been reported to be the largest producer of this crop (Ojuederie et al., 2021; Popoola et al., 2022b; Arise et al., 2022). Tubers of Zombi peas are crispy, rich in protein (15%) and can be consumed raw (Tripathi et al., 2020; Von Wettberg et al., 2021). The seeds and tubers of many of the NULs are also rich in protein. For instance, AYB seeds contain 19.5% protein, while the tubers hold about 15.5% protein (Ojuederie and Balogun, 2017; Ojuederie and Balogun, 2019; Abberton et al., 2020a). In winged bean, the protein content of the seeds and tubers are 29.8% to 42.5% and 20% respectively (Abberton et al., 2020b). Negi and Gaur (1994) stated that Zombi peas contain 14.5% protein when their roots are dried. Nevertheless, a more recent study conducted by Tripathi et al. (2020) to analyze the nutritional content of seven different Zombi peas accessions found that the protein content of their tubers ranged from 7.64% to 9.93%. This is remarkable because it was seven to nine times higher than the protein content found in sweet potato and cassava tubers (Tripathi et al., 2020). Although Zombi peas is not considered in this review, its rich nutritional contents particularly the tubers call for more research attention (Tripathi et al., 2020). The edible tubers of AYB and winged beans are still propagated at the subsistence level with no genetic improvement. The mechanism behind tuberization in AYB is yet to be understood.

Bioactive compounds have been identified in NULs, but with little or no impact on the nutritional and food security in Africa. In plants, bioactives perform several functions, ranging from protection against herbivores and insect pest to attraction of pollinators during pollination and induction of essential functions (Chandrasekara and Josheph Kumar, 2016; Divekar et al., 2022). These bioactive compounds also exhibit pharmacological properties in humans and animals (Chandrasekara and Josheph Kumar, 2016) which forms a major part of this review. The bioactive components of the NULs are yet to be fully harnessed for improved health and well-being as many consumers in Africa are unaware of their nutritional and health benefits. In our previous review, we emphasized the need to integrate

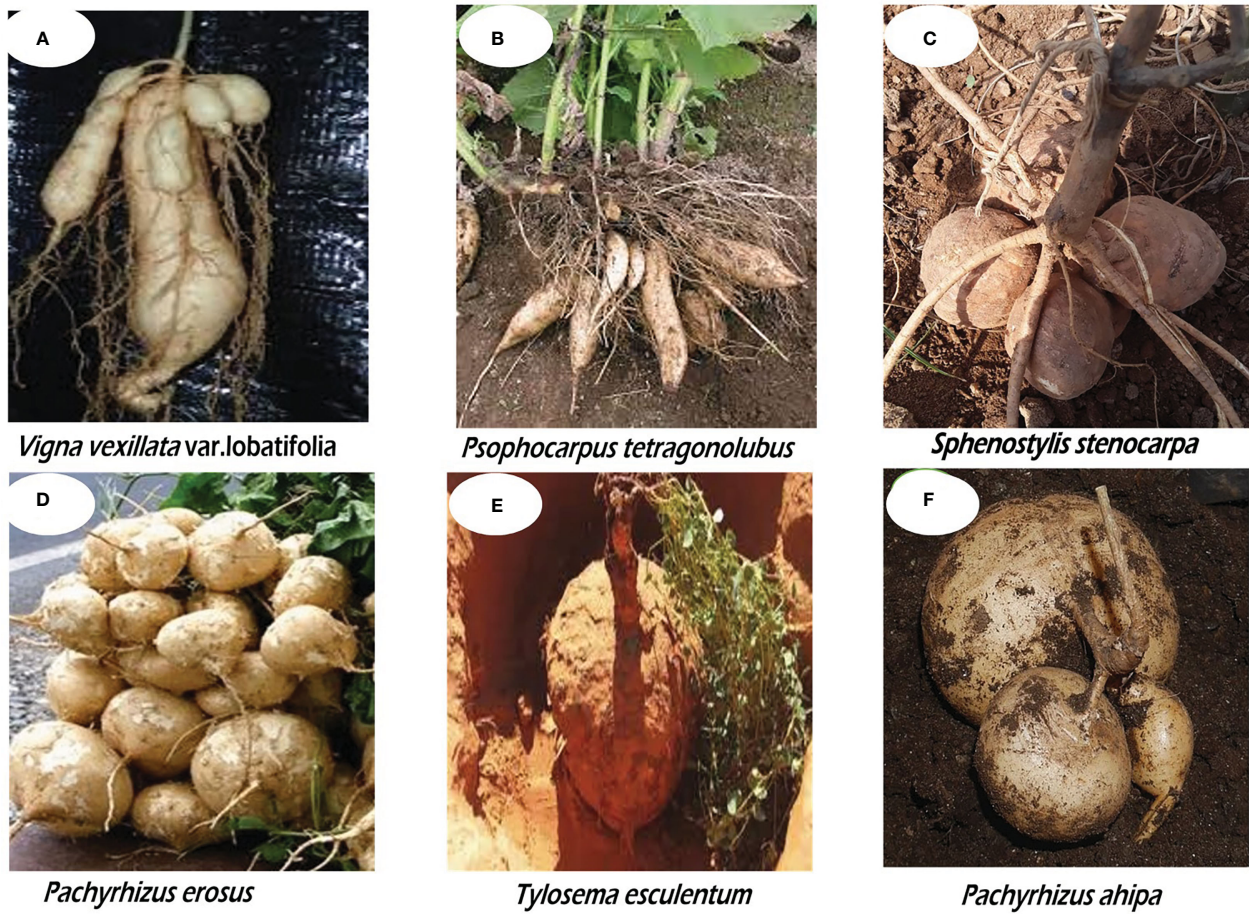


FIGURE 1
Different tuber shapes and sizes of some underutilized legumes: **(A)** Zombi pea (*Vigna vexillata*), **(B)** Winged bean (*Psophocarpus tetragonolobus*), **(C)** African yam bean (*Sphenostylis stenocarpa*), **(D)** Mexican yam bean (*Pachyrhizus erosus*), **(E)** Marama bean (*Tylosema esculentum*), **(F)** Ahipa (*Pachyrhizus aphipa*).

the NULs into food systems in sub-Saharan Africa (SSA) to cushion the negative effects of climate change, soil degradation, poverty, food insecurity, and malnourishment (Popoola et al., 2022b). This article attempts to present a broad review of some selected NULs, and their nutritional and functional properties. The selected NULs include Adzuki beans (*Vigna angularis*), African yam bean (*Sphenostylis stenocarpa*), Bambara groundnut (*Vigna subterranea*), Jack bean (*Canavalia ensiformis*), Kidney bean (*Phaseolus vulgaris*), Lima bean (*Phaseolus lunatus*), Marama bean (*Tylosema esculentum*), Mung bean, (*Vigna radiata*), Rice bean (*Vigna Umbellata*), and Winged bean (*Psophocarpus tetragonolobus*). In Africa, these NULs have been relegated to the status of "poor man's food" with abysmally low level of cultivation, production, consumption, and utilization compared to the mainstream legumes. Consequently, the need to create awareness about their potential utility, health and nutritional benefits becomes imperative. Also, the relevance of the untapped bioactive compounds inherent in the seeds of these potential food and nutrition security crops are discussed. Furthermore, we highlight the prospects and current challenges associated with the utilization of these NULs and present strategies to facilitate their exploitation as not only sources of vital nutrients, but also integration for the development of cheap and accessible functional foods. The plant products and distribution of the selected underutilized legumes and center of diversity in Africa are presented in Table 1 and Figure 2.

2 Neglected and underutilized legumes

The term "neglected" or "underutilized" alludes to a class of legumes that are climate-smart, adapted to marginal areas, indigenously propagated with fewer or no *ex-situ* collections, and have not given priority by policymakers. The term also refers to legumes that have received little research attention, possess local significance in production and consumption, traded regionally or internationally and are usually cultivated on a small scale by rural families for subsistence, particularly under adverse environment conditions (Cullis and Kunert, 2017; Yang et al., 2018; Rathi et al., 2021; Popoola et al., 2022b). While these crops have received relatively little research and funding, their potential is well recognized (Cullis and Kunert, 2017; Yang et al., 2018). The NULs exhibit an array of genetic diversity and exist as wild or

TABLE 1 Plant products and distribution of the selected underutilized legumes in Africa.

S/N	Common name	Botanical name	Seeds	Tuber	African Countries
1.	Adzuki beans	*Vigna angularis*	Present	Absent	DR Congo, Kenya, Angola, Zambia, Madagascar, Seychelles
2.	African yam bean	*Sphenostylis stenocarpa*	Present	Present	Nigeria, Ghana, Benin Republic, Cameroun, Togo, Niger, Kenya, Ethiopia, Mozambique, Tanzania.
3.	Bambara groundnut	*Vigna subterranea*	Present	Absent	Nigeria, Ghana, Niger, Mali, Côte d'Ivoire, Benin Republic, South Africa, Kenya.
4.	Jack bean	*Canavalia ensiformis*	Present	Absent	Western, Eastern, and Northern Africa
5.	Kidney bean	*Phaseolus vulgaris*	Present	Absent	Western, Eastern and Northern Africa
6.	Lima bean	*Phaseolus lunatus*	Present	Absent	Western, Eastern and Northern Africa
7.	Marama bean	*Tylosema esculentum*	Present	Present	Botswana, Namibia, Mozambique, Zambia, and northern South Africa
8.	Mung bean	*Vigna radiata*	Present	Absent	Nigeria, Liberia, Sierra Leone, Ghana, Côte d'Ivoire, and DR Congo
9.	Rice bean	*Vigna Umbellata*	Present	Absent	Egypt, Kenya, Tanzania, Burundi, Somalia, Rwanda
10.	Winged Bean	*Psophocarpus tetragonolobus*	Present	Present	Papua New Guinea Highland, northern Ghana, and Burma Nigeria, Togo, Benin

cultivated species across different regions of the world (Agbolade et al., 2019). These crops are primarily grown by traditional farmers in SSA, Asia, and North America (Alvarado-Lopez et al., 2019). The NULs are marked by unique characteristics such as ethno-uses, seed sizes, growth habits, and fruiting patterns that distinguish them from the common pea. Furthermore, they are of agricultural importance owing to their capability to augment soils *via* symbiotic nitrogen fixation (Agbolade et al., 2019; Hunter et al., 2019; Popoola et al., 2022b). Underutilized legumes are a great source of essential nutrients such as dietary fiber, vitamins, polyunsaturated fatty acids (PUFAs), proteins with essential amino acids, minerals, and bioactive compounds. These legumes are therefore considered functional foods that can have positive effects on our health (Popoola et al., 2020; Rai et al., 2021).

African yam bean, Bambara groundnut, lablab bean, lima bean, and the winged bean are examples of some commonly cultivated underutilized leguminous species in SSA (Agbolade et al., 2019; Popoola et al., 2020). These legumes have the potential to drive sustainable agri-food systems in the region given their diversity, climate resilience, nutrient-dense nature, and cultural attachment to the regional food habits of the communities of origin (Paliwal et al., 2021). A recent analysis of the utilization and cultivation of lesser known legumes brought attention to the significant traits and potential prospects of some of the crops (Popoola et al., 2022b).

FIGURE 2
Center of the diversity of some underutilized legumes in Africa Source: http://www.zipcodezoo.com/Plants/S/Sphenostylis%5Fstenocarpa/Default.asp (February 19, 2023).

3 Nutritional properties of NULs

Nutritionally, the seeds of the selected NULs are rich in proteins, carbohydrates, minerals (calcium, manganese, phosphorus amongst others), and a wide range of vitamins (Table 2). Nutritional information on the tubers is scanty and has not been studied in detail compared to the seeds (Ojuederie and Balogun, 2019; Tripathi et al., 2020). Nevertheless, a few reports on AYB, Zombi pea, and winged beans indicate that the tubers are nutritionally rich with varied carbohydrates, proteins, ash, dietary fiber, minerals, and vitamin contents when compared to those of cassava, potato, and yam (Adegboyega et al., 2019; Konyeme et al., 2020; Ojuederie et al., 2020; Tripathi et al., 2020). The studies of Konyeme et al. (2020) emphasized the rich nutritional value of the tubers, while the investigation of Ojuederie et al. (2020) confirmed their safe consumption by humans and livestock. Globally, the demand for food legumes is ever-increasing as one of the essential nutritional and conventional food with health and pharmacological relevance (Tadele and Bartels, 2019). The nutritional contents of the seeds are diverse and have been widely reported and discussed by many researchers (Gagné-Bourque et al., 2016; Huang et al., 2018; Omotayo and Aremu, 2021; Popoola et al., 2022a). The nutritional profile of the seeds varies in different accessions of the same and different species. This can be exploited by breeders to enhance yield, taste and value chain for the food and confectionary industries. The amount of carbohydrates found in the selected underutilized

legumes range from 18.90g in Marama bean (MB) to 70.48g in Kersting's groundnut (KG) (Table 2). In legumes, carbohydrates usually contain resistant starch sugars such as stachyose, raffinose, and fructooligosaccharides. These sugars have the ability to improve the microbial environment in our gastrointestinal tract and promote gut metabolism (Johnson et al., 2020). Adding such components to food systems can greatly improve health and ensure nutritional quality. The proteins are of high quality and range from 7.80g in Lima beans (LB) to 29.60g in winged beans (WG) while others also contain a good quantity of protein. The amino acids found in these legumes are valuable in boosting the immune system, regulating metabolic processes, and enhancing glucose and fatty-acid metabolism (Tjahjadi et al., 1988; Semba et al., 2021; Ayilara et al., 2022). Numerous studies, including those conducted by Gohara et al. (2016), Nnamani et al. (2017), Baiyeri et al. (2018), Adegboyega et al. (2020), and Abberton et al. (2022), have documented the functions of various vitamins and minerals. Selected underutilized legumes (NULs) have been found to contain

TABLE 2 Nutritional composition of the selected underutilized legumes' raw, mature seeds, with values per 100g.

Nutrient contents	ADB	AYB	BG	JB	KB	LB	MB	MGB	RB	WB
Protein (g)	20.36	22.46	18.8	20.90	23.00	7.80	34.71	23.80	20.50	29.60
Carbohydrate (g)	62.26	53.68	61.30	60.61	70.48	20.90	18.9	61.00	51.31	41.70
Moisture (g)	13.07	9.53	2.10	2.19	Nr	Nr	2.80	9.80	Nr	Nr
Ash (g)	3.85	4.28	2.40	3.45	5.13	1.15	3.19	3.51	Nr	3.98
Fat (g)	0.45	3.59	6.20	1.59	1.38	0.38	40.06	1.22	0.60	16.30
Total Dietary Fibre	7.30	7.30	5.50	3.98	20.93	7.00	50.81	4.57	13.10	25.90
Water (g)	13.40	61.50	10.30	Nr	59.00	69.80	Nr	Nr	Nr	8.34
Energy (kcal/100 g)	334.80	333.67	367.00	338.00	386.39	115.00	544.57	344	318.00	409.0
Folates (B9) μg	0.62	0.10	0.25	0.40	1.11	83.00	0.14	0.62	Nr	45
Thiamine (B1) (mg/100g)	0.46	0.19	0.61	0.51	0.50	0.161	0.38	0.62	Nr	1.03
Niacin (B3) (mg/100g)	2.63	0.07	1.80	1.54	0.51	0.421	0.06	2.25	Nr	3.09
Riboflavin (B2) (mg/100g)	0.22	0.20	0.31	0.20	0.03	0.055	0.06	0.23	Na	0.45
Vitamin B6 (mg/100g)	0.35	0.10	0.44	0.51	4.67	0.161	9.21	0.38	Na	0.175
Vitamin A (mg/100g)	11.39	Nr	Nr	Nr	0.00	0.00	0.27	200	Na	Nr
Vitamin C (mg/100g)	0	12.97	0.27	8.10	0.55	0.00	0.00	4.80	Na	0
Vitamin D (mg/100)	0.00	0.00	3.42	0.00	0.00	0.00	132.9	0.00	Na	Nr
Vitamin E (mg/100g)	Nr	0.19	Nr	Nr	0.10	0.18	6.27	0.51	Na	Na
Vitamin K ((mg/100g)	Nr	Nr	0.001	Nr	14.90	2.00	0.22	9.00	Na	Na
Pantothenic acid (B5) (mg/100g)	1.47	Nr	1.80	Nr	0.40	0.422	Nr	1.91	Na	0.795
Sodium (mg/100g)	5.00	1.00	3.60	2.53	53.48	2.00	63.75	15.00	32.00	38
Calcium (mg/100g)	66.00	15.00	1.60	3.21	104.12	17.00	241.00	216	340.00	440
Copper (mg/100g)	1.09	0.29	0.09	0.43	0.40	0.235	1.04	1.27	1.12	2.88
Iron (mg/100g)	4.98	1.50	5.52	0.83	7.00	2.39	3.95	6.74	5.80	13.40
Magnesium (mg/100g)	127	69.00	7.58	1.95	118.95	43.00	274.50	204		179
Manganese (mg/100g)	1.73	3.35	0.26	0.35	0.80	0.516	1.85	1.23	0.68	3.72
Phosphorus (mg/100g)	381	99.00	32.50	1500	251.30	111.00	454.00	374	Na	451
Potassium (mg/100g)	1254	419.00	183.00	5.93	1517.36	508.00	895.00	1443	Na	977
Zinc (mg/100g)	5.04	0.78	0.27	2.90	2.38	0.95	6.20	1.88	3.39	4.48
Selenium (mg/100g)	3.10	150.00	Nr	Nr	2.10	4.50	0.08	8.2	Nr	8.20
Beta-carotin (μg)	Nr	7.00	0.47	Nr	Nr	0.00	Nr	68	Nr	Nr

ADB, Adzuki beans (Vigna angularis); AYB, African yam bean (Sphenostylis stenocarpa); BG, Bambara groundnut; (Vigna subterranea), JB, Jack bean (Canavalia ensiformis); KB, Kidney bean (Phaseolus vulgaris); LB, Lima bean (Phaseolus lunatus); MB, Marama bean (Tylosema esculentum); MGB, Mung bean (Vigna radiata); RB, Rice bean (Vigna Umbellata); WB, Winged Bean (Psophocarpus tetragonolobus); Nr, Not reported; Na, Not available.
Values adopted from the United States Department of Agriculture (USDA), Baiyeri et al. (2018), Charrondière et al. (2020) and Nnamani et al. (2018).

thiamine, niacin, riboflavin, vitamins A, B6, C, D, E, K, and pantothenic acid, according to these studies (Table 2). Underutilized legumes have been found to contain various mineral elements such as sodium (Na), calcium (Ca), copper (Cu), magnesium (Mg), manganese (Mn), phosphorus (Ph), potassium (K), and zinc (Zn) (Baiyeri et al., 2018; Nnamani et al., 2018; Adegboyega et al., 2019; Adedayo et al., 2021). The vitamins and minerals are required for optimal health and growth, improved memory, and blood circulation. Nevertheless, allergenicity, digestibility, and antinutritional factors (ANF) are major constraints to their functional utilization. However, various methods such as steaming, boiling, fermentation, irradiation, and high-pressure cooking have been found to overcome these challenges, as indicated in studies conducted by (Maphosa and Jideani, 2017; Bessada et al., 2019; Tan et al., 2020). Despite this, underutilized legumes, especially tubers, have not been fully utilized and their nutritional content has not been fully exploited (Ojuederie et al., 2020; Ayilara et al., 2022; Popoola et al., 2022a). The nutritional composition of the selected underutilized legumes' raw, mature seeds, with values per 100g, is presented in Table 2.

4 Functional properties and probiotics of underutilized legumes

4.1 Functional properties of underutilized legumes

Underutilized legumes possess noteworthy functional properties which are beneficial to food systems. Functional properties such as solubility, hydration, emulsification, foaming stability, gel-forming index, and pasting properties govern the utilization of legumes as protein-rich gluten-free food additives. The functional properties of these substances should be taken into account when formulating and processing food, to develop innovative food products (Bessada et al., 2019). Moreover, the functionality of legumes is affected by protein including its molecular size, structure, and charge distribution as well as non-protein molecules such as carbohydrates, lipids, and salts.

The functional properties of the underutilized legumes considered in this review are presented in Table 3. The hydration properties of flour are swelling power, solubility, water, and oil absorption capacities. These properties influence the structural, rheological, thermal, and sensory characteristics of foods. Swelling power is a measure of both intragranular and intergranular water present in flour/starch under excess water and high thermal conditions (Iuga and Mironeasa, 2020). According to recent research by (Palupi et al., 2021), lima bean flour has a swelling power of 6.88 g/g, a solubility of 18.68%, a water absorption capacity of 1.93 g/g, and an oil absorption capacity of 1.56 g/g when exposed to a temperature of 87°C. The African yam bean seed flour displayed good gelation properties, while protein solubility varied with pH, with high solubilities in acid and alkali (Oshodi et al., 1997) (Table 2). Ratnawati et al. (2019) revealed that at 95°C, mung bean and red bean flours had significantly higher swelling powers (10.5 and 10.1 g/g, respectively) than soybean flour (4.8 g/g). The research of Yadav et al. (2018) indicated that adzuki beans had a hydration capacity range of 0.05 to 0.12 g/seed, a swelling capacity of 0.04 to 0.15 mL/seed when the cooking time is reduced (48.67 to 74.33 min). Typically, flour with high swelling power elicits high gelatinization and paste properties which is vital for the structural and textural development of baked foods (Onyango, 2016). In addition, the solubility of flour or starch provides an index of the hydrophilic behavior of amylose molecules under high moisture and thermal conditions (Dudu et al., 2019). Flour solubility has an impact on the clarity of drinks as well as foam formation and stabilization. It also affects emulsification, gelation, and retrogradation which influences crumb grain formation, texture, sensory properties, and staling of baked foods. Furthermore, water and oil absorption capacities are the maximum amount of intragranular water or oil present in flour under excess moisture and ambient temperature conditions (Iuga and Mironeasa, 2020). Emulsification properties consist of emulsifying activity and stability which are modulated by the ratio of hydrophobic to hydrophilic amino acids present in the legume flour. Emulsifying activity is a measure of the ability of flour to form a stable emulsion (oil-water interaction) by protein dispersion in the presence of oil. On the other hand, stability measures the strength of the emulsion formed (Nawaz et al., 2021). Most of the underutilized legumes have been reported to show good functional properties of solubility, emulsification, oil absorption capacity, gelation and forming properties (Okezie and Bello, 1988; Onimawo et al., 1998; Barac et al., 2015; Arogundade et al., 2016; Mubaiwa et al., 2018; Diedericks et al., 2020; Yang et al., 2022).

Foaming properties include foaming capacity and stability. Foaming capacity is the ability of protein or flour to add air when whipped, and foam stability is the capacity to stabilize foams (volume) over time (≤30 min) (Bessada et al., 2019). Iwe et al. (2016) revealed that African yam bean flour has a foaming capacity and stability of 18% and 92.6%, respectively. Pasting properties, mainly pasting temperature, peak, breakdown, and setback viscosities provide insights into swelling capacity, structural stability, and amylose retrogradation tendency of flour under combined high mechanical shearing and hydrothermal conditions. This is typically carried out in a rapid viscosity analyzer or a Brabender Visco-Amylo-Graph. These properties are critical to the functionality of flour in food and industrial systems. Pasting temperature is the thermal energy required to destroy flour granule structures leading to the onset of paste development. Associations between flour composition, pasting, viscosity, and bulk density of some underutilized legumes have been investigated (Du et al., 2014). Furthermore, some studies on legume flour highlighted the relationship of flour microstructure with pasting even though such studies are scanty for the underutilized legumes (Shevkani et al., 2021). The pasting properties of the processed lima bean flour showed a peak of 1172 cP, a breakdown of 83 cP, final of 2377 cP, and a setback viscosity of 1288 cp (Palupi et al., 2021). Ratnawati et al. (2019) revealed that the pasting temperature of red kidney bean flour (81.7 °C) was higher than that of mungbean flour (77.9 °C). Flours with low breakdown viscosity can serve as structuring agents in food where a minimal structural breakdown is required. Setback viscosity

TABLE 3 Functional properties of the selected underutilized legumes considered in this review.

Legume flour	Swelling power	Solubility	Water absorption capacity	Oil absorption capacity	Foam capacity	Emulsifying capacity	Gel Formation	Pasting Temperature	Peak Viscosity	Breakdown Viscosity	Setback Viscosity	Bulk Density	References
Adzuki bean	0.04 to 0.15 mL/seed	–	281.35%	252.27%	–	–	–	–	–	–	–	0.76 - 1.00 g/mL	(Yadav et al., 2018)
African yam bean	4.98 g/mL		2.01 g/mL	2.07 g/mL	5.01%	2.61 g/mL		83.55 °C	1108 BU	65.50 BU	768 BU		(Nonye and Chinasa, 2022)
	–	–	0.70%, 131.9% - 218.8%	1.48%	18.0%, 40.2%	56.67%, 50.7%	14.2%	–	–	–	–	0.63-0.87 g/mL	(Iwe et al., 2016; Obatolu et al., 2001)
Bambara groundnut			1.62-2.38 g/g	2.29-2.82 g/g		46%-55%		84-85.63 °C	892.50-1320.50 cP	98-157 cP	1414-1647 cP	0.58-0.71 g/mL	(Falade and Nwajei, 2015)
Jack bean		Temperature-dependent	Temperature-dependent	–				60 – 70°C					(Akinyemi et al., 2020)
Kidney bean			1.7- 2.7 g/g	1.4-1.7 g/g				79.2- 84.3°C	372-1015 cP	30- 67 cP	658- 1428 cP	1088-2385 cP	(Shevkani et al., 2022)
Lima bean	6.88 g/g	18.68 g/g	1.93 g/g	1.56 g/g					1172cP	83 cP	2377 cP	1288 cP	(Palupi et al., 2021).
Marama bean	–	6.6 g/g	1.50 g/g	2.7 g/g	31.1%	59.9%							(Maruatona et al., 2010)
Mung bean	10.50 g/g	18.80 g/g						77.9 °C	90.9 cP	40 cP	518.7 cP		(Ratnawati et al., 2019)
Rice bean	–	–	–	–	–	–	–	–	–	–		–	–
Winged bean	–	–	0.66 g/g	1.63g/g	13.67%	3.42 m2/g	–	–	–	–	–	–	

10.3389/fpls.2023.1105364

is the recovery of viscosity during the cooling of flour after being subjected to combined high mechanical shearing and hydrothermal conditions (Cornejo-Ramírez et al., 2018). It is an index of the retrogradation ability of flour, which is an important prerequisite for staling activity during product storage. Flours with low setback viscosity may be utilized in delaying retrogradation activity in cooked infant formulas, breakfast foods, and pasta products as well as prevent staling of baked products.

Bulk density is an index of the structural integrity of granules that relates to the packaging and raw material handling of flour (Adedayo et al., 2021). Iwe et al. (2016) revealed that AYB has a loose bulk density of 0.63 g/mL, repacked bulk density of 0.87g/mL, water absorption of 0.70%, and oil absorption of 1.48%, all of which are comparable to that of cowpea and rice. Bulk density in adzuki beans range 0.76 to 1.00 g/mL (Yadav et al., 2018). Nwajagu et al. (2021) showed that *M. flagellipes* seed flour has a bulk density of 0.8 mg/100g. Okechukwu-Ezike et al. (2020) revealed that black-eyed beans and black beans have higher bulk densities (0.6 g/cm3) than brown beans (0.4 g/cm3). Sattar et al. (2017) showed that black gram (*Vigna mungo*) flour, green gram (*Vigna radiata*) flour, and lentil (*Lens culinaris*) flour have bulk densities of 0.5 g/cm3, 0.5 g/cm3, and 0.6 g/cm3, respectively. Flours with high bulk density are suitable as thickeners, while those with low bulk density can be in complementary food formulas. The functional properties of the underutilized legumes considered in this review are presented in Table 3.

4.2 Probiotics and prebiotics potential of underutilized legumes

According to The International Scientific Association for Probiotics and Prebiotics (ISAPP), "Probiotics are live microorganisms, which when administered in adequate amounts, confer a health benefit on the host" (Marco et al., 2021). Moreover, prebiotics are "substrates that are selectively utilized by host microorganisms conferring a health benefit" (Sanders et al., 2019). In time past, the health benefits of probiotics were realized from the consumption of milk, soybean, and other dairy products. However, the problem of short shelf-life, allergenic milk proteins, high cholesterol content, lactose intolerance, consumer inclination towards veganism, and economic considerations for developing countries, have compelled the exploration of non-dairy alternatives with good nutritional profile and health-promoting factors (Panghal et al., 2018; Chaturvedi and Chakraborty, 2021; Chaturvedi and Chakraborty, 2022). In the last decade, the non-dairy food products market has received positive acceptance and is projected to reach an estimated 26 billion USD by the year 2025. The underutilized legumes hold an exceptional capacity to be utilized as probiotic carriers (Rasika et al., 2021; Chaturvedi and Chakraborty, 2022). These legumes constitute appropriate matrices for the production of non-dairy alternatives like plant-based beverages due to the presence of natural prebiotics including resistant starch, oligosaccharides, isoflavones, and polyphenols. These prebiotics exert a wide range of physiological functions such as immune system modulation, metabolic regulation, and anti-inflammatory and anti-cancer properties, and therefore offer enormous potential for the development of symbiotic foods (a blend of prebiotics and probiotics) using lactic acid bacteria (LAB) (Chaturvedi and Chakraborty, 2022; Cichońska and Ziarno, 2022).

Generally, research findings have shown that underutilized legumes such as adzuki beans, African yam bean, Bambara groundnut, and mung beans exhibit a low glycemic index due to their high resistant starch and amylose contents and have been shown to reduce the risks of high blood pressure and type-2 diabetes (Barac et al. 2015; Johnson et al., 2020; Adedayo et al., 2021; Johnson et al., 2022). Adeoye et al. (2021) reported the probiotic nutraceutical potential of Bambara groundnut. In the study, *Lactobacillus delbrueckii*, *L. casei* and *L. brevis* were the preponderant LAB found in isolates of fermented Bambara groundnut. The authors further demonstrated the *in vitro* antagonistic properties of the LAB isolates against pathogenic namely *Salmonella sp, Escherichia coli, Staphylococcus sp, Shigella sp and Pseudomonas* sp. Very recently, Chaturvedi and Chakraborty (2022) conducted a study to evaluate the prebiotic characteristics of synbiotic drinks made from legumes, specifically red kidney beans and green mung beans. The results indicated that these drinks had a considerable impact on promoting the growth of the probiotic Lactobacillus casei ATCC 335, while simultaneously hindering the colonization of the enteric pathogen *Escherichia coli*. The study established that the formulated beverage showed prebiotic and probiotic potentials that could serve as a veritable alternative to dairy symbiotic beverages. Thus, the functional and probiotic properties elicited by the above-mentioned underutilized legumes serve as a point of reference for their suitability and/or exploitation in food, industrial and pharmaceutical systems. However, more research is required in this regard to provide more insight into the probiotic and prebiotic potentials of underutilized legumes.

5 Bioactive components of underutilized legumes

Bioactives are compounds that when ingested by humans/animals have some physiological contributions that could enhance healthy living and support a decrease in the occurrence of illness (James et al., 2020; Cui et al., 2021). Commonly, legumes including the underutilized ones are rich in polyphenols, alkaloids, saponins, carotenoids, terpenoids, omega-3 fatty acids, flavonoids, and anthocyanins amongst others (Figure 3). These substances have varying degrees of abilities to act as antioxidants, antimicrobials, anticancer agents, anti-tumor agents, anti-inflammatory agents, and neuroprotective agents (Zheng et al., 2019; Shevkani et al., 2022; Popoola et al., 2022a). Furthermore, it has been suggested that polyphenols, alkaloids, and saponins play a vital role in protecting the plant from herbivores and pathogens by serving as defense mechanisms. They also act as signaling molecules between the plant and its biotic environment (Divekar et al., 2022).

Several authors such as Oboh et al. (2015); Ajibola and Olapade (2016); Soetan (2017); Soetan and Atanda (2018); Soetan and Adeola (2018); Soetan et al. (2018); Adegboyega et al. (2019); Mayes et al. (2019); Ojuederie and Balogun (2019); Adegboyega

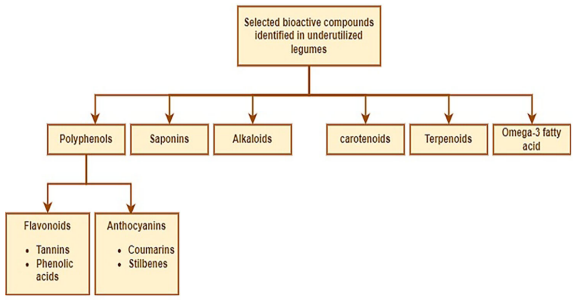


FIGURE 3
Selected bioactive compounds in underutilized legumes.

et al. (2020); Adegboyega et al. (2021); Ikhajiagbe et al. (2021); Ojuederie et al. (2021) and Popoola et al. (2022a) have attempted to unravel the bioactive ingredients available in underutilized legume crops such as African yam bean, Bambara groundnut, Kersting's groundnut, and winged bean which can be exploited as nutraceuticals. In general, the seeds contain bioactive compounds that can improve human health and provide several benefits, such as aiding digestion, promoting weight loss, and reducing the risk of heart diseases and type 2 diabetes (Alcázar-Valle et al., 2020; Bhadkaria et al., 2021). In addition to the dietary fiber, polyphenols, and natural antioxidants embedded in the seeds are of vital benefits to defend against free radicals (Amarowicz and Pegg, 2008; Xu et al., 2017). These components are urgently needed to be exploited for the benefit of man and animals particularly in the management of degenerative infections (Silva et al., 2007; Singh et al., 2017). The promising nature of these compounds (Figure 3) has afforded researchers to look into the possibilities of developing food-based therapy for disease management but more actions are still required in this direction (James et al., 2020).

Polyphenols and their derivatives such as flavonoids, anthocyanins, tannins, and tocopherol are among the essential bioactive compounds found in underutilized legumes. In a comparative study of antioxidants produced by Bambara groundnut (BG) using methanolic extract, Salawu (2016) identified the presence of several polyphenols in the raw and cooked BG seeds after quantification using HPLC-DAD. A recent study identified catechin, epicatechins, rutin, quercetin, isoquercetin, kaempferol, luteolin, gallic acid, chlorogenic acid, caffeic acid, and ellagic acid as polyphenols found in BG (Okafor et al., 2022). Likewise, Harris et al. (2018) compared the level of flavonoids in red and brown BG hulls and observed that the brown hull had the highest amount of rutin (24.46 ± 0,23 mg g^{-1}) and myricetin (1.80 ± 0.77 mg g^{-1}) while the phytochemicals chlorogenic acid and ellagic acids which are tannins, had their highest concentrations in red BG hulls (0.12 ± 0.19 mg g^{-1}) and brown BG (0.11 ± 0.08 mg g^{-1}) respectively (Harris et al., 2018). Their findings revealed that the best source of flavonoids and tannins were found in the brown and red hulls rather than in the whole or dehulled BG seeds.

Apart from producing antioxidants, polyphenols are also known to possess anti-microbial, anti-viral, anti-inflammatory, anti-allergic as well as anti-mutagenic effects, scavenging free radicals that cause cell degeneration and death. The anti-cancer properties of some of these polyphenols have been previously tested and confirmed. For instance, ellagic acid, quercetin, catechin, and phenolic acid prevented several kinds of cancer that affect the skin, stomach, duodenum, mouth, colon, liver, lung, and mammary glands (Yang et al., 2001; Okafor et al., 2022). Ade-Omowaye et al. (2015) and Soetan et al. (2018) independently confirmed the presence of antioxidant-related phytochemicals: phenolics and flavonoids in AYB. Out of nine underutilized legumes studied by Ade-Omowaye et al. (2015), AYB had very high total polyphenolic contents of 293.23 mg 100 g^{-1} and 288.68 mg 100 g^{-1} with higher antioxidant activities, (1.00 mmoleTE 100 g^{-1} and 0.67 mmolTE 100 g^{-1}), respectively, including a variety of BG (0.88 mmolTE 100 g^{-1}). Akinyemi et al. (2020) confirmed the presence of flavonoids, tannins, alkaloids, saponins, and cardiac glycosides in Jack beans. According to several studies (Sowndhararajan et al., 2011; Solomon et al., 2018; Akinyemi et al., 2020), Jack beans have high levels of antioxidants that have been shown to possess numerous health benefits such as reducing the risk of type 2 diabetes, cancer, and inflammation, improving lipid metabolism, lowering bad cholesterol levels, preventing metabolic syndrome, and reducing the incidence of cardiovascular diseases. Legume seeds contain significant amounts of antioxidants due to their high phenolic, flavonoid, and anthocyanin contents. Consuming products made from these legumes may help prevent and manage various chronic and degenerative diseases, as well as address protein-calorie malnutrition (Foyer et al., 2016; Tsamo et al., 2020; Popoola et al., 2022a). Studies by Adefegha and Oboh (2012); Kennedy (2014); Adefegha et al. (2017); Gupta and Prakash (2019), and Adefegha (2018) have confirmed that these bioactive components of NULs could play significant roles in increasing the immune level and support prevention of common diseases such as malnutrition (severe and acute particularly in infants), sexual enhancers, obesity, diabetes, heart-related diseases amongst others.

In terms of potentials as nutraceuticals, African yam bean (*Sphenostylis stenocarpa*), and Horse gram (*Macrotyloma uniflorum*) are legumes worthy of note, due to their anti-diabetic property, anti-urolithiasis effect, and role in the prevention and management of cardiovascular diseases, kidney stones, gastritis, pile, and urinary tract disease (Sharma et al., 2019; Vijayakumar, 2021). Extracts from mung bean, adzuki bean, black bean, rice bean, and lima bean have been documented to exert hepato-protective effects due to the presence of antioxidant and anti-inflammatory compounds (Vijayakumar, 2021). The bioactive components and health benefits of the selected NULs are presented in Table 4 and Figure 3.

Bioactive proteins and peptides are abundant in legume seeds (Kortt, 1984; Mojica and González de Meija, 2015; Makeri et al., 2016; Hussein et al., 2020). A notable bioactive protein in legume seeds is a lectin. Lectins possess anticancer and immunostimulatory activities. Lectins also help to reduce the risk of cardiovascular diseases in obsessed individuals (Roy et al., 2010; Carbonaro and Nucara, 2022).

Haemagglutinins from Jack bean possess anticancer and immunostimulatory properties (Carbonaro and Nucara, 2022). Jack bean produces a well-known lectin called Concanavalin A (Con. A) which has an extremely high anti-hepatoma activity

TABLE 4 Bioactive components of the selected neglected and underutilized legumes.

S/N	NULs Food Sources	Phytochemicals/Bio-active Contents	Health Benefits	References
1.	Adzuki beans	Phenol (tocopherols) and Flavonoids	Antioxidant activities, anti-atherogenic, anti-thrombogenic, and hypochloremia effects	(Gohara et al., 2016; Luo et al., 2016; Johnson et al., 2022)
2.	African yam bean	Phenols and flavonoids Resistant starch, slowly digestible starch (SDS), and non-starch polysaccharides	Antioxidant activities as nutraceuticals Stabilizes glucose metabolism and insulin levels improves mental performance and modulates appetite	(Huang et al., 2018; Soetan et al., 2018; George et al., 2020; Lin Tan et al., 2020)
3.	Bambara groundnut,	Tannins and flavonoids	Neuroprotective, cardioprotective, antitumor, and antioxidant properties	(Lin Tan et al., 2020; Adedayo et al., 2021)
4.	Jack bean	Polyphenols: flavonoids, tannins, alkaloids, saponins, and cardiac glycosides	Antioxidant with anti-diabetic, anti-cancer, and anti-inflammatory properties. Improves lipid metabolism, prevents metabolic syndrome, lowers bad cholesterol levels, and reduces cardiovascular diseases incidence and cancer risk	(Akpapunam and Sefa-Dedeh, 1997; Sowndhararajan et al., 2011; Akinyemi et al., 2020)
		Haemagglutinins	Anticancer and Immunostimulant activities	(Carbonaro et al., 2015)
5.	Kidney bean	Phenolic (tocopherol, total phenolics, total flavonoids and antioxidant activities	Promotes weight loss, anti-cholesterol, and anti-diabetic, and hepatoprotective properties.	(Kan et al., 2017; Idoko et al., 2020)
6.	Lima bean	Polyphenols, Flavonoids and Tannins	Anti-diabetic, antifungal, antiproliferative properties, hepaprotective activity, antioxidant effects, trypsin, hypocholesterolemia activities	(Agostini-Costa et al., 2014; Drago et al., 2016)
7.	Marama bean	Oleic acid Stearic acid Palmitic acid Polyphenol	Enhances glucose homeostasis and anti-inflammatory activity. Reduces blood pressure and atherosclerosis risk, improves heart function Hypolipidemic and anti-inflammatory properties associated with the prevention of cardiovascular disease, metabolic syndrome, and diabetes-related insulin resistance. Antioxidant, anti-bacterial, anti-fungal, anti-inflammatory, antihyperglycemic and pro-apoptotic properties; protects against free radical-induced erythrocyte hemolysis; represses rotavirus-induced inflammation	(Omotayo and Aremu, 2021; Omotayo and Aremu, 2019)
8.	Mung bean	Tannins and phytic acid	Attenuates blood glucose level and insulin responses to plasma cholesterol and starchy foods reduces cancer risks	(Gemede and Ratta, 2014; Ganesan and Xu, 2018)
9.	Rice bean	Phenols and Flavonoids Phenolic acids (*p*-coumaric acid, ferulic acid, and sinapic acid) Flavonoids (catechin, epicatechin, vitexin, isovitexin and quercetin)	Antidiabetic properties including α-glucosidase inhibition and advanced glycation end-product formation inhibitory activities.	(Yao et al., 2012; Bhagyawant et al., 2019; Kaur et al., 2021)
10.	Winged Bean	Polyphenol	Anticarcinogenic, antioxidant, anti-inflammatory, antitumoral, antimicrobial, antimutagenic, anti-ischemic and anti-allergic properties	(Mohanty et al., 2013; Bassal et al., 2020)
		Lectin	Antiploliferative activity	(Kortt, 1984)

arising from its resistance and structural stability to *in vitro* proteolysis and denaturation (Carbonaro et al., 2015; Huldani et al., 2022). The high level of lectins in winged bean has also been shown to have antiproliferative activity on human cancer cell lines. Two lectins (B2 and B3) were identified in winged bean which exhibited the same amino-terminal sequences and the sequence of lectin B3 to residue 40 reflected extensive homology with other legume lectins such as soybean lectin (Kortt, 1984).

Saponins are often regarded as antinutritional factors (ANTs) in grain legumes as they inhibit active transport and simultaneously increase the general permeability of enterocyte barrier (Bennetau-Pelissero, 2018). Thus, saponins increase the permeability of the small intestinal mucosal cells, facilitating the uptake of substances to which the gut would normally be impermeable. It also reduces the bioavailability of nutrients and decreases enzyme activity, resulting in an inhibition of growth. Notwithstanding, saponins have some positive health benefits as they contain a triterpenoid aglycone (sapogenin) linked to one or more oligosaccharide groups with the ability to absorb free radicals and activate antioxidant enzymes (Hai et al., 2021). Saponins in legume seeds contain two major components soya-saponin I (approximately 630 to 900 mg/kg) and dehydrosoyasaponin I (approximately 650 to 1300 mg/kg). The hilum portion of legume seeds has been identified as having the highest saponin content compared to the cotyledons (Hai et al., 2021). Research findings reveal that Japanese and Chinese populations have a lower risk for breast, colon, corpus uterine,

and prostate cancers due to their high intakes of legumes and legume products, which are good sources of saponins (Messina et al., 1994; Shi et al., 2004). Thus, they tend to have a longer life span than Africans (Lu et al., 2017). Terpenoids are a sub-group of triterpenoids and have been implicated to reduce bad cholesterol level and possess anti-cancer and antimicrobial properties (Marrelli et al., 2016). To our knowledge, little or no research has been done on the benefits of saponins and terpenoids derived from African underutilized legumes. This is an aspect of research that should be promoted to enhance the livelihood of the African population and would aid in the attainment of the Sustainable Development Goal of the United Nation on better health and well-being.

Alkaloids just like saponins are considered ANTs, and have been reported in a few underutilized legumes but not in detail (Konyeme et al., 2020; Popoola et al., 2022a). Alkaloids are naturally essential as defense agents which make up approximately 20% of the known secondary metabolites available in plants (Kaur and Arora, 2015). Therapeutically, alkaloids are particularly well-known as antioxidants, anti-inflammatories, anesthetics, and cardioprotective agents (Kurek, 2019; Heinrich et al., 2021). The presence of alkaloids in some underutilized legumes (winged bean, AYB, and Kersting's groundnut) suggests their potential application as anti-cancer, anti-inflammatory, antimicrobial, and analgesic agents amongst others (Kaur and Arora, 2015; Popoola et al., 2022a). Jack bean and AYB have been found to exhibit higher contents of alkaloids (0.645g/100g) and (22.195-183g\100g) (Konyeme et al., 2020). More investigations are needed on underutilized alkaloids, especially about their contents and variability as regards quinolizidine (QA) and pyrrolidone (PA).

Carotenoids such as β-carotene, lutein, and cryptoxanthin have been detected in most legumes though much lower compared to that of fruits and vegetables (Tee et al., 1995). Carotenoids are widely distributed in legumes and have been reported to exhibit health-promoting benefits such as antioxidants, better visual function, and reduction of cardiovascular diseases (Voutilainen et al., 2006; Ku et al., 2020; Maoka, 2020). In a study to evaluate bioactive components of selected underutilized legumes indigenous to Nigeria, James et al. (2020) found out that, fermentation and germination reduced carotenoid, anthocyanin, tannin, and flavonoid contents of the legumes.

Lipid profiles of underutilized legumes has been fairly documented but their effects on blood lipid levels are limited (Zhang et al., 2010; Adebowale et al., 2011). However, research has linked the consumption of these diets to a decreased risk of heart disease and obesity, according to (Hossain et al., 2016). Furthermore, Yao et al. (2015) discovered that the Ci12 landrace of Bambara groundnut from Côte d'Ivoire contained a high concentration of n-6 fatty acids, which are classified as polyunsaturated fatty acids (PUFAs) and include Omega-6 linoleic acid (C18:2, ώ-6) and Omega-3 alpha-linoleic acid (C18:3, ώ-3). These acids cannot be produced by the body and must be obtained through diet. Studies have also shown that consuming diets rich in Omega-6 fatty acids can reduce the incidence of cardiovascular disease and obesity, as noted by (Patterson et al., 2012; Djuricic and Calder, 2021). Oleic, stearic and palmitic acids have been recorded for Maraba bean (Omotayo and Aremu, 2021). The fatty acids components such as palmitic, palmitoleic, oleic, arachidonic, eicosapentaenoic, docosapentaenoic, lignoceric, docosahexaenoic and nervonic acids have not been studied extensively in African underutilized legumes. Further studies are required to unravel the PUFAs available in these lesser-known legumes.

6 Antimicrobial properties of underutilized legumes

Infections resulting from microbial sources are a great source of threat to plants, animals, and human health, which have necessitated the use of effective, safe, and sustainable biocontrol methods. This is particularly important due to the resistance of microbes to antibiotics and other control mechanisms as well as the search for novel antimicrobial agents (Udeh et al., 2020). Underutilized legumes are embedded with inherent antimicrobial abilities through the presence of different phytochemicals which include phytate, tannins, anthocyanin, flavonoids, etc. (Ayilara et al., 2022). These phytochemicals have been reported to be capable of controlling both gram-positive and gram-negative pathogenic bacteria (Ramatsetse et al., 2023). For instance, Bambara groundnut has been reported to inhibit the growth of different human pathogenic organisms which include *Klebsiella pneumonia, Escherichia coli, Bacillus aureus, Pseudomonas aeruginosa, Candida albicans, Klebsiella aerogenes, Aspergillus niger* and *Staphylococcus aureus* (Klompong and Benjakul, 2015; Wanyama et al., 2017; Oyeyinka et al., 2021) (Table 5). Antimicrobial properties of other selected underutilized legumes are shown in Table 5. The mechanism of action of the underutilized legumes as antimicrobial agents includes disruption of the microbial activity, chelation of crucial micro mineral elements (zinc and iron), suppression of the cell surface microbial enzymes, hydrophobic and electrostatic interaction with the cell membrane and cell wall (leading to the production of large pores and consequently its disintegration), induction of morphological changes in bacteria cells, increase in the permeability of cell wall which results to cell lysis and death, penetration of the cytoplasmic membrane, reduced intracellular ATP concentration and the prevention of spore germination and mycelial growth in fungi (Sitohy et al., 2013; Lopes and Brandelli, 2018; Udeh et al., 2020; Jia et al., 2021). On antifungal potentials, an array of proteins/peptides from mungbean, kidney bean, African yam bean, lima beans, brown kidney, winged beans), have elicited antifungal effect against plant and human pathogens including *Fusarium oxysporum* and *Coprinus comatus, Verticillium dahlia, Botrytis cinerea, Setosphaeria turcica, Rhizoctonia solani, Mycosphaerella arachidicola, Helminthosporium maydis, Candida albicans, Gibberalla sanbinetti, Sclerotinia sclerotiorum*, etc. (Mani-López et al., 2021). More research focus is required on arrays of antimicrobial properties of underutilized legumes of Africa which will possibly lead to cheaper means of drug discovery and good health care in Africa.

The species of legumes, the concentration of the extract, and the type of extractant (solvent) used are essential factors that affect the activity of underutilized legumes as antimicrobial agents (Shelke et al., 2022). Kaundal et al. (2019) reported that when different solvents (dichloromethane, 1-butanol, water and ethyl acetate) were used to assess the antimicrobial properties of Horse gram against human pathogenic organisms (*Bacillus* sp., *E.coli, Shigella* sp., *Staphylococcus* sp. *and Salmonella* sp.), ethyl acetate and

TABLE 5 Antimicrobial properties of underutilized legumes.

Plant	Botanical name	Extractant	Part used	Microbe or disease control	References
Horse gram	*Dolichos biflorus*	Water	Seeds	*Escherichia coli, Pseudomonas aeruginosa, Bacillus subtilis* and *Staphylococcus aureus*	(Basu et al., 2017)
Mung bean Bengal gram	*Vigna radiata Cicer arietinum*	Water	Hull	*Bacillus cereus*	(Kanatt et al., 2011)
Lablab bean	*Lablab purpureus*	Peptide	Seeds	*Bacillus cereus*	(Bai-Ngew et al., 2021)
Bambara groundnut	*Vigna subterranea*	Water	Hull, seeds	*Klebsiella pneumoniae, Aspergillus niger, Pseudomonas aeruginosa, Candida albicans Staphylococcus aureus, Escherichia coli* and *Bacillus cereus*	(Udeh et al., 2020)
Adzuki bean	*Vigna angularis*	Ethanol	Seed coat	*E. coli* and *Staph aureus*	(Jia et al., 2021)
Kidney bean	*Phaseolus vulgaris*	Methanol	Seeds	Multidrug-resistant *Enterobacterales*	(Ebrahim et al., 2022)
Marama bean	*T. esculentum*	Water	Testae	*Campylobacter jejuni*, *Staphylococcus* sp., *Escherichia coli*, *Shigella* sp., *Yersinia* sp., MRSA, and *Salmonella* sp.	(Chingwaru et al., 2015)
Mung bean	*Vigna radiata*	Ethanol	Seed Flour	*L. monocytogens, C. jejuni, S.aureus, E.coli, B. subtilis* and *Pseudomonas aeruginosa*	(Keawpeng et al., 2022)
Winged bean	*Psophocarpus tetragonolobus*	Methanol	Leaves	*Pseudomonas aeruginosa*	(Latha et al., 2007)

dichloromethane extracts revealed antibacterial activities while the aqueous and 1-butanol extract showed no antibacterial properties (Kaundal et al., 2019). Hence, it is essential to carry out further research on different underutilized legumes to unravel the best extractant that can be used to extract the active ingredients in different NULs to promote their potential in the discovery of new drugs.

In addition, different plant parts are used in the production of plant extracts, these include the pods, seeds, flowers, hall, root, stem, tuber, and leaf, where different types and different forms of phytochemicals can be found (Henciya et al., 2017; Zhong et al., 2022). More research should be carried out on underutilized legumes to unravel the different parts of each NULs that can give a maximum recovery and variety of antimicrobial active compounds.

7 Prospects in harnessing the benefits of underutilized legumes

Too much reliance on a few staple crops to meet the food and nutritional needs of man is a potential threat to the global fight against food insecurity and to ensure that the zero hunger sustainable development goals (SDGs) are achieved by 2030. Traditional or indigenous food crops in Africa have major roles to play in realizing the SDGs 2 and 3 of the United Nations if given the utmost attention and necessary improvement for human consumption. The African populace needs to be sensitized to the benefits derived from her indigenous legumes. Furthermore, researchers in Africa must embark on collaborative research and give priority to these legumes in crop improvement programs using a holistic approach. Cellular oxidative stress has been implicated in the development of chronic diseases such as cardiovascular disease, cancer, arthritis, diabetes, and degenerative diseases in humans. Nevertheless, antioxidants in foods regulate and reduce oxidative destruction by inhibiting oxidation caused by reactive oxygen species (ROS), and improve the shelf-life and quality of these foods (Ames et al., 1993; Altemimi et al., 2017). The bioactive components of legume seeds possess antioxidant activity which could mitigate the effects of oxidative stress. However, these bioactive forms a small percentage of the nutritional components of legume seeds.

Singh et al. (2021) revealed that bioactive compounds are concentrated in different parts of the seeds of legumes. For instance, phenolic compounds such as flavonoids and dietary fibers occur in the seed coat while non-flavonoids such as oligosaccharides and dietary fiber occur in the cotyledons (Singh et al., 2021). Including dietary fibers from legume seeds and cotyledons in one's diet has several positive effects on human health. These fibers can assist with digestion in the gastrointestinal tract by increasing water-holding capacity, viscosity, bulk, fermentability, and the ability to bind bile acids, as noted by Stosh and Yada (2010). In addition, it is known to reduce serum cholesterol in hypercholesterolemic people and postprandial glycemia. The dietary fiber in the seed testa has been reported to be significantly higher than the quantity in the cotyledons (Sergio et al., 2020). These bioactive compounds could also be used to design functional food products. Unfortunately, most of the bioactive components of the seeds of underutilized grain legumes are unknown. Therefore, there is a need for these bioactives present in Africa's indigenous legumes to be extracted, purified, and characterized using biochemical approaches with the chemical structures elucidated with the aid of high-tech equipment

(Figure 4). To extract the bioactive compounds from these NULs, modern extraction techniques can be employed for maximum efficiency. Once extracted, these components are further purified using chromatographic methods, such as high performance liquid chromatography (HPLC) or column chromatography, before being profiled using analytical techniques such as nuclear magnetic resonance (NMR), gas-chromatography-mass spectroscopy (GC-MS), GC-time of flight-MS (GC-TOF-MS) which improves resolved peaks, LC-MS, or Fourier Transform Infrared spectroscopy (FTIR) used to detect different functional groups of metabolites. These techniques help to identify and quantify a variety of primary and secondary metabolites in the purified bioactive compounds (Pandey et al., 2016).

Different methods have been used for the extraction of bioactive compounds; these include microwave-assisted extraction (MAE) soxhlet extraction (SE) using different solvents as well as ultrasonic/ultrasound-assisted extraction (UAE). Other modern methods also in use include supercritical fluid extraction (SFE), solid-phase extraction (SPE), enzyme-assisted extraction, pressurized liquid extraction (PLE) or accelerated solvent extraction (ASE), and extraction assisted by a pulsed electric field. These recent methods are efficient in the removal of flavonoids from plant products. Current reviews have discussed in detail, the use of these modern methods for the extraction of essential bioactives from plants (Selvamuthukumaran and Shi, 2017; Chaves et al., 2020).

The MAE is one of the modern extraction methods that has received much patronage due to several merits such as a reduction in utilization of solvents, enhanced yield recovery, better selectivity, reproducibility, reduced operation time, and less sample manipulation (Alara and Abdurahman, 2019; Alara et al., 2019; Nour et al., 2021). MAE utilizes microwave energy for faster heating which results from a range of electromagnetic spectrums of light (300 MHz to 300 GHz) with short wavelengths usually between 1 cm^{-1} and 1 m^{-1}. The combined effect of increased temperature within the extraction medium and the effect of microwave electromagnetic radiation on vibrations of both the extraction solvents and the analytes being extracted enhances the extraction yield (Ameer et al., 2017; Chaves et al., 2020). It is essential to note that the efficiency of the extraction process is dependent on several factors such as the temperature and particle size, solvent-solid ratio, as well as the nature of the solvent used for extraction.

MAE utilizes microwave energy for rapid heating which results from a range of electromagnetic spectrums of light (300 MHz to 300 GHz) with short wavelengths usually between 1 cm^{-1} and 1 m^{-1}. The combined effect of increased temperature within the extraction medium and the effect of microwave electromagnetic radiation on vibrations of both the extraction solvents and the analytes being extracted enhances the extraction yield (Ameer et al., 2017; Chaves et al., 2020). It is noteworthy that the efficiency of the extraction process is dependent on several factors such as the temperature and particle size, solvent-solid ratio, as well as the nature of the solvent used for extraction. The extraction efficiency is dependent on factors such as the nature of the solvent, solvent–solid ratio, temperature, as well as particle size. To enhance the extraction process, optimization of modern extraction methods is required. Synergistic effects could be obtained if there is a combination of the above novel extraction methods. Tsamo et al. (2020) used an ultra-performance liquid chromatographic system with a PDA detector coupled to a mass spectrophotometer detector (UPLC-qTOF-MS) to identify phenolic compounds in the seeds of Kersting's groundnut. A total of 57 potential compounds were identified, among which were flavonoids; catechin, gallocatechin, quercetin, rutin, naringin, kaempferol, 7-rutinoside, and eriodictyol 7-rutinoside.Gallocatechin is known to enhance lipid metabolism and aids the prevention of metabolic syndrome as has been reported in some other legume seeds such as pea and lentils (Mirali et al., 2017; Jha et al., 2019; Tsamo et al., 2020). Table 6 shows different modern methods used in extraction of bioactive components of plants, including seeds.

With the advancement in sequencing techniques and omics technologies such as genomics, proteomics, transcriptomics, metabolomics, and the genome editing tools such as the

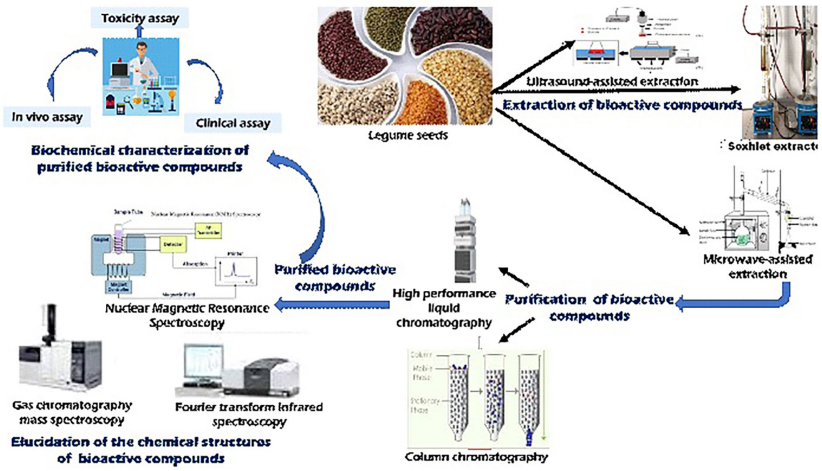


FIGURE 4
Extraction, characterization, and purification methods for legume seed bioactives.

TABLE 6 Bioactives from legumes and their extraction conditions.

Bioactive compound	Extraction method	Plant source	Extraction conditions	References
Phenolic phytochemicals	MAE	*Phaseolus vulgaris* (L.).	Effective extraction of polyphenols at a temperature of 150°C using 50% ethanol	(Sutivisedsak et al., 2010)
inositols and α-galactooligosaccharides	Optimized MAE	Mung bean (Vigna radiata)	0.5 g dry sample, 2 cycles of 3 min, 50°C, 10 mL 50:50 (ethanol: water, v: v), resulted in extraction of bioactive carbohydrates between 74.1 and 104.2 mg.g–1 dry sample	(Carrero-Carralero et al., 2018)
inositols,α-galactooligosaccharides (GOS)	MAS	Alfalfa (*Medicago sativa* L.) leaves and stems	Optimal extraction temperatures of 40°C (leaves), and 80° C (seeds) resulted in higher yields of inositols (2x) and α-GOS (7 x) with more Pinitol in leaves and stems (24.2–31.0 mg_g^{-1} and 15.5–22.5 mg_g^{-1}, respectively) while seed extracts were rich in α-GOS, mainly in stachyose (48.8–84.7 mg_g1).	(Solarte et al., 2021)
Quercetin flavonoid	Sequential MAS	Red kidney bean	The extraction efficiency of quercetin was enhanced yielding 35.8 mg quercetin/g kidney bean	(Aghajanian et al., 2020)
Saponins	UAE	red lentils (*Lens culinaris*)	The ethanol extraction efficiency of total saponin content in red lentil seeds was increased (11 g 100 g^{-1}).	(Del Hierro et al., 2018)

MAE, Microwave assisted extraction; UAE, Ultrasound-assisted extraction.

CRISPER-Cas9 or TALEN, underutilized legumes can be genetically improved for better utilization and acceptance as these legumes are currently faced with some production constraints such as high antinutritional factors in the seeds which reduce the bioavailability of minerals, prolonged cooking time due to hardness of the seed coat of some pulses, and photoperiod sensitivity which also affects tuberization in tuberous legumes. The utilization of plant-based functional foods can be enhanced by the use of innovative technologies for the extraction and microencapsulation of bioactive compounds using novel technologies in metabolomics (Nayak et al., 2021; Pattnaik et al., 2021). Metabolomic studies can be used to identify rich value-added compounds from different parts of underutilized legumes such as the seeds, leaves, stems, or tubers, listing the main bioactive metabolites identified and the factors affecting their production. Metabolomics finds its usefulness in the identification of metabolites after the bioactives have been extracted using one of the modern methods discussed above. It highlights the expressions and changes of metabolites, as well as their interactions and resulting phenotypic traits in plants subjected to harsh environmental conditions. Under such stress, plants must adapt their metabolomic pathways to maintain metabolic homeostasis, a process referred to as acclimation (Joshi et al., 2021; Makhumbila et al., 2022).

Chen et al. (2020) studied the metabolomic profile of common bean and identified major findings related to amino acids, flavonoids, isoflavonoids, purines, and proline metabolism. These pathways enhanced the plant's potential for defense against pathogens like *Fusarium solani* (FS). The study combined RNA sequencing and metabolomics techniques to investigate changes in gene expression and metabolic processes in common bean infected with FS. The results showed that metabolic pathways were enriched, leading to an increase in metabolites involved in plant defense response. Infected common bean seedlings responded with modifications to their cell walls, the generation of reactive oxygen species, and a synergistic hormone-driven defense response. The study also found that infected plants induced energy metabolism, nitrogen mobilization, accumulation of sugars, and arginine and proline metabolism (Chen et al., 2020; Makhumbila et al., 2022). Reliable software tools such as GCMS, LC-MS, and NMR are required to analyze the vast amounts of data generated by metabolomic technologies. These tools should be capable of visualizing, detecting peaks, normalizing/transforming sample data, annotating, identifying, quantifying, and statistically analyzing targeted and untargeted metabolite variations using algorithms for univariate and multivariate analysis (Sun and Weckwerth, 2012; Junot et al., 2014; Makhumbila et al., 2022). There are now several metabolomic pathway databases available online that group metabolites with similar functions. These databases include the Kyoto Encyclopedia of Genes and Genomes (KEGG), Cytoscape, MapMan, and iPath, which are relevant to plants. Cytoscape is an open-source software platform used to visualize complex networks and integrate them with any type of attribute data. MapMan is a user-driven tool that displays large datasets onto diagrams of metabolic pathways while iPath is a relevant tool for plants (Fukushima and Kusano, 2013; Patel et al., 2021).

With the advent of Next-generation Sequencing (NGS), the cost of sequencing has plummeted, making it possible to sequence large and complex genomes in a shorter period (Hamilton and Robin Buell, 2012; Kumar et al., 2021). Several whole genome sequencing studies are underway for some underutilized crops, and some have been completed. Once these sequences are available, they can be applied for in-depth structural and functional genomic studies to characterize and annotate the genes. Furthermore, the availability of the whole genome sequence will accelerate the development of genetic linkage maps of genomic regions that control particular traits of the plant, as well as the accumulation of bioactive compounds. The coupling of sequencing technologies with bioinformatics and high-through put phenotyping techniques genomic studies and bioinformatics tools could facilitate the improvement of the genetic pathways for the production of

bioactive compounds and identification of genes that regulate essential agronomic traits relevant to the quality of NULs (Mochida and Shinozaki, 2011; Steinwand and Ronald, 2020; Kumar et al., 2021). Utilization of various genomics approaches such as genome-wide association studies (GWAS); marker-assisted selection (MAS) and genomic selection (GS) have been used to identify useful markers linked to nutritional traits and bioactives in various crops. For instance, a study on 94 chickpea genotypes from a diverse population using GWAS resulted in the identification of eight single nucleotide polymorphisms (SNPs) associated with Fe and Zn content in chickpea seeds (Diapari et al., 2014), while two closely associated SNPs markers for Fe and Zn were identified by GWAS in lentils (Khazaei et al., 2017). Following the identification of SNPs linked to these trace elements, marker-assisted selection can be applied for the introgression of these traits into underutilized legumes through the process of biofortification. This is particularly relevant in the current global pandemic as Zn is known to be an immune booster. Shreds of evidence have shown that Zn deficiency increases the risk of infectious diseases, autoimmune disorders, and cancer (Roohani et al., 2013; Haase and Schomburg, 2019; Wessels and Rink, 2020; Wessels et al., 2020). Although most of the NULs especially the African yam bean are good sources of Zn and Fe, the levels of these elements could be enhanced through biofortification. Legumes with high Fe and Ze levels could be harnessed to boost the immunity of risk groups mainly the elderly and patients with inflammatory or autoimmune diseases. With the use of modern breeding methods, an international organization located in Malaysia, Crops for the Future is spearheading research on some underutilized species such as bambara groundnut and winged bean to enhance food and nutritional security.

Omic technologies should be employed to enhance the functional components of NULs towards ensuring food and nutritional security. The use of transcriptomic analysis for the identification of regulatory genes in biochemical pathways can assist researchers to gain significant insight into the functional mechanisms of plant's biosynthetic pathways especially those involved in secondary metabolite synthesis. The transcriptomic analysis is usually carried out to study gene expression in plants. This is done using microarray technology or RNAseq analysis. Available transcriptome data of some model legumes could be applied to study NULs when the transcripts have been obtained as has been reported for *Medicago trunculata* using theCTDB (RNASeq) and MtGEA (Mocroarray) (Garg and Jain, 2013). Integrated use of omics technologies methods to enhance the nutrient potential of any crop, could influence nutritional security if applied in food processing and formulations (Tian et al., 2016; Nayak et al., 2021). If this is not done the rich bioactive compounds inherent in most of these indigenous legumes will remain unknown and untapped. Molecular studies should therefore be used for genetic dissection of antioxidant activities in NULs and nutrient-related traits. This has not received much research attention thus far. Apart from enhancing the nutritional contents of these legumes, omics technologies coupled with genome editing could aid the reduction of antinutrients like oxalate and phytic acid which affects the bioavailability of vital minerals, thereby enhancing human health.

TThe CRISPR/Cas 9 method of genetic manipulation in plants has gained a lot of attention and acceptance for crop improvement because it is straightforward, adaptable, and accurate (Bhowmik et al., 2021). It would be very useful for enhancing genetic improvement in NULs. This technology is constantly evolving and has a wide range of applications, such as producing knockouts, precise modifications, multiplex genome engineering, or controlling gene expression (Arora and Narula, 2017; Baloglu et al., 2022). CRISPR/Cas9 relies on two key components: a Cas9 endonuclease and a guide RNA (gRNA) consisting of two small RNA molecules: the CRISPR RNA (crRNA) which is a 20-nucleotide sequence that matches the target DNA, and the transactivating crRNA (tracrRNA), which serves as a binding scaffold (Bhowmik et al., 2021). The CRISPR/Cas9 technology is highly useful for improving genetic traits in plants. Though it has been utilized in gene editing of known legumes such as soybean, cowpea, and the model legume *Medicago trunculata* (Curtin et al., 2018; Al Amin et al., 2019; Bao et al., 2019; Ji et al., 2019; Bao et al., 2020; Juranic et al., 2020; Chen et al., 2020; Wolabu et al., 2020; Bhowmik et al., 2021), very few attempts have been made on genome editing in underutilized legumes. The bottleneck has been in getting the whole genome sequence of the underutilized legumes.

Meng et al. (2017) developed an efficient CRISPR/Cas9 system for inducing targeted mutations in the MtPDS gene in the model legume *M. truncatula*. Among the 309 T0 transgenic plants, 32 displayed the albino phenotype. To determine if the albino phenotype was due to the targeted mutation, 16 out of the 32 transgenic plants were randomly selected for sequence analysis. Results revealed that all the albino plants had mutations at the targeted site of the MtPDS gene. These findings were supported by Baloglu et al. (2022). Efforts are ongoing by the African Orphan Crops Consortium (AOCC) through a network of international to regional public-private partnerships and collaborators, to generate genomic sequences of some underutilized legumes (Faba bean, Mungbean, Bambara groundnut, Marama bean, Mungbean, and *Lablab purpureus*). The complete genome sequence of Mungbean published in 2014 permitted genomic research and molecular breeding of mung bean (Kang et al., 2014; Baloglu et al., 2022). Schafleitner et al. (2015) evaluated agronomic traits in 1481 Mungbean collections based on the availability of its whole genome sequence. This paves the way for genome editing to be used for the genetic improvement of this species to enhance its yield, nutritional content, and resistance to diseases. The lack of whole genome sequences in most of these beneficial underutilized legumes poses challenges to their improvement using CRISPER/Cas9. Despite the absence of a genome reference sequence for Faba bean, significant advancements have been made in genetic and genomic resources to aid molecular breeding. The robust synteny shared with the model legume *M. trunculata* allows for the use of omic technologies like transcriptomics and comparative genomics. These methods help identify single-nucleotide polymorphisms (SNPs), develop high-density consensus genetic maps, and predict the candidate genes responsible for various desirable traits. Bhowmik et al. (2021) discussed these approaches in their study. The genome editing technology no doubt has the potential to enhance crop improvement in various ways. However, some

individuals are against its global acceptance pushing for regulations on its use. Nevertheless, genome editing technology differs from genetic engineering or modification which requires novel genes to be inserted into another organism of a different species. Across many countries and regions in the world, different regulatory approaches are being sought. In Africa, the National Biosafety Management Agency of Nigeria released the first gene editing guidelines which paves the way for its utilization in the improvement of economic crops in the country. Other African governments should take a cue from Nigeria and develop a regulatory framework for gene editing.

8 Conclusion

There is no doubt that underutilized legumes are rich sources of micronutrients and bioactive compounds with a great capacity to achieve zero hunger of the Sustainable Development Goals (SDG) by 2030. The rich bioactive compounds inherent in underutilized legumes which are yet to be tapped have great health benefits for man. Most phytochemicals in legumes are regarded as antinutritional components as they have no nutritional value. However, recent studies have shown that these non-nutrients such as tannins, glycosides, and saponins possess hypocholesterolemic and anticarcinogenic activity while flavonoids possess antioxidant activities which are essential for scavenging reactive oxygen species which cause oxidative stress in diseased conditions. If the abundant bioactive compounds in these underutilized legumes are identified and employed as therapeutics or used in the development of functional food products, it will greatly enhance human health, reduce the over utilization of the common legumes as well as help to increase food, protein, and nutrition security in Africa. The renewed efforts in this direction will be to evolve strong research and development between industries (pharma and foods) and the academia/research for appropriate food-based or pharmaceutical product developments. Thus, incorporating NULs rich in bioactive compounds into the diet of man will boost achieving the Sustainable Development Goal 3 of the United Nations on good health and well-being.

Author contributions

JP and OO – conceived the idea, wrote the first draft, searched the literature, and reviewed the manuscript. OA, LO, TA, MA, AO, SD, and AA – contributed to the writing and review of the manuscript. PA, SO, MA, and CO – contributed to the writing. JP – edited, fine-tuned, and approved the final draft. All authors contributed to the article and approved the submitted version.

Funding

The authors acknowledge Bowen University's Directorate of Research and Strategic Partnerships (DRSPs) for defraying the publication fees.

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

Abberton, M., Adegboyega, T. T., Faloye, B., Paliwal, R., and Oyatomi, O. (2020a). African Yam bean, *Sphenostylis stenocarpa*. *Legume Perspect.* 14. doi: 10.1155/2020/6569420

Abberton, M., Adegboyega, T. T., Faloye, B., Paliwal, R., and Oyatomi, O. (2020b). Winged bean, *Psophocarpus tetragonolobus*. *Legume Perspect.* 27-28. doi: 10.1155/2020/3439620

Abberton, M., Paliwal, R., Faloye, B., Marimagne, T., Moriam, A., and Oyatomi, O. (2022). Indigenous African orphan legumes: Potential for food and nutrition security in SSA. *Front. Sustain. Food Syst.* 6. doi: 10.3389/fsufs.2022.708124

Adebowale, Y. A., Schwarzenbolz, U., and Henle, T. (2011). Protein isolates from bambara groundnut (Voandzeia subterranean l.): Chemical characterization and functional properties. *Int. J. Food Prop* 14, 758–775. doi: 10.1080/10942910903420743

Adedayo, B. C., Anyasi, T. A., Taylor, M. J. C., Rautenbauch, F., Le Roes-Hill, M., and Jideani, V. A. (2021). Phytochemical composition and antioxidant properties of methanolic extracts of whole and dehulled bambara groundnut (*Vigna subterranea*) seeds. *Sci. Rep.* 11, 14116. doi: 10.1038/s41598-021-93525-w

Adefegha, S. A. (2018). Functional foods and nutraceuticals as dietary intervention in chronic diseases; Novel perspectives for health promotion and disease prevention. *J. Diet. Suppl.* 15, 977–1009. doi: 10.1080/19390211.2017.1401573

Adefegha, S. A., and Oboh, G. (2012). Acetylcholinesterase (AChE) inhibitory activity, antioxidant properties and phenolic composition of two *Aframomum species*. *J. Basic Clin. Physiol. Pharmacol.* 23, 153–161. doi: 10.1515/jbcpp-2012-0029

Adefegha, S. A., Oboh, G., Oyeleye, S. I., Dada, F. A., Ejakpovi, I., and Boligon, A. A. (2017). Cognitive enhancing and antioxidative potentials of velvet beans (*Mucuna pruriens*) and horseradish (*Moringa oleifera*) seeds extracts: A comparative study. *J. Food Biochem.* 41, e12292. doi: 10.1111/jfbc.12292

Adegboyega, T. T., Abberton, M. T., AbdelGadir, A. H., Dianda, M., Maziya-Dixon, B., Oyatomi, O. A., et al. (2020). Evaluation of nutritional and antinutritional properties of African yam bean (*Sphenostylis stenocarpa* (Hochst ex. a. rich.) harms.) seeds. *J. Food Qual.* 2020, 6569420. doi: 10.1155/2020/6569420

Adegboyega, T. T., Abberton, M. T., Abdelgadir, A. H., Dianda, M., Maziya-Dixon, B., Oyatomi, O. A., et al. (2019). Nutrient and antinutrient composition of winged bean (*Psophocarpus tetragonolobus* (L.) DC.) seeds and tubers. *J. Food Qual.* 2019, 3075208. doi: 10.1155/2019/3075208

Adegboyega, T. T., Mansurat, O. S., and Anas, A. T. (2021). Proximate and antinutrient composition of selected West African bambara groundnut (Vigna subterranea (L.) verdc.) accessions. *J. Underut. Leg* 3, 13–25.

Ade-Omowaye, B., Tucker, G., and Smetanska, I. (2015). Nutritional potential of nine underexploited legumes in southwest Nigeria. *Int. Food Res. J.* 22, 798.

Adeoye, B. K., Aransiola, E. F., Alebiowu, G., Bisi-Johnson, M. A., Olorunmola, F. O., and Adepoju, O. A. (2021). The characterization and microbiological evaluation of probiotic isolated from bambara groundnut. *Int. J. Appl. Sci. Biotechnol.* 9 (1), 54–64. doi: 10.3126/ijasbt.v9i1.32959

Adewale, B. D., and Nnamani, C. V. (2022). Introduction to food, feed, and health wealth in African yam bean, a locked-in African indigenous tuberouslegume. *Front. Sustain. Food Syst.* 6:72645. doi: 10.3389/fsufs.2022.726458

Agbolade, J., Olakunle, T., Popoola, K., Idowu, J., Isiaka, A., and Aasa-Sadique, A. (2019). Genetic variability and diversity analysis in pod and seed characters of some neglected and underutilized legumes (NULs). *Asian J. Biochem. Genet. Mole. Biol.* 2, 1–8. doi: 10.9734/ajbgmb/2019/v2i330059

Aghajanian, S., Kazemi, S., Esmaeili, S., Aghajanian, S., and Moghadamnia, A. A. (2020). Sequential microwave-assisted extraction for isolation of quercetin from red kidney bean. *Int. J. Eng.* 33, 12–17. doi: 10.5829/ije.2020.33.01a.02

Agostini-Costa, T., da, S., Teodoro, A. F. P., Alves, R., de B. das, N., Braga, L. R., et al. (2014). Total phenolics, flavonoids, tannins and antioxidant activity of lima beans conserved in a Brazilian genebank. *Cienc. Rural* 45, 335–341. doi: 10.1590/0103-8478cr20140030

Ajibola, G. O., and Olapade, A. A. (2016). Physical, proximate and anti-nutritional composition of African yam bean (*Sphenostylis stenocarpa*) seeds varieties. *J. Food Res.* 5, 67. doi: 10.5539/jfr.v5n2p67

Akinyemi, F. A., Orishadipe, A. T., Ebun-Oluwa, O., and Aladesanmi, O. A. (2020). Physico-chemical properties and functional characteristics of jack beans (*Canavalia ensiformis*) starch. *World J. Biol. Pharm. Health Sci.* 3 (2), 012–022. doi: 10.30574/wjbphs.2020.3.2.0018

Akpapunam, M. A., and Sefa-Dedeh, S. (1997). Jack bean (*Canavalia ensiformis*): Nutrition related aspects and needed nutrition research. *Plant Foods Hum. Nutr.* 50, 93–99. doi: 10.1007/BF02436029

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. doi: 10.1186/s12896-019-0501-2

Alara, O. R., and Abdurahman, N. H. (2019). Microwave-assisted extraction of phenolics from *Hibiscus sabdariffa* calyces: Kinetic modelling and process intensification. *Ind. Crop Prod.* 137, 528–535. doi: 10.1016/j.indcrop.2019.05.053

Alara, O. R., Abdurahman, N. H., and Abdul Mudalip, S. K. (2019). Optimizing microwave-assisted extraction conditions to obtain phenolic-rich extract from *Chromolaena odorata* leaves. *Chem. Eng. Technol.* 42, 1733–1740. doi: 10.1002/ceat.201800462

Alcázar-Valle, M., Lugo-Cervantes, E., Mojica, L., Morales-Hernández, N., Reyes-Ramírez, H., Enriquez-Vara, J. N., et al. (2020). Bioactive compounds, antioxidant activity, and antinutritional content of legumes: A comparison between four phaseolus species. *Molecules* 25, 3528. doi: 10.3390/molecules25153528

Altemimi, A., Lakhssassi, N., Baharlouei, A., Watson, D. G., and Lightfoot, D. A. (2017). Phytochemicals: Extraction, isolation, and identification of bioactive compounds from plant extracts. *Plants* 6, 42. doi: 10.3390/plants6040042

Alvarado-López, A. N., Gómez-Oliván, L. M., Heredia, J. B., Baeza-Jiménez, R., Garcia-Galindo, H. S., and Lopez-Martinez, L. X. (2019). Nutritional and bioactive characteristics of Ayocote bean (*Phaseolus coccienus L.*): An underutilized legume harvested in Mexico. *CyTA-J. Food* 17 (1), 199–206. doi: 10.1080/19476337.2019.1571530.

Amarowicz, R., and Pegg, R. B. (2008). Legumes as a source of natural antioxidants. *Eur. J. Lipid Sci. Technol.* 110, 865–878. doi: 10.1002/ejlt.200800114

Ameer, K., Shahbaz, H. M., and Kwon, J. H. (2017). Green extraction methods for polyphenols from plant matrices and their byproducts: A review. *Compr. Rev. Food Sci. Food Saf.* 16, 295–315. doi: 10.1111/1541-4337.12253

Ames, B. N., Shigenaga, M. K., and Hagen, T. M. (1993). Oxidants, antioxidants, and the degenerative diseases of aging. *Proc. Natl. Acad. Sci. U.S.As* 90, 7915–7922. doi: 10.1073/pnas.90.17.7915

Arise, A. K., Malomo, S. A., Awaw, A. A., and Arise, R. O. (2022). Quality attributes and consumer acceptability of custard supplemented with bambara groundnut protein isolates. *App. Food Res.* 2, 100056. doi: 10.1016/j.afres.2022.100056

Arogundade, L. A., Mu, T.-H., and Akinhanmi, T. F. (2016). Structural, physicochemical and interfacial stabilisation properties of ultrafiltered African yam bean (*Sphenostylis stenocarpa*) protein isolate compared with those of isoelectric protein isolate. *LWT - Food Sci. Tech.* 69, 400–408. doi: 10.1016/j.lwt.2016.01.049

Arora, L., and Narula, A. (2017). Gene editing and crop improvement using CRISPR-Cas9 system. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01932

Ayilara, M. S., Abberton, M., Oyatomi, O. A., Odeyemi, O., and Babalola, O. O. (2022). Potentials of underutilized legumes in food security. *Front. Soil Sci.* 2. doi: 10.3389/fsoil.2022.1020193

Bai-Ngew, S., Chuensun, T., Wangtueai, S., Phongthai, S., Jantanasakulwong, K., Rachtanapun, P., et al. (2021). Antimicrobial activity of a crude peptide extract from lablab bean (*Dolichos lablab*) for semi-dried rice noodles shelf-life. *Qual. Assur. Saf. Crop Food* 13, 25–33. doi: 10.15586/qas.v13i2.882

Baiyeri, S. O., Uguru, M. I., Ogbonna, P. E., Samuel-Baiyeri, C. C. A., Okechukwu, R., Kumaga, F. K., et al. (2018). Evaluation of the nutritional composition of the seeds of some selected African yam bean (*Sphenostylis stenocarpa* hochst ex. a. rich (harms)) accessions. *Agroscience* 17, 37–44. doi: 10.4314/as.v17i2.5

Baloglu, M. C., Celik, A. Y., Baloglu, P., Yildiz, A. B., Türkölmez, N., Özden, et al. (2022). Gene-editing technologies and applications in legumes: Progress, evolution, and future prospects. *Front. Genet.* 13. doi: 10.3389/fgene.2022.859437

Bao, A., Chen, H., Chen, L., Chen, S., Hao, Q., Guo, W., et al. (2019). CRISPR/Cas9-mediated targeted mutagenesis of GmSPL9 genes alters plant architecture in soybean. *BMC Plant Biol.* 19, 131. doi: 10.1186/s12870-019-1746-6

Bao, A., Zhang, C., Huang, Y., Chen, H., Zhou, X., and Cao, D. (2020). Genome editing technology and application in soybean improvement. *Oil Crop Sci.* 5 (1), 31–40. doi: 10.1016/j.ocsci.2020.03.001

Barac, M. B., Pesic, M. B., Stanojevic, S. P., Kostic, A. Z., and Bivolarevic, V. (2015). Comparative study of the functional properties of three legume seed isolates: Adzuki, pea and soy bean. *J. Food Sci. Technol.* 52, 2779–2787. doi: 10.1007/s13197-014-1298-6

Barman, A., Marak, C. M., Barman, R. M., and Sangma, C. S. (2018). "Nutraceutical properties of legume seeds and their impact on human health," in *Legume seed nutraceutical research* (United Kingdom (UK): IntechOpen). doi: 10.5772/intechopen.78799

Bassal, H., Merah, O., Ali, A. M., Hijazi, A., and El Omar, F. (2020). *Psophocarpus tetragonolobus*: An underused species with multiple potential uses. *Plants* 9, 1730. doi: 10.3390/plants9121730

Basu, S., Ghosh, M., Bhunia, R. K., Ganguly, J., and Banik, B. K. (2017). Polysaccharides from *Dolichos biflorus* Linn and trachyspermum ammi Linn seeds: Isolation, characterization and remarkable antimicrobial activity. *Chem. Cent. J.* 11, 1–10. doi: 10.1186/s13065-017-0349-2

Bennetau-Pelissero, C. (2018). "Plant proteins from legumes," in *Bioactive molecules in food, reference series in phytochemistry*. Eds. J. M. Mérillon and K. G. Ramawat (Springer Nature). doi: 10.1007/978-3-319-54528-8_3-1

Bessada, S. M. F., Barreira, J. C. M., and Oliveira, M. B. P. P. (2019). Pulses and food security: Dietary protein, digestibility, bioactive and functional properties. *Trends Food Sci. Technol.* 93, 53–68. doi: 10.1016/j.tifs.2019.08.022

Bhadkaria, A., Srivastava, N., and Bhagyawant, S. S. (2021). A prospective of underutilized legume moth bean (*Vigna aconitifolia* (Jacq.) marechàl): Phytochemical profiling, bioactive compounds and *in vitro* pharmacological studies. *Food Biosci.* 42, 101088. doi: 10.1016/j.fbio.2021.101088

Bhagyawant, S. S., Bhadkaria, A., Narvekar, D. T., and Srivastava, N. (2019). Multivariate biochemical characterization of rice bean (*Vigna umbellata*) seeds for nutritional enhancement. *Biocatal. Agric. Biotechnol.* 20, 101193. doi: 10.1016/j.bcab.2019.101193

Bhowmik, P., Konkin, D., Polowick, P., Hodgins, C. L., Subedi, M., Xiang, D., et al. (2021). CRISPR/Cas9 gene editing in legume crops: Opportunities and challenges. *Legume Sci.* 3 (3), e96. doi: 10.1002/leg3.96

Carbonaro, M., Maselli, P., and Nucara, A. (2015). Structural aspects of legume proteins and nutraceutical properties. *Food Res. Int.* 76, 19–30. doi: 10.1016/j.foodres.2014.11.007

Carbonaro, M., and Nucara, A. (2022). Legume proteins and peptides as compounds in nutraceuticals: A structural basis for dietary health effects. *Nutrients* 14 (6), 1188. doi: 10.3390/nu14061188

Carrero-Carralero, C., Mansukhani, D., Ruiz-Matute, A. I., Martínez-Castro, I., Ramos, L., and Sanz, M. L. (2018). Extraction and characterization of low molecular weight bioactive carbohydrates from mung bean (*Vigna radiata*). *Food Chem.* 266, 146–154. doi: 10.1016/j.foodchem.2018.05.114

Chandrasekara, A., and Josheph Kumar, T. (2016). Roots and tuber crops as functional foods: A review on phytochemical constituents and their potential health benefits. *Int. J. Food Sci.* 2016, 3631647. doi: 10.1155/2016/3631647

Charrondière, U. R., Vincent, A., and Grande, F. (2020). *"FAO/INFOODS Food Composition Table for Western Africa (2019): User Guide & Condensed Food Composition Table = Table de Composition Des Aliments FAO/INFOODS Pour l'Afrique de l'Ouest (2019): Guide d'utilisation & Table de Composition Des Aliments Condensée." AGRIS: International Information System for the Agricultural Science and Technology*. Available at: https://agris.fao.org/agrissearch/search.do?ecordID=XF2020000998. (Accessed March 10, 2023).

Chaturvedi, S., and Chakraborty, S. (2021). Review on potential non-dairy synbiotic beverages: A preliminary approach using legumes. *Int. J. Food Sci. Technol.* 56 (5), 2068–2077. doi: 10.1111/ijfs.14779

Chaturvedi, S., and Chakraborty, S. (2022). Evaluation of prebiotic properties of legume-based synbiotic beverages. *J. Food Proc. Preserv.* 46, e16685. doi: 10.1111/jfpp.16685

Chaves, J. O., De Souza, M. C., Da Silva, L. C., Lachos-Perez, D., Torres-Mayanga, P. C., Da Fonseca Machado, A. P., et al. (2020). Extraction of flavonoids from natural sources using modern techniques. *Front. Chem.* 8. doi: 10.3389/fchem.2020.507887

Chen, L., Wu, Q., He, T., Lan, J., Ding, L., Liu, T., et al. (2020). Transcriptomic and metabolomic changes triggered by fusarium solani in common bean (*Phaseolus vulgaris* l.). *Genes* 11, 177. doi: 10.3390/genes11020177

Chingwaru, W., Vidmar, J., Kapewangolo, P. T., Mazimba, O., and Jackson, J. (2015). Therapeutic and prophylactic potential of morama (Tylosema esculentum): A review. *Phytother. Res.* 29, 1423–1438. doi: 10.1002/ptr.5419

Cichońska, P., and Ziarno, M. (2022). Legumes and legume-based beverages fermented with lactic acid bacteria as a potential carrier of probiotics and prebiotics. *Microorganisms* 10 (1), 91. doi: 10.3390/microorganisms1001009

Cornejo-Ramírez, Y. I., Martínez-Cruz, O., Del Toro-Sánchez, C. L., Wong-Corral, F. J., Borboa-Flores, J., and Cinco-Moroyoqui, F. J. (2018). The structural characteristics of starches and their functional properties. *CyTA-J. Food* 16, 1003–1017. doi: 10.1080/19476337.2018.1518343

Cui, H., Abdel-Samie, M. A.-S., Lin, L., and Jafari, S. M. (2021). "Application of antimicrobial-loaded nano/microcarriers in different food products," in *Application of*

Nano/Microencapsulated ingredients in food products (Elsevier), 469–517. doi: 10.1016/B978-0-12-815726-8.00012-X

Cullis, C., and Kunert, K. J. (2017). Unlocking the potential of orphan legumes. *J. exper. Bot.* 68, 1895–1903. doi: 10.1093/jxb/erw437

Curtin, S. J., Xiong, Y., Michno, J.-M., Campbell, B. W., Stec, A. O., Čermák, T., et al. (2018). CRISPR/Cas9 and TALENs generate heritable mutations for genes involved in small rna processing of *Glycine max* and *Medicago truncatula*. *Plant Biotechnol. J.* 16, 1125–1137. doi: 10.1111/pbi.12857

Del Hierro, J. N., Herrera, T., García-Risco, M. R., Fornari, T., Reglero, G., and Martin, D. (2018). Ultrasound-assisted extraction and bioaccessibility of saponins from edible seeds: Quinoa, lentil, fenugreek, soybean and lupin. *Food Res. Int.* 109, 440–447. doi: 10.1016/j.foodres.2018.04.058

Diapari, M., Sindhu, A., Bett, K., Deokar, A., Warkentin, T. D., and Tar'an, B. (2014). Genetic diversity and association mapping of iron and zinc concentrations in chickpea (*Cicer arietinum* l.). *Genome* 57, 459–468. doi: 10.1139/gen-2014-0108

Diedericks, C. F., Shek, C., Jideani, V. A., Venema, P., and van der Linden, E. (2020). Physicochemical properties and gelling behaviour of bambara groundnut protein isolates and protein-enriched fractions. *Food Res. Inter.* 138, 109773. doi: 10.1016/j.foodres.2020.109773

Divekar, P. A., Narayana, S., Divekar, B. A., Kumar, R., Gadratagi, B. G., Ray, A., et al. (2022). Plant secondary metabolites as defense tools against herbivores for sustainable crop protection. *IJMS* 23, 2690. doi: 10.3390/ijms23052690

Djuricic, I., and Calder, P. C. (2021). Beneficial outcomes of omega-6 and omega-3 polyunsaturated fatty acids on human health: An update for 2021. *Nutrients* 13, 2421. doi: 10.3390/nu13072421

Drago, S. R., Franco-Miranda, H., Cian, R. E., Betancur-Ancona, D., and Chel-Guerrero, L. (2016). Bioactive properties of *Phaseolus lunatus* (Lima bean) and *Vigna unguiculata* (Cowpea) hydrolyzates incorporated into pasta. residual activity after pasta cooking. *Plant Foods Hum. Nutr.* 71, 339–345. doi: 10.1007/s11130-016-0565-2

Du, S., Jiang, H., Yu, X., and Jane, J. (2014). Physicochemical and functional properties of whole legume flour. *LWT - Food Sci. Tech.* 55, 308–313. doi: 10.1016/j.lwt.2013.06.001

Dudu, O. E., Li, L., Oyedeji, A. B., Oyeyinka, S. A., and Ma, Y. (2019). Structural and functional characteristics of optimised dry-heat-moisture treated cassava flour and starch. *Int. J. Biol. Macromol.* 133, 1219–1227. doi: 10.1016/j.ijbiomac.2019.04.202

Ebrahim, A. E., Abd El-Aziz, N. K., Elariny, E. Y., Shindia, A., Osman, A., Hozzein, W. N., et al. (2022). Antibacterial activity of bioactive compounds extracted from red kidney bean (*Phaseolus vulgaris* l.) seeds against multidrug-resistant enterobacterales. *Front. Microbiol.* 13. doi: 10.3389/fmicb.2022.1035586

Falade, K. O., and Nwajei, C. P. (2015). Physical, proximate, functional and pasting properties of four non-and γ-irradiated bambara groundnut (Vigna subterranean) cultivars. *Int. J. Food Sci. Technol.* 50 (3), 640–651.

Foyer, C. H., Lam, H.-M., Nguyen, H. T., Siddique, K. H., Varshney, R. K., Colmer, T. D., et al. (2016). Neglecting legumes has compromised human health and sustainable food production. *Nat. Plants* 2, 16112. doi: 10.1038/nplants.2016.112

Fukushima, A., and Kusano, M. (2013). Recent progress in the development of metabolome databases for plant systems biology. *Front. Plant Sci.* 4, 73. doi: 10.3389/fpls.2013.00073

Gagné-Bourque, F., Bertrand, A., Claessens, A., Aliferis, K. A., and Jabaji, S. (2016). Alleviation of drought stress and metabolic changes in timothy (*Phleum pratense* l.) colonized with bacillus subtilis B26. *Fron. Plant Sc.* 7. doi: 10.3389/fpls.2016.00584

Ganesan, K., and Xu, B. (2018). A critical review on phytochemical profile and health promoting effects of mung bean (*Vigna radiata*). *Food Sci. Hum. Well.* 7, 11–33. doi: 10.1016/j.fshw.2017.11.002

Garg, R., and Jain, M. (2013). Transcriptome analyses in legumes: A resource for functional genomics. *Plant Genome* 6. doi: 10.3835/plantgenome2013.04.0011

Gemede, H. F., and Ratta, N. (2014). Antinutritional factors in plant foods: Potential health benefits and adverse effects. *Int. J. Nutrit. Food Sci.* 3, 284–289. doi: 10.11648/j.ijnfs.20140304.18

George, T. T., Obilana, A. O., and Oyeyinka, S. A. (2020). The prospects of African yam bean: past and future importance. *Heliyon* 6, e05458. doi: 10.1016/j.heliyon.2020.e05458

Gohara, A. K., Souza, A. H. P., de Gomes, S. T. M., Souza, N. E., de, Visentainer, J. V., and Matsushita, M. (2016). Nutritional and bioactive compounds of adzuki beans cultivars using chemometric approach. *Ciênc. agrotec.* 40, 104–113. doi: 10.1590/S1413-70542016000100010

Gupta, C., and Prakash, D. (2019). "Nutraceuticals from microbes of marine sources," in *Nutraceuticals-past, present and future* (United Kingdom (UK): Intech Open), 99. doi: 10.5772/intechopen.82369

Haase, H., and Schomburg, L. (2019). "You'd better zinc–trace element homeostasis in infection and inflammation".). *Nutrient* 11 (9), 2078. doi: 10.3390/nu11092078

Hai, H., Thi, P., and Nguyen, T. L. (2021). "Nutraceutical properties of legume seeds," in *Phytochemical compounds legume research*, vol. 2. (United Kingdom (UK): Intech Open).

Hamilton, J. P., and Robin Buell, C. (2012). Advances in plant genome sequencing. *Plant J.* 70, 177–190. doi: 10.1111/j.1365-313X.2012.04894.x

Harris, T., Jideani, V., and Le Roes-Hill, M. (2018). Flavonoids and tannin composition of bambara groundnut (*Vigna subterranea*) of mpumalanga, south Africa. *Heliyon* 4, e00833. doi: 10.1016/j.heliyon.2018.e00833

Heinrich, M., Mah, J., and Amirkia, V. (2021). Alkaloids used as medicines: Structural phytochemistry meets biodiversity–an update and forward look. *Molecules* 26, 1836. doi: 10.3390/molecules26071836

Henciya, S., Seturaman, P., James, A. R., Tsai, Y.-H., Nikam, R., Wu, Y.-C., et al. (2017). Biopharmaceutical potentials of *Prosopis* spp.(mimosaceae, leguminosa). *J.f Food Drug Anal.* 25, 187–196. doi: 10.1016/j.jfda.2016.11.001

Hossain, S., Ahmed, R., Bhowmick, S., Mamun, A. A., and Hashimoto, M. (2016). Proximate composition and fatty acid analysis of *Lablab purpureus* (L.) legume seed: implicates to both protein and essential fatty acid supplementation. *SpringerPlus* 5(1), 1899. doi: 10.1186/s40064-016-3587-1

Huang, J., Yang, Q., and Pu, H. (2018). "Slowly digestible starch," in *Functional starch and applications in food* (United Kingdom (UK): Springer), 27–61.

Huldani, H., Rashid, A. I., Turaev, K. N., Opulencia, M. J. ,. C., Abdelbasset, W. K., Bokov, D. O., et al. (2022). Concanavalin a as a promising lectin-based anti-cancer agent: the molecular mechanisms and therapeutic potential. *Cell Commun. Signal* 20, 167. doi: 10.1186/s12964-022-00972-7

Hunter, D., Borelli, T., Beltrame, D. M., Oliveira, C. N., Coradin, L., Wasike, V. W., et al. (2019). The potential of neglected and underutilized species for improving diets and nutrition. *Planta*, 1–21. doi: 10.1007/s00425-019-03169-4

Hussein, H., Awad, S., El-Sayed, I., and Ibrahim, A. (2020). Impact of chickpea as prebiotic, antioxidant and thickener agent of stirred bio-yoghurt. *Annal. Agric. Sci.* 65, 49–58. doi: 10.1016/j.aoas.2020.03.001

Idoko, A., Onyinye, A. O., Blessing, N. O., Ayomide, T., Phillip, O. C., and Nwali, O. N. (2020). Heating effect on phytochemical and proximate contents of cooked aqueous extract of *Phaseolus vulgaris* (Kidney beans). *Univ J. Pharm. Res* 4 (6), 35-41. doi: 10.22270/ujpr.v4i6.334

Ikhajiagbe, B., Ogwu, M. C., Ogochukwu, O. F., Odozi, E. B., Adekunle, I. J., and Omage, Z. E. (2021). The place of neglected and underutilized legumes in human nutrition and protein security in Nigeria. *Crit. Rev. Food Sc. Nutr.*, 1–10. doi: 10.1080/10408398.2020.1871319

Iuga, M., and Mironeasa, S. (2020). A review of the hydrothermal treatments impact on starch based systems properties. *Crit. Rev.Food Sci. Nutr.*, 1–26. doi: 10.1080/10408398.2019.1664978

Iwe, M. O., Onyeukwu, U., and Agiriga, A. N. (2016). Proximate, functional and pasting properties of FARO 44 rice, African yam bean and brown cowpea seeds composite flour. *Cogent Food Agric.* 2 (1), 1142409.

James, S., Nwabueze, T., Ndife, J., Onwuka, G., Usman, M., and Audu, Y. (2020). Effects of treatments on some bioactive components of selected lesser known legumes indigenous to Nigeria. *J. Food Chem. Nanotechnol.* 6, 197–206. doi: 10.17756/jfcn.2020-102

Jha, A. B., Purves, R. W., Elessawy, F. M., Zhang, H., Vandenberg, A., and Warkentin, T. D. (2019). Polyphenolic profile of seed components of white and purple flower pea lines. *Crop Sci.* 59, 2711–2719. doi: 10.2135/cropsci2019.04.0279

Ji, J., Zhang, C., Sun, Z., Wang, L., Duanmu, D., and Fan, Q. (2019). Genome editing in cowpea (*Vigna unguiculata* (L.)) using CRISPR-Cas9. *Int. J. Mol. Sci.* 20 (10), 2471. doi: 10.3390/ijms20102471

Jia, R., Ge, S., Ren, S., Luo, Y., Xiu, L., Liu, H., et al. (2021). Antibacterial mechanism of adzuki bean seed coat polyphenols and their potential application in preservation of fresh raw beef. *Int. J.f Food Sci. Technol.* 56, 5025–5039. doi: 10.1111/ijfs.15292

Johnson, N., Johnson, C. R., Thavarajah, P., Kumar, S., and Thavarajah, D. (2020). The roles and potential of lentil prebiotic carbohydrates in human and plant health. *Plants People Planet* 2, 310–319. doi: 10.1002/ppp3.10103

Johnson, J. B., Neupane, P., Bhattarai, S. P., Trotter, T., and Naiker, M. (2022). Partitioning of nutritional and phytochemical constituents in nine adzuki bean genotypes from Australia. *J. Agric. Food Res.* 10, 100398. doi: 10.1016/j.jafr.2022.100398

Joshi, J., Hasnain, G., Logue, T., Lynch, M., Wu, S., Guan, J., et al. (2021). A core metabolome response of maize leaves subjected to long-duration abiotic stresses. *Metabolites* 11, 797. doi: 10.3390/metabo11110797

Junot, C., Fenaille, F., Colsch, B., and Becher, F. (2014). High-resolution mass spectrometry-based techniques at the crossroads of metabolic pathways. *Mass Spectrom. Rev.* 33, 471–500. doi: 10.1002/mas.21401

Juranić, M., Nagahatenna, D. S., Salinas-Gamboa, R., Hand, M. L., Sánchez-León, N., Leong, W. H., et al. (2020). A detached leaf assay for testing transient gene expression and gene editing in cowpea (Vigna unguiculata [L.] walp.). *Plant Methods* 16 (1), 1–17.

Kan, L., Nie, S., Hu, J., Wang, S., Cui, S. W., Li, Y., et al. (2017). Nutrients, phytochemicals and antioxidant activities of 26 kidney bean cultivars. *Food Chem. Toxicol.* 108, 467–477. doi: 10.1016/j.fct.2016.09.007

Kanatt, S. R., Arjun, K., and Sharma, A. (2011). Antioxidant and antimicrobial activity of legume hulls. *Food Res. Int.* 44, 3182–3187. doi: 10.1016/j.foodres.2011.08.022

Kang, Y. J., Kim, S. K., Kim, M. Y., Lestari, P., Kim, K. H., Ha, B.-K., et al. (2014). Genome sequence of mungbean and insights into evolution within *Vigna species*. *Nat. Commun.* 5 (1), 1–9. doi: 10.1038/ncomms6443

Kaundal, S. P., Sharma, A., Kumar, R., Kumar, V., and Kumar, R. (2019). Exploration of medicinal importance of an underutilized legume crop, *Macrotyloma uniflorum* (Lam.) Verdc.(Horse gram): A review. *Int. J. Pharm. Sci. Res.* 10, 3178–3186. doi: 10.13040/IJPSR.0975-8232.10(7).3178-86

Kaur, R., and Arora, S. (2015). Alkaloids–important therapeutic secondary metabolites of plant origin. *J. Crit. Rev.* 2, 1–8.

Kaur, D., Rasane, P., Dhawan, K., Singh, J., Kaur, S., Gurumayum, S., et al. (2021). Rice bean (*Vigna umbellata*) based ready-to-eat geriatric premix: Optimization and analysis. *J. Food Process. Preserv.* 45. doi: 10.1111/jfpp.16075

Keawpeng, I., Paulraj, B., and Venkatachalam, K. (2022). Antioxidant and antimicrobial properties of mung bean phyto-film combined with longkong pericarp extract and sonication. *Membranes* 12, 379. doi: 10.3390/membranes12040379

Kennedy, D. O. (2014). Polyphenols and the human brain: Plant "secondary metabolite" ecologic roles and endogenous signaling functions drive benefits. *Adv. Nutr.* 5, 515–533. doi: 10.3945/an.114.006320

Khazaei, H., Podder, R., Caron, C. T., Kundu, S. S., Diapari, M., Vandenberg, A., et al. (2017). Marker–trait association analysis of iron and zinc concentration in lentil (*Lens culinaris* medik.) seeds. *Plant Genome* 10(2). doi: 10.3835/plantgenome2017.02.0007

Khoury, C. K. (2015). *The conservation and use of crop genetic resources for food security* (United Kingdom (UK): Wageningen University and Research Centre).

Klompong, V., and Benjakul, S. (2015). Antioxidative and antimicrobial activities of the extracts from the seed coat of bambara groundnut (*Voandzeia subterranea*). *RSC Adv.* 5, 9973–9985. doi: 10.1039/C4RA10955D

Konyeme, T., Nyananyo, B., and Tanee, F. (2020). Diversity in proximate analysis of tubers of some African yam bean (*Sphenostylis stenocarpa*)(Hochst ex. a. rich.) harms (Fabaceae) accessions. *J. Appl. Sci. Environ. Manage.* 24, 1787–1793. doi: 10.4314/jasem.v24i10.12

Kortt, A. A. (1984). Purification and properties of the basic lectins from winged bean seed [*Psophocarpus tetragonolobus* (L.) DC]. *Eur. J. Biochem.* 138, 519–525. doi: 10.1111/j.1432-1033.1984.tb07946.x

Ku, Y.-S., Contador, C. A., Ng, M.-S., Yu, J., Chung, G., and Lam, H.-M. (2020). The effects of domestication on secondary metabolite composition in legumes. *Front. Genet.* 11. doi: 10.3389/fgene.2020.581357

Kumar, A., Anju, T., Kumar, S., Chhapekar, S. S., Sreedharan, S., Singh, S., et al. (2021). Integrating omics and gene editing tools for rapid improvement of traditional food plants for diversified and sustainable food security. *Int. J. Mol. Sci.* 22, 8093. doi: 10.3390/ijms22158093

Kurek, J. (2019). "Introductory chapter: Alkaloids - their importance in nature and for human life," in *Alkaloids - their importance in nature and human life*. Ed. J. Kurek (United Kingdom (UK): IntechOpen). doi: 10.5772/intechopen.85400

Latha, L. Y., Sasidharan, S., Zuraini, Z., Suryani, S., Shirley, L., and Sangetha, S. (2007). Antibacterial activity and toxicity of psophocarpus tetragonolobus. *Pharm. Biol.* 45 (1), 31–36. doi: 10.1080/13880200601026317

Lin Tan, X., Azam-Ali, S., Goh, E. V., Mustafa, M. A., Chai, H. H., Kuan Ho, W., et al. (2020). Bambara groundnut: An underutilized leguminous crop for global food security and nutrition. *Front. Nutr.* 7, 276. doi: 10.3389/fnut.2020.601496

Lopes, N. A., and Brandelli, A. (2018). Nanostructures for delivery of natural antimicrobials in food. *Crit. Rev. Food Sci. Nutr.* 58, 2202–2212. doi: 10.1080/10408398.2017.1308915

Lu, S., Qian, Y., Huang, X., Yu, H., Yang, J., Han, R., et al. (2017). The association of dietary pattern and breast cancer in jiangsu, China: A population-based case-control study. *PloS One* 12 (9), e0184453. doi: 10.1371/journal.pone.0184453

Luo, J., Cai, W., Wu, T., and Xu, B. (2016). Phytochemical distribution in hull and cotyledon of adzuki bean (*Vigna angularis* l.) and mung bean (Vigna radiate l.), and their contribution to antioxidant, anti-inflammatory and anti-diabetic activities. *Food Chem.* 201, 350–360. doi: 10.1016/j.foodchem.2016.01.101

Makeri, M. U., Karim, R., Adbulkarim, M. S., Ghazali, H. M., Miskandar, M. S., and Muhammad, K. (2016). Comparative analysis of the physiochemical, thermal, and oxidative properties of winged bean and soybean oils. *Int. J. Food Prop.* 19, 2769–2278. doi: 10.1080/10942912.2015.1031246

Makhumbila, P., Rauwane, M., Muedi, H., and Figlan, S. (2022). Metabolome profiling: A breeding prediction tool for legume performance under biotic stress conditions. *Plants* 11, 1756. doi: 10.3390/plants11131756

Mani-López, E., Palou, E., and López-Malo, A. (2021). Legume proteins, peptides, water extracts, and crude protein extracts as antifungals for food applications. *Trend. Food Sci. Technol.* 112, 16–24. doi: 10.1016/j.tifs.2021.03.035

Maoka, T. (2020). Carotenoids as natural functional pigments. *J. Nat. Med.* 74, 1–16. doi: 10.1007/s11418-019-01364-x

Maphosa, Y., and Jideani, V. A. (2017). "The role of legumes in human nutrition," in *Functional food-improve health through adequate food*, vol. 1. (United Kingdom (UK): Intech Open), 13. doi: 10.5772/intechopen.69127

Marco, M. L., Sanders, M. E., Gänzle, M., Arrieta, M. C., Cotter, P. D., De Vuyst, L., et al. (2021). The international scientific association for probiotics and prebiotics (ISAPP) consensus statement on fermented foods. *Nat. Rev. Gastro. Hepatol.* 18 (3), 196–208. doi: 10.1038/s41575-020-00390-5

Marrelli, M., Conforti, F., Araniti, F., and Statti, G. A. (2016). Effects of saponins on lipid metabolism: a review of potential health benefits in the treatment of obesity. *Molecules* 21, 1404. doi: 10.3390/molecules21101404

Maruatona, G. N., Duodu, K. G., and Minnaar, A. (2010). Physicochemical, nutritional and functional properties of marama bean flour. *Food Chem.* 121, 400–405. doi: 10.1016/j.foodchem.2009.12.054

Mayes, S., Ho, W. K., Chai, H. H., Gao, X., Kundy, A. C., Mateva, K. I., et al. (2019). Bambara groundnut: An exemplar underutilized legume for resilience under climate change. *Planta* 250, 803–820. doi: 10.1007/s00425-019-03191-6

Meng, Y., Hou, Y., Wang, H., Ji, R., Liu, B., Wen, J., et al. (2017). Targeted mutagenesis by CRISPR/Cas9 system in the model legume medicago truncatula. *Plant Cell Rep.* 36, 371–374. doi: 10.1007/s00299-016-2069-9

Messina, M. J., Persky, V., Setchell, K. D. R., and Barnes, S. (1994). Soy intake and cancer risk: a review of the *in vitro* and *in vivo* data. *Nutr. Cancer* 21, 113–131.

Mirali, M., Purves, R. W., and Vandenberg, A. (2017). Profiling the phenolic compounds of the four major seed coat types and their relation to color genes in lentil. *J. Nat. Prod* 80, 1310–1317. doi: 10.1021/acs.jnatprod.6b00872

Mochida, K., and Shinozaki, K. (2011). Advances in omics and bioinformatics tools for systems analyses of plant functions. *Plant Cell Physiol.* 52, 2017–2038. doi: 10.1093/pcp/pcr153

Mohanty, C. S., Verma, S., Singh, V., Khan, S., Gaur, P., Gupta, P., et al. (2013). Characterization of winged bean (*Psophocarpus tetragonolobus* (L.) DC.) based on molecular, chemical and physiological parameters. *AJMB* 03, 187–197. doi: 10.4236/ajmb.2013.34025

Mojica, L., and González de Meija, E. (2015). Characterization and comparison of protein and peptide profiles and their biological activities of improved common bean cultivars (*Phaseolus vulgaris* l.) from Mexico and Brazil. *Plant Foods Hum. Nutr.* 70, 105–112. doi: 10.1007/s11130-015-0477-6

Mubaiwa, J., Fogliano, V., Chidewe, C., and Linnemann, A. R. (2018). Bambara groundnut (*Vigna subterranea* (L.) verdc.) flour: A functional ingredient to favour the use of an unexploited sustainable protein source. *PloS One* 13, e0205776. doi: 10.1371/journal.pone.0205776

Nawaz, H., Aslam, M., Rehman, T., and Mehmood, R. (2021). Modification of emulsifying properties of cereal flours by blending with legume flours. *Asian J. Dairy Food Res.*, 1–6. doi: 10.18805/ajdfr.DR-223

Nayak, S., Malavalli, S. S., Bharati, P., Hefferon, K. L., Kole, C., and Puppala, N. (2021). Omics technologies to enhance plant based functional foods: An overview. *Front. Genet.* 12. doi: 10.3389/fgene.2021.742095

Negi, K. S., and Gaur, R. D. (1994). Principle wild food plants of western himalaya, uttar pradesh, India. *Hishar Plants Indian Subconti-nent* (India) pp, 1–78.

Nnamani, C. V., Ajayi, S. A., Oselebe, H. O., Atkinson, C. J., Adewale, D. B., Igwe, D. O., et al. (2018). Updates on nutritional diversity in *Sphenostylis stenocarpa* (Hoechst ex. a. rich.) harms. for food security and conservation. *Am. J. Agric. Biol. Sci.* 13, 38–49. doi: 10.3844/ajabssp.2018.38.49

Nnamani, C. V., Ajayi, S. A., Oselebe, H. O., Atkinson, C. J., Igboabuchi, A. N., and Ezigbo, E. C. (2017). *Sphenostylis stenocarpa* (ex. a. rich.) harms., a fading genetic resource in a changing climate: prerequisite for conservation and sustainability. *Plan. Theory* 6, 30. doi: 10.3390/plants6030030

Nonye, H. U. H., and Chinasa, O. R. (2022). Evaluation of pasting and functional properties of flour blends made from African yam bean (Sphenostylis stenocarpa) and corn (Zea mays) seed. *World J. Innovat. Res.* 12 (3), 19–24.

Nour, A. H., Oluwaseun, A. R., Nour, A. H., Omer, M. S., and Ahmed, N. (2021). "Microwave-assisted extraction of bioactive compounds," in *Microwave heating*. Ed. G. L. Churyumov (United Kingdom (UK): IntechOpen), (Ed). doi: 10.5772/intechopen.96092

Nwajagu, I. U., Garba, A., Nzelibe, H. C., Chukwuekezie, N. E., Abah, C. R., Umar, A. T., et al. (2021). Effect of processing on the nutrient, anti-nutrient and functional properties of *Mucuna flagellipes* (Ox-eyed bean) seed flour; An underutilized legume in Nigeria. *Ame. J. Food Nutr.* 9, 49–59. doi: 10.12691/ajfn-9-1-7

Obatolu, V. A., Fasoyiro, S. B., and Ogunsumi, L. (2001). Effect of processing on functional properties of yam beans (Sphenostylis stenocarpa). *Food Sci. Techn. Res.* 7 (4), 319–322. doi: 10.3136/fstr.7.319

Oboh, G., Agunloye, O. M., Adefegha, S. A., Akinyemi, A. J., and Ademiluyi, A. O. (2015). Caffeic and chlorogenic acids inhibit key enzymes linked to type 2 diabetes (*in vitro*): a comparative study. *J. Basic Clin. Physiol. Pharmacol.* 26, 165–170. doi: 10.1515/jbcpp-2013-0141

Ojuederie, O. B., Ajiboye, J. A., and Babalola, O. O. (2020). Biochemical and histopathological studies of key tissues in healthy Male wistar rats fed on African yam bean seed and tuber meals. *J. Food Qual.* 2020, 8892618. doi: 10.1155/2020/8892618

Ojuederie, O., and Balogun, M. (2017). Genetic variation in nutritional properties of African yam bean *Sphenostylis stenocarpa* (Hochst ex. a. rich. harms) accessions. *Nigerian J. Agric. Food Environ.* 13, 180–187.

Ojuederie, O., and Balogun, M. (2019). African Yam bean (*Sphenostylis stenocarpa*) tubers for nutritional security. *J. Underut. Leg.* 1, 56–68.

Ojuederie, O. B., Popoola, J. O., Aremu, C., and Babalola, O. O. (2021). "Harnessing the hidden treasures in African yam bean (Sphenostylis stenocarpa), an underutilized grain legume with food security potentials," in *Food security and safety*. Ed. O. O. Babalola (Cham: Springer), 1–20. doi: 10.1007/978-3-030-50672-8_1

Okafor, J. N., Jideani, V. A., Meyer, M., and Le Roes-Hill, M. (2022). Bioactive components in bambara groundnut (*Vigna subterraenea* (L.) verdc) as a potential source of nutraceutical ingredients. *Heliyon* 8(3), e09024. doi: 10.1016/j.heliyon.2022.e09024

Okechukwu-Ezike, N., Munonye, J., and Esiegwu, A. (2020). Comparative evaluation of the proximate composition, anti-nutrients and functional properties of some underutilized pulses. *Food Sci. Qual. Manage.* 95, 72–76.

Okezie, B.O., and Bello, A. B. (1988). Physicochemical and functional properties of winged bean flour and isolate compared with soy isolate. *J. Food Sci.* 53, 450–454. doi: 10.1111/j.1365-2621.1988.tb07728.x

Omotayo, A. O., and Aremu, A. O. (2019). "Nutritional, phytochemical and economic potential of marama bean," in *The 'green gold' of southern Africa* (Boston, USA: Harvard University). doi: 10.13140/RG.2.2.22809.36961

Omotayo, A. O., and Aremu, A. O. (2021). Marama bean [*Tylosema esculentum* (Burch.) a. schreib.]: an indigenous plant with potential for food, nutrition, and economic sustainability. *Food Funct.* 12, 2389–2403. doi: 10.1039/D0FO01937B

Onimawo, I. A., Momoh, A. H., and Usman, A. (1998). Proximate composition and functional properties of four cultivars of bambara groundnut (*Voandezeia subterranea*). *Plant Foods Hum. Nutri.* 53, 153–158. doi: 10.1023/A:1008027030095

Onyango, C. (2016). Starch and modified starch in bread making: A review. *Afr. J.f Food Sci.* 10, 344–351. doi: 10.5897/AJFS2016.1481

Oshodi, A. A., Ipinmoroti, K. O., and Adeyeye, E. I. (1997). Functional properties of some varieties of African yam bean (*Sphenostylis stenocarpa*) flour — III. *Int. J. Food Sci. Nutr.* 48, 243–250. doi: 10.3109/09637489709028568

Oyeyinka, S. A., Abdulsalam, A. O., Ahmed El-Imam, A. M., Oyeyinka, A. T., Olagunju, O. F., Kolawole, F. L., et al. (2021). Total phenolic content, antioxidant, anti-inflammatory and anti-microbial potentials of bambara groundnut (*Vigna subterranea* l.) seed extract. *Brit. Food J.* 123, 3421–3435. doi: 10.1108/BFJ-07-2020-0637

Paliwal, R., Adegboyega, T. T., Abberton, M., Faloye, B., and Oyatomi, O. (2021). Potential of genomics for the improvement of underutilized legumes in sub-Saharan Africa. *Legume Sci.* 3 (2), 1–16. doi: 10.1002/leg3.69

Palupi, H. T., Estiasih, T., Yunianta, and Sutrisno, A. (2021). Characterization of nutritional and functional properties of Lima bean flour (*Phaseolus lunatus* l.). *IOP Conf. Ser.: Earth Environ. Sci.* 924, 12033. doi: 10.1088/1755-1315/924/1/012033

Pandey, M. K., Roorkiwal, M., Singh, V. K., Ramalingam, A., Kudapa, H., Thudi, M., et al. (2016). Emerging genomic tools for legume breeding: current status and future prospects. *Front. Plant Sci.* 7, 455. doi: 10.3389/fpls.2016.00455

Panghal, A., Janghu, S., Virkar, K., Gat, Y., Kumar, V., and Chhikara, N. (2018). Potential non-dairy probiotic products–a healthy approach. *Food Biosci.* 21, 80–89. doi: 10.1016/j.fbio.2017.12.003

Patel, M., Pandey, S., Kumar, M., Haque, M., Pal, S., and Yadav, N. (2021). Plants metabolome study: Emerging tools and techniques. *Plants* 10, 2409. doi: 10.3390/plants10112409

Patterson, E., Wall, R., Fitzgerald, G. F., Ross, R. P., and Stanton, C. (2012). Health implications of high dietary omega-6 polyunsaturated fatty acids. *J. nutri. Metabol.* 2012, 539426. doi: 10.1155/2012/539426

Pattnaik, M., Pandey, P., Martin, G. J., Mishra, H. N., and Ashokkumar, M. (2021). Innovative technologies for extraction and microencapsulation of bioactives from plant-based food waste and their applications in functional food development. *Foods* 10, 279. doi: 10.3390/foods10020279

Popoola, J. O., Ajani, O. C., Oyesola, O. L., and Oniha, M. (2022a). Comparative analysis of phytochemical and proximate contents of three orphan legume species. *Asian J. Plant Sci.* 21, 520–528. doi: 10.3923/ajps.2022.520.528

Popoola, J. O., Aworunse, O. S., Ojuederie, O. B., Adewale, B. D., Ajani, O. C., Oyatomi, O. A., et al. (2022b). The exploitation of orphan legumes for food, income, and nutrition security in Sub-Saharan Africa. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.782140

Popoola, J., Ojuederie, O., Omonhinmin, C., and Adegbite, A. (2020). "Neglected and underutilized legume crops: Improvement and future prospects," in *Recent advances in grain crops research*. Eds. F. Shah, Z. Khan, A. Iqbal, M. Turan and M. Olgun (United Kingdom (UK): IntechOpen). doi: 10.5772/intechopen.87069

Rai, K. K., Pandey, N., Meena, R. P., and Rai, S. P. (2021). Biotechnological strategies for enhancing heavy metal tolerance in neglected and underutilized legume crops: A comprehensive review. *Ecotoxicol. Environ. Safe.* 208, 111750. doi: 10.1016/j.ecoenv.2020.111750

Ramatsetse, K. E., Ramashia, S. E., and Mashau, M. E. (2023). A review on health benefits, antimicrobial and antioxidant properties of bambara groundnut (*Vigna subterranean*). *Int. J. Food Prop.* 26, 91–107. doi: 10.1080/10942912.2022.2153864

Rasika, D. M. D., Vidanarachchi, J. K., Rocha, R. S., Balthazar, C. F., Cruz, A. G., Sant'ana, A. S., et al. (2021). Plant-based milk substitutes as emerging probiotic carriers. *Curr. Opin. Food Sci.* 38, 8–20. doi: 10.1016/j.cofs.2020.10.025

Rathi, D., Chakraborty, S., and Chakraborty, N. (2021). Grasspea, a critical recruit among neglected and underutilized legumes, for tapping genomic resources. *Curr. Plant Biol.*, 100200. doi: 10.1016/j.cpb.2021.100200

Ratnawati, L., Desnilasari, D., Surahman, D. N., and Kumalasari, R. (2019). Evaluation of physicochemical, functional and pasting properties of soybean, mung bean and red kidney bean flour as ingredient in biscuit. *IOP Conf. Ser.: Earth Environ. Sci.* 251, 12026. doi: 10.1088/1755-1315/251/1/012026

Roohani, N., Hurrell, R., Kelishadi, R., and Schulin, R. (2013). Zinc and its importance for human health: An integrative review. *J. Res. Med. Sci.* 18 (2), 144–157.

Roy, F., Boye, J. I., and Simpson, B. K. (2010). Bioactive proteins and peptides in pulse crops: Pea, chickpea and lentil. *Food Res. Int.* 43, 432–442. doi: 10.1016/j.foodres.2009.09.002

Salawu, S. (2016). Comparative study of the antioxidant activities of methanolic extract and simulated gastrointestinal enzyme digest of bambara nut (*Vigna subterranea*) *FUTA. J. Res.Sci.* 1, 107–120.

Sanders, M. E., Merenstein, D. J., Reid, G., Gibson, G. R., and Rastall, R. A. (2019). Probiotics and prebiotics in intestinal health and disease: from biology to the clinic. *Nat. Rev. Gastro. Hepatol.* 16 (10), 605–616. doi: 10.1038/s41575-019-0173-3

Sattar, D., Ali, T. M., and Hasnain, A. (2017). Effect of germination on enzymatic, functional and bioactive attributes of different Pakistani legume cultivars. *J. Food Meas. Charact.* 11, 2076–2086. doi: 10.1007/s11694-017-9591-5

Schafleitner, R., Nair, R. M., Rathore, A., Wang, Y.-w., Lin, C.-y., Chu, S.-h., et al. (2015). The AVRDC–the world vegetable center mungbean (*Vigna radiata*) core and mini core collections. *BMC Genomics* 16, 344. doi: 10.1186/s12864-015-1556-7

Selvamuthukumaran, M., and Shi, J. (2017). Recent advances in extraction of antioxidants from plant by-products processing industries. *Food Qual. Safe.* 1, 61–81. doi: 10.1093/fqs/fyx004

Semba, R. D., Ramsing, R., Rahman, N., Kraemer, K., and Bloem, M. W. (2021). Legumes as a sustainable source of protein in human diets. *Global Food Secur.* 28, 100520. doi: 10.1016/j.gfs.2021.100520

Sergio, O., Saldívar, S., and Hernández, D. S. (2020). "Dietary fiber in cereals, legumes,pseudocereals and other seeds," in *Science and technology of fibers in food systems, food engineering series*. Ed. J. Welti-Chanes, O. S Sergio, C. Osvaldo and T. Viridiana (Switzerland AG: Springer Nature). doi: 10.1007/978-3-030-38654-2_5

Sharma, N., Bisht, S. S., Gupta, S., Rana, M., and Kumar, A. (2019). Nutraceutical evaluation of horse gram (*Macrotyloma uniflorum*) cultivated in high altitudes of uttarakhand himalaya, India. *Indian J. Pure Appl. Biosci.* 7, 190–202.

Shelke, D. B., Tayade, S., Gawande, P., and Sonawane, H. B. (2022). GC-MS analysis and antioxidant potential of wild underutilized medicinally important legume, velvet bean (*Mucuna pruriens* l. DC.). *Not. Sci. Biol.* 14, 11098–11098. doi: 10.15835/nsb14111098

Shevkani, K., Kaur, M., and Singh, N. (2021). Composition, pasting, functional, and microstructural properties of flours from different split dehulled pulses (dhals). *J. Food Process. Preserv.* 45(6). doi: 10.1111/jfpp.15485

Shevkani, K., Singh, N., Patil, C., Awasthi, A., and Paul, M. (2022). Antioxidative and antimicrobial properties of pulse proteins and their applications in gluten-free foods and sports nutrition. *Int. J. Food Sci. Tech* 57, 5571–5584. doi: 10.1111/ijfs.15666

Shi, J., Arunasalam, K., Yeung, D., Kakuda, Y., Mittal, G., and Jiang, Y. (2004). Saponins from edible legumes: Chemistry, processing, and health benefits. *J. Med. Food* 7 (1), 67–78. doi: 10.1089/109662004322984734

Silva, E., Souza, J., Rogez, H., Rees, J.-F., and Larondelle, Y. (2007). Antioxidant activities and polyphenolic contents of fifteen selected plant species from the Amazonian region. *Food Chem.* 101, 1012–1018. doi: 10.1016/j.foodchem.2006.02.055

Singh, J. P., Singh, B., and Kaur, A. (2021). Bioactive compounds of legume seeds. *Bioactive Compounds Underutilized Vegetables Legumes*, 645–665. doi: 10.1007/978-3-030-57415-4_33

Singh, B., Singh, J. P., Shevkani, K., Singh, N., and Kaur, A. (2017). Bioactive constituents in pulses and their health benefits. *J. Food Sci. Technol.* 54, 858–870. doi: 10.1007/s13197-016-2391-9

Sitohy, M., Mahgoub, S., Osman, A., El-Masry, R., and Al-Gaby, A. (2013). Extent and mode of action of cationic legume proteins against *Listeria monocytogenes* and *Salmonella enteritidis*. *Probiotics Antimicrob. Proteins* 5, 195–205. doi: 10.1007/s12602-013-9134-2

Soetan, K. O. (2017). Preliminary studies on the nutritional and antinutritional constituents of six accessions of African yam bean (*Sphenostylis stenocarpa*): An underutilized legume. *J.Appl. Res. Child.* 8.

Soetan, K., and Adeola, A. (2018). Comparative nutritional and functional properties of selected underutilized legumes. *Nigerian J. Anim. Product.* 45, 96–106-196–106.

Soetan, K., and Atanda, T. (2018). Proximate composition, phytochemical screening and some antinutritional factors of three accessions of lima beans (*Phaseolus lunatus*) an underutilized legume. *Nigerian J. Anim. Product.* 45, 71–78-71–78.

Soetan, K. O., Olaiya, C. O., and Karigidi, K. O. (2018). Comparative *in vitro* antioxidant activities of six accessions of African yam beans (*Sphenostylis stenocarpa* l.). *Ann. Food Sci. Technol.* 19, 455–461.

Solarte, D. A., Ruiz-Matute, A. I., Chito-Trujillo, D. M., Rada-Mendoza, M., and Sanz, M. L. (2021). Microwave assisted extraction of bioactive carbohydrates from different morphological parts of alfalfa (*Medicago sativa* l.). *Foods* 10, 346. doi: 10.3390/foods10020346

Solomon, S. G., Okomoda, V. T., and Oguche, O. (2018). Nutritional value of raw *Canavalia ensiformis* and its utilization as partial replacement for soybean meal in the diet of *Clarias gariepinus* (Burchell 1822) fingerlings. *Food Sci. Nutr.* 6, 207–213. doi: 10.1002/fsn3.548

Sowndhararajan, K., Siddhuraju, P., and Manian, S. (2011). Antioxidant activity of the differentially processed seeds of jack bean (*Canavalia ensiformis* l. DC). *Food Sci. Biotechnol.* 20, 585–591. doi: 10.1007/s10068-011-0083-9

Sriwichai, S., Monkham, T., Sanitchon, J., Jogloy, S., and Chankaew, S. (2021). Dual-purpose of the winged bean (*Psophocarpus tetragonolobus* (L.) DC.), the neglected tropical legume, based on pod and tuber yields. *Plants* 10, 1746. doi: 10.3390/plants10081746

Steinwand, M. A., and Ronald, P. C. (2020). Crop biotechnology and the future of food. *Nat. Food* 1, 273–283. doi: 10.1038/s43016-020-0072-3

Stosh, S. M., and Yada, S. (2010). Dietary fibres in pulse seeds and fractions: Characterization, functional attributes, and applications. *Food Res. Int.* 43, 450–460. doi: 10.1016/j.foodres.2009.09.005

Sun, X., and Weckwerth, W. (2012). COVAIN: A toolbox for uni- and multivariate statistics, time-series and correlation network analysis and inverse estimation of the differential Jacobian from metabolomics covariance data. *Metabolomics* 28, 81–93. doi: 10.1007/s11306-012-0399-3

Sutivisedsak, N., Cheng, H., Willett, J., Lesch, W., Tangsrud, R., and Biswas, A. (2010). Microwave-assisted extraction of phenolics from bean (*Phaseolus vulgaris* l.). *Food Res. Int.* 43, 516–519. doi: 10.1016/j.foodres.2009.09.014

Tadele, Z., and Bartels, D. (2019). Promoting orphan crops research and development. *Planta* 250, 675–676. doi: 10.1007/s00425-019-03235-x

Tan, S. S., Tan, C. X., and Tan, S. T. (2020). "Cold pressed peanut (Arachis hypogaea l.) oil," in *Cold pressed oils* (United States: Elsevier), 357–364. doi: 10.1016/B978-0-12-818188-1.00032-3

Tee, E. S., Goh, A. H., and Khor, S. C. (1995). Carotenoid composition and content of legumes, tubers and roots by HPLC. *Malaysian J. nutri.* 1 (1), 51–61.

Tian, J., Bryksa, B., and Yada, R. (2016). Feeding the world into the future–food and nutrition security: the role of food science and technology. *Front. Life Sci.* 9 (3), 155–166. doi: 10.1080/21553769.2016.1174958

Tjahjadi, C., Lin, S., and Breene, W. M. (1988). Isolation and characterization of adzuki bean (*Vigna angularis* cv takara) proteins. *J. Food Sci.* 53, 1438–1443. doi: 10.1111/j.1365-2621.1988.tb09294.x

Tripathi, K., Gore, P. G., Pandey, A., Nayar, E. R., Gayacharan, C., Pamarthi, R. K., et al. (2020). Morphological and nutritional assessment of *Vigna vexillata* (L.) a. rich.: a potential tuberous legume of India. *Genet. Resour Crop Evol.* 68, 397–408. doi: 10.1007/s10722-020-01023-1

Tsamo, A. T., Mohammed, M., and Dakora, F. D. (2020). Metabolite fingerprinting of kersting's groundnut [*Macrotyloma geocarpum* (Harms) maréchal & baudet] seeds using UPLC-qTOF-MS reveals the nutraceutical and antioxidant potentials of the orphan legume. *Front. Nutr.* 7, 306. doi: 10.3389/fnut.2020.593436

Udeh, E. L., Nyila, M. A., and Kanu, S. A. (2020). Nutraceutical and antimicrobial potentials of bambara groundnut (*Vigna subterranean*): A review. *Heliyon* 6, e05205. doi: 10.1016/j.heliyon.2020.e05205

Vadivel, V., Patel, A., and Biesalski, H. K. (2012). Effect of traditional processing methods on the antioxidant, α-amylase and α-glucosidase enzyme inhibition properties of sesbania sesban Merrill seeds. *CyTA-J. Food* 10, 128–136. doi: 10.1080/19476337.2011.601427

Vijayakumar, V. (2021). "Nutraceutical legumes: A brief review on the nutritional and medicinal values," in *Sustainable agriculture reviews*, vol. 51 . Eds. P. Guleria, V. Kumar and E. Lichtfouse (Cham: Springer), 1–28. doi: 10.1007/978-3-030-68828-8_1

Von Wettberg, E. B., Naito, K., Kur, A., and Ludidi, N. (2021). The unusual biogeography of zombi pea, *Vigna vexillata*. *Legumes* 8-9.

Voutilainen, S., Nurmi, T., Mursu, J., and Rissanen, T. H. (2006). Carotenoids and cardiovascular health. *Amer. J. Clin. Nutr.* 83, 1265–1271. doi: 10.1093/ajcn/83.6.1265

Wanyama, A., Orwa, J., Njenga, P., and Irungu, B. (2017). Evaluation of cytotoxicity, antimicrobial activities and minerals composition of *Vigna subterranea* (L.) verdc.(Bambara groundnut) extracts. *Afr. J. Health Sci.* 30, 88–104.

Wessels, I., and Rink, L. (2020). Micronutrients in autoimmune diseases: Possible therapeutic benefits of zinc and vitamin d. *J. Nutrit. Biochem.* 77, 108240. doi: 10.1016/j.jnutbio.2019.108240

Wessels, I., Rolles, B., and Rink, L. (2020). The potential impact of zinc supplementation on COVID-19 pathogenesis. *Front. Immunol.* 11, 1712. doi: 10.3389/fimmu.2020.01712

Wolabu, T. W., Park, J. J., Chen, M., Cong, L., Ge, Y., Jiang, Q., et al (2020). Improving the genome editing efficiency of CRISPR/Cas9 in *Arabidopsis* and *Medicago truncatula*. *Planta* 252, 15. doi: 10.1007/s00425-020-03415-0

Xu, D.-P., Li, Y., Meng, X., Zhou, T., Zhou, Y., Zheng, J., et al. (2017). Natural antioxidants in foods and medicinal plants: Extraction, assessment and resources. *IJMS* 18, 96. doi: 10.3390/ijms18010096

Yadav, U., Singh, N., Kaur, A., and Thakur, S. (2018). Physico-chemical, hydration, cooking, textural and pasting properties of different adzuki bean (*Vigna angularis*) accessions. *J. Food Sci. Technol.* 55, 802–810. doi: 10.1007/s13197-017-2994-9

Yang, J., de Wit, A., Diedericks, C. F., Venema, P., van der Linden, E., and Sagis, L. M. C. (2022). Foaming and emulsifying properties of extensively and mildly extracted bambara groundnut proteins: A comparison of legumin, vicilin and albumin protein. *Food Hydrocolloids* 123, 107190. doi: 10.1016/j.foodhyd.2021.107190

Yang, S., Grall, A., and Chapman, M. A. (2018). Origin and diversification of winged bean (*Psophocarpus tetragonolobus* (L.) DC.), a multipurpose underutilized legume. *Am. J. Bot.* 105, 888–897. doi: 10.1002/ajb2.1093

Yang, C. S., Landau, J. M., Huang, M.-T., and Newmark, H. L. (2001). Inhibition of carcinogenesis by dietary polyphenolic compounds. *Ann. Rev. Nutrit.* 21, 381. doi: 10.1146/annurev.nutr.21.1.381

Yao, Y., Cheng, X.-Z., Wang, L.-X., Wang, S.-H., and Ren, G. (2012). Major phenolic compounds, antioxidant capacity and antidiabetic potential of rice bean (*Vigna umbellata* l.) in China. *IJMS* 13, 2707–2716. doi: 10.3390/ijms13032707

Yao, D. N., Kouassi, K. N., Erba, D., Scazzina, F., Pellegrini, N., and Casiraghi, M. C. (2015). Nutritive evaluation of the bambara groundnut Ci12 landrace [*Vigna subterranea* (L.) verdc. (Fabaceae)] produced in côte d'Ivoire. *Int. J. Mol. Sci.* 16, 21428–21441. doi: 10.3390/ijms160921428

Yao, X., Zheng, X., Zhao, R., Li, Z., Shen, H., Li, T., et al. (2021). Quality formation of adzuki bean baked: From acrylamide to volatiles under microwave heating and drum roasting. *Foods* 10, 2762. doi: 10.3390/foods10112762

Zhang, Z., Lanza, E., Kris-Etherton, P. M., Colburn, N. H., Bagshaw, D., Rovine, M. J., et al. (2010). A high legume low glycemic index diet improves serum lipid profiles in men. *Lipids* 45 (9), 765–775. doi: 10.1007/s11745-010-3463-7

Zheng, Z., Li, J., Li, J., Sun, H., and Liu, Y. (2019). Physicochemical and antioxidative characteristics of black bean protein hydrolysates obtained from different enzymes. *Food Hydrocolloids* 97, 105222. doi: 10.1016/j.foodhyd.2019.105222

Zhong, J., Lu, P., Wu, H., Liu, Z., Sharifi-Rad, J., Setzer, W. N., et al. (2022). Current insights into phytochemistry, nutritional, and pharmacological properties of prosopis plants. evid-based complement. *Altern. Med.* 2022. doi: 10.1155/2022/2218029

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 02 June 2023
DOI 10.3389/fpls.2023.1164555

OPEN ACCESS

EDITED BY
David Okeh Igwe,
Cornell University, United States

REVIEWED BY
Oluwaseyi Samuel Olanrewaju,
North-West University, South Africa
Huan Liu,
Beijing Genomics Institute (BGI), China

*CORRESPONDENCE
Ingrid Pinheiro Machado
ingridpin@alu.ufc.br

RECEIVED 13 February 2023
ACCEPTED 10 May 2023
PUBLISHED 02 June 2023

CITATION
Machado IP, DoVale JC, Sabadin F and Fritsche-Neto R (2023) On the usefulness of mock genomes to define heterotic pools, testers, and hybrid predictions in orphan crops.
Front. Plant Sci. 14:1164555.
doi: 10.3389/fpls.2023.1164555

On the usefulness of mock genomes to define heterotic pools, testers, and hybrid predictions in orphan crops

Ingrid Pinheiro Machado[1*], Júlio César DoVale[1], Felipe Sabadin[2] and Roberto Fritsche-Neto[3]

[1]Department of Crop Science, Federal University of Ceará, Fortaleza, Brazil, [2]School of Plant and Environmental Sciences, Virginia Tech: Virginia Polytechnic Institute and State University, Blacksburg, VA, United States, [3]LSU AgCenter, Louisiana State University Agricultural Center, Baton Rouge, LA, United States

The advances in genomics in recent years have increased the accuracy and efficiency of breeding programs for many crops. Nevertheless, the adoption of genomic enhancement for several other crops essential in developing countries is still limited, especially for those that do not have a reference genome. These crops are more often called orphans. This is the first report to show how the results provided by different platforms, including the use of a simulated genome, called the mock genome, can generate in population structure and genetic diversity studies, especially when the intention is to use this information to support the formation of heterotic groups, choice of testers, and genomic prediction of single crosses. For that, we used a method to assemble a reference genome to perform the single-nucleotide polymorphism (SNP) calling without needing an external genome. Thus, we compared the analysis results using the mock genome with the standard approaches (array and genotyping-by-sequencing (GBS)). The results showed that the GBS-Mock presented similar results to the standard methods of genetic diversity studies, division of heterotic groups, the definition of testers, and genomic prediction. These results showed that a mock genome constructed from the population's intrinsic polymorphisms to perform the SNP calling is an effective alternative for conducting genomic studies of this nature in orphan crops, especially those that do not have a reference genome.

KEYWORDS

genotyping-by-sequencing, SNP-array, formation of heterotic groups, genomic prediction of single-crosses, minor crops, underused crops, simulated genome

1 Introduction

Molecular markers have been used to develop genomic tools to improve economically important crops (Mammadov et al., 2012; Thomson, 2014). Currently, single-nucleotide polymorphism (SNP) markers are the most used in genomic studies (Fritsche-Neto et al., 2021), as they provide higher resolution due to their frequent occurrence and uniformity

throughout the genome (Gupta et al., 2008). Rapid advances in next-generation sequencing (NGS) technologies, combined with high levels of diversity in SNP, have made it possible to develop high-throughput genotyping platforms (Bachlava et al., 2012).

There are several genotyping platforms for obtaining SNPs throughout the genome, which have provided an infinity of sequencing information with remarkable improvements in coverage, time, and costs, making it possible to genotype thousands of samples with many markers (Bevan and Uauy, 2013), with SNP array and NGS platforms being the most appropriate for this purpose (Rasheed et al., 2017). There are many array-based genotyping platforms available in major crops such as maize (Unterseer et al., 2014), wheat (Winfield et al., 2016), rice (Singh et al., 2015), and soybean (Lee et al., 2015). These platforms have many advantages, such as fast scans with high call rates and density. However, they present an investigation bias when the set of individuals does not faithfully represent the genetic diversity explored in the study panel. Furthermore, it has a high cost and is inaccessible to small breeding programs (Messing and Dooner, 2006; Frascaroli et al., 2013), especially those of unprofitable crops.

Beyond crop-specific SNP arrays, NGS-based platforms are adaptable to various crops, regardless of prior knowledge of genomics, genome size, organization, or ploidy (Rasheed et al., 2017). Genotyping-by-sequencing (GBS) appears as an alternative to overcome the verification bias since it is based on sequencing and, therefore, allows the discovery of alleles in the diversity panel analyzed, in addition to having a lower cost when compared to SNP array. However, GBS generates many low-quality markers with a high rate of lost data (Heslot et al., 2013). The advances in genomics in recent years have increased the accuracy and efficiency of breeding programs for many crops. However, adopting genomic enhancement for several other staple crops essential in developing countries is still limited, especially for traits under complex genetic control, which are crucial to crop performance (Varshney et al., 2012). This is because most studies use the array-based and GBS-based SNP marker approach, which depends on a reference genome for SNP discovery. Crops that do not have a reference genome cannot take advantage of biotechnological tools to improve their genetic gain and develop modern cultivars faster (Armstead et al., 2009).

There are many crops of unique relevance to developing countries, essential for the food, nutritional, and economic security of these countries, which still do not have a reference genome (Baldermann et al., 2016; Hendre et al., 2019). These crops are more often called orphans. The term "orphan" is derived from the condition of neglect and helplessness of these crops by the scientific community, leading to the designation of such species as underused, neglected, or minor crops (Tadele and Assefa, 2012). GBS also appears as an option for genomic studies in these crops, especially when they do not have a reference genome (Sabadin et al., 2022). With these data, it is possible to build a mock reference to perform the SNP calling, where the discovery of polymorphisms will be intrinsic to the study population without using an external genome (Melo et al., 2016). This pipeline has already been successfully used in several genomic studies (Adhikari et al., 2018; Holloway et al., 2018; Munjal et al., 2018; Matias et al., 2019; Sabadin et al., 2022). Adopting this technology in poorly studied crops has a tangible impact on the progress of the breeding process (Ye and Fan, 2021).

Recent studies have compared the performance of genotyping platforms and how this choice affects genomic studies regarding genetic diversity studies (Elbasyoni et al., 2018; Darrier et al., 2019), genome-wide association study (GWAS) (Negro et al., 2019), and genomic prediction (Chu et al., 2020). Only one report compares the performance of the mock reference pipeline with standard genotyping approaches in genomic prediction studies (Sabadin et al., 2022). However, this study used a relatively small germplasm panel. It did not consider the effect of genotyping platforms on population structure, the formation of heterotic groups, and the choice of testers, which is fundamental for a breeding program that wants to synthesize single crosses.

Although some reports are available, there still needs to be a consensus on how different platforms provide the results. Thus, we conducted a full study on this topic with a robust germplasm panel. For this, we compared different genotyping scenarios from the beginning of the breeding process with the approach of genetic diversity and population structure, advancing to the formation of heterotic groups and choice of testers to the synthesis and prediction of single crosses. Thus, this information will be valuable to leverage genomic studies and accelerate the development of cultivars in minor crops without a reference genome. Therefore, our goals were to verify whether the source of SNP can influence the assessment of the population structure of parental lines, to ascertain if the source of SNP can affect the determination of heterotic groups and the prediction of single crosses performance, and check if the GBS and the mock genome efficiently performs the SNP calling in orphan crops (without reference genome).

2 Materials and methods

To facilitate the understanding of the analysis carried out in this study, a workflow is described in Figure 1. Each stage of the analysis is detailed in the following sections.

2.1 Species model

We used maize as a model species in this study because it is already well-established regarding SNP array, with several array options available (Ganal et al., 2011; Unterseer et al., 2014; Xu et al., 2017), and GBS protocols are also well-established for this species (Crossa et al., 2013; Li et al., 2015; Wang et al., 2020). In addition, maize has a diverse, complex, and dynamic genome (Schnable et al., 2009), which is suitable for this study. We used a public panel of tropical maize germplasm containing 360 parental lines (Yassue et al., 2021a). The genomic and phenotypic information about this panel is available on the Mendeley platform (https://data.mendeley.com/datasets/5rtc89t7v5/1).

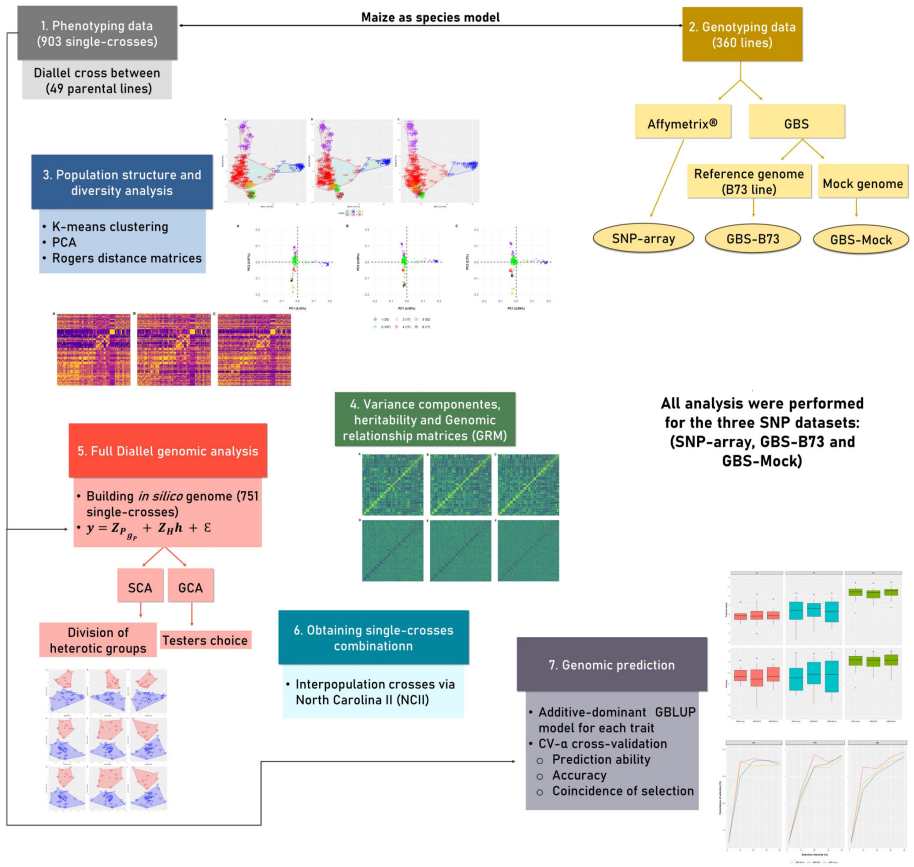


FIGURE 1
The workflow of analysis performed in the study. The different colors represent different stages of the analysis. Genomic information from the three SNP datasets was used in all analyses. SNP, single-nucleotide polymorphism.

2.2 Phenotypic data

The phenotypic dataset consists of 903 maize single crosses (Fritsche-Neto et al., 2019) derived from a diallel cross between 49 parental lines to a public tropical maize diversity, selected based on nitrogen use efficiency (Mendonça et al., 2017). Field trials were carried out in Anhembi (22°50′51″S, 48°01′06″W) and Piracicaba (22°42′23″S, 47°38′12″W), in the State of São Paulo, during the second growing season, from January to June 2016 and 2017. Single crosses were evaluated in an augmented block design, where each block consisted of 16 single crosses and two checks (commercial single crosses). In both locations and years, the single crosses were evaluated under two nitrogen (N) conditions, low N with 30 kg N ha^{-1} and ideal N with 100 kg N ha^{-1}. Each location × year × N level combination was defined as an environment.

Each plot consisted of 7 m rows spaced 0.50 m apart. Conventional fertilization and weed and pest control were carried out. The traits evaluated were grain yield (GY, mg ha^{-1}), plant height (PH, cm), and ear height (EH, cm). The plots were harvested manually, and the grains were harvested with a moisture content of approximately 18%. Subsequently, grain yield was corrected for 13% moisture, according to Mulvaney and Devkota (2020). More details on the phenotypic dataset's experimental design and cultivation practices were previously reported by Fritsche-Neto et al. (2018) and Galli et al. (2020).

2.3 Genetic-statistical model for obtaining BLUEs

The joint analysis of each trait was performed to estimate the means of the single crosses across the environments. Thus, an equation was adjusted to obtain the Best Linear Unbiased Estimator (BLUE) for each genotype, and later, the adjusted means of these across the environments evaluated by the following mixed model were estimated:

$$\boldsymbol{y}=\boldsymbol{Ql}+\boldsymbol{Sb}+\boldsymbol{Tc}+\boldsymbol{Ug}+\boldsymbol{Vi}+\boldsymbol{\epsilon},$$

where $\boldsymbol{y}$ is the vector of phenotypic values of single crosses and checks; $\boldsymbol{l}$ is the vector of fixed effects of the environment (site × year × N level combination); b is the vector of random effects of block nested within environments, where $b \sim N(0, I\sigma^2_b)$; c is the vector of fixed effects of checks; g is the vector of fixed effects of single crosses; i is the vector of fixed effects of the interaction checks × environments; ε is the vector of random residual effects, where $\epsilon \sim N(0, De)$. An unstructured covariance matrix across

environments was assumed for the residual term (*De*) due to the contrasting doses of N. *Q*, *S*, *T*, *U*, and *V* are the incidence matrices for *l*, *b*, *c*, *g*, and *i*. The analysis was performed using the *ASReml-R* (Butler et al., 2018).

2.4 Genotypic data and analysis

The 360 lines belonging to the public tropical maize diversity mentioned above were genotyped using two SNP high-density genotyping platforms: 1) Affymetrix® Axiom Maize Genotyping Array (SNP array) and 2) GBS method following the protocol described by Poland et al. (2012). In this last method, individual samples of genomic DNA were digested by two restriction enzymes, *Pst*I and *Mse*I, to reduce the genome complexity uniformly. Subsequently, the samples were included in a sequencing plate, performed on the Illumina NextSeq 500 platform (Illumina Inc., San Diego, CA, USA).

The raw GBS data were used for two purposes: the first was to perform the SNP calling using the B73 line of temperate germplasm as a reference genome. The second purpose was to build a simulated reference genome (mock genome) according to the GBS-SNP-CROP pipeline proposed by Melo et al. (2016) and use it to perform the SNP calling. This pipeline aggregates custom analysis and filtering procedures with bioinformatics tools on raw GBS readings. The method employs a variant calling strategy based on patterns of polymorphisms within the individual or cluster and across populations or clusters to identify sequencing or PCR errors. Finally, the pipeline uses a reading grouping strategy based on similarity to generate representative sequences, that is, a simulated reference of GBS fragments. Details of each stage of mock genome building can be found in Melo et al. (2016).

Further analysis was performed considering three SNP datasets: 1) SNP array, 2) GBS with SNP call using B73 as the reference genome (GBS-B73), and 3) GBS with the simulated genome being used as the reference genome (GBS-Mock). For GBS datasets, according to standard parameters, SNPs were scored from raw data using the TASSEL 5.0 GBSv2 pipeline (Glaubitz et al., 2014). With the use of the Burrows-Wheeler Alignment tool (BWA) (Li and Durbin, 2009), the tags were aligned against the reference genome (GBS-B73 and GBS-Mock).

As two genotyping platforms (SNP array and GBS) were performed, the parental lines that showed a very contrasting genotypic profile between the two platforms were removed from the analysis to obtain a fair comparison. Thus, between sequencing errors and divergences in genotypic profiles between platforms, 330 parental lines remained, among which 45 parental lines make up the diallel, which generated 751 single crosses. The number of markers concerning the raw data was 18,413 (SNP array), 131,350 (GBS-B73), and 46,9126 (GBS-Mock). All SNP sets underwent quality control, in which low call rates (<90%) and non-biallelic markers were removed from the datasets. The remaining missing data were imputed by the Beagle 5.0 algorithm (Browning et al., 2018). Pairwise linkage disequilibrium was calculated as the correlation of allele frequencies squared (r^2), and values greater than 0.99 were removed from the datasets using the *SNPRelate* package (Zheng et al., 2012), resulting in 12,704 (SNP array), 11,153 (GBS-B73, and 4,935 (GBS-Mock) SNP markers.

Subsequently, new quality control was performed, in which heterozygous loci in at least one individual were removed. High-quality polymorphic SNPs from the parental lines were combined (*in silico*) to build an artificial single-cross genomic matrix. In addition, duplicate markers between chromosomes were removed to avoid overparameterization caused by multicollinearity. Finally, markers with minor allele frequency (MAF)<0.05 were removed from the single-cross genomic matrices, resulting in 11,884 (SNP array), 10,361 (GBS-B73), and 4,801 (GBS-Mock) SNP markers to perform the remaining analysis.

2.5 Analysis of population structure and genetic diversity

The three SNP datasets (SNP array, GBS-B73, and GBS-Mock) from the 330 parental lines were used to assess the population structure of the panel. In these analyses, in particular, heterozygous loci and rare variants (MAF< 0.05) were considered to capture all diversity and variability to perform principal component analysis (PCA) and determine the relatedness between parental lines.

K-means clustering was applied, using the total within-cluster sum of square (WSS) method to determine the optimal number of clusters so that the total intra-cluster variation is minimized (Kassambara, 2017). The factoextra package (Kassambara and Mundt, 2020) was used for this. Subsequently, Kendall's method determined the coincidence in forming clusters among the different datasets (Kendall, 1938). Kendall's tau correlation coefficient was tested at a probability level of 0.01. PCA was performed, and biplots were constructed to assess population structure.

The genetic distances between the parental lines were calculated for each SNP dataset using the Rogers distance (Rogers, 1972). Subsequently, to measure the correlation among the kinship matrices, the Mantel correlation test (Mantel, 1967) was applied to detect significance. The Mantel correlation test is non-parametric and computes the significance of the correlation similarity measures using 1,000 permutations of the rows and columns of one distance matrix. The heatmaps of the genetic distance matrices were obtained using the *superheat R* package (Barter and Yu, 2018). Correlations were obtained using the *vegan* package (Oksanen et al., 2019), and each analysis was performed for each SNP dataset scenario.

2.6 Full diallel genomic analysis

To find out how diversity and population structure can influence the formation of heterotic groups, it was necessary to construct *in silico* genome of the 751 single crosses from parental lines. Therefore, at this stage, we combined phenotypic and genotypic information from these individuals to estimate general (GCA) and specific combining abilities (SCA). For this, the following diallel model was adjusted:

$$\boldsymbol{y}=\boldsymbol{Z}_{\boldsymbol{P}}\boldsymbol{g}_{\boldsymbol{P}}+\boldsymbol{Z}_{\boldsymbol{H}}\mathbf{h}+\boldsymbol{\epsilon},$$

where $\boldsymbol{y}$ is the adjusted phenotypic data vector of the single crosses for the trait, g_P is the random effects vector of the GCA captured by the markers of the parental lines, and h is the random effects vector of the SCA that denotes the interaction effects across the parental lines. Z_P and Z_H are incidence matrices that relate y to g_P and h to $g_P \sim N(0, \sigma^2_P G_P)$ and $h \sim N(0, \sigma^2_H H)$, where σ^2_P and σ^2_H are variance components associated with GCA and SCA, respectively. G_P and H are relationship matrices for the parental lines and single crosses, respectively. Finally, $\varepsilon \sim N(0, \sigma^2_\varepsilon I)$, where σ^2_ε is the variance associated with the residuals.

The G_P relationship matrix was calculated using the SNP markers according to (VanRaden, 2008), where W_P is the matrix of centered and patterned markers. Therefore, $G_P = \frac{W_P W_P'}{p}$ (Technow et al., 2014; Lopez-Cruz et al., 2015), where p is the number of markers. This resulted in an average diagonal G_P value of ~1; therefore, σ^2_P was defined on the same scale as σ^2_ε. The elements of the H matrix were obtained directly from the G_P (Bernardo, 2002; Technow et al., 2014). The matrix H for all possible crosses was obtained with the Kronecker product between G_P's, $H = G_P \otimes G_P$ (Covarrubias-Pazaran, 2016). A model was built with their respective kernels for each SNP marker source. Analyses were performed using the *ASReml-R* package (Butler et al., 2018).

2.7 Heterotic groups and testers

The determination of heterotic groups was performed based on SCA estimates for each trait. These estimates corresponded to a matrix of genetic distances. According to Falconer and Mackay (1996), the genetic distance between parents positively affects heterosis. This association depends on dominance effects or differences in the frequency of the alleles that control the trait considered (Falconer, 1960). Burstin et al. (1994) also found that SCA variance is an indicator for predicting hybrid performance by genetic distance between parents. According to this information, it was assumed that the higher the SCA estimates, the greater the distance between the parents and the more significant the heterosis. From this, the 45 lines were divided into two heterotic groups.

The SCA estimates were submitted to a clustering algorithm, K-means, which grouped them according to the SCA estimates. To estimate the correlation between the heterotic groups formed for the different genotyping methods, Pearson's correlation was applied and tested at a probability level of 0.01 by Student's t-test. Subsequently, the identification of the best tester in each group was performed according to the GCA estimates. The best tester of a given group was the line that showed the highest GCA with the other group. Based on this, the coincidence of testers between the scenarios was evaluated.

2.8 Obtaining single-cross combinations and genomic prediction

After the parental lines were divided into heterotic groups, only the single crosses corresponding to interpopulation crosses *via* North Carolina II (NCII) design were considered for the following analyses. The number of single crosses changed according to the configuration of heterotic groups for each trait in the three SNP scenarios.

For the genomic prediction of the single crosses, an additive–dominance genomic best linear unbiased prediction (GBLUP) model was used, as described below:

$$\hat{\boldsymbol{y}}=\boldsymbol{1\mu}+\boldsymbol{Za}+\boldsymbol{Zd}+\boldsymbol{\epsilon},$$

where $\hat{y}$ is the adjusted means vector of the single crosses for the trait; μ is the mean (intercept); a is the vector of additive genetic effects of individuals, where $a \sim N(0, G_a\sigma^2_a)$; d is the vector of dominance effects, where $d \sim N(0, G_d\sigma^2_d)$; ε is the random effects vector of the residuals, where $\varepsilon \sim N(0, I\,\sigma^2_\varepsilon)$. Z is the incidence matrix for a and d. σ^2_a is the additive genomic variance, σ^2_d is the dominance genomic variance, and σ^2_ε is the residual variance. G_a and G_d are the additive and dominance genomic relationship matrices, respectively, of the single crosses, where $G_a = \frac{W_A W_A'}{2\sum_{i=1}^{n} p_i(1-p_i)}$ and $G_d = \frac{W_D W_D'}{4\sum_{i=1}^{n}(p_i(1-p_i))^2}$, where p_i is the frequency of an allele at locus i and W is the matrix incidence of markers (VanRaden, 2008). The W_A matrix was encoded as 0 for A_1A_1, 1 for A_1A_2 heterozygote, and 2 for A_2A_2 homozygote. For W_D, genotypes were coded as 0 for both homozygotes and 1 for the heterozygote. The genomic relationship matrices were built using the *snpReady* package (Granato et al., 2018). The genomic prediction models were performed using the *sommer* package (Covarrubias-Pazaran, 2016). It is worth noting that all three sets of markers were used to build the kernels. The Mantel correlation test (Mantel, 1967) was applied to detect the significance between the additive and dominance genomic relationship matrices.

To evaluate the model performance, we used the CV-α cross-validation with five folds and four replicates (Yassue et al., 2021b). The predictive ability was estimated by Pearson's correlation between predicted genotypic and observed values from the validation set. The prediction accuracy was estimated by the correlation between predictive ability and heritability, according to Mrode (2014). Correspondence between phenotypic and genotypic selection was calculated for each set of markers through the percentage of common genotypes selected by their adjusted means from the phenotypic analysis and their genomic estimated breeding values (GEBVs) from the genomic prediction model concerning different intensities of selection (1%, 10%, 20%, 30%, and 40%). The heritability in the broad-sense (H^2) and the narrow-sense (h^2) was also estimated by the equations below:

$$H^2 = \frac{\hat{\sigma}^2_a + \hat{\sigma}_d{}^2}{(\hat{\sigma}^2_a + \hat{\sigma}^2_d + \hat{\sigma}^2_\epsilon)} \text{ and } h^2 = \frac{\hat{\sigma}^2_a}{(\hat{\sigma}^2_a + \hat{\sigma}^2_d + \hat{\sigma}^2_\epsilon)},$$

where $\hat{\sigma}^2_a$ is the additive genetic variance, $\hat{\sigma}^2_d$ is the dominance genetic variance, and $\hat{\sigma}^2_\epsilon$ is the residual variance.

3 Results

3.1 Genetic diversity and population structure

According to the WSS method, for all datasets, the optimal number of clusters among the 330 parental lines that minimized

within-group variance and maximized between-group variance was six (Figure S1). Subsequently, the K-means clustering method showed a remarkable similarity in the arrangement of clusters among datasets (Figure 2). This similarity is confirmed by the coincidence values in the clustering (Table S1), with correlation coefficients above 0.95.

Concerning PCA, the SNP datasets showed similar performances regarding the variance explained by the principal components. The first principal components hold the highest percentage of explained variance (Figure S2A). When considering the first 10 main components, SNP array showed the highest value of cumulative explained variance (27.3%). At the same time, GBS-B73 and GBS-Mock presented discounts of 24.1% and 16.8%, respectively (Figure S2B).

In general, PCA revealed that the first eigenvectors exhibited similar patterns of variance in all combinations between datasets, supported by the coefficient of determination (R^2). However, the other eigenvectors showed less similarity between the captured variance patterns (Figure 3). The first four eigenvectors of SNP array and GBS-B73 showed high values of R^2 (Figure 3A). In contrast, for SNP array and GBS-Mock, the highest values of R^2 were concentrated in the first three eigenvectors (Figure 3B). For GBS-B73 and GBS-Mock, all eigenvectors showed high magnitude R^2, with the former being slightly higher than the others (Figure 3C).

Biplots were constructed to visualize the spatial distribution of lines in all SNP datasets (Figure 4; S3, S4). For this, the first three PCs were used, together with the information obtained by the K-means clustering method (Figure 2). All datasets showed the same pattern of dispersion among the lines, in agreement with the cluster analysis, which suggests that the SNP datasets capture similar patterns of variance (Figure 4).

Rogers distance matrices (GD) from all SNP datasets sampled similar groups and subgroups, with slight differences between them (Figure S5). Regarding the Mantel correlations between the GDs, high magnitude correlations (>0.83) were observed involving different scenarios (Table 1).

3.2 Variance components, genomic heritability, and genomic relationship matrices

Broad- and narrow-sense heritabilities were higher for EH, followed by PH and GY (Table S2). GY showed broad-sense heritability for all SNP datasets, on average, 36% higher than narrow-sense heritability. This difference is significantly smaller for the other traits, 15% and 6%, for PH and EH, respectively. The narrow-sense heritability for all SNP datasets was practically the same for GY. As for PH, there was a slight difference in SNP array, and GBS-Mock presented heritability slightly higher than that of GBS-B73. For EH traits, narrow-sense heritability varied little among SNP datasets, with GBS-Mock and SNP array showing the highest heritabilities. The heritabilities in the broad-sense (H^2) followed the same tendency.

Regarding the additive genomic relationship matrices (*Ga*) across the single-crosses, SNP array, GBS-B73, and GBS-Mock showed high Mantel correlations (Table 1; Figures S6A-C). However, the genomic dominance relationship matrices (*Gd*) showed lower correlations than *Ga*. The correlations between the dominance relationship matrices (*Gd*) were lower but still from medium to high. GBS-Mock stands out with a correlation of 0.72 with GBS-B73.

3.3 Heterotic groups and testers

Based on SCA estimates, the 45 parental lines were divided into heterotic groups as the genetic distance between them for the evaluated

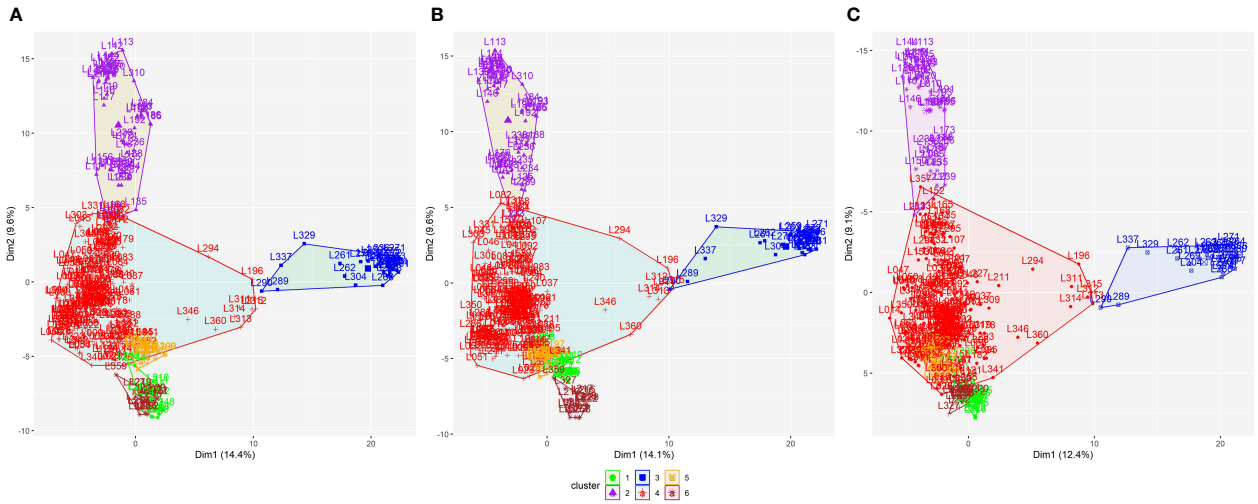


FIGURE 2
K-means clustering method for dividing the 330 parental lines into subpopulations applied to all SNP datasets: **(A)** SNP array, **(B)** GBS-B73, and **(C)** GBS-Mock. SNP, single-nucleotide polymorphism.

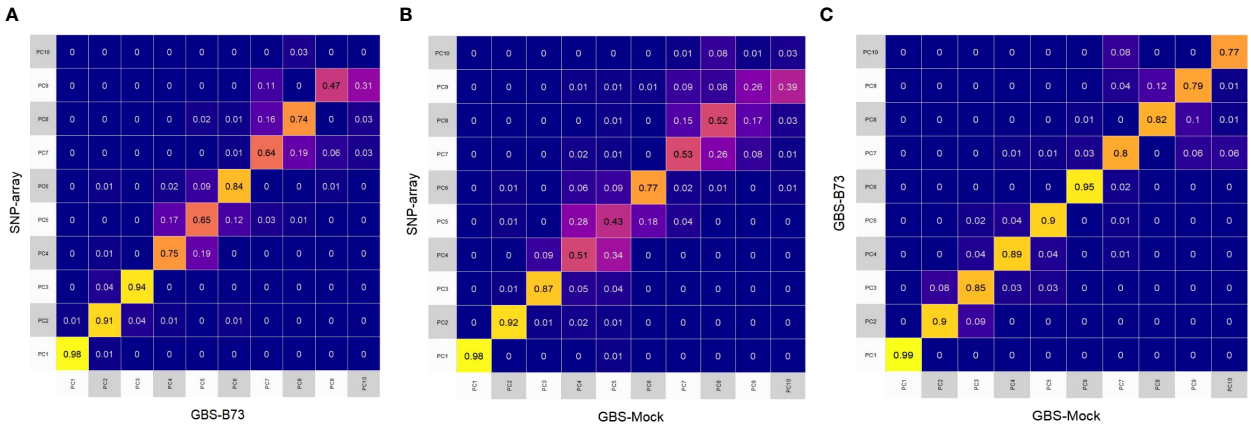


FIGURE 3
Heatmap of the coefficient of determination (R^2) of the 10 first eigenvectors built from the Rogers distance among all SNP datasets: **(A)** SNP array and GBS-B73, **(B)** SNP array and GBS-Mock, and **(C)** GBS-B73 and GBS-Mock. SNP, single-nucleotide polymorphism.

traits, GY, PH, and EH. Accordingly, two heterotic groups were formed for all SNP datasets (Figure 5). The formation of heterotic groups among the SNP datasets was quite similar, with high correlations, higher than 0.94 for GY and 0.87 for PH and EH (Table S3). There was, at most, a change in the allocation of two parental lines between heterotic groups in different SNP datasets. Likewise, the SCA correlations of the parental lines among the SNP datasets were higher than 0.96 (Table S4).

GCA estimates from each parental line, trait, and SNP dataset were used to choose the best tester in each group (Table 2). Thus, the testers matched among SNP datasets in the respective heterotic groups for each trait. Based on GY, the tester chosen for heterotic group one (HP_1) was L023, and for heterotic group two (HP_2), it was L006. As for PH and EH, L001 was elected as the HP_1 tester and L003 as the HP_2 tester. The correlations between the GCAs confirm this result, with maximum correlations (Table S4).

3.4 Genomic prediction

The predictive ability estimated by the additive–dominance model for all traits and following the same trend as the other results did not vary significantly among SNP datasets. The mean values of PA were 0.58 for GY, 0.64 for PH, and 0.83 for EH. Prediction accuracy also did not vary significantly between SNP datasets. It showed a high magnitude for all traits, with a mean value of 76% for GY, 78% for PH, and 90% for EH (Figure 6). The coincidence between selected individuals based on the adjusted means of the phenotypic analysis and the GEBVs of the genomic prediction model was generally satisfactory. It increased with rising selection intensity (Figure 7). Although GY is considered the most complex, the selection coincidence levels of this one were similar to the other traits. SNP array showed slightly higher coincidence values for almost all selection intensities. However, the different datasets showed approximate coincident values.

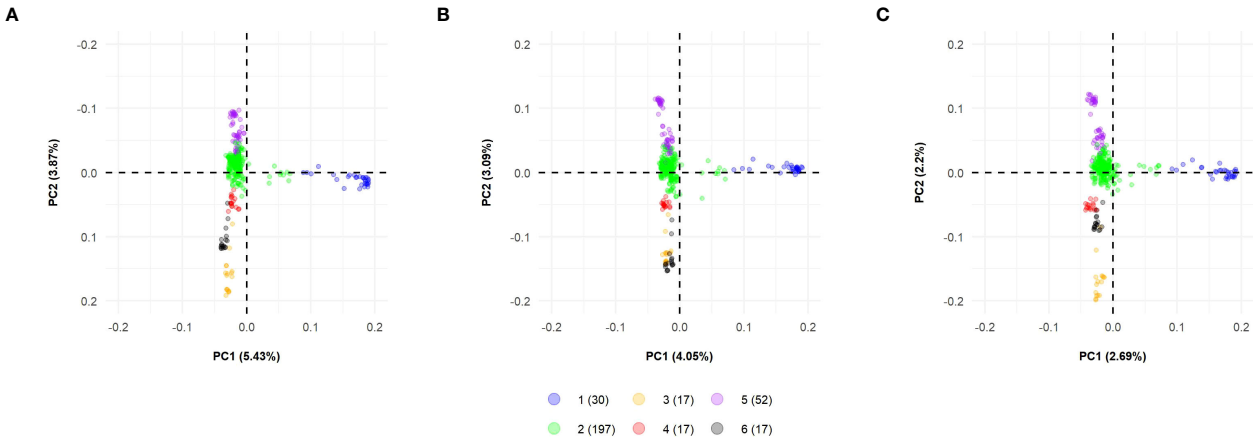


FIGURE 4
Biplot among two first principal components using all datasets for 330 tropical parental lines: **(A)** SNP array, **(B)** GBS-B73, and **(C)** GBS-Mock. Explained variance percentages of each principal component are in parentheses. Clusters were used to color-code parental lines. SNP, single-nucleotide polymorphism.

TABLE 1 Mantel correlation of Rogers genetic distance (GD) matrices for 330 parental lines and of additive genomic relationship (*Ga*) and dominance genomic relationship (*Gd*) matrices for 751 maize single crosses estimated from SNP array, GBS-B73, and GBS-Mock markers.

		GBS-B73	GBS-Mock
GD	SNP array	0.91**	0.83**
	GBS-B73	–	0.91**
G_a	SNP array	0.97**	0.96**
	GBS-B73	–	0.99**
G_d	SNP array	0.78**	0.58**
	GBS-B73	–	0.72**

Rogers genetic distance (GD) matrices were computed with markers from 330 parental lines data. *Ga* and *Gd* matrices were computed with markers from 751 maize singles-crosses.
SNP array, Affymetrix® Axiom Maize Genotyping array; GBS-B73, genotyping-by-sequence with SNP calling using B73 as reference genome; GBS-Mock, genotyping-by-sequence with SNP calling using the mock reference built with all parental lines.
SNP, single-nucleotide polymorphism.
**Empirical significance level from permutations.
The symbol "-" means that the correlation of a value with itself is maximum, that is "1".

4 Discussion

Recent crop genetics and genomics advances have gained remarkable attention and offered genotyping technologies (Chakradhar et al., 2017). Various genotyping platforms are available to meet the most diverse needs regarding costs per sample and different marker densities (Thomson, 2014). GBS, in particular, has emerged as a cost-effective strategy for genome-wide SNP discovery and population genotyping due to the simple library preparation and the robust approach to genome reduction (Elshire et al., 2011).

All this progress is focused on a small group of crops (Tester and Langridge, 2010) to the detriment of smaller agricultural species, considered orphans, historically poorly researched (Mayes et al., 2012), in that the large majority do not have a reference genome. Sabadin et al. (2022) showed that using mock genomes could be a worthy strategy that permits using SNP markers for genomic selection in orphan crops. However, orphan crop breeding programs focused on single-cross development must also determine heterotic groups to maximize the heterosis. Our study aims to go forward and verify the usefulness of mock genomes as a method to permit reliable heterotic group clustering.

4.1 Influence of genotyping methods on population structure and diversity

The study of the characterization of genetic diversity, population structure, and genetic relationships among elite germplasm parents, based on molecular markers, can accelerate genetic gains in breeding programs (Romay et al., 2013; Adu et al., 2019). This study helps understand how the germplasm is organized in selecting parents that present effective contributions and in the designation of heterotic groups (Wu et al., 2016). Thus, genomic data not only allow the estimation of genetic diversity but also combine them with phenotypic information to find new functional genes and build prediction models (Milner et al., 2019). However, in this topic, the focus is on whether, with the simulated reference genome, there is the discovery of the same polymorphisms and how it reflects on the population structure of the lines.

The WSS method indicated the optimal number of clusters by locating a curve on the plot, which is generally considered an indicator of the optimal number of groups (Kassambara, 2017). With this information and the results of the K-means clustering, the parental lines were partitioned into subpopulations, where the SNP datasets showed similar behavior (Figure S1; Figure 2; Table S1), in agreement with the spatial distributions obtained in the biplot graphs (Figure 4), in which all SNP datasets showed the same dispersion pattern between lines. This suggests that the SNP datasets capture similar patterns of variance, despite the difference in the number of markers between them, where GBS-Mock has a lower number and the difference in the genotyping platform itself (array and GBS). Thus, SNP array, GBS-B73, and GBS-Mock revealed similar performances concerning genetic diversity and the population structure of parental lines. Darrier et al. (2019) compared the performance of SNP array and GBS to investigate the extent and pattern of genetic variance in barley and observed that the two methodologies selectively access the informative polymorphism in different portions of the genome. Despite this, their results showed a strong positive correlation between the matrices of both genotyping approaches, supporting their similarity and validity.

PCA shows that these variance patterns captured by the SNP datasets are more similar concerning the first eigenvectors (Figure 3). However, the captured variance is more consistent when comparing GBS-B73 and GBS-Mock (Figure 3C). This can be explained by the verification bias existing in SNP array since this bias arises when the markers are not obtained from a random sample of the polymorphisms of the population of interest since the matrix is constructed using temperate maize lines (Frascaroli et al., 2013; Heslot et al., 2013; Unterseer et al., 2014), and the lines in the study are from tropical germplasm.

The matrices of genetic distances among the parental lines revealed similar patterns, showing the formation of subpopulations between the lines (Figure S5). When using wheat as a model species to test for verification bias and investigate its impact on genetic diversity estimates, Chu et al. (2020) observed a tendency for SNP array, leading to an underestimation of molecular diversity within the population. These results agree with a previous study on wheat lines (Elbasyoni et al., 2018) and maize lines (Frascaroli et al., 2013). Despite the verification bias mentioned above and the difference between the reference genome used, the temperate B73 genome, or the mock genome, the population structure between the lines did not show a significant difference, as the correlations between the matrices of genetic distances were of high magnitude. Even though GBS-Mock uses a different reference genome from SNP array and GBS-B73, their correlation was high (Table 1). Elbasyoni et al. (2018), investigating the influence of SNPs from different genotyping platforms on genomic prediction, observed a high correlation (r = 0.77) between SNP array and GBS genetic distance matrices. These high-magnitude correlations suggest that the broad sampling of diversity is well represented by the approaches used in the study. This is supported by the GWAS by

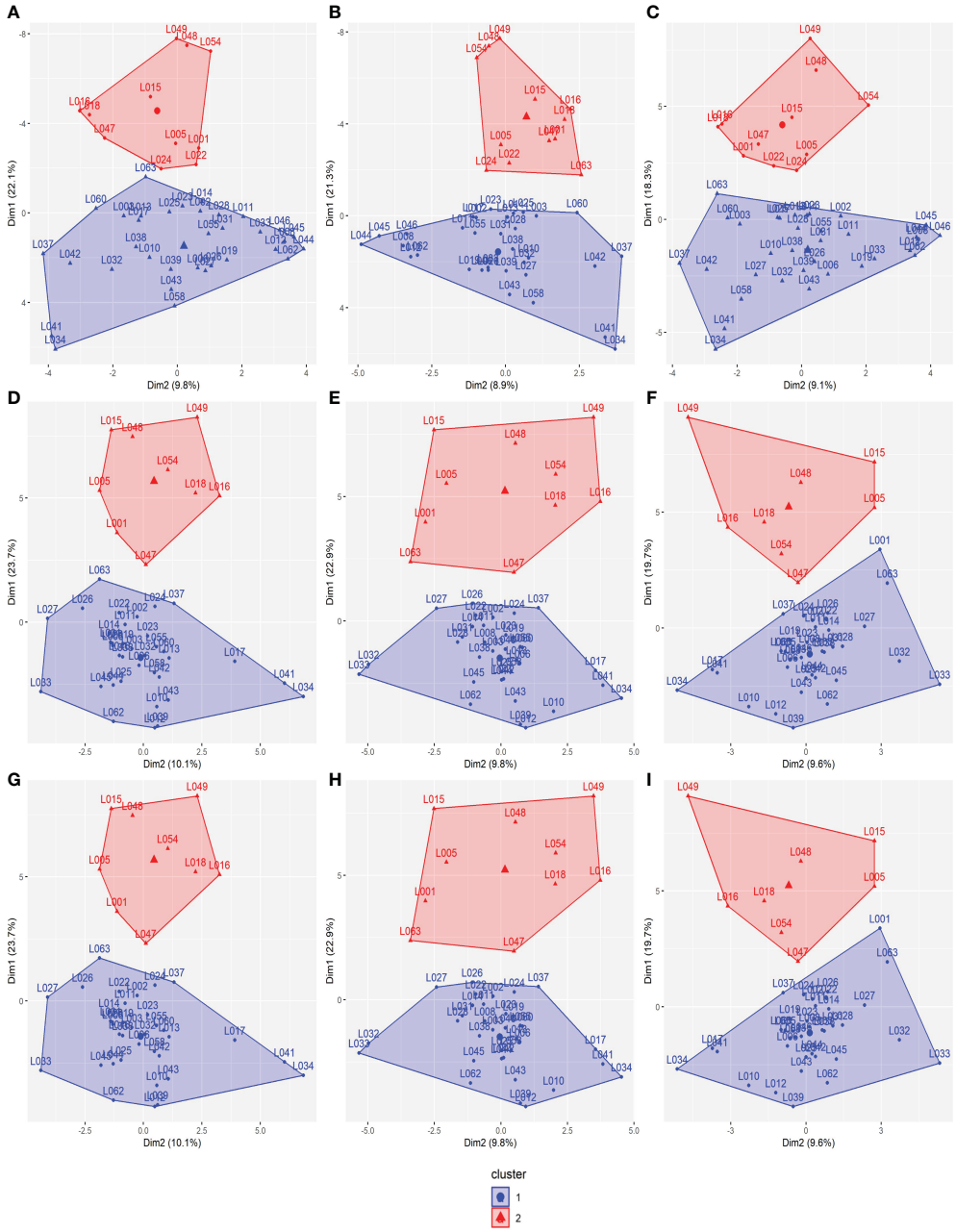


FIGURE 5
Heterotic groups among the 45 tropical parental lines for all traits: **(A)** SNP array (GY), **(B)** GBS-B73 (GY), **(C)** GBS-Mock (GY), **(D)** SNP array (PH), **(E)** GBS-B73 (PH), **(F)** GBS-Mock (PH), **(G)** SNP array (EH), **(H)** GBS-B73 (EH), and **(I)** GBS-Mock (EH). GY, grain yield; PH, plant height: EH, ear height; SNP, single-nucleotide polymorphism.

Darrier et al. (2019). They indicated that SNP array and GBS methods could detect markers closely associated with genes that control key phenotypic traits.

4.2 Influence of genotyping methods in the determination of heterotic groups in the choice of testers

Heterosis is a fundamental phenomenon in obtaining superior single crosses. Establishing heterotic groups to exploit them effectively throughout the breeding cycles is necessary. These, in turn, are made up of genetically related parental lines, which generate little or no heterosis when crossed with each other. Crossing with lines from another heterotic group tends to result in vigorous single crosses (Lee, 1995). Therefore, genetic diversity among heterotic groups tends to increase the level of heterosis detected in hybrid combinations (Falconer and Mackay, 1996; Fu et al., 2014). Badu-Apraku et al. (2011) reported in their diallel study between maize lines that their genetic diversity was small, and because of this, distinct heterotic groups could not be identified. Significant genetic diversity was found in a similar study with other

10.3389/fpls.2023.1164555

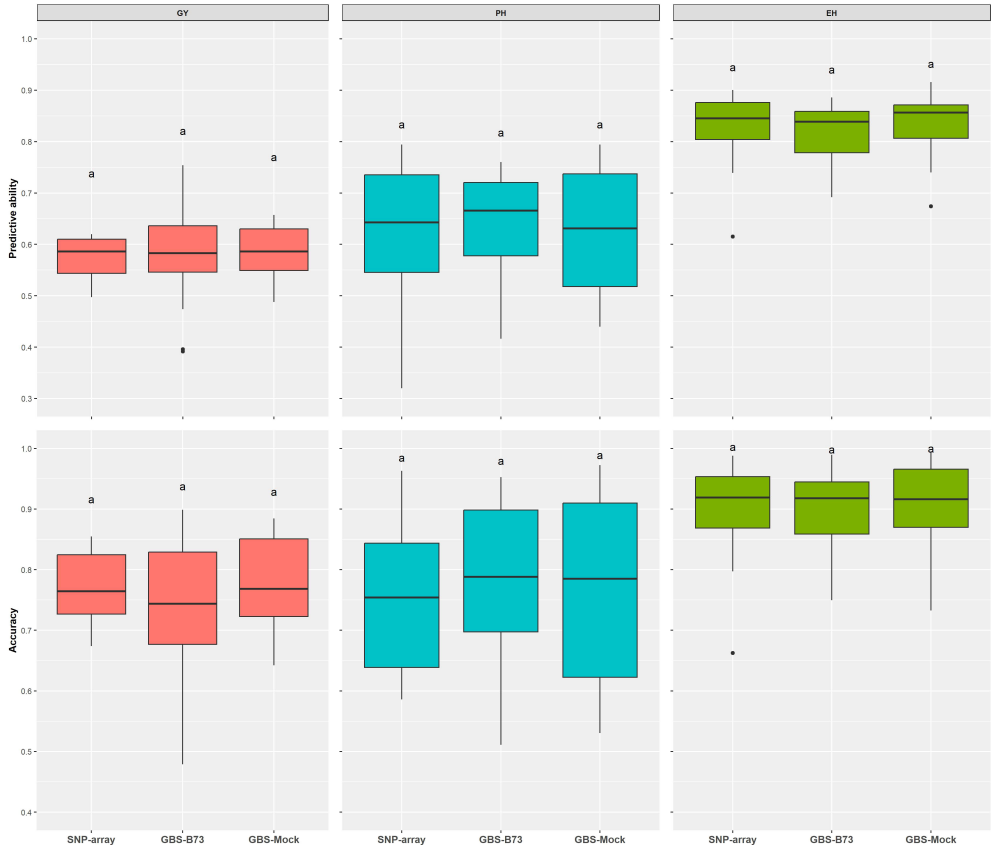


FIGURE 6
Predictive ability via additive–dominance GBLUP model estimated by Pearson's correlation between predicted and observed genotypic values of the validation set for all SNP datasets (SNP array, GBS-B73, and GBS-Mock). GBLUP, genomic best linear unbiased prediction; SNP, single-nucleotide polymorphism. Equal letters indicate no significant differences between groups (Tukey's post hoc test, $P < 0.05$).

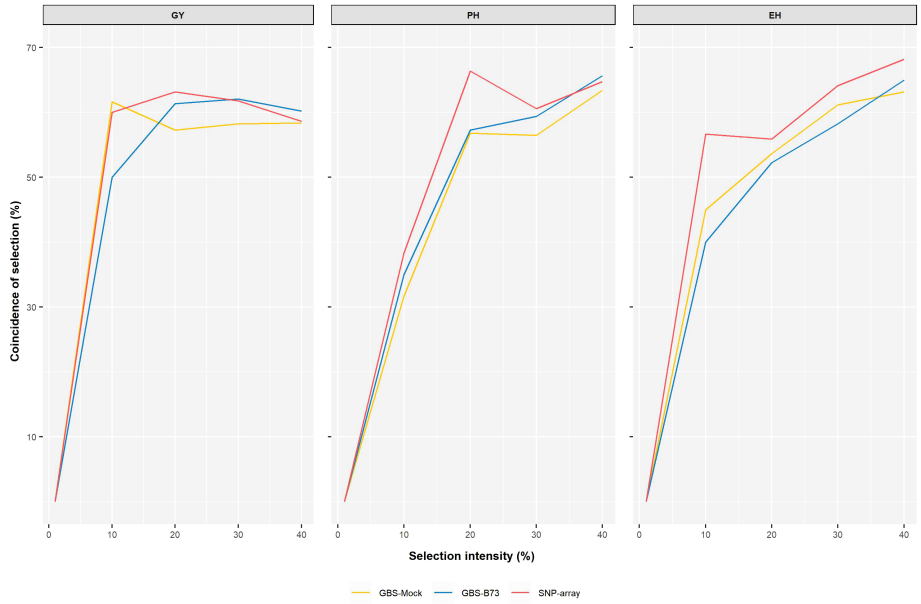


FIGURE 7
Coincidence between phenotypic and genotypic selection for each set of markers from the genomic prediction model concerning different selection intensities. The coincidence of selection percentage (y-axis) under a series of continuous selection intensities (1%–40%) (x-axis).

maize lines, and two clear heterotic groups were identified. The type of predominant gene action in the parents under investigation is another factor that affects heterotic clustering. When additive and non-additive effects are significant, and there is a predominance of additive gene action over non-additive gene action, heterotic groups are easily identified (Badu-Apraku et al., 2015; Badu-Apraku et al., 2016a; Badu-Apraku et al., 2016b).

The PH and EH traits showed higher proportions of additive variance captured by the *Ga* matrices than GY (Table S2). Although these traits have polygenic inheritance, GY is the most complex trait and most influenced by dominance deviations (Fischer et al., 2008; Hallauer et al., 2010). According to Hallauer et al. (2010), most of the loci involved with GY in maize are due to the occurrence of dominance. This is reflected in a greater difference between H^2 and h^2 for GY than for the other traits, confirming the greater influence of dominance deviations on this trait. The additive genomic relationship matrices of the single crosses (*Ga*) showed high correlations among SNP array, GBS-B73, and GBS-Mock, indicating that these approaches capture similar additive variance patterns. GBS-Mock captures additive relationships in single crosses similar to standard procedures, SNP array, and GBS-B73 (Table 1; Figures S6A–C). However, the correlations between the dominance relationship matrices (*Gd*) were lower but still from medium to high. In both *Ga* and *Gd*, the correlations between SNP array and GBS-Mock were lower, which can be explained by the fact that these SNP datasets use different reference genomes to perform SNP calling.

SCA reflects the action of non-additive gene effects, indicating intra-allelic interactions, is one of the most important parameters in identifying superior single crosses, and is an indicator of genetic distance between parents (Sprague and Tatum, 1942; Carvalho, 1993). Thus, using the SCA estimates as the genetic distance between the lines to identify the panel structure, two heterotic groups were formed, in which the distance between them is maximized. The correlations between the SCA estimates were almost perfect (Table S4). In other words, SNP array, GBS-B73, and GBS-Mock presented equivalent SCA estimates. Thus, the composition of heterotic groups practically did not change from one SNP dataset to another. Therefore, the determination of heterotic groups was similar regardless of the platform used (Figure 5; Table S3).

In addition to presenting distinct heterotic groups, a well-established breeding program also offers good testers. When crossed with parental lines, these provide information about the genetic value of the lines when evaluating the ability to combine between them since it is associated with the additive effects of alleles and additive-type epistatic actions (Cruz and Vencovsky, 1989; Albrecht et al., 2014). The correct choice of a tester can have great significance in the expectation of a successful selection process (Miranda Filho, 2018). According to Hallauer and Martinson (1975), a good tester presents simplicity in use, information that correctly classifies the relative merit of the lines, and the potential for maximizing genetic gain. Thus, based on the GCA estimates between the lines, testers were elected for each heterotic group based on the evaluated traits and the SNP datasets. As expected, there were no differences in tester choice between SNP datasets, as the correlations between GCA estimates across rows were perfect (Table 2; Table S5).

The genotyping approaches produced very similar results but not the same as previous results regarding the study of the population structure of parental lines; it was expected that this would somehow influence the formation of heterotic groups and the choice of testers. However, given the results, the genotyping platform and, more specifically, the approach that uses the simulated genome as a strategy, the GBS-Mock, produced similar results to the standard procedures.

4.3 Influence of genotyping methods on genomic prediction of single crosses

Assessing the performance of all single-cross combinations of parental lines that excel in a breeding program is impractical in most cases, given that the number of combinations grows exponentially as the number of elite parents increases. Thus, obtaining estimates of the genetic values of single crosses not evaluated became viable with the increased availability of molecular markers and genomic prediction models (Hallauer et al., 2010). Therefore, to accelerate genetic gain with limited resources, the prediction of single-cross performance is highly important in modern breeding programs (Basnet et al., 2019).

However, few studies still address how genotyping platforms influence single crosses' prediction and, more specifically, regarding the mock genome as a tool for more sophisticated analyses, such as genomic prediction. Only one recent study shows the mock

TABLE 2 Choice of the best tester according to the SNP datasets (SNP array, GBS-B73, and GBS-Mock), evaluated traits (GY, PH, and EH), and heterotic groups (HP_1 and HP_2).

		Testers	
		HP_1	HP_2
GY	**SNP array**	L023	L006
	GBS-B73	L023	L006
	GBS-Mock	L023	L006
PH	**SNP array**	L001	L003
	GBS-B73	L001	L003
	GBS-Mock	L001	L003
EH	**SNP array**	L001	L003
	GBS-B73	L001	L003
	GBS-Mock	L001	L003

SNP array: Affymetrix® Axiom Maize Genotyping array: GBS-B73, genotyping-by-sequence with SNP calling using B73 as reference genome; GBS-Mock, genotyping-by-sequence with SNP calling using the mock reference built with all parental lines.
SNP, single-nucleotide polymorphism.

genome's efficiency in predicting maize single crosses, which may be an alternative for crops that do not yet have a reference genome (Sabadin et al., 2022). However, our study is more complete and more representative because obtaining approaches from the population structure phase is crucial for the intended use of germplasm through the division of heterotic groups, the definition of testers, and, finally, the genomic prediction of single crosses.

Predictive ability and prediction accuracy are closely related measures. Therefore, we will only discuss it based on predictive ability. GY showed the lowest predictive abilities in all SNP datasets, and EH has the highest in the additive–dominance GBLUP prediction model (Figure 6, Table S2). Combs and Bernardo (2013) suggested that genomic predictions are more accurate for traits with higher heritability. In the results of Hayes et al. (2010), complex traits controlled by many small effect loci, such as GY, showed lower predictive abilities than less complex traits. Although GBS-Mock has a lower number of markers, this approach presented a similar performance to the other SNP datasets for all traits, corroborating the hypothesis that it is possible to substantially reduce the number of markers and maintain a high predictive ability (Tayeh et al., 2015; Ma et al., 2016; Sousa et al., 2019), except for long-term breeding cycles without updating the training population that would demand high marker densities (DoVale et al., 2022). In addition, the genetic distance estimates between the SNP datasets were very similar (Figure S5).

Selection intensity must be chosen thoughtfully, as genetic variability can be drastically reduced with high selection pressure. The choice of appropriate selection intensities depends on the size of the population and the duration of the breeding program, whether short-term or long-term. In general, selection intensities ranging from 10% to 40% are used in plant breeding, the highest being applied at the beginning of a breeding program (Hallauer et al., 2010). For the coincidence of individuals by phenotypic selection and genomic selection, the SNP datasets showed similar behavior as the selection intensity was increased, being more pronounced from 1% to 10% of selection intensity. From then on, observing the coincidence of selection gains smaller increments (Figure 7). Our results for predictive ability and coincidence of selection agree with the results of Sabadin et al. (2022). It is valid to consider that those different intensities modify the response rates. Thus, this coincidence between phenotypic and genomic selections is expected to reach a plateau and subsequently decrease.

Despite the apparent differences between SNP datasets, the general message is that these approaches perform comparably in the analyses performed in this study, even accessing different types of genomic sequences. While SNP array is derived from exome capture and therefore focused on coding sequence variation, the GBS data represent a wider diversity survey in genomic regions associated with low levels of DNA methylation, which may also include many genes and gene regulatory regions (Darrier et al., 2019; Negro et al., 2019). However, the physical distribution of markers reveals higher frequencies of SNPs at the gene-rich telomeric ends of each of the chromosomes for both approaches, with this frequency being more pronounced in SNP array (Bayer et al., 2017). The platforms probably capture nearby markers in linkage disequilibrium with quantitative trait loci (QTLs). In this sense, using different platforms can be advantageous, as it allows the identification of additional QTLs.

4.4 Possible applications of the mock genome in plant breeding

Until recently, only the main commercial crops benefited from state-of-the-art technologies. However, the development of the GBS platform emerged as an alternative for using such technologies to be viable for orphan crops. Approaches like this can convert orphan crops into crops rich in genomic resources and substantially reduce the breeding process (Varshney et al., 2009; Varshney and May, 2012; Varshney et al., 2012).

Previously, this process was much slower than nowadays. Rice, for example, took almost 20 years to stop being an orphan crop and become a basic model for cereals (Varshney et al., 2009). Introducing these crops into the genomic era also accelerates the identification of genes underlying important agronomic traits and improves our understanding of the evolution of these species (Ye and Fan, 2021). However, many minor crops are becoming rich in genetic resources as a result of investments from various public and private initiatives, such as the African Orphan Crops Consortium (AOCC) (Hendre et al., 2019), which is a global partnership that is generating resources genomics for 101 African orphans. One of the objectives of this Consortium is to create reference genomes for these cultures. Although some efforts are being made to pay greater attention to these crops (Chiurugwi et al., 2019; Gregory et al., 2019; Jamnadass et al., 2020), the ideal is still far from being achieved with a view to several species relevant to local diets around the world that are understudied.

Despite initiatives and investments, not all crops will benefit, so they cannot take advantage of modern breeding tools. While these advances are being consolidated, mock genomes can be an alternative, where the absence of a reference genome presents a barrier to the efficient use of GBS data (Melo et al., 2017; Hale et al., 2018). In the meantime, the present study has shown that using a population-tailored mock reference to perform SNP discovery is a valid alternative. With this approach, it was possible to carry out investigations to outline a breeding program, from studies of diversity and population to genomic prediction studies. However, it is important to emphasize that a population with maximum representativeness must be considered when building a mock reference to capture all the population polymorphisms (Sabadin et al., 2022).

These advantages of using a mock genome in genomic studies must consider some caveats; for example, diploid crops with smaller genomes are preferred over cross-pollinated or polyploid orphan crops, as they have genomes that are too complex to be sequenced. However, genome size will become less of a barrier with advances in sequencing technologies and bioinformatics tools (Armstead et al., 2009). Another challenge is in the SNP calling due to the limitations of GBS, which can lead to incorrect identification of homozygotes and heterozygotes because of the low coverage of NGS reads, in addition to a large number of lost and low-quality data (Heslot et al., 2013). According to Sabadin et al. (2022), the mock genomes do not present the physical position of the markers in a constant reference, which hinders studies such as GWAS and candidate gene discovery.

Negro et al. (2019) stated that SNP array and GBS are complementary to detect QTLs tagging different haplotypes in association studies. In this sense, using other platforms can be advantageous, as it allows the identification of additional QTLs. However, no studies still demonstrate the performance of mock genomes for these purposes. When looking for these larger effect marks, the results will probably differ from those obtained with SNP array due to changes in coverage between platforms.

Given what has been shown, it is possible to infer and recommend that a mock genome constructed from the population's polymorphisms to perform the SNP calling is an excellent strategy to support plant breeders in studies of diversity, population structure, the definition of heterotic groups, choice of testers, and genomic prediction in species that still do not have a reference genome available, which is an alternative for the rapid advancement of orphan crop improvement. This approach will play a key role in improving the genetic potential of orphan crops and helping develop sustainable food systems.

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://data.mendeley.com/datasets/5rtc89t7v5/1, DOI:10.17632/5rtc89t7v5.1.

Author contributions

IM was responsible for writing the report, analysing data, interpreting results, and creating tables/figures. JD was responsible for supervising and analyzing data regarding genomic prediction. FS contributed to the manipulation of the raw genomic data and mock genome assembly. RF-N contributed extracting data and designed the work that led to the submission. JD, FS and RF-N played an important role in interpreting the results and contributed to the report review. All authors contributed to the article and approved the submitted version.

Funding

This work was financially supported by Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq)—Process: 409096/2021-0 CNPq/MCTI/FNDCT.

Acknowledgments

We thank the Coordination for the Improvement of Higher Education Personnel—Brazil (CAPES)—Finance Code 001, Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP)—Process 2017/24327-0, for the financial support and Limagrain Field Seeds for the technical support.

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/fpls.2023.1164555/full#supplementary-material

References

Adhikari, L., Lindstrom, O. M., Markham, J., and Missaoui, A. M. (2018). Dissecting key adaptation traits in the polyploid perennial medicago sativa using GBS-SNP mapping. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00934

Adu, G. B., Badu-Apraku, B., Akromah, R., Garcia-Oliveira, A. L., Awuku, F. J., and Gedil, M. (2019). Genetic diversity and population structure of early-maturing tropical maize inbred lines using SNP markers. *PloS One* 14. doi: 10.1371/journal.pone.0214810

Albrecht, J., Gertrud Berens, D., Jaroszewicz, B., Selva, N., Brandl, R., and Farwig, N. (2014). Correlated loss of ecosystem services in coupled mutualistic networks. *Nat. Commun.* 5. doi: 10.1038/ncomms4810

Armstead, I., Huang, L., Ravagnani, A., Robson, P., and Ougham, H. (2009). Bioinformatics in the orphan crops. *Brief Bioinform.* 10, 645–653. doi: 10.1093/bib/bbp036

Bachlava, E., Taylor, C. A., Tang, S., Bowers, J. E., Mandel, J. R., Burke, J. M., et al. (2012). Snp discovery and development of a high-density genotyping array for sunflower. *PloS One* 7. doi: 10.1371/journal.pone.0029814

Badu-Apraku, B., Fakorede, M. A. B., Gedil, M., Annor, B., Talabi, A. O., Akaogu, I. C., et al. (2016a). Heterotic patterns of IITA and CIMMYT early-maturing yellow maize inbreds under contrasting environments. *Agron. J.* 108, 1321–1336. doi: 10.2134/agronj2015.0425

Badu-Apraku, B., Fakorede, M. A. B., Gedil, M., Talabi, A. O., Annor, B., Oyekunle, M., et al. (2015). Heterotic responses among crosses of IITA and CIMMYT early white maize inbred lines under multiple stress environments. *Euphytica* 206, 245–262. doi: 10.1007/s10681-015-1506-0

Badu-Apraku, B., Fakorede, M. A. B., Talabi, A. O., Oyekunle, M., Akaogu, I. C., Akinwale, R. O., et al. (2016b). Gene action and heterotic groups of early white quality protein maize inbreds under multiple stress environments. *Crop Sci.* 56, 183–199. doi: 10.2135/cropsci2015.05.0276

Badu-Apraku, B., Fontem, L. A., Akinwale, R. O., and Oyekunle, M. (2011). Biplot analysis of diallel crosses of early maturing tropical yellow maize inbreds in stress and nonstress environments. *Crop Sci.* 51, 173–188. doi: 10.2135/cropsci2010.06.0366

Baldermann, S., Blagojević, L., Frede, K., Klopsch, R., Neugart, S., Neumann, A., et al. (2016). Are neglected plants the food for the future? *CRC Crit. Rev. Plant Sci.* 35, 106–119. doi: 10.1080/07352689.2016.1201399

Barter, R. L., and Yu, B. (2018). Superheat: an r package for creating beautiful and extendable heatmaps for visualizing complex data. *J. Comput. Graphical Stat* 27, 910–922. doi: 10.1080/10618600.2018.1473780

Basnet, B. R., Crossa, J., Dreisigacker, S., Pérez-Rodríguez, P., Manes, Y., Singh, R. P., et al. (2019). Hybrid wheat prediction using genomic, pedigree, and environmental covariables interaction models. *Plant Genome* 12, 180051. doi: 10.3835/plantgenome2018.07.0051

Bayer, M. M., Rapazote-Flores, P., Ganal, M., Hedley, P. E., Macaulay, M., Plieske, J., et al. (2017). Development and evaluation of a barley 50k iSelect SNP array. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01792

Bernardo, R. (2002). *Breeding for quantitative traits in plants* (Woodbury: Stemma Press).

Bevan, M. W., and Uauy, C. (2013). Genomics reveals new landscapes for crop improvement. *Genome Biol* 14, 206. doi: 10.1186/gb-2013-14-6-206

Browning, B. L., Zhou, Y., and Browning, S. R. (2018). A one-penny imputed genome from next-generation reference panels. *Am. J. Hum. Genet.* 103, 338–348. doi: 10.1016/j.ajhg.2018.07.015

Burstin, J., Vienne, D., Dubreuil, P., and Damerval, C. (1994). Molecular markers and protein quantities as genetic descriptors in maize. I. Genetic diversity among 21 inbred lines. theoret. app. *Genetics* 89, 943–950. doi: 10.1007/BF00224522

Butler, D. G., Cullis, B. R., Gilmour, A. R., Gogel, B. J., and Thompson, R. (2018) *ASReml-r reference manual version 4 ASReml estimates variance components under a general linear mixed model by residual maximum likelihood (REML)*. Available at: http://www.vsni.co.uk/.

Carvalho, G. R. (1993). Evolutionary aspects of fish distribution: genetic variability and adaptation. *J. Fish Biol.* 43, 53–73. doi: 10.1111/j.1095-8649.1993.tb01179.x

Chakradhar, T., Hindu, V., and Reddy, P. S. (2017). Genomic-based-breeding tools for tropical maize improvement. *Genetica* 145, 525–539. doi: 10.1007/s10709-017-9981-y

Chiurugwi, T., Kemp, S., Powell, W., and Hickey, L. T. (2019). Speed breeding orphan crops. *Theor. Appl. Genet.* 132, 607–616. doi: 10.1007/s00122-018-3202-7

Chu, J., Zhao, Y., Beier, S., Schulthess, A. W., Stein, N., Philipp, N., et al. (2020). Suitability of single-nucleotide polymorphism arrays versus genotyping-By-Sequencing for genebank genomics in wheat. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00042

Combs, E., and Bernardo, R. (2013). Accuracy of genomewide selection for different traits with constant population size, heritability, and number of markers. *Plant Genome* 6. doi: 10.3835/plantgenome2012.11.0030

Covarrubias-Pazaran, G. (2016). Genome-assisted prediction of quantitative traits using the r package sommer. *PloS One* 11. doi: 10.1371/journal.pone.0156744

Crossa, J., Beyene, Y., Semagn, K., Pérez, P., Hickey, J. M., Chen, C., et al. (2013). Genomic prediction in maize breeding populations with genotyping-by-sequencing. *G3: Genes Genomes Genet.* 3, 1903–1926. doi: 10.1534/g3.113.008227

Cruz, C. D., and Vencovsky, R. (1989). Comparação de alguns metodos de análise dialélica. *Rev. Bras. Genética* 12, 425–438.

Darrier, B., Russell, J., Milner, S. G., Hedley, P. E., Shaw, P. D., Macaulay, M., et al. (2019). A comparison of mainstream genotyping platforms for the evaluation and use of barley genetic resources. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00544

DoVale, J. C., Carvalho, H. F., Sabadin, F., and Fritsche-Neto, R. (2022). Genotyping marker density and prediction models effects in long-term breeding schemes of cross-pollinated crops. *Theor. Appl. Genet.* 135, 4523–4539. doi: 10.1007/s00122-022-04236-3

Elbasyoni, I. S., Lorenz, A. J., Guttieri, M., Frels, K., Baenziger, P. S., Poland, J., et al. (2018). A comparison between genotyping-by-sequencing and array-based scoring of SNPs for genomic prediction accuracy in winter wheat. *Plant Sci.* 270, 123–130. doi: 10.1016/j.plantsci.2018.02.019

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One* 6. doi: 10.1371/journal.pone.0019379

Falconer, D. S. (1960). *Introduction to quantitative genetics* (Edinburgh: Oliver and Boyd Ltd).

Falconer, D. S., and Mackay, T. F. C. (1996). *Introduction to quantitative genetics. 4th ed* (Harlow: Addison Wesley Longman).

Fischer, S., Möhring, J., Schön, C. C., Piepho, H. P., Klein, D., Schipprack, W., et al. (2008). Trends in genetic variance components during 30 years of hybrid maize breeding at the university of hohenheim. *Plant Breed.* 127, 446–451. doi: 10.1111/j.1439-0523.2007.01475.x

Frascaroli, E., Schrag, T. A., and Melchinger, A. E. (2013). Genetic diversity analysis of elite European maize (Zea mays l.) inbred lines using AFLP, SSR, and SNP markers reveals ascertainment bias for a subset of SNPs. *Theor. Appl. Genet.* 126, 133–141. doi: 10.1007/s00122-012-1968-6

Fritsche-Neto, R., Akdemir, D., and Jannink, J. L. (2018). Accuracy of genomic selection to predict maize single-crosses obtained through different mating designs. *Theor. Appl. Genet.* 131, 1153–1162. doi: 10.1007/s00122-018-3068-8

Fritsche-Neto, R., Galli, G., Borges, K. L. R., Costa-Neto, G., Alves, F. C., Sabadin, F., et al. (2021). Optimizing genomic-enabled prediction in small-scale maize hybrid breeding programs: a roadmap review. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.658267

Fritsche-Neto, R., Galli, G., Mendonça, L., de, F., Vidotti, M. S., Matias, F. I., et al. (2019). USP Tropical maize hybrid panel. *Mendeley Data* 3.

Fu, D., Xiao, M., Hayward, A., Fu, Y., Liu, G., Jiang, G., et al. (2014). Utilization of crop heterosis: a review. *Euphytica* 197, 161–173. doi: 10.1007/s10681-014-1103-7

Galli, G., Alves, F. C., Morosini, J. S., and Fritsche-Neto, R. (2020). On the usefulness of parental lines GWAS for predicting low heritability traits in tropical maize hybrids. *PloS One* 15. doi: 10.1371/journal.pone.0228724

Ganal, M. W., Durstewitz, G., Polley, A., Bérard, A., Buckler, E. S., Charcosset, A., et al. (2011). A large maize (zea mays l.) SNP genotyping array: development and germplasm genotyping, and genetic mapping to compare with the B73 reference genome. *PloS One* 6. doi: 10.1371/journal.pone.0028334

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). TASSEL-GBS: a high capacity genotyping by sequencing analysis pipeline. *PloS One* 9. doi: 10.1371/journal.pone.0090346

Granato, I. S. C., Galli, G., de Oliveira Couto, E. G., e Souza, M. B., Mendonça, L. F., and Fritsche-Neto, R. (2018). snpReady: a tool to assist breeders in genomic analysis. *Mol. Breed.* 38. doi: 10.1007/s11032-018-0844-8

Gregory, P. J., Mayes, S., Hui, C. H., Jahanshiri, E., Julkifle, A., Kuppusamy, G., et al. (2019). Crops for the future (CFF): an overview of research efforts in the adoption of underutilised species. *Planta* 250, 979–988. doi: 10.1007/s00425-019-03179-2

Gupta, P. K., Rustgi, S., and Mir, R. R. (2008). Array-based high-throughput DNA markers for crop improvement. *Heredity (Edinb)* 101, 5–18. doi: 10.1038/hdy.2008.35

Hale, I., Melo, A. T. O., and Gustafson, H. (2018). Sex-linked molecular markers for two cold-hardy kiwifruit species, actinidia arguta and a. kolomikta. *Eur. J. Hortic. Sci.* 83, 236–246. doi: 10.17660/eJHS.2018/83.4.4

Hallauer, A. R., Carena, M. J., and Filho, J. B. M. (2010). Quantitative Genetics in Maize Breeding. In *Handbook of Plant Breeding*, vol 6. New York, NY: Springer. doi: 10.1007/978-1-4419-0766-0_1

Hallauer, A. R., and Martinson, C. A. (1975). Maternal effects in maize hybrids infected with bipolaris maydis (Nisikado) shoemaker, race T 1. *Crop Sci.* 15, 686–689. doi: 10.2135/cropsci1975.0011183x001500050021x

Hayes, B. J., Pryce, J., Chamberlain, A. J., Bowman, P. J., and Goddard, M. E. (2010). Genetic architecture of complex traits and accuracy of genomic prediction: coat colour, milk-fat percentage, and type in holstein cattle as contrasting model traits. *PloS Genet.* 6. doi: 10.1371/journal.pgen.1001139

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

Heslot, N., Rutkoski, J., Poland, J., Jannink, J. L., and Sorrells, M. E. (2013). Impact of marker ascertainment bias on genomic selection accuracy and estimates of genetic diversity. *PloS One* 8. doi: 10.1371/journal.pone.0074612

Holloway, H. M. C. P., Yu, X., Dunne, J. C., Schwartz, B. M., Patton, A. J., Arellano, C., et al. (2018). A SNP-based high-density linkage map of zoysiagrass (Zoysia japonica steud.) and its use for the identification of QTL associated with winter hardiness. *Mol. Breed.* 38. doi: 10.1007/s11032-017-0763-0

Jamnadass, R., Mumm, R. H., Hale, I., Hendre, P., Muchugi, A., Dawson, I. K., et al. (2020). Enhancing African orphan crops with genomics. *Nat. Genet.* 52, 356–360. doi: 10.1038/s41588-020-0601-x

Kassambara, A. (2017) *Multivariate analysis I practical guide to cluster analysis in r unsupervised machine learning*. Available at: http://www.sthda.com.

Kassambara, A., and Mundt, F. (2020). *Extract and Visualize the Results of Multivariate Data Analyses*. Package "factoextra" [WWW Document]. URL http://www.sthda.com/english/rpkgs/factoextra. (Accessed 11.12.23).

Kendall, M. G. (1938). A new measure of rank correlation. *Biometrika* 30, 81–93. doi: 10.1093/biomet/30.1-2.81

Lee, M. (1995). Dna markers and plant breeding programs. *Adv. Agronomy* 55, 265–344. doi: 10.1016/S0065-2113(08)60542-8

Lee, Y. G., Jeong, N., Kim, J. H., Lee, K., Kim, K. H., Pirani, A., et al. (2015). Development, validation and genetic analysis of a large soybean SNP genotyping array. *Plant J.* 81, 625–636. doi: 10.1111/tpj.12755

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, C., Li, Y., Bradbury, P. J., Wu, X., Shi, Y., Song, Y., et al. (2015). Construction of high-quality recombination maps with low-coverage genomic sequencing for joint linkage analysis in maize. *BMC Biol.* 13. doi: 10.1186/s12915-015-0187-4

Lopez-Cruz, M., Crossa, J., Bonnett, D., Dreisigacker, S., Poland, J., Jannink, J. L., et al. (2015). Increased prediction accuracy in wheat breeding trials using a marker × environment interaction genomic selection model. *G3: Genes Genomes Genet.* 5, 569–582. doi: 10.1534/g3.114.016097

Ma, Y., Reif, J. C., Jiang, Y., Wen, Z., Wang, D., Liu, Z., et al. (2016). Potential of marker selection to increase prediction accuracy of genomic selection in soybean (Glycine max l.). *Mol. Breed.* 36. doi: 10.1007/s11032-016-0504-9

Mammadov, J., Aggarwal, R., Buyyarapu, R., and Kumpatla, S. (2012). SNP markers and their impact on plant breeding. *Int. J. Plant Genomics* 2012. doi: 10.1155/2012/728398

Mantel, N. (1967). The detection of disease clustering and a generalized regression approach. *Cancer Res.* 27, 209–220.

Matias, F. I., Xavier Meireles, K. G., Nagamatsu, S. T., Lima Barrios, S. C., Borges do Valle, C., Carazzolle, M. F., et al. (2019). Expected genotype quality and diploidized marker data from genotyping-by-Sequencing of urochloa spp. tetraploids. *Plant Genome* 12, 190002. doi: 10.3835/plantgenome2019.01.0002

Mayes, S., Massawe, F. J., Alderson, P. G., Roberts, J. A., Azam-Ali, S. N., and Hermann, M. (2012). The potential for underutilized crops to improve security of food production. *J. Exp. Bot.* 63, 1075–1079. doi: 10.1093/jxb/err396

Melo, A. T. O., Bartaula, R., and Hale, I. (2016). GBS-SNP-CROP: a reference-optional pipeline for SNP discovery and plant germplasm characterization using variable length, paired-end genotyping-by-sequencing data. *BMC Bioinf.* 17. doi: 10.1186/s12859-016-0879-y

Melo, A. T. O., Guthrie, R. S., and Hale, I. (2017). GBS-based deconvolution of the surviving north American collection of cold-hardy kiwifruit (Actinidia spp.) germplasm. *PloS One* 12. doi: 10.1371/journal.pone.0170580

Mendonça, L., de, F., Granato, Í.S.C., Alves, F. C., Morais, P. P. P., Vidotti, M. S., et al. (2017). Accuracy and simultaneous selection gains for n-stress tolerance and n-use efficiency in maize tropical lines. *Sci. Agric.* 74, 481–488. doi: 10.1590/1678-992x-2016-0313

Messing, J., and Dooner, H. K. (2006). Organization and variability of the maize genome. *Curr. Opin. Plant Biol.* 9, 157–163. doi: 10.1016/j.pbi.2006.01.009

Milner, S. G., Jost, M., Taketa, S., Mazón, E. R., Himmelbach, A., Oppermann, M., et al. (2019). Genebank genomics highlights the diversity of a global barley collection. *Nat. Genet.* 51, 319–326. doi: 10.1038/s41588-018-0266-x

Miranda Filho, J. B. (2018). "Testadores e dialelo," in *Melhoramento de milho*. Eds. R. DeLIMA and A. BORÉM (Viçosa, MG: Editora UFV), 130–158.

Mrode, R. A. (2014). *Linear models for the prediction of animal breeding values. 3rd ed* (Edinburgh, UK: CABI).

Mulvaney, M. J., and Devkota, P. J. (2020). Adjusting crop yield to a standard moisture content. *EDIS* 2020. doi: 10.32473/edis-ag442-2020

Munjal, G., Hao, J., Teuber, L. R., and Brummer, E. C. (2018). Selection mapping identifies loci underpinning autumn dormancy in alfalfa (Medicago sativa). *G3: Genes Genomes Genet.* 8, 461–468. doi: 10.1534/g3.117.300099

Negro, S. S., Millet, E. J., Madur, D., Bauland, C., Combes, V., Welcker, C., et al. (2019). Genotyping-by-sequencing and SNP-arrays are complementary for detecting quantitative trait loci by tagging different haplotypes in association studies. *BMC Plant Biol.* 19. doi: 10.1186/s12870-019-1926-4

Oksanen, J., Simpson, G. L., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., et al. (2019) *Vegan: community ecology package*. Available at: https://github.com/vegandevs/vegan (Accessed November 2, 2022).

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. *PloS One* 7. doi: 10.1371/journal.pone.0032253

Rasheed, A., Hao, Y., Xia, X., Khan, A., Xu, Y., Varshney, R. K., et al. (2017). Crop breeding chips and genotyping platforms: progress, challenges, and perspectives. *Mol. Plant* 10, 1047–1064. doi: 10.1016/j.molp.2017.06.008

Rogers, J. S. (1972). "Measures of genetic similarity and genetic distance," in *Studies in genetics VII*, vol. 7213. (Texas, USA: University of Texas Publication), 145–153.

Romay, M. C., Millard, M. J., Glaubitz, J. C., Peiffer, J. A., Swarts, K. L., Casstevens, T. M., et al. (2013). Comprehensive genotyping of the USA national maize inbred seed bank. *Genome Biol.* 14. doi: 10.1186/gb-2013-14-6-r55

Sabadin, F., Carvalho, H. F., Galli, G., and Fritsche-Neto, R. (2022). Population-tailored mock genome enables genomic studies in species without a reference genome. *Mol. Genet. Genomics* 297, 33–46. doi: 10.1007/s00438-021-01831-9

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. *Sci. (1979)* 326, 1112–1115. doi: 10.1126/science.1178534

Singh, N., Jayaswal, P. K., Panda, K., Mandal, P., Kumar, V., Singh, B., et al. (2015). Single-copy gene based 50 K SNP chip for genetic studies and molecular breeding in rice. *Sci. Rep.* 5. doi: 10.1038/srep11600

Sousa, M. B., Galli, G., Lyra, D. H., Granato, Í.S.C., Matias, F. I., Alves, F. C., et al. (2019). Increasing accuracy and reducing costs of genomic prediction by marker selection. *Euphytica* 215. doi: 10.1007/s10681-019-2339-z

Sprague, G. F., and Tatum, L. A. (1942). General vs combining ability in single crosses of corn. *Agron. J.* 34, 923–932. doi: 10.2134/agronj1942.00021962003400100008x

Tadele, Z., and Assefa, K. (2012). Increasing food production in africa by boosting the productivity of understudied crops. *Agronomy* 2, 240–283. doi: 10.3390/agronomy2040240

Tayeh, N., Klein, A., le Paslier, M. C., Jacquin, F., Houtin, H., Rond, C., et al. (2015). Genomic prediction in pea: effect of marker density and training population size and composition on prediction accuracy. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00941

Technow, F., Schrag, T. A., Schipprack, W., Bauer, E., Simianer, H., and Melchinger, A. E. (2014). Genome properties and prospects of genomic prediction of hybrid performance in a breeding program of maize. *Genetics* 197, 1343–1355. doi: 10.1534/genetics.114.165860

Tester, M., and Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Sci. (1979)* 327, 818–822. doi: 10.1126/science.1183700

Thomson, M. J. (2014). High-throughput SNP genotyping to accelerate crop improvement. *Plant Breed Biotechnol.* 2, 195–212. doi: 10.9787/pbb.2014.2.3.195

Unterseer, S., Bauer, E., Haberer, G., Seidel, M., Knaak, C., Ouzunova, M., et al. (2014). A powerful tool for genome analysis in maize: development and evaluation of the high density 600 k SNP genotyping array. *BMC Genomics* 15. doi: 10.1186/1471-2164-15-823

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

Varshney, R. K., Close, T. J., Singh, N. K., Hoisington, D. A., and Cook, D. R. (2009). Orphan legume crops enter the genomics era! *Curr. Opin. Plant Biol.* 12, 202–210. doi: 10.1016/j.pbi.2008.12.004

Varshney, R. K., and May, G. D. (2012). Next-generation sequencing technologies: opportunities and obligations in plant genomics. *Brief Funct. Genomics* 11, 1–2. doi: 10.1093/bfgp/els001

Varshney, R. K., Ribaut, J.-M., Buckler, E. S., Tuberosa, R., Rafalski, J. A., and Langridge, P. (2012). Can genomics boost productivity of orphan crops? *Nat. Biotechnol.* 30, 1172–1176. doi: 10.1038/nbt.2440

Wang, N., Yuan, Y., Wang, H., Yu, D., Liu, Y., Zhang, A., et al. (2020). Applications of genotyping-by-sequencing (GBS) in maize genetics and breeding. *Sci. Rep.* 10. doi: 10.1038/s41598-020-73321-8

Winfield, M. O., Allen, A. M., Burridge, A. J., Barker, G. L. A., Benbow, H. R., Wilkinson, P. A., et al. (2016). High-density SNP genotyping array for hexaploid wheat and its secondary and tertiary gene pool. *Plant Biotechnol. J.* 14, 1195–1206. doi: 10.1111/pbi.12485

Wu, Y., San Vicente, F., Huang, K., Dhliwayo, T., Costich, D. E., Semagn, K., et al. (2016). Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theor. Appl. Genet.* 129, 753–765. doi: 10.1007/s00122-016-2664-8

Xu, C., Ren, Y., Jian, Y., Guo, Z., Zhang, Y., Xie, C., et al. (2017). Development of a maize 55 K SNP array with improved genome coverage for molecular breeding. *Mol. Breed.* 37. doi: 10.1007/s11032-017-0622-z

Yassue, R. M., Carvalho, H. F., Gevartosky, R., Sabadin, F., Souza, P. H., Bonatelli, M. L., et al. (2021a). On the genetic architecture in a public tropical maize panel of the symbiosis between corn and plant growth-promoting bacteria aiming to improve plant resilience. *Mol. Breed.* 41. doi: 10.1007/s11032-021-01257-6

Yassue, R. M., Sabadin, F., Galli, G., Alves, F. C., and Fritsche-Neto, R. (2021b). CV-α: designing validations sets to increase the precision and enable multiple comparison tests in genomic prediction. *Euphytica* 217. doi: 10.1007/s10681-021-02831-x

Ye, C. Y., and Fan, L. (2021). Orphan crops and their wild relatives in the genomic era. *Mol. Plant* 14, 27–39. doi: 10.1016/j.molp.2020.12.013

Zheng, X., Levine, D., Shen, J., Gogarten, S. M., Laurie, C., and Weir, B. S. (2012). A high-performance computing toolset for relatedness and principal component analysis of SNP data. *Bioinformatics* 28, 3326–3328. doi: 10.1093/bioinformatics/bts606

frontiers | Frontiers in Plant Science

TYPE Review
PUBLISHED 06 June 2023
DOI 10.3389/fpls.2023.1179547

OPEN ACCESS

EDITED BY
Omena Bernard Ojuederie,
Kings University, Nigeria

REVIEWED BY
Oluwaseyi Samuel Olanrewaju,
North-West University, South Africa
Chandra Mohan Singh,
Banda University of Agriculture and Technology, India

*CORRESPONDENCE
Rahul Chandora
rahul1@icar.gov.in
Jai Chand Rana
j.rana@cgiar.org

RECEIVED 04 March 2023
ACCEPTED 02 May 2023
PUBLISHED 06 June 2023

CITATION
C. KR, G, T. B, Chandora R and Rana JC (2023) Moth bean (*Vigna aconitifolia*): a minor legume with major potential to address global agricultural challenges. *Front. Plant Sci.* 14:1179547. doi: 10.3389/fpls.2023.1179547

Moth bean (*Vigna aconitifolia*): a minor legume with major potential to address global agricultural challenges

Kanishka R. C.[1], Gayacharan[2], Basavaraja T.[3], Rahul Chandora[1*] and Jai Chand Rana[4*]

[1]Indian Council of Agricultural Research (ICAR)-National Bureau of Plant Genetic Resources, Regional Station, Shimla, India, [2]Indian Council of Agricultural Research (ICAR)-National Bureau of Plant Genetic Resources, New Delhi, India, [3]Indian Council of Agricultural Research (ICAR)-Indian Institute of Pulses Research, Kanpur, India, [4]Alliance of Bioversity International and International Center for Tropical Agriculture (CIAT), India Office, National Agricultural Science Complex, New Delhi, India

Moth bean (*Vigna aconitifolia*) is an orphan legume of *Vigna* genus, exhibiting wide adaptability and has the potential to grow well in arid and semi-arid areas, predominantly across different eco-geographical regions of Asia, particularly the Indian subcontinent. The inherent adaptive attributes of this crop have made it more tolerant towards a diverse array of abiotic and biotic stresses that commonly restrain yield among other *Vigna* species. Additionally, the legume is recognized for its superior nutritional quality owing to its high protein content as well as amino acid, mineral and vitamin profile and is utilized as both food and fodder. Moth bean can play a vital role in sustaining food grain production, enhancing nutritional security as well as provide a source of income to resource-poor farmers amid rise in global temperatures and frequent drought occurrences, particularly in rain-fed cropping systems which accounts for about 80% of the world's cultivated land. However, this minor legume has remained underutilized due to over-exploitation of major staple crops. With the exception of a few studies involving conventional breeding techniques, crop improvement in moth bean for traits such as late maturity, indeterminate growth habit, shattering and anti-nutritional factors has not garnered a lot of attention. Recent advances in sequencing technologies, modern breeding approaches and precision phenotyping tools, in combination with the available crop gene pool diversity in gene banks, can accelerate crop improvement in moth bean and lead to the development of improved cultivars. Considering the recent surge in awareness about the development of climate-smart crops for sustainable agricultural future, collective effort towards effective utilization of this hardy, neglected legume is the need of the hour.

KEYWORDS

moth bean, genetic diversity, crop improvement, stress tolerance, climate change, sustainable agriculture

1 Introduction

Agricultural sustainability is stymied by excessive reliance on major staple crops, climate change as well as deterioration of cultivable land. These limitations extend a challenge to global food security while simultaneously accentuating rural poverty and malnutrition in developing and under-developed nations. Many crops are missing from agricultural fields while others have seen a slump in both cultivation and usage. Evolution of new pests and pathogens along with enhanced crop uniformity in farmers' field are posing a serious threat to crop production. Tackling global food demand and hidden hunger warrants a radical shift from current unsustainable agricultural practices and call attention to alternative potential future crops. Introduction of new crop species in cropping systems and enhancing crop diversity through the utilization of diverse germplasm can play vital role in global food and nutritional security. Underutilized plants are the potential genetic resources to be exploited to address the decreasing nutritional quality and stress resilience in the already mainstreamed cultivars. 'Underutilized' plant species are those that had been grown widely in the past or have the capability to be grown extensively in the future, however, are cultivated restrictively owing to agronomic, economic or genetic causes (Gruère et al., 2006). They are also known as 'orphan' or 'neglected' crops owing to insufficient information available about them as a result of sparse consideration from research and development (Eyzaguirre et al., 1999). These crops remain 'underexploited' or 'minor' since their economic potential has not been ascertained (Padulosi and Hoeschle-Zeledon, 2004).

Pulse crops are exceptional sources of quality protein, in addition to being utilized for their medicinal properties, as fodder for livestock, for enriching soil as a result of symbiotic relationship with nitrogen-fixing bacteria and mitigating greenhouse emissions (Lemke et al., 2007; Singh et al., 2007; Stagnari et al., 2017). The current scenario of booming food insecurity signals an urgency to reformulate and steer crop improvement and production agronomy strategies, in the coming decades, towards grain legumes to successfully determine climate-resilient species having enhanced grain attributes (Considine et al., 2017). Crop diversification, alternative cropping systems and value-added products' development are significant contributions made by underutilized legumes, to the life of local communities. Mostly, these minor pulses are adapted to marginal lands. The genetic resources of these orphan plants face rapid destruction owing to changes in traditional food habits and depreciation of traditional farming culture along with the introduction and acclimatization of high yielding crops. Therefore, mainstreaming of these underutilized crop plants, essentially minor legumes, should be set in motion in order to impede global food concerns.

Moth bean [*Vigna aconitifolia* (Jacq.) Maréchal syn. *Phaseolus aconitifolius* Jacq. (2n= 2x=22)] is an underutilized, minor grain legume. It is also known as mat bean, math, mattenbohne, matki, dew bean, Turkish gram or haricot papillon (Heuzé et al., 2020). Moth bean is an annual herbaceous trailing legume of warm and dry habitats, particularly the semi-arid and arid regions of Indian subcontinent. It is principally grown for its protein-rich seeds, sprouts and green pods that are used as vegetable, apart from being equally important as feed and fodder for livestock and as a cover crop as well as green manuring, hence serves as multipurpose crop (Saravanan and Ignacimuthu, 2015) (Figure 1). Integrated consumption of moth bean like pulse legume crops with cereals is expedient as they have a complementary relationship, for moth is rich in lysine and leucine and cereals supply sulphur-containing amino acids, thereby supplementing mutual amino acid deficiencies.

As a prospective major legume crop, *Vigna aconitifolia* is quite notable for being the most drought hardy and heat tolerant pulse among Asian *Vigna*s (Tomooka et al., 2011). It is usually grown on marginal and sub-marginal lands with poor moisture holding capacity. Having a deep, fast penetrating tap root system and a dense low-lying leaf cover that resembles a mat, it can withstand lack of water and drying, hot temperatures (~ 40°C), by keeping the soil moist as well as lowering soil temperature, thereafter also reducing the possibility of soil erosion. The multi-adaptive nature of moth bean, in addition to requiring little to no care and input, has led it to being recognized as an arid legume. Notwithstanding the attributes of moth bean that makes it a future legume for sustainable agriculture, it is only grown limitedly and there has been a precipitous reduction in the cultivable area and production of the legume over the past decade (Figure 2) owing to agronomic bottlenecks *viz.*, low variability, slow growth, longer maturation duration, substandard yields due to poor seed set and below par response to fertilizer treatments (Harsh et al., 2016). Its competence, with respect to agronomy, genetics and nutrition has been vastly overlooked. This review attempts to encapsulate the work done in *Vigna aconitifolia*, in an effort to harness its potentialities to curb the unprecedented challenges to global agriculture.

2 Origin and distribution

Moth bean is considered to be a native of India, Pakistan, Myanmar and Sri Lanka (Piper and Morse, 1914; Purseglove, 1974; Rachie and Roberts, 1974; Jain and Mehra, 1980). Vavilov (1926) and De Candolle (1986) reported India as the center of origin for wild and cultivated forms of *V. aconitifolia* (Jacq), and Maréchal (1978) proposed that Sri Lanka and Pakistan are the centers of diversity of this minor legume. Dana (1976) proffered that the wild forms of moth bean were distributed from Sonoran Desert of Mexico to Tapachula, near the Guatemalan border in Central America.

Vigna aconitifolia is principally cultivated in India. It is also grown in Sri Lanka, Myanmar, China, Pakistan, Malaysia, Africa and is also grown in the South-Western states of the USA (Jain and Mehra, 1980; Munro and Small, 1998; Brink and Jansen, 2006; Kochhar, 2016). In India, it is distributed from the North-West Himalayas (up to the height of 6000 m) to Karnataka in south and from the foot hills of North-East Himalayas in the east to Saurashtra in the west (Arora, 1985). Moth bean is essentially grown in the arid and semi-arid regions, particularly in the North-Western states. Rajasthan, Maharashtra, Gujarat, Punjab, Haryana, Jammu and Kashmir, Madhya Pradesh and Uttar Pradesh are the prime moth

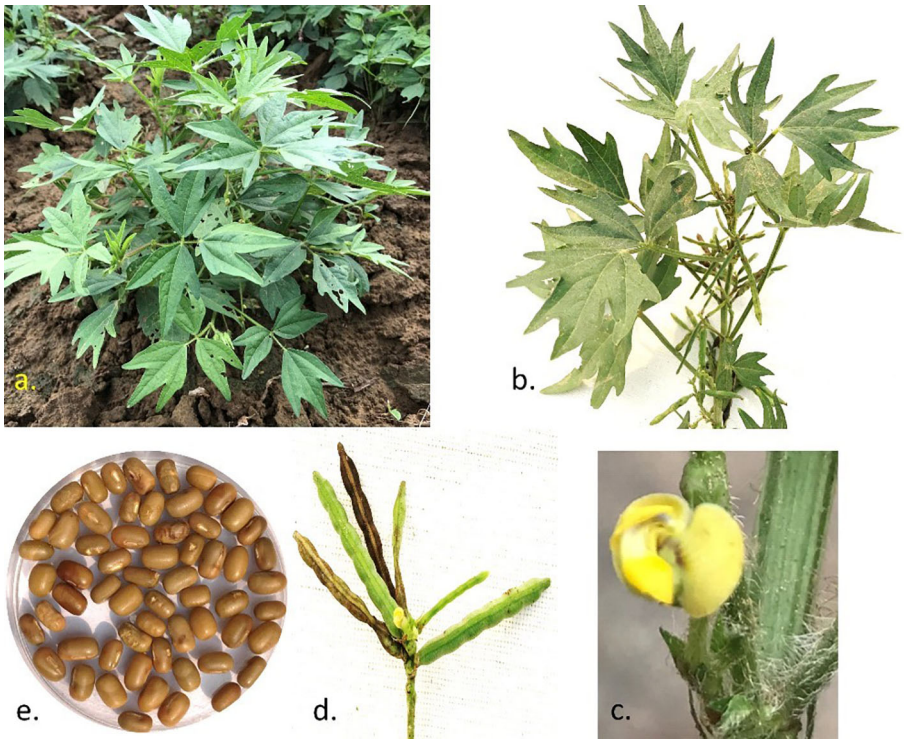


FIGURE 1
Highlights of the morphological features of Moth bean: A typical plant type with central branch and trailing primary branches **(A)**, Inflorescence and pods in a branch **(B)**, papilionaceous flower **(C)**, semi-erect pods arrangement in peduncle **(D)** and seed shape and size **(E)**. Seeds are placed in a 2.5 cm dish.

bean growing states (Arora, 1985) (Figure 3). Rajasthan, the driest state in India, contributes the most, both in terms of production as well as area, at national level (Gupta et al., 2016; Viswanatha et al., 2016).

3 Taxonomy and phylogenetic relationship

Moth bean belongs to the genus *Vigna* and sub-family Papilionoideae which comes under the family Fabaceae. It is considered as one of the most primitive species among the *Vigna* genus, with respect to its evolution (Smartt, 1985). Earlier, moth bean or Dew gram (*Phaseolus aconitifolius* Jacq. syn. *Vigna aconitifolia* (Jacq.) Marechal) was a part of the genus *Phaseolus*. Thereafter, it was shifted to the *Vigna* genus of the tribe Phaseolae. *Vigna* and *Phaseolus*, together, form a very elaborate taxonomic group called the *Phaseolus-Vigna* complex. Verdcourt (1970) proposed that the genus *Phaseolus* be restricted to exclusively include those American species that have a tightly coiled style and pollen grains without course reticulation, thereby significantly promoting the concept of *Vigna*. In the taxonomic revision made by Verdcourt (1970) in Phaseolus Linn. and Vigna savi, species with yellow-colored flowers, under subgenus *Ceratotropis* Piper, were transferred to *Vigna* genus, *Ceratotropis* Verdc. Subgenus. This led to an increase in the number of species included in the genus *Vigna*. Maréchal (1978) followed Verdcourt by presenting a monograph on the *Phaseolus-Vigna* complex and also proposed nomenclatural changes, following which, *Phaseolus aconitifolius* (Jacq.) became *Vigna aconitifolia* (Jacq.) Marechal. This taxonomic system is generally accepted now.

The wild, conspecific forms as well as cultigens of moth bean are not recognized (Maréchal, 1978). The precursor of moth bean is thought to be *Phaseolus trilobata* (L.) (syn. *Phaseolus trilobus*), which is a wild species endemic to India, with both being diploid and having chromosome number 2n=22 (Darlington and Wylie, 1955; Biswas and Dana, 1976a). *Phaseolus trilobata* and *Phaseolus aconitifolia* were considered to be equivalent by some while others believed *Phaseolus trilobata* to be the wild form of *Phaseolus aconitifolia*, however, studies have established both of them to be different (Sampson, 1936; Whyte et al., 1953; Klozová, 1965; West and Garber, 1967). Similarly, *Vigna aconitifolia* was classified with *Vigna radiata* and *Vigna mungo*, nonetheless, RAPD analysis to study genetic variability showed interspecific differences among them (Dana, 1980; Kaga et al., 1996). Takahashi et al. (2016) have performed molecular phylogenetic analysis of *Vigna* species using rDNA-ITS and *atpB-rbcL* sequences which divided the *Aconitifoliae* section into various branches, of which *V. aconitifolia* is one (Figure 4).

4 Nutritional profile

Moth bean is primarily consumed by people in the arid and semi-arid regions of South Asian countries, particularly India

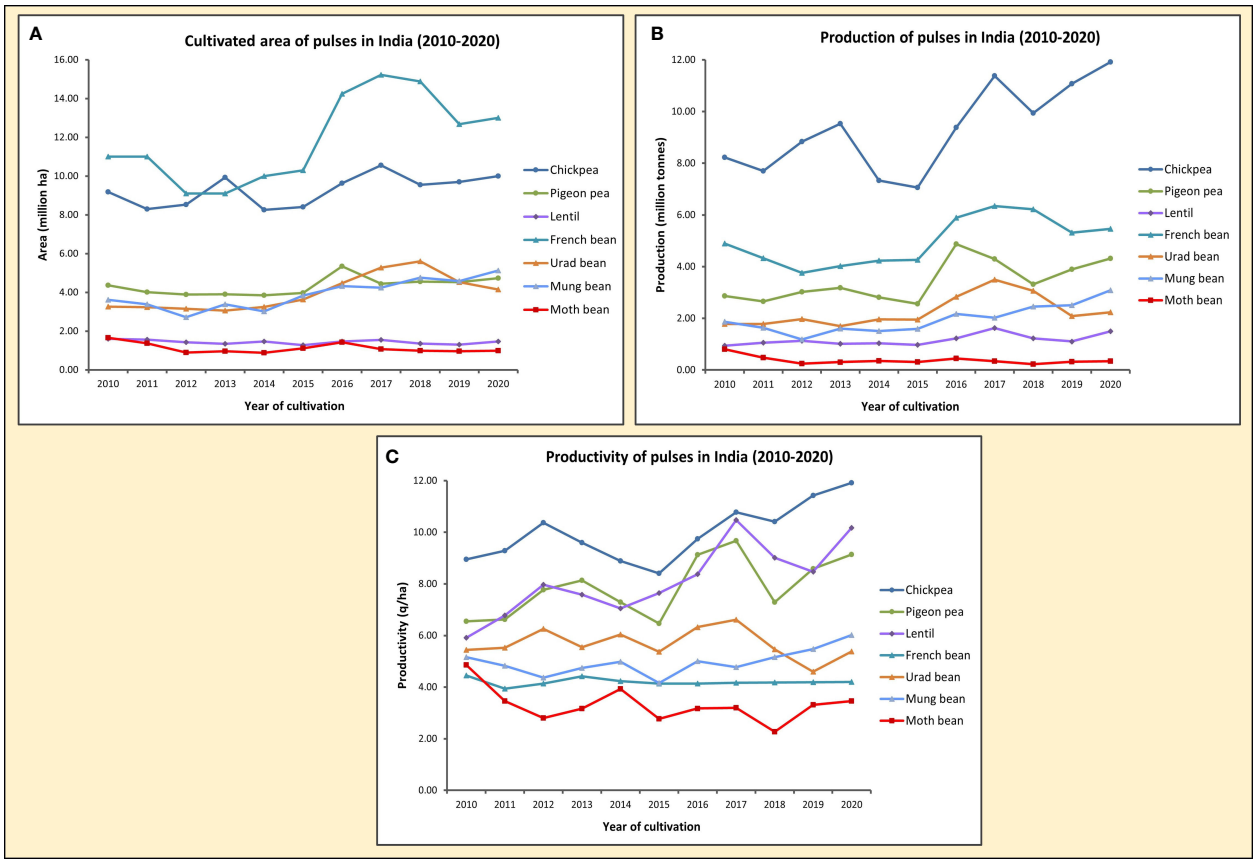


FIGURE 2
Graphs showing the cultivated area in million ha **(A)**, production in million tonnes **(B)** and productivity in q/ha **(C)** of seven pulse crops, including Moth bean, in India from 2010-2020. [Source: Directorate of Economics and Statistics, India (eands.dacnet.nic.in); Food and Agriculture Organization Corporate Statistical Database (FAOSTAT) (www.fao.org/faostat)].

(Adsule, 1996; Takahashi et al., 2016). Ripe seeds are often consumed whole or split and can be cooked, roasted or fried. In India, a popular use of moth bean seeds is in the form of '*dal*' (bean stew) or '*bhujia*' (crispy snack), in addition to being used in the preparation of *papad, mangori, rabri and vada* (moth bean fritters) (Senthilkumar and Ngadi, 2020). The seeds are also ground into flour and often mixed with other flours to make unleavened bread (Brink and Jansen, 2006). Sprouted as well as cooked seeds are a popular breakfast dish while fried, split moth bean can be readily consumed (Medhe et al., 2019). Boiled immature green pods are consumed as a vegetable. As a medical use, the seeds are used to treat fevers. The roots are found to be narcotic (Brink and Jansen, 2006). Other pharmacological properties of Moth bean include anti-hypertensive, anti-oxidant, anti-cancerous, anti-bacterial, diuretic and hypocholesteromic activities (Adsule, 1996; Ma et al., 2010; Saravanan and Ignacimuthu, 2015; Panicker and Hamdule, 2021; Verma et al., 2021). Moth bean also contributes indirectly to human nutrition by being used as fodder for livestock.

In terms of the nutritional aspects of moth bean, it is considered as a good and inexpensive source of protein in cereal-based vegetarian diets among the people of developing nations. The protein content in moth bean seeds ranges from 20 to 24% (Gopalan et al., 1989; Adsule, 1996; Longvah et al., 2017). It is rich in the essential amino acids, lysine (6.63 g/100 g protein) and leucine (7.85 g/100 g) that are deficient in cereals and thus complements the amino acid profile of cereals. Among the total soluble protein present in moth, the glutelin fraction was found to be the highest at 63.93 g/100 g of total soluble protein, followed by glutelin, albumin and prolamin (Sathe and Venkatachalam, 2007). Additionally, the total carbohydrate content in moth bean was found to range from 52 to 68%, the total fat ranged from 1.1 to 3.9% while the crude fiber content ranged from 3.9 to 4.5%. Moth bean is also found to be rich in minerals such as calcium, magnesium, iron, phosphorus and potassium as well as vitamins like niacin, thiamine and riboflavin (Longvah et al., 2017; USDA FoodData Central, 2019). A comparison of the nutritional profile of moth bean with other pulses is illustrated in Table 1.

Moth bean, like other pulses, also contains flatulence-producing oligosaccharides such as raffinose, stachyose and verbascose. Also, like other legumes, it contains anti-nutritional and phytonutrient factors that affect the nutrient bioavailability as well as palatability and digestibility of moth bean. These factors include trypsin inhibitors, phytic acid, saponins and phyto-haemagglutinins (lectins). Trypsin and chymotrypsin inhibitors have been involved in reducing protein digestibility and thereafter in pancreatic hypertrophy (Liener, 1976). Phytic acid interferes with the

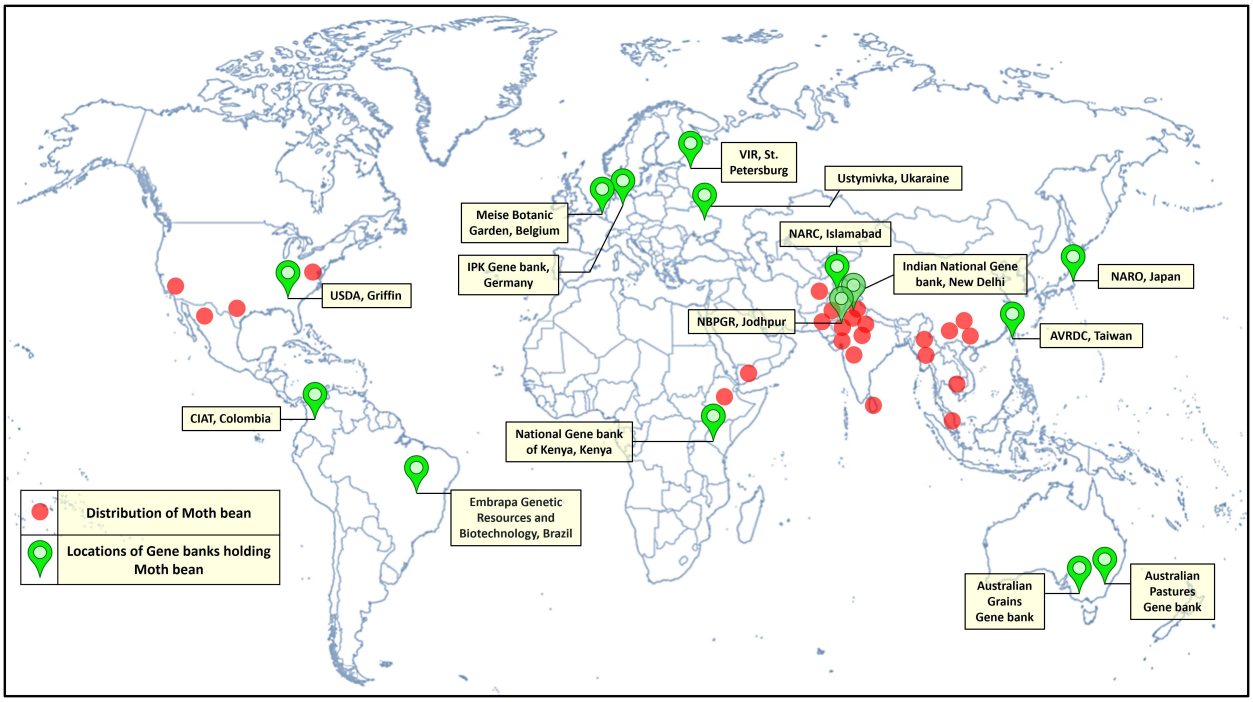


FIGURE 3
Distribution and locations of moth bean germplasm holdings across the world.

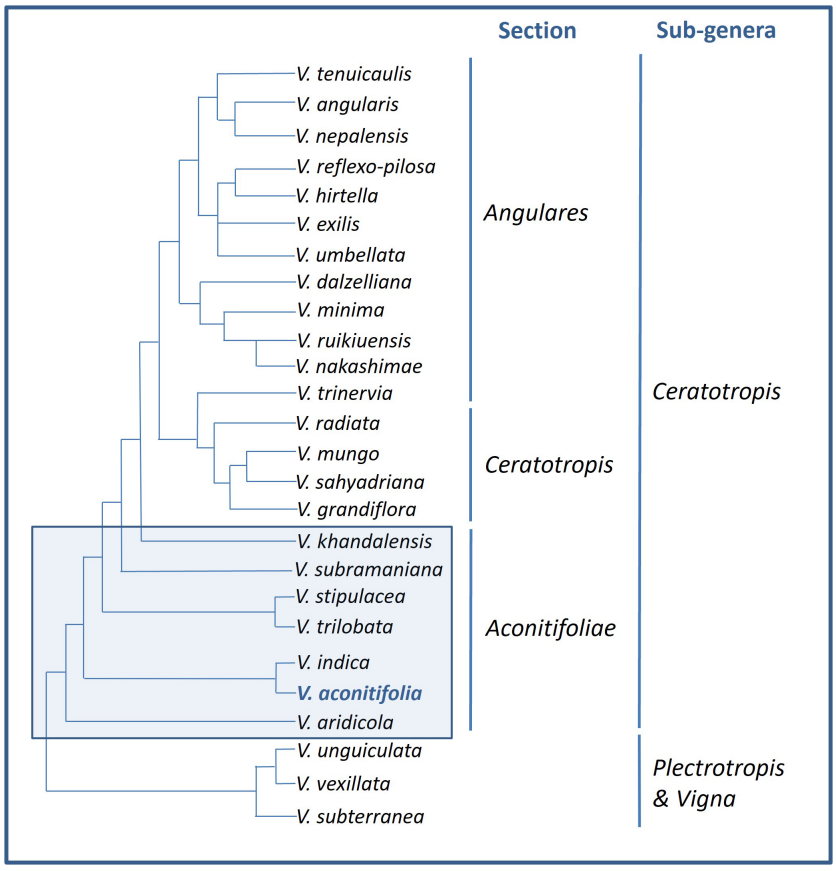


FIGURE 4
Taxonomic relationships among moth bean and other species of the genus *Vigna* [Source: Adapted from Takahashi et al. (2016)].

TABLE 1 Comparative analysis of the nutritional composition of moth bean with other pulses.

Composition	Chickpea (*Cicer arietinum*)	Lentil (*Lens culinaris*)	Kidney bean (*Phaseolus vulgaris*)	Dry peas (*Pisum sativum*)	Moth Bean (*Vigna aconitifolia*)	Black gram (*Vigna mungo*)	Green gram (*Vigna radiata*)	Bambara groundnut (*Vigna subterranea*)	Rice bean (*Vigna umbellata*)
Energy (KJ)	1201 ± 9	1251 ± 23	1252 ± 14	1269 ± 13	1291 ± 16	1219 ± 5	1229 ± 10	1514.78[a]	1265
Carbohydrates (g)	39.56 ± 0.16	48.47 ± 1.12	48.61 ± 0.65	48.93 ± 0.45	52.09 ± 0.96	43.99 ± 0.76	46.13 ± 0.64	55 ± 0.50[a]	51.26
Protein (g)	18.77 ± 0.42	22.49 ± 0.58	19.91 ± 1.44	20.43 ± 0.79	19.75 ± 0.38	21.97 ± 0.63	22.53 ± 0.43	22.46 ± 0.02[a]	19.97
Fat (g)	5.11 ± 0.11	0.64 ± 0.02	1.77 ± 0.04	1.89 ± 0.06	1.76 ± 0.09	1.58 ± 0.06	1.14 ± 0.17	5.80 ± 0.02[a]	0.74
Moisture (g)	8.56 ± 0.37	9.20 ± 0.77	9.87 ± 0.30	9.33 ± 0.61	8.14 ± 0.49	8.70 ± 0.33	9.95 ± 0.42	9.82 ± 0.02[a]	11.12
Total dietary fiber (g)	25.22 ± 0.39	16.82 ± 1.30	16.57 ± 0.63	17.01 ± 0.63	15.12 ± 0.49	20.41 ± 0.06	17.04 ± 0.38	4.50 ± 0.01[a]	13.37
Minerals and trace elements									
Iron (mg)	6.78 ± 0.75	7.57 ± 0.67	6.13 ± 0.77	5.09 ± 0.45	7.90 ± 0.1 7	5.97 ± 0.56	4.89 ± 0.46	18.51 ± 0.10[a]	4.76
Magnesium (mg)	160 ± 17.5	101 ± 13.9	173 ± 9.7	123 ± 8.1	205 ± 13.5	190 ± 19.1	198 ± 39.2	347.15 ± 0.01[a]	201
Manganese (mg)	2.17 ± 0.33	1.55 ± 0.26	1.24 ± 0.11	1.08 ± 0.09	1.07 ± 0.02	1.83 ± 0.34	1.05 ± 0.08	10.46 ± 0.05[a]	1.68
Phosphorus (mg)	267 ± 21.9	274 ± 27.4	409 ± 32.4	334 ± 18.3	362 ± 31.7	345 ± 36.5	353 ± 33.6	738.04 ± 0.15[a]	270
Potassium (mg)	935 ± 37.9	756 ± 63.6	1324 ± 195	922 ± 67.4	1356 ± 53.2 1	1093 ± 24.5	1177 ± 74.3	1702.10 ± 0.50[a]	1196
Sodium (mg)	26.56 ± 10.12	11.20 ± 0.08	10.45 ± 0.05	23.40 ± 0.07	26.34 ± 0.10	26.80 ± 3.77	12.48 ± 0.07	135.30 ± 0.05[a]	10.62
Calcium (mg)	150 ± 18.3	76.13 ± 7.23	126 ± 8.1	75.11 ± 13.93	154 ± 17.0	86.18 ± 8.99	92.43 ± 10.68	256.56 ± 0.05[a]	200
Zinc (mg)	3.37 ± 0.26	3.60 ± 0.23	2.69 ± 0.34	3.10 ± 0.14	1.92 ± 0.08	3.05 ± 0.24	2.67 ± 0.13	1.9 ± 0.1[b]	2.29
Vitamins									
Thiamine (mg)	0.37 ± 0.40	0.40 ± 0.073	0.30 ± 0.020	0.56 ± 0.049	0.45 ± 0.070	0.32 ± 0.024	0.45 ± 0.027	0.28[c]	0.46
Riboflavin (mg)	0.24 ± 0.011	0.22 ± 0.026	0.19 ± 0.018	0.16 ± 0.013	0.09 ± 0.005	0.11 ± 0.08	0.27 ± 0.011	0.12[c]	0.14
Niacin (mg)	2.10 ± 0.06	2.54 ± 0.12	2.42 ± 0.15	2.69 ± 0.15	1.87 ± 0.08	1.85 ± 0.13	2.16 ± 0.13	2.1[c]	2.32
Biotin (µg)	0.93 ± 0.07	1.74 ± 0.16	0.77 ± 0.18	0.53 ± 0.12	2.12 ± 0.21	1.28 ± 0.18	1.35 ± 0.16	–	2.65
Total folic acid (µg)	233 ± 12.9	132 ± 6.7	316 ± 20.1	110 ± 9.3	349 ± 10.8	134 ± 14.2	145 ± 5.4	–	122
Vitamin E (mg)	1.72 ± 0.07	0.19 ± 0.02	0.23 ± 0.01	0.32 ± 0.02	0.79 ± 0.05	0.23 ± 0.02	0.33 ± 0.02	8.22 ± 1.29[d]	1.06

(Source: Longvah et al. (2017); Oyeleke et al. (2012)[a]; N'Dri Yao et al. (2015)[b]; Hardy (2016)[c]; Baptista et al. (2017)[d]).

10.3389/fpls.2023.1179547

bioavailability and absorption of minerals in the body and inhibits proteolytic enzymes and amylases while saponins are gastric irritants (Davies and Nightingale, 1975; Erdman, 1979; Singh and Krikorian, 1982; Deshpande and Cheryan, 1984; Pugalenthi et al., 2005; Bouchenak and Lamri-Senhadji, 2013). However, food preparation techniques, such as soaking, sprouting, boiling, pressure cooking as well as fermentation of moth, aid in improving the taste as well as the bioavailability of nutrients as they deactivate these anti-nutritional factors besides allowing the digestion and assimilation of its starch and protein (Khokhar and Chauhan, 1986; Kigel, 1999; Xu and Chang, 2008; Deshmukh and Pawar, 2020).

Regardless of being nutritionally dense, the present consumption levels of moth bean are rather low which can be attributed to its underutilized status. While there are indirect evidences that support its role in reduction of diseases, they are more based on its nutritional composition, therefore, focused research on the clarification of the health effects of moth bean should be undertaken.

5 Crop improvement

5.1 Plant genetic resources (PGR): status and trait discovery

Major moth bean *ex-situ* collections are maintained by the Indian National Gene bank, ICAR-National Bureau of Plant Genetic Resources, New Delhi, India. In India, systematic exploration and collection of the germplasm work commenced as back early as the 1940s, with the establishment of the Plant Introduction unit in the Division of Botany, IARI (Indian Agricultural Research Institute), New Delhi. Intensive collection efforts, carried out by National Bureau of Plant Genetic Resources, from 1971 till date, have resulted in the assembly of 3422 accessions of moth bean. These accessions include primitive cultivars/ landraces, primarily from the states of Rajasthan, Gujarat, Maharashtra, Karnataka, western Uttar Pradesh, Punjab, Haryana and Madhya Pradesh. Variations in growth habit, leaf location as well as pod and seed color were observed in major collections of *V. aconitifolia*. The collections made from Rajasthan and Gujarat have turned out to be more promising and are currently being utilized in extensive characterization and evaluation programmes, to identify superior genotypes (Bisht and Singh, 2013). Besides germplasm collections from within the country, NBPGR further plays a crucial role in the augmentation of the germplasm from other countries as well. 41 accessions of moth bean have been introduced from Sri Lanka, USA (19), Mexico, USSR, Thailand (5) and others (14). Other small collections are being held by institutions worldwide are University of Georgia, USDA (56 accessions), VIR (48), Australian Grain Gene bank (35), AVRDC-The World Vegetable Centre (26), CIAT, Columbia (8) and Leibniz Institute of Plant Genetics and Crop Plant Research (6) (Table 2; Figure 3). Moth bean has orthodox seeds that can be dried and stored for a longer period with minimum to no loss of viability. A limited range of germplasm accessions can also be found in countries such as Bangladesh, Belgium and Kenya. Active collections are being maintained at NBPGR, Regional Station, Jodhpur (Rajasthan). The working collections of a range of *Vigna* species are also maintained at the Indian Institute of Pulses Research (IIPR), Kanpur, India and its coordinating centers (Asthana, 1998). More than 2000 accessions of moth bean have been characterized and evaluated at NBPGR, Regional Station, Jodhpur (Singh et al., 2001). Yield and other growth characteristics have shown a wide range of variation, as

TABLE 2 Moth bean germplasm holdings at some key centers worldwide.

Countries	Institutes/Centers	Accessions
India	National Bureau of Plant Genetic Resources, Jodhpur	3422
Pakistan	National Agricultural Research Centre (NARC), Islamabad	66
Japan	NARO (National Agriculture and Food Research Organization)	43
Kenya	Kenya National Gene Bank of Kenya, Crop Plant Genetic Resources	50
Russian Federation	N.I. Vavilov Research Institute of Plant Industry Russian VIR, St Petersburg	64
Taiwan	World Vegetable Centre, Taiwan AVRDC, Taiwan	26
USA	Plant Genetic Resources Conservation Unit, Southern Regional Plant Introduction Station, USDA-ARS, Griffin, GA	58
Australia	Australian Grains Genebank	36
Australia	Australian Pastures Genebank, Australia	28
Ukraine	Ustymivka Experimental Station of Plant Production, Ukraine	35
Belgium	Botanic Garden Meise, Belgium	7
Colombia	Centro Internacional de Agricultura Tropical, Colombia	8
Germany	Gene bank, Leibniz Institute of Plant Genetics and Crop Plant Research, Germany	7
Brazil	Embrapa Recursos Genéticos e Biotecnologia, Brazil	7

(Source: Agrawal et al. (2019); Zahoor (2007), www.genesys-pgr.org, www.gene.affrc.go.jp).

demonstrated in Table 3. Nodulation and nitrogenase activity as well as resistance to insect pests and diseases have shown extensive varietal divergence in moth bean (Rao et al., 1984; Dabi and Gour, 1988). Quality characters have also demonstrated a substantial amount of variability (Figure 5).

5.2 Conventional breeding

Initially, moth bean breeders focused mainly on the maintenance of crop genetic resources, germplasm evaluation, and genetic diversity assessment (Cullis and Kunert, 2017). Owing to the difficulty in handling the tiny flowers that drop off during crossing, hybridization programmes were unsuccessful in minor legumes like moth bean resulting in lack in genetic studies. However, studies on germplasm evaluation resulted in the identification of promising genotypes in addition to providing information on diversity, heritability, and other genetic parameters as well as the correlation between traits and environmental influences that could be useful in expediting genetic gains (Vir and Singh, 2015; Kohakade et al., 2017; Sahoo et al., 2019; Pal et al., 2020; Chaudhary et al., 2021). A few promising accessions identified through germplasm evaluation, are IC36245, IC36555, IC36667, IC36577 and IC36604 (Vir and Singh, 2015). Likewise, Sihag et al. (2004) reported that in terms of the traits studied, MH 65, MH 34/66, MH 45, and MH 66 were the most promising genetic stocks, and their use in the moth bean varietal improvement programme could be propitious. Similarly, Kohakade et al. (2017) identified five promising genotypes (DHMB-32, DHMB-26, DHMB-31, DHMB-30 and DHMB-16) under the hybridization programme based on inter-cluster distances, cluster mean and per se performance. Systematic germplasm evaluation studies by various workers have identified valuable trait-specific accessions for moth bean improvement (Table 4). In moth bean breeding programmes, these accessions could be a valuable resource for improving moth bean resistance to major biotic and abiotic stresses and for enhancement of grain yield. Variety development programme in moth bean was probably initiated in the late 1960 or early 1970s (Singh and Mann, 1976). Traditional germplasm-selected varieties were of spreading types that covered the ground like a mat, grew slowly, and matured in 100-120 days. These varieties (Type-I, Type-3, MG-l, Baleswar-12, and so on) were introduced prior to 1970. However, under ideal conditions, these fodder types barely give about 200-300 kg/ha of yield (Kumar, 2002). Further selections from the germplasm resulted in identification of high yielding varieties with long and medium duration (Maru moth) maturity before the inception of AICRP on Arid legumes in 1992 (Table 5). Work on arid legumes *viz.*, cluster bean, moth bean, cowpea and horse gram started in 1992. High-yielding varieties that mature in 75-90 days have thus been developed as a result of continuous, sincere, and deliberate efforts to improve this crop's genetics and yield. As a result, the newer cultivars (Jadia, Jawala, IPCMO-912, and CAZRI moth-1) could produce higher yield while also withstanding the extreme heat and limiting soil moisture conditions. These varieties were also better suited to areas with high levels of rain (250- 350 mm) (Kumar, 2002). In 2003, CAZRI moth-2 (CZM-2) was developed by Central Arid Zone Research Institute, which was the first variety to be released through hybridization and has gained wide popularity in recent years owing to its high yield potential and considerable

TABLE 3 A comparative account of agronomic traits of moth bean.

Traits	Sihag et al., 2004	Singh et al., 2006	Kohakade et al., 2017	Sahoo et al., 2019	Overall range
Days to 50% flowering	37.85 (29.5-47.5)	63.73 (32.0–84.0)	84.93 (43.7-106.7)	46.29 (36.3-50.7)	29.50-106.67
Day to 80% maturity	73.78 (54.5-89.5)	82.84 (57.0-105.0)	135.09 (90.3-158.7)	72.57 (61.7-91.0)	54.50-158.67
Plant height (cm)	51.35 (37.5-94.3)	26.45 (11.7–49.3)	107.44 (11.6-153.1)	32.07 (19.5-40.4)	11.63-153.13
Primary branches/plant	–	6.09 (1.6-12.0)	5.01 (2.3-8.8)	04.37 (3.2-6.0)	1.58-12.00
Clusters per plant	22.23 (4.4-55.1)	20.98 (9.0-73.7)	35.08 (14.8-58.8)	–	4.40-73.67
Pods per plant	62.57 (20.0-129.0)	49.52 (24.8-128.3)	66.05 (20.8-145.2)	98.66 (65.3-148.8)	20.00-148.83
Pod length (cm)	3.51 (2.4-4.0)	3.95 (2.2–5.3)	3.37 (2.9-3.8)	3.99 (3.3-5.5)	2.20-5.51
Seeds per pod	5.74 (4.85-7.0)	6.75 (2.0–10.0)	5.67 (3.2-6.9)	4.96 (3.6-6.9)	2.00-10.00
100-seed weight (g)	2.39 (1.6-3.2)	2.90 (1.5–4.6)	2.32 (1.5-2.9)	2.60 (1.7-3.3)	1.50-4.60
Yield/plant (g)	5.74 (0.9-12.1)	6.17 (1.6–19.1)	6.94 (1.8-20.6)	11.41 (4.3-20.5)	0.94-20.64

10.3389/fpls.2023.1179547

FIGURE 5
Variability in the shape, size and color of *Vigna aconitifolia* seeds conserved in the Indian National Gene bank, New Delhi.

TABLE 4 Potential trait-specific accessions identified in germplasm for utilization in moth bean improvement.

Traits	Materials used	Trait attributes considered	Promising entries	References
Agronomic traits	49 genotypes	Pods per plant, seeds per pod, pod length, hundred seed weight (g), biological yield, harvest index, and seed yield/ plant	IC370469, GP387, IC415116, IC310670, IC329040, RMO225, RMO-257, IC983	Pal et al., 2020
Early maturity	Mutant of moth bean variety RMO-40	Phenological traits like days to initial flowering, days to 50% flowering, days to maturity,	IC432859	Bisht and Singh, 2013
	62 genotypes	Pod length and seeds/pods	MH 65, MH 34, MH 66, and MH 45	Sihag et al., 2004
Heat tolerance	10 germplasm lines & varieties	Relative water content (RWC), membrane stability index (MSI), and proline, protein, and chlorophyll (Chl)	MO 40, Jadia, IC36157, Jwala, and Maru moth	Tiwari et al., 2018
Drought tolerance	Mutant of moth bean variety Jadia	Relative water content (RWC), membrane stability index (MSI), and proline, protein, and chlorophyll (Chl)	IC296803	Bisht and Singh, 2013
	32 accessions	Plant height, root length, branches plant dry mass, root dry mass, chlorophyll & relative water content	IC129177, IC103016, IC415139, IC415155 and IC36157	Malambane and Bhatt, 2014
YMV resistance	204 accessions	Disease intensity (%), Disease incidence (%) & Plant Disease Index (%)	PLMO 12, IC36096, IC415152, IC129177, IC129177,	Meghwal et al., 2015
	11 grain type & 10 fodder type germplasm	YMV disease incidence was recorded on a single-plant basis	013393-C, DMB-118-E	Yaqoob et al., 2015
Leaf crinkle virusresistance	44 accessions	Leaf crinkle disease scoring on a single plant basis	IC39786	Vir and Singh, 2015
Phytochemicals	25 accessions	Protease inhibitors, phytic acid, radical scavenging activity, and tannins	IC39784, IC8851, IC39711 and IC39774	Gupta et al., 2016
	60 accessions	Comprehensive biochemical profiling includes higher contents of magnesium and iron	IC36361, IC-311448, IC415152 and RMO-257	Bhadkaria et al., 2021

TABLE 5 List of moth bean varieties developed through conventional breeding approaches in India.

Variety	Maturity (Days)	Grain yield (Kg/ha)	Salient features
Type-1	120	200-300	It was selected from a local collection, medium-sized brownish-red seed with an average forage production of 10-14 q per ha
Type3	120	370-375	Forage type, selected from a local collection, produces 22-25 q/ha of green forage, has narrowly lobed leaves, and has a canopy that extends outward and trails horizontally.
MG-l	110-115	350-450	Selected from a local collection, extremely vulnerable to YMV, plants are taller (45-55 cm), harvest Index is poor (10-12 percent)
Baleshwar-12	110-115	400-475	Selected from a local collection; plants are taller; green fodder yields 15–17 q/ha; seeds are brown and medium size; extremely vulnerable to YMV.
Jadia	85-90	450-500	Selected from a local collection, has a spreading growth habit, dark brown, medium bold seeds (l00-seed weight of 2.5-2.5g), is prone to yellow mosaic virus, and produces 10–12 q/ha of green fodder production.
Jawala	80-90	500-550	Selected from a local collection, resistant to YMV, and with an average fodder yield of 17-18 q/ha.
Maru moth	80-85	500-550	Selected from the local collection, this variety is a semi-spreading type that performs well in planting situations that are less affected by the *Cercospora* leaf spot disease. It is also suitable for intercropping.
IPCMO-800	80-85	450-500	Selected from a local collection; vining type; has broad, deeply lobed leaves; seed protein is 22–24 percent; harvest index is 20–25.
IPCMO 880	80-85	450-500	Pure line selection from a local collection in the Jhunjhunu district; matures in 90-100 days; produces 3-4 pods per cluster with medium to bold seed (2.8-3.1 g/100 seeds); seed yield: 4-5 q/ha.
IPCMO-912	75-85	400-500	This variety shows field tolerance to YMV and Bacterial blight, narrow leaflets.
CAZRI Moth-1	72-75	400-500	It is semi-spreading and dual-purpose variety responding to inputs, Field resistance to YMV, bold grain with 25% protein
CAZRI Moth-2	75-85	1000-1200	This is the first variety produced through hybridization (RMO-40 x Jadia); erect plant habit and profuse podding.

(Source: Kumar (2002); Solanki et al. (2018)).

podding (Solanki et al., 2018). In addition to these, its introduction to the marginal irrigated farming system, development of early maturing, photosynthetically efficient plants, breed of input responsive genotypes, development of varieties suitable for inter/mixed cropping system, identification of 'dual types' (high grain as well as fodder yielder) and development of varieties resistant to biotic and abiotic factors are some of the priority areas which should be taken care of in future genetic improvement programmes on moth bean.

5.3 Interspecific hybridization

In India and other Asian countries, *Vigna* species constitute an economically significant group of 10 cultivated crops (domesticated species) and more than 100 wild species (Schrire et al., 2005; Bisht and Singh, 2013; Takahashi et al., 2016). Moth bean is the most interesting of the Asian Vigna species that have been domesticated as it is highly resistant to drought and heat (Tomooka et al., 2014). The wild ancestor of the moth bean hasn't been precisely described, but it is said to have occurred in India (Arora and Nayar, 1984). Understanding the genetic basis of crop domestication aids in the identification of beneficial genes (or genes) in wild relatives of cultivated crops that can be used in plant breeding. In addition, as a reliable strategy for expanding a crop species' genetic base, interspecific hybridization has successfully introduced beneficial traits from wild relatives into closely related crops (Singh, 1990). Varying degrees of success in interspecific hybridization of *Vigna* has been reported (Chen et al., 1983; Singh, 1990; Pandiyan et al., 2012; Tomooka et al., 2014; Basavaraja et al., 2019). Cross-compatibility between *V. aconitifolia* and other relative species hasn't been studied much, but it has been crossed as a seed parent with *V. trilobata*, and *V. trilobata* has been crossed with success, as a pollen parent, with *V. mungo*, *V. radiata*, and *V. aconitifolia*, but the reciprocal hasn't reportedly worked (Bisht et al., 2005). The interspecific hybrids between *V. aconitifolia* and *V. trilobata* produced viable seeds. The F_1 hybrids had 5.7% pollen fertility, but complete seed sterility. Colchicine-induced amphiploids had 89.7% pollen fertility and such plants had fairly regular meiosis, suggesting the possibility of such crosses. Hybrid sterility is segregational in nature (Biswas and Dana, 1976b).

5.4 Tissue culture approach

Through interspecific hybridization, efforts have been made in the recent past to develop disease-resistant as well as high-yielding varieties of moth bean. However, interspecific cross incompatibility and hybrid sterility are some significant hurdles for moth bean variety development programme. Consequently, it is conceivable that moth bean quality and yield can be improved by using a combination of tissue culture techniques and traditional breeding

methods. Tissue culture is a potent method that enables scientists to cultivate and manipulate plants in a sterile environment. It is an advantageous method for plant breeding. Therefore, cell and tissue culture application to moth bean assumes a special significance. However, the lack of or low rate of plant differentiation from cultured cells is one of the major obstacles towards the utilization of cell and tissue culture techniques in grain legume breeding programmes (Bajaj and Gosal, 1981; Gresshoff and Mohapatra, 1982; Mroginski and Kartha, 1984; Saxena, 2004). Several studies have reported that many legumes have now been successfully cultured and regenerated as whole plants. Bhargava and Chandra (1983) initiated callus cultures of two cultivars of *Vigna aconitifolia* (moth bean) (IPCM0-926, RDM-120) and studied their growth and differentiation. Likewise, Godbole et al. (1984) optimized the cultural conditions for whole moth bean plant regeneration from explant shoot apices and cotyledons, callus cultures derived from explant shoot apices, and field transfer of rooted plantlets. As a practical issue, the development of shoots and subsequent plantlets from excised roots is of interest because it provides a new source of somaclonal variation. Gill and Eapen (1986) attempted to develop cell and tissue cultures in moth bean, whereby a successful attempt was made in moth bean plant regeneration from hypocotyl protoplasts. Eapen and Gill (1986) were able to be regenerate plants from root explants of moth bean by culturing on Murashiage and Skoog's basal medium without any phytohormone, but the addition of cytokinins increased the frequency of plant regeneration.

Protoplasts, on the other hand, are also an excellent source of inducing variation in *Vigna*. Protoplasts from this important group of economic crops have been isolated and regenerated into cell colonies and callus tissues. Callus cultures derived from shoot apices of moth bean provided a high yield of viable moth bean protoplasts as a starting point for the development of an efficient and reproducible moth bean protoplast isolation and regeneration method (Krishnamurthy et al., 1984). Similarly, Arya et al. (1990) regenerated plants from leaves' protoplasts of *Vigna aconitifolia* 'Jadia' and reported that protoplasts from highly homozygous legumes are of application in moth bean improvement. Kumar et al. (1988) demonstrated successful plant regeneration from cell suspension cultures lines of *Vigna aconitifolia* in L-6 medium containing 44.5 M 2,4-D, which resulted in the highest growth rate. Till date, no systematic study has been reported in moth bean on the effects of cultivar, basal medium, different medium combinations (Murashiage and Skoog medium (MS), 2,4-D & 6-Benzylaminopurine (BAP) and interactions on callus induction, callus propagation and subsequent plant regeneration (Bhargava and Chandra, 1983; Bhargava and Chandra, 1989). Furthermore, Jangid et al. (2010) studied factors affecting callus induction in moth bean and found significant differences, for both callus initiation days and callus fresh weight, between the varieties, explants, medium combinations, and their two- and three-way interactions while also reporting that MS medium supplemented with 1.0 mg-1, 2-D was the best medium for maintaining callus in all of the tested types. There have been numerous attempts made to develop moth bean regeneration protocols *in vitro* (Arya et al., 1990; Chandra and Pental, 2003; Kaviraj et al., 2006). When tested on Indian moth varieties, the *in vitro* regeneration techniques developed by these researchers failed to deliver the expected outcomes. Hence, Choudhary et al. (2009) developed an effective method for regenerating moth bean using somatic embryos and *in vitro*-grown plantlets that resulted in well-formed roots being successfully hardened and established in soil. Recently, regeneration potential, gene expression, and genetic stability were found to be affected by heat treatment in moth bean (Jangid et al., 2010). Additionally, the response of tissue to callus formation was studied in explants that had been heated at 37, 42, or 47°C for 10 minutes (Sharma et al., 2018). The results showed that most of the heat treatments slowed down regeneration, and a few polypeptides in the protein profile of the callus were both up-regulated and down-regulated by the heat treatments. Understanding the mechanism of flowering regulation in *in-vitro* cultures is practical and theoretically relevant. The effect of growth regulators such as abscisic acid and proline on *in vitro* flowering manipulation in moth bean has been investigated by Saxena et al., 2008. On the contrary, moth bean presents an efficient system to be used for *in vitro* mutagenesis because it exhibits an efficient regenerability from various explants and bear seeds under *in vitro* conditions (Saxena et al., 2006). Likewise, Narayan et al. (2015) demonstrated that gamma radiation exposure increased the number of shoot bud primordia and the number of shoots per explant in callus generated from primary leaves of seedlings of moth bean seeds. Such explants also resulted in a higher number of flowering plantlets under *in vitro* growth conditions. Considerable success has been achieved in moth bean *in vitro* culture plant regeneration, and this progress offers an alternative option, over conventional breeding or interspecific hybridization, for the improvement of moth bean.

5.5 Mutation breeding

Mutation is a highly effective technique to generate the desired variation in crop plants. It has been extensively used in many of the world's most important crops, including wheat, rice, pulses, millets, and oilseeds, and several original and review studies show its effectiveness (Khadke and Kothekar, 2011; Raina et al., 2016; Jegadeesan and Reddy, 2018; Pandit et al., 2021). Work on the varietal improvement of moth bean crop is minimal; a mutation breeding programme in moth bean was taken up in 1983, to isolating mutants with high yield potential, early maturing and at the same time possessing tolerance to drought conditions. Henry and Daulay (1983) researchers examined the performance of induced mutants in moth bean, treating seeds of the Jadia variety with aqueous solutions of EMS at doses ranging from 0 to 3 percent. In the M_2 generation, they found 25 mutants with higher pods than their parents. The mutants were further carried forward up to M_5 and M_6 generation and tested under dry land conditions for yield performance, along with other high-yielding varieties. On the contrary, to improve arid legumes, a mutation breeding technique

utilizing chemical and physical mutagens was undertaken, (Ramkrishna et al., 2008; Jain et al., 2013). The most commonly used mutagens, namely, Ethyl methane sulphoanate (EMS), Methyl methane sulphonate (MMS), Sodium azide (SA) and Hydroxylamine (HA) were used in moth mutation breeding and the effectiveness of the mutagens was tested on three different bean crops. Moth bean genotypes had the highest efficiency, followed by cluster bean, while mung bean genotypes had the lowest efficiency (Sharma and Kakani, 2002; Ramkrishna et al., 2008). Four moth bean varieties, RMO-40, RMO-257, Jwala and CZM-1, were tested for their response to EMS, MMS, SA, and HA, induced biological damage, polygenic variability, and comparative mutagenic effectiveness and efficiency. This was performed to better understand how different moth bean varieties respond to different types and treatments of chemical mutagens. Subsequently, the mutagenic progenies, concerning yield and yield attributes in M_2 and M_3 generation, were evaluated. Treatment efficacy declined in the following order: MMS > SA> EMS > HA. The moth bean variety, RMO-40, was shown to have the maximum mutagenesis effectiveness with a concentration of 1.0 mM Sodium azide. In M_3 generation, 15 progenies showed higher seed yield with superior magnitude of yield contributing traits than the best check, RMO-225 (Jain et al., 2013).

Many national and state agricultural universities, including Bhabha Atomic Research Centre (BARC), Mumbai (India), have been doing intensive mutation research since decades by employing X-rays, gamma rays, fast and thermal neutrons in order to generate genetic variability in these arid legume crops (D'Souza et al., 2009). As a result, research on induced mutations in oilseeds and legumes has remained at the leading edge of Indian agricultural research for developing popular varieties with higher productivity in these crops. Khadke and Kothekar (2018) studied the effect of EMS and SA on moth bean trypsin inhibitor content. On electrophoresis, various viable and micro mutants of moth bean were found to have between three and seven iso-inhibitors of trypsin. Some viable mutants had considerable differences in their TI profiles when compared to the control. The TI content of these mutants reduced by 25 to 45 percent, and the seed protein content of micro mutants and viable mutants differed significantly. In 1994, the release of the first early maturing mutant variety, RMO-40, sparked an interest in mutagenesis, leading to the development of a series of short-duration varieties, including RMO-257, RMO-225, RMO-435, RMO-423, and RMO-2251 (Table 6). All of these cultivars were early maturing (60-67 days), exhibited completely transformed plant types, were semi-erect to erect, and had synchronous maturation. Due to short growing seasons, these varieties could successfully fend off drought problems and circumvent devastating diseases like YMD (yellow mosaic disease) and *Cercospora* leaf spot. Such varieties are more suited to low rainfall (200-250 mm) and shorter growing season. Varieties developed through mutation breeding could be direct mutants with good agro-morphological traits or the resultants of hybridization with mutants for mutant trait introgression in more adaptable and elite genetic backgrounds Pulses with narrow genetic bases have enormous potential for increasing genetic variability through a combination of mutation breeding and conventional breeding approaches. This could lead to the development of elite varieties that are suitable for various agro-climatic zones, thereby improving pulses production subsequently leading to nutritional security.

TABLE 6 Popular high yielding moth bean mutant varieties under cultivation in India.

Mutant variety	Maturity (days)	Yield (kg/ha)	Special features	References
RMO-40 (Rajasthan Moth 40)	62-65	600 to 900	This is an extra early maturing, escapes drought, short stature, and non-spreading variety with synchronous maturity.	Sharma and Kakani, 2002
RMO-225 (Maru vardan)	64-67	600 to 750	This variety has a short duration and a high level of resistance to the yellow mosaic virus.	Kumar, 2005
RMO-435	67-70	600 to 800	This variety is erect with medium duration, has good yield potential, leaves green colored, mutant from RMO-40and has fodder value. Therefore, it is suitable for all moth-growing areas of the country.	Kumar, 2005
RMO-257	65-67	500 to 550	This mutant from RMO-225 variety is utilized for both grain and fodder production. It is a variety that matures early and has a moderate level of resistance to bacterial blight and YMV. Its flowers are arranged in clusters and each one has a little petiole.	Jain et al., 2013; Sharma et al., 2015
RMO-423	67-70	550 to 600	The cultivar can be used for either seed or fodder cultivation. It can also resistant attacks from insects and the yellow mosaic virus.	Jain et al., 2013
RMB 25	67-70	600 to 700	It gives a higher seed (28.6%) and fodder yield (15.4%) than RMO-225 and RMO-257. The variety has field resistance to yellow mosaic virus and high protein content.	Sharma et al., 2015
RMO-2251 (Maru dhar)	63-67	600 to 650	It is a short duration variety with early maturity, erect plant type and average incidence of YMV.	Solanki et al., 2018
CAZRI Moth-3	62-64	550-750	Escapes YMV. Early maturing and erect. Heavy pod-bearing. Drought tolerant.	Solanki et al., 2018

6 Abiotic and biotic stress tolerance and adaptive traits in moth bean

6.1 Heat stress tolerance

Moth bean is a highly adaptable annual legume that thrives in arid and warm climates, making it one of the most temperature-resistant legumes in arid climates (Sachdeva et al., 2016). It has a deep and fast-penetrating root system, can thrive in open fields for up to 30-40 days, and can withstand temperatures of more than 45°C. By virtue of its adaptive features, whereby it can sustain harsh environmental conditions, moth bean is widely recognized as an arid legume. To mitigate heat stress, moth bean has developed morphological accommodations such prostrate development thereby shading the soil to reduce temperature, large leaves with dense canopy for transpirational cooling, higher biomass but poor partitioning, indeterminate growth, and a fast-penetrating tap root system (Sharma et al., 2014; Harsh et al., 2016). The crop appears to have a genetic buffer that allows it to quickly adapt to rapidly changing moisture conditions as well as deprived and hot environments. Temperature stress, contrarily, has been shown to speed up plant growth in moth bean begetting reduction in both vegetative and grain yields (Khadke et al., 2011; Sharma et al., 2014). Similar results were observed in a study based on the induction of thermo-tolerance for deciphering the efficaciousness of heat acclimation in nine different varieties of moth bean, which showed that all nine varieties could tolerate a sudden temperature rise of 42°C, and that the development of thermo-tolerance was linked to the induction of peroxidase (POX), ascorbic peroxidase (APOX), and catalase (CAT) enzymes (Sharma et al., 2014; Harsh et al., 2016). Likewise, the accumulation of total sugar and proline, as well as an increase in the activity of CAT, GPOX (glutathione peroxidase), and SOD (superoxide dismutase), was seen in thirty-seven genotypes of moth bean under heat stress in earlier investigations. (Harsh et al., 2016; Tiwari et al., 2018) and by that mutagenesis was suggested as an effective means of improving thermo-tolerance by altering the osmo-protectants and anti-oxidative enzymes that go along with it.

6.2 Drought tolerance

As a drought-tolerant crop, moth bean has had a long history of success in rain-fed arid environments. Pre-meditated and need-based efforts have been made for decades to increase the productivity and adaptability of this drought-resistant crop in harsher and more hostile conditions. There has been paucity in systematic research on the physiological basis of yield in rain-fed conditions and the mechanism for drought resistance in this crop, despite its economic importance. Physiological investigations of yielding ability in moth bean types under rain-fed conditions established that the crop is a poor seed producer despite producing enough dry matter. Leaf water potential, osmotic potential, and pressure potential were measured at regular intervals over a period of time. There were noticeable changes in transpiration amongst the different varieties (Srivastava and Soni, 1995). It is quite difficult to predict the outcome of crop production in arid and semi-arid areas. However, the choice of proper plant type and adopting an appropriate cropping system may provide yield stability in the region. The importance of the adoption of suitable genotypes and cropping systems has been emphasized by several investigators (Singh and Reddy, 1988; Lawn, 1989). Higher photosynthetic rates and more effective water use have been linked to the increased performance of mixed crops under low-rainfall conditions. Kathju et al. (2003) suggested that under both low and high rainfall situations, planting early and late genotypes in alternate rows at a 1:1 ratio is an ideal strategy for maintaining the production stability of moth bean. Priyanka et al. (2011) investigated the effect that a low water potential, caused by PEG (polyethylene glycol), has on growth, sugar content, and enzymes related to stress (such as catalase, GPOX). Additionally, SDS-PAGE (Sodium dodecyl-sulfate polyacrylamide gel electrophoresis) analysis was performed on seedlings of twelve moth bean genotypes and traits such as seed set and abscisic acid concentration in the pods as well as root traits have been proven to increase the tolerance of moth bean to drought (Kumar et al., 2019). Accordingly, Garg et al. (2001) reported that on rise in water stress in three moth bean genotypes (RMO-40, Maru moth, and CZM-32 E), net photosynthetic rate and starch and soluble protein content dropped besides a marked decline in the nitrate reductase activity of the plant. Discordantly, there was a build-up of reducing and soluble sugars as well as proline content at pre-flowering phase of the genotypes under study. A germplasm collection comprising of 32 genotypes was evaluated under both dry (terminal drought) and wet (rain-fed) environments. The results exhibited significant differences for genotypes, treatment (stress), and the GxS effects. Also, the differences in the amount of chlorophyll were significant for genotype and combination of genotype and stress treatment. The treatment effect showed a significant difference and led to the identification of two varieties (Maru moth and Jadia) and five accessions (IC 129177, IC 103016, IC 415139, IC 415155, and IC 36157) that were drought tolerant (Malambane and Bhatt, 2014). In another study, Sachdeva et al. (2016) reported that three cultivars (IC103016, IC36011 and IC36157) were more drought tolerant because they sustained greater RWC (Relative Water Content) and MSI (Membrane Stability Index) along with plant height, leaf area, seed weight and plant dry mass besides higher levels of proline accumulation. Recognized as having a primitive plant type, moth bean has evolved to survive but not for increased productivity. In this context, early partitioning, early maturation, and semi-erect to erect growth habit types may be preferred to the conventional sort. Consequently, such advantageous plant traits will not only lead to an increase in yield but also encourage the cultivation of this crop in new and unexplored areas.

6.3 Salinity tolerance

Regulation of proline biosynthesis as well as its degradation, uptake and transport are one of the key mechanisms of the plant defense system, which plays a compelling role in the survival of

osmotically stressed plants (Stewart and Lee, 1981; Verma et al., 1992). The gene encoding the enzyme, ornithine delta-aminotransferase (delta-OAT) that is involved in proline biosynthesis pathway, was isolated from moth bean cDNA library by complementation of *Escherichia coli* proBA mutant auxotroph (Delauney et al., 1993). Zhu et al. (1998) transformed rice var. Nipponbare with P5SC gene obtained from moth bean and increased biomass in salt and water stress conditions. Similarly, when the rice cultivar, ADT 43, was transformed with Δ1-pyrroline-5-carboxylate synthetase (P5CS) gene obtained from moth bean and it was observed that under 200 mM NaCl, the transformed plants showed better plant growth and biomass production while the control plants died within 10 days (Karthikeyan et al., 2011). In another study, the overexpression of P5CS gene derived from moth bean in rice cultivar IR-50, showed tolerance to high salt (200 mM NaCl) conditions (Anoop and Gupta, 2003). Surekha et al. (2014) used *P5CSF129A*, a mutagenized P5CS gene obtained from moth bean, in combination with CaMV35S, a constitutive promoter, for the genetic transformation of *Cajanus cajan* to enhance proline accumulation and salt tolerance. Moreover, the P5CS gene from moth bean was transplanted into the hybrid larch (Larix x leptoeuropaea (Dengler)) to improve the productivity of Larch, a tree species that thrives under stresses like cold, salt, and frost, by which the transgenic tissues exhibited a thirtyfold increase in proline concentration (Gleeson et al., 2005).

6.4 YMV disease resistance

Biotic stresses such as Yellow Mosaic Disease (YMD) and the pod borer extend major production challenges for various *Vigna* crops. Like other *Vigna* crops, moth bean has not received the proper scientific attention towards management of biotic stresses, and among many, the crop requires particular consideration of YMV and Bacterial Leaf Spot disease. The *Mung bean yellow mosaic virus* (MYMV) is resistant to mung bean's morphological and biochemical diversity, according to recent investigations. In terms of resistance, morphological characteristics such as leaf thickness and trichome density are varied. Leaf thickness was found to be higher in the MYMV-resistant genotypes compared to the highly sensitive ones. Similarly, high levels of trichome density were found in resistant genotypes, compared to low levels seen in the most susceptible ones (Patel et al., 2013; Mantesh et al., 2020). In another study, 204 moth bean lines were screened for resistance to the *Mung bean yellow mosaic virus* (MYMV). Thirteen of the 204 accessions studied showed resistance or tolerance to MYMV in the field (Meghwal et al., 2015). Stress tolerance genes must be identified which requires substantial information on the control of gene expression as well as the biochemical activity of specific proteins involved in the process of tolerance. According to Tiwari et al. (2018), the genes for stress-enzymes such as Catalase, Cyt P450 monooxygenase, heat shock proteins (HSP 90 and HSP 70), oxidoreductase, protein kinases, dehydration-responsive protein (DRP), universal stress proteins, and ferridoxin NADH oxidase were over-expressed in stressed samples. Additionally, expression profiling revealed ten transcripts to be up-regulated and 41 to be down-regulated, whereas 490 exhibited no significant change under moisture stress (Tiwari et al., 2018).

6.5 Insect resistance

Specific management measures to reduce yield losses due to insect pests, such as Jassids, white flies, pulse beetles, white grubs, and other storage pests is imperative *Callosobruchus chinensis* L. and *Callosobruchus maculatus* L. are major bruchid pests of most *Vigna* species, particularly cultivated ones. In light of this, bruchid resistance enhancement is prioritized in all *Vigna* crop breeding programmes. Plant breeders have long aimed to improve the bruchid resistance of cowpeas, mung beans, adzuki beans, and black gram. These *Vigna* crops have sources of bruchid resistance, but the resistant germplasm is limited. A wild moth bean accession, TN67, was recently discovered to be extremely resistant to *C. chinensis* (Somta et al., 2018). For *Vigna* crop breeding, wild moth bean's resistance gene(s), including moth bean with *C. chinensis* resistance could be highly favorable (Somta et al., 2018).

7 Biotechnological interventions

7.1 Genetic linkage map and QTL mapping for agronomic traits

The work done in moth bean, with respect to breeding, is exiguous, therefore, mapping populations, core sets or trait specific reference sets are lacking in the crop. In addition, studies analyzing the genetic variability in *V. aconitifolia* germplasm, by way of molecular techniques like DNA markers, have been sparse (Sharma et al., 2021). Nevertheless, recently, there has been some progress. Despite having the potentiality to be a new food crop of the future owing to being a source of genes for resistance against biotic and abiotic stresses, the genetics of the domestication process in moth bean is not known. Its domestication will principally involve phenotypic changes, including reduction of seed dormancy and pod shattering, increased organ size and earlier flowering and maturity, which needs extensive knowledge of its genomic resources that are currently scarce. Recently, Yundaeng et al. (2019) constructed a genetic linkage map for moth bean, based on an F_2 population of 188 individuals produced from a cross of wild moth bean (TN67) and cultivated moth bean (ICPMO056), and utilized it for the identification of quantitative trait loci (QTL) for domestication-related traits that can be used for genetic improvement of the moth bean and related *Vigna* species. The genetic linkage map comprised of 11 linkage groups (LGs) of 172 simple sequence repeat (SSR) markers that spanned a total length of 1016.8 centimorgan (cM), with an average marker distance of 7.34 cM. Additionally, a high genome synteny was observed between moth bean and other orphan legumes; mung bean (*Vigna radiata*), adzuki bean (*Vigna angularis*), rice bean (*Vigna umbellata*), and yard long bean (*Vigna unguiculata*) based on comparative genome analysis. A total of 50 QTLs and 3 genes associated with 20 domestication-related traits were identified.

Most of the QTLs belonged to five LGs (1, 2, 4, 7, and 10). Pod shattering, trailing plant canopy, seed dormancy, poor yield potential, less responsive to input resources, late flowering and maturity are the key attributes of moth bean which indicates that the crop is still under active domestication. In another study, Somta et al. (2018) developed QTL mapping in F_2 population (188 plants) of moth bean by crossing resistant accession, 'TN67' (male parent) and susceptible accession, 'IPCMO056' (female parent) for seed resistance to adzuki bean weevil (*Callosobruchus chinensis* L.). Segregation analysis suggested that *C. chinensis* resistance in TN76 is controlled by a single dominant gene, designated as *Rcc*. QTL analysis revealed one principal and one modifying QTL for the resistance, named *qVacBrc2.1* and *qVacBrc5.1* respectively. *qVacBrc2.1* was located on linkage group 2 between simple sequence repeat markers CEDG261 and DMB-SSR160 and accounted for 50.41 to 64.23% of resistance-related traits, depending on the trait and population, while *qVacBrc5.1* was detected in only one population. Somta et al. (2018), suggested that markers CEDG261 and DMB-SSR160 should be useful for marker-assisted selection for *C. chinensis* resistance in moth bean. A list of genetic linkage maps in *Vigna aconitifolia* is depicted in Table 7. Oliveira et al. (2020) presented the first physical map of moth bean and compared chromosomes with other *Vigna* and *Phaseolus* species. As with Yundaeng et al. (2019), a high magnitude of genome synteny was observed between moth bean and other related pulse crops such as mung bean (*Vigna radiata*), adzuki bean (*V. angularis*), rice bean (*V. umbellata*), yard long bean (*V. unguiculata*) as well as common bean (*Phaseolus vulgaris*) (Oliveira et al., 2020). As a result, molecular markers and genomic resources developed in these crops will be highly useful in moth bean crop improvement programmes.

7.2 Transcriptome analysis

An extensive study was done to identify genes associated with moisture stress tolerance utilizing differential transcriptome assembly under stressed and non-stressed environments (Tiwari et al., 2018). They have identified 51 differentially expressing transcripts along with 1287 useful SSRs (Simple sequence repeats). Another study was done to identify genes associated with heat stress utilizing forward suppression subtractive hybridization (SSH) cDNA library of heat tolerant cultivar RMO-40 (Gurjar et al., 2014). A total 125 unigenes, out of which 21 were novel to moth bean, were identified. Functional annotation of ESTs (expressed sequence tags) led to the identification of enzymes and heat-shock proteins involved in plant defense against heat stress. In a first ever attempt to identify heat stress responsive genes in moth bean, an elevated temperature of 42°C for 5 min was given to a heat tolerant moth bean cultivar RMO-40 (Rampuria et al., 2012). Utilizing SSH cDNA library, 488 unigenes were identified that were constructed from 738 ESTs. Annotation and semi-quantitative PCR analysis indicated 20 signaling genes and 16 transcription factors associated with heat stress in moth bean. Recently, a number of transcriptome studies in other *Vigna* species have been added to the public data bases. Transcriptome derived resources, particularly EST-SSR markers, from other *Vigna* species can be utilized in genetic mapping, marker trait association and marker assisted selection for improvement programmes in moth bean.

7.3 Genetic transformation

Direct transformation was found to be an efficient transformation method in moth bean (Köhler et al., 1987a). Protoplast was used to transform moth bean by using PEG (polyethylene glycol) treatment and electroporation method. The plant genotype also has the role in the transformation success rate (Köhler et al., 1987b). In another transformation method, 60% plating efficiency was observed using co-cultivation of protoplast cells with *Agrobacterium tumefaciens* containing the Ti plasmid derivative pGV38501103 (Eapen et al., 1987). Particle bombardment method was also successfully used for direct gene transfer in moth bean using mature embryos (Bhargava and Smigocki, 1994) and hypocotyl explants (Kamble et al., 2003). A stable transformation was reported by co-cultivation of primary leaves and cotyledonary nodes, as well as vacuum infiltration into cotyledonary nodes of moth bean (cultivars Jadia and Jawala), with *Agrobacterium tumefaciens* strain EHA105 accommodating the binary vector p35SGUSINT (Kumar et al., 2008; Kumar et al., 2010). Unlike the earlier studies, genotypic effect was not observed in transformation frequency in this study. However, other studies show that the methodology and genotypes influence the transformation frequency (Kamble et al., 2003).

TABLE 7 List of genetic linkage maps in *Vigna aconitifolia*.

Name of population	Trait	Gene/QTLs	PV explained by the QTLs (%)	Linkage group	References
IPCMO056 × TN67	*Callosobruchus chinensis* L. resistance	*Rcc*	-	-	Somta et al., 2018
		qVacBrc2.1	50.41 - 64.23	2	
		qVacBrc5.1	12.19	5	
TN67 × ICPMO056	Domestication related quantitative traits	*Pdt1.1-, Npdd1.1-, Sdwa1.1+, Sd100wt2.1+, Sdl2.1+*, etc.	3.72 - 68.67	1, 2, 3, 4, 6, 7, 8, 9, 10	Yundaeng et al., 2019

10.3389/fpls.2023.1179547

8 Future perspectives

The impact of drastic changes in global environment on crop productivity has been recognized by the agricultural sector, resulting in increased awareness about Climate Smart Agriculture (CSA) approach in order to increase food security and mitigate the effects of climate change. Orphan crops, though neglected, are well acclimatized to current conditions. Moth bean has evolved and is adapted to hot arid and semi-arid agro-climatic conditions, which makes it a relevant crop amid such challenging situations, with the potential to serve as an alternative to major legume crops. However, owing to its agronomic constraints, cultivation of the crop is still confined to traditional growing areas. Moth bean has not been fully exploited and has also lagged behind other pulses in terms of the

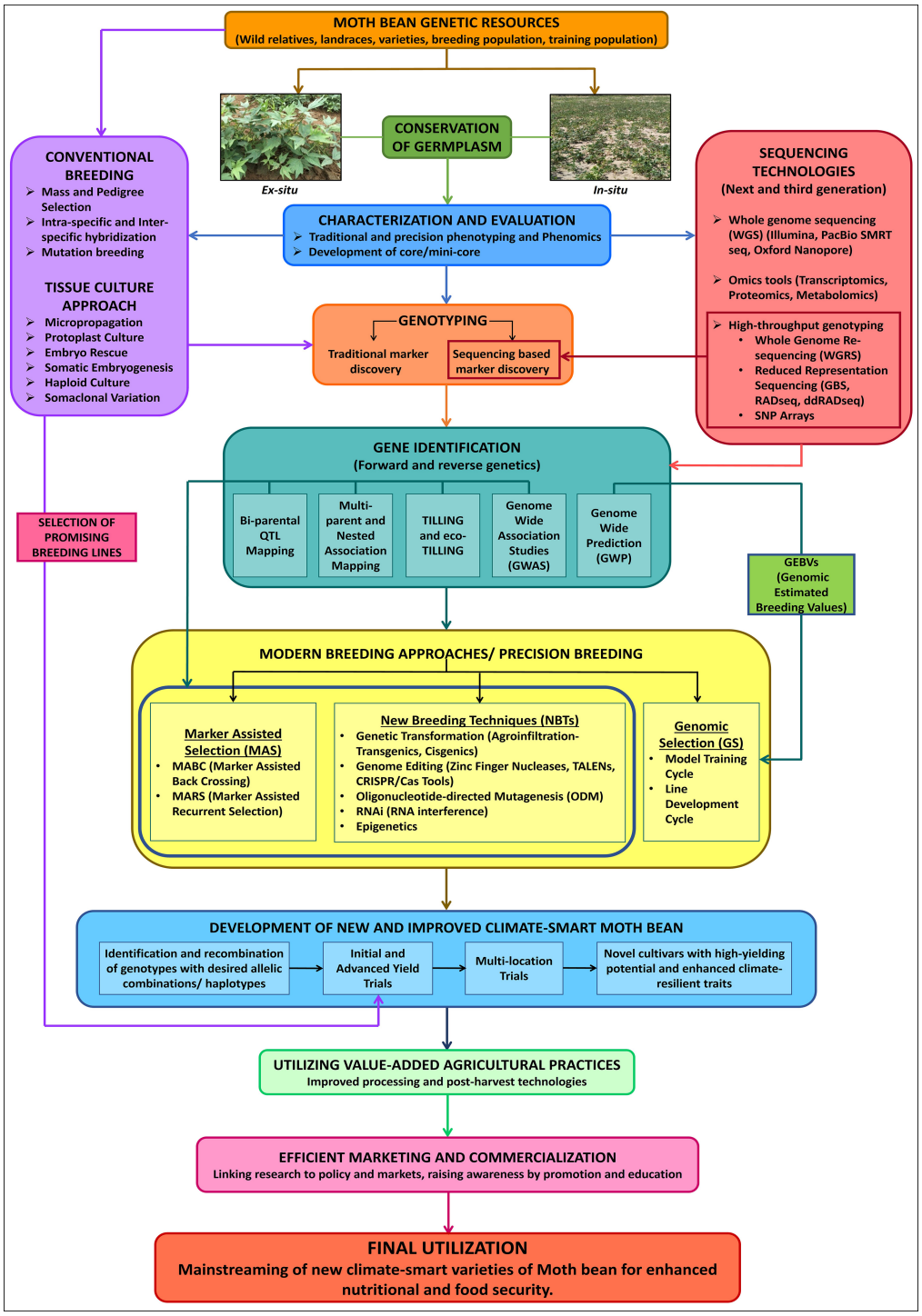


FIGURE 6
A blueprint for the mainstreaming of moth bean to develop climate-smart varieties.

quantum of developed genetic and genomic resources. Since there is limited number of genetic resources, considerable efforts are needed to carry out extensive explorations as well as characterization and evaluation of the collected moth bean germplasm for employing the genetic variability for high yield potential, yield stability, improved salinity tolerance, multi-disease and pest resistance as well as enhanced nutrition. With respect to genomic resources, extensive efforts are needed for development of reference gold standard genome since it is fundamental to understand the genome composition and gene repertory of the crop. Further, whole genome resequencing (WGRS), transcripts and genome-wide SSR and SNP markers are required for the identification and annotation of novel genetic polymorphism/candidate genes which are responsible for different agro-morphological, biotic and abiotic stresses as well as biofortification traits, along with being crucial for the construction of genetic linkage maps and physical maps in the future. Therefore, by utilizing recent advances in breeding methods such as haplotype-based breeding, precision-based phenotyping and high throughput genotyping, accompanied with bioinformatic resources, in association with available genetic diversity can help in accelerating domestication in moth bean and expediting its productivity. Identifying pan-genome sequence variation associated with advantageous traits and utilizing them for development of varieties will have a greater impact. Neo-domestication of stress adapted wild species using mutation breeding and TILLING and *de novo* domestication through new breeding techniques (NBTs) including genome editing tools like CRISPR/Cas could also help in development of novel plants with desired traits (Tomooka et al., 2014; Zsögön et al., 2018). Additionally, utilization of innovative value-added agricultural practices besides adopting effective marketing strategies can be instrumental in the promotion of moth bean as a potential major legume crop (Figure 6). Concentrated, continuous and coordinated efforts are required, from both the scientific community and policy-makers, for implementing these approaches which will not only aid in the development of high yielding and climate-resilient varieties, but also in the mainstreaming of moth bean for re-diversification of global food systems. Finally, when these challenges will be addressed, moth bean, being nutritionally sound and environmentally hardy, is going to be a hopeful climate-smart legume crop for sustainable food and nutritional security across the globe.

Author contributions

Conceptualization: KC and RC. Writing – Original draft preparation: KC, G, BT, and RC. Writing – Review and editing: RC and JR. All authors contributed to the article and approved the submitted version.

Funding

This research was conducted with funding provided under the UN Environment implanted GEF project 'Mainstreaming agricultural biodiversity conservation and utilization in agricultural sector to ensure ecosystem services and reduce vulnerability in India (Project Code: A 1265)' being executed jointly by Alliance of Bioversity International and CIAT, and the Indian Council of Agricultural Research (ICAR).

Acknowledgments

The authors duly acknowledge the ICAR-National Bureau of Plant Genetic Resources, New Delhi; the ICAR-Indian Institute of Pulses Research, Kanpur, India; and the Alliance of Biodiversity International and CIAT, India Office, NAS Complex, New Delhi for providing the necessary support.

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

Adsule, R. N. (1996). Moth bean (*Vigna aconitifolia* (Jacq.) marechal). *Food feed legumes oilseeds* (New York, NY: Springer), 203–205. doi: 10.1007/978-1-4613-0433-3_21

A. Agrawal, S. Rajkumar, J. C. Rana, R. K. Tyagi, O. P. Yadav, K. Singh and R. S. Paroda (Eds.) (2019). *Proceedings and recommendations of the satellite symposium on 'Dryland agrobiodiversity for adaptation to climate change* Vol. 87 (Jodhpur, Rajasthan, India: Indian Society of Plant Genetic Resources, New Delhi, Bioversity International, New Delhi, Asia Pacific Association of Agricultural Research Institutes, Bangkok).

Anoop, N., and Gupta, A. K. (2003). Transgenic indica rice cv IR-50 over-expressing *Vigna aconitifolia* Δ1-pyrroline-5-carboxylate synthetase DNA shows tolerance to high salt. *J. Plant Biochem. Biotechnol.* 12 (2), 109–116. doi: 10.1007/BF03263170

Arora, R. K. (1985). Diversity and collection of wild vigna species in India. *Plant Genet. Resour. Newslett.* (IBPGR/FAO). (63), 26–33.

Arora, R. K., and Nayar, E. R. (1984). *Wild relatives of Crop Plants in India*. New Delhi: National Bureau of Plant Genetic Resources.

Arya, I. D., Arya, S., Rao, D. V., and Shekhawat, N. S. (1990). Variation amongst protoplast-derived moth bean *Vigna aconitifolia* plants. *Euphytica* 47 (1), 33–38. doi: 10.1007/BF00040360

Asthana, A. N. (1998). Pulse crops research in India. *Indian J. Agric. Sci.* 68 (8), 448–452.

Bajaj, Y. P., and Gosal, S. S. (1981). "Induction of genetic variability in grain-legumes through tissue culture," in *Tissue culture of economically important plants: proceedings of the international symposium held at the botany department, national university of Singapore* (Singapore: COSTED and ANBS (Asian Network for Biological Sciences), 25–41.

Baptista, A., Pinho, O., Pinto, E., Casal, S., Mota, C., and Ferreira, I. M. (2017). Characterization of protein and fat composition of seeds from common beans (*Phaseolus vulgaris* l.), cowpea (*Vigna unguiculata* l. walp) and bambara groundnuts (*Vigna subterranea* l. verdc) from Mozambique. *J. Food Meas. Charact.* 11, 442–450. doi: 10.1007/s11694-016-9412-2

Basavaraja, T., Murthy, N., Kumar, L. V., and Mallikarjun, K. (2019). Studies on cross compatibility in interspecific crosses of *Vigna radiata* × *Vigna umbellata* species. *Legume Res.* 42 (5), 699–704.

Bhadkaria, A., Srivastava, N., and Bhagyawant, S. S. (2021). A prospective of underutilized legume moth bean (*Vigna aconitifolia* (Jacq.) marechàl): phytochemical profiling, bioactive compounds and *in vitro* pharmacological studies. *Food Biosci.* 42, 101088.

Bhargava, S., and Chandra, N. (1983). *In vitro* differentiation in callus cultures of moth bean, *Vigna aconitifolia* (JACQ) marechal. *Plant Cell Rep.* 2 (1), 47–50. doi: 10.1007/BF00269235

Bhargava, S., and Chandra, N. (1989). Factors affecting regeneration from leaf explants of moth bean *Vigna aconitifolia* (Jacq.) marechal. *Indian J. Exp. Biol.* 27 (1), 55–57.

Bhargava, S. C., and Smigocki, A. C. (1994). Transformation of tropical grain legumes using particle bombardment. *Curr. Sci.* 66 (6), 439–442.

Bisht, I. S., Bhat, K. V., Lakhanpaul, S., Latha, M., Jayan, P. K., Biswas, B. K., et al. (2005). Diversity and genetic resources of wild vigna species in India. *Genet. Resour. Crop Evol.* 52 (1), 53–68. doi: 10.1007/s10722-005-0286-0

Bisht, I. S., and Singh, M. (2013). "Asian Vigna," in *Genetic and genomic resources of grain legume improvement* (London, United Kingdom: Elsevier), 237–267.

Biswas, M. R., and Dana, S. (1976a). *Phaseolus aconitifolius* x *Phaseolus trilobatus*. *Indian J. Genet.* 36, 125–131.

Biswas, M. R., and Dana, S. (1976b). Allopolyploid of tetraploid *Phaseolus* species× *P. calcaratus* cross. *Biol. Plant* 18 (5), 335–338. doi: 10.1007/BF02922458

Bouchenak, M., and Lamri-Senhadji, M. (2013). Nutritional quality of legumes, and their role in cardiometabolic risk prevention: a review. *J. Med. Food.* 16 (3), 185–198. doi: 10.1089/jmf.2011.0238

Brink, M., and Jansen, P. C. M. (2006). "(Jacq.) maréchal," in *PROTA (Plant resources of tropical Africa/Ressources végétales de l'Afriquetropicale)*. Eds. M. Brink and G. Belay(Wageningen, Netherlands: PROTA Foundation).

Chandra, A., and Pental, D. (2003). Regeneration and genetic transformation of grain legumes: an overview. *Curr. Sci.* 84 (3), 381–387.

Chaudhary, N. B., Patel, A. M., Chaudhari, R. H., Viradiya, Y. A., and Patel, H. S. (2021). Genetic analysis for yield and quality related traits in moth bean [*Vigna aconitifolia* (Jacq.) marechal]. *Int. J. Chem. Stud.* 9 (1), 2658–2661. doi: 10.22271/chemi.2021.v9.i1ak.11629

Chen, N. C., Baker, L. R., and Honma, S. (1983). Interspecific crossability among four species of vigna food legumes. *Euphytica* 32 (3), 925–937. doi: 10.1007/BF00042175

Choudhary, K., Singh, M., Rathore, M. S., and Shekhawat, N. S. (2009). Somatic embryogenesis and *in vitro* plant regeneration in moth bean [*Vigna aconitifolia* (Jacq.) marechal]: a recalcitrant grain legume. *Plant Biotechnol. Rep.* 3 (3), 205–211.

Considine, M. J., Siddique, K. H., and Foyer, C. H. (2017). Nature's pulse power: legumes, food security and climate change. *J. Exp. Bot.* 68 (8), 1815–1818. doi: 10.1093/jxb/erx099

Cullis, C., and Kunert, K. J. (2017). Unlocking the potential of orphan legumes. *J. Exp. Bot.* 68 (8), 1895–1903.

Dabi, R. K., and Gour, H. N. (1988). Field screening of moth bean (*Vigna aconitifolia*) for susceptibility to insect pests and diseases. *Indian J. Agric. Sci.* 58 (11), 843–844.

Dana, S. (1976). Origin, evolution and distribution of some grain legumes. *Indian J. Genet.* 36, 125–131.

Dana, S. (1980). Genomic relationship in the genus vigna and its implications in the breeding programme. In *Breeding methods for the improvement of pulse crops* (Ludhiana, India: Punjab Agr Univ), 357–367.

Darlington, C. D., and Wylie, A. P. (1955). *Chromosome atlas of flowering plants. chromosome atlas of flowering plants, 2nd Ed.* (London: Allen & Unwin)

Davies, N. T., and Nightingale, R. (1975). The effects of phytate on intestinal absorption and secretion of zinc, and whole-body retention of zn, copper, iron and manganese in rats. *Br. J. Nutr.* 34 (2), 243–258. doi: 10.1017/S0007114575000293

De Candolle, A. (1986). *Origin of cultivated plants* Vol. 235 (USA: Hafner Publ. Co).

Delauney, A. J., Hu, C. A., Kishor, P. B., and Verma, D. P. (1993). Cloning of ornithine delta-aminotransferase cDNA from *Vigna aconitifolia* by trans-complementation in *Escherichia coli* and regulation of proline biosynthesis. *J. Biol. Chem.* 268 (25), 18673–18678. doi: 10.1016/S0021-9258(17)46682-8

Deshmukh, B. A., and Pawar, V. S. (2020). Effects of different pretreatments on physicochemical and anti-nutritional quality of moth bean. *J. Pharmacogn. phytochem.* 9 (1), 1965–1968.

Deshpande, S. S., and Cheryan, M. (1984). Effects of phytic acid, divalent cations, and their interactions on α-amylase activity. *J. Food Sci.* 49 (2), 516–519. doi: 10.1111/j.1365-2621.1984.tb12456.x

Directorate of Economics & Statistics, DAC&FW India. (Accessed Accessed: July 16, 2022). eands.dacnet.nic.in.

D'Souza, S. F., Reddy, K. S., Badigannavar, A. M., Manjaya, J. G., and Jambhulkar, S. J. (2009). Mutation breeding in oilseeds and grain legumes in India: accomplishments and socio-economic impact. In *Induced plant mutations in the genomics era. Proceedings of an International Joint FAO/IAEA Symposium* (Rome, Italy: Food and Agriculture Organization of the United Nations), 55–57.

Eapen, S., and Gill, R. (1986). Regeneration of plants from cultured root explants of moth bean (*Vigna aconitifolia* l. jacq. marechal). *Theor. Appl. Genet.* 72 (3), 384–387. doi: 10.1007/BF00288577

Eapen, S., Köhler, F., Gerdemann, M., and Schieder, O. (1987). Cultivar dependence of transformation rates in moth bean after co-cultivation of protoplasts with *Agrobacterium tumefaciens*. *Theor. Appl. Genet.* 75 (1), 207–210. doi: 10.1007/BF00249165

Erdman, J. W. (1979). Oilseed phytates: nutritional implications. *J. Am. Oil Chem. Soc* 56 (8), 736–741. doi: 10.1007/BF02663052

Eyzaguirre, P. B., Padulosi, S., and Hodgkin, T. (1999). IPGRI's strategy for neglected and underutilized species and the human dimension of agrobiodiversity. *Priority-setting underutilized Negl. Plant species Mediterr. region. Report of the IPGRI Conference 9-11 February 1998 – ICARDA, Aleppo, Syria*. (Rome, Italy: International Plant Genetic Resources Institute), 1–19.

Food and Agriculture Organization of the United Nations Statistics Division. Available at: www.fao.org/faostat/en (Accessed Accessed: July 16, 2022).

Garg, B. K., Kathju, S., and Burman, U. (2001). Influence of water stress on water relations, photosynthetic parameters and nitrogen metabolism of moth bean genotypes. *Biol. Plant* 44 (2), 289–292. doi: 10.1023/A:1010215812791

Genebank Project, NARO (National Agriculture and Food Research Organization). Available at: www.gene.affrc.go.jp (Accessed Accessed: September 30, 2022).

Genesys. Available at: www.genesys-pgr.org (Accessed Accessed: September 30, 2022).

Gill, R., and Eapen, S. (1986). Plant regeneration from hypocotyl protoplasts of moth bean (*Vigna aconitifolia*). *Curr. Sci.* 55 (2), 100–102.

Gleeson, D., Lelu-Walter, M. A., and Parkinson, M. (2005). Overproduction of proline in transgenic hybrid larch (Larix x leptoeuropaea (Dengler)) cultures renders them tolerant to cold, salt and frost. *Mol. Breed.* 15 (1), 21–29. doi: 10.1007/s11032-004-1363-3

Godbole, D. A., Kunachgi, M. N., Potdar, U. A., Krishnamurthy, K. V., and Mascarenhas, A. F. (1984). Studies on a drought resistant legume: the moth bean, *Vigna aconitifolia* (Jacq) marechal. II. morphogenetic studies. *Plant Cell Rep.* 3 (2), 75–78. doi: 10.1007/BF00270976

Gopalan, C., Ramasastri, B. V., and Balasubramanian, S. C. (1989). *Nutritive Value of Indian Foods*. (Hyderabad: National Institute of Nutrition)

Gresshoff, P. M., and Mohapatra, S. S. (1982). "Legume cell and tissue culture," in *Tissue culture of economically important plants: proceedings of the international symposium held at the botany department, national university of Singapore, Singapore, 28-30 April 1981/edited by AN rao* (Bangalore: COSTED and ANBS (Asian Network for Biological Sciences).

Gruère, G., Giuliani, A., and Smale, M. (2006). Marketing underutilized plant species for the benefit of the poor: a conceptual framework (No. 154). *International Food Policy Research Institute (IFPRI)*, 1–36.

Gupta, N., Shrivastava, N., Singh, P. K., and Bhagyawant, S. S. (2016). Phytochemical evaluation of moth bean (*Vigna aconitifolia* l.) seeds and their divergence. *Biochem. Res. Int.* 2016, 128–133.

Gurjar, K., Rampuria, S., Joshi, U., Palit, P., Bhatt, K. V., Mohapatra, T., et al. (2014). Identification of heat-related ESTs in moth bean through suppression subtraction hybridization. *Appl. Biochem. Biotechnol.* 173 (8), 2116–2128. doi: 10.1007/s12010-014-1011-2

Hardy, Z. (2016). *Functional and nutritional characteristics of bambara groundnut milk powder as an ingredient in yoghurt* [Doctoral dissertation] (Cape Town Western Province: Cape Peninsula University of Technology).

Harsh, A., Sharma, Y. K., Joshi, U., Rampuria, S., Singh, G., Kumar, S., et al. (2016). Effect of short-term heat stress on total sugars, proline and some antioxidant enzymes in moth bean (*Vigna aconitifolia*). *Ann. Agric. Sci.* 61 (1), pp.57–pp.64. doi: 10.1016/j.aoas.2016.02.001

Henry, A., and Daulay, H. S. (1983). Performance of induced mutants in moth bean. *Indian J. Genet. Plant Breed.* 43 (3), 342–344.

Heuzé, V., Tran, G., and Lebas, F. (2020). Moth bean (*Vigna aconitifolia*). Feedipedia, a programme by INRAE, CIRAD, AFZ and FAO. Available at: https://www.feedipedia.org/node/237 (Accessed: July 16, 2022).

Jain, H. K., and Mehra, K. L. (1980). *Evolution, adaptation, relationships, and uses of the species of vigna cultivated in India*. (Richmond: Royal Botanical Gardens, Kew) 459–468.

Jain, U. K., Ramkrishana, K., and Jain, S. K. (2013). Comparative mutagenic efficiency, effectiveness and induced polygenic variability in mothbean (*Vigna aconitifolia* l.). *Indian J. Genet. Plant Breed.* 73 (1), 57–63. doi: 10.5958/j.0019-5200.73.1.008

Jangid, R. K., Sharma, R., Sudarsan, Y., Eapen, S., Singh, G., and Purohit, A. K. (2010). Microwave treatment induced mutations and altered gene expression in *Vigna aconitifolia*. *Biol. Plant* 54 (4), 703–706. doi: 10.1007/s10535-010-0124-x

Jegadeesan, S., and Reddy, S. K. (2018). *Radiation Induced Mutations in Genetic Enhancement and Development of New Varieties in Black Gram [Vigna mungo (L.) Hepper]* (No. IAEA-CN–263). Austria: International Atomic Energy Agency (IAEA)

Kaga, A., Tomooka, N., Egawa, Y., Hosaka, K., and Kamijima, O. (1996). Species relationships in the subgenus ceratotropis (genus vigna) as revealed by RAPD analysis. *Euphytica* 88 (1), 17–24. doi: 10.1007/BF00029261

Kamble, S., Misra, H. S., Mahajan, S. K., and Eapen, S. (2003). A protocol for efficient biolistic transformation of mothbean (*Vigna aconitifolia* l. jacq. marechal). *Plant Mol. Biol. Rep.* 21 (4), 457–458. doi: 10.1007/BF02772595

Karthikeyan, A., Pandian, S. K., and Ramesh, M. (2011). Transgenic indica rice cv. ADT 43 expressing a Δ1-pyrroline-5-carboxylate synthetase (P5CS) gene from *Vigna aconitifolia* demonstrates salt tolerance. *Plant Cell Tissue Organ Cult.* 107 (3), 383–395.

Kathju, S., Garg, B. K., Vyas, S. P., and Lahiri, A. N. (2003). Sustainable production of moth bean through genotype management under arid environments. *J. Arid Environ.* 53 (1), 137–143. doi: 10.1006/jare.2002.1026

Kaviraj, C. P., Kiran, G., Venugopal, R. B., Kishor, P. K., and Rao, S. (2006). Somatic embryogenesis and plant regeneration from cotyledonary explants of green gram [*Vigna radiata* (L.) Wilezesk.]–a recalcitrant grain legume. *In Vitro Cell. Dev. Biol. Plant* 42, 134–138. doi: 10.1079/IVP2005739

Khadke, S. G., and Kothekar, V. S. (2011). Genetic improvement of moth bean (Vigna acconitifolia (Jacq.) marechal) through mutation breeding. In: S. Khan and M. I. Kozgar (Eds.), *Breeding of Pulse Crops* (Ludhiana, India: Kalyani Publishers), 34–54.

Khadke, S., and Kothekar, V. (2018). *Effect of EMS and SA on Trypsin Inhibitor Content in Moth Bean (Vigna aconitifolia (Jacq.) Marechal)*, (No. IAEA-CN–263). Austria: International Atomic Energy Agency (IAEA)

Khokhar, S., and Chauhan, B. M. (1986). Antinutritional factors in moth bean (*Vigna aconitifolia*): varietal differences and effects of methods of domestic processing and cooking. *J. Food Sci.* 51 (3), 591–594. doi: 10.1111/j.1365-2621.1986.tb13887.x

Kigel, J. (1999). *Culinary and nutritional quality of phaseolus vulgaris seeds as affected by environmental Biotechnol. Agron. Soc. Environ* 3(4):205–209 .

Klozová, E. (1965). Interrelations among several Asiatic species of the genus *Phaseolus* studied by immunochemical methods. In *Proceedings of the Symposium on the Mutational Process, 9-11 August 1965 - Prague, Czech Republic*. (Prague, Czech Republic: Academia), 485–487.

Kochhar, S. L. (2016). "Mat or moth bean [*Phaseolus aconitifolius*]," in *Economic botany: a comprehensive study*, *5th edition*, vol. 430 . Ed. Kochhar, (Cambridge Univ. Press), 167–195.

Kohakade, S. N., Bhavsar, V. V., and Pawar, V. Y. (2017). Research article genetic divergence in moth bean (*Vigna aconitifolia* (Jacq) marechal). *Electron. J. Plant Breed.* 8 (4), 1148–1158. doi: 10.5958/0975-928X.2017.00167.3

Köhler, F., Golz, C., Eapen, S., Kohn, H., and Schieder, O. (1987a). Stable transformation of moth bean vigna aconitifolia via direct gene transfer. *Plant Cell Rep.* 6 (4), 313–317. doi: 10.1007/BF00272007

Köhler, F., Golz, C., Eapen, S., and Schieder, O. (1987b). Influence of plant cultivar and plasmid-DNA on transformation rates in tobacco and moth bean. *Plant Sci.* 53 (1), 87–91. doi: 10.1016/0168-9452(87)90182-8

Krishnamurthy, K. V., Godbole, D. A., and Mascarenhas, A. F. (1984). Studies on a drought resistant legume: the moth bean, *Vigna aconitifolia* (Jacq) marechal. i. protoplast culture and organogenesis. *Plant Cell Rep.* 3 (1), 30–32.

Kumar, D. (2002). *Production technology for moth bean in India* (Jodhpur, India: Indian Turisum Publication).

Kumar, D. (2005). Status and direction of arid legumes research in India. *Indian J. Agric. Sci.* 75, 375–391.

Kumar, J., Choudhary, A. K., Gupta, D. S., and Kumar, S. (2019). Towards exploitation of adaptive traits for climate-resilient smart pulses. *Int. J. Mol. Sci.* 20 (12), 2971. doi: 10.3390/ijms20122971

Kumar, A. S., Gamborg, O. L., and Nabors, M. W. (1988). Plant regeneration from cell suspension cultures of *Vigna aconitifolia*. *Plant Cell Rep.* 7 (2), 138–141. doi: 10.1007/BF00270124

Kumar, S., Saxena, S. N., Jat, R. S., and Sharma, R. (2008). Agrobacterium tumefaciens mediated genetic transformation of moth bean (*Vigna aconitifolia* (Jacq) marechal). *Indian J. Genet. Plant Breed.* 68 (3), 327.

Kumar, S., Singh, G., Jat, R. S., Saxena, S. N., and Sharma, R. (2010). Transformation of moth bean (*Vigna aconotifolia*) cotyledonary nodes through vacuum infiltration of. *Agrobacterium tumefaciens. Green Farm.* 1(6), 551–554.

Lawn, R. J. (1989). Agronomic and physiological constraints to the productivity of tropical grain legumes and prospects for improvement. *Exp. Agric.* 25(4), 509–528. doi: 10.1017/S0014479700015143

Lemke, R. L., Zhong, Z., Campbell, C. A., and Zentner, R. (2007). Can pulse crops play a role in mitigating greenhouse gases from north American agriculture? *Agron. J.* 99 (6), 1719–1725.

Liener, I. E. (1976). Legume toxins in relation to protein digestibility-a review. *J. Food Sci.* 41 (5), 1076–1081. doi: 10.1111/j.1365-2621.1976.tb14391.x

Longvah, T., Anantan, I., Bhaskarachary, K., Venkaiah, K., and Longvah, T. (2017). *Indian Food composition tables* (Hyderabad: National Institute of Nutrition, Indian Council of Medical Research), 2–58.

Ma, D. Z., Sun, J., Zhang, G. Q., Wang, H. X., and Ng, T. B. (2010). A storage protein-like trypsin inhibitor from the moth bean (*Phaseolus acutifolius*) with antiproliferative activity toward lymphoma cells. *Protein Pept. Lett.* 17 (6), 782–788. doi: 10.2174/092986610791190408

Malambane, G., and Bhatt, K. V. (2014). Agro-morphological performance of moth bean (*Vigna aconitifolia*) lines in severe drought stressed and rain-fed condition. *Direct Res. J. Agric. Food Sci. (DRJAFS)* 2, 107–112.

Mantesh, M., Venkatesh,, and Pankaja, N. S. (2020). The studies on the morphological variability and biochemical changes induced by mungbean yellow mosaic virus (MYMV) in mungbean [*Vigna radiata* (L.) wilczek]. *Indian Phytopathol.* 73 (3), 543–553. doi: 10.1007/s42360-020-00238-7

Maréchal, R. (1978). Taxonomic study of a complex group of species of the genera *Phaseolus* and *Vigna* (Papilionaceae) on the basis of morphological and pollen data, processed by computer analysis. *Boissiera* 28, 1–273.

Medhe, S., Jain, S., and Anal, A. K. (2019). Effects of sprouting and cooking processes on physicochemical and functional properties of moth bean (*Vigna aconitifolia*) seed and flour. *J. Food Sci. Technol.* 56, 2115–2125. doi: 10.1007/s13197-019-03692-y

Meghwal, R., Joshi, U., Kumar, S., and Sharma, R. (2015). Screening of moth bean (*Vigna aconitifolia*) core collection against yellow mosaic virus. *Indian J. Agric. Sci.* 854, 571–575.

Mroginski, L. A., and Kartha, K. K. (1984). Tissue culture of legumes for crop improvement: phaseolus. *Plant Breed. Rev.* 2, 215–264. doi: 10.1002/9781118060995.ch7

Munro, D. B., and Small, E. (1998). *Vegetables of Canada* (NRC Research Press).

Narayan, S., Saxena, S. N., Jakhar, M. L., and Sharma, R. (2015). Enhancement of regeneration in moth bean [*Vigna aconitifolia* (Jacq) marechal] through gamma irradiation. *Legume Res.* 38 (4), 519–523. doi: 10.5958/0976-0571.2015.00086.7

N'Dri Yao, D., Kouassi, K. N., Erba, D., Scazzina, F., Pellegrini, N., and Casiraghi, M. C. (2015). Nutritive evaluation of the bambara groundnut Ci12 landrace [*Vigna subterranea* (L.) verdc. (Fabaceae)] produced in côte d'Ivoire. *Int. J. Mol. Sci.* 16 (9), 21428–21441.

Oliveira, A. R. D. S., Martins, L. D. V., Bustamante, F. D. O., Muñoz-Amatriaín, M., Close, T., da Costa, A. F., et al. (2020). Breaks of macrosynteny and collinearity among moth bean (*Vigna aconitifolia*), cowpea (V. unguiculata), and common bean (*Phaseolus vulgaris*). *Chromosome Res.* 28 (3), 293–306.

Oyeleke, G. O., Afolabi, O., and Isola, A. D. (2012). Some quality characteristics and carbohydrate fractions of bambara groundnut (*Vigna subterranea* l.) seed flour. *J. Appl. Chem.* 2 (4), 16–19.

Padulosi, S., and Hoeschle-Zeledon, I. (2004). Underutilized plant species: what are they? *Leisa-Leusden* 20, 5–6.

Pal, S., Sharma, N. K., Sharma, A. K., Kumar, A., and Meena, S. (2020). Study of genetic variability parameters for seed yield and component traits in mothbean [*vigna aconitifolia* (jacq) marechal]. *Int. J. Chem. Stud.* 8, 2638–2641. doi: 10.22271/chemi.2020.v8.i4ae.10038

Pandit, R., Bhusal, B., Regmi, R., Neupane, P., Bhattarai, K., Maharjan, B., et al. (2021). Mutation breeding for crop improvement: a review. *Rev. Food Agr.* 2 (1), 31–35. doi: 10.26480/rfna.01.2021.31.35

Pandiyan, M., Senthil, N., Suersh, R., Chakravarthy, N., Packiaraj, D., and Jagadeesh, S. (2012). Interspecific hybridization of *Vigna radiata*× *Vigna trilobata*. *Euphytica* 35, 1017–1022.

Panicker, S., and Hamdule, A. (2021). Analysis of medicinal properties of *Vigna aconitifolia* (Matki). *Asian J. Pharm. Clin. Res.* 14(5), 78–81. doi: 10.22159/ajpcr.2021.v14i5.41186

Patel, H., Kalaria, R., Mahatma, M., Chauhan, D. A., and Mahatma, L. (2013). Physiological and biochemical changes induced by mungbean yellow mosaic virus (MYMV) in mungbean [*Vigna radiata* (L.) wilczek]. *Cell Tissue Res.* 13 (3), 3927.

Piper, C. V., and Morse, W. J. (1914). Five Oriental Species of Beans. *U.S. Dept. Agric. Bull.* 119, 1–32.

Priyanka, S., Rizwan, M., Bhatt, K. V., Mohapatra, T., and Govind, S. (2011). *In-vitro* response of *Vigna aconitifolia* to drought stress induced by peg 6000. *J. Stress Physiol. Biochem.* 7 (3), 108–121.

Pugalenthi, M., Vadivel, V., and Siddhuraju, P. (2005). Alternative food/feed perspectives of an underutilized legume *Mucuna pruriens* var.utilis–a review. *Plant foods Hum. Nutr.* 60 (4), 201–218. doi: 10.1007/s11130-005-8620-4

Purseglove, J. W. (1974). *In: tropical crops: dicotyledons* (London: Longman), 290–294.

Rachie, K. O., and Roberts, L. M. (1974). Grain legumes of the lowland tropics. *Adv. Agron.* 26, 1–132. doi: 10.1016/S0065-2113(08)60869-X

Raina, A., Laskar, R. A., Khursheed, S., Amin, R., Tantray, Y. R., Parveen, K., et al. (2016). Role of mutation breeding in crop improvement-past, present and future. *Asian Res. J. Agric.* 2 (2), 1–13. doi: 10.9734/ARJA/2016/29334

Ramkrishna, K., Jain, U., Kumawat, S. C., and Balai, O. P. (2008). *Mutation breeding in arid legumes in India (No.IAEA-CN–167)*. Austria: International Atomic Energy Agency (IAEA).

Rampuria, S., Joshi, U., Palit, P., Deokar, A. A., Meghwal, R. R., Mohapatra, T., et al. (2012). Construction and analysis of an SSH cDNA library of early heat-induced genes of *Vigna aconitifolia* variety RMO-40. *Genome* 55 (10), 783–796.

Rao, A. V., Venkateswarlu, B., and Henry, A. (1984). Genetic variation in nodulation and nitrogenase activity in guar and moth [bean]. *Indian J. Genet. Plant Breed.* 44, 425–428.

Sachdeva, S., Sharma, V., Bhat, K. V., and Sen, S. (2016). Morphological and physiological response studies of moth bean (*Vigna aconitifolia* l.) genotypes under drought stress. *Int. J. IT Eng. Appl. Sci. Res.* 5 (7), 2319–4413.

Sahoo, S., Sanadya, S. K., Kumari, N., and Baranda, B. (2019). Estimation of the various genetic variability parameters for seed yield and its component traits in mothbean germplasm [*Vigna aconitifolia* (Jacq.) marechal]. *J. Pharmacognosy Phytochem.* 8 (3S), 49–52.

Sampson, H. C. (1936). Cultivated crop plants of the British empire and the Anglo-Egyptian Sudan tropical and sub-tropical. *Kew Bulletin*. London: H.M. Stationery Office.

Saravanan, M., and Ignacimuthu, S. (2015). Hypocholesterolemic effect of Indian medicinal plants-a review. *Med. Chem.* 5 (1), 040–049. doi: 10.4172/2161-0444.1000241

Sathe, S. K., and Venkatachalam, M. (2007). Fractionation and biochemical characterization of moth bean (*Vigna aconitifolia* l.) proteins. *LWT-Food Sci. Technol.* 40 (4), 600–610. doi: 10.1016/j.lwt.2006.03.021

Saxena, S. N. (2004). *Physio-molecular analysis of somaclonal variation in cultures of moth bean (Vigna aconitifolia) derived from different levels of cytokinins [Doctoral dissertation]* (Jobner (RAJ)]: Sri Karan Narendra Agriculture University).

Saxena, S. N., Kaushik, N., and Sharma, R. (2008). Effect of abscisic acid and proline on *in vitro* flowering in *Vigna aconitifolia*. *Biol. Plant* 52 (1), 181–183. doi: 10.1007/s10535-008-0040-5

Saxena, S. N., Sharma, R., and Mahla, H. R. (2006). Analysis of induced and naturally available genetic diversity in moth bean (*Vigna aconitifolia* (Jacq) merachal) using RAPD markers. *J. Arid Environ.* 3 (1), 60–64.

Schrire, B. D., Lavin, M. A. T. T., and Lewis, G. P. (2005). "Global distribution patterns of the leguminosae: insights from recent phylogenies," in *Plant diversity and complexity patterns: local, regional and global dimensions. proceedings of an international symposium held at the royal Danish academy of sciences and letters in Copenhagen, Denmark, 25-28 may 2003* (Copenhagen, Denmark: The Royal Danish Academy of Sciences and Letters), 375–422.

Senthilkumar, T., and Ngadi, M. (2020). "Moth bean," in *Pulses: Processing and Product Development* (Switzerland: Springer Nature), pp.205–pp.212.

Sharma, R., Jain, M., Kumar, S., and Kumar, V. (2014). Evaluation of differences among *Vigna aconitifolia* varieties for acquired thermotolerance. *Agric. Res.* 3 (2), 104–112. doi: 10.1007/s40003-014-0108-8

Sharma, R. C., and Kakani, R. K. (2002). "Moth bean genetic improvement," in *Nat. Symp.on Arid Legume for Sustainable Agriculture and Trade* (Jodhpur, India: Scientific Publishers), 9–30.

Sharma, R., Kumar, S., Mahla, H. R., Khandelwal, V., Roy, P. K., and Sundria, M. M. (2021). "Moth bean," in *The beans and the peas* (Cambridge, United Kingdom: Woodhead Publishing), 67–88.

Sharma, N. K., Kumawat, N., and Panwar, P. K. (2015). Evaluation of high yielding varieties and demonstration of production technologies in moth bean at farmers' fields in Western rajasthan. *Ann. Arid Zone.* 54(1&2), 51–53.

Sharma, R., Sharma, P., Kumar, S., Saxena, S. N., Khandelwal, V., and Rizwan, M. (2018). Heat treatment affects regeneration, protein expression and genetic make-up of *Vigna aconitifolia* (Jacq.) marechal. *Ann. Agrar. Sci.* 16 (2), 116–120. doi: 10.1016/j.aasci.2017.12.001

Sihag, S. K., Khatri, R. S., and Joshi, U. N. (2004). Genetic variability and heritability for grain yield and its components in mothbean [*Vigna aconitifolia* (Jacq.) marechal]. *Ann. Biol. Res.* 20, 219–222.

Singh, D. P. (1990). Distant hybridization in genus vigna-a review. *Indian J. Genet. Plant Breed.* 50 (3), 268–276.

Singh, M., Bisht, I. S., Sardana, S., Gautam, N. K., Husain, Z., Gupta, S., et al. (2006). Asiatic Vigna. *Plant Genetic Resources: Food Grain Crops*. (New Delhi, India: Narosa Publishing House), 275–301.

Singh, R. J., Chung, G. H., and Nelson, R. L. (2007). Landmark research in legumes. *Genome* 50 (6), 525–537. doi: 10.1139/G07-037

Singh, M., and Krikorian, A. D. (1982). Inhibition of trypsin activity *in vitro* by phytate. *J. Agric. Food Chem.* 30 (4), 799–800. doi: 10.1021/jf00112a049

Singh, M., Kumar, G., Abraham, Z., Phogat, B. S., Sharma, B. D., and Patel, D. P. (2001). "Germplasm evaluation," in *National bureau of plant genetic resources: a compendium of achievements*. Eds. B. S. Dhillon, K. S. Varaprasad, M. Singh, S. A. U. Srivastava and G. D. Sharma (New Delhi: NBPGR), 116–165.

Singh, R. P., and Mann, H. S. (1976). *Improved dryland agriculture for Western rajasthan* (Hisar, India: Haryana Agricultural University Press).

Singh, R. P., and Reddy, G. S. (1988). Identifying crops and cropping systems with greater production stability in water-deficit environments. *Drought Res. priorities dryland tropics.* 77, 85.

Smartt, J. (1985). Evolution of grain legumes. III. pulses in the genus *Vigna*. *Exp. Agric.* 21 (2), 87–100.

Solanki, R. K., Mahla, H. R., Kakani, R. K., Choudhary, K. B., and Sharma, R. (2018). Stress tolerant crop varieties of major arid zone crops to promote resilience to climatic stresses. *Indian Farm.* 68(09), 41–45.

Somta, P., Jomsangawong, A., Yundaeng, C., Yuan, X., Chen, J., Tomooka, N., et al. (2018). Genetic dissection of azuki bean weevil (*Calloso bruchuschinensis* l.) resistance in moth bean (*Vigna aconitifolia* [Jaqc.] maréchal). *Genes* 9 (11), 555.

Srivastava, J. P., and Soni, C. P. (1995). Physiological basis of yield in moth bean [*Vigna aconitifolia* (Jacq.) marechal] under rainfed conditions. *J. Agron. Crop Sci.* 174 (2), 73–77.

Stagnari, F., Maggio, A., Galieni, A., and Pisante, M. (2017). Multiple benefits of legumes for agriculture sustainability: an overview. *Chem. Biol. Technol. Agric.* 4 (1), 1–13. doi: 10.1186/s40538-016-0085-1

Stewart, C. R., and Lee, J. A. (1981). "Proline accumulation: biochemical aspects," in *Physiology and biochemistry of drought resistance in plants*, vol. 243. (Sydney, Australia: Academic Press), 243–259.

Surekha, C. H., Kumari, K. N., Aruna, L. V., Suneetha, G., Arundhati, A., and Kavi Kishor, P. B. (2014). Expression of the *Vigna aconitifolia* P5CSF129A gene in transgenic pigeonpea enhances proline accumulation and salt tolerance. *Plant Cell Tissue Organ Cult.* 116 (1), 27–36. doi: 10.1007/s11240-013-0378-z

Takahashi, Y., Somta, P., Muto, C., Iseki, K., Naito, K., Pandiyan, M., et al. (2016). Novel genetic resources in the genus *Vigna* unveiled from gene bank accessions. *PloS One* 11 (1), p.e0147568. doi: 10.1371/journal.pone.0147568

Tiwari, B., Kalim, S., Bangar, P., Kumari, R., Kumar, S., Gaikwad, A., et al. (2018). Physiological, biochemical, and molecular responses of thermotolerance in moth bean (*Vigna aconitifolia* (Jacq.) marechal). *Turk. J. Agric. For.* 42 (3), 176–184. doi: 10.3906/tar-1709-1

Tomooka, N., Kaga, A., Isemura, T., and Vaughan, D. (2011). Vigna. In *Wild Crop Relatives: Genomic and Breeding Resources: Legume Crops and Forages* (Heidelberg, Berlin: Springer), 291–311.

Tomooka, N., Naito, K., Kaga, A., Sakai, H., Isemura, T., Ogiso-Tanaka, E., et al. (2014). Evolution, domestication and neo-domestication of the genus *Vigna*. *Plant Genet. Resour.* 12 (S1), S168–S171. doi: 10.1017/S1479262114000483

United States Department of Agriculture (USDA). *Food data central* (Accessed Accessed: July 16, 2022). fdc.nal.usda.gov [fdc Published: April 1.

Vavilov, N. L. (1926). Centers of origin of cultivated plants. *Inst. Appl. Bot. Plant Breed.* 16 (2), 1–248.

Verdcourt, B. (1970). Studies in the leguminosae-papillionoideae for the 'Flora of tropical East africa'. *IV. Kew Bull.* 24, 507–570. doi: 10.2307/4102859

Verma, D. P. S., Hu, C. A., and Delauney, A. J. (1992). Genetic manipulation for proline overproduction and the control of osmoregulation in plants. In *AVRDS Symposium* (Tainan, Taiwan: Asian Vegetable Research and Development Center), 47–57.

Verma, N., Sehrawat, K. D., Sehrawat, A. R., and Pandey, D. (2021). Effective green synthesis, characterization and antibacterial efficacy of silver nanoparticles from seaweed treated sprouts of moth bean (*Vigna aconitifolia* jacq.). *Regen. Eng. Transl. Med.* 8, 152–165.

Vir, O., and Singh, A. K. (2015). Moth bean [*Vigna aconitifolia* (Jacq.) marechal] germplasm: evaluation for genetic variability and inter characters relationship in hot arid climate of western rajasthan, India. *Legume Res.* 38 (6), 748–752.

Viswanatha, K. P., Kumar, D., Ramavtar, S., and Kumar, D. (2016). Improvement of minor pulses and their role in alleviating malnutrition. *Indian J. Genet. Plant Breed.* 76 (4), 593–607. doi: 10.5958/0975-6906.2016.00075.4

West, N. B., and Garber, E. D. (1967). Genetic studies of variant enzymes. i. an electrophoretic survey of esterases and leucine aminopeptidases in the genus phaseolus. *Can. J. Genet. Cytol.* 9 (3), 640–645.

Whyte, R. O., Nilsson-Leissner, G., and Trumble, H. C. (1953). *Legumes agriculture*. *Soil Sci.* 76 (5), 403. doi: 10.1097/00010694-195311000-00017

Xu, B. J., and Chang, S. K. C. (2008). Total phenolic content and antioxidant properties of eclipse black beans (*Phaseolus vulgaris* l.) as affected by processing methods. *J. Food Sci.* 73 (2), H19–H27.

Yaqoob, M., Khan, N. U., Mansoor, M., Gul, S., Khalil, I. A., and Afridi, K. (2015). Moth bean germplasm screening against yellow mosaic virus, II. developmentof moth bean high-yielding seed and fodder cultivars. *Turk. J. Agric. For.* 39(2), 212–226.

Yundaeng, C., Somta, P., Amkul, K., Kon gjaimun, A., Kaga, A., and Tomooka, N. (2019). Construction of genetic linkage map and genome dissection of domestication-related traits of moth bean (*Vigna aconitifolia*), a legume crop of arid areas. *Mol. Genet. Genom.* 294 (3), 621–635. doi: 10.1007/s00438-019-01536-0

Zahoor, A. (2007). Country report on plant genetic resources for food and agriculture. *Pakistan Agric. Res. Council.* 86, 27–34.

Zhu, B., Su, J., Chang, M., Verma, D. P. S., Fan, Y. L., and Wu, R. (1998). Overexpression of a Δ1-pyrroline-5-carboxylate synthetase gene and analysis of tolerance to water-and salt-stress in transgenic rice. *Plant Sci.* 139 (1), 41–48. doi: 10.1016/S0168-9452(98)00175-7

Zsögön, A., Čermák, T., Naves, E. R., Notini, M. M., Edel, K. H., Weinl, S., et al. (2018). *De novo* domesticati.on of wild tomato using genome editing. *Nat. Biotechnol.* 36 (12), 1211–1216.

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 10 July 2023
DOI 10.3389/fpls.2023.1193635

OPEN ACCESS

EDITED BY
Beckley Ikhajiagbe,
University of Benin, Nigeria

REVIEWED BY
Ravish Choudhary,
Indian Agricultural Research Institute (ICAR), India
Fatümetüzzehra Küçükbay,
İnönü University, Türkiye

*CORRESPONDENCE
Dinesh Kumar
dineshscn@gmail.com

RECEIVED 25 March 2023
ACCEPTED 19 June 2023
PUBLISHED 10 July 2023

CITATION
Kumar D, Ladaniya MS, Gurjar M, Mendke S, Kumar S and Ghosh D (2023) Elucidation of flavanones, phenols and antioxidant capacity influenced by drying methods from physiologically dropped underutilized *Citrus grandis* fruits. *Front. Plant Sci.* 14:1193635. doi: 10.3389/fpls.2023.1193635

Elucidation of flavanones, phenols and antioxidant capacity influenced by drying methods from physiologically dropped underutilized *Citrus grandis* fruits

Dinesh Kumar*, M. S. Ladaniya, Manju Gurjar, Sachin Mendke, Sunil Kumar and Dilip Ghosh

Indian Council of Agricultural Research- Central Citrus Research Institute, Nagpur, Maharashtra, India

Introduction: Nutritional content in citrus fruit is enormous. *Citrus grandis* (L.) Osbeck is underutilised citrus crop that receives little attention due to the lack of knowledge regarding its nutritional value. Citrus waste disposal poses a problem due to economic and environmental factors.

Methods: The metabolites flavonoids, phenols and antioxidant capacity in the dropped fruits of the underutilised citrus species pomelo (*Citrus grandis* (L.) Osbeck) were examined.

Results and discussion: Hesperidin varied from 1.22 to 2.83% and 1.08 to 1.16% from 10 mm to 14 mm whereas naringin dominates in fruits measuring 10 mm and 12mm with 60.61%, 60.77%, and 47.76%, 45.87% in freeze dried (FD) and hot air oven dried (HAOD) samples. According to the results of the antioxidant assays, the highest concentrations of ABTS azino-bis (3-ethylbenzthiazoline-6-sulfonic acid) and DPPH (2, 2-diphenyl-1-picrylhydrazyl radical) were found in freeze dried samples, ranging from 9.679 to 10.416 mmol L^{-1} Trolox and 14.825 to 16.432 mmol L^{-1} Trolox, respectively. However, the Ferric Reducing Antioxidant Power (FRAP) assay revealed higher content in samples of both FD and HAOD that were 10mm in size (4.578 mmol L^{-1} Trolox and 3.730 mmol L^{-1} Trolox). Total phenol content was measured, and the highest concentrations were found in fruits with a diameter between 10 mm and 18 mm. It ranged from 48.479 to 54.498 mg GAE L^{-1} in FD samples and from 45.757 to 51.159 mg GAE L^{-1} in HAOD samples. The smallest fruits, or those that were still in the immature stage, had the highest content. It was found that when the immature dropped fruits were dried by HAOD, the content decreased. At $p<0.01$ and $p<0.05$, there was a significant positive correlation between the flavonoids, antioxidants, and total phenols. The results showed that the immature dropped immature fruits of lesser known underutilised citrus sp. *Citrus grandis* can act as potential source of flavonoids, total phenol concentration, and antioxidant potential. Freeze drying

can be recommended to recover the most bioactive substances from physiologically dropped fruits of *Citrus grandis* for use in the pharmaceutical and nutraceutical sectors. This study will help in reducing the environmental impact caused due to citrus dropped fruits and its responsible management.

KEYWORDS

pomelo (*Citrus grandis*), immature dropped fruits, phytochemicals, waste utilization, nutraceutical source, freeze drying

1 Introduction

Fruit consumption has gained increasing interest among consumers due to the existence of several bioactive and its ability to protect from diseases such as diabetes, cancer, neurodegenerative, among others (Lapuente et al., 2019). The presence of bioactive chemicals neutralises dangerous reactive oxygen species for the body, preventing oxidation of vital macromolecules like DNA, RNA, and proteins and minimizing the occurrence of diseases (Zou et al., 2016). Citrus, a member of the *Rutaceae* family and the *Aurantioideae* subfamily, is one of the most significant fruit crops. *Citrus reticulata*, *Citrus sinensis*, *Citrus limon*, *Citrus aurantium*, and *Citrus paradisi* are among the citrus species grown for commercial purposes (Turner and Burri, 2013). China, Brazil, India, Mexico and United States of America are the top countries that produce citrus fruits (Marti et al., 2009).

Citrus fruits are liked all around the world for their pleasant taste, distinctive flavour, and nutritive value. Citrus fruits composed of a large number of phytochemicals and bioactive substances, such as ascorbic acid, carotene, flavonoids, antioxidants, phenolic compounds, minerals, etc. Citrus fruits have found use in the manufacturing of numerous cosmetics, functional foods, pharmaceuticals, and nutraceutical medications due to its antioxidant, anti-inflammatory, anti-cancer, and anti-fungal properties (Hayat et al., 2010; Kumar et al., 2021). Due to the diverse climatic conditions, India has a vast array of citrus genetic diversity and is also the home to numerous underutilized citrus species that are still unexplored (Kumar et al., 2021).

Citrus grandis (L.) Osbeck sparsely cultivated and underutilised citrus crop that receives little attention due to the lack of knowledge regarding its nutritional value. Pomelo comes in white and pink colour segments and is commonly known by other common names like Pummelo, shaddock, or Chinese grapefruit. Three main categories of citrus flavonoids include flavanones, flavones, and flavonols in which flavanones can be found as aglycones or glycosides. Hesperitin, narirutin, and didymin belong to the rutinosides group, whereas naringin, neohesperidin and neoeriocitrin belong to the neohesperidosides group. Naringenin and hesperitin comes under the aglycone forms (Tripoli et al., 2007). Reactive oxygen species that are produced under stressful circumstances can lead to the oxidation of biomolecules, which can interfere with the healthy cells' normal metabolism and operation. Oxidative stress leads to development of several diseases, including cancer, atherosclerosis, and Alzheimer's disease. Citrus fruits contain natural antioxidants that scavenge or neutralise dangerous free radicals, lowering the risk of disease (Kumar et al., 2019; Kumar et al., 2021). Phenolic compounds are responsible for flavour and colour of the food and have many health promoting and antioxidant properties (Gasecka et al., 2020).

Immature citrus fruits that are green in colour, drop from the stem-branch or ovary-stem junction due to physiological reasons, food deficiencies, insufficient pollination, ovule dysplasia, degeneration, or changes in endogenous hormones, etc. This phenomenon is well known as physiological dropping (Sun et al., 2015). The immature fruits that have fallen to the ground due to physiological dropping are typically dumped in the field or treated as waste. If carefully explored, these rejected dropped citrus fruits can provide a low-cost and environmentally friendly platform for the formulation of nutraceuticals or value-added food supplements. The dropped fruits can also be sold in dried form. In comparison to the conventional sun-drying process, oven drying and freeze drying are more appealing due to their easy control, industrial use, availability during off-season, higher retention of nutritional value, and low temperature and pressure operation (Sun et al., 2015; Bhatta et al., 2020).

Currently, researchers have made attempts to examine the flavonoids and antioxidants in juvenile dropped fruits of commercially cultivated citrus species, but no study has been done to elucidate the nutritional content from the dropped underutilised citrus species, i.e. pomelo. Moreover, there is relatively little knowledge about how drying methods affect the phytochemical and antioxidant content. An experiment was conducted in light of the significance and health advantages of pomelo; to understand how drying procedures affect flavonoids, antioxidant capacity, and phenolic content, as well as how effective they are at producing the highest yield.

2 Materials and methods

2.1 Chemicals and reagents

Standards of flavonoid compounds, such as hesperidin, narirutin, naringin, quercitin, and naringenin (97% purity), antioxidant standard Trolox (97% purity), radical cation $ABTS^{+}$ azino-bis [3-ethylbenzthiazoline-6-sulfonic acid], 2, 2-diphenyl-1-

picrylhydrazyl radical (DPPH), 2, 4, 6-Tri (2-pyridyl)-s-triazine (TPTZ), and gallic acid were purchased from Sigma–Aldrich (Mumbai, India). Chemicals such as sodium acetate trihydrate, ammonium acetate, acetonitrile, dimethyl sulphoxide, and acetonitrile were of the HPLC grade for use in the extraction procedure and in HPLC technique (Himedia, India). Additional chemical substances and reagents utilised in the study, such as methanol, ferric chloride, folin-Ciocalteu reagent, sodium carbonate, manganese dioxide, and 37% hydrochloric acid, were of analytical grade (Himedia, India).

2.2 Plant materials

ICAR- Central Citrus Research Institute experimental blocks with the geographic coordinate's latitude: 21°9'0"N and longitude: 79°9'0"E, respectively, were the site of collection for the immature dropped pomelo (*Citrus grandis*) fruits, which are oblate spheroid in shape (Figure 1). The region's average ambient temperature at the collection time was 26.7°C, with relative humidity 60%. Its average relative humidity ranges from 13.8% to 99.3%, and its temperature ranges from 9.3°C to 43.6°C. Pummelo belongs to the *Rutaceae* family. The tree typically stands between 6 and 15 metres tall, with a 10 to 30 cm thick, crooked trunk. Spines of upto 5 cm can be found. Fruits are around 10 to 30 cm wide; the peel is clingy or more or less readily removed with greenish-yellow or pale-yellow in color. The albedo is soft, white, or pink and is divided into 11 to 18 segments and contains few, large, yellowish-white and white seeds. Samples were taken in accordance with all applicable institutional rules and regulations.

2.3 Drying treatment and extraction method

The collected immature dropped fruits from 8mm to 24mm were segregated and thoroughly washed with tap water in order to remove of the dirt. The samples were then sliced into 0.5 cm thickness and divided into two separate portions. One portion was stored in microwave-oven (RIVOTEK, Riviera Glass Pvt. Ltd., Mumbai, India) for 24h to 36h at 45-50°C for the hot air drying (HAOD) procedure. For the freeze drying (FD) process, a set of additional portion were placed in an ultra-low deep freezer (NEW BRUNSWICKTM, Eppendorf, India) for 12 to 24h at - 80°C, and they were lyophilized in a vacuum freeze dryer (iGene Labserve Pvt. Ltd., New Delhi, India) for 24 to 48h at – 50°C to – 55°C and at 14 to 20 Pa of pressure. The HAOD and FD samples were ground into a fine powder in a blender and passed through a sieve of 50 microns before examination. The resulting powder was kept until further study at -20°C in a deep freezer (Blue Star Ltd., Mumbai, India).

2.4 Determination of antioxidant capacity

Tecan Infinite M200 Pro 96-well microplate reader was used to measure antioxidant capacity (Tecan Group Ltd, Switzerland). The nitrogen radical scavenging activities i.e. ABTS and DPPH are determined as reported (Mena et al., 2011). The reaction was allowed to run at a temperature of 25°C and a wavelength of 414 nm for the ABTS assay and 515 nm for the DPPH test for 50 minutes. In both the assays, water and methanol respectively were used as blanks. Minor modifications were made to the procedure (Benzie and Strain, 1996) in order to assess the antioxidant capacity using the FRAP assay. In this

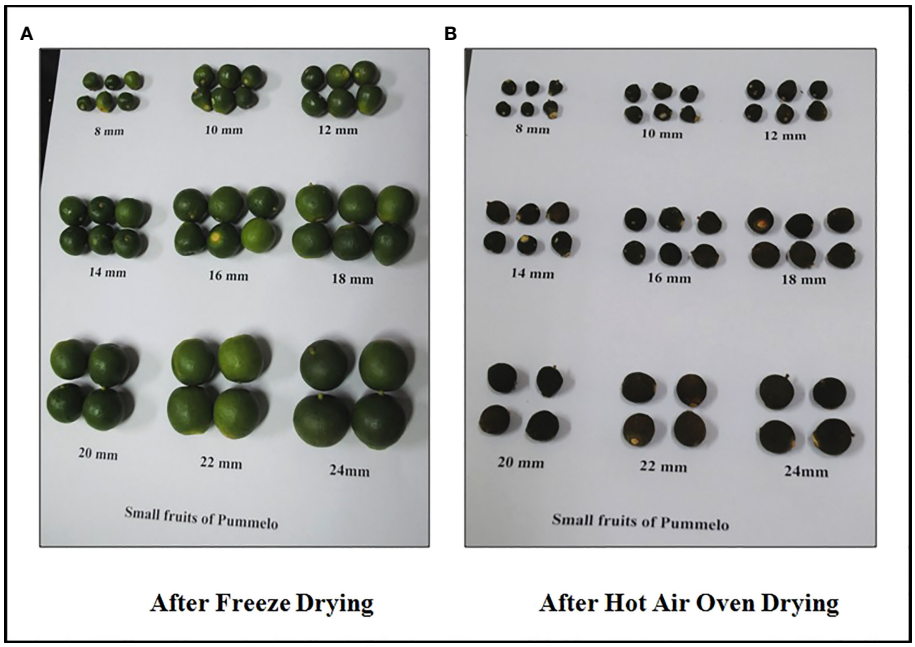


FIGURE 1
Immature Pomelo (*Citrus grandis*) dropped fruits varying from size 8 mm to 24 mm after **(A)** FD and **(B)** HAOD.

assay, the FRAP reagent was prepared freshly using 300 mM acetate buffer, TPTZ solution, and ferric chloride solution in 10:1:1 ratio. A sample extract of 2 µl was then mixed with the FRAP reagent. After 40 min of reaction time at about 25°C, the absorbance was recorded at 593 nm. The results were quantified from the standard curve prepared using trolox and the results obtained were expressed as mmol L^{-1} Trolox. Three replicated trails were carried out during the experimental trials.

2.5 Determination of total phenol content

Total phenolic content was determined by mixing about 10 µL of sample extract with 790 µL milli-Q water, 50 µL Folin–Ciocalteu reagent and 150 µL of 20% sodium carbonate solution (Singleton and Rossi, 1965). The eppendorf tube containing the reaction mixture was shaken to agitate properly and was kept at room temp 23.5°C for 1 hr. The absorbance of the sample solution was recorded at 750 nm. The findings are reported in terms of mg GAE L^{-1} and were measured using gallic acid as the standard.

2.6 Determination of flavonoids composition

The Agilent Model No. 1260 Infinity System (M/s. Agilent Technologies Pvt. Ltd., United States) containing UV detector were used to analyze the flavonoid content of dropped immature pomelo fruits. Only the hesperidin and naringin flavonoids which are primarily present and persistent biomarkers in citrus fruits were found in the dropped immature fruits, despite the study carried out used the standards of flavonoids such as narirutin, hesperidin, naringin, quercetin, and naringenin respectively. The reverse phase column Nucleosil C-18 of 4.6 mm in diameter and 100 mm in length and mobile phase containing 5 mM ammonium acetate as solvent A and acetonitrile as solvent B in 75:25 (v/v) ratio was used in the analysis. The pH of the mobile phase was adjusted using acetic acid. 3 mg of the powdered sample material was mixed with 5 mL of dimethyl sulphoxide (DMSO) for extraction, and the mixture was then sonicated in a 2K1008008 series sonicator (Life-Care Equipments Pvt. Ltd., Mumbai, India). 5 µL each of sample solution as well as standard was then injected into an HPLC system for measurement after being filtered via a 0.45 µ nylon filter. The flow rate of the mobile phase was 1.0 mL/min, and the column temperature was kept at 20°C. Hesperidin and naringin, two flavonoids, were detected and quantified at 284 nm from their corresponding peak areas and calibrated against each standard (stock solution-600 ppm) diluted with the help of the mobile phase (Omidbaigi and Nasiri, 2004; Marten, 2007). The obtained results were reported as a percentage (%).

2.7 Statistical analysis

For each analytical parameter, three replicated measurements were conducted, and the findings were expressed as mean standard deviation (SD). For comparison and to identify significant differences in the data, Tukey's honestly significant difference (HSD) test (multiple range test) and one-way analysis of variance (ANOVA) were performed. The correlation between flavonoids, antioxidant capacity, and total phenol concentration in the sample extracts was evaluated using Pearson correlation coefficients. The probability values (p) <0.01 was considered significantly different. After comparisons, the means in the table with the different-letter superscripts are determined to be statistically significant.

3 Results

3.1 Antioxidant capacities of dropped pomelo fruits

Free radicals from oxygen are known for damaging the human body and leads to conditions like cancer, cardiovascular disease, and problems associated with ageing (Bellocco et al., 2009). The ABTS, DPPH, and FRAP assays were assessed to determine the antioxidant capacity. Figure 2 illustrates the results of antioxidant capacity of dropped fruits of *Citrus grandis* that ranging from size 8mm to 24mm as measured by the ABTS and DPPH assay. When measured using the ABTS assay, the antioxidant capacity varied from 9.679 to 10.416 mmol L^{-1} trolox in FD and from 9.460 to 10.093 mmol L^{-1} trolox in HAOD samples, whereas the DPPH assay measured 14.825 to 16.432 mmol L^{-1} trolox in FD and 13.458 to 15.914 mmol L^{-1} trolox in HAOD samples. Fruit's DPPH content decreases as it attain a higher level of maturity (Rekha et al., 2012). Lower values were obtained with the FRAP assay, which assess ferric-reducing activity, but it exhibited the same trend as the ABTS and DPPH assays (Figure 2). The FD samples of the pomelo dropped fruits resulted in the retention of the FRAP content values ranging from 3.803 to 4.578 mmol L^{-1} Trolox. The results from HAOD fruits ranged from 3.066 to 3.780 mmol L^{-1} Trolox. The greatest concentration was recorded in fruits of 10 mm size (4.578 mmol L^{-1} Trolox in FD and 3.730 mmol L^{-1} Trolox in HAOD). The lowest amount was quantified in dropped fruits of size 24 mm, with concentrations of 3.103 mmol L^{-1} Trolox in FD and 3.803 mmol L^{-1} Trolox in HAOD.

3.2 Total phenol content of dropped pomelo fruits

Phenols are regarded as one of the primary components of citrus fruits. It shields the fruit from detrimental effects of UV radiation and pathogens, as well as from predators. Figure 3 depicts the total phenol content found in the various-sized dropped pomelo fruits. The findings indicate that the 12 mm sample had the highest total phenol content (52.403 mg GAE L^{-1} in FD and 50.530 mg GAE L^{-1} in HAOD), followed by the 18 mm and 10 mm samples, which had amounts of 53.096 mg GAE L^{-1} in FD and 49.338 mg GAE L^{-1} in HAOD and 52.403 mg GAE L^{-1} in FD and 50.530 mg GAE L^{-1} in HAOD, respectively. The findings were consistent with those of vacuum FD Citrus reticulata Blanco dropped fruits with TPC values 50.50–54.19 mg GAE L^{-1} (Kumar et al., 2021). Similarly, the concentration of total phenol decreases in Citrus sinensis L. Osbeck dropped fruits with maturity (Kumar et al., 2022a). Their

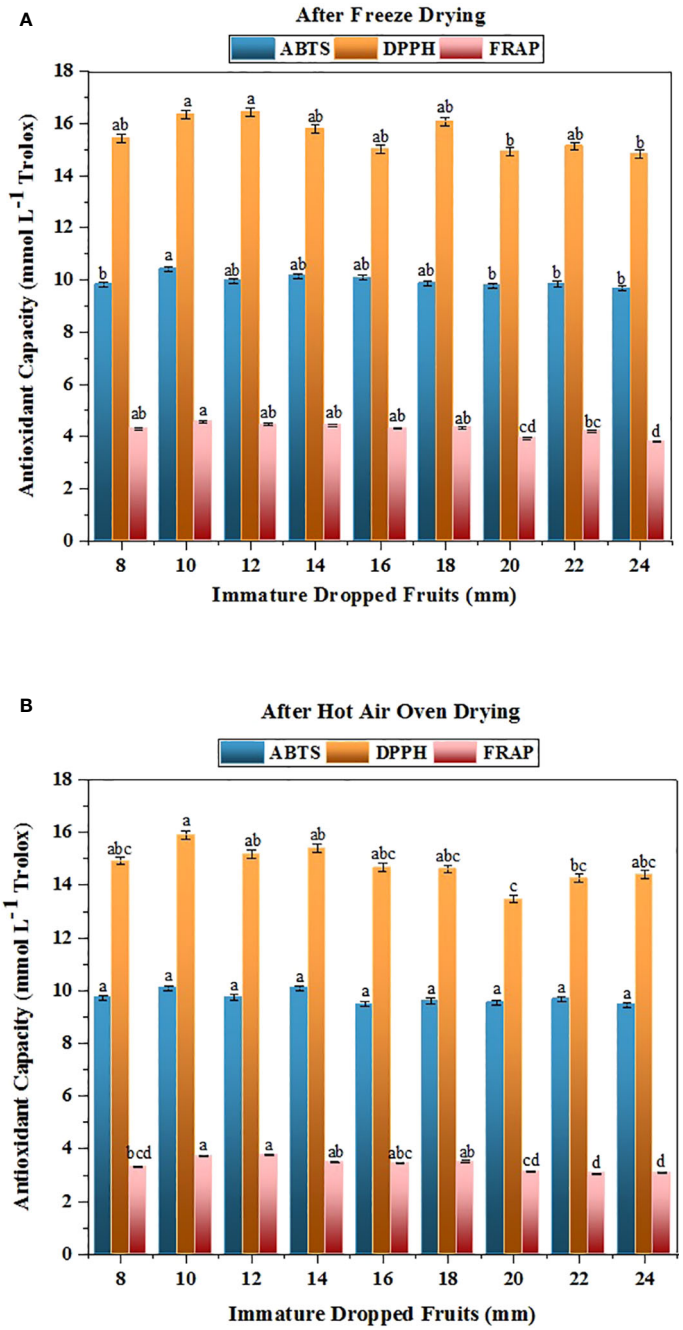


FIGURE 2
Changes in the antioxidant capacity in immature dropped fruits of Pomelo assessed by ABTS, DPPH and FRAP assay after **(A)** FD and **(B)** HAOD. The data with different superscripts are statistically significant at p<0.01 as per Tukey's honestly significant difference (HSD) multiple range test.

findings showed that TPC content ranged from 41.736 to 55.161 mg GAE L^{-1}, which was lower than our findings. The variance in the results can be attributable to different citrus species examined.

3.3 Flavonoid content of dropped pomelo fruits

One of the main secondary metabolites, citrus flavanones, is commonly found in the diglycoside form (Tripoli et al., 2007). Statistical analysis revealed substantial significant differences in the amounts of flavonoids in the various sizes of immature dropped fruits of the pomelo (*Citrus grandis*) as shown in Table 1. Different fruit sizes (12 mm, 14 mm, 16 mm, and 18 mm) were compared for difference in the flavonoid compounds using HPLC chromatograms, detected at a wavelength of 284 nm (Figure 4). The flavonoids hesperidin and naringin were quantified with the peaks against those from standards. It was found that hesperidin has longer retention duration than naringin. Dropped fruits when assess reported naringin as the predominant flavonoid. The fruit of 10

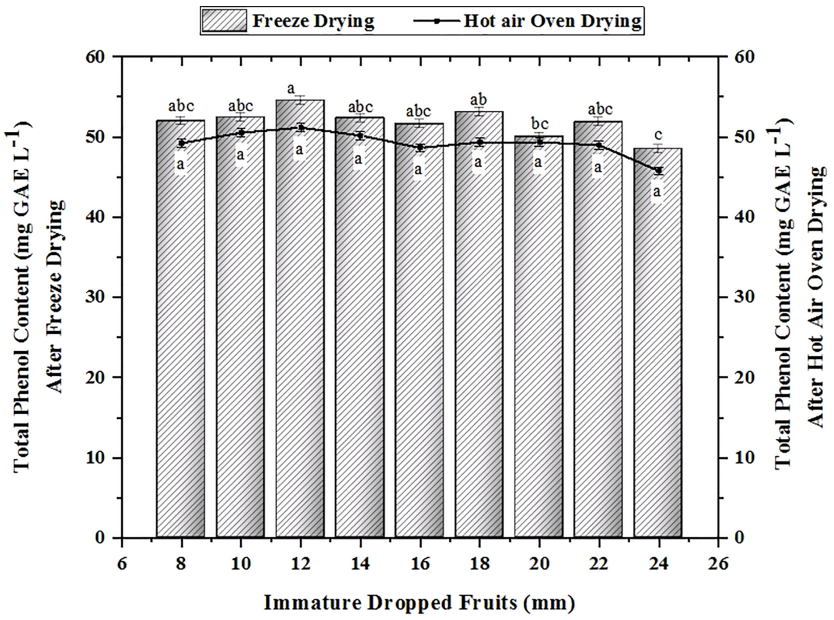


FIGURE 3
Changes in the total phenol content in immature dropped fruits of Pomelo after FD and HAOD. The data with different superscripts are statistically significant at p<0.01 as per Tukey's honestly significant difference (HSD) multiple range test.

TABLE 1 Flavonoid contents (hesperidin and naringin) in immature dropped fruits of Pomelo after freeze drying and hot air oven drying.

Sr. No.	Fruit size (mm)	Drying methods	Hesperidin (%)	Naringin (%)
1	8	FD	$1.210^{bc} \pm 0.351$	$52.870^{abc} \pm 2.339$
		HAOD	$0.930^{a} \pm 0.396$	$42.680^{abc} \pm 2.450$
2	10	FD	$1.220^{bc} \pm 0.108$	$60.610^{a} \pm 4.665$
		HAOD	$1.080^{a} \pm 0.114$	$47.760^{a} \pm 2.027$
3	12	FD	$2.830^{a} \pm 0.173$	$60.770^{a} \pm 1.947$
		HAOD	$1.160^{a} \pm 0.185$	$45.870^{a} \pm 0.830$
4	14	FD	$1.610^{b} \pm 0.139$	$55.600^{ab} \pm 1.057$
		HAOD	$0.880^{ab} \pm 0.321$	$45.040^{ab} \pm 2.933$
5	16	FD	$0.880^{c} \pm 0.230$	$49.990^{bcd} \pm 2.256$
		HAOD	$0.520^{abc} \pm 0.171$	$41.510^{abc} \pm 1.527$
6	18	FD	$1.230^{bc} \pm 0.072$	$45.500^{cde} \pm 1.778$
		HAOD	$0.470^{abc} \pm 0.026$	$37.510^{bcd} \pm 3.468$
7	20	FD	$0.630^{cd} \pm 0.226$	$43.130^{de} \pm 2.278$
		HAOD	$0.480^{abc} \pm 0.122$	$31.060^{de} \pm 2.230$
8	22	FD	$0.150^{d} \pm 0.053$	$37.010^{e} \pm 2.886$
		HAOD	$0.229^{bc} \pm 0.026$	$35.890^{cde} \pm 1.047$
9	24	FD	$0.760^{cd} \pm 0.044$	$41.240^{de} \pm 1.835$
		HAOD	$0.200^{bc} \pm 0.036$	$29.390^{e} \pm 2.629$
Tukey's HSD at 1%		FD	0.6389	8.8562
		HAOD	0.6951	8.001

Where, FD, Freeze drying and HAOD, Hot air oven drying.
Values are of three replicated trials ± standard deviation.
The data with different superscripts are statistically significant at p<0.01 as per Tukey's honestly significant difference (HSD) multiple range test.

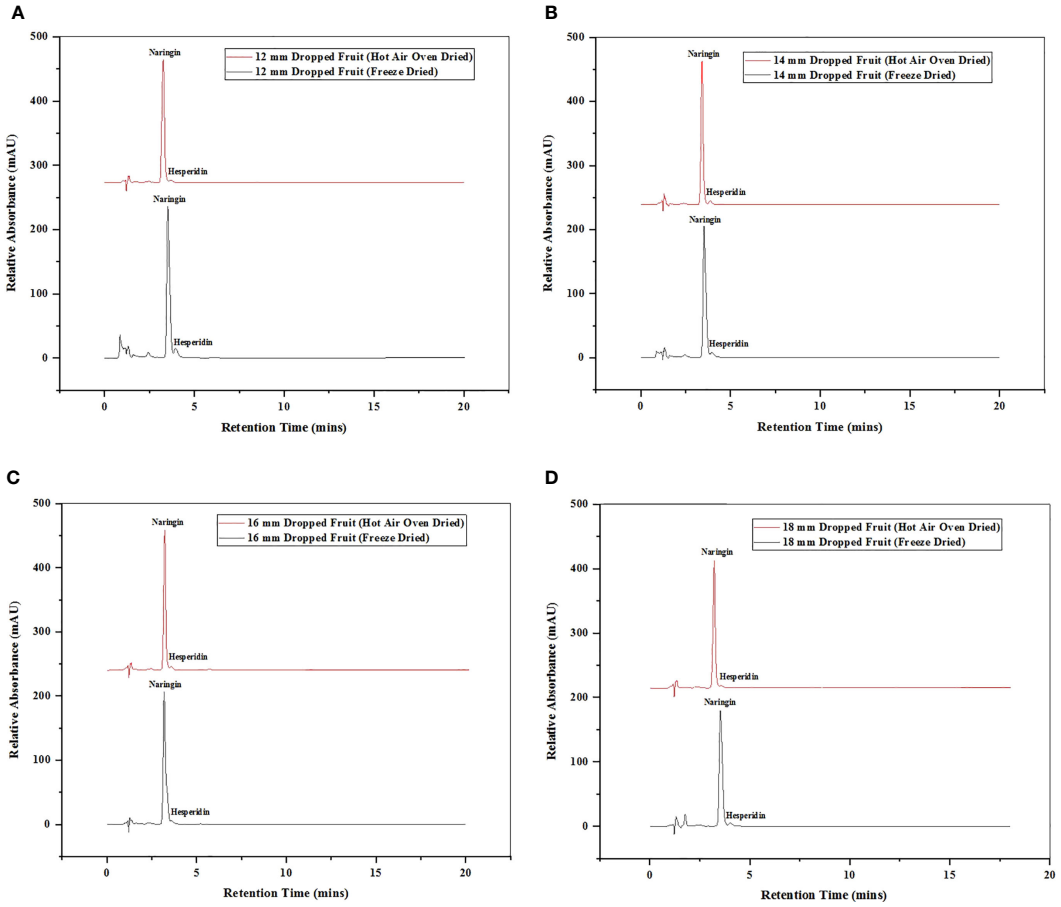


FIGURE 4
Chromatogram of HPLC of hesperidin and narinign flavonoid quantified in immature dropped fruits of Pomelo of sizes **(A)** 12 mm, **(B)** 14 mm, **(C)** 16 mm **(D)** 18 mm after FD and HAOD.

mm, 12 mm, and 14 mm size had highest concentrations of 60.61%, 60.77%, and 55.60% in FD samples and 47.76%, 45.87%, and 45.04% in HAOD samples, respectively. When hesperidin levels were determined, they varied from 0.15 to 2.83% (FD dropped fruits) and 0.20 to 1.16% (HAOD dropped fruits). Dropped fruits ranging from 8 mm to 14 mm had the highest hesperidin content but when compared to naringin, the content was lower.

3.4 Correlation between antioxidants, total phenol and flavonoids

According to the parameters that were examined, the results of FD were much better than those of the HAOD approach for bioactive chemical and antioxidant capacity. The parameters of the freeze-dried samples listed in Table 2 were correlated using

TABLE 2 Pearson's correlation coefficient of flavonoids, antioxidants and total phenol content in dropped pomelo fruits after FD treatment.

	Naringin	Hesperidin	ABTS	DPPH	FRAP	TPC
Naringin	–					
Hesperidin	0.804**	–				
ABTS	0.716*	0.284	–			
DPPH	0.764*	0.757*	0.594	–		
FRAP	0.778*	0.559	0.805**	0.832**	–	
TPC	0.628	0.688*	0.472	0.840**	0.875**	–

**Correlation coefficient (r) values significant at $p < 0.01$.
*Correlation coefficient (r) values significant at $p < 0.05$.
Correlation coefficient (r) values not marked with any asterisk means not significant.

Pearson's coefficient. Citrus fruits contain larger concentrations of flavanone glycosides during the early developmental phases, or in immature fruits, than other categories of flavonoids (Omidbaigi and Nasiri, 2004; Ye et al., 2011; Lou and Ho, 2016). Naringin and hesperidin, two flavonoids, were shown to be substantially associated at $p<0.01$ and $p<0.05$ with correlation coefficients (r) of 0.804. The antioxidant tests ABTS (r= 0.716), DPPH (r= 0.764), as well as FRAP (r= 0.778) also demonstrated a statistically significantly positive correlation with naringin flavonoid at $p<0.05$. However, hesperidin correlated only with the DPPH with r= 0.757 at $p<0.05$. Hesperidin flavonoid also showed a positive correlation with TPC (r= 0.688 at $p<0.05$). Between the antioxidant capacity measured by the ABTS and DPPH assays and that of the FRAP assay, the correlation coefficient (r) in dropped pomelo fruits was 0.805 and 0.832, which were deemed significant at $p<0.01$ and $p<0.05$. At both $p<0.01$ and $p<0.05$, the total phenol showed a significant correlation with both DPPH (r= 0.840) and FRAP (r= 0.875). The correlation coefficient in immature calamondin peel and pulp was 0.7911 at $p<0.01$ (Lou et al., 2014a).

4 Discussion

In the ABTS assay, the radical cation 2, 2'-azino-bis (3-ethylbenzothiazoline-6-sulfonic acid) ($ABTS^{\bullet+}$) can be found as an antioxidant. Higher quenching ability shows the sample's higher antioxidant capability (Barecca et al., 2011; Kumar et al., 2021; Kumar et al., 2022b). In the assay of DPPH, the DPPH is stable free radical. In this assay, the purple DPPH solution transforms into a colourless product indicating the presence of antioxidants. An increase level of discolouration indicates a greater antioxidant capacity (Almeida et al., 2011). The DPPH assay is used to measure both hydrophilic and lipophillic antioxidants, whereas ABTS assay reveals only the hydrophobic antioxidants (Floegel et al., 2011). The observation reveals two important features i.e. (i) the FD treatment were found to retain antioxidants more effectively than the HAOD method and (ii) the immature fruits were found to have greater antioxidant capacity in comparison to mature fruits. According to the ABTS assay, dropped fruits between the sizes of 10 mm and 16 mm had higher antioxidant capacities in both FD and HAOD fruits than mature fruits between the sizes of 18 mm and 24 mm, with a p-value of 0.01. After FD, sour cherries had a higher ABTS scavenging activity than with the other convective procedures that were used in the study (Wojdylo et al., 2013). Moreover, immature kumquats showed increased radical scavenging activity (Lou and Ho, 2017). Similar to this, fruits with a diameter between 10 and 14 mm and dropped fruits with a diameter of 18 mm had greater DPPH concentration. According to research, the juice of Chinotto fruits reported higher DPPH radical scavenging activity in immature fruits than in mature fruit juice (Barecca et al., 2010). A similar pattern was seen in the extract of thinned immature *Citrus unshiu* fruits (Kim and Kim, 2017). The ABTS and DPPH assay results are consistent with those of other researchers (Gasecka et al., 2020; Kumar et al., 2021; Kumar et al., 2022a), who also discovered a higher antioxidant content when freeze-drying the *Leccinum scabrum* (Bull.) Gray and *Hericium erinaceous* (Bull.); citrus fruits namely *Citrus reticulata* Blanco and *Citrus sinensis* L. Osbeck respectively. The findings of FRAP assay pointed out a direct relationship between the content and the drying method used. Similar results were found while studying the effect of drying process on Nagpur mandarin (Kumar et al., 2021), immature mandarin fruits (Ye et al., 2011), and physiological drop citrus fruits (Sun et al., 2013; Kumar et al., 2021), respectively. In contradiction to the results obtained, the FRAP activity was found to be lower in the unripe chinotto fruits (Barecca et al., 2010). Due to prolonged exposure to hot air, the HAOD method caused oxidation. On the other hand, the FD approach operated at lower air pressure for a shorter period of time and thus lessened this oxidation effect (Wojdylo et al., 2013).

The amount of total phenolic content was greatly impacted by the drying process. The largest phenolic concentration was found in FD samples operated at temperatures between −50°C to −55°C for 24 to 48 h compared to samples dried in hot air ovens at 45-50°C for 24 to 36 h, which had a lower phenolic content. The primary enzyme in the phenylpropanoid pathway for the biosynthesis of phenolic compounds is phenylalanine ammonia lyase. The decrease in the activity of this enzyme during the citrus fruit developmental stages and simultaneous increase in the activity of polyphenol oxidase enzyme is attributed to the decrease in the content of total phenols (Gupta et al., 2021).The varying concentration could also be due to the higher temperature of hot air oven (Wojdylo et al., 2013). The findings showed that the content of the dropped immature fruits get influenced during their maturation stages. Phenolic compounds acts as antioxidants (Rice-Evans and Miller, 1996). As a result, samples with greater phenol contents had higher antioxidant potential (Buyukkormaz and Kucukbay, 2022). Similar observations were seen with vacuum-drying of kumquats (Ozcan-Sinir et al., 2019). The findings are consistent with the research done using rose hip (Rosa rubiginosa) and persimmon leathers (Karaman et al., 2014; Ruiz et al., 2014). FD samples when assessed recorded minimum degradation of phenolic content. Several researchers observed similar results when drying sour cherries (Wojdylo et al., 2013), kumquat (Lou et al., 2015); studying maturity stages with Citrus unshiu (Kim et al., 2022), Citrus aurantium (Mansour, 2018) and thinned immature Citrus unshiu (Kim and Kim, 2017) respectively. On the other hand, hot air drying enhanced the total phenol content (TPC) of the aqueous extract of dried lemon (Citrus limon) pomace (Papoutsis et al., 2017). The total phenol concentration observed in our investigation is in accordance with earlier studies done on mandarins (Ye et al., 2011), immature citrus fruits, green and ripe Chinotto (Citrus x myrtifolia Raf.) fruits (Barecca et al., 2010).

Hesperidin is the main flavonoid in mandarin, sweet orange, and lemon, whereas naringin is mostly present in the citrus species of sour orange, pummelo, and grapefruit (Dhuique-mayer et al., 2005; Vanamala et al., 2006). The amount of flavonoids varies greatly amongst citrus species. The variable content is caused by a number of factors, including genetic and environmental, geographic origin, meteorological conditions, soil qualities, time of fruit collection, storage, portions of the fruit, etc (Lu et al., 2006). According to the experimental findings, drying methods and flavonoid content are interconnected. The content was found maximum in the vacuum FD samples. Immature calamondin peels, an underutilised citrus species, showed

the same behaviour (Lou et al., 2014b). Due to increased PPO activity and decreased chalcone synthase gene expression, there is a drop in flavonoid concentration during citrus fruit development (Gupta et al., 2021). The flavonoid content of *Citrus unshui* decreased with increase in maturity (Kim and Lim, 2020; Kim et al., 2022). Calamondin, *Citrus grandis* Osbeck, Chinotto (*Citrus myrtifolia* Raf.), and other immature citrus fruit extracts likewise showed a shifting pattern in flavonoid content (Barecca et al., 2010; Lou et al., 2014b; Kim and Kim, 2017). The study's findings were consistent with those of experimental tests done on grape skin and immature physiologically dropped citrus fruits (Sun et al., 2013; Sun et al., 2015; Kumar et al., 2021; Kumar et al., 2022a). Finally, it can be said with certainty that the FD approach should be used to retain the flavanone glycosides hesperidin and naringin.

Significant correlation was observed between antioxidants, total phenol and flavonoids as per the results obtained. The study with *C. aurantium* citrus fruits found similar results with significant correlation (Mansour, 2018). Flavonoids are responsible for antioxidant capacity (Sun et al., 2013). Other researchers have also found positive correlation between total phenol and antioxidant compounds in Satsuma mandarin and Ponkan, immature kumquat and *Citrus sinensis L.* Osbeck fruits (Xu et al., 2008; Lou et al., 2015; Kumar et al., 2022a). Phenolic molecules are thought to contribute significantly to antioxidant capability (Xu et al., 2008). More is the total phenol content; more is the antioxidant capacity (Rice-Evans and Miller, 1996). The findings are consistent with research undertaken with physiological drops in citrus fruits, respectively (Sun et al., 2013; Kumar et al., 2021).

The findings of the current study will increase consumption of the little-known, underutilised citrus fruit pomelo in light of the growing consumer interest in items with authentic nutritional content. At the same time, the study will encourage industrial applications of dropped fruits towards nutraceutical formulations, herbals, etc. The study will also address the issue of citrus dropped fruits as waste and will contribute in valorization. The socioeconomic status of the citrus growing region will also improve as a result of the study.

5 Conclusions

In this paper, we investigated the bioactive components, primarily flavonoids, antioxidants, and total phenol of immature dropped fruits of underutilised pomelo species and determined their correlation. Furthermore, we studied the impact of FD and HAOD on the assessed components. We found that Naringin was the main flavanone glycoside present in all the different sized fruits with the highest level in the 10 mm and 12 mm sizes. Next, in comparison to other dropped fruits, immature fruits with sizes ranging from 10 mm to 16 mm were found to have abundant antioxidant capacity measured by ABTS, DPPH, and FRAP assay. In case of total phenol content, fruits between 20 and 24 mm in diameter had the lowest levels and those between 12 and 18 mm in diameter had the greatest levels. Flavonoids and total phenol contributed well to the antioxidant capacity and significant correlated at $p<0.01$ and 0.05. The results of the study highlights that (1) Drying effect had a substantial impact on the flavonoids, antioxidant potential, and total phenol content of dropped fruits and (2) FD performed at – 50°C to – 55°C for 24h to 48h was found to be more efficient for obtaining maximum recovery than hot air oven drying at 45-50°C for 24h to 36h. This kind of approach appeared to very useful and offer crucial information regarding the bioactive components of underutilised citrus spp. *Citrus grandis* focusing mainly the dropped fruits which are usually unexplored due to the lack of information. The FD examined reveals great potential applications which can be adapted in the citrus industry in the future.

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

DK conceptualized, designed and wrote the first manuscript. MG performed the experiment, supported data analysis, assisted with the manuscript's writing and editing. ML and DG administered and supervised the research, SM performed the experiment. SK gathered the resources. All authors contributed to the article and approved the submitted version.

Acknowledgments

Authors are thankful to the guidance and the facilities provided by DG (Director, ICAR-CCRI, Nagpur) for carrying out the experimental trials.

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/fpls.2023.1193635/full#supplementary-material

References

Almeida, M. M. B., Machado de Sousa, P. H., Arriaga, A. M. C., Matias do Prado, G., Emanuel de Carvalho Magalhaes, C., Maia, G. A., et al. (2011). Bioactive compounds and antioxidant activity of fresh exotic fruits from northeastern Brazil. *Food Res. Int.* 44, 2155–2159. doi: 10.1016/j.foodres.2011.03.051

Barecca, D., Bellocco, E., Caristi, C., Leuzzi, U., and Gattuso, G. (2010). Flavonoid composition and antioxidant activity of juices from chinotto (*Citrus* x *myrtifolia* raf.) fruits at different ripening stages. *J. Agric. Food. Chem.* 58, 3031–3036. doi: 10.1021/jf9044809

Barecca, D., Bellocco, E., Caristi, C., Leuzzi, U., and Gattuso, G. (2011). Elucidation of the flavonoid and furocoumarin composition and radical-scavenging activity of green and ripe chinotto (*Citrus myritifolia* raf.) fruit tissues, leaves and seeds. *Food Chem.* 129, 1504–1512. doi: 10.1016/j.foodchem.2011.05.130

Bellocco, E., Barreca, D., Laganà, G., Leuzzi, U., Tellone, E., Kotyk, A., et al. (2009). Influence of l-rhamnosyl-D-glucosyl derivatives on properties and biological interaction of flavonoids. *Mol. Cell. Biochem.* 321 (1–2), 165–171. doi: 10.1007/s11010-008-9930-2

Benzie, I. F. F., and Strain, J. J. (1996). The ferric reducing ability of plasma (FRAP) as a measure of "antioxidant power": the FRAP assay. *Anal. Biochem.* 239 (1), 70–76. doi: 10.1006/abio.1996.0292

Bhatta, S., Janezic, T. S., and Ratti, C. (2020). Freeze drying of plant based foods. *Foods* 9 (87), 1–22. doi: 10.3390/foods9010087

Buyukkormaz, C., and Kucukbay, F. Z. (2022). Kumquat fruit and leaves extracted with different solvents: phenolic content and antioxidant activity. *Foods Raw Mater.* 10 (1), 51–66. doi: 10.21603/2308-4057-2022-1-51-66

Dhuique-mayer, C., Caris-veyrat, C., Ollitrault, P., Curk, F., and Amiot, M. J. (2005). Varietal and interspecific influence on micronutrient contents in citrus from the Mediterranean area. *J. Agric. Food Chem.* 53 (6), 2140–2145. doi: 10.1021/jf0402983

Floegel, A., Kim, D. O., Chung, S. J., Koo, S. I., and Chun, O. K. (2011). Comparison of ABTS/DPPH assays to measure antioxidant capacity in popular antioxidant-rich US foods. *J. Food Compost. Anal.* 24, 1043–1048. doi: 10.1016/j.jfca.2011.01.008

Gasecka, M., Siwulski, M., Magdziak, Z., Budzynska, S., Stuper-Szablewska, K., Niedzielski, P., et al. (2020). The effect of drying temperature on bioactive compounds and antioxidant activity of *Leccinum scabrum* (Bull.) Gray and *Hericium erinaceus* (Bull.) pers. *J. Food Sci. Technol.* 57, 513–525. doi: 10.1007/s13197-019-04081-1

Gupta, A. K., Dhua, S., Sahu, P. P., Abate, G., Mishra, P., and Mastinu, A. (2021). Variation in phytochemical, antioxidant and volatile composition of pomelo fruit (*Citrus grandis* (L.) osbeck) during seasonal growth and development. *Plants (Basel)* 10 (9), 1941. doi: 10.3390/plants10091941

Hayat, K., Zhang, X., Chen, H., Xia, S., Jia, C., and Zhong, F. (2010). Liberation and separation of phenolic compounds from citrus mandarin peels by microwave heating and its effect on antioxidant activity. *Separation Purification Technol.* 73 (3), 371–376. doi: 10.1016/j.seppur.2010.04.026

Karaman, S., Toker, O. S., Çam, M., Hayta, M., Doğan, M., and Kayacier, A. (2014). Bioactive and physicochemical properties of persimmon as affected by drying methods. *Drying Technol.* 32 (3), 258–267. doi: 10.1080/07373937.2013.821480

Kim, J. H., and Kim, M. Y. (2017). Phytochemical and antioxidant characterization of thinned immature *Citrus unshui* fruits. *Int. J. Pharm. Pharm. Sci.* 9 (12), 293–297. doi: 10.22159/ijpps.2017v9i12.22971

Kim, S. S., Kim, H.-J., Parl, K.-J., Kang, S.-B., Park, Y.-S., Han, S.-G., et al. (2022). Metabolomic profiling of *Citrus unshui* during different stages of fruit development. *Plants* 11 (967), 1–15. doi: 10.3390/plants11070967

Kim, D.-S., and Lim, S.-B. (2020). Extraction of flavanones from immature *Citrus unshui* pomace: process optimization and antioxidant evaluation. *Sci. Rep.* 10, 19950. doi: 10.1038/s41598-020-76965-8

Kumar, D., Ladaniya, M. S., and Gurjar, M. (2019). Underutilized citrus sp. pomelo (*Citrus grandis*) and kachai lemon (*Citrus jambhiri*) exhale in phytochemicals and antioxidant potential. *J. Food Sci. Technol.* 56, 217–223. doi: 10.1007/s13197-018-3477-3

Kumar, D., Ladaniya, M. S., Gurjar, M., and Kumar, S. (2022a). Impact of drying methods on natural antioxidants, phenols and flavanones of immature dropped *Citrus sinensis* l. osbeck fruits. *Sci. Rep.* 12(), 1–12. doi: 10.1038/s41598-022-10661-7

Kumar, D., Ladaniya, M. S., Gurjar, M., Kumar, S., and Mendke, S. (2021). Quantification of flavonoids, phenols and antioxidant potential from dropped *Citrus reticulata* blanco fruits influenced by drying techniques. *Molecules* 26 (4159), 2–13. doi: 10.3390/molecules26144159

Kumar, D., Ladaniya, M. S., Gurjar, M., Kumar, S., and Mendke, S. (2022b). Metabolic diversity of flavonoids and antioxidant potential for characterization of underutilized citrus species for nutritional security. *Plants* 11 (862), 1–12. doi: 10.3390/plants11070862

Lapuente, M., Estruch, R., Shahbaz, M., and Casas, R. (2019). Relation of fruits and vegetables with major cardiometabolic risk factors, markers of oxidation, and inflammation. *Nutrients* 11 (10), 1–45. doi: 10.3390/nu11102381

Lou, S.-N., and Ho, C.-T. (2016). Phenolic compounds and biological activities of small-size citrus: kumquat and calamondin. *J. Food Drug Anal.* 25, 162–175. doi: 10.1016/j.jfda.2016.10.024

Lou, S.-N., and Ho, C.-T. (2017). Phenolic compounds and biological activities of small-size citrus: kumquat and calamondin. *J. Food Drug Anal.* 25, 162–175. doi: 10.1016/j.jfda.2016.10.024

Lou, S. N., Hsu, Y. S., and Ho, C. T. (2014a). Flavonoid compositions and antioxidant activity of calamondin extracts prepared using different solvents. *J. Food Drug Anal.* 1-6. doi: 10.1016/j.jfda.2014.01.020

Lou, S.-N., Lai, T.-C., Huang, J.-D., Ho, C.-T., Ferng, L.-H. A., and Chang, Y.-C. (2015). Drying effect on flavonoid composition and antioxidant activity of immature kumquat. *Food Chem.* 171, 356–363. doi: 10.1016/j.foodchem.2014.08.119

Lou, S. N., Lin, Y. S., Hsu, Y. S., Chiu, E. M., and Ho, C. T. (2014b). Soluble and insoluble phenolic compounds and antioxidant activity of immature calamondin affected by solvents and heat treatment. *Food Chem.* 161, 246–253. doi: 10.1016/j.foodchem.2014.04.009

Lu, Y., Zhang, C., Bucheli, P., and Wei, D. (2006). Citrus flavonoids in fruit and traditional Chinese medicinal food ingredients in China. *Plant Foods Hum. Nutr. (Dordrecht Netherlands).* 61 (2), 57–65. doi: 10.1007/s11130-006-0014-8

Mansour, R. (2018). Determination of nutritional composition in citrus fruits (*C. aurantium*) during maturity. *Nutr. Food Sci.* 1–20. doi: 10.1108/NFS-05-2018-0141

Marten, S. (2007). *Determination of naringin and hesperidin in fruit juice (Application note ID-VDD2).* Berlin, Germany: Knauer.

Marti, N., Mena, P., Canovas, J. A., Micol, V., and Saura, D. (2009). Vitamin c and the role of citrus juices as functional food. *Nat. Prod. Commun.* 4 (5), 677–700. doi: 10.1177/1934578X0900400506

Mena, P., García-Viguera, C., Navarro-Rico, J., Moreno, D. A., Bartual, J., Saura, D., et al. (2011). Phytochemical characterisation for industrial use of pomegranate (*Punica granatum* l.) cultivars grown in Spain. *J. Sci. Food Agric.* 91 (10), 1893–1906. doi: 10.1002/jsfa.4411

Omidbaigi, R., and Nasiri, M. F. (2004). Quantitative distribution of hesperidin in *Citrus* species, during fruit maturation and optimal harvest time. *Nat. Prod. Rad.* 3 (1), 12–15.

Ozcan-Sinir, G., Ozkan-Karabacak, A., Tamer, C. E., and Copur, O. U. (2019). The effect of hot air, vacuum and microwave drying on drying characteristics, rehydration capacity, color, total phenolic content and antioxidant capacity of kumquat (*Citrus japonica*). *Food Sci. Technol. (Campinas).* 39 (2), 475–484. doi: 10.1590/fst.34417

Papoutsis, K., Pristijono, P., Golding, J. B., Stathopoulos, C. E., Bowyer, M. C., Scarlett, C. J., et al. (2017). Effect of vacuum-drying, hot air-drying and freeze-drying on polyphenols and antioxidant capacity of lemon (*Citrus limon*) pomace aqueous extracts. *Int. J. Food Sci. Technol.* 52, 880–887. doi: 10.1111/ijfs.13351

Rekha, C., Poornima, G., Manasa, M., Abhipsa, V., Pavithra Devi, J., Vijay Kumar, H. T., et al. (2012). Ascorbic acid, total phenol content and antioxidant activity of fresh juices of four ripe and unripe citrus fruits. *Chem. Sci. Trans.* 1 (2), 303–310. doi: 10.7598/cst2012.182

Rice-Evans, C. A., and Miller, N. J. (1996). Antioxidant activities of flavonoids as bioactive components of food. *Biochem. Soc Trans.* 24 (3), 790–795. doi: 10.1042/bst0240790

Ruiz, N. A. Q., Demarchi, S. M., and Giner, S. A. (2014). Effect of hot air, vacuum and infrared drying methods on quality of rose hip (*Rosa rubiginosa*) leathers. *Int. J. Food Sci. Technol.* 49 (8), 1799–1804. doi: 10.1111/ijfs.12486

Singleton, V. L., and Rossi, J. A. (1965). Colorimetry of total phenolics with phosphomolybdic-phosphotungstic acid reagents. *Am. J. Enol. Vitic.* 16, 144–158. doi: 10.5344/ajev.1965.16.3.144

Sun, Y., Qiao, L., Shen, Y., Jiang, P., Chen, J., and Ye, X. (2013). Phytochemical profile and antioxidant activity of physiological drop of citrus fruits. *J. Food Sci.* 78 (1), C37–C42. doi: 10.1111/j.1750-3841.2012.03002.x

Sun, Y., Shen, Y., Liu, D., and Ye, X. (2015). Effects of drying methods on phytochemical compounds and antioxidant activity of physiologically dropped unmatured citrus fruits. *LWT Food Sci. Technol.* 60, 1269–1275. doi: 10.1016/j.lwt.2014.09.001

Tripoli, E., La Guardia, M., Giammanco, S., Di Majo, D., and Giammanco, M. (2007). Citrus flavonoids: molecular structure, biological activity and nutritional properties: a review. *Food Chem.* 104, 466–479. doi: 10.1016/j.foodchem.2006.11.054

Turner, T., and Burri, B. J. (2013). Potential nutritional benefits of current citrus consumption. *Agriculture* 3, 170–187. doi: 10.3390/agriculture3010170

Vanamala, J., Reddivari, L., Yoo, K. S., Pike, L. M., and Patil, B. S. (2006). Variation in the content of bioactive flavonoids in different brands of orange and grapefruit juices. *J. Sci. Food Agric.* 19 (2), 157–166. doi: 10.1016/j.jfca.2005.06.002

Wojdylo, A., Figiel, A., Lech, K., Nowicka, P., and Oszmianski, J. (2013). Effect of convective and vacuum-microwave drying on the bioactive compounds, color, and antioxidant capacity of sour cherries. *Food Bioprocess Technol.* 1–13. doi: 10.1007/s11947-013-1130-8

Xu, G. H., Chen, J. C., Liu, D. H., Zhang, Y. H., Jang, P., and Ye, X. Q. (2008). Minerals, phenolic compounds, and antioxidant capacity of citrus peel extract by hot water. *J. Food Sci.* 73 (1), C11–C18. doi: 10.1111/j.1750-3841.2007.00546.x

Ye, X.-Q., Chen, J.-C., Liu, D.-H., Jiang, P., Shi, J., Xue, S., et al. (2011). Identification of bioactive composition and antioxidant activity in young mandarin fruits. *Food Chem.* 124, 1561–1566. doi: 10.1016/j.foodchem.2010.08.013

Zou, Z., Xi, W., Hu, Y., Nie, C., and Zhou, Z. (2016). Antioxidant activity of citrus fruits. *Food Chem.* 196, 885–896. doi: 10.1016/j.foodchem.2015.09.072

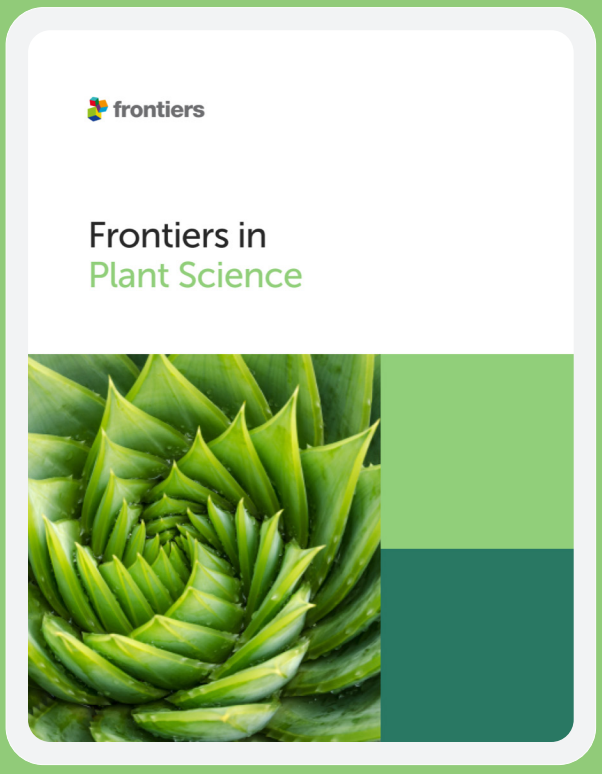
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

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