frontiers | Research Topics

Agro-morphological and nutritional profiling of crops

Edited by
Sapna Langyan, Tarun Belwal, Chunpeng Wan,
Pranjal Yadava and Tanushri Kaul

Published in
Frontiers in Nutrition
Frontiers in Food Science and Technology
Frontiers in Plant Science

FRONTIERS EBOOK COPYRIGHT STATEMENT

The copyright in the text of individual articles in this ebook is the property of their respective authors or their respective institutions or funders. The copyright in graphics and images within each article may be subject to copyright of other parties. In both cases this is subject to a license granted to Frontiers.

The compilation of articles constituting this ebook is the property of Frontiers.

Each article within this ebook, and the ebook itself, are published under the most recent version of the Creative Commons CC-BY licence. The version current at the date of publication of this ebook is CC-BY 4.0. If the CC-BY licence is updated, the licence granted by Frontiers is automatically updated to the new version.

When exercising any right under the CC-BY licence, Frontiers must be attributed as the original publisher of the article or ebook, as applicable.

Authors have the responsibility of ensuring that any graphics or other materials which are the property of others may be included in the CC-BY licence, but this should be checked before relying on the CC-BY licence to reproduce those materials. Any copyright notices relating to those materials must be complied with.

Copyright and source acknowledgement notices may not be removed and must be displayed in any copy, derivative work or partial copy which includes the elements in question.

All copyright, and all rights therein, are protected by national and international copyright laws. The above represents a summary only. For further information please read Frontiers' Conditions for Website Use and Copyright Statement, and the applicable CC-BY licence.

ISSN 1664-8714
ISBN 978-2-8325-3955-2
DOI 10.3389/978-2-8325-3955-2

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

Agro-morphological and nutritional profiling of crops

Topic editors
Sapna Langyan — National Bureau of Plant Genetic Resources, Indian Council of Agricultural Research (ICAR), India
Tarun Belwal — Texas A and M University, United States
Chunpeng Wan — Jiangxi Agricultural University, China
Pranjal Yadava — Indian Agricultural Research Institute (ICAR), India
Tanushri Kaul — International Centre for Genetic Engineering and Biotechnology, India

Citation
Langyan, S., Belwal, T., Wan, C., Yadava, P., Kaul, T., eds. (2023). *Agro-morphological and nutritional profiling of crops*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3955-2

Table of contents

frontiers | Frontiers in Food Science and Technology

TYPE Editorial
PUBLISHED 02 November 2023
DOI 10.3389/frfst.2023.1297763

Check for updates

OPEN ACCESS

EDITED AND REVIEWED BY
Isabel Sousa,
University of Lisbon, Portugal

*CORRESPONDENCE
Sapna Langyan,
singh.sapna06@gmail.com
Pranjal Yadava,
pranjal.yadava@icar.gov.in

RECEIVED 20 September 2023
ACCEPTED 13 October 2023
PUBLISHED 02 November 2023

CITATION
Langyan S, Belwal T, Wan CC, Yadava P and Kaul T (2023), Editorial: Agro-morphological and nutritional profiling of crops.
Front. Food. Sci. Technol. 3:1297763.
doi: 10.3389/frfst.2023.1297763

COPYRIGHT
© 2023 Langyan, Belwal, Wan, Yadava and Kaul. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Editorial: Agro-morphological and nutritional profiling of crops

Sapna Langyan[1]*, Tarun Belwal[2], Chunpeng Craig Wan[3], Pranjal Yadava[4]* and Tanushri Kaul[5]

[1]Indian Council of Agricultural Research-National Bureau of Plant Genetic Resources (ICAR-NBPGR), New Delhi, India, [2]Department of Horticultural Science, Texas A&M University, College Station, TX, United States, [3]School of Agricultural Sciences, Jiangxi Agricultural University, Jiangxi, China, [4]Indian Agricultural Research Institute (ICAR), New Delhi, India, [5]International Centre for Genetic Engineering and Biotechnology (India), New Delhi, India

KEYWORDS

agro-morphologic traits, nutrition, anti-nutritional, diversity, *germplasm* (genetic) resources

Editorial on the Research Topic
Agro-morphological and nutritional profiling of crops

The 21th century is witnessing rapid population growth, climate change, and shifting dietary patterns, leading to a great challenge to feed and nourish billions of people sustainably. The 2030 Agenda for Sustainable Development of the United Nations' aim is to mitigate hunger, achieve food security, improve nutrition, and promote sustainable agriculture and livelihoods. Science-based interventions leading to enhanced yield and quality of crops are central to this aim.

The spectrum is wide and captivating. The comprehensive review of Kaur et al. on the diversity of *Linum* genetic resources housed in global genebanks navigates through the agro-morphological characterization of *Linum* varieties and explores how novel genomic technologies are being harnessed to unlock the hidden potential of this versatile crop. *Linum*, commonly known as flax, has been cultivated for centuries for its fibre and oil. At present, approximately 61,000 germplasm accessions of *Linum*, including 1,127 wild accessions, are conserved worldwide.

Apart from the conservation of genetic resources, it is equally important to understand how to assess the nutritional content of foods accurately. Diverse bioanalytical methods, including chromatography, microscopic techniques, molecular assays, and metabolomics, are necessary for accurate nutritional profiling and the development of robust models. Underscoring this important need, Mondal et al. have reviewed the different bioanalytical techniques, the various protocols, and their application for the development and refinement of nutritional profiling models. Evidence-based studies show that the classification of foods depends not only on the nutrition composition but also on the distribution of food in our total diet.

The Research Topic also presents ten original research papers on various aspects of morphological and nutritional profiling across diverse crops. The Food and Agriculture Organization and United Nations have recognised 2023 as the International Year of Millets to raise awareness about the health and nutritional benefits of millets-a highly diverse group of small-seeded grasses grown mainly in stressed ecologies for fodder and human food. Wang et al. investigate how nitrogen deficiency affects the folate content of foxtail millet, shedding light on the nutritional consequences of nutrient limitations. The content of folate derivatives was studied in 29 diverse foxtail millet cultivars under two soil nitrogen regimes

(0 and 150 kg N ha^{-1}) to explore folate potential grown under low nitrogen soils. Another study on millets employs molecular docking and simulation-based approaches to probe the potential of bioactive compounds in millets.

Pulses are crucial for combating hidden hunger and malnourishment and enriching the cropping systems. Lentil (*Lens culinaris*) and mungbean (*Vigna radiata*) are two important pulse crops and a rich source of protein and essential nutrients. Aski et al. used advanced genetical statistics like Genotype, Genotype × Environment interactions and Additive Main effects and Multiplicative Interaction models to study genetic stability and Genotype × Environment interactions in 16 Indian lentil cultivars for important nutrition-related traits like iron, zinc, aluminium, and phytic acid content in grains. Their study identified promising cultivars that can be promoted for biofortification programs. The results also indicated potential for simultaneously increasing iron and zinc in lentils. In another study, Bansal et al. examine how different moisture regimes affect the seed nutritional quality of lentils. Water stress not only diminished lentil yield but also had an adverse impact on the quality of lentil grains. Consequently, when breeding for environments with limited water availability, it becomes imperative to factor in considerations for grain quality as well. Taking this kind of nutritional analysis in field-grown crops forward, Sinha et al. attempted to identify genomic regions and DNA variations associated with nutritional traits in mungbeans. They dissected the genetic architecture of grain iron, zinc, phytic acid, and tannin content in an association mapping panel of 145 diverse mungbean genotypes. Genotyping by sequencing identified genome-wide single nucleotide polymorphisms and candidate genes associated with these traits in a diverse selected panel of mungbean genotypes.

Rice is a staple food for over half of the world's population, making it a critical crop for global food security. The research article by Gautam et al. explores intra-varietal variation in a popular rice landrace grown in geographical isolation in the Andaman Islands of India. The variation was studied by employing 22 agro-morphological and biochemical traits. By harnessing the diversity within a single variety, researchers can devise unique strategies to improve agro-morphological and nutritional traits to ensure sustainable food production on tropical islands. The existence of intra-varietal variations could also be important from an evolutionary biology perspective.

Apart from the islands, the far-flung high-altitude regions of Indian Himalayas are also rich in unique bioresources. *Selinum carvifolia* is a medicinal plant useful in managing ailments such as hysteria and seizures. Srivastava et al. collected *S. carvifolia* samples from different altitudes (2,150 m–3,178 m) from the Chopta region of Uttarakhand, in the Indian Himalayas. Chromatographic analysis revealed different phenolic compounds, like chlorogenic acid, gallic acid, rutin, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid in *S. carvifolia.* They also analysed the cytotoxic effects and antibacterial and antifungal activity from different extracts. *S. carvifolia* extracts reduced the cell viability, indicating the anticancer potential, of this less explored medicinal plant.

Tropical pastures and the Indian grasslands system are species-rich and make up rangelands, forests, community lands, etc. that serve as one of the major roughage sources for ruminants. In India, *Cenchrus* is an important component of such grasslands. Singh et al. evaluated 79 accessions of *Cenchrus* genus belonging to six species and report wide variability for protein fibre, energy, sugar, and other nutritional traits. They also identified promising genotypes for good silage-making quality.

Apart from India, China is also a bioresource-dense country. This Research Topic presents two interesting studies from China spanning green tea and Chinese citrus varieties. Xiao et al. studied the possible mechanisms behind tea quality that change with harvest time in Wufengshan green tea grown in high-attitude mountains. They identified flavone and flavonol biosynthesis and phenylalanine metabolism as key determinants of green tea quality. Like tea, Citrus fruits are also renowned for their refreshing flavors and health benefits. Sun et al. evaluated the biological characteristics, phylogeny, and phytochemical profile, including antioxidant activity, of the two local citrus varieties.

Our journey through this Research Topic has been nothing short of illuminating, a testament to the relentless pursuit of knowledge and innovation in the world of agriculture and nutrition. The seeds of scientific inquiry sown in these articles hold the potential to yield a bountiful harvest, nourishing billions and sustaining our planet in the face of myriad challenges.

Author contributions

SL: Conceptualization, Data curation, Investigation, Methodology, Project administration, Supervision, Validation, Writing–original draft, Writing–review and editing. TB: Conceptualization, Data curation, Investigation, Methodology, Supervision, Writing–review and editing. CW: Conceptualization, Data curation, Investigation, Methodology, Software, Writing–review and editing. PY: Conceptualization, Data curation, Formal Analysis, Investigation, Methodology, Project administration, Software, Supervision, Validation, Writing–original draft, Writing–review and editing. TK: Investigation, Methodology, Software, Supervision, Validation, Writing–review and editing.

Conflict of interest

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

The author(s) declared that they were an editorial board member of Frontiers, at the time of submission. This had no impact on the peer review process and the final decision.

Publisher's note

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

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 06 January 2023
DOI 10.3389/fnut.2022.1079325

Check for updates

OPEN ACCESS

EDITED BY
Tarun Belwal,
Zhejiang University, China

REVIEWED BY
Chunpeng Wan,
Jiangxi Agricultural University, China
Yong-Quan Xu,
Tea Research Institute (CAAS), China

*CORRESPONDENCE
Haiyan Zhou
✉ xieyijie@stu.hunau.edu.cn

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

RECEIVED 25 October 2022
ACCEPTED 08 December 2022
PUBLISHED 06 January 2023

CITATION
Xiao H, Yong J, Xie Y and Zhou H (2023) The molecular mechanisms of quality difference for Alpine Qingming green tea and Guyu green tea by integrating multi-omics. *Front. Nutr.* 9:1079325. doi: 10.3389/fnut.2022.1079325

COPYRIGHT
© 2023 Xiao, Yong, Xie and Zhou. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

The molecular mechanisms of quality difference for Alpine Qingming green tea and Guyu green tea by integrating multi-omics

Hongshi Xiao[1,2], Jie Yong[1], Yijie Xie[1] and Haiyan Zhou[1]*

[1]College of Bioscience and Biotechnology, Hunan Agricultural University, Changsha, China,
[2]Agricultural and Rural Bureau of Hefeng County, Hefeng, China

Introduction: Harvest time represents one of the crucial factors concerning the quality of alpine green tea. At present, the mechanisms of the tea quality changing with harvest time have been unrevealed.

Methods: In the current study, fresh tea leaves (qmlc and gylc) and processed leaves (qmgc and gygc) picked during Qingming Festival and Guyu Festival were analyzed by means of sensory evaluation, metabolomics, transcriptomic analysis, and high-throughput sequencing, as well as their endophytic bacteria (qm16s and gy16s).

Results: The results indicated qmgc possessed higher sensory quality than gygc which reflected from higher relative contents of amino acids, and soluble sugars but lower relative contents of catechins, theaflavins, and flavonols. These differential metabolites created features of light green color, prominent freshness, sweet aftertaste, and mild bitterness for qmgc.

Discussion: Flavone and flavonol biosynthesis and phenylalanine metabolism were uncovered as the key pathways to differentiate the quality of qmgc and gygc. Endophytic bacteria in leaves further influence the quality by regulating the growth of tea trees and enhancing their disease resistance. Our findings threw some new clues on the tea leaves picking to pursue the balance when facing the conflicts of product quality and economic benefits.

KEYWORDS

Hefeng tea, transcriptomics, *Camellia sinensis*, endophytic bacteria, green tea

Introduction

Green tea is a well-known beverage native with health benefits and pleasant taste (1). The tea plant [*Camellia sinensis* (L.) O. Kuntze] grows widely in tropical and subtropical regions around the world, primarily in China, Japan, Argentina, Vietnam, India, and Kenya (2). With the growth of tea tree cultivation areas, a surge in processing products and exploitation is expected in the near future. In order to cater to the healthy

food market demand, it is necessary to develop tea products with local characteristics and high quality.

Hefeng green tea is regarded as the specialty of Enshi (Hubei Province, China), the selenium capital of the world. This particular tea is known for its excellent taste, color, and aroma, attributed to the ecological environment which is far from industrial pollution and has selenium-rich soil. Moreover, Hefeng tea also possesses some properties including anti-cancer, anti-aging, immunity, and fertility-enhancing.

The tea quality is positively correlated with the altitude of cultivation (3), and largely depends on the content of secondary metabolites such as flavonoids, phenolic acids, and alkaloids (4). Tea polyphenols, citric acid, theanine, and sucrose normally increase with altitude, while proanthocyanidin C1, theanine B, and catechins show a decrease with altitude (3). The high-altitude slows down the growth rate of Hefeng tea relative to other varieties in the same latitude, and definitely leads to delayed marketing dates and a competitive disadvantage. Moreover, some traditional planting patterns without pruning, pesticides, and fertilizers also further delayed its market timing, while preserving the original flavor of Hefeng green tea. All of these traditional methods enhance the quality of green tea. For instance, shade directly leads to the down-regulation of epigallocatechin gallate, catechin gallate, anthocyanins, and proanthocyanidins, resulting in marked enhancement in tea quality (5). It is also noteworthy that green tea processed from unpruned tea plants has better aroma and taste (6).

Furthermore, some researchers have claimed that the harvest time exerts a significant impact on the tea quality. One previous study indicated that in early spring tea leaves, the concentrations of amino acids (L-glutamine and L-tryptophan), (S)-(-)-limonene, catechins, and flavonol/flavone glycosides were higher, while the concentrations of proanthocyanidins (proanthocyanidin A1, protofibronectin A1, and protofibronectin A2 3′-gallate) were diminished compared to the control. There are also vast differences in the metabolic profiles of young tea leaves in early spring and late spring, which can be attributed to certain close-related biosynthetic pathways like flavonoids, phenylpropanoids, flavonoids, and flavonols, phenylalanine, tyrosine, and tryptophan (7). In one experiment (8), all of the early, middle, and late spring green teas at low altitudes were analyzed by gas chromatography-time-of-flight mass spectrometry, the results of which revealed that with decreasing concentration of amino acids, there was a strong enhancement in the concentration of carbohydrates, flavonoids, and their glycosides in the late spring season, which was feedbacked from the sensory quality of the tea leaves made.

In Hefeng (Enshi), fresh tea leaves picked at the Qingming Festival are generally used to make high-grade green tea, while fresh tea leaves picked at the Guyu Festival are generally used to make ordinary green tea. Both the Qingming Festival and the Guyu Festival belong to the 24 solar terms in the Chinese lunar calendar. The Qingming Festival falls around the 5th of April each year and 15 days earlier than Guyu Festival. This 15-day waves down the price of green tea sensitive.

In brief, it is necessary to probe into the relationship between the harvest time and the variation of key differential metabolites relative to the quality of alpine green tea. The Wufengshan green tea grown in high-attitude mountains was selected as the focus of the current study. Its fresh leaves collected at the Qingming Festival and Guyu Festival were processed into commercial products through a traditional process. To explain the changes in key metabolic pathways with harvest time, metabolomics was utilized to detect differential metabolites (DEMs) in the samples, and transcriptomics was adopted to identify differentially expressed genes (DEGs). Meanwhile, the effect of symbiotic bacteria on the quality of alpine green tea was discussed by high-throughput sequencing technology. Our study aims to objectively evaluate the quality difference between Qingming green tea and Guyu green tea at the molecular level to shed new light on the optimization of tea processing technology.

Materials and methods

Plant materials

Firstly, six Wufengshan green tea plants with 15-year-olds were divided into two groups. Fresh leaves were collected from Wufengshan tea plantation (110.091363◦E, 29.845475◦N, Hefeng, Hubei Province, China) at an altitude of 1,450 m on 5 April 2022 (11:00 a.m.) and 20 April 2022 (11:00 a.m.), respectively. Samples collected under aseptic conditions were wrapped in tin foil and labeled. A portion of the samples was rapidly frozen with liquid nitrogen for 15 min and stored at −80°C for subsequent experiments. The remaining portion of the samples was prepared directly to make green tea according to the local traditional process as follows: spreading → killing → kneading → initial drying → shaping and extracting hairs → full drying → aroma. One part of processed samples was stored at 4°C for sensory evaluation, and the other part was stored at −80°C for subsequent experiments.

Sensory evaluation

Sensory evaluation of Qingming green tea (qmgc) and Gu Yu green tea (gygc) was performed independently by five professional tasters with methodological reference (9). The score was 100 points as follows: appearance color (25%), brew color (10%), aroma (25%), taste (30%), and infused leaves (10%).

Metabolomic analysis

Twenty milligrams of freeze-dried samples were added to 1,000 μl of extract (containing 70% methanol and isotope-labeled internal standard mixture), ground at 35 Hz for 4 min,

and sonicated in an ice-water bath for 5 min. The samples were subsequently incubated at −40°C for 1 h and then centrifuged at 4°C for 15 min at 12,000 rpm. The supernatant and QC samples (an equal mixture of all samples) were collected for metabolomics assay.

The chromatographic separation of the target compounds was performed using a Vanquish (Thermo Fisher Scientific) ultra-performance liquid chromatograph and a Waters ACQUITY UPLC HSS T3 (2.1 mm × 100 mm, 1.8 μm) liquid chromatographic column. A phase of the liquid chromatography was aqueous (containing 5 mmol/L ammonium acetate and 5 mmol/L acetic acid), and B phase was acetonitrile. Sample tray temperature: 4°C, injection volume: 2 μl. An Orbitrap Exploris 120 mass spectrometer (Xcalibur, version: 4.4, Thermo Fisher Scientific) was adopted for mass spectrometry data acquisition. The raw data were processed by peak identification, peak extraction, peak alignment, and integration and then matched with BiotreeDB (V2.1) secondary mass spectrometry database for substance annotation.

Transcriptomic analysis

Total RNA content was extracted from fresh tea samples using the Trizol reagent (Thermo Fisher Scientific, 15596018). The obtained RNA quantity and purity were analyzed with Bioanalyzer 2100 and RNA 6000 Nano LabChip Kit (Agilent, CA, USA, 5067-1511), respectively. Next, the mRNA was purified using Dynabeads Oligo (Thermo Fisher Scientific, CA, USA). Library construction was subsequently performed with the VAHTS™ Stranded mRNA-seq Library Prep Kit for Illumina® (Nanjing Novaseq NR601-01). Afterward, the cDNA libraries were sequenced using the Illumina Novaseq™ 6000 system.

The raw images obtained from high-throughput sequencing were transformed into raw sequences by means of base calling analysis, and the clean data were collected by filtering out the unqualified sequences with the Cutadapt tool and pre-processed. The pre-processed valid data was then compared with the reference genome ("shuchazao") using Hisat2. Based on the results of Hisat2 alignment, transcripts were reconstructed using the StringTie transcript assembler, followed by calculation of the expression levels of all genes in each sample. DESeq2 software was adopted for differential gene expression analysis between the two groups.

Transcripts of genes related to flavonoid biosynthesis were randomly selected for qRT-PCR validation with GAPDH serving as the internal reference gene. qRT-PCR reaction parameters were as follows: 95°C for 10 min, 94°C for 10 s, 58°C for 15 s, for a total of 45 cycles. Fluorescence intensity was detected using the LightCycler 480 system (Roche, Sussex, UK), and then the relative expression values of genes were calculated.

Diversity analysis of endophytic bacteria

Genomic DNA content was extracted using the PowerSoil® DNA isolation kit under aseptic conditions, and the quality of DNA extraction was determined by 1% agarose gel electrophoresis. PCR amplification was performed using primers specific for the V7-V9 region of the 16S rRNA gene: 799F-1193R 5′-AACMGGA TTAGATACCCKG-3′ and 5′-ACGTCATCCCCACCTTCC-3′. The PCR reaction conditions

FIGURE 1
Sensory evaluation of Qingming tea and Guyu tea. **(A)** Appearance of qmlc. **(B)** Appearance of qmgc. **(C)** Brew color of qmgc. **(D)** Infused leaves of qmgc. **(E)** Appearance of gylc. **(F)** Appearance of gygc. **(G)** Brew color of gygc. **(H)** Infused leaves of gygc.

were as follows: 95°C for 5 min; 95°C for 1 min, 50°C for 1 min, 72°C for 1 min, for a total of 35 cycles; 72°C for 7 min. The amplification products were detected using 1% agarose gel electrophoresis and sent to the Illumina NovaSeq platform for sequencing.

Following sequencing, the data were spliced, quality-controlled, and chimera filtered by overlap to obtain high-quality clean data. Single-base precision representative sequences were obtained using the DADA2 algorithm, and then ASVs (Amplicon Sequence Variants) were adopted to construct class OTUs (Operational Taxonomic Units), and to obtain the final ASV feature table as well as the feature sequences for further diversity analysis, species taxonomic annotation, and difference analysis, etc.

Data analysis and statistics

Statistical analyses were processed using the SPSS software. Diagrams were drawn by OriginPro 2018 and Adobe Illustrator CC 2019.

Results and discussion

Differences between qmlc and gylc products from sensory perception

As illustrated in **Figure 1** and **Supplementary Table 1**, there were obvious phenological differences between the two kinds of green tea. Most of the qmlc leaves presented with white fluffy buds or a small amount of one leaf and one bud in the length of 30 mm, otherwise the green leaves were slightly spreading or not spreading. Meanwhile, the vast majority of gylc leaves had two leaves and one bud or a small amount of one leaf and one bud in the length of 60 mm, and the leaves appeared translucent yellow-green coloration in sunlight.

Furthermore, in regard to appearance color, aroma, taste, brew color, and infused leaves of gygc and qmgc, it was found that qmgc was of better quality with a higher total score, especially in appearance color. In addition, qmgc retained the straight shape and was further covered by silvery-white hairs with a light green shiny color. Existing studies indicate that these hairs contribute to the defense of the tea plant, the flavor, and nutritional quality of leaves (10). Gygc possesses a curved shape, and is covered by a few silvery white hairs with a dark green oily color. Meanwhile, in regard to the aroma, gygc was found to be better than qmgc. qmgc was dominated by soft floral and fruity aromas, while gygc exhibited a strong and persistent honeysuckle aroma and chestnut aroma. The latter differences are important as aroma is regarded as one of the most crucial factors in evaluating tea quality, and high-quality green tea often emits clean or chestnut aromas (11). Moreover, in terms of taste, qmgc was pure, tasty, and sweet with a lighter flavor. On the other hand, gygc had a strong flavor with sweetness, but also more pronounced bitterness. Further in regard to brew color, qmgc was bright green, while gygc was bright greenish-yellow. On the infused leaves, qmgc could stand upright in the glass for a short time and spread naturally after fully absorbing water without producing crumbs. Although, gygc leaves could not fully spread after absorbing water and some leaves were mutilated, while some exhibited yellowish coloration. The above findings indicated that both green teas could meet the standard of Hefeng green tea production,[1] and qmlc was more suitable for processing as high-grade handmade green tea.

Differences of secondary metabolism products from metabolomics analysis

Metabolites serve as reflectors of the physiological state of tea plants, and DEMs are the direct cause of quality differences between qmgc and gygc. A prior study observed 20,971 peaks from the data of the Q Exactive LC-MS/MS platform in positive ion mode (POS). PCA analysis in our study (**Figure 2A**) illustrated that all samples fell into the 95% confidence interval. When samples were concentrated within groups and dispersed between groups, QC samples were tightly clustered, emphasizing that the above experimental method is reliable. In addition, the reliability of the model was further validated with the OPLS-DA tools (**Figures 3A, B, D, E**). Subsequent results demonstrated that the samples were all within the 95% confidence interval, indicating that the study model had good predictability and repeatability with no overfitting. The classification of 1,374 metabolites identified in this experiment comprised 478 lipids and lipid-like molecules, 209 organoheterocyclic, 208 phenylpropanoids and polyketides, 140 organic acids and derivatives, 114 benzenoids, 93 organic oxygen compounds, and others (**Figure 2C**).

The screening conditions for the differential metabolites in the current study were VIP greater than 1 and *P*-value less than 0.05. As illustrated in **Figure 2D**, the characteristic differential metabolites of the three groups compared (gygc vs. qmgc, qmgc vs. qmlc, and gylc vs. qmlc) were 78, 86, and 132, respectively, suggesting that the samples underwent very active chemical reactions in each of three conditions, and the picking time and processing would have an impact on the quality of "Hefeng tea."

Quality differences between gygc and qmgc from DEMs

In the volcano plot of 592 differential metabolites of gygc and qmgc (**Figure 3C**), top 15 differential metabolites with up- and down-regulation folds can be adopted as marker

1 https://www.cnki.com.cn/Article/CJFDTotal-SPZH201222038.htm

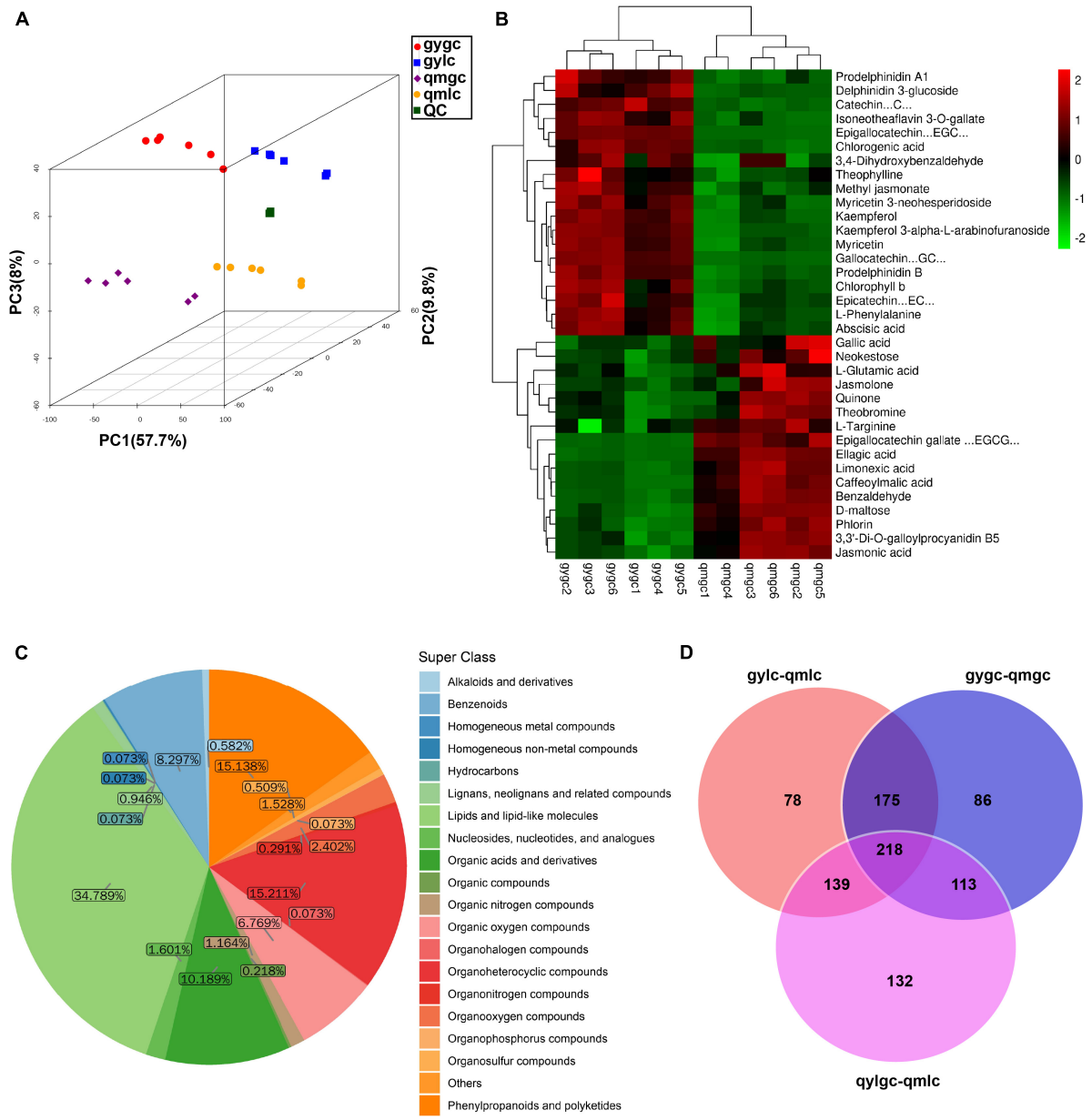


FIGURE 2
The metabolomic analysis of tea samples. **(A)** Score scatter plot for PCA model total with QC. **(B)** Key DEMs determining the difference in quality between gygc and qmgc. **(C)** Metabolites detected in all tea samples. **(D)** Venn diagram of metabolite statistics for differences between tea sample groups.

compounds for the identification of qmgc and gygc (Figure 3G). It is noteworthy that the chemical classification of 30 DEMs was diverse, which established that the leaves-picking time was responded to multiple metabolic pathways. Six flavonoids exhibited a noticeable alteration with the up-regulation of luteolin 7-glucoside, maysin, theadibenzotropolone A, in addition to the down-regulation of comosin, daidzein, and egonol glucoside. Meanwhile, three terpenoids significantly altered the up-regulation of alpha-dihydroartemisinin, an artemisinin that treats malaria. Terpene volatiles has been previously identified as the main contributors to floral aroma, and are further known to influence the sensory quality of green tea (12). Furthermore, enilconazole, a fungicide widely used in agriculture, especially in the cultivation of citrus fruits, was previously documented to be markedly reduced in gygc (13).

The significant variations between the 30 DEMs provided an initial insight into the differences between the two green teas. Figure 2B further illustrates the mechanism how the DEMs

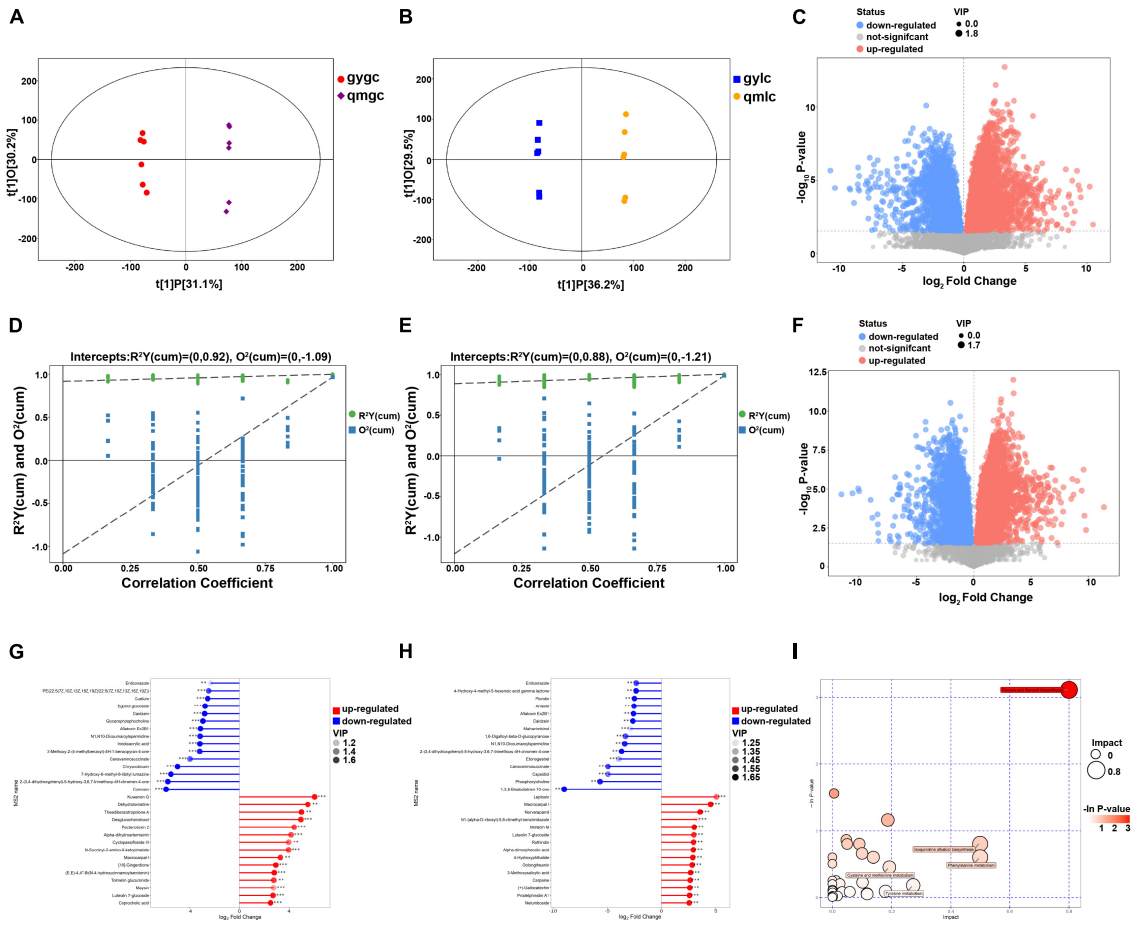


FIGURE 3
Metabolomic analysis of gylc and qmlc. **(A)** Score scatter plot of OPLS-DA model for group gygc vs. qmgc. **(B)** Score scatter plot of OPLS-DA model for group gylc vs. qmlc. **(C)** Volcano plot for group gylc vs. qmlc. **(D)** Permutation plot test of OPLS-DA model for group gygc vs. qmgc. **(E)** Permutation plot test of OPLS-DA model for group gylc vs. qmlc. **(F)** Volcano plot for group gylc vs. qmlc. **(G)** Matchstick analysis for group gygc vs. qmgc. **(H)** Matchstick analysis for group gylc vs. qmlc. **(I)** Pathway analysis for group gylc vs. qmlc.

(which included catechins, anthocyanins, theaflavin, phenolic acids, etc.) determine green tea quality.

Catechins

Catechins are a type of phenolic compounds very abundant in green tea and have a bitter and astringent taste. In our study, among the five monomeric catechins, namely catechin (C), epicatechin (EC), epigallocatechin (EGC), and gallocatechin (GC), were up-regulated, while epigallocatechin gallate (EGCG) was down-regulated. Accumulating works have shared similar conclusions with our study (5). For instance, young leaves possess more catechins compared to mature leaves. Moreover, a high level of catechins was previously associated with the strong floral aroma of "Xinyang Maojian." Meanwhile, the longer daylight and increased light intensity from Qingming Festival to Guyu Festival (15 days) are known to promote catechin content in tea leaves, whereas gallocatechin, epicatechin gallate, and catechin gallate content in tea plant healing tissues decreases. Furthermore, catechin and epicatechin also contribute to the sweet aftertaste of green tea (14).

Anthocyanins

Additionally, we found that the relative content of anthocyanins, prodelphinidin A1, prodelphinidin B, delphinidin 3-glucoside, and 3,3′-di-O-galloylprocyanidin B5 were all significantly down-regulated in gygc. This was in contrast to previous studies, wherein the levels of proanthocyanidin A1, prodelphinidin A1, and prodelphinidin A2 3′-gallate were much higher in late spring tea compared to early spring tea (15). The latter could be attributed to the fact that our samples were collected from high-altitude rather than low-altitude areas. Nevertheless, some of our findings are in accordance with previous reports (16). The high-grade "Huangshan Maofeng" exhibited more proanthocyanidins than the low-grade kind. Existing reports suggest that the accumulation of anthocyanins gave the leaves a purple color,

and also greatly enhances the bitterness of green tea (17). We speculate that for qmgc, the dark greenish oily color and bright green soup may be attributed to the high anthocyanin content.

Theaflavin

Theaflavin is responsible for the yellowing of the tea broth, and further contribute to the astringency and aftertaste of tea leaves (18). Our findings revealed that gygc presented with up-regulation of isoneotheaflavin 3-O-gallate, theaflavin, in addition to down-regulation of quinone. Meanwhile, quinone is an intermediate product in the oxidation of catechins to theaflavin, while we learned that the tea broth of gylc was yellowish, presumably the catechins were being converted from quinone to theaflavin at that time.

Phenolic acids

Gallic acid, ellagic acid, and chlorogenic acid are essential for the synthesis of flavonols and catechins in tea plants. In our study, we found that the acids of gallic, chlorogenic, and ellagic were all present at higher levels in qmgc. Gallic acid is further associated with astringency, sourness, and sweet aftertaste of green tea (19), while young and tender parts normally possess higher gallic acid content especially in early spring teas (7, 15). Chlorogenic acid has also been identified as a flavor modifier for high-quality products and to improve sensory quality (20). Furthermore, ellagic acid is regarded as a marker of high-quality white tea (21).

Alkaloids

Major well-known alkaloids in tea include caffeine, theobromine, and theophylline, of which caffeine is the most dominant alkaloid in tea, accounting for more than 90% of the total alkaloids and the primary source of bitterness in green tea (22). Surprisingly, it has been uncovered that caffeine is not the cause of the bitterness difference between qmlc and gylc since statistical analysis revealed that caffeine was not present in the DEMs of gylc vs. qmlc. This particular finding is inconsistent with one previous report (15), wherein higher levels of caffeine were documented in the young parts of tea plants. In addition, theobromine and theophylline were up-regulated and down-regulated respectively.

Flavonols and flavonol glycosides

Flavonols and flavonol glycosides are astringent compounds and serve to enhance the bitterness in green tea (14). The relative contents of kaempferol, kaempferol 3-alpha-L-arabinofuranoside, myricetin, and myricetin 3-neohesperidoside were found to be markedly up-regulated in gygc. Fermentation process helps reduce the content of kaempferol-o-glucosides, leading to astringency reduction (22). One previous study also found that kaempferol-glucose-rhamnose-glucose in low-grade green tea was present at a much higher concentration than in the high-grade kind (16).

Free amino acids

Free amino acids that are involved in the formation of aroma substances impart green tea with a refreshing taste, such that their presence can be an indicator of tea quality (23). L-theanine is one such amino acid in tea and is known to decrease in content with shoot maturation (15). In our study, we uncovered that the content of L-theanine remained at a stable level. Meanwhile, the up-regulation of L-phenylalanine is regarded as the precursor for flavonoid synthesis, which eventually produces catechins, anthocyanins, and flavonol glycosides. In addition, L-phenylalanine can augment the astringency and bitterness (18), but also positively correlated with the freshness of green tea (24). Additionally, in our study, more freshness was detected when the content of L-glutamic acid was up-regulated in qmgc. This is in accordance with a previous study in which higher levels of L-glutamic acid were noted in early spring tea than late spring tea (7). On the other hand, when L-targinine was down-regulated in gygc, there was a decrease in the bitterness of tea broth (25).

Soluble sugars

Additional experimentation in our study revealed that the contents of D-maltose, neokestose, and phlorin were all down-regulated in gygc. These soluble sugars have a sweet taste, and have also previously been shown to be effective in alleviating the bitterness of tea broths (14). A similar study found that early spring tea accumulated a large number of sugars and sugar alcohols, which contribute to an increase in the quality of green tea (26). Herein, we hypothesized that the decrease in soluble sugar of gygc was caused by unfolded leaves. In that case, net photosynthesis was negative and soluble sugars were heavily consumed for cellular energy supply leading to decreased soluble sugar of gygc. Furthermore, the sensory analysis support that the sweetness was more pronounced in qmgc.

Organic acids

Organic acids including caffeoylmalic acid, limonexic acid, and jasmonic acid were all down-regulated in gygc, and further contributed to the sour and fruity flavors of qmgc, which is much in accordance with a recent study (7). Moreover, our findings also pointed out that the accumulation of organic acids in early spring tea could effectively resist the invasion of pathogenic bacteria and diminish the use of pesticides to improve the quality of tea leaves. Meanwhile, abscisic acid, an important phytohormone, was significantly up-regulated, which affected the quality of green tea by acting on lipid and flavonoid metabolism (27).

Other compounds

The contents of methyl jasmonate and jasmolone were all found to be down-regulated in gygc. Methyl jasmonate is a phytohormone involved in plant defense under adversity conditions, and possesses a floral, creamy aroma (23).

Meanwhile, Jasmolone is associated with the enhanced sweetness of green tea (1).

Furthermore, the contents of benzaldehyde and 3,4-dihydroxybenzaldehyde were both diminished in gygc. Normally, these compounds are associated with a specific almond odor in tea products (28). Benzaldehyde has also been previously shown to create a distinct herbal aroma (29).

The content of chlorophyll b was found to be up-regulated in gygc, and further contributed significantly to the transition from light green to yellow-green color in qmlc and gylc. Chlorophyll b represents a yellow-green photosynthetic pigment and its up-regulation is associated with enhanced photosynthesis in tea plants.

The above findings highlighted that qmgc and gygc possess varying metabolomic profiles, and the quality differences between them were determined by a variety of complex interactions of compounds including catechins, free amino acids, alkaloids, and flavanols. In addition, the metabolic profile of qmgc described high-quality green tea with light green soup, prominent fresh flavor, sweet aftertaste, and low bitterness and astringency, which was highly consistent with the sensory evaluation of qmgc and gygc. However, in our study, the pattern of metabolite variation from the samples collected at different picking times did not exactly match the previous Low altitude area research.

Metabolic differences between gylc and qmlc

The processing of green tea must be fried at high temperatures and the chemical reactions are mediated non-enzymatically (23). Therefore, it is helpful to understand the quality differences of corresponding commercial green tea by discussing the differential metabolites and key metabolic pathways of gylc and qmlc. As illustrated in **Figures 3F, H**, a total of 30 significant DEMs were directly related to green tea quality such as flavonoids, alkaloids, and amino acids. Among them, flavonoids and alkaloids usually aid plants adapt to their environment, and large amounts of secondary metabolites may be detected from samples grown in an abnormal environment which adds certain stressors to the plants.

Furthermore, we found that picrotin was down-regulated and its content was positively correlated with altitude. The reason may be attributed to the increased temperatures during rainy season, and that tea plants have not to produce large amounts of picrotin to combat low-temperature stress.

Additionally, previous studies have shown that gallocatechin and oolongtheanin, a dimeric catechin formed from epigallocatechin and epigallocatechin gallate with a bitter taste, are present in much higher levels in gylc (30). Moreover, we found that the content of prodelphinidin A1 was significantly lower in gylc, which is in accordance with the works of Zeng et al. (7).

Interestingly, the contents of some compounds that functioned as antibiotics had significant changes. Capsaidiol is a natural fungicide found in peppers formed *via* the isoprenoid pathway from 5-epi-aristolochene, and exerts an antibacterial effect on *Helicobacter pylori* (31). Moracin M is a phytoantitoxin isolated from *Morus alba* infected *Fusarium solani* (32). In addition, enilconazole and capsaidiol were down-regulated in gylc, whereas moracin M was up-regulated. These changes attracted the notice of the symbiotic microorganisms which appear to respond to the timing of fresh tea picking. Additional analysis of the DEMs of gylc and qmlc (**Supplementary Table 1**) revealed that lycoperdic acid and citrinin were up-regulated, while nivalenol and validamycin B were down-regulated in gylc. The synthesis of lycoperdic acid imparts a pungent odor and antibacterial properties to the sporophore of *Lycoperdon perlatum* (33). The study performed by Bryla et al. found that Nivalenol produced by *Fusarium graminearum* (34) was down-regulated greatly. Meanwhile, validamycin B, an agricultural antibiotic isolated from *Streptomyces hygroscopicus* (35), has also been shown to be down-regulated. On the other hand, citrinin, a fungal toxin with a wide range of biological activities *in vitro* (36) was up-regulated.

As illustrated in **Figure 3I**, a comprehensive analysis (including enrichment analysis and topological analysis) for the pathways of DEMs was performed by MetPA. Five key pathways in qmlc and gylc were screened with a high correlation to DEMs as follows: flavone and flavonol biosynthesis (p = 0.044592, impact = 0.8), isoquinoline alkaloid biosynthesis (p = 0.44741, impact = 0.5), phenylalanine metabolism (p = 0.54683, impact = 0.5), cysteine and methionine metabolism (p = 0.63443, impact = 0.19231), and tyrosine metabolism (p = 0.83274, impact = 0.27273). These key metabolic pathways involved the synthesis of flavonoids, alkaloids, and amino acids of tea plants and the biosynthesis of flavonoid, phenylpropanoids, flavone, flavonol, tyrosine, and tryptophan, which were important pathways in early and late spring teas (7), especially flavone and flavonol biosynthesis serve as a key pathway to differentiate tea quality from early, mid and late spring timing (8). From the above findings, the metabolic profiles provided some insight into the relevant physiological properties of fresh tea leaves and the foundation to select the best pick timing.

Transcriptomic information of gylc and qmlc

To further elucidate the molecular mechanisms underlying the quality differences between gygc and qmgc, RNA-Seq analysis was utilized to investigate the differences between gylc and qmlc. In the high-quality data (**Supplementary Table 2**) obtained from transcriptome sequencing, the raw database and

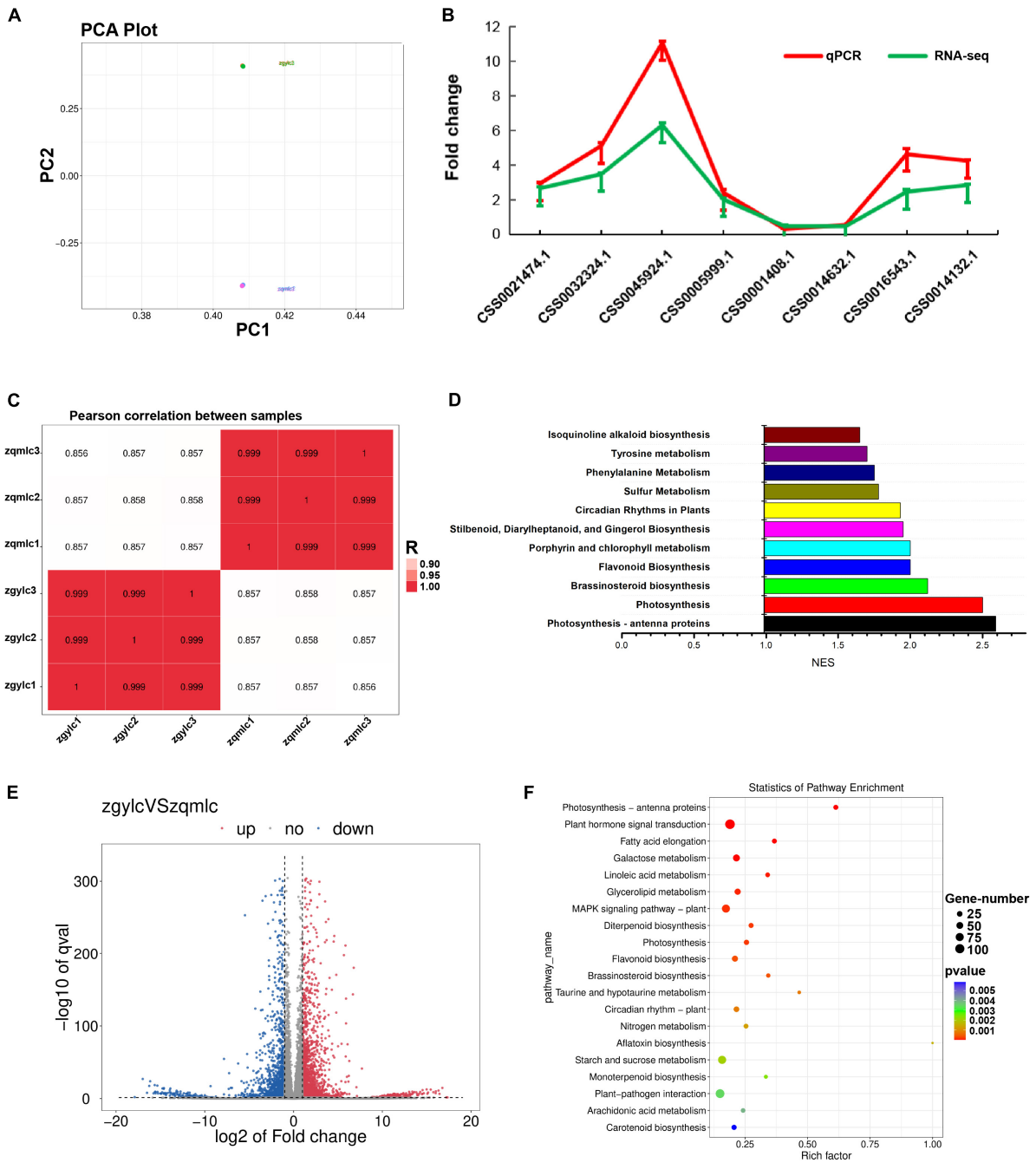


FIGURE 4
Results of transcriptomic analysis of gylc vs. qmlc. **(A)** PCA plots of qmlc and gylc transcriptomic analyses. **(B)** Results of qPCR validation of transcriptomic data. **(C)** Pearson correlation coefficient plot of zgylc vs. zqmlc. **(D)** GSEA enrichment results for zgylc vs. zqmlc. **(E)** Gene expression volcano plot of zgylc vs. zqmlc. **(F)** KEGG enrichment results of zgylc vs. zqmlc.

valid database were 43.04 G and 42.06 G, while the accuracy of sequencing Q20 and Q30 were 99.96 and 97.64%. The clean reads were compared with the genome of "Shuchazao" and the valid data covered 87.94% of the reference genome, indicating that the transcriptome data were reliable and could be used for subsequent analyses.

Results of PCA analysis, which could prove the biological replicates of qmlc and gylc, indicated that gene expression was closely related to the phenotypic changes in the samples (Figure 4A). From the Pearson correlation coefficient plot, the correlation of the within-group samples was higher than that of the between-group samples (Figure 4C). For the differential

gene expression analysis (| log2FC| ≥ 1 and $q < 0.05$ as the criteria), a total of 4,018 genes were considered differentially expressed and of which 2,327 genes were up-regulated, while 1,791 genes were down-regulated (Figure 4E). To validate the reliability of the RNA-Seq results, eight DEGs related to flavonoid metabolism were utilized for RT-qPCR analysis (Figure 4B, Supplementary Figures 1, 2, and Supplementary Table 3) and the results of which revealed that the expression patterns of these genes were highly consistent with those of RNA-Seq analysis.

According to the differential gene KEGG enrichment analysis (Figure 4F), a large number of pathways associated with tea plant growth and development were enriched, which was in perfect agreement with the phenotypes of qmlc and gylc (Figure 1). From Qingming to Guyu Festival (15 days), sunlight duration became longer and tea plants responded to circadian rhythm—plant. Moreover, the activation of the MAPK signaling pathway helped tea leaves with spreading and greening. Plant hormone signal transduction and brassinosteroid biosynthesis provided hormones for the growth of tea shoots. Photosynthesis, photosynthesis-antenna proteins, and carotenoid biosynthesis are further known to enhance photosynthesis activity in young leaves. Some researchers (37) have suggested that the gradual increase of chlorophyll is involved in carbon fixation, and the photosynthesis would differ from the accumulation of flavor metabolites and the tea quality. Meanwhile, enhanced photosynthesis is also known to facilitate the accumulation of free amino acids and aroma components in tea leaves (6). Besides, previous studies have shown that the biosynthesis of flavonoid, diterpenoid, and monoterpenoid, which is closely related to secondary metabolism and flavor quality of tea leaves, were enriched. In addition, pathways related to plant disease resistance, such as plant-pathogen interaction, aflatoxin biosynthesis were also strengthened, which highlighted that the effect of endophytic bacteria on the quality of gygc and qmgc cannot be ignored.

Since KEGG enrichment analyses results were conflicting with the data from DEMs. GSEA (gene set enrichment analysis) was utilized to compensate for the lack of effective information mining of micro-effective genes by traditional enrichment analysis. The screening criteria for GSEA analysis were | NES| > 1, NOM p-val < 0.05, and FDR q-val < 0.25. Subsequent results (Figure 4D) showed that flavonoid biosynthesis (| NES| = 2, NOM p-val = 0.00, FDR q-val = 0.00), isoquinoline alkaloid biosynthesis (| NES| = 1.65, NOM p-val = 0.014, FDR q-val = 0.184), tyrosine metabolism (| NES| = 1.73, NOM p-val = 0.00, FDR q-val = 0.00) and phenylalanine metabolism (| NES| = 1.75, NOM p-val = 0.00, FDR q-val = 0.00), all of which were significantly enriched. Together, these findings highlighted the metabolic pathways (phenylalanine metabolism and flavone, flavonol biosynthesis) as the key pathways for the quality difference of gylc and qmlc.

Herein, for further exploring more about the key metabolism relative to quality differences of qmgc and gygc, the transcriptomic and metabolomic data of the samples were combined. Subsequently, correlations were calculated by randomly selecting differential genes and differential metabolites using the "spearman" algorithm. As shown by the heat map (Supplementary Figure 3), the differential metabolites of catechin, procyanidin, and gallic acid were strongly associated with cinnamic acid hydroxylase (C4H), flavonol synthase (FLS), and hydroxycinnamoyltransferase (HCT).

Additionally, Figure 5 illustrates the DEMs and DEGs changes in phenylalanine metabolism, flavone, and flavonol biosynthesis, and it can be figured out that the biosynthesis of flavonoids begins with phenylalanine. With the participation of phenylalanine ammonialyase (PAL), cinnamic acid hydroxylase (C4H), and coumadin CoA ligase (4CL), phenylalanine, there was a production of the important intermediate 4-Coumaroyl-CoA which provides the precursor for the subsequent biosynthesis of flavonoids. Phenylalanine and all eight DEGs encoding PAL, C4H, and 4CL were significantly up-regulated in gylc, and it can be speculated that this may contribute to the increase in flavonoid abundance of gylc. PAL has also been reported to serve as the key enzyme for catalyzing the production of cinnamic acid from phenylalanine and its expression corresponds to the catechin content in tea plants (38). 4-Coumaroyl-CoA produced naringenin in the presence of chalcone synthase (CHS) and chalcone isomerase (CHI). Naringenin is a known stable intermediate in flavone and flavonol biosynthesis, and naringenin provides the basic carbon skeleton for flavonoid synthesis (39). Naringenin was down-regulated in gylc and all four DEGs encoding CHS and CHI were up-regulated, which indicated gene expression had opposite trending to the metabolites. The reason may be the increased subsequent flavonoid synthesis led to naringenin depletion.

Naringenin would further produce dihydrokaempferol with the action of flavonoid 3-hydroxylase (F3H), while dihydrokaempferol produced three flavonols along three separate pathways. The flavonoid 3′-hydroxylase (F3′H) and flavonoid 3′ 5′-hydroxylase (F3′5′H) catalyze the biosynthesis of dihydroquercetin and dihydromyricetin respectively. In addition, they further produced quercetin, kaempferol, and myricetin in the presence of flavonol synthase (FLS). These three flavonols would transfer into flavonol glycosides with the action of UDP glucose-flavonoid 3-o-glcosyl-transferase (UFGT). When kaempferol was up-regulated, its flavonol glycosides kaempferol 3-alpha-L-arabinofuranoside and kaempferol 3-O-alpha-L-rhamnofuranoside were down-regulated. On the other hand, when quercetin was down-regulated, its flavonol glycoside quercetin-3′-glucuronide was up-regulated and when myricetin was up-regulated, its flavonol glycoside myricetin 3-galactoside was up-regulated. From the encoding genes, there were no significant changes in the F3′H encoding gene.

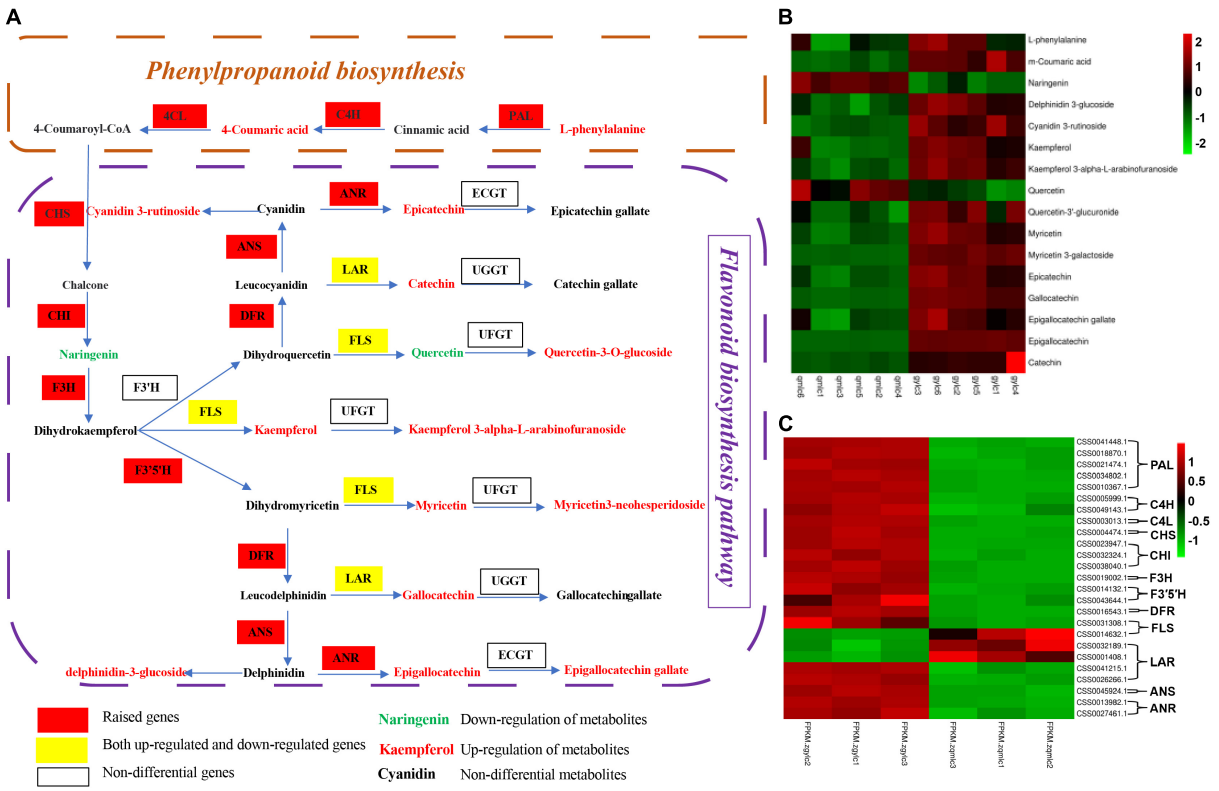


FIGURE 5
The key metabolic pathways of gylc and qmlc. **(A)** Diagram of the key metabolic pathways of qmlc and gylc. PAL, phenylalanine ammonia-lyase; C4H, cinnamic acid hydroxylase; 4CL, coumadin CoA ligase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavonoid 3-hydroxylase; F3′H, flavonoid 3′-hydroxylase; F3′5′H, flavonoid 3′ 5′-hydroxylase; DRF, dihydroflavonol 4-reductase; LAR, leucoanthocyanidin reductase; UFGT, UDP glucose-flavonoid 3-o-glcosyl-transferase; ANR: anthocyanidin reductase; ANS: anthocyanidin synthase; FLS: flavonol synthase. **(B)** Key metabolic pathways involve DEMs. **(C)** Key metabolic pathways involve DEGs.

Genes of *CSS0014132.1* and *CSS0043644.1* encoding F3′5′H were up-regulated and *CSS0048887.1* was down-regulated. Genes of *CSS0045924.1* and *CSS0031308.1* encoding FLS were up-regulated but *CSS0014632.1* was down-regulated. Meanwhile, there were no significant expression differences in the UFGT encoding genes.

Dihydroquercetin and dihydromyricetin form colorless anthocyanins in the presence of dihydroflavonol 4-reductase (DFR), and subsequently, produce colored anthocyanins by anthocyanidin synthase (ANS) catalysis and are finally converted to stable anthocyanins. Cyanidin 3-(6″-p-coumarylsambubioside) and delphinidin 3-glucoside were up-regulated and Cyanidin 3-(6-feruloylglucoside) 5-(6-malonylglucoside) was down-regulated in our study. Moreover, the genes of *CSS0016543.1* encoding DFR and *CSS0045924.1* encoding ANS were significantly up-regulated.

Colorless anthocyanins can be directly reduced to non-epi-type catechins C and GC by leucoanthocyanidinreductase (LAR). Additionally, colorless anthocyanins can be successively catalyzed by the anthocyanidinsynthase (ANS) and anthocyanidinreductase (ANR) to produce the phenotypic catechins EC and EGC. These catechins can be further synthesized into ester catechins by the action of UDG-galloyl-1-o-β-D-glucosetransferase (UGGT) and EC-1-O-galloyl-B-D-gallicacyltransferase (ECGT). Herein, the genes of *CSS0032189.1* and *CSS0001408.1* encoding LAR were down-regulated and *CSS0041215.1* and *CSS0026266.1* were up-regulated. Additionally, the genes of *CSS0013982.1* and *CSS0027461.1* encoding ANR were up-regulated. Moreover, five catechins including catechin (C), epicatechin (EC), epigallocatechin (EGC), gallocatechin (GC), and epigallocatechin gallate (EGCG) all up-regulated, indicating the up-regulated genes exerted a dominant role in the differential genes of LAR. There were no significant expression differences between the genes encoding UGGT and ECGT.

Through flow charts, key pathways directly related to gylc and qmlc quality were regulated by DEGs. Our work lays the foundation for future research into the relationships and molecular mechanisms of secondary metabolite accumulation in tea. However, there may be more complex regulatory mechanisms explaining this phenomenon, and the same requires further elaboration in future studies.

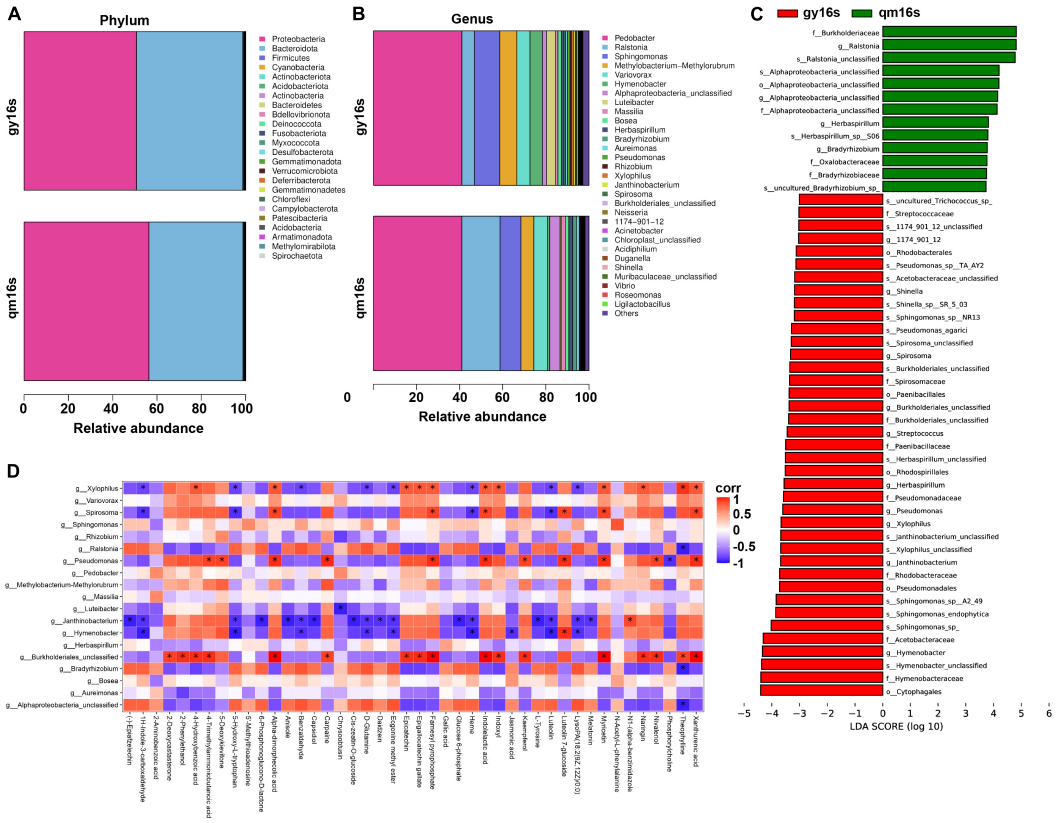


FIGURE 6
Symbiotic bacterial diversity analysis of qmlc and gylc. **(A)** Differences between gy16s vs. qm16s at the phylum level. **(B)** Differences between gy16s vs. qm16s at the genus level. **(C)** Results of EfSe analysis of gy16s vs. qm16s. **(D)** Heat map of association analysis of DEMs and commensal bacteria for gylc and qmlc. $^{*}P < 0.05$.

Endophytic bacteria of gygc and qmgc and their effects on the tea quality

Both metabolomic and transcriptomic data from tea samples suggest that the quality differences between gygc and qmgc may be attributed to symbiotic microorganisms. Endophytic bacteria are a special class that can colonize healthy plant tissues for a long time, and further establish a harmonious association with the plant in symbiotic relationships. This hypothesis was determined by high-throughput sequencing of endophytic bacteria (qm16s and gy16s) from gylc and qmlc. Species richness, evenness, and sequencing depth of gy16s and qm16s were evaluated by alpha diversity analysis (**Supplementary Table 4**). Our findings revealed that goods coverage was 1 for all samples, indicating high completeness of sequencing data and all bacteria in the samples could be detected. Chao1 and observed species were significantly lowered for gy16s, which meant that the bacterial diversity of gy16s expanded.

As illustrated in **Figures 6A, B**, the diversity of the samples was very low at the phylum level and *proteobacteria* and *bacteroidota* were the dominant species. At the genus level, gy16s exhibited a more complex colony structure. Apart from a significant down-regulation of the relative abundance of ralstonia and an up-regulation of the relative abundance of spirosoma, the composition and proportions of dominant species were similar in both samples, which suggested that the two samples shared a core bacterial flora. Surprisingly, ralstonia was reported to be the causal agent of cyanobacteria in plants (40). Moreover, infection of tea plants can lead to poor quality tea and cause economic losses. The inappropriate use of synthetic chemicals to control pests and diseases has become a major problem in the tea industry (41). For this reason, tea samples from the plants without any human intervention were focused on for follow-up observations. Until 26 July 2022, there were no signs of disease in these tea plants. As tea plants on Wufeng mountain were grown for 15 years without pesticides or chemical fertilizers intervention and even the metabolomic data did not contain any evidence of chemical pesticides, the symbiotic microorganisms, the tea plants and its surrounding have produced a stable and harmonious ecosystem. The mechanism by which tea plants infected with the cyanobacteria produce

"self-healing" warrants further exploration, especially given that this mechanism could be developed into a novel type of pesticide to improve tea quality by reducing the use of traditional pesticides.

EfSe analysis serves as a tool for the discovery and interpretation of biomarkers for high-dimensional data, and functions as a combination of non-parametric testing and linear discriminant analysis. As illustrated in **Figure 6C**, the relative abundances of most bacteria were up-regulated in gy16s. Similar to the results in **Figure 6B**, *Ralstonia*, the pathogen of bacterial wilt, was found to be down-regulated at the genus and species levels. Similarly, *Alphaproteobacteria* was observed to be down-regulated in order, family, genus, and species, and also serves other many functions like phosphate solubilization, IAA production, siderophore production, and ammonia production (42). The relative abundance of s__*Herbaspirillum*_sp__S06, g__*Herbaspirillum* was also down-regulated. *Herbaspirillum* sp., a class of nitrogen-fixing bacteria, and further reported to enhance the effective use of selenate and selenite by tea plants and promote the growth of branch lateral shoots after pruning (43). The relative abundance of s__*Sphingomonas*_sp, s__*Sphingomonas_endophytica*, and s__*Sphingomonas* sp. A2-49 was significantly up-regulated in gy16s. *Sphingomonas* is a well-known common bacterium in tea gardens (44). o__*Pseudomonadales* were significantly up-regulated and their representative strain *Pseudomonas* sp. strain GN6 exerts the functions of phosphate solubilization, IAA production, siderophore production, and ammonia production (45). s__*Janthinobacterium*_unclassified and g__*Janthinobacterium* are significantly up-regulated, such that *Janthinobacterium* has been reported to possess the ability to resist fungi such as *Alternaria brassicicola* (46). Furthermore, J*anthinobacterium* has previously reported as being a coldness-resistant and low-nutrient-needed bacterium, and these characteristics facilitate tea plants to adapt alpine plantations and form a stable colonization (47).

According to correlation analysis between DEMs of gylc and qmlc and endophytic bacteria (**Figure 6D**), *Burkholderiales* and *Pseudomonas* exhibited positive correlations with DEMs, whereas there were negative correlations between *Janthinobacterium* and *Hymenobacter* with DEMs. Meanwhile, Epigallocatechin gallate, epicatechin were significantly positively correlated with *Xylophilus* and *Burkholderiales*. Catechin and proanthocyanidin biosynthesis have been reported to activate poplar defense against *Melampsora laricipopulina* (48). Additional experimentation revealed that theophylline was negatively correlated with *Alphaproteobacteria*, *Ralstonia*, but positively correlated with *Xylophilus*, *Burkholderiales*. Additionally, benzaldehyde was negatively correlated with *Janthinobacterium*, *Hymenobacter*, and *Xylophilus*. Myricetin was-positively correlated with *Burkholderiales*, *Pseudomonas*, *Spirosoma*, and *Xylophilus*. Kaempferol was positively correlated with *Burkholderiales* and *Pseudomonas*. Jasmonic acid was negatively correlated with *Hymenobacter*. D-Glutamine and L-Tyrosine were negatively correlated with *Janthinobacterium*. All these findings collectively suggested that the dynamics of the symbiotic bacterial communities were closely associated with DEMs that determined the quality of qmgc and gygc. Nivalenol was positively correlated with *Pseudomonas* and *Burkholderiales*, therefore it can be inferred that the synthesis of nivalenol might come from these two bacteria and the symbiotic microbes protect tea plants from pathogens by synthesizing antibiotics.

Briefly speaking, the endophytic bacteria inside fresh tea leaves can directly or indirectly influence the quality of Hefeng tea. These bacteria tend to colonize healthy tea plant tissues for a long time and establish a harmonious symbiotic environment. In addition to being directly associated with the DEMs, endophytic bacteria can also regulate the growth rate of tea plants by fixing nitrogen and synthesizing phytohormones, which in turn improve metabolites related to tea quality. From a certain point of view, endophytic bacteria address the root problem of pesticide residues by protecting tea plants from pathogens and reducing pesticides and fertilizers.

Conclusion

The pick timing of green tea leaves exerts a detrimental effect on their sensory quality from broth color, aroma, and flavor. At a molecular level, the metabolites including catechins, free amino acids, alkaloids, flavonols, soluble sugars, and organic acids vary with the pick timing, leading to teas with different aromas and flavors. In our findings, the variation pattern of differential metabolites associated with green tea quality at different picking times did not exactly match the results of previous studies carried out at lower altitudes. The combined analysis of metabolomic and transcriptomic from fresh tea samples highlighted that flavonoid biosynthesis and phenylalanine metabolism serve as the key pathways responsible for the quality differences of gygc and qmgc. In this pathway, the content of catechins, flavonols, and anthocyanins was regulated by the expression of DEGs and exhibited an up-regulation, which in turn enhanced the bitterness of gygc.

On the other hand, both transcriptomic and metabolomic analysis data on the tea samples revealed that endophytic bacteria indirectly promote tea quality by strengthening plants defense ability and augmenting their growth. Additionally, the microbial diversity analysis verified that endophytic bacteria directly correlate DEMs to influence tea product quality. Metabolites were significantly altered in tea products pre- and post-processing.

Previous studies tend to test the differential metabolites of fresh tea leaves or the differential metabolites of processed

tea leaves to analyze the tea quality, which was insufficient. Our study sought to provide a novel theoretical basis for how different pick timing affects the abundance and mechanism of tea metabolites for subsequent research. This study also guided high altitude area (e.g., Hefeng) green tea to find the optimal solution between quality and economic benefits.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Author contributions

HX: methodology, investigation, software, data curation, and writing—original draft. JY and YX: software and data curation. HZ: funding acquisition, supervision, methodology, and writing—review and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Science and Technology Department of Hunan Province (2018NK1030).

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/fnut.2022.1079325/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Melting curve.

SUPPLEMENTARY FIGURE 2
Amplification curve.

SUPPLEMENTARY FIGURE 3
Results of association analysis of DEMs and DEGs for qylc and qmlc. $^{*}P < 0.05$.

References

1. Guo X, Ho CT, Schwab W, Wan X. Aroma profiles of green tea made with fresh tea leaves plucked in summer. *Food Chem.* (2021) 363:130328. doi: 10.1016/j.foodchem.2021.130328

2. Zhao M, Zhang N, Gao T, Jin J, Jing T, Wang J, et al. Sesquiterpene glucosylation mediated by glucosyltransferase UGT91Q2 is involved in the modulation of cold stress tolerance in tea plants. *New Phytol.* (2020) 226:362–72. doi: 10.1111/nph.16364

3. Wang H, Hua J, Yu Q, Li J, Wang J, Deng Y, et al. Widely targeted metabolomic analysis reveals dynamic changes in non-volatile and volatile metabolites during green tea processing. *Food Chem.* (2021) 363:130131. doi: 10.1016/j.foodchem.2021.130131

4. Fan FY, Huang CS, Tong YL, Guo HW, Zhou SJ, Ye JH, et al. Widely targeted metabolomics analysis of white peony teas with different storage time and association with sensory attributes. *Food Chem.* (2021) 362:130257. doi: 10.1016/j.foodchem.2021.130257

5. Shi J, Zhang X, Zhang Y, Lin X, Li B, Chen Z. Integrated metabolomic and transcriptomic strategies to understand the effects of dark stress on tea callus flavonoid biosynthesis. *Plant Physiol Biochem.* (2020) 155:549–59. doi: 10.1016/j.plaphy.2020.07.048

6. Chen Y, Zhou B, Li J, Tang H, Zeng L, Chen Q, et al. Effects of long-term non-pruning on main quality constituents in 'Dancong' Tea (*Camellia sinensis*) leaves based on proteomics and metabolomics analysis. *Foods.* (2021) 10:2649. doi: 10.3390/foods10112649

7. Zeng C, Lin H, Liu Z, Liu Z. Metabolomics analysis of *Camellia sinensis* with respect to harvesting time. *Food Res Int.* (2020) 128:108814. doi: 10.1016/j.foodres.2019.108814

8. Liu J, Zhang Q, Liu M, Ma L, Shi Y, Ruan J. Metabolomic analyses reveal distinct change of metabolites and quality of green tea during the short duration of a single spring season. *J Agric Food Chem.* (2016) 64:3302–9. doi: 10.1021/acs.jafc.6b00404

9. Li Y, Ran W, He C, Zhou J, Chen Y, Yu Z, et al. Effects of different tea tree varieties on the color, aroma, and taste of Chinese Enshi green tea. *Food Chem X.* (2022) 14:100289. doi: 10.1016/j.fochx.2022.100289

10. Xue J, Liu P, Guo G, Wang W, Zhang J, Wang W, et al. Profiling of dynamic changes in non-volatile metabolites of shaken black tea during the manufacturing process using targeted and non-targeted metabolomics analysis. *LWT.* (2022) 156:113010.

11. Zhu Y, Lv HP, Shao CY, Kang S, Zhang Y, Guo L, et al. Identification of key odorants responsible for chestnut-like aroma quality of green teas. *Food Res Int.* (2018) 108:74–82. doi: 10.1016/j.foodres.2018.03.026

12. Xu Q, He Y, Yan X, Zhao S, Zhu J, Wei C. Unraveling a crosstalk regulatory network of temporal aroma accumulation in tea plant (*Camellia sinensis*) leaves by integration of metabolomics and transcriptomics. *Environ Exp Bot.* (2018) 149:81–94.

13. Ruiz-Rodriguez L, Aguilar A, Diaz AN, Sanchez FG. Enantioseparation of the fungicide imazalil in orange juice by chiral HPLC. Study on degradation rates and extractive/enrichment techniques. *Food Chem.* (2015) 178:179–85. doi: 10.1016/j.foodchem.2015.01.004

14. Zhang L, Cao Q-Q, Granato D, Xu Y-Q, Ho C-T. Association between chemistry and taste of tea: A review. *Trends Food Sci Technol.* (2020) 101:139–49.

15. Lee JE, Lee BJ, Hwang JA, Ko KS, Chung JO, Kim EH, et al. Metabolic dependence of green tea on plucking positions revisited: a metabolomic study. *J Agric Food Chem.* (2011) 59:10579–85. doi: 10.1021/jf202304z

16. Han Z, Wen M, Zhang H, Zhang L, Wan X, Ho CT. LC-MS based metabolomics and sensory evaluation reveal the critical compounds of different grades of Huangshan Maofeng green tea. *Food Chem.* (2022) 374:131796. doi: 10.1016/j.foodchem.2021.131796

17. Jiang L, Shen X, Shoji T, Kanda T, Zhou J, Zhao L. Characterization and activity of anthocyanins in Zijuan tea (Camellia sinensis var. kitamura). *J Agric Food Chem.* (2013) 61:3306–10. doi: 10.1021/jf304860u

18. Zhu MZ, Li N, Zhou F, Ouyang J, Lu DM, Xu W, et al. Microbial bioconversion of the chemical components in dark tea. *Food Chem.* (2020) 312:126043. doi: 10.1016/j.foodchem.2019.126043

19. Cao QQ, Zou C, Zhang YH, Du QZ, Yin JF, Shi J, et al. Improving the taste of autumn green tea with tannase. *Food Chem.* (2019) 277:432–7. doi: 10.1016/j.foodchem.2018.10.146

20. Sittipod S, Schwartz E, Paravisini L, Peterson DG. Identification of flavor modulating compounds that positively impact coffee quality. *Food Chem.* (2019) 301:125250. doi: 10.1016/j.foodchem.2019.125250

21. Yue W, Sun W, Rao RSP, Ye N, Yang Z, Chen M. Non-targeted metabolomics reveals distinct chemical compositions among different grades of Bai Mudan white tea. *Food Chem.* (2019) 277:289–97. doi: 10.1016/j.foodchem.2018.10.113

22. Ye J-H, Ye Y, Yin J-F, Jin J, Liang Y-R, Liu R-Y, et al. Bitterness and astringency of tea leaves and products: Formation mechanism and reducing strategies. *Trends Food Sci Technol.* (2022) 123:130–43. doi: 10.1016/j.tifs.2022.02.031

23. Wang H, Cao X, Yuan Z, Guo G. Untargeted metabolomics coupled with chemometrics approach for Xinyang Maojian green tea with cultivar, elevation and processing variations. *Food Chem.* (2021) 352:129359. doi: 10.1016/j.foodchem.2021.129359

24. Zhang J, Sun-Waterhouse D, Su G, Zhao M. New insight into umami receptor, umami/umami-enhancing peptides and their derivatives: a review. *Trends Food Sci Technol.* (2019) 88:429–38. doi: 10.1016/j.tifs.2019.04.008

25. Yang C, Hu Z, Lu M, Li P, Tan J, Chen M, et al. Application of metabolomics profiling in the analysis of metabolites and taste quality in different subtypes of white tea. *Food Res Int.* (2018) 106:909–19. doi: 10.1016/j.foodres.2018.01.069

26. Shen J, Wang Y, Ding Z, Ding S, Wang H, Bi C, et al. Metabolic analyses reveal growth characteristics of young tea shoots in spring. *Sci Hortic.* (2019) 246:478–89. doi: 10.1016/j.scienta.2018.11.022

27. Gai Z, Wang Y, Ding Y, Qian W, Qiu C, Xie H, et al. Exogenous abscisic acid induces the lipid and flavonoid metabolism of tea plants under drought stress. *Sci Rep.* (2020) 10:12275. doi: 10.1038/s41598-020-69080-1

28. Xiao Z, Wang H, Niu Y, Liu Q, Zhu J, Chen H, et al. Characterization of aroma compositions in different Chinese congou black teas using GC-MS and GC-O combined with partial least squares regression. *Flav Frag J.* (2017) 32:265–76.

29. Chen Q, Weidong Dai YZ, Lv H, Mu B, Li P, Tan J, et al. Aroma formation and dynamic changes during white tea processing. *Food Chem.* (2019) 274:915–24.

30. Hirose S, Kamatari YO, Yanase E. Mechanism of oolongtheanin formation via three intermediates. *Tetrahedron Lett.* (2020) 61:151601.

31. Marino SD, Borbone N, Gala F, Zollo F, Iorizzi M. New constituents of sweet *Capsicum annuum* L. fruits and evaluation of their biological activity. *J Agric Food Chem.* (2006) 54:7508–16. doi: 10.1021/jf061404z

32. Mann IS, Widdowson DA, Clough JM. Transition metal directed synthesis of moracin M, a phytoalexin of *Morus alba* Linn. *Tetrahedron.* (1991) 47:7981–90.

33. Morokuma K, Irie R, Oikawa M. Total synthesis of lycoperdic acid and its C4-epimer. *Tetrahedron Lett.* (2019) 60:2067–9.

34. Bryla M, Ksieniewicz-Wozniak E, Waskiewicz A, Szymczyk K, Jedrzejczak R. Natural occurrence of nivalenol, deoxynivalenol, and deoxynivalenol-3-Glucoside in Polish Winter Wheat. *Toxins.* (2018) 10:81. doi: 10.3390/toxins10020081

35. Plabutong N, Ekronarongchai S, Niwetbowornchai N, Edwards SW, Virakul S, Chiewchengchol D, et al. The inhibitory effect of validamycin A on aspergillus flavus. *Int J Microbiol.* (2020) 2020:3972415. doi: 10.1155/2020/3972415

36. de Oliveira Filho JWG, Islam MT, Ali ES, Uddin SJ, Santos JVO, de Alencar MVOB, et al. A comprehensive review on biological properties of citrinin. *Food Chem Toxicol.* (2017) 110:130–41. doi: 10.1016/j.fct.2017.10.002

37. Mei X, Lin C, Wan S, Chen B, Wu H, Zhang L. A comparative metabolomic analysis reveals difference manufacture suitability in "Yinghong 9" and "Huangyu" Teas (*Camellia sinensis*). *Front Plant Sci.* (2021) 12:767724. doi: 10.3389/fpls.2021.767724

38. Singh K, Kumar S, Rani A, Gulati A, Ahuja PS. Phenylalanine ammonia-lyase (PAL) and cinnamate 4-hydroxylase (C4H) and catechins (flavan-3-ols) accumulation in tea. *Funct Integr Genomics.* (2009) 9:125–34. doi: 10.1007/s10142-008-0092-9

39. Gu H, Wang Y, Xie H, Qiu C, Zhang S, Xiao J, et al. Drought stress triggers proteomic changes involving lignin, flavonoids and fatty acids in tea plants. *Sci Rep.* (2020) 10:15504. doi: 10.1038/s41598-020-72596-1

40. Mansfield J, Genin S, Magori S, Citovsky V, Sriariyanum M, Ronald P, et al. Top 10 plant pathogenic bacteria in molecular plant pathology. *Mol Plant Pathol.* (2012) 13:614–29.

41. Bag S, Mondal A, Banik A. Exploring tea (*Camellia sinensis*) microbiome: Insights into the functional characteristics and their impact on tea growth promotion. *Microbiol Res.* (2022) 254:126890. doi: 10.1016/j.micres.2021.126890

42. Bhattacharyya C, Banerjee S, Acharya U, Mitra A, Mallick I, Haldar A, et al. Evaluation of plant growth promotion properties and induction of antioxidative defense mechanism by tea rhizobacteria of Darjeeling, India. *Sci Rep.* (2020) 10:15536. doi: 10.1038/s41598-020-72439-z

43. Xu X, Cheng W, Liu X, You H, Wu G, Ding K, et al. Selenate reduction and selenium enrichment of tea by the *Endophytic herbaspirillum* sp. Strain WT00C. *Curr Microbiol.* (2020) 77:588–601. doi: 10.1007/s00284-019-01682-z

44. Arafat Y, Ud Din I, Tayyab M, Jiang Y, Chen T, Cai Z, et al. Soil sickness in aged tea plantation is associated with a shift in microbial communities as a result of plant polyphenol accumulation in the tea gardens. *Front Plant Sci.* (2020) 11:601. doi: 10.3389/fpls.2020.00601

45. Dutta J, Thakur D. Evaluation of multifarious plant growth promoting traits, antagonistic potential and phylogenetic affiliation of rhizobacteria associated with commercial tea plants grown in Darjeeling, India. *PLoS One.* (2017) 12:e0182302. doi: 10.1371/journal.pone.0182302

46. Lyakhovchenko NS, Nikishin AI, Gubina ED, Pribylov DA, Senchenkov VY, Sirotin AA, et al. Assessment of the antifungal activity of the violacein-forming strain Janthinobacterium sp. B-3515 against the mould fungus *Alternaria brassicicola* F-1864. *IOP Conf Ser Earth Environ Sci.* (2021) 908:012006.

47. Yang M, Lu D, Qin B, Liu Q, Zhao Y, Liu H, et al. Highly efficient nitrogen removal of a coldness-resistant and low nutrient needed bacterium, Janthinobacterium sp. M-11. *Bioresour Technol.* (2018) 256:366–73. doi: 10.1016/j.biortech.2018.02.049

48. Ullah C, Tsai CJ, Unsicker SB, Xue L, Reichelt M, Gershenzon J, et al. Salicylic acid activates poplar defense against the biotrophic rust fungus Melampsora larici-populina via increased biosynthesis of catechin and proanthocyanidins. *New Phytol.* (2019) 221:960–75. doi: 10.1111/nph.15396

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 18 January 2023
DOI 10.3389/fnut.2023.1035739

Check for updates

OPEN ACCESS

EDITED BY
Pranjal Yadava,
Indian Agricultural Research Institute (ICAR), India

REVIEWED BY
Rajkumar U. Zunjare,
Indian Council of Agricultural Research (ICAR), India
Rajan Sharma,
Punjab Agricultural University, India

*CORRESPONDENCE
Xiao-yan Jiao
✉ xiaoyan_jiao@sxau.edu.cn
Xian-min Diao
✉ diaoxianmin@caas.cn

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

RECEIVED 03 September 2022
ACCEPTED 02 January 2023
PUBLISHED 18 January 2023

CITATION
Wang Y, Wang J-s, Dong E-w, Liu Q-x, Wang L-g, Chen E-y, Jiao X-y and Diao X-m (2023) Foxtail millet [*Setaria italica* (L.) P. Beauv.] grown under nitrogen deficiency exhibits a lower folate contents. *Front. Nutr.* 10:1035739. doi: 10.3389/fnut.2023.1035739

COPYRIGHT
© 2023 Wang, Wang, Dong, Liu, Wang, Chen, Jiao and Diao. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Foxtail millet [*Setaria italica* (L.) P. Beauv.] grown under nitrogen deficiency exhibits a lower folate contents

Yuan Wang[1], Jin-song Wang[1], Er-wei Dong[1], Qiu-xia Liu[1], Li-ge Wang[1], Er-ying Chen[2], Xiao-yan Jiao[1]* and Xian-min Diao[3]*

[1]College of Resources and Environment, Shanxi Agricultural University, Taiyuan, Shanxi, China, [2]Institute of Crop Research, Shandong Academy of Agricultural Sciences, Jinan, China, [3]Institute of Crop Sciences, Chinese Academy of Agricultural Sciences, Beijing, China

Foxtail millet [*Setaria italica* (L.) P. Beauv.], as a rich source of folates, has been cultivated on arid infertile lands, for which N deficiency is one of the major issues. Growing environments might have a significant influence on cereal folate levels. However, little is known whether N deficiency modulates cereal folate levels. In order to obtain enriched folate foxtail millet production in nutrient-poor soil, we conducted a study investigating the content of folate derivatives of 29 diverse foxtail millet cultivars under two N regimes (0 and 150 kg N ha^{-1}) for 2 years to explore folate potential grown under low N. The contents of total folate and most derivatives were reduced by N deficiency. The effect on total folate content caused by N was stronger than cultivar genotype did. Folate content of enriched folate cultivars was prone to be reduced by N deficiency. Structural equation models (SEMs) revealed that N fertilization had a positive indirect effect on grain folate content through influencing plant N and K accumulation. Collectively, the results indicate much more attention should be paid to N management when foxtail millet is cultivated in infertile soil, to improve foxtail millet folate contents.

KEYWORDS

foxtail millet, cultivars, folates, nitrogen deficiency, plant NPK accumulation, nitrogen recovery efficiency

1. Introduction

Folates (vitamin B9), consisting of tetrahydrofolate (THF) and its derivatives, are essential for nucleotide synthesis and cofactors in one-carbon units. Humans lack the capability of *de novo* folate biosynthesis (1). For this reason, folates are essential micronutrients, and must be supplied through balanced diet. Folate efficiently prevents the neural tube defects in the developing fetus, cardiovascular disease, and stroke, and folate deficiency may lead to megaloblastic anemia (2, 3). It has been reported that the prevalence of folate deficiency in women of reproductive age is more than 40% in most countries, resulting in approximately 300,000 newborns with neural tube defects per year. To prevent neural tube defects, the US Public Health Service has recommended that all women capable of becoming pregnant consume 400 μg per day of folic acid (4). Maruvada et al. (5) reported that food fortification and medical supplementation are suggested as complementary ways to alleviate folate deficiency. However, folic acid supplementation may lead to adverse effects of elevated folate status, such as potential cancers. Thus, food fortification is a cost-effective and efficient alternative. Endogenous folate in wholegrain cereals is readily bioavailable and may improve folate status (6).

Genotypes and environment significantly influence the folate content in cereal. For instance, total folate content among wheat and soybean genotypes shows a remarked variation (7, 8). Moreover, most of the folate in cereal products also differs markedly according to the growing conditions (climate, soil type, weather conditions, etc.) (9). Therefore, genotypes with high folate by careful selection and the means to enhance natural folate contents by improving environmental factors need to be studied.

Nitrogen (N) is an essential, often limiting, factor in plant growth and grain development (10). Upon N limitation, plants develop physiological alterations, including folate synthesis processes (11). According to Jiang et al. (12), N deficiency decreased the expression of most of the genes involved in folate synthesis and C1 units in *Arabidopsis* seedlings. Furthermore, folates are assembled from pterin, ρ-aminobenzoate, and glutamate precursors. Most folates are conjugated to a c-linked polyglutamyl tail of up to eight residues. These polyglutamyl tails may help to protect folates from oxidative breakdown, and folates tend to be stabilized by polyglutamylation (13). In *Arabidopsis*, low N typically results in significant decreases in glutamine and asparagines (14) and might exhibit a less conjugation of polyglutamylated folates. Folate levels correlate positively with polyglutamate tail length (15). These observations indicate that N limitation in the soil might have a detrimental effect on folate levels of grain produced.

Foxtail millet [*Setaria italica* (L.) P. Beauv.], as the second millet in terms of worldwide production after pearl millet (16), has valuable nutritional and medicinal properties (17). It comprises a wide range of health-benefiting components, including phenolic (18), protein hydrolysates (19), carotenoids (20), polysaccharides (21), and other antioxidants, which have also been shown to possess several health benefits like prevention of cancer, hypoglycemic, and hypolipidemic effects (17). These functional components make foxtail millet as an unique nutritional crop among the cereal categories. A few of previous studies revealed that folate contents in foxtail millet were much higher than that in other cereals such as rice, wheat, and maize (22, 23). While several studies have investigated the changes of both folate metabolite content and the expression patterns of folate metabolite-related genes in foxtail millet grain during it developed (24), fewer evaluation concerning the effect of a specific environmental factor on cereal crops folate levels has been carried out, especially for the effect of N deficiency on foxtail millet. Foxtail millet has been cultivated primarily on infertile lands in arid and semi-arid areas of Asia. It is prone to adapt to adverse soils than most other crops (25). Being sessile in nutrient-poor soil, foxtail millet encounters some environmental challenges while obtaining the nutrients necessary for development and biomass production, especially in N deficiency. Therefore, the variation in foxtail millet folate contents under N-deficiency could be used to elucidate foxtail millet folate potential grown in nutrient-poor soil and reveal the effects of agronomic fertilization measures on grain folate content.

In order to survey the natural variation of folate contents among foxtail millet cultivars and their folate potential levels under nutrient-poor soil in China, a total of 29 Chinese foxtail millet cultivars were grown with or without N fertilization in 2020 and 2021 and were analyzed for folate contents by using HPLC-MS/MS. The aims of this study were: (1) to investigate the folate content variation of the leading Chinese foxtail millet cultivars; (2) to assess the effect of N deficiency on folate content; (3) to evaluate the association of foxtail millet folate levels with nutrient accumulation, which was affected by N regimes. The findings of this study would provide essential information for folate improvement in foxtail millet and other cereals. In particular, much attention should be paid to fertilization management when cultivating elite folate cultivars, especially in infertile soil conditions.

2. Materials and methods

2.1. Plant materials and design of experiment

Field experiments were conducted at the Dongyang Experimental Station of Shanxi Agricultural University, Shanxi, China (37°33′21″N, 112°40′2″E) in 2020 and 2021. This site generally experiences a temperate continental climate with a mean annual air temperature of 9.7°C and 440.7 mm of rainfall. The precipitation during the foxtail millet growth period was 411.8 mm and 310.8 mm in 2020 and 2021, respectively (Figure 1). This site possesses sandy loam soil with the pH of 8.45 and contained 0.93 g kg^{-1} total N, 6.87 mg kg^{-1} Olsen-P (P), 138.2 mg kg^{-1} available potassium (K), and 20.33 g kg^{-1} organic matter.

The field experiment was a twenty-nine by two [most widely used and representative twenty-nine (29) foxtail millet cultivars and two levels of N rates] factorial design with 58 combinations, in which each combination had three plots as replicates. Two N rates were included 0 kg N ha^{-1} application (N-) and 150 kg N ha^{-1} application (N+). Nitrogen was supplied as slow-release urea (46%-N) before sowing. For all treatments, 75 kg ha^{-1} P_2O_5 and 75 kg ha^{-1} K_2O were applied before sowing as well. Each plot area was 30 m^2 (5 × 6 m). Plants on the borders of the plots were treated for protection and not harvested. The information of sowing date, plant density, plant height with N applied, and harvest date (the mature stage of the selected cultivars) of every selected cultivar is shown in Supplementary Table 1. At the three-leaf stage, seedlings were thinned to the suitable density of every selected cultivar. A total of 120 mm of water was applied through irrigation, half applied before fertilizer application and the rest irrigated before heading. Weeds, diseases, and insect pests were controlled adequately; No factor other than the N level limited growth.

2.2. Sampling

When panicles and grains were fully matured and plant individuals were harvested from 20 m^2 of each plot, then panicle weight, straw weight, and grain weight were estimated, and the thousand-grain weight was determined. The total weight of panicles and straw was used to compute aboveground biomass. Some representative plant samples (grain, panicle, and straw) were used to collect tissue for N, P, and K analysis after drying at 75°C to maintain a constant weight. After being weighted, the samples were finely grounded to pass through a 250 μm sieve. Another portion of representative grains from each cultivar was stored in a paper bag after threshing on time. About 5 g grains were manually ground to fine powder below 4°C, and the powder was completely filtered through the screen mesh (100 μm sieve). Then powder was temporarily stored at −80°C until further analyses of the folate levels.

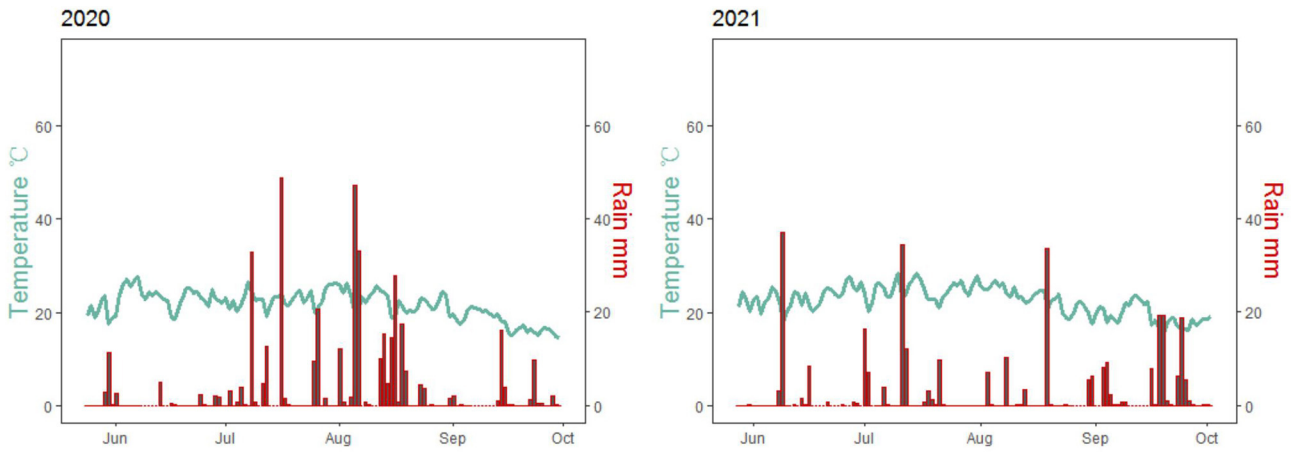


FIGURE 1
Average daily temperature and rainfall during growing seasons in 2020 and 2021.

2.3. Determination of N, P, and K accumulation in plant tissues

According to the micro-Kjeldahl method, N content in plant tissues was determined with automatic Kjeldahl Apparatus (FOSS-8400, Sweden) (26). Furthermore, about 0.5 g ground sample was digested by microwave-assisted acid digestion (27), followed by total P and K analysis using a vanadium-molybdenum yellow colorimetric method and flame photometer (28), respectively. Grain N, P, and K accumulation per hectare were calculated as grain N, P, or K concentration times grain dry weight per hectare. Total NPK accumulation per hectare during harvest was calculated as the product of NPK concentration and yield of above-ground parts of the foxtail millet plant on a dry matter basis per hectare. Nitrogen recovery efficiency (NRE) was estimated based on the difference in the total plant N uptake between two N treatments in 2 years divided by the total N rate applied in 2 years.

2.4. Extraction of folate

2.4.1. Chemical compounds

The folate standards: 10-formyl-folic acid (10–CHO-PteGlu; 10–CHO–FA), 5,10-methenyl-5,6,7,8-tetrahydrofolate (5,10–CH=H_4PteGlu; 5,10–CH=THF), 5-formyl-tetrahydrofolate (5–CHO-H_4PteGlu; 5–CHO–THF), 5-methyl-tetrahydrofolate (5–CH_3-H_4PteGlu; 5–CH_3-THF), dihydrofolate (H_2PteGlu; DHF), folic acid (PteGlu; FA), tetrahydrofolate (H_4PteGlu; THF), and methotrexate (MTX) were purchased from Schircks Laboratories (Jona, Switzerland) and MeFox, an oxidation product of 5–CH_3-THF, was obtained from Toronto Research Chemicals (Toronto, Canada). Sodium phosphate dibasic (Na_2HPO_4), sodium phosphate monobasic (NaH_2PO_4), sodium ascorbate, and β-mercaptoethanol were obtained from the Sigma-Aldrich Chemical. The endogenous folates in rat serum were removed by incubation with one-tenth (w/w) of activated charcoal for 1 h on ice, followed by centrifuging at 13,000 rpm at 4°C for 30 min (Sigma 3K15, Osterode am Harz, Germany), and the supernatant was used for the following incubation experiment. Acetonitrile and formic acid (LC-MS grade) were obtained from Fisher Scientific (Geel, Belgium). The HPLC analytical column (Kromasil 100-5 C18, 2.1 × 50 mm, 2.5 μM particle size) was purchased from Akzo Nobel (Stockholm, Sweden), and an Agilent SB-C18 pre-column (2.1 × 5 mm, 2.7 μM particle size) was acquired from Agilent Technologies (California, USA).

2.4.2. Folate extraction and deglutamylation

Both extraction and measurement of folates were conducted as previously described for wheat (7), with slight modifications. Briefly, folate extraction was performed under subdued light to minimize degradation. The moisture content of each sample was measured after oven overnight at 75°C. Another 50 mg of powder was transferred into a 1.5 mL screw-cap tube (Axygen, ST-150) and 1 mL of freshly prepared extraction solution [5 mM phosphate buffer, pH 7.2; 0.5% sodium ascorbate; and 0.2% β-mercaptoethanol] was added.

Post homogenization, the mixture was immediately boiled in a water bath for 10 min using an electromagnetic oven, cooled on ice for 10 min, then centrifuged for 10 min at 13,000 rpm at 4°C. The supernatant (0.5 mL) was transferred to a fresh tube. Then 35 mL of rat serum was added, and the polyglutamylated tails were deconjugated by incubation at 37°C for 4 h. Consequently, the samples were boiled for 10 min, cooled on ice for 10 min, and centrifuged at 13,000 rpm at 4°C for 10 min. The supernatant was moved to 3 kDa ultra-filtration tubes (Millipore) for cleanup and centrifuged at 13,000 rpm at 4°C for 25 min. Finally, the resulting solution was collected and 100 μL was transferred into new tubes for direct folate detection. The remaining solution was stored at −80°C. Each sample was repeated three times.

2.5. Folate standard solutions

The chemical powder of folate standards was dissolved as 0.1 mg/mL in a solution of 20 mM ammonium acetate in methanol and water (1:1, v/v) containing 1% (w/v) L-ascorbic acid, and 0.5% (v/v) β-mercaptoethanol (pH 6.2). 5,10–CH=THF standard was prepared in a pH 4.5 buffer. All standard solutions were kept at

−80°C. For spiking and calibration, working solutions were diluted with folate extraction buffer, which was made fresh and used the same day.

2.6. Folate determination by HPLC–MS/MS

Chromatographic studies were carried out at a flow rate of 0.30 mL/min using an Agilent 1260 HPLC system (Palo Alto, CA) with an Akzo Nobel analytical column (Kromasil 100-5 C18, 50*2.1 mm). The injection volume was 15.0 μL. The injector and column oven temperature were separately maintained at 4 and 25°C, respectively. The mobile phases were 0.1% (v/v) formic acid in water (phase A) and 0.1% (v/v) formic acid in acetonitrile (phase B). The gradient program ran for a total of 19 min. The proportion of mobile phase B increased linearly from 5 to 9% over 2 min. In the following 6 min, phase B increased to 9.6% and then sharply increased to 20% in over 0.2 min. After holding steady at 20% for 3 min, the proportion of phase B decreased to 5% in 0.2 min, followed by a subsequent equilibration.

An Agilent 6420 triple, quadruple tandem MS coupled with an ESI (electron spray ionization) interface was used for mass analyses and quantification of target analytes. The mass spectrometer was used in positive ion mode. The parameters were optimized for the target analytes with a gas temperature of 350°C, drying gas flow at 11 L/min, nebulizer pressure at 35 psi, and capillary voltage at 3500 V (+). The parameters for folate standards were m/z 456–412, 30 eV for 5,10–CH=THF, m/z 460–313, 20 eV for 5–CH3–THF, m/z 474–327, 20 eV for 5–CHO–THF, m/z 446–299, 20 eV for THF, and m/z 455–308, 30 eV for the internal control MTX. System operation and data acquisition were performed with MassHunter software. The sum of contents of different folate derivatives represents as the total folate levels. The seven folate derivatives chromatograms of representative cultivar, Jingu21, under two N treatments were exhibited in Supplementary Figure 1.

2.7. Method validation

The precision, linearity, sensitivity, recovery and matrix effect were evaluated for method validation, for which a series of dilution of folate standards was prepared in the blank foxtail millet solutions. To prepare a blank foxtail millet matrix, 1 g fine sample of Jingu21 was mixed with 20 mL phosphate buffer without sodium ascorbate and β-mercaptoethanol. The mixture was boiled for 1 h, then exposed to direct sunlight to degrade endogenous folates. Ten percent activated charcoal was added to the supernatant and was incubated with shaking for 1 h, and then centrifuged at 13,000 rpm for 30 min at 4°C. The supernatant was filtered through 3KDa MWCO ultra-filtration tubes (29).

Precision was assessed based on the relative standard deviation (RSD) of peak areas of Jingu 21 ($n = 6$) on the same and different days. Intra-day repeatability and inter-day repeatability was evaluated by running samples on three different days (Supplementary Table 2). The calibration curves was evaluated by preparing nine-point (1, 2, 5, 10, 50, 100, 200, 500, 1,000 ng/mL) blank foxtail millet solution for MTX and each folate derivative ($n = 3$). The linearity and correlation coefficients were calculated by plotting the peak area at different concentrations. Data analyses were performed with MassHunter software. The correlation coefficients (R^2) for all these folates were approximately > 0.99 in the folate derivatives matrices (Supplementary Table 2). Sensitivity was evaluated by determining the limit of detection (LOD) and limit of quantification (LOQ). The LOD and LOQ values of the amount of analyte were estimated for spiked samples on signal/noise ratios of 3:1 and 10:1, respectively. The LOD and LOQ of the established method in folate derivatives and MTX were in the range of 0.05–0.25 μg per 100 g and 0.16–0.85 μg per 100 g, respectively (Supplementary Table 2). Recovery and matrix effect were calculated as described by Matuszewski et al. (30). The recoveries and matrix effects were illuminated in Supplementary Table 2.

2.8. Statistical analysis

The data of folate derivatives in cereal samples were shown as means ± SE of three biological replicates in micrograms per 100 g of grains, with the exception of the contents of folate derivatives under different NRE levels (Table 2). Data were exposed to the analysis of variance (ANOVA) with agricolae package and boxplots with the ggplot 2 package in R 4.0.1 (R Foundation for TUNA Team, Tsinghua University, China, Beijing). Pearson correlation was performed on the R statistical software (Performance Analytics and ggplot2).

Structural equation models (SEMs) were established to test the direct and indirect effects of N regimes, plant NPK accumulation, and grain NPK accumulation on folate levels (5–CHO–THF, 5–CH_3-THF, THF, 10–CHO–FA, and 5,10–CH=THF) in foxtail millets, in which this composite variable did not alter the underlying SEM model. A direct path from N treatment to plant NPK accumulation was used to account for the effects of nutrient utilization efficiency on the plant. A direct path from plant NPK accumulation to grain NPK accumulation was used to account for the effects of crop growth and nutrient distribution on grain nutrient uptake. Folate level was assumed to be affected by all other components both directly and indirectly. We started with a priori models that included all plausible pathways between these factors. Subsequently, the significance of each path coefficient was tested by calculating its critical ratio ($P < 0.05$). The SEMs was performed using the software Amos Graphics v22 (IBM Corp., Armonk, NY, USA). The capacity of SEM to separate the direct and indirect effects of a variable on dependent variables is considered one of the most important advantages of SEM (31). The overall fit of the final model was evaluated with the goodness-of-fit index (GFI), Bentler comparative fit index (CFI), Chi-square test, and root mean square error of approximation (RMSEA) (32).

3. Results and discussion

3.1. Effects of cultivars and N deficiency on total folate contents in foxtail millets

When N (N+) was supplied, the folate levels of 29 foxtail millet cultivars ranged from 42.36 ± 2.91 to 72.89 ± 0.50 μg per 100 g grains and 42.84 ± 1.60 to 69.20 ± 2.13 μg per 100 g

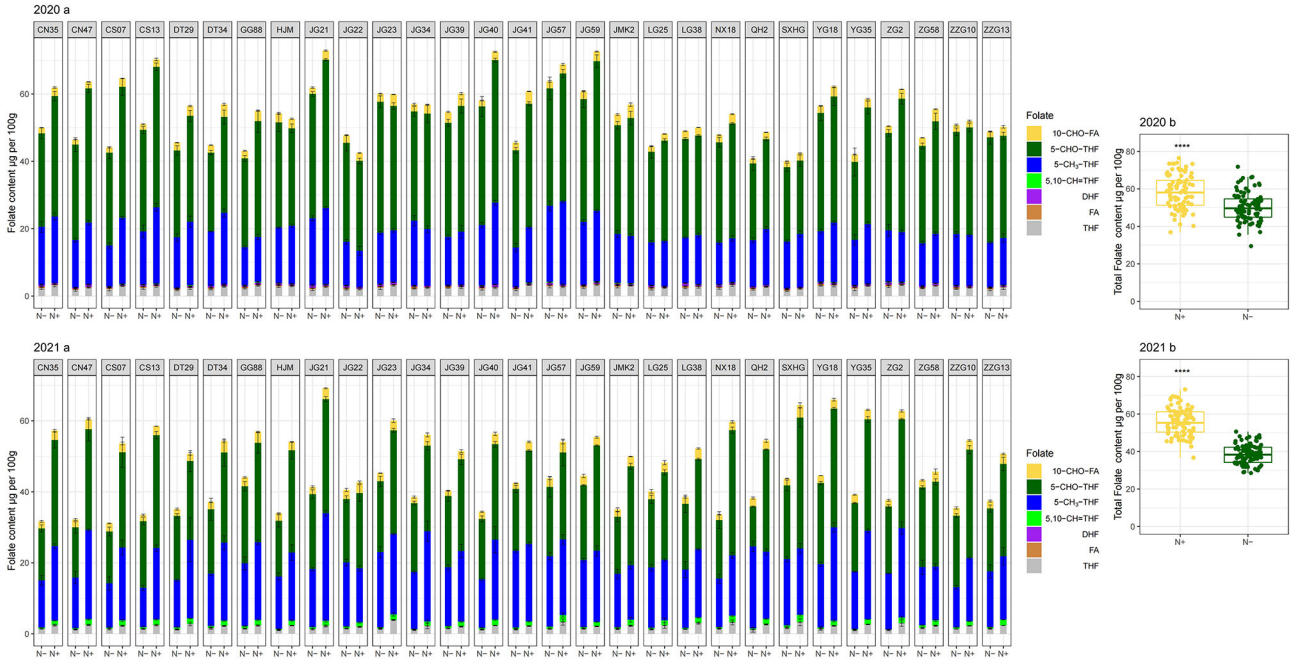


FIGURE 2
Folate profiling in folate derivatives of the selected cultivars (2020a in 2020; 2021a in 2021) and the total folate of all cultivars (2020b in 2020; 2021b in 2021) grown with or without N fertilization. CN35, Changnong35; CN47, Changnong47; CS07, Changsheng07; CS13, Changsheng13; DT29, Datong29; DT34, Datong34; GG88, Gonggu88; HJM, Huangjinmiao; JG21, Jingu21; JG22, Jigu22; JG23, Jiugu23; JG34, Jingu34; JG39, Jigu39; JG40, Jingu40; JG41, Jigu41; JG57, Jingu57; JG59, Jingu59; JMK2, JinmiaoK2; LG25, Longgu25; LG38, Longgu38; NX18, Nenxuan18; QH2, Qinhuang2; SXHG, shanxihonggu; YG18, Yugu18; YG35, Yugu35; ZG2, Zhonggu2; ZG58, Zhaogu58; ZZG10, Zhangzagu10; ZZG13, Zhangzagu13; N−, N deficiency; N+, N applied treatment. Error bars represent the standard error of folate derivatives. ****Indicate significant differences at 0.0001.

grains for 2020 and 2021, respectively (Figure 2 2020b and 2021b). Among the 29 foxtail millet cultivars, only one cultivar (Jingu 21) had mean total folate content of more than 70 μg per 100 g grains for 2 years. About 31% of cultivars (Jingu40, Changsheng13, Yugu18, Jingu59, Zhonggu2, Changnong47, Jingu57, Jiugu23, and Yugu35) had mean total folate content that ranged from 60–70 μg per 100 g for 2 years. A total of 17 cultivars (Changnong35, Changsheng07, Jigu41, Nenxuan18, Jingu34, Gonggu88, Jigu39, Datong34, Datong29, Shanxihonggu, JinmiaoK2, Huangjinmiao, Zhangzagu10, Qinhuang2, Longgu38, Zhaogu58, and Zhangzagu13) fell in the 50-60 μg per 100 g range, and two cultivars (Longgu25, and Jigu22) measured less than 50 μg per 100 g, but over 42 μg per 100 g grains for 2 years (Figure 2 2020a and 2021a). The average folate content of the richest cultivar (e.g. Jingu21) was 1.7-fold higher than that of the poorest cultivar (e.g. Jigu22) under N-sufficient growth conditions. Thus there were obvious differences in folate content among cultivars. Averagely, the total folate content of 2 years for 29 cultivars ranged from 42.59 to 71.03 μg per 100 g. Relative to the folates content in wheat (38–43 μg per 100 g) and rice (6–8 μg per 100 g) reported by Bekaert et al. (23), our study further demonstrated that the foxtail millet cultivars rich in folates could serve as health-benefiting cereals. The total folate contents of the cultivars evaluated in the present study fall within the range of contents reported in previous studies (40.35–111.1 μg per 100 g) (24).

Two-way analysis of variance showed that both two N regimes and cultivars produced highly significant variation in the total folate contents in 2 years, with N factors affecting folates contents much more strongly than that of the cultivar genotypes (mean squares of 2685 and 271 in 2020; mean squares of 13032 and 91 in 2021, respectively) (Supplementary Table 3). Among the 29 foxtail millet cultivars, the folate levels ranged from 40.02 ± 1.74 to 63.72 ± 5.58 μg per 100 g grains when N was deficient in 2020. And the year of 2021, low-N stress affected total folate content much more than that of 2020. During the year of 2021, the folate levels ranged from 31.12 ± 1.57 to 45.25 ± 2.07 μg per 100 g grains under low-N (Figure 2 2020b and 2021b). This demonstrated that increased N stress diminishes folate content seriously. Nitrogen deficiency significantly reduced the total folate contents in the foxtail millet. This may be due to that, typically low N results in large decreases in glutamine and asparagines (14), and exhibits a less conjugation of polyglutamylated folates. These polyglutamyl tails may help to protect folates from oxidative breakdown. As a result, folates tend to be stabilized by polyglutamylation (13). Folate levels correlate positively with polyglutamate tail length (15). As previously stated, folates play an important role in various metabolic processes, including amino acid synthesis (33). This indicates that N status in plant tissues might strongly link with folates. However, further studies will be necessary to explore the mechanism of detrimental effects on folates induced by low-N stress.

The responses of total folate content of 29 cultivars to N deficiency differed significantly. Some foxtail millet cultivars (such as Jingu21, Jingu59, Jingu57, and so on) were found to have relatively high folate contents compared with others for each N regime, but obvious differences in folate content have appeared between two

N treatments for these cultivars, in which the reduction ratio of Jingu21 was 47% in 2021 (Figure 2 2020a and 2021a). It indicates that much attention should be paid to N fertilization management when elite folate-rich cultivars are cultivated, especially in infertile soil conditions. Sufficient N should be supplied to ensure elite folate-rich foxtail millet's potential. There were no significant differences in total folate content of Zhaogu58 and Jigu22 detected between two N treatments for both years; whereas both had a relatively low folate content among 29 cultivars (Figure 2 2020a and 2021a). Further, low N stress decreased the difference of folates level among 29 cultivars. Thus, high folate cultivar selection should be conducted in N sufficient conditions in practice.

3.2. Effects of N deficiency on folate derivatives contents in foxtail millets

Folates and their derivatives occur as polyglutamates in nature. The extreme low concentrations of folate derivatives limit the applicability of the separation technique and emphasize the demand for sensitive detection techniques (34). Owing to the sensitive detection techniques, the folate levels were analyzed by HPLC-MS/MS in this study, which is supported by Upadhyaya et al. (35). The content of seven folate forms, including 5–CHO–THF, 5–CH_3-THF, THF, 5,10–CH=THF, 10–CHO–FA, DHF and FA were detected in present study. Other folate forms failed to be detected due to their trace quantities (Figure 2 2020 and 2021).

The distribution pattern of the seven derivatives showed that 5–CHO–THF and 5–CH_3-THF were the predominant folate forms in the foxtail millet cultivars. They contributed to quantifying more than 85% of the total folates, in which 5–CHO–THF contributed more than 45%. Meanwhile, total content of 10–CHO–FA and THF accounted for about 5% of total folate. While FA and DHF were the least abundant folate vitamers, collectively contributing about 1% of total folate in 2020, and failed to be detected because of their trace quantities in 2021 (Figure 2 2021b and Figure 3 2021). Studies on folate vitamer distribution in foxtail millet are scanty. This research indicated that 5–CHO–THF, was the most stable derivative, and the major folate form in foxtail millet; which is also consistent with a previous study in wheat (9, 36), where the content of folate decreases in the following order: 5–CHO–THF > 5–CH_3-THF > THF (37). This might be related to 5–CHO–THF acting as a storage form of one-carbon groups in seeds (38, 39). Furthermore, 5–CH_3-THF is one of the folate forms with great bioavailability for animals (40). Therefore, the foxtail millet cultivars with higher 5–CH_3-THF content can be considered the preferred folate form of food cereals. The present study showed that the highest amounts of 5–CH_3-THF were 30.29 μg per 100 g grains (Jingu 21 in 2021) when N was supplied (Figure 2 2021a).

Previous studies demonstrated that crop genotypes differ in NRE in the same N-supply environments (41). As mentioned above (Figures 2, 3), N modulated folates content of foxtail millet. Consequently, exploring the relationship between folate levels and N use efficiency is essential. Nitrogen recovery efficiency is an important parameter for N use efficiency. Indeed, Pearson correlation coefficients demonstrated a strong positive relationship between NRE and 5–CH_3-THF ($P < 0.01$), 5–CHO–THF ($P < 0.01$) or total folate ($P < 0.01$) under sufficient N conditions in two successive years, respectively (Table 1). When the cultivars had more than 45% NRE, its contents of 5–CHO–THF, 5–CH_3-THF and total folate were obviously high. Their average contents were increased by 17, 19, and 16%, respectively, compared to the contents of cultivars with a low NRE level (<20%) (Table 2). This probably indicated the foxtail millet cultivars with higher levels of 5–CHO–THF, 5–CH_3-THF, and total folate might have the ability to absorb more N in soils at the same N application rates. Such high N-use efficiency foxtail millet genotypes probably have the relative high folates potential. Therefore high folate content cultivars might be selected from high N-use efficient genotypes.

Low-N stress had a great negative effect on 5–CHO–THF, 5–CH_3-THF, THF, 10–CHO–FA and 5,10–CH=THF for 2 years. Compared with the values when N was applied, N stress resulted in diminished the contents of 5–CHO–THF, by 13.56% and 34.67% in 2020 and 2021, and the content of 5–CH_3-THF by 12.71% and 22.07% in 2020 and 2021, respectively. Also, the decreased contents of THF, 10–CHO–FA and 5,10–CH=THF by 19.62%, 23.68%, and 39.05% in 2020, and 46.71%, 23.65%, and 71.54% in 2021, respectively, induced by N stress, were observed as well (Figure 3). Further, the negative effects of low-N stress, in 2021, on folate derivatives contents were much more obvious than those in 2020. Nitrogen deficiency did not significantly affect the content of DHF, but the content of FA enhanced by 168.54% in 2020. These two folate derivatives were in the range of <1.5 μg per 100 g grains (Figure 3). The study to evaluate the effect of N regimes on folate level in cereals is sanity. Thus, the effect of N deficiency on grain folate derivatives in this study can provide more precise information regarding how to modulate natural folate components by agronomic fertilization measures.

The obvious effects of N deficiency on the correlation between each folate derivative content and total folate content were not noticed, usually. All folate derivatives were positively associated with total folate content with the exception of 5,10–CH=THF for both N regimes. Among these seven folate derivatives, the highest correlation was observed between 5–CHO–THF and total folate ($r = 0.804^{***}$ with N; $r = 0.927^{***}$ without N $r = 0.90^{***}$), which may have been caused by its high abundance in foxtail millet. The 5,10–CH=THF and total folate content had a negative correlation ($r = -0.260^{***}$) under N deficiency, while had no correlation when N was supplied (Figure 4). Consistent with the previous study (7), presently a negative correlation ($r = -0.380^{***}$ with N; $r = -0.391^{***}$ without N) was also observed between 5–CHO–THF and 5,10- CH=THF (Figure 4), which may result from the interconversion relationship between these two components (23).

3.3. Effect of N deficiency on crop production and nutrient accumulation in foxtail millet

Low N stress had negative effects on grain yield and aboveground biomass. In compared with the N application, the average grain yield decreased by 13.24% and 15.03% in 2020 and 2021, respectively (Figure 5 2020a and 2021a), and biomass decreased by 7.58% and 14.78% in 2020 and 2021, respectively (Figure 5 2020b and 2021b),

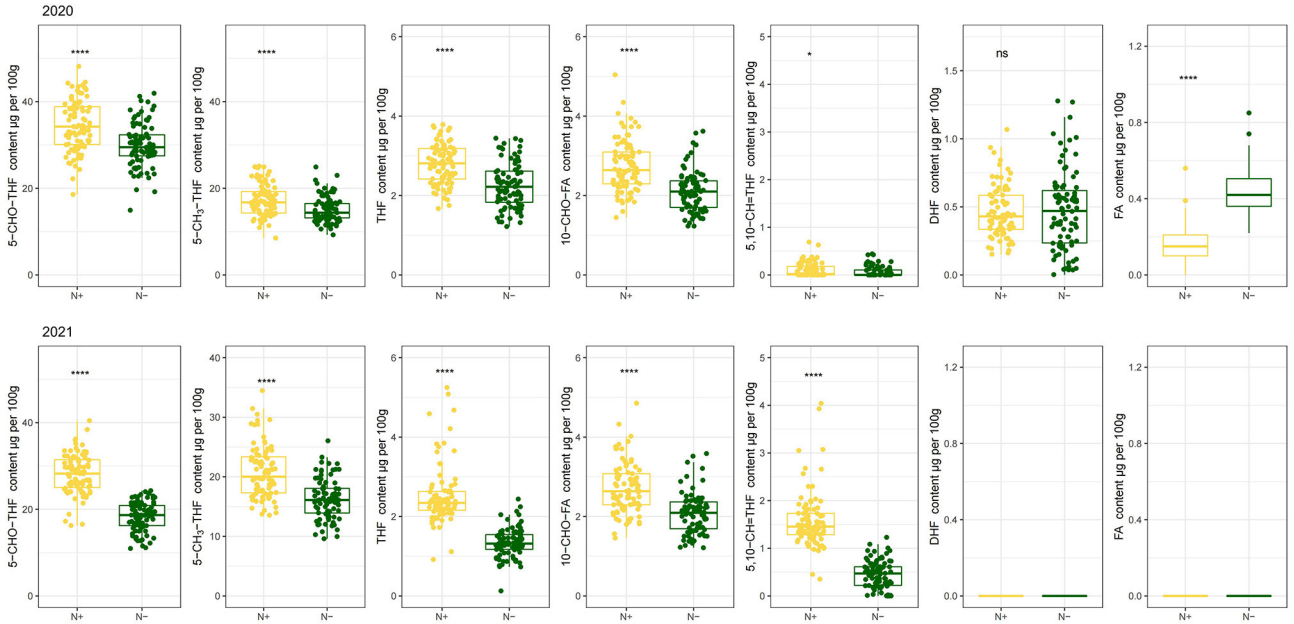


FIGURE 3
The distribution of each folate derivative 5–CHO–THF, 5–CH_3-THF, THF, 10–CHO–FA, 5, 10–CH=THF, DHF and FA grown with or without N fertilization. N–, N deficiency; N+, N applied treatment. *, **** and ns indicate significant differences at 0.05, 0.0001 and no significant difference, respectively.

TABLE 1 Correlation between NRE and different folate derivatives with N application.

Correlations with	Folate derivatives							
	5–CHO–THF	5–CH_3-THF	THF	10–CHO–FA	5,10–CH=THF	DHF	FA	Total
	0.391**	0.280**	0.032	−0.097	−0.019	0.118	0.106	0.384**

**Indicate significant differences at 0.01.

TABLE 2 The contents of different folate derivatives with different NRE levels of cultivars.

NRE (%)	Folate content (μg/100 g)							
	5–CHO–THF	5–CH_3-THF	THF	10–CHO–FA	5,10–CH=THF	DHF	FA	Total
<20	30.13 ± 2.87a	17.59 ± 2.63a	2.63 ± 0.62a	2.62 ± 0.38ab	0.80 ± 0.38a	0.22 ± 0.09a	0.07 ± 0.04a	54.06 ± 4.30a
20-30	30.30 ± 3.60a	18.71 ± 3.51a	2.67 ± 0.37a	2.86 ± 0.56b	0.93 ± 0.34a	0.23 ± 0.10a	0.08 ± 0.05a	55.80 ± 5.31ab
30-45	31.87 ± 4.20a	19.70 ± 3.82ab	2.65 ± 0.46a	2.78 ± 0.32b	0.79 ± 0.23a	0.24 ± 0.08a	0.08 ± 0.06a	58.11 ± 7.87b
>45	35.35 ± 2.46b	20.94 ± 1.39b	2.74 ± 0.22a	2.48 ± 0.27a	0.82 ± 0.16a	0.25 ± 0.12a	0.10 ± 0.05a	62.67 ± 3.38c

Data are means of three replications ± standard deviation (Mean, SD). The different lowercase letters indicate statistical differences at $P < 0.05$.

which indicated that the grain yield and aboveground biomass are known to be directly associated with N supply. Lower biomass and grain yield, induced by N deficiency, were observed in the present and other studies (42). The effects of N deficiency on grain weight are conflicting. Low N could lead to increases in potential grain weight for wheat due to grain abortion (43). However, it is also found that N deficiency also had negative effects on wheat grain weight due to reductions in source strength (44). Whereas, N application had no significant effects on thousand-grain weights in this study of foxtail millet (Figure 5 2020c and 2021c). This probably could have led to a balance between these two effects under N deficient conditions.

The plant and grain nutrient accumulation were strongly limited by N deficiency. Low-N induced clearly decreases in N accumulation by 27.14% and 36.62% on average in 2020 and 2021, respectively, for the whole plant (Figure 5 2020d and 2021d). Accordingly, N accumulation decreased by 22.36% and 30.71% for grains in 2020 and 2021, respectively (Figure 5 2020g and 2021g). This finding was consistent with other studies in foxtail millet (45). Consistently, a similar trend was also observed in the phosphorus (P) and potassium (K) accumulation for whole plant and grain tissues for 2 years, although there was no significant difference in grain K accumulation between two N treatments in 2020 (Figure 5 2020i). Perhaps it is due to the low proportion for K of crop reproductive organs (46). Moreover, N stress, which was not serious in 2020, did diminish K accumulation in grain. These negative effects might be due to decreased P and K absorption by N deficiency in foxtail millets (47).

Applied N can increase N accumulation at the whole-plant level (48), which is in line with the present results. Nitrogen, helpful for a relative massive root system development, improves P and K accumulation in plants. The enhanced P

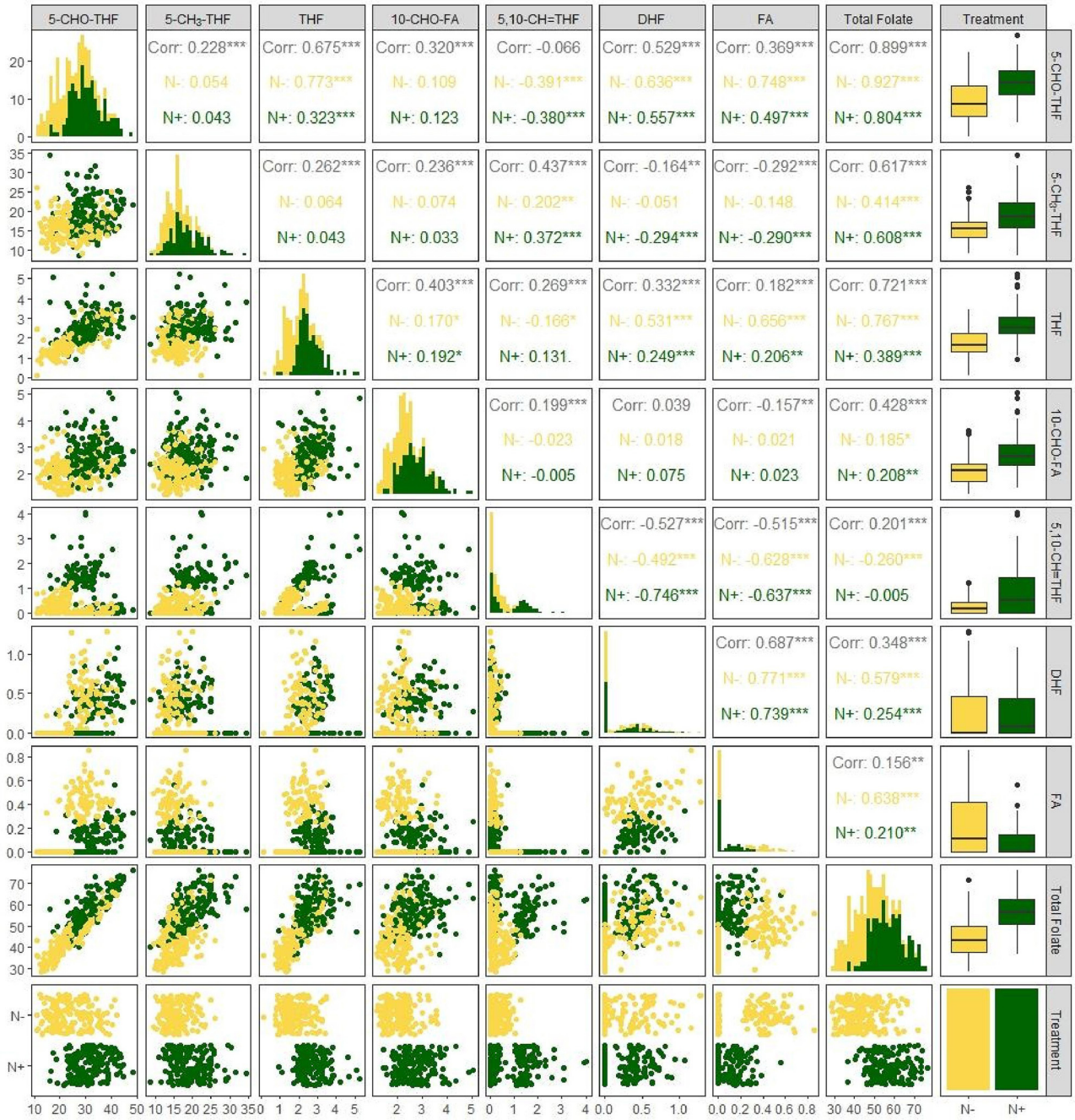


FIGURE 4
Correlation among the content of folate derivatives 5–CHO–THF, 5–CH_3-THF, THF, 10–CHO–FA, 5,10–CH=THF, DHF, FA, and total folate grown with or without N fertilization. N–, N deficiency; N+, N applied treatment. **, *** Indicate significant differences at 0.01 and 0.001.

and K accumulation might be associated with increased biomass and yield.

3.4. Association of foxtail millet folate contents with grain weight and nutrient accumulation

According to Giordano et al. (49), folates were unevenly distributed in wheat grains, and wheat germs had a higher concentration of folates than their outer layer. Usually, thousand-grain weight and folate content had a negative correlation in Wheat (7, 9). However, in the present study, there was no statistically significant relationship appeared between thousand-grain weights and total folate in 2 years under two N treatments (Figure 6 2020 and 2021). The effect of seed morphological traits on foxtail millet folates is unknown. It is noticeable that the embryo of millet represented a larger proportion of the grain weight than in other cereal grains (50). And the foxtail millet grains are much smaller than other cereals. The wheat thousand-grain weight, for

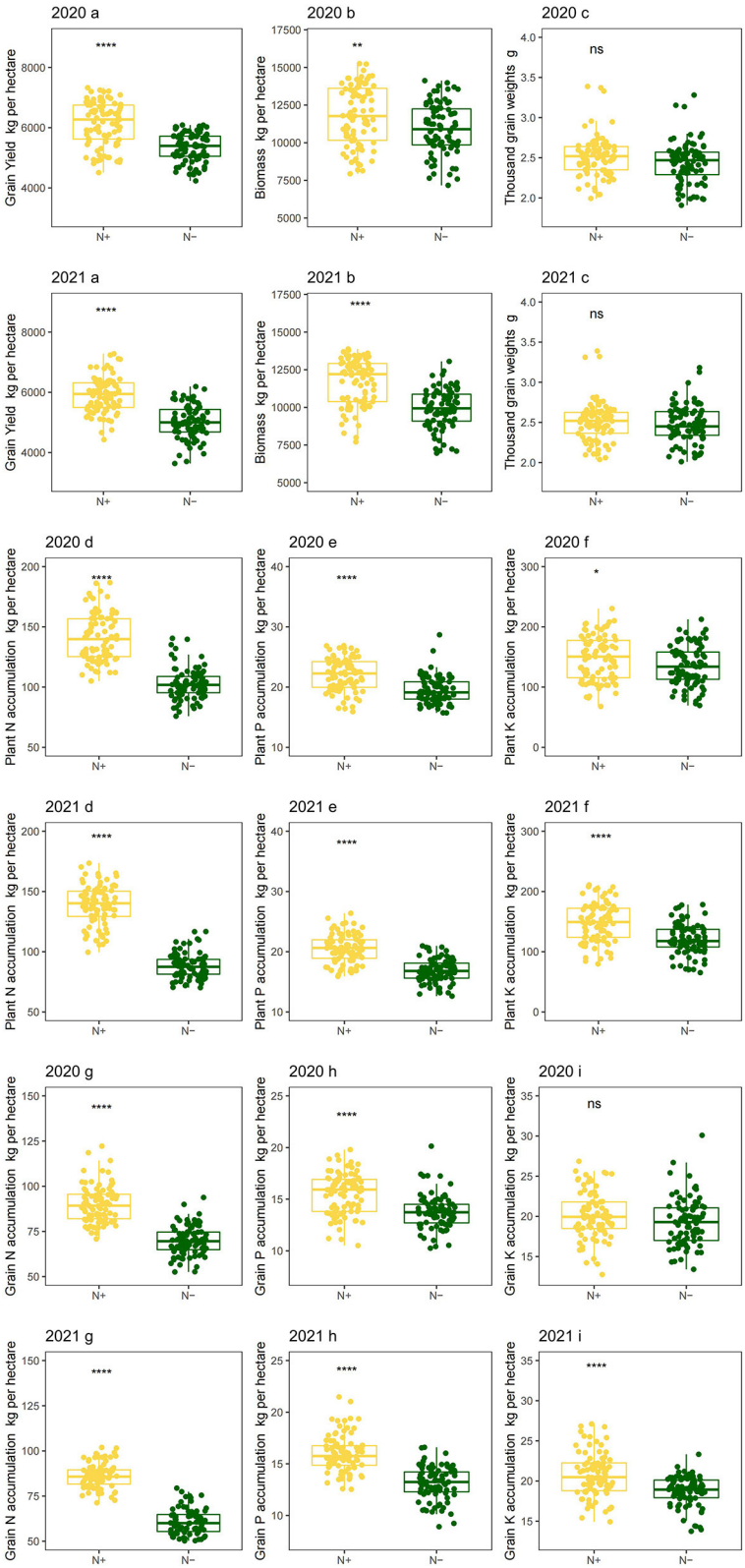


FIGURE 5

Crop production and nutrient accumulation of foxtail millet under low-N stress in 2 years. Grain yield per hectare (2020a in 2020; 2021a in 2021), aboveground biomass per hectare (2020b in 2020; 2021b in 2021), thousand grain weights (2020c in 2020; 2021c in 2021), plant N accumulation per hectare (2020d in 2020; 2021d in 2021), plant P accumulation per hectare (2020e in 2020; 2021e in 2021), plant K accumulation per hectare (2020f in 2020; 2021f in 2021), grain N accumulation per hectare (2020g in 2020; 2021g in 2021), grain P accumulation per hectare (2020h in 2020; 2021h in 2021) and grain K accumulation per hectare (2020i in 2020; 2021i in 2021). N−, N deficiency; N+, N applied treatment. *, **, **** and ns indicate significant differences at 0.05, 0.01, 0.0001 and no significant difference, respectively. Error bars represent the standard error of folate derivatives.

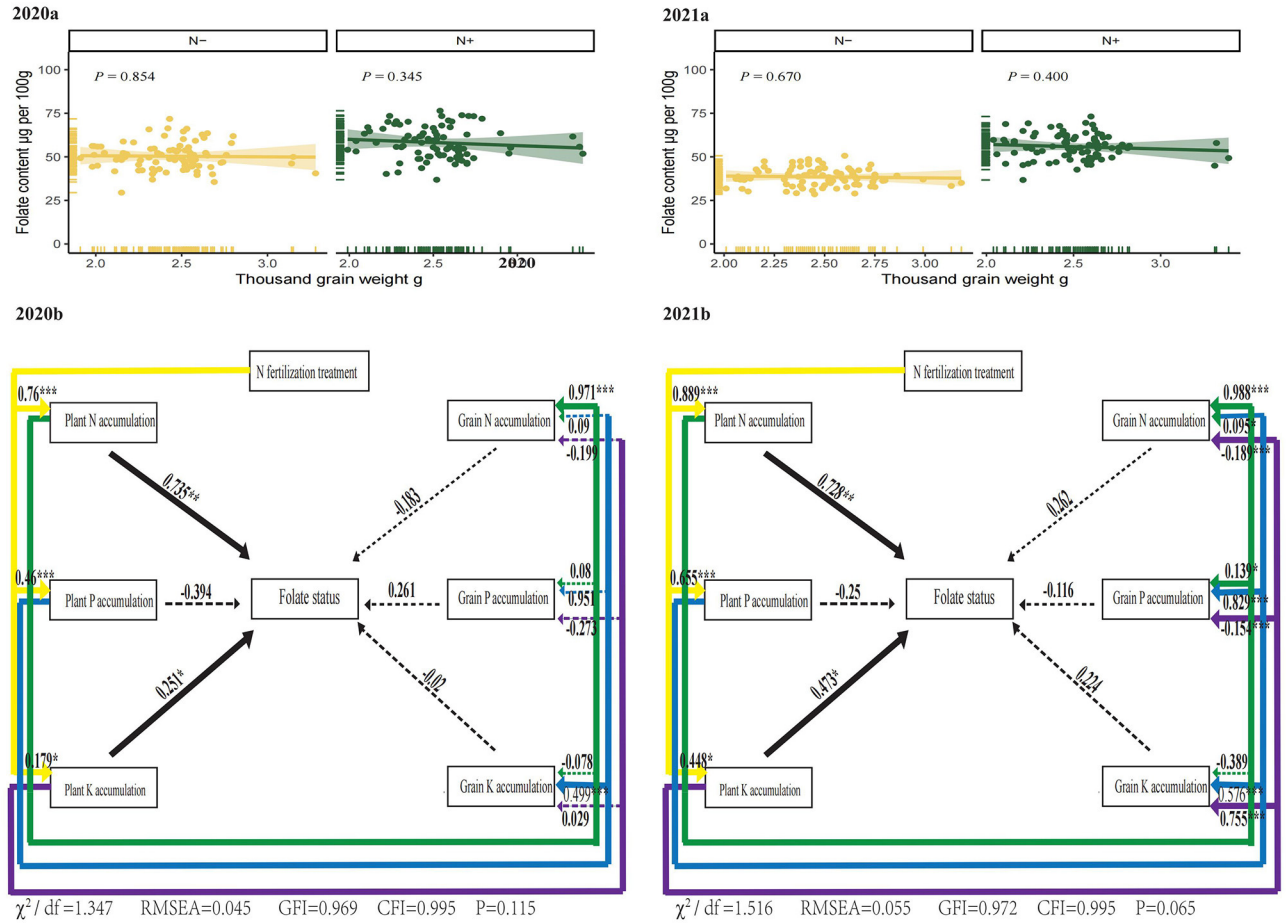


FIGURE 6
Correlation of foxtail millet thousand grain weights with total folate in 2 years. Structural equation models (SEM) to evaluate the direct and indirect effects of plant nutrient accumulation, grain nutrient accumulation, above ground biomass and grain yield on folate levels in foxtail millet. N−, N deficiency; N+, N applied treatment. The goodness-of-fit index (GFI), Bentler comparative fit index (CFI), Chi-square test and root mean square error of approximation (RMSEA) indicate the goodness-of-fit of the models to the original data. * and *** indicate the standard of significance at 0.05 and 0.001 level, respectively.

example, is about 20-fold higher than that of foxtail millet. Therefore, the effects of dilution on folates in foxtail millet were much lower than that of wheat, which probably supported no significant relationship appeared between thousand grain weights and folates in foxtail millet.

Structural equation models (SEMs) are statistical procedures for testing measurement, functional, predictive, and causal hypotheses. We further constructed SEMs to explore the direct and indirect relationships among N Fertilization, plant nutrient uptake, grain nutrient uptake and folate levels in foxtail millets (Figure 6), in terms of the effects of N deficiency on nutrient accumulation described in Figure 5. Applied N had a strong positive and indirect effect on the foxtail millet folate through positively influencing plant N accumulation and plant K accumulation in 2 years (Figure 6). The interactions between folate levels in foxtail millets (a latent variable) and plant N accumulation or plant K accumulation (observed variables) revealed that plant N and K accumulation affected folate levels more obviously than grains. Plant N and K status, which were influenced by N regimes, had a strong positive effect on folates.

4. Conclusion

We used HPLC-MS/MS techniques to investigate the folate derivatives of 29 different foxtail millet cultivars for 2 years under two N treatment rates (0 and 150 kg N ha^{-1}). With the application of N, the cultivar Jingu21 recorded the highest mean folate content among the 29 foxtail millet cultivars (71.03 μg per 100 g of grains) for 2 years. Whereas folate contents of Jigu22 and Longgu25 were <50 μg per 100 g grains. The folate content for remaining foxtail millet cultivars ranged from 50 to 70 μg per 100 g grains. The contents of total folate and derivatives in foxtail millet were significantly reduced by N deficiency. The effects of N regimes on folate contents were much more evident compared with the impacts induced by cultivars. SEMs demonstrated N fertilization affected folate content of foxtail millet positively, brought about by enhanced N and K accumulation in plant aboveground. Furthermore, foxtail millet cultivars with higher level folate might have the ability to accumulate more N at the same N application rates. Folate content of high folate enriched cultivars was prone to be reduced by N deficiency, which indicated that much attention

should be paid to N management when elite folate cultivars were cultivated, especially in infertile soil conditions, to ensure foxtail millet grain quality.

Chemical compounds studied in this article

5-Formyl-tetrahydrofolate (PubChem CID: 135403648); 5-Methyl-tetrahydrofolate (PubChem CID: 135483998); Tetrahydrofolate (PubChem CID: 135444742); 10-Formyl-folic acid (PubChem CID: 135405023); 5,10-Methenyl-tetrahydrofolate (PubChem CID: 135398657); Dihydrofolate (PubChem CID: 135398604); Folic acid (PubChem CID: 135398658); Methotrexate (PubChem CID: 126941); Sodium phosphate monobasic (PubChem CID: 23672064); Sodium phosphate dibasic (PubChem CID: 24203); Sodium ascorbate (PubChem CID: 23667548); β-Mercaptoethanol (PubChem CID: 1567); Acetonitrile (PubChem CID: 6342); Formic acid (PubChem CID: 284).

Data availability statement

The data that support the findings of this study are available from the corresponding author, X-yJ, upon reasonable request.

Author contributions

YW: data curation, formal analysis, investigation, methodology, and writing-original draft. J-sW: supervision and validation. E-wD: validation. Q-xL: visualization. L-gW: project administration. E-yC: investigation and resources. X-yJ and X-mD: conceptualization, funding acquisition, project administration, resources, and writing—review and editing. All authors contributed to the article and approved the submitted version.

Funding

This work was financially supported by the China Agriculture Research System of MOF and MARA (Grant No. CARS-06-14.5-A20) and the National Key Research and Development Program of China (Grant No. 2020YFD1000801-2).

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/fnut.2023.1035739/full#supplementary-material

References

1. Hanson AD, Gregory IIIJF. Folate biosynthesis, turnover, and transport in plants. *Annu Rev Plant Biol.* (2011) 62:105–25. doi: 10.1146/annurev-arplant-042110-103819

2. Czeizel AE, Dudas I. Prevention of the first occurrence of neuraltube defects by periconceptional vitamin supplementation. *New Engl J Med.* (1992) 327:1832–5. doi: 10.1056/NEJM199212243272602

3. MRC Vitamin Study Research Group. Prevention of neural tube defects: results of the Medical Research Council Vitamin Study. *Lancet.* (1991) 338:131–7. doi: 10.1016/0140-6736(91)90133-A

4. World Health Organization. *Nutritional Anaemias: Tools for Effective Prevention and Control* (2015). Geneva: World Health Organization.

5. Maruvada P, Stover JP, Mason JB, Bailey RL, Davis CD, et al. Knowledge gaps in understanding the metabolic and clinical effects of excess folates/folic acid: a summary, and perspectives, from an NIH workshop. *Am J Clin Nutr.* (2020) 112:1390–403. doi: 10.1093/ajcn/nqaa259

6. Saini RK, Nile SH, Keum YS. Folates: chemistry, analysis, occurrence, biofortification and bioavailability. *Food Res Int.* (2016) 89:1–13. doi: 10.1016/j.foodres.2016.07.013

7. Riaz B, Liang Q, Wan X, Wang K, Zhang C, Ye X, et al. Folate content analysis from wheat cultivars developed in the North China Plain. *Food Chem.* (2019) 289:377–83. doi: 10.1016/j.foodchem.2019.03.028

8. Agyenim-Boateng KG, Zhang S, Islam MS, Gu Y, Li B, Azam M, et al. Profiling of naturally occurring folates in a diverse soybean germplasm by HPLC-MS/MS. *Food Chem.* (2022) 384:132520. doi: 10.1016/j.foodchem.2022.132520

9. Kariluoto S, Edelmann M, Piironen V. Effects of environment and genotype on folate contents in wheat in the HEALTHGRAIN diversity screen. *J Agri Food Chem.* (2010) 58:9324–31. doi: 10.1021/jf100251j

10. Xu G, Fan X, Miller AJ. Plant nitrogen assimilation and use efficiency. *Annu Rev Plant Biol.* (2012) 63:153–82. doi: 10.1146/annurev-arplant-042811-105532

11. Li W, Liang Q, Mishra RC, Sanchez-Muñoz R, Wang H, Chen X, et al. The 5-formyl-tetrahydrofolate proteome links folates with C/N metabolism and reveals feedback regulation of folate biosynthesis. *Plant Cell.* (2021) 33:3367–85. doi: 10.1093/plcell/koab198

12. Jiang L, Liu Y, Sun H, Han Y, Li J, Li C, et al. The mitochondrial folylpolyglutamate synthetase gene is required for nitrogen utilization during early seedling development in Arabidopsis. *Plant Physiol.* (2013) 161:971–89. doi: 10.1104/pp.112.203430

13. Suh JR, Herbig AK, Stover PJ. New perspectives on folate catabolism. *Annu Rev Nutr.* (2001) 21:255–82. doi: 10.1146/annurev.nutr.21.1.255

14. Tschoep H, Gibon Y, Carillo P, Armengaud P, Szecowka M, Nunes-Nesi A, et al. Adjustment of growth and central metabolism to a mild but sustained nitrogen-limitation in Arabidopsis. *Plant Cell Environ.* (2009) 32:300–18. doi: 10.1111/j.1365-3040.2008.01921.x

15. Akhtar TA, Orsomando G, Mehrshahi P, Lara-Núñez A, Bennett MJ, Gregory IIIJF, et al. A central role for gamma-glutamyl hydrolases in plant folate homeostasis. *Plant J.* (2010) 64:256–66. doi: 10.1111/j.1365-313X.2010.04330.x

16. Bhatt D, Fairos M, Mazumdar A. Millets: nutritional composition, production and significance: a review. *J Pharm Innov.* (2022) 11:1577–82.

17. Sharma N, Niranjan K. Foxtail Millet: Properties, processing, health benefits, and uses. *Food Rev Int*. (2018) 34:329–63. doi: 10.1080/87559129.2017.1290103

18. Chandrasekara A, Shahidi F. Content of insoluble bound phenolics in millets and their contribution to antioxidant capacity. *J Sci Food Agri*. (2010) 58:6706–14. doi: 10.1021/jf100868b

19. Chandrasekara A, Shahidi F. Bioaccessibility and antioxidant potential of millet grain phenolics as affected by simulated in vitro digestion and microbial fermentation. *J Funct Foods*. (2012) 4:226–37. doi: 10.1016/j.jff.2011.11.001

20. Shen R, Yang S, Zhao G, Shen Q, Diao X. Identification of carotenoids in foxtail millet (*Setaria italica*) and the effects of cooking methods on carotenoid content. *J Cereal Sci*. (2015) 61:86–93. doi: 10.1016/j.jcs.2014.10.009

21. Zhu A, Tang L, Fu Q, Xu M, Chen J. Optimization of alkali extraction of polysaccharides from foxtail millet and its antioxidant activities in vitro. *J Food Biochem*. (2015) 39:708–17. doi: 10.1111/jfbc.12183

22. Shao L, Wang L, Bai W, Liu Y. Evaluation and analysis of folic acid content in millet from different ecological regions in Shanxi Province. *Sci Agric Sin*. (2014) 47:1265–72.

23. Bekaert S, Storozhenko S, Mehrshahi P, Bennett MJ, Lambert W, Gregory IIIJF, et al. Folate biofortification in food plants. *Trends Plant Sci*. (2008) 13:28–35. doi: 10.1016/j.tplants.2007.11.001

24. Hou S, Man X, Lian B, Ma G, Sun Z, Han L, et al. Folate metabolic profiling and expression of folate metabolism-related genes during panicle development in foxtail millet (*Setaria italica* (L.) P. Beauv). *J Sci Food Agri*. (2021) 102:268–279. doi: 10.1002/jsfa.11355

25. Panaud O. *Foxtail Millet Cereals and Millets, Genome Mapping and Molecular Breeding in Plants* Berlin, Heidelberg: Springer Science and Business Media. (2006) 1:325–32. doi: 10.1007/978-3-540-34389-9_9

26. Yoshida S, Coronel V. Nitrogen nutrition, leaf resistance, and leaf photosynthetic rate of the rice plant. *Soil Sci Plant Nutr*. (1975) 4:207–11. doi: 10.1080/00380768.1976.10432983

27. Santos HM, Coutinho JP, Amorim FC, Lôbo IP, Moreira LS, Nascimento MM, et al. (2017). Microwave-assisted digestion using diluted HNO_3 and H_2O_2 for macro and microelements determination in guarana samples by ICP-OES. *Food Chem*. (2019) 273:159–65. doi: 10.1016/j.foodchem.2017.12.074

28. Wu SC, Cao ZH, Li ZG, Cheung KC, Wong MH. Effects of biofertilizer containing N-fixer, P and K solubilizers and AM fungi on maize growth: a greenhouse trial. *Geoderma*. (2005) 125:155–66. doi: 10.1016/j.geoderma.2004.07.003

29. Shahida M, Lian T, Wan X, Jiang L, Han L, Zhang C, et al. Folate monoglutamate in cereal grains: Evaluation of extraction techniques and determination by LC-MS/MS. *J Food Compos Anal*. (2020) 91:103510. doi: 10.1016/j.jfca.2020.103510

30. Matuszewski, B, Constanzer, M, and Chavez-Eng, C. Strategies for the assessment of matrix effect in quantitative bioanalytical methods based on HPLC– MS/MS. *Anal Chem*. (2003) 75:3019–30. doi: 10.1021/ac020361s

31. Grace JB. *Structural Equation Modeling and Natural Systems* (2006). Cambridge, UK: Cambridge University Press.

32. Schermelleh-Engel K, Moosbrugger H. Evaluating the fit of structural equation models: tests of significance and descriptive goodness-of-fit me*asures*. *Methods Psychol Res*. (2003) 8:23–74.

33. Puthusseri B, Divya P, Veeresh L, Kumar G, Neelwarne B. Evaluation of folate-binding proteins and stability of folates in plant foliages. *Food Chem*. (2018) 242:555–9. doi: 10.1016/j.foodchem.2017.09.049

34. Vahteristo LT, Ollilainen V, Varo PHPLC. determination of folate in liver and liver products. *J Food Sci*. (1996) 61:524–6. doi: 10.1111/j.1365-2621.1996.tb13148.x

35. Upadhyaya P, Tyagi K, Sarma S, Tamboli V, Sreelakshmi Y, Sharma R, et al. (2016). Natural variation in folate levels among tomato (*Solanum lycopersicum*) accessions. *Food Chem*. (2017) 217:610–9. doi: 10.1016/j.foodchem.2016.09.031

36. Gujska E, Kuncewicz A. Determination of folate in some cereals and commercial cereal-grain products consumed in Poland using trienzyme extraction and high-performance liquid chromatography methods. *Eur Food Res Technol*. (2005) 221:208–13. doi: 10.1007/s00217-004-1122-z

37. Delchier N, Herbig AL, Rychlik M, Renard CMGC. Folates in fruits and vegetables: Contents, processing, and stability. *Compr Rev Food Sci F*. (2016) 15:506–28. doi: 10.1111/1541-4337.12193

38. Kruschwitz HL, McDonald D, Cossins EA, Schirch V. 5-Formyltetrahydropteroylpolyglutamates are the major folate derivatives in Neurospora crassa conidiospores. *J Biol Chem*. (1994) 269:28757–63. doi: 10.1016/S0021-9258(19)61970-8

39. Goyer A, Collakova E, Garza RD, Quinlivan EP, Williamson J, Gregory JF, et al. 5-Formyltetrahydrofolate is an inhibitory but well tolerated metabolite in Arabidopsis leaves. *J Biol Chem*. (2005) 280:26137–42. doi: 10.1074/jbc.M503106200

40. Scott, J, Rébeillé, F, and Fletcher, J. (2000). Folic acid and folates: The feasibility for nutritional enhancement in plant *foods*. *J* Sci Food Agri. 80, 795–824. doi: 10.1002/(SICI)1097-0010(20000515)80:7<795::AID-JSFA599>3.0.CO;2-K

41. Hitz K, Clark AJ, Van Sanford DA. Identifying nitrogen-use efficient soft red winter wheat lines in high and low nitrogen environments. *Field Crop Res*. (2017) 200:1–9. doi: 10.1016/j.fcr.2016.10.001

42. Gaju O, Allard V, Martre P, Snape JW, Heumez E, LeGouis, et al. Identification of traits to improve the nitrogen-use efficiency of wheat genotypes. *Field Crop Res*. (2011) 123:139–52. doi: 10.1016/j.fcr.2011.05.010

43. Ferrante A, Savin R, Slafer GA. Relationship between fruiting efficiency and grain weight in durum wheat. *Field Crop Res*. (2015) 177:109–16. doi: 10.1016/j.fcr.2015.03.009

44. Slafer GA, Savin R. Source-sink relationship and grain mass at different positions within the spike in wheat. *Field Crop Res*. (1994) 37:39–49. doi: 10.1016/0378-4290(94)90080-9

45. Nadeem F, Ahmad Z, Wang R, Han J, Shen Q, Chang F, et al. Foxtail Millet [*Setaria italica* (L.) Beauv.] grown under low nitrogen shows a smaller root system. Enhanced biomass accumulation, and nitrate transporter expression. *Front Plant Sci*. (2018) 9:205. doi: 10.3389/fpls.2018.00205

46. Osaki M, Morikawa K, Shinano T, Urayama M, Tadano T. Productivity of high-yielding crops. *Soil Sci Plant Nutr*. (1991) 37:445–54. doi: 10.1080/00380768.1991.10415057

47. Guan R, Pan H, He W, Sun M, Wang H, Cui X, et al. (2022). Fertilizer recommendation for foxtail millet based on yield response and nutrient accumulation. *J Plant Nutr*. 45, 332–345. doi: 10.1080/01904167.2021.1943679

48. Gaju O, DeSilva J, Carvalho P, Hawkesford MJ, Griffiths S, Greenland A, et al. Leaf photosynthesis and associations with grain yield, biomass and nitrogen-use efficiency in landraces, synthetic-derived lines and cultivars in wheat. *Field Crop Res*. (2016) 193:1–15. doi: 10.1016/j.fcr.2016.04.018

49. Giordano D, Reyneri A, Blandino M. Folate distribution in bution in barley (*Hordeum vulgare* L.). common wheat (*Triticumaestivum* L) and durum wheat (*Triticum turgidum durum Desf*) pearled fractions. *J Sci Food Agri*. (2016) 96:1709–15. doi: 10.1002/jsfa.7276

50. Abdelrahman A, Hoseney RC, Varriano-Marston E. The proportions and chemical compositions of hand-dissected anatomical parts of pearl millet. *J Cereal Sci*. (1984) 2:127–33. doi: 10.1016/S0733-5210(84)80025-9

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 25 January 2023
DOI 10.3389/fnut.2023.1103041

Check for updates

OPEN ACCESS

EDITED BY
Sapna Langyan,
National Bureau of Plant Genetic Resources (ICAR), India

REVIEWED BY
Yue Wang,
Zhejiang University, China
Sanjay Guleria,
Sher-e-Kashmir University of Agricultural Sciences and Technology, India

*CORRESPONDENCE
Fuzhi Ke
fuzkgjs@sohu.com

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

RECEIVED 19 November 2022
ACCEPTED 09 January 2023
PUBLISHED 25 January 2023

CITATION
Sun L, Xu J, Nasrullah, Wang L, Nie Z, Huang X, Sun J and Ke F (2023) Comprehensive studies of biological characteristics, phytochemical profiling, and antioxidant activities of two local citrus varieties in China.
Front. Nutr. 10:1103041.
doi: 10.3389/fnut.2023.1103041

COPYRIGHT
© 2023 Sun, Xu, Nasrullah, Wang, Nie, Huang, Sun and Ke. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Comprehensive studies of biological characteristics, phytochemical profiling, and antioxidant activities of two local citrus varieties in China

Lifang Sun[1,2], Jianguo Xu[1,2], Nasrullah[3], Luoyun Wang[1,2], Zhenpeng Nie[1,2], Xiu Huang[1,2], Jianhua Sun[1,2] and Fuzhi Ke[1,2]*

[1]Institute of Citrus Research, Zhejiang Academy of Agricultural Sciences, Taizhou, China, [2]National Center for Citrus Variety Improvement, Taizhou, China, [3]Department of Plant Biology and Ecology, College of Life Sciences, Nankai University, Tianjin, China

Citrus is widely grown all over the world, and citrus fruits have long been recognized for their nutritional and medical value for human health. However, some local citrus varieties with potentially important value are still elusive. In the current study, we elucidated the biological characteristics, phylogenetic and phytochemical profiling, antioxidants and antioxidant activities of the two local citrus varieties, namely Zangju and Tuju. The physiological and phylogenetic analysis showed that Zangju fruit has the characteristics of wrinkled skin, higher acidity, and phylogenetically closest to sour mandarin *Citrus sunki*, whereas, Tuju is a kind of red orange with vermilion peel, small fruit and high sugar content, and closely clustered with *Citrus erythrosa*. The phytochemical analysis showed that many nutrition and antioxidant related differentially accumulated metabolites (DAMs) were detected in the peel and pulp of Zangju and Tuju fruits. Furthermore, it was found that the relative abundance of some key flavonoids and phenolic acids, such as tangeritin, sinensetin, diosmetin, nobiletin, and sinapic acid in the peel and pulp of Zangju and Tuju were higher than that in sour range Daidai and satsuma mandarin. Additionally, Zangju pulp and Tuju peel showed the strongest ferric reducing/antioxidant power (FRAP) activity, whereas, Tuju peel and pulp showed the strongest DPPH and ABTS free radical scavenging activities, respectively. Moreover, both the antioxidant activities of peel and pulp were significantly correlated with the contents of total phenols, total flavonoids or ascorbic acid. These results indicate that the two local citrus varieties have certain nutritional and medicinal value and potential beneficial effects on human health. Our findings will also provide an important theoretical basis for further conservation, development and medicinal utilization of Zangju and Tuju.

KEYWORDS

Zangju, Tuju, biological characteristics, phytochemicals, antioxidant activity

Introduction

Citrus belong to the family Rutaceae and are considered as one of the largest fruit plant species broadly dispersed in the temperate, tropical, and subtropical areas with potential socio-economic influence (1). Oranges, grapefruits, mandarins, pummelos, lemons, and limes are popular for nutritional value and are the main industrialized citrus crops (2). Citrus fruits, not only their delicious flavors, are also the most abundant fruits containing valuable beneficial phytochemicals and are rich sources of natural antioxidants, which are now widely accepted as being beneficial to human health (2–4). Antioxidants in food have been shown to have a suppressive effect on oxidative stress *in vivo*, and are thought to play a role in the prevention of atherosclerosis and complications from diabetes (5, 6). The antioxidant activity and medicinal value of citrus fruits has been researched and reported in many literatures (3, 7, 8).

Antioxidant activity denotes the capability of a bioactive compound to clear free radicals and inhibit oxidative degradation, such as lipid peroxidation, for preventing the oxidative damage (9, 10). It is the foundation of many biological functions and associated with the prevention of many chronic diseases (3). Therefore, natural antioxidants from fruits play a predominant role in stabilizing the health of human. In recent years, the citrus fruits are attracting more attention due to their potential health-promoting functions, which contain a number of secondary metabolites with antioxidant activity, such as flavonoids, carotenoids, phenolic acids, vitamins, alkaloids, coumarins, limonoids, and essential oils (11). These active secondary metabolites have a wide variety of biological activities of importance to human health, including anti-oxidative, anti-microbial, anti-inflammatory, anti-cancer, anti-proliferative, anti-mutagenicity, anti-carcinogenicity and anti-aging, as well as cardiovascular protective effects, hypoglycemic and insecticidal activities, neuroprotective effects, etc. (3, 11–15). In addition, attributed to the presence of these medicinally active secondary metabolites, citrus fruits have been used as traditional medicinal herbs for relieving stomachache, fever, cardiac diseases, edema, snakebite, bronchitis, and asthma (16) in several Asian countries for a long time, such as China, Japan, and Korea (17).

Many clinical and animal studies have shown that some citrus metabolites can help protect against the effects of reactive oxygen species (ROS), improve digestive function, and prevent cardiovascular diseases, inflammation, diabetes, and neurological diseases (18–20). Different varieties of fruits exhibit a great diversity in secondary metabolite constituents, and antioxidants with different components in citrus fruit extracts contribute unequally to their total antioxidant ability. Much of the total antioxidant activity of fruits is related to their phenolic content, and a close correlation exists between the polyphenol content and antioxidant activities (21). One previous research suggests that many flavonoids are more potential antioxidants than vitamins (22). Accordingly, it has been reported that citrus peels in traditional medicine exhibit important pharmacological and nutraceutical properties, and these bioactivities are significantly related to the amounts of active polyphenols (phenolic), especially phenolic acids and flavonoids (23). The phenolic composition and antioxidant activity of fruit tissues from different mandarin cultivars were reported by two researchers, and it was found that the peel and juice are the main tissues with higher total phenolic content and total antioxidant activity compared to pulp and seeds (24). Logically, DPPH or ABTS values of antioxidant activities showed higher correlation with the phenolic content in different fruit tissues (25). Furthermore, the phenolic compounds and antioxidant activities of different fruit parts and species of citrus were widely evaluated. For example, it was found that the polyphenol contents and antioxidant activities for the different fruit parts of nine grapefruits varieties varied as the following order: flavedo > segment > membrane > juice vesicle > albedo > seed, with a diversity among all the varieties (26). Besides the difference between different tissues of fruit, a remarkable diversity exists in the content of polyphenols and antioxidant activity among citrus varieties. Several previous reports showed that the highest levels of total phenols and total flavonoids were found in mandarin (27). Naringin was the dominant compound in the peel of pummelo (*Citrus grandis*), while mandarin (*Citrus reticulata*) was rich in hesperidin (25, 28, 29). Grapefruits (Cocktail and Rio Red) were more precious than the pummelo in flavonoids (26, 28). The highest phenolic acid content, dominated by protocatechuic acid, was found in kumquat (27). Additionally, Zhang et al. reported that the content and composition of phenolic compounds, including flavonoids and phenolic acid, and antioxidant capacities of 14 native wild mandarin genotypes showed clear differences in the grapefruit or mandarin group (8).

China is one of the important center origins for the genus *Citrus*, and many important citrus genotypes originated from here (30, 31). Over the past few years, many local citrus genotypes in China have been researched and documented. For example, the content and composition of bioactive compounds in 14 Chinese wild mandarin genotypes and 27 local citrus cultivars had been determinated and their antioxidant activities were also evaluated (8, 32). In this study, for the first time we studied the biological characteristics, phylogenetic relationships, phytochemical profiles, and antioxidant activities of the two local citrus varieties in China, Zangju (ZG) and Tuju (TG). The aim of our study is to explore the nutritional and medicinal value and provide an important theoretical basis for the conservation and medicinal utilization of these two local citrus varieties.

Materials and methods

Biological characteristics and phylogenetic analysis of Zangju and Tuju

Mature fruits collected from Zangju (Derong county of Sichuan province), Tuju (Chunan county of Zhejiang province), Daidai, satsuma mandarin trees were divided into peel (flavedo and albedo) and pulp (segment epidermis and juice vesicle). Five similar fruits were collected as one repeat, three repeats for each group. For each variety, peel and pulp of more than five fruits collected from different trees were sampled as one repeat and frozen in liquid nitrogen for metabolites extraction and subsequent RNA extractions, three repeats for each group.

Both of the two total genomic DNA were extracted from 100 mg of fresh leaves frozen in liquid nitrogen using a modified CTAB method. The DNA concentration (>50 ng μl^{-1}) was measured using a NanoDrop spectrophotometer, and fragmentation was achieved using sonication, and integrity was evaluated using 0.8% agarose gel. Sequencing was performed using an Illumina NovaSeq 6000 platform (Genepioneer Biotechnologies Co. Ltd., Nanjing, China) with PE250 based on Sequencing by Synthesis (SBS) technology. Except for Zangju and Tuju, the chloroplast genomic sequences of the other 20 citrus varieties and 2 close genus were downloaded from NCBI is used

for and manually annotated for phylogenetic tree analysis. MAFFTv.5 (33) was utilized to align the cp genomes of the 24 species. Then we constructed a maximum likelihood (ML) tree using MEGA 7, and a bootstrap test was performed with 1,000 repetitions.

Extraction, identification, and analysis of metabolites

Samples of the peel and pulp of mature fruits of Zangju, Tuju, Daidai, and satsuma mandarin stored at −80°C were used for metabolites extraction. Metabolome analysis was performed by Biomarker Technologies Co., Ltd. (Beijing, China) using non-target liquid chromatography–mass spectrometry (LC-MS; Biomarker Technologies) to identify differences in the metabolite profile among treatments. The LC/MS system for metabolomics analysis is composed of Waters Acquity I-Class PLUS ultra-high performance liquid tandem Waters Xevo G2-XS QT of high resolution mass spectrometer. The column used is purchased from Waters Acquity UPLC HSS T3 column (1.8 μm 2.1 × 100 mm). The experimental methods were as follows. The sample (50 mg) was added to 1 ml extract buffer containing an internal standard (methanol/acetonitrile, 2:2, v/v; internal standard concentration 20 mg/L), swirled for 30 s. Then mixture was treated with ultrasound for 10 min in an ice water bath and standing for 1 h at −20°C. The homogenate was centrifuged at 4°C and 12,000 × *g* for 15 min. A sample (500 μl) of the supernatant was carefully removed to the tube and dried in a vacuum concentrator. An aliquot (160 μl) of the extract (acetonitrile/water, 1:1, v/v) was added to the dried metabolites for resolution. After vortexing for 30 s being treated with ultrasound for 10 min in an ice water bath, the solution was centrifuged at 4°C and 12,000 × *g* for 15 min. A sample (120 μl) of the supernatant was transferred to a 2 ml injection flask and mixed with 10 μl from each sample for the QC sample analysis. Kyoto Encyclopedia of Genes and Genomes (KEGG) database was used for metabolites annotation and enrichment analysis.

Determination of total phenols, flavonoids, and ascorbic acid content

To measure the total phenols in peel and pulp, 100 mg of freeze-dried flesh samples was homogenized with 1.5 ml of 60% aqueous ethanol and vibrated for 2 h under sonication, then centrifuged at 25°C and 12,000 × *g* for 10 min. Total phenols content was determined using the Folin-Denis method described previously with some modifications (34). The sample extract (10 μl) was mixed briefly with 50 μl Folin-Ciocalteu phenol reagent, 10 μl distilled water and 50 μl reagent were prepared as the blank, and the mixture were kept in the dark for 3 min. Then 50 μl of 10% Na_2CO_3 and 90 μl distilled water was added, adjust the total volume to 200 μl. The sample mixture was incubated at room temperature for 60 min, and then its absorbance was measured at 760 nm by NanoDrop 2000C (Thermo Scientific, USA). Gallic acid was used as a standard and total phenols was expressed as mg GAE/g FW extract.

Total flavonoids content was determined according to the method described by Kim et al. (35). After extraction, 15 μl of 5% $NaNO_2$ were added to a 50 μl extract in a volumetric flask, and the mixture was kept for 6 min under dark at room temperature. Then, 30 μl of 10% Al $(NO_3)_3$ was added to the mixture and incubated for 6 min again. At last, 105 μl of 1 M NaOH was added. After incubating for 15 min at room temperature, the absorbance was measured at 510 nm. Results were expressed as mg of rutin equivalents (RE) per gram of FW (mg RE/g FW) for the total flavonoids content.

Measurement of ascorbic acid content was performed according to Chiaiese et al. (36). Fresh sample of 0.5 g was homogenized in 5 ml of 5% (w/v) TCA (trichloroacetic acid). After centrifugation at 6,000 rpm for 15 min, 1 ml of the supernatant was incubated with the mixture of 1 ml of TCA, 0.5 ml of 0.4% phosphoric acid-ethanol (w/v), 1 ml of 0.5% bathophenanthroline-ethanol (w/v) and 0.5 ml of 0.03% $FeCl_3$-ethanol (w/v) for 60 min at 30°C. The ascorbic acid content was calculated at 534 nm and expressed in $mg{\cdot}g^{-1}$ FW.

Assays of antioxidant activity

Ferric reducing/antioxidant power, DPPH (1,1-diphenyl-2-picrylhydrazyl radical) and ABTS (2,2′-azinobis 3-ethylbenzothiazoline-6-sulfonic acid) assays were conducted by the method in a previous report (32). FRAP reagent was prepared by mixing 200 ml acetate buffer (300 mM, pH 3.60), 20 ml of $FeCl_3{\cdot}6H_2O$ solution (20 mM) and 20 ml TPTZ solution (10 in 40 mM HCl). Then 3.80 ml FRAP reagent was added to 200 μl of peel or pulp extract. After 30 min at room temperature, the absorbance of extracts and reference substances Trolox was detected at a wavelength of 590 nm by NanoDrop 2000C (Thermo Scientific, USA). For DPPH assay, 3.50 ml of DPPH was added into 500 μl ethanol extracts. After 30 min for darkness, the absorbance was detected at wavelength of 517 nm. For ABTS assay, fruit extracts (40 μl) were allowed to react with 3.90 ml of the ABTS radical solution under dark conditions for 10 min, and then the absorbance at 734 nm was measured. The results of ABTS, DPPH, and FRAP were represented as μmol of Trolox equivalent per gram fresh weight (μmol Trolox/g FW) for peels and pulps. Three repeats were performed for one sample.

Statistical analysis

All tests were conducted in triplicate. Statistical analysis was performed using IBM SPSS Statistics 22.0. Significant differences of the contents of total phenols, total flavonoids, ascorbic acid, and the antioxidants activities of FRAPs, DPPH and ABTS in peel and pulp of the four different citrus varieties were calculated. Tukey's test was performed by using one-way analysis of variance (ANOVA) at the 5% level ($P < 0.05$). The Pearson correlation analysis was also performed by SPSS 22.0 at $P < 0.05$ for determination of the correlations among the contents of total phenols, total flavonoids, and vitamin C and the antioxidant activities of citrus peel and pulp, respectively.

Results

Biological characteristics and phylogenetic analysis of two local citrus varieties

Zangju (*Citrus sunki* Hort. ex Sakurai cv. Zangju), also known as Dengrongzhoupigan, is mainly distributed in Derong county of

Ganzi Tibetan Autonomous Prefecture of Sichuan province, with a longitude of 99° 30′ 15″ and a latitude of 28° 61′ 27″. The planting area of Zangju in the county is about 1,000 μ, mainly distributed in Bendu and Guxue towns. The biological characteristics of Zangju are as follows (Figure 1A and Table 1): trees of Zangju are moderately large and vigorous, with semicircular crowns. Young shoots of Zangju are light green with simple leaves, wedge-shaped leaf base, linear wing leaves with an acuminate tip, entire leaf margins. Zangju flowers occur singly or in clusters, and are complete flowers with 5–6 white petals, medium pollen, filaments partially united, erect light-yellow style, and the length yellow anthers are longer than stigma. Zangju fruit is slightly flattened in shape with large fruit size for a mandarin, with deep radiating grooves and rough in the rind, dense oil cells. The single weight of fruit is 150 g, the transverse and longitudinal diameter is 7.89 and 5.70 cm, and the fruit shape index is 0.72; the rind is orange-yellow, with the thickness 6–8 mm, rather large open core, easy to peel, segments 9–10, numerous oval monoembryonic seeds (13–15). The flesh is deep orange in color with a moderately fine texture, fragrant and rich in juice. Fruit quality characteristics were measured for Zangju, including the soluble solids content (11.3°Brix) and total organic acid content (1.03%), solid-acid ratio of 11.07, slightly sour taste. The fruit of Zangju matures at late November, and can be picked until March.

Tuju (*Citrus erythrosa* Hort. ex Tanaka cv. Tuju), is mainly distributed in Jiukeng town, Chunan county of Zhejiang province, with a longitude of 118° 38′ 42″ and a latitude of 29° 38′ 15″. There is little amount of cultivation of Tuju in the local area, and it has been developed and utilized in recent years to process tangerine peel, small green tangerine, tea wine, etc. Hence, the cultivation of Tuju has been gradually expanded. The biological characteristics of Tuju are as follows (Figure 1B and Table 1): trees of Tuju are evergreen, moderate tree with vigor good cold resistance. Young shoots of Tuju are light green with simple leaves, showing similar leaf characteristics with Zangju. Tuju flowers occur singly or in clusters, and are complete flowers with 5 white petals, medium pollen, filaments partially united, erect green-yellow style, and the length yellow anthers are longer than stigma. Tuju fruit is globe in shape, no radiating grooves and smooth in the rind, dense, and fine oil cells. The single weight of fruit is 47 g, the transverse and longitudinal diameter is 4.43 and 3.76 cm, and the fruit shape index is 0.85; the rind is red and orange, with the thickness 2-3 mm, rather little open core, easy to peel, segments 10–11, numerous oval polyembryonic seeds (12–15). The flesh is red and orange in color with a moderately fine texture, fragrant and rich in juice. The soluble solids content of fruit is 14.5°Brix, total organic acid content is 0.65%, and solid-acid ratio of juice is 22.64. The fruit of Tuju matures at early December with sweet and moderate sour taste.

Chloroplast genomes of fruit trees play an important role in phylogenetic studies. To understand the evolutionary relationships of Zangju and Tuju with other species, the complete chloroplast genome sequences of 22 species of the genus *Citrus*, including typical mandarin, sweet orange, lemon, pomelo, sour orange, and citron, were used to construct a phylogenetic relationship tree, with *Anacardium occidentale* and *Dimocarpus longan* as the outgroups (Figure 1C). The phylogenetic analysis revealed that Zangju belonged to *C. reticulata* and was phylogenetically closest with sour mandarin *C. sunki*, whereas Tuju was a kind of red orange that closely clustered with *C. erythrosa*. The phylogenetic analysis also indicated that Zangju and Tuju from the south and north of China have a distant phylogenetic relationship.

Phytochemical metabolites analysis in the peel and pulp of ZG and TG

Metabolites profiles in the peel and pulp

To better understand the nutritional and medicinal differences among Zangju (ZG), Tuju (TG), Daidai (DD), and satsuma mandarin (WZMG), we performed widely targeted UPLC-MS/MS-based non-targeted metabolite profiling of these four species. The satsuma mandarin WZMG, the most widely cultivated variety in China (37) and DD fruit (*Citrus aurantium* L. var. *daidai*) contained the higher amounts of flavonoids and medicinal components than other citrus species according (38). Hence, here we used the sour orange DD and satsuma mandarin WZMG as controls (Figure 2A).

To obtain a clear overview of the clear separation among the fruit peel and pulp of the four citrus varieties, we conducted an unsupervised principal component analysis (PCA) (Figure 2B), revealing that the first and second principal components (PC1 and PC2) displayed 36.79 and 22.17% of the variation, respectively. PCA analysis of metabolomics data showed that 24 samples of fruit peel and pulp were significantly separated into 8 groups (Figure 2B), indicating that the fruit samples from every group exhibited different metabolic characters and the metabolomics data were reliable. We performed an intragroup correlation analysis and found that biological replicates of the samples from the same variety were highly correlated ($r^2 > 0.7$), whereas peel and pulp samples showed significant differences (Figure 2C). As shown in Figure 2D and Supplementary Table 1, 1,126 metabolites were totally identified, of which there were a large number of secondary metabolites likely to contribute to the medicinal value, and some primary metabolites are related to the nutritional quality. Similar to the previous studies (24, 25), it was also shown that the relative abundance of most of the detected metabolites in the peel of TG and ZG fruits were significantly higher than that in the pulp (Figure 2D).

A total of 1,042 differentially accumulated metabolites (DAMs) were detected in the peel (Figure 3A), and 985 DAMs were detected in the pulp (Figure 3B). TG peel contains most secondary metabolites with higher abundance, followed by DD, WZMG, and ZG (Figure 3A and Supplementary Table 1). Similarly, TG pulp contains most abundant secondary metabolites, followed by DD, ZG, and WZMG (Figure 3B and Supplementary Table 1). Additionally, among all the DAMs, carboxylic acids and derivatives, benzene and substituted derivatives, organooxygen compounds, flavonoids, fatty acyls glycerophospholipids, prenol lipids, coumarins and derivatives, phenols, etc. were listed in the top 20 classification (Figure 3C). As shown in Figure 3D, based on FPKM (fragments per kilobase million) data, there were 831, 642, 564, and 429 DAMs were identified in the groups of TG_P vs. DD_P, TG_P vs. WZMG_P, TG_R vs. DD_R, TG_R vs. WZMG_R, respectively. Furthermore, 751, 628, 543, and 491 DAMs were identified in the groups of ZG_P vs. DD_P, ZG_P vs. WZMG_P, ZG_R vs. DD_R, and ZG_R vs. WZMG_R, respectively (Figure 3D). Compared with DD peel, TG peel contained more distinct DAMs than ZG peel, 652 DAMs in common, and 450 conserved DAMs were found between TG and ZG peels vs. WZMG peel (Figure 3E). Compared with DD pulp, there were 418 DAMs in common between TG and ZG pulp, whereas, compared with WZMG pulp, ZG pulp contained more distinct DAMs than TG pulp, with 314 conserved DAMs (Figure 3F). These results also revealed significant diversity and specificity of the metabolites in the peel and pulp of TG and ZG fruits.

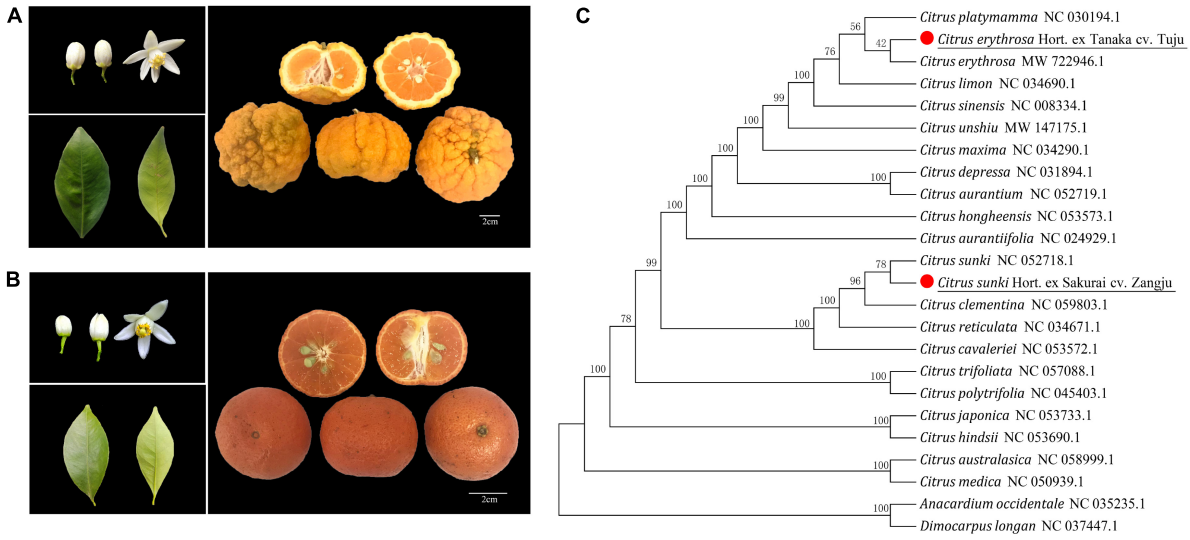


FIGURE 1
Biological characteristics and phylogenetic analysis of two local citrus varieties. **(A)** Biological characteristics of leaves, flowers, and fruits of Zangju. **(B)** Biological characteristics of leaves, flowers, and fruits of Tuju. **(C)** Phylogenetic analysis of Zangju and Tuju with other citrus species. The phylogenetic tree of 24 citrus and related species was conducted using chloroplast DNA sequences data.

TABLE 1 Fruit characteristics and quality of Zangju and Tuju.

Index	Segment number	Seed number	Single fruit weight (g)	Equatorial diameter (cm)	Fruit height (cm)	Fruit shape index	TSS (Brix)	Titratable acidity (%)	TSS/TA ratio
Zangju	9.83 ± 0.43	13.75 ± 0.96	147.01 ± 7.10	7.89 ± 0.42	5.70 ± 0.39	0.72 ± 0.05	11.30 ± 0.54	1.03 ± 0.09	11.07 ± 1.15
Tuju	10.67 ± 1.03	13.25 ± 1.26	47.04 ± 4.88	4.43 ± 0.2	3.76 ± 0.01	0.85 ± 0.05	14.48 ± 0.78	0.65 ± 0.06	22.64 ± 3.22

Data are expressed as mean ± SE.

Analysis of the nutrition and antioxidant related DAMs in TG and ZG

In order to explore the metabolite function of TG and ZG fruits, we further analyzed the KEGG enrichment pathway of the DAMs. The results showed that, compared with the peel and pulp of DD and WZMG, the DAMs in TG and ZG fruit were mainly enriched in ascorbate and aldarate metabolism, tyrosine metabolism, flavonoid biosynthesis, isoflavonoid biosynthesis, biosynthesis of alkaloids derived from terpenoid and polyketide, flavone and flavonol biosynthesis, biosynthesis of phenylpropanoids, phenylalanine metabolism, pyruvate metabolism, tryptophan metabolism, starch and sucrose metabolism, biosynthesis of terpenoids and steroids, xylene degradation, and citrate cycle (TCA cycle) pathway (Figures 4A–D and Supplementary Figure 1). According to the KEGG enrichment pathway analysis of the DAMs in the peel (Figures 4A–D) and pulp (Supplementary Figure 1), relative abundance of main DAMs in the key pathways was displayed in the heat map. The major DAMs in the pulp and peel included some primary and secondary metabolites related to nutrition, such as sugars, amino acids, organic acids, and antioxidant activity, such as flavonoids, flavanones, isoflavonoids, vitamin C, and phenolic acids (Figure 4E). As shown in the heat map, both TG and ZG peel contains several DAMs with significantly higher abundance, whereas, less DAMs with high abundance were found in the pulp of TG and ZG.

The antioxidant activity and medicinal value of citrus fruits are mainly related to flavonoids, phenolic acids, and vitamin C (17, 21). Based on the metabolome data, we analyzed the relative abundance of ascorbic acid and major flavonoids and phenolic acids in the peel and pulp of ZG and TG. The relative abundances of ascorbic acid and key flavonoids and phenolic acids are listed in Table 2. The relative abundance of ascorbic acid in ZG and TG was higher than that in DD and WZMG, except for the content in the pulp of TG (Table 2). Compared with DD and WZMG peel, the relative abundance of some key flavonoids and phenolic acid, such as diosmetin, sinensetin, tangeritin, nobiletin, quercetin (quercetin 5,7,3′,4′-tetramethyl ether), scutellarein (scutellarein 6,7-dimethyl ether 4′-glucoside), sinapic acid, etc. were significantly higher in TG peel (Table 2). Furthermore, ZG peel showed relatively higher abundance of flavonoids and phenolic acid, such as diosmetin, sinensetin, tangeritin, nobiletin and caffeic acid, gallic acid, sinapic acid, etc. compared with DD and WZMG peel (Table 2). Additionally, both the relative abundance of naringin and neohesperidin in ZG peel and pulp are higher than that in WZMG. Moreover, TG pulp recorded higher relative abundance of sinensetin, tangeritin, heptamethoxyflavone (3,5,6,7,8,3′,4′-heptamethoxyflavone), nobiletin, caffeic acid, sinapic acid, chlorogenic acid, etc. than DD and WZMG pulp (Table 2). ZG pulp showed relatively higher abundance of key flavonoid and phenolic acid, such as tangeritin, caffeic acid, gallic acid, sinapic acid, chlorogenic acid, etc. compared to DD and WZMG pulp. Additionally, ZG and TG peel and pulp recorded relatively higher content of hesperidin than DD, but lower than the previously reported the highest hesperidin content in WZMG (39). In

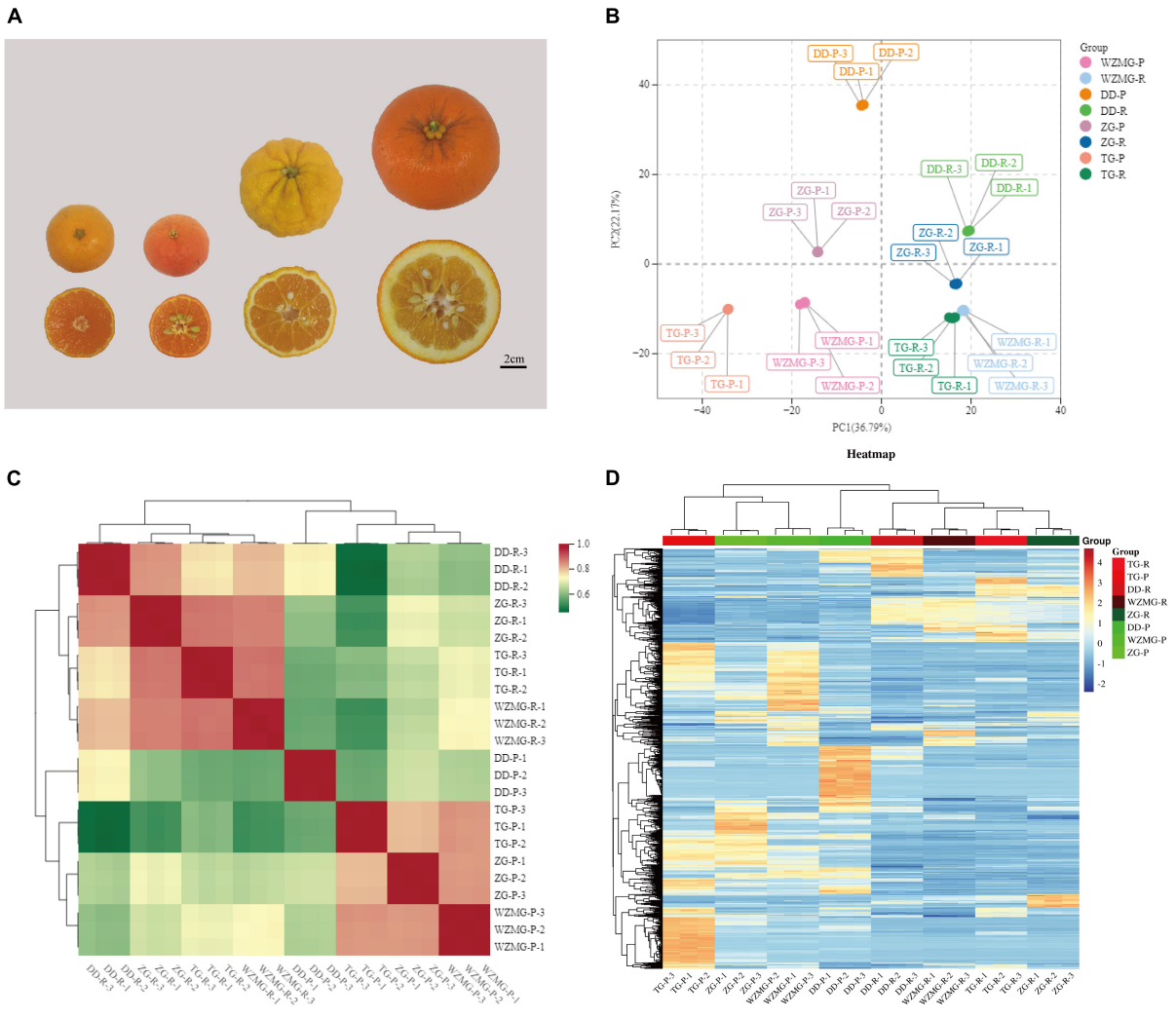


FIGURE 2
Samples and metabolites profiles in the peel and pulp. **(A)** Samples of Zangju (ZG), Tuju (TG), Daidai (DD), and satsuma mandarin (WZMG). **(B)** Principal component analysis (PCA) score plot of the first and second principal components of all the samples. **(C)** Heatmap of biological replicate sample correlation analysis. **(D)** Heatmap of the total 1,126 detected metabolites of all groups.

conclusion, these findings suggest that the peel and pulp of these two local varieties ZG and TG are rich of some key flavonoids and phenolic acid.

Importantly, besides the polyphenols listed in Table 2, metabolite profiling showed that some secondary metabolites, such as bergenino was only detected in ZG and TG, with the highest relative content in the peel of TG (Supplementary Table 1). Furthermore, 5-hydroxyferulate, harpagoside, 1-naphthol, and methylenedioxycinnamic acid were only found in the peel of ZG (Supplementary Table 1). Therefore, further research on the extraction and application of these special metabolites in ZG and TG is needed in the future.

Comparative analysis of the contents of total phenols, total flavonoids, and vitamin C in TG and ZG fruit

Citrus peel is a rich source of naturally occurring antioxidants (39). Antioxidant activity of citrus peel is due to the abundance of phenols, flavonoids, and ascorbic acid (25, 39). Here, as shown in Figure 5A, satsuma mandarin WZMG peel has the highest total phenols content of 5.61 mg GAE/g FW, followed by TG peel (5.40 mg GAE/g FW), ZG peel (4.58 mg GAE/g FW), DD peel (3.50 mg GAE/g FW). Furthermore, TG peel has highest content of total flavonoids of 3.40 mg RE/g FW, followed by ZG (2.85 mg RE/g FW), WZMG (2.18 mg RE/g FW), and DD (1.78 mg RE/g FW). The TG peel recorded the highest total ascorbic acid content (0.31 mg/g FW), followed by ZG (0.22 mg/g FW), WZMG (0.20 mg/g FW), and DD (0.16 mg/g FW).

Moreover, the TG pulp contained the highest content of total phenols of 1.64 mg GAE/g FW (Figure 5B), followed by DD pulp (1.41 mg GAE/g FW), ZG pulp (1.16 mg GAE/g FW), WZMG pulp (0.77 mg GAE/g FW). ZG pulp recorded the highest total flavonoids content (0.57 mg RE/g FW), followed by DD pulp (0.54 mg RE/g FW), TG pulp (0.38 mg RE/g FW) and WZMG pulp (0.20 mg RE/g FW). The contents of total flavonoids and phenols in the peel of the four different citrus varieties were significantly higher than that in the pulp, which is consistent with the previous studies (39). Additionally, ZG pulp has the highest ascorbic acid content (0.36 mg/g FW), followed by DD pulp, WZMG pulp, and TG pulp (Figure 5B).

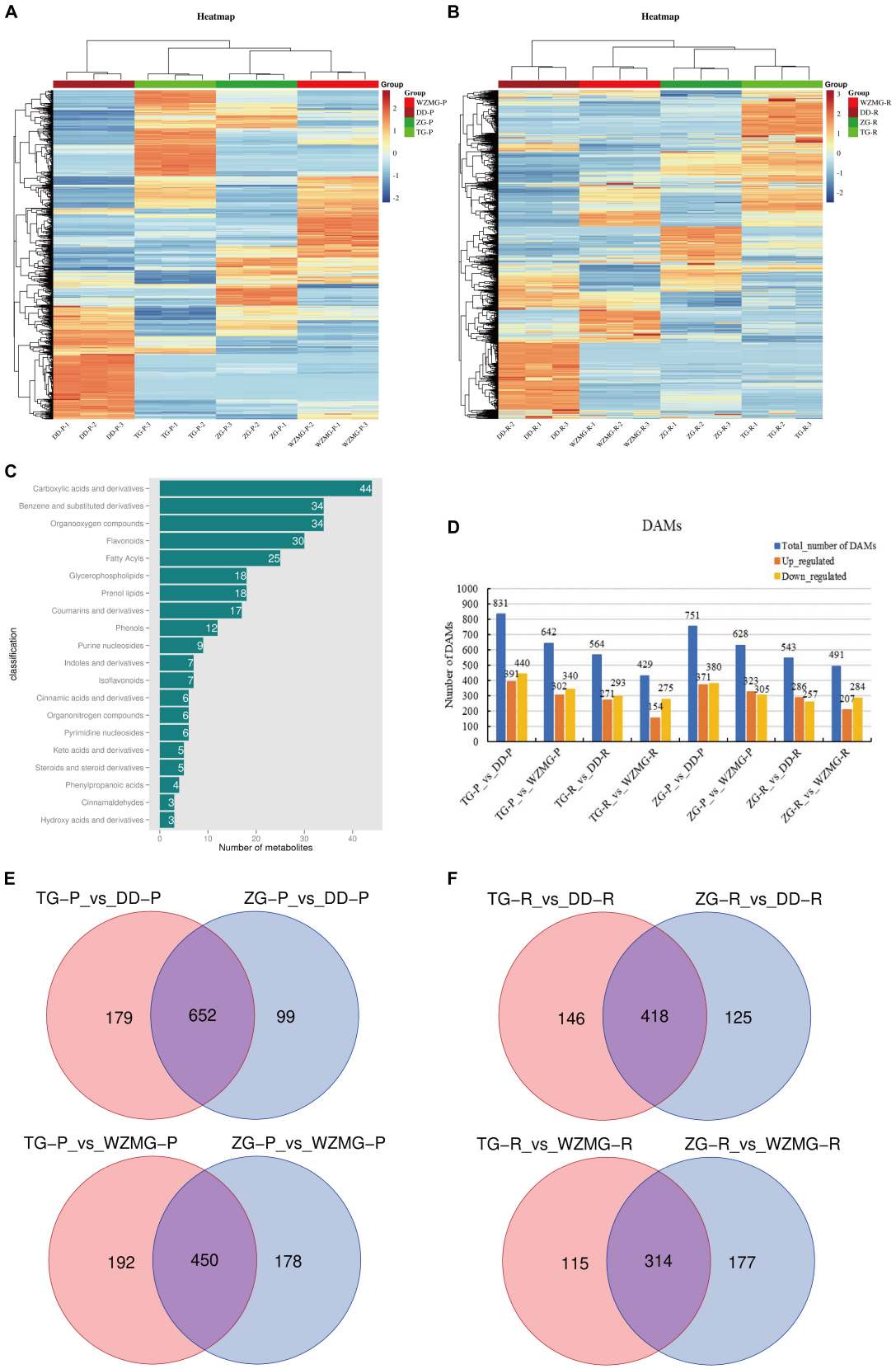


FIGURE 3

Comparative analysis of differentially accumulated metabolites (DAMs) in the pulp and peel of ZG, TG, DD, and WZMG. **(A)** Heatmap of the different metabolites among the peel (P) of ZG, TG, DD, and WZMG. **(B)** Heatmap of the different metabolites among the pulps (R) of ZG, TG, DD, and WZMG. **(C)** Top 20 classifications of all different metabolites. **(D)** The number of up- or down-regulated DAMs in different comparison groups. **(E)** The DAMs number of TG-P (ZG-P) vs. DD-P and TG-P (ZG-P) vs. WZMG-P shown in Venn diagrams. **(F)** The DAMs number of TG-R (ZG-R) vs. DD-R and TG-R (ZG-R) vs. WZMG-R shown in Venn diagrams.

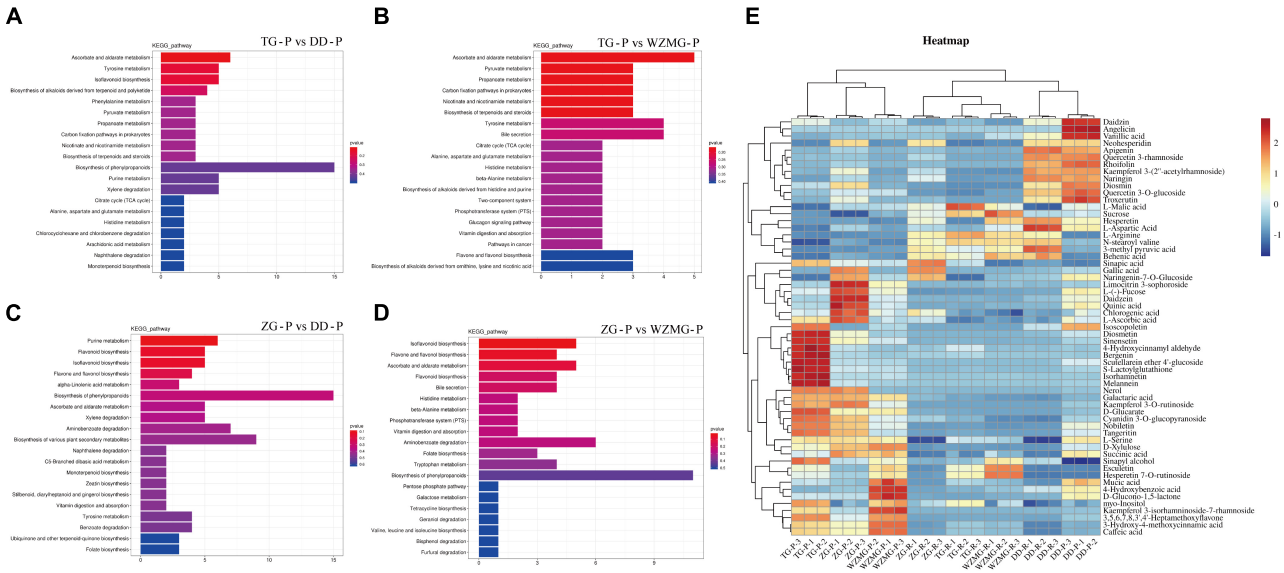


FIGURE 4
Analysis of the DAMs and nutrition and antioxidant metabolites in ZG and TG. KEGG pathway enrichment analysis of DAMs in groups TG-P vs. DD-P **(A)**, TG-P vs. WZMG-P **(B)**, ZG-P vs. DD-P **(C)**, and ZG-P vs. WZMG-P **(D)**. **(E)** Heatmap of the key DAMs in the peel and pulp.

TABLE 2 Represents the log_2FC of the main flavonoids, phenolic acid, and vitamin C in fruit peel and pulp between different groups.

Metabolites name	Category	Log2 of fold change of different groups of fruit peel				Log2 of fold change of different groups of fruit pulp			
		TG-P vs. DD-P	TG-P vs. WZMG-P	ZG-P vs. DD-P	ZG-P vs. WZMG-P	TG-R vs. DD-R	TG-R vs. WZMG-R	ZG-R vs. DD-R	ZG-R vs. WZMG-R
Ascorbic acid	Vitamin C	1.61	1.29	0.93	1.25	−0.17	−0.31	1.11	0.97
Apigenin	Flavonoids	−5.20	−0.03	−2.63	2.54	−7.50	0.78	−4.00	4.28
Diosmetin	Flavonoids	2.37	3.34	0.85	1.82	−1.80	0.81	−2.39	0.23
Diosmin	Flavonoids	−1.65	1.17	−0.47	2.35	−4.39	−0.30	−3.05	1.05
Naringin	Flavonoids	−4.97	−1.40	−1.85	1.72	−6.40	−1.14	−1.33	3.92
Hesperetin	Flavonoids	−2.30	0.50	−2.03	0.77	−3.66	−3.39	−0.84	−0.57
Hesperidin	Flavonoids	3.48	−0.48	2.75	−1.21	4.13	−0.74	1.91	−2.96
Neohesperidin	Flavonoids	−3.53	−0.61	−0.03	2.89	−4.52	−1.22	−0.08	3.21
Sinensetin	Flavonoids	3.15	2.15	1.79	0.79	2.16	3.00	−0.62	0.22
Tangeritin	Flavonoids	1.51	1.84	1.19	1.51	2.88	3.57	1.39	2.08
Heptamethoxyflavone	Flavonoids	2.74	0.04	1.34	−1.37	2.50	0.23	−0.10	−2.37
Kaempferol	Flavonoids	−2.91	1.10	−0.92	3.09	−5.64	3.84	−1.61	7.86
Nobiletin	Flavonoids	1.11	1.30	0.75	0.94	1.68	3.13	−0.10	1.35
Quercetin	Flavonoids	4.32	2.67	1.54	−0.98	7.12	3.52	0.00	−6.79
Scutellarein	Flavonoids	2.25	3.37	−0.80	0.31	−2.98	−1.43	−0.68	0.87
Troxerutin	Flavonoids	−4.33	−0.45	−1.43	2.45	−7.59	0.92	−2.58	5.93
Caffeic acid	Phenolic acid	1.67	−0.42	1.36	−0.73	1.29	0.10	0.65	−0.54
Gallic acid	Phenolic acid	−4.87	−1.48	2.48	5.87	−2.35	−0.64	3.96	5.67
Vanillic acid	Phenolic acid	−6.50	−2.92	−3.81	−0.24	−6.11	−4.96	−2.38	−1.23
Sinapic acid	Phenolic acid	2.01	0.79	1.40	0.18	0.76	1.31	1.85	2.39
Chlorogenic acid	Phenolic acid	0.04	−0.03	0.01	0.26	0.73	0.66	0.62	0.87
4-Hydroxybenzoic acid	Phenolic acid	−2.49	−3.47	−1.79	−2.76	−1.59	−0.63	−2.37	−1.41

Comparative analysis of antioxidant activities of the peel and pulp of TG and ZG fruits

Phenolic compounds, including flavonoids and phenolic acids, are known to be responsible for antioxidant activity of different fruits, and fruits with higher total phenolic content generally showed stronger antioxidant capacity (40). The antioxidants activities of peel and pulp of ZG and TG were evaluated through FRAPs, DPPH, and ABTS assays. The FRAP activity of the four citrus varieties peels ranged from 13.60 to 26.31 μmol Trolox/g FW. The TG peel displayed the strongest activity, followed by satsuma mandarin WZMG peel, DD peel, and ZG peel (**Figure 6A**). The DPPH value of the four citrus varieties peels ranged from 4.62 to 11.04 μmol Trolox/g FW, where TG peel showed the strongest DPPH scavenging activity, followed by WZMG peel, ZG peel, and DD peel (**Figure 6A**). Furthermore, ABTS scavenging activity of the four citrus varieties peels ranged from 32.95 to 61.51 μmol Trolox/g FW, where WZMG peel exhibited the strongest activity, and the rest were in descending order: TG peel > DD peel > ZG peel (**Figure 6A**). In conclusion, these results suggest that TG peel exhibits the stronger antioxidant activity.

Further, FRAP activity of the four citrus varieties pulp ranged from 4.39 to 7.98 μmol Trolox/g FW, where ZG Pulp showed the strongest FRAP activity, followed in the order TG pulp > DD pulp > WZMG pulp (**Figure 6B**). The DPPH scavenging activities of the four citrus pulps ranged from 0.97 to 2.41 μmol Trolox/g FW, where DD pulp displayed strongest activity, followed by ZG pulp, TG pulp, and WZMG pulp (**Figure 6B**). The value of ABTS scavenging activity were between 3.50 and 10.41 μmol Trolox/g FW, where TG pulp showed the significantly higher activity, followed by ZG pulp, DD pulp, and WZMG pulp (**Figure 6B**). These results indicate that TG and ZG pulps exhibit relatively stronger antioxidant activities.

Correlation analysis between antioxidant activities and contents of total phenols, flavonoids, and vitamin C in peel and pulp

Recently, few studies concluded that total antioxidant activity of citrus fruits are mainly attributed to phenolic compounds and vitamin C, though there is disagreement as to which compound is the major contributor (28). To identify the chemical compounds which contribute to the antioxidant activity of the four citrus fruits, Pearson's correlation coefficients between the contents of total phenols, total flavonoids, and vitamin C and the antioxidant activity of citrus peel and pulp were analyzed.

As shown in **Table 3**, the antioxidant activity (FRAP) of the peels of the four citrus varieties had the strongest correlation with the total phenols content ($r = 0.873$, $P < 0.01$), followed by the vitamin C content ($r = 0.782$, $P < 0.01$) and total flavonoid content ($r = 0.673$, $P < 0.05$). The correlation between DPPH value and vitamin C content was the strongest ($r = 0.829$, $P < 0.01$), followed by the correlation with total flavonoids ($r = 0.740$, $P < 0.01$), but there was no significant correlation with total phenols content. However, ABTS value in peel showed no significant correlation with the content of total flavonoids, total phenols, and vitamin C. For pulp, FRAP of the four citrus varieties had the strongest correlation with the content of total flavonoids ($r = 0.726$, $P < 0.01$), but displayed no significant correlation with the content of total phenols and vitamin C. DPPH value of pulp also had the strongest correlation with the content of total flavonoids ($r = 0.857$, $P < 0.01$), followed by the content of vitamin C ($r = 0.735$, $P < 0.01$), but had no significant correlation with the content of total phenols. However, ABTS value of pulp showed the strongest significant correlation with total phenols content ($r = 0.731$, $P < 0.01$). These results suggest that the main antioxidant components in citrus fruits are phenols, flavonoids, and vitamin C, which is consistent with previous reports (8).

Discussion

In the present study, the morphological characteristics, phylogenetic relationships, phytochemical profiling, and antioxidant activities of two local citrus varieties, Zangju (ZG) and Tuju (TG), were systematically analyzed for the first time. We found that ZG fruit grown in Sichuan province has a little acidic fruit flavor and phylogenetically closest relationship with *C. sunki*, whereas TG grown in Zhejiang province has sweet fruit taste and phylogenetically closest with *C. erythrosa*. Fruits qualities and flavors of ZG and TG are consistent with the dietary habits of local residents, which is also one of the reasons why these two citrus varieties have been cultivated so far. Both Sichuan and Zhejiang provinces are the main producing areas of citrus in China and rich in citrus varieties (41). These two local varieties are still be preserved and cultivated in citrus producing areas for a long time, indicating that their fruits may contain some metabolites with beneficial effects on human health besides the characteristic flavors.

Citrus fruits are highly nutritious, containing many primary metabolites, which are important sources of essential nutrients, and secondary metabolites, which form an excellent source of bioactive substances and exhibit potent health-promoting effects (42). Chemical profiles, such as primary and secondary metabolites related to taste, color, and health benefits, are significantly different depending on the citrus varieties, leading to different general quality parameters and antioxidant activities (43–45). Several previous reports showed that the highest levels of total phenols and total flavonoids were found in mandarin (25–29). The phytochemical metabolites analysis showed that TG peel and pulp recorded higher abundance of some important secondary metabolites compared with the controls DD and mandarin WZMG. The relative abundance of flavonoids, such as tangeritin, sinensetin, diosmetin, and nobiletin, in the peel and pulp of ZG and TG were higher than that in sour range DD and mandarin WZMG. Nobiletin and tangeritin are the primary and most widely distributed flavonoids for their bioactivities in the peels of different citrus species (46, 47). Mandarin peel is rich in nobiletin as compared to other citrus species, and the concentration of isolated nobiletin varied significantly among the peel extracts of mandarin, sweet orange, white grapefruit and lime with striking different values of 202.91, 73.15, 18.13, and 0.09 μg/ml, respectively (48). However, our results showed that the contents of nobiletin and tangeritin in both the peel and pulp of ZG and TG are relative higher than that in mandarin WZMG, especially in the pulp. Furthermore, both the relative content of naringin and neohesperidin in ZG peel and pulp are significantly higher than that in WZMG, similar with that in two wild zhoupigan varieties (8). These relatively abundant flavonoids in ZG and TG, including naringin, neohesperidin, tangeritin, sinensetin, diosmetin, and nobiletin, are the main flavonoids in most citrus, which is consistent with previous reports (27, 32). Hence, higher contents of these flavonoids with important medicinal value (49) will furtherly increase the economic value of ZG and TG fruits. Phenolic acid,

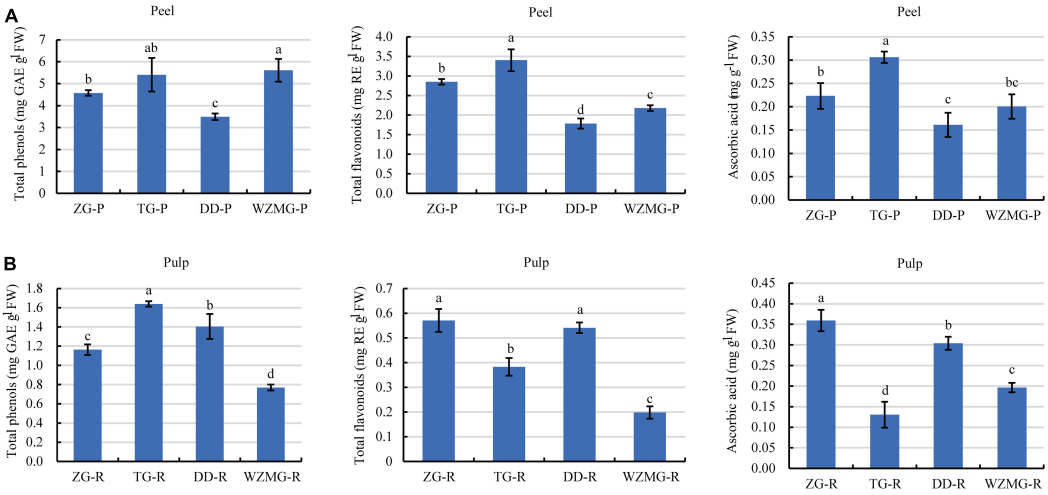


FIGURE 5
Comparative analysis of the contents of total phenols, total flavonoids, and ascorbic acid in peel and pulp. **(A)** Contents of total phenols, total flavonoids, and ascorbic acid in ZG-P, TG-P, DD-P, and WZMG-P. **(B)** Contents of total phenols, total flavonoids, and ascorbic acid in ZG-R, TG-R, DD-R, and WZMG-R. The different letters in each bar indicate significant differences by Tukey's test ($P < 0.05$). Total phenols and total flavonoids were expressed as gallic acid equivalents (GAE) and rutin equivalents (RE), respectively. FW, fresh weight.

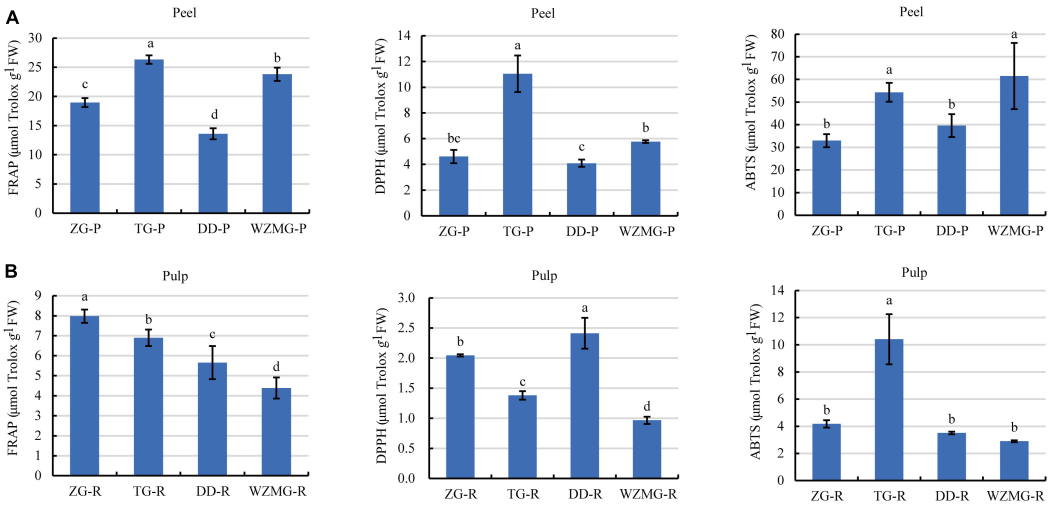


FIGURE 6
Comparative analysis of antioxidant activity of peel and pulp. **(A)** FRAP activity, DPPH scavenging activity, and ABTS scavenging activity of ZG-P, TG-P, DD-P, and WZMG-P. **(B)** FRAP activity, DPPH scavenging activity, and ABTS scavenging activity of ZG-R, TG-R, DD-R, and WZMG-R. The different letters in each bar indicate significant differences by Tukey's test ($P < 0.05$). FRAP, DPPH, and ABTS value were expressed as μmol Trolox equivalents/g FW. FW, fresh weight.

such as caffeic acid, chlorogenic acid, and sinapic acid, contents are abundant in ZG and TG, except for gallic acid, which is contrary to some previously published reports (50–52). For example, ferulic acid was identified as the most abundant and caffeic acid as the least abundant phenolic acid in kinnow peel extract (53). Ferulic acid was quantified as a major phenolic acid and caffeic acid as minor in peels of citrus fruits including lemons, oranges and grapefruits, the levels of which were significantly higher than those of peeled fruits (50). The level of chlorogenic, caffeic, and ferulic acid were the highest of phenolic acids in citrus hybrids peels from China (51). Similarly, chlorogenic acid is the phenolic acid with highest abundance in ZG, TG, DD, and WZMG (**Supplementary Table 1**). However, gallic acid is identified as the major phenolic acid in all grapefruit tested (52). Besides the major flavonoids and phenolic acid listed above, some secondary metabolites, such as bergenino was only detected in ZG and TG, with the highest relative content in the peel of TG. Furthermore, 5-hydroxyferulate, harpagoside, 1-naphthol, and methylenedioxycinnamic acid were only found in the peel of TG. Hence, it is necessary to further study the extraction and application of these special metabolites in ZG and TG in the future.

In general, citrus fruits are considered as one of the natural resources of antioxidants, which contain an appreciable amount of ascorbic acid, flavonoids, and phenols compounds (54–56). These active antioxidants have a wide variety of biological activities of importance to human health (3, 11–15). However, due to the dissimilarities in composition of antioxidants, antioxidant activities of fruit varies among different citrus species and tissues (39, 50). All the four citrus peels presented higher total flavonoids and phenols

TABLE 3 Correlation analysis of total phenols, total flavonoids, and vitamin C content and antioxidant activity in peel and pulp.

Citrus tissues	Compounds	FRAP	DPPH	ABTS
Peel	Total flavonoids	0.673*	0.740**	0.088
	Total phenols	0.873**	0.485	0.472
	Vitamin C	0.782**	0.850**	0.172
Pulp	Total flavonoids	0.726**	0.857**	0.155
	Total phenols	0.467	0.437	0.731**
	Vitamin C	0.341	0.735**	−0.602*

*Correlation is significant at the 0.05 level. **Correlation is significant at the 0.01 level (two-tailed).

content than pulps, which is consistent with the previous reports (57, 58). Though antioxidants, such as flavonoids, phenols, in DD and WZMG are higher than some citrus varieties (27), our results showed that the peel or pulp of ZG and TG have higher contents of total flavonoids, total phenols, and vitamin C than DD and WZMG, indicating that both ZG and TG have much more richer antioxidants that provide human health benefits, such as antioxidative, anti-inflammatory, anticancer, and cardiovascular protective activities (59). Additionally, it is previously reported that DD and WZMG peels or pulps have stronger antioxidant activity than some other citrus varieties (25, 60), and it has also been reported that DD has the strongest antioxidant activity in several sour oranges (61). However, compared to DD and WZMG, TG peel and ZG pulp have the strongest FRAP activity and a higher DPPH scavenging activity, and TG pulp has the highest ABTS scavenging activity. This may be due to the relatively higher content of some major flavonoids in TG and ZG, such as nobiletin and tangeritin. Previously, Chen et al. reported that the orange peel collected from China contained the highest content of nobiletin and tangeritin compared to orange peel collected from USA and showed highest ABTS activity and DPPH free radical scavenging activity (62). All the four citrus peels showed much more higher antioxidant activities than pulps, due to the significantly higher contents of total flavonoids, total phenols and ascorbic acid in the peels, and these results are similar with the results observed by Xi et al. (58) and Nogata et al. (63). In summary, based on the higher amount of antioxidant compounds and stronger antioxidant activities in the fruits, we inferred that these two local varieties could offer some medicinal value for human health.

In the present study, we found that both FRAP value of the peels and pulp of the four citrus varieties displayed the strongest correlation with the total phenols content and total flavonoids, respectively, and this observation is in accordance with the previous reports that a positive correlation between the polyphenols contents and the antioxidant activities of different citrus germplasms (64). However, our results showed that the DPPH value of peel and pulp recorded strong correlation with total flavonoids and vitamin C content, but no significant correlation with total phenols content. The antioxidant activity may not always strongly correlate with phenols compounds. This is due to the correlation between DPPH value and antioxidants is depending on the types of fruit (65). Furthermore, a total of 39 flavonoids were identified and quantified from 35 varieties of five types of citrus fruit by Wang et al. (66), and they revealed that the correlation between DPPH value and total phenolics is also depending on the tissues of citrus fruit. In another study, Toh et al. reported that ABTS activity of two varieties of pomelo fruit positively correlated with total phenolic content and total flavonoid content, except for ascorbic acid (52). Interestingly, we found that ABTS value in peel exhibited no significant correlation with the content of total flavonoids, total phenols, and vitamin C, and only displayed the significant correlation with total phenols content in pulp. However, Arena et al. (67) stated that phenolic compounds in citrus fruits contributed less than vitamin C in establishing the antioxidant power. Whereas, others reported that the antioxidant power is mainly governed mainly by phenolic compounds than ascorbic acid (68, 69). These distinctive differences in results may be due to cultivar types, maturity of fruit or the analytical methods used for estimation of antioxidant activity (70).

Our results show that the two local citrus varieties, ZG and TG, are rich in antioxidant related secondary metabolites and have strong antioxidant activities. Based on these results, we proposed that fruits of ZG and TG have certain medicinal value. Moreover, our findings will provide an important theoretical basis for the conservation and utilization of these two local citrus germplasm resources. Further research is needed to develop and utilize the medicinal value of these two local varieties to improve the economic benefits of citrus growers.

Conclusion

In the present study, the biological, phylogenetic characteristics, and phytochemical profiles, antioxidants contents and antioxidant activities of two local citrus varieties, ZG and TG, were systematically demonstrated for the first time. The results showed that Zangju fruit had the characteristics of wrinkled skin and acidity, and grouped with *C. reticulata*, which had the closest phylogenetic relationship with sour orange *C. sunki*. Tuju is a kind of red orange with vermilion peel, small fruit and high sugar content, which is closely clustered with *C. erythrosa*. Phytochemical metabolites analysis showed that the relative content of some flavonoids and phenolic acid, such as tangeritin, sinensetin, diosmetin, nobiletin, and sinapic acid in the peel and pulp of both Zangju and Tuju were higher than that in sour range Daidai and satsuma mandarin. The contents of total flavonoids, total phenols, antioxidant activity (FRAP), DPPH and ABTS free radical scavenging capacity of peels of Zangju and Tuju were significantly higher than that in pulp. Compared with Daidai and satsuma mandarin, Zangju pulp and Tuju peel showed the strongest FRAP activity, and Tuju peel and pulp had the highest DPPH and ABTS value, respectively. Moreover, both the antioxidant activity in peel and pulp were significantly correlated with the contents of total phenols, total flavonoids, and ascorbic acid. The results suggested that Zangju and Tuju are rich in key antioxidants and have stronger antioxidant activity, indicating that they may have certain medicinal value and potential beneficial effects on human health.

Data availability statement

The original contributions presented in this study are included in the article/**Supplementary material**, further inquiries can be directed to the corresponding author.

Author contributions

LS, JX, and FK conceived and designed the experiments. JX, ZN, and JS collected the fruits. LS, JX, ZN, JS, and XH

analyzed the fruit characteristics and quality. LS, JX, LW, and ZN performed the experiments of antioxidant activities. LS, LW, and N analyzed the data and elaborated the figures. LS wrote the manuscript. LS and N revised the manuscript. FK supervised the project. All authors read and approved the final manuscript.

Funding

This research was funded by sub-project of "Breeding New Fruit Cultivars Major Project of Zhejiang Province" (grant no. 2021C02066-1), China Agriculture Research System of MOF and MARA (CARS-26-01), and Science and Technology Plan Project of Taizhou (22nya06).

Acknowledgments

We thank Lingjiang Yang in Derong county of Sichuan province and Xiaoguo Fang in Chunan county of Zhejiang province for the help of sample collecting.

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/fnut.2023.1103041/full#supplementary-material

References

1. Khan U, Sameen A, Aadil R, Shahid M, Sezen S, Zarrabi A, et al. *Citrus genus* and its waste utilization: a review on health-promoting activities and industrial application. *Evid Based Complement Altern Med.* (2021) 2021:2488804. doi: 10.1155/2021/2488804

2. Satari B, Karimi K. *Citrus* processing wastes: environmental impacts, recent advances, and future perspectives in total valorization. *Resour Conserv Recycl.* (2018) 129:153–67. doi: 10.1016/j.resconrec.2017.10.032

3. Zou Z, Xi W, Hu Y, Nie C, Zhou Z. Antioxidant activity of citrus fruits. *Food Chem.* (2016) 196:885–96. doi: 10.1016/j.foodchem.2015.09.072

4. Hou H, Bonku E, Zhai R, Zeng R, Hou Y, Yang Z, et al. Extraction of essential oil from *Citrus reticulate* Blanco peel and its antibacterial activity against *Cutibacterium acnes* (formerly *Propionibacterium acnes*). *Heliyon.* (2019) 5:e02947. doi: 10.1016/j.heliyon.2019.e02947

5. Yagi K. Lipid peroxides and human diseases. *Chem Phys Lipids.* (1987) 45:337–51. doi: 10.1016/0009-3084(87)90071-5

6. Saita E, Kondo K, Momiyama Y. Anti-inflammatory diet for atherosclerosis and coronary artery disease: antioxidant foods. *Clin Med Insights Cardiol.* (2014) 8:61–5. doi: 10.4137/CMC.S17071

7. Bocco A, Cuvelier M, Richard H, Berset C. Antioxidant activity and phenolic composition of citrus peel and seed extracts. *J Agric Food Chem.* (1998) 46:2123–9. doi: 10.1021/jf9709562

8. Zhang Y, Sun Y, Xi W, Shen Y, Qiao L, Zhong L, et al. Phenolic compositions and antioxidant capacities of Chinese wild mandarin (*Citrus reticulata* Blanco) fruits. *Food Chem.* (2014) 145:674–80. doi: 10.1016/j.foodchem.2013.08.012

9. Roginsky V, Lissi E. Review of methods to determine chain-breaking antioxidant activity in food. *Food Chem.* (2005) 92:235–54. doi: 10.1016/j.foodchem.2004.08.004

10. Bravo L. Polyphenols: chemistry, dietary sources, metabolism, and nutritional significance. *Nutr Rev.* (1998) 56:317–33. doi: 10.1111/j.1753-4887.1998.tb01670.x

11. Dalia I, Maged E, Assem M. *Citrus reticulata* Blanco cv. Santra leaf and fruit peel: a common waste products, volatile oils composition and biological activities. *J Med Plants Res.* (2016) 10:457–67.

12. Adenaike O, Abakpa G. Antioxidant compounds and health benefits of citrus fruits. *Eur J Nutr Food Saf.* (2021) 13:65–74. doi: 10.9734/ejnfs/2021/v13i230376

13. Ke Z, Xu X, Nie C, Zhou Z. *Citrus* flavonoids and human cancers. *J Food Nutr Res.* (2015) 3:341–51. doi: 10.12691/jfnr-3-5-9

14. Rajendran P, Nandakumar N, Rengarajan T, Palaniswami R, Gnanadhas E, Lakshminarasaiah U, et al. Antioxidants and human diseases. *Clin Chim Acta.* (2014) 436:332–47. doi: 10.1016/j.cca.2014.06.004

15. Zhang H, Xi W, Yang Y, Zhou X, Liu X, Yin S, et al. An on-line HPLC-FRSD system for rapid evaluation of the total antioxidant capacity of citrus fruits. *Food Chem.* (2015) 172:622–9. doi: 10.1016/j.foodchem.2014.09.121

16. Yabesh J, Prabhu S, Vijayakumar S. An ethnobotanical study of medicinal plants used by traditional healers in silent valley of Kerala India. *J Ethnopharmacol.* (2014) 154:774–89. doi: 10.1016/j.jep.2014.05.004

17. Makni M, Jemai R, Kriaa W, Chtourou Y, Fetoui H. *Citrus limon* from Tunisia: phytochemical and physicochemical properties and biological activities. *Biomed Res Int.* (2018):2018:6251546. doi: 10.1155/2018/6251546

18. Tundis R, Loizzo M, Menichini F. An overview on chemical aspects and potential health benefits of limonoids and their derivatives. *Crit Rev Food Sci Nutr.* (2014) 54:225–50. doi: 10.1080/10408398.2011.581400

19. Gabriele M, Frassinetti S, Caltavuturo L, Montero L, Dinelli G, Longo V, et al. *Citrus bergamia* powder: antioxidant, antimicrobial and anti-inflammatory properties. *J Funct Foods.* (2017) 31:255–65. doi: 10.1016/j.jff.2017.02.007

20. Rafiq S, Kaul R, Sofi S, Bashir N, Nazir F, Nayik G. citrus peel as a source of functional ingredient: a review. *J Saudi Soc Agric Sci.* (2018) 17:351–8.

21. Kaushik R, Narayanan P, Vasudevan V, Muthukumaran G, Usha A. Nutrient composition of cultivated stevia leaves and the influence of polyphenols and plant pigments on sensory and antioxidant properties of leaf extracts. *J Food Sci Technol.* (2010) 47:27–33. doi: 10.1007/s13197-010-0011-7

22. Oboh G, Ademosun A. Characterization of the antioxidant properties of phenolic extracts from some citrus peels. *J Food Sci Technol.* (2012) 49:729–36. doi: 10.1007/s13197-010-0222-y

23. Olfa T, Gargouri M, Akrouti A, Brits M, Gargouri M, Ben Ameur R, et al. A comparative study of phytochemical investigation and antioxidative activities of six citrus peel species. *Flavour Fragr J.* (2021) 36:564–75. doi: 10.1002/ffj.3662

24. Costanzo G, Vitale E, Iesce M, Naviglio D, Amoresano A, Fontanarosa C, et al. Antioxidant properties of pulp, peel and seeds of phlegrean mandarin (*Citrus reticulata* Blanco) at different stages of fruit ripening. *Antioxidants.* (2022) 11:187. doi: 10.3390/antiox11020187

25. Zhang H, Yang Y, Zhou Z. Phenolic and flavonoid contents of mandarin (*Citrus reticulata* Blanco) fruit tissues and their antioxidant capacity as evaluated by DPPH and ABTS methods. *J Integr Agr.* (2018) 17:256–63. doi: 10.1016/S2095-3119(17)61664-2

26. Xi W, Zhang G, Jiang D, Zhou Z. Phenolic compositions and antioxidant activities of grapefruit (*Citrus paradisi* Macfadyen) varieties cultivated in China. *Int J Food Sci Nutr.* (2015) 66:858–66. doi: 10.3109/09637486.2015.1095864

27. Chen Y, Pan H, Hao S, Pan D, Wang G, Yu W. Evaluation of phenolic composition and antioxidant properties of different varieties of Chinese *Citrus*. *Food Chem.* (2021) 364:130413. doi: 10.1016/j.foodchem.2021.130413

28. Xi W, Fang B, Zhao Q, Jiao B, Zhou Z. Flavonoid composition and antioxidant activities of Chinese local pummelo (*Citrus grandis* Osbeck.) varieties. *Food Chem.* (2014) 161:230–8. doi: 10.1016/j.foodchem.2014.04.001

29. Zhang M, Duan C, Zang Y, Huang Z, Liu G. The flavonoid composition of flavedo and juice from the pummelo cultivar (*Citrus grandis* (L.) Osbeck) and the grapefruit cultivar (*Citrus paradisi*) from China. *Food Chem.* (2011) 129:1530–6. doi: 10.1016/j.foodchem.2011.05.136

30. Rao M, Zuo H, Xu Q. Genomic insights into citrus domestication and its important agronomic traits. *Plant Commun.* (2021) 2:100138. doi: 10.1016/j.xplc.2020.100138

31. Wu G, Terol J, Ibanez V, López-García A, Pérez-Román E, Borredá C, et al. Genomics of the origin and evolution of citrus. *Nature.* (2018) 554:311–6. doi: 10.1038/nature25447

32. Chen Q, Wang D, Tan C, Hu Y, Sundararajan B, Zhou Z. Profiling of flavonoid and antioxidant activity of fruit tissues from 27 Chinese local citrus cultivars. *Plants.* (2020) 9:196. doi: 10.3390/plants9020196

33. Katoh K, Kuma K, Toh H, Miyata T. MAFFT version 5: improvement in accuracy of multiple sequence alignment. *Nucleic Acids Res.* (2005) 33:511–8.

34. Xu G, Liu D, Chen J, Ye X, Ma Y, Shi J. Juice components and antioxidant capacity of citrus varieties cultivated in China. *Food Chem.* (2008) 106:545–51. doi: 10.1016/j.foodchem.2007.06.046

35. Kim D, Jeong S, Lee C. Antioxidant capacity of phenolic phytochemicals from various cultivars of plums. *Food Chem.* (2003) 81:321–6. doi: 10.1016/S0308-8146(02)00423-5

36. Chiaiese P, Corrado G, Minutolo M, Barone A, Errico A. Transcriptional regulation of ascorbic acid during fruit ripening in pepper (*Capsicum annuum*) varieties with low and high antioxidants content. *Plants.* (2019) 8:206. doi: 10.3390/plants8070206

37. Liu Y, Deng X. citrus breeding and genetics in China. *Asian Australas J Plant Sci Biotechnol.* (2007) 1:23–8.

38. Zeng L, Chen D, Huang Q, Huang Q, Lian Y, Cai W, et al. Isolation of a new flavanone from *Daidai* fruit and hypolipidemic activity of total flavonoids extracts. *Nat Prod Res.* (2015) 29:1521–8. doi: 10.1080/14786419.2014.992344

39. Singh B, Singh J, Kaur A, Singh N. Phenolic composition, antioxidant potential and health benefits of citrus peel. *Food Res Int.* (2020) 132:109114.

40. Rice-Evans C, Miller N. Antioxidant activities of flavonoids as bioactive components of food. *Biochem Soc Trans.* (1996) 24:790–5. doi: 10.1042/bst0240790

41. Wang S, Xie W, Yan X. Effects of future climate change on citrus quality and yield in China. *Sustainability.* (2022) 14:9366. doi: 10.3390/su14159366

42. Liu S, Lou Y, Li Y, Zhang J, Li P, Yang B, et al. Review of phytochemical and nutritional characteristics and food applications of *Citrus* L. fruits. *Front Nutr.* (2022) 9:968604. doi: 10.3389/fnut.2022.968604

43. Kim D, Lee S, Park S, Yun S, Gab H, Kim S, et al. Comparative metabolomics analysis of citrus varieties. *Foods.* (2021) 10:2826.

44. Wang S, Tu H, Wan J, Chen W, Liu X, Luo J, et al. Spatio-temporal distribution and natural variation of metabolites in citrus fruits. *Food Chem.* (2016) 199:8–17. doi: 10.1016/j.foodchem.2015.11.113

45. Feng S, Niu L, Suh J, Hung W, Wang Y. Comprehensive metabolomics analysis of mandarins (*Citrus reticulata*) as a tool for variety, rootstock, and grove discrimination. *J Agric Food Chem.* (2018) 66:10317–26. doi: 10.1021/acs.jafc.8b03877

46. Gao Z, Gao W, Zeng S, Li P, Liu E. Chemical structures, bioactivities and molecular mechanisms of citrus polymethoxyflavones. *J Funct Foods.* (2018) 40:498–509. doi: 10.1021/jf400562p

47. Choi S, Hwang J, Ko H, Park J, Kim S. Nobiletin from citrus fruit peel inhibits the DNA-binding activity of NF-κB and ROS production in LPS-activated RAW 264.7 cells. *J Ethnopharmacol.* (2007) 113:149–55. doi: 10.1016/j.jep.2007.05.021

48. Fayek N, El-Shazly A, Abdel-Monem A, Moussa M, Abd-Elwahab S, El-Tanbouly N. Comparative study of the hypocholesterolemic, antidiabetic effects of four agro-waste citrus peels cultivars and their HPLC standardization. *Rev Bras Farmacogn.* (2017) 27:488–94. doi: 10.1016/j.bjp.2017.01.010

49. Gupta S, Rahman M, Sundaram S. *Citrus* fruit as a potential source of phytochemical, antioxidant and pharmacological ingredients. *J Sci Healthc Explor.* (2021) 2581:8473.

50. Gorinstein S, Martín-Belloso O, Park Y, Haruenkit R, Lojek A, Ciž M, et al. Comparison of some biochemical characteristics of different citrus fruits. *Food Chem.* (2001) 74:309–15. doi: 10.1016/S0308-8146(01)00157-1

51. He D, Shan Y, Wu Y, Liu G, Chen B, Yao S. Simultaneous determination of flavanones, hydroxycinnamic acids and alkaloids in citrus fruits by HPLC-DAD–ESI/MS. *Food Chem.* (2011) 127:880–5. doi: 10.1016/j.foodchem.2010.12.109

52. Toh J, Khoo H, Azrina A. Comparison of antioxidant properties of *Pomelo* [*Citrus grandis* (L) Osbeck] varieties. *Int Food Res J.* (2013):20:1661–8.

53. Safdar M, Kausar T, Jabbar S, Mumtaz A, Ahad K, Saddozai A. Extraction and quantification of polyphenols from kinnow (*Citrus reticulate* L.) peel using ultrasound and maceration techniques. *J Food Drug Anal.* (2017) 25:488–500. doi: 10.1016/j.jfda.2016.07.010

54. Al-Juhaimi F, Ghafoor K. Bioactive compounds, antioxidant and physico-chemical properties of juice from lemon, mandarin and orange fruits cultivated in Saudi Arabia. *Pak J Bot.* (2013) 45:1193–6.

55. Fernandez-Lopez J, Zhi N, Aleson-Carbonell L, Pérez-Alvarez J, Kuri V. Antioxidant and antibacterial activities of natural extracts: application in beef meatballs. *Meat Sci.* (2005) 69:371–80. doi: 10.1016/j.meatsci.2004.08.004

56. Jayaprakasha G, Patil B. *In vitro* evaluation of the antioxidant activities in fruit extracts from citron and blood orange. *Food Chem.* (2007) 101:410–8. doi: 10.1016/j.foodchem.2005.12.038

57. Dong X, Hu Y, Li Y, Zhou Z. The maturity degree, phenolic compounds and antioxidant activity of Eureka lemon [*Citrus limon* (L.) burm. f.]: a negative correlation between total phenolic content, antioxidant capacity and soluble solid content. *Sci Hortic.* (2019) 243:281–9.

58. Xi W, Lu J, Qun J, Jiao B. Characterization of phenolic profile and antioxidant capacity of different fruit part from lemon (*Citrus limon* burm.) cultivars. *J Food Sci Technol.* (2017) 54:1108–18. doi: 10.1007/s13197-017-2544-5

59. Dhanavade M, Jalkute C, Ghosh J, Sonawane K. Study antimicrobial activity of lemon (*Citrus lemon* L.) peel extract. *Br J Pharmacol Toxicol.* (2011) 2:119–22.

60. Tanizawa H, Ohkawa Y, Takino Y, Miyase T, Ueno A, Kageyama T, et al. Studies on natural antioxidants in citrus species. I. Determination of antioxidative activities of citrus fruits. *Chem Pharm Bull.* (1992) 40:1940–2. doi: 10.1248/cpb.40.1940

61. Miyake Y. Characteristics of flavonoids in niihime fruit-a new sour citrus fruit. *Food Sci Technol Res.* (2006) 12:186–93.

62. Chen X, Tait A, Kitts D. Flavonoid composition of orange peel and its association with antioxidant and anti-inflammatory activities. *Food Chem.* (2017) 218:15–21. doi: 10.1016/j.foodchem.2016.09.016

63. Nogata Y, Sakamoto K, Shiratsuchi H, Ishii T, Yano M, Ohta H. Flavonoid composition of fruit tissues of citrus species. *Biosci Biotechnol Biochem.* (2006) 70:178–92. doi: 10.1271/bbb.70.178

64. Thieme C, Westphal A, Malarski A, Böhm V. Polyphenols, vitamin C, *in vitro* antioxidant capacity, α-Amylase and COX-2 inhibitory activities of citrus samples from Aceh, Indonesia. *Int J Vitam Nutr Res.* (2019) 89:337–47. doi: 10.1024/0300-9831/a000481

65. Chang S, Tan C, Frankel E, Barrett D. Low-density lipoprotein antioxidant activity of phenolic compounds and polyphenol oxidase activity in selected clingstone peach cultivars. *J Agric Food Chem.* (2000) 48:147–51. doi: 10.1021/jf9904564

66. Wang Y, Qian J, Cao J, Wang D, Liu C, Yang R, et al. Antioxidant capacity, anticancer ability and flavonoids composition of 35 citrus (*Citrus reticulata* Blanco) varieties. *Molecules.* (2017) 22:1114. doi: 10.3390/molecules22071114

67. Arena E, Fallico B, Maccarone E. Evaluation of antioxidant capacity of blood orange juices as influenced by constituents, concentration process and storage. *Food Chem.* (2001) 74:423–7.

68. Manganaris G, Goulas V, Vicente A, Terry L. Berry antioxidants: small fruits providing large benefits. *J Sci Food Agric.* (2014) 94:825–33. doi: 10.1002/jsfa.6432

69. Silva F, O'callagahan Y, O'brien N, Netto F. Antioxidant capacity of flaxseed products: the effect of *in vitro* digestion. *Plant Foods Hum Nutr.* (2013) 68:24–30. doi: 10.1007/s11130-012-0329-6

70. Sir Elkhatim K, Elagib R, Hassan A. Content of phenolic compounds and vitamin C and antioxidant activity in wasted parts of Sudanese *Citrus* fruits. *Food Sci Nutr.* (2018) 6:1214–9. doi: 10.1002/fsn3.660

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 02 February 2023
DOI 10.3389/fnut.2022.1094763

Check for updates

OPEN ACCESS

EDITED BY
Sapna Langyan,
National Bureau of Plant Genetic Resources (ICAR), India

REVIEWED BY
Chirag Maheshwari,
Indian Agricultural Research Institute (ICAR), India
Aalok Shiv,
Indian Institute of Sugarcane Research (ICAR), India
Amar Kant Kushwaha,
Central Institute for Subtropical Horticulture (ICAR), India

*CORRESPONDENCE
Sultan Singh
✉ Sultan.Singh1@icar.gov.in
Tejveer Singh
✉ tejveersinghbhu@gmail.com

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

RECEIVED 10 November 2022
ACCEPTED 13 December 2022
PUBLISHED 02 February 2023

CITATION
Singh S, Singh T, Singh KK, Srivastava MK, Das MM, Mahanta SK, Kumar N, Katiyar R, Ghosh PK and Misra AK (2023) Evaluation of global *Cenchrus* germplasm for key nutritional and silage quality traits. *Front. Nutr.* 9:1094763.
doi: 10.3389/fnut.2022.1094763

COPYRIGHT
© 2023 Singh, Singh, Singh, Srivastava, Das, Mahanta, Kumar, Katiyar, Ghosh and Misra. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Evaluation of global *Cenchrus* germplasm for key nutritional and silage quality traits

Sultan Singh[1]*, Tejveer Singh[2]*, Krishan Kunwar Singh[1], Manoj Kumar Srivastava[2], Madan Mohan Das[1], Sanat Kumar Mahanta[1], Neeraj Kumar[2], Rohit Katiyar[1], Probir Kumar Ghosh[1] and Asim Kumar Misra[1]

[1]Plant Animal Relationship Division, ICAR-Indian Grassland and Fodder Research Institute, Jhansi, India, [2]Crop Improvement Division, ICAR-Indian Grassland and Fodder Research Institute, Jhansi, India

Cenchrus is important genera of grasses inhabiting tropical pastures and the Indian grasslands system. Its forage value is well established to sustain nomadic livestock and wildlife. This study deals with the evaluation of the representative set of global *Cenchrus* germplasm collection with 79 accessions belonging to six species (*C. ciliaris, C. setigerus, C. echinatus, C. myosuroides, C. pennisetiformis*, *and C. biflorus)* at flowering stage. Crude protein (CP), neutral detergent fiber (NDF), acid detergent fiber (ADF), cellulose, and lignin values were in the range of 61.1–136, 640–749, 373–490, 277–375, and 35.6–75.50 g kg^{-1}DM, respectively, while sugar contents varied from 11.6 to 101 mg g^{-1} DM. From the evaluated germplasm, 14 accessions of *C. ciliaris* having >70 mg g^{-1} DM sugar contents were selected and further evaluated for protein, fiber, carbohydrate and protein fractions, palatability indices, *in vitro* CH_4 production, and ensiling traits. Protein contents were lower in EC397323 (61.8) and higher in IG96-96 (91.5), while the NDF, ADF, cellulose, and lignin contents varied between 678–783, 446–528, 331–405, and 39.6–62.0 g kg^{-1}DM, respectively. The carbohydrate and protein fractions of selected accessions differed ($p < 0.05$), and the sugar contents varied ($p < 0.05$) between 74.6 and 89.6 mg^{-1}g DM. Dry matter intake (DMI) and relative feed value (RFV) of accessions varied ($p < 0.05$) and were in the range of 1.53–1.77% and 58.2–73.8 g kg^{-1} DM, respectively. The total digestible nutrients (TDNs), digestible energy (DE), and metabolizable energy (ME) of selected accessions varied between 362–487 g kg^{-1} DM, 6.62–8.90, and 5.42–7.29 Mj kg^{-1} DM, respectively. *In vitro* gas and CH_4 production (24 h) varied ($p < 0.05$) between 73.1 to 146 and 7.72 to 21.5 ml/g, respectively, while the degraded dry matter (g kg^{-1} DM) and CH_4 (ml/g DDM) ranged between 399–579 and 17.4–47.2, respectively. The DM contents at ensiling, silage pH, and lactic acid contents of accessions differed ($p < 0.05$) and ranged between 185–345 g kg^{-1} DM, 5.10–6.05, and 1.39–23.3 g kg^{-1} DM, respectively. Wide genetic diversity existed in germplasm and selected *C. ciliaris* accessions for protein fiber, energy, sugar,

and other nutritional traits. Silage prepared from EC397366, IG96-96, IG96-50, and EC397323 had pH and lactic acid contents acceptable for moderate to good quality silage of tropical range grasses.

KEYWORDS

Cenchrus germplasm, ensiling, energy value, methane, sugar

1. Introduction

Grasses constitute up to 48% of all biomass fed to livestock globally (1), and the habitat of natural pastures, rangelands, forests, community lands, etc. serves as one of the major roughage sources for ruminants across the globe and usually constitutes more than 60% of the diet for small ruminants. Tropical grasses are nutritionally poor than temperate grasses, and other cereal forage crops [oat, maize, sorghum, barley, etc., Minson (2)] and their yield and nutritive value vary with species, growth stage, season, soil nitrogen status, and fertilizer application. Grasses are usually fed as green and hay and hardly conserved as silage primarily due to low dry matter, less water-soluble carbohydrate (WSC) contents, higher buffering capacity, and low energy contents (3, 4), which restrict the fermentation process and the subsequent adoption of tropical grass silage technology (5). Success in ensiling of grasses is governed by their readily available carbohydrate, and it is sufficiently high to promote the lactic acid bacteria to produce lactic acid to reduce pH during fermentation for subsequent preservation (6). According to Haigh (7), fresh grass should contain a minimum of 37.0 g kg^{-1} water-soluble carbohydrates or about 150 g kg^{-1} WSC on a dry weight basis to prepare good quality silage without silage additives. In addition, fodder species and their developmental stage are also important pre-ensiling factors responsible for silage quality (8).

The *Cenchrus* genus of the grass family has many species, which can tolerate a wide range of soil types and moisture conditions found globally including Asia, Africa, Australia, and the United States of America (9). In India, *Cenchrus* is an important component of *Dichanthium-Cenchrus-Lasiurus*-type grassland cover with coverage of >436,000 km^2 (10). *Cenchrus species* mainly *Cenchrus ciliaris* (Buffel grass) and *Cenchrus setigerus* are important pasture grasses in the tropics (11), which are commonly used as a forage grass in India (12). It is drought-tolerant and a well-fertilized *C. ciliaris* crop may yield up to 24 t DM ha^{-1} (13) with a yield range of 2–18 t DM ha^{-1} without fertilizer. At the early flowering stage, hay prepared is of medium quality and rarely made into silage due to lower sugar contents and usually low moisture contents in the semi-arid regions. Efforts have been made to breed its cultivars for improved nutritive value and higher fodder yield, particularly in Australia. In India, a global collection of *Cenchrus* species are maintained at ICAR-Indian Grassland and Fodder Research Institute, and also, few varieties have been developed for higher biomass (14). However, no research efforts have been put to evolve the varieties for silage making (ensiling properties) but the need for Indian tropical grasslands due to the climatic situation, which favor surplus availability of fodder during monsoon months (mid-September–mid-November) and growth dormancy afterward. Keeping this in view, a multidisciplinary project was initiated on the evaluation of *Cenchrus* germplasm for higher sugar contents (>70 mg g^{-1} DM) required to initiate fermentation. So, in the present study, a 79 *Cenchrus* spp. genotypes were evaluated for yield, protein, and cell wall contents including sugar contents, while the sugar-rich (>70 mg g^{-1} DM) selected genotypes were evaluated for various nutritional parameters and ensiling properties.

2. Materials and methods

2.1. Location, germplasm maintenance, and multiplication of sugar-rich accessions

The study was carried out at ICAR-Indian Grassland and Fodder Research Institute, Central Research Farm, Jhansi, India (25°31′ N, 78°32′ E; 237 masl). The experimental site has a prevalence of semi-arid climatic conditions with extreme winter (as low as 2°C) and summer (43–46°C) temperatures. The edaphic/soil parameter consisted of deep, moderately well drained, and brown to dark grayish brown with a fine loamy texture. The optimum dose of fertilizers such as nitrogen (80 kg N ha^{-1}), phosphorus (60 kg P/ha), and farmyard manure (30 t ha^{-1}) was applied at sowing. Seeds of 79 accessions of *Cenchrus* spp. (*Cenchrus ciliaris* 53, *Cenchrus setigerus* 20, *Cenchrus echinatus* 3, *Cenchrus myosuroides* 1, *Cenchrus pennisetiformis* 1, *and Cenchrus biflorus* 1; Supplementary Table 1) representing corsets developed from over 600 global germplasm of *Cenchrus* spp. maintained in the Institute Gene Bank (14). When seedlings that reached the height of around 30 cm were transplanted with three checks (cv. IGFRI727, IGFRI3108, and IG-96-83) in an augmented randomized complete block design (15) during the rainy season. Each accession was transplanted in 1 × 3 m plots with 2 rows of plants/plot. Line-to-line and plant-to-plant

distances were maintained at 50 × 50 cm, with a 1-m distance between two plots. Out of the 79, 14 accessions of *C. ciliaris* with >70 mg g^{-1} DM sugar contents were transplanted in RCBD during the rainy season in three replications. Each accession was planted in 4 × 3 m plots with six rows of plants/plots. Line-to-line and plant-to-plant distances were maintained at 50 × 50 cm with 1-m spacing between two plots.

2.2. Sample collection, processing, and drying

Samples of each accession from *Cenchrus* germplasm and selected sugar-rich accessions of *C. ciliaris* were harvested at the flowering stage in the rainy season from each row for nutritional evaluation. Immediately after harvesting, fresh forage yield was recorded using a digital portable balance. For the dry matter (DM) and dry matter yield (DMY), estimation samples were dried at 100°C for 72 h and at 60°C for 72 h for chemical/biochemical estimations (16). Dried samples were stored in the plastic sample containers (Tarson make) after fine grounding through a 1-mm sieve using a Willey mill for further nutritional and *in vitro* analyses.

2.3. Ensiling of sugar-rich accessions

Samples of IG99-124, IG97-379, IG-97-377, IG97-403, EC397323, IG96-87, EC400605, IG97-378, EC397366, EC397379, CC-14-1, IG96-96, IG96-89, and IG96-50 accessions were harvested in the forenoon (September 2016) and wilted for 2 h. Samples were chaffed (1–1.5 cm) through a manually operated chaffing machine and filled in the plastic containers (25.5 cm long × 13 cm diameter wide 5 kg volume) in triplicate for each accession. The chaffed samples filled were pressed manually with a hand and broad-based wooden rod to exclude as much air as possible, and then, the containers were capped and sealed with adhesive tape for ensiling. After 45 days of ensiling silage, containers were opened and representative samples were analyzed for silage DM, pH, lactic acid, and chemical composition.

2.4. Chemical analyses

2.4.1. Chemical composition

Dry matter (930.15), N (976.05), ether extract (EE, 920.39), and ash (932.05) contents of *Cenchrus* genotypes and selected sugar-rich accessions were determined as per the standard protocol of AOAC (17). The obtained nitrogen values were multiplied by 6.25 to get CP-values. Samples of neutral detergent fiber (NDF), acid detergent fiber (ADF), cellulose, and lignin (sa) were estimated sequentially (18) using the fiber analyzer (Fibra Plus FES 6, Pelican, Chennai, India). Both NDF and ADF were expressed inclusive of residual ash. Heat stable α-amylase and sodium sulfite were not used in NDF determination. Lignin (sa) was determined by the solubilization of cellulose with 72% sulfuric acid in the ADF residue (18). Cellulose was calculated as the difference between ADF and lignin (sa) in the sequential analysis. Hemicellulose was calculated as the difference between NDF and ADF.

2.4.2. Sugar contents

Total sugar contents of germplasm and selected accessions were estimated by the Anthrone method using glucose as standard (19). For this, 100 mg of ground sample (1-mm sieve) was treated with 10 ml of 80% ethanol in the water bath (80°C) for 30 min. The contents were centrifuged at 10,000 rpm for 10 min, the supernatant was collected in a volumetric flask, and the volume was made up to 25 ml. From this, 1 ml was further diluted to 50 ml, from this, 0.25 ml was taken in a tube, and 2.25 ml of distilled water and 5 ml of 0.2% Anthrone reagent were added. The mixture was boiled for 7 min and cooled, and the blue color developed was measured using a UV spectrophotometer (LABINDIA3000) at 630 nm.

2.4.3. Carbohydrate fractions

Cornell Net Carbohydrate and Protein (CNCP) system (20) was used to determine the carbohydrate fractions of sugar-rich *C. ciliaris* accessions. This system further divides the carbohydrate components into four fractions based on their degradation rate; C_A: rapidly degradable sugars; C_{B1}: intermediately degradable starch and pectin; C_{B2}: slowly degradable cell wall; and C_C: unavailable/lignin bound cell wall.

Total carbohydrate (tCHO g kg^{-1} DM) was determined by subtracting CP, EE, and ash contents from 1,000. The difference between NDF and neutral detergent-insoluble protein (NDIP) was used to calculate the structural carbohydrates (SCs), and the difference between tCHO and SC (21) was estimated to calculate the non-structural carbohydrate (NSC). Starch was determined by extracting grass samples in 80% ethyl alcohol to solubilize free sugars, lipids, pigments, and waxes. The residue rich in starch was solubilized with perchloric acid and the extract was treated with anthrone-sulfuric acid to determine glucose calorimetrically using the glucose standard (19).

2.4.4. Protein fractions

The CP fractions of sugar-rich accessions were partitioned into five fractions according to the Cornell Net Carbohydrate and Protein System [CNCPS; (20)] as modified previously (22). These are fraction P_A and non-protein N, which are calculated as the difference between total N and true CP N precipitated with sodium tungstate (0.30 M) and 0.5 M sulfuric acid; fraction P_{B1},

buffer-soluble protein, determined as the difference between true protein and buffer-insoluble protein, estimated with borate-phosphate buffer (pH 6.7–6.8) and freshly prepared 0.10 sodium azide solution. Fraction P_{B2}, neutral detergent-soluble protein, was estimated as the difference in buffer-insoluble protein and ND-insoluble protein, whereas fraction P_{B3}, acid detergent-soluble CP, was estimated as the difference between ND-insoluble protein and acid detergent-insoluble CP. Fraction P_C is assumed to be indigestible.

Neutral detergent-insoluble protein (NDIP), acid detergent-insoluble protein (ADIP), and non-protein nitrogen (NPN) were determined following the standard method (22). For NDIP and ADIP, samples extracted with neutral detergent and acid detergent solutions, respectively, were analyzed as Kjeldahl N × 6.25 using a semi-auto analyzer (Kel Plus Classic-DX Pelican India). For NPN estimation, samples were treated with sodium tungstate (0.30 M) and filtered, and residual nitrogen was determined by the Kjeldahl procedure. Non-protein nitrogen of the sample was calculated by subtracting residual nitrogen from total nitrogen. Soluble protein (SP) was estimated by treating the samples in borate-phosphate buffer, pH 6.7–6.8, consisting of monosodium phosphate ($Na_2PO_4.H_2O$) 12.2 g L^{-1}, sodium tetra borate ($Na_2B_4O_7.10H_2O$) 8.91 g L^{-1}, and tertiary butyl alcohol 100 mL $L^{-1}$and freshly prepared 10% sodium azide solution (23). The N estimated in the residue gives the insoluble protein fraction. The SP was calculated by subtracting the insoluble protein from the total CP.

2.4.5. Gross energy estimation and calculations for DDM, DMI, RFV, and energy

Cenchrus germplasm and selected sugar-rich accessions, dry matter intake (DMI), digestible dry matter (DDM), relative feed value (RFV), total digestible nutrients (TDN), and net energy for different animal functions, i.e., lactation (NE_L), gain (NE_G), and maintenance (NE_M), were calculated using the equations [DMI = 120/NDF; DDM = 88.9–0.779*ADF; RFV = (DDM*DMI)*0.775; TDN = 104.97–(1.302*ADF); NE_L =(TDN*0.0245)−0.012; NE_G = (TDN*0.029)−1.01; NE_M = (TDN*0.029)−0.29] of Undersander et al. (24). Digestible energy (DE, KJ g^{-1} DM; DE = TDN*0.04409) and metabolizable energy (ME, KJ g^{-1} DM) values were calculated using the equations of Fonnesbeck et al. (25) and Khalil et al. (26), respectively. Metabolizable energy was calculated as DE × 0.821.

2.5. *In vitro* incubation

2.5.1. Donor animals and inoculum preparation

Overall, four adult male *Jalauni* sheep with a mean body weight of 38.7 ± 0.473 kg were used as inoculum donors. These animals were maintained on a sole berseem hay diet and had free access to clean drinking water. Rumen liquor was collected in a pre-warmed thermos from each animal before feeding using a perforated tube from the stomach with the help of a vacuum pressure pump. Rumen liquor collected from each animal was filtered through four layers of muslin cloth and mixed well to have the composite sample, kept at 39°C in a water bath, and gassed with CO_2 till used for mixing with incubating buffer media.

In vitro gas production was estimated as per the pressure transducer technique (27). The incubation medium was formulated by sequential mixing of buffer solution (NH_4HCO_3 and $NaHCO_3$), macro-mineral solution, micro-mineral solution, and resazurin solution (28). Samples (1.0 g) of air-dry *Cenchrus* sugar-rich accessions were weighed into three serum bottles (150 ml of capacity). In total, three serum bottles without substrate were used as blank cultures. Sample and control serum bottles were gassed briefly with CO_2 before adding 65 ml of medium. Bottles were continuously fluxed with CO_2, and then, 3 ml of reducing solution was added to each bottle. The gassing of bottles with CO_2 continued till the pink color turned colorless. Before inoculation, the gas pressure transducer was used to adjust the head-space gas pressure in each bottle (to adjust the zero reading on the LED display). Serum bottles were inoculated with 8 ml of ruminal fluid inoculum using a 10-ml syringe. Inoculated bottles were sealed and incubated at 39°C. Samples were incubated in triplicates and gas production (ml) was measured at 24 h of incubation. The whole process was repeated on a different day.

2.5.2. Methane measurements

At 24 h of incubation, methane in total gas was measured from three bottles incubated for each of the *Cenchrus* accession by gas chromatography (Nucon 5765 Microprocessor controlled gas chromatograph, Okhla, New Delhi, India) equipped with a stainless-steel column packed with Porapak-Q and a Flame Ionization Detector. Gas (1 ml) was sampled from gas produced using a Hamilton syringe and injected manually (pull and push methods of sample injection) into a gas chromatograph calibrated with standard methane and CO_2. Methane was also measured from three serum bottles used as blanks for the correction of methane produced from the rumen inoculum. Methane measured was related to total gas to estimate its concentration (29). Short-chain fatty acids (SCFA) were calculated using 24 h gas production as described by Getachew et al. (30). Microbial mass (MBM) and partitioning factor (PF) were calculated as described by Blümmel et al. (31).

2.6. Silage analysis

For DM estimation, 100 g of the fresh silage sample was dried in a hot air oven at 60°C till the constant weight is achieved

and then corrected for DM using the equation of Kaiser and Kerr (32) as estimated true DM (%) = 4.686 + (0.89 × oven DM %). For silage pH and lactic acid estimation, a 20 g of fresh silage sample was put in a beaker, and to this, 100 ml of tepid water was added. Beaker was kept in a water bath shaker (30°C) for 30 min, and contents were agitated manually and filtered through filter paper. The filtrate was mixed well, a portion of it was used to measure pH using a digital pH meter (Systronic 360), and the remaining filtrate was used for lactic acid estimation as described by Barker and Summerson (33). For this, 1 ml of extract was added to the tubes and 0.05 ml of 4% CuSO4 and 6 ml of concentrated H_2SO_4 were added drop by drop with continuous shaking. Tubes were kept in a boiling water bath for 5 min and cooled at room temperature, and then, 0.1 ml of P-hydroxyphenyl reagent was added drop by drop and incubated in a shaker water bath at 30°C for 30 min. The blue color developed was measured at 560 nm using UV-spectrophotometer (LABINDIA3000).

2.7. Statistical analysis

To describe the variability among the accessions, univariate statistics including means and ranges were used, which were obtained for each trait based on the accessions. Data on dry matter yield were subjected to statistical analysis using the descriptive statistics on adjusted means estimated by the R package for augmented design (34). Data were subjected to a one-way analysis of variance of SPSS 17.0 to test the differences between *Cenchrus* accessions for chemical composition, sugar contents, carbohydrate and protein fractions, energy values, digestibility and *in vitro* gas and methane production, and silage quality (pH, lactic acid, and DM contents). Variable means were compared for significance ($p < 0.05$) level using Duncan's multiple-range test (35). Euclidean distance as a measure of dissimilarity and incremental sums of squares as a grouping strategy was utilized for clustering the accessions based on their morphological traits using the "cluster" package of SAS statistical software (36). Dendrograms were constructed based on the fusion level to examine the similarities in the pattern of performance among the accessions.

3. Results

3.1. Biomass and nutritional variability in *Cenchrus* spp. germplasm

The variance and range showed that sufficient variability exits in germplasm evaluated for DMY, chemical composition (CP, NDF, ADF, cellulose, and lignin), sugar contents, and other nutritional traits (Table 1), and the values for these traits for individual accession are given in Supplementary Table 1. The DMY of evaluated *Cenchrus* species germplasm varied from 1.85 to 34.27 t/ha; CP, NDF, ADF, cellulose, and lignin contents varied between 61.1–136, 640–750, 370–510, 250–400, and 31.0–97.0 g kg^{-1} DM, with their mean values of 88.0, 694, 426, 321, and 53.2 g kg^{-1} DM, respectively. Soluble sugar contents of *Cenchrus* germplasm varied widely in the range of 11–101mg/g with a mean value of 57.07 mg/g. Mean values of total digestible nutrients (TDNs), digestible energy (DE), metabolizable energy (ME), net energy for lactation, net energy for maintenance, and net energy for the gain of the *Cenchrus* germplasm were 495 g kg^{-1} DM, 9.02, 7.41, 4.51, 5.42, and 1.82 Mj/Kg, respectively. The mean values of dry matter intake (DMI), digestible dry matter (DDM), and relative feed values (RFVs) for the *Cenchrus* germplasm were 1.73%, 553 g kg^{-1} DM, and 72.75%, respectively. The cluster analysis placed 79 accessions into five clusters ($R^2 = 0.4$). Clusters one and two included single accession each, cluster three included five accessions of *C. ciliaris* and *C. setigerus*, cluster four contain 19 accessions of *C. ciliaris*, *C. setigerus*, *C. echinatus*, *C. myosuroides,* and *C. pennisetiformis*, and cluster five included 53 accessions of *C. ciliaris*, *C. setigerus*, *C. echinatus*, and *C. biflorus* (Figure 1).

3.2. Sugar contents and chemical composition of *C. ciliaris* accessions

Sugar contents of 14 sugar-rich *C. ciliaris* accessions, *viz.* IG99-124, IG97-379, IG-97-377, IG97-403, EC397323, IG96-87, EC400605, IG97-378, EC397366, EC397379, IG96-96, IG96-89, IG96-50, and CC-14-1 were more than 70 mg g^{-1} DM required for ensiling (Table 2). Sugar contents of accessions differed ($p < 0.05$) and varied between 74.6 (IG99-127) and 89.6 mg g^{-1} DM (EC397323). Accessions of CP, NDF, ADF, cellulose, and lignin contents varied ($p < 0.05$), and their mean values were 76.4, 737, 478, 351, and 49.1 g kg^{-1} DM, respectively (Table 2). The OM and EE contents of sugar-rich accessions varied ($p < 0.05$) with mean values of 856 and 20.5 g kg^{-1} DM, respectively.

3.3. Protein and carbohydrate fractions of *C. ciliaris* accessions

Total carbohydrate (tCHO), non-structural carbohydrates (NSCs), and structural carbohydrates (SCs) of *C. ciliaris* accessions differed ($p < 0.05$) from 730 to 795, 27.1 to 67.9, and 662 to 765 g kg^{-1} DM, respectively (Table 3). Similarly, the carbohydrate fractions, namely, C_A C_{B1}, C_{B2}, and C_c varied ($p < 0.05$) across the accessions, and their mean values were 46.5, 56.1, 742, and 155 g kg^{-1} tCHO, respectively. *C. ciliaris* accessions protein fractions, *viz.* P_A, P_{B1}, P_{B2}, P_{B3}, and

TABLE 1 Mean performance of *Cenchrus* species for chemical components and nutritional quality traits.

Species[†]	*C. ciliaris*			*C. setigerus*			*C. echinatus*			*C. biflorus*	*C. pennisetiformis*	*C. myosuroides*
Traits[1]	Mean ± SE	Variance	Range	Mean ± SE	Variance	Range	Mean ± SE	Variance	Range	Mean	Mean	Mean
DMY	12.28±0.97	50.02	4.02–34.27	7.38±0.68	9.15	2.48–12.27	4.14±0.97	2.81	3.15–6.08	7.80	4.20	1.85
CP	88.5±2.0	2.20	61.1–136	86.3±1.5	0.48	75.0–103	93.1±10.1	3.08	73.2–106	105	80.3	75.0
NDF	695±2.9	4.43	640–749	692±4.6	4.27	652–728	712±15.1	6.85	689–740	689	671	665
ADF	427±2.9	4.44	373–490	426±4.2	3.49	390–474	424±19.3	11.22	386–451	446	411	388
Cellulose	326±2.7	3.97	277–375	309±2.8	1.58	285–334	305±9.2	2.54	287–319	341	309	304
Lignin	51.8±0.9	0.46	35.6–70.7	56.5±1.6	0.49	45.6–75.5	61.1±4.7	0.68	53.6–69.9	55.8	52.2	37.8
Sugar	60.36±2.92	452	11.65–101.47	53.84±2.48	123	33.59–72.84	27.74±8.58	221	10.69–37.92	58.3	56.9	62.0
DMI	1.73±0.01	0.00	1.60–1.85	1.74±0.01	0.00	1.65–1.84	1.69±0.04	0.00	1.62–1.74	1.74	1.79	1.8
DDM	552±2.2	2.65	502–595	552±3.3	2.18	515–581	554±15.2	6.96	533–584	537	564	583
RFV	72.57±0.47	11.67	64.21–81.04	72.77±0.79	12.49	67.25–82.16	70.71±3.46	35.97	64.55–76.53	69.58	76.7	80.79
TDN	493±3.7	7.23	411–564	494±5.4	5.91	432–542	498±25.2	19.04	462–547	469	514	544
DE	9.05±0.07	0.24	7.55–10.34	9.07±0.10	0.20	7.93–9.94	9.13±0.46	0.64	8.48–10.02	8.6	9.43	9.99
ME	7.43±0.06	0.16	6.20–8.49	7.45±0.08	0.13	6.51–8.16	7.50±0.38	0.43	6.96–8.23	7.06	7.74	8.2
NE_L	4.53±0.04	0.07	3.70–5.25	4.54±0.06	0.06	3.91–5.02	4.57±0.26	0.20	4.21–5.07	4.28	4.74	5.05
NE_M	5.41±0.04	0.11	4.43–6.27	5.43±0.07	0.09	4.68–6.00	5.47±0.31	0.28	5.04–6.06	5.12	5.66	6.03
NE_G	1.80±0.05	0.13	0.58–2.67	1.84±0.06	0.07	1.18–2.35	1.91±0.31	0.29	1.52–2.52	1.62	2.07	2.38

[†] *Cenchrus ciliaris, Cenchrus setigerus, Cenchrus echinatus, Cenchrus myosuroides, Cenchrus pennisetiformis, and Cenchrus biflorus.*

[1] DMY, Dry matter yield t/ha; CP, Crude protein g kg^{-1} DM; NDF, Neutral detergent fibre g kg^{-1} DM; ADF, Acid detergent fibre g kg^{-1} DM; Cellulose, g kg^{-1} DM; Lignin, g kg^{-1} DM; Sugar = mg g^{-1} DM; DMI, Dry matter intake%; DDM, Digestible dry mater g kg^{-1} DM; RFV, Relative feed value %; TDN, Total digestible nutrients g kg^{-1} DM; DE, Digestible energy Mj kg^{-1} DM; ME, Metabolizable energy Mj kg^{-1} DM; $NE_{L,}$ Net energy for lactation Mj kg^{-1} DM; $NE_{G,}$ Net energy for growth/gain Mj kg^{-1} DM; $NE_{M,}$ Net energy for maintenance Mj kg^{-1} DM.

10.3389/fnut.2022.1094763

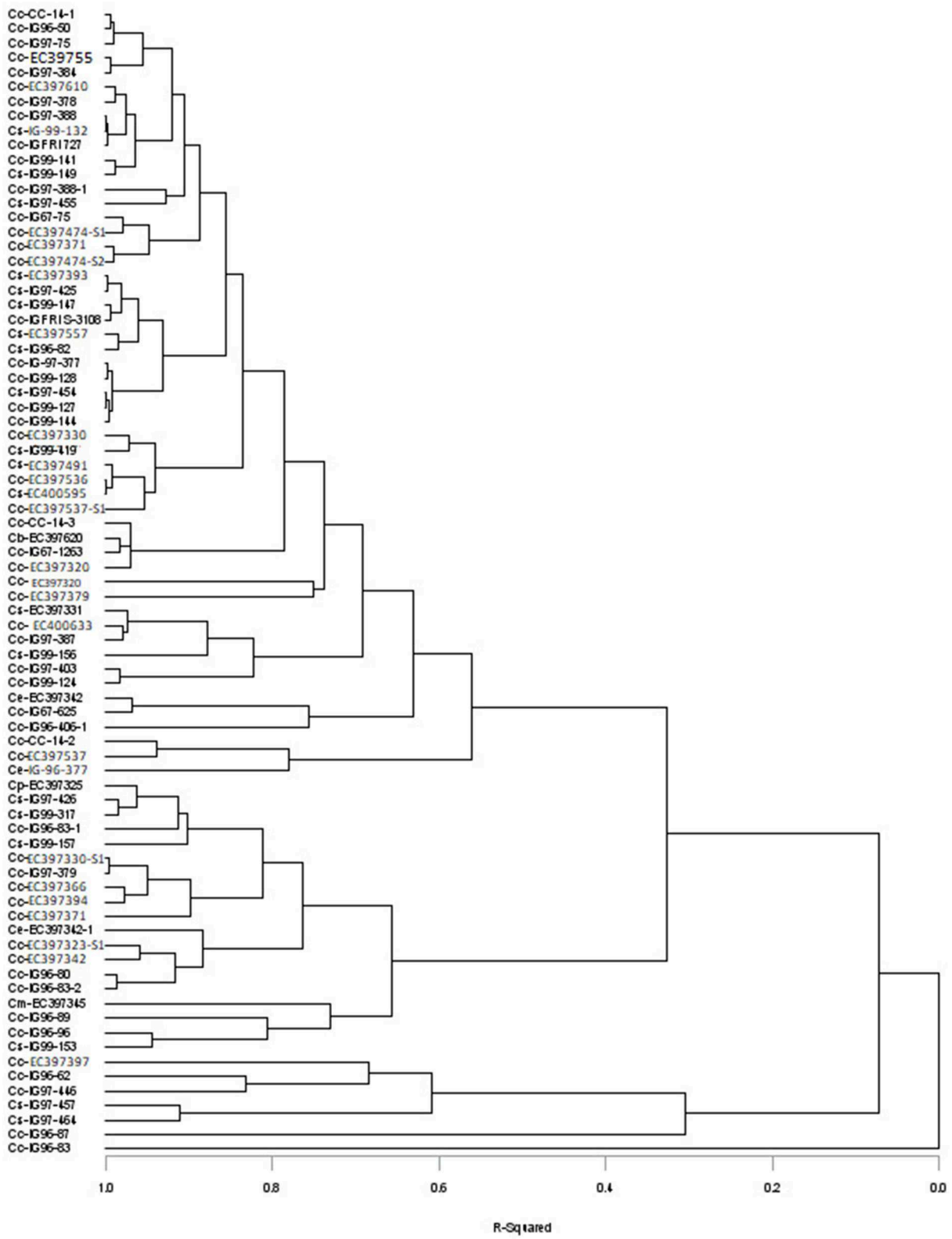


FIGURE 1
Distribution of different *Cenchrus* species accessions into different clusters based on nutritional and biochemical traits (Cc, *Cenchrus ciliaris*; Cs, *Cenchrus setigerus*; Ce, *Cenchrus echinatus*; Cm, *Cenchrus myosuroides*; Cp, *Cenchrus pennisetiformis*; Cb, *Cenchrus biflorus*).

TABLE 2 Sugar contents (mg g^{-1}DM) and chemical composition (g kg^{-1} DM) of sugar-rich *Cenchrus ciliaris* genotypes.

Accessions	Sugar	OM	CP	EE	NDF	ADF	Cellulose	Lignin	Hemi cellulose
IG99-124	80.1^{abc}	849^{de}	68.7^{bc}	20.2^{cd}	735^{def}	432^{a}	326^{a}	39.8^{a}	303^{h}
IG97-379	78.6^{a}	865^{g}	77.8^{de}	21.3^{de}	722^{bc}	507^{ef}	343^{cdef}	60.0^{d}	215^{ab}
IG-97-377	76.0^{a}	881^{i}	72.4^{cd}	20.9^{cde}	738^{ef}	4724^{c}	348^{efg}	39.6^{a}	266^{f}
IG97-403	74.6^{a}	875^{h}	73.5^{cd}	16.4^{a}	777^{h}	516^{f}	355^{g}	51.7^{c}	260^{def}
EC397323	79.1^{ab}	866^{g}	61.8^{a}	17.3^{ab}	763^{g}	455^{b}	339^{bcde}	46.8^{abc}	308^{h}
IG96-87	88.9^{bc}	841^{b}	75.9^{de}	21.1^{cde}	717^{b}	471^{c}	342^{cdef}	42.1^{ab}	245^{cd}
EC400605	81.3^{abc}	851^{de}	83.9^{f}	22.7^{e}	733^{cde}	469^{c}	353^{fg}	62.0^{d}	264^{ef}
IG97-378	81.3^{abc}	853^{e}	77.5^{de}	21.0^{cde}	746^{f}	488^{d}	346^{defg}	50.9^{c}	258^{def}
EC397366	81.9^{abc}	847^{cd}	76.8^{de}	22.0^{de}	735^{def}	528^{g}	395^{i}	48.8^{bc}	207^{a}
EC397379	78.8^{ab}	882^{i}	65.2^{ab}	21.6^{de}	783^{h}	505^{e}	405^{j}	48.5^{bc}	281^{g}
IG96-96	82.5^{abc}	840^{b}	91.5^{g}	18.9^{bc}	678^{a}	451^{b}	334^{abcd}	59.6^{d}	227^{b}
IG96-89	82.5^{abc}	860^{f}	85.5^{f}	22.3^{de}	740^{ef}	446^{b}	331^{abc}	51.3^{c}	294^{gh}
IG96-50	89.6^{c}	842^{bc}	82.1^{ef}	20.7^{cde}	726^{bcd}	483^{d}	372^{h}	46.2^{abc}	243^{c}
CC-14-1	81.9^{abc}	832^{a}	76.6^{de}	21.2^{de}	724^{bcd}	473^{c}	328^{ab}	40.4^{ab}	250^{cde}
Mean	81.2^{abc}	856	76.4	20.5	737	478	351	49.1	258
SEM	0.85	2.44	1.27	0.32	4.08	4.28	3.71	1.27	4.75
P-value	0.075	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

CP, Crude protein; OM, Organic matter; EE, Ether extract; NDF, Neutral detergent fiber; ADF, Acid detergent fiber. Superscripts a–h within columns differ significantly between rows at ($P < 0.05$).

TABLE 3 Carbohydrate fractions of sugar-rich *Cenchrus ciliaris* genotypes.

Accessions	tCHO	NSC	SC	C_A	C_{B1}	C_{B2}	C_C
IG99-124	760^{g}	46.8^{d}	713^{cd}	68.8^{cd}	32.5^{a}	773^{c}	126^{ab}
IG97-379	766^{h}	61.0^{e}	706^{bc}	86.8^{d}	50.0^{b}	675^{b}	188^{ef}
IG-97-377	788^{i}	64.6^{e}	723^{de}	68.3^{cd}	46.1^{b}	765^{c}	120^{a}
IG97-403	785^{i}	27.1^{a}	758^{g}	29.1^{ab}	47.2^{b}	766^{c}	158^{cd}
EC397323	787^{i}	42.2^{bcd}	745^{f}	46.9^{bc}	45.2^{ab}	765^{c}	143^{abcd}
IG96-87	742^{cd}	43.2^{de}	699^{b}	45.3^{b}	60.0^{bcd}	759^{c}	136^{abcd}
EC400605	744^{cd}	24.7^{a}	719^{de}	33.6^{ab}	67.4^{de}	699^{b}	200^{f}
IG97-378	755^{fg}	30.5^{abc}	725^{e}	17.7^{a}	75.0^{e}	745^{c}	162^{de}
EC397366	748^{de}	29.4^{ab}	719^{de}	15.3^{a}	76.7^{e}	751^{c}	156^{cd}
EC397379	795^{j}	30.3^{abc}	765^{g}	17.8^{a}	58.0^{bcd}	778^{c}	146^{abcd}
IG96-96	730^{a}	67.9^{e}	662^{a}	110^{e}	55.4^{bcd}	638^{a}	196^{f}
IG96-89	752^{ef}	27.8^{a}	724^{de}	25.8^{ab}	65.1^{cde}	745^{c}	1647^{de}
IG96-50	739^{bc}	33.8^{abc}	705^{bc}	42.6^{b}	54.8^{bcd}	752^{c}	150^{bcd}
CC-14-1	735^{ab}	33.0^{abc}	702^{b}	43.1^{b}	52.0^{bc}	773^{c}	132^{abc}
Mean	759	40.2	719	46.5	56.1	742	155
SEM	3.32	2.38	3.99	4.53	2.07	6.60	4.20
P-value	<0.0001	<0.0001	0.051	<0.0001	<0.0001	<0.0001	<0.0001

tCHO, Total carbohydrates g kg^{-1} DM; NSC, Non-structural carbohydrates g kg^{-1} DM; SC, Structural carbohydrates g kg^{-1} DM; C_A, Rapidly degradable sugars g kg^{-1} tCHO; C_{B1}, Intermediately degradable starch and pectins g kg^{-1} tCHO; C_{B2}, Slowly degradable cell wall g kg^{-1} tCHO; C_C, Unavailable/lignin-bound cell wall g kg^{-1} tCHO. Superscripts a–j within columns differ significantly between rows at ($P < 0.05$).

TABLE 4 Protein fractions (g kg^{-1} CP) of sugar-rich *Cenchrus ciliaris* accessions.

Genotype	P_A	P_{B1}	P_{B2}	P_{B3}	P_C
IG99-124	234^{abc}	362^{bcde}	86.5^{a}	222	95.4^{ab}
IG97-379	193^{ab}	450^{efg}	139^{cde}	92.7^{bc}	124^{cde}
IG-97-377	290^{bc}	289^{b}	206^{f}	108^{bcd}	107^{bc}
IG97-403	236^{abc}	387^{cdef}	137^{bcde}	143^{de}	97.8^{ab}
EC397323	314^{c}	312^{bc}	85.8^{a}	136^{cde}	152^{f}
IG96-87	254	396^{def}	120^{abcd}	96.4^{bcd}	132^{def}
EC400605	227^{abc}	455^{fg}	165^{ef}	44.3^{a}	109^{bc}
IG97-378	215^{abc}	408^{defg}	100^{abcd}	171^{ef}	106^{bc}
EC397366	306^{bc}	325^{bcd}	157^{de}	77.0^{ab}	134^{def}
EC397379	204^{abc}	414^{defg}	96.0^{abc}	140^{cde}	146^{ef}
IG96-96	238^{abc}	410^{defg}	178^{ef}	92.8^{bc}	81.6^{a}
IG96-89	283^{bc}	414^{defg}	115^{abcd}	71.0^{ab}	11.71^{bcd}
IG96-50	163^{a}	492^{g}	101^{abcd}	144^{de}	99.9^{ab}
CC-14-1	425^{d}	17.2^{a}	93.0^{ab}	194^{fg}	116^{bcd}
Mean	256	377	127	124	115
SEM	12.4	13.7	6.42	8.18	3.41
P-value	0.002	<0.0001		<0.0001	<0.0001

P_A, Non-protein nitrogen; P_{B1}, Buffer-soluble protein; P_{B2}, Neutral detergent-soluble protein; P_{B3}, Acid detergent-soluble protein; P_C, Indigestible protein. Superscripts a–g within columns differ significantly between rows at ($P < 0.05$).

P_C differed ($p < 0.05$) and ranged between 163–425, 172–492, 85.8–206, 44.3–194, and 81.6–152 g kg^{-1} CP, respectively (Table 4).

3.4. Energy contents of *C. ciliaris* accessions

Energy contents of sugar-rich *C. ciliaris* accessions in terms of TDN, DE, and ME differed ($p < 0.05$), and their mean values were 427 g kg^{-1} DM, 7.78 Mj kg^{-1} DM, and 6.42 Mj kg^{-1} DM, respectively (Table 5). Accessions net energy efficiency for different animal functions, *viz.* maintenance (NE_M), lactation (NE_L), and growth ($NE_{G)}$ varied ($p < 0.05$) from 3.81 to 5.30, 3.19 to 4.43, and 0.16 to 1.66 Mj kg^{-1} DM, respectively.

3.5. Palatability attributes of *Cenchrus* accessions

The DMI, DDM, and RFV for sugar-rich accessions also differed ($p < 0.05$) and their mean values were 1.63%, 516 g kg^{-1} DM, and 65.28%, respectively (Table 6).

3.6. Gas and methane production from sugar-rich *Cenchrus* accessions

In vitro gas and methane production from *C. ciliaris* accessions differed ($p < 0.05$) with the mean values of 108 ml g^{-1} DM and 14.8 ml g^{-1} DM, respectively (Table 7). Methane production ml g^{-1} DDM and DDM (g kg^{-1} DM) varied ($P < 0.05$) between 17.4–47.2 and 399–579, respectively, across the accessions. Gas fermentation parameters, namely, partition factor (PF), short-chain fatty acid (SCFA), microbial mass (MBM), and efficiency for microbial mass production (EMBM) varied ($p < 0.05$) across the accessions. The values of PF and EMBM were highest and SCFA were lowest for accession G96-50.

3.7. Silage composition

Silage pH and lactic acid contents for evaluated *C. ciliaris* accessions differed ($p < 0.05$) and ranged between 5.11 (EC397366) to 6.07 (EC397379) and 3.71 (IG97-403) to 23.7 g kg-^{1} DM (EC397366), respectively (Table 8).

4. Discussion

4.1. Chemical composition

The chemical composition of feed/fodder is one of the important determinants of its nutritive value. *C. ciliaris* germplasm and sugar rich accessions had CP more than 70.0g kg^{-1} DM required for sustained rumen microbial activity (37). CP content is a measure of nutritional quality (38), our germplasm and sugar rich accessions had CP similar to 80.0 g kg^{-1} DM which is considered adequate for the maintenance of beef cattle (39).

Information on the nutritive value of *C. echinatus*, *C myosuroides,* and *C. pennisetiformis* is not available; however, the CP, NDF, ADF, cellulose, and lignin along with energy values have been reported for *C. biflorus* and *C. setigerus* (40, 41), and our values are more or less within the range of their reported values. The OM, CP, NDF, ADF, cellulose, and lignin of five new genotypes of *CC* differed ($p < 0.05$) and were in the range of 881–904, 80–96, 687–738, 485–519, 367–432, and 34–60 g kg^{-1} DM, and the mean values of 893, 87, 713, 492, 400, and 43 g kg^{-1} DM (42) were more or less similar to the values recorded for the accessions evaluated in the present study. In the study, mean values of OM, CP, NDF, cellulose, and lignin contents of 78 new genotypes of *CC* evaluated in Mexico (43) were 861, 82, 734, 413, and 31 g kg^{-1} DM, respectively. Melesse et al. (44) reported that C ciliaris grass at pre flowering growth had CP, EE, NDF, ADF, cellulose, and lignin contents of 82, 14.5, 601, 373, 342, and 26.7 g kg^{-1} DM, respectively. Ashraf et al. (45) reported higher

TABLE 5 Energy contents of sugar-rich *Cenchrus ciliaris* accessions.

Accessions	TDN	DE	ME	NE_L	NE_M	NE_G
IG99-124	487^{g}	8.90^{g}	7.29^{h}	4.43^{h}	5.30^{g}	1.66^{g}
IG97-379	389^{bc}	7.12^{c}	5.84^{bc}	3.44^{bc}	4.14^{bc}	0.50^{c}
IG-97-377	434^{e}	7.95^{e}	6.50^{ef}	3.89^{ef}	4.68^{e}	1.04^{e}
IG97-403	377^{b}	6.87^{b}	5.67^{b}	3.31^{b}	3.97^{b}	0.33^{b}
EC397323	457^{f}	8.32^{f}	6.83^{g}	4.14^{g}	4.97^{f}	1.32^{f}
IG96-87	436^{e}	7.95^{e}	6.54^{f}	3.93^{f}	4.72^{e}	1.08^{e}
EC400605	439^{e}	8.03^{e}	6.58^{f}	3.97^{f}	4.72^{e}	1.08^{e}
IG97-378	415^{d}	7.58^{d}	6.21^{d}	3.73^{d}	4.43^{d}	0.79^{d}
EC397366	362^{a}	6.62^{a}	5.42^{a}	3.19^{a}	3.81^{a}	0.17^{a}
EC397379	395^{c}	7.20^{c}	5.92^{c}	3.52^{c}	4.22^{c}	0.58^{c}
IG96-96	463^{f}	8.45^{f}	6.91^{g}	4.18^{g}	5.01^{f}	1.37^{f}
IG96-89	469^{f}	8.57^{f}	7.04^{g}	4.26^{g}	5.09^{f}	1.45^{f}
IG96-50	420^{d}	7.66^{d}	6.29^{de}	3.77^{de}	4.51^{d}	0.87^{d}
CC-14-1	433^{e}	7.91^{e}	6.50^{ef}	3.89^{ef}	4.68^{e}	1.04^{e}
Mean	427	7.78	6.42	3.85	4.60	0.95
SEM	5.56	0.025	0.020	0.014	0.030	0.016
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

TDN, Total digestible nutrients g kg^{-1} DM; DE, Digestible energy Mj Kg^{-1} DM; ME, Metabolizable energy Mj Kg^{-1} DM; NE_L, Net energy for lactation Mj Kg^{-1} DM; NE_G, Net energy for growth/gain Mj Kg^{-1} DM; NE_M, Net energy for maintenance Mj Kg^{-1} DM. Superscripts a–h within columns differ significantly between rows at ($P < 0.05$).

EE (27.0–53.0) and protein contents (132–175 g kg^{-1} DM) of 10 *C. ciliaris* accessions from the Cholistan desert of Pakistan than our EE and CP contents. Saini et al. (46) evaluated six cultivars/species of *CC* for 2 years (2003–2004) and found that CP contents ranged from 94.1 to 157 g kg^{-1} DM in 2003 and 37.2 to 101 g kg^{-1} DM in 2004 during 1st and 2nd cut. Cenchrus ciliaris from light and heavy grazed rangeland (Gemeda and Hassan (47) had mean value of 873, 41.0, 19.5, 682, 418, 360, and 55.0 g kg^{-1} DM for OM, CP, EE, NDF, ADF, cellulose, and lignin, respectively, lies within the range of our values except for CP, which was lower than our values. Bezabih et al. (48) reported that *C. ciliaris* collected from six transects of grazing areas of semi-arid savanna grassland in the Mid Rift Valley of Ethiopia had mean values of 889, 563, and 96.0 g kg?1 DM for OM, NDF and CP, respectively. Coelho et al. (49) reported that the mean values of CP, NDF, ADF, and lignin were 98.0, 685, 337, and 26.0 and 111, 688, 349, and 37.0 g kg^{-1} DM for *C. ciliaris* harvested four times at 60-90 days growth under stockpiled and grazing conditions. CP and ash contents of 11 ecotypes of *C. ciliaris* grass grown in the semi-arid lands of Kenya ranged between 66.4–109 and 112–152 g kg^{-1}1 DM, respectively (50). Jonathan et al. (51) reported that C. ciliaris grass hay fed to sheep had CP, NDF, ADF, OM, and EE of 46.0, 725, 542, 898, and 86.5 g kg^{-1}1 DM, respectively.

4.2. Sugar contents

The sugar content is a measure of rapidly fermentable energy available from a forage/feed and plays an important role in ruminant nutrition as required for both efficient rumen microbial fermentation and lactic acid production during the ensiling process. Sugar contents are of significant importance for ensiling (52) as it is the main source of nutrients for microbes to produce lactic acid and a sugar content level of 80 mg g^{-1} DM, which is desirable. *C. ciliaris* germplasm showed ($p < 0.05$) that the differences in sugar contents were lowest in IG67-625 (11.0 mg g^{-1} DM) and highest in EC397323 (101 mg g^{-1}). The *CC* germplasm mean soluble sugar contents (57.1 mg g^{-1} DM) were lower than the desired level while sugar-rich accession mean sugar contents were more than 70 mg/g DM required for ensiling. Low levels of sugar and water-soluble carbohydrates in tropical grasses limit the fermentative capacity and result in low-quality silage (53). Aminah et al. (54) reported that the sugar content (WSC) of six types of tropical grass ranged between 12.6 and 98.8 mg g^{-1} DM, which are in agreement with the evaluated germplasm and selected accession sugar contents. Water-soluble carbohydrate contents of *Kikuyu grass, Seteria, Rhodes, Signal, Napier, Guinea, and Paspalum grass* were 45, 48, 30, 86, 99, 30, and 31 mg g^{-1} DM, respectively (55).

TABLE 6 Palatability attributes of sugar-rich *Cenchrus ciliaris* accessions.

Accessions	DMI	DDM	RFV
IG99-124	1.63^{cde}	552^{g}	69.9^{h}
IG97-379	1.66^{fg}	494^{bc}	63.6^{c}
IG-97-377	1.62^{cd}	521^{e}	65.6^{de}
IG97-403	1.54^{a}	486^{b}	58.2^{a}
EC397323	1.57^{b}	534^{f}	65.1^{d}
IG96-87	1.67^{g}	522^{e}	67.7^{fg}
EC400605	1.64^{efg}	524^{e}	66.4^{def}
IG97-378	1.61^{c}	509^{d}	63.4^{c}
EC397366	1.63^{cde}	477^{a}	60.4^{b}
EC397379	1.53^{a}	497^{c}	59.0^{a}
IG96-96	1.77^{h}	538^{f}	73.8^{i}
IG96-89	1.62^{cd}	541^{f}	68.0^{g}
IG96-50	1.65^{efg}	512^{d}	65.7^{de}
CC-14-1	1.66^{fg}	520^{e}	66.8^{efg}
Mean	1.63	516	65.3
SEM	0.006	3.32	0.641
P-values	<0.0001	<0.0001	<0.0001

DMI, Dry matter intake %; DDM, Digestible dry mater g kg^{-1} DM; RFV, Relative feed value%. Superscripts a–i within columns differ significantly between rows at ($P < 0.05$).

4.3. Carbohydrate and protein fractions

The mean tCHO of seven types of tropical grass at 56 days of cutting age varied from 730 to 836 g kg^{-1} DM (56), and our values of sugar-rich accessions (730–795 g kg^{-1} DM) are within this range. Brandstetter et al. (57) reported that carbohydrate fractions C_{A+B1}, C_{B2}, and C_C of Jiggs Bermuda grass in different seasons (fall, winter, spring, and summer) varied between 240–376, 539–650, and 84.9–128 g kg^{-1} tCHO, respectively. Sa et al. (58) reported that carbohydrate fractions C_{A+B1}, C_{B2}, and Cc of *Cynodon dactylon, Brachiaria brizantha,* and *Panicum maximum* grasses at 28, 35, and 54 days of cutting age ranged between 165–255, 346–449, and 110–263 g kg^{-1} tCHO. These workers further reported that tCHO and NFC contents of *Cynodon dactylon, Brachiaria brizantha,* and *Panicum maximum* grasses ranged between 728–827 and 20–90 g kg^{-1} DM, respectively. The NSCs of 15 types of tropical grass from 87 to 223 g kg^{-1} DM (59) were higher than our values (24.7–67.9 NSC g kg^{-1} DM) and similar to 42.0 g kg^{-1} DM recorded by Jonathan et al. (51) for *C. ciliaris.* Higher C_C and lower C_{B2} for IG97-379, EC400605, and IG96-96 accessions may be attributed to their higher lignin and lower NDF contents as the forages with high NDF have a higher proportion of C_{B2} fraction, and the increase in the fraction C_C can be partly attributed to the increased lignin concentration in NDF (60).

C. ciliaris high sugar accessions P_C fraction varied ($p < 0.05$) with the mean value of 115 g kg^{-1} CP, which is in the agreement with the range of 100–150 g kg^{-1} CP as reported earlier (61, 62). Fraction C is the insoluble N in acid detergent solution (ADIN) and is associated with lignin, tannin–protein complexes, and Maillard products. This fraction represents the unavailable protein and is assumed to have zero ruminal and intestinal digestibility. In our study, P_C fraction varied ($p < 0.05$) between 81.6 and 152 g kg^{-1} CP is within the range of 90–180 g kg^{-1} CP as reported by Sanderson and Wedin (63) and Hoekstra et al. (64). Jonathan et al. (51) reported that *C. ciliaris* P_A, P_{B1}, P_{B2}, P_{B3}, and P_C fractions of protein were 345, 152, 137, 195, and 480 g kg^{-1} CP, respectively, which were inconsistent with our protein fraction values except the mean value of P_{B3} (124 g kg^{-1} CP). Braga et al. (65) reported ($p < 0.05$) differences in protein fractions in grass species and their harvesting age. Protein fraction P_A composed of NPN has a higher rate of ruminal degradation, which was lower in *Andropogon* (120–130 g kg^{-1} CP) than in *C. ciliaris* and Massai (160–170 g kg^{-1} CP) at 63 days of cutting age. Grasses P_{B2}, P_{B3}, and P_C protein fractions ranged between 280–340, 270–310, and 213–273 g kg^{-1} CP, respectively, at 63 days of cutting age. Brandstetter et al. (57) reported that the protein fractions P_A, P_{B1}, P_{B2}, P_{B3}, and P_C of Jiggs Bermuda grass in different seasons (fall, winter, spring, and summer) between 407–550, 138–139, 100–157, 125–148, and 80–115 g kg^{-1} CP, respectively, and their P_{B2}, P_{B3}, and P_C values corroborate with our values of sugar-rich accessions for these fractions. Sa et al. (58) reported that protein fractions P_A, P_{B1+B2}, P_{B3}, and P_C of *Cynodon dactylon, Brachiaria brizantha,* and *Pancum maximum* grasses at 28, 35, and 54 days of cutting age ranged between 164 to 287, 251 to 538, 116 to 349, and 91 to 182 g kg^{-1} CP, respectively, which are more or less similar to our observations.

4.4. Energy contents

The calculated mean ME values of *Cenchrus* germplasm (7.41 Mj kg^{-1} DM) and sugar-rich accessions (6.42 Mj kg^{-1} DM) were lower than those reported by Getachew et al. (66) for 17 grass samples (7.7–13.6 Mj kg^{-1} DM) and are inadequate to fulfill the energy requirements for the maintenance of growing cattle (8.8 Mj kg^{-1} DM; 37). Mlay et al. (67) reported that TDN, DE, and ME contents of 10 types of tropical grass ranged between 342–609 g kg^{-1} DM, 5.92–11.26, and 4.85–9.23 Mj kg^{-1} DM, respectively, and our values of *Cenchrus* germplasm (411–549 g kg^{-1} DM, 7.99–10.31, and 6.17–8.44 Mj kg^{-1} DM) and sugar-rich accessions (362–487 g kg^{-1} DM, 6.62–8.90, and 5.42-7.29 Mj kg^{-1} DM) lie within these values. The higher TDN and DE values of IG99-124 may be due to lower ADF and lignin contents (432 and 39.8 g kg^{-1} DM) as higher ADF and lignin contents reduce the nutrient utilization present in forages (68). Yigzaw (69) reported that the ME contents of *C. ciliaris* varied in the range of 7.65–9.02 Mj kg^{-1} DM during

TABLE 7 *In vitro* gas and methane production from sugar-rich *Cenchrus ciliaris* accessions.

Genotypes	Gas ml g^{-1}	CH_4 ml g^{-1}	DDM g kg^{-1} DM	CH_4 ml g^{-1} DDM	PF mg DDM ml^{-1}	SCFA mmol g^{-1} DM	MBM mg g^{-1} DM	EMBM mg mg^{-1}
IG96-87	97.6^{abcd}	7.72^{a}	439^{ab}	17.4^{a}	4.50^{abcd}	2.16^{abcd}	224^{bc}	0.51^{bcd}
EC400605	111^{bcd}	11.2^{abc}	501^{bcde}	22.8^{ab}	4.59^{bcd}	2.46^{bcd}	256^{c}	0.51^{bcd}
IG97-378	104^{bcd}	10.6^{ab}	458^{abcd}	23.1^{ab}	4.40^{abcd}	2.32^{bcd}	228^{bc}	0.49^{bc}
EC397366	104^{bcd}	15.3^{bcd}	531^{ef}	28.5^{abc}	5.16^{de}	2.30^{bcd}	303^{cde}	0.57^{cde}
EC397379	122^{def}	17.3^{cde}	509^{cde}	33.9^{bcd}	4.18^{abcd}	2.72^{def}	239^{bc}	0.47^{bc}
IG96-96	101^{bcd}	14.3^{bcd}	530^{ef}	26.6^{abc}	5.37^{de}	2.23^{bcd}	309^{cde}	0.58^{cde}
IG96-89	104^{bcd}	15.3^{bcd}	482^{bcde}	31.4^{bc}	4.69^{cd}	2.31^{bcd}	253^{c}	0.52^{bcd}
IG96-50	73.1^{a}	13.6^{abcd}	517^{def}	26.2^{abc}	7.07^{f}	1.62^{a}	356^{de}	0.69^{e}
CC-14-1	116^{cde}	19.8^{de}	524^{def}	37.7^{cde}	4.67^{cd}	2.57^{cde}	269^{cd}	0.51^{bcd}
IG99-124	110^{bcd}	15.3^{bcd}	399^{a}	38.5^{cde}	3.64^{abc}	2.44^{bcd}	156^{ab}	0.39^{ab}
IG97-379	141^{ef}	19.1^{de}	447^{abc}	43.5^{de}	3.30^{ab}	3.13^{ef}	136^{a}	0.30^{a}
IG-97-377	146^{f}	21.5^{e}	457^{abcd}	47.2^{e}	3.24^{a}	3.23^{f}	137^{a}	0.30^{a}
EC397323	87.4^{ab}	12.7^{abc}	544^{ef}	23.4^{ab}	6.34^{ef}	1.94^{ab}	352^{de}	0.65^{de}
IG97-403	93.8^{abc}	14.3^{bcd}	579^{f}	24.8^{ab}	6.39^{ef}	2.08^{abc}	373^{e}	0.64^{de}
Mean	108	14.9	493	30.5	4.79	2.40	254	0.51
SEM	3.21	0.66	7.97	1.44	0.176	0.07	11.9	0.02
P-value	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001	<0.0001

DDM, Degraded dry matter; PF, Partition factor mg of *in vitro* degraded dry matter to ml of gas thereby produced; SCFA, Short-chain fatty acid; MBM, Microbial mass; EMBP, Efficiency of microbial protein production. Superscripts a–f within columns differ significantly between rows at ($P < 0.05$).

60 to 120 days of crop growth. The effect of growth stage, season, and species on ME contents of tropical grasses has been reported earlier (51, 70). The net energy system recommends an animal energy requirement for different physiological functions, *viz.* tissue maintenance, tissue growth, and lactation (71). The NE_L contents of 4.38 and 4.85 Mj kg^{-1} DM for Napier and Pangola grasses, respectively, reported by Tikam et al. (70) are similar to the mean NE_L values of *Cenchrus* germplasm (4.51 Mj kg^{-1} DM) and higher than the mean NE_L contents of sugar-rich *C. ciliaris* accessions (3.85 Mj kg^{-1} DM). The NE_L values of IG99-124. EC397323, IG96-96, and IG96-98 accessions are almost similar to the values of Tikam et al. (70). *Cenchrus* germplasm and high sugar accessions had the adequate NE_M levels recommended for a mature beef cow [4.92–5.30 Mj kg^{-1} DM (71)].

4.5. Palatability attributes of *Cenchrus* accessions

Many indices to predict the forage quality for feeding ruminants have been developed (72) based on the chemical constituents of forages. Intake is one of the important indices to measure the nutrient availability of animals influenced by both diet and livestock species. Differences in DMI of *Cenchrus* germplasm (1.61–1.88%) and sugar-rich accessions (1.53–1.77%) may be attributed to the variation in their NDF contents. The mean NDF contents of both germplasm (694 g kg^{-1} DM) and sugar-rich accessions (737 g kg^{-1} DM) were beyond the range of 600–650 g kg^{-1} DM level, which is considered to influence the intake negatively (18). The DMI of tropical grasses at the flowering stage varied ($p < 0.05$) and ranged between 30.0 for *Panicum maximum* and 54.0g/kg w^{075} for *Brachiaria ruziziensis* and *Pennisetum purpureum* in sheep fed *ad lib* (73). Aguir et al. (74) reported DMI of 2.33% and 2.25% for Sudan grass and Elephant grass, respectively, in goats. Assoumaya et al. (75) reported that the voluntary intake of tropical forages is lower (1.95) than that of temperate forages (2.03%).

Forages with digestibility values of 500 g kg^{-1} DM or more can meet the energy requirements for the maintenance of grazing ruminants (76). The significant differences in DDM values of *C. ciliaris* germplasm (533–585 g kg^{-1} DM) and sugar-rich accessions (477–552 g kg^{-1} DM) may be attributed to the differences in their ADF and lignin contents as both nature and quantity of cell wall contents and cell contents of forages influence the DM degradability (37). Like our observations, Dessommes et al. (42) reported ($p < 0.05$) the differences in effective DM degradability of six *C. ciliaris* genotypes (550–663 g kg^{-1} DM). The *IVDMD* contents of 11 ecotypes of *C. ciliaris*

TABLE 8 *Cenchrus ciliaris* accessions silage composition.

Accessions	Silage composition†			
	pH	Lactic acid g kg^{-1} (Fresh)	Lactic acid g kg^{-1} DM)	DM at ensiling g kg^{-1} DM
IG96-87	5.54^{bcd}	5.63^{ef}	21.6^{ef}	185^{a}
EC400605	5.68^{cde}	1.08^{a}	5.01^{ab}	212^{ab}
IG97-378	5.77^{def}	2.98^{abcde}	5.90^{abc}	228^{abc}
EC397366	5.11^{a}	7.15^{f}	23.7^{f}	305^{de}
EC397379	6.07^{g}	3.15^{abcde}	9.22^{bcd}	343^{e}
IG96-96	5.48^{bc}	4.30^{bcdef}	17.8^{e}	239^{abc}
IG96-89	5.66^{cde}	4.45^{cdef}	19.3^{ef}	232^{abc}
IG96-50	5.46^{bc}	2.85^{abcde}	11.1^{cd}	254^{bcd}
CC-14-1	5.95^{fg}	2.15^{abc}	8.71^{bcd}	246^{bc}
IG99-124	5.77^{def}	2.63^{abcde}	12.5^{d}	213^{ab}
IG97-379	5.70^{cde}	3.83^{abcde}	7.70^{abcd}	274^{cd}
IG-97-377	5.54^{bcd}	5.30^{def}	6.91^{abc}	258^{bcd}
EC397323	5.41^{b}	2.50^{abcd}	14.3^{de}	233^{abc}
IG97-403	5.48^{bc}	1.28^{ab}	3.71^{a}	245^{bc}
IG3158	5.77^{def}	1.80^{abc}	7.90^{abcd}	225^{abc}
Mean	5.60	3.40	11.6	246
SEM	0.034	0.30	0.89	6.16
P-value	<0.0001	<0.0001	<0.0001	<0.0001

† LA, Lactic acid; DM, Dry matter. Superscripts a–g within columns differ significantly between rows at ($P < 0.05$).

ranged between 456 and 550 g kg^{-1} DM (50). Coelho et al. (49) reported *IVDMD* of 452 and 500 g kg^{-1} DM for *C. ciliaris* under stockpiled and grazing conditions.

Relative feed value (RFV) forage quality index combines the intake and digestibility into one unit (77). Alike our observations on RFV for *Cenchrus* germplasm (64.21–83.87%) and sugar-rich accessions (58.24–73.78%), Hackman et al. (78) reported a wide variability in RFV of 11 cool-season (71.5–130.0%) and four warm-season types of grass (88.0–165.0%), respectively. Suhaimi et al. (79) evaluated more than 900 samples of *Brachiaria decumbens* grass over 5 years (1999–2003) and observed that their RFV values ranged between 74.83 and 84.17%. Cinar and Hatipoglu (80) reported that the RFV of Dallis, Bermuda, and Rhodes grasses varied between 68.0–82.1, 75.0–93.7, and 72.8–86.7%, during 3 years of growth (2009–2011) and our RFV lies within these values.

4.6. Gas and methane production from sugar-rich *Cenchrus* accessions

Cenchrus ciliaris accessions differ ($p < 0.05$) in gas production with the mean value of 108 ml g^{-1} DM and are on the pattern of Ley de Coss et al. (81) who recorded ($p < 0.05$) the differences for *in vitro* gas production from four types of tropical grass (122–170 ml g^{-1} DM) at 24 h of incubation in bovine rumen liquor. Garcia and Dessommes (42) reported CH_4 production of 4.38 ml g^{-1} DM of *C. ciliaris* at 24 h of fermentation while CH_4 production for 16 types of grass ranged between 4.02 and 11.70 ml g^{-1} DM, which was lower than our CH_4 values (7.72–19.80 ml g^{-1} DM). In contrast, Bezabiah et al. (48) reported higher gas and CH_4 (202 and 40 ml g^{-1} DM) of *C. ciliaris* fermented for 24 h in rumen liquor of Holstein Friesian cattle. Similarly, Melesse et al. (44) recorded higher gas and methane production (204.5 and 34.5 ml g^{-1} DM) for *C. ciliaris* than our values. The authors further reported that the total gas and CH_4 of 24 grass species ranged between 94 to 232 ml g^{-1} DM and 26 to 43 ml g^{-1} DM, respectively. The variation in methane production among *C. ciliaris* accessions may be partially attributed to their significant differences in chemical constituents such as CP, ash, ether extract, ADF, NDF, ADL, NDIN, ADIN, and NFC concentration. The ratio between methane to total gas production indicates that the methane emission potential per unit of OM degraded from forages, and in the present study, this ratio varied widely from 0.079 to 0.170 across the sugar-rich accessions, which shows the opportunity to select the accessions with low methane potential.

Accessions gas fermentation parameters, *viz.* partition factor (PF), short-chain fatty acid (SCFA), microbial mass (MBM), and efficiency for microbial mass production (EMBM) differed ($p < 0.05$). Accession IG96-50 higher values for PF and EMBM and lower for SCFA were consistent with the previous report, and the microbial mass and SCFA are inversely related (82, 83). Higher PF recorded for IG96-50 resulted in greater microbial mass as PF is the measure of the efficiency of microbial production. The amount of short-chain fatty acid produced is related to OMD and the energy content of the feed.

4.7. Silage quality

Typical concentrations of lactic acid in commonly fed silages range from 20.0 to 40.0 g kg^{-1} DM, but can be considerably higher in silages with low concentrations of DM (<300 g kg^{-1} DM). The final pH of silage is affected by many factors but is most related to the concentration of lactic acid and buffering capacity of the crop. Silage prepared from *C. ciliaris* accessions had pH values (5.11–6.07) above the 3.8 to 4.2 ideal ranges (52) usually observed in corn or sorghum or oat silages. Silage pH values of EC397366 (5.11), IG96-96 (5.46), IG96-50 (5.46), and EC397323 (5.41) accessions are more or less in the acceptable range of tropical grasses with lactic acid contents of 23.7, 17.8, 11.1, and 14.3 g kg^{-1} DM, respectively. Harrison et al. (84) recommended that good grass silage should have pH <4.47 and lactic acid between 40.0 and 70.0 g kg^{-1} DM, respectively. The pH values recorded for sugar-rich *Cenchrus* accessions are acceptable and consistent with the values reported for grasses (85, 86). Aminah et al. (54) reported that silages from *Seteria splendid* and *Pennisetum purpureum* had lower pH (4.07 and 3.96) and more lactic acid (24.7 and 25.3 g kg^{-1} DM) than other evaluated tropical grasses (4.71–5.32) and 10.4–18.4 g kg^{-1} DM) partially agree with our values of IG96-96, EC397366, EC397323, and IG96-50 accessions. The DM content of evaluated accessions except EC397366 was below the range of 300 g kg^{-1} DM desirable for ensiling grasses. Pitt (87) suggested that grasses ensiled below 300 g kg^{-1} D should have >100 g kg^{-1} DM soluble sugar for adequate fermentation to achieve the desired pH. Accessions having higher pH might have failed to provide adequate substrate (sugar) to lactic acid bacteria to produce lactic acid. Pinho et al. (88) recorded the pH of Buffel grass silage between 4.6 and 5.4 at 30 days of fermentation harvested at different heights. Li et al. (89) reported the pH and LA contents of *Paspalum plicatulum* grass 5.2 and 11.0 g kg^{-1} DM and 5.2 and 18.0 g kg^{-1} DM during 30 days of fermentation at 28 and 40°C temperature and 5.2 and 17.0 g kg^{-1} DM and 5.1 and 20.0 g kg^{-1} DM during 60 days of ensiling at 28 and 40°C temperature, respectively. Yahaya et al. (90) showed that silage of tropical grass (*Pennisetum purpurum*) had higher pH and lower lactic acid (5.45 and 9.00 g kg^{-1} DM) than temperate rye grass silage (3.86 and 19.0 g kg^{-1} DM). In another study, Arroquy et al. (91) recorded lower pH (4.04–4.47) and higher lactic acid (39.1–76.5 g kg^{-1} DM) for six warm season types of grass than our pH and lactic acid values. However, Vendramini et al. (92) reported higher pH (6.5–8.6) and lower lactic acid (1.00–19.0 g kg^{-1} DM) for warm season grasses except for Limpo grass (26.0 g kg^{-1} DM) than our pH and lactic acid values.

5. Conclusions

The results revealed wide genetic variability in *Cenchrus* germplasm and sugar-rich accessions for dry matter yield, protein, fiber, energy, and soluble sugar contents. Sugar-rich accessions also differ ($p < 0.05$) for carbohydrate fractions, protein fractions, *in vitro* gas and methane production, and silage quality (pH and lactic acid). Silage prepared from EC397366, IG96-96, IG96-50, and EC397323 accessions had pH and lactic acid contents acceptable for tropical range grasses. Nutritional evaluation of silage prepared from selected accessions may be undertaken using *in vivo* studies. The present global subset having wide variability could be utilized for the identification of genomic regions associated with key forage nutritional traits for future breeding programs. Selected accessions need to be introduced in rangelands and pastures to enhance their yield and quality for sustainable livestock production.

Data availability statement

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

Ethics statement

The animal study was reviewed and approved by Institute Animal Ethics Committee.

Author contributions

SS, TS, and SKM: conceptualized the study. SS, TS, SKM, MMD, KKS, RK, and AKM: methodology and laboratory work. PKG, AKM, SS, and TS: resources and supervision. SS, TS, and MKS: investigation. SS, TS, and NK: data analysis and writing. All authors contributed to the article and approved the submitted version.

Funding

This study was funded by the Indian Council of Agricultural Research project no. CRSCIGFRICOL20130702.

Acknowledgments

The authors are grateful to the Indian Council of Agricultural Research for funding and to the Director, ICAR-Indian Grassland and Fodder Research Institute for providing the facilities necessary to conduct these studies.

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/fnut.2022.1094763/full#supplementary-material

References

1. Herrero M, Pavlik P, Valin H, Notenbaert A, Rufino M, Thornton PK, et al. Global livestock systems: Biomass use, production, feed efficiency and green house gas emissions. *Proc National Academy Sci.* (2013) 110:20888–93. doi: 10.1073/pnas.1308149110

2. Minson DJ. *Forage in Ruminant Nutrition*. San Diego: Academic Press (1990). p.483.

3. Coan RM, Reis RA, Garcia GR, Schocken-Inturrino RP, Souza-Ferreira D, Dutra RF, et al. Microbiological and fermentative dynamics of Tanzania grass and *B.* bryzanta grass silage using pelleted citrus pulp as an additive. *R Bras Zootec.* (2007) 36:1502–11. doi: 10.1590/S1516-35982007000700007

4. Keady TWJ, Kilpatrick DJ, Mayne CS, Gordon FJ. Effects of replacing grass silage with maize silages, differing in maturity, on performance and potential concentrate sapring effect of dairy cows offered two feed value grass silages. *Livest Sci.* (2008) 119:1–11. doi: 10.1016/j.livsci.2008.02.006

5. Spitaleri RF, Sollenberger LE, Staples CR, Schank SC. Harvest management effects on ensiling characteristics and silage nutritive value of seeded *Pennisetum* hexaplois hybrids. *Postharvest Biol Technol.* (1995) 5:353–62. doi: 10.1016/0925-5214(94)00033-O

6. Humphreys MO. "Variation in the carbohydrate and protein content of ryegrass: Potential for genetic manipulation," In: *Proceeding of the 19th EUCARPIA Fodder Crops Section Meeting*. Brugge: Belgium (1994). p. 165–71.

7. Haigh PM. Effect of herbage water soluble carbohydrate content and weather conditions at ensilage on the fermentation of grass silages made on commercial farms. *Grass Forage Sci.* (1990) 45:263–71. doi: 10.1111/j.1365-2494.1990.tb01949.x

8. Mc Eniry J, King C, O'Kiely P. Silage fermentation characteristics of three common grassland species in response to advancing stage of maturity and additive application. *Grass Forage Sci*. (2013) 69:393–404. doi: 10.1111/gfs.12038

9. Syamaladevi DP, Meena SS, Nagar RP. Molecular understandings on 'the never thirsty' and apomictic Cenchrus grass. *Biotechnol Lett*. (2016) 38:369–76. doi: 10.1007/s10529-015-2004-0

10. Kumar VV, Mahato AK, Patel R. "Ecology and management of Banni grassland of Kuchchh Gujrat *in* Ecology and management of grassland habitats in India," In: Rawat GS, Adhikari BS (Eds) *ENVIS Bulletin Wildlife and Protected Areas*. Dehradun: Wildlife Institute of India (2015). p. 43–53.

11. Ecocrop. *Ecocrop database*. Rome: FAO (2010).

12. Duke JA. *Handbook of Energy Crops*. West Lafayette, IN: New CROPS web site, Purdue University (1983).

13. Osman AE, Makawi M, Ahmed R. Potential of the indigenous desert grasses of the Arabian Peninsula for forage production in a water-scarce region. *Grass Forage Sci.* (2008) 63:495–503. doi: 10.1111/j.1365-2494.2008.00656.x

14. Sahay G, Dikshit N, Singh T, Tyagi VC, Dheeravathu SN, Ahmed S, et al. *FORAGE Genetic Resources at ICAR-IGFRI Jhansi, Accomplishments and Future Strategies*. Jhansi: ICAR-Indian Grassland and Fodder Research Institute (2019). p. 34.

15. Federer WT. Augmented designs with one-way elimination of heterogeneity. *Biometrics*. (1961) 17:447–73. doi: 10.2307/2527837

16. VDLUFA (Verband Deutscher Landwirtschaftlicher Untersuchungs- und Forschungsanstalten). *Die chemische Untersuchung von Futtermitteln. VDLUFA-Methodenbuch. Band III, Ergänzungslieferungen von 2007*. Darmstadt: VDLUFA-Verlag (2007). Available online at: https://www.vdlufa.de/Methodenbuch/index.php?option=com_content&view=article&id=4&Itemid=111&lang=de&lang=en

17. AOAC. *Association of Official Analytical Chemists, 18th edn*. Washington, DC: Official Methods of Analysis of AOAC International (2005).

18. Van Soest PJ, Robertson JB, Lewis BA. Methods for dietary fiber, neutral detergent fiber and non starch polysaccharides in relation to animal nutrition. *J Dairy Sci.* (1991) 74:3583–97. doi: 10.3168/jds.S0022-0302(91)78551-2

19. Sastry VRB, Kamra DN, Pathak NN. *Laboratory Manual of Animal Nutrition*. Izatnaar, UP: Centre of Advance Studies, Division of Animal Nutrition Indian Veterinary Research Institute (1991). p. 116–7.

20. Sniffen CJ, O'Connor JD, Van Soest PJ, Fox DG, Russel JB. A net carbohydrate and protein system for evaluating cattle diets II carbohydrate and protein availability. *J Anim Sci.* (1992) 70:3562–77. doi: 10.2527/1992.70113562x

21. Caballero R, Alzueta C, Ortiz LT, Rodriguez ML, Barro C, Rebolé A, et al. Carbohydrate and protein fractions of fresh and dried common vetch at three maturity stages. *Agron. J.* (2001) 93:1006–13. doi: 10.2134/agronj2001.9351006x

22. Licitra G, Harnandez TM, Van Soest PJ. Standardizations of procedures for nitrogen fractionation of ruminant feeds. *Anim Feed Sci Technol.* (1996) 57:347–58. doi: 10.1016/0377-8401(95)00837-3

23. Krishnamoorthy U, Muscato TV, Sniffen CJ, Van Soest PJ. Nitrogen fractions in selected feedstuffs. *J Dairy Sci.* (1982) 65:217–25. doi: 10.3168/jds.S0022-0302(82)82180-2

24. Undersander DJ, Howard WT, Shaver RD. Milk per acre spreadsheet for combining yield and quality into a single term. *J Prod Agric.* (1993) 6:231–223. doi: 10.2134/jpa1993.0231

25. Fonnesbeck PV, Clark DH, Garret WN, Speth CF. Predicting energy utilization from alfalfa hay from the Western region. *Proc Amer Soc Anim Sci W Sect.* (1984) 35:305–8.

26. Khalil JK, Sawayaw N, Hyder SZ. Nutrient composition of Atriplex leaves grown in Saudi Arabia. *J Range Manag.* (1986) 39:104–7. doi: 10.2307/3899277

27. Theodorou MK, Williams AB, Dhanoa MS, Mc Allan AB, France J. a simple gas production method using pressure transducer to determine the fermentation kinetics of ruminant feeds. *Anim Feed Sci Technol.* (1994) 48:185–97. doi: 10.1016/0377-8401(94)90171-6

28. Menke KH, Steingass H. Estimation of the energetic feed value obtained from chemical analysis and *in vitro* gas production using rumen fluid. *Anim Res Dev.* (1988) 28:7–55.

29. Tavendale MH, Meahger LP, Pacheco D, Walker N, Attwood GC, Sivakumaran S, et al. Methane production from *in vitro* rumen incubations with

Lotus penduculatus and *Medicago sativa* and effects of extractable condensed tannin fractions on methanogenesis. *Anim. Feed Sci. Technol.* (2005) 123:403–19. doi: 10.1016/j.anifeedsci.2005.04.037

30. Getachew G, Makkar HPS, Becker K. Tropical browses: contents of phenolic compounds, *in vitro* gas production and stoichiometric relationship between short chain fatty acid and *in vitro* gas production. *J Agric Sci.* (2002) 139:341–52. doi: 10.1017/S0021859602002393

31. Blümmel M, Makkar HPS, Becker K. *In vitro* gas production: a technique revisited. *J Anim Physiol Anim Nutr.* (1997) 77:24–34. doi: 10.1111/j.1439-0396.1997.tb00734.x

32. Kaiser AG, Kerr KL. More accurate laboratory tests for assessing silage quality, Final Report for DRDC Project DAN 100 and RIRDC Project DRD-4A (2003). p. 50.

33. Barker SB, Summerson WH. The colorimetric determination of lactic acid in biological material. *J Biol Chem.* (1941) 138:535–54. doi: 10.1016/S0021-9258(18)51379-X

34. Aravind J, Mukesh S, Wankhede D, Kaur V. Augmented RCBD: analysis of augmented randomised complete block designs. *R Package Vers 0.1.5.900.* (2021) 2:1–54.

35. Snedecor GW, Cochran WG. *Statistical Methods, 8th ed*. Calcutta, India: Oxford and IBH Publishing Company. (1994).

36. SAS Institute. *SAS/STAT 9, 3. User's Guide*. Cary, NC: SAS Institute Inc. (2011).

37. Van Soest PJ. *Nutritional Ecology of the Ruminant*. Znd ed. Ithaca, New York, USA Comstock Publishing Associates and Cornell University Press (1994).

38. Ganskopp D, Bohnert D. Nutritional Dynamics of 7 Northern Great Basin Grasses. *J Range Manag.* (2001) 54:640–7. doi: 10.2307/4003664

39. NRC. *Nutrient Requirements of Domestic Animals, Nutrient Requirements of Beef Cattle. 7th ed.* National Academy of Sciences. National Research Council, Washington, USA (2000).

40. Peerzada AM, Naeem M, Ali HH, Tanveer A, Javaid MM, Chauhan BS. *Cenchrus biflorus* Roxb. (Indian sandbur), a blessing or curse in arid ecosystems: a review. *Grass Forage Sci.* (2016) 72, 179–192. doi: 10.1111/gfs.12215

41. Kumar K, Soni A. Nutrient evaluation of common vegetation of Rajasthan: *Pennisetum typholdenum*, *Cenchrus ciliaris*, *Cenchrus setigerus* and *Lasiurus sindicus*. *Int J Plant Anim Environ Sci.* (2014) 4:177–84. Available online at: www.ijpaes.com

42. García-Dessommes GJ, Ramírez-Lozano RG, Morales Rodríguez R, García-Díaz G. Ruminal digestion and chemical composition of new genotypes of buffelgrass (*Cenchrus ciliaris* L.) under irrigation and fertilization. *Interciencia.* (2007) 32:349–53. Available online at: https://www.redalyc.org/articulo.oa?id=33932511

43. Morales-Rodríguez R, Ramírez RG, García-Dessommes GJ, González-Rodríguez H. Nutrient Content and *in-situ* Disappearance in Genotypes of Buffelgrass (*Cenchrus ciliaris* L.). *J. Appl. Anim. Res.* (2006) 29:17–22. doi: 10.1080/09712119.2006.9706562

44. Melesse A, Steingass H, Schollenberger M, Rodehutscord M. Screening of common tropical grass and legume forages in Ethiopia for their nutrient composition and methane production profile *in vitro*. *Tropical Grasslands Forrajes Tropicales.* (2017) 5:163–75. doi: 10.17138/TGFT(5)163-175

45. Ashraf MA, Mahmood K, Yusoff I, Qureshi AK. Chemical constituents of *Cenchrus ciliaris L.* from the Cholistan desert, Pakistan. *Arch Biol Sci Belgr.* (2013) 65:1473–8. doi: 10.2298/ABS1304473A

46. Saini ML, Jain P, Joshi UN. Morphological characteristics and nutritive value of some grass species in an arid ecosystem. *Grass Forage Sci.* (2007) 62:104–8. doi: 10.1111/j.1365-2494.2007.00563.x

47. Gemeda BS, Hassan A. *In vitro* fermentation, digestibility and methane production of tropical perennial grass species. *Crop Pasture Sci.* (2014) 65:479–88. doi: 10.1071/CP13450

48. Bezabih M, Pellikaan WF, Tolera A, Khan NA, Hendriks WH. Chemical composition and *in vitro* total gas and methane production of forage species from the Mid Rift Valley grasslands of Ethiopia. *Grass Forage Sci.* (2014) 69:635–43. doi: 10.1111/gfs.12091

49. Coêlho JJ, Leão AC, Santos MVF, Junior BD, Cunha MV, Lira A. Prediction of the nutritional value of grass species in the semiarid region by repeatability analysis. *Pesqui. Agropecu. Bras.* (2018) 53:378–385. doi: 10.1590/s0100-204x2018000300013

50. Kirwa EC, Njoroge K, Chemining'wa G, Mnene WN. Nutritive composition of *Eragrostis egund* Peyr and *Cenchrus ciliaris* L. collections from the ASALs of Kenya. *Livestock Res Rural Develop.* (2015) 27:144. Available online at: https://www.lrrd.cipav.org.co/lrrd27/8/kirw27144.html

51. Jonathan N, Avilés N, José L, Valle C, Francisco CP, Sergio AC, et al. Digestibility of Buffel grass (*Cenchrus ciliaris*)-based diets supplemented with four levels of *Gliricidia sepium* hay in hair sheep lambs. *Proc. New Zealand Grassland Assoc.* (2013) 45:1357–62. doi: 10.1007/s11250-013-0369-4

52. McDonald P, Henderson AR, Heron SJE. *The Biochemistry of Silage. 2nd ed*. Mallow, Bucks: Chalcombe Publications (1991).

53. Gallaher RN, Pitman WD. "Conservation of forages in the tropics and subtropics," In: Sotomayor-Rios A, Pitman WD, eds *Tropical Forage Plants: Development and Use*. Boca Raton: CRC Press LLC. (2001). p. 2136–2145. doi: 10.1201/9781420038781.ch14

54. Aminah A, Abu BC, Izham A. *Silages From Tropical Forages: Nutritional Quality and Milk Production*. FAO Electronic Conference on Tropical Silage. (1999). Available online at: http://ftpmirror.your.org/pub/misc/cd3wd/1005/_ag_lstock_silage_4P1_en_111610.pdf (accessed March 12, 2020).

55. Piltz JW, Kaiser AG. "Principles of silage preservation," In: Piltz JW, Kaiser AG, Burns HM, Griffith NW. *Successful Silage* The State of New South Wales: Department of Primary industries and Dairy Australia (2004). p. 26–56.

56. Campos SS, Silva JFC, Vásquez HM, Vittori A, Silva MA. Fractions of carbohydrates and of nitrogenous compounds of tropical grasses at different cutting ages. *R Bras Zootec.* (2010) 39:1538–47. doi: 10.1590/S1516-35982010000700021

57. Brandstetter EV, Costa KAP, Santos DC, Souza WF, Silva VC, Dias MBC, et al. Protein and carbohydrate fractionation of Jiggs Bermudagrass in different seasons and under intermittent grazing by Holstein cows. *Acta Scientiarum. Acta Sci. Anim. Sci.* (2019) 41:e43363. doi: 10.4025/actascianimsci.v41i1.43363

58. Sá JF, Pedreira MS, Silva FF, Bonomo P, Figueiredo MP, Menezes DR, et al. Carbohydrates and proteins fractions of tropical grasses cut at three ages. *Arq Bras Med Vet Zootec.* (2010) 62:667–76. doi: 10.1590/S0102-09352010000300023

59. Juarez Lagunes FI, Fox DG, Blake RW, Pell AN. Evaluation of Tropical Grasses for Milk Production by Dual-Purpose Cows in Tropical Mexico. *J Dairy Sci.* (1999) 82:2136–45. doi: 10.3168/jds.S0022-0302(99)75457-3

60. Silva SP, Silva MMC. Fracionamento de carboidrato e egundo egundo o sistema CNCPS. *Vet Notícias.* (2013) 19:95–108. Available online at: https://seer.ufu.br/index.php/vetnot/article/view/23169

61. Elizalde JC, Merchen NR, Faulkner DB. Fractionation of fiber and crude protein in fresh forages during the spring growth. *J Anim Sci.* (1999) 77:476–84. doi: 10.2527/1999.772476x

62. Bryant RH, Gregorin IP, Edwards GR. Effects of N fertilisation, leaf appearance and time of day on N fractionation and chemical composition of *Lolium perenne* cultivars in spring. *Anim Feed Sci Tech.* (2012) 173:210–9. doi: 10.1016/j.anifeedsci.2012.02.003

63. Sanderson MA, Wedin WF. Nitrogen in the detergent fibre fractions of temperate legumes and grasses. *Grass Forage Sci.* (1989) 44:159–68. doi: 10.1111/j.1365-2494.1989.tb01923

64. Hoekstra NJ, Struik PC, Lantinga EA, Amburgh ME, Schulte RPO. Can herbage nitrogen fractionation in *Lolium perenne* be improved by herbage management? *NJAS-Wagen. J Life Sci.* (2008) 55:167–80. doi: 10.1016/S1573-5214(08)80035-8

65. Braga AP, Assis LC, Lucena JA, Lima JSS, Barreto FB, Amâncio AF, et al. Fractionation of nitrogen compounds and carbohydrates in forages of different ages. *Semina: Ciências Agrárias.* (2018) 39:819–32. doi: 10.5433/1679-0359.2018v39n2p819

66. Getachew G, DePeters EJ, Robinson PH, Fadel JG. Use of an *in vitro* rumen gas production technique to evaluate microbial fermentation of ruminant feeds and its impact on fermentation products. *Anim Feed Sci Technol.* (2005) 124:547–59. doi: 10.1016/j.anifeedsci.2005.04.034

67. Mlay PS, Pereka A, Phiri EC, Balthazary S, Igusti J, Hvelplund T, et al. Feed value of selected tropical grasses, legumes and concentrates. *Vet Arh.* (2006) 76:53–63. Available online at: https://hrcak.srce.hr/file/8188

68. Aydin N, Mut Z, Mut H, Ayan I. Effect of autumn and spring sowing dates on hay yield and quality of oat (*Avena sativa* L.) genotypes. *J. Anim. Vet. Adv.* (2010) 9:1539–1545. doi: 10.3923/javaa.2010.1539.1545

69. Yigzaw GW. Effect of harvesting stage on yield and nutritive value of Buffel Grass (*Cenchrus ciliaris* linn*)* under irrigation at Gewane district, North Eastern Ethiopia. *J Sci Innov Res.* (2019) 8:7–12. Available online at: www.jsirjournal.com

70. Tikam K, Phatsara C, Mikled C, Vearasilp T, Phunphiphat W, Chobtang J, et al. Pangola grass as forage for ruminant animals: a review. *Springer Plus*. (2013) 2:604. doi: 10.1186/2193-1801-2-604

71. NASEM. *National Academies of Sciences, Engineering, and Medicine. Nutrient Requirements of Beef Cattle, Eighth Revised Edition.* Washington, DC: The National Academies Press (2016). doi: 10.17226/19014

72. Moore JE. "Forage quality indices: development and application," In: Fahey GC, ed *Forage Quality, Evaluation, and Utilization*. Madison, WI: ASA, CSSA, SSSA (1994). p. 977–98. doi: 10.2134/1994.foragequality.c24

73. Babatoundé S, KaKaï RG, Alkoiret I, Mensah GA. Intake and digestibility of native and exotic grasses fed *ad libitum* to Djallonke sheep in South Benin. *J. Agric. Sci. Technol.* (2011) 5, 513–523. doi: 10.17265/2161-6256/2011.04A.019 Available online at: https://www.researchgate.net/publication/233780926

74. Aguiar EM, Lima GFC, Santos MVF, Carvalho FFR, Medeiros HR, Maciel FC, et al. Intake and apparent digestibility of chopped grass hays fed to goats. *R Bras Zootec.* (2006) 35:2219–25. doi: 10.1590/S1516-35982006000800004 Available online at: https://www.researchgate.net/publication/262650759

75. Assoumaya C, Sauvant D, Archimède H. Comparative study of ingestion and digestion tropical and temperate forages. *INRA Prod Anim.* (2007) 20:383–92. Available online at: https://www.researchgate.net/publication/285665975

76. Kearl LC. *Nutrient Requirements of Ruminants in Developing Countries*. International Feedstuffs Institute, Utah Agricultural Experimental Station Utah State University, Logan Utah, USA (1982). p. 67–9.

77. Linn JG, Martin NP. Forage quality tests and interpretations. *Vet Clin North Am Food Anim Pract.* (1999) 7:509–23. doi: 10.1016/S0749-0720(15)30790-8

78. Hackmann TJ, Sampson JD, Spain JN. Comparing relative feed value with degradation parameters of grass and legume forages. *J Anim Sci.* (2008) 86:2344–56. doi: 10.2527/jas.2007-0545

79. Suhaimi D, Sharif S, Normah MA, Norain NM, Wan SH. Estimating relative feed value of local Brachiaria decumbens. *Malaysian J Vet Res.* (2017) 8:78–82. Available online at: http://www.dvs.gov.my/dvs/resources/u

80. Cinar S, Hatipoglu R. Quality characteristics of the mixtures of some warm season perennial grasses with alfalfa (*Medicago sativa* L.) under irrigated conditions of Cukurova. *Turk. J. Field Crops.* (2015) 20:31–37. doi: 10.17557/.36336

81. De Ley Coss A, Guerra-Medina C, Montañez-Valdez O, Guevara HF, Pinto RR, Reyes-Gutiérrez J. *In vitro* production of gas methane by tropical grasses. *Rev. MVZ Córdoba*. (2018) 23:6788–98. doi: 10.1007/s11250-020-02393-5

82. Hungate RE. *Variations in the Rumen, The Rumen and Its Microbes*. Cambridge, MA: Academic Press (1966). p. 376–418. doi: 10.1016/B978-1-4832-3308-6.50013-9

83. Anele UY, Südekum KH, Hummel J, Arigbede OM, Oni AO, Olanite JA, et al. Chemical characterization, *in vitro* dry matter and ruminal crude protein degradability and microbial protein synthesis of some cowpea (*Vigna unguiculata* L. Walp) haulm varieties. *Anim Feed Sci Technol.* (2011) 163:161–9. doi: 10.1016/j.anifeedsci.2010.11.005

84. Harrison JH, Blauwiekel R, Stokes MR. Symposium: Utilization of Grass Silage- Fermentation and Utilization of Grass Silage. *J Dairy Sci.* (1994) 77:3209–35. doi: 10.3168/jds.S0022-0302(94)77264-7

85. Pereira OG, Rocha KD, Ferreira CL. Chemical composition, characterization, and population of microorganisms on elephantgrass "Cameroon" (*Pennisetum purpureum*, Schum) and its silages. *R Bras Zootec.* (2007) 36:1742–50. doi: 10.1590/S1516-35982007000800006

86. Santos EM, Pereira OG, Garcia R, Lucas CL, Ferreira F, Oliveira JS, et al. Microbial populations, fermentative profile and chemical composition of signal grass silages at different regrowth ages. *R Bras Zootec.* (2011) 40:747–55. doi: 10.1590/S1516-35982011000400007

87. Pitt RE. "The biology of silage preservation," In: *Silage and Hay Preservation* Ithaca, NY: National Resource, Agriculture, and Engineering Service (1990). p. 5–20.

88. Pinho MA, Santos EM, Carvalho GP, Silva PG, Silva TC, Campos FS, et al. Microbial and fermentation profiles, losses and chemical composition of silages of buffel grass harvested at different cutting heights. *R Bras Zootec.* (2013) 42:850–6. doi: 10.1590/S1516-35982013001200003

89. Li D, Ni K, Zhang Y, Lin Y, Yang F. Fermentation characteristics, chemical composition and microbial community of tropical forage silage under different temperatures. *Asian-Australas J Anim Sci.* (2019) 32:665–74. doi: 10.5713/ajas.18.0085

90. Yahaya MS, Goto M, Yimiti W, Smerjail B, Kawamoto Y. Evaluation of fermentation quality of a tropical and temperate forage crops ensiled with additives of fermented juice of epiphytic lactic acid bacteria (FJLB). *Asian-Aust J Anim Sci.* (2004) 17:942–6. doi: 10.5713/AJAS.2004.942

91. Arroquy JI, Cornacchione MV, Colombatto D, Kunst JC. Chemical composition and *in vitro* ruminal degradation of hay and silage from tropical grasses. *Can J Anim Sci.* (2014) 94:705–15. doi: 10.4141/cjas-2014-014

92. Vendramini MB, Desogan AA, Silveira ML, Sollenberger LE, Queiroz OC, Anderson WF. Nutritive value and fermentation parameter of warm-season grass silage. *Professional Anim. Scientist*. (2010) 26:193–200. doi: 10.15232/S1080-7446(15)30580-5

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 07 February 2023
DOI 10.3389/fnut.2023.1099004

Check for updates

OPEN ACCESS

EDITED BY
Sapna Langyan,
National Bureau of Plant Genetic Resources (ICAR), India

REVIEWED BY
Satinder Kaur,
Punjab Agricultural University, India
Mahendar Thudi,
Dr. Rajendra Prasad Central Agricultural University, India

*CORRESPONDENCE
Muraleedhar S. Aski
✉ murali2416@gmail.com
Gyan Prakash Mishra
✉ gyan.gene@gmail.com
Harsh Kumar Dikshit
✉ harshgeneticsiari@gmail.com

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

RECEIVED 15 November 2022
ACCEPTED 16 January 2023
PUBLISHED 07 February 2023

CITATION
Sinha MK, Aski MS, Mishra GP, Kumar MBA, Yadav PS, Tokas JP, Gupta S, Pratap A, Kumar S, Nair RM, Schafleitner R and Dikshit HK (2023) Genome wide association analysis for grain micronutrients and anti-nutritional traits in mungbean [*Vigna radiata* (L.) R. Wilczek] using SNP markers.
Front. Nutr. 10:1099004.
doi: 10.3389/fnut.2023.1099004

COPYRIGHT
© 2023 Sinha, Aski, Mishra, Kumar, Yadav, Tokas, Gupta, Pratap, Kumar, Nair, Schafleitner and Dikshit. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Genome wide association analysis for grain micronutrients and anti-nutritional traits in mungbean [*Vigna radiata* (L.) R. Wilczek] using SNP markers

Mayank Kumar Sinha[1], Muraleedhar S. Aski[1]*, Gyan Prakash Mishra[1]*, M. B. Arun Kumar[2], Prachi S. Yadav[1], Jayanti P. Tokas[3], Sanjeev Gupta[4], Aditya Pratap[5], Shiv Kumar[6], Ramakrishnan M. Nair[7], Roland Schafleitner[8] and Harsh Kumar Dikshit[1]*

[1]Division of Genetics, ICAR - Indian Council of Agricultural Research– Indian Agricultural Research Institute, New Delhi, India, [2]Division of Seed Science and Technology, ICAR – Indian Agricultural Research Institute, New Delhi, India, [3]Division of Biochemistry, Chaudhary Charan Singh Haryana Agricultural University, Hissar, India, [4]Krishi Bhavan, Indian Council of Agricultural Research, New Delhi, India, [5]Division of Crop Improvement, ICAR – Indian Institute of Pulses Research, Kanpur, India, [6]International Center for Agricultural Research in the Dry Areas (ICARDA), New Delhi, India, [7]World Vegetable Center, South and Central Asia, Hyderabad, India, [8]World Vegetable Center Head Quarters (HQ), Tainan, Taiwan

Mungbean is an important food grain legume for human nutrition and nutritional food due to its nutrient-dense seed, liked palatability, and high digestibility. However, anti-nutritional factors pose a significant risk to improving nutritional quality for bio-fortification. In the present study, genetic architecture of grain micronutrients (grain iron and zinc concentration) and anti-nutritional factors (grain phytic acid and tannin content) in association mapping panel of 145 diverse mungbean were evaluated. Based on all four parameters genotypes PUSA 1333 and IPM 02-19 were observed as desired genotypes as they had high grain iron and zinc concentration but low grain phytic acid and tannin content. The next generation sequencing (NGS)-based genotyping by sequencing (GBS) identified 14,447 genome-wide SNPs in a diverse selected panel of 127 mungbean genotypes. Population admixture analysis revealed the presence of four different ancestries among the genotypes and LD decay of ∼57.6 kb kb physical distance was noted in mungbean chromosomes. Association mapping analysis revealed that a total of 20 significant SNPs were shared by both GLM and Blink models associated with grain micronutrient and anti-nutritional factor traits, with Blink model identifying 35 putative SNPs. Further, this study identified the 185 putative candidate genes. Including potential candidate genes *Vradi07g30190*, *Vradi01g09630*, and *Vradi09g05450* were found to be associated with grain iron concentration, *Vradi10g04830* with grain zinc concentration, *Vradi08g09870* and *Vradi01g11110* with grain phytic acid content and *Vradi04g11580* and *Vradi06g15090* with grain tannin content. Moreover, two genes *Vradi07g15310* and *Vradi09g05480* showed significant variation in protein structure between native and mutated versions. The identified SNPs and candidate genes are potential powerful tools

to provide the essential information for genetic studies and marker-assisted breeding program for nutritional improvement in mungbean.

KEYWORDS

micronutrients, marker trait association, anti-nutrients, bio-fortification, tannins

Introduction

Mungbean is one of nearly 150 species in the *Vigna* genus, with 22 endemic to India and 16 to Southeast Asia. Africa, on the other hand, is home to the most species. Mungbean [*Vigna radiata* (L.) Wilczek] is a diploid legume with a genome size of 0.60 pg/1C (579 Mbp) and a genome size of 0.60 pg/1C (579 Mbp) (1). Mungbean is a warm-season legume that grows between 40 and 10 degrees north in the tropics and subtropics. India, China, Pakistan, Bangladesh, Sri Lanka, Thailand, Myanmar, Vietnam, Indonesia, Australia, and the Philippines are the top mungbean producers (2). India is the world leader in mungbean production, with 4.53 million hectares yielding 2.08 million tons of grain (AICRP on MULLaRP PC Report 2020–21).

The recommended dietary allowance (RDA), which for adult women is roughly 0.06 g day^{-1} with a low-iron-bioavailability (5%) diet and 0.02 g day^{-1} with a high-iron-bioavailability (15%) diet, is not met by a sizeable fraction of the population in underdeveloped nations (3). The most vulnerable are women of reproductive age and children. Anemia affects around 88 percent of pregnant women and 63 percent of children aged 5–14 years in South Asia [ACC/SCN, 2000; (4)]. Micronutrient deficiency in humans is referred to as "hidden hunger." Bio-fortification is a genetic enhancement method that boosts mineral absorption while lowering anti-nutritive components and balances mineral concentrations in edible plant parts and seeds (5, 6). Iron deficiency is the most prevalent micronutrient problem worldwide. Iron deficiency reduces the amount of oxygen delivered to cells, resulting in tiredness, poor work performance, lowered immunity, and mortality (7).

Zinc is an essential nutrient for plant growth and production due to its participation in over 300 enzymes involved in the metabolism of glucose, DNA, protein synthesis and digestion, and bone development. Zinc deficiency can lead to stunted growth, skin blemishes, and an increased risk of infection (8). For male adults, the RDA for zinc is 0.011 g per day, while for female adults, it is 0.008 g per day.[1] A typical main organic form of phosphorus (P) storage in plants is phytic acid (PA) (chemically called as myo-inositol hexaphosphoric acid). Because it is an effective chelator of positively charged cations, PA binds to nutritionally important mineral cations such as calcium, iron, and zinc, as well as inhibiting trypsin (9). Furthermore, humans and other non-ruminants lack the phytase enzyme, which prohibits them from digesting PA and excreting a large part of these salts.

Tannins are polyphenolic compounds with a molecular weight more than 500 kD that can form a complex with proteins (10–12). Tannins are classified into two groups based on their structure: condensed tannins and hydrolysable tannins. The majority of tannins are found in the seed coat, with only a few residues in the cotyledons (13, 14). Condensed tannins seem to bind proteins very tightly, lowering protein digestibility in pulses *in vitro* (15, 16). Tannins have the potential to bind with proteins and prevent them from being absorbed by the body. Because tannins are extremely reactive, processing alters their profiles and amounts in foods, potentially affecting their anti-oxidant activity and nutritional value (17).

Quantitative trait loci (QTL) analyses through genome-wide association studies (GWAS) using molecular markers and high throughput sequencing techniques can be used to identify the genes underlying nutritional (grain minerals, protein content, and anti-oxidant capacity) and anti-nutritional (phytic acid and tannins) traits (18). The GWAS studies have been reported in many other legume crops also (19). The two major methods for finding genes or QTLs are linkage mapping and association mapping based linkage disequilibrium (LD). The capacity to examine only two alleles at any given locus in biparental crosses and low mapping resolution are two major constraints of linkage mapping (20), whereas association mapping promises to overcome these limitations (21). Furthermore, association mapping uncovers QTLs, allowing to take advantage of natural variation and find beneficial genes in the genome using modern genetic tools.

In the present experiment, we have utilized single nucleotide polymorphism (SNP) markers discovered by sequencing of diverse mungbean germplasm and used to test nutritional and anti-nutritional attributes to characterize the diverse mungbean panel for grain micronutrient (Fe, Zn) concentration and anti-nutritional factors (Phytic acid, Tannins). To study the genetic diversity among mungbean genotypes using SNP markers. And to identify the linked SNPs with grain micronutrient concentration, tannins and phytic acid content using association mapping approach.

Materials and methods

Plant materials and experimental conditions

The grain micronutrient concentrations (Fe, Zn) and anti-nutritional components (Phytic acid, Tannins) were investigated using 145 different mungbean genotypes and 5 checks (**Supplementary Table 1**), which included released varieties, advanced breeding lines, and exotic germplasm lines from the World Vegetable Center. The experiment was carried out at the Indian Agricultural Research Institute (IARI) in New Delhi at the experimental field (28.638690, 77.156046). The genotypes were planted in the Augmented block design (22) with five checks in each 5 blocks. However, only one hundred and twenty-seven lines were selected for association mapping, as shown in **Supplementary Table 2**.

Iron and zinc estimation

The hand-threshed grains were then carefully placed in a clean plastic tray (using contaminant-free gloves). Each sample's grains were rinsed in double distilled water and dried for 5 days at 35°C

1 http://ods.od.nih.gov/factsheets/zinc-HealthProfessional/

in a contamination-free, non-corroded oven. Using a mortar and pestle, 10 g of grains from each sample were manually ground into a fine powder. A microwave digestion device was used to digest the grain powder sample (1 g) according to the modified diacid technique (23). The Fe and Zn concentrations in three technical replications per biological sample (in mg/kg seed) were determined using an AAS (Atomic Absorption Spectrophotometer) (ElementAS, Electronics Corporation of India Ltd., Model- AAS4141).

Phytic acid estimation

Megazyme Phytic acid assay kit (2022 Megazyme Ltd.) was used to estimate phytic acid content. The total phosphate emitted is calculated as grams of phosphorus per 100 g of sample material using a modified colorimetric approach. Using a spectrophotometer with a wavelength of 655 nm, the total phosphate emitted during the procedure was measured.

Estimation of condensed tannin (pro-anthocyanidin) content (CTC)

The approach provided by de Camargo et al. (24) was used to evaluate condensed tannins in lentils.

Statistical analysis

The data recorded on investigated traits were subjected to descriptive statistics including mean, standard deviation, range, coefficient of variation, and broad sense heritability. The analysis of variance (ANOVA), Pearson's correlation analysis, cluster analysis and principal component analysis for recoded data were performed in R (Version 4.1.2). Frequency distribution graphs of recorded traits were developed using MS-EXCEL program. The ANOVA was carried out for the Augmented Block Design Using R agricolae package 1.4.0.

Genomic DNA extraction, purification, and quantification

According to Doyle and Doyle (25), genomic DNA was isolated from immature mungbean leaves using the CTAB (CetylTrimethyl Ammonium Bromide) method. The DNA was quantified using an agarose gel electrophoresis method. This was accomplished by dissolving 0.8 g of agarose powder in 100 ml of 1X TBE buffer to obtain a 0.8 percent final concentration of agarose. The DNA integrity was further evaluated using a double beam UV spectrophotometer at 260 and 280 nm wavelengths. The concentration of DNA was calculated using the following formula:

$$\text{Concentration of ds DNA} = 50\mu g/ml \times OD(260) \times \text{dilution factor.}$$

The protein contamination of samples was determined using the optical density (OD) ratio at 260/280 nm. If the ratio is less than 1.8, the sample has been contaminated with protein. The ratio of pure DNA is 1.8.

Genome wide discovery of GBS based SNP markers

127 mungbean genotypes of AM panel, used for molecular characterization and association mapping (AM) (**Supplementary Table 2**).

Then samples are digested with APeKI restriction enzyme and ligated to adaptors with unique barcodes to create 127-plex GBS libraries, then pooled (26). The generated libraries were single end sequenced (150 bp) using the Illumina HiSeq 4000 NGS platform as per Bastien et al. (27) and Kujur et al. (28). The GBS assay's repeatability was tested using a non-template control and biological duplicates with three accessions. For quality assessment of sequence reads, the resulting FASTQ sequence files were processed using the STACKS v1.0[2] sliding window technique (29). Sequence readings with a quality of 90% below confidence were eliminated, as were sequence reads with long decreases in quality (30). The FASTQ sequence reads were mapped and aligned to the mungbean reference genome (31) using Bowtie v2.1.0 after demultiplexing using unique barcodes (32). Furthermore, the SNPs were accurately identified by processing the resulting SAM (sequence alignment map) files of 127 genotypes utilizing a reference based GBS technique of the STACKS v1.0 approach. The STACKS *de novo* based GBS technique was used to process the unmapped sequence reads on the reference genome yet again. The sampled SNPs were reconstructed into STACKS from sequencing reads of each genotype for the detection of probable SNPs, as described by Kujur et al. (28) and Hohenlohe et al. (33). Structure and functional annotation were carried out according to the mungbean genome (31) annotation to determine the precise position of GBS-based SNPs in different variations of the genome.

Molecular diversity and population structure

The indices representing molecular diversity including $\theta\pi$ (nucleotide diversity based on substitution of nucleotide in any two randomly selected DNA sequences at a particular site), $\theta\omega$ (Watterson's estimator of segregating sites, based on mutation rate estimates from loci that are segregating in population) and Tajima's D (to test the null hypothesis of selective neutrality within the population) were estimated using a TASSEL v5.0 sliding window approach as suggested by Xu et al. (34) and Varshney et al. (35). The population structure among the 127 genotypes was identified by evaluating the obtained SNP data with ADMIXTURE version 1.3.0's model-based program, utilizing (36) method. Furthermore, the *ad-hoc*, delta K technique was used to calculate the optimum population number (K) value, as described by Evanno et al. (37). The collected SNP data of 127 genotypes were examined using TASSEL v5.0 (38) software to construct an unrooted neighbor-joining (NJ)-based phylogenetic tree (with 1,000 bootstrap replicates).

Linkage disequilibrium (LD) measurement

The correlation between pairs of SNP sites on the chromosome is mostly determined by LD (39). As a result, the correct measurement of LD is the square of correlation (r^2) between pairs of alleles (40). The degree of LD and its degradation in the population largely affected the identification of markers connected to trait loci and the resolution of association analyses (41). Several statistics for LD assessment were derived based on the effect of sample size and marginal allelic frequencies (42). To assess patterns of LD (r^2) and LD decay, the produced SNP data were analyzed using TASSEL v5.0 (sliding window technique) and R (Version 4.1.2) in the current study [following Remington et al. (43)].

2 http://creskolab.uoregon.edu/stacks

Association mapping (AM) of investigated traits

Using the HapMap file including genotypic data for 127 different genotypes as well as phenotypic data, GAPIT (Genomic Association and Prediction Integrated Tool) was utilized to run a GWAS on seed Iron and Zinc concentration as well as Phytic acid and Tannin content. GAPIT (version 3) was used to conduct GAWAS, which used MLM and BLINK models. BLINK (Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway) is a Genome Wide Association Study (GWAS) Method (44). (BLINK). BLINK stands for "Fixed and random model Circulating Probability Unification" and is an improved version of the FarmCPU GWAS approach. BLINK uses a multi-locus model for evaluating markers across the genome, similar to the Multi-loci Mixed Linear Model (MLMM). BLINK iteratively runs two fixed effect models. To account for population stratification, one model tests each marker one at a time, with many associated markers fitted as covariates. The other model uses covariate markers instead of kinship to directly control spurious association, removing the confounding between testing marker and kinship. To boost statistical power, BLINK eliminates the requirement that genes underlying a characteristic be scattered evenly across the genome. To improve processing speed, BLINK substitutes the REstricted Maximum Likelihood (REML) in a mixed linear model with Bayesian Information Content (BIC) in a fixed effect model in FarmCPU. The first three main components produced from all of the markers, as well as the origin-group, are included in the covariate variables. To eliminate linear dependency, the origin group was coded as an indicator (0/1) for each of the origin groups except the last one. The default GAPIT settings were utilized, as well as a Bayesian Information Criterion (BIC) model selection, which determines the degree of population structure that should be accounted for in a model to minimize overfitting. According to the BIC analysis, none of the models required the use of PCs. To account for population stratification, a mixed linear model with a kinship matrix was chosen for analysis. The following formula was used to fit the mixed linear models:

$$y = X\beta + Z\mu + e$$

According to the GAPIT user manual, y is a vector of observed phenotypes, b is an unknown vector containing fixed effects that account for the genetic marker, population structure (Q), and intercept, u is an unknown vector of random additive effects from background QTLs and individuals, X and Z are the known design matrices, and e is an unobserved vector of residuals. For GWAS of nutritional traits, the Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (BLINK) method was used because it has high statistical power and does not assume that causal genes are distributed normally across the genome, which can lead to false positives and exclusion of causal genes (45). Only the most important markers are reported since BLINK utilizes BIC to eliminate markers based on linkage disequilibrium (LD) (45). The following formula was used to suit the BLINK models:

$$y = s_i + S + e$$

Where y is a vector of observed phenotypes; s_i is a testing marker; S is a pseudo quantitative trait nucleotide (QTN), and e is the unobserved vector of residuals according to the GAPIT user manual. A Bonferroni correction was used to avoid false positives and identify significant SNPs ($a^{1/4}0.05$) for each trait. The Bonferroni correction was calculated as $-\log_{10}$ (0.05/n), where n equals the number of SNPs used in the GWAS for each trait.

The threshold of significance, threshold probability of –log 10 (p-value) > 3.0 was used as cut off to identify significant markers associated with grain iron and grain zinc concentration while threshold probability of –log 10 (p-value) > 4.0 was used as cut off to identify significant markers associated with grain phytic acid and grain tannin content. For multiple comparisons, the significance threshold of the adjusted p-value was corrected according to the false discovery rate (FDR) with cut off $\leq$ 0.05 (46). The p-value distribution of significant SNP markers related with examined attributes was depicted using Manhattan plots. The adequacy of controlling type I error was assessed by plotting observed and expected –log10 (p) values using quantile-quantile (Q-Q) plots following Diapari et al. (47).

Delineation of putative candidate genes for investigated traits

Initially, genes with SNP variations were identified by functional annotation with the mungbean reference genome to identify the likely candidate genes affecting the characteristics in the study (31). A window of 57,679 kb in the vicinity of SNPs in the genomic area was investigated to identify candidate genes affecting the attributes. The interval sequences were retrieved and mapped on the mungbean genome using the mungbean reference genome (31), and the candidate genes were found using the reference genome location generated by blast. The SNPs within respective LD decay range of respective chromosome were considered as the same locus and SNP sites were considered as significantly associated. Following on, a legume information system[3] was used to retrieve the CDS sequences of all protein-coding genes. SNPs found in the CDS section of prospective candidate genes were evaluated for type of SNP using TASSEL software, and the matching CDS sequences were translated to their expressed amino acid sequences using the EXPASY website's facilities.[4] In the I-TASSER platform,[5] the differing amino acid sequences acquired from EXPASY were utilized to estimate protein structure (48).

Results

The ANOVA augmented block design revealed the presence of highly significant variation among the genotypes for all tested traits (**Supplementary Table 3**). The study revealed the highly significant interaction between iron and phytic acid content being negatively correlated. Among the studied traits, the coefficient of variation was 21.23% for grain iron concentration, 29.90% for grain zinc concentration, 28.80 for grain phytic acid content and 26% for grain tannin content. The mean values obtained for iron was 74.15 mg Kg^{-1}, for zinc it was 32.20 mg Kg^{-1}, for phytic acid it was 7.35 mg g^{-1} and for tannins the value was 3.8 g 100 g^{-1}) (**Supplementary Table 4**).

3 https://legacy.legumeinfo.org/genomes/jbrowse/?data=Vr1.0
4 https://web.expasy.org/translate/
5 https://zhanggroup.org/I-TASSER/

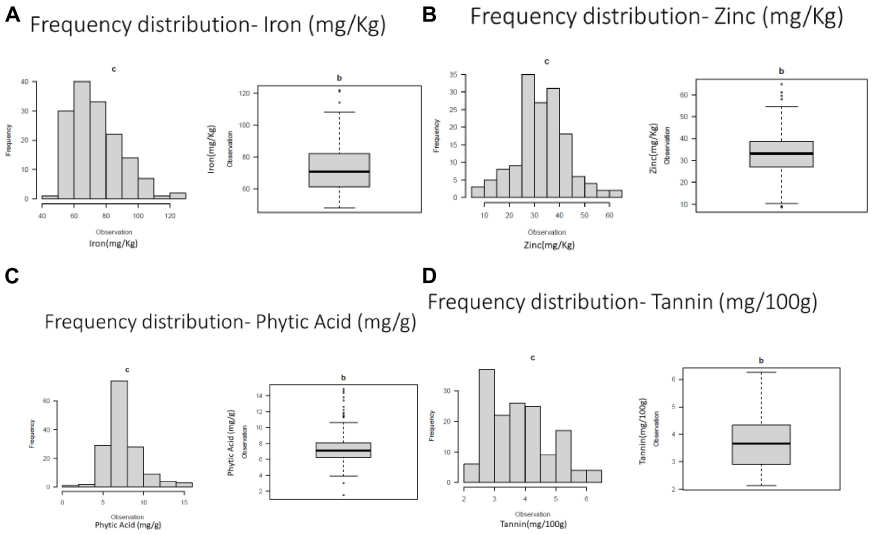


FIGURE 1
Frequency distribution of variation for **(A)** grain iron concentration **(B)** grain zinc concentration **(C)** grain phytic acid content and **(D)** grain tannin content.

The genotypes showed variation for iron content in the range of 48.2–121.85 mg Kg^{-1} where genotypes GANGA 8 (121.85 mg Kg^{-1}), ML 818 (121.20 mg Kg^{-1}), KM 16–69 (114.20 mg Kg^{-1}) showed highest iron content. The zinc content for which the values ranged from 8.6 to 61.05 mg Kg^{-1}. The genotypes BASANTI (61.05 mg Kg^{-1}), IC 325828 (59.45 mg Kg^{-1}), KM 16–75 (52.25 mg Kg^{-1}) showed highest zinc content. The phytic acid values ranged from 1.5 to 14.85 mg g^{-1} in the genotypes. Some of the lowest phytic acid values were observed in the genotypes IPM 02–19 (1.5 mg g^{-1}), GANGA 8 (3 mg g^{-1}), PUSA 1333 (3.8 mg g^{-1}), M1209 (4.19 mg g^{-1}), IPM 02–14 (4.4 mg g^{-1}), MH 1442 (4.46 mg g^{-1}), and IC 436637 (4.7 mg g^{-1}). Considerable variations were also seen in the tannin content in the studied genotypes which varied from 2.14 g 100 g^{-1} to 6.25 g 100 g^{-1}. IPM 288 (2.14 g 100 g^{-1}), PUSA 1333 (2.33 g 100 g^{-1}), KM 2241 (2.33 g 100 g^{-1}) showed the lowest values for the tannin content (Supplementary Table 5).

Frequency distribution, genetic correlations, principal component analysis, and cluster analysis

Frequency distribution of variation for studies traits were presented in Figure 1.

Pearson's correlation coefficients indicated significant negative correlations between grain iron concentration and grain phytic acid content. While non-significant negative correlation was observed between grain zinc concentration and grain phytic acid content and grain phytic acid content and grain tannin content (Supplementary Table 6 and Figure 2).

Principal component analysis (PCA) was carried out to identify the most contributing traits of variation in the studied genotypes. The first principal component explained 32.95% of total variation, the second principal component explained 27.60% of total variation, the third principal component explained 21.9% of total variation and the fourth principal component explained 17.55 and all totaling 100% of total variation (Supplementary Table 7 and Figure 3).

The genetic diversity among mungbean genotypes using SNP markers

The reference genome based GBS approach resulted in the detection of 14,447 high quality SNPs with a read depth of 10, 0% missing data, 10% heterozygosity, and 1% minor allele frequency (MAF).

Among all the 11 chromosomes, the maximum number of SNPs was observed on chromosome 1 (2,437 SNPs), whereas the minimum number of SNPs was observed on chromosome 3 (652 SNPs). The average SNP density (SNPs per 50 kb) was found to be high on chromosome 1 (3.33) and low on chromosome 5 (1.5) (Table 1).

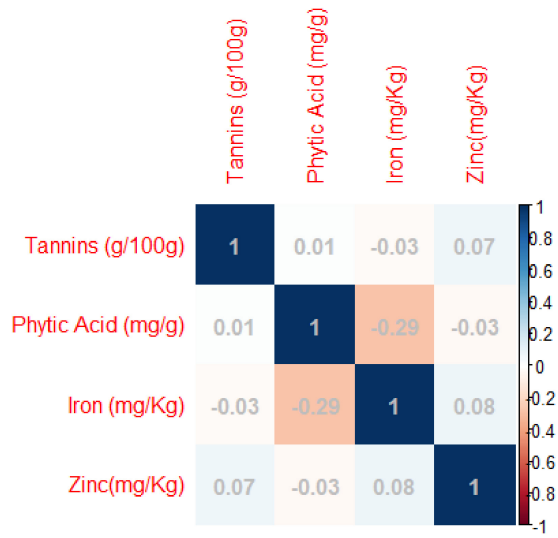


FIGURE 2
Pearson's correlation coefficients between investigated traits-Diagrammatic view PC1 = 32.95%, PC2 = 27.60%.

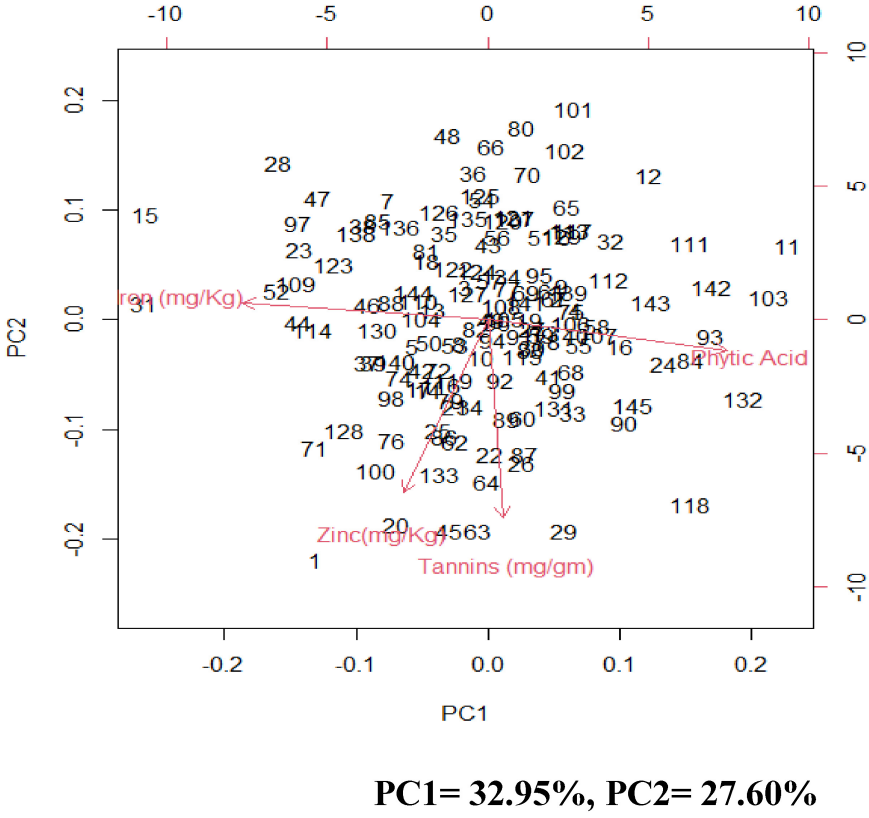


FIGURE 3
Biplots of the first two principal components (PC) showing variation among grain iron concentration, grain zinc concentration, grain phytic acid content, and grain tannin content.

Molecular diversity and population structure analysis

ADMIXTURE version 1.3.0. software (49) produces a Q matrix containing estimates of ancestry for each individual tested. The corresponding Q matrix with the lowest cross-validation error was chosen as the most representative of the study population, which was at $K^{1/4}4$, corresponding to 4 distinct subpopulations (Figure 4 and Supplementary Figure 1).

The Q matrix was then sorted by the ancestry coefficients for each subpopulation, assigning individuals with coefficients > 50% to the corresponding sub population (50). These grouping of genotypes into three subpopulations were further confirmed by the distinct differentiation of genotypes into four clusters by unrooted neighbor-joining phylogenetic tree construction (Supplementary Figure 2).

TABLE 1 Details of the number of SNPs and their distribution on 11 mungbean chromosomes.

Chromosome	Chromosome size (bp)	SNPs per chromosome	Average density (SNPs per 50 kb)
Chromosome 1	36,501,346	2,437	3.34
Chromosome 2	25,360,630	1,199	2.36
Chromosome 3	12,950,713	652	2.52
Chromosome 4	20,812,224	1,015	2.44
Chromosome 5	37,180,910	1,124	1.51
Chromosome 6	37,436,759	1,527	2.04
Chromosome 7	55,601,358	2,030	1.83
Chromosome 8	45,727,239	1,702	1.86
Chromosome 9	21,008,463	1,119	2.66
Chromosome 10	20,996,616	815	1.94
Chromosome 11	19,732,206	827	2.10
Total	**333,308,464**	**14,447**	**2.17**

Linkage disequilibrium analysis

The identified 14,447 SNPs were analyzed to estimate the LD patterns (r^2) and LD decay extent across 11 chromosomes of mungbean. The LD patterns in a population of 127 AM panel genotypes showed that the LD decay was to be between the physical distances of 0–100 kb (around 57 kb) in mungbean chromosomes (Figure 5). The high resolution LD patterns resulting from a large number of SNP markers facilitate the higher mapping resolution in marker trait association analysis.

Marker trait association (MTA) analysis

The default GAPIT parameters were used, as well as a model selection with Bayesian Information Criterion (BIC), which determines the degree of population structure that should be accounted for in a model to avoid overfitting. Further, in this study, the Bonferroni correction threshold value of –log10 > 3.0

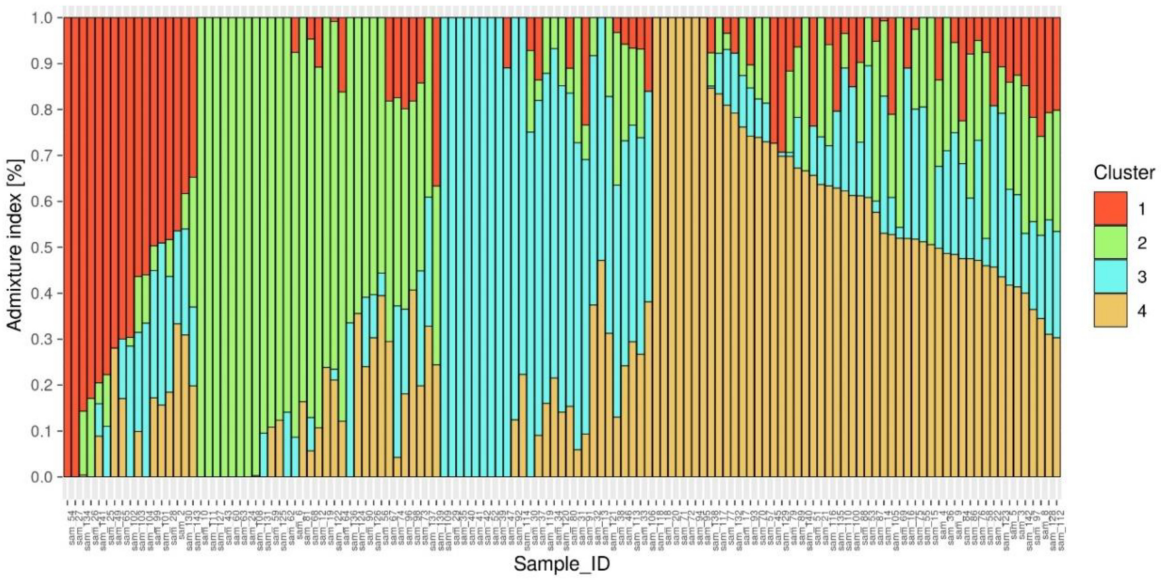


FIGURE 4
Population genetic structure plot of 127 diverse mungbean genotypes with optimal population number k = 4 with four different colors.

(p-value) was used as cut off to identify the significant SNPs associated with the grain iron and grain zinc concentration while –log10 > 4.0 (p-value) was used as cut off to identify the significant SNPs associated with for grain phytic acid and grain tannin content. The markers considered to be significantly associated with tested traits were represented by illustrating the Manhattan plots. Significant SNPs were identified from the BLINK model for (a) grain iron concentration (Figure 6) (b) grain zinc concentration (Figure 7) (c) grain phytic acid content (Figure 8) and (d) grain tannin content (Figure 9) across all chromosomes. The Fe showed strong association with SNPs present on chromosomes 1 and 9 exclusively. While, Zn exhibited significant linkage with SNPs present on chromosomes 5 and 7. In case of Phytic acid chromosome 8, having significant SNPs linked and Tannin content indicated strong association with SNPs found on chromosomes 6, 4, and 9. These SNPs were in local LD with multiple candidate genes. Summary table for studied traits and respective SNP were indicated in Table 2.

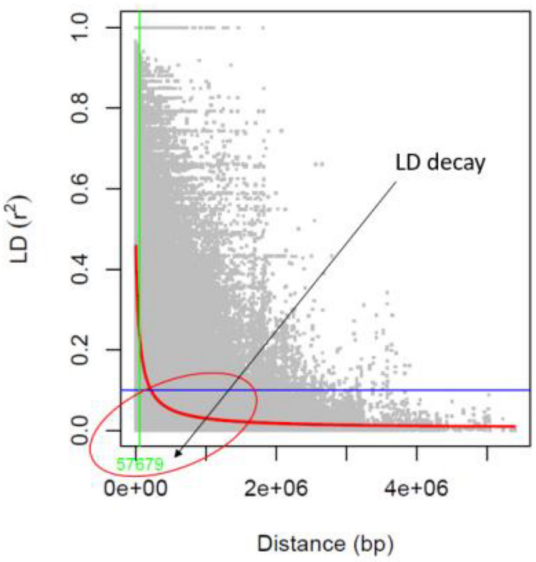


FIGURE 5
Linkage disequilibrium (LD) decay value (bp) in the association mapping panel.

Delineation of putative candidate genes

Total of 15 SNPs were found to be associated with the grain iron concentration. There were total 38 protein coding genes in the LD region of these SNPs. These genes were found to be involved in various protein formation, some of which are homeobox leucine zipper protein, WRKY family transcription factor, Pentatricopeptide repeat (PPR-like) superfamily protein, stress upregulated protein, Cytochrome P450 superfamily protein, iron ion binding and heme binding protein. Total of 10 SNPs were found to be associated with the grain zinc concentration and 59 genes were present in the haplotype of these SNPs. These were found to be associated with protein and enzyme formation like adenylate cyclase, zinc finger family protein, magnesium ion binding protein and protein kinase family protein. 5 SNPs which had 27 putative candidate genes in their haplotype were detected to be associated with grain phytic acid content (gene description are presented in Supplementary Table 10) which were found to be involved in several protein and enzyme formation with examples of Serine/threonine protein phosphatase family protein, Tubby like protein, Serine/Threonine kinase family protein and Phosphatidyl inositol kinase (PIK-G1)n. The study also found 4 SNPs having 26 genes in haplotype, associated with grain tannin content which were involved in synthesis of serine/threonine-protein phosphatase, triacylglycerol lipase, heat shock transcription factor B4, Sugar transporter SWEET n etc. proteins. *Vradi08g09870* was found to be significantly associated with grain phytic acid

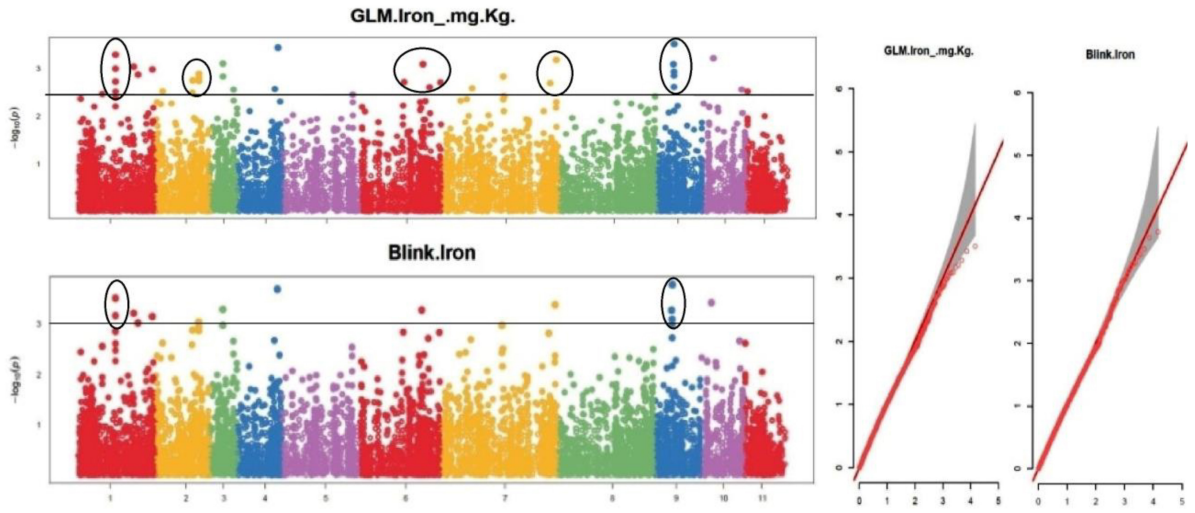


FIGURE 6
Manhattan plots and Quantile-Quantile plots depicting the significant association of SNP markers with grain iron concentration.

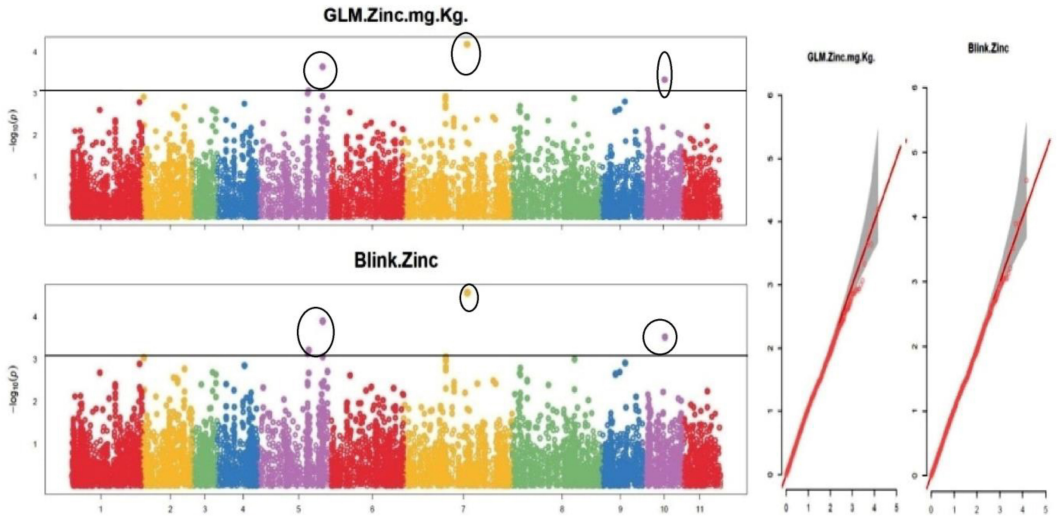


FIGURE 7
Manhattan plots and Quantile-Quantile plots depicting the significant association of SNP markers with grain zinc concentration.

content, *Vradi04g09970* was found to be significantly associated with grain iron concentration, *Vradi07g13710* was found to be significantly associated with gain zinc concentration and *Vradi06g15120* was found to be significantly associated with grain tannin content (Supplementary Tables 8–11). Furthermore, among the putative candidate genes found for grain iron concentration three genes were found to be containing missense SNP in their CDS region. These missense SNPs were found to be involved in changing of the amino acids from C/T; valine to alanine, A/C; serine to tyrosine, G/T; serine to alanine, respectively, for genes *Vradi04g09970, Vradi06g11980,* and *Vradi09g05480*. Also there was one SNP in CDS region of the gene *Vradi07g30210*, but it caused a same sense mutation resulting in no structural change. So these genes may regulate the iron concentration in the mungbean grain by affecting the protein structure at tertiary level. Circular diagram depicting a summarized view of the significant association of SNP markers with all the four traits in the study along with SNP density in the outer ring (Supplementary Figure 3).

Interestingly, a change in protein structure was observed for one of these genes *Vradi07g15310* by I-TASSER (see text footnote 5) (Figure 10). The structural analysis of these genes revealed a significant variation in native and mutated versions at protein level.

Therefore, it may be concluded that these are potential candidate genes involved in the regulation of tannin content in mungbean grains. These results showed that the identified SNPs and candidate genes are useful and worthy of being used in developing lines having low tannin content. In case of other traits, study did not observe any conformational changes with respect to CDS domain.

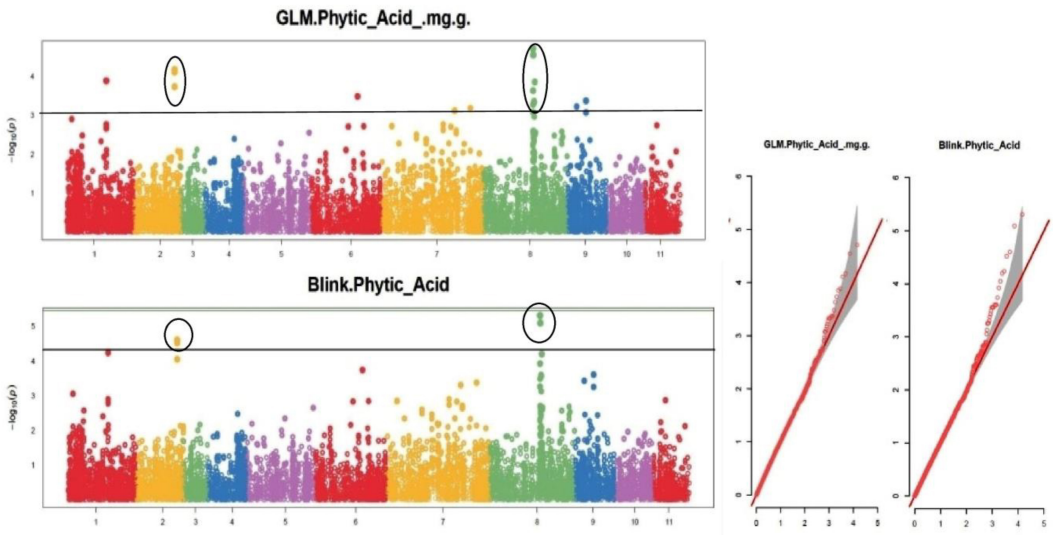


FIGURE 8
Manhattan plots and Quantile-Quantile plots depicting the significant association of SNP markers with grain phytic acid content.

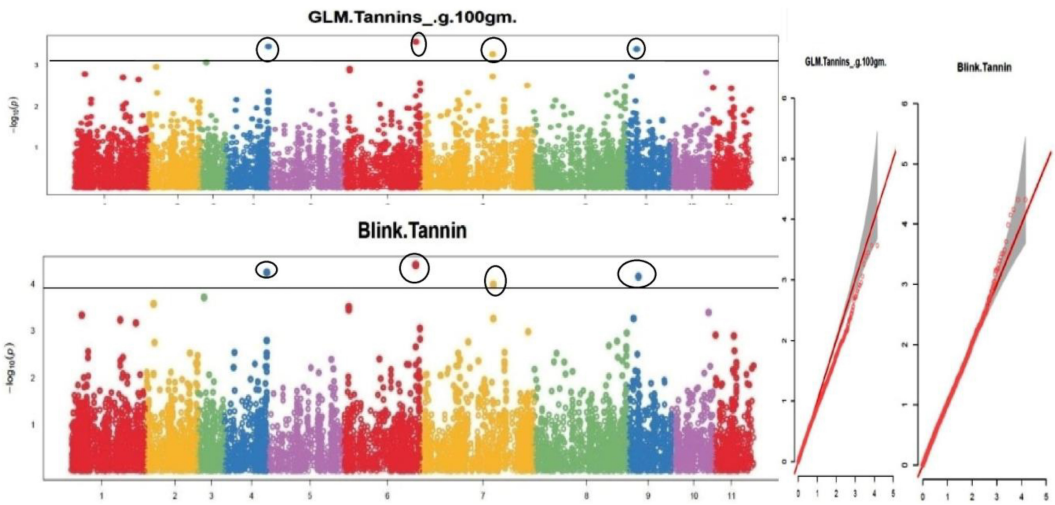


FIGURE 9
Manhattan plots and Quantile-Quantile plots depicting the significant association of SNP markers with grain tannin content.

Discussion

Mungbean is grown by resource-poor farmers since it only requires minimal irrigation and other inputs. It also replenishes soil fertility through symbiotic nitrogen fixation, is a drought-tolerant crop, and can survive high temperatures (average 35°C). Mungbean, being a substantial source of protein, is crucial for the country's vegetarian population. It contains 25–31% of crude protein (51), 4–6 mg/100 g of iron (52), 355–375 Kcal /100 g of energy, and 1–5% crude fiber (53). Considering mungbean is a nutrient-dense legume with the potential to be mineral-dense, it really is a significant issue. Iron and zinc are vital minerals, and anemia caused by a lack of iron is a big issue. The results in NFHS 2019–21, the fifth in the series, show that across all age groups, children aged 6–59 months experienced the greatest increase in anemia, rising to 67.1% (NFHS-5) from 58.6% (NFHS-4, 2015–16). According to the information, the number was larger in rural India (68.3%) than in urban India (64.2%). Anemia affects 59.1% of females aged 15–19 years (NFHS-5), up from 54.1 percent in the previous year (NFHS-4). In this group as well, rural India had a higher percentage (58.7%) than urban India (54.1 per cent). 52.2% of pregnant women aged 15–49 years were found to be anemic, up from 50.4% in the previous survey. However, there is a significant gap between metropolitan areas (45.7%) and countryside India in this group (54.3%). According to the NFHS-5 data, 35.5% of kids under the age of five are stunted (height-for-age), compared to 38.8% in the NFHS-4. In 2007, a survey was done in Hisar-1 and Barwala block of Haryana state to determine the incidence of iron deficiency anemia and its relationship to dietary intake patterns of local communities (54). In these areas, 58% of the schoolchildren were anemic, with 49% of them lacking sufficient iron. The food quality in these areas was poor, with low iron bioavailability

TABLE 2 Details of SNPs on different chromosomes and their corresponding putative genes associated with traits studied.

Sl. No.	Trait	Chromosome	No. of SNP	Blink model	GLM model			No. of putative candidate genes
				Log_{-10P} range	(Log_{-10P}) range	R square range	Phenotypic variation explained (%)	
1	Grain iron	1	6	3.019–3.510	2.8674–3.2844	0.118–0.12656	11.8–12.6	18
2	Grain iron	2	1	3.038	2.883721	0.111	11.1	1
3	Grain iron	3	1	3.288	3.09	0.119	11.9	12
4	Grain iron	4	1	3.688	3.43155	0.132	13.2	1
5	Grain iron	6	1	3.273	3.08	0.122	12.2	5
6	Grain iron	7	1	3.380	3.17	0.122	12.2	11
7	Grain iron	9	3	3.090–3.7789	2.298–3.506	0.1131–0.13506	11.31–13.50	28
8	Grain iron	10	1	3.420	3.209	0.124	12.40	3
9	Grain Zn	2	1	3.033	2.910	0.102	10.20	12
10	Grain Zn	5	5	3.89–3.050	2.93–3.644	0.1025–0.1296	10.25–12.96	37
11	Grain Zn	7	3	3.050–4.568	2.929–4.186	0.1024–0.1515	10.24–15.15	1
12	Grain Zn	10	1	3.520	3.332	0.118	11.8	9
13	Grain PA	1	1	4.522	3.885	0.154	15.4	9
14	Grain PA	2	2	4.601	4.110–4.170	0.163–0.165	16.30–16.50	14
15	Grain PA	8	2	5.090–5.306	4.54–4.70	0.1805–0.1870	18.05–18.70	4
16	Grain TAN	4	1	4.243	3.451	0.120	12	6
17	Grain TAN	6	2	4.404	3.567	0.124	12.40	20
18	Grain TAN	7	1	4.001	3.265	0.113	11.30	1
19	Grain TAN	9	1	4.155	3.387	0.117	11.70	1

ranging between 3.1 and 4.6 percent, compared to healthy adult iron absorption of 10–15%.

Additionally, the discovery of QTLs/genes for grain iron and zinc concentrations, as well as grain phytic acid and grain tannin content features, allows for marker assisted selection to improve micronutrient content and bioavailability to consumers. The advancement of next-generation sequencing (NGS) technology in recent years has allowed for the effective characterization of genotypes at the molecular level (55). It also offers the greatest platform for studies such as genome-wide association mapping, which identifies SNP markers and candidate genes that are significantly related with attributes with high resolution (56).

The di-acid digestion method has been proven to be a reliable approach for determining micronutrients such as iron and zinc in organic samples (23). The Megazyme kit's phytic acid estimation also delivers an accurate measurement of phytic acid content. A significant negative connection (R^2 = −0.28) was identified between grain iron concentration and grain phytic acid content in this study. Akond et al. (57) found a similar trend in common bean. While there was absolutely no association between grain iron and grain zinc (R^2 = 0.07), there was a loose positive link between Tannins and zinc (R^2 = 0.06) and phytic acid (R^2 = 0.07) (0.01). Furthermore, there was no significant relationship between zinc and grain phytic acid levels (R^2 = −0.03). Despite the fact that (58) identified a substantial positive association between grain iron and grain zinc concentration, this investigation discovered a non-significant but positive correlation. This could be due to the diverse genotypes utilized in the study and the geographic positions where the experiments were conducted. The findings indicate that the accumulation and augmentation of one mineral has no effect on the concentration of others, and that they are inherited separately in the mungbean genome, which is consistent with Welch and Graham's (59) findings. The findings of House et al. (60) in common bean further support the lack of a link between grain zinc concentration and grain phytic acid level. Furthermore, the levels of Fe, Zn, Phytic acid, and Tannin in this investigation were comparable to prior studies (61–64).

PUSA 1333 and IPM 02–19 genotypes were identified to exhibit high grain iron and zinc concentrations while having low grain phytic acid and tannin content. GANGA 8, IC 436637, KM 16–82, MH 1442, and TM 96–2 genotypes had high grain iron content, moderate grain zinc content, and low grain phytic acid and tannin content.

The genetic diversity among mungbean genotypes using SNP markers

Molecular markers have been extensively utilized in the mungbean for molecular characterization, genetic diversity, and gene tagging (65). Molecular characterization of 127 different mungbean genotypes was performed in this study utilizing SNP markers. In other legumes like Chickpea (28), common bean (66), and in cereals like rice (67), wheat (68), and maize (69) have all employed SNP markers developed through GBS for diversity study. GBS approach is successfully used for the generation of SNP

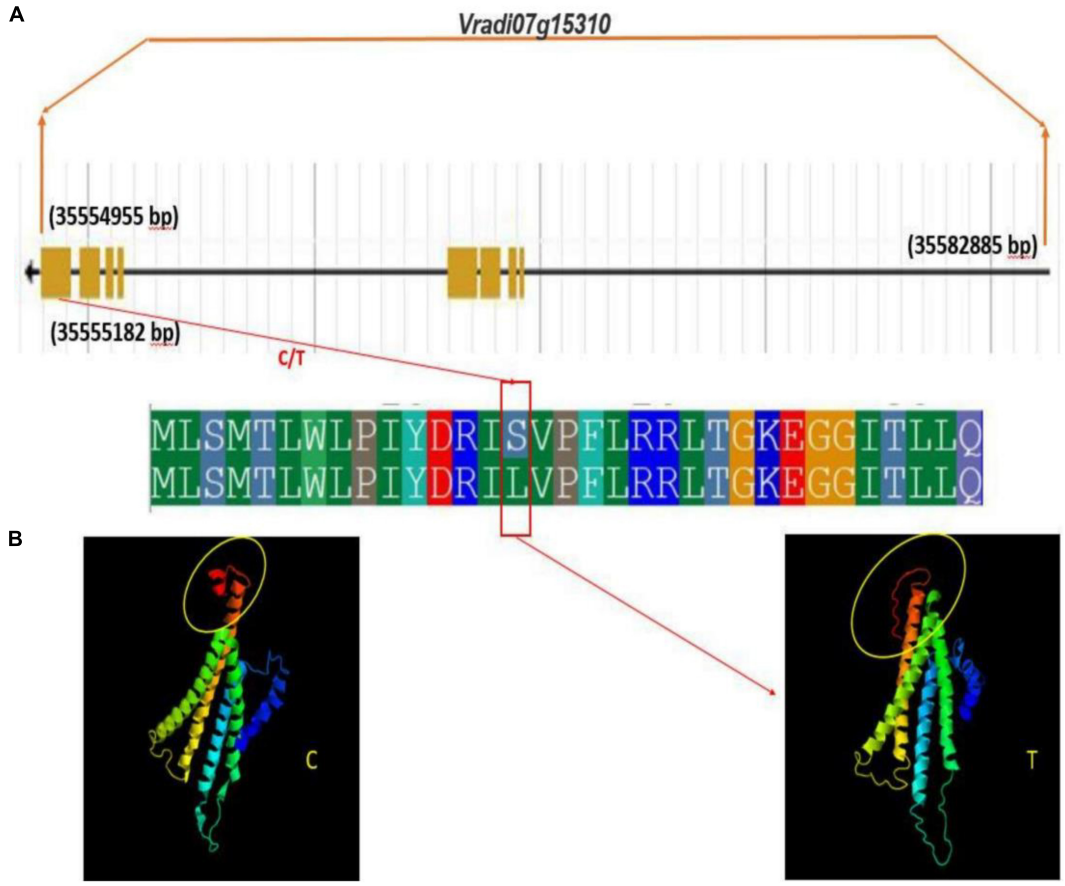


FIGURE 10
Gene and protein structure of potential candidate gene *Vradi07g15310* **(A)** gene structure **(B)** conformational changes in protein structure due to sense SNP (C/T) at the CDS domain.

markers in mungbean (70) after the development of the reference genome (31). Noble et al. (71) and Breria et al. (72) used GBS technology to develop 22,230 and 24,870 SNP markers in mungbean, respectively. A total of 14,447 high-quality and non-erroneous SNPs were generated with comprehensive genome coverage in this study. As a result, SNPs with an average sequence read length of more than 150 bp were found in longer, high-quality sequence reads in our investigation. The maximum and minimum of SNPs were observed in ours study was in line with Noble et al. (71) and Breria et al. (72), on chromosome 1 (2,437 SNPs) and chromosome 3 (652 SNPs), respectively.

The population structure study with ADMIXTURE v 1.3.0 software revealed the presence of four subpopulations in analyzed 127 genotypes. Although the phylogenetic tree built by TASSEL v5.0 software utilizing neighbor end joining strategy revealed the presence of three subpopulations, one subpopulation obtained with this approach was particularly vast and was further separated into two groups. Noble et al. (71) and Breria et al. (72) previously reported the occurrence of four subpopulations among the 466 different mungbean accessions and 297 mungbean minicore collections, respectively. Versha et al. (73) recently discovered four subpopulations in an association mapping analysis comprising 80 genotypes. The genomic resources developed in this study will pave the way for the discovery of SNPs/candidate genes linked to agronomic traits in mungbean.

Researchers must first evaluate the degree of linkage disequilibrium (LD) and its degradation before conducting a genome-wide association mapping analysis in a population. A high resolution LD pattern in a population of 127 genotypes was observed in this study, with an LD estimate of 0.62 r^2-value in a population of 127 genotypes. In mungbean chromosomes, the LD decline (decrease of r^2-value to half of its highest) was seen between 0 and 100 kb (about 57.67 kb). LD degradation was seen at 60 and 100 kb physical distances for wild and cultivated mungbean genotypes, respectively, in a previous study (71). Furthermore, in a population of 297 mungbean minicore collections, Breria et al. (72) showed LD decline at a physical distance of 350 kb. The LD degradation seen in this study was likely similar to that observed in other legume crops such as soybean (74), but differed from that found in chickpea (28) at a physical distance of 1,000 kb.

The mineral bioavailability to the consumer is determined by complex features such as grain iron, zinc concentration, and grain phytic acid and tannin content. As a result, dissecting the genetic architecture of these quantitative features in crop plants is critical. For dissecting complex features in crop plants, the association mapping (AM) technique has evolved as a strong and alternative tool to biparental mapping. This method has been used to successfully identify markers/candidate genes associated with grain iron and grain zinc traits in a variety of crop plants, including chickpea (75), wheat (76, 77), pearlmillet (78), and the mungbean itself (58). Till date, just one study in mungbean has used the association mapping approach,

and that study was conducted in the USDA core collection (58). For the first time in mungbean, a variety of Indian and exotic lines were used in an association mapping technique for grain iron and zinc.

Correspondingly, the association mapping strategy has been successful in identifying markers/candidate genes associated with grain phytic acid in a variety of crop plants, including *Brassica rapa* (79), rice (80), common bean (81), and most recently pea (82). However, association mapping has not been used to find markers linked with grain phytic acid levels in the mungbean crop. SNPs/Candidate genes linked with grain phytic acid content were found for the first time in mungbean using an association mapping approach in this study.

In numerous crop plants, including sorghum (83, 84) and rape seed (85), the association mapping approach has been successful in identifying the markers/candidate genes linked with grain tannin concentration. However, association mapping has not been used to find markers linked with grain tannin content in the mungbean crop. SNPs/Candidate genes linked with grain tannin concentration in mungbean were identified for the first time in this study using an association mapping approach.

To create the AM panel for this investigation, 127 different mungbean genotypes were selected and genotyped using 14,447 SNPs. The general linear model (GLM) and Bayesian-information and Linkage-disequilibrium Iteratively Nested Keyway (Blink) techniques were used to analyze the associations. Furthermore, a cut-off value of –log10 > 3.0 was used to identify significant SNPs linked to grain iron and zinc concentrations, while a cut-off value of –log10 > 4.0 was used to identify significant SNPs linked to grain phytic acid and tannin content. The GLM found 9, 6, 5, and 0 SNPs linked to grain iron concentration, zinc concentration, phytic acid content, and tannin content, respectively. While Blink identified 15, 10, 5, and 5 SNPs linked with grain iron, grain zinc, grain phytic acid content, and grain tannin content, correspondingly.

The multigenic regulation of nutrient accumulation in mungbean seeds found in this work coincides with (86) findings of quantitative inheritance. In a mung bean RIL population, these researchers discovered 17 QTLs for Fe and Zn, including seven QTLs on linkage groups LG 6 and LG 7 for Zn and one QTL shared with Fe. In this study, LG 6 and LG 7 each had one QTL for iron, but LG7 had three QTLs for zinc. On the LG 11 map, Singh (87) discovered a potential QTL (qFe-11-1) for iron. One QTL for iron was found at LG11 in this study. In 2020, Wu et al. discovered SNPs associated with grain iron content on LG 06, and an SNP was discovered on LG 06 in this study as well. Wu et al. (58) discovered SNPs linked with grain zinc content on LG 07 in 2020, and three SNPs were located on LG 07 in this study as well.

The 35 SNPs detected by Blink were shown to be linked to 170 protein-coding genes and 11 unidentified genes. Iron ion/heme binding proteins, Glutathione-S transferase (GST), major intrinsic (MIP) protein family, WRKY family transcription factors, squamosal promoter binding proteins, and ATP dependent metalloproteases are among the proteins coding genes for grain iron content. In flowering plants, the WRKY gene family encodes a vast number of transcription factors (TFs) that are involved in a variety of root development, stress responses, developmental, and physiological activities (88). Plant-specific transcription factors encoded by Squamosa Promoter-Binding Protein-Like (SPL) genes serve critical roles in plant phase transition, flower and fruit development, and plant architecture (89). Aquaporin proteins, which are members of the big major intrinsic (MIP) protein family, are the primary facilitators of water transport activity through plant cell membranes. These proteins appear to govern the transcellular route of water (90) and play a critical role in delivering a high volume of water with minimal energy expenditure (91). A variety of stress-response genes are up-regulated in an ATP-dependent metalloprotease with a high level of reactive oxygen species (ROS) (92). Glutathione-S transferases (GST) have been used in a variety of plant functions, including xenobiotic detoxification, secondary metabolism, growth and development, and, most importantly, protection against biotic and abiotic stimuli (93).

Adenylate cyclase, Pentatricopeptide repeat (PPR) superfamily protein, zinc finger (Ran-binding) family protein, magnesium ion binding protein, copper ion binding protein, major intrinsic protein (MIP) family transporter, and protein kinase family protein are the most important protein coding genes for grain zinc concentration.

Protein coding genes for grain phytic acid content consists of Serine/threonine protein phosphatase family protein, serine/threonine kinase, transmembrane amino acid transporter family protein, callose synthase, Phosphatidyl inositol kinase (PIK-G1) n, and Cytochrome P450 superfamily protein. Serine/threonine protein phosphatase family protein plays a prominent role in the regulation of specific signal transduction cascades, as witnessed by its presence in a number of macromolecular signaling modules, where it is often found in association with other phosphatases and kinases (94). The network of protein serine/threonine kinases in plant cells act as a "central processor unit" (cpu), accepting input information from receptors that sense environmental conditions, phytohormones, and other external factors, and converting it into appropriate outputs such as changes in metabolism, gene expression, and cell growth and division (95). According to Lee et al. (96), phosphatidylinositol 3-kinase is essential for vacuole reorganization and nuclear division during pollen development. Phosphatidylinositol 3-phosphate (PtdInsP) is made by the enzyme phosphatidylinositol 3-kinase (PI3K), which phosphorylates phosphoinositides at the D-3 position. PtdIns(3)P is required for normal plant growth (97) and has been linked to a number of physiological processes, including root nodule formation (98), auxin-induced production of reactive oxygen species and root gravitropism (99), root hair curling and Rhizobium infection in *Medicago truncatula* (100), increased plasma membrane endocytosis and the intracellular production of reactive oxygen species in salt tolerance response (101), stomatal closing movement (102, 103), and root hair elongation (100). The cytochrome P450 (CYP) superfamily is the largest enzymatic protein family in plants. Members of this superfamily are involved in multiple metabolic pathways with distinct and complex functions, playing important roles in a vast array of reactions. As a result, numerous secondary metabolites are synthesized that function as growth and developmental signals or protect plants from various biotic and abiotic stresses (104).

While the protein coding genes for grain tannin content consists of serine/threonine-protein phosphatase, heat shock transcription factor B4, triacylglycerol lipase, serine/threonine kinase, Sugar transporter SWEET n, and Nitrate transporter. As discussed earlier the serine/threonine-protein phosphatase plays a prominent role in the regulation of specific signal transduction cascades and control

the changes in metabolism, gene expression, and cell growth and division (95). The enhanced heat shock gene expression in response to various stimuli is regulated by heat shock transcription factors (HSFs) (105) which may have some correlation to the tannin content in grain considering tannins are related to stress response. SWEET (Sugars Will Eventually Exported Transporters) proteins are one of the biggest sugar transporter families in the plant kingdom, and they play an important role in plant growth and stress responses (106). SWEET genes' various functions in critical developmental and physiological processes including as growth, senescence, and flower/seed/pollen formation are similarly explained in higher plants. They are also known to have a role in abiotic and biotic stress adaptation, as well as host-pathogen interactions (107–113). This gives some hints on the relation of the SWEET gene and the tannin content in the mungbean grain. The inclusion of nitrogen transporters in the list of associated genes with tannins suggests possibility of some relation between nitrogen assimilation and tannin content in mungbean grain.

There were 11 uncharacterized genes detected in relation with all of the attributes studied, necessitating more research to determine their function and potential impact on the phenotypic of the trait in question in mungbean grain.

Among these putative candidate genes, genes namely *Vradi04g09970, Vradi07g30210, Vradi06g11980, Vradi09g05480,* and *Vradi07g15310* were identified with missense SNPs in their CDS region. Further, structural changes at protein level due to missense SNPs in their CDS region were observed for two genes namely *Vradi07g15310 and Vradi09g05480.* The allelic variation between native and mutant versions of these genes reveals the discrepancy in protein structure and domains for modification at post transcriptional level. The variation in protein-protein interactions or signal integration leads to a difference in transcriptional modulations that result in an observed phenotypic difference in grain iron concentration and tannin content.

After Wu et al. (58), the current study is the first to report on association mapping of grain phytic acid and tannin content in mungbean, as well as the second to report on association mapping of grain iron and grain zinc concentration in mungbean, however, this study looked at different genotypes. In mungbean breeding programs focusing on bio-fortification and increased nutritional availability, the found SNP markers and candidate genes are useful resources. Furthermore, this research demonstrates that association mapping, particularly using the Blink model, is a powerful tool for dissecting complex traits such as grain iron concentration, grain zinc concentration, grain phytic acid content, and grain tannin content, and provides high resolution mapping at a low cost and in a short amount of time.

Conclusion

- The genotypes PUSA 1333 and IPM 02–19 were identified as desired genotypes as they had high grain iron and zinc concentration but low grain phytic acid and tannin contents.
- The study generated 14,447 genome wide SNPs by employing next generation sequencing (NGS) based genotyping by sequencing (GBS) methodology.
- Population admixture analysis revealed the presence of four different ancestry among the 127 genotypes and LD decay of ∼57.6 kb physical distance was observed in mungbean chromosomes.
- Association mapping analysis revealed that a total of 20 significant SNPs were shared by both GLM and Blink models associated with grain micronutrient and anti-nutritional factor traits.
- The study identified the 185 putative candidate genes including potential candidate genes *Vradi07g30190*, *Vradi01g09630*, and *Vradi09g05450* were found to be associated with grain iron concentration, *Vradi10g04830* with grain zinc concentration, *Vradi08g09870* and *Vradi01g11110* with grain phytic acid content and *Vradi04g11580* and *Vradi06g15090* with grain tannin content.
- Two genes *Vradi07g15310 and Vradi09g05480* showed significant variation in protein structure between native and mutated versions. The identified SNPs and candidate genes are potential powerful tools for nutritional improvement in mungbean breeding program.

Data availability statement

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

Author contributions

MA, GM, and HD: conceptualization and supervision. MS, JT, PY, MK, MA, and RN: methodology. MK, MA, HD, and AP: formal analysis. RN, SK, RS, and HD: resources. MS and MA: data curation. MS, MA, and RN: writing—original draft preparation. MA, MS, RS, RN, AP, and SG: writing—review and editing. All authors contributed to the article and approved the submitted version.

Funding

Funding for this research was provided by the Australian Centre for International Agricultural Research (ACIAR) for the International Mungbean Improvement Network Project (project no. CROP/2019/144).

Acknowledgments

We are thankful to the ICAR - Indian Council of Agricultural Research-IARI, New Delhi, Division of Genetics, IARI, New Delhi and Dr. Y. S. Shivay of the Division of Agronomy, IARI, New Delhi for providing the necessary facilities for the smooth conduct of research.

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/fnut.2023.1099004/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Determination of Q matrix with the lowest cross-validation error (in this case $k = 4$) (X-axis has the values of k while the Y-axis contains the values of CV at corresponding k).

SUPPLEMENTARY FIGURE 2
Phylogenetic tree depicting the genetic relations among 127 diverse mungbean genotypes based on Nei's genetic distance using 14,447 high quality GBS based SNPs.

SUPPLEMENTARY FIGURE 3
Circular diagram depicting a summarized view of the significant association of SNP markers with all the four traits in the study along with SNP density in the outer ring.

References

1. Somta P, Srinives P. Genome research in mungbean [*Vigna radiata* (L.) Wilczek] and blackgram [*V. mungo* (L.) Hepper]. *Sci Asia.* (2007) 33: 69–74.

2. Raturi A, Singh SK, Sharma V, Pathak R. Stability and environmental indices analyses for yield attributing traits in Indian *Vigna radiata* genotypes under arid conditions. *Asian J Agric Sci.* (2012) 4:126–33.

3. World Health Organization. *Vitamin and Mineral Requirements in Human Nutrition*. 2nd ed. Geneva: WHO (2004). p. 246–72.

4. United Nations System Standing Committee on Nutrition. *Sixth Report on the World Nutrition Situation*. Geneva: SCN (2010).

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

6. Dwivedi SL, Sahrawat KL, Rai KN, Blair MW, Andersson MS, Pfeiffer WH. *Nutritionally enhanced staple food crops*. Hoboken, NJ: JohnWiley & Sons Inc (2012). doi: 10.1002/9781118358566.ch3

7. Gopala Krishna AG, Prabhakar JV, Aitzetmüller K. Tocopherol and fatty acid composition of some Indian pulses. *J Am Oil Chem Soc.* (1997) 74:1603–6. doi: 10.1007/s11746-997-0084-2

8. Jones G. Minerals. In: Wahlquist M editor. *Food and Nutrition*. Sydney: Allen & Unwin (1997). p. 249–54.

9. Singh M, Krikorian AD. Inhibition of trypsin activity in vitro by phytate. *J Agric Food Chem.* (1982) 30:799–800. doi: 10.1021/jf00112a049

10. Liener IE. Implications of antinutritional components in soybean foods. *Crit Rev Food Sci Nutr.* (1994) 34:31–67. doi: 10.1080/10408399409527649

11. Haslam E. *Plant polyphenols: vegetable tannins revisited*. Cambridge: CUP Archive (1989).

12. Reed JD. Nutritional toxicology of tannins and related polyphenols in forage legumes. *J Anim Sci.* (1995) 73:1516–28. doi: 10.2527/1995.7351516x

13. Ravindran V, Ravindran G. Nutritional and anti-nutritional characteristics of mucuna (*Mucuna utilis*) bean seeds. *J Sci Food Agric.* (1988) 46:71–9. doi: 10.1002/jsfa.2740460108

14. Josephine RM, Janardhanan K. Studies on chemical composition and antinutritional factors in three germplasm seed materials of the tribal pulse, *Mucuna pruriens* (L.) DC. *Food Chem.* (1992) 43:13–8. doi: 10.1016/0308-8146(92)90235-T

15. Laurena AC, Den Truong V, Mendoza EMT. Effects of condensed tannins on the in vitro protein digestibility of cowpea [*Vigna unguiculata* (L.) Walp.]. *J Agric Food Chem.* (1984) 32:1045–8. doi: 10.1021/jf00125a025

16. Sathe SK, Deshpande SS, Salunkhe DK, Rackis JJ. Dry beans of *Phaseolus*. A review. Part 2. Chemical composition: carbohydrates, fiber, minerals, vitamins, and lipids. *Crit Rev Food Sci Nutr.* (1984) 21:41–93. doi: 10.1080/10408398409527396

17. Dlamini NR, Dykes L, Rooney LW, Waniska RD, Taylor JR. Condensed tannins in traditional wet-cooked and modern extrusion-cooked sorghum porridges. *Cereal Chem.* (2009) 86:191–6. doi: 10.1094/CCHEM-86-2-0191

18. Blair MW. Mineral biofortification strategies for food staples: the example of common bean. *J Agric Food Chem.* (2013) 61:8287–94. doi: 10.1021/jf400774y

19. Kwon S-J, Brown AF, Hu J, Mcgee R, Watt C, Kisha T, et al. Genetic diversity, population structure and genome-wide marker-trait association analysis emphasizing seed nutrients of the USDA pea (*Pisum sativum* L.) core collection. *Genes Genom.* (2012) 34:305–20. doi: 10.1007/s13258-011-0213-z

20. Flint-Garcia SA, Thornsberry JM, Buckler E IV. Structure of linkage disequilibrium in plants. *Annu Rev Plant Biol.* (2003) 54:357–74. doi: 10.1146/annurev.arplant.54.031902.134907

21. Kraakman AT, Niks RE, Van den Berg PM, Stam P, Van Eeuwijk FA. Linkage disequilibrium mapping of yield and yield stability in modern spring barley cultivars. *Genetics.* (2004) 168:435–46. doi: 10.1534/genetics.104.026831

22. Federer WT, Raghavarao D. On augmented designs. *Biometrics.* (1975) 31:29–35.

23. Singh D, Chonkar P, Dwivedi B. *Manual on soil, plant and water analysis*. New Delhi: Westville Publishers (2005).

24. De Camargo AC, de Souza Vieira TMF, Regitano-D'Arce MAB, Calori-Domingues MA, Canniatti-Brazaca SG. Gamma radiation effects on peanut skin antioxidants. *Int J Mol Sc.* (2012) 13:3073–84.

25. Doyle JJ, Doyle JL. A rapid DNA isolation procedure for small quantities of fresh leaf tissue. *Phytochem Bull.* (1987) 19:11–5.

26. Elshire RJ, Glaubitz JC, Sun Q, Poland JA, Kawamoto K, Buckler ES, et al. A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PLoS One.* (2011) 6:e19379. doi: 10.1371/journal.pone.0019379

27. Bastien M, Sonah H, Belzile F. Genome wide association mapping of *Sclerotinia sclerotiorum* resistance in soybean with a genotyping-by-sequencing approach. *Plant Genome.* (2014) 7:1–13. doi: 10.3835/plantgenome2013.10.0030

28. Kujur A, Bajaj D, Upadhyaya HD, Das S, Ranjan R, Shree T, et al. A genome-wide SNP scan accelerates trait-regulatory genomic loci identification in chickpea. *Sci Rep.* (2015) 5:1–20. doi: 10.1038/srep11166

29. Catchen J, Hohenlohe PA, Bassham S, Amores A, Cresko WA. Stacks: an analysis tool set for population genomics. *Mol Ecol.* (2013) 22:3124–40. doi: 10.1111/mec.12354

30. Ewing B, Hillier L, Wendl MC, Green P. Base-calling of automated sequencer traces usingPhred. I. Accuracy assessment. *Genome Res.* (1998) 8:175–85. doi: 10.1101/gr.8.3.175

31. Kang YJ, Kim SK, Kim MY, Lestari P, Kim KH, Ha BK, et al. Genome sequence of mungbean and insights into evolution within *Vigna* species. *Nat Commun.* (2014) 5:1–9. doi: 10.1038/ncomms6443

32. Langmead B, Salzberg SL. Fast gapped-read alignment with Bowtie 2. *Nat Methods.* (2012) 9:357–9.

33. Hohenlohe PA, Bassham S, Etter PD, Stiffler N, Johnson EA, Cresko WA. Population genomics of parallel adaptation in threespine stickleback using sequenced RAD tags. *PLoS Genet.* (2010) 6:e1000862. doi: 10.1371/journal.pgen.1000862

34. Sun S, Gu M, Cao Y, Huang X, Zhang X, Ai P, et al. A constitutive expressed phosphate transporter, OsPht1; 1, modulates phosphate uptake and translocation in phosphate-replete rice. *Plant Physiol.* (2012) 159:1571–81.

35. Varshney RK, Saxena RK, Upadhyaya HD, Khan AW, Yu Y, Kim C, et al. Whole-genome resequencing of 292 pigeonpea accessions identifies genomic regions associated with domestication and agronomic traits. *Nat Genet.* (2017) 49:1082–8. doi: 10.1038/ng.3872

36. Alexander DH, Novembre J, Lange K. Fast model-based estimation of ancestry in unrelated individuals. *Genome Res.* (2009) 19:1655–64. doi: 10.1101/gr.094052.109

37. Evanno G, Regnaut S, Goudet J. Detecting the number of clusters of individuals using the software structure: a simulation study. *Mol Ecol.* (2005) 14:2611–20.

38. Bradbury PJ, Zhang Z, Kroon DE, Casstevens TM, Ramdoss Y, Buckler ES. TASSEL: software for association mapping of complex traits in diverse samples. *Bioinformatics.* (2007) 23:2633–5. doi: 10.1093/bioinformatics/btm308

39. Rafalski JA. Novel genetic mapping tools in plants: SNPs and LD-based approaches. *Plant Sci.* (2002) 162:329–33. doi: 10.1016/S0168-9452(01)00587-8

40. Hill WG, Robertson A. Linkage disequilibrium in finite populations. *Theor Appl Genet.* (1968) 38:226–31.

41. Caldwell KS, Russell J, Langridge P, Powell W. Extreme population-dependent linkage disequilibrium detected in an inbreeding plant species. *Hordeum vulgare. Genetics.* (2006) 172:557–67. doi: 10.1534/genetics.104.038489

42. Warde-Farley D, Donaldson SL, Comes O, Zuberi K, Badrawi R, Chao P, et al. The GeneMANIA prediction server: biological network integration for gene prioritization and predicting gene function. *Nucleic Acids Res.* (2010) 38(Suppl. 2):W214–20.

43. Remington DL, Thornsberry JM, Matsuoka Y, Wilson LM, Whitt SR, Doebley J, et al. Structure of linkage disequilibrium and phenotypic associations in the maize genome. *Proc Natl Acad Sci USA.* (2001) 98:11479–84. doi: 10.1073/pnas.201394398

44. Huang M, Liu X, Zhou Y, Summers RM, Zhang Z. BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience* (2019) 8:giy154.

45. Huang M, Liu X, Zhou Y, Summers RM, Zhang Z. BLINK: a package for the next level of genome-wide association studies with both individuals and markers in the millions. *Gigascience.* (2019) 8:giy154. doi: 10.1093/gigascience/giy154

46. Benjamini Y, Hochberg Y. Controlling the false discovery rate: a practical and powerful approach to multiple testing. *J R Stat Soc Ser B (Methodological).* (1995) 57:289–300. doi: 10.1111/j.2517-6161.1995.tb02031.x

47. Diapari M, Sindhu A, Warkentin TD, Bett K, Tar'an B. Population structure and marker-trait association studies of iron, zinc and selenium concentrations in seed of field pea (*Pisum sativum* L.). *Mol Breed.* (2015) 35:30. doi: 10.1007/s11032-015-0252-2

48. Zheng W, Zhang C, Li Y, Pearce R, Bell EW, Zhang Y. Folding non-homologous proteins by coupling deep-learning contact maps with I-TASSER assembly simulations. *Cell Rep Methods.* (2021) 1:100014.

49. Alexander DH, Lange K. Enhancements to the ADMIXTURE algorithm for individual ancestry estimation. *BMC Bioinform.* (2011) 12:1–6.

50. Boatwright JL, Brenton ZW, Boyles RE, Sapkota S, Myers MT, Jordan KE, et al. Genetic characterization of a *Sorghum* bicolor multiparent mapping population emphasizing carbon-partitioning dynamics. *G3 Bethesda.* (2021) 11:jkab060.

51. Anwar F, Latif S, Przybylski R, Sultana B, Ashraf M. Chemical composition and antioxidant activity of seeds of different cultivars of mungbean. *J Food Sci.* (2007) 72:S503–10. doi: 10.1111/j.1750-3841.2007.00462.x

52. Vijayalakshmi P, Amirthaveni M, Devadas RP. *Possibilities of increasing bioavailability of iron from mungbean and study on the effects of its supplementation on children and women. Project Report*. Coimbatore: Avinashilingam Institute for Home Science and Higher Education for Women, Coimbatore (2001).

53. Afzal MA, Haque MM, Shanmugasundaram S. Random amplified polymorphic DNA (RAPD) analysis of selected mungbean [*Vigna radiata* (L.) Wilczek] cultivars. *Asia J Plant Sci.* (2004) 3:20–4.

54. Singh V. *Genotypic response and QTL identification for micronutrient (iron and zinc) contents in mungbean [Vigna radiata (L.) Wilczek]. Department of Genetics and Plant Breeding, College of Agriculture*. (Vol. 125). Hisar: CCS Haryana Agricultural University (2013). 96 p.

55. Vlk D, Řepková J. Application of next-generation sequencing in plant breeding. *Czech J Genet Plant Breed.* (2017) 53:89–96. doi: 10.17221/192/2016-CJGPB

56. Huang X, Han B. Natural variations and genome-wide association studies in crop plants. *Ann Rev Plant Biol.* (2014) 65:531–51.

57. Akond AGM, Heath Crawford JB, Talukder ZI, Hossain K. Minerals (Zn, Fe, Ca and Mg) and antinutrient (phytic acid) constituents in common bean. *Am J Food Technol.* (2011) 6:235. doi: 10.3923/ajft.2011.235.243

58. Wu X, Islam AF, Limpot N, Mackasmiel L, Mierzwa J, Cortés AJ, et al. Genome-wide Snp identification and association mapping for seed mineral concentration in mung bean (*Vigna Radiata* L.). *Front Genet.* (2020) 11:656. doi: 10.3389/fgene.2020.00656

59. Welch RM, Graham RD. Breeding for micronutrients in staple food crops from a human nutrition perspective. *J Exp Bot.* (2004) 55:353–64. doi: 10.1093/jxb/erh064

60. House WA, Welch RM, Beebe S, Cheng Z. Potential for increasing the amounts of bioavailable zinc in dry beans (*Phaseolus vulgaris* L) through plant breeding. *J Sci Food Agricult.* (2002) 82:1452–7.

61. Nair RM, Yang RY, Easdown WJ, Thavarajah D, Thavarajah P, Hughes JDA, et al. Biofortification of mungbean (*Vigna radiata*) as a whole food to enhance human health. *J Sci Food Agric.* (2013) 93:1805–13. doi: 10.1002/jsfa.6110

62. Nair RM, Pandey AK, War AR, Hanumantharao B, Shwe T, Alam AKMM, et al. Biotic and abiotic constraints in mungbean production—progress in genetic improvement. *Front Plant Sci.* (2019) 10:1340. doi: 10.3389/fpls.2019.01340

63. Souframanien J, Dhole V, Reddy K. Breeding for low phytates and oligosaccharides in mungbean and blackgram. In *Breeding for Enhanced Nutrition and Bio-Active Compounds in Food Legumes*. Berlin: Springer (2021). p. 99–130. doi: 10.1007/978-3-030-59215-8_5

64. Khandelwal S, Udipi SA, Ghugre P. Polyphenols and tannins in Indian pulses: Effect of soaking, germination and pressure cooking. *Food Res Int.* (2010) 43:526–30. doi: 10.1016/j.foodres.2009.09.036

65. Kim SK, Nair RM, Lee J, Lee SH. Genomic resources in mungbean for future breeding programs. *Front Plant Sci.* (2015) 6:626. doi: 10.3389/fpls.2015.00626

66. Lioi L, Zuluaga DL, Pavan S, Sonnante G. Genotyping-by-sequencing reveals molecular genetic diversity in Italian common bean landraces. *Diversity.* (2019) 11:154. doi: 10.3390/d11090154

67. Mgonja EM, Park CH, Kang H, Balimponya EG, Opiyo S, Bellizzi M, et al. Genotyping-by-sequencing-based genetic analysis of African rice cultivars and association mapping of blast resistance genes against *Magnaporthe oryzae* populations in Africa. *Phytopathology.* (2017) 107:1039–46. doi: 10.1094/PHYTO-12-16-0421-R

68. Alipour H, Bihamta MR, Mohammadi V, Peyghambari SA, Bai G, Zhang G. Genotyping-by-sequencing (GBS) revealed molecular genetic diversity of Iranian wheat landraces and cultivars. *Front Plant Sci.* (2017) 8:1293. doi: 10.3389/fpls.2017.01293

69. Wu Y, San Vicente F, Huang K, Dhliwayo T, Costich DE, Semagn K, et al. Molecular characterization of CIMMYT maize inbred lines with genotyping-by-sequencing SNPs. *Theor Appl Genet.* (2016) 129:753–65. doi: 10.1007/s00122-016-2664-8

70. Das A, Poornima KN, Thakur S, Singh NP. The mungbean genome sequence: a blueprint for *Vigna* improvement. *Curr Sci India.* (2016) 111:1144–5.

71. Noble TJ, Tao Y, Mace ES, Williams B, Jordan DR, Douglas CA, et al. Characterization of linkage disequilibrium and population structure in a mungbean diversity panel. *Front plant Sci.* (2018) 8:2102. doi: 10.3389/fpls.2017.02102

72. Breria CM, Hsieh CH, Yen JY, Nair R, Lin CY, Huang SM, et al. Population structure of the world vegetable center mungbean mini core collection and genome-wide association mapping of loci associated with variation of seed coat luster. *Trop Plant Biol.* (2020) 13:1–12.

73. Rohilla V, Yadav RK, Poonia A, Sheoran R, Kumari G, Shanmugavadivel PS, et al. Association Mapping for yield attributing traits and yellow mosaic disease resistance in mung bean [*Vigna radiata* (L.) Wilczek]. Accelerating genetic gains in pulses. *Front Plant Sci.* (2022) 12:749439. doi: 10.3389/fpls.2021.749439

74. Copley TR, Duceppe MO, O'Donoughue LS. Identification of novel loci associated with maturity and yield traits in early maturity soybean plant introduction lines. *BMC Genom.* (2018) 19:167. doi: 10.1186/s12864-018-4558-4

75. Diapari M, Sindhu A, Bett K, Deokar A, Warkentin TD, Tar'an B. Genetic diversity and association mapping of iron and zinc concentrations in chickpea (*Cicer arietinum* L.). *Genome.* (2014) 57:459–68. doi: 10.1139/gen-2014-0108

76. Alomari DZ, Eggert K, Von Wiren N, Alqudah AM, Polley A, Plieske J, et al. Identifying candidate genes for enhancing grain Zn concentration in wheat. *Front Plant Sci.* (2018) 9:1313. doi: 10.3389/fpls.2018.01313

77. Alomari DZ, Eggert K, Von Wirén N, Polley A, Plieske J, Ganal MW, et al. Whole-genome association mapping and genomic prediction for iron concentration in wheat grains. *Int J Mol Sci.* (2018) 20:76. doi: 10.3390/ijms20010076

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

79. Zhao J, Paulo MJ, Jamar D, Lou P, Van Eeuwijk F, Bonnema G, et al. Association mapping of leaf traits, flowering time, and phytate content in *Brassica rapa. Genome.* (2007) 50:963–73. doi: 10.1139/G07-078

80. Perera I, Fukushima A, Arai M, Yamada K, Nagasaka S, Seneweera S, et al. Identification of low phytic acid and high Zn bioavailable rice (*Oryza sativa* L.) from 69 accessions of the world rice core collection. *J Cereal Sci.* (2019) 85:206–13. doi: 10.1016/j.jcs.2018.12.010

81. Blair MW, Herrera AL, Sandoval TA, Caldas GV, Filleppi M, Sparvoli F. Inheritance of seed phytate and phosphorus levels in *Common bean* (L.) and association with newly-mapped candidate genes. *Mol Breed.* (2012) 30:1265–77. doi: 10.1007/s11032-012-9713-z

82. Powers S, Boatwright JL, Thavarajah D. Genome-wide association studies of mineral and phytic acid concentrations in pea (*Pisum sativum* L.) to evaluate biofortification potential. *G3.* (2021) 11:jkab227. doi: 10.1093/g3journal/jkab227

83. Habyarimana E, Dall'Agata M, De Franceschi P, Baloch FS. Genome-wide association mapping of total antioxidant capacity, phenols, tannins, and flavonoids in a panel of Sorghum bicolor and S. bicolor× S. halepense populations using multi-locus models. *PLoS One.* (2019) 14:e0225979. doi: 10.1371/journal.pone.0225979

84. Habyarimana E, Piccard I, Catellani M, De Franceschi P, Dall'Agata M. Towards predictive modeling of sorghum biomass yields using fraction of absorbed photosynthetically active radiation derived from sentinel-2 satellite imagery and supervised machine learning techniques. *Agronomy.* (2019) 9:203. doi: 10.3390/agronomy9040203

85. Rezaeizad A, Wittkop B, Snowdon R, Hasan M, Mohammadi V, Zali A, et al. Identification of QTLs for phenolic compounds in oilseed rape (*Brassica napus* L.) by association mapping using SSR markers. *Euphytica.* (2011) 177:335–42. doi: 10.1007/s10681-010-0231-y

86. Singh V, Yadav RK, Yadav NR, Yadav Rajesh MR, Singh J. Identification of genomic Regions/genes for high iron and zinc content and cross transferability of SSR markers in mungbean (*Vigna radiata* L.). *J Article.* (2017) 6:1004–11. doi: 10.18805/lr.v40i04.9006

87. Singh R. *Development of iron and zinc enriched mungbean (Vigna radiata L.) cultivars with agronomic traits in consideration*. Wageningen: Wageningen University and Research (2013).

88. Chen F, Hu Y, Vannozzi A, Wu K, Cai H, Qin Y, et al. The WRKY transcription factor family in model plants and crops. *Crit Rev Plant Sci.* (2017) 36:311–35. doi: 10.1080/07352689.2018.1441103

89. Chen X, Zhang Z, Liu D, Zhang K, Li A, Mao L. SQUAMOSA promoter-binding protein-like transcription factors: Star players for plant growth and development. *J Integr Plant Biol.* (2010) 52:946–51. doi: 10.1111/j.1744-7909.2010.00987.x

90. Maurel C. Aquaporins and water permeability of plant membranes. *Ann Rev Plant Biol.* (1997) 48:399–429. doi: 10.1146/annurev.arplant.48.1.399

91. Tyerman SD, Bohnert HJ, Maurel C, Steudle E, Smith JAC. Plant aquaporins: their molecular biology, biophysics and significance for plant water relations. *J Exp Bot.* (1999) 50:1055–71. doi: 10.1093/jxb/50.Special_Issue.1055

92. Liu Q, Galli M, Liu X, Federici S, Buck A, Cody J, et al. NEEDLE1 encodes a mitochondria localized ATP-dependent metalloprotease required for thermotolerant maize growth. *Proc Natl Acad Sci USA*. (2019) 116:19736–42. doi: 10.1073/pnas.1907071116

93. Vaish S, Gupta D, Mehrotra R, Mehrotra S, Basantani MK. Glutathione S-transferase: A versatile protein family. *3 Biotech.* (2020) 10:1–19. doi: 10.1007/s13205-020-02312-3

94. Janssens V, Goris J. Protein phosphatase 2A: a highly regulated family of serine/threonine phosphatases implicated in cell growth and signalling. *Biochem J.* (2001) 353:417–39. doi: 10.1042/bj3530417

95. Hardie DG. Plant protein serine/threonine kinases: classification and functions. *Ann Rev Plant Biol.* (1999) 50:97–131. doi: 10.1146/annurev.arplant.50.1.97

96. Lee Y, Kim ES, Choi Y, Hwang I, Staiger CJ, Chung YY, et al. The *Arabidopsis* phosphatidylinositol 3-kinase is important for pollen development. *Plant Physiol.* (2008) 147:1886–97. doi: 10.1104/pp.108.121590

97. Welters P, Takegawa K, Emr SD, Chrispeels MJ. AtVPS34, a phosphatidylinositol 3-kinase of *Arabidopsis thaliana*, is an essential protein with homology to a calcium-dependent lipid binding domain. *Proc Natl Acad Sci USA*. (1994) 91:11398–402. doi: 10.1073/pnas.91.24.11398

98. Hong Z, Verma DP. A phosphatidylinositol 3-kinase is induced during soybean nodule organogenesis and is associated with membrane proliferation. *Proc Natl Acad Sci.* (1994) 91:9617–21. doi: 10.1073/pnas.91.20.9617

99. Joo JH, Yoo HJ, Hwang I, Lee JS, Nam KH, Bae YS. Auxin-induced reactive oxygen species production requires the activation of phosphatidylinositol 3-kinase. *FEBS Lett.* (2005) 579:1243–8. doi: 10.1016/j.febslet.2005.01.018

100. Lee Y, Bak G, Choi Y, Chuang WI, Cho HT, Lee Y. Roles of phosphatidylinositol 3-kinase in root hair growth. *Plant Physiol.* (2008) 147:624–35. doi: 10.1104/pp.108.117341

101. Leshem Y, Seri L, Levine A. Induction of phosphatidylinositol 3-kinase-mediated endocytosis by salt stress leads to intracellular production of reactive oxygen species and salt tolerance. *Plant J.* (2007) 51:185–97. doi: 10.1111/j.1365-313X.2007.03134.x

102. Jung JY, Kim YW, Kwak JM, Hwang JU, Young J, Schroeder J I, et al. Phosphatidylinositol 3-and 4-phosphate are required for normal stomatal movements. *Plant Cell.* (2002) 14:2399–412. doi: 10.1105/tpc.004143

103. Park KY, Jung JY, Park J, Hwang JU, Kim YW, Hwang I, et al. A role for phosphatidylinositol 3-phosphate in abscisic acid-induced reactive oxygen species generation in guard cells. *Plant Physiol.* (2003) 132:92–8. doi: 10.1104/pp.102.016964

104. Jun XU, Wang XY, Guo WZ. The cytochrome P450 superfamily: Key players in plant development and defense. *J Integr Agric.* (2015) 14:1673–86. doi: 10.1016/S2095-3119(14)60980-1

105. Pirkkala L, Nykänen P, Sistonen LEA. Roles of the heat shock transcription factors in regulation of the heat shock response and beyond. *FASEB J.* (2001) 15:1118–31. doi: 10.1096/fj00-0294rev

106. Gautam T, Saripalli G, Gahlaut V, Kumar A, Sharma PK, Balyan HS, et al. Further studies on sugar transporter (SWEET) genes in wheat (*Triticum aestivum* L.). *Mol Biol Rep.* (2019) 46:2327–53. doi: 10.1007/s11033-019-04691-0

107. Chen LQ, Lin IW, Qu XQ, Sosso D, McFarlane HE, Londoño A, et al. A cascade of sequentially expressed sucrose transporters in the seed coat and endosperm provides nutrition for the *Arabidopsis* embryo. *Plant Cell.* (2015) 27:607–19. doi: 10.1105/tpc.114.134585

108. Durand M, Porcheron B, Hennion N, Maurousset L, Lemoine R, Pourtau N. Water deficit enhances C export to the roots in *Arabidopsis thaliana* plants with contribution of sucrose transporters in both shoot and roots. *Plant Physiol.* (2016) 170:1460–79. doi: 10.1104/pp.15.01926

109. Seo P, Park J, Kang S, Kim S, Park C. An Arabidopsis senescence-associated protein SAG29 regulates cell viability under high salinity. *Planta.* (2011) 233:189–200. doi: 10.1007/s00425-010-1293-8

110. Sosso D, Luo D, Li Q, Sasse J, Yang J, Gendrot G, et al. Seed filling in domesticated maize and rice depends on SWEET-mediated hexose transport. *Nat Genet.* (2015) 47:1489–93. doi: 10.1038/ng.3422

111. Zhou Y, Liu L, Huang W, Yuan M, Zhou F, Li X, et al. Overexpression of OsSWEET5 in rice causes growth retardation and precocious senescence. *PLoS One.* (2014) 9:94210. doi: 10.1371/journal.pone.0094210

112. Le Hir R, Spinner L, Klemens PA, Chakraborti D, de Marco F, Vilaine F, et al. Disruption of the sugar transporters AtSWEET11 and AtSWEET12 affects vascular development and freezing tolerance in *Arabidopsis*. *Mol Plant.* (2015) 8:1687–90. doi: 10.1016/j.molp.2015.08.007

113. Chen L. SWEET sugar transporters for phloem transport and pathogen nutrition. *New Phytol.* (2014) 201:1150–5. doi: 10.1111/nph.12445

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 23 February 2023
DOI 10.3389/fnut.2023.1088208

Check for updates

OPEN ACCESS

EDITED BY
Ivan Salmerón,
Autonomous University of Chihuahua, Mexico

REVIEWED BY
Nitish Ranjan Prakash,
Central Soil Salinity Research Institute (ICAR), India
Shanmugavadivel P. S.,
Indian Institute of Pulses Research (ICAR), India

*CORRESPONDENCE
R. K. Gautam
✉ raj.gautam@icar.gov.in
Sapna Langyan
✉ singh.sapna06@gmail.com

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

RECEIVED 03 November 2022
ACCEPTED 19 January 2023
PUBLISHED 23 February 2023

CITATION
Gautam RK, Singh PK, Venkatesan K, Rakesh B, Sakthivel K, Swain S, Srikumar M, Zamir Ahmed SK, Devakumar K, Rao SS, Vijayan J, Ali S and Langyan S (2023) Harnessing intra-varietal variation for agro-morphological and nutritional traits in a popular rice landrace for sustainable food security in tropical islands. *Front. Nutr.* 10:1088208.
doi: 10.3389/fnut.2023.1088208

COPYRIGHT
© 2023 Gautam, Singh, Venkatesan, Rakesh, Sakthivel, Swain, Srikumar, Zamir Ahmed, Devakumar, Rao, Vijayan, Ali and Langyan. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Harnessing intra-varietal variation for agro-morphological and nutritional traits in a popular rice landrace for sustainable food security in tropical islands

Raj Kumar Gautam[1,2]*, Pankaj Kumar Singh[1], Kannan Venkatesan[1], Bandol Rakesh[1], Krishnan Sakthivel[1,3], Sachidananda Swain[1], Muthulingam Srikumar[1], S. K. Zamir Ahmed[1], Kishnamoorthy Devakumar[1,4], Shyam Sunder Rao[1], Joshitha Vijayan[1,5], Sharik Ali[2] and Sapna Langyan[2]*

[1]ICAR-Central Island Agricultural Research Institute, Port Blair, Andaman and Nicobar Islands, India, [2]ICAR-National Bureau of Plant Genetic Resources, Pusa, New Delhi, India, [3]ICAR-Indian Institute of Oilseed Research, Hyderabad, Telangana, India, [4]ICAR-Sugarcane Breeding Institute, Coimbatore, Tamil Nadu, India, [5]ICAR-National Institute of Plant Biotechnology, Pusa, New Delhi, India

Introduction: Rice crop meets the calorie and nutritional requirements of a larger segment of the global population. Here, we report the occurrence of intra-varietal variation in a popular rice landrace C14-8 traditionally grown under the geographical isolation of the Andaman Islands.

Methods: Based on grain husk color, four groups were formed, wherein the extent of intra-varietal variation was studied by employing 22 agro-morphological and biochemical traits.

Results: Among the traits studied, flavonoid and anthocyanin contents and grain yield exhibited a wider spectrum of variability due to more coefficients of variation (>25%). The first five principal components (PCs) of principal components analysis explained a significant proportion of the variation (91%) and the first two PCs explained 63.3% of the total variation, with PC1 and PC2 explaining 35.44 and 27.91%, respectively. A total of 50 highly variable SSR (HvSSR) markers spanning over 12 chromosomes produced 314 alleles, which ranged from 1 to 15 alleles per marker, with an average of 6.28. Of the 314 alleles, 64 alleles were found to be rare among the C14-8 selections. While 62% of HvSSR markers exhibited polymorphism among the C14-8 population, chromosomes 2, 7, 9, and 11 harbored the most polymorphic loci. The group clustering of the selections through HvSSR markers conformed to the grouping based on grain husk coloration.

Discussion: Our studies on the existence and pertinence of intra-varietal variations are expected to be of significance in the realms of evolutionary biology and sustainable food and nutritional security under the changing climate.

KEYWORDS

rice landraces, open florets, biochemical traits, intra-varietal variations, sustainable food system

1. Introduction

Although the rice crop contributes immensely to the world's food supply, especially in Asia, its history of origin, domestication, and evolutionary genetics remain elusive (1, 2). It is believed that rice was domesticated approximately 10,000 years ago from its wild ancestor, *Oryza rufipogon*, a widely distributed Asian native species (3, 4). Rice landraces constitute an important gene pool for

genetic improvement for the present as well as future needs (5). However, the genetic diversity of domesticated rice has been reduced to the tune of 80% compared with wild progenitor right from domestication to the development of modern cultivars (6, 7). If the present trend continues, it may deprive mankind of the genetic advantage of rice diversity in meeting the unforeseen challenges of diseases, pests, agronomic adaptation, and climatic change (8). Traditional farmers in different parts of the world prefer landraces due to their time-tested adaptations to local climates and specific traits (9, 10), and therefore, detailed profiling of these landraces with respect to agro-morphological, biochemical, and genomic characterization is very essential (11, 12). Singh (13) opined that the genetic diversity of the heterogenous population confers resilience and adaptive traits to the species.

Traditional landraces having distinct morphological traits and locally popular names based on their novelties and representing the intermediate stage between a wild ancestor and modern cultivars serve as reservoirs of various useful genes (14). Although landraces have a lower yielding ability, these play an important role in maintaining yields in traditional and stress-prone agricultural systems (15, 16). Such ancient varieties have evolved under a series of climatic, management, and cultural events over the long period of domestication, developing genetic resilience to climatic upheavals (17). Kyratzis et al. (18) opined that the intra-varietal variation found within landraces offers the scope of genetic plasticity, and thus, landraces have the ability to adapt to local field conditions and marginal environments. Numerous studies confirm that locally adapted crops and their landraces are time-tested and deeply seated in the local livelihood system, which requires low inputs, provides better nutrition, and is resistant to prevailing abiotic and biotic stresses, particularly in marginal areas (16, 19). Therefore, it is worthwhile to improve farmers' genetic resources not only for indigenous requirements but also for trait improvement of other ecosystems (9, 20). However, it is significant to mention that all the landraces, no matter how useful these could be in future, may be expected to be conserved under on-farm conditions by farmers themselves even if these are low yielding in heterogeneous form, and seed purity and availability are the concerning issues (21).

It is recently reported that natural evolution and artificial selection may cause speciation in rice (22). While traditional landraces have been reconstituted through the interplay between adaptation to the local environments and selection exercised by farmers as per their agro-ecological and cultural requirements and preferences, wild rice populations proliferate owing to their invasive attributes, and outcompeting ability under natural conditions (23). However, it is not understood clearly how these events interacted to shape and influence the population genetics and intra-varietal structure of a landrace over time and space (18, 24). Although landraces constitute genetically a dynamic system, there are scanty reports on the systematic characterization and quantification of such intra-varietal diversity in their hot spot areas (10). Their detailed characterization and exposition for a series of traits and selection of desirable types may be vital for understanding evolutionary trajectory, trait discovery, and crop improvement (25).

A traditionally popular C14-8 rice landrace has carved out a historical niche in the agricultural landscape of the tropical Andaman and Nicobar Islands due to its advantageously adaptive traits albeit its precise origin and introduction in the islands are not clearly known (26). Londo et al. (7) and Lu (27) concluded that the domestication of rice occurred at least twice, once in the south of Northeastern India, Myanmar, and Thailand and again in Southern China. Here, it is pertinent to mention that the Andaman Islands where C14-8 has been traditionally grown since 1945 or before are also geographically located within the aforementioned first center of domestication [(28); Figure 1]. The tropical climate of the islands is also understood to favor higher diversification rate during the evolutionary process due to biological process of reproductive isolation, faster genetic evolution, and dynamic abiotic and biotic pressures (28, 29).

Although rice is a cleistogamous crop with almost complete autogamy, a peculiar tendency for an open floret was observed in C14-8 (Figure 2A). The genetic and molecular bases of an open floret trait in rice have been unraveled and mostly attributed to jasmonic acid pathways manipulations (30). Furthermore, the detection of different grain color types (Figures 2B,C) occurring naturally within promiscuous C14-8 landrace population prompted us to further investigate the causes, nature, and extent of genetic diversity prevailing in a traditional rice landrace grown under geographical and phenological isolation through the use of morphological, biochemical, and molecular markers. Furthermore, it was intriguing to know if the classification for grain husk color conformed to the cluster grouping based on the agro-morphological, biochemical, and molecular markers. Our study also revealed the traits, which underwent relatively wider divergence as a result of natural selection and adaptation under low input agronomic management for over 7 decades under geographical isolation of islands. It is also imperative to understand the dynamics of inter-trait correlation coefficients in such a biologically unique population for their utilization during indirect selection.

Although there is a good number of studies on the diversity between varieties, pure lines, local cultivars, and/or landraces [(31–38); Vannirajan et al., 2012], systematic studies on intra-varietal variation in landraces are far more inadequate. Therefore, the novelty of our study is the documented multi-trait attempt to understand the promiscuity-induced intra-varietal genetic diversity for 22 agro-morphological, biochemical, and molecular traits in an open floret tropical *japonica* rice variety being cultivated over 7 decades under the closed system of tropical islands. The model on understanding the biological basis and evolutionary genetics of these findings could be relevant and extrapolated to other self-pollinated crops with similar traits and conditions as well. It is also important to know which traits or combinations of traits contributed to maximum variation in a landrace population. The information, thus, generated could be utilized for pinpointing responsible traits useful in direct selection for genetic improvement of such a population. In addition, it is also interesting to look for promising lines in the germplasm pool that possess superiority for the maximum number of agronomic and biochemical traits. The incremental gain in the production potential of socially popular C14-8 rice landrace through our selection interventions will be realized without additional cost to the farmers and with no adverse impact on the local environment and the prevailing food system. In addition to their use for a basic understanding of evolutionary genetic perspectives and mapping studies, potential genotypes emanating from such landraces are strongly likely expected in realizing the Sustainable Development Goals (SDG) by 2030.

2. Methodology

2.1. About the study locale

The tropical Andaman and Nicobar Islands encompassing more than 500 islands located in the Bay of Bengal are approximately 1,200 km

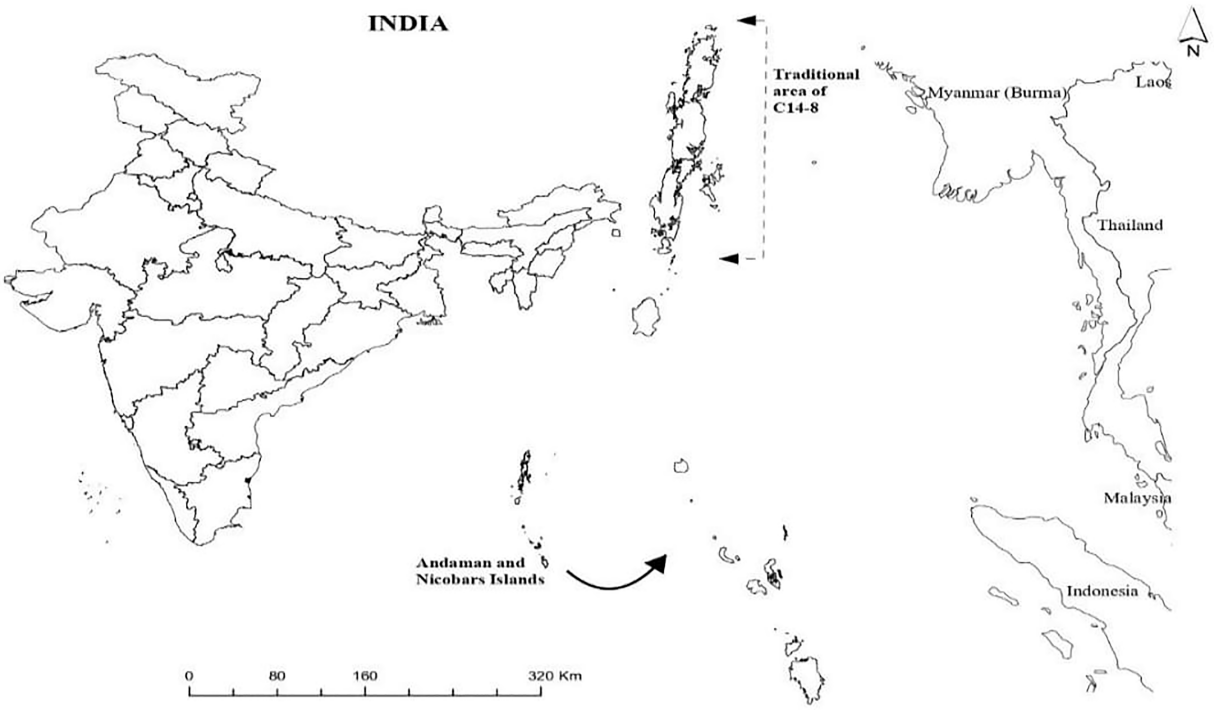


FIGURE 1
Map showing Andaman and Nicobar Islands.

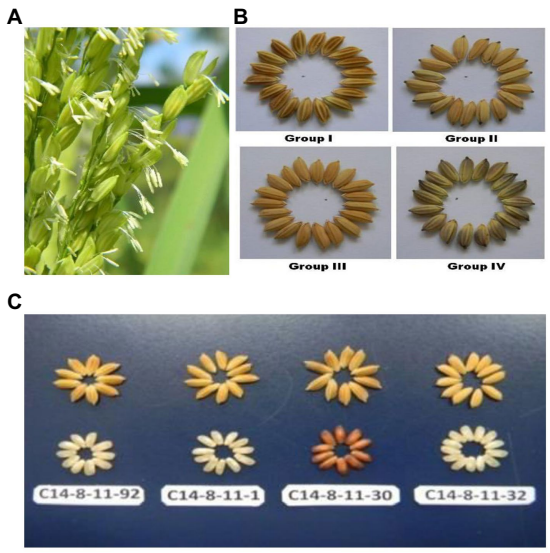


FIGURE 2
(A) Rice florets expressing wider opening with extra-long filaments and protruding anthers. **(B)** Rice grains having different husk color in Group I (brown husk with white apiculus), Group II (yellow husk with black apiculus), Group III (yellow husk with Yellow apiculus) and Group IV (Golden furrows on brown husk with Golden furrows on brown husk with black apiculus). **(C)** Classification of C14-8 selections into 4 types of colors in husk and corresponding decorticated rice grains.

far from mainland India (Figure 1). In view of pristine oceanic and terrestrial life forms, Andaman and Nicobar Islands are the first in India to have been listed as the new "hope spots" by the International Union for Conservation of Nature.[1] The Andaman and Nicobar archipelago is positioned between two major biodiversity areas of long Island Arch extending from the Arakan Yoma hill region of Myanmar to the Sumatran range of Indonesia (38) and rich in tropical plant diversity representing Indian, Burmese, Thai, Malaysian, and Indonesian floras (39). These islands are also listed as one of the 22 agrobiodiversity hot spots in India.

2.2. Salient features of C14-8 landrace

C14-8, locally called *Aath Number Dhan* is a highly popular rice landrace owing to its agronomic merits and ability to fit in the geo-climatic and socioeconomic conditions of the Bay Islands. It is suitable for high rainfall conditions (3,000 mm annual rainfall) of tropical islands and marshy areas and fresh ponds due to its taller height and good culm strength. It is also stringently photosensitive maturing in January month, which enables its safe harvest and threshing when rainfall ceases to occur in these islands. This landrace expressed the "open floret" trait, which is documented as a novel biological feature with breeding ramifications in self-pollinated cereals such as rice (40), and it is imperative to note that all the 20 lines under investigation exhibited open floret tendency. Due to this unique feature and novel trait for hybrid breeding, this germplasm has also been registered in the Indian national genebank (IC0613963, INGR15014). Due to its luxuriant and clumpy growth of C14-8, a smaller number of seedlings for planting

1 https://www.downtoearth.org.in/news/andamans-lakshwadeep-declared-hope-spots-by-iucn-42556

per unit area is required, which saves labor as well as seed rate. It has medium bold non-chalky grains that do not break even due to conventional hulling structures and has palatable and non-sticky cooked rice. This variety fits well under these conditions due to minimal management needs including less work involved for transplanting, weeding, and cultural operations, which also suits the organic mandate of the Islands. It has also been observed to have a strong resurgence and kneeing ability after disease and pests and cyclonic stresses. Additional straw due to its tall stature is also used for cattle feeding (after lopping during early stages and after threshing) or plowing into the soil, traditional houses thatching, barnyard covers, and mushroom cultivation. C14-8 attains a quite tall stature (approximately 200 cm height), matures very late (180 days), has 7–8 tillers/plant with longer panicles (30 cm), good spikelet fertility (75%) with short bold grains having test weight (1,000 grains) of approximately 26 g.

Our interaction with some octogenarian farmers revealed that conventionally C14-8 had yellow-colored grains but subsequently other grain color types got intruded in it, although the red grain type is traditionally perceived as more adaptable to swampy/marshy areas. Recently, it has been found that C14-8 grains also possess higher contents of zinc (Zn ~30 ppm) than the standard check varieties Swarna (14 ppm) and IR64 (15 ppm). Similarly, C14-8 grains showed a higher amount of iron (Fe ~16 ppm) than Swarna (14 ppm) and IR64 (9 ppm; our unpublished data). Understandably, all these traits might have accorded socioeconomic popularity to this landrace and thus served the food and nutritional requirements of the geographically isolated Andaman and Nicobar Islands for a long time.

2.3. Plant materials, characterization, and evaluation

Approximately 150 panicles showing grain husk color variation in C14-8 landraces were collected from diverse fields in Andaman districts at the time of maturity in the month of January 2012. During the subsequent years (2012–2014), all these panicle-to-row progenies were grown and evaluated for both qualitative and quantitative traits. Based on grain yield and grain husk color, 20 lines were chosen from these panicle-to-row progenies so that five representative lines from each of the four grain husk color types were selected for further investigation. The 20 lines were classified into four groups, such as Group I: brown husk with white apiculus, Group II: yellow husk with black apiculus, Group III: yellow husk with yellowish apiculus, and Group IV: golden furrows on brown husk with black apiculus (Table 1; Figures 2B,C). These four groups, henceforth, will be referred to as basic classification groups. During 2012 and 2014, these 20 selections of C14-8 were characterized and evaluated in a randomized block design (RBD) with three replications at Bloomsdale Farm, ICAR-CIARI, and Port Blair for morphological markers, such as basal leaf sheath color, leaf sheath anthocyanin, stem length, anthocyanin pigment on nodes, apiculus color, grain husk color, panicle secondary branching, leaf senescence, and decorticated grain color, and the data on the recorded traits were averaged across 2 years, which is more representative. The lines were also evaluated for quantitative traits such as plant height (PH, cm), days to 50% flowering (DF), ear bearing tillers per plant (EBT), panicle length (PL, cm), 1,000-grain weight (TGW, gm), grain length (GL, mm), grain width (GW, mm), grain yield (GY, t/ha), brown rice recovery (BRR, %), milled rice recovery (MRR, %), and head rice recovery (HRR, %), as mentioned in Table 2. The chemical test to confirm whether C14-8 belongs to *indica* or *japonica* sub-specific group revealed C14-8 to be *japonica* type due to yellow grain husk color retention as per the method of Sanni et al. (35).

The classification for stem length was carried out as per National Guidelines for the conduct of tests for distinctness, uniformity, and stability (DUS) published by ICAR-Indian Institute of Rice Research (IIRR), Hyderabad, India. Therefore, the following scale was followed for stem length: very short (<91 cm), short (91–110 cm), medium (111–130 cm), long (131–150 cm), and very long (>150 cm). Next, when panicles have relatively more secondary branches, it is classified as "strong." When panicles have a greater number of tertiary branches, it is classified as "clustered." Furthermore, leaf senescence is classified as "light" (observed only in two selections) and medium (18 selections) as per leaf color appearance at maturity according to Shobha et al. (41).

2.4. Estimation of biochemical traits

Rice starch (complex carbohydrate) has attracted attention for its use in foods, extruded products, soups, and dressings due to its small size of starch granules, neutral taste, and soft mouth feel. Similarly, the market value of rice is influenced by its cooking qualities (amylose content, gel consistency, and alkali spreading value) apart from differential amounts of phenolics, flavonoids, antioxidants, head rice yield, etc.

2.4.1. Amylose content

Amylose content and total and reducing sugar were determined for all 20 single-panicle progenies using the standard protocol described by Sadasivam and Manickam (42).

2.4.2. Alkali digestibility

A duplicate set of six whole-milled kernels without cracks was selected and placed in a plastic box (5 cm × 5 cm × 2.5 cm). Approximately, 10 mL of 1.7% (0.3035 M) potassium hydroxide (KOH) solution was added. The samples were arranged to provide enough space between kernels to allow for spreading. The boxes were covered and incubated for 23 h in a 30°C oven. Starchy endosperm was rated visually based on a seven-point numerical spreading scale (43).

2.4.3. Gel consistency

The gel consistency (GC) is based on the consistency of a cold 4.4% milled rice paste in 0.2 M KOH (44). The GC was measured by the length of the cold gel in the culture tube held horizontally for 0.5 to 1 h. The GC of rice with less than 24% amylose is usually soft. The test separated high-amylose rice into three categories as follows:

1. Very flaky rice with hard GC (length of gel 40 mm or less).
2. Flaky rice with medium GC (length of gel 41 to 60 mm).
3. Soft rice with soft GC (length of gel more than 61 mm).

2.4.4. Total phenolic content

Total phenolic content in fresh samples was determined with the Folin–Ciocalteu reagent by the method described by Singleton et al. (45), with some modifications. In brief, 0.2 mL of a sample extract (1 mg/mL) was mixed with 1 mL of a 10-fold dilution of the Folin–Ciocalteu reagent and 0.8 mL of 15% (w/v) sodium bicarbonate solution and allowed to stand at room temperature for 30 min. The absorbance was measured at 765 nm, using a UV–visible spectrophotometer (Elico

TABLE 1 Morphological descriptors of 20 lines derived from C14-8 rice population*.

Group	Genotypes	Leaf sheath intensity of anthocyanin color	Stem length	Anthocyanin color on nodes	Apiculus color	Grain husk color	Panicle: secondary branching	Leaf: senescence	Decorticated grain: color
I	C14-8-11-1	Absent	V. long	Present	White	Brown husk	Clustered	Light	White
	C14-8-11-30	Absent	V. long	Present	White	Brown husk	Clustered	Light	Light red
	C14-8-11-91	Absent	Long	Present	White	Brown husk	Clustered	Medium	White
	C14-8-11-92	Absent	Long	Present	White	Brown husk	Clustered	Medium	White
	C14-8-11-93	Absent	Long	Present	White	Brown husk	Clustered	Medium	White
II	C14-8-11-31	Present	Long	Absent	Purple	Yellow husk	Strong	Medium	White
	C14-8-11-32	Present	Long	Absent	Purple	Yellow husk	Strong	Medium	White
	C14-8-11-43	Present	V. Long	Absent	Purple	Yellow husk	Strong	Medium	White
	C14-8-11-59	Present	V. Long	Absent	Purple	Yellow husk	Strong	Medium	White
	C14-8-11-60	Present	V. Long	Absent	Purple	Yellow husk	Strong	Medium	White
III	C14-8-11-61	Absent	Long	Absent	Yellowish	Yellow husk	Clustered	Medium	White
	C14-8-11-90	Absent	Long	Absent	Yellowish	Yellow husk	Clustered	Medium	White
	C14-8-11-108	Present	Long	Absent	Yellowish	Yellow husk	Clustered	Medium	White
	C14-8-11-113	Present	Long	Absent	Yellowish	Yellow husk	Clustered	Medium	White
	C14-8-11-143	Present	Long	Absent	Yellowish	Yellow husk	Clustered	Medium	White
IV	C14-8-11-114	Present	Long	Absent	Yellowish	Golden furrows on brown husk	Clustered	Medium	White
	C14-8-11-115	Present	Long	Absent	Black	Golden furrows on brown husk	Clustered	Medium	White
	C14-8-11-116	Present	Long	Absent	Black	Golden furrows on brown husk	Clustered	Medium	White
	C14-8-11-117	Present	Long	Absent	Black	Golden furrows on brown husk	Clustered	Medium	White
	C14-8-11-118	Present	Long	Absent	Black	Golden furrows on brown husk	Clustered	Medium	White

*Classification method: Shobha et al. (41).

10.3389/fnut.2023.1088208

TABLE 2 Range and statistical parameters for agro-morphological and biochemical traits of C14-8 derived lines.

Trait	Range	Mean±std. error	Variance	Coeff. var
PH	190.0–213.0	202.6 ± 1.3	35.5	2.9
DF	174.0–179.0	176.6 ± 0.3	2.6	0.9
EBT	7.0–9.0	7.3 ± 0.1	0.3	7.9
PL	28.0–33.0	29.7 ± 0.3	1.6	4.3
TGW	23.1–27.4	26.46 ± 0.3	1.7	7.2
GL	7.7–8.8	8.2 ± 0.1	0.1	3.3
GB	2.9–3.4	3.1 ± 0.0	0.0	4.5
GL/GB	2.4–2.9	2.61 ± 0.1	0.1	3.9
GY	1.6–4.3	2.6 ± 0.2	0.5	27.6
BRR	70.4–78.8	75.4 ± 0.5	4.8	2.9
MRR	58.8–73.5	67.4 ± 0.8	11.9	5.1
HRR	58.1–71.1	63.5 ± 0.9	16.4	6.4
AC	6.4–13.0	9.3 ± 0.3	1.7	14.0
RSC	1.7–2.6	2.0 ± 0.0	0.0	10.9
TSC	50.8–77.3	74.1 ± 1.6	55.4	10.0
GC	3.3–3.9	3.7 ± 0.0	0.0	4.2
AD	4.0–6.0	5.0 ± 0.1	0.1	7.6
PC	14.2–18.2	16.5 ± 0.2	0.8	5.4
FC	2.4–6.1	3.4 ± 0.2	0.9	27.7
AnC	−0.2-1.2	0.5 ± 0.1	0.1	64.4
ASA1	72.4–92.9	88.7 ± 0.9	16.3	4.6
ASA2	72.4–93.6	89.1 ± 0.9	17.1	4.6

PH, Plant height (cm); DF, Days to 50% flowering; EBT, Ear bearing tillers per plant; PL, Panicle length (cm); TGW, 1,000-grain weight (gm); GL, Grain length (mm); GW, Grain width (mm); GY, Grain yield (t/ha); BRR, Brown Rice Recovery (%); MRR, Milled rice recovery (%); HRR, Head Rice Recovery (%); AC, Amylose content (mg/100 mg); RSC, Reducing sugar content (mg/100 mg); TSC, Total sugar content (mg/100 mg); GC, Gel consistency; AD, Alkali digestibility; PC, Phenolic content (mg/100 mg); FC, Flavonoid content (mg/100 mg); AnC, Anthocyanin content (mg/100 mg); ASA1, % Antioxidant scavenging activity/1 g by DPPH; ASA2, % Antioxidant scavenging activity/1 g by ABTS.

SL-164, Elico Ltd., Hyderabad, India), and total phenolic content was expressed as gallic acid equivalent (mg/100 g fresh weight).

2.4.5. Total flavonoid content

Total flavonoid content was determined by a colorimetric method (46). In brief, 0.5 mL extracts were added to 15 mL polypropylene conical tubes containing 2 ml ddH_2O and mixed with 0.15 mL of 5% $NaNO_2$. After reacting for 5 min, 0.15 mL of 10% $AlCl_3.6H_2O$ solution was added. After another 5 min, 1 mL of 1 M NaOH was added. The reaction solution was mixed well and kept for 15 min, and the absorbance was determined at 415 nm. Quantification was performed using Rutin as standard, and the results were expressed as milligrams of Rutin equivalent (mg RE) per 100 g of flour weight.

2.4.6. Total anthocyanin content

Total anthocyanin content was analyzed using the pH differential method (42). Cyanidin-3-glucoside was used as the reference, and results were expressed as mg of cyanidin-3-glucoside equivalent per 100 g fresh weight, using the following equation: TAC = DA × MW × DF × 1,000)/ε where TAC is total anthocyanin content, DA is difference in optical density (OD) value of extract at pH 1.0 and pH 4.5, MW is the molecular weight of cyanidin-3-glucoside, DF is dilution factor, and ε is molar absorbance coefficient of cyanidin-3-glucoside.

2.4.7. 2,2-diphenyl-1-picrylhydrazyl activity

Total antioxidant activity was obtained by the 2,2-diphenyl-2-picrylhydrazyl (DPPH) method (47) with some modifications. The working solution of DPPH was freshly prepared by diluting 3.9 mg of DPPH with 95% ethanol to get an absorbance of 0.856 ± 0.05 at 517 nm. The different concentration of the extract was mixed with 1.5 mL of working DPPH, and the absorbance of the mixture was immediately measured spectrophotometrically after 10 min. The total antioxidant activity of the extracted rice was expressed as mg BHA/g sample equivalent, obtained from the calibration curve.

$$\%\text{Inhibition of DPPH radical} = \left(A_{\text{control}} - A_{\text{sample}} \,/\, A_{\text{control}}\right) \times 100.$$

where $A_{control}$ is the absorbance of the control (without extract), and A_{sample} is the absorbance in the presence of the extract/standard.

2.4.8. 2,2′-azino-bis (3-ethylbenzothiazoline-6-sulphonic acid) activity

The total antioxidant capacity was determined by a colorimetric method (48) with a little modification. First, $ABTS^+$ solution was prepared and adjusted with pH 0.784 ± 0.01 with 80% ethanol at 734 nm. Then, 3.9 mL $ABTS^+$ cation solution was added to 1 mL of extracts and mixed thoroughly. The mixture was incubated for 6 min at room temperature and tested for absorbance at 734 nm. The results were expressed in terms of Trolox equivalent antioxidant capacity (TEAC, mM Trolox equivalents per 100 g dry weight).

$$\%\text{Inhibition of ABTS radical} = \left(A_{\text{control}} - A_{\text{sample}} \,/\, A_{\text{control}}\right) \times 100$$

2.5. Genomic DNA extraction and molecular profiling with SSR markers

The total genomic DNA was extracted from leaves of selections and the original mixed population of C14-8 as described by Murray and Thompson (49). PCR amplification was performed with 50 pairs of HvSSR markers spanning all 12 chromosomes (50) using a thermal cycler (Bio-Rad, United States). The thermal profile of PCR reactions was set as follows: initial denaturation at 94°C for 5 min followed by 35 cycles of denaturation at 94°C for 1 min, annealing temperature of 55°C for 1 min, extension at 72°C for 2 min, and a final extension at 72°C for 7 min. The PCR reaction volume of 10 µL was constituted using 50 ng genomic DNA, 1 × PCR buffer, 0.1 mM dNTP, 5 pmole each of forward and reverse primers, 2 mM $MgCl_2$, 0.2 units of Taq polymerase, and nuclease-free water. The amplified product was resolved in 3% metaphor agarose gel in 1X TAE buffer along with a 1 kb ladder for 1 h at 80 V (51). The gel was analyzed using a gel documentation system (Bio-Rad, United States).

2.5.1. Allele scoring

Different markers amplified a different number of bands for each genotype. Image Lab software (Bio-Rad, United States) was used to

determine the size of the bands based on differential migration relative to standard molecular weight markers (50–1,000 bp ladder, Thermo-Scientific, United States). Alleles were numbered from one to as many alleles as obtained for a particular marker x genotype combination based on their molecular weight from lowest to highest. The presence of a band for a particular allele was scored as "1," while the absence was scored as "0." The most polymorphic markers were determined based on their PIC value of ≥0.70 and polymorphic alleles of ≥6 as per Babu et al. (52).

3. Data analysis

Mean data of agro-morphological and biochemical traits were used for calculating descriptive statistics, cluster analysis, principal component analysis (PCA), and character association studies. PCA and character associations were carried out using the software package PAST [paleontological statistics software package for education and data analysis (53)]. Clustering of genotypes was performed based on the Euclidean distance and Ward's method, and the analysis was performed using SPSS 17Q15 software (IBM, United States), and the heatmap was generated using the R software package "gplots" (54). SSR marker-based clustering of genotypes based on the Euclidean distance and neighbor-joining method was performed using the software package PAST (53).

The Polymorphic Information Content (PIC) of each SSR marker was estimated using the following formula as per Botstein et al. (55).

$$1-\left(\sum_{i=1}^{n} p_{i^2}\right)-\sum_{i-1}^{n-1}\sum_{j=i+1}^{n} 2p_{i^2}p_{j^2}$$

where P_i = frequency of ith allele, and P_j = frequency of jth allele. Alleles whose frequency is less than 0.05 were considered rare alleles (34).

4. Results

4.1. Characterization for qualitative traits

A discernible variation was recorded for the qualitative traits in C14-8 selections especially for grain husk color (Figures 2B,C; Table 1). All genotypes exhibited light purple colored basal leaf sheath except for two selections such as 1 and 30 from group I, which showed distinct green basal leaf sheath color as well as light red decorticated grains. Thirteen accessions in group II exhibited higher intensity of anthocyanin coloration in their leaf sheaths, which was absent in the remaining selections. Only group I selections showed anthocyanin coloration on nodes, which was absent in all remaining groups. Out of the 20 selections, five were classified as very long-stemmed and the remaining 15 as long stemmed. Selections of groups I and II had white and purple apiculi, respectively, while groups III and IV exhibited yellowish and black apiculi, respectively. Grains of groups II and III were yellow husked, and grains of groups I and IV were brown husked, while the latter was golden furrowed. Selections of group II exhibited strong secondary branching while selections in the remaining groups had clustered secondary branching. Leaf senescence was medium in all except for two selections 1 and 30 from group I, which showed light leaf senescence. Similarly, the decorticated grain color of all selections was white except for selection C14-8-11-30 from group I, which expressed light red color.

4.2. Descriptive statistics and character associations

Our study assumes cultural and agro-evolutionary significance because the diverse grain color selections descended from a common ancestral traditional landrace, which has been on-farm maintained under the geographical isolation of the Andaman and Nicobar Islands. It is also imperative to understand the inter-trait correlation dynamics in such biologically unique populations for their utilization during indirect selection for trait improvement.

In general, the agro-morphological traits studied displayed a low coefficient of variation (%) except for the economically important GY (27.6), as shown in Table 2; Figure 3A, 4A,B. Among the biochemical traits, FC (27.7) and AnC (64.4) were highly variable, and traits AC (14.0), RSC (10.9), and TSC (10.0) were moderately variable while the remaining traits were least variable. It was also interesting to note that all the 22 traits using their mean values followed the normal distribution pattern, wherein the C14-8 mixed population possessed the central value (Figure 3B), indicating the polygenic nature of the population. Character association study of significant correlations revealed that PH was negatively correlated with four morphological traits of grain, such as TGW (−0.41), BRR (−0.44), MRR (−0.55), and HRR (−0.45), as shown in Table 3. DF was positively correlated with GY (0.49) and FlaC (0.47) but negatively correlated with RSC (−0.41). EBT was positively correlated with GL (0.67) and GB (0.67). TGW was found negatively correlated with RSC (−0.40). GL was positively correlated with GB (0.46), but GB was negatively correlated with GL/GB (−0.67) and FlaC (−0.41). BRR was positively associated with MRR (0.40), and in turn, MRR was positively correlated with HRR (0.75). AC was positively correlated with GC (0.41) and PhC (0.42). However, a negative correlation was observed between PhC and FlaC (−0.55; Table 3).

4.3. Diversity based on agro-morphological and biochemical traits

Data on agro-morphological and biochemical traits of 20 C14-8 selections and the original C14-8 population were subjected to cluster analysis based on the Euclidean distance-based genetic dissimilarity, and the heatmap generated is presented in Figure 3A. It grouped these 21 rice genotypes into three major clusters with 10 (C14-8-11-43, C14-8-11-59, C14-8-11-108, C14-8-11-143, C14-8-11-60, C14-8-11-31, C14-8 mix, C14-8-11-113, C14-8-11-61, C14-8-11-30), 9 (C14-8-11-32, C14-8-11-116, C14-8-11-118, C14-8-11-114, C14-8-11-115, C14-8-11-93, C14-8-11-1, C14-8-11-91, C14-8-11-92), and 2 (C14-8-11-90, C14-8-11-117) genotypes in Clusters I, II, and III, respectively. Cluster I was mainly represented by the C14-8 selections belonging to groups II and III of our classification of these 20 genotypes on the basis of grain husk color and apiculus color, whereas, Cluster II was mainly represented by the C14-8 selections belonging to groups I and IV. The first five PCs (principal components) of PCA explained a significant proportion of the variation (90.95%) present in agro-morphological and biochemical traits (Table 4). The first two PCs explained 63.3% of the total variation with PC1 and PC2 explaining 35.44% and 27.91%, respectively. Traits such as TSC, PH, BRR, FlaC, AC, EBT, PhC, GL/GB, HRR, MRR, TSC, BRR, ASA by DPPH, TGW, PhC, ASA by 2,2′-azino-bis (3-ethylbenzothiazoline-6-sulphonic acid; ABTS), AC, AD, EBT, and RSC contributed to PC1 and PC2, respectively (Figure 5). Distribution of genotypes and in-turn classification of the same by

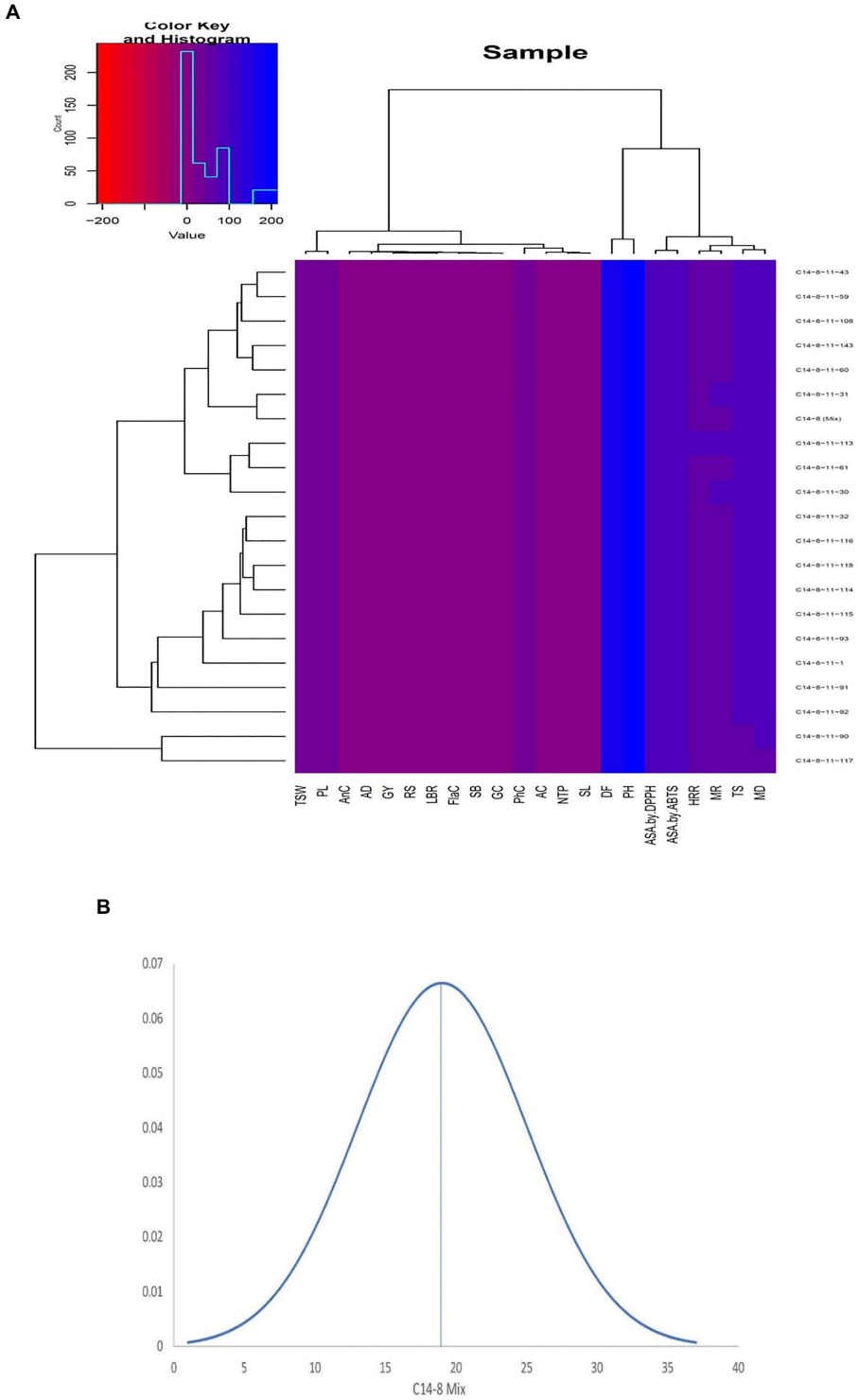


FIGURE 3
(A) Agro-morphological and bio-chemical traits heatmap of the 21 intra-varietal C14-8 selections in rice based on Euclidean distances. **(B)** Normal distribution curve of the means of all traits of C14-8 selections with respect to C14-8 mix population.

plotting the first two PCs against one another resulted in a similar classification as that of cluster analysis with genotypes under Cluster I (green colored) and Cluster II (red colored), as shown in Figure 5. Here, we can see that most of the genotypes are clustered in green and red oval groups, and only two are in the blue oval groups and hence explained 63.3% of the total variation with PC1 and PC2.

4.4. Cooking qualities and biochemical properties

In a bid to reveal the variability of C14-8 genotypes for quality attributes, physicochemical properties were studied and classified based on hierarchical cluster analysis. Cluster analysis is a statistical method

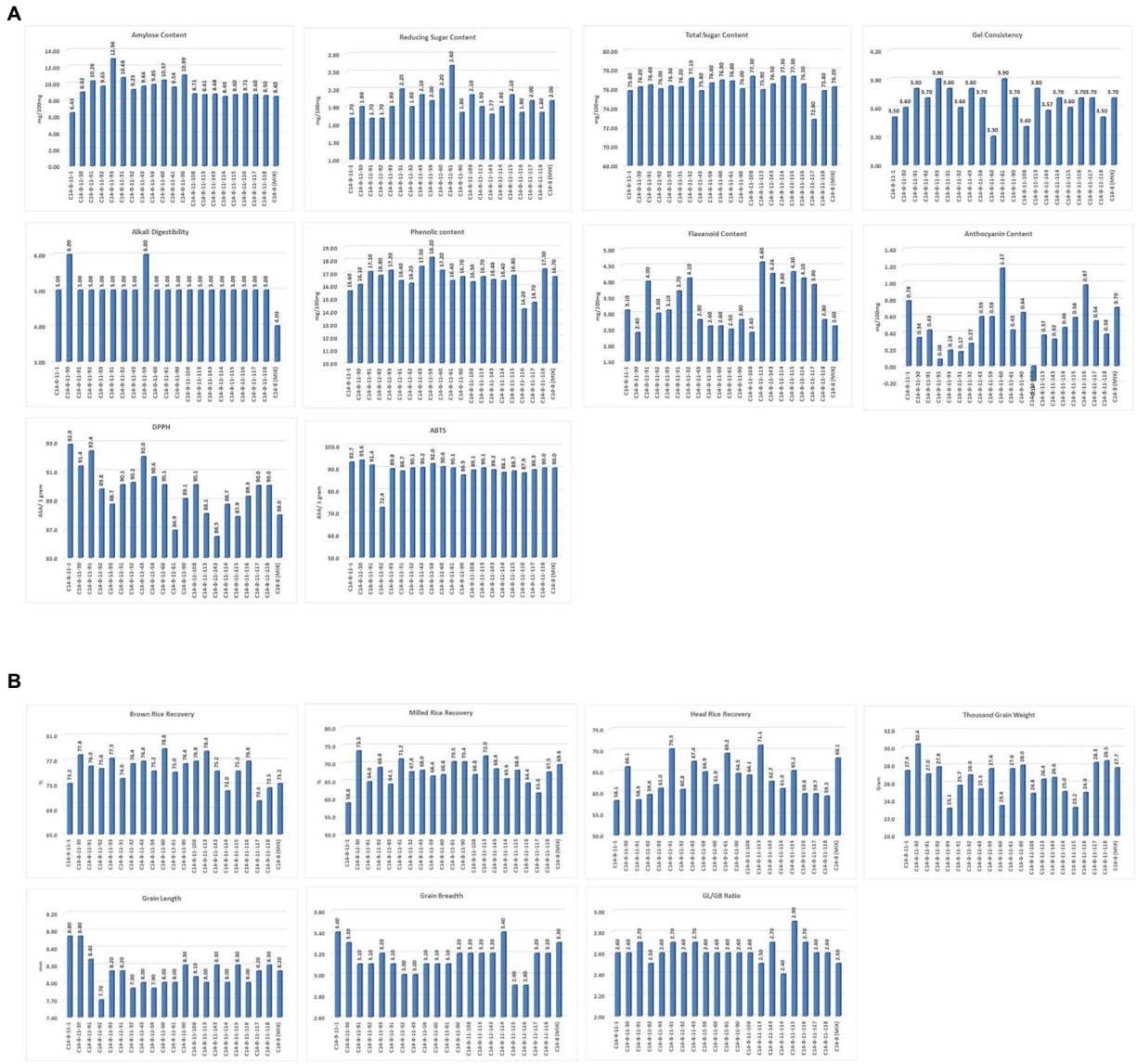


FIGURE 4
(A) Intra-Varietal variation for biochemical and cooking parameters in C14-8 selections. **(B)** Intra-Varietal variation for grain and milling parameters in C14-8 selections.

to convert various characteristics of objects to quantitative measures also called similarity distance and correspondingly to cluster them at relatively close distances into a category. We hypothesized that cluster analysis can be used to categorize the progenies having similar properties into a single group. As shown in Table 5, the 20 selections are classified based on factors influencing cooking qualities, phytochemical content, and carbohydrate and amylose content. On an overall basis, selections C14-8-11-108, C14-8-11-114, and C14-8-11-115 have been separated into different groups.

4.5. Molecular diversity

A total of 314 alleles were produced by 50 SSR markers, which ranged from 1 to 15 alleles per marker with an average of 6.28 (Supplementary Table 1). Out of 314 alleles, 64 alleles were found to be rare among the C14-8 selections. There were five markers with 0 PIC value. PIC values for the remaining markers ranged from 0.35 to 0.93, with an average of 0.67. Among the 50 HvSSR markers employed, 62% (31 markers) were most polymorphic with an average of 2.58 per linkage group. Among the linkage groups, chromosomes 2, 7, 9, and 11 harbored the most polymorphic loci (four markers) in the C14-8 population. On the other hand, none of them was present on chromosome 12. Clustering of genotypes based on Euclidean dissimilarity infers those selections of group II of the basic classification groups, which formed a separate group. Of the remaining, selections of group I and group IV formed separate clusters but interspersed with selections of group III. While three selections of group III aligned with group I, two selections got aligned with group IV. A similar classification was obtained except a genotype C14-8-11-59, which formed a solitary cluster when the dendrogram was constructed based on Jaccard's similarity coefficient. Notably, the molecular clustering of the selections through SSR markers almost conformed to the grouping based on grain husk coloration (Figure 6).

TABLE 3 Pearson's correlation coefficients among morphological and biochemical traits in C14-8 population.

	PH	DF	EBT	PL	TGW	GL	GB	GL/GB	GY	BRR	MRR	HRR	AC	RSC	TSC	GC	GC	AD	PhC	FlaC	AnC	ASA by DPPH
PH	1.00																					
DF	0.12	1.00																				
EBT	−0.09	0.05	1.00																			
PL	0.06	0.00	0.16	1.00																		
TGW	−0.41*	−0.14	0.33	0.01	1.00																	
GL	−0.06	0.12	0.67*	0.19	0.32	1.00																
GB	−0.15	0.02	0.47*	0.28	0.38	0.46*	1.00															
GL/GB	0.19	0.10	0.00	−0.23	−0.31	0.25	−0.67*	1.00														
GY	0.00	0.49*	−0.03	−0.05	−0.13	0.00	0.28	−0.22	1.00													
BRR	−0.44*	−0.22	−0.07	0.09	−0.30	−0.13	−0.37	0.16	−0.06	1.00												
MRR	−0.55*	−0.23	0.11	0.00	0.24	−0.14	−0.12	−0.04	−0.27	0.40*	1.00											
HRR	−0.45*	−0.31	−0.04	−0.03	0.01	−0.12	−0.08	0.04	−0.15	0.25	0.75*	1.00										
AC	0.08	−0.09	−0.19	0.25	−0.26	−0.33	−0.32	0.10	−0.11	0.33	0.22	0.21	1.00									
RSC	−0.07	−0.41*	−0.24	−0.31	−0.40*	−0.20	−0.18	0.12	−0.33	0.24	0.18	0.33	0.16	1.00								
TSC	0.18	−0.19	0.18	−0.11	−0.33	−0.12	−0.13	0.06	−0.11	0.29	0.11	0.10	0.05	−0.12	1.00							
GC	0.11	0.14	−0.04	0.30	−0.31	−0.11	−0.08	−0.06	0.13	0.26	0.15	0.11	0.41*	0.12	−0.20	1.00	1.00					
AD	−0.23	0.03	0.38	0.03	0.10	0.21	−0.15	0.24	−0.03	0.36	0.17	0.06	0.10	−0.03	0.04	−0.06	−0.06	1.00				
PhC	−0.26	−0.29	−0.14	−0.09	−0.05	−0.13	−0.01	0.01	−0.07	0.23	0.40*	0.32	0.42*	0.12	0.14	0.05	0.05	0.16	1.00			
FlaC	0.35	0.47*	−0.02	−0.06	−0.26	−0.13	−0.41*	0.29	0.22	−0.02	−0.19	−0.17	−0.17	−0.31	0.03	0.20	0.20	−0.14	−0.55*	1.00		
AnC	0.05	−0.05	−0.14	−0.17	0.12	0.02	−0.14	0.11	−0.38	−0.20	−0.26	0.00	−0.17	0.02	−0.09	−0.21	−0.21	−0.12	−0.13	0.20	1.00	
ASA by DPPH	−0.12	−0.24	−0.16	0.40	0.13	−0.03	0.16	−0.20	−0.30	−0.12	0.09	0.21	−0.22	0.16	−0.19	−0.31	−0.31	0.17	−0.15	−0.24	0.10	1.00
ASA by ABTS	−0.03	−0.09	0.29	−0.20	0.00	0.58	0.19	0.23	−0.11	0.07	−0.10	0.16	−0.13	0.23	−0.05	−0.10	−0.10	0.20	0.02	−0.06	0.27	−0.06

PH, Plant height (cm); DF, Days to 50% flowering; EBT, Ear bearing tillers per plant; PL, Panicle length (cm); TGW, 1,000-grain weight (gm); GL, Grain length (mm); Grain width (mm); GY, Grain yield (t/ha); BRR, Brown Rice Recovery (%); MRR, Milled rice recovery (%); HRR, Head Rice Recovery (%); AC, Amylose content (mg/100 mg); RSC, Reducing sugar content (mg/100 mg); TSC, Total sugar content (mg/100 mg); GC, Gel consistency; AD, Alcohol digestibility; PC, Phenolic content; FC, Flavonoid content (mg/100 mg); AnC, Anthocyanin content (mg/100 mg); ASA by DPPH, % Antioxidant scavenging activity/1 g by DPPH; ASA by ABTS, % Antioxidant scavenging activity/1 g by ABTS.

*Significance at 0.05 probability.

10.3389/fnut.2023.1088208

TABLE 4 Percent variance of first five principal components for quantitative traits of 20 rice genotypes selected from C14-8 population.

Principal components	Eigenvalue	Variation (%)	Cumulative variation (%)	Characters involved
1	60.28	35.44	35.44	TSC, PH, BRR, FlaC, AC, EBT, PhC, GL/GB,
2	47.47	27.91	63.35	HRR, MRR, TSC, BRR, ASA by DPPH, TGW, PhC, ASA by ABTS, AC, AD, EBT, RSC
3	18.27	10.74	74.09	ASA by ABTS, BRR, HRR, EBT, GL, PhC, AnC, DF, GY, AD, FlaC, RS, GL/GB, GB, GC
4	17.14	10.07	84.16	ASA by DPPH, HRR, ASA by ABTS, PH, PL, MRR, RSC, TSC, AnC, AD, GL, GB, GL/GB
5	11.54	6.79	90.95	HRR, MRR, PH, AC, PhC, FlaC, GC, DF, RSC, GY, GL/GB,

PH, Plant height (cm); DF, Days to 50% flowering; EBT, Ear bearing tillers per plant; PL, Panicle length (cm); TGW, 1,000-grain weight (gm); GL, Grain length (mm); GB, Grain breadth (mm); GY, Grain yield (t/ha); BRR, Brown Rice Recovery (%); MRR, Milled rice recovery (%); HRR, Head Rice Recovery (%); AC, Amylose content (mg/100 mg); RSC, Reducing sugar content (mg/100 mg); TSC, Total sugar content (mg/100 mg); GC, Gel consistency; AD, Alcohol digestibility; PC, Phenolic content; FC, Flavonoid content (mg/100 mg); AnC, Anthocyanin content (mg/100 mg); ASA1, % Antioxidant scavenging activity/1 g by DPPH; ASA2, % Antioxidant scavenging activity/1 g by ABTS.

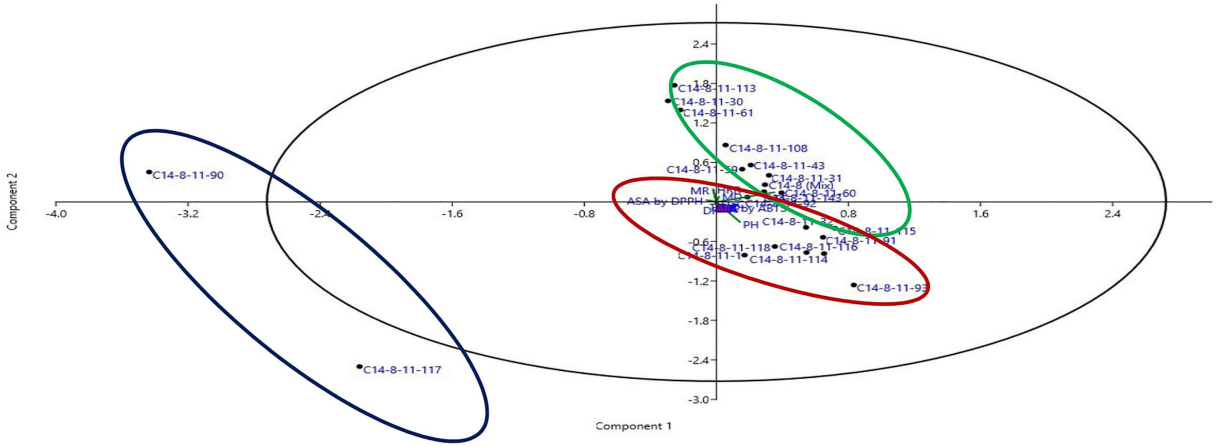


FIGURE 5
Two-dimensional scatter plot of PCA ordination using two most informative axis of 21 selected C14-8 rice lines obtained through PCA analysis.

5. Discussion

It is worth mentioning that the old landraces being grown and selected by farmers in their areas of adaptation for specific needs over the millennia have retained important traits and genes for crop improvement (9, 32). However, over the centuries, these traditional varieties have developed varying levels of heterogeneity (56, 57) plausibly due to manual admixtures as well as genetic events of recombination and mutations. In this context, it is interesting as well as imperative to understand the causes and types of intra-varietal variation for different phenotypic, biochemical, and molecular traits for genetic studies and breeding applications. Although C14-8 is grown as a single landrace variety, we observed typical open floret features and intra-varietal variation for the grain husk color, which prompted us to investigate the extent and pattern of diversity for different traits in these progenies and to project their implications in plant breeding. Although there are good numbers of studies on the diversity between varieties, pure lines, local cultivars, and/or landraces (31–36), systematic studies on intra-varietal variation are far more inadequate. Therefore, the dilution of genetic fidelity of landrace population due to inter-crossing over several decades of cultivation might reduce their productivity potential and popularity. Hence, the collection and genetic differentiation of landrace undertaken in the present study gain the importance to retain its on-farm diversity, varietal signatures, sustainability, and use as a donor for improvement of other varieties due to its salient traits, especially in light of socioeconomic requirements and emerging environmental challenges. In this context, Marone et al. (9) have also reviewed that traditional landraces have substantially contributed genes for stress adaptation in cereals.

This study perhaps reports a first-ever systematic and elaborate effort on the collection of diverse lines of a popular and promiscuous rice C14-8 followed by isolation, grouping, and identification of different types under geographical isolation of the Andaman Islands. The occurrence of distinct grain color types in the C14-8 population pool may be attributed to the manual admixture, spontaneous mutants, and/or occurrence of open florets, which indirectly permitted varying levels of intra-varietal as well as inter-varietal outcrossing resulting in a heterogeneous population undergoing natural selection over several decades. Here, it is pertinent to add that there is negligible chance of inter-varietal outcrossing due to significant phenological differences in flowering time between C14-8 (December month of the season) and other prominent varieties (before the November month of the season) owing to stringently photosensitive nature and significantly taller stature of tropical *japonica* C14-8 compared with all other *indica* rice varieties

TABLE 5 Cluster analysis of C14-8 variants based on cooking and nutritional quality parameters.

Parameters	Group-I	Group-II	Group-III	Group-IV	Group-V
Cooking qualities	C14-8-11-59	C14-8-11-93, C14-8-11-1	C14-8-11-30, C14-8-11-32, C14-8-11-115, C14-8-11-114, C14-8-11-116, C14-8-11-117, C14-8-11-118, C14-8-11-108	C14-8-11-91, C14-8-11-90, C14-8-11-31, C14-8-11-61, C14-8-11-92, C14-8-11-43, C14-8-11-60, C14-8-11-113, C14-8-11-143	-
Phytochemical content	C14-8-11-116	C14-8-11-1, C14-8-11-30, C14-8-11-91, C14-8-11-59	C14-8-11-92, C14-8-11-61, C14-8-11-108, C14-8-11-114, C14-8-11-115, C14-8-11-31, C14-8-11-93, C14-8-11-60, C14-8-11-118, C14-8-11-143	C14-8-11-32, C14-8-11-43, C14-8-11-117, C14-8-11-90, C14-8-11-113	-
Carbohydrate and amylose	C14-8-11-117	C14-8-11-93, C14-8-11-1	C14-8-11-30, C14-8-11-32, C14-8-11-118, C14-8-11-113, C14-8-11-116, C14-8-11-60	C14-8-11-91, C14-8-11-92, C14-8-11-31, C14-8-11-90, C14-8-11-43, C14-8-11-59, C14-8-11-61	C14-8-11-108, C14-8-11-114, C14-8-11-115, C14-8-11-143

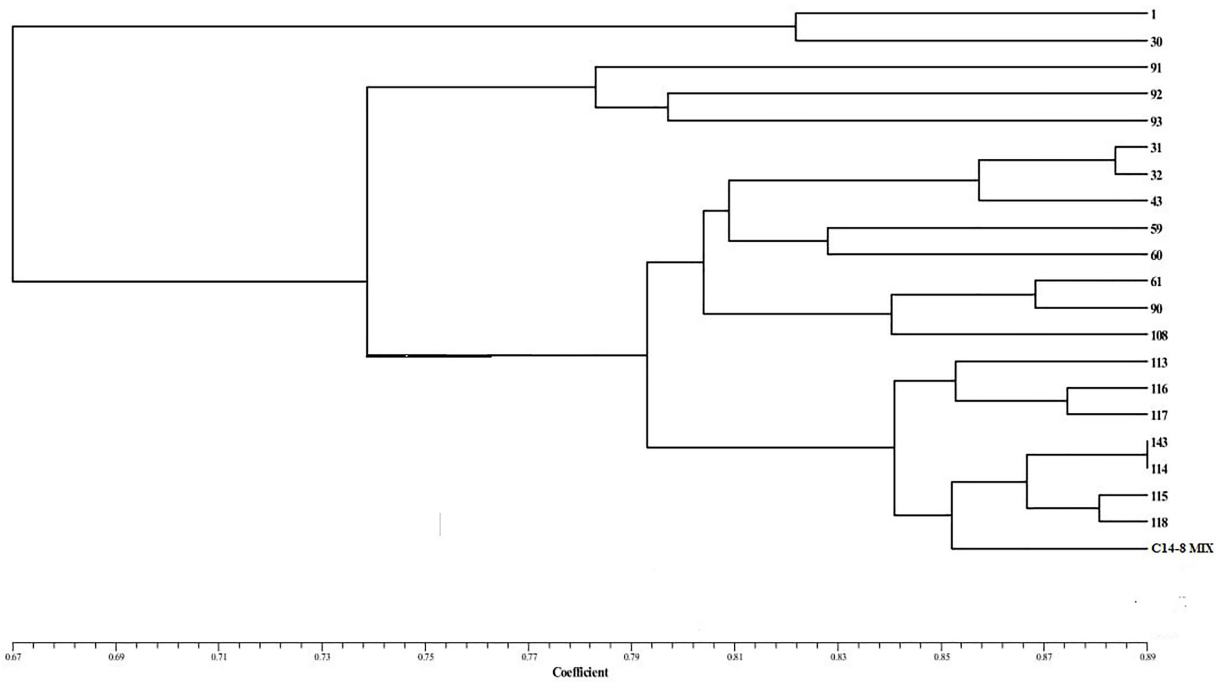


FIGURE 6
Dendrogram showing molecular diversity.

in these islands. In addition, the likely wide incompatibility at the *S5* locus (22) might have further reduced the probability of hybridization and recombination between C14-8 (tropical *japonica*) and local *indica* varieties. Therefore, it is quite probable that some spontaneous mutation or recombination at a few loci controlling grain color occurred in this open floret landrace, which could have floated in the population to cause intra-varietal variation. In support of our findings, according to Fujino et al. (58), while studying the ancestral implications of a Japanese rice variety, Kitaake has also concluded that the fast accumulation of pre-existing mutations in landraces might facilitate the adaptability of rice varieties to the local regions.

For rice grain husk color, Liu et al. (59) have fine-mapped *Pa-6*, a dominant gene on chromosome 6 governing the purple color of the apiculus, leaf sheath, and pericarp. Saitoh et al. (60) further reported that anthocyanin coloration in rice caused by the interplay of three basic genes, C (chromogen), A (activator), and P (distributor), is not only a morphological marker but also implicated in understanding of rice domestication from wild to cultivated type. Hence, our study was an endeavor to delve into these myriads and quantify genetic alteration and agro-cultural adaptation over decades in a promiscuous landrace population C14-8 under natural isolation of geographical remoteness. Similar to our attempt, a traditional Thai rice landrace Bue Chomee

exhibited significant genetic differentiation among Karen villages as revealed by microsatellite markers (61). Due to common genetic background but variation for a few loci, the landraces can also serve as ideal mapping populations for molecular mapping of useful traits. While the process of domestication has reduced the stigma exsertion rate and length of stigma and anther, it caused thicker and wider lemma and palea in cultivated varieties than the wild rice accessions (62). Therefore, a shift from outcrossing to selfing occurred during the domestication of cereal crops except for maize, pearl millet, and sorghum (63–65). C14-8 was also found to exhibit perennial nature due to a considerably longer growth duration of approximately 8 months and better regeneration after ratooning as also reported in wild rice accessions (66).

Open floret is a rare reproductive trait documented in self-pollinated cereals such as rice, which can permit inter-varietal and intra-varietal out-crossing resulting in genetic variation. However, this rare trait if present in the restorer or male sterile lines can also be a boon for achieving a higher seed set in hybrid rice breeding. We have previously reported the occurrence of this deviant floral feature in C14-8, which got approved for registration (IC0613963, INGR15014) in the Indian national genebank (40). The swelling of lodicules adjacent to the ovary causes flower opening in cereals at anthesis in autogamous cereals such as rice, wheat, and barley (67–69). Some studies associated with flower opening and ovary size in wheat indicated that the bigger the size of the ovary, the wider will be floret opening (70). It is also perceived that sometimes lack of self-pollination in self-pollinated cereals has influenced flower opening to facilitate cross-pollination for setting seed, wider genetic base, and adaptation (71). The association of purple apiculus with purple basal leaf sheath color in group II as found in our study was also reported earlier by Liu et al. (59).

It is important to mention that the diverse lines descended from a common ancestral traditional landrace, which has been on-farm maintained under the geographical isolation of the Andaman and Nicobar Islands. Furthermore, the genetic fidelity of the C14-8 population in relation to other rice varieties was also maintained because C14-8 is stringently photosensitive, extremely late flowering (150 days), and very tall (190 cm) rice varieties unlike all other photo-insensitive and early maturing rice varieties grown in these islands. Therefore, it is opined that the varying levels of diversification of traits recorded by us could be plausibly due to selection pressure and development of spontaneous variants/mutants arising in the C14-8 followed by the floating of the variants in the entire population facilitated by open floret nature.

The variation in grain yield and related traits is not only important for analysis but also equally important for grain milling and nutritional characteristics. Because the C14-8 rice variety has long served the remote and tropical warm and humid islands, which are perceived to be vulnerable to geo-climatic aberrations such as the devastating tsunami in 2004, dry spells, as well as frequent cyclonic storms. Furthermore, the island population requires to consume a cheaper source of calories such as rice nutritionally rich in terms of antioxidants, flavonoids, anthocyanins, etc., due to the tropical hot, radiant, and humid climate. Although we have found that C14-8 is nutritionally rich in terms of higher zinc and iron contents, it was imperative to understand the status and variation of these selections for nutritional parameters in the popular C14-8 rice. Interestingly, Eigenvalue revealed that most of the percentage genetic variation among the representative lines is explained by the invisible but important grain milling and nutritional traits.

Adaptation of japonicas to tropical islands has been revealed by Thomson et al. (23) who found most rice landraces belonging to *japonica* types in the isolated Island of Borneo, Indonesia. This region is also geographically and climatically near the tropical Andaman and Nicobar Islands where C14-8 tropical *japonica* rice has carved out its cultural niche among traditional rice farmers due to their resource constraints, especially during the initial years of settlement in these islands. Since C14-8 has been adapted for long and survived the biotic and abiotic pressures in the geographically isolated islands, efforts could be directed to fish out useful genes underlying multiple stress adaptation traits for their likely utilization for improving stress tolerance of rice under changing climate (9). It was also interesting to find that C14-8 selected lines in our study planted in a zinc deficient field exhibited differential symptoms to this abiotic stress (unpublished data), which also indicated the probable variation for zinc deficiency adaptation, which however needs further validation through systematic studies. Furthermore, intra-varietal variants could also serve as a good population for genetic and molecular mapping of useful traits due to near common genetic background. It is also important to know that the traits or combination of traits exhibited maximum contribution to the total genetic variation in C14-8. The information, thus, generated could be utilized for pinpointing responsible traits for direct selection from a farmer's perspective or else for identifying donors for genetic improvement of such a population. It is, therefore, interesting to look for such genotypes in the germplasm pool, which possess superiority for a maximum number of agronomic and biochemical traits. C14-8-11-33, C14-8-11-113, and C14-8-11-93 were identified as superior lines based on milling and biochemical traits. Our study also assumes cultural and agro-evolutionary significance because the diverse grain color selections descended from a common ancestral traditional landrace, which has been on-farm maintained under the geographical isolation of the Andaman and Nicobar Islands.

Descriptive statistics revealed good variability for traits such as GY, AC, FC, and AnC. Similarly, variability for biochemical traits can also be exploited for the improvement in grain quality. Such association studies between various traits have also been conducted previously (72–74), and the useful and easily traceable associations could be employed as an indirect selection for crop improvement.

Furthermore, we were also curious to know if the classification for grain husk color conformed to the cluster grouping based on the agro-morphological, biochemical, and molecular markers. Cluster analysis (CA) and PCA based on agro-morphological traits and biochemical traits resulted in broadly two groups except for the two selections forming a separate cluster. One of the two groups includes group I and group IV selections of the basic classification group, and the other group includes group II and group III selections. CA and PCA based on agro-morphological and biochemical traits support the basic classification, although they were grouped solely based on grain husk and apiculus color. A "correlated response" between the quantitative and qualitative traits and/or non-deliberated selection for phenotypic similarity during the selection process might have led to this correspondence. On the other hand, SSR marker profile-based CA grouped selections of group II into a separate cluster, and group I and group IV selections formed separate groups but interspersed with the selections of group III. It means the variability present in groups I, II, and IV is adequately represented by 50 SSR markers to classify them into separate clusters. However, employing additional markers would have possibly resulted in a separate group III cluster.

The extent of genetic diversity and degree of gene flow is influenced by anthropogenic activities, farming practices, prevailing climatic conditions (75), and geographical isolation (61). It is also

pertinent to add here that crop landraces may possess considerable trait richness due to genetic diversity because these represent an intermediate stage between wild species and cultivated crops, and therefore, these become a natural choice for crop improvement (25). In general, farmers are known for maintaining the purity and perpetuity of local varieties/landraces through various traditional practices (76). However, the pattern and extent of diversity in the C14-8 variety, particularly for distinct grain husk color as well as qualitative and quantitative traits are surprising. This leaves the question of whether the variety itself was a mixture of pure lines when originally introduced in the islands or if different pure lines might have been brought initially but eventually turned into a mixture later. One of the causes of high variation recorded in the tropical *japonica* C14-8 selections may be spontaneous mutation with an intended or unintended selection of rare alleles for different traits over the years. While working in 148 traditional and modern cultivars of *indica* and *japonica* rice, Hour et al. (21) in Taiwan have also observed that despite less variation among cultivars, *japonica* landraces exhibited higher genetic variation than *indica* landraces. In addition, since the C14-8 cultivar exhibits an open floret tendency, there is a high probability of outcrossing among rare variants for grain husk color and other cryptic traits. The promiscuous tendency in C14-8 predisposes it to constant introgression of different alleles from various landraces and/or varieties (76) albeit the stringent photosensitive nature and *japonica* status of this landrace, further reinforced by geographical isolation eliminates the probability of its genetic recombination with other rice varieties in these islands. Gao et al. have also concluded that intra-varietal variation among landraces was more pronounced due to a greater number of diverse alleles, especially at farmers' fields than the modern varieties.

Therefore, varying levels of diversification of traits recorded by us could be plausibly due to the genetic events of open floret-mediated recombination, mutations, and selection drift in C14-8 followed by selection pressure caused by climatic variations, marginal input conditions, and differing agronomic practices practiced by farmers' groups across scattered island ecosystem. It may also be pertinent to add that the expression of this unique floral attribute of this landrace might be confined to the specific microclimatic conditions of the Andaman and its expression beyond its natural habitat needs to be validated. Furthermore, the economic benefits for the commercial hybrid industry if any arising out of the genetic transfer of this novel trait may also be shared with the custodian farming community of these islands who are on farms conserving the landrace for the past several decades to ensure "conservation through compensation."

6. Conclusion

The intra-varietal genetic diversity revealed was through 22 agro-morphological and biochemical traits, and molecular markers can be attributed to either accumulated mutation coupled with intended or unintended human selection, which might have increased the frequency of rare alleles over time in the C14-8 population. The findings could be further useful in rice and other self-pollinated crops in understanding the evolutionary significance and relevance of naturally occurring promiscuous behavior from genomic and biological perspectives. Futuristically, it is also worthwhile to explore this population for genetic mapping of its agronomically adaptive traits because the individuals, although may be contrasting in few traits, share almost common genetic background. Although rice is a self-pollinated crop, rare landraces such as C14-8 due to open floret tendency may represent a heterogeneous gene pool having both agronomically favorable and unfavorable alleles, which need further validation. Our observations and findings could be a precursor to identifying, quantifying, and utilizing the intra-varietal genetic variation in self-pollinated landrace populations exhibiting allogamous behavior. The useful variability can be captured, fixed, and purified in the form of new varieties or can serve as a source of allele mining for climatic resilience and nutritional traits, especially for marginal areas. While the open floret trait in self-fertilizing cereals could be potentially useful for transfer in hybrid parental lines for achieving efficient hybrid seed production, it may pose bottlenecks for obtaining molecular signatures for variety identity and purity for practical purposes. In view of the above, the pure line selection varieties derived from such landraces might necessitate temporal and spatial isolation for maintaining their genetic fidelity for sustainable conservation and production system. Therefore, identification, purification, and popularization of agronomically and nutritionally superior selections from this landrace population will aid in the sustainable rice production and livelihood security of the island population in terms of healthy food and improved economic and social system without impacting the natural environment. This will lead to the food sustainability and nutritional security of a large population in the remote Andaman and Nicobar Islands.

Data availability statement

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

Author contributions

RG: conceptualization, methodology, project administration, and editing. PS: methodology (designed and conducted field experiment). KV: data analysis using spatial software. BR: molecular data analysis and original draft writing. KS: methodology and conducted lab experiment (molecular work). SS: methodology and conducted lab experiment (biochemical analysis). MS: methodology (molecular analysis using SSR markers). SZ: methodology (social interaction with farmers). KD: field experiment and collection of rice samples. SR: methodology, conducted field experiment, and validation of the data. JV: tabulation and data curation. SA: writing the original draft. SL: writing the original draft, reviewing, and editing. All authors contributed to the article and approved the submitted version.

Funding

The authors thank the Director, ICAR-Central Island Agricultural Research Institute, Port Blair for financial support to the project HORTCARISIL201200100146. We are also thankful to the Department of Biotechnology (DBT), Government of India for the financial grant (BT/PR6575/AG11/106/892/2012) for the "Application of molecular markers under accelerated crop improvement programme." We also acknowledge the funding support from the Bill and Melinda Gates

Foundation and IRRI, Philippines for the "Stress tolerant rice for poor farmers of Africa and South Asia" project.

Acknowledgments

The authors acknowledge the Indian Council of Agricultural Research for all-time 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/fnut.2023.1088208/full#supplementary-material

References

1. Gatto, M, de Haan, S, Laborte, A, Bonierbale, M, Labarta, R, and Hareau, G. Trends in varietal diversity of Main staple crops in Asia and Africa and implications for sustainable food systems. *Front Sustain Food Syst*. (2021) 5:626714. doi: 10.3389/fsufs.2021.626714

2. Lu, Y, Xu, Y, and Li, N. Early domestication history of Asian Rice revealed by mutations and genome-wide analysis of gene genealogies. *Rice*. (2022) 15:11–20. doi: 10.1186/s12284-022-00556-6

3. He, W, Chen, C, Xiang, K, Wang, J, Zheng, P, Tembrock, LR, et al. The history and diversity of rice domestication as resolved from 1464 complete plastid genomes. *Front Plant Sci*. (2021) 12:2669. doi: 10.3389/fpls.2021.781793

4. Oka, HI. *Origin of cultivated Rice*. Tokyo: Japan Scientific Societies Press/Elsevier (1988).

5. Wang, C, and Han, B. 20 years of Rice genomics research: from sequencing, functional genomics to quantitative genomics. *Mol Plant*. (2022) 15: 593–619. doi: 10.1016/j.molp.2022.11.016

6. Allaby, RG, Ware, RL, and Kistler, L. A re-evaluation of the domestication bottleneck from archaeogenomic evidence. *Evol Appl*. (2019) 12:29–37. doi: 10.1111/eva.12680

7. Londo, JP, Chiang, YC, Hung, KH, Chiang, TY, and Schaal, BA. Phylogeography of Asian wild rice, Oryza rufipogon, reveals multiple independent domestications of cultivated rice, Oryza sativa. *Proc Natl Acad Sci*. (2006) 103:9578–83. doi: 10.1073/pnas.0603152103

8. Raza, A, Razzaq, A, Mehmood, SS, Zou, X, Zhang, X, Lv, Y, et al. Impact of climate change on crops adaptation and strategies to tackle its outcome: A review. *Plan Theory*. (2019) 8:34. doi: 10.3390/plants8020034

9. Marone, D, Russo, MA, Mores, A, Ficco, D, Laidò, G, Mastrangelo, AM, et al. Importance of landraces in cereal breeding for stress tolerance. *Plan Theory*. (2021) 10:1267. doi: 10.3390/plants10071267

10. Roy, S, Marndi, BC, Mawkhlieng, B, Banerjee, A, Yadav, RM, Misra, AK, et al. Genetic diversity and structure in hill rice (Oryza sativa L.) landraces from the north-eastern Himalayas of India. *BMC Genet*. (2016) 17:107. doi: 10.1186/s12863-016-0414-1

11. Bouargalne, Y, Ben Mrid, R, Bouchmaa, N, Zouaoui, Z, Benmrid, B, Kchikich, A, et al. Genetic diversity for agromorphological traits, phytochemical profile, and antioxidant activity in Moroccan sorghum ecotypes. *Sci Rep*. (2022) 12:5895–13. doi: 10.1038/s41598-022-09810-9

12. Choudhury, B, Khan, ML, and Dayanandan, S. Genetic structure and diversity of indigenous rice (Oryza sativa) varieties in the eastern Himalayan region of Northeast India. *Springerplus*. (2013) 2:1–10. doi: 10.1186/2193-1801-2-228

13. Singh, YT. Genetic diversity and population structure in upland rice (Oryza sativa L.) of Mizoram, north East India as revealed by morphological, biochemical and molecular markers. *Biochem Genet*. (2019) 57:421–42. doi: 10.1007/s10528-018-09901-1

14. Elkelish, A., (Ed.). (2021). *Landraces-traditional variety and natural breed*. United Kingdom: Intechopen.

15. Fukuoka, S, Suu, TD, Ebana, K, Trinh, LN, Nagamine, T, and Okuno, K. Diversity in phenotypic profiles in landrace populations of Vietnamese rice: a case study of agronomic characters for conserving crop genetic diversity on farm. *Genet Resour Crop Evol*. (2006) 53:753–61. doi: 10.1007/s10722-004-4635-1

16. Langyan, S, Bhardwaj, R, Kumari, J, Jacob, SR, Bisht, IS, Pandravada, SR, et al. Nutritional diversity in native Germplasm of maize collected from three different fragile ecosystems of India. *Front Nutr*. (2022) 16:878269. doi: 10.3389/fnut.2022.812599

17. Henry, R. Genomics strategies for germplasm characterization and the development of climate resilient crops. *Front Plant Sci*. (2014) 5:68. doi: 10.3389/fpls.2014.00068

18. Kyratzis, AC, Nikoloudakis, N, and Katsiotis, A. Genetic variability in landraces populations and the risk to lose genetic variation. The example of landrace 'Kyperounda'and its implications for ex situ conservation. *PLoS One*. (2019) 14:e0224255. doi: 10.1371/journal.pone.0224255

19. Rao, VR, and Hodgkin, T. Genetic diversity and conservation and utilization of plant genetic resources. *Plant Cell Tissue Organ Cult*. (2002) 68:1–19. doi: 10.1023/A:1013359015812

20. Teklu, Y, and Hammer, K. Farmers' perception and genetic erosion of tetraploid wheats landraces in Ethiopia. *Genet Resour Crop Evol*. (2006) 53:1099–113. doi: 10.1007/s10722-005-1145-8

21. Hour, AL, Hsieh, WH, Chang, SH, Wu, YP, Chin, HS, and Lin, YR. Genetic diversity of landraces and improved varieties of rice (*Oryza sativa* L.) in Taiwan. *Rice*. (2020) 13:82–12. doi: 10.1186/s12284-020-00445-w

22. Mi, J, Li, G, Xu, C, Yang, J, Yu, H, Wang, G, et al. Artificial selection in domestication and breeding prevents speciation in rice. *Mol Plant*. (2020) 13:650–7. doi: 10.1016/j.molp.2020.01.005

23. Thomson, MJ, Polato, NR, and Prasetiyono, J. Genetic diversity of isolated populations of Indonesian landraces of Rice (Oryza sativa L.) collected in East Kalimantan on the island of Borneo. *Rice*. (2009) 2:80–92. doi: 10.1007/s12284-009-9023-1

24. Islam, MZ, Siddique, MA, Akter, N, Prince, MFRK, Islam, MR, Anisuzzaman, M, et al. Morpho-molecular divergence of restorer lines for hybrid rice (Oryza sativa l.) development. Cereal res. *Cereal Res Commun*. (2019) 47:531–40. doi: 10.1556/0806.47.2019.29

25. Li, FP, Lee, YS, Kwon, SW, Li, G, and Park, YJ. Analysis of genetic diversity and trait correlations among Korean landrace rice (Oryza sativa L.). *Genet Mol Res*. (2014) 13:6316–31. doi: 10.4238/2014.April.14.12

26. Gautam, RK, Singh, PK, Sakthivel, K, Srikumar, M, Kumar, N, Kumar, K, et al. Analysis of pathogenic diversity of the rice bacterial blight pathogen (Xanthomonas oryzae pv. Oryzae) in the Andaman Islands and identification of effective resistance genes. *J Phytopathol*. (2015) 163:423–32. doi: 10.1111/jph.12338

27. Lu, TL. The occurrence of cereal cultivation in China. *Asian Perspect*. (2006) 45:129–58. doi: 10.1353/asi.2006.0022

28. Sakthivel, K, Kumar, A, Gautam, RK, Manigundan, K, Laha, GS, Velazhahan, R, et al. Intra-regional diversity of rice bacterial blight pathogen, Xanthomonas oryzae pv. Oryzae, in the Andaman Islands, India: revelation by pathotyping and multilocus sequence typing. *J Appl Microbiol*. (2020) 130:1259–72. doi: 10.1111/jam.14813

29. Mittelbach, GG, Schemske, DW, Cornell, HV, Allen, AP, Brown, JM, Bush, MB, et al. Evolution and the latitudinal diversity gradient: speciation, extinction and biogeography. *Ecol Lett*. (2007) 10:315–331. doi: 10.1111/j.1461-0248.2007.01020.x

30. Liu, L, Zou, Z, Qian, K, Xia, C, He, Y, Zeng, H, et al. Jasmonic acid deficiency leads to scattered floret opening time in cytoplasmic male sterile rice Zhenshan 97A. *J Exp Bot*. (2017) 68:4613–25. doi: 10.1093/jxb/erx251

31. Jasim Aljumaili, S, Rafii, MY, Latif, MA, Sakimin, SZ, Arolu, IW, and Miah, G. Genetic diversity of aromatic rice germplasm revealed by SSR markers. *Biomed Res Int*. (2018) 2018:1–11. doi: 10.1155/2018/7658032

32. Cui, D, Tang, C, Li, J, A, X, Yu, T, Ma, X, et al. Genetic structure and isolation by altitude in rice landraces of Yunnan, China revealed by nucleotide and microsatellite marker polymorphisms. *PLoS One*. (2017) 12:e0175731–17. doi: 10.1371/journal.pone.0175731

33. Das, B, Sengupta, S, Parida, SK, Roy, B, Ghosh, M, Prasad, M, et al. Genetic diversity and population structure of rice landraces from eastern and north eastern states of India. *BMC Genet*. (2013) 14:1–14. doi: 10.1186/1471-2156-14-71

34. Jain, S, Jain, RK, and McCouch, SR. Genetic analysis of Indian aromatic and quality rice (Oryza sativa L.) germplasm using panels of fluorescently-labeled microsatellite markers. *Theor Appl Genet*. (2004) 109:965–77. doi: 10.1007/s00122-004-1700-2

35. Mathure, S, Shaikh, A, Renuka, N, Wakte, K, Jawali, N, Thengane, R, et al. Characterisation of aromatic rice (Oryza sativa L.) germplasm and correlation between their agronomic and quality traits. *Euphytica*. (2011) 179:237–46. doi: 10.1007/s10681-010-0294-9

36. Sanni, KA, Fawole, I, Ogunbayo, SA, Tia, DD, Somado, EA, Futakuchi, K, et al. Multivariate analysis of diversity of landrace Rice Germplasm. *Crop Sci*. (2012) 52:494–504. doi: 10.2135/cropsci2010.12.0739

37. Zeven, AC, and Zhukovsky, PM. *Dictionary of cultivated plants and their Centres of diversity—excluding ornamentals, Forest trees and lower plants*. Wageningen: Centre for Agricultural Publishing and Documentation (1975).

38. Vanniarajan, C, Vinod, KK, and Pereira, A. Molecular evaluation of genetic diversity and association studies in rice (*Oryza sativa L.*). *J Genet*. (2012) 91:9–19. doi: 10.1007/s12041-012-0146-6

39. Balakrishnan, NP, and Ellis, JL. Andaman and Nicobar Islands In: NP Hajra Balakrishnanet al, editors. *Flora of India, part 1. Botanical Kolkota: survey of India*, vol. *1*. Calcutta, India: (1996). 523–38.

40. Gautam, RK, Singh, PK, Zamir Ahmed, SK, Kumar, N, Singh, AK, Sakthivel, K, et al. ANR38 (IC 0613963, INGR15014), a rice (Oryza sativa) germplasm with open floret. *Ind J Plant Genet Resources*. (2016) 29:201.

41. Shobha, NR, Subba Rao, LV, and Viraktamath, BC. *National Guidelines for the conduct of tests for distinctness, uniformity and stability: Rice (Oryza sativa L), Directorate of Rice Research*. Hyderabad: Hyderabad, Andhra Pradesh, India: (2006). 33 p.

42. Sadasivam, S., and Manikam, A., (1992). *Biochemical methods*. Wiley Eastern Ltd., Madras, pp. 246–248.

43. Juliano, BO. Criteria and tests for rice grain qualities In: BO Juliano, editor. *Rice chemistry and technology*. *2nd* ed. MN: American Association of Cereal Chemists Inc (1985). 43–524.

44. Cagampang, CB, Perez, CM, and Juliano, BO. A gel consistency test for eating quality of rice. *J Sci Food Agric*. (1973) 24:1589–94. doi: 10.1002/jsfa.2740241214

45. Singleton, VL, Orthofer, R, and Lamuela-Raventos, R. Analysis of total phenols and other oxidation substrates and antioxidants by means of FC reagent. *Methods Enzymol*. (1999) 299:152–78. doi: 10.1016/S0076-6879(99)99017-1

46. Bao, JS, Cai, YS, Wang, G, and Corke, H. Anthocyanins, flavonols, and free radical scavenging activity of Chinese bayberry (Myrica rubra) extracts and their color properties and stability. *J Agri Food Chem*. (2005) 53:2327–32. doi: 10.1021/jf048312z

47. Rattanachitthawat, S, Suwannalert, P, Riengrojpitak, S, Chaiyasut, CB, and Pantuwatana, S. Phenolic content and antioxidant activities in red unpolished Thai rice prevents oxidative stress in rats. *J Med Plant Res*. (2010) 4:796–801. doi: 10.5897/JMPR10.067

48. Re, R, Pellengrini, N, Proteggente, A, Pannala, A, Yang, M, and Rice-Evan, C. Antioxidant activity applying on improved ABTS radical cation decolorization assay. *Free Radic Biol Med*. (1999) 26:1231–7. doi: 10.1016/S0891-5849(98)00315-3

49. Murray, MG, and Thompson, WF. Rapid isolation of high molecular weight plant DNA. *Nucleic Acids Res*. (1980) 8:4321–6. doi: 10.1093/nar/8.19.4321

50. Singh, H, Deshmukh, RK, Singh, A, Singh, AK, Gaikwad, K, Sharma, TR, et al. Highly variable SSR markers suitable for rice genotyping using agarose gels. *Mol Breed*. (2010) 25:359–64. doi: 10.1007/s11032-009-9328-1

51. Sambrook, J, Fritschi, EF, and Maniatis, T. *Molecular cloning: A laboratory manual*. New York: Cold Spring Harbor Laboratory Press (1989).

52. Babu, BK, Meena, V, Aggarwal, V, and Aggarwal, PK. Population structure and genetic diversity of Indian and exotic rice (Oryza sativa L) accessions using SSR markers. *Mol Biol Rep*. (2014) 41:4329–39. doi: 10.1007/s11033-014-3304-5

53. Hammer, Ø, Harper, DA, and Ryan, PD. PAST: paleontological statistics software package for education and data analysis. *Palaeontol Electron*. (2001) 4:9.

54. Warnes, G.R., Bolker, B., Bonebakker, L., Gentleman, R., Huber, W., Liaw, A, et al, (2009). *Gplots: Various R programming tools for plotting data. R package version*. 2, GitHub.

55. Botstein, D, White, RL, Skolnick, M, and Davis, RW. Construction of a genetic linkage map in man using restriction fragment length polymorphisms. *Am J Hum Genet*. (1980) 32:314–31. PMID: 6247908

56. Frankel, OH, Soulé, ME, and Swingland, I. *Conservation and evolution*, vol. *16* United United Kingdom: Cambridge University Press, Hardback£ 25. Paperback£ 7.95. Oryx (1982). 358 p.

57. Frankel, OH, Brown, AH, and Burdon, JJ. *The conservation of plant biodiversity* United Kingdom: Cambridge University Press (1995).

58. Fujino, K, Obara, M, Ikegaya, T, and Tamura, K. Genetic shift in local rice populations during rice breeding programs in the northern limit of rice cultivation in the world. *Theor Appl Genet*. (2015) 128:1739–46. doi: 10.1007/s00122-015-2543-8

59. Liu, X, Sun, X, Wang, W, Ding, H, Liu, W, Li, G, et al. Fine mapping of pa-6 gene for purple apiculus in rice. *J Plant Biol*. (2012) 55:218–25. doi: 10.1007/s12374-011-0276-z

60. Saitoh, K, Onishi, K, Mikami, I, Thidar, K, and Sano, Y. Allelic diversification at the C (OsC1) locus of wild and cultivated rice: nucleotide changes associated with phenotypes. *Genetics*. (2004) 168:997–1007. doi: 10.1534/genetics.103.018390

61. Pusadee, T, Jamjod, S, Chiang, YC, Rerkasem, B, and Schaal, BA. Genetic structure and isolation by distance in a landrace of Thai rice. *Proc Natl Acad Sci*. (2009) 106:13880–5. doi: 10.1073/pnas.0906720106

62. Uga, Y, Fukuta, Y, Ohsawa, R, and Fujimura, T. Variations of floral traits in Asian cultivated rice (Oryza sativa L.) and its wild relatives (O. rufipogon Griff.). *Breed Sci*. (2003) 53:345–52. doi: 10.1270/jsbbs.53.345

63. Bandaru, R. Stigma Exsertion in Rice (Oryza Sativa L.). *Acta Scientific Agric*. (2018) 2:12–22.

64. Evans, L.T., (1993). *The domestication of crop plants. In crop evaluation, adaptation and yield*. Cambridge University Press, New York. P. 62–112.

65. Zheng, Z, Xiao, M, Lin, F, Ge, T, and Sun, H. Genetic analysis and gene mapping of a low stigma exposed mutant gene by high throughput sequencing. *PLoS One*. (2018) 13:e0186942. doi: 10.1371/journal.pone.0200397

66. Oka, HI, and Sano, Y In: YP Altukhov, editor. *Differentiation of Oryza perennis populations in adaptive strategy*. Moscow: Mir Publications (1981). 68–85.

67. De Vries, AP. Flowering biology of wheat, particularly in view of hybrid seed production—a review. *Euphytica*. (1971) 20:152–70. doi: 10.1007/BF00056076

68. Frankel, R, and Galun, E. Allogamy In: . *Pollination mechanisms, reproduction and plant breeding*. Berlin, Heidelberg: Springer (1977). 79–234.

69. Virmani, SS. *Heterosis and hybrid rice breeding*. Berlin: Springer-Verlag (1994).

70. Guo, L, Zhuang, J, Qiu, F, Gandhi, H, Kadaru, S, Jon, E, et al. Genome-wide association study of outcrossing in cytoplasmic male sterile lines of rice. *Sci Rep*. (2017) 7:3223. doi: 10.1038/s41598-017-03358-9

71. Okada, T, Ridma, JEA, Jayasinghe, M, Nansamba, M, Baes, M, Warner, P, et al. Unfertilized ovary pushes wheat flower open for cross-pollination. *J Exp Bot*. (2018) 69:399–412. doi: 10.1093/jxb/erx410

72. Li, F, Xie, J, Zhu, X, Wang, X, Zhao, Y, Ma, X, et al. Genetic basis underlying correlations among growth duration and yield traits revealed by GWAS in rice (Oryza sativa L.). *Front Plant Sci*. (2018) 9:650. doi: 10.3389/fpls.2018.00650

73. Oladosu, Y, Rafii, MY, and Magaji, U. Genotypic and phenotypic relationship among yield components in Rice under tropical conditions. *Biomed Res Int*. (2018) 2018:1–10. doi: 10.1155/2018/8936767

74. Zia-Ul-Qamar, AAC, Muhammad, A, Muhammad, R, and Ghulam, RT. Association analysis of some yield influencing traits in aromatic and non-aromatic rice. *Pak J Bot*. (2005) 37:613–27.

75. Roy, S, Banerjee, A, Mawkhlieng, B, Misra, AK, Pattanayak, A, Harish, GD, et al. Genetic diversity and population structure in aromatic and quality rice (Oryza sativa L.) landraces from north-eastern India. *PLoS One*. (2015) 10:e0129607. doi: 10.1371/journal.pone.0129607

76. Zeven, AC. Traditional maintenance breeding of landraces: 1 data by crop. *Euphytica*. (2000) 116:65–85. doi: 10.1023/A:1004089816030

frontiers | Frontiers in Nutrition

TYPE Review
PUBLISHED 21 March 2023
DOI 10.3389/fnut.2023.1124409

Check for updates

OPEN ACCESS

EDITED BY
Sapna Langyan,
National Bureau of Plant Genetic Resources (ICAR), India

REVIEWED BY
Isabel Castanheira,
Instituto Nacional de Saúde Doutor Ricardo Jorge (INSA), Portugal
Balaram Mohapatra,
Gujarat Biotechnology Research Centre (GBRC), India
Shubha Rani Sharma,
Birla Institute of Technology, Mesra, India

*CORRESPONDENCE
Debasish Kar
✉ debasish.bios@gmail.com

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

RECEIVED 15 December 2022
ACCEPTED 27 February 2023
PUBLISHED 21 March 2023

CITATION
Mondal DD, Chakraborty U, Bera M, Ghosh S and Kar D (2023) An overview of nutritional profiling in foods: Bioanalytical techniques and useful protocols. *Front. Nutr.* 10:1124409. doi: 10.3389/fnut.2023.1124409

COPYRIGHT
© 2023 Mondal, Chakraborty, Bera, Ghosh and Kar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

An overview of nutritional profiling in foods: Bioanalytical techniques and useful protocols

Deb Duhita Mondal[1], Ushashi Chakraborty[1], Manotosh Bera[1], Subhrojyoti Ghosh[1] and Debasish Kar[2]*

[1]Department of Biotechnology, Heritage Institute of Technology, Kolkata, India, [2]Department of Biotechnology, Ramaiah University of Applied Sciences, Bengaluru, India

Maintaining a nutritious diet is essential for humans if they want to live a healthier life. Several food businesses and food safety organizations play a significant role and offer useful ways for improving nutritional quality that assists consumers in making informed selections. Making poor food choices and consuming unhealthy meals are the main causes of non-communicable diseases (NCDs). Nutritional profiling (NP) models are developed to evaluate the nutritional value, calorie content, and the amount of micronutrients and macronutrients contained in a given food accompanied by additional details on the nutritional anomaly provided by published standard nutrients and nutritional databases. To construct an ideal nutritional model that can facilitate food consumption, bioanalytical methods such as chromatography, microscopic techniques, molecular assays, and metabolomics can be applied. With the use of these technologies, one can learn more about the health advantages of nutrition and how to prevent disease. A wider element of NP is also provided by the developing technologies in the area of nutrition research, such as nanotechnology, proteomics, and microarray technology. In this review, we are focusing on the different bioanalytical techniques and the various protocols of NP and their application and refinement of the models. We have evaluated various NP techniques currently used in the food industry for the detection of different components present in food items.

KEYWORDS
nutrition, nutritional profiling, chromatography, microscopic, health, consumers

1. Introduction

The primary cause of any non-communicable diseases (NCDs) across the world, with most cases in European countries, is an unhealthy diet, among other risk factors. An unbalanced diet could be established by taking into account many factors. In the last few decades, there has been a noticeable rise in overeating, and by 2025, it is predicted that one-fifth of persons worldwide will be obese. Finding and obtaining healthy foods will be challenging due to a variety of circumstances, including finances, individual preferences, different cultural traditions, geographical regions, and other environmental concerns such as climate change. The vast population lacks a proper dietary plan and instead consumes processed foods that are rich in fats, carbohydrates, and sodium (1). According to the WHO European Food and Nutrition Action Plan, food consumption that is high in energy or low in micronutrients and non-alcoholic drinks ought to be limited to a balanced diet to meet food targets for the vast population. At a conference held in 2020, health promotion and prevention of NCDs were promised, along with the development of common policies and implementation of standardized nutrition profile models (Figure 1). The nutritional value of a given product has been evaluated using a variety of nutritional profiling (NP) models based on its nutrient composition.

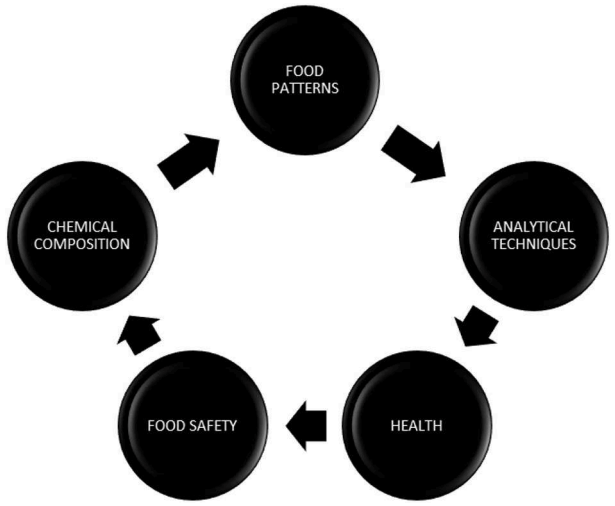


FIGURE 1
Nutritional profiling. The interrelationship of all the components that involve the development of the nutritional profiling model shows their interdependence with each other.

Several food firms have formed their nutritional standards by developing various models to inform the development of their products or highlight their healthy options (2). The nutrient profiling models, created for evaluating the nutritional content of meals, have developed into a crucial instrument for public policy. The energy density and the nutritional density of food are typically inversely correlated. NP models seek to discover nutritious foods that are high in nutrients and classify them based on their nutritional value. In developed nations, the nutrition profile models provide the scientific groundwork for a variety of educational, labeling, regulatory, and tax policies. Numerous front-of-pack icons and logos that convey a specific message are based on NP models (3).

The Food and Agriculture Organization of the UN (2007) strongly approves the consumption of certain food groups that have been associated with health-protective benefits, including fruits and vegetables, whole grains, legumes, and nuts. In contrast, studies on nutrition over a decade confirm the increased intake of beneficial nutrients such as vitamins, minerals, and fibers. The reason for this is the nutrient profile and the most common labeling schemes that provide nutritional information on the front of food packages, reducing the consumption of such unfavorable foods (4). EFSA's Panel on Nutrition Products and Nutrition and Allergies found that the NP of certain foods did not comply with reference intakes of nutrients. The comparison of the nutritional value of all foods and communicating their importance is made possible by the declaration of energy, macronutrient and micronutrient content that is established, and its relevant reference values among the general population (5). Finally, the strategy of using nutrient profiles is detrimental to some specific foods for which it is impractical to reconstitute in order to meet the threshold set for the same profile, resulting in being labeled "unhealthy", and despite being present in significant amounts, it plays a beneficial nutritional role.

2. Need for nutritional profiling

Several definitions have been developed over the years due to the potential of NP as a tool for evaluating specific foods in terms of their contribution to healthy eating patterns. The generally accepted definition, put forth by WHO in 2015, is "the science of classifying or ranking foods following their nutritional makeup for reasons linked with disease prevention and health promotion." NP is a phrase that has been connected to the composition of a food or diet in a context that is a helpful tool to help customers select healthier foods for themselves (6). The NP model can be used to improve the nutritional quality of diets by identifying foods that may constitute healthy and unhealthy diets and assessing the nutritional quality of foods rather than diets. Nutrient profiling is now used in a variety of nutrition policy applications worldwide, and the diversity of NP models has increased significantly in recent years. NP models are being used for a variety of purposes, such as helping customers choose healthier foods through food labeling systems, deciding which food items should be sold in schools, developing regulations for health or nutrition claims, and limiting the marketing of food to children (Table 1). The proliferation of NP models around the world, together with their numerous applications and specificities, might raise the likelihood of model differences and confuse regulators, producers, and consumers. The main goals of this article are to describe the process of developing an NP model and the various methods used for its validation, as well as to identify the potential role of nutrient profiling applications in promoting healthier food choices. This is done while also taking into account initiatives established in recent years concerning the development and application of various nutrient profiling tools. This cutting-edge review aims to improve the knowledge of NP models and assist policymakers in selecting an appropriate model once the implementation of nutrition-related policies would benefit from the use of such models. NP models serve as an excellent tool for decision-making in the area of public health nutrition interventions in order to help consumers choose healthier foods.

3. Analytical techniques for nutritional profiling

3.1. Chromatographic techniques

The discrimination of proteins with extreme physicochemical features is hampered by poor throughput, very restricted dynamic range, and other severe restrictions of conventional analytical methods. As a result, methods based on chromatography have been developed for the pre-separation of proteins and peptides. Chromatography is a laboratory technique for the separation of a mixture into its components. The combination passes through a system (a column, a capillary tube, a plate, or a sheet) in which a substance known as the stationary phase is fixed after being dissolved in a fluid solvent (gas or liquid) termed the mobile phase. The name of the method is either due to gas–solid chromatography or gas–liquid chromatography depending on the nature of the stationary phase used. The components of the mixture travel at varying apparent velocities in the mobile fluid, which causes them to separate because they tend to have different affinities for the

TABLE 1 Several schemes developed for nutritional model implementation across the world.

Schemes	Food categories	Aim
Pan American Health Organization model (USA)	Mostly processed and ultra-processed foods	Classifies foods and beverages based on excessive sugar content and fats and serves various strategies
Health Star Rating (Australia and New Zealand government)	Retail foods and beverages	Allow food comparison
Multiple traffic Light (United Kingdom)	Processed foods	Provides clarity on food characteristics
Mexican Committee of Nutrition Experts (Mexico)	Basic and non-basic foods	Selection of healthier food options

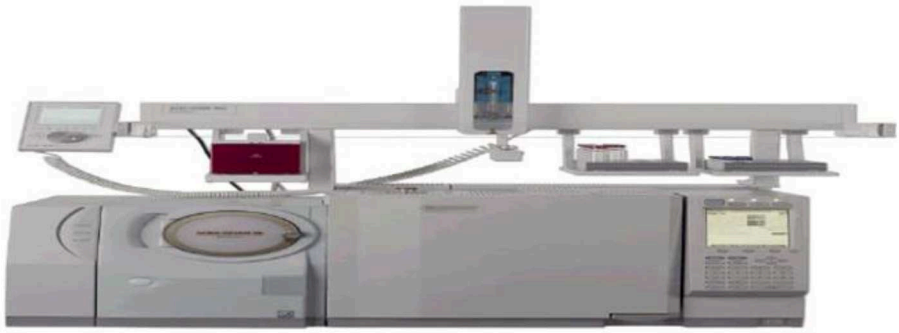


FIGURE 2
Gas chromatography. A typical GC-MS tool is used for the analysis of food for determining the food quality and evaluating food's nutritional benefits.

stationary phase and are held for various amounts of time based on their interactions with its surface sites. The differential partitioning between the mobile and stationary phases serves as the foundation for the separation. This partition coefficient can be expressed as $K_x = [C]_s/[C]_m$, where $[C]_s$ and $[C]_m$ are concentrations of stationary and mobile phases, respectively. Chromatography can be used at different points in the food chain to assess food quality and find additives, pesticides, and other dangerous contaminants. Chromatography enables the food business to deliver precise details about the nutrients in a specific product as well as many other things (7). Many food businesses use various additives and preservatives to try to improve their manufacturing process, which necessitates additional testing procedures to guarantee the safety of their goods. However, chromatography is by far the most adaptable technique, and many businesses can perform the majority of the necessary tests using chromatography equipment.

3.1.1. Gas chromatography

The earliest established chromatographic technique, gas chromatography (GC) analysis, which is being used today, forms the foundation of a traditional method for evaluating the quality of food. GC is a popular method of chromatography used in analytical chemistry for separating and studying substances that may be evaporated without decomposing. It is frequently used to determine a substance's purity or to separate the various ingredients in a combination. GC may be used in preparative chromatography to separate pure substances from a mixture. An inert or nonreactive gas continuously flows down a small tube known as the column, which is the foundation of a gas chromatograph, carrying the vaporized sample through it. Depending on their chemical and physical characteristics and the interactions they have with the stationary phase, the filling or lining of the column, various components of the sample move through it at different speeds. The popularity of gas chromatography is due to the unique combination of its high sensitivity, wide dynamic concentration range, excellent selectivity, and resolution (8). GC is widely utilized to research compounds such as sterols, oils, low-chain fatty acids, aroma components, and other contaminants such as different pesticides, pollutants from the industry, and particular classes of medications in food (Figure 2). Food quality is determined by its origin, chemical composition, adequate physical properties (e.g., texture, color, and tenderness), unique sensory evaluation, and several precautionary norms regarding toxic and microbiological contamination (9). In GC, the components present in the sample to be injected are constantly pushed through the column by a mobile gas phase, allowing them for separation and eluting through the column outlet. The retention times vary with phase loading, temperature, and flow; each probe's retention index is calculated by infusing the probe compound with a group of regular hydrocarbons that will span the compound's retention time. The retention index can be calculated from the following equation:

$$I = 100z + 100\frac{\log(t^1_{R(x)} - t^1_{R(z)})}{\log(t_{R(z+1)} + t^1_{R(z)})} \quad (1)$$

In the abovementioned equation, t'_R is the adjusted retention time and both x and z are the number of carbon that are present in the hydrocarbon before the sample elution; (z+1) denotes

the hydrocarbon after elution. Both the gas and the sample then pass *via* a detector through the column. The apparatus produces an electrical signal and measures the sample's volume. A chromatogram is produced by a signal of the detector and is used for gathering and analyzing samples for qualitative and quantitative data (10). Food samples must be homogenized as part of sample preparation to effectively extract the metabolites and minimize experimental mistakes. To extract the non-volatile, hydrophilic, low-molecular-weight molecules, such as amino acids, sugars, and organic acids, an extraction solvent is next added to the sample. Methanol is frequently used in methods to extract polar molecules. In some research, liquid extraction with methanol and chloroform is frequently used (11). Water impedes the derivatization reaction; therefore, after extraction, the samples are freeze-dried to eliminate it. The separated compounds are derivatized to increase the polar molecules' volatility for GC. The derivatization procedure uses oximation based on methylamine, silylation based on N, O-bis(trimethylsilyl)trifluoroacetamide, N-methyl-N-trimethylsilylacetamide (MSTA), and/or trimethylchlorosilane. After being derivatized, the sample is placed into a GC vial for GC/MS analysis. Multiple peaks are present in the GC/MS raw data, which must be deconvoluted and recognized. Deconvolution distinguishes between the true metabolite peaks and noise signals, and the numerous peaks indicate the strengths of distinct metabolites. The annotation of identified peaks is then performed by comparing them with known spectra from a metabolite spectral library (12). Following data processing, both univariate and multivariate analyses will be possible using a data matrix comprising the identification and intensity of the metabolite. These studies can currently be carried out by many universal applications, including the MSDIAL software and the XCMS online platform.

3.1.2. High-performance liquid chromatography

This method was initially an abbreviation for high-pressure liquid chromatography since early columns produced high operating pressures. High-performance liquid chromatography (HPLC), stressing the successful separations attained, had taken over as the favored phrase by the late 1970s. A discrete small amount (usually a few microliters) of the sample combination to be separated and studied is added to the stream of mobile phase percolating through the column. It has been widely recognized that polyphenolic compounds are a key dietary component for antioxidants that occur naturally and provide numerous health benefits. From a nutritional perspective, the need for quick, precise and sensitive techniques for food sample detection is in demand. For instance, creating food composition databases is the primary prerequisite for determining the daily intake of polyphenolic compounds. Phenolic compounds such as flavonoids, phenolic acids, tannins, stilbenes, and lignans are extraordinarily complex chemicals with a variety of health benefits and draw attention when they are examined in diverse dietary samples. The most popular separation technique for these uses is HPLC. Plant-based analysis of sugars with conventional extraction processes can create challenges due to the excessive absorption of water by these samples for subsequent separation by HPLC.

The sample's components flow through the column at various speeds as a result of their various physical interactions with the adsorbent (also called the stationary phase). Each component's velocity is influenced by its chemical makeup, the characteristics of the stationary phase (column), and the makeup of the mobile phase. The retention time of a particular analyte is the time at which it elutes (emerges from the column). The strength of elution is measured by the polarity index denoted by using P. The more the value of P, the higher the eluent strength. For example, the polarity index of a solvent P_m composed of solvents a and b, respectively, is $P_m = P_a^* X_a + P_b * X_b$, where P_a and P_b are polarity indexes of solvents a and b, respectively, and X_a and X_b are the fractions of their volumes. The effect of eluent polarity on the capacity factor k' of a compound is given by the equation:

$$\frac{K'_2}{k'_1} = \frac{10(P'_2 - P'_1)}{2} \quad (2)$$

$$\frac{K'_2}{k'_1} = \frac{10(P'_1 - P'_2)}{2} \quad (3)$$

where P_1' and P_2' are the polarity indices of the two eluent mixtures. HPLC has a few advantages over traditional low-pressure column liquid chromatography methods such as speed, as many studies suggest that an operation can be completed within 30 min; diversification, as different detectors can be employed to improve resolution and sensitivity; and lastly recovery of a sample due to the low volume of eluent (13). Another application of HPLC, which is frequently employed for the separation and purification of macromolecules like proteins and polysaccharides, is the evaluation of tiny molecules and ions such as sugars, vitamins, and amino acids (14).

The characterization of major and minor sugars relevant to nutrition can now be accomplished using HPLC. A minute change in the mobile phase flow rate can characterize the presence of sugar concentration in food (15). In addition, HPLC is utilized in distinct domains of carbohydrate research, including enzyme investigations on polysaccharides and their further analysis providing food makers with a quick quantification method with good accuracy and reproducibility. Screening for the presence of sucrose, maltose, glucose, and fructose in high-protein ingredients derived from legumes, pseudocereals, and cereals, such as quinoa, amaranth, and buckwheat as well as from soy, pea, lupin, lentil, carob, chickpea, and fava beans, confirmed the suitability of the selected extraction technique (16).

There is a significant need for accurate information about the amount of vitamin K present in food and feed products, as well as in human and animal blood, to better understand the vitamin's nutritional role. We concentrate on the vitamin phylloquinone, often known as vitamin K1, which is produced in plants due to its significance. The function of the so-called menaquinones, which are present in microbes, is currently unknown. These days, HPLC defeats time-consuming bioassays and thin-layer chromatography as the best analytical approach. Electrochemical or fluorescence detection provides the necessary results for the food components followed by the cleaning of samples. In the case of fluorescence detectors, the equation for dilute solutions can be written as $I_f = I_o \varphi_f$ (2.3abC), where I_f is measured emission intensity; I_o is excited beam intensity; φ_f is the number of photons emitted; a is the

molar absorption coefficient; b is the cell path length; and C is the concentration of the sample. A normal-phase HPLC is sometimes employed as a solid-phase extraction step. A straightforward, quick, and adaptable HPLC assay was introduced to replace time-consuming and labor-intensive methods for determining vitamin K1 (phylloquinone) (17). It should be stressed that affordable chemicals and generally accessible laboratory equipment were used. Materials of various origins were examined, and it was determined that the process was suitable for these ends. In the future, further research on NP will provide insight into a better-determining factor for the composition of foods and nutritional databases.

3.1.3. Chromatography in determining vitamins in food

Chromatographic methods developed over the past few decades have the capability of determining the vitamins having diverse chemical properties that are present in food, which are responsible for different biological activities in humans. The unstable nature of the target analytes makes routine analysis of vitamins problematic. Numerous elements, including exposure to heat, light, and air, as well as interactions with other dietary ingredients, can impact the stability of vitamins. Two qualitative techniques for the detection of water-soluble and fat-soluble vitamins were quickly developed combining reverse-phase high-pressure liquid chromatography with diode array detection (DAD). Due to the instability of different vitamins, whose breakdown frequently takes place during sample preparation, distinct HPLC procedures are advised for quantitative analysis.

Vitamins A, E, and D generally considered fat-soluble vitamins can be detected using HPLC-UV present in butter and vegetable oils. Vitamin C content, high in citrus fruit, is estimated using the reverse-phase HPLC method with the presence of UV at a certain wavelength, and optimized pH can also simultaneously determine citric acid, malic acid, and quinic acid present in fruits. Through the analytical procedure, vitamin B6 compounds can be determined in intact forms followed by extractions and assays, thereby performing a routine analysis in meat, fish, and other dairy food products (18).

Foods are subjected to vitamin analyses for a variety of reasons, such as regulatory compliance, nutrient labeling, and determining how food processing, packaging, and storage affect variations in vitamin content. HPLC is becoming a more significant method because of its automation capabilities employing autosamplers and robotics. Since the vast majority of vitamins are found in foods in tiny amounts, detection and sensitivity are important factors to take into account.

Fluorescence and electrochemical detection are also utilized in some circumstances, even though UV absorbance is the most typical detection technique. Due to its intrinsic lack of specificity and sensitivity, refractive index detection is not frequently utilized for vitamin identification.

Smaller particles, shorter columns, and microbore columns are being employed more frequently to increase speed and sensitivity in HPLC procedures that typically involve bonded phases, notably RP packing materials (19). Although UV and Fluoresence Detection (FLD) are frequently used, Electrochemical Detection (ED) is becoming more significant since it has improved sensitivity and selectivity for the detection of extremely minute levels of vitamins. The sample preparation process is made easier, and some oxidizable vitamins are shielded from oxidation by the use of HPLC column-switching procedures. For routine work in food control laboratories, where quick analysis and straightforward sample preparation are required, as well as a trustworthy and repeatable chromatographic test, these automatic techniques are becoming more crucial (20).

3.1.4. Chromatographic techniques are useful for identifying adulterations

Food adulteration is the deliberate lowering of a food's quality by the addition or substitution of unapproved alternative ingredients, the removal of important ingredients, or both. This is typically done to reduce the price or boost the volume of a certain food product. Chromatography-based techniques may identify various food adulterations. Among the most popular analytical detection techniques were HPLC and GC. With the use of these procedures, which are able to undertake qualitative and quantitative analyses of many classes of food elements, almost all foods may be evaluated.

3.2. Spectroscopic techniques

The dispersion of light into its individual colors is referred to as spectroscopy. It is a technique used for determining how much light is absorbed by a chemical material and at what intensity light flows through it. Spectroscopy is regarded as a crucial technique employed in the NP of food, both qualitatively as well as quantitatively. Spectroscopic methods help determine protein interactions and can be used with or without the combination of other analytical methods. The key principle is supported by electromagnetic light radiation being absorbed, transmitted, and emitted and the nature of its interaction with molecules, following the theory of effective collision. Since light itself is electromagnetic radiation, it can be useful for detecting the presence of microbes, harmful pathogens, and other compounds present in food and their analysis thereafter. Hence, food quality is ascertained through spectroscopic techniques, which can quickly distinguish the carbohydrate, fats, and water available in a variety of foods.

3.2.1. NMR spectroscopy

An important spectroscopic method for determining and calculating water content in food items is nuclear magnetic resonance (NMR), an experimental procedure that calls for meticulous calibration. The interaction between the magnetic field used and the magnetic characteristics of atoms and molecules is utilized by NMR spectroscopy (Figure 3). The fact that the resonance frequency of a given sample substance is precisely proportional to the intensity of the applied magnetic field is a crucial aspect of NMR. The resonance frequencies of a sample's nuclei depend on where in the field they are situated if it is placed in a non-uniform magnetic field, and imaging techniques make use of this fact. Three successive processes typically make up the NMR principle:

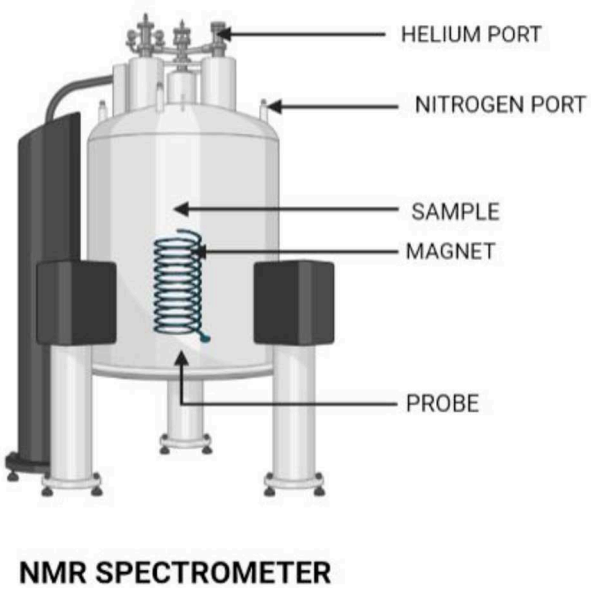


FIGURE 3
NMR spectrometer. An NMR spectrometer has a wide application in the field of research. This tool is also used in identifying genetic traits of foods.

- The polarization of magnetic nuclear spins in a static magnetic field (B_0) that is being applied.
- The nuclear spin alignment is disturbed by a radio frequency (RF) pulse, which is a weakly oscillating magnetic field. The static magnetic field (B_0) and the observational nuclei affect the oscillation frequency necessary for meaningful disruption.
- The precession of the nuclear spins around B_0 causes a voltage to be created in a detecting coil, which in turn causes the NMR signal to be detected during or after the RF pulse. Precession after an RF pulse often takes place at the nuclei's intrinsic Larmor frequency and does not, in and of itself, require changes in energy or spin states.

Food ingredients can be measured using this specific type of spectroscopic technique for a variety of applications. NMR is capable of identifying the genotypic features of grapes that are preserved to grow wine. In addition, it can also examine when food products have been adulterated, ripened, or dried out. In instances where sample composition is unknown, NMR can reveal the metabolic components of the sample. To identify biotic and abiotic stress in plants as well as genetic variations, the NMR spectroscopy technique is performed in conjunction with a multivariate analysis. This particular method is frequently used to detect metabolites present in fruit juices, wine, tomatoes, and beer as well as flavonoids, organic acids, and soluble sugars. An added advantage here is the fact that NMR equipment used in NP is portable. Due to the NMR spectra generating a significant quantity of data, its results are evaluated through Student's *t*-test and variate analysis. However, the identification of a molecule is at times proven to be difficult owing to the overlapping of spectra.

3.2.2. Atomic absorption spectroscopy

Atomic absorption spectroscopy (AAS) is an instrumental analytical technique used for quantifying trace elemental concentrations in a sample. This method employs light absorption of free atoms that are in gaseous form, and the measurement of the concentration of a particular analyte is done based on how much specific light has been absorbed in a certain wavelength. It is one of the foremost valuable means to identify nutritional constituents in food items such as iron, sodium, and calcium. AAS is essential as the analysis of trace elements forms an integral part of the labeling and quality control of food items.

The most recent AAS makes use of fiber optic technology, which results in an entirely enclosed optical system. Better light throughput for higher detection thresholds is provided by the optical system. The instrument's size is similarly decreased by the improved light path. In addition, it employs a layered architecture that enables the employment of a graphite furnace and flame on a single instrument. It makes use of a burner made of titanium which is simple to remove for various analyses. For quick startup and stability over time without the requirement for recalibration, it has a twin-beam design. Analyzing food products by AAS has novel applications, like the identification of infant food samples and formula, vegetables and oils derived from them, and meat and filets of fish (21).

3.2.3. Mass spectrometry

It is an experimental method for calculating the atomic or molecular weight of materials. At present, modern mass spectroscopic detectors used in mass spectrometry (MS) are the greatest ways to identify a wide variety of compounds. All molecules have a mass. As MS distinguishes chemicals present as ions in the gaseous phase, it is difficult to achieve an accurate measurement of a chemical spectrum. For ionization in the gas phase, the requisites vary greatly among chemicals. The energy needed to produce ions can cause chemical species to change as a result of interactions involving metabolites. Furthermore, because of the presence of other molecules, a single chemical can produce a variety of ionic forms with varying relative abundances. Standardized separation methods, such as liquid and gas chromatography and capillary electrophoresis, which reduce the complexity of the chemical mixtures entering the mass spectrometer, and techniques for ionization, such as atmospheric pressure chemical ionization (APCI), electrospray ionization (ESI), and desorption electrospray ionization (DESI), all of which ionize various chemicals, control these limitations. For the purpose of trace element detection and chemical speciation, HPLC is frequently used in combination with inductively coupled plasma-mass spectrometry (ICP-MS) owing to its efficient separation technique. HPLC-ICP-MS coupling is currently an indispensable screening method employed for the identification of unknown metal species as it quantitatively reacts to every molecule where a certain heteroelement is present, notwithstanding its coordination environment.

Absolute quantification by MS necessitates normalization concerning real standards and is frequently easily accomplished only for a few compounds. Nutritional metabolomics continue

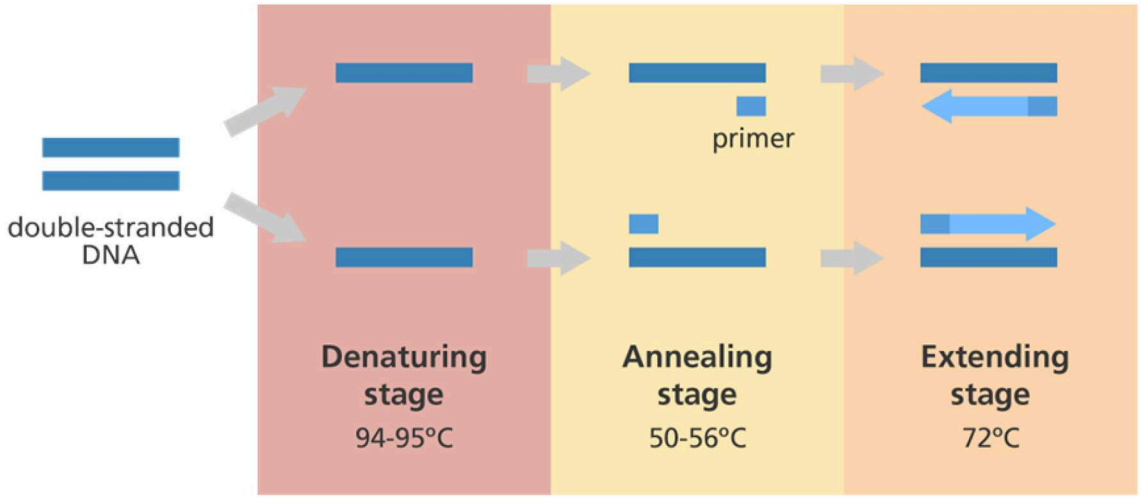


FIGURE 4
Polymerase chain reaction (PCR). A typical schematic representing the workflow for the PCR reaction. The double-stranded DNA undergoes three major steps, namely denaturation, annealing, and extension before the amplification step.

to face significant challenges due to the limited capacity to obtain the absolute calculation of a high number (>2,000) of metabolites. Based on mass resolution and accuracy, the high-resolution MS method has been utilized for estimating a huge number of compounds. Using precise mass/charge (m/z) values, these properties enable the elemental composition of a chemical to be predicted (22).

3.3. Polymerase chain reaction

A handful or a singular nucleic acid sample can be amplified using the polymerase chain reaction (PCR) technique to produce thousands or even millions of copies of that same nucleic acid (Figure 4). This makes characterizing and comparing genetic material from various people and species simpler. Altogether, it is regarded as a machine for duplicating DNA on a molecular level. The fundamentals of the PCR technique are founded on heat cycling, which makes use of thermodynamics in interactions between nucleic acids. Thermal cycling of the PCR samples after a specified set of temperature increments is now used by most PCR machines. These thermal cycling stages are initially necessary for the physical separation of two strands in a dsDNA double helix before the high-temperature process of DNA replication. The DNA polymerase enzyme then helps in synthesizing the new DNA strands that are complementary to each other by using each of them as a template for dsDNA synthesis at lower temperatures. Traditional methods for detecting infections and other microbes rely on culturing techniques, but these are laborious and time-intensive, no longer meeting the expectations of diagnostic laboratories and quality control processes to deliver outcomes quickly. Direct identification of microbes in a food item is greatly expanded by the specificity and sensitivity of PCR, which can deliver precise results in roughly 24 h.

Some pathogenic DNA or RNA are mostly the main targets in the food products for evaluating food authenticity such as microbes causing spoilage, molds producing mycotoxins, toxin-generating bacterial DNA, and DNA containing unwanted trace components. Several issues can arise when PCR is used for the identification of pathogens in food products such as toxicogenic fungi, although many of them may be resolved by using appropriate sample preparation techniques. To detect dietary allergens like peanuts, buckwheat, and wheat, precise and sensitive real-time PCR approaches have been developed. These methods can balance instrument effects and reduce the possibility of false-positive and false-negative results (23).

Real-time PCR has emerged as a viable alternative method for food diagnostics. It has a variety of benefits over traditional culturing methods, including speed, exceptional analytical sensitivity and selectivity, and the ability to quantify. However, the actual application of it for food monitoring and control is being hampered by the use of expensive tools and reagents, the requirement for knowledgeable personnel, and the absence of established protocols.

For the identification of dietary allergens like wheat, buckwheat, and peanuts, precise real-time qualitative PCR approaches have been developed. These methods can balance instrument effects and reduce the possibility of false-positive and false-negative results. Using reference plasmids that contained known copies of the target sequences, the cutoff for identifying positive samples was established in each run of the real-time PCR assay. The allergenic components in highly processed foods (cooked for longer than 30 min at 122°C) corresponding to 10 ppm (w/w) protein were identified using the copy counts of the plasmids. A reference plasmid analysis for each real time PCR is run in reduced instrument and variability. In addition, it assisted in preventing false positives brought on by minute quantities of agricultural- or laboratory-related pollutants. Using 79 frequently consumed food items and some of their relatives, the specificity of the real-time PCR approach was confirmed. In various types of samples, the approach was found to be perceptive enough to identify allergenic components matching 10 ppm (w/w).

3.4. ELISA

The test employs a solid-phase form of enzyme immunoassay (EIA), utilizing antibodies against the target protein to find the presence of a ligand (often a protein) in a liquid sample. Antigens from the sample to be examined are coupled to a surface in the simplest ELISA. The surface is then covered with a corresponding antibody so it can bind the antigen. Any unbound antibodies are then taken out when this antibody is coupled to an enzyme. A substance containing the enzyme's substrate is introduced in the last stage. If there was binding, the next reaction generates a discernible signal, most frequently, a change in color. An ELISA requires at least one antibody that is specific for a given antigen. The sample containing an unknown quantity of antigen is either non-specifically (through adsorption to the surface) or specifically immobilized on a solid substrate (often a polystyrene microtiter plate) (*via* capture by another antibody specific to the same antigen, in a "sandwich" ELISA). The detecting antibody is added and forms a complex with the antigen once the antigen has been immobilized. The detection antibody may be bioconjugated to an enzyme or may be covalently attached to an enzyme in order to be detected by a secondary antibody. To get rid of any non-specifically bound proteins or antibodies, the plate is routinely cleaned with a mild detergent solution between each step. After the final wash process, an enzymatic substrate is added to create the plate, which provides a visual signal that indicates the quantity of antigen in the sample.

The serological methods like precipitation and agglutination frequently used in food microbiology to isolate fungal strains from bacterial ones are gradually being replaced by contemporary, precision immunoassays based on a format similar to ELISA. Although occasionally poisons, enzymes, or polysaccharides present extracellularly are detected, the majority of immunoassays are predicted by certain antigens being present in cell membranes or the cytoplasm. Several microbial species and strains for which there are commercially accessible analytical kits and immunoassays have been developed to provide details on the format of the analysis. When it comes to food microbiology, immunochemical approaches are more laborious and time-efficient than microbiological methods, especially in strains where extraction and detection are laborious and time-consuming. By using various techniques, immunochemical processes enable a significant reduction of these needs.

Mycotoxins, the most dangerous and difficult-to-analyze group of food-related toxins, are routinely identified using immunoassays. Major mycotoxins have maximum permissible levels defined globally in several commodities due to their toxicity and their ubiquity in processed as well as raw food. As a result, immunoassays to detect controlled mycotoxins have been developed. Depending on the need, ELISAs and other tests are commercially available, with the list being updated every day.

ELISAs are frequently used in mycotoxin analysis to screen a lot of samples, while chromatographic techniques form the basis for confirmatory results. The development of immunoanalytical platforms enables the simultaneous identification of several chemicals and has been pushed by the introduction of MS and its multi-residual analytical features. Due to the great specificity and precision of immunoassays, it is frequently probable to completely omit time-consuming purification processes and reduce sample size as well as the volume of extracting solvents. However, immunochemical cross-reactivity may partially affect the assay's selectivity in the specific situation of tiny molecules with identical chemical properties, making it more ideal for screening than for a single compound's quantification. Immunoassays for pollutants found in industries, phytopharmaceuticals, and pesticides have recently gained acceptance as ways to supplement conventional analytical techniques in the food analysis domain. For pesticides, kits are available which are both quantitative and semi-quantitative. An antibody made against a chemical's protein conjugate may cross-react with other structurally related molecules in the same class, however, with varying degrees of success. This can be used to quickly estimate the total number of pollutants in a particular food item (24).

3.5. Solid phase microextraction

Solid phase microextraction (SPME) is a solid phase extraction sampling method that uses a fiber coated with an extracting phase, which may be a liquid (polymer) or a solid (sorbent), to extract various analytes (e.g., both volatile and non-volatile ones) from various media, which may be in the liquid or gas phase. As long as equilibrium is attained, or in the case of a brief period of pre-equilibrium, with the aid of convection or agitation, the amount of analyte extracted by the fiber is proportional to its concentration in the sample. Ultra-performance liquid chromatography incorporated with high-resolution MS and SPME incorporated with MS have been used for phytonutrient and aroma profiling followed by a nutrient overview by GC-MS. A holistic insight into breast milk composition is important for health benefits that derive from breastfeeding and the composition of infant formulas based on a novel analytical approach, holistic profiling of Human Breast Milk (HBM) lipidomes. SPME and LC-MS have been developed to improve microextraction (25). A new extraction method allows a wide range of lipids to be extracted directly from his HBM samples quickly and easily. A lipid extraction protocol provides high lipidome without using toxic solvents such as chloroform. Searching of lipid database by quadrupole time-of-flight MS detects lipids in HBM. A headspace solid-phase microextraction-gas chromatography-time-of-flight-mass spectrometry method was developed for the profiling of apple volatile metabolites. The selected SPME method was applied to the profiling of four different apple cultivars using GC-EI-TOF-MS. The developed headspace solid phase differential combined with gas chromatography-mass spectrometry (HS-SPME) sampling method is fully automated and is useful for obtaining volatile substance fingerprints in fruit (26). An improved method based on HS-SPME/GC-MS has been proposed for the semi-quantitative determination of volatile compounds. Polydimethylsiloxane or divinylbenzene fibers were used to extract unsettled particles from the headspace of bread dough samples that were dispersed in an aqueous sodium chloride solution (20%) and stored for 60 min at a temperature of 50°C in surrounding conditions. By calibrating with extracts tailored to the matrix, the method's excellent linearity for the selection of volatiles from various chemical groups has been confirmed.

3.6. Microscopic techniques

The majority of foods represent heterogeneous systems, and their several structural components determine their mechanical, biological, and functional properties along with their stability. Using different microscopic methods, a variety of meals and their structural properties can be estimated on a large scale. All these methods allow researchers to gain better clarity regarding the relationship between food components. With the influence of certain specific structural constituents on the different functionality, the impact of composition and processing on a substance's physical properties, the scattering of microbes and their food surface interactions, as well as changes in the accessibility and structural characteristics of different compounds are bioactive during digestion. Multiple microscopic techniques are used to describe matrices of food. Optic, electron, and probe microscopy are the most common forms. The most effective microscope technique will be based on the food type and the objective of the investigation. The ability to discriminate between two nearby objects, the degree of magnification, the sample preparation requirements, and the method's impact on the structure of the food should all be considered when selecting an appropriate methodology. The utilization of a combined imaging approach, such as light microscopy and electron microscopy, often makes data analysis and interpretation simpler and more efficient (27).

Early uses of light microscopy in the food industry focused mostly on ensuring food safety, such as identifying microbiological contamination or adulterants in food products. With the development of optical microscopy by Olga Flint, it is now possible to employ LM to determine the microstructures and macrostructures of meals using methods such as selective staining and optical contrast. LM continues to be important in the field of food research because of its availability, low cost, ability to distinguish between food color, and versatility in the testing area, which permits the analysis of food samples, especially certain wet specimens in ambient nature. The different advanced techniques like correlation compared to methods have expanded their applicability (e.g., scanning electron microscopy [SEM]). Its main disadvantages are its poor resolution and depth of focus, especially in the settings of a microscope. The different image processing methods used digitally such as focus or z-stacking are widely used in confocal microscopy and are utilized in wide-field digital microscopy to improve the focusing parameter, albeit this frequently results in picture acquisition periods. In food applications, LM techniques such as bright-field microscopy, polarized microscopy, and fluorescence microscopy are widely used. Apart from these, the main apparatus used is a conventional bright field microscope. Most modern microscopes are accompanied by a digital camera, which facilitates the analysis of the micrograph images (28). It is simple to attach polarizing and fluorescence accessories to them. To determine the presence, structural organization, and spatial distribution of particular food components in a product, fluorescent microscopy uses the autofluorescence method on compounds inside a specimen or the addition of selective fluorescent probes. If the fluorophores are chosen properly, their radiation may be widely recognized in the black backdrop when light with a specific wavelength is utilized to excite the intrinsic or additional fluorophores. Many dietary components are autofluorescent such as some aromatic amino acids, pigments, phenolic compounds, antioxidants, and tastes. It is well known that fluorophores are frequently found in food compounds; synthetic fluorescent dyes are frequently chosen for microscopy because of their higher quantum yields and brightness, environmental sensitivity (i.e., how the environment's physical and chemical characteristics impact emission), and selective interactions with particular food components. Rhodamine B and fluorescein isothiocyanate (FITC) are used to non-covalently label starch. To improve selectivity toward food components, fluorescent probes are linked to either lectins or antibodies. They constitute lectins that are fluorescently labeled like Concanavalin A and can effectively stain a-linked carbohydrates. For carbohydrate labeling, green fluorescent protein (GFP) can be used in conjunction with carbohydrate-binding modules that have high selectivity for carbohydrates such as cellulose. Nile Red is a widely used fluorophore that is soluble in lipids for assessing the lipid distribution in solid, semi-liquid, and liquid foods (29). Biomolecules such as proteins, carbohydrates, and phospholipids, tagged with fluorophores, are linked covalently. When selecting the different strains, it is crucial to take into account potential autofluorescence sources in samples because overlapping in the emission makes data analysis more difficult.

Currently, the function of important structural components like protein aggregates, fat crystals, or polysaccharide fibers inside the food can be understood using electron microscopy, which is a crucial tool for assessing the external or internal composition of meals. Since electron microscopy produces images by using an accelerated electron beam rather than photons, it has a better resolution than LM. Due to their shorter wavelengths than visible light, electron beams can identify tiny structures. To reduce electron beam scatter, conventional electron microscopy operates in a high vacuum. Therefore, thorough sample preparation is necessary to remove water or other volatile compounds from the food samples. Artifacts are frequently introduced as a result, which affects the final photographs. The most used EM method for studying foods is SEM. A low-energy electron beam is used to scan the specimen's surface in an x-y direction, and the sample's electrons are detected using SEM. SEM gives a distinct three-dimensional perspective of the specimens enabling structural alterations to compositional or processing-related variances in bulk attributes. To test wet samples with cryogenic SEM, the specimens must be frozen before being tested while in a cryo condition connected with SEM. Thus, the images obtained are more accurate depictions of food systems than images made with conventional methods of SEM. For samples with high moisture content, like dairy products, this approach has been widely utilized. Transmission electron microscopy (TEM) images are produced by the abrasion of high and intense electron beams as it passes thin specimens that have undergone careful preparation. It can be difficult to interpret the data from the 2D TEM micrographs; instead, they need to be viewed in the context of the 3D environment being investigated. In the TEM, components with different electron densities can be estimated such as the myofibrils of pig muscle rupture as a result of ultrasonic cavitation during the curing process. The development

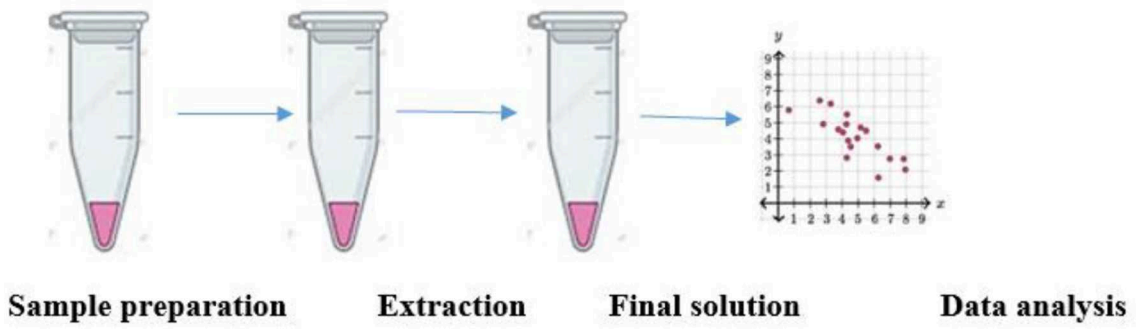


FIGURE 5
Nutritional metabolomics. A typical flowchart describing the application of metabolomics in the field of nutrition. The most important step of sample preparation is followed by the extraction of the cell and leading to the formation of the final solution. An *in silico* data analysis generates results, thereby providing valuable output.

of environmental scanning electron microscopy (ESEM) and environmental transmission electron microscopy (ETEM), which permits the testing of hydrated samples in their natural state in a gaseous environment with less sample preparation, has significantly increased the use of electron microscopy in food applications in recent decades. Compared to conventional SEM and TEM methods, environmental EM generates images utilizing minimal artifacts but lower resolution. E-SEM is used successfully for tracking the formation of colloidal aggregates and films in food models. Conventional optical microscopy that is fitted with lasers and Raman detectors can be used to perform Raman microspectroscopy. For example, the mapping of the chemical composition of a food sample and material distribution as to how they are altered by processing is made possible by combining microscopy and Raman detection. The spectrum is captured at each position along the specimen's spatial range as the surface is scanned point-by-point or line-by-line. The banding pattern within a region of the spectrum allows the identification of proteins and carbohydrates, thereby utilizing sophisticated chemometric methods. For each inspected spot, the intensity of a band is plotted and related to its concentration to generate images. With the advent of affordable and simple equipment with specialized software for processing spectra and images, Raman microspectroscopy becomes more accessible. The minute examinations of food samples have substantially advanced in recent years. The application of molecular modeling approaches will continue to support this ongoing trend. Chemical imaging methods will advance since they are non-invasive and relatively easy to use. They will be able to link microstructural alterations to compositional changes as a result, which will aid in the logical design of innovative meals and the optimization of processing techniques.

3.7. Nutritional metabolomics

Metabolomics studies consistently show that alterations in fatty acid, lipid, and tryptophan pathways are common and associated with disease state and outcome in critically ill patients. Metabolomics offers many opportunities to advance nutritional cancer epidemiology. This statement overview summarizes recent research, challenges, and prospects in nutritional metabolomics and epidemiological cancer research (Figure 5). Additional metabolomics research examining the relationship between food exposures and cancer risk, prognosis, and survival is required, in addition to methodological research, longitudinal analysis, and studies validating biomarkers. Although the objective remains, metabolomics offers a potential route for future research in nutritional cancer. In addition, nutrients that directly and indirectly affect and regulate gene activity play important roles in the prevention and treatment of chronic degenerative diseases. A growing body of research has found that obesity and altered metabolic responses to low-calorie weight-loss diets can be altered by genetic variants associated with obesity, metabolic status, and nutritional preferences. Survival depends on the intake of essential nutrients, and dietary components influence both disease prevention and promotion. Metabolomics is the study of all low molecular weight metabolites in a system. Nutrients and dietary components are key environmental factors that interact with genomes, transcriptomes, proteomes, metabolomes, and microbiota, and this lifelong interaction defines an individual's health and disease. In genetically predisposed people who are exposed to environmental factors, rheumatoid arthritis is a chronic autoimmune disease that causes a systemic immune-inflammatory response. In recent years, increasing evidence suggests that dietary factors and gut microbiota play a central role in the risk and progression of rheumatoid arthritis. Plasma samples obtained from lactating women participating in a study in Samar, Philippines, with low (vitamin A–) or adequate (vitamin A+) (plasma retinol <0>) vitamin A status were 1.05 μmol/was chosen). A total of 28 metabolites were altered in vitamin A– and vitamin A+ status groups, and 24 were lipid mediators ($P = 0.05$). Low quantities of oxylipins produced from arachidonic acid, eicosapentaenoic acid, and the VA group's lysophospholipids and sphingolipids were among these lipid mediators ($P < 0.05$). Reduced lipid mediator concentrations were found in multi-assay dietary profiles of low and adequate vitamin A status in breastfeeding women. Diet has an impact on how the gut microbiota and its mammalian host interact. Small chemicals produced by microbiota may be ingested by the host and affect a variety of crucial physiological functions (30).

4. Protocol for nutritional profiling

A very-low-calorie ketogenic diet contains low carbohydrates. The cardiovascular risk profile is distinguished by the presence of abdominal obesity, high total cholesterol, high triglycerides, and abnormal fasting blood glucose levels. Irritable bowel syndrome (IBS) is a chronic functional gastrointestinal disorder characterized by abdominal pain associated with bowel movements or changes in bowel habits. How different diets modulate gut microbiota profile such as a low-FODMAP diet, which is effective in people with IBS, contributes to changes in the gut microbiota. The purpose of this review was to examine different dietary protocols (conventional dietary advice, low FODMAP diets, gluten-free diets, etc.). Although there is no ideal nutritional protocol for patients with IBS-D, we investigated the impact of different nutritional approaches on gut microbiota composition to better define efficient strategies to treat these disorders. It seems important to consider. Alterations in the gut microbiome have been shown to contribute to the progression of metabolic diseases such as prediabetes and type 2 diabetes. Studies suggest that *in vivo* modulation of the gut microbiota by certain probiotic microbes can improve insulin sensitivity and glycemic control, and prevent or delay the onset of type 2 diabetes. However, further research is needed to understand the efficacy of probiotics as a treatment for metabolic diseases. Evidence-based multi-probiotics are designed to shift a cohort of gastrointestinal bacteria from prone to balanced in order to improve metabolic markers associated with type 2 diabetes. A total of 60 adults with prediabetes or type 2 diabetes (diagnosed within the last 12 months) and BMI $\geq$25 kg/m^2 will be enrolled in a double-blind, placebo-controlled pilot study. Participants will be randomized to multiple probiotics or placebos for 12 weeks. Both groups receive lifestyle and nutrition advice. The primary endpoint was the change in fasting plasma glucose between groups from baseline to his 12th week. Secondary outcome parameters include changes in lipid profile, systemic inflammation, intestinal permeability, and stool microbial and metabolomic profiles. Blood and stool samples will be collected at baseline and 12 weeks after treatment. Research on the role of vitamin C in the prevention and treatment of pneumonia and sepsis has been ongoing for decades (31). This study provided a strong basis for extrapolating these findings to patients with severe COVID-19. Studies have shown that patients with pneumonia and sepsis have low vitamin C status and increased oxidative stress. Giving vitamin C to patients with pneumonia can reduce the severity and duration of the disease. Severely ill patients with sepsis should be given intravenous gram doses of the vitamin to normalize plasma levels. This is an intervention that several studies have found to reduce mortality. Vitamins have multiple physiological functions, many of which are relevant to COVID-19. These include antioxidant, anti-inflammatory, antithrombotic, and immunomodulatory functions. Preliminary observational studies have shown low vitamin C status in critically ill patients with COVID-19. Several randomized controlled trials (RCTs) are currently registered worldwide, investigating intravenous vitamin C monotherapy in patients with COVID-19. Studies were conducted in populations with chronic vitamin C deficiency because vitamin C deficiency is common in low- and middle-income settings, and many of the risk factors for vitamin C deficiency overlap with those for COVID-19, which may have a greater effect. This is particularly relevant to global research efforts, as COVID-19 disproportionately affects lower-middle-income countries and low-income populations around the world. A small study from China was completed prematurely, and the results are now being reviewed by experts. Patients who received vitamin C therapies tended to be more seriously unwell. Mortality was significantly reduced. Future results from a large ongoing RCT will provide more conclusive evidence for the optimization of intervention protocols in future studies. Early and sustained dosing is warranted as it may improve efficacy. Patients with vitamin C deficiency are affected by respiratory infections due to its superior safety profile, low cost, and potential for rapid scale-up of production (Table 2).

5. Case studies of nutritional profiling using chromatography

5.1. Case study on edible vegetable oils

One of the main categories of food items are oils and fats, with edible vegetable oils generally replacing animal fats in processed food compositions. Their validity has, therefore, become a crucial concern (32). The identification of their adulteration can be facilitated by fat and oil authentication using the examination of several component lipid classes by both HPLC and GC analyses. Triacylglycerols, fatty acids, sterols, and other minor substances are found in edible vegetable oils (33). The whole fatty acid family of edible fat, or "fatty acidomics," is a crucial indicator and quality measure for these food matrices. After being converted into fatty acid methyl esters (FAMEs), they are often subjected to GC-MS analysis (33). Their identification is performed by comparing their mass spectra to the retention indices of standards and/or by utilizing mass spectra libraries. To determine if there has been adulteration, the percentage values for each fatty acid are compared to those accepted by national and international organizations. An increase or reduction in the quantities of certain fatty acids may be indicative of this. Even if they are not very unique, fatty acid patterns may be beneficial for determining authenticity, particularly when examining pure oils. For instance, the presence of palm oil is indicated by the presence of significant percentages of palmitic acid, whereas the presence of rapeseed oil is indicated by the presence of trace levels of erucic acid. The FA composition is too complicated to be utilized for verification in blends including oils from several botanical sources, in part because of chromatographic peak merging. It is acknowledged that the fatty acid profile only makes up a small portion of the composition of oils; occasionally, further examination of other macrocomponents and microcomponents of oils is required (34).

5.2. Case study on milk and dairy products

By analyzing the lipids in dairy products, both qualitatively and quantitatively, it is possible to identify foreign fat in milk fat. Butter is the most significant milk fat product due to its extensive usage. Its quality is crucial because of this; it has always been

TABLE 2 Techniques currently used in nutritional profiling of foods with advantages and disadvantages.

Technique	Use in nutritional profiling	Advantages	Disadvantages
Gas Chromatography (GC)	Detects compounds like sterols, oils, low-chain fatty acids and contaminants like pesticides and pollutants	Allows separation of the components of complex mixtures in a reasonable time owing to high level of efficiency	Limited to thermally stable and volatile compounds
High-Performance Liquid Chromatography (HPLC)	Identifies major and minor sugars in carbohydrate research as well as vitamins	Rapid and precise method for identification of specific chemical compounds	Expensive and complex process which is not applicable to work for all samples
Nuclear Magnetic Resonance (NMR) Spectroscopy	Determines water content, metabolic composition including presence of flavonoids, organic acids and soluble sugars	Aids in obtaining clear structural information of molecules in their natural environment	It is not a cost effective method and determination of higher molecular weight structures poses a problem
Atomic Absorption Spectroscopy (AAS)	Quantifies trace elements such as iron, sodium, calcium	Easy operation and high level of sensitivity as well as accuracy	Non-metals cannot be identified by this method
Mass Spectrometry (MS)	Measures protein concentration, trace elements and unknown metal species	Automated technique that can be employed on a large scale	Identification of hydrocarbons having similar ions is difficult
PCR	Detects dietary allergens like peanuts, buckwheat and microbes and unwanted contaminants	Highly sensitive technique producing quick results	Cannot be utilized for amplification of unknown targets
ELISA	Analysis of mycotoxins and other pollutants present in food	High precision and rapid results	Quite a labor-intensive as well as expensive method to carry out
Solid Phase Microextraction (SPME)	Used for phytonutrient and aroma profiling as well as determination of volatile compounds	It is a simple and economic method to implement	Limited choice of selectivity
Microscopic Techniques	Estimation of structural properties of food components like protein and fats	High resolution helps in better viewing	Quite expensive instruments
Nutritional Metabolomics	Identifies nutrients, dietary fibers and metabolites and their interaction with genomes defining human health and diseases	Useful tool for premature disease diagnosis providing quick analysis	Low in sensitivity

tainted by the addition of less expensive vegetable or animal fats. By examining fatty acids (FAs), triacylglycerols, and minor lipid components of the unsaponifiable fraction, foreign fat in milk fat can be found (35). Following GC analysis, the addition of vegetable oils to milk fat can be identified by measuring individual FAs or the concentration ratio of two or more FAs. FA ratios have also been applied in the past to distinguish between various types of milk fat (36). Classes of substances with equal numbers of acyl-C atoms have been suggested for and put to use in TAG analysis by utilizing GC analysis (35). For authentication reasons, analysis of small chemicals can be quite helpful. By using silica column GC to analyze free sterols or trimethylsilyl derivatives, it was possible to identify the presence of vegetable oils in milk fat that had been tampered with (37). For the qualitative and quantitative study of the unsaponifiable fraction of milk lipids (cow butter, buffalo, sheep, and goat milk), a thorough GC technique with dual MS/FID detection was developed and refined. The identity of chemicals that were found was confirmed using a GC-high-resolution TOF-MS analysis. The efficacy of such a technology for dairy product authenticity tests was illustrated by the high number of chemicals that were isolated and identified (38).

5.3. Case study on honey

The adulteration of honey is a complicated issue that currently has negative effects on nutrition and organoleptic quality as well as considerable economic impact. Numerous forms of economically driven adulteration have been discovered in the honey business; chromatographic techniques, less-priced syrups, and misleading claims about the botanical and geographic origins of honey have all been implicated. Numerous writers have suggested using chromatographic methods to analyze the primary components of honey, which are carbohydrates. Corn syrup (CS), high-fructose corn syrup (HFCS), invert syrup (IS), and high-fructose inulin syrup are examples of common, affordable sweeteners that are added to honey as adulterants (HFIS). A high-performance anion-exchange chromatography pulsed amperometric detection (HPAEC-PAD) system (39), HPLC equipped with a common RID detector, and simultaneously GC-FID and HPAEC-PAD analyses were used to obtain a fingerprint profile of honey oligosaccharides (40, 41). This analysis used a straightforward HPAEC-PAD method for quantitative analysis of maltooligosaccharides to identify honey adulterated with CS and HFCS. The technique was based on fractionating honey carbohydrates using activated charcoal sample treatment before analyzing the results. The amount of oligosaccharides was measured after analysis of several samples of honey, CS, and HFCS. A honey sample was made into adulterated samples by adding 5, 10, and 20% of syrup to it. The technique proved to be a helpful tool for identifying honey tainted with commercial syrups. Fructose and glucose were measured using the HPAEC-PAD technique, and the full profile of disaccharides and trisaccharides was determined using the GC-FID method (41). With the use of statistical processing, the outcomes of the two

procedures were integrated. This strategy was developed for a better understanding of the effect of syrup when added to acacia, chestnut, and lavender honeys. Using genuine honey samples, an oligosaccharide database was produced. Numerous French monofloral commercial honeys were examined. For the examined monofloral honeys, the limits of detection were quite good, ranging between 5 and 10%.

5.4. Case studies with fruit juices

Due to their taxonomic uniqueness, phenolic chemicals are particularly promising indicators for determining the authenticity of the food. Orange juice adulteration has been a particular focus of their chromatographic investigation. Juices from sweet oranges (*Citrus sinensis*), whose consumption has greatly grown recently, might also contain grapefruit (*Citrus paradisi*), tangerine (*Citrus reticulata*), or lemon (*Citrus limon*). HPLC-DAD/ESI-MS/MS and HPLC-DAD were developed by Abad-Garcia et al. (40) for the characterization and quantification of phenolic chemicals in citrus juices, respectively (42). These phenolic chemical profiles were examined with the goal of distinguishing citrus liquids according to the species utilized for their elaboration: sweet orange, tangerine, lemon, or grapefruit. They are typical of the Spanish citrus fruit juice production. For the purpose of creating classification models and identifying potential markers, statistical and chemometric techniques were used. All delicious orange and tangerine liquids were successfully recognized by classification models. Despite needing an outside validation, the proposed model appears to be effective at identifying sweet orange juice tainted with tangerine juice.

6. Case study of nutritional profiling with NMR

Despite being a relatively new use of NMR, food authentication accounts for the bulk of NMR's applications in the field of food science (43). Beverages, fruits and vegetables, honey, fats and oils, spices, dietary supplements, as well as meat, fish, and dairy products, are some examples of the applications. As these are the most frequent causes of fraudulent claims for food authenticity, the elements that are analyzed most frequently include variety, geographic origin, harvest season and/or agronomic techniques, and adulteration with items of lower price and quality (44). The typical nucleus used for analysis is the proton. Exceptions include the application of ^{13}C NMR-based metabolomics for coffee authentication to avoid interactions between caffeine and chlorogenic acids that cause issues related to chemical shift alignment and the application of ^{13}C NMR coupled with discriminant analysis for the classification of olive oils (45). Numerous uses of ^{31}P NMR spectroscopy in food authenticity exist, primarily in the field of olive oil analysis (46). Food adulteration by melamine, adulteration of honey by syrups (high fructose corn, maltose, or jaggery syrup), adulteration of olive oil by vegetable oils or lampante/pomace olive oils, mixing of ground black pepper with buckwheat and millet, adulteration of culinary spices by Sudan I dye, and meat adulteration are typical examples of spectroscopic methodologies used for food authentication. There have been reports of the authenticity of milk, olive oils, honeys, wines, spirits, spices, and other culinary items, as well as saffron and lentil seeds (47, 48).

Moreover, using SNIF-NMR, which is site-specific, it is possible to accurately fingerprint natural compounds. The identification of the geographic origin of the wine, pioneered by the EU in 1990, is a well-known use of SNIF-NMR. Food's geographic origin has been evaluated using profiling techniques including non-targeted ^{1}H-NMR analysis. NMR analysis has been used to determine adulteration, including the addition of cane or corn sugar to maple syrup, the adulteration of red wine with anthocyanins, and artificial tastes marketed as natural. Wines, coffees, olive oils, honeys, fish, spirits, vinegars, and saffron are among the foods that may be differentiated by their metabolic profiles using NMR (49).

7. Metrology in nutritional profiling of food

It is often necessary to include metrological principles with chemical and biological measures in food analysis. Based on the definition of metrology as "the science of measurement, embracing both experimental and theoretical determinations at any level of uncertainty in any field of science and technology," its concepts include the definition and implementation of globally recognized units of measurement and (metrological) traceability by evaluating uncertainty in relation to national and global reference standard measurements. In addition, metrology offers the resources needed to make measurement findings consistent and comparable.

Only a small number of techniques can use the RMs that are now available since they are only approved for a small number of parameters and small numbers of matrices. To meet the new requirements related to the analytical determination of nutraceutical substances and natural substances with protective effects on health, new RMs are either required or closely related to the new analytical needs and the emerging challenges of food safety (e.g., application of nanotechnologies or biotechnologies). The adoption of new RMs to be used for ensuring the origin and traceability of raw materials and products as well as for identifying frauds and adulterations is associated with a particularly urgent necessity (50). According to this perspective, new RMs are becoming more important in order to recognize genetic markers and chemical compositions and to confirm the geographical and/or biological (botanical, zoological, and genetic) origin of raw materials and finished goods. The ability to create multiparameter RMs for use in multi-parametric determinations, qualitative analyses, and identity studies through the definition of elemental, isotopic, molecular, and/or genetic markers or patterns for the traceability of food products is particularly intriguing, especially with regard to assessing the foodome (51). In addition to promoting multiparametric characterization standardization, a more thorough characterization of the foodome present in these materials would be of additional value for laboratories to avoid the costs of acquiring different materials for the required set of analytes and for RM producers to reduce the effort, for example, for the testing of homogeneity.

In addition to the use of RMs, accurate calibration of equipment, particularly the important components, is necessary to ensure the quality and traceability of measurement data. As a result, it is important to develop and maintain the traceability chain, which is described as a series of standards and calibrations that connect a measurement's result to a reference.

There is a significant dispersion in Europe and around the globe when it comes to the condition of food analysis and research today. Despite the fact that food production, marketing, and fraud occur on a multi-national and even international scale, every European member state approaches the overall goal of verifying food integrity differently. This is in addition to the common European Commission regulations on official food controls and common notifications within the EU Rapid Alert System for Food and Feed (RASFF) (52). Furthermore, a wide range of scientific fields and organizations are involved. It is worse that the information produced by these control and verification procedures is dispersed and entered into a wide range of diverse databases. Many initiatives seem to be unnecessary since they are being repeated and gaps are not being consistently filled.

8. Conclusion

This review mainly provides an outline that focuses on the nutritional model and the several bioanalytical tools that are used for detection and their potential applications. Recent scientific literature and evidence show that the classification of foods depends not only on the nutrition composition but also on the distribution of food in our total diet. To take into account the significance of the food in the diets of the general population, children, and other particular groups, nutrient profile models are also required. NP models are developed to prevent diseases that are still prevalent such as obesity, vitamin and mineral deficiencies, etc. Based on research, GC/MS and HPLC are used for the characterization of the food of interest and can also be used to evaluate the quality, aroma, and authentication of food samples. In addition, the GC method also has some drawbacks which can be further improved by optimizing parameters. Microscopic analysis showed significant advancements in detecting food quality and nutrition values and providing food security globally and high-quality standards of foods. PCR and different immunoassay techniques facilitated the analysis of food offering possibilities for the detection of allergens and toxins, thereby increasing food hygiene. On the other hand, nutritional metabolomics was proposed across populations addressing nutritional phenotypes and health and diseases but the rapid development of improved biosystem models and bioanalytical techniques helped to advance nutrition research. There are currently numerous initiatives underway to create, verify, and test nutrient profile models. The Food Standards Agency (FSA), the UK's counterpart of the Food and Drug Administration (FDA) in the US, has released interim and final findings, and they are now accessible online. France and the Netherlands have both developed additional NP models. These models often include both nutrients and dietary groups and are based on certain combinations of macronutrients, vitamins, and minerals. The relationship between nutrition and health is validated by the existence of nutritional databases. In the future, governments, industries, and other non-profit organizations should develop such advanced models with additional studies and research for more appropriateness making the domain of global health a priority.

Author contributions

DM, UC, MB, and SG contributed equally in writing the manuscript. DK revised, corrected, and edited the manuscript. All authors contributed to the article and approved the submitted version.

Conflict of interest

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

Publisher's note

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

References

1. Visioli F, Marangoni F, Poli A, Ghiselli A, Martini D. Nutrition and health or nutrients and health? *Int J Food Sci Nutr.* (2022) 73:141–8 doi: 10.1080/09637486.2021.1937958

2. Drewnowski A, Amanquah D, Gavin-Smith B. Perspective: how to develop nutrient profiling models intended for global use: a manual. *Adv Nutr.* (2021) 12:609–20. doi: 10.1093/advances/nmab018

3. Santos M, Rito AI, Matias FN, Assunção R, Castanheira I, Loureiro I. Nutrient profile models a useful tool to facilitate healthier food choices: a comprehensive review. *Trends Food Sci Technol.* (2021) 110:120–31. doi: 10.1016/j.tifs.2021.01.082

4. Margaritelis NV, Paschalis V, Theodorou AA, Kyparos A, Nikolaidis MG. Antioxidants in personalized nutrition and exercise. *Adv Nutr.* (2018) 9:813–23. doi: 10.1093/advances/nmy052

5. Martini D, Godos J, Bonaccio M, Vitaglione P, Grosso G. Ultra-processed foods and nutritional dietary profile: a meta-analysis of nationally representative samples. *Nutrients.* (2021) 13:3390. doi: 10.3390/nu13103390

6. McClements DJ, Grossmann L. A brief review of the science behind the design of healthy and sustainable plant-based foods. *NPJ Sci Food.* (2021) 5:17 doi: 10.1038/s41538-021-00099-y

7. Lehotay SJ, Hajšlová J. Application of gas chromatography in food analysis. *TrAC.* (2002) 21:686–97 doi: 10.1016/S0165-9936(02)00805-1

8. Putri SP, Ikram MM, Sato A, Dahlan HA, Rahmawati D, Ohto Y, Fukusaki E. Application of gas chromatography-mass spectrometry-based metabolomics in food science and technology. *J Biosci Bioeng.* (2022) 133:425–35 doi: 10.1016/j.jbiosc.2022.01.011

9. Farag MA, Khattab AR, Ehrlich A, Kropf M, Heiss AG, Wessjohann LA. Gas chromatography/mass spectrometry-based metabolite profiling of nutrients and antinutrients in eight lens and lupinus seeds (Fabaceae). *J Agric Food Chem.* (2018) 66:4267–80 doi: 10.1021/acs.jafc.8b00369

10. Feng T, Sun M, Song S, Zhuang H, Yao L. *Gas Chromatography for Food Quality Evaluation*. New York, NY: Elsevier Inc. (2019).

11. Stilo F, Bicchi C, Reichenbach SE, Cordero C. Comprehensive two-dimensional gas chromatography as a boosting technology in food-omic investigations. *J Sep Sci.* (2021) 44:1592–611. doi: 10.1002/jssc.202100017

12. Marret E, Beloeil H, Lejus C. What are the benefits and risk of non-opioid analgesics combined with postoperative opioids? *Annal Franc D'anesthesie Reanimation.* (2009) 28:e135–51. doi: 10.1016/j.annfar.2009.01.006

13. Pyrzynska K, Sentkowska A. Recent developments in the HPLC separation of phenolic food compounds. *Crit Rev Anal Chem.* (2015) 45:41–51 doi: 10.1080/10408347.2013.870027

14. Hurst WJ, Martin RA, Zoumas BL. Application of HPLC to characterization of individual carbohydrates in foods. *J Food Sci.* (1979) 44:892–5. doi: 10.1111/j.1365-2621.1979.tb08529.x

15. Hoehnel A, Salas García J, Coffey C, Zannini E, Arendt EK. Comparative study of sugar extraction procedures for HPLC analysis and proposal of an ethanolic extraction method for plant-based high-protein ingredients. *J Food Agric.* (2022) 102:5055–64 doi: 10.1002/jsfa.11204

16. Jakob E, Elmadfa IJ. Rapid and simple HPLC analysis of vitamin K in food, tissues and blood. *Food Chem.* (2000) 68:219–21. doi: 10.1016/S0308-8146(99)00158-2

17. Omeje KO, Ozioko JN, Ezema BO, Eze SO. Tiger nut (Cyperus esculentus): Nutrient profiling using HPLC and UV-spectroscopic techniques. *S Afr J Sci.* (2022) 118:1–4. doi: 10.17159/sajs.2022/11783

18. Rizzolo A, Polesello S. Chromatographic determination of vitamins in foods. *J Chromat.* (1992) 624:103–52. doi: 10.1016/0021-9673(92)85676-K

19. Stöggl WM, Huck CW, Scherz H, Popp M, Bonn GK. Analysis of vitamin E in food and phytopharmaceutical preparations by HPLC and HPLC-APCI-MS-MS. *Chromatographia.* (2001) 54:179–85. doi: 10.1007/BF02492241

20. Romero Rodriguez MA, Vazquez Oderiz ML, Lopez Hernandez J, Lozano JS. Determination of vitamin C and organic acids in various fruits by HPLC. *J Chromat Sci.* (1992) 30:433–7. doi: 10.1093/chromsci/30.11.433

21. Noh MF, Gunasegavan RD, Mustafa Khalid N, Balasubramaniam V, Mustar S, Abd Rashed A. Recent techniques in nutrient analysis for food composition database. *Molecules.* (2020) 25:4567. doi: 10.3390/molecules25194567

22. Jones DP, Park Y, Ziegler TR. Nutritional metabolomics: progress in addressing complexity in diet and health. *Annu Rev Nutr.* (2012) 32:183–202. doi: 10.1146/annurev-nutr-072610-145159

23. Klancnik A, Kovac M, Toplak N, Piskernik S, Jersek B. PCR in food analysis. *Polymerase Chain Reaction.* (2012). doi: 10.5772/38551

24. Anfossi L. *Immunoassays, Food Applications*. New York, NY: Elsevier Inc. (2019).

25. Sofi F, Dinu M, Pagliai G, Cesari F, Marcucci R, Casini A. Mediterranean versus vegetarian diet for cardiovascular disease prevention (the CARDIVEG study): study protocol for a randomized controlled trial. *Trials.* (2016) 17:1–8. doi: 10.1186/s13063-016-1353-x

26. Palacios T, Vitetta L, Coulson S, Madigan CD, Denyer GS, Caterson ID. The effect of a novel probiotic on metabolic biomarkers in adults with prediabetes and recently diagnosed type 2 diabetes mellitus: study protocol for a randomized controlled trial. *Trials.* (2017) 18:1–8 doi: 10.1186/s13063-016-1762-x

27. Zaffer S, Al-obadi M, Ibrahim RM, Al-khalifa KN. *Measuring Food Quality and Safety for Food Security Related Assessments?: A Microscopic Review* (2022). p. 5064–72.

28. Corradini MG, McClements DJ. Microscopy, food applications. *Encycl Anal Sci.* (2019) 6:47–56. doi: 10.1016/B978-0-12-409547-2.14314-8

29. Shappirio JR, Cook CF. Modern analytical techniques for failure analysis. *Solid State Technol.* (1979) 22:89–94. doi: 10.1016/0038-1101(79)90177-1

30. Carr AC, Rowe S. The emerging role of vitamin C in the prevention and treatment of COVID-19. *Nutrients.* (2020) 12:3286. doi: 10.3390/nu12113286

31. Altomare A, Di Rosa C, Imperia E, Emerenziani S, Cicala M, Guarino MP. Diarrhea predominant-irritable bowel syndrome (IBS-D): effects of different nutritional patterns on intestinal dysbiosis and symptoms. *Nutrients.* (2021) 13:1506. doi: 10.3390/nu13051506

32. Aparicio R, Aparicio-Ruiz R. Authentication of vegetable oils by chromatographic techniques. *J Chromatograph.* (2000) 881:93e104. doi: 10.1016/S0021-9673(00)00355-1

33. Tsimidou M, Boskou D. Adulteration of foods: detection. In: L Trugo, PM Finglas, editors *Encyclopedia of Food Sciences and Nutrition.* New York, NY: Acadamic Press (2003).

34. Osorio MT, Haughey SA, Elliott CT, Koidis A. Evaluation of methodologies to determine vegetable oil species present in oil mixtures: proposition of an approach to meet the EU legislation demands for correct vegetable oils labelling. *Food Res Int.* (2014) 60:66–75. doi: 10.1016/j.foodres.2013.12.013

35. De La Fuente MA, Juárez M. Authenticity assessment of dairy products. *Crit Rev Food Sci Nutr.* (2005) 45:563–85 doi: 10.1080/10408690490478127

36. Ulberth F. Milk and dairy products. In: M Lees, editor *Food Authenticity and Traceability.* Boca Raton, FL: CRC Press (2003).

37. Alonso L, Fontecha J, Lozada L, Juarez M. Determination of mixtures in vegetable oils and milk fat by analysis of sterol fraction by gas chromatography. *J Am Oil Chem Soc.* (1997) 74:131–5. doi: 10.1007/s11746-997-0157-2

38. Kamm W, Dionisi F, Hischenhuber C, Schmarr HG, Engel KH. Rapid detection of vegetable oils in milk fat by on-line LC-GC analysis of β-sitosterol as marker. *Eur J Lipid Sci Technol.* (2002) 104:756–61. doi: 10.1002/1438-9312(200211)104:11<756::AID-EJLT756>3.0.CO;2-F

39. Morales V, Corzo N, Sanz ML. HPAEC-PAD Oligosaccharide analysis to detect adulterations of honey with sugar syrups. *Food Chem.* (2008) 107:922–8. doi: 10.1016/j.foodchem.2007.08.050

40. Wang S, Guo Q, Wang L, Lin L, Shi H, Cao H, Cao B. Detection of honey adulteration with starch syrup by high performance liquid chromatography. *Food Chem.* (2015) 172:669–74 doi: 10.1016/j.foodchem.2014.09.044

41. Cotte JF, Casabianca H, Chardon S, Lheritier J, Grenier-Loustalot MF. Application of carbohydrate analysis to verify honey authenticity. *J Chromato.* (2003) 1021:145–5. doi: 10.1016/j.chroma.2003.09.005

42. Abad-García B, Garmón-Lobato S, Sánchez-Ilárduya MB, Berrueta LA, Gallo B, Vicente F, et al. Polyphenolic contents in Citrus fruit juices: authenticity assessment. *Eur Food Res Technol.* (2014) 238:803–18. doi: 10.1007/s00217-014-2160-9

43. Ellis DI, Brewster VL, Dunn WB, Allwood JW, Golovanov AP, Goodacre R. Fingerprinting food: Current technologies for the detection of food adulteration and contamination. *Chem Soc Rev.* (2012) 41:5706–27. doi: 10.1039/c2cs35138b

44. Hatzakis E. Nuclear magnetic resonance (NMR) spectroscopy in food science: a comprehensive review. *Food Sci Safety.* (2018) 18:189–200. doi: 10.1111/1541-4337.12408

45. Ramakrishnan V, Luthria DL. Recent applications of NMR in food and dietary studies. *J Sci Food Agric.* (2017) 97:33–42. doi: 10.1002/jsfa.7917

46. Wei F, Furihata K, Koda M, Hu F, Kato R, Miyakawa TM. 13C NMR-based metabolomics for the classification of green coffee beans according to variety and origin. *J Agric Food Chem.* (2012) 60:10118–25. doi: 10.1021/jf3033057

47. Zamora R, Navarro JL, Hidalgo FJ. Identification and classification of olive oils by high-resolution13C nuclear magnetic resonance. *J Am Chemists' Soc.* (1994) 71:361–4. doi: 10.1007/BF02540514

48. Dais P, Spyros A, Christophoridou S, Hatzakis E, Fragaki G, Agiomyrgianaki A, et al. Comparison of analytical methodologies based on 1H and 31P NMR spectroscopy with conventional methods of analysis for the determination of some olive oil constituents. *J Agric Food Chem.* (2007) 55:577–84. doi: 10.1021/jf061601y

49. Danezis GP, Tsagkaris AS, Camin F, Brusic V, Georgiou CA. Food authentication: techniques, trends and emerging approaches. *TrAC.* (2016) 85:123–32. doi: 10.1016/j.trac.2016.02.026

50. Tsimidou MZ, Ordoudi SA, Mantzouridou FT, Nenadis N, Stelzl T, Rychlik M, et al. Strategic priorities of the scientific plan of the European Research Infrastructure METROFOOD-RI for promoting metrology in food and nutrition. *Foods.* (2022) 11:599. doi: 10.3390/foods11040599

51. Zappa G, Zoani C. *Reference Materials in Support to Food Traceability. "Frontiers in food science for feeding the world". Abstract from the Pre-conference workshop of the 'Food technology 2015 conference'*—Università di Pisa. Agrochimica—Pisa University Press (2015).

52. Zoani C, Caprioli R, Gatti R, Zappa G. *Development of Innovative Reference Materials for the Agrofood Sector*. Abstract from the International Conference 1st IMEKOFOODS Metrology promoting objective and measurable food quality and safety. Rome (2014).

frontiers | Frontiers in Nutrition

TYPE Review
PUBLISHED 01 June 2023
DOI 10.3389/fnut.2023.1165580

Check for updates

OPEN ACCESS

EDITED BY
Miroslav Griga,
Agritec (Czech Republic), Czechia

REVIEWED BY
Petr Smýkal,
PalackýUniversity in Olomouc, Czechia
Roberto Rodríguez Ramírez,
Instituto Tecnológico de Sonora (ITSON),
Mexico

*CORRESPONDENCE
Vikender Kaur
✉ vikender.kaur@icar.gov.in

†PRESENT ADDRESSES
Boopathi Thangavel,
Crop Protection, Indian Council of Agricultural Research-Indian Institute of Oilseeds Research, Hyderabad, Telangana, India
Kuldeep Singh,
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), Patancheru, Telangana, India

RECEIVED 14 February 2023
ACCEPTED 27 April 2023
PUBLISHED 01 June 2023

CITATION
Kaur V, Singh M, Wankhede DP, Gupta K, Langyan S, Aravind J, Thangavel B, Yadav SK, Kalia S, Singh K and Kumar A (2023) Diversity of *Linum* genetic resources in global genebanks: from agro-morphological characterisation to novel genomic technologies – a review. *Front. Nutr.* 10:1165580.
doi: 10.3389/fnut.2023.1165580

COPYRIGHT
© 2023 Kaur, Singh, Wankhede, Gupta, Langyan, Aravind, Thangavel, Yadav, Kalia, Singh and Kumar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Diversity of *Linum* genetic resources in global genebanks: from agro-morphological characterisation to novel genomic technologies – a review

Vikender Kaur[1]*, Mamta Singh[1], Dhammaprakash Pandhari Wankhede[1], Kavita Gupta[1], Sapna Langyan[1], Jayaraman Aravind[1], Boopathi Thangavel[1†], Shashank Kumar Yadav[1], Sanjay Kalia[2], Kuldeep Singh[1†] and Ashok Kumar[1]

[1]Division of Germplasm Evaluation, Indian Council of Agricultural Research-National Bureau of Plant Genetic Resources, New Delhi, India, [2]Department of Biotechnology, Ministry of Science and Technology, Government of India, New Delhi, India

Linseed or flaxseed is a well-recognized nutritional food with nutraceutical properties owing to high omega-3 fatty acid (α-Linolenic acid), dietary fiber, quality protein, and lignan content. Currently, linseed enjoys the status of a 'superfood' and its integration in the food chain as a functional food is evolving continuously as seed constituents are associated with lowering the risk of chronic ailments, such as heart diseases, cancer, diabetes, and rheumatoid arthritis. This crop also receives much attention in the handloom and textile sectors as the world's coolest fabric linen is made up of its stem fibers which are endowed with unique qualities such as luster, tensile strength, density, bio-degradability, and non-hazardous nature. Worldwide, major linseed growing areas are facing erratic rainfall and temperature patterns affecting flax yield, quality, and response to biotic stresses. Amid such changing climatic regimes and associated future threats, diverse linseed genetic resources would be crucial for developing cultivars with a broad genetic base for sustainable production. Furthermore, linseed is grown across the world in varied agro-climatic conditions; therefore it is vital to develop niche-specific cultivars to cater to diverse needs and keep pace with rising demands globally. Linseed genetic diversity conserved in global genebanks in the form of germplasm collection from natural diversity rich areas is expected to harbor genetic variants and thus form crucial resources for breeding tailored crops to specific culinary and industrial uses. Global genebank collections thus potentially play an important role in supporting sustainable agriculture and food security. Currently, approximately 61,000 germplasm accessions of linseed including 1,127 wild accessions are conserved in genebanks/institutes worldwide. This review analyzes the current status of *Linum* genetic resources in global genebanks, evaluation for agro-morphological traits, stress tolerance, and nutritional profiling to promote their effective use for sustainable production and nutrition enhancement in our modern diets.

KEYWORDS

characterisation, genetic resources, genebank collections, genomics, linseed, flax, nutrition enrichment

Introduction

Linseed or flax (*Linum usitatissimum* L.) has been cultivated for seed oil and stem fiber since ancient times across the world. Globally, six countries (Canada, Kazakhstan, Russia, China, USA, and India) are the major producers (1). At present, flax (fiber-type) ranks as the third largest textile crop, whereas linseed (oil-type) ranks the fifth among the oil crops in the world (1). True to the species name '*usitatissimum*' meaning 'very useful,' linseed/flax stem, seeds, and seed oil have a wide range of applications in the preparation of food, nutritional and industrial products. The oil has wide industrial utility in paints and varnishes because of its unique drying properties attributable to its distinctive fatty acid composition (2). Other diverse uses of linseed oil include the manufacturing of hardboards, brake linings, printing ink, linoleum, and soaps. Fiber flax is used in handloom, textiles, and polymeric composites owing to its natural fibers endowed with unique luster and tensile strength (3, 4). The utilization of linseed plants for food, feed, fiber, and value added products has been reviewed comprehensively in the recent past (5, 6).

Linseed has emerged as a well-recognized nutrition-rich food because of its high omega-3 fatty acid (α-linolenic acid; ALA) content, dietary fiber, high quality protein, and lignans. With respect to omega-3 fatty acids, linseed is considered one of the richest plant-based sources and has around 55% ALA of the total fatty acids (7, 8). It also has an impressive omega 6/omega 3 fatty acid ratio of 0.3:1. As the modern-day diet predominates high omega 6 fats, nutritionists recommend a higher intake of essential omega 3 fats in food, which offer tremendous health benefits. Linseed has high fiber content (27.4%) along with high protein content (18.29%) (5). The seeds contain up to 800 times more lignans than other plant foods (9) and are principally composed of secoisolariciresinol diglucoside (SDG) (294–700 mg/100 g) (10–12). ALA together with SDG and dietary fiber has been reported to lower the risk of multiple chronic ailments such as heart diseases, stroke, hormonal disorders, cancer, diabetes, and rheumatoid arthritis (13–16). Currently, cancer is ranked as the second most important cause of death worldwide leading to a heavy global economic toll estimated at $1.16 trillion *per annum*. Over 2 million new cases were diagnosed globally in the year 2020, of which 24.5% accounted for breast cancer (17). Daily consumption of 25 g ground flax or 50 mg SDG has been found to be associated with a decreased risk of breast cancer, its recurrence, and mortality risk among survivors (10, 18). A survey reported increased intake of flaxseed in regular diets in cancer patients (19) since people prefer to use natural therapies in addition to conventional medical treatments for the management of the disease. The integration of linseed into the food chain as a functional food is evolving continuously, and it is popularly being called a 'nutritional punch'. Whole seeds, milled powder, extracted oil, and mucilage are extensively used as nutritional additives in the preparation of baked/ready-to-eat cereal products, bars, salad dressings, bread, muffins, sweets, processed meat and spaghetti, etc. (5). Renewed interest has led to an increase in consumer demand for linseed-based products not only for culinary use but also for novel industrial applications, such as geotextiles, biopolymers, and biofuels. However, unpredictable environmental stresses, such as drought, salinity, heat, diseases, and pest pressure, result in huge losses in yield and quality of oil/fiber (20–23). Hence, in the near future, climate change may exert strong pressure on flax breeding programs to adapt cultivars to changing conditions at an accelerating rate. Furthermore, the narrow genetic base of modern cultivars poses a major constraint to achieving sustainable yields to cater to diverse needs. Genebank collections of cultivated and wild species of the genus *Linum* from natural diversity-rich areas are important sources of genetic diversity for finding new alleles to meet new environmental challenges. Keeping this in view, the present review article is intended to apprise (1) an overview of the origin, domestication, taxonomy, and gene pool; (2) the status of global *ex-situ* collections of *Linum* sp., (3) trait-specific accessions identified through germplasm evaluation for agro-morphological traits, nutrition profiling, and biotic and abiotic stress tolerance in target environments to promote the effective use of these materials; (4) enhancement of the use of trait specific germplasm through the application of recent advances in genomic resources, molecular and biotechnological tools; and (5) documentation and access to germplasm/information through a global exchange system to facilitate transboundary flow.

Origin and domestication

Vavilov (24) described two centers of origin based on two distinct morphotypes, linseed (oil) type (short stature, bushy, profusely branched, and small seeded for the seed or oil purpose) originated in south-western Asia comprising India, Afghanistan, and Turkey and the fiber flax type (erect, long, pliable, unbranched stem with few branches restricted to top and bold seeded for fiber purpose) originated in the Mediterranean region including Asia Minor, Egypt, Algeria, Spain, Italy, and Greece. Later, Zeven and de Wet (25) and Damania et al. (26) reported Central Asia as the primary center of origin and the Mediterranean region as the secondary center of origin.

The domestication and subsequent spread of linseed across continents have not been clearly delineated. However, the archaeological evidence suggests that it was domesticated around 8,000 years ago in the Fertile Crescent and consequently spread to Europe and the rest of the world (27). Other dominant regions of diversity include the Indian subcontinent, Abyssinia, and the Mediterranean (24), where the present-day domesticated form, *L. usitatissimum*, originated from the wild ancestor *L. bienne* in geographic isolation. An independent domestication event in Central Asia (Indo-Afghan region) resulted in the development and production of fiber varieties as reported by Duk et al. (28). Crops, such as linseed, having multiple utilities may carry variable domestication signatures. Therefore, multiple domestication events have been suggested to be associated with different regions and included lineages that contain oil, fiber, and winter varieties (29). The long-term domestication of stem fiber and seed oil has diversified it into different types, including fiber, oil, and intermediate (dual-purpose) type (30–33). The first flax varieties were reported to be oil type, and later domestication for fiber use was reported, and multiple paths of flax domestication have been suggested (29, 34, 35). Very recently, Guo et al. (36) also confirmed that oil flax is the ancestor of cultivated flax and revealed signatures of artificial selection during the oil-to-fiber type transition.

Taxonomy, gene pool, and interspecific hybridization

Linum is the largest genus in the family *Linaceae* and consists of nearly 230 species with diverse chromosome numbers ranging from

2n = 16, 18, 30, 36, and 60 or more (37, 38). The scientific names and classifications used for *Linum* taxa in local, national, or regional floras are not consistent. The Global Biodiversity Information Facility lists only 252 accepted species names and 22 doubtful species names in the genus *Linum* (39), whereas Turkish and European floras list more species. The database "Plant List" (40) includes 460 scientific plant names of species rank for the genus *Linum* of which 108 are accepted species names and 229 are unresolved. Many synonyms exist and are used in scientific communications without cross-references. For example, *L. bienne* and *L. angustifolium* are considered synonymous in several databases, such as GRIN, Flora Europaea, The Plant List, and NCBI taxonomy; however, genome analysis based on molecular markers showed that *L. bienne* is a subspecies of *L. usitatissimum* rather than a separate species (41, 42). Several morphological, cytological, and molecular characterizations have revealed that *L. bienne* is the wild progenitor of cultivated flax (35, 43–52).

Despite many taxonomic, cytological, and evolutionary studies conducted for the genus *Linum,* flax genetic resources could not be classified into distinct gene pools as proposed by Harlan and de Wet (53) or modified by Gepts and Papa (54). In addition to classification, there are ambiguities in the identification of species, mainly *L. perenne, L. lewisii*, *L. flavum,* and *L. africanum;* hence, taxonomic revisions had been suggested (55, 56). Jhala et al. (57) also emphasized that species delimitation needs clarification for classification and efficient conservation of genetic diversity in the genus *Linum*. Thus, the absence of a coherent taxonomic review for reference to communicate interspecific diversity has made *Linum* systematics and taxonomy unclear. Understanding this requirement, a recent conspectus was presented by Fu (58) using flax as a case, and different flax species were assigned to various gene pools based on the existing literature. *L. usitatissimum* and *L. bienne* were categorized under the primary gene pool as they are interfertile and share the same chromosome number. Interspecific crosses between *L. usitatissimum* and other species, such as *L. africanum, L. angustifolium, L. corymbiferum, and L. decumbens,* having n = 15 chromosomes have been reported as successful (45). In addition, successful crosses with *L. nervosum*, *L. pallescens*, *L. africanum*, *L. corymbiferum*, *L. decumbens*, *L. hirsutum*, *L. floccosum*, and *L. tenue* have been reported in either direction based on crossability studies of cultivated flax with wild species (38, 59). The tertiary gene pool includes all 200 other species of the genus *Linum* that cannot hybridize with cultivated flax but could be exploited using advanced biotechnological tools (Figure 1). Several *Linum* species have the potential for beneficial trait introgression such as for lowering the linolenic acid content (*L. tenuifolium*, *L. sulcatum, L. hudsonoides*); drought and cold hardiness (*L. perenne*); resistance to linseed bud fly and *Alternaria* blight (*L. grandiflorum, L. bienne*); oil, fiber quality, and yield improvement (*L. bienne*), number of tillers (*L. strictum*); resistance to rust (*L. grandiflorum, L. bienne, L. africanum, L. creptans, L. flocossum, L. gallicum, L. marginale, L. perenne, L. strictum, L. tenue, L. trigyna L. alpinum, L. corymbiferum, L. hispidum); and* medicinal use as purgative (*L. cartharticum* L.) (48, 57, 60–62). The latest interest has been to explore the cut flower potential of perennial flax (*L. austriacum*, *L. perenne*, and *L. lewisii*) for ornamental use in floral arrangements (62). Thus, despite its huge potential, interspecific hybridization in *Linum* is hitherto unexplored to a large extent. A well-defined breeding program aided by biotechnological and molecular biology techniques is required to

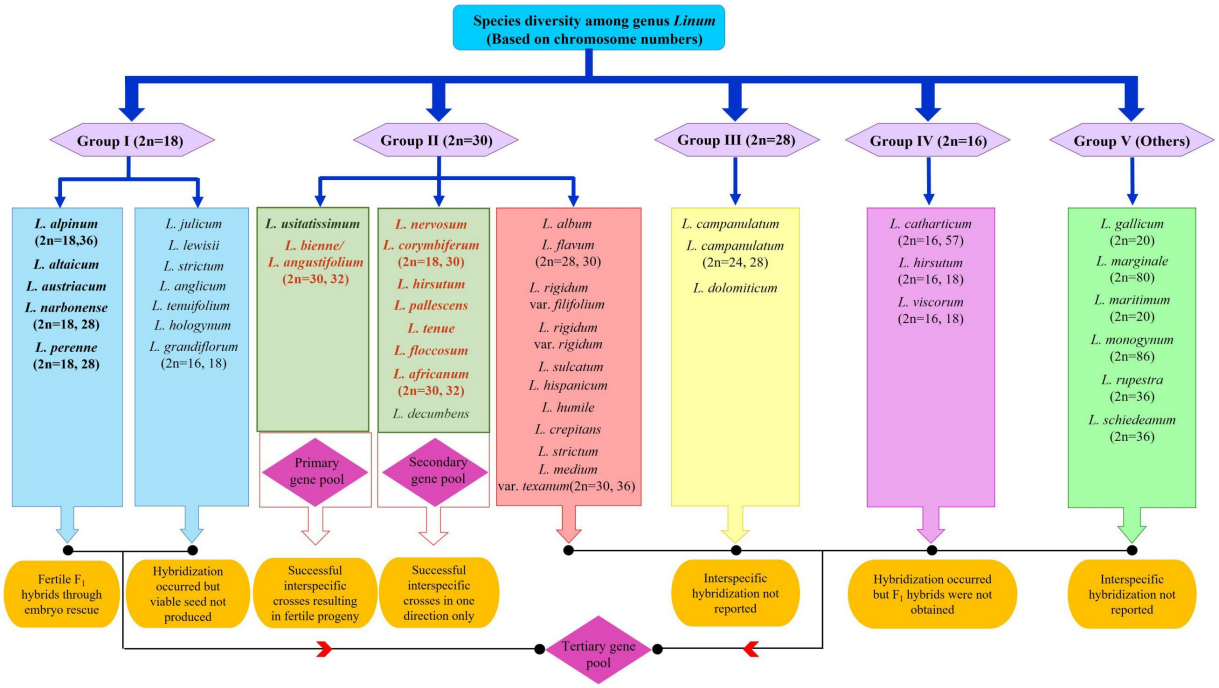


FIGURE 1
Species diversity in the genus *Linum* and status of interspecific hybridization among different gene pools. Grouping is based on chromosome numbers as proposed by Gill (37). Bold letters indicate the greatest potential to hybridize with cultivated linseed. The colored font indicates most potential wild species for utilization in breeding programs as donors for specific traits of economic importance such as oil and fiber quality, disease resistance, and abiotic stress tolerance.

harness the potential of the wild gene pool and support germplasm enhancement.

Global *ex situ* holdings of *Linum* genetic resources

Earlier reviews on *ex situ* collections were presented by Maggioni et al. (63) who provided the status of *ex situ* germplasm of linseed in Europe and Diederichsen (55) for global collections. The N.I. Vavilov Research Institute for Plant Industry (VIR) in St. Petersburg, Russia, and the All-Russian Flax Research Institute (VNIIL) at Torzhok, constitute the largest collections of approximately 6,000 accessions at each institute (55, 64, 65). The Indian National GeneBank (INGB) at the Indian Council of Agricultural Research-National Bureau of Plant Genetic Resources (ICAR-NBPGR), New Delhi, holds about 2,900 accessions (66). In addition, around 2,942 accessions are being maintained by the All India Coordinated Research Project on Linseed, ICAR-Indian Institute of Oilseed Research, Hyderabad, India (https://aicrp.icar.gov.in/linseed/ (67)). The Plant Gene Resources of Canada (PGRC) holds around 3,551 accessions of cultivated and 152 accessions of 25 wild species assembled from 72 countries. Currently, a total of 59,786 germplasm accessions of *L. usitatissimum* and 1,129 accessions belonging to different wild *Linum* species are conserved worldwide (Figure 2; Supplementary Table 1).

Evaluation of *Linum* germplasm collections

Agro-morphological traits

Many diversity assessment studies were conducted in PGRC flax collection and a wide range of variations for important traits, such as

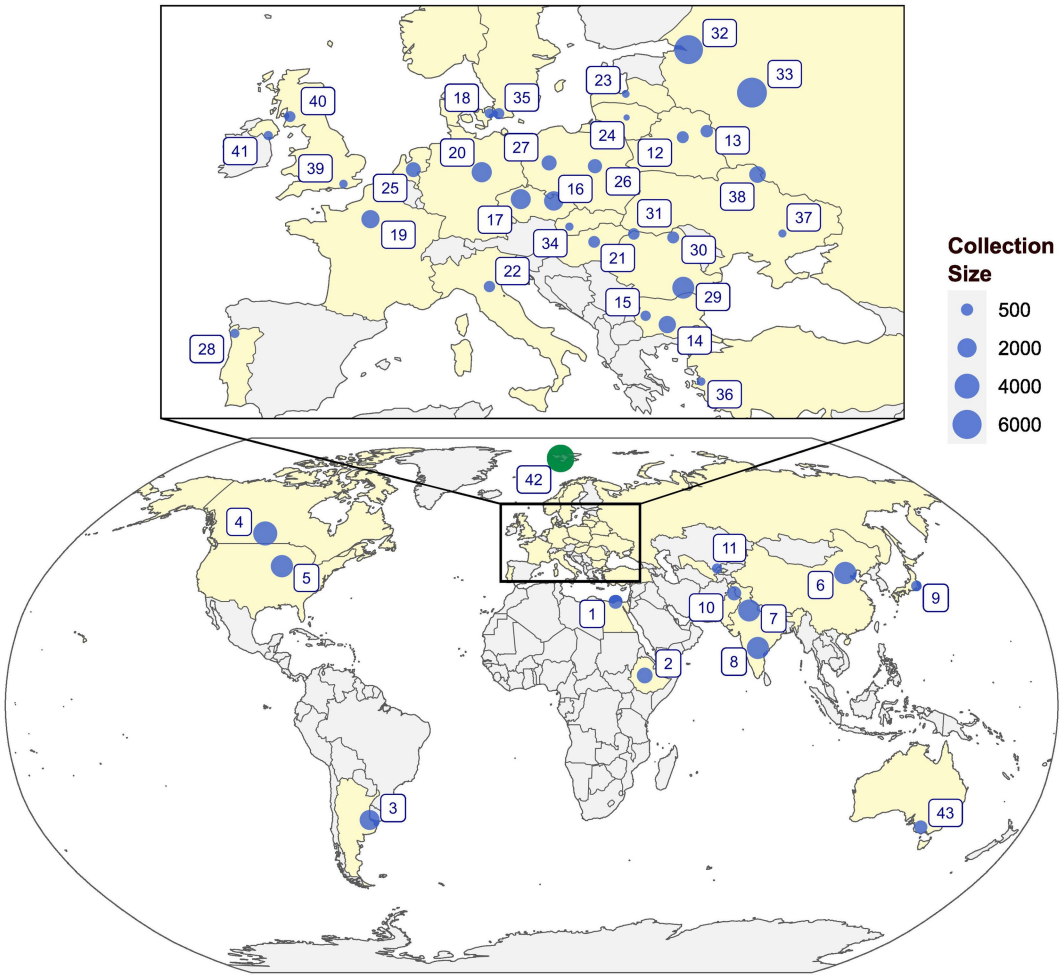


FIGURE 2
An overview of the global genebanks and institutes holding major collections of *Linum* genetic resources. 1: NGBGR, Egypt; 2: EBI, Ethiopia; 3: BG-CNIA, Argentina; 4: PGRC, Canada; 5: NCRPIS, United States; 6: ICS-CAAS, China; 7: NGB, ICAR-NBPGR, India; 8: AICRP, ICAR-IIOR, India; 9: NARO, Japan; 10: BCI, Pakistan; 11: UzRIPI, Uzbekistan; 12: NA, Belarus; 13: NA, Belarus; 14: IPGR, Bulgaria; 15: ABI, Bulgaria; 16: AGRITEC, Czechia; 17: CRI, Czechia; 18: NA, Denmark; 19: INRAE-VERSAILLES, France; 20: IPK, Germany; 21: NODiK, Hungary; 22: CREA-CI-BO, Italy; 23: LSFRI, Latvia; 24: LIA, Lithuania; 25: CGN, Netherlands; 26: IHAR, Poland; 27: IWNIRZ, Poland; 28: BPGV-INIAV, Portugal; 29: NARDI, Romania; 30: BRGV, Romania; 31: SCDA Livada, Romania; 32: VIR, Russia; 33: VNIIL, Russia; 34: SVKPIEST, Slovakia; 35: NordGen, Sweden; 36: AARI, Turkey; 37: IOK, Ukraine; 38: IBC, Ukraine; 39: RBGK, United Kingdom; 40: SRUC, United Kingdom; 41: NA, United Kingdom; 42: SGSV, Norway; and 43: AGG, Australia.

the onset of flowering (37–69 days), plant height (17–130 cm), and thousand seed weight (2.8 to 11.5 g), were reported (68–71). You et al. (72) phenotyped the PGRC flax core collection for agronomic, seed quality, fiber, and disease resistance traits (Pasmo, Powdery mildew, and *Fusarium* wilt) in multilocation-year environments and reported significant phenotypic variation in both fiber and oil accessions. Similarly, Worku et al. (73) evaluated Ethiopian collection for agronomic traits and Zhuchenko and Rozhmina (65) evaluated flax landraces for fiber quality and enlisted many accessions with superior trait value. In India, agro-morphological characterization and diversity analyses have shown broad range of phenotypic expression and many trait-specific accessions for early flowering, early maturity, oil content, bold seededness, high-test weight, and ALA were identified (5, 74–79). In pale flax, phenotypic diversity was studied by Diederichsen and Hammer (43) in Canadian and Uysal et al. (80) in Turkish germplasm. A high range of variation in pale flax was recorded for traits, such as plant height (38.4–123.3 cm), no. of days until the start of flowering (97–129), days to maturity (178–207), thousand seed weight (1.10–2.70 g), and seed color. Researchers are exploring advanced technologies, such as genomic tools and marker-assisted breeding, to accelerate the development of high-quality linseed varieties with desirable traits.

Germplasm evaluation for nutritional and nutraceutical traits

Biochemical characterization and evaluation of linseed for nutritional traits have been performed around the world. Green and Marshall (81) studied 214 accessions of linseed for the content of seed oil, seed weight, and fatty-acid composition and observed that the oil concentration in flaxseed was in the range of 33.3–46.4%, which positively correlated with high seed weight. Earlier Zimmerman and Klosterman (82) also reported a similar oil content in 1,175 linseed germplasm accessions. Variation in seed oil content, as well as the fatty acid composition of cultivated flax germplasm accessions at PGRC grown in western Canada, was studied by Diederichsen and Fu (68), Diederichsen and Raney (69), Diederichsen et al. (71, 83) in 2,934 germplasm accessions and a min (31.4%) to max (45.7%) oil content was found to be contributed by brown seeded linseed (2730) and yellow seeded accessions (84). Various Indian researchers (5, 76, 77, 79, 85) evaluated small subsets of linseed germplasm for qualitative and quantitative traits and reported oil content between 29.4 and 42.6%. For ALA content, a range of 48.08 to 57.58% was reported by Bayrak et al. (86) in Turkish and Romanian genotypes, while a similar variation (48.9 to 59.9%) was recorded in a Polish flax collection by Silska (87). In the Indian germplasm, the ALA content varied from 39.5% in germplasm accession IC564687 to 57.1% in IC564631 (85) and to as low as 33.14% (88). The proportion of ALA in a huge set of PGRC flax collection (2,243 accessions) was reported to be 39.6 to 66.7% with a mean of 52.6% (55).

Linseed also has high protein content which generally varies from 10.5 to 31% (7, 11). Flaxseed meal is a significant by-product (de-oiled) extracted from the processing of flaxseed and is generally rich in protein content of up to 40%. Protein contentmay vary with the environmental, genetic factors, processing techniques and is negatively correlated with oil content. Oomah et al. (89) evaluated 109 accessions to determine the composition of carbohydrates, protein, and oil content. Until recently, sparse information has been reported on germplasm evaluation for protein content in linseed. Similarly, the evaluation of linseed germplasm for SDG content has been performed to a limited extent only. A varied amount of SDG content ranging from 12.9 to 14.3 mg/g (90, 91) and 11.9–25.9 mg/g in whole linseed (92) has been reported. Thus, genotypic variation had a profound effect on the seed SDG content.

Recently, seed coat mucilage content has been recognized as a new character of interest for industrial applications. Large scale evaluation for mucilage content (assessed as Mucilage indicator value; MIV) was carried out by Diederichsen et al. (83), who screened 1,689 accessions from the PGRC collection and identified potential genetic resources. Canadian germplasm had been reported to have higher MIV (22.1 to 343.4 cSt mL g^{-1}) than North American cultivars (90.6–246.1 cSt mL g^{-1}). The scarcity of information on the chemical composition and functional properties of flax mucilage has limited its use in industries (93). However, recently, linseed mucilage has been recognized as having high technological value as a retardant polymer in pharmaceutical applications as well as in the food industry as a cost-effective natural polymer. Table 1 summarizes the major studies reporting variability for agro-morphological, phenological, fiber, and nutritional traits in flax germplasm and sources of trait-specific superior accessions. The health-related properties of linseed and their rejuvenated importance in animal and human nutrition have stimulated research and breeding programs for novel traits in germplasm collections. The current status of this research area is promising, with several studies identifying linseed varieties with high levels of desirable nutritional and nutraceutical traits. The future of linseed germplasm evaluation for nutritional and nutraceutical traits seems promising with an increasing focus on developing high-quality varieties and functional food products that can provide health benefits beyond basic nutrition.

Evaluation for major biotic stresses

Among biotic stresses, wilt caused by *Fusarium oxysporum* f. sp. *lini* has been recognized as the most devastating disease worldwide resulting in 80 to 100% loss in yield (115, 116). Differential diversity in *Fusarium* wilt tolerance in flax with higher resistance in accessions from East Asia, moderate resistance in American accessions, lower than average resistance in accessions from Europe and the Indian subcontinent was reported by Diederichsen et al. (117). Although most domesticated varieties are moderately resistant to *Fusarium* wilt (118, 119), the low genetic diversity of flax varieties and climate change may lead to increased aggressiveness of pathogenic breeds, substantially increasing the risk of disease. Another important fungal pathogen is *Alternaria linicola* Groves and Skolko, which is mostly prevalent in northwestern Europe and causes seedling blight (120), whereas *A. lini* Dey is found predominantly in the Indian subcontinent and causes flower and stem blight. Field evaluations under artificial epiphytotic conditions resulted in the identification of either no or very few lines as resistant (84, 121–124).

Powdery mildew, caused by the fungus *Oidium lini* Skoric, is another important disease of flax, which is reported to cause yield losses between 12–38% in India (125), upto 18% in the United Kingdom (126) and 10–20% in Canada (127). Sources of resistance to powdery mildew have been reported (121, 128, 129) in

TABLE 1 Potential accessions identified in linseed and fiber flax germplasm for various agro-morphological and nutritional traits.

Traits	Trait value	No. of accessions studied	Promising accessions* (value)	References
Early flowering	<60 days for 50% flowering	111	IC345409 (55.0); IC345397 (55.0), IC345425 (58.5)	(76)
		50	Shival (41.47)	(94)
			EC704 (47.61), EC41741 (55.47)	(95)
		103	IC397953 (58)	(77)
		220	IC0525939 (57.28), IC0523807 (58.44), IC096539 (59.95)	(79)
		198	59 accessions (<57 days)	(73)
		191	IC0096539 (58), IC0096496 (60)	(5)
		58	IC15888, IC113156, IC11310, IC426932, IC113105 (<60 days)	(96)
Early maturity	<120 days	50	Shival (76.30)	(94)
		191	IC0096539 (102), IC0096496 (97)	(5)
		220	IC0523807 (118.26), IC0525939 (116.76)	(79)
		111	IC268349 (89), IC345425 (92)	(76)
		198	30 accessions (<112 days)	(73)
Number of capsules per plant	>150	220	IC0053278 (267.52), IC0384578 (280.26)	(79)
		103	EC541212	(77)
		191	IC0426935 (168.55)	(5)
Thousand seed weight	>8 g		EC41741 (8.679 g), Ruchi (8.393 g)	(95)
		220	EC0041469 (9.51), EC0041700 (10.93), EC0041720 (9.49 g)	(79)
		191	IC0096490, IC0096488, IC0096489, IC0096543 (>8 g)	(5)
		200	CIli 2,719 (~10 g)	(36)
		1,689	CN98192 (10.87 g)	(97)
		36	Bekoki-14 (8.60 g) (Ethiopian germplasm)	(98)
Plant height	Tall (>110 cm)	3,087	CN101419 (130 cm)	(70)
	Dwarf (<45 cm)	3,087	CN95176 (20 cm)	(70)
		111	IC345425 (41.1 cm)	(76)
Primary branches/plant	>10	50	IC54970 (16.20 cm)	(94)
		36	Acc. 13,507 (11.90 cm), 207,789 (13.80 cm), 208,794 (11 cm), 212,512 (13.50 cm) (Ethiopian germplasm)	(98)
Seeds per capsule	>8	50	IC56363 (8.70)	(94)
		198	195 accessions (>9)	(73)
		191	EC0718845, IC0267682	(5)
		60	Acc. No. 10007, 10,008, 10,061, 10,111, 10,119, 10,162, 10,169, 10,185, 10,192, 10,235, 10,064, 10,256, 10,260 (>10) (Ethiopian germplasm)	(99)
Bold capsules	>50mm^2 capsule area	220	EC0041700 (52.58 mm^2)	(79)
		191	IC0094487, IC0096488	(5)
Large seed length, width, and seed area			EC41741, EC704, Ruchi	(95)
		200	Cili 2,719 (PI 523932)	(36)
		191	IC0054949, IC0054954, IC0096490 (15.13 mm^2 seed area)	(5)
		220	EC0041469 (14.04 mm^2), EC0041700 (14.63 mm^2 seed area)	(79)
Seed yield/plant		49	Litwania-9 (Litwanian variety), Evelen (French variety)	(100)
		64	Acc. 243,807, 243,810, 244,809, 231,457, 230,822 (Ethiopian germplasm)	(101)
		81	PGRC/E 10104, 10,120 (Ethiopian germplasm)	(102)
		151	Shweta, Gaurav	(103)

(Continued)

TABLE 1 (Continued)

Traits	Trait value	No. of accessions studied	Promising accessions* (value)	References
Technical stem height	>35 cm	198	45 accessions (<36.50 cm) (Ethiopian germplasm)	(73)
		7	Ariane (97.4 cm), Viking (93.3 cm) (French cultivars)	(104)
		103	EC5412149 (Taller stalk)	(77)
Cell wall (%)			Upto 80% (recombinant inbred line population derived from CDC Bethune/Macbeth)	(105)
Straw yield/plant	>3 g/plant	7	Giza 5 (3.32 g/plant) and Giza 6 (3.16 g/plant) (Egyptian cultivars)	(104)
		7	INA, Emilin, Rolin, Daniela, Madras, Istru	(106)
		49	Litwania-5 (3.29 g/plant)	(100)
Oil content	>40%	191	IC0096490 (42.9), IC0268345 (42.7)	(5)
		111	IC345425 (41.5), IC345447 (41.4), IC345417 (41.4), IC345423 (41.3)	(76)
		103	IC567363	(77)
			IC564681 (42.6)	(107)
		120	CN18973, CN18979, CN19003, CN19005, CN30861, CN97306, CN97307, CN97308, CN97334, CN97366, CN97396, CN97430, CN97430B (>44)	(108)
		243	Upto 48.5% (recombinant inbred line population derived from CDC Bethune/Macbeth)	(105)
		49	Vaiko (45)	(100)
		36	Acc. No. 208425 (40.05) (Ethiopian germplasm)	(98)
		84	IC564681 (42.6)	(85)
		151	Shubhra (45.09), Laxmi-27 (45.06), Mukta (44.94), Shweta (44.25)	(103)
High α-linolenic acid (ALA)	>57%		NuLin™ 50 (68%) PI523353 (57.61%)	(109)
		50	RLC 92 (58.71), Janki (57.45)	(94)
		198	25 accessions (>57%)	(73)
			IC564631 (57.1%)	(107)
			UGG5-5 (63–66%) (Breeding line)	(110)
		120	CN18979, CN18980, CN18989, CN19004, CN19157, CN19158, CN30861, CN97214, CN97393, CN97406, CN97402, CN97424 (>57%)	(108)
		243	Upto 61.7% (recombinant inbred line population derived from CDC Bethune/Macbeth)	(105)
		84	IC564631 (57.1%)	(85)
			EC541221 (66%)	All India Coordinated Research Project on Linseed#, IIOR, Hyderabad, India
Low alpha-linolenic acid (ALA)			Linola™ or Solin™ (~3%)	(111) (112);
			SP2047 (Solin™ line) (2–4%)	(110)
			<29%	(113)
		84	IC564687 (39.5%)	(85, 107)
			Kiran (33.14%)	(88)
High Oleic acid content		84	IC564627 (32%)	(85)
		198	19 accessions (>19%)	(73)
		120	CN96958, CN96974, CN97083, CN97176, CN97238, CN97312, CN97064	(108)
		64	Acc. No. 13545 (21.4%) (Ethiopian germplasm)	(114)

(Continued)

TABLE 1 (Continued)

Traits	Trait value	No. of accessions studied	Promising accessions* (value)	References
Protein content	>25%	243	Upto 27% (recombinant inbred line population derived from CDC Bethune/Macbeth)	(105)
Mucilage content	Mucilage indicator value (MIV) >200 cSt mL g^{-1}	14	CN19004 (246.06), CN18973 (244.05), CN52732 (222.48)	(83)
		1,689	CN98100 (343.4), CN98254 (342.2)	

CN = Canadian number, PGRC (https://pgrc-rpc.agr.gc.ca/gringlobal/search). IC = Indigenous collection; EC = Exotic collection, Indian National GeneBank (http://www.nbpgr.ernet.in:8080/PGRPortal/(S(lzvbzn55u4enjtfhv3dybl3b))/default.aspx). PI = Passport information, Genesys (https://www.genesys-pgr.org/). Ethiopian germplasm (The Ethiopian Institute of Biodiversity Conservation (IBC/ETH); formerly The Plant Genetic Resources Centre Ethiopia, PGRC/E) (http://www.ebi.gov.et/). VNIIL = Russian Genebank (http://www.vir.nw.ru/; http://vniil.narod.ru/). *Only the accessions having National id/released cultivars are mentioned. #https://aicrp.icar.gov.in/linseed/.

natural hotspot areas (Table 2). Genes conferring resistance to local *O. lini* isolates have been identified in Indian flax breeding lines (152, 153) and Canadian and European oil cultivars (127). Recently, several powdery mildew resistance QTLs have been identified (154, 155).

In the case of linseed rust, caused by the fungus, *Melompsora lini,* field evaluations to identify sources of resistance, physiologic races, and inheritance of resistance have been reported (137, 138, 156–159) (Table 2). Misra (160) and Singh et al. (116) reported that during screening of wild *Linum* species for rust resistance, the species *L. africanum, L. augustifolium, L. creptans, L. flocossum, L. gallicum, L. marginale, L. perenne, L. strictum, L. tenue,* and *L. trigyna* were found resistant to all the races, while *L. mysorense* and *L. pallecum* were found susceptible.

Among insect pests, defoliators, such as semilooper (*Plusia orichalcea* Fabr), Lucern caterpillar (*Spodoptera exigua* Hubn.), tobacco caterpillar (*Spodoptera litura* Fabr.), and Bihar hairy caterpillar (*Spilarctia obliqua* Walk.), appear as sporadic and region-specific pests with high incidence in some cropping seasons. Bud fly (*Dasyneura lini* Barnes) is a key pest of this crop in Asia, particularly India, Pakistan, and Bangladesh, while it appears as a pest of less economic significance for flaxseed/fiber flax in Europe (161). The linseed incidence of *Dasyneura lini* was first recorded in India by Pruthi and Bhatia (162) and has been reported as the main pest causing yield losses upto 90% in central India (163–165). A few accessions were recorded as resistant (below 10% bud fly infestation score) or moderately resistant germplasm under field conditions in deliberate late planting (139, 165–169) (Table 2). Studies have reported that varietal attributes, such as short flowering periods, flower shape, thin sepals, and higher polyphenol content, render resistance or escape from midges (139, 140, 170, 171). However, limited breeding efforts have been taken to improve this complex polygenic trait.

Evaluation for major abiotic stresses

Drought, salinity, and heat are the major environmental stressors that negatively affect flowering initiation, plant height, seed yield, straw, and fiber yield in linseed. Heller et al. (172) observed that the fiber formation process in the stem is the most intensive in the vegetative growth period till the end of flowering and therefore weather conditions in this period determine the fiber content and quality. Heller and Byczyńska (173) reported a reduction in fiber yield from 35–50% while screening 51 flax cultivars under water limitation to identify superior yielding lines. In most developing countries, linseed is extensively cultivated in rain-fed or moisture-scarce *utera* conditions and therefore, root system architecture plays an important role in determining water and nutrient acquisition from soil. Linseed has a shallower and less aggressive tap-rooted system compared to other oilseeds such as canola, sunflower, and safflower (174, 175). A deeper and dense root system could extract water more efficiently from deep soil and is thus advantageous, particularly in rainfed production areas (144, 176). Differential performance among flax genotypes under different moisture regimes attributed to variations of their root systems has been reported (97, 177, 178). Linseed is predominantly a rainfed crop in India and therefore Indian genotypes are likely to be more drought tolerant owing to long-time divergent selection under moisture-scarce conditions as also corroborated by several studies (97, 142, 147, 179).

Soil salinization significantly affects the growth and distribution of linseed (151). Only a few studies have reported screening of linseed germplasm against salinity-alkalinity stresses to identify salinity-tolerant lines using germination, biomass and K^+/Na^+ ratio (149, 150, 180–183), and the underlying biochemical and physiological mechanisms (151) (Table 2). Yu et al. (184) reported differentially expressed genes (DEGs) and saline-alkaline tolerant miRNAs in flax (Lus-miRNAs) for the first time. Wu et al. (185) identified genes that might enhance salt tolerance by increasing root length and improving membrane injury and ion distribution. Lately, Li et al. (148) screened 200 accessions of flax germplasm for salt tolerance and revealed that stress tolerance indices for germination and early root and shoot traits of the oil flax subpopulation were significantly higher than those of the fiber flax subpopulation. They suggested that oil flax may contain more resistance sites related to abiotic stress, and the screening of resistant germplasm and resistance genes should focus more on oil flax.

Heat stress has a significant impact on the climatic adaptation of linseed, particularly superior quality fiber cultivars as it is a cool-season crop. A rise in temperature during flowering and seed filling may lead to necrosis of the ovules, resulting in a reduced seed set (186). Gusta et al. (187) studied genetic variability among flax cultivars in response to temperature stress and reported negative effects on flowering and seed yield. Cross et al. (188) studied the adverse effects of heat stress (>40°C for 5 days) on reproductive organ functioning and reported a decline in boll formation and seed setting although the composition of the oil was not affected. Fofana et al. (189) reported that warmer and drier environmental conditions can cause a lowering of ALA by 5% owing to thermos-sensitivity of the linked enzyme, fatty

TABLE 2 Evaluation of linseed germplasm for tolerance to major biotic and abiotic stresses and sources of promising accessions.

Traits	No. of accessions studied	Traits value	Promising accessions*	References
Biotic stress resistance				
Fusarium wilt		Resistant (up to 5% wilt)	OL 1–3, EC1392, JRF 5, Padmini, JRF-3, RLC-2, RLC 33, R-7	(121)
			Ayogi, EC41656, JLS-9, RLC 46, Nagarkot, Padmini, Rashmi, R-552, Surabhi, Sweta, T-397	(130, 131, 132)
	297	Disease Severity Index (DSI)	25 accessions (DSI = 0)	(133)
Powdery mildew	294	Resistant (0.1 to 10% leaf area affected)	EC41562, EC1465, IC16392, Neelum	(121)
		Resistant (0.1 to 10% leaf area affected)	EC41656, EC322646, Meera	(134)
	150	Highly resistant (0% leaf area affected)-21 accessions	Parvati, Jeevan, R-17, RLC-148, RLC-151	(75, 128)
Alternaria blight	250	Resistant (1–10% leaf area infected)	EC22672, EC22704, EC41623	(135)
	200	Resistant	NP-8, NP-48	(84)
			Kiran, Jeevan	(136)
Rust	200		EC384154, EC1497, Baner, Nagarkot, JRF-3, Rashmi	(61)
			Hira, Mukta, Neelum, KL-1 (Surbhi)	(137)
			EC77959, JLS (J)-1, Jawahar 17, Jawahar-17, Garima, Himalini, Padmini, Rashmi, Sheela, Shweta, Meera, Surabhi, Nagarkot, Kiran	(138)
Budfly	288	Resistant (10% bud infestation)	EC1392, EC1424, IC15888, JRF-5	(139)
	60		Neela	(140)
	250		EC22596, EC22672, EC22704, EC22823, EC41528, EC41551, EC41593, EC41715, EC41581, EC41595, EC41623, EC26006, EC41404, EC41580, EC41690	(135)
			IC16382	(141)
Abiotic stress tolerance				
Drought stress	96	Thousand seed weight	CN98566	(97)
	115	Thousand seed weight	CN98712, CN98566	(142)
		High grain yield, Stress tolerance index	CN101595	
		Bundle weight	CN101052, CN101419	
	105	Improved root and shoot traits	CN98193	(143)
	105	Yield, mean productivity, and improved stress indices	CN19004, CN19003	(144)
	41	Superior plant height and root trait stability	CN33393, CN18987	(145)
	41	Total root length	CN98946	(145)
	119	Higher mean productivity and Stress tolerance index (STI)	13 lines from population of KO37 (Iranian breeding line) × SP1066 (Canadian breeding line)	(146)
	120	Total root length (Stability index >3)	CN101348, CN30861, CN18994	(147)
		Total root volume (Stability index >4)	CN101348, CN98056, CN18994	
		Root surface area (Stability index >3)	CN101348, CN30861, CN18994, CN33399	
		Yield under drought stress (>95 gm^{-2})	CN19003, CN19004, CN52732, CN100674	
	200	Germination, early root and shoot growth	VNIIL-180, 1,270, Torzhokij 4, VNIIL-409, Crepitam Tabor, XLB, BELADI Y 6903, CIli 1919 (PI 249989), CIli 2038, CIli 2047, ALSEE, MESSENIAS, OLEIFERA, 62/125–4, CIli 1832, TJK04-72 (PI 649756), TJK04-348 (PI 649760)	(148)

(Continued)

TABLE 2 (Continued)

Traits	No. of accessions studied	Traits value	Promising accessions*	References
Salt stress tolerance	2	More root length, shoot length, and higher proline and peroxidase activity under salt stress	NL-97	(149)
	10	Germination, seedling growth, and ion content	Sar_1-85 (Turkish cultivar)	(150)
	5	High antioxidant enzymes peroxidase (POD), superoxide dismutase (SOD), ascorbate peroxidase (APX), and malondialdehyde (MDA) content	Ariane (French) and Sakha-1 (Egyptian)	(151)

CN = Canadian number, PGRC (https://pgrc-rpc.agr.gc.ca/gringlobal/search). IC = Indigenous collection; EC = Exotic collection, Indian National GeneBank, (http://www.nbpgr.ernet.in:8080/PGRPortal/(S(lzvbzn55u4enjtfhv3dybl3b))/default.aspx). PI = Passport information, Genesys (https://www.genesys-pgr.org/). VNIIL = Russian Genebank (http://www.vir.nw.ru/; http://vniil.narod.ru/). *Only the accessions having National id/released cultivars are mentioned.

acid desaturase (FAD2). The molecular mechanism underlying heat stress tolerance in flax is largely unknown. Recently, Saha et al. (20, 190) attempted to describe the role of heat shock factors (HSFs) and DNA hypomethylation with regard to heat stress adaptation in flax.

Genomic resources in linseed

Molecular marker resources in linseed

The PCR-based dominant marker systems, such as RAPD, AFLP, and ISSR, have been employed in flax for linkage mapping and genetic diversity studies since the 1990s until the recent past (99, 191–201). In addition, retrotransposon-based inter-retrotransposon amplified polymorphism (IRAP) markers have been developed in flax and utilized for genetic diversity studies (202). The most preferred PCR-based markers, simple sequence repeats (SSRs) (110, 203), were identified in linseed to the tune of a few hundred until 2011 (50, 51, 204–207). From 2012 onward, there was a substantially higher number of SSRs reported in linseed by different research groups largely due to the advances in next-generation sequencing (NGS) technologies. Cloutier et al. (110) reported 1,164 and 342 SSRs from BAC-end genomic sequences (BESs) and ESTs, respectively. Kale et al. (208) used 454 GS-FLX platform for the sequencing of PCR amplicons and designed 290 SSRs. Furthermore, a reduced representation genome sequencing (RRGS) approach was also used to identify 1,574 SSR loci (209). More recently, 24,375 SSR motifs were identified employing a pseudomolecule-scale genome-wide scan of the flax genome for the development of SSR markers (210). From the identified SSR markers, the polymorphic loci have also been identified in different studies which were used for distinguishing flax and linseed cultivars (206, 207). In terms of functional marker resources, 580 regulatory gene-derived simple sequence repeat (ReG-SSR) markers from transcription factor-coding genes and long non-coding RNAs have been identified (211). In wild species of flax, *Linum bienne*, 44 microsatellite loci have been identified using genome skimming; of which, 16 have been used for genotyping six *L. bienne* populations (212). More recently, a few to hundreds of thousands of single nucleotide polymorphic (SNP) marker loci have been unraveled by an array of techniques such as reduced representation sequencing approaches and whole genome resequencing in linseed accessions (36, 78, 143, 213, 214). Detailed information on the major marker resources available in linseed has been shown (Supplementary Table 2). These resources, especially SNP markers, have enabled the construction of high-density linkage maps, identification of QTLs, and thereby a better understanding of the genetic architecture of several complex traits in linseed (78, 215–219).

Genetic diversity in linseed germplasm elucidated by molecular markers

Different marker systems, such as AFLP, RAPD, ISSR, IRAP, SSR, and SNPs, have been employed in linseed specifically for studying the genetic diversity (33, 79, 195, 197, 198, 202, 220–222). The genetic diversity studies with sufficiently large germplasm accessions (>100) have been considered in this review. The genetic diversity of 708 accessions of cultivated flax and 10 wild species was studied using polymorphic IRAP markers (202). The 708 accessions of cultivated flax comprised 143 landraces, 387 varieties, and 178 breeding lines from 36 countries. The robust 141 reproducible data points per accession were obtained from 10 polymorphic IRAP primers with 52% polymorphism and 0.34 Shannon diversity index. The study showed the highest genetic diversity in wild *Linum* species (polymorphism: 100% and Jaccard similarity: 0.57), followed by landraces (58%, 0.63), breeding lines (48%, 0.85), and cultivars (50%, 0.81).

Genetic diversity analysis of 168 linseed accessions predominantly of Indian origin was conducted using 50 SSR loci, which unraveled a total of 337 alleles (74). The mean Shannon's information index for all three populations was 0.23. Similarly, in the case of Ethiopian linseed landraces, IRAP and ISSR markers showed a comparable level of molecular diversity (PIC, 0.16; GD, 0.19) in 203 accessions (222). The genetic diversity of the PGRC flax core collection was studied using 448 microsatellite markers (33). The genetic structure of the core set showed two major groups with six sub-groups having weak population differentiation, weak relatedness (mean = 0.287), and abundant genetic diversity in the total panel (5.32 alleles per locus). The sub-groups were found to have a high proportion of private alleles. Similarly, 350 globally distributed flax genotypes were studied using 6,200 SNP

markers, which clustered the 350 accessions into seven sub-populations with moderate genetic diversity (average H = 0.22 and I = 0.34) (221). There was a significant positive correlation (r = 0.30 and $p < 0.01$) between genetic and geographic distances in the whole collection. Overall, the moderate to low genetic diversity reported in different studies on linseed using varied marker systems are as anticipated and could be accounted for the self-pollinated nature of the crop.

QTL mapping and genome-wide association analysis

In one of the earliest QTL mappings for *Fusarium* wilt resistance in flax using a doubled-haploid population, two AFLP markers (afB13 and afXR6) were found tightly linked to flax rust resistance (223). A linkage map-based QTL study with 329 SNP and 362 SSR markers using a RILs population of a cross of two Canadian varieties, CDC Bethune and Macbeth, had unraveled 20 QTLs for 14 traits comprising oil, fatty acid, iodine, and protein content as well as fiber quality traits (105). Interestingly, one SSR marker Lu2031 on LG4 (chromosome 4, coordinates: 14489225–14489333) was found to be linked to QTLs of five different traits including cell wall (fiber components), straw weight, seeds per boll, days to maturity, and yield (105, 217). Plant height and technical length are crucial traits for fiber flax. A total of 19 QTLs were identified for both traits following linkage map-based QTL analysis using two RIL populations and a total of 4,497 SNPs were anchored on 15 linkage groups with an average marker density of 2.71 cM (224). A high-density linkage map of flax was constructed using 112 F_2 plants and 2,339 specific-locus amplified fragment (SLAF) markers on 15 linkage groups with a total length of 1483.25 cM and a mean distance of 0.63 cM between two adjacent markers (216). The study also helped map 12 QTLs for 6 flax fiber-related traits. In biotic stress, three QTLs have been identified for powdery mildew resistance located on LG1, 7, and 9 accounting for 97% of the phenotypic variation and suggesting a dominant gene action (225). Until 2020, 313 QTLs for quantitative traits comprising seed yield & agronomic traits (155 QTLs), seed quality (75 QTLs), fiber (11 QTLs), and diseases (72 QTLs) have been identified in flax, most of which were mapped on chromosome-scale pseudomolecules (217).

In addition to biparental QTL mapping, genome-wide association studies (GWAS) have been extensively employed in flax for the genetic dissection of complex traits. In the early years of last decades, SSR markers have been used for association mapping of seed weight trait in linseed, which identified five SSR markers associated with the trait (50). Later, considerably high numbers of SNPs identified by reduced representation sequencing and whole genome resequencing approaches have been used in GWAS. The GWAS strategy was reported for genetic dissection of seed weight trait in flax by two independent research groups, unraveling the associated SNP markers and important candidate genes such as *PHO1*, *cytochrome P450,* and *ubiquitin-proteosome pathway genes* (36, 214). GWAS was also employed for genetic dissection of flowering time in linseed on 200 accessions of the Canadian flax core collection using 70,935 curated SNPs by single and multi-locus methods (ML-GWAS). A total of 40 quantitative trait nucleotides (QTNs) associated with 27 QTL were identified for flowering time accounting for 3.06–14.71% of trait variation (215). ML-GWAS was also used on a panel of germplasm accessions of the National Genebank of India for dissecting flowering time, maturity, and plant height trait using 68,925 SNPs identified by genotyping by sequencing. The study identified 53, 30, and 27 stable QTNs for flowering time traits, days to maturity, and plant height, respectively (78). GWAS was also successfully employed for identifying genomic regions associated with pasmo resistance (PR) in flax in a panel of 370 core collections. In total, 10 different statistical models identified a total of 692 unique QTNs associated with 500 putative QTLs from six phenotypic PR datasets. Interestingly, 45 of the identified QTLs spanned 85 resistance gene analogs including a large toll interleukin receptor, nucleotide-binding site, and leucine-rich repeat (TNL) type gene cluster (213). In a similar fashion, GWAS have been deployed in the genetic dissection of other complex traits in linseed including fiber-related traits, fatty acid biosynthesis, mucilage and seed hull content, capsule numbers, branch numbers, and other important agronomic traits (21, 214, 226–228). Detailed information on GWAS in linseed has been presented (Supplementary Table 3). The identified QTLs/QTNs/associated markers/candidate genes are expected to facilitate linseed researchers for the crop improvement and tailoring niche specific, stress tolerant cultivars through genomics/marker-assisted breeding.

Whole genome sequence

In 2012, a *de novo* assembly of the flax genome sequence of the cultivar CDC Bethune was made available, having an estimated 81% genome coverage from 302 Mb non-redundant sequences (229). However, there was a misassembly observed in several regions at the genome level. This genome assembly was further improved using the BioNano genome (BNG) optical map of CDC Bethune in 2018 (230). The refined scaffold sequences and validation of the BAC-based physical map were performed followed by further scaffolding of BNG contigs to the super BNG contigs. These super BNG contigs were then assigned to the 15 flaxseed chromosomes with the help of the genetic maps. These pseudomolecules constituted total 316 Mb sequences covering 97% of the annotated genes (230). These resources were pivotal in localizing the earlier identified and new QTLs/markers for important traits to the chromosome-scale pseudomolecules (78, 215, 217). More recently, the genome of fiber flax cultivar Atlant was sequenced using the Oxford Nanopore and Illumina platforms reporting the complete assembly with a total length of 361.7 Mb (N50 = 350 kb) and 97.40% completeness (231). Furthermore, a chromosome-scale high-quality genome sequence of another fiber flax cultivar YY5 is reported with HiFi and Hi-C technology, which helped substantially improved the earlier assembly of fiber flax (232). Besides these, scaffold-level genome assemblies of the other three cultivars including linseed 'Longya-10', fiber purpose cultivar 'Heiya-14', and pale seed color flax are also available (233).

The plastid genome of *L. usitatissimum* was sequenced and assembled showing a typical circular DNA molecule of 156,721 bp length (234). The assembled plastid genome showed a total of 109 unique genes, 2 pseudogenes, 176 SSRs, 20 tandem repeats, and 39 dispersed repeats.

These whole genome sequence assemblies would be valuable resources for further fine mapping of causal genetic variants, precise positioning of markers, and thereby in genomics/marker-assisted breeding in linseed.

Effective utilization of *Linum* genetic resources and trait information

Linum core collections

The large size of germplasm collections conserved at genebanks poses restrictions for the evaluation, utilization, and maintenance of these resources (235). Therefore, the concept of "core collections" (containing approximately 5–10% of the whole collection) was introduced while preserving the maximum genetic diversity of the entire collection with minimum repetitiveness to promote utilization (236, 237). Diverse germplasm in the form of the core collection has stimulated the research to have more and important insights into the genetic variability conserved in huge collections in genebanks. Approximately 10,000 genetically diverse flax accessions in the form of core collections have been maintained in global genebanks (55). Initially, small collections representing donors for specific quality traits and disease resistance were assembled by Kutuzova (238) and Brutch (64) at The N.I. Vavilov Institute (St. Petersburg). Later van Soest and Bas (239) developed a core set of 84 accessions from 506 fiber flax accessions at The Centre for Genetic Resources at Wageningen in the Netherlands. Diederichsen et al. (70) assembled a PGRC flax core collection comprising 407 accessions encompassing both the oil and fiber morphotypes from a world collection of around 3,500 accessions. This core set has been characterized phenotypically for agronomic, fiber quality, and disease resistance traits (72) and to identify the candidate genes and QTLs for seed quality traits and drought tolerance (33, 50, 142, 147), thus contributing to the efficient utilization of genetic resources.

Documentation and access to germplasm and trait information

Documentation of data on germplasm accessions and their characterization and evaluation is of pivotal importance to facilitate the sustainable use of genetic resources. World Information and Early Warning System (WIEWS) maintained by the Plant Genetic Resources of Food and Agriculture Organization, Rome, Italy hosts information on major flax collections.[1] Based on the recommendations by IPGRI (240), most of the genebanks have developed their own documentation systems, wherein the passport data are linked to germplasm accessions. These data are often accessible to the users through genebank databases/websites (Supplementary Table 1). Figshare repository (Federal Research Center for Bast Fiber Crops, Russia) holds data on quantitative phenotypes of flax.[2] PGRC has provided full access to the passport, characterization, and evaluation of data of the whole flax collection along with photographs of each accession of the core set.[3] For European flax collections, the passport data is available in International Flax Database.[4] The PGR information of Indian linseed collection has been duly documented by ICAR-NBPGR in user-friendly databases such as PGR Portal and mobile/desktop apps Genebank, PGRMap (http://www.nbpgr.ernet.in/PGR_Databases.aspx). The evaluated germplasm and trait-specific accessions have been published through catalogs, workshop proceedings, and notification of germplasm registration (241–246). The Inventory of elite germplasm/genetic stocks is available at the NBPGR website.[5] The catalogs listing 50 linseed accessions as donors for specific traits including disease resistance (238) and fiber characterization data of 250 flax accessions (64) were published by the Vavilov Institute. The comprehensive dataset of flax accessions from the Russian Federal Research Center for Bast Fiber has been recently published by Rozhmina et al. (247). FIBexDB is a database based on transcriptomic data for flax where the expression pattern for genes and co-expression network are available and pairwise comparison can be made (248).

Exchange of germplasm and information

Universal accessibility to information through unified record maintenance in the system is essential to promote the wider use of genetic resources in breeding programs. With the aid of modern information technology tools, a well-developed global network system through inter-operatable data sets is in place now to facilitate the exchange of information and trait discovery. An overview of the evaluation of *Linum* genebank collections through comprehensive phenomics and genomics for the identification of germplasm traits relevant to climate resilience and sustainability is presented in Figure 3. Global genebank systems and databases facilitate the exchange of germplasm and the flow of information for enhanced utilization. Institutional, regional, and global platforms provide inter- operatable data sets and create customized data analysis tools to address the needs of end users and different stake holders (e.g., genebank managers, curators, researchers, breeders, students, and farmers). Accession-level information on plant genetic resources secured in genebanks can be retrieved through FAO WIEWS.[6] To ease the search of accession level information, a transition from individual genebank website access to multi-institutional/regional portals, such as the European Plant Genetic Resources Search catalogue (EURISCO; http://eurisco.ecpgr.org), the CGIAR system Wide Information Network for Genetic Resources (SINGER; http://www.singer.cgiar.org/), and the Germplasm Resources Information Network (GRIN; http://www.ars-grin.gov/) of United states Department of Agriculture (USDA), has helped a lot. On a global scale, GENESYS (https://www.genesys-pgr.org/) portal integrates data from the website as a single data entry point for users to mine genetic diversity and order germplasm of interest. Currently, it contains information on around 50% of the global accessions; of which, around 24,555 accessions belong to *Linum* sp., and this information is continuously evolving over time (249).

Conclusion and future perspectives

Germplasm evaluation is a critical aspect of linseed research, as it involves the identification and characterization of genetic resources

1 http://apps3.fao.org/wiews/
2 https://figshare.com/s/86a68ecfacf6872ef239
3 http://pgrc3.agr.gc.ca/
4 http://www.ecpgr.cgiar.org/database/crops/flax
5 http://www.nbpgr.ernet.in:8080/registration/InventoryofGermplasm.aspx
6 https://www.fao.org/wiews/en/

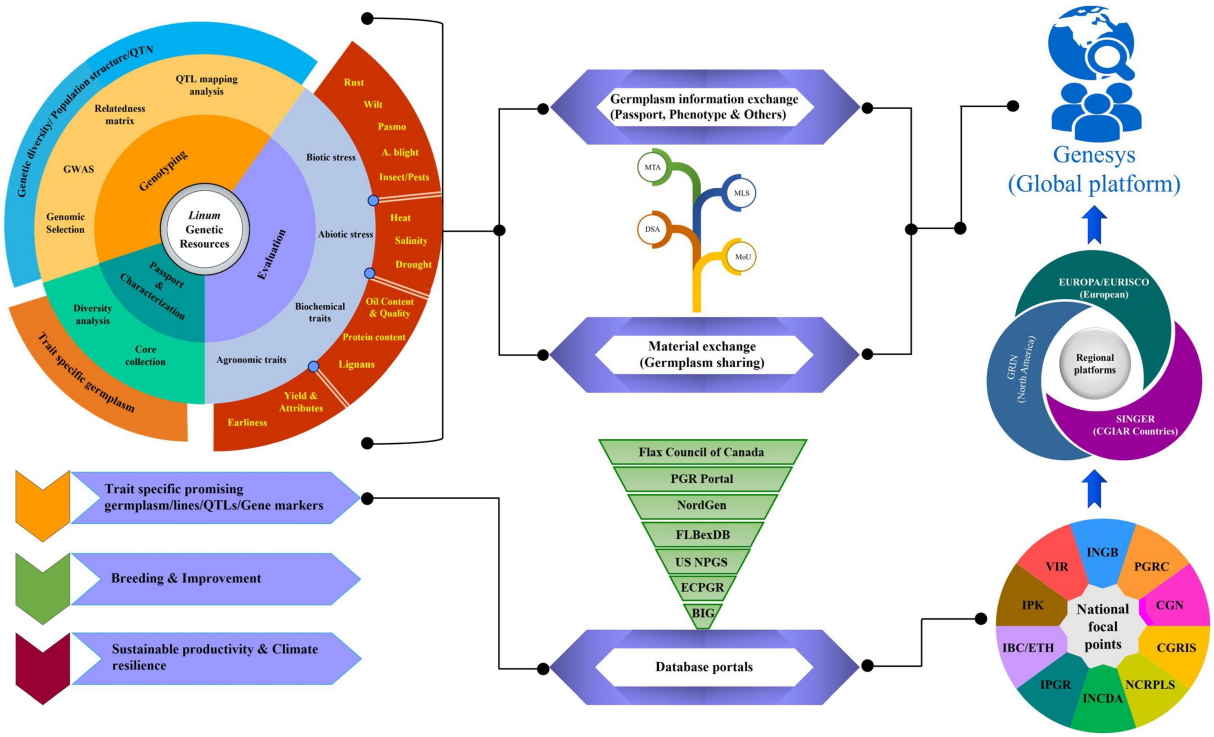


FIGURE 3
A comprehensive overview of the global network for evaluation, conservation, and exchange of *Linum* genetic resources. The requisite information on conserved genetic resources in genebanks at the national, regional, and global levels is accessible at the respective focal point. Some of the genebanks have been transitioned and progressed from individual institute-level limited operations to multi-institutional, regional, as well as global platforms for access to genetic resources and information. CGIAR, Consultative Group on International Agricultural Research; SINGER, System-wide Information Network for Genetic Resources; EURISCO, European Plant Genetic Resources Search Catalogue; GRIN, Germplasm Resources Information Network; ECPGR, European Central Crop Databases; BIG, Germany's Federal Information System on Genetic Resources; US NPGS, US National Plant Germplasm System; NordGen, Nordic Genetic Resources Centre; FIBexDB, Plant Fiber Expression Database; INGB-Indian National Genebank; CGN, Centre for Genetic Resources; the Netherlands; VIR, N.I. Vavilov Research Institute of Plant Industry; INCDA, Research Institute for Cereals and Industrial Crops; IBC/ETH, The Ethiopian Institute of Biodiversity Conservation; IPK, The Leibniz Institute of Plant Genetics and Crop Plant Research; Gatersleben; Germany; NCRPLS, North Central Regional Plant Introduction Station; USA; CGRIS, Institute of Crop Germplasm Resources; China; PGRC, Plant Gene Resources of Canada; Saskatoon; MLS, Multilateral System; SMTA, Standard Material Transfer Agreement; MoU, Memorandum of Understanding; DSA, Data Sharing Agreement.

for the improvement of the crop. Research in linseed germplasm evaluation has been focused on the identification and characterization of diverse germplasm collections for traits, such as oil content, fatty acid composition, seed size, and resistance to biotic and abiotic stresses. This has been accomplished through the use of molecular markers, biochemical analysis, and field evaluations. The present work attempts to review and systematize the existing scientific knowledge about *Linum* genetic resources with respect to the present status of the collection, evaluation, and utilization of major traits of economic importance. The conservation and systematic evaluation of *Linum* genetic diversity is essential to overcome environmental challenges and produce resilient cultivars. Amid the alarming rate of climate change, potential genetic resources from diversity rich and environmentally adapted areas need to be identified for sustainable production. Many studies have reported value-rich and trait-specific promising linseed accession from India, especially high salt and drought tolerant genotypes (68, 142, 147, 148, 179). Therefore, the identification of promising donors for economically important traits and novel molecular tags/genes for important traits in the Indian material will add new information to the domain of knowledge, as of now, only PGRC flax core germplasm has been extensively evaluated for various traits. Moreover, several reports have indicated a narrow genetic base in Canadian linseed cultivars (196, 197, 204), which is an impediment to further breeding progress. In addition, the scarce availability of compatible wild species to incorporate novel variation and the limited molecular breeding have hampered flax yield and quality improvements, thus limiting the competitiveness of the species. Owing to these facts, conventional breeding could result in very low and slow yield gains in linseed. Hence, the utilization of advanced genomic tools to identify key genes and pathways associated with important traits can facilitate the development of molecular breeding approaches for targeted futuristic improvement. Furthermore, the integration of omics technologies in future research and breeding should be emphasized to ensure sustainability in yield and climate resilience. In view of these facts, a mega initiative 'Leveraging genetic resources of linseed through comprehensive phenotyping and genotyping approaches' under the Mission program on 'Minor oilseed of Indian origin' is being taken up presently in India with financial support from the Department of Biotechnology, Govt. of India, wherein whole INGB linseed collection is being phenotyped and genotyped extensively for key agronomic, quality traits, and major biotic and abiotic stresses. The access to evaluation data, genome-wide availability of molecular markers, and QTLs/QTNs/SNPs/candidate genes underlying important traits will ensure further genetic gains in

linseed through markers/genomics-assisted and/or haplotype-based breeding.

Author contributions

VK, AK, SK and KS conceptualized the theme. VK wrote the initial draft and edited the manuscript. MS, DW, KG, SL, BT, and JA helped in the preparation of manuscript in the respective area of expertise. SY and JA helped in the preparation of figures and tables and editing of the manuscript. AK, KS and VK reviewed the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by funding for the project (No. BT/Ag/Network/Linseed/2019–20) from Department of Biotechnology (DBT), Government of India.

Acknowledgments

The authors thank the Director, ICAR-National Bureau of Plant Genetic Resources (NBPGR), New Delhi, India, and the Head, Division of Germplasm Evaluation, ICAR-NBPGR for guidance and 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/fnut.2023.1165580/full#supplementary-material

References

1. FAO (2022) https://www.fao.org/faostat/en/#data/QCL [Accessed 5 June, 2022].

2. Przybylski R. Flax oil and high linolenic oils In: *Bailey's industrial oil and fat products* (2005). 281–301. doi: 10.1002/047167849x.bio010

3. Foulk JA, Fuqua MA, Ulven CA, Alcock MM. Flax fiber quality and influence on interfacial properties of composites. *Int J Sustain Eng*. (2010) 3:17–24. doi: 10.1080/19397030903348710

4. Yan L, Chow N, Jayaraman K. Flax fiber and its composites- a review. *Compos Part B Eng*. (2014) 56:296–317. doi: 10.1016/j.compositesb.2013.08.014

5. Kaur V, Kumar S, Yadav R, Wankhede DP, Aravind J, Radhamani J, et al. Analysis of genetic diversity in Indian and exotic linseed germplasm and identification of trait specific superior accessions. *J Environ Biol*. (2018) 39:702–9. doi: 10.22438/jeb/39/5/MRN-849

6. Zuk M, Richter D, Matuła J, Szopa J. Linseed, the multipurpose plant. *Ind Crop Prod*. (2015) 75:165–77. doi: 10.1016/j.indcrop.2015.05.005

7. Rabetafika HN, Van Remoortel V, Danthine S, Paquot M, Blecker C. Flaxseed proteins: food uses and health benefits. *Int J Food Sci Technol*. (2011) 46:221–8. doi: 10.1111/j.1365-2621.2010.02477.x

8. Rubilar M, Gutiérrez C, Verdugo M, Shene C, Sineiro J. Flaxseed as a source of functional ingredients. *J Plant Nutr Soil Sci*. (2010) 10:373–7. doi: 10.4067/S0718-95162010000100010

9. Westcott NA, Muir AD. Flax seed lignan in disease prevention and health promotion. *Phytochem Rev*. (2003) 2:401–17. doi: 10.1023/B:PHYT.0000046174.97809.b6

10. De Silva SF, Alcorn J. Flaxseed lignans as important dietary polyphenols for cancer prevention and treatment: chemistry, pharmacokinetics, and molecular targets. *Pharmaceuticals*. (2019) 12:68. doi: 10.3390/ph12020068

11. Ganorkar PM, Jain RK. Flaxseed - a nutritional punch. *Int Food Res J*. (2013) 20:519–25.

12. Touré A, Xueming X. Flaxseed lignans: source, biosynthesis, metabolism, antioxidant activity, bio-active components, and health benefits. *Compr Rev Food Sci Food Saf*. (2010) 9:261–9. doi: 10.1111/j.1541-4337.2009.00105.x

13. Adolphe JL, Whiting SJ, Juurlink BH, Thorpe LU, Alcorn J. Health effects with consumption of the flax lignan secoisolariciresinol diglucoside. *Br J Nutr*. (2010) 103:929–38. doi: 10.1017/S0007114509992753

14. Kajla P, Sharma A, Sood DR. Flaxseed-a potential functional food source. *J Food Sci Technol*. (2015) 52:1857–71. doi: 10.1007/s13197-014-1293-y

15. Tarpila A, Wennberg T, Tarpila S. Flaxseed as a functional food. *Curr Top Nutraceutical Res*. (2005) 3:167–88. doi: 10.18231/j.ijnmhs.2020.008

16. Westcott ND, Paton D. Complex containing lignan, phenolic and aliphatic substances from flax process for preparing. *PubChem Patent Summary*. (2001) 6:264–853.

17. GLOBOCAN (2020) Available at: https://gco.iarc.fr/today/home [Accessed 29 July, 2022].

18. Flower G, Fritz H, Balneaves LG, Verma S, Skidmore B, Fernandes R, et al. Flax and breast cancer: a systematic review. *Integr Cancer Ther*. (2014) 13:181–92. doi: 10.1177/1534735413502076

19. Boucher BA, Cotterchio M, Curca IA, Kreiger N, Harris SA, Kirsh VA, et al. Intake of phytoestrogen foods and supplements among women recently diagnosed with breast cancer in Ontario. *Canada Nutr Cancer*. (2012) 64:695–703. doi: 10.1080/01635581.2012.687426

20. Saha D, Mukherjee P, Dutta S, Meena K, Sarkar SK, Mandal AB, et al. Genomic insights into *HSFs* as candidate genes for high-temperature stress adaptation and gene editing with minimal off-target effects in flax. *Sci Rep*. (2019) 9:5581. doi: 10.1038/s41598-019-41936-1

21. Yadav B, Kaur V, Narayan OP, Yadav SK, Kumar A, Wankhede DP. Integrated omics approaches for flax improvement under abiotic and biotic stress: current status and future prospects. *Front Plant Sci*. (2022) 13:931275. doi: 10.3389/fpls.2022.931275

22. Zare S, Mirlohi A, Saeidi G, Sabzalian MR, Ataii E. Water stress intensified the relation of seed color with lignan content and seed yield components in flax (*Linum usitatissimum* L.). *Sci Rep*. (2021) 11:23958. doi: 10.1038/s41598-021-02604-5

23. Saha D, Shaw AK, Datta S, Mitra J. Evolution and functional diversity of abiotic stress-responsive NAC transcription factor genes in *Linum usitatissimum* L. *Environ Exp Bot*. (2021) 188:104512. doi: 10.1016/j.envexpbot.2021.104512

24. Vavilov NI. The origin, variation, immunity and breeding of cultivated plants. *Chron Bot*. (1951) 13:1–366. doi: 10.2134/agronj1952.00021962004400020016x

25. Zeven AC, JMJ D-W. *Dictionary of cultivated plants and their regions of diversity*. Wagenigen: The Netherlands: Centre for Agricultural Publishing and Documentation (1982).

26. Damania A, Valkoun J, Willcox G, Qualset CO. (1998). The origin of agriculture and crop domestication.

27. van Zeist W, Bakker-Heeres JAH. Evidence for linseed cultivation before 6,000 B.C. *J Archaeol Sci*. (1975) 2:215–9. doi: 10.1016/0305-4403(75)90059-X

28. Duk M, Kanapin A, Rozhmina T, Bankin M, Surkova S, Samsonova A, et al. The genetic landscape of fiber flax. *Front Plant Sci*. (2021) 12:764612. doi: 10.3389/fpls.2021.764612

29. Fu YB. Genetic evidence for early flax domestication with capsular dehiscence. *Genet Resour Crop Evol*. (2011) 58:1119–28. doi: 10.1007/s10722-010-9650-9

30. Diederichsen A, Ulrich A. Variability in stem fibre content and its association with other characteristics in 1177 flax (*Linum usitatissimum* L.) genebank accessions. *Ind Crop Prod*. (2009) 30:33–9. doi: 10.1016/j.indcrop.2009.01.002

31. Liu FH, Chen X, Long B, Shuai RY, Long CL. Historical and botanical evidence of distribution, cultivation and utilization of *Linum usitatissimum* L. (flax) in China. *Veget Hist Archaeobot*. (2011) 20:561–6. doi: 10.1007/s00334-011-0311-5

32. Povkhova LV, Melnikova NV, Rozhmina TA, Novakovskiy RO, Pushkova EN, Dvorianinova EM, et al. Genes associated with the flax plant type (oil or fiber) identified based on genome and transcriptome sequencing data. *Plan Theory*. (2021) 10:2616. doi: 10.3390/plants10122616

33. Soto-Cerda BJ, Diederichsen A, Ragupathy R, Cloutier S. Genetic characterization of a core collection of flax (*Linum usitatissimum* L.) suitable for association mapping studies and evidence of divergent selection between fiber and linseed types. *BMC Plant Biol*. (2013) 13:78. doi: 10.1186/1471-2229-13-78

34. Allaby RG, Peterson GW, Andrew DM, Fu YB. Evidence of the domestication history of flax (*Linum usitatissimum*) from genetic diversity of the Sad2 locus. *Theor Appl Genet*. (2005) 112:58–65. doi: 10.1007/s00122-005-0103-3

35. Fu YB, Allaby RG. Phylogenetic network of *Linum* species as revealed by non-coding chloroplast DNA sequences. *Genet Resour Crop Evol*. (2010) 57:667–77. doi: 10.1007/s10722-009-9502-7

36. Guo D, Jiang H, Yan W, Yang L, Ye J, Wang Y, et al. Resequencing 200 flax cultivated accessions identifies candidate genes related to seed size and weight and reveals signatures of artificial selection. *Front Plant Sci*. (2020) 10:1682. doi: 10.3389/fpls.2019.01682

37. Gill KS. *Linseed*. New Delhi, India: Indian Council of Agricultural Research (1987).

38. Seetharam A. Interspecific hybridization in *Linum*. *Euphytica*. (1972) 21:489–95. doi: 10.1007/BF00039344

39. GBIF (2022) Global biodiversity information facility. Available at: http://www.gbif.org/ (Accessed November 13, 2022).

40. Royal botanic gardens Kew and Missouri Botanical Garden (2022) The plant list, a working list of all plant species. Available at: http://www.theplantlist.org/ (Accessed November 13, 2022).

41. Muravenko OV, Lemesh VA, Samatadze TE, Amosova AV, Grushetskaya ZE, Popov KV, et al. Genome comparisons with chromosomal and molecular markers for three closely related flax species and their hybrids. *Russ J Genet*. (2003) 39:414–21. doi: 10.1023/A:1023309831454

42. Vromans J. *Molecular genetic studies in flax (Linum usitatissimum L.)* (2006). 143 p.

43. Diederichsen A, Hammer K. Variation of cultivated flax (*Linum usitatissimum* L. subsp. *usitatissimum*) and its wild progenitor pale flax (subsp. angustifolium (Huds.) Thell.). *Genet Resour Crop Evol*. (1995) 42:263–72. doi: 10.1007/BF02431261

44. Fu YB, Peterson GW. Characterization of expressed sequence tag derived simple sequence repeat markers for 17 *Linum* species. *Botany*. (2010) 88:537–43. doi: 10.1139/B10-019

45. Gill KS, Yermanos DM. Cytogenetic studies on the genus *Linum* I. hybrids among taxa with 15 as the haploid chromosome number. *Crop Sci*. (1967) 7:623–7. doi: 10.2135/cropsci1967.0011183X000700060021x

46. Gill KS, Yermanos DM. Cytogenetic studies on the genus *Linum* II. Hybrids among taxa with nine as the haploid chromosome number. *Crop Sci*. (1967) 7:627–31. doi: 10.2135/cropsci1967.0011183X000700060022x

47. Soto-Cerda BJ, Carrasco RJ, Aravena GA, Urbina HA, Navarro CS. Identifying novel polymorphic microsatellites from cultivated flax (*Linum usitatissimum* L.) following data mining. *Plant Mol Biol Report*. (2011) 29:753–9. doi: 10.1007/s11105-010-0270-5

48. Soto-Cerda BJ, Diederichsen A, Duguid S, Booker H, Rowland G, Cloutier S. The potential of pale flax as a source of useful genetic variation for cultivated flax revealed through molecular diversity and association analyses. *Mol Breed*. (2014) 34:2091–107. doi: 10.1007/s11032-014-0165-5

49. Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S. Genomic regions underlying agronomic traits in linseed (*Linum usitatissimum* L.) as revealed by association mapping. *J Integr Plant Biol*. (2014) 56:75–87. doi: 10.1111/jipb.12118

50. Soto-Cerda BJ, Duguid S, Booker H, Rowland G, Diederichsen A, Cloutier S. Association mapping of seed quality traits using the Canadian flax (*Linum usitatissimum* L.) core collection. *Theor Appl Genet*. (2014) 127:881–96. doi: 10.1007/s00122-014-2264-4

51. Soto-Cerda BJ, Saavedra HU, Navarro NC, Ortega PM. Characterization of novel genic SSR markers in *Linum usitatissimum* (L.) and their transferability across eleven *Linum* species. *Electron J Biotechnol*. (2011) 14:4. doi: 10.2225/vol14-issue2-fulltext-6

52. Tammes T. The genetics of the genus *Linum* In: Lotsy JP, Goddijn WA, editors. *Bibliographia Genetica*. The Netherlands: Martinus Nijhoff (1928). 1–34.

53. Harlan JR, de Wet JMJ. Toward a rational classification of cultivated plants. *Taxon*. (1971) 20:509–17. doi: 10.2307/1218252

54. Gepts P, Papa R. Possible effects of (trans) gene flow from crops on the genetic diversity from landraces and wild relatives. *Environ Biosaf Res*. (2003) 2:89–103. doi: 10.1051/ebr:2003009

55. Diederichsen A. *Ex situ* collections of cultivated flax (*Linum usitatissimum* L.) and other species of the genus Linum L. *Genet Resour Crop Evol*. (2007) 54:661–78. doi: 10.1007/s10722-006-9119-z

56. Diederichsen A. A taxonomic view on genetic resources in the genus *Linum* L. for flax breeding In: Cullis CA, editor. *Plant genetics and genomics: Crops and models*. Springer (2019). 23:1–15

57. Jhala AJ, Hall LM, Hall JC. Potential hybridization of flax with weedy and wild relatives: An avenue for movement of engineered genes? *Crop Sci*. (2008) 48:825–40. doi: 10.2135/cropsci2007.09.0497

58. Fu YB. A molecular view of flax gene pool In: Cullis C, editor. *Plant genetics and genomics: Crops and models*. Springer (2019). 16–36.

59. Bari G, Godward MBE. Interspecific crosses in *Linum*. *Euphytica*. (1970) 19:443–6. doi: 10.1007/BF01902918

60. Hegi G. *Illustrierte Flora von Mitteleuropa, illustrated flora of Central Europe*. Munchen: Lehmanns Verlag (1925). 5 p 3-38.

61. Kumar A, Ekka S, Tripathi UK. Evaluation of linseed germplasm for resistance against rust (*Melampsora lini*). *J Oilseeds Res*. (2016) 33:56–61.

62. Tork DG, Anderson NO, Wyse DL, Betts KJ. Perennial flax: a potential cut flower crop. *HortScience*. (2022) 57:221–30. doi: 10.21273/HORTSCI16098-21

63. Maggioni L, Pavelek M, van Soest LJM, Lipman E. (2002). "Flax genetic resources in Europe" in Ad hoc meeting: Prague, Czech Republic. International Plant Genetic Resources Institute, Rome, Italy.

64. Brutch NB. The flax genetic resources collection held at the N.I. Vavilov institute, Russian Federation In: *Flax genetic resources*. IPGRI, Maccarese Rome: Europe (2002). 61–5.

65. Zhuchenko AA, Rozhmina TA. Mobilization of flax genetic resources In: *VILAR and VNIIL*. Russia: Starica (2000)

66. Kaur V, Yadav R, Wankhede DP. Linseed genetic resources for climate change intervention and future breeding. *J Appl Nat Sci*. (2017) 9:1112–8. doi: 10.31018/jans.v9i2.1331

67. FAO. *The second report on the state of the world's plant genetic resources for food and agriculture*. Rome (2010).

68. Diederichsen A, Fu Y. Phenotypic and molecular (RAPD) differentiation of four infraspecific groups of cultivated flax (*Linum usitatissimum* L. subsp. *usitatissimum*). *Genet Resour Crop Evol*. (2006) 53:77–90. doi: 10.1007/s10722-004-0579-8

69. Diederichsen A, Raney JP. Seed colour, seed weight and seed oil content in *Linum usitatissimum* accessions held by plant gene resources of Canada. *Plant Breed*. (2006) 125:372–7. doi: 10.1111/j.1439-0523.2006.01231.x

70. Diederichsen A, Kusters PM, Kessler D, Bainas Z, Gugel RK. Assembling a core collection from the flax world collection maintained by plant gene resources of Canada. *Genet Resour Crop Evol*. (2013) 60:1479–85. doi: 10.1007/s10722-012-9936-1

71. Diederichsen A, Raney JP, Fu YB, Richards KW. (2002). "Diversity in the flax collection at plant gene resources of Canada" in *Proceedings of the 59th Flax Institute of the United States. Doublewood Inn, Fargo, North Dakota, USA. Flax Institute, Fargo, USA*. pp. 138–143.

72. You FM, Jia G, Xiao J, Duguid SD, Rashid KY, Booker HM, et al. Genetic variability of 27 traits in a core collection of flax (*Linum usitatissimum* L.). front. *Plant Sci*. (2017) 8:1636. doi: 10.3389/fpls.2017.01636

73. Worku N, Heslop-Harrison JS, Adugna W. Diversity in 198 Ethiopian linseed (*Linum usitatissimum*) accessions based on morphological characterization and seed oil characteristics. *Genet Resour Crop Evol*. (2015) 62:1037–53. doi: 10.1007/s10722-014-0207-1

74. Chandrawati SN, Kumar R, Kumar S, Singh PK, Yadav VK, et al. Genetic diversity, population structure and association analysis in linseed (*Linum usitatissimum* L.). *Physiol Mol Biol Plants*. (2017) 23:207–19. doi: 10.1007/s12298-016-0408-5

75. Dhirhi N, Mehta N. Estimation of genetic variability and correlation in F_2 segregating generation in linseed (*Linum usitatisimum* L.). *Plant Arch*. (2019) 19:475–84.

76. Dikshit N, Sivaraj N. Analysis of agro-morphological diversity and oil content in Indian linseed germplasm. *Grasas Aceites*. (2015) 66:e060. doi: 10.3989/gya.0689141

77. Nizar MA, Mulani RM. Genetic diversity in indigenous and exotic linseed germplasm (*Linum usitatissimum* L.). Electron. *J Plant Breed*. (2015) 6:848–54.

78. Saroha A, Pal D, Gomashe SS, Akash KV, Ujjainwal S, et al. Identification of QTNs associated with flowering time, maturity, and plant height traits in *Linum usitatissimum* L. using genome-wide association study. *Front Genet*. (2022) 13:811924. doi: 10.3389/fgene.2022.811924

79. Saroha A, Pal D, Kaur V, Kumar S, Bartwal A, Aravind J, et al. Agro-morphological variability and genetic diversity in linseed (*Linum usitatissimum* L.) germplasm

accessions with emphasis on flowering and maturity time. *Genet Resour Crop Evol*. (2022) 69:315–33. doi: 10.1007/s10722-021-01231-3

80. Uysal H, Kurt O, Fu YB, Diederichsen A, Kusters P. Variation in phenotypic characters of pale flax (*Linum bienne* mill.) from Turkey. *Genet Resour Crop Evol*. (2012) 59:19–30. doi: 10.1007/s10722-011-9663-z

81. Green AG, Marshall DR. Variation for oil quantity and quality in linseed (*Linum usitatissimum*). *Aust J Agric Res*. (1981) 32:599–607. doi: 10.1071/AR9810599

82. Zimmerman DC, Klosterman HJ. The distribution of fatty acids in linseed oil from the world collection of flax varieties. *Proc North Dakota Acad Sci*. (1959) 13:71–5.

83. Diederichsen A, Raney JP, Duguid SD. Variation of mucilage in flax seed and its relationship with other seed characters. *Crop Sci*. (2006) 46:365–71. doi: 10.2135/cropsci2005.0146

84. Kumar N, Tripathi UK. Screening of linseed germplasm for resistance against *Alternaria* spp. cause blight disease in linseed (*Linum usitatissimum* L.). *J Pharmacogn Phytochem*. (2018) 7:1389–92.

85. Sivaraj N, Sunil N, Pandravada SR, Kamala V, Babu A, Vinod K, et al. Variability in linseed (*Linum usitatissimum*) germplasm collections from peninsular India with special reference to seed traits and fatty acid composition. *Indian J Agric Sci*. (2012) 82:102–5.

86. Bayrak A, Kiralan M, Ipek A, Arslan N, Cosge B, Khawar KM. Fatty acid compositions of linseed (*Linum usitatissimum* L.) genotypes of different origin cultivated in Turkey. *Biotechnol Biotechnol Equip*. (2010) 2:1836–42. doi: 10.2478/V10133-010-0034-2

87. Silska G. Genetic resources of flax (*Linum usitatissimum* L.) as very rich source of α-linolenic acid. *Herba Pol*. (2017) 63:26–33. doi: 10.1515/hepo-2017-0022

88. Pali V, Mehta N. Evaluation of oil content and fatty acid compositions of flax (*Linum usitatissimum* L.) varieties of India. *J Agric Sci*. (2014) 6:198–207. doi: 10.5539/jas.v6n9p198

89. Oomah BD, Kenaschuk EO, Cui W, Mazza G. Variation in the composition of water-soluble polysaccharides in flaxseed. *J Agric Food Chem*. (1995) 43:1484–8. doi: 10.1021/jf00054a013

90. Madhusudhan B. Potential benefits of flaxseed in health and disease - a perspective. *Agric Conspec Sci*. (2009) 74:67–72.

91. Morris DH. *Flax-a health and nutrition primer*. Winnipeg: Flax council of Canada (2007).

92. Eliasson C, Kamal-Eldin A, Andersson R, Aman P. High-performance liquid chromatographic analysis of secoisolariciresinol diglucoside and hydroxycinnamic acid glucosides in flaxseed by alkaline extraction. *J Chromatogr A*. (2003) 1012:151–9. doi: 10.1016/s0021-9673(03)01136-1

93. Kaewmanee T, Bagnasco L, Benjakul S, Lanteri S, Morelli CF, Speranza G, et al. Characterisation of mucilages extracted from seven Italian cultivars of flax. *Food Chem*. (2014) 148:60–9. doi: 10.1016/j.foodchem.2013.10.022

94. Kumar M, Patel M, Chauhan R, Tank C, Solanki S. Delineating multivariate divergence, heritability, trait association and identification of superior omega-3-fatty acid specific genotypes in linseed (*Linum usitatissimum* L.). *Genetika*. (2021) 53:825–36. doi: 10.2298/GENSR2102825K

95. Hussain ME, Goyal VK, Paul PJ, Yadav Y, Jha UC, Moitra PK. Assessment of genetic variability, diversity, and identification of promising lines in linseed germplasm for harnessing genetic gain in central plain of the Indian subcontinent. *J Plant Breed Crop Sci*. (2022) 14:12–20. doi: 10.5897/JPBCS2021.0990

96. Chandrawati SN, Yadav VK, Kumar R, Kumar S, Yadav HK. Genetic variability and interrelationship among morphological and yield traits in linseed (*Linum usitatissimum* L.). *Genetika*. (2016) 48:881–92. doi: 10.2298/GENSR1603881C

97. Diederichsen A, Rozhmina TA, Zhuchenko AA, Richards KW. Screening for broad adaptation in 96 flax (*Linum usitatissimum* L.) accessions under dry and warm conditions in Canada and Russia. *PGR Newsl*. (2006) 146:9–16.

98. Dabalo DY, Singh BCS, Weyessa B. Genetic variability and association of characters in linseed (*Linum usitatissimum* L.) plant grown in Central Ethiopia region. Saudi. *J Biol Sci*. (2020) 27:2192–206. doi: 10.1016/j.sjbs.2020.06.043

99. Adugna W, Labuschagne MT, Viljoen CD. The use of morphological and AFLP markers in diversity analysis of linseed. *Biodivers Conserv*. (2006) 15:3193–205. doi: 10.1007/s10531-005-6970-8

100. Bakry AB, Ibrahim OM, Abd El-Fattah Elewa T, El-Karamany MF. Performance assessment of some flax (*Linum usitatissimum* L.) varieties using cluster analysis under sandy soil conditions. *Agric Sci*. (2014) 5:677–86. doi: 10.4236/as.2014.58071

101. Ali MF, Mekbib F, Wakjira A. Morphological diversity of Ethiopian linseed (*Linum usitatissimum* L.) landrace accessions and non-native cultivars. *J Plant Breed Genet*. (2014) 2:115–24.

102. Tadesse T, Parven A, Singh H, Weyessa B. Estimates of variability and heritability in linseed germplasm. *Int J Sustain Crop Prod*. (2010) 5:8–16.

103. Siddiqui A, Shukla SK, Rastogi A, Bhargava A, Niranjan A, Lehri A. Relationship among phenotypic and quality traits in indigenous and exotic accessions of linseed. *Pesq Agropec Bras*. (2016) 51:1964–72. doi: 10.1590/s0100-204x2016001200007

104. El-Hariri DM, Al-Kordy MA, Hassanein MS, Ahmed MA. Partition of photosynthates and energy production in different flax cultivars. *J Nat Fibers*. (2005) 1:1–15. doi: 10.1300/J395v01n04_01

105. Kumar S, You FM, Duguid S, Booker H, Rowland G, Cloutier S. QTL for fatty acid composition and yield in linseed (*Linum usitatissimum* L.). *Theor Appl Genet*. (2015) 128:965–84. doi: 10.1007/s00122-015-2483-3

106. Ottai MES, Al-Kordy MAA, Hussein RM, Hassanein MS. Genetic diversity among Romanian fiber flax varieties under Egyptian conditions. *Aust J Basic Appl Sci*. (2012) 6:162–8.

107. Biradar SA, Ajithkumar K, Rajanna B, Savitha AS, Shubha GV, Shankergoud I, et al. Prospects and challenges in linseed (*Linum usitatissimum* L.) production: a review. *J Oilseeds Res*. (2016) 33:1–13.

108. Thambugala D, Duguid S, Loewen E, Rowland G, Booker H, You FM, et al. Genetic variation of six desaturase genes in flax and their impact on fatty acid composition. *Theor Appl Genet*. (2013) 126:2627–41. doi: 10.1007/s00122-013-2161-2

109. Friedt W, Bickert C, Schaub H. *In vitro* breeding of high linolenic, doubled haploid lines of linseed (*Linum usitatissimum* L.) *via* androgenesis. *Plant Breed*. (1995) 114:322–6. doi: 10.1111/J.1439-0523.1995.TB01242.X

110. Cloutier S, Ragupathy R, Miranda E, Radovanovic N, Reimer E, Walichnowski A, et al. Integrated consensus genetic and physical maps of flax (*Linum usitatissimum* L.). *Theor Appl Genet*. (2012) 125:1783–95. doi: 10.1007/s00122-012-1953-0

111. Green AG. Genetic control of polyunsaturated fatty acid biosynthesis in flax (*Linum usitatissimum*) seed oil. *Theor Appl Genet*. (1986) 72:654–61. doi: 10.1007/BF00289004

112. Rowland GG. An EMS-induced low-linolenic-acid mutant in McGregor flax (*Linum usitatissimum* L.). *Can J Plant Sci*. (1991) 71:393–6. doi: 10.4141/cjps91-054

113. Green AG, Marshall AR. Isolation of induced mutants in linseed (*Linum usitatissimum* L.) having reduced linolenic acid content. *Euphytica*. (1984) 33:321–8. doi: 10.1007/BF00021128

114. Fikere M, Mekbib F, Wakjira A. Seed oil diversity of Ethiopian linseed (*Linum usitatissimum* L.) landraces accessions and some exotic cultivars. *Afr J Biochem Res*. (2013) 7:76–85. doi: 10.5897/AJBR.9000214

115. Dmitriev AA, Krasnov GS, Rozhmina TA, Novakovskiy RO, Snezhkina AV, Fedorova MS, et al. Differential gene expression in response to *fusarium oxysporum* infection in resistant and susceptible genotypes of flax (*Linum usitatissimum* L.). *BMC Plant Biol*. (2017) 17:253. doi: 10.1186/s12870-017-1192-2

116. Singh J, Singh PK, Srivastava RL. Diseases of linseed (*Linum usitatissium* L.) in India and their management a review. *J Oilseeds Res*. (2017) 34:52–69.

117. Diederichsen A, Rozhmina T, Kudrjavceva L. Variation patterns within 153 flax (*Linum usitatissimum* L.) genebank accessions based on evaluation for resistance to fusarium wilt, anthracnose and pasmo. *Plant Genet Resour*. (2008) 6:22–32. doi: 10.1017/S1479262108913897

118. Rozhmina TA. *Identification of genes for resistance to fusarium wilt in flax, breeding seed production agrotechnology, economics and primary processing of flax*. (2002) *30*:48–52. Torzhok, Russia

119. Rozhmina TA, Loshakova NI. Samples of spinning and oil flax (*Linum usitatissimum* L.). sources of effective genes for resistance to *Fusarium* wilt and its dependence on temperature. *Agric Biol*. (2016) 51:310–7.

120. Evans N, Mcroberts N, Hitchcock D, Marshall G. Assessing linseed (*Linum usitatissimum*) resistance to *Alternaria linicola* using a detached cotyledon assay. *Ann Appl Biol*. (1997) 130:537–47. doi: 10.1111/j.1744-7348.1995.tb06671.x

121. Dash J, Naik BS, Mohapatra UB. Field screening of linseed genotypes for resistance to powdery mildew (*Oidium lini* Skoric) in the north central plateau zone of Odisha. *Int J Adv Res*. (2016) 4:961–2. doi: 10.21474/IJAR01/204

122. Naik BS, Sahoo KC. Survey of diseases of linseed in Mayurbhanj district of Odisha. *Int J Adv Res*. (2016) 4:1407–10. doi: 10.21474/IJAR01/235

123. Pandey MK, Singh R, Singh A, Mishra BK. Influence of dates of sowing, disease incidence and crop yield against *Alternaria* blight of linseed. *J Pharmacogn Phytochem*. (2019) 8:1459–61.

124. Singh RB, Singh AK, Srivastava RK. Assessment of yield losses due to *Alternaria* blight of linseed. *J Oilseeds Res*. (2003) 20:168–9.

125. Pandey RN, Mishra DP. Assessment of yield loss due to powdery mildew of linseed. *Indian Bot Rep*. (1992) 11:62–4.

126. Beale RE. Studies of resistance in linseed cultivars to *Oidium lini* and *Botrytis cinerea*. In: Fraud Williams RJ, editor. *Production and protection of linseed. Aspects of applied biology series, association of applied biologists*. Wellesbourne, Warwick, UK: Horticulture Research International (1991). 85–90.

127. Rashid K, Duguid S. Inheritance of resistance to powdery mildew in flax. *Can J Plant Pathol*. (2005) 27:404–9. doi: 10.1080/07060660509507239

128. Dhirhi N, Mehta N, Singh S. Screening of powdery mildew tolerance in linseed (*Linum usitatisimum* L.). *J Plant Dev Sci*. (2017) 9:153–6.

129. Dutt S, Chahal AS, Lobana KS, Singh M. Evaluation of genetic stock of linseed (*Linum usitatissimum* L.) against powdery mildew in Punjab. *J Res*. (1975) 12:6–8.

130. Kamthan KP, Shukla AK, Misra DP. Independent genetic resistance to wilt, rust and powdery mildew in linseed I. *J Agric Biol Sci*. (1991) 7:56–60.

131. Kishore R, Pandey M, Tripathi UK, Singh J. Evaluation of elite genotypes of linseed against *Fusarium* wilt. *Indian Phytopathol*. (2011) 64:209.

132. Singh RB, Singh RN. Date of sowing and varieties for root-rot and wilt complex in linseed (*Linum usitatissimum* L.). *Indian J Agric Sci*. (2011) 81:287–9.

133. Kanapin A, Bankin M, Rozhmina T, Samsonova A, Samsonova M. Genomic regions associated with *Fusarium* wilt resistance in flax. *Int J Mol Sci*. (2021) 22:12383. doi: 10.3390/ijms222212383

134. Ajithkumar K, Biradar SA, Rajanna B. Screening of linseed germplasms for resistance against powdery mildew caused by Oidium lini Skoric. *J Mycopathol Res*. (2015) 53:247–51.

135. Ramakant CMP, Singh RB. Screening of linseed genotypes against *Alternaria* blight and bud fly. *Crop Res*. (2008) 35:124–7.

136. Singh RB, Singh HK, Parmar A. Yield loss assessment due to *Alternaria* blight and its management in linseed. *Pak J Biol Sci*. (2014) 17:511–6. doi: 10.3923/pjbs.2014.511.516

137. Sinha JN, Singh AP, Kumar P. Reaction of linseed germplasm to rust (*Melampsora lini*). *J Appl Biol*. (1993) 3:80–2.

138. Saharan GS. Linseed and flax rust. *Indian J Mycol Pl Pathol*. (1991) 21:119–37.

139. Pal R, Malik YP. Histological basis of resistance in linseed against bud fly (*Dasyneura lini* Barnes) in Central Uttar Pradesh. *Int J Agri Invent*. (2020) 5:25–30. doi: 10.46492/IJAI/2020.5.1.3

140. Ekka RK, Mandal S, Meena RS, Padamshali S. Some physical basis of host plant resistance in linseed (*Linum usitatissimum* L.), against bud fly, *Dasyneuralini* (Barnes). *J Pharmacogn Phytochem*. (2018) 7:2010–5.

141. Annual Report Linseed. (2019-20) ICAR-Indian Institute of pulses research, Kanpur-208 024.

142. Sertse D, You FM, Ravichandran S, Soto-Cerda BJ, Duguid S, Cloutier S. Loci harboring genes with important role in drought and related abiotic stress responses in flax revealed by multiple GWAS models. *Theor Appl Genet*. (2021) 134:191–212. doi: 10.1007/s00122-020-03691-0

143. Sertse D, You FM, Ravichandran S, Cloutier S. The complex genetic architecture of early root and shoot traits in flax revealed by genome-wide association analyses. *Front Plant Sci*. (2019) 10:1483. doi: 10.3389/fpls.2019.01483

144. Soto-Cerda BJ, Cloutier S, Gajardo HA, Aravena G, Quian R. Identifying drought-resilient flax genotypes and related-candidate genes based on stress indices, root traits and selective sweep. *Euphytica*. (2019) 215:41–16. doi: 10.1007/s10681-019-2362-0

145. Soto-Cerda BJ, Cloutier S, Gajardo HA, Aravena G, Quian R, You FM. Drought response of flax accessions and identification of quantitative trait nucleotides (QTNs) governing agronomic and root traits by genome-wide association analysis. *Mol Breed*. (2020) 40:1–24. doi: 10.1007/s11032-019-1096-y

146. Asgarinia P, Mirlohi A, Saeidi G, Mirik MAA, Gheysari M, Razavi VS. Selection criteria for assessing drought tolerance in a segregating population of flax (*Linum usitatissimum* L.). *Can J Plant Sci*. (2017) 97:424–37. doi: 10.1139/cjps-2016-0140

147. Soto-Cerda BJ, Larama G, Gajardo H, Inostroza-Blancheteau C, Cloutier S, Fofana B, et al. Integrating multi-locus genome-wide association studies with transcriptomic data to identify genetic loci underlying adult root traits responses to drought stress in flax (*Linum usitatissimum* L.). *Environ Exp Bot*. (2022) 202:105019. doi: 10.1016/j.envexpbot.2022.105019

148. Li X, Guo D, Xue M, Li G, Yan Q, Jiang H. Genome-wide association study of salt tolerance at the seed germination stage in flax (*Linum usitatissimum* L.). *Genes*. (2022) 13:486. doi: 10.3390/genes13030486

149. Patil N, Datir S, Shah P. Salt-induced physiological and biochemical changes in two varieties of *Linum usitatissimum* L. *Int J Curr Microbiol App Sci*. (2015) 4:296–304.

150. Kaya MD, Day S, Cikili Y, Arslan N. Classification of some linseed (*Linum usitatissimum* L.) genotypes for salinity tolerance using germination, seedling growth, and ion content. *Chil J Agric Res*. (2012) 72:27–32. doi: 10.4067/s0718-58392012000100005

151. El-Beltagi HS, Salama ZA, El Hariri DM. Some biochemical markers for evaluation of flax cultivars under salt stress conditions. *J Nat Fibers*. (2008) 5:316–30. doi: 10.1080/15440470802252487

152. Basandrai D, Basandrai AK. Genetics of field resistance to powdery mildew in flax. *Indian Phytopathol*. (2000) 53:224–6.

153. Singh BM, Saharan GS. Inheritance of resistance to *Oidium lini* Skoric in linseed (*Linum usitatissimum* L.). *Euphytica*. (1979) 28:531–2. doi: 10.1007/BF00056614

154. Asgarinia P, Cloutier S, Duguid S, Rashid K, Mirlohi A, Banik M, et al. Mapping quantitative trait loci for powdery mildew resistance in flax (*Linum usitatissimum* L.). *Crop Sci*. (2013) 53:2462–72. doi: 10.2135/cropsci2013.05.0298

155. Speck A, Trouvé JP, Enjalbert J, Geffroy V, Joets J, Moreau L. Genetic architecture of powdery mildew resistance revealed by a genome-wide association study of a worldwide collection of flax (*Linum usitatissimum* L.). *Front Plant Sci*. (2022) 13:871633. doi: 10.3389/fpls.2022.871633

156. Goray SC, Khosla HK, Upadhyaya YM, Naik SL, Mandoli SL. Inheritance of wilt resistance in linseed. *Indian J Agric Sci*. (1987) 57:625–7. doi: 10.1093/oxfordjournals.jhered.a111281

157. Misra DP, Pandey RN. Linkage of resgenes for powdery-mildew and rust resistance in linseed. *Indian J Agric Sci*. (1981) 51:559–62.

158. Misra DP, Shukla AK. Genetic identity for rust resistance in some linseeds. *Indian J Agric Sci*. (1981) 27:444–52. doi: 10.1177/0019556119810209

159. Saharan GS, Singh BM. Inheritance of resistance in linseed to rust caused by *Melampsora lini* under field conditions. *Indian J Mycol Pl Pathol*. (1980) 10:136–41.

160. Misra DP. Rust resistance in the species and varieties of *Linum* against Indian physiological races of linseed rust. *J Oilseeds Res*. (1964) 8:277–9.

161. Biswas GC, Das GP. Insect and mite pests diversity in the oilseed crops ecosystems in Bangladesh. *Bangladesh J Zool*. (2011) 39:235–44. doi: 10.3329/bjz.v39i2.10594

162. Pruthi HS, Bhatia HL. A new cecidomyiidae pest of linseed in India. *Indian J Agric Sci*. (1937) 7:797–808.

163. Malik YP. Estimation of economic threshold level for bud fly, *Dasyneura lini* Barnes in linseed, *Linum usitatissimum* L. *J Oilseeds Res*. (2005) 22:100–2.

164. Malik YP. Yield losses due to bud fly, *Dasyneura lini* Barnes in linseed. *J Oilseeds Res*. (2006) 23:363.

165. Malik YP, Husain K, Singh SV, Srivastava RL. Development of management module for bud fly, *Dasyneura lini* in linseed. *Indian J Entomol*. (2000) 62:260–9.

166. Gupta MP. Field screening of linseed, *Linum usitatissimum* Linn. Genotypes against the incidence of budfly, *Dasyneura lini* Barnes and blight, *Alternaria lini* Dey. *J Oilseeds Res*. (2004) 21:108–11.

167. Pal R, Malik YP. Screening of linseed germplasm against bud fly *Dasyneura lini* (Barnes) under field condition. *Int J Pure Appl Biosci*. (2018) 6:196–201. doi: 10.18782/2320-7051.6542

168. Pal RK, Singh R. Screening of linseed germplasms against bud fly, *Dasyneura lini* Barnes. *Int J Plant Prot*. (2010) 3:410–1.

169. Reddy N, Maheshwari JJ, Patil PV, Ghorapade PB, Gohokar RT. Simplified triple test cross analysis for bud fly resistance and yield components in linseed, *Linum usitatissimum* L. *J Oilseed*. (2009) 26:166–8.

170. Gupta AK. Biophysical basis of resistance in linseed against bud fly, *Dasyneura lini* (Barnes). *Indian J Entomol*. (2015) 77:345–52. doi: 10.5958/0974-8172.2015.00073.5

171. Sood NK, Pathak SC. Thickness of sepal: a factor for resistance in linseed against *Dasyneura lini* Barnes. *Indian J Plant Prot*. (1990) 52:28–30.

172. Heller K, Konczewicz W, Byczyńska M, Łukaszewska N, Praczyk M. (2008). "The effect of fiber flax growing technologies on plants ontogenesis and cultivars yielding capacity" in *Conference on flax and other bast plants. Saskatoon: Saskatchewan, Canada*. pp. 315–325.

173. Heller K, Byczyńska M. The impact of environmental factors and applied agronomy on quantitative and qualitative traits of flax fiber. *J Nat Fibers*. (2015) 12:26–38. doi: 10.1080/15440478.2013.879088

174. Hall LM, Booker H, Siloto RMP, Jhala AJ, Weselake RJ. Flax (*Linum usitatissimum* L.) In: McKeon T, Hayes D, Hildebrand D, Weselake R, editors. *J.Industrial oil crops*. AOCS Press, New York: Elsevier (2016). 157–94.

175. Kar G, Kumar A, Martha M. Water use efficiency and crop coefficients of dry season oilseed crops. *Agric Water Manag*. (2007) 87:73–82. doi: 10.1016/j.agwat.2006.06.002

176. Dash P, Cao Y, Jailani A, Gupta P, Venglat P, Xiang D, et al. Genome-wide analysis of drought induced gene expression changes in flax (*Linum usitatissimum*). *GM Crops Food*. (2014) 5:106–19. doi: 10.4161/gmcr.29742

177. Foster R, Pooni H, Mackay I. The impact of water deprivation on the performance of *Linum usitatissimum* cultivars. *J Genet Breed*. (1998) 52:63–71.

178. Lisson S, Mendham N. Agronomic studies of flax (*Linum usitatissimum* L.) in South-Eastern Australia. *Aust J Exp Agric*. (2000) 40:1101–12. doi: 10.1071/EA00059

179. Diederichsen A, Raney JP. Pure-lining of flax (*Linum usitatissimum* L.) genebank accessions for efficiently exploiting and assessing seed character diversity. *Euphytica*. (2008) 164:255–73. doi: 10.1007/s10681-008-9725-2

180. EL Afry MM, EL OSA, EL KESA, EL YGSA. Exogenous application of ascorbic acid for alleviation the adverse effects of salinity stress in flax (*Linum usitatissimum* L.). *Middle East J*. (2018) 7:716–39.

181. Guo R, Zhou J, Hao W, Gu F, Liu Q, Li H, et al. Germination, growth, chlorophyll fluorescence and ionic balance in linseed seedlings subjected to saline and alkaline stresses. *Plant Prod Sci*. (2014) 17:20–31. doi: 10.1626/pps.17.20

182. Nasri N, Maatallah S, Saidi I, Lachâal M. Influence of salinity on germination, seedling growth, ion content and acid phosphatase activities of *Linum usitatissimum* L. *J Anim Plant Sci*. (2017) 27:517–21.

183. Qayyum MA, Bashir F, Maqbool MM, Ali A, Bashir S, Abbas Q. Implications of saline water irrigation for linseed on seed germination, seedling survival and growth potential. *Sarhad J Agric*. (2019) 35:1289–97. doi: 10.17582/journal.sja/2019/35.4.1289.1297

184. Yu Y, Wu G, Yuan H, Cheng L, Zhao D, Huang W, et al. Identification and characterization of miRNAs and targets in flax (*Linum usitatissimum*) under saline, alkaline, and saline-alkaline stresses. *BMC Plant Biol*. (2016) 16:124. doi: 10.1186/s12870-016-0808-2

185. Wu J, Zhao Q, Wu G, Yuan H, Ma Y, Lin H, et al. Comprehensive analysis of differentially expressed unigenes under NaCl stress in flax (*Linum usitatissimum* L.) using RNA-Seq. *Int J Mol Sci*. (2019) 20:369. doi: 10.3390/ijms20020369

186. Kraft JM, Kommedahl T, Linck AL. Histological study of malformation in flaxseed after exposure to 31°C. *Bot Gaz*. (1963) 125:367–71. doi: 10.1086/336221

187. Gusta LV, O'Conner BJ, Bhatty RS. Flax (*Linum usitatissimum* L.) response to chilling and heat stress on flowering and seed yield. *Can J Plant Sci*. (1997) 77:97–9. doi: 10.4141/P95-205

188. Cross RH, Mckay SAB, McHughen AG, Bonham-Smith PC. Heat-stress effects on reproduction and seed set in *Linum usitatissimum* L. (flax). *Plant Cell Environ*. (2003) 26:1013–20. doi: 10.1046/j.1365-3040.2003.01006.x

189. Fofana B, Cloutier S, Duguid S, Ching J, Rampitsch C. Gene expression of stearoyl-ACP desaturase and D_{12} fatty acid desaturase $_2$ is modulated during seed development of flax (*Linum usitatissimum*). *Lipids*. (2006) 41:705–20. doi: 10.1007/s11745-006-5021-x

190. Saha D, Shaw AK, Datta S, Mitra J, Kar G. DNA hypomethylation is the plausible driver of heat stress adaptation in *Linum usitatissimum*. *Physiol Plant*. (2022) 174:e13689. doi: 10.1111/ppl.13689

191. Aldrich J, Cullis CA. RAPD analysis in flax: optimization of yield and reproducibility using klenTaq1 DNA polymerase, Chelex 100, and gel purification of genomic DNA. *Plant Mol Biol Report*. (1993) 11:128–14. doi: 10.1007/BF02670471

192. Cullis CA, Oh TJ, Gorman MB. (1995), "Genetic mapping in flax (Linum usitatissimum)" in Proc. 3rd Meeting Int. Flax Breed Res Group. St. Vale´ ryencaux. pp. 161–169.

193. Cullis CA, Swami S, Song Y. RAPD polymorphisms detected between flax genotrophs. *Plant Mol Biol*. (1999) 41:795–800. doi: 10.1023/a:1006385606163

194. Everaert I, Riek JD, Loose MD, Waes JV, Bockstaele EV. Most similar variety grouping for distinctness evaluation of flax and linseed (*Linum usitatissimum* L.) varieties by means of AFLP and morphological data. *Plant Var Seeds*. (2001) 14:69–87.

195. Fu Y, Guerin S, Peterson GW, Diederichsen A, Rowland GG, Richards KW. RAPD analysis of genetic variability of regenerated seeds in the Canadian flax cultivar CDC normandy. *Seed Sci Technol*. (2003) 31:207–11. doi: 10.15258/sst.2003.31.1.22

196. Fu YB, Diederichsen A, Richards KW, Peterson G. Genetic diversity within a range of cultivars and landraces of flax (*Linum usitatissimum* L) as revealed by RAPDs. *Genet Resour Crop Evol*. (2002) 49:167–74. doi: 10.1023/A:1014716031095

197. Fu YB, Rowland GG, Duguid SD, Richards K. RAPD analysis of 54 north American flax cultivars. *Crop Sci*. (2003) 43:1510–5. doi: 10.2135/cropsci2003.1510

198. Kumari A, Paul S, Sharma V. Genetic diversity analysis using RAPD and ISSR markers revealed discrete genetic makeup in relation to fibre and oil content in *Linum usitatissimum* L. genotypes. *Nucleus*. (2018) 61:45–53. doi: 10.1007/s13237-017-0206-7

199. Rajwade AV, Arora RS, Kadoo NY, Harsulkar AM, Ghorpade PB, Gupta VS. Relatedness of Indian flax genotypes (*Linum usitatissimum* L.): an inter-simple sequence repeat (ISSR) primer assay. *Mol Biotechnol*. (2010) 45:161–70. doi: 10.1007/s12033-010-9256-7

200. Uysal H, Fu YB, Kurt O, Peterson GW, Diederichsen A, Kusters P. Genetic diversity of cultivated flax (*Linum usitatissimum* L.) and its wild progenitor pale flax (*Linum bienne* mill.) as revealed by ISSR markers. *Genet Resour Crop Evol*. (2010) 57:1109–19. doi: 10.1007/s10722-010-9551-y

201. Wiesnerova D, Wiesner I. ISSR-based clustering of cultivated flax germplasm is statistically correlated to thousand seed mass. *Mol Biotechnol*. (2004) 26:207–14. doi: 10.1385/MB:26:3:207

202. Smykal P, Kerteszova NB, Kalendar R, Corander J, Schulman AH, Pavelek M. Genetic diversity of cultivated flax (*Linum usitatissimum* L.) germplasm assessed by retrotransposon-based markers. *Theor Appl Genet*. (2011) 122:1385–97. doi: 10.1007/s00122-011-1539-2

203. Kumari R, Wankhede DP, Bajpai A, Maurya A, Prasad K, Gautam D, et al. Genome wide identification and characterization of microsatellite markers in black pepper (*Piper nigrum*): a valuable resource for boosting genomics applications. *PLoS One*. (2019) 14:e0226002. doi: 10.1371/journal.pone.0226002

204. Cloutier S, Niu Z, Datla R, Duguid S. Development and analysis of EST-SSRs for flax (*Linum usitatissimum* L.). *Theor Appl Genet*. (2009) 119:53–64. doi: 10.1007/s00122-009-1016-3

205. Deng X, Long S, He D, Li X, Wang Y, Hao D, et al. Isolation and characterization of polymorphic microsatellite markers from flax (*Linum usitatissimum* L.). *Afr J Biotechnol*. (2011) 10:734–9.

206. Deng X, Long S, He D, Li X, Wang Y, Liu J, et al. Development and characterization of polymorphic microsatellite markers in *Linum usitatissimum*. *J Plant Res*. (2010) 123:119–23. doi: 10.1007/s10265-009-0271-3

207. Roose-Amsaleg C, Cariou-Pham E, Vautrin DA, Tavernier R, Solignac M. Polymorphic microsatellite loci in *Linum usitatissimum*. *Mol Ecol Notes*. (2006) 6:796–9. doi: 10.1111/j.1471-8286.2006.01348.x

208. Kale SM, Pardeshi VC, Kadoo NY, Ghorpade PB, Jana MM, Gupta VS. Development of genomic simple sequence repeat markers for linseed using next-generation sequencing technology. *Mol Breed*. (2012) 30:597–606. doi: 10.1007/s11032-011-9648-9

209. Wu J, Zhao Q, Wu G, Zhang S, Jiang T. Development of novel SSR markers for flax (*Linum usitatissimum* L.) using reduced-representation genome sequencing. Front. *Plant Sci*. (2017) 7:2016. doi: 10.3389/fpls.2016.02018

210. Pan G, Chen A, Li J, Huang T, Chang L, Zhao L, et al. Genome-wide development of simple sequence repeats database for flax (*Linum usitatissimum* L.) and its use for genetic diversity assessment. *Genet Resour Crop Evol*. (2020) 67:865–74. doi: 10.1007/s10722-020-00882-y

211. Saha D, Rana RS, Das S, Datta S, Mitra J, Cloutier SJ, et al. Genome-wide regulatory gene-derived SSRs reveal genetic differentiation and population structure in fiber flax genotypes. *J Appl Genet*. (2019) 60:13–25. doi: 10.1007/s13353-018-0476-z

212. Landoni B, Viruel J, Gómez R, Allaby RG, Brennan AC, Picó FX, et al. Microsatellite marker development in the crop wild relative *Linum bienne* using genome skimming. *Appl Plant Sci*. (2020) 8:e11349. doi: 10.1002/aps3.11349

213. He L, Xiao J, Rashid KY, Yao Z, Li P, Jia G, et al. Genome-wide association studies for pasmo resistance in flax (*Linum usitatissimum* L.). *Front Plant Sci*. (2018) 9:1982. doi: 10.3389/fpls.2018.01982

214. Xie D, Dai Z, Yang Z, Sun J, Zhao D, Yang X, et al. Genome-wide association study identifying candidate genes influencing important agronomic traits of flax (*Linum usitatissimum* L.) using SLAF-seq. *Front Plant Sci*. (2018) 8:8–2232. doi: 10.3389/fpls.2017.02232

215. Soto-Cerda BJ, Aravena G, Cloutier S. Genetic dissection of flowering time in flax (*Linum usitatissimum* L.) through single- and multi-locus genome-wide association studies. *Mol Gen Genomics*. (2021) 296:877–91. doi: 10.1007/s00438-021-01785-y

216. Wu J, Zhao Q, Zhang L, Li S, Ma Y, Pan L, et al. QTL mapping of fiber-related traits based on a high-density genetic map in flax (*Linum usitatissimum* L.). *Front Plant Sci*. (2018) 9:9–885. doi: 10.3389/fpls.2018.00885

217. You FM, Cloutier S. Mapping quantitative trait loci onto chromosome-scale pseudomolecules in flax. *Methods Protoc*. (2020) 3:28. doi: 10.3390/mps3020028

218. You FM, Rashid KY, Zheng C, Khan N, Li P, Xiao J, et al. Insights into the genetic architecture and genomic prediction of powdery mildew resistance in flax (*Linum usitatissimum* L.). *Int J Mol Sci*. (2022) 23:4960. doi: 10.3390/ijms23094960

219. You FM, Xiao J, Li P, Yao Z, Jia G, He L, et al. Genome-wide association study and selection signatures detect genomic regions associated with seed yield and oil quality in flax. *Int J Mol Sci*. (2018) 19:2303. doi: 10.3390/ijms19082303

220. Chandrawati MR, Singh PK, Ranade SA, Yadav HK. Diversity analysis in Indian genotypes of linseed (*Linum usitatissimum* L.) using AFLP markers. *Gene*. (2014) 549:171–8. doi: 10.1016/j.gene.2014.07.067

221. Hoque A, Fiedler JD, Rahman M. Genetic diversity analysis of a flax (*Linum usitatissimum* L.) global collection. *BMC Genomics*. (2020) 21:557. doi: 10.1186/s12864-020-06922-2

222. Mhiret WN, Heslop-Harrison JS. Biodiversity in Ethiopian linseed (*Linum usitatissimum* L.): molecular characterization of landraces and some wild species. *Genet Resour Crop Evol*. (2018) 65:1603–14. doi: 10.1007/s10722-018-0636-3

223. Spielmeyer W, Green AG, Bittisnich D, Mendham N, Lagudah ES. Identification of quantitative trait loci contributing to fusarium wilt resistance on an AFLP linkage map of flax (*Linum usitatissimum*). *Theor Appl Genet*. (1998) 97:633–41. doi: 10.1007/s001220050939

224. Zhang J, Long Y, Wang L, Dang Z, Zhang T, Song X, et al. Consensus genetic linkage map construction and QTL mapping for plant height-related traits in linseed flax (*Linum usitatissimum* L.). *BMC Plant Biol*. (2018) 18:160. doi: 10.1186/s12870-018-1366-6

225. Wang YH, Wu DH, Huang JH, Tsao SJ, Hwu KK, Lo HF. Mapping quantitative trait loci for fruit traits and powdery mildew resistance in melon (*Cucumis melo*). *Bot Stud*. (2016) 57:19. doi: 10.1186/s40529-016-0130-1

226. Akhmetshina AO, Strygina KV, Khlestkina EK, Porokhovinova EA, Brutch NB. High-throughput sequencing techniques to flax genetics and breeding. *Ecol Gen*. (2020) 18:103–24. doi: 10.17816/ecogen16126

227. Soto-Cerda BJ, Cloutier S, Quian R, Gajardo HA, Olivos M, You FM. Genome-wide association analysis of mucilage and hull content in flax (*Linum usitatissimum* L.) seeds. *Int J Mol Sci*. (2018) 19:2870. doi: 10.3390/ijms19102870

228. Xie D, Da Z, Yang Z, Tang Q, Sun J, Yang X, et al. Genomic variations and association study of agronomic traits in flax. *BMC Genomics*. (2018) 19:512. doi: 10.1186/s12864-018-4899-z

229. Wang Z, Hobson N, Galindo L, Zhu S, Shi D, McDill J, et al. The genome of flax (*Linum usitatissimum*) assembled *de novo* from short shotgun sequence reads. *Plant J*. (2012) 72:461–73. doi: 10.1111/j.1365-313x.2012.05093.x

230. You FM, Xiao J, Li P, Yao Z, Jia G, He L, et al. Chromosome-scale pseudomolecules refined by optical, physical, and genetic maps in flax. *Plant J*. (2018) 95:371–84. doi: 10.1111/tpj.13944

231. Dmitriev AA, Pushkova EN, Novakovskiy RO, Beniaminov AD, Rozhmina TA, Zhuchenko AA, et al. Genome sequencing of fiber flax cultivar Atlant using oxford

nanopore and illumina platforms. *Front Genet*. (2021) 11:590282. doi: 10.3389/fgene.2020.590282

232. Sa R, Yi L, Siqin B, An M, Bao H, Song X, et al. Chromosome-level genome assembly and annotation of the fiber flax (*Linum usitatissimum*) genome. *Front Genet*. (2021) 12:735690. doi: 10.3389/fgene.2021.735690

233. Zhang J, Qi Y, Wang L, Wang L, Yan X, Dang Z, et al. Genomic comparison and population diversity analysis provide insights into the domestication and improvement of flax. *iScience*. (2020) 23:100967. doi: 10.1016/j.isci.2020.100967

234. de Santana LA, Pacheco TG, Santos KGD, Vieira LDN, Guerra MP, Nodari RO, et al. The *Linum usitatissimum* L. plastome reveals atypical structural evolution, new editing sites, and the phylogenetic position of Linaceae within Malpighiales. *Plant Cell Rep*. (2018) 37:307–28. doi: 10.1007/s00299-017-2231-z

235. Kaur V, Aravind J, Manju JSR, Kumari J, Panwar BS, et al. Phenotypic characterization, genetic diversity assessment in 6,778 accessions of barley (*Hordeum vulgare* L. ssp. *vulgare*) germplasm conserved in National Genebank of India and development of a core set. *Front Plant Sci*. (2022) 13:771920. doi: 10.3389/fpls.2022.771920

236. Brown AHD. Core collections: a practical approach to genetic resources management. *Genome*. (1989) 31:818–24. doi: 10.1139/g89-144

237. Frankel OH. Genetic perspectives of germplasm conservation In: *In genetic manipulation: Impact on man and society*. Cambridge: Cambridge University Press (1984). 161–70.

238. Kutuzova. *Catalogue of the world collection at the VIR, donors of economically important characters for breeding of fiber flax*, vol. *714*. St. Petersburg: VIR (2000). 50 p.

239. Van Soest LJM, Bas N. Current status of the CGN Linum collection In: *Flax genetic resources in Europe*. Maccarese Rome: IPGRI (2002). 44–8.

240. Painting KA, Perry MC, Dennin RA, Ayad WG. *Guidebook for genetic resources documentation, a self-teaching approach to the understanding, analysis and development of genetic resources documentation*. Rome: International Plant Genetic Resources Institute (1993).

241. Gupta K, Archak S, Pradheep K, Kumar S, Jacob SR, Tyagi V, et al. ICAR-NBPGR: bridging science and service In: *ICAR-National Bureau of plant genetic resources*. India: New Delhi (2012-2018). 74.

242. Kak A, Gupta V. (2015) Inventory of registered crop germplasm (2015–2017). ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, Indian Council of Agricultural Research, New Delhi, India. pp. 74.

243. Kak A, Tyagi RK. (2017). Inventory of registered crop germplasm (2010–2014). ICAR-National Bureau of Plant Genetic Resources, Pusa Campus, Indian Council of Agricultural Research, New Delhi, India. pp. 103.

244. Kak A, Srinivasan K, Sharma SK. (2009). Plant germplasm registration National Bureau of Plant Genetic Resources, Pusa campus, Indian Council of Agricultural Research, New Delhi, India. pp. 75.

245. Srivastava RL, Singh K, Singh J, Kerkhi SA. Catalogue on linseed germplasm *Project coordinating unit (linseed)*. Kanpur, UP, India: CSAUA&T (2010)

246. Verma VD, Mandel S, Patel DP, Loknathan TR, Singh B, Sapra RL, et al. Evaluation of linseed germplasm. *National Bureau of plant genetic resources*. Akola, Maharashtra, India: Regional Station (1993)

247. Rozhmina T, Mikhail B, Anastasia S, Alexander K, Maria S. A comprehensive dataset of flax (*Linum uitatissimum* L.) phenotypes. *Data Br*. (2021) 37:107224. doi: 10.1016/j.dib.2021.107224

248. Mokshina N, Gorshkov O, Takasaki H, Onodera H, Sakamoto S, Gorshkova T, et al. FIBexDB: a new online transcriptome platform to analyze development of plant cellulosic fibers. *New Phytol*. (2021) 231:512–5. doi: 10.1111/nph.17405

249. GENESYS (2023) Available at: https://www.genesys-pgr.org/ [Accessed 31 January, 2023].

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 16 June 2023
DOI 10.3389/fnut.2023.1141040

Check for updates

OPEN ACCESS

EDITED BY
Tanushri Kaul,
International Centre for Genetic Engineering and Biotechnology, India

REVIEWED BY
Suresh Kumar Paramasivam,
National Research Centre for Banana (ICAR), India
Parameswaran Chidambaranathan,
Indian Council of Agricultural Research (ICAR), India

*CORRESPONDENCE
Ruchi Bansal
✉ ruchibansal06@gmail.com

†These authors have contributed equally to this work and share first authorship

RECEIVED 09 January 2023
ACCEPTED 24 May 2023
PUBLISHED 16 June 2023

CITATION
Bansal R, Bana RS, Dikshit HK, Srivastava H, Priya S, Kumar S, Aski MS, Kumari NKP, Gupta S and Kumar S (2023) Seed nutritional quality in lentil (*Lens culinaris*) under different moisture regimes.
Front. Nutr. 10:1141040.
doi: 10.3389/fnut.2023.1141040

COPYRIGHT
© 2023 Bansal, Bana, Dikshit, Srivastava, Priya, Kumar, Aski, Kumari, Gupta and Kumar. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Seed nutritional quality in lentil (*Lens culinaris*) under different moisture regimes

Ruchi Bansal[1,2*†], Ram Swaroop Bana[2†], Harsh K. Dikshit[2], Harshita Srivastava[1], Swati Priya[1], Sunil Kumar[1], Muraleedhar S. Aski[2], N. K. Prasanna Kumari[3], Sanjeev Gupta[4] and Shiv Kumar[5]

[1]ICAR-National Bureau of Plant Genetic Resources, New Delhi, India, [2]ICAR-Indian Agricultural Research Institute, New Delhi, India, [3]National Institute of Science Communication and Policy Research, New Delhi, India, [4]Indian Council of Agricultural Research, New Delhi, India, [5]International Center for Agriculture in Dryland Areas, Rabat, Morocco

The world's most challenging environmental issue is climate change. Agricultural productivity and nutritional quality are both substantially threatened by extreme and unpredicted climate events. To develop climate resilient cultivars, stress tolerance along with the grain quality needs to be prioritized. Present study was planned to assess the effect of water limitation on seed quality in lentil, a cool season legume crop. A pot experiment was carried out with 20 diverse lentil genotypes grown under normal (80% field capacity) and limited (25% field capacity) soil moisture. Seed protein, Fe, Zn, phytate, protein and yield were recorded in both the conditions. Seed yield and weight were reduced by 38.9 and 12.1%, respectively, in response to stress. Seed protein, Fe, Zn, its availability as well as antioxidant properties also reduced considerably, while genotype dependent variation was noted with respect to seed size traits. Positive correlation was observed between seed yield and antioxidant activity, seed weight and Zn content and availability in stress. Based on principal component analysis and clustering, IG129185, IC559845, IC599829, IC282863, IC361417, IG334, IC560037, P8114 and L5126 were promising genotypes for seed size, Fe and protein content, while, FLIP-96-51, P3211 and IC398019 were promising for yield, Zn and antioxidant capacity. Identified lentil genotypes can be utilized as trait donors for quality improvement in lentil breeding.

KEYWORDS

biofortification, food, lentil, legume, nutrition, protein, water stress, yield

1. Introduction

Climate change is detrimental to all the dimensions of food and nutritional security. Climate unpredictability has disturbed the global food production, accessibility, utilization as well as food system stability (1). During the past 40 years, agricultural productivity has suffered a significant setback as a result of climatic variabilities like extreme temperatures, flooding, drought, and an increase in the occurrence of pests and diseases (2). The United Nations has set Sustainable Development Goals (SDGs) to achieve a better and sustainable future for all till 2030. The accomplishment of "Zero Hunger" and eradication of poverty is the most important goal among SDGs. To achieve these goals, agriculture and food systems must be sustainable, resource- efficient, nutrition-sensitive, and climate-smart.

Legumes play an important role in food and nutritional security and contribute roughly 10% of daily protein consumption and 5% of daily energy intake (3). They also contain considerable levels of vitamins (thiamin, riboflavin, pyridoxine, vitamin K, E, B and folic acid) and minerals (Ca, Fe, Zn, Mg, and lysine). Though the food legumes grow in a diverse range of environments, abiotic stresses such as drought, heat/temperature, salinity, and heavy metals adversely affect grain yield and quality (4, 5). Water stress is among the most critical factors limiting the production of legumes, particularly in the arid and semi-arid tropics. Water limitation during the flowering/grain filling is highly detrimental to grain yield and quality (6). High temperature/water stress may lead to early senescence, shorten seed filling duration and affect remobilization of assimilates from source to sink (7). Grain development is mostly limited by stress-induced reductions in assimilate supply (8, 9). Poor soil moisture deteriorated grain quality in wheat by affecting protein composition and dietary fibre content (10), while, grain N, P, Fe, and Zn levels along with the total grain protein in chickpea (11). In rice, grain length, width, and total milling recovery decreased, and chalkiness increased under the water deficit (12). Terminal stress altered the fatty acid composition of soybean seeds, which depreciated the oil content, quality, and stability (13). Though different studies have shown the negative impact of water stress on yield in major cereals and legumes, nutritional aspects have not gained much attention.

Lentil (*Lens culinaris* L.) is an important legume crop, which is primarily cultivated in Canada, India, and Turkey (14). Lentil seeds are highly rich in protein (20–30%), low digestible carbohydrates (20%), fat (1.0%), and vitamins (15). The high concentrations of prebiotic or low-digestible carbohydrates in lentils, such as resistant starch (75 mg g^{-1}), raffinose-family oligosaccharides (40.7 mg g^{-1}), sugar alcohols (14.2 mg g^{-1}) and fructo-oligosaccharides (0.62 mg g^{-1}) contribute to its health benefits (16). In black gram and green gram, water stress during the post-flowering growth phase may reduce up to 70% of grain yield (17). Heat individually and in combination with water stress declined grain Fe, Zn and crude protein content in lentils. Combined stress was more detrimental to lentil yield and quality compared to heat individually (18). Since lentil is a highly nutritious legume crop, there is a need to generate information on effect of water limitation on grain quality to ensure nutritional security amid climatic variability. Therefore, we hypothesized that water limitation may deteriorate the seed quality in lentils and identification of genotypes with better yield and quality under stress should be targeted in lentil improvement breeding.

Keeping this in view, the present study aimed to (i) determine the effect of limited soil moisture on nutritional quality in diverse lentil genotypes (ii) analyze the relationship among yield and quality traits in different water regimes (iii) selection of superior lentil genotypes in terms of seed yield and quality in response to stress.

2. Materials and methods

2.1. Experimental material

A set of 20 lentil accessions (P3211, IC560037, IC398019, P8110, L5126, IC278791, IC559845, IG129185, IG334, IC282863, IC201678, IC361417, IC559829, IC279627, IC201676, IC208327, EC78391, P8114, FLIP-96-51, JL3) was assessed for their yield and quality response to water stress in controlled conditions during 2019–20 and 2020–21. The seeds were obtained from National Genebank, ICAR-National Bureau of Plant Genetic Resources, New Delhi, India.

2.2. Experimental conditions

The experiment was conducted in a randomized control block design with four replications (10 pots per replication) and two treatments during rabi season at ICAR-National Bureau of Plant Genetic Resources, New Delhi (28.6331°N, 77.1525°E). Pots with 14″ diameter were filled with the top field soil (sandy loam, pH 7.0) and farmyard manure in a ratio of 1:1. The potting mixture was supplemented with Tricalcium phosphate fertilizer (10 mg kg^{-1}) before pot filling. Plants were maintained at 80% field capacity till flowering. Field capacity was maintained gravimetrically by measuring the pots regularly and supplying only the required amount of water (19). At the onset of flowering, water stress was implemented in one set of genotypes by restricting the water supply till 25% field capacity is reached. Thereafter, stressed plants were maintained at the same field capacity till harvesting. Normal plants were maintained at 80% field capacity till harvesting. Before sowing, seeds were treated with 1% Sodium hypochlorite solution followed by thorough washing. Seed germination was carried out in dark at 22°C. After emergence, five seedlings were transferred to pots. Later on, two plants were maintained in each pot. Plants from normal and stress were harvested at maturity for recording the yield and quality parameters.

2.3. Grain yield and test weight

At maturity, the plants were harvested. Grain yield was measured by thrashing 10 plants from each replication. Test weight was recorded from three replications for 100 seeds per replication.

2.4. Estimation of seed quality traits

Seed samples from three replications were taken for quantifying Fe and Zn content from both water regimes. Samples were digested using the standard diacid digestion method. Total Fe and Zn were measured by using atomic absorption spectroscopy (20). N content was determined in seed samples using the Kjeldahl method (21). Seed protein was calculated by multiplying the N content with 6.25 as a conversion factor.

Phytic acid (PA) was analyzed in the seed samples using the Megazyme kit (22) as per standard assay procedure for P issued by phytase and alkaline phosphatase. The inositol phosphates are acid extracted, then treated with a phytase that is selective for PA (IP6), and the lower myo-inositol phosphate forms. Further reaction with alkaline phosphatase produces the final phosphate from myo-inositol phosphate (IP1) which is relatively phytase resistant. A modified colorimetric method was used to determine the total released phosphate. Inorganic phosphate was quantified as P from a calibration curve developed using standards of known P concentration. PA and Zn contents were converted to moles and the ratio was calculated accordingly. DPPH radical scavenging activity assay was carried out in a methanolic extract of lentil genotypes

spectrophotometrically (23). The activity was calculated using the below- mentioned equation:

$$\text{DPPH free radical scavenging activity}\,(\%) = \left(1 - \text{Absorbance of sample / Absorbance of control}\right) \times 100$$

2.5. Seed size traits

Seed size (area, length and breadth) was studied from each replication by scanning the seeds of each genotype with a flatbed scanner (Canon LiDE 110 version 1.3.00). The scanned images were analyzed using Grain Size & Shape Properties, a MATLAB based software developed by ICAR-Central Institute of Agriculture Engineering, Bhopal, India.

2.6. Statistical analysis

Data was analyzed using R software. The least significant difference was calculated at 5 and 1% p level. Pearson's correlation coefficient analysis and principal component analysis were done to study the correlation and identification of traits contributing to yield and quality under normal and stress conditions. The clustering of the genotypes was done using Ward's method in Microsoft Excel.

3. Results

In both normal and stress conditions, lentil genotypes were assessed for yield (seed plant^{-1}, test weight), seed size (length, width, and area), and quality parameters (Fe, Zn, protein, phytic acid, PA:Zn, DPPH radical scavenging activity). The effects of genotype (G), environment (E), and their interaction (G × E) were highly significant for all examined characteristics at $p < 0.001$, with the exception of test weight (TW) at $p < 0.01$.

3.1. Effect of water stress on yield traits

Seed yield plant^{-1} (SY) and test weight (TW) decreased significantly because of stress (Table 1). Effect of stress was more on yield compared seed weight as shown by reduction in mean SY (38.9%) and TW (12.1%). Coefficient of variation was 6.3% for yield and 7.34% for seed weight under normal condition, which reduced under stress environment (Table 1). G, E and G × E interaction effects were highly significant for SY at $p < 0.001$ and TW at $p < 0.01$ (Table 2).

3.2. Effect of water stress on grain quality

Fe content ranged 31.3–84.1 ppm with a mean value of 52.8 ppm and reduced by 24.2% on exposure to stress (Table 1). Similarly, Zn content recorded the depreciation of 31.1% compared to the control. Zn content varied from 36.6–66.8 ppm under normal and 20.4–49.6 ppm under stress conditions. Grain protein ranged from

TABLE 1 Summary statistics for seed quality, size and yield traits of lentil genotypes.

	Trt	Fe (ppm)	Zn (ppm)	PRT (%)	PA (mgg^{-1})	PA:Zn	DPPH (%)	AREA (mm^2)	LEN (mm)	BRD (mm)	SY gm plant^{-1}	TW (gm)
Mean	N S	52.79 40.02	50.63 34.90	23.26 17.37	8.50 11.25	12.35 25.08	6.81 3.75	9.69 12.94	3.79 4.51	3.22 3.79	5.17 3.16	2.57 2.26
Max	N S	84.05 52.65	66.80 49.60	25.81 22.50	13.30 13.90	18.08 32.56	8.40 5.60	29.32 18.37	6.28 6.88	6.15 4.55	6.92 4.34	3.57 3.37
Min	N S	31.30 27.30	36.60 20.40	17.84 12.66	5.10 9.70	8.05 19.20	4.20 1.40	6.37 8.90	3.15 3.68	2.58 3.13	3.13 2.17	2.21 1.47
SD	N S	12.61 6.74	8.62 8.98	2.02 2.26	1.98 1.36	0.05 0.12	1.25 0.87	4.77 2.61	0.63 0.68	0.73 0.42	0.82 0.52	0.35 0.36
CV (%)	N S	4.17 5.93	5.87 3.89	11.51 7.67	4.29 8.27	3.40 2.83	5.45 4.31	2.03 4.96	6.32 6.63	4.41 9.02	6.30 6.07	7.34 6.28

Mean, Average value; SD, standard deviation; Min, minimum value; Max, maximum value; Fe, Iron; Zn, Zinc; PRT, protein; PA, phytic acid; DPPH, DPPH scavenging activity; AREA, seed area; LEN, length; BRD, breadth; SY, seed yield; TW, test weight; Trt, treatment; N, normal; S, water stress.

TABLE 2 Analysis of variance (ANOVA) for seed size, quality and yield traits of lentil genotypes.

	Fe	Zn	PRT	PA	PA:Zn	DPPH	AREA	LEN	BRD	SY	TW
G	501.51***	344.62 ***	18.99 ***	13.96 ***	3.50 ***	5.47 ***	62.98 ***	1.51 ***	1.50 ***	2.38 ***	0.65 ***
E	219.36***	9216.92***	1039.35***	226.82 ***	9.93 ***	281.39 ***	316.14 ***	15.48 ***	9.84 ***	120.98 ***	3.05***
GxE	318.62***	239.65 ***	8.61 ***	3.34 ***	1.50 ***	1.51 ***	25.80 ***	1.09 ***	0.66 ***	0.43 ***	0.10**

*$p<0.05$, **$p<0.01$,***$p<0.001$. Fe, Iron; Zn, Zinc; PRT, protein; PA, phytic acid; DPPH, DPPH scavenging activity; AREA, seed area; LEN, length; BRD, breadth; SY, seed yield; TW, test weight; G, genotype; E, environment.

17.8–25.8% among different genotypes and. it reduced severely by 25.3% (Table 1).

Mean PA content ranged from 5.1–13.3 in normal with a mean of 8.5 (Table 1). Stress increased PA by 32.4%. Genotypic variation was the highest for PA (11.5%) among all the recorded traits in no stress condition. Similarly, the mean PA:Zn ratio became two -fold in water stress. PA:Zn ratio ranged from 8.05–28.08 in normal conditions. DPPH free radical scavenging activity ranged 4.2–8.4 in normal and declined remarkably by 44.9% on exposure to stress (Table 1). G, E and G×E effects were significant for all the quality traits at $p<0.001$ (Table 2).

3.3. Effect of water stress on seed size

Limited moisture availability resulted in a significant increase in mean seed length, breadth, and area by 18.9, 17.7 and 33.5%, respectively, (Table 1). Variation was the highest for BRD (9.02%) under drought compared to all observed traits. The range for seed LEN was 3.2–6.3 mm, BRD 2.6–6.2 mm and AREA 6.4–29.3 mm^2 under normal condition. ANOVA showed the existence of highly signification G, E, G×E interaction effects for all the size traits (Table 2).

3.4. Association between yield, seed size and quality traits

The correlations between yield and other quality traits are depicted in Figure 1. Seed yield had negative correlation to Fe ($r=0.49$, $p<0.05$) and a positive correlation to DPPH activity ($r=0.54$, $p<0.05$) under normal condition (Figure 1A). Grain weight was positively associated with seed length ($r=0.63$ $p<0.05$), breadth ($r=0.62$ $p<0.05$) and area ($r=0.66$ $p<0.05$). No correlation between TW and seed nutritional quality traits existed in normal condition.

Under stress, yield showed a significant positive correlation to DPPH activity only ($r=0.55$ $p<0.05$; Figure 1B). Seed weight exhibited a positive association with Zn ($r=0.66$ $p<0.05$) and negative correlation to PA ($r=0.51$ $p<0.05$), PA/Zn ($r=0.71$, $p<0.05$) in stress conditions.

3.5. Principal component analysis

Principal component analysis (PCA) was used to determine the major traits accountable for genotypic variability under both the treatments. Eigenvalues, variance, and cumulative variances are shown in Table 3. We identified four principal components explaining 81.21% variability under normal and 80.26% variability under stress conditions.

The PCA analysis revealed that PC1 contributed 33.40% of total variability and it was positively associated with grain quality parameters while grain size traits were negatively correlated in normal conditions (Table 3). The PC2 accounted for 19% and was associated positively with Zn, protein, and phytic acid. Similarly, PC3 contributed to 16.52% of the variability and had a significant association with Zn, Fe, and PA, TW and a low association with grain yield. PC4 explained a 12.30% highly positive association with Zn, protein, and other studied traits except for Fe and grain weight (Table 3). PCA biplot analysis considering PC1, and PC2 identified three trait groups among studied genotypes (Figure 2A). Seed size, quality (Except Fe, PA and PA:Zn ratio) and yield traits were the major contributors to PC1 and were highly correlated in the genotypes present in group I, and IV. Seed Fe, phytate and PA:Zn ratio contributed to PC2 and was correlated in genotypes present in group II. Genotypes in group III were associated to yield and quality traits contributing to PC1.

In stress, PC1 explained 35.02% variability and was positively correlated to all the grain yield, quality parameters and morphometric parameters except phytic acid and PA/Zn (Table 3). PC2 contributed 17.94% and phytate and its ratio to Zn contributed mainly to component 2. Similarly, PC3 accounted for 15.58% variation and had a profound positive association with Zn and Fe. PC4 depicted 11.72% variability with a significant positive association with Zn, seed protein, and phytate. PCA biplot based on PC1 and PC2 accounting for 52.96% variation showed the presence of two groups under stress (Figure 2B). All the seed size, yield, and quality traits contributed for PC1 and group II genotypes had a great association with the traits referring to PC1. PA and its ratio to Zn were the major traits for PC2 and group I genotypes were strongly correlated to these traits.

3.6. Identification of potential genotypes by clustering

The hierarchical cluster analysis was performed following Ward's method considering all the recorded traits under normal and water limited conditions. The genotypes were grouped into 4 clusters each in normal and stress environment (Figure 3). In no stress condition, the genotypes were put together into two major clusters at a distance of 38% under normal condition (Figure 3A). Cluster I had FLIP-96-51. Second cluster was further divided into 3 sub-clusters., cluster II IC278791, IC208327, IC559845 and IC282863, cluster III IG129185, C279627, IG334, IC201678, P8110, P8114, IC398019, IC361417, IC559829, EC78391 and cluster IV L5126, JL3, P3211, IC560037 and IC201676 (Figure 3A).

In stress, genotypes were classified into two major clusters (Figure 3B). Cluster I consisted of JL3, IC201676, IC208327 and Cluster 2 had the rest of the genotypes and was further sub divided

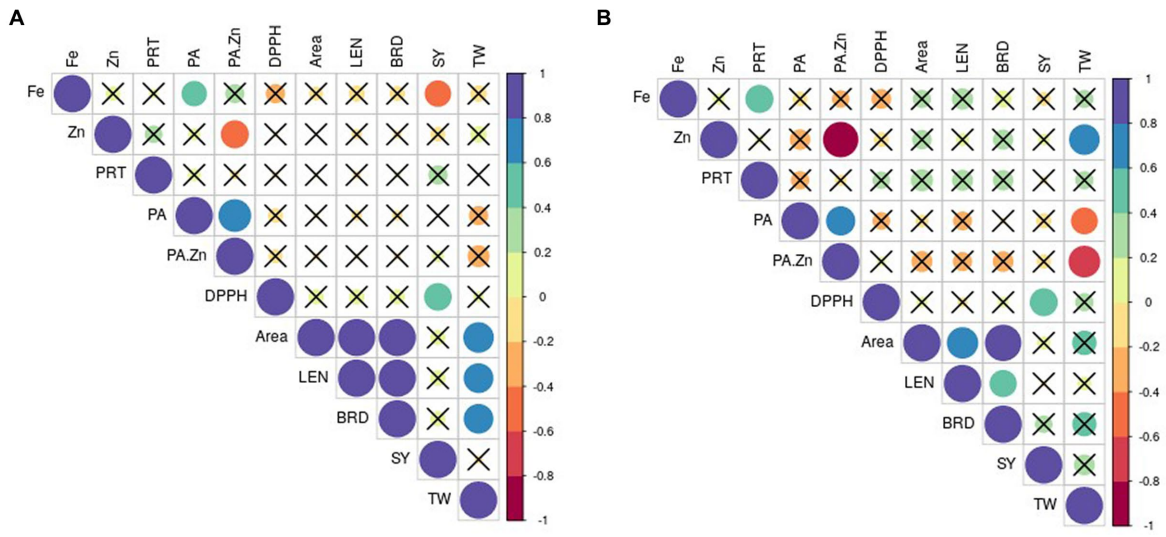


FIGURE 1
Pearson correlation's coefficient between selected traits in **(A)** normal and **(B)** stress conditions. The non-significant correlations ($p<0.05$) are indicated with a cross in the individual cells. PRT, protein; PA, phytic acid; DPPH, DPPH scavenging activity; LEN, length; BRD, breadth; SY, seed yield; TW, test weight.

TABLE 3 Extracted Eigenvalues and vectors associated with the first four principal components (PC) under normal and stress condition.

Particulars	Treatment	PC1	PC2	PC3	PC4
Eigenvalues	Normal Stress	3.67 3.85	2.09 1.97	1.82 1.72	1.35 1.29
variance (%)	Normal Stress	33.40 35.02	19.00 17.93	16.52 15.58	12.30 11.72
Cumulative variance (%)	Normal Stress	33.40 35.02	52.40 52.95	68.92 68.53	81.22 80.25
Traits		Coefficient vectors			
Zn	Normal Stress	0.19 0.21	0.03 0.32	0.52–0.25	0.76–0.45
Fe	Normal Stress	0.36 0.34	−0.17 -0.30	0.35–0.30	−0.12 0.22
Protein	Normal Stress	0.17 0.23	0.01 0.28	−0.65 -0.31	0.45–0.57
Phytic acid	Normal Stress	−0.34 -0.27	0.17 0.28	0.22 0.06	0.26 0.38
PA/Zn ratio	Normal Stress	−0.39 -0.39	0.24 0.32	−0.15 0.29	0.16–0.06
DPPH scavenging activity	Normal Stress	0.37 0.06	−0.10 -0.23	−0.19 0.62	0.14–0.24
Area	Normal Stress	−0.25 0.39	−0.51 0.33	−0.03 0.20	0.13 0.29
Length	Normal Stress	−0.28 0.30	−0.42 0.38	−0.05 0.04	0.14 0.08
Breadth	Normal Stress	−0.27 0.37	−0.48 0.28	−0.05 0.23	0.10 0.34
Seed yield	Normal Stress	0.37 0.13	−0.09 -0.29	−0.24 0.51	0.13 0.02
Test weight	Normal Stress	0.22 0.40	−0.45 -0.29	0.13–0.01	−0.16 -0.05

into small subclusters. The most diverse genotypes, JL3 and P8114, were identified along the dendrogram's edge. Accessions within each cluster showed less variation, however among the clusters a significant difference was observed with respect to grain yield and quality traits. Cluster 1 (JL3, IC201676, IC208327) recorded poor performance with respect to seed size, yield and quality traits and reported higher values for phytate and PA/Zn ratio under stress (Figure 2B). In cluster II, IG129185, IC559845, IC599829, IC282863, IC361417, IG334, IC560037, P8114 and L5126 were promising with respect to seed size, Fe and protein content (Figures 2B, 3B). FLIP-96-51, P3211, IC398019, and P8110 performed better in terms of yield, Zn, and DPPH radical scavenging activity (Figures 2B, 3B).

4. Discussion

Lentil is sensitive to water limitation during the seedling and flowering stage. Severe water stress may reduce the crop yield by 50% depending on the stage (18). In the study, the effect of water stress was studied on yield, seed size and quality traits in lentils. Stress reduced seed Fe, Zn, protein, and DPPH radical scavenging capacity in different lentil genotypes (Table 1). In contrast, there was an increase in mean length, breadth, area, phytate, and PA/Zn ratio in all the tested genotypes.

Seed yield and test weight reduced in genotypes on exposure to stress and G, E, and their interaction effects were significant at

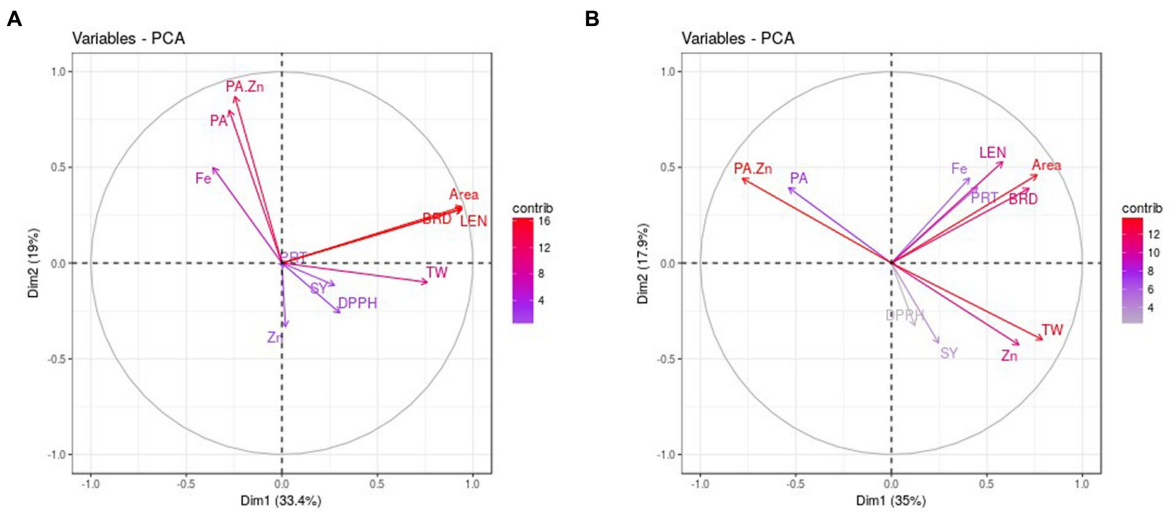


FIGURE 2
PCA biplots of studied traits in normal **(A)** and water stress **(B)** conditions. PRT: protein, PA: phytic acid, DPPH: Total antioxidant activity, LEN: length, BRD: breadth, SY: seed yield, TW: test weight.

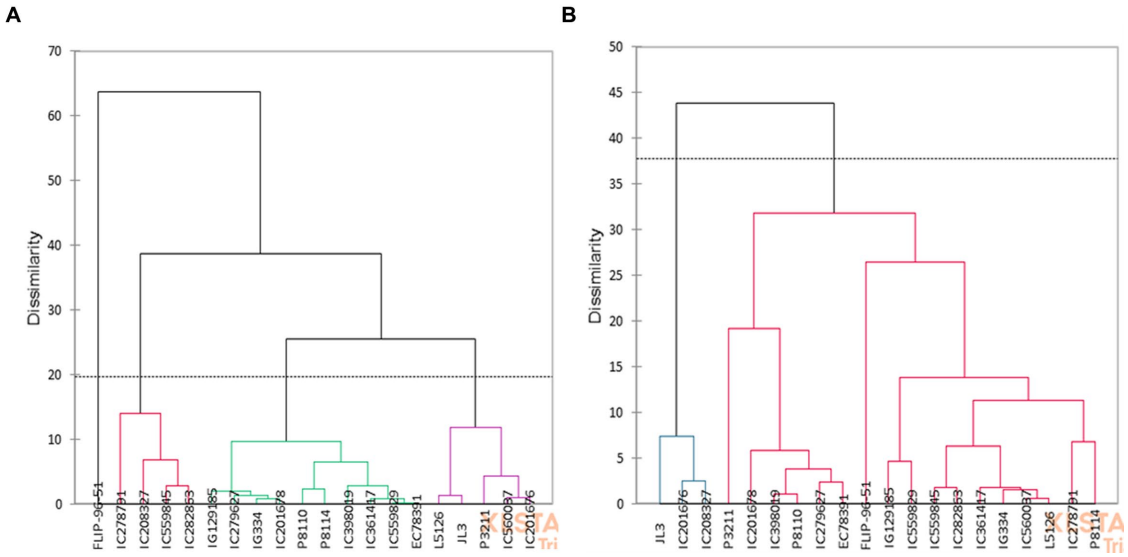


FIGURE 3
Dendrogram representing the clustering in **(A)** normal and **(B)** water stress conditions.

$p < 0.001$. As already reported in different studies, impairment of physiological mechanism, photosynthate mobilization, and loss of pods was associated with the yield loss in lentil under stress condition (24–26).

Genetic constitution and environmental factors affect the seed composition in legumes significantly (27). Grain mineral and protein content have a profound correlation to environmental conditions. Present reported significant reduction in Fe and Zn content in different lentil genotypes in stress. Water limitation induced decrease in grain Fe and Zn content may be due to hampered nutrient uptake, availability, transport and unloading mechanism (28, 29). In addition, nutrient absorption and utilization efficiency may further decline due to slow transpiration rate under water stress (26, 30). Though we reported low mineral content in lentil, on the contrary, water deficit had no effect on seed mineral content in common bean (31).

The PA is an important form of phosphorus storage in legumes. It is required during germination and the early stages of plant growth. It is a potential chelator of cations and can bind with minerals in the digestive tract available by food consumption. Therefore, PA poses a constraint to nutrient absorption and may lead to deficiency. Though the lentil germplasm had significant variability

for Fe and Zn content, but the presence of antinutrients like PA may limit the intake of micronutrients. Therefore, in the present study PA and its ratio to Zn were quantified under normal and water stress conditions. We observed a highly significant increase in PA and a reduction in Zn availability because of stress. Dumschott et al. (32) reported the accumulation of phytic acid like myo-inositol hexa phosphoric acid during water scarcity observed in the present study as well. High PA concentration was reported in lentils and common beans under the high temperature regimes (33, 34) and also under combined heat and water stress (24, 26). Water limitation downregulates the protein synthesis. In addition, poor N fixation and partitioning also result in low protein in legumes under drought conditions (18). We also observed mean reduction of 25.3% in the protein content of the lentil genotypes (18) and chickpeas (11). However, the highest protein content was 25.81 and 22.5% under normal and stress conditions, respectively, (Table 1), which shows that response was genotype dependent. Lentil is rich in polyphenols, which have antioxidant, anti-inflammatory, and nephro-protective properties. Lentil consumption is recommended to decrease the risk of diabetes, obesity, and other cardio-vascular diseases. Mean antioxidant activity quantified using DPPH asaay declined dramatically under stress. Reduction in antioxidant capacity may be attributed to a decrease in phenol concentration as reported in cow pea (35). Correlation study showed that grain yield was positively related to antioxidant properties in stress. Though mean antioxidant activity reduced, genotypes with higher yield may had comparatively more antioxidant capacity. Seed size is indirectly determined by measuring the test weight of seeds in majority of the studies (36). But seed weight does not represent the seed size precisely. In the present study, seeds were scanned, and images were analyzed to measure seed length, breadth and area under both the conditions. We noticed an increase in seed shape parameters in stress conditions. Mean length and breadth increased slightly, while the change in seed area was the most significant (Table 1). The range for different parameters revealed that minimum seed length, breadth, and area increased slightly, while maximum reduced (Except length) in water- limited environment (Table 1). Correlation analysis showed that seed size traits were correlated to grain weight under normal condition and no correlation existed in water stress. Shrestha et al. (37) reported 40 and 30% reduction in flower formation and seed abortion, respectively, in lentil under water deficit. Since these parameters were not recorded in the present study, there is a possibility that due to high seed abortion, seed length might have increased. However, there was no significant difference between control and water stress conditions for pod related traits in the present investigation, intensity of water stress imposed in both the studies may be the reason for the difference in the trait response. Another possible reason for increase in grain size may be genotypic adaptation in response to drier conditions as observed in wild lupin in contrast to cultivated lupin species (38). In this regard, a more detailed study considering the seed coat thickness and role of pod wall may be carried out. Similar findings were documented in lentils under heat (39) and soybean (40). As hypothesized, we reported reduced seed yield along with poor quality in different lentil genotypes in response to water limitation.

Correlation analysis showed presence of positive correlation between antioxidant properties and grain yield under water stress, which may be due to tolerance in lentil genotypes for water stress. Lentil genotypes with higher Zn content and availability were able to maintain higher grain weight by maintain better plant water relations, antioxidative potential, stomatal regulation and photosynthesis in stress as reflected by its positive correlation to TW in stress. PCA analysis defined major traits contributing to total variability in the studied genotypes, which was mainly explained by yield and nutritional traits as reported in previous studies in lentils (18, 39). Clustering identified four groups under normal and two groups under stress conditions. Genotypes present in Cluster I had greater decline in seed quality traits (Fe, Zn, PRT and DPPH radical scavenging activity) compared to cluster II and its sub-clusters (Figures 2, 3). Cluster I genotype showed more susceptibility to water stress. Similar findings were previously reported in lentils (41, 42). Based on clustering, genotypes IG129185, IC559845, IC599829, IC282863, IC361417, IG334, IC560037, P8114 and L5126 were promising for seed size, Fe and protein content and FLIP-96-51, P3211, IC398019 promising for yield, Zn and antioxidant properties. By focusing on seed quality traits along with seed yield, we could identify promising genotypes, which can be target for quality breeding in drought prone environment. Identified genotypes can also be used for further detailed study of Fe/Zn uptake, transport and availability during water scarcity in lentil.

5. Conclusion

Our findings revealed that water stress not only reduced yield but also deteriorated grain quality in lentils, therefore breeding for water limited environment should also take grain quality into consideration. Genotypic responses were significantly different with respect to the yield and quality in different moisture regimes, which may be attributed to the diversity prevailing in the studied genotypes. Positive correlation among yield and antioxidant property, grain Zn and TW in stress may be investigated and validated further with large number of genotypes.Identified genotypes IG129185, IC559845, IC599829, IC282863, IC361417, IG334, IC560037, P8114 and L5126, FLIP-96-51, P3211, IC398019 may serve as potential donor for different yield and quality traits in lentil genetic improvement or biofortification.

Data availability statement

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

Author contributions

RB: conceptualization, resources, supervision, manuscript writing, and finalization. RSB: resources, manuscript writing, and editing. HD: resources and manuscript editing. HS and SP: investigation. SuK: data analysis and manuscript writing. NK: data analysis. MA: resources and manuscript writing. SG: manuscript editing and supervision. ShK: resources and supervision. All authors contributed to the article and approved the submitted version.

Funding

The study was supported by the Department of Biotechnology, India (BT/PR30790/BIC/101/1185/2018) and the CGIAR Research program on grain legumes and dryland cereals.

Acknowledgments

The authors acknowledge the support provided by the Director, ICAR-National Bureau of Plant Genetic Resources, New Delhi, and the Director, ICAR- Indian Agricultural Research Institute, New Delhi to carry out the study.

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

1. De Moura Ariza Alpino T, Mazoto ML, De Barros DC, De Freitas CM. The impacts of climate change on food and nutritional security: a literature review. *Ciên Saúde Colet*. (2022) 27:273–86. doi: 10.1590/1413-81232022271.05972020

2. Fan S. *Food policy in 2018–2019: Growing urgency to address the SDGs*. International food policy research institute (IFPRI), 2019 global food policy report. (2019).

3. Devirian TA, Volpe SL. The physiological effects of dietary boron. *Crit Rev Food Sci Nutr*. (2003) 43:219–31. doi: 10.1080/10408690390826491

4. Chen PX, Bozzo GG, Freixas-Coutin JA, Marcone MF, Pauls PK, Tang Y, et al. Free and conjugated phenolic compounds and their antioxidant activities in regular and non-darkening cranberry bean (*Phaseolus vulgaris* L.) seed coats. *J Funct Foods*. (2015) 18:1047–56. doi: 10.1016/j.jff.2014.10.032

5. Qados AMA. Effect of salt stress on plant growth and metabolism of bean plant *Vicia faba* (L.). *J Saudi Soc Agric Sci*. (2011) 10:7–15. doi: 10.1016/j.jssas.2010.06.002

6. Sehgal A, Sita K, Bhandari K, Kumar S, Kumar J, Vara Prasad PV, et al. Influence of drought and heat stress, applied independently or in combination during seed development, on qualitative and quantitative aspects of seeds of lentil (*Lens culinaris* Medikus) genotypes, differing in drought sensitivity. *Plant Cell Environ*. (2019) 42:198–211. doi: 10.1111/pce.13328

7. Plaut Z, Butow BJ, Blumenthal CS, Wrigley CW. Transport of dry matter into developing wheat kernels and its contribution to grain yield under post-anthesis water deficit and elevated temperature. *Field Crops Res*. (2004) 86:185–98. doi: 10.1016/j.fcr.2003.08.005

8. Farooq M, Gogoi N, Barthakur S, Baroowa B, Bharadwaj N, Alghamdi SS, et al. Drought stress in grain legumes during reproduction and grain filling. *J Agron Crop Sci*. (2017) 203:81–102. doi: 10.1111/jac.12169

9. Farooq M, Nadeem F, Gogoi N, Ullah A, Alghamdi SS, Nayyar H, et al. Heat stress in grain legumes during reproductive and grain-filling phases. *Crop Pasture Sci*. (2017) 68:985–1005. doi: 10.1071/CP17012

10. Rakszegi M, Darko E, Lovegrove A, Molnár I, Lang L, Bedő Z, et al. Drought stress affects the protein and dietary fiber content of wholemeal wheat flour in wheat/*Aegilops* addition lines. *PLoS One*. (2019) 14:e0211892. doi: 10.1371/journal.pone.0211892

11. Ashrafi V, Pourbozorg H, Kor NM, Ajirloo AR, Shamsizadeh M, Shaaban M. Study on seed protein and protein profile pattern of chickpea (*Cicer arietinum* L.) by SDS-PAGE under drought stress and fertilization. *Int J Life Sci*. (2015) 9:87–90. doi: 10.3126/ijls.v9i5.12704

12. Mumtaz MZ, Saqib M, Abbas G, Akhtar J, Ul-Qamar Z. Drought stress impairs grain yield and quality of rice genotypes by impaired photosynthetic attributes and K nutrition. *Rice Sci*. (2020) 27:5–9. doi: 10.1016/j.rsci.2019.12.001

13. Ardakani LG, Farajee H, Kelidari A. The effect of water stress on grain yield and protein of spotted bean (*Phaseolus vulgaris* L.), cultivar Talash. *Int J Adv Biol Biomed Res*. (2013) 1:940–9.

14. FAOSTAT. *Statistics database*. Food and Agriculture Organization of the United Nations, Rome, Italy. (2020). Available at: https://www.fao.org/faostat.

15. Thavarajah D, Johnson CR, McGee R, Thavarajah P. Phenotyping nutritional and antinutritional traits. In: *Phenomics of crop plants: trends, and options and limitations*. New Delhi: Springer (2015) 223–233.

16. Johnson CR, Combs GF Jr, Thavarajah P. Lentil (*Lens culinaris* L.): a prebiotic-rich whole food legume. *Food Res Int*. (2013) 51:107–13. doi: 10.1016/j.foodres.2012.11.025

17. Baroowa B, Gogoi N. Biochemical changes in black gram and green gram genotypes after imposition of drought stress. *J Food Legumes*. (2014) 27:350–3.

18. Choukri H, Hejjaoui K, El-Baouchi A, El Haddad N, Smouni A, Maalouf F, et al. Heat and drought stress impact on phenology, grain yield and nutritional quality of lentil (*Lens culinaris* Medikus). *Front Nutr*. (2020) 7:596307. doi: 10.3389/fnut.2020.596307

19. Pang J, Turner NC, Khan T, Du YL, Xiong JL, Colmer TD, et al. Response of chickpea (*Cicer arietinum* L.) to terminal drought: leaf stomatal conductance, pod abscisic acid concentration, and seed set. *J Exp Bot*. (2017) 68:1973–85. doi: 10.1093/jxb/erw153

20. AOAC. *Official methods of analysis. 15th* ed. Arlington, VA: Association of Official Analytical Chemist (1990).

21. Kjeldahl C. A new method for the determination of nitrogen in organic matter. *Anal Chem*. (1983) 22:366–82. doi: 10.1007/BF01338151

22. Megazyme Phytic Acid Assay Kit. (2016). Available at: https://www.megazyme.com/phytic-acid-assay-kit.

23. Brand-Williams W, Cuvelier ME, Berset C. Use of a free radical method to evaluate antioxidant activity. *LWT*. (1995) 28:25–30. doi: 10.1016/S0023-6438(95)80008-5

24. Awasthi R, Kaushal N, Vadez V, Turner NC, Berger J, Siddique KH, et al. Individual and combined effects of transient drought and heat stress on carbon assimilation and seed filling in chickpea. *Funct Plant Biol*. (2014) 41:1148–67. doi: 10.1071/fp13340

25. Delahunty A, Nuttall J, Nicolas M, Brand J. Response of lentil to high temperature under variable water supply and carbon dioxide enrichment. *Crop Pasture Sci*. (2018) 69:1103–12. doi: 10.1071/CP18004

26. Sehgal A, Sita K, Kumar J, Kumar S, Singh S, Siddique KHM, et al. Effects of drought, heat and their interaction on the growth, yield and photosynthetic function of lentil (*Lens culinaris* Medikus) genotypes varying in heat and drought sensitivity. *Front Plant Sci*. (2017) 8:1776. doi: 10.3389/fpls.2017.01776

27. Grusak MA. Enhancing mineral content in plant food products. *J Am Coll Nutr*. (2002) 21:178S–83S. doi: 10.1080/07315724.2002.10719263

28. Huang B, Rachmilevitch S, Xu J. Root carbon and protein metabolism associated with heat tolerance. *J Exp Bot*. (2012) 63:3455–65. doi: 10.1093/jxb/ers003

29. Sánchez-Rodríguez E, del Mar Rubio-Wilhelmi M, Cervilla LM, Blasco B, Rios JJ, Leyva R, et al. Study of the ionome and uptake fluxes in cherry tomato plants under moderate water stress conditions. *Plant Soil*. (2010) 335:339–47. doi: 10.1007/s11104-010-0422-2

30. Bista DR, Heckathorn SA, Jayawardena DM, Mishra S, Boldt JK. Effects of drought on nutrient uptake and the levels of nutrient-uptake proteins in roots of drought-sensitive and-tolerant grasses. *Plan Theory*. (2018) 7:28. doi: 10.3390/plants7020028

31. Smith MR, Veneklaas E, Polania J, Rao IM, Beebe SE, Merchant A. Field drought conditions impact yield but not nutritional quality of the seed in common bean (*Phaseolus vulgaris* L.). *PLoS One*. (2019) 14:e0217099. doi: 10.1371/journal.pone.0217099

32. Dumschott K, Richter A, Loescher W, Merchant A. Post photosynthetic carbon partitioning to sugar alcohols and consequences for plant growth. *Phytochemistry*. (2017) 144:243–52. doi: 10.1016/j.phytochem.2017.09.019

33. Hummel M, Hallahan BF, Brychkova G, Ramirez-Villegas J, Guwela V, Chataika B, et al. Reduction in nutritional quality and growing area suitability of common bean under climate change induced drought stress in Africa. *Sci Rep*. (2018) 8:1–11. doi: 10.1038/s41598-018-33952-4

34. Thavarajah P, See CT, Vandenberg A. Phytic acid and Fe and Zn concentration in lentil (*Lens culinaris* L.) seeds is influenced by temperature during seed filling period. *Food Chem*. (2010) 122:254–9. doi: 10.1016/j.foodchem.2010.02.073

35. Carvalho M, Gouvinhas I, Castro I, Matos M, Rosa E, Carnide V, et al. Drought stress effect on polyphenolic content and antioxidant capacity of cowpea pods and seeds. *J Agron Crop Sci*. (2021) 207:197–207. doi: 10.1111/jac.12454

36. Tullu A, Kusmenoglu I, McPhee KE, Muehlbauer FJ. Characterization of core collection of lentil germplasm for phenology, morphology, seed and straw yields. *Genet Resour Crop Evol*. (2001) 48:143–52. doi: 10.1023/A%3A1011254629628

37. Shrestha R, Turner NC, Siddique KHM, Turner DW. Physiological and seed yield responses to water deficits among lentil genotypes from diverse origins. *Aust J Agric Res*. (2006) 57:903–15. doi: 10.1071/AR05204

38. Berger JD, Shrestha D, Ludwig C. Reproductive strategies in mediterranean legumes: trade-offs between phenology, seed size and vigor within and between wild and domesticated *Lupinus* species collected along aridity gradients. *Front Plant Sci*. (2017) 8:548. doi: 10.3389/fpls.2017.00548

39. Choukri H, El Haddad N, Aloui K, Hejjaoui K, El-Baouchi A, Smouni A, et al. Effect of high temperature stress during the reproductive stage on grain yield and nutritional quality of lentil (*Lens culinaris* Medikus). *Front Nutr*. (2022) 9:857469. doi: 10.3389/fnut.2022.857469

40. Wang X, Fu G, Yuan S, Zhang H. Genetic diversity analysis of seed appearance quality of Chinese soybean mini core collection. *Proc Eng*. (2011) 18:392–7. doi: 10.1016/j.proeng.2011.11.063

41. Kumar H, Dikshit HK, Singh A, Jain N, Kumari J, Singh AM, et al. Characterization of grain iron and zinc in lentil ('*Lens culinaris*' Medikus' culinaris') and analysis of their genetic diversity using SSR markers. *Aust J Crop Sci*. (2014) 8:1005–12.

42. Thavarajah D, Thavarajah P, Sarker A, Vandenberg A. Lentils (*Lens culinaris* Medikus subspecies culinaris): a whole food for increased iron and zinc intake. *J Agric Food Chem*. (2009) 57:5413–9. doi: 10.1021/jf900786e

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 13 July 2023
DOI 10.3389/fpls.2023.1102879

Check for updates

OPEN ACCESS

EDITED BY
Emmanuel Oladeji Alamu,
International Institute of Tropical Agriculture, Zambia

REVIEWED BY
Manish Kumar Pandey,
International Crops Research Institute for the Semi-Arid Tropics (ICRISAT), India
Vadivalagan Chithravel,
China Medical University, Taiwan
Ummed Singh,
Agriculture University, Jodhpur, India

*CORRESPONDENCE
Muraleedhar S. Aski
murali2416@gmail.com
Gyan Prakash Mishra
gyan.gene@gmail.com
Harsh Kumar Dikshit
harshgeneticsiari@gmail.com

RECEIVED 19 November 2022
ACCEPTED 13 June 2023
PUBLISHED 13 July 2023

CITATION
Aski MS, Mishra GP, Tokkas JP, Yadav PS, Rai N, Bansal R, Singh A, Gupta S, Kumar J, Parihar A, Kumar S, Kumar V, Saxsena AK, Das TR, Kumar A and Dikshit HK (2023) Strategies for identifying stable lentil cultivars (*Lens culinaris* Medik) for combating hidden hunger, malnourishment, and climate variability. *Front. Plant Sci.* 14:1102879. doi: 10.3389/fpls.2023.1102879

COPYRIGHT
© 2023 Aski, Mishra, Tokkas, Yadav, Rai, Bansal, Singh, Gupta, Kumar, Parihar, Kumar, Kumar, Saxsena, Das, Kumar and Dikshit. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Strategies for identifying stable lentil cultivars (*Lens culinaris* Medik) for combating hidden hunger, malnourishment, and climate variability

Muraleedhar S. Aski[1]*, Gyan Prakash Mishra[1]*, Jayanti P. Tokkas[2], Prachi S. Yadav[1], Neha Rai[1], Ruchi Bansal[3], Akanksha Singh[4], Sanjeev Gupta[5], Jitendra Kumar[6], Ashok Parihar[6], Shiv Kumar[7], Vinod Kumar[8], Ashok Kumar Saxsena[9], Tapas Ranjan Das[1,10], Anil Kumar[11] and Harsh Kumar Dikshit[1]*

[1]Division of Genetics, Indian Council of Agricultural Research (ICAR)-Indian Agricultural Research Institute, New Delhi, India, [2]Division of Biochemistry, Chaudhary Charan Singh (CCS), Hissar Agricultural University, Hissar, Haryana, India, [3]Division of Plant Physiology, Indian Council of Agricultural Research (ICAR)-Indian Agricultural Research Institute, New Delhi, India, [4]Amity Institute of Organic Agriculture, Amity University, Noida, India, [5]Krishi Bhavan, Indian Council of Agricultural Research (ICAR), Delhi, India, [6]Division of Crop Improvement, Indian Council of Agricultural Research (ICAR) – Indian Institute of Pulses Research, Kanpur, UP, India, [7]Regional Research Station, International Center for Agricultural Research in the Dry Areas (ICARDA), New Delhi, India, [8]Regional Agricultural Research Station, Jawaharlal Nehru Krishi Vishwavidyalaya (JNKVV), Sagar, Madhya Pradesh, India, [9]RVSKVV-Rajmata Vijayaraje Scindia Krishi Vishwa Vidyalaya, Department of Plant Breeding & Genetics, College of Agriculture, Sehore, MP, India, [10]Indian Council of Agricultural Research (ICAR)-Indian Agricultural Research Institute (IARI) Regional Research Station, Pusa Samastipure, Bihar, India, [11]Department of Plant Breeding & Genetics, Bihar Agricultural College, Bihar Agriculture University, Bhagalpur, Bihar, India

Iron and zinc malnutrition is a global humanitarian concern that mostly affects newborns, children, and women in low- and middle-income countries where plant-based diets are regularly consumed. This kind of malnutrition has the potential to result in a number of immediate and long-term implications, including stunted growth, an elevated risk of infectious diseases, and poor development, all of which may ultimately cause children to not develop to the fullest extent possible. A determination of the contributions from genotype, environment, and genotype by environment interactions is necessary for the production of nutrient-dense lentil varieties that offer greater availability of iron and zinc with a high level of trait stability. Understanding the genotype and environmental parameters that affect G x E (Genotype x Environment) interactions is essential for plant breeding. We used GGE(Genotype, Genotype x Environment interactions) and AMMI (Additive Main effects and Multiplicative Interaction) models to study genetic stability and GE(Genotype x Environment interactions) for grain Fe, Zn, Al, and anti-nutritional factors like phytic acid content in sixteen commercially produced lentil cultivars over several different six geographical locations across India. Significant genetic variability was evident in the Fe and Zn levels of different genotypes of lentils. The amounts of grain iron, zinc, and phytic acid varied from 114.10 to 49.90 mg/kg, 74.62 to 21.90 mg/kg, and 0.76 to 2.84 g/100g (dw) respectively. The environment and GE (Genotype x Environment interactions) had an impact on the concentration of grain Fe, Zn,

and phytic acid (PA). Heritability estimations ranged from low to high (53.18% to 99.48%). The study indicated strong correlation between the contents of Fe and Zn, a strategy for simultaneously increasing Fe and Zn in lentils may be recommended. In addition, our research revealed that the stable and ideal lentil varieties L4076 (Pusa Shivalik) for Fe concentration and L4717 (Pusa Ageti) for Zn content, which have lower phytic acid contents, will not only play an essential role as stable donors in the lentil bio-fortification but will also enable the expansion of the growing area of bio-fortified crops for the security of health and nutrition.

KEYWORDS

micro-nutrients, hidden hunger, anti-nutrients, stability, bio-fortification, lentil grain iron, grain zinc, phytic acid

Introduction

Micronutrient malnutrition, which affects more than a quarter of the world's population, is a serious global health issue. (Gonmei and Toteja, 2018). The most typical cause of anemia, affecting 27% of the global population (Ning and Zeller, 2019), is iron deficiency. Symptoms of anemia and iron deficiency include mental impairment, lowered immunity, fatigue, early birth in neonates, and higher rates of morbidity and mortality (Cappellini et al., 2019; Jamnok et al., 2020). The oxygen-carrying proteins myoglobin and hemoglobin need iron to function (Fava et al., 2019). According to a study (Lynch et al., 1984), the amount of Fe consumed from diets based on legumes ranges from 0.8 to 1.9 percent.

According to scientific forecasts, between 17.6% and 29.6% of the world's population will have poor zinc consumption, with Sub-Saharan Africa, South and Southeast Asia, and Central America showing the largest prevalences (Gupta et al., 2020a). Infections brought on by insufficient zinc intake have been linked to a major part of child mortality. In biological processes, zinc functions as a catalyst, structural ion, and regulatory ion (Grüngreiff et al., 2020). Due to its involvement in several metabolic processes, zinc is necessary for healthy body growth and development. Zinc insufficiency results in immune system issues, epidermal issues, hypogonadism, problems with the central nervous system, and growth retardation. (Chasapis et al., 2020). The only way to correct a zinc deficiency is through continuous zinc intake because the human body cannot store zinc. Phytic acid (PA) serves as the most potent absorption inhibitor (Sandberg, 2002). Phytate has been shown to decrease Ca, Zn, and Fe absorption in humans in a dose-dependent way (Fredlund et al., 2002). The reduction of Fe and Zn absorption by inositol pentaphosphate has also been demonstrated (Sandberg et al., 1999).The National Institute of Health (NIH) specifies that the Recommended Dietary Allowances (RDAs) for iron is 8 mg for men and 18 mg for women, but the RDA for zinc is 11 mg and 8 mg for men and women, correspondingly. The grain legume lentil (*Lens culinaris* Medikus. *culinaris*) is rich in macro- and micronutrients as well as protein, vitamins, and prebiotic carbohydrates (Taleb et al., 2013; Gupta et al., 2018; Khazaei et al., 2019). Daily consumption of 100 grams of lentils can supply significant amounts of Zn and Fe (Thavarajah et al., 2009; Kumar et al., 2018). In many poor nations, lentils are supplemented to cereals in low-income people's daily meals (Ozer et al., 2010). Lentil is a promising grain for micronutrient biofortification since it is high in Zn and Fe and may be grown in areas where people are malnourished and face economic issues. (Singh et al., 2017a; Singh et al., 2017b).

Biofortification, a traditional or molecular breeding-based method, can boost the nutritional content of food crops by enhancing bioavailability (Garg et al., 2018). Significant genetic diversity in the genetic pool for the desired characteristic is necessary for breeding micronutrient-rich crops. It has been demonstrated that the levels of Fe and Zn in lentil germplasm vary significantly (Thavarajah et al., 2011; Karakoy et al., 2012 and Sen Gupta et al., 2013). At Indian and international level breeding efforts are made to identify genetic variability and stability in grain minerals in food grain legumes like chickpea (Erdemci, 2018; Misra et al., 2020), mungbean (Ullah et al., 2011), lentil (Darai et al., 2020), soybean (Mwiinga, 2018), faba bean (Fikere et al., 2008; Tekalign et al., 2017). In addition, IARI New Delhi delivered its first iron-rich lentil variety, Pusa Agethi Masoor, while IIPR Kanpur presented IPL 220 for commercial cultivation to farming community (Yadava et al., 2020).

Its heredity is complex since the environment heavily regulates grain micronutrient concentration (Kumar et al., 2018). The GGE model aids in identifying winning genotypes suitable for various environments and ranking them in tested environments in terms of performance, albeit the AMMI model aids in understanding the structure of GEI (Genotype Environment Interaction), trying to predict the total deviation of interaction, and distinguishes the main interactions from each other (Ebdon and Gauch, 2022a; Ebdon and Gauch, 2022b). Genotypes with stable micronutrient concentration performance across conditions can be exploited when breeding bio-fortified lentil cultivars.

This study focused on the genotype by environment interaction (G x E) across various environments in order to identify stable Fe and Zn-rich genotypes with lower phytic acid concentration.

Material and methods

Plant material and field experiment

In the present study, 16 commercially available lentil cultivars generated in different lentil breeding facilities in India were utilized (Table 1). These 16 cultivars were grown in six distinct geographical locations; each location is representative of a unique lentil growing zone, officially demarked by the All India Co-ordinated Research Project (AICRP) on MULLaRP (Mungbean, Uradbean, Lentil, Lathyrus, and Pea (field) in India.: i) ICAR-IARI New Delhi (North-West Plain Zone); (ii) ICAR-IIPR Kanpur, Uttar Pradesh (UP) (North-East Plain Zone); (iii) Sehore, Madhya Pradesh (MP) (Central Zone); (iv) Sabour, Bihar (BR) (North-East Plain Zone); (v) Samastipure, Bihar (BR) (North-East Plain Zone); and (vi) Sagar Madhya Pradesh (MP) (Central Zone) (Table 2).

The soil's properties, including pH, EC, organic carbon (OC), accessible nitrogen, phosphorus (P), and potassium (K), as well as soil texture, are presented in Supplementary Table 1. In three regions (Delhi, Kanpur, and Sehore), mungbean was previously planted; in the other three, blackgram was cultivated. No basal fertilization or micronutrient spraying has taken place. DAP (Di Ammonium Phosphate) was the only fertilizer used, and it was applied at a rate of 100 kg/ha. To allow adequate uniformity, the topsoil was carefully shredded and mixed, and the land was leveled in each location. The plants were planted in a randomized block design (RBD) with a plant to plant spacing of 5 cm, a row to row distance of 30 cm, and a row length of 5 m, for a total of three repetitions per entry (6 rows each replication). Crop cultivation was carried out using standard agronomic methods. Employing recognized techniques, the amounts of Fe and Zn in soil were determined (Singh et al., 2005).

TABLE 1 Information regarding the lentil genotypes used in the study.

S.No	Genotype	Pedigree	Developing center	Average yield(Q/ha)	Days to maturity	Reaction to Major diseases
1	DPL 62 (Sheri)	JLS-1 x LG 171	ICAR-Indian Institute of Pulses Research, Kanpur, India	17-18	130-135	Resistant to rust and tolerance wilt
2	L 4596 (Pusa Masoor-6)	LC 68-17-3-5 x L 4602	ICAR-Indian Agricultural Research Institute, New Delhi, India	20-22	120-124	Resistant to rust
3	IPL 321	DPL-62 x K 75	ICAR-Indian Institute of Pulses Research, Kanpur, India	9-10	130-135	Resistant to rust and wilt
4	L4147(Pusa Vaibhav)	(L 3875 x P4) x PKVL	ICAR-Indian Agricultural Research Institute, New Delhi, India	17-18	130-135	Resistant to rust
5	DPL 58	PL 639 X PRECOZ	ICAR-Indian Institute of Pulses Research, Kanpur, UP, India	15-18	130-135	Resistant to wilt
6	JL 3	Land race selection from Sagar MP	Jawaharlal Nehru Krishi Vishwa Vidyalaya (JNKVV) Sehore, MP, India	14-15	110-115	Resistant to wilt
7	WBL 77	ILL7723 x BLX88176	Bidhan Chandra Krishi Vidyalaya (BCKV), Berhampore, WB, India	14-15	115-120	Resistant to rust
8	L 4076	PL 234 x PL 639	ICAR-Indian Agricultural Research Institute, New Delhi, India	14-15	135-140	Resistant to rust
9	BM 4	ILL 5888 x ILL 5782	ICARDA Syria for BARI Bangladesh	20-23	116-120	Resistance to lentil rust and Stemphylium blight
10	K 75 (Malika)	Selection from Bundelkhand region	C.S. Ajad University of Agriculture and Technology, Kanpur, India	13-14	130-135	–
11	VL 520	DPL 15 x SEHORE 74-3	ICAR-Vivekananda Parvatiya Krishi Anusandhan Shansthan, Almora, India	14-15	118-120	Resistant to rust
12	PL7	L 4076 x DPL 15	G.B. Pant University of Agriculture and Technology, Pantnagar, India	16-18	125-145	Resistant to rust
13	PL 6	Pant L 4 x DPL 55	G.B. Pant University of Agriculture and Technology, Pantnagar, India	16-18	125-145	Resistant to rust
14	L 4717	ILL 7617 x 91516	ICAR-Indian Agricultural Research Institute, New Delhi, India	12-13	96-106	Resistant to wilt and AB
15	PL 406(IPL 406)	DPL 35 x EC 157634/382	G.B. Pant University of Agriculture and Technology, Pantnagar, India	13-14	120-155	Resistant to rust and wilt
16	PL 639	L 9-12 x T 8	G.B. Pant University of Agriculture and Technology, Pantnagar, India	20-22	140-150	Resistant to rust

TABLE 2 Descriptions regarding test locations in India (2018-2019) during Rabi (Oct-March).

S.No	Location Details	Biplot Name	Elevation (msl)	Latitude and longitude	Total Rain in season (mm)	RH (%)	Temp(O^C)		
							Max	Mini	Mean
1	ICAR- Indian Agricultural Research Institute (IARI) New Delhi	Delhi	235	28.65401 77.17172	48.32	77.36	25.87	14.89	20.38
2	ICAR- Indian Institute of Pulses Research (IIPR) Kanpur, Uttar Pradesh	Kanpur	130	26.49222 80.27682	49.18	75.66	25.38	15.26	20.32
3	Jawaharlal Nehru Krishi Vishwa Vidyalaya (JNKVV) Regional Agricultural Research Station (RARS), Sagar, Madhya Pradesh(MP)	Sagar	542	23.84083 78.74582	39.57	69.89	32.11	15.85	23.98
4	Rajmata Vijayaraje Scindia Krishi Vishwavidyalaya (RKVV) Sehore, Madhya Pradesh(MP)	Sehore	502	23.21235 77.08011	40.35	70.49	31.21	18.35	24.78
5	IARI-Regional Research Station(RAS) Samstipore, Bihar	Samstipore	47	25.84522 85.78377	42.41	79.65	31.47	15.60	23.53
6	Bihar Agricultural University (BAU) Sabour, Bihar	Sabour	87	32.803056 74.061389	51.1	78.42	30.52	14.30	22.41

Estimation of seed iron (Fe), zinc (Zn) and aluminum (Al) concentration

Physiologically matured seeds were plucked and dried in the shade. The seeds are given two ethanol rinses to remove dust particles. To avoid metal and dust contamination, 10 g of seeds from each entry were ground into a fine powder (approximately average diameter size of 10 microns (10^{-3} cm)) using a mortar and a pestle. The microwave

digestion apparatus (Anton Parr: Multiwave ECO) was used to process the 0.5 g sample of ground grain powder in line with the modified di-acid technique (Singh et al., 2005). Fe, Zn, and Al concentrations (in ppm) were measured using self-sampling techniques with inductively coupled plasma mass spectrometry (ICP-MS) (, Model: NexION 300, ICP-MS, manufactured in USA by Perkin Elmer inc.) Aluminum (Al) was measured at 167.000 nm, Fe at 238.204nm and Zn at 213.856 nm. ICP-MS has the lower detection limit can extend to parts per trillion (ppt), while the linear range of ICP-MS is 10-11 orders of magnitude. The kits for organic solvents used in as fallows Nitric acid (HNO_3) and hydrochloric acid (HCl) Sigma Aldrich, Merck KGaA, Darmstadt, Germany) were purified in perfluoralkoxy-polymer (PFA) sub-boiling units (DST-4000, Savillex corporation, Eden Prairie, MN 55344-3446 USA). Hydrogen peroxide solution (H_2O_2) Merck KGaA, Darmstadt, Germany) and tetrafluoroboric acid (HBF4) Sigma Aldrich, Merck KGaA, Darmstadt, Germany). Aluminium (Al) was identified as an indicator element in global research efforts. Fe and Zn quantification was not performed on samples that had an Al concentration of more than 5 ppm. These samples were re-washed with 70% ethanol to remove any dust contamination before being reanalyzed. Aluminum (Al) worked as an indicator element for possible potential dust contamination in this investigation as per HarvestPlus guidelines (Pfeiffer and McClafferty, 2007).

Determination of seed phytic acid (PA) content

Phytic acid (PA) was estimated using the K-PHYTA (Megazyme International) standard assay methodology as the phosphorus produced by phytase and alkaline. Inositol phosphates are used to extract the acid, which is then processed using phytases that are specific to phytic acid (IP6) and lower myo-inositol phosphate types (i.e. IP2, IP3, IP4, IP5). Alkaline phosphatase treatment causes the final phosphate, which is very resistant to phytase activity, to be released from myo-inositol phosphate (IP1). The total amount of phosphate released is calculated and expressed in grams of phosphorus per 100 g sample using a modified colorimetric method. A calibration curve with predetermined phosphorus content standards is used to convert Pi to phosphorus. Standard phosphorus concentration curve, Standard curve: y = 0.00461 + 0.16857x Linearity: R 2 = 0.99. Concentration range: standard assay procedure this corresponds to a phosphorus concentration of ~ 2.82 mg to ~ 11.29 mg/100 g (or phytic acid concentration of ~ 10 mg to ~ 40 mg/100 g). Finally, the percentage of phytic acid is computed on the presumptions that phytic acid accounts for all of the observed phosphorus and that 28.2% of phytic acid is present (Singh et al., 2017a).

$$\textbf{Phytic acid}\ (g/100g) = phosphorus(g/100g)/0.282$$

Construction of GGE biplot

The GGE biplot was designed depending on the first two principal components (PCs) produced *via* singular value decomposition (SVD) after computing each component of the

matrix using the suggested equation. (Yan et al., 2000; Yan and Falk, 2002, and Yan and Kang, 2022). The model used is given in Eq 1:

$$Y_{ij} = \mu + e_j + \sum_{(n-1)}^{N} \lambda_n \gamma_{in} \quad \delta_{jn} + \epsilon_{ij}$$

Where,

Y_{ij} = mean response of ith genotype (i = 1,…,I) in the jth environment (j = 1,.,J).

μ = grand mean.

e_j = environment deviations from the grand mean.

λ_n = the eigen value of PC analysis axis.

γ_{in} & δ_{jn} = genotype and environment PCs scores for axis n.

N = number of PCs retained in the model.

ϵ_{ij} = residual effect_ N (0,s2).

A "average environment coordination" (AEC) viewpoint of the GGE biplot has been constructed for genotype evaluation and stability determination, enabling genotype comparisons based on mean values of grain minerals (iron, zinc), phytic acid content, and stability between locations within a "mega-environment." (Yan, 2001; Yan and Rajcan, 2002). The average performance of the genotype was determined using a performance line that passed through the origin of the biplot. The performance line's arrow indicates a decline in genotype stability (Yan and Falk, 2002). The "ideal" test environment should be both genetically discriminating and representative of the "mega-environment," according to the "discriminating power vs. representativeness" viewpoint of the GGE biplot, which was developed for the evaluation of test environments (Yan et al., 2007). The "repeatability" of a test environment was assessed using the average rating of the genetic correlations across years within the settings for sustaining stability in genotypic performances (Yan et al., 2011). The AEC has also been used to create a "desirability index" for the test sites that considers the relationship between environmental factors and ideal genotype lengths as well as genotypic stability and adaptability (Yan and Holland, 2010). To evaluate the relationship between test sites and surrounding environments, angles within different location vectors were used (Yan and Kang, 2002). In order to determine genotype dominance across several testing scenarios and to combine testing environments into separate "mega environments," a "which won-where" GGE biplot viewpoint was also developed (Yan and Rajcan, 2002). Bootstrapping, a nonparametric re-sampling approach, was used to construct CL at the 95 percent level for each principal component value of both genotypes and environments in order to assess the validity of the GGE biplot (Yang et al., 2009).

Data analysis

Analysis of variance (ANOVA) was used to evaluate the effects of environments, genotypes, and their interactions across sites and for each individual genotype using mixed model analysis in R software. A combined analysis of variance (ANOVA) across locations was carried out after an error variance homogeneity test based on Bartlett's test. Stability was investigated using the AMMI and genotype + genotype x environment (GGE) models. The AMMI1 biplot was plotted using the mean of the main effect vs. the first interaction principal component (IPC1) score (Zobel et al., 1988). The ANOVA demonstrated how the variance distribution was impacted by genotypes, environment, and their interaction. The LSD test was used to calculate the mean significant difference between genotypes and environments at the P = 0.05 level of probability. A box plot was used to show how the mineral (Fe & Zn) and PA content varied among genotypes and locations. The Ward method was utilized to establish the hierarchical cluster that represented the genetic and environmental relatedness. Using R software (R Core Team, Vienna, Austria, https://www.R-project.org), the GGE biplot analysis was performed out.

Results

Analysis of variance

The study finds that several lentil genotypes responded differently with seed iron, zinc, and phytic acid. The pooled ANOVA showed that the genotypes under investigation were significantly influenced by genotype, environment, and genotype x environment interactions. For seed iron and zinc, the genotype and environment interaction produced a high estimate of the sum of squares (SS). On the other hand, phytic acid (PA) revealed more about genotype. The relative contribution of each source of variation to the total variation was estimated for seed iron (29.30%), zinc (40.99%), environment (61.85%), and GEI (18.06%) (Table 3). This showed an unexpected environmental influence on the mineral content of seed among genotypes tested in diverse locations. For seed iron, zinc, and phytic acid, genotype and genotype x environment interactions were significant across all genotypes examined at the various testing sites. Environmental variations showed that the habitats were unique, and they may explain a sizable portion of the variation in Fe, Zn, and PA. The illustrations for the biplot analysis were produced using these data. Biplot analysis was carried out and presented by plots to make distinctions between these environments, to assess stable and wide adaptive lines, and to assess the environments to determine whether a particular graph depicts the ideal environment to choose genotypes based on these parameters. Genotype Environment Interaction (GEI) was clearly evident in the AMMI 1 model when the interaction was divided among the first three Interaction Principal Component Axis (IPCA). Each and every PCA had statistically significant results (PCA 1, PCA 2, and PCA 3). Grain Fe, PC1 is responsible for 62.6 of the total variation. PC2 is in responsibility of 21.03 percent of the total variation, while IPC3 is in charge of 6.01 percent of the variation and has a Pr. F value above 0.005. PC1 and PC2 may be responsible for 83.63 percent of the variance in the Fe study. While PC1 and PC2 accounted for 86.35 and 79.76 percent, respectively, of the variation in Zn and PA (Figure 1).

The AMMI1 Biplots (Means vs PC1) indicate the genotypes DPL 62, L 4596, and L 4147 for iron content, JL 3, PL 406, and L 4147 for zinc content, and L 4596, PL7, and L 4147 for PA content (Figure 2). The pattern of mineral (Fe & Zn) and PA content

TABLE 3 Analysis of variance for seed iron, zinc and phytic acid content in lentil cultivars evaluated at six locations in India during (2018–2019).

Source of variation	Degrees of freedom DF	Seed Iron (Fe)			Seed Zinc (Zn)			Phytic Acid (PA)		
		SS	MSS	% TSS	SS	MSS	% TSS	SS	MSS	% TSS
ENV	5	13589.15	2717.83*** (0.000)	29.30	7296.073	1459.21*** (0.000)	20.10	4.17	0.83*** (0.000)	3.74
GEN	15	13774.91	918.33***(0.000)	29.70	6556.436	437.10***(0.000)	18.06	99.63	6.64*** (0.000)	93.00
GEN *ENV	75	19008.78	253.45***(0.000)	40.99	22454.46	299.39***(0.000)	61.85	7.81	0.10*** (0.000)	100.00
PC1	19	7581.14	399.01***(0.000)	39.88	10254.75	539.72***(0.000)	45.67	3.17	0.17*** (0.000)	40.52
PC2	17	4951.10	291.24***(0.000)	26.05	6535.867	384.46***(0.000)	29.11	2.36	0.14*** (0.000)	70.69
PC3	15	4080.40	272.03***(0.000)	21.47	2697.802	179.85***(0.000)	12.01	1.81	0.12*** (0.000)	93.81
PC4	13	1536.91	118.22***(0.000)	8.09	2107.734	162.13***(0.000)	9.39	0.33	0.03* (0.245)	98.00
PC5	11	859.22	78.11***(0.000)	4.52	858.3005	78.03** (0.007)	3.82	0.16	0.01*(0.275)	100.00
PC6	9	0.0	0.00 (1.00)	0.00	0	0.00 (1.00)	0.00	0.00	0.00 (1.00)	100.00
Residuals	192	1616.75	8.42	0.00	6669.495	34.74	0.00	4.68	0.02	0.00

* Significant at P ≤ 0.05 respectively; ** Significant at P ≤ 0.01 respectively; *** Significant at P ≤ 0.001respectively.

distribution among genotypes and locations was depicted using a box plot (Figure 3).

Character association analysis

The association between test locations and seed Fe, Zn, and PA was investigated using Spearman's correlation analysis (Supplementary Figure 1). When it came to seed Fe concentration, it was revealed that Delhi and Kanpur showed a positive significant association. When it related to seed zinc concentration, Kanpur and Sagar showed a negative correlation, but when it got to PA, all six places had a positive correlation. According to the significant correlation between Sehore and Sagar, the examined genotypes had a lot in common when it comes to seed iron and zinc levels. Spearman's correlation analysis was used to explore the relationship between seed iron, zinc, and PA. Seed Fe and seed Zinc levels were found to have a significant positive relationship (Supplementary Figure 2). Both attributes can be increased as a result of this relationship.

Evaluation of genotypes

Using a "AEC" perspective of the biplot, the genotype's average performance and consistency across places were graphically represented (Figure 4). The single arrow-head line on the graph known as "AEC abscissa," which crosses through biplot origin, indicates more seed iron. Seed iron concentration was higher in L

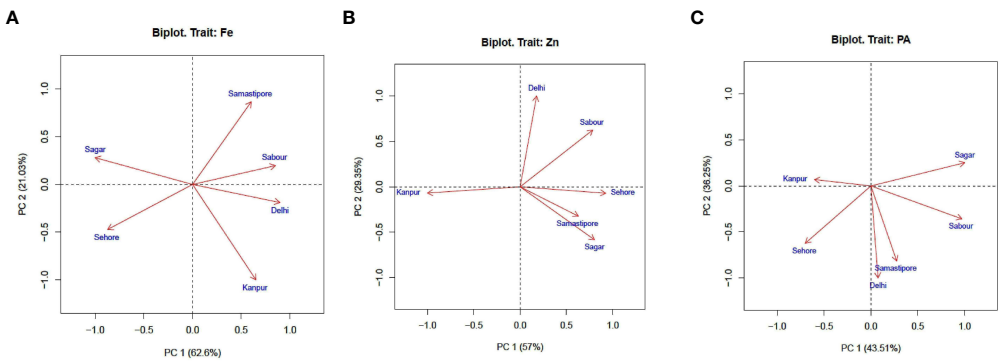


FIGURE 1
Principal component analysis (PCA) illustrating significant difference among test environments and for **(A)** Fe, **(B)** Zn, and **(C)** PA.

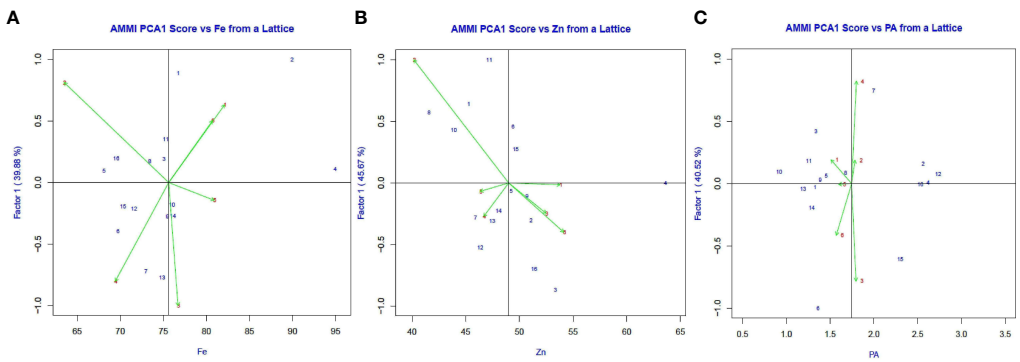


FIGURE 2
The AMMI1 Biplots (Means vs PC1) the first principal component (PC1) and mean values for **(A)** Fe, **(B)** Zn, and **(C)** PA in six environments.

4147 (Pusa Vaibhav) (4), L4596 (2), DPL 62 (Sheri) (1), and K 75 (Mallika) (10) types, as shown in Figure 4A. The length of a genotype's projection in absolute terms is commonly used to determine genotypic stability. The genotypes with the highest stability, i.e., a projection on AEC close to zero, and the highest seed iron content (a bigger negative projection on AEC) would be the best performers. As a result, the most "ideal" genotype was identified to be L 4076 (Pusa Shivalik) (8), with short projection from the "AEC abscissa" and optimal Iron levels. Genotypes that are more "desirable" are those that are closer to the "ideal" genotype. As a result, K-75 (10) and DPL 62(Sheri) (1) were designated as "desirable" genotypes because they were closer to the "ideal" genotype, with optimal iron and consistent performance.

The seed zinc concentration is higher in L4717 (Pusa Ageti) (14), L4596 (2), BM 4(9), DPL 58 (5), and PL 639 (16) (Figure 4B). AEC close to zero rated L4717 (Pusa Ageti) the "ideal genotype" for seed zinc concentration. L4596 (2), BM-4, and DPL 58 (5) were classified "ideal" genotypes because they were close to the "ideal" genotype, with optimum seed zinc concentration and consistent performance.

Figure 4C shows that seed phytic acid concentrations were higher in IPL 406 (15), PL639 (16), L4596 (2), and PL 7 (12). AEC found WBL 77 to be the "ideal genotype" for seed PA content because it was close to zero. IPL 406 (15), PL639 (16), L4596 (2), and PL 7 (12) were classified "ideal" genotypes because they were close to the "ideal" genotype, with optimum seed PA content and consistent performance.

Evaluation of the environments

Among the test locations for seed iron concentration, Kanpur had the longest environmental vector, followed by Sehore, Delhi, and Sabour, with Samastipur having the shortest projection (Figure 5A). As a result, Kanpur was chosen as having the most "discriminating locations" with potential for genetic discrimination.

Delhi had the longest environmental vector among the test locations for seed zinc concentration throughout the year (2018-19), followed by Kanpur and Sabour, and Samastipur had the shortest projection (Figure 5B). As a result, in terms of genotype discrimination power, Delhi was categorized as one of the most "discriminating environments."

During the year (2018-19), Delhi had the longest environmental vector for seed PA content, followed by Samastipur and Sehore,

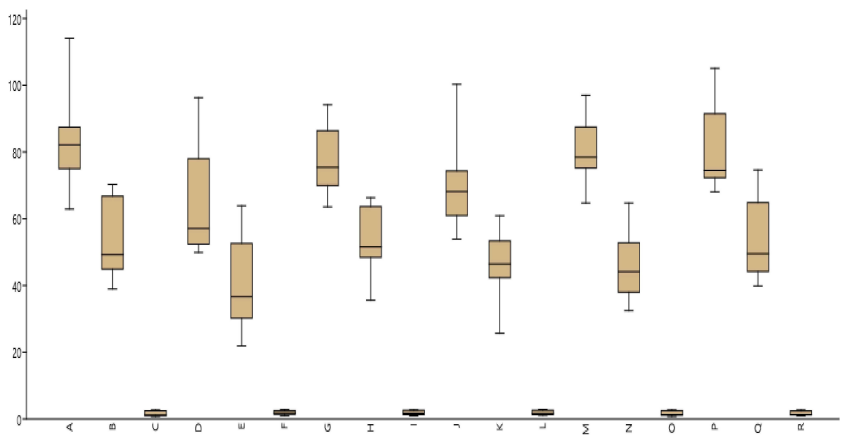


FIGURE 3
Box plot mean illustrating significant difference among test environments and for Fe, Zn and PA.

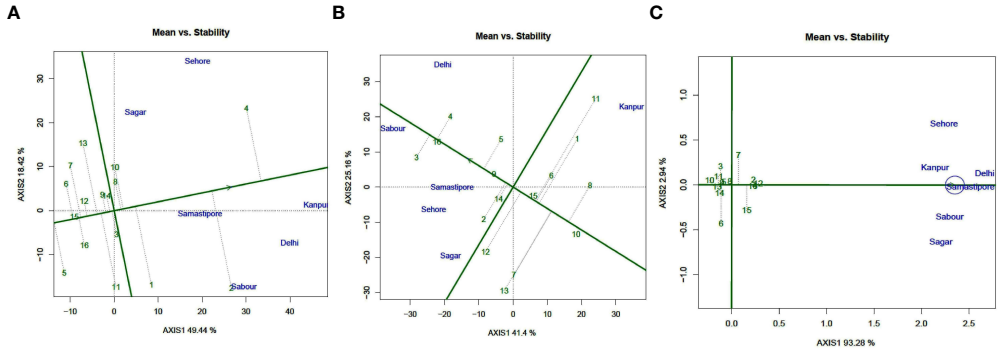


FIGURE 4
ean vs. Stability view of the GGE biplot of 16 lentil genotypes across 6 testing locations. **(A)** Fe, **(B)** Zn, and **(C)** PA). There was no transformation of data (transform = 0), and data were centered by means of the environments (centring = 2). The biplot was based on "row metric preserving." Numbers correspond to genotypes as listed in Table 1.

with Kanpur having the shortest projection (Figure 5C). Delhi was thus designated as one of the most "discriminating locations" in terms of genotype discrimination power. The solitary arrow-head line in the graph is labelled "AEC abscissa." The stronger the "representative" power of the place, the smaller the angle between the environment vectors and the "AEC abscissa."

Kanpur, followed by Samstipur, had the shortest angle with the AEC during the testing year, and were thus chosen as the most "Representative" test locations for seed iron concentration, whereas Sabour and Delhi were chosen as the most "Representative" test locations for seed zinc content. Samastipur was found to be the most "Representative" test location for seed PA content, followed by Delhi. Locations with high "discrimination" power but low "representativeness," such as Sehore and Sagar, should be investigated for finding stable genotypes for seed iron and zinc content.

Mega environments

GGE biplot employs a two-dimensional polygon visualization in the form of a "which won-where" polygon to detect genotypes for a certain production environment. Perpendicular lines were drawn from the biplot's origin to each side of the polygon to partition the biplot into numerous sectors, with one "winning" genotype placed at the polygon's vertex for each sector. L4147 (Pusa Vaibhav) (4) was revealed to have a substantially greater iron concentration and to be far from the origin, indicating that the performance was constant (Figure 6A). PL 6 (13), L4596 (2), L4717 (Pusa Ageti) (14), JL 3 (6), and WBL 77 (7) also had significant seed iron concentrations. DPL 58 (5), on the other hand, was identified downstream from the origin, exactly opposite L4147 (Pusa Vaibhav) (4), and was thus identified as the genotype with the lowest seed iron concentration. L4717(Pusa Ageti) (14) exhibited the most consistent performance of all the genotypes with moderate to medium iron content when placed near to the "AEC abscissa" with the least projection onto the "AEC ordinate." The equality lines divided the plot into seven pieces. These sectors could be labelled "Mega Environment," meaning that there is environmental unpredictability and G x E interaction.

VL 520 (11) had the highest mean zinc content and was far from the origin, indicating that its performance was consistent, according to this analysis (Figure 6B). Seed zinc content was also high in L 4076 (8), K 75 (Mallika) (8), IPL 321 (3), and L 4147 (Pusa Vaibhav)

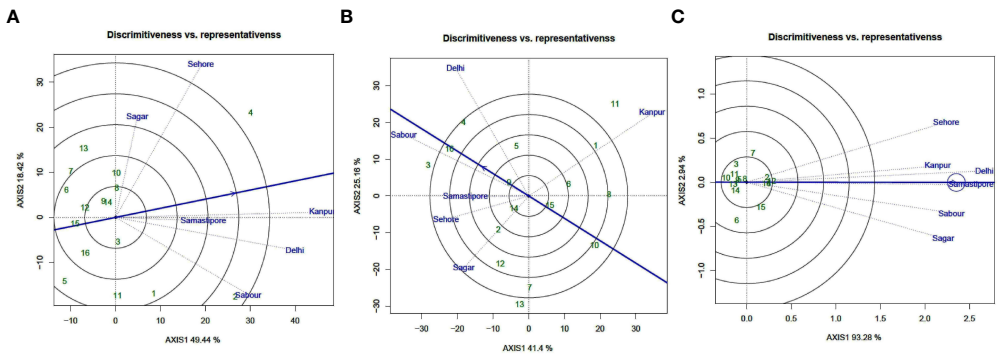


FIGURE 5
"Discrimitiveness vs. Representativeness" view of test locations based on GGE biplot of 16 lentil genotypes across 6 testing locations. **(A)** (Fe), **(B)** (Zn), and **(C)** (PA). There was no transformation of data (transform = 0), and data were centered by means of the environments (centring = 2). The biplot was based on "row metric preserving." Numbers correspond to genotypes as listed in Table 1.

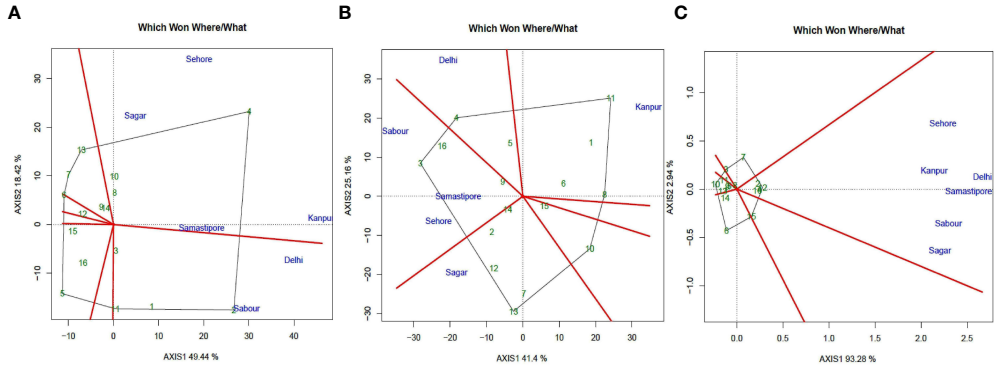


FIGURE 6
"Which-won-where" view of the GGE biplot of 16 lentil genotypes across 6 testing locations. **(A)** (Fe), **(B)** (Zn), and **(C)** (PA). There was no transformation of data (transform = 0), and data were centered by means of the environments (centring = 2). The biplot was based on "row metric preserving." Numbers correspond to genotypes as listed in Table 1. Locations are: A, Delhi; B, Kanpur; C, Sehore; D, Sagar; E, Sabour and F, Samstipore.

(4). PL 6 (13) was discovered downstream from the origin, right across from VL 520 (11), and was thus recognized as the low seed zinc content genotype. Sagar, Sehore, Samastipur, and Kanpur were single "mega environments" with different ecological features and genotypic reactions to seed iron concentration. The two "mega-environments" were located in Delhi and Sabour. The concentration of zinc in seeds was divided into four different "mega habitats." Kanpur is part of a single "mega environment," while Sabour, Samastipur, and Sehore are out of their own. The third and fourth "mega environments," respectively, are Delhi and Sagar. All of the locations were grouped into a single habitat for phytic acid content, showing that the feature was less varied.

GGE biplots by site regression (SREG) analysis

Figures 7A–C show GGE biplots for seed iron, zinc, and PA content obtained by the SREG model. The discriminating power of the sites was determined by their proximity to the origin of the vertices between PC1 and PC2, and the scores of the cultivars furthest from the origin were joined to construct a polygon. The polygon encompassed all other genotypes, indicating which cultivars were the most stable based on their correlation with site scores. The genotypes that made up the polygon were the most responsive to their environment and were reflective of the greatest or worst performance. The L 4076 and K75 for Fe, JL 3 and DPL 62 for Zn, and WBL 77 for PA were stable as they were in polygon of GGE and their values were close to zero on the Y-axis. Concentric circles rippling around the average environmental coordinate (AEC) of a genotype focused GGE biplots encompass genotypes that are relatively similar in their overall desirability. Based on this criterion L4147 (Pusa Vaibhav) for seed iron content, IPL 321, L 4147 and PL 639 for seed zinc content and all genotypes for PA were under the desirable genotypes for wider adaptation.

Genotype ranking based on their mean performance and stability

Using the average environment coordinate, ranking biplots were utilized to rate the genotypes according to their performance

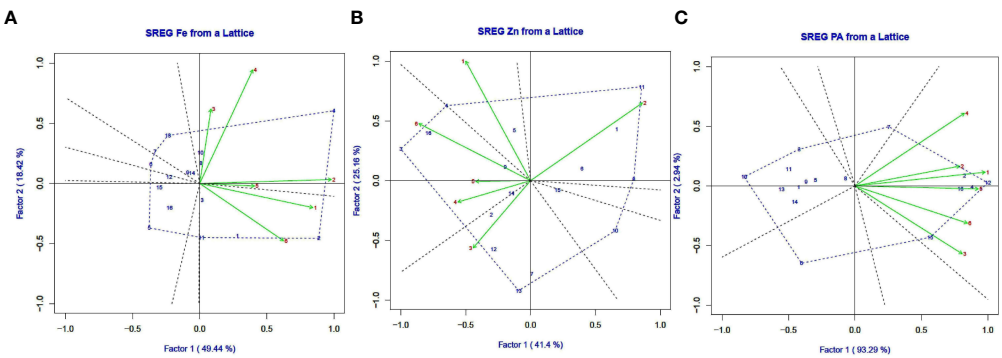


FIGURE 7
GGE biplots generated using SREG as an indication of seed iron content **(A)**, seed zinc content **(B)** and seed PA content **(A)** (Fe), **(B)** (Zn), and **(C)** (PA). for stability in lentil genotypes grown in six environments in India during 2018-2019. Genotype (G) and environment (E) codes are given in Tables 1 and 2, respectively.

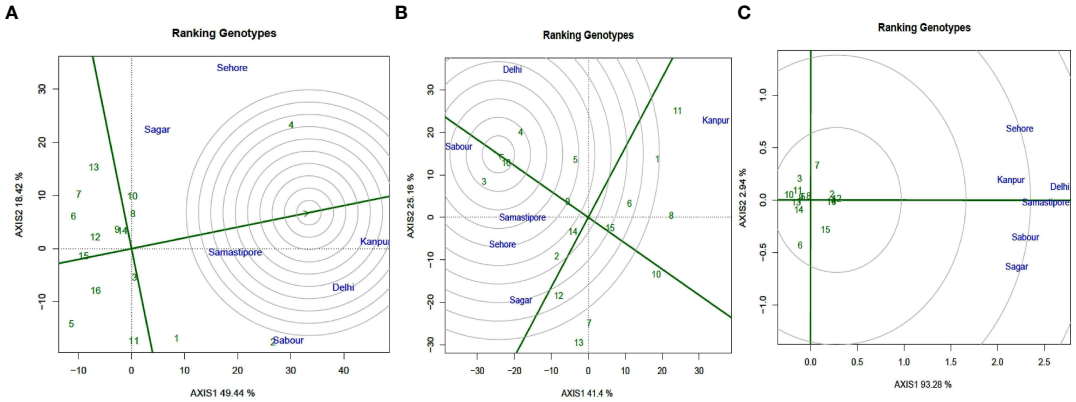


FIGURE 8
GGE-biplot based on genotype-focused scaling for comparison of the genotypes with the ideal genotype for **(A)** (Fe), **(B)** (Zn), and **(C)** (PA). Green and violet numbers stand for Genotypes and environments, respectively.

and stability (AEC). In the ranking biplot, an average environment axis (AEA) depicted by a single arrowhead line passing through the origin indicates that a genotype's mean performance is superior. The ranking biplot AEC revealed that genotypes L 4147 (4), L4076 (8), IPL 321 (3), and DPL 62(1) had high mean Fe content and genotypes DPL 58 (5), BM-4 (9) and Pl 639 (16) had high mean Zn content in this study. Genotypes DPL 58 (5) and PL 7 (12) exhibited the lowest Fe and Zn levels, respectively (Figure 8). In PA, the majority of genotypes were close to the AEC, but WBL 77 (7) and JL 3 (6) were far away. The length of the vector between the genotype positions and the AEA in ranking biplot was used to assess genotype stability. Genotypes that are remote from the origin but on the AEA or near to it have the best performance and stability. As a result, L 4147 was the most stable genotype for Fe, while IPL 321(3), L 4147 (4), and PL639 were the most stable genotypes for PA, with a high mean and shorter vector from AEA.

Discussion

Substantial genetic variations for grain mineral concentration

Lentil grains typically contain higher Fe and Zn attributable to breeding since genotypes have substantial genetic variation. According to other investigations (Thavarajah et al., 2009; Thavarajah et al., 2011; Karakoy et al., 2012; Sen Gupta et al., 2013), the lentil gene pool comprises a wide range of genetic diversity for these micronutrients. Iron levels in red and green lentil genotypes ranged from 43 to 132 mg/kg, while zinc levels ranged from 22 to 78 mg/kg, according to an investigation of 1,600 genotypes of lentils (Sarker et al., 2007). In a multi-location, multi-year experiment in Saskatoon, Canada (Thavarajah et al., 2009), significant genetic diversity was found in 19 lentil genotypes for grain iron (73-90 mg/kg) and zinc (44-54 mg/kg) levels. In 1,000 core collection common bean germplasm at the CIAT (International Center for Tropical Agriculture), the levels of zinc (21 to 54 mg/kg) and iron (34 to 89 mg/kg) demonstrated a wide range of variability (Welch & Graham, 2002; Welch, 2002).

The Fe, Zn, and phytic acid contents between lentil cultivars were clearly distinguishable, according to the current study. Across the study locations, Fe and Zn levels varied from 114.10 to 49.90 mg/kg and 74.62 to 21.90 mg/kg, respectively. The range of phytic acid concentrations in present investigations ranged 0.76 to 2.84 g/100g (dw) (Table 4). While, lentils had phytic acid contents of 0.40 to 1.29 g/100 g, 0.43 to 0.77 g/100 g in pea (*Pisum sativum*), and 1.17 to 1.70 g/100 g in soybean (*Glycine max*) (Vojtíšková et al., 2010), chickpeas 0.28–1.60 g/100 g, kidney beans 0.61–2.38 g/100 g, (Lehrfeld, 1994) and peanuts 0.17–4.47 g/100g (Venkatachalam and Sathe, 2006). Cereal crops, such as maize germ (6.39 g/100g), wheat bran (2.1–7.3 g/100g) (Harland et al., 1986), and rice bran (2.56–8.7 g/100g) (Kasim and Edwards, 1998), have a much higher range than pulses. Plant breeding and molecular approaches have convincingly demonstrated that lentil crops can be used as bio-fortified crops Thavarajah et al., 2011).

Genotype, genotype × environment interaction

The levels of seed iron and zinc in nineteen lentil lines produced over a two-year period in eight locations in Saskatchewan, Canada, demonstrated that genotype x location interactions had a significant effect primarily on zinc content, but not on iron content (Thavarajah et al., 2009). In wheat, genotype x location interactions were found to be significant for Zn and Fe levels in both wild and modified cultivars (Gomez-Becerra et al., 2010). In the instance of durum wheat, 46 genotypes were tested for Fe, Zn, and phytic acid content in two habitats, and the genotype, environment, and their interaction revealed highly significant impacts (G x E). The effect of the environment was particularly strong in the case of phytic acid: Fe ratio and phytic acid (Magallanes-López et al., 2017). The genotype, environment, and genotype x environment interactions for iron, zinc, and phytic acid

TABLE 4 Mean of seed iron, zinc and phytic acid in 16 genotypes of lentil at six locations during 2018–2019.

S.No	Delhi			Kanpur			Sagar			Sehore			Samastipore			Sabour			Mean		
	Fe	Zn	PA	Fe	Zn	PA	Fe	Zn	PA	Fe	Zn	PA	Fe	Zn	PA	Fe	Zn	PA	Fe	Zn	PA
1	82.14	54.70	1.08	80.73	53.42	1.41	63.55	35.56	1.47	65.26	46.19	1.48	70.38	34.96	1.27	98.21	46.81	1.23	76.71	45.27	1.33
2	114.10	47.62	2.49	80.01	36.66	2.52	76.57	66.29	2.65	66.77	42.31	2.82	97.02	48.70	2.47	105.10	64.88	2.42	89.93	51.08	2.56
3	87.43	66.35	1.03	66.32	21.90	1.83	75.43	51.57	1.31	60.93	53.45	1.61	87.46	52.19	1.18	72.88	74.62	1.06	75.07	53.35	1.34
4	96.72	70.29	2.57	96.30	53.93	2.79	87.71	64.24	2.62	100.29	60.91	2.61	94.19	64.15	2.57	94.46	68.17	2.54	94.94	63.61	2.62
5	64.68	66.76	1.25	57.09	38.29	1.49	72.00	48.42	1.69	53.90	48.40	1.75	75.21	35.91	1.25	85.79	57.58	1.29	68.11	49.23	1.45
6	65.80	53.17	0.94	56.89	52.60	1.49	73.42	53.91	2.11	71.65	45.61	1.21	78.35	43.49	0.94	72.28	47.57	1.47	69.73	49.39	1.36
7	62.89	42.44	1.96	50.61	32.53	1.96	86.37	64.98	1.96	75.11	55.78	2.64	76.10	32.46	1.79	86.78	47.26	1.67	72.98	45.91	2.00
8	82.96	47.42	1.52	77.98	47.43	1.79	75.43	46.78	1.82	71.38	25.65	1.89	64.69	38.02	1.52	68.06	44.18	1.52	73.42	41.58	1.68
9	89.80	67.73	1.15	55.48	34.73	1.78	86.51	52.69	1.45	65.27	52.43	1.42	83.24	52.81	1.26	72.46	43.86	1.25	75.46	50.71	1.39
10	82.66	38.94	0.76	68.21	47.15	0.98	79.71	49.32	0.95	73.95	45.78	1.13	82.02	42.21	0.76	69.36	39.88	0.93	75.98	43.88	0.92
11	83.40	58.32	0.92	60.11	63.92	1.79	74.11	38.15	1.24	55.68	32.65	1.30	86.90	40.03	1.30	91.46	50.01	0.98	75.28	47.18	1.26
12	81.54	44.92	2.71	51.39	25.16	2.71	69.93	54.50	2.71	73.70	53.40	2.84	78.46	45.55	2.71	74.46	54.51	2.71	71.58	46.34	2.73
13	80.23	38.96	1.06	49.90	30.18	1.06	94.13	63.68	1.28	74.35	50.73	1.35	74.96	57.39	1.18	75.48	43.81	1.22	74.84	47.46	1.19
14	79.15	49.23	1.16	55.46	35.26	1.16	76.32	51.68	1.43	74.65	46.43	1.34	88.48	44.13	1.34	82.53	61.61	1.31	76.10	48.06	1.29
15	75.02	46.69	2.20	52.40	46.25	2.23	69.47	49.44	2.61	68.14	41.20	2.01	82.71	64.72	2.55	74.11	49.54	2.20	70.31	49.64	2.30
16	86.33	68.24	2.46	57.86	24.25	2.68	67.76	49.61	2.55	60.69	47.07	2.52	75.45	45.97	2.54	69.05	72.99	2.43	69.53	51.35	2.53
Mean	82.18	53.86	1.58	63.55	40.23	1.85	76.78	52.55	1.87	69.48	46.75	1.87	80.98	46.42	1.66	80.78	54.21	1.64	75.62	49.00	1.75
Variance	156.45	121.39	0.48	184.21	150.80	0.35	67.84	79.09	0.36	113.71	74.62	0.38	72.61	95.30	0.45	133.79	123.47	0.37	51.01	24.28	0.37
SD(E)	3.13	2.75	0.17	3.39	3.07	0.15	2.06	2.22	0.15	2.67	2.16	0.15	2.13	2.44	0.17	2.89	2.78	0.15	1.79	1.23	0.15
SD(D)	12.51	11.02	0.69	13.57	12.28	0.59	8.24	8.89	0.60	10.66	8.64	0.62	8.52	9.76	0.67	11.57	11.11	0.61	7.14	4.93	0.61
CV	15.22	20.46	43.89	21.36	30.53	31.87	10.73	16.92	32.23	15.35	18.48	33.03	10.52	21.03	40.13	14.32	20.50	37.10	9.44	10.05	34.73

10.3389/fpls.2023.1102879

concentration were also quite significant in the current analysis (Tables 3), implying that the environmental conditions as stated in Table 2 are important. Micronutrient mobility from root to seed is likely to be influenced by growth seasons, in addition to soil conditions. These data reveal that crop mineral properties are influenced by both heredity and environmental factors. To improve mineral content through breeding, it is vital to examine the location of certain environmental circumstances as well as the genetic make-up of the genotype.

The current study aimed to shed light on the impact of environmental and genotype-by-environment interactions on the responsiveness of lentil genotypes to nutritional and anti-nutritional attributes. The incoherent response of genotypes and locations across sites revealed the impact of the environment on the volatility of the parameters investigated. Quantitative traits control the traits under consideration. Many genes with lesser, comparable, and cumulative effects influence the expression of quantitative features. The genotypes, as well as the interaction of all of these variables with genotypes, contributed to superior/inferior performance and genotype stability across sites. The genotype with lesser effect of G X E interactions performs stable way in the expression of the traits. These mineral-rich stable cultivars grown in the ideal environment would not only increase lentil output but also productivity among marginal and small farmers. Furthermore, a large number of genotypes, many sites, and multiple years will produce robust results. In addition to the aforementioned points, *in vivo* and *in vitro* studies in lentils must be conducted so that the benefit of the high Fe method can be substantiated as one of the best approaches in biofortification initiatives.

Role of broad sense heritability (H^2) in mineral bio-fortification

Heritability is critical in the genetic improvement of quantitatively inherited characteristics through selection. Estimates of trait heritability distinguish the amount of total phenotypic variance caused by genotypes and environmental factors, and they tell us how much response may be obtained by selecting any plant population over the initial genetic pool (Lynch and Walsh, 1998). Understanding trait heredity is crucial for plant breeders. In our study, the broad sense heritability (H^2) for Fe (0.94 to 0.99) and Zn (0.53 to 0.99) content was moderate to high, indicating that genotypic impacts across contexts accounted for a large percentage of the variability in the character (Table 5). However, strong estimates of broad sense heritability (H^2) were found in phytic acid levels across regions (0.96 to 0.99 percent). The genotypes analyzed have genetic potential, as evidenced by moderately high to high levels of heritability in our tested samples. As a result, it can be utilized to develop lentil cultivars with increased iron and zinc content while lowering phytic acid levels. Despite the fact that both genotype and genotype x environment interactions accounted for a significant amount of overall phenotypic variance, a sufficient fraction of genetic variability is proven to be heritable.

In lentils, previous research found moderate heredity for Fe concentration, but low to relatively high heritability for Zn content (Thavarajah et al., 2009; Kumar et al., 2018). Very high estimates of broad-sense heritability (h^2bs) for Fe and Zn concentration were found in black gram (*Vigna mungo* (L) Hepper), whereas lower heredity in phytic acid indicated a substantial environmental influence (Singh J. et al., 2017). Broad-sense heritability (h^2bs) for both grain Fe and Zn ranged from moderate to high values in pearl millet (Velu et al., 2007; Gupta et al., 2009).

Grain iron and zinc can be increased simultaneously

Positive trait relationships encourage breeders to simultaneous improvement of two or more traits. Our findings demonstrated that iron and zinc levels had a strong and favorable association. However, both minerals had a favorable but non-significant relationship with phytic acid (Figure 6). The concentration of iron revealed a non-significant and positive connection with the content of zinc in the lentil studies mentioned before (Karakoy et al., 2012; Kumar et al., 2018). Fe content in black gram was shown to have a high positive correlation with Zn content. There was a clear association between phytic acid concentration and the minerals (Fe and Zn) (Singh et al., 2017a). Rice (Kabir et al., 2003; Inabangan-Asilo et al., 2019), wheat (Gomez-Becerra et al., 2010), maize (Long et al., 2004; Mallikarjuna et al., 2015), pearl

TABLE 5 Calculations heritability estimates based on the geography for grain iron (Fe), zinc (Zn), and phytic acid (PA).

S.No	Environment	Board sense heritability		
		Fe	Zn	PA
1	Delhi	0.9824	0.9912	0.9939
2	Kanpur	0.9902	0.9925	0.9770
3	Sabour	0.9695	0.5318	0.9819
4	Sagar	0.9855	0.9884	0.9868
5	Samastipore	0.9481	0.9897	0.9948
6	Sehore	0.9874	0.9901	0.9600
	Mean	**0.9772**	**0.9140**	**0.9824**

The bold values are the mean values of the respective traits.

millet (Pucher et al., 2014), and sorghum (Reddy et al., 2005) have all been found to exhibit positive relationships between these two minerals. These findings show that simultaneous selection for high iron and zinc levels in particular crops is possible. Correlations between seed iron sites in Delhi and Kanpur, Sehore and Sagar were positive and significant. In terms of seed zinc concentration, there was a substantial negative association between Kanpur and Sagar, but a large positive correlation between Sagar and Sehore. Grain mineral micronutrients (Zn and Fe) between two locations were highly and positively associated in milled and brown rice (Bollinedi et al., 2020).

Genetic improvement has previously increased the concentration of iron and zinc in crops such as wheat, rice, and common bean (Welch and Graham, 2002; Welch, 2002). Our findings revealed a wide range of Fe and Zn content genetic variability that can be leveraged to create nutritionally dense Fe and Zn lentil cultivars. It may thus be a feasible method for treating micronutrient deficiency in human beings who consume lentils on a daily basis. The frequency of the association varies based on the situation. These findings imply that both genetic and environmental factors influence mineral association. Mineral association in grain may have a genetic the basis due to mineral transporter genes co-segregating in genotypes and/or the availability of common transporters for many minerals (Schaaf et al., 2005).

Ideal and desirable genotypes

Plant breeders tend to identify genotypes that have the least interacting influence with a broad adaptation environment in their extensive plant breeding program. Multi-environmental studies (Kang, 2002) can uncover minor geographical characteristics with consistent performance across sites, as well as small temporal variables with consistency over years. In the GGE biplot's "Mean vs. Stability" view, the "AEC ordinates" show a larger GE interaction effect in both directions and poor stability (Yan and Tinker, 2006), whereas the vector projection of the genotype to the "AEC abscissa" indicates mean performance (Yan & Falk, 2002) [34]. In addition, in the current study, K-75 (10), as well as DPL 62 (Sheri) (1), were identified as "desirable" genotypes, and were found to be closer to the ideal genotype, L4076 (Pusa Shivalik) (8). For Zn concentration, L4717 (Pusa Ageti) was deemed the "ideal genotype," whereas L4596 (2), BM-4, and DPL 58 (5) were deemed "desirable." For phytic acid content, WBL 77 was deemed "ideal," while IPL 406 (15), PL639 (16), L4596 (2), and PL 7 (12) were deemed "desirable." Those "ideal" genotypes had more mineral content, indicating robust stability (Yan et al., 2007), with higher negative projection on the ATC abscissa and less projection on AEC ordinates. These techniques have been used to successfully identify stable genotypes in chickpea (Erdemci, 2018; Misra et al., 2020), mungbean (Ullah et al., 2011), lentil (Darai et al., 2020), soybean (Mwiinga, 2018), faba bean (Fikere et al., 2008; Tekalign et al., 2017), and maize (Beyene et al., 2011; Tonk et al., 2011). The "mega environment" can be successfully depicted using GGE biplot methods in a "which-won-where" approach (Gauch and Zobel, 1997; Yan and Kang, 2002; Yan et al., 2007). The goal of mega-environment identification is to grasp the region's complicated GEI pattern in order to exploit specific adaptability and increase selection response (Yan et al., 2011). Earlier studies defined a "mega environment" as a collection of places with consistent genotypic responses (Yan et al., 2000; Yan and Rajcan, 2002; Yan & Tinker, 2006). Many investigations, including lentil (Singh et al., 2017a; Jeberson et al., 2019), chickpea (Erdemci, 2018), uradbean (Gupta et al., 2020b), mungbean (Asfaw et al., 2012), pea(Rana et al., 2020), pigeonpea (Kumar et al., 2021), and soybean (El-Harty et al., 2018), used these methodologies to depict mega environments.

In the current study, the genotypic response to grain minerals and phytic acid content was shown to be identical in each "mega environment" tested. It is critical to control the synchrony of study locations and convergent breeding activities in a location-specific manner in order to improve the precision of lentil bio-fortification. The goal of this study was to find out more about how environmental and genotype-by-genotype interactions influence lentil genotype responses to grain mineral and phytic acid concentrations.

Lentil as bio-fortification tool

The inconsistent response of genotypes and locations to the environmental influence on mineral and phytic acid content reflected the environmental effect on mineral and phytic acid content. "Ideal" and "desirable" genotypes for grain iron and zinc content were successfully discriminated against in our study. Not only the stable cultivars like K-75 (for Fe) and L4596 (for Zn) but also "desirable" genotypes with consistent performance like L4076 (Pusa Shivalik) (for Fe) and L4717 (Pusa Ageti) (for Zn) genotypes were recommended for use. In terms of determining the levels of phytic acid content and its stability across locations, our study adds to the current knowledge. The decreased inhibitor concentration, such as phytic acid, will improve the bioavailability of grain minerals in legumes. Furthermore, genetic variation in iron and zinc concentrations can aid in the identification of genes/quantitative trait loci (QTL) linked to iron and zinc consumption and accumulation. Furthermore, a genetic examination of iron and zinc levels in seeds demonstrated the impact of environmental variables. Thus, location testing or region-specific breeding can aid in the generation of lentil varieties that are high in iron and zinc.

The most prevalent problem among Asian and African women and preschool children is iron and zinc deficiency. For males and women, the RDA for iron is 8 mg/day and 18 mg/day, respectively, whereas the RDA for zinc is 11 mg/day for men and 8 mg/day for women aged 19 and up (https:/ods.od.nih.gov/professional/factsheets/IronHealth). The lentil genotypes studied were able to deliver a significant amount of RDA for Iron and Zinc in our study. For example, genotypes of L4147 (Pusa Vaibhav) had the highest average Fe content in their seeds (94.94 mg/kg), which could be enough to supply 168.91 and 136.24 percent of RDA Fe intake for adult males and females, respectively. The same line L4147 (Pusa

Vaibhav) contained 63.61 mg/kg of seed Zn, which is sufficient to give 141.02 and 155.31 percent of RDA (in case of Zn) for adult men and females, respectively.

The iron-rich nature of lentil variety L4147 (Pusa Vaibhav) has been established by numerous earlier investigations (Kumar et al., 2014 and Kumar et al., 2019). Our findings also show that suitable and stable mineral-rich lentils like L 4076 and L4717 can be used as donors for further mapping and molecular analysis. The decreased phytic acid content of these discovered types naturally increases the bioavailability of grain micronutrients in poor people's diets. These cultivars are critical trait donors for future mapping and tagging investigations. They have the added benefit of being able to directly release or notify other zones, improving lentil yield and productivity, because they are newly released cultivars. Transcriptomics studies using these lines could provide insight into the paths for grain mineral absorption, transport, and storage in lentils and other pulses. The mapping populations developed through these parents make it much easier to find the genes and QTLs involved in grain mineral uptake, transportation, and regulation. The proposed trait-specific desirable genotypes, as well as large environments like Sagar and Sehore, will revolutionize lentil cultivation by enhancing productivity and production. Specific labeling and marketing methods must be developed in order to popularize bio-fortified crops. Direct production will be profitable, and immediate inclusion in the normal diet through the public distribution system (PDS) will boost micronutrient consumption in poor families, reducing micronutrient deficiency. In order to address the issue of hidden hunger, investigations on the bioavailability of these plant-based Fe and Zn must be investigated.

Conclusions

- The grain iron (Fe), grain zinc (Zn), and grain phytic acid (PA) concentrations in commercially cultivated lentil genotypes showed significant genetic variations in different locations.
- The environment (E) and G x E (Genotype x Environment interactions) had an impact on the concentration of grain Fe, Zn, and phytic acid (PA).
- Our research identified strong positive correlation between the contents of Fe and Zn, a strategy for simultaneously increasing Fe and Zn in lentils may be recommended.
- In addition, our study found that the stable and ideal lentil varieties L4076 (Pusa Shivalik) for Fe concentration and L4717 (Pusa Ageti) for Zn content, with lower phytic acid contents, will not only play a crucial role as stable donors in lentil bio-fortification but will also enable the expansion of bio-fortified crops to achieve health and nutrition security.
- The lentil genotypes identified in our study were able to deliver a significant amount of Recommended Dietary Allowance (RDA) for Iron and Zinc.
- In case of genotypes of L4147 (Pusa Vaibhav) had the highest average Fe content in their seeds (94.94 mg/kg), which could be enough to supply 168.91 and 136.24 percent of RDA Fe intake for adult males and females, respectively.
- The same line L4147 (Pusa Vaibhav) contained 63.61 mg/kg of seed Zn, which is sufficient to give 141.02 and 155.31 percent of RDA (in case of Zn) for adult men and females, respectively.
- The ideal and stable mineral-rich lentils such as L 4076 and L4717 can serve as donors for further mapping and molecular dissection.
- Direct production of L 4147 (Pusa Vaibhav) and L 4717 not only profitable, but direct inclusion in the normal diet through the public distribution system (PDS) will boost micronutrient consumption in poor families, reducing micronutrient deficiency.
- Furthermore, a large number of genotypes, more umber of environments, and many years will yield reliable data. In addition to this, *in vivo* and *in vitro* studies in lentils are required to validate the high Fe method as one of the best approaches in biofortification initiatives.

Data availability statement

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

Author contributions

MA, HD, GM, PY, and JT planned and designed the research. MA, TD, AK, VK, AP, and AS performed the experiment. VR, P, TB, KT, and PG helped in data recording. MA, NR, RB, AK, SG, JK, and AP prepared the manuscript. MA, SK, PY, and VK edited the manuscript for publication. All authors contributed to the article and approved the submitted version.

Funding

This research work was carried out from the grants received from DST, SERB (ECR/2017/003178). The funds from DST were utilized for field, lab and phenotyping work.

Acknowledgments

The authors are thankful to DST, SERB, and ECR for financial support to this research study. The authors acknowledge the facilities and guidance provided by the Head, Division of Genetics, Joint Director Research and Director, ICAR-IARI. The authors are grateful to the Head, Division of Plant Physiology and Division of Genetics for help and guidance during the investigations.

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.2023.1102879/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Pearson's correlation between six test locations for seed Fe, Zn, and phytic acid during 2018-19. _P < 0.05 are boxed. Locations are: A: Delhi, B Kanpur; C: Sehore; D: Sagar; E: Sabour and F: Samstipore. (p <0.05 are boxed).

SUPPLEMENTARY FIGURE 2
Pearson's correlation between seed Fe, Zn, and phytic acid A: Iron (Fe); B: Zinc (Zn), C: Phytic Acid (PA) (p <0.05 are boxed).

References

Asfaw, A., Gurum, F., Alemayehu, F., and Rezene, Y. (2012). Analysis of multi-environment grain yield trials in mung bean *Vigna radiate* (L.) wilczek based on GGE bipot in southern Ethiopia. *J. Agric. Sci. Tech.* 14, 389–398.

Beyene, Y., Mugo, S., Mutinda, C., Tefera, T., Karaya, H., Ajanga, S., et al. (2011). Genotype by environment interactions and yield stability of stem borer resistant maize hybrids in Kenya. *Afr. J. Biotechnol.* 10 (23), 4752–4758. doi: 10.5897/AJB10.2710

Bollinedi, H., Yadav, A. K., Vinod, K. K., Krishnan, S. G., Bhowmick, P. K., Nagarajan, M., et al. (2020). Genome-wide association study reveals novel marker-trait associations (MTAs) governing the localization of fe and zn in the rice grain. *Front. Genet.* 11. doi: 10.3389/fgene.2020.00213

Cappellini, M. D., Musallam, K. M., and Taher, A. T. (2019). Iron deficiency anaemia revisited. *J. Intern. Med.* 287, 153–170. doi: 10.1111/joim.13004

Chasapis, C. T., Ntoupa, P. A., Spiliopoulou, C. A., and Stefanodou, M. E. (2020). Recent aspects of the effects of zinc on human health. *Arch. Toxicol.* 94, 1443–1460. doi: 10.1007/s00204-020-02702-9

Darai, R., Dhakal, K. H., Sarker, A., Pandey, M. P., Agrawal, S. K., Ghimire, S. K., et al. (2020). Effect of genotype by environment interaction (GEI), correlation, and GGE biplot analysis for high concentration of grain iron and zinc biofortified lentils and their agronomic traits in multi-environment domains of Nepal. *Fields Interests* 6 (6), 24. doi: 10.5281/zenodo.3931172

Ebdon, J. S., and Gauch, H. G. (2002a). Additive main effect and multiplicative interaction analysis of national turfgrass performance trials i. interpretation of genotype x environment interaction. *Crop Sci.* 42, 489–496. doi: 10.2135/cropsci2002.4890

Ebdon, J. S., and Gauch, H. G. (2002b). Additive main effect and multiplicative interaction analysis of national turfgrass performance trials II. cultivar recommendations. *Crop Sci.* 42, 497–506. doi: 10.2135/cropsci2002.4970

El-Harty, E. H., Alghamdi, S. S., Khan, M. A., Migdadi, H. M., and Farooq, M. (2018). Adaptability and stability analysis of different soybean genotypes using biplot model. *Int. J. Agric. Biol.* 20 (10), 2196–2202. doi: 10.17957/IJAB/15.0760

Erdemci, I. (2018). Investigation of genotype× environment interaction in chickpea genotypes using AMMI and GGE biplot analysis. *Turkish J. Field Crops.* 23 (1), 20–26. doi: 10.17557/tjfc.414846

FAOSTAT. Available at: http://www.fao.org/faostat/en/#data/QC (Accessed 5 June 2021).

Fava, C., Piepoli, M., and Villani, G. Q. (2019). Heart failure and iron deficiency. *G. Ital. Cardiol.* 20, 126–135. doi: 10.1714/3108.30962

Fikere, M., Tadesse, T., and Letta, T. (2008). Genotype-environment interactions and stability parameters for grain yield of faba bean (*Vicia faba* l.) genotypes grown in south Eastern Ethiopia. *Int. J. Sustain. Crop Prod.* 3 (6), 80–87.

Fredlund, K., Rossander-Hulthe, N. L., Isaksson, M., Almgren, A., and Sandberg, A. S. (2002). Absorption of zinc and calcium: dosedependent inhibition by phytate. *J. Appl. Microbiol.* 93, 197–204. doi: 10.1046/j.1365-2672.2002.01676.x

Garg, M., Sharma, N., Sharma, S., Kapoor, P., Kumar, A., Chunduri, V., et al. (2018). Biofortified crops generated by breeding, agronomy, and transgenic approaches are improving lives of millions of people around the world. *Front. Nutr.* 5, 12. doi: 10.3389/fnut.2018.00012

Gauch, H. G.Jr., and Zobel, R. W. (1997). Identifying mega-environments and targeting genotypes. *Crop Sci.* 37 (2), 311–326. doi: 10.2135/cropsci1997.0011183X003700020002x

Gomez-Becerra, H. F., Yazici, A., Ozturk, L., Budak, H., Peleg, Z., Morgounov, A., et al. (2010). Genetic variation and environmental stability of grain mineral nutrient concentrations in triticum dicoccoides under five environments. *Euphytica* 171 (1), 39–52. doi: 10.1007/s10681-009-9987-3

Gonmei, Z., and Toteja, G. S. (2018). Micronutrient status of Indian population. *Indian J. Med. Res.* 148, 511–521. doi: 10.4103/ijmr.IJMR_1768_18

Grüngreiff, K., Gottstein, T., and Reinhold, D. (2020). Zinc deficiency-an independent risk factor in the pathogenesis of haemorrhagic stroke? *Nutrients* 12, 3548. doi: 10.3390/nu12113548

Gupta, S., Brazier, A. K. M., and Lowe, N. M. (2020a). Zinc deficiency in low- and middle-income countries: prevalence and approaches for mitigation. *J. Hum. Nutri. Diet.* 33, 624–643. doi: 10.1111/jhn.12791

Gupta, S., Singh, A., Dikshit, H. K., Aski, M., Mishra, G. P., and Kumar, J. (2018). Assessment of total phenol content, total flavonoid content and anti-oxidant capacity in exotic lentil germplasm. *Chem. Sci. Rev. Lett.* 7, 459–463.

Gupta, D. S., Singh, U., Kumar, J., Shivay, Y. S., Dutta, A., Sharanagat, V. S., et al. (2020b). Estimation and multi-variate analysis of iron and zinc concentration in a diverse panel of urdbean (*Vigna mungo* l. hepper) genotypes grown under differing soil conditions. *J. Food Composition Anal.* 93, 103605. doi: 10.1016/j.jfca.2020.103605

Gupta, S. K., Velu, G., Rai, K. N., and Sumalini, K. (2009). Association of grain iron and zinc content with grain yield and other traits in pearl millet (*Pennisetum glaucum* (L.) r. br.). *Crop Improvement* 36 (2), 4–7. doi: 10.1186/2193-1801-3-763

Harland, B. F., Oberleas, D., Ellis, R., Gelroth, J., Gordon., D., Phillips, K., et al. (1986). Anion-exchange method for determination of phytate in foods: collaborative study. *J. Assoc. Off. Analytical Chem.* 69 (4), 667–670. doi: 10.1093/jaoac/69.4.667

https:/ods.od.nih.gov/professional/factsheets/IronHealth. Available at: https:/ods.od.nih.gov/professional/factsheets/IronHealth.

Inabangan-Asilo, M. A., Swamy, B. P. M., Amparado, A. F., Descalsota-Empleo, G. I. L., Arocena, E. C., and Reinke, R. (2019). Stability and G × E analysis of zinc-biofortified rice genotypes evaluated in diverse environments. *Euphytica* 215, 61. doi: 10.1007/s10681-019-2384-7

Jamnok, J., Sanchaisuriya, K., Sanchaisuriya, P., Fucharoen, G., Fucharoen, S., and Ahmed, F. (2020). Factors associated with anaemia and iron deficiency among women of reproductive age in northeast Thailand: a cross-sectional study. *BMC Public Health* 20, 102. doi: 10.1186/s12889-020-8248-1

Jeberson, M. S., Shashidhar, K. S., Wani, S. H., Singh, A. K., and Dar, S. A. (2019). Identification of stable lentil (Lens culinaris medik) genotypes through GGE biplotand AMMI analysis for north hill zone of India. *Legume Res.: Int. J.* 42 (4), 467–472. doi: 10.18805/LR-3901

Kabir, K. A., Haque, M., Hossain, M. A., Dipti, S. S., and Tetens, I. (2003). Breeding for iron-dense rice in bangladesh. rice science: innovations and impact for livelihood. Metro Manila. In T. W. Mew, D. S. Brar, S. Peng, D. Dawe and B. Hardy *Proceedings of the International Rice Research Conference*, Beijing, China, 16-19 September 2002. International Rice Research Institute (IRRI), 397–402.

Kang M. S. (2002). *Quantitative genetics, genomics and plant breeding*. Cabi international Publishing, 219. doi: 10.1079/9780851996011.0221

Karakoy, T., Erdem, H., Baloch, F. S., Toklu, F., Eker, S., Kilian, B., et al. (2012). Diversity of macro- and micronutrients in the seeds of lentil landraces. *Sci. World J.* 2012, 1–9. doi: 10.1100/2012/710412

Kasim, A. B., and Edwards, H. M.Jr. (1998). The analysis for inositol phosphate forms in feed ingredients. *J. Sci. Food Agric.* 76 (1), 1–9. doi: 10.1002/(SICI)1097-0010(199801)76:1<1::AID-JSFA922>3.0.CO;2-9

Khazaei, H., Subedi, M., Nickerson, M., Martínez-Villaluenga, C., Frias, J., and Vandenberg, A. (2019). Seed protein of lentils: current status, progress, and food applications. *Foods* 8, 391. doi: 10.3390/foods8090391

Kumar, H., Dikshit, H. K., Singh, A., Jain, N., Kumari, J., Singh, A. M., et al. (2014). Characterization of grain iron and zinc in lentil ('Lens culinaris' medikus' culinaris') and analysis of their genetic diversity using SSR markers. *Aust. J. Crop Sci.* 8 (7), 1005.

Kumar, J., Kumar, S., Sarker, A., and Singh, N. P. (2018). Analysis of genetic variability and genotype × environment interactions for iron and zinc content among diverse genotypes of lentil. *J. Food Sci. Technol.* 55 (9), 3592–3605. doi: 10.1007/s13197-018-3285-9

Kumar, M. N., Ramya, V., Sameer Kumar, C. V., Raju, T., Sunil Kumar, N. M., and Seshu, G. (2021). Identification of pigeonpea genotypes with wider adaptability to rainfed environments through AMMI and GGE biplot analyses. *Indian J. Genet.* 81 (1), 63–73. doi: 10.31742/IJGPB.81.1.7

Kumar, H., Singh, A., Dikshit, H. K., Mishra, G. P., Aski, M., Meena, M. C., et al. (2019). Genetic dissection of grain iron and zinc concentrations in lentil (*Lens culinaris* medik.). *J. Genet.* 98 (3), 66. doi: 10.1007/s12041-019-1112-3

Lehrfeld, J. (1994). HPLC separation and quantitation of phytic acid and some inositol phosphates in foods: problems and solutions. *J. Agric. Food Chem.* 42 (12), 2726–2731. doi: 10.1021/jf00048a015

Long, J. K., Bänziger, M., and Smith, M. E. (2004). Diallel analysis of grain iron and zinc density in southern African-adapted maize inbreds. *Crop Sci.* 44 (6), 2019–2026. doi: 10.2135/cropsci2004.2019

Lynch, S. R., Beard, J. L., Dassenko, S. A., and Cook, J. D. (1984). Iron absorption from legumes in humans. *Am. J. Clin. Nutr.* 40, 42–47. doi: 10.1093/ajcn/40.1.42

Lynch, M., and Walsh, B. (1998). *Genetics and analysis of quantitative traits* (Sunderland, MA: Sinauer).

Magallanes-López, A. M., Hernandez-Espinosa, N., Velu, G., Posadas-Romano, G., Ordoñez-Villegas, V. M., Crossa, J., et al. (2017). Variability in iron, zinc and phytic acid content in a worldwide collection of commercial durum wheat cultivars and the effect of reduced irrigation on these traits. *Food Chem.* 15, 237:499–505.

Mallikarjuna, M. G., Thirunavukkarasu, N., Hossain, F., Bhat, J. S., Jha, S. K., Rathore, A., et al. (2015). Correction: stability performance of inductively coupled plasma mass spectrometry-phenotyped kernel minerals concentration and grain yield in maize in different agro-climatic zones. *PloS One* 10 (10), e0140947. doi: 10.1371/journal.pone.0140947

Misra, G., Joshi-Saha, A., Salaskar, D., Reddy, K. S., Dixit, G. P., Srivastava, A. K., et al. (2020). Baseline status and effect of genotype, environment and genotype × environment interactions on iron and zinc content in Indian chickpeas (Cicer arietinum l.). *Euphytica* 216 (9), 1–6. doi: 10.1007/s10681-020-02673-z

Mwiinga, B. (2018). *Genotypey by environment interaction, genetic variability and path analysis for grain yield in elite soybean [Glycine max (L.) Merrill] lines 2018 (Doctoral dissertation)*. (South Africa: University of KwaZulu-Natal Pietermaritzburg) PP-1–123.

Ning, S., and Zeller, M. P. (2019). Management of iron deficiency. *Hematol. Am. Soc Hematol. Educ. Program.* 1, 315–322. doi: 10.1182/hematology.2019000034

Ozer, S., Karakoy, T., Toklu, F., Baloch, F. S., Kilian, B., and Ozkan, H. (2010). Nutritional and physicochemical variation in Turkish *kabuli* chickpea (*Cicer arietinum* L.) landraces. *Euphytica* 175, 237–249. doi: 10.1007/s10681-010-0174-3

Pfeiffer, W. H., and McClafferty, B. (2007). Biofortification: breeding micronutrient-dense crops. *Breed. Major Food Staples* pp, 61–91. doi: 10.1002/9780470376447.ch3

Pucher, A., Høgh-Jensen, H., Gondah, J., Hash, C. T., and Haussmann, B. I. (2014). Micronutrient density and stability in West African pearl millet–potential for biofortification. *Crop Sci.* 54 (4), 1709–1720. doi: 10.2135/cropsci2013.11.0744

Rana, C., Sharma, A., Sharma, K. C., Mittal, P., Sinha, B. N., Sharma, V. K., et al. (2020). Stability analysis of garden pea (*Pisum sativum* l.) genotypes under north Western Himalayas using joint regression analysis and GGE biplots. *Genet. Resour. Crop Evol.* 19, 1–2. doi: 10.1007/s10722-020-01040-0

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

Sandberg, A. S. (2002). Bioavailability of minerals in legumes. *Br. J. Nutr.* 3, S281–S285. doi: 10.1079/BJN/2002718

Sandberg, A. S., Brune, M., Carlsson, N. G., Hallberg, L., Skoglund, E., and Rossander-Hulthe, N. L. (1999). Inositol phosphates with different number of phosphate groups influence iron absorption in humans. *Am. J. Clin. Nutr.* 70, 240–246. doi: 10.1093/ajcn.70.2.240

Sarker, A., El-Askhar, F., Uddin, M. J., Million, E., Yadav, N. K., Dahan, R., et al. (2007). "Lentil improvement for nutritional security in the developing world," in *merican Society of Agronomy (ASA), Crop Science Society of America (CSSA) – Soil Science Society of America (SSSA)*. (New Orleans, LA, USA: International Annual Meetings), 4–8.

Schaaf, G., Schikora, A., Haberle, J., Vert, G., Ludewig, U., Briat, J. F., et al. (2005). A putative function for the arabidopsis fe-phytosiderophore transporter homolog AtYSL2 in fe and zn homeostasis. *Plant Cell Physiol.* 46, 762–774. doi: 10.1093/pcp/pci081

Sen Gupta, D., Thavarajah, D., Thavarajah, P., McGee, R., Coyne, C. J., and Kumar, S. (2013). Lentils (*Lens culinaris* l.), a rich source of folates. *J. Agric. Food Chem.* 61, 7794–7799. doi: 10.1021/jf401891p

Singh, D., Chonkar, P. K., and Dwivedi, B. S. (2005). *Manual on soil, plant and water analysis* (New Delhi: Westville Publishers).

Singh, J., Kanaujia, R., Srivastava, A. K., Dixit, G. P., and Singh, N. P. (2017). Genetic variability for iron and zinc as well as antinutrients affecting bioavailability in black gram (*Vigna mungo* (L.) hepper). *J. Food Sci. Technol.* 54 (4), 1035–1042. doi: 10.1007/s13197-017-2548-1

Singh, A., Sharma, V. K., Dikshit, H. K., Singh, D., Aski, M., Prakash, P., et al. (2017a). Microsatellite marker-based genetic diversity analysis of elite lentil lines differing in grain iron and zinc concentration. *J. Plant Biochem. Biotechnol.* 26 (2), 199–207. doi: 10.1007/s13562-016-0382-6

Singh, A., Sharma, V., Meena, J., Gupta, S., Dikshit, H. K., Aski, M., et al. (2017b). Genetic variability for grain iron and zinc concentration in lentil. *Chem. Sci. Rev. Lett.* 6, 1327–1332.

Taleb, M. H., Khodambashi, M., and Karimi, M. (2013). Study of physical and nutritional quality properties in segregating generations of lentil cross. *Int. J. Agric. Crop Sci.* 5, 2740–2742.

Tekalign, A., Sibiya, J., Derera, J., and Fikre, A. (2017). Analysis of genotype× environment interaction and stability for grain yield and chocolate spot (*Botrytis fabae*) disease resistance in faba bean (*Vicia faba*). *Aust. J. Crop Sci.* 11 (10), 1228–1235. doi: 10.21475/ajcs.17.11.10.pne413

Thavarajah, D., Thavarajah, P., Sarker, A., and Vandenberg, A. (2009). Lentils (*Lens culinaris* Medikus subspecies *culinaris*): a whole food for increased iron and zinc intake. *J. Agric. Food Chem.* 57, 5413–5419. doi: 10.1021/jf900786e

Thavarajah, D., Thavarajah, P., Wejesuriya, A., Rutzke, M., Glahn, R. P., Combs, J. G. F., et al. (2011). The potential of lentil *(Lens culinaris* l.) as a whole food for increased selenium, iron, and zinc intake: preliminary results from a 3 year study. *Euphytica* 180, 123–128. doi: 10.1007/s10681-011-0365-6

Tonk, F. A., Ilker, E., and Tosun, M. (2011). Evaluation of genotype x environment interactions in maize hybrids using GGE biplot analysis. *Crop Breed. Appl. Biotechnol.* 11, 01–09. doi: 10.1590/S1984-70332011000100001

Ullah, H., Khalil, I. H., Khalil, I. A., and Khattak, G. S. (2011). Performance of mungbean genotypes evaluated in multi environmental trials using the GGE biplot method. *Atlas J. Biotechnol.* 1 (1), 1–8. doi: 10.5147/ajbtch.2011.0024

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 Breeding.* 126 (2), 182–185. doi: 10.1111/j.1439-0523.2007.01322.x

Venkatachalam, M., and Sathe, S. K. (2006). Chemical composition of selected edible nut seeds. *J. Agric. Food Chem.* 54 (13), 4705–4714. doi: 10.1021/jf0606959

Vojtišková, P., Kráčmar, S., and Hoza, I. (2010). "Content of phytic acid in selected sorts of legumes," in *Mendel University in Brno Acta Universitatis Agriculturae et Silviculturae Mendelianae Brunensis*. (Brno, Czech Republic: Mendel University)

Welch, R. M. (2002). The impact of mineral nutrients in food crops on global human health. *Plant Soil.* 247 (1), 83–90. doi: 10.1023/A:1021140122921

Welch, R. M., and Graham, R. D. (2002). "Breeding crops for enhanced micronutrient content," in *Food security in nutrient-stressed environments: exploiting plants' genetic capabilities* (Dordrecht: Springer), 267–276.

Yadava, D. K., Choudhury, P. R., Hossain, F., Kumar, D., Sharma, T. R., and Mohapatra, T. (2020). *Biofortified varieties: sustainable way to alleviate malnutrition* (New Delhi: Indian Council of Agricultural Research), 19.

Yan, W. (2001). GGEbiplot–a windows application for graphical analysis of multienvironment trial data and other types of two-way data. *Agron. J.* 93, 1111–1118. doi: 10.2134/agronj2001.9351111x

Yan, W., and Falk, D. E. (2002). Biplot analysis of host-by-pathogen data. *Plant Dis.* 86 (12), 1396–1401. doi: 10.1094/PDIS.2002.86.12.1396

Yan, W., and Holland, J. B. (2010). A heritability-adjusted GGE biplot for test environment evaluation. *Euphytica* 171, 355–369. doi: 10.1007/s10681-009-0030-5

Yan, W., Hunt, L. A., Sheng, Q., and Szlavnics, Z. (2000). Cultivar evaluation and mega-environment investigation based on the GGE biplot. *Crop Sci.* 40 (3), 597–605. doi: 10.2135/cropsci2000.403597x

Yan, W., and Kang, M. S. (2002). *GGE biplot analysis: a graphical tool for breeders, geneticists, and agronomists* (Boca Raton, Florida USA: CRC press).

Yan, W., Kang, M. S., Ma, B., Woods, S., and Cornelius, P. L. (2007). GGE biplot vs. AMMI analysis of genotype-by-environment data. *Crop Sci.* 47, 643–653. doi: 10.2135/cropsci2006.06.0374

Yan, W., Pageau, D., Frégeau-Reid, J., and Durand, J. (2011). Assessing the representativeness and repeatability of test locations for genotype evaluation. *Crop Sci.* 51 (4), 1603–1610. doi: 10.2135/cropsci2011.01.0016

Yan, W., and Rajcan, I. (2002). Biplot evaluation of test locations and trait relations for breeding superior soybean cultivars in Ontario. *Crop Sci.* 42 (1), 11–20. doi: 10.2135/cropsci2002.1100

Yan, W., and Tinker, N. A. (2006). Biplot analysis of multi-environment trial data: principles and applications. *Can. J. Plant Sci.* 86 (3), 623–645. doi: 10.4141/P05-169

Yang, R. C., Crossa, J., Cornelius, P. L., and Burgueño, J. (2009). Biplot analysis of genotype _ environment interaction: proceed with caution. *Crop Sci.* 49, 1564–1576. doi: 10.2135/cropsci2008.11.0665

Zobel, R. W., Wright, M. J., and Gauch, H. G. (1988). Statistical analysis of a yield trial. *Agron. J.* 80, 388–393. doi: 10.2134/agronj1988.00021962008000030002x

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 13 July 2023
DOI 10.3389/fnut.2023.1180225

Check for updates

OPEN ACCESS

EDITED BY
Sapna Langyan,
Indian Council of Agricultural Research (ICAR), India

REVIEWED BY
Saumendu Deb Roy,
Mata Gujri University, India
Pavan Kumar,
Guru Angad Dev Veterinary and Animal Sciences University, India
Indrani Chandra,
University of Burdwan, India

*CORRESPONDENCE
Gauri Saxena
gaurigupta72@yahoo.com

RECEIVED 05 March 2023
ACCEPTED 25 May 2023
PUBLISHED 13 July 2023

CITATION
Srivastava RP, Kumar S, Singh L, Madhukar M, Singh N, Saxena G, Pandey S, Singh A, Devkota HP, Verma PC, Shiva S, Malik S and Rustagi S (2023) Major phenolic compounds, antioxidant, antimicrobial, and cytotoxic activities of *Selinum carvifolia* (L.) collected from different altitudes in India. *Front. Nutr.* 10:1180225.
doi: 10.3389/fnut.2023.1180225

COPYRIGHT
© 2023 Srivastava, Kumar, Singh, Madhukar, Singh, Saxena, Pandey, Singh, Devkota, Verma, Shiva, Malik and Rustagi. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Major phenolic compounds, antioxidant, antimicrobial, and cytotoxic activities of *Selinum carvifolia* (L.) collected from different altitudes in India

Ravi Prakash Srivastava[1], Sachin Kumar[2], Lav Singh[3], Mayank Madhukar[4], Nitesh Singh[5], Gauri Saxena[1]*, Shivaraman Pandey[1], Arpit Singh[1], Hari Prasad Devkota[6], Praveen C. Verma[7], Shatrughan Shiva[7], Sumira Malik[8] and Sarvesh Rustagi[9]

[1]Department of Botany, University of Lucknow, Lucknow, Uttar Pradesh, India, [2]Food, Drug and Chemical Toxicology Group, CSIR-Indian Institute of Toxicology Research (CSIR-IITR), Lucknow, Uttar Pradesh, India, [3]Forest Training Institute, Ministry of Environment, Forest and Climate Change, Govt. of Uttar Pradesh, Kanpur, India, [4]PG Department of Zoology, RD and DJ College, Munger University, Bihar, India, [5]Faculty of Agricultural Sciences, Shree Guru Gobind Singh Tricentenary University, Gurugram, Haryana, India, [6]Graduate School of Pharmaceutical Sciences, Kumamoto University, Kumamoto, Japan, [7]Plant Molecular Biology and Genetic Engineering Laboratory, Council of Scientific and Industrial Research, National Botanical Research Institute (CSIR-NBRI), Lucknow, Uttar Pradesh, India, [8]Amity Institute of Biotechnology, Amity University, Ranchi, Jharkhand, India, [9]School of Applied and Life Sciences, Uttaranchal University, Dehradun, Uttarakhand, India

Antibiotic resistance poses a serious threat to public health, raising the number of diseases in the community. Recent research has shown that plant-derived phenolic compounds have strong antimicrobial, antifungal, and cytotoxic properties against a variety of microorganisms and work as great antioxidants in such treatments. The goal of the current work is to evaluate the anticancerous, antibacterial, antifungal, antioxidant, and cytotoxicity activities in the extracts of the different plant parts (leaves, stems, and roots) of *S. carvifolia* (L.) L. This is a medicinally important plant and has been used for different kinds of diseases and ailments such as hysteria and seizures. The phenolic compounds from the different plant parts were analyzed using HPLC and the following were found to be present: chlorogenic acid, gallic acid, rutin, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid. Gallic acid was found to have the highest concentration (13.93 mg/g), while chlorogenic acid (0.25 mg/g) had the lowest. The maximum TPC value, which ranged from 33.79 to 57.95 mg GAE/g dry extract weight, was found in the stem. Root extract with 9.4 mg RE/g had the greatest TFC level. In the leaf and stem extracts, the RSC ranged from 0.747 mg/mL to 0.734 mg/1 mL GE/g dry extract weight, respectively. The 2,2-diphenyl-1-picrylhydrazyl (DPPH) assay was used to measure *in vitro* antioxidant activity. In a concentration-dependent way, promising antioxidant activity was reported. Moreover, 3,5-dinitrosalicylic acid (DNSA) and the Folin–Ciocalteu phenol reagent technique were used to determine reducing sugar content and total phenolic content, respectively. Antibacterial activity against eight strains (MIC: 250–1,000 μg/mL) was analyzed, and the stem extract exhibited maximum activity. Antifungal activity was also assessed, and potent activity was reported especially in the extract obtained from the stem. Cytotoxicity was evaluated using an MTT assay in the A549 cell line, where different doses (0.0625, 0.125, 0.25, 0.5, and 1 mg/mL) of leaf, root, and stem extracts

were used. Treatment with these extracts reduced the cell viability, indicating that *S. carvifolia* may possess anticancer potential, which can be of great therapeutic value.

KEYWORDS

antimicrobial, antioxidant, HPLC, MTT assay, pharmacology, *Selinum carvifolia*

Introduction

A serious concern to the public's health is the emergence of pathogenic bacteria and fungi that become resistant to synthetic antibacterial treatments over a period of time. This may result in an increased load of several antibiotic resistant bacteria and illnesses caused by them (1).

Most of the conventional antibiotics currently available for purchase have significant adverse effects on patients and have been responsible for the emergence of multiple drug resistance in pathogenic bacteria (2). Previous research has shown that bactericidal antibiotics such as quinolones and aminoglycosides have negative side effects, while lactams cause mitochondrial dysfunction and excessive production of reactive oxygen species (ROS) in mammalian cells, resulting in oxidative damage to DNA, proteins, and membrane lipids (3). Free radicals have also been linked to a number of diseases, including ischemic heart disease, diabetes, atherosclerosis, cancer, inflammation, and aging (4).

The defense against free radicals and harmful microbes is greatly aided by secondary plant metabolites such as phenols, flavonoids, and terpenoids (5). In order to tackle these resistant infections while avoiding or limiting the side effects associated with the consumption of synthetic antibiotics, there has been increasing interest in the identification of novel natural antimicrobial agents (6). As a result, combining antioxidant therapy with antibiotic medication appears to be a strategy to reduce or avoid these side effects. Plant-derived medicines cause few side effects, and they have a long history of use in folk medicine for the treatment of infectious diseases and oxidative stress conditions (7, 8).

The use of antioxidant and antibacterial drugs together helps to boost their antioxidant and antibacterial capabilities. It has also been observed that phenols, flavonoids, and terpenoids can increase the sensitivity of various bacteria to particular antibiotics (9–11).

The antibacterial and antioxidant properties of natural compounds from higher plants may prove to be a new source with a novel mechanism of action (12). As a result, there are three different levels of interaction at play: interaction with the outside cellular components, engagement with the cytoplasmic membrane, and interaction with cytoplasmic components. To exercise their antibacterial effects, natural compounds might interact with bacterial cells on one level or all three levels.

They may find innovative active principles to circumvent resistance mechanisms in multidrug-resistant microbes as a result of their thorough and systematic screening (13).

Many medicinally important plants can be found in the high altitude, alpine, and sub-alpine regions of Uttarakhand, Himachal Pradesh, and the surrounding areas of the Western Himalayas in India. The richness of medicinal plants in the Himalayas reflects its distinct geological, topographical, ecological, climatic, and physiographic position (14, 15). These plants have ethnobotanical significance and are often utilized in traditional medicine (16, 17). Medicinal plants are known to contain specific bioactive compounds that can inhibit microbial growth in the environment; therefore, they play a significant role in the creation of effective therapeutic treatments (18–20). Pharmacological industries have created a large number of antibiotics (21–23). Researchers have been investigating the antibacterial activity of medicinal plants as a result of the acceptance of traditional medicine as an alternative form of healthcare and the development of microbial resistance to current antibiotics (24–26).

Recent research has revealed that the antioxidant activity of plant products is mostly due to the presence of phenolic components including flavonoids, phenolic acids, tannins, and other phenolic compounds (27–29). Studies report that excessive free radicals are responsible for various pathological conditions such as asthma, arthritis, inflammation, neuro-degeneration, Parkinson's disease, and diabetes. Natural antioxidants have become the focus of a multitude of studies aimed at identifying sources of potentially safe, effective, and low-cost antioxidants (30, 31). Herbal medications that include free radical scavengers have been shown to have therapeutic value (32). To avoid the oxidation of the vulnerable substrate, plants create an astonishing array of antioxidant molecules such as carotenoids, flavonoids, ascorbic acid, and others. Antioxidants are commonly used in the food industry to prevent lipid peroxidation (33, 34). Plants include different iso-quinoline alkaloids that have antibacterial, anticancer, anti-inflammatory, adrenolytic, sympatholytic, and anti-acetylcholine esterase effects, according to chemical and pharmacological research (35, 36). The most efficient method for quantifying phenolic and flavonoid chemicals is HPLC analysis (37).

Selinum carvifolia plants have shown useful therapeutic effects, which are mostly dependent on the presence of phenols, flavonoids, and terpenoids. This is a perennial herb native to the high altitudes of Uttarakhand, Himachal Pradesh, and the neighboring regions of Indian Western Himalayas (38). This important species grows in humus-rich mountainous regions of the Himalayas between 2,200–4,000 m in states of Himachal Pradesh, Uttarakhand, and Sikkim in India and in neighboring countries such as Nepal and China (39). In India, the genus *Selinum*, also known as Bhutkeshi, is used as an insecticide, a nervine sedative with anti-spasmodic stimulating effects, and in the cure of constipation, menstruation, and digestion, among other things (40). A paste made from the leaves of *S. carvifolia* has been ethno-medically used for wound healing for centuries. The smoke produced from burning the roots of *S. carvifolia* is reported to repel insects, and the root decoction contains antimicrobial properties that can be used to cure coughs, colds, fevers, and body problems (41–43). However, due to their indiscriminate usage, the plants now need to be grown *in vitro* for their cultivation and preservation (44).

Previous studies on *Selinum* mainly focused on the volatile constituents of the different plant parts (45–49), but there are hardly any detailed reports on the non-volatile constituents in the leaves, stems, and roots of *Selinum carvifolium*. Recently, two novel compounds, bhutkesoside A (1) and bhutkesoside B, and 10 known compounds from the roots of the same plant have been reported from the leaves of *Ligusticopsis wallichianum* (50). The present study, therefore, aimed at quantifying and identifying the phytochemical constituents of the leaves, stems, and roots of *S. carvifolia* collected from three different altitudes (2,150, 2,593, and 3,178 m) in the Chopta region of Uttarakhand, India, and probing their antioxidant activity, reducing sugar capacity (RSC), and antibacterial and cytotoxicity properties in order to establish the therapeutic potential of this plant.

Materials and methods

Collection of plant material

The samples were collected in September 2018 from three different altitudes: 2,150 m (latitude 22°21′12.4″N75°12′01.4″E), 2,593 m (latitude 25°23′14.9”N75°11′01.3″E), and 3,178 m (latitude 28°24′25.9″N 75°10′57.6″E). Five to six whole plants were collected from each altitude in the Chopta region of Uttarakhand, India. A GPS positioning system was used for geographic coordinates, and the elevation information of each measurement point was recorded. The plants were identified by Dr. L. B. Chaudhary, Principal Scientist, Herbarium (Angiosperm Taxonomy), National Botanical Research Institute, Lucknow, India, and the samples were deposited in the NBRI repository as voucher specimen LWG 105543.

HPLC

Analysis sample preparation

Ten grams of dry weight of raw material (roots, leaves, and stems) were powdered and poured into a conical flask. One hundred milliliters of 70% MeOH was added to it, and the extract thus formed was left for 48 h. Fractionation was performed in different solvents in the order of hexane (Hx), petroleum ether (PTE), and chloroform (Chl), as per the extraction methods of Abubakar and Haque (51). The residue was returned to the conical flask after the extracts from various plant sections were filtered through Whatman filter paper No.1 into a 250 mL volumetric flask. The extraction technique was performed two more times, with 70% methanol used to make up the difference in volume. Before the HPLC injection samples could be analyzed using a standardized HPLC procedure, each sample solution was filtered through a 0.2 μm membrane filter into a HPLC sample vial.

Instrumentation and conditions

The principal ingredients of all *S. carvifolia* extracts were analyzed by HPLC (Shimadzu, Japan) with PDA SPD M 20. A 20 μL sample loop was included with the LC-20 AD dual pump system and SIL-20 AC auto-injector (with cooler). Compounds were separated on a Shimadzu RP-C18 column with an internal diameter of 250 × 4.6 mm and a pore size of 5 m, which was protected by a guard column with the same packing. Subsequent elution of the column with hexane and a 2% ethyl acetate/hexane mixture yielded 100 mg of pure compounds infraction. The crude extract of leaf, stem, and roots was fractionated by 70% MeOH for 48 h before being injected three times with a 20 μL sample loop, and this ran for 25 min. Shimadzu Lab solution software was used to combine the data with the detection of peaks at 510 nm, and the findings were obtained by comparing them to accessible standards. For the detection of blank peaks, the plain mobile phase was employed as a control.

Total phenolic content

The total phenolic content (TPC) was determined from the leaf, stem, and roots of *S. carvifolia* extracts from different altitudes (2,150, 2,593, and 3,178 m) using the Folin–Ciocalteu method (52). Briefly, 1 mL of extract (100–500 μg/mL) solution was mixed with 2.5 mL of 10% (w/v) Folin–Ciocalteu reagent. After 5 min, 2.0 mL of Na_2CO_3 (75%) was subsequently added to the mixture and incubated at 50°C for 10 min with intermittent agitation. Afterwards, the sample was cooled, and the absorbance was measured utilizing a UV spectrophotometer (Shimazu, UV-1800) at 765 nm against a blank without extract. The outcome data were expressed as mg/g of gallic acid equivalents in milligrams per gram (mg GAE/g) of dry extract.

Total flavonoid content

The flavonoid contents of the leaf, stem, and roots of *S. carvifolia* extracts from different altitudes (2,150; 2,593; 3,178 m) were measured as per the aluminum chloride (AlCl3) assay (colorimetry). An aliquot of 1 mL of extract solution (25–200 μg/mL) or rutin (25–200 μg/mL) was mixed with 0.2 mL of 10% (w/v) $AlCl_3$ solution in methanol, 0.2 mL (1 M) potassium acetate, and 5.6 mL distilled water. The mixture was incubated for 30 min at room temperature followed by the measurement of absorbance at 415 nm against the blank. The outcome data were expressed as mg/g of quercetin equivalents in milligrams per gram (mg RE/g) of dry extract (53, 54).

Statistical analysis

The data were reported as the mean ± standard deviation. The linear regression coefficient (R^2) for phenolic and flavonoid content with antioxidant activity was analyzed by Graph Pad Prism for Windows, Version 7 (Graph Pad Software, San Diego, CA, United States).

A linear regression analysis was used to obtain the IC50 values.

DPPH free radical scavenging assay

The free radical scavenging ability of the leaf stem and root extracts of *S. carvifolia* from different altitudes were tested by a DPPH radical scavenging assay (55). The hydrogen atom-donating ability of the plant extractives was determined by the decolorization of the

methanol solution of 2,2-diphenyl-1-picrylhydrazyl (DPPH). DPPH produces violet/purple color in methanol solution and fades to shades of yellow color in the presence of antioxidants. A solution of 0.1 mM DPPH in methanol was prepared, and 2.4 mL of this solution was mixed with 1.6 mL of extract in methanol at different concentrations (12.5–150 μg/mL). The reaction mixture was vortexed thoroughly and left in the dark at RT for 30 min. The absorbance of the mixture was measured spectrophotometrically at 517 nm. Ascorbic acid was used as a reference. The percentage DPPH radical scavenging activity was calculated by the following equation:

$$\text{Free radical scavenging activity}(\%) = (A0)\,\text{control} - (A1)\,\text{sample} / (A0)\,\text{control} * 100$$

where A0 is the control reaction absorbance, and A1 is the testing specimen absorbance. IC50 was determined graphically from the graph plotting the inhibition percentage (%) against the extracts (56).

Sugar-reducing capacity assay

The reducing sugar content (RSC) was determined using the 3,5-dinitrosalicylic acid (DNSA) method. The measurement was performed according to the procedure of Krivorotova and Sereikaite with slight modifications (57).

The DNSA reagent was prepared by dissolving 1 g of DNSA and 30 g of sodium–potassium tartaric acid in 80 mL of 0.5 N NaOH at 45° C. After dissolution, the solution was cooled down to room temperature and diluted to 100 mL with the help of distilled water. For the measurement, 2 mL of DNSA reagent was pipetted into a test tube containing 1 mL of plant extract (1 mg/mL) and kept at 95° C for 5 min. After cooling, 7 mL of distilled water was added to the solution, and the absorbance of the resulting solution was measured at 540 nm using a UV-VIS spectrophotometer (Shimadzu UV-1800). The reducing sugar content was calculated from the calibration curve of standard D-glucose (200–1,000 mg/L), and the results were expressed as mg D-glucose equivalent (GE) per gram of dry extract weight.

Antibacterial assay

Sample preparation and extraction

After cleaning, the different plant parts of the *S. carvifolia* plants collected from three altitudes (2,150, 2,593, and 3,178 m) were cut into small pieces using scissors and stored for drying. Drying was performed in a room for about 2 weeks in shade. They were then powdered using a grinder to enhance the surface area for a better extraction process. Twenty grams of finely ground powder from each plant part was used for making extracts. The 70% methanolic and aqueous extracts were prepared using a Soxhlet apparatus and simple maceration, and then the extracts were preserved for further studies.

Antibacterial activity

The antibacterial activity of the *S. carvifolia* leaf, stem, and root extracts from three different altitudes (2,150, 2,593, and 3,178 m) was evaluated against eight strains [3 Gram +ve *Staphylococcus*: *Staphylococcus aureus* (MTCC 96), *Staphylococcus epidermidis* (MTCC 435), and *Streptococcus mutans* (MTCC 890); 5 Gram –ve: *Klebsiella pneumoniae* (MTCC 109), *Escherichia coli* (MTCC 723), *Escherichia coli* (DH5α), *Salmonella typhimurium* (MTCC 98), and *Pseudomonas aeruginosa* (MTCC741); obtained from the Microbial Type Culture Collection Centre (MTCC), Institute of Microbial Technology (IMT), Chandigarh, India] by a micro-dilution broth assay using 96-well flat bottom microtiter plates according to the CLSI guidelines, according to which the antibiotic norfloxacin was used as a standard drug, and DMSO was used as a negative control (58).

Antifungal activity

By using the agar well diffusion method, the antifungal activity of all the *S. carvifolia* leaf, stem, and root extracts collected at different altitudes was investigated against four fungal strains: *Candida albicans* (CA-3010), *Candida albicans* (CA-227), *Creptococcus neoformans* (CN), and *Sporothrix schenckii* (SS) (59). All the fungal isolates were grown on a potato dextrose agar (PDA) at 28°C for proper sporulation. Subsequently, the fungal spores were harvested in sterile distilled water and evenly spread on PDA plates using a sterile glass spreader. Using a sterile cork borer (diameter 5 mm), wells were drilled into the agar media and filled with a sufficient amount of plant extract, oil, and water (control). The plates were allowed to rest at room temperature for 1 h to allow the extract to properly diffuse into the media before being incubated at 28°C for 96 h. The zones of inhibition around the wells were measured and recorded after incubation. The antifungal screening was performed in triplicates.

Cytotoxicity assay

Sample preparation and extraction

Selinum carvifolia leaves, stems, and roots were collected from three altitudes and air-dried for 10 days at room temperature before being milled into a powder. Following this, 70% methanol was used to extract the dry powder for roughly 7 days at room temperature. After evaporating the solvent, dry methanolic extracts were produced. To obtain suitable solutions for the extracts, dry methanolic extracts were dissolved in dimethyl sulfoxide (DMSO).

Cell culture

The study's A549 cell line (adenocarcinoma human alveolar basal epithelial cells) was obtained from the American Type Culture Collection (ATCC, Manassas, Virginia, United States). The cells were grown in Dulbecco's modified Eagle medium (Sigma) supplemented with 10% FBS (fetal bovine albumin) and 1% antibiotic/antimycin cocktail at 37°C in humidified air containing 5% CO2. The cells were allowed to develop to 80–90% confluence before being harvested with 0.25% trypsin/EDTA solution, and they were sub-cultured in 96-well plates according to the experiment's instructions.

Cytotoxicity assay

With minor changes (in concentration), the cytotoxicity assay based on MTT (3-(4, 5-dimethylthiazol-2-yl)-2, 5-diphenyl tetrazolium bromide) established by Mosman in 1983 was employed. A 5 mg/mL fresh stock solution of MTT was prepared in phosphate-buffered saline (PBS, pH 7.2) and filtered prior to initiating the experiment. Briefly, A549 cells (1×10^4 cells/well) were seeded in a 96-well flat-bottom cell culture plate and cultured overnight (60). Furthermore, the cells were then treated with ethanolic extract of the plant at five different concentrations (1, 0.5, 0.25, 0.125, and 0.0625 mg/mL), as previously described (61). As a control, the cells were cultured in regular media (without extract and same culture conditions). After the treatment period was completed, 10 µL of MTT stock solution (1/10th volume of the total media in the well) was added to each well and incubated for 3 h under standard culture conditions. After the incubation period, the medium was withdrawn, and each experimental well was filled with 200 µL of DMSO (Sigma Aldrich), which was then incubated for 20 min at room temperature. Finally, MTT activity was measured using an ELISA plate reader set at 492 nm, and the absorbance intensity was recorded at a 550 mm wavelength (Biotek, PowerwaveXS2). The following formula was used to compute the percentage of growth inhibition:

$$\%\text{Cell inhibition} = 100 - \left[\left(\text{At} - \text{Ab}\right) / \left(\text{Ac} - \text{Ab}\right)\right] \times 100$$

where At represents the absorbance value of the test substance, Ab represents the absorbance value of the blank, and Ac represents the absorbance value of the control.

IC50 values were used to express the effects of the extracts (the drug concentration reducing the absorbance of treated cells by 50% for untreated cells).

Results

HPLC analysis

To obtain an accurate quantification of chemicals, a sufficient resolution is required. The selection of the column, mobile phase composition, gradient flow parameters, and temperature was carried out by HPLC in order to obtain chromatograms with baseline separation of the 10 marker chemicals (ferulic acid, quercetin, gallic acid, rutin, chlorogenic acid, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid) in *S. carvifolia*. To design a separation procedure for the isolates from the *S. carvifolia* extract solutions, a range of solvent systems were initially tested, starting with pure methanol and gradually adding the aqueous phase. The best-performing solvent system, consisting of methanol and water, was chosen for the current study. Eventually, all 10 phenolic marker chemicals were eluted in less than 40 min using a straightforward gradient approach based on water and methanol. With a fixed wavelength of 250 nm, the most effective detection was noted. The approach that was devised showed a good deal of specificity. The specificity of the developed analytical methods was determined by comparing the spectra's peaks and retention durations. Consequently, we created a single stock solution in a standard manner before diluting it to the various concentration levels above and below the nominal amounts for each analyte. Based on their retention lengths and standard calibration curves, gallic acid, rutin, chlorogenic acid, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid were identified and quantified from all plant components at all altitudes. In the herbal drug field, the standardization and characterization of herbal drugs are ongoing research interests together with formulations. With the development of contemporary chromatographic techniques, there is an increasing desire to create and develop simple, quick, practical, and affordable procedures for standardization. HPLC is a sensitive and precise tool that is frequently used for the quality assessment of plant extracts and the products that are made from them for standardizing the methanolic extract of plant leaves, stems, and roots (62, 63). The limitation of the study is that we have 10 phenolic compounds as markers, and these phenolic compounds have potent microbial and anticancerous activity; therefore, we only attempted to validate the plants that have these kinds of potent compounds that have been used for herbal drug formulations.

The HPLC fingerprints of *S. carvifolia* from three different altitudes (2,150 m, 2,593 m, 3,178 m) are given in Figures 1–3. The results of the HPLC analysis of *S. carvifolia* methanolic extract at 400 nm revealed eight different important chemicals found in various plant parts, as shown in Figure 1. Chlorogenic acid (0.25 mg/g) was identified in the smallest concentration in (PTE) stem extract, while gallic acid (13.93 mg/g) was discovered in the highest concentration in (Hx) stem extract in plants growing at an altitude of 3,178 m. Figure 2 summarizes the extract composition of different plant parts of *S. carvifolia* at an altitude of 2,593 m. The lowest concentration of chlorogenic acid (0.01 mg/gm) was discovered in the dry weight basis of (PTE) stem extract and (Hx) leaf extract, while the highest concentration of gallic acid (3.11 mg/gm) was discovered in the dry weight of (PTE) leaf. Similarly, at an altitude of 2,150 meters, Figure 3 summarizes the extract composition of plant parts of *S. carvifolia*. The results revealed that chlorogenic acid and syringic acid (0.01 mg/gm) were found in the lowest concentrations on dry weight in (PTE) and (Hx) leaf extracts, respectively, while chlorogenic acid (15.38 mg/gm) was found in the highest concentration in (PTE) stem extract and protocatechuic acid (15.12 mg/gm) in (PTE) stem extract. A single chromatogram of mixed standards and a single chromatogram of samples have been provided as Supplementary Data for reference.

Total phenolic content

The total phenolic content (TPC) of standard gallic acid was estimated using the calibration curve ($Y = 0.0032x + 0.0528$, $R2 = 0.9983$) and represented as mg GAE/g dry extract weight. The TPC of the investigated plant samples ranged from 33.79 to 57.95 mg GAE/g dry extract weights (Figure 4). The highest TPC level was found in stem extract at an altitude of 3,178 m, followed by stem extract at 2,593 m (57.95 and 51.85 mg GAE/g dry extract weight), while the lowest content was found in root extract at an altitude of 2,150 m (33.79 mg GAE/g dry extract weight) among the nine plant

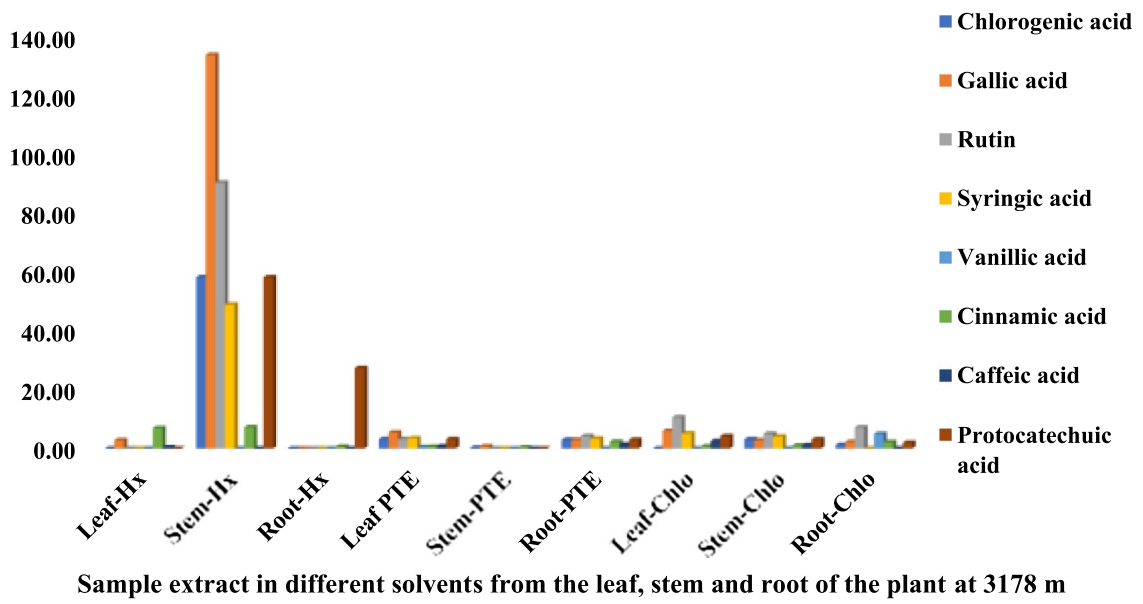


FIGURE 1
Chemical composition of leaf, stem, and root extracts [(Chl), (Hx), and (PTE)] of *Selinum carvifolia* plants (in mg/gm) growing at an altitude of 3,178 m using HPLC showing the presence of chlorogenic acid, gallic acid, rutin, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid.

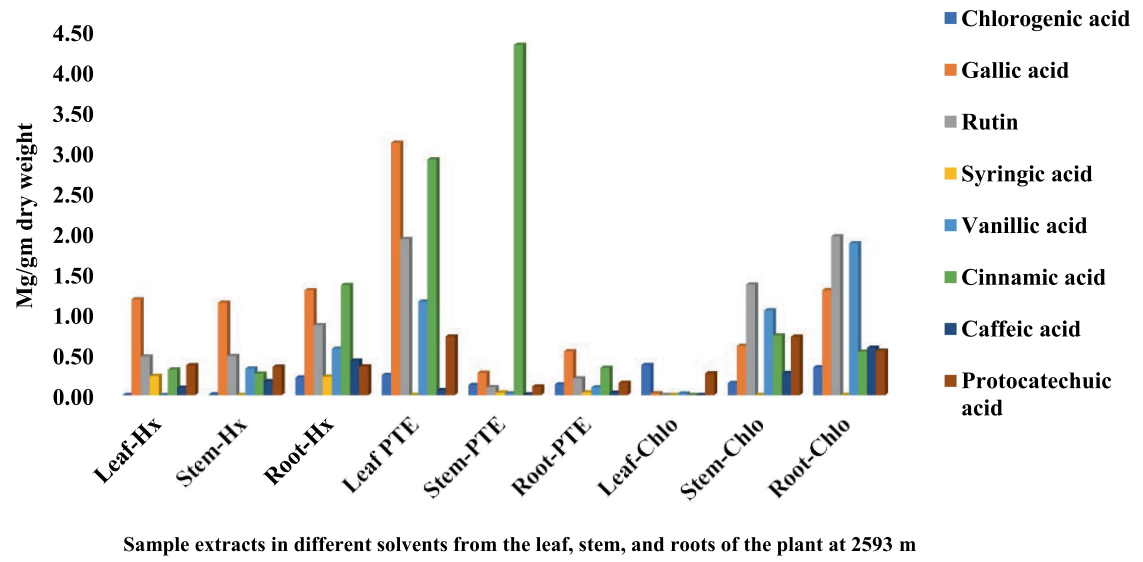


FIGURE 2
Chemical composition of leaf, stem, and root extracts [(Chl), (Hx), and (PTE)] of *Selinum carvifolia* plants (in mg/gm) growing at an altitude of 2,593 m using HPLC showing the presence of chlorogenic acid, gallic acid, rutin, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid.

samples analyzed. This is the first study of TPC in *S. carvifolia* leaves, stems, and roots and their variation at different altitudes (2,150, 2,593, and 3,178 m).

Total flavonoid content

The equation obtained from the standard rutin (R) graph was used to calculate the total flavonoid content (TFC) in the organic extract of the analyzed plant parts. TFC was found in considerable amounts in all plant parts (Figure 5). The highest TFC content was found in root extract at 3,178 m (9.4 mgRE/g) and the lowest in root extract at 2,150 m (3.8 mgRE/g).

Antioxidant activity

As shown in Figure 6, the antioxidant activity of methanolic extracts of *S. carvifolia* plant parts collected at various altitudes was assessed using the DPPH free radical scavenging test. In a

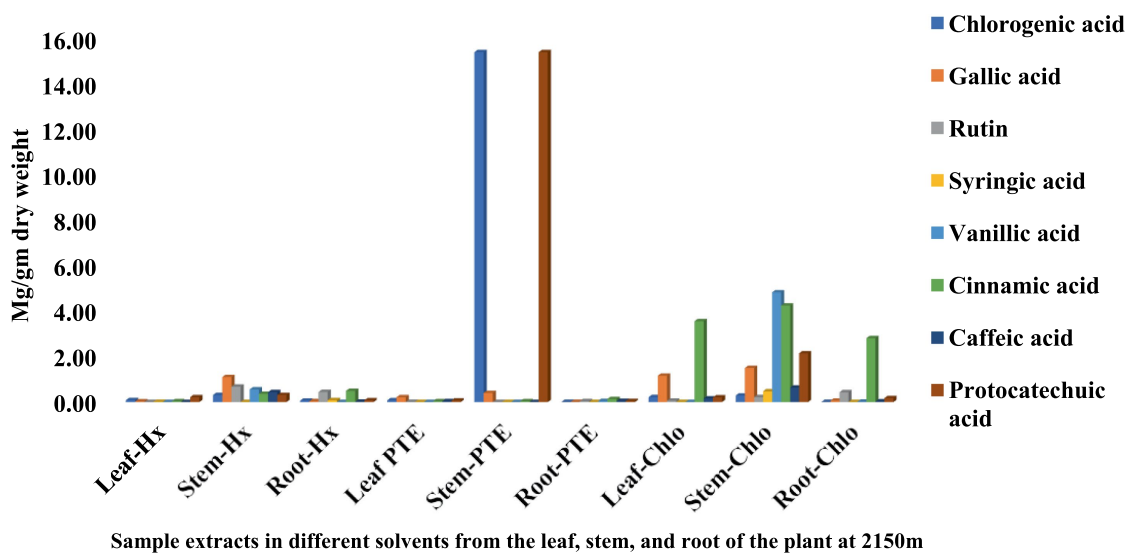


FIGURE 3
Chemical composition of leaf, stem, and root extracts [(Chl), (Hx), and (PTE)] of *Selinum carvifolia* plants (in mg/gm) growing at an altitude of 2,150 m using HPLC showing the presence of chlorogenic acid, gallic acid, rutin, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid.

concentration-dependent way, all the extracts show promising antioxidant activity. The IC50 value of the plant extracts and ascorbic acid as a positive control was graphically derived from the graph of the DPPH inhibition percentage. We found excellent efficiency of scavenging activity in extracts of the leaves, stems, and roots of *S. carvifolia* from different altitudes (2,150, 2,593, and 3,178 m; Figure 7).

Reducing sugar activity

The reducing sugar capacity (RSC) of different plant part extracts (leaves, stems, and roots) from different altitudes (2,150, 2,593, and 3,178 m) in *S. carvifolia* was studied using the known standard of ascorbic acid (AA), with the concentration level ranging from 250 ug/mL to 1 mg/mL. RSC was determined using the standard D-glucose calibration curve and represented as µg/mL GE/g dry extract weight. The RSC of the investigated plant samples ranged from 0.747 to 0.734 µg/mL GE/g dry extract weight. The RSC was highest in the leaf and stem extracts at an altitude of 3,178 m (0.747 and 0.734 ug/1 mL GE/g dry extract weight, respectively) among the nine plant samples examined. The lowest RPC was found in the root extract at an altitude of 2,150 m (0.495 µg/mL of GE/g dry extract).

Antibacterial activity

The plant extracts obtained from different plant parts collected from three altitudes (2,150, 2,593, and 3,178 m) were tested for antibacterial activity on strains of *Staphylococcus aureus* (SA), *Staphylococcus epidermidis* (SE), *Streptococcus mutans* (SM), *Klebsiella pneumoniae* (KP), *Escherichia coli* (EC), *Escherichia coli* (DH5α), *Salmonella typhimurium* (STM), and *Pseudomonas aeruginosa* (PA). Root extract of the plant collected from all altitudes was alone found to be active against SA-96, while other plant parts failed to show any kind of activity against the bacterial strains listed. Methanolic extract of the root collected from an altitude of 2,150 m showed the highest activity (<250 µg/mL) on the SA-96 strain. Similar activity was also recorded from the roots collected at 2,593 m on SA-96 and moderate activity (<500 µg/mL) on KP and PA. On the other hand, extracts from roots collected at 3,178 m showed moderate activity (<500 µg/mL) on both strain SA-96 and SE strains (Table 1).

Aqueous extracts from leaves, stems, and roots from plants at three different altitudes (2,150, 2,593, and 3,178 m) were also tested for antibacterial activity against same bacterial strains. However, no activity was found against any bacterial strain except that of stem growing at an altitude of 3,178 m with moderate activity (>500 µg/mL) against PA (Table 2).

Antifungal activity

Methanolic and aqueous plant part extracts (leaves, stems, and roots) collected from three different altitudes (2,150, 2,593, and 3,178 m) were tested for activity against the fungal strains *Candida albicans* (CA-3010), *Candida albicans* (CA-227), *Creptococcus neoformans* (CN), and *Sporothrix schenckii* (SS). The root extracts belonging to plants growing at altitudes of 2,593 and 3,178 m were found to be active against the SS (<500 µg/mL). The activity (<500 µg/mL) against CA-227 was also reported in the extract of plant collected at an altitude of 3,178 m. However, no significant activity was reported against any fungal strains in the extract from plants growing at 2,150 m altitude (Table 3). Similar activity was also recorded in the aqueous extracts of roots growing at altitudes of 2,593 and 3,178 m. On the other hand,

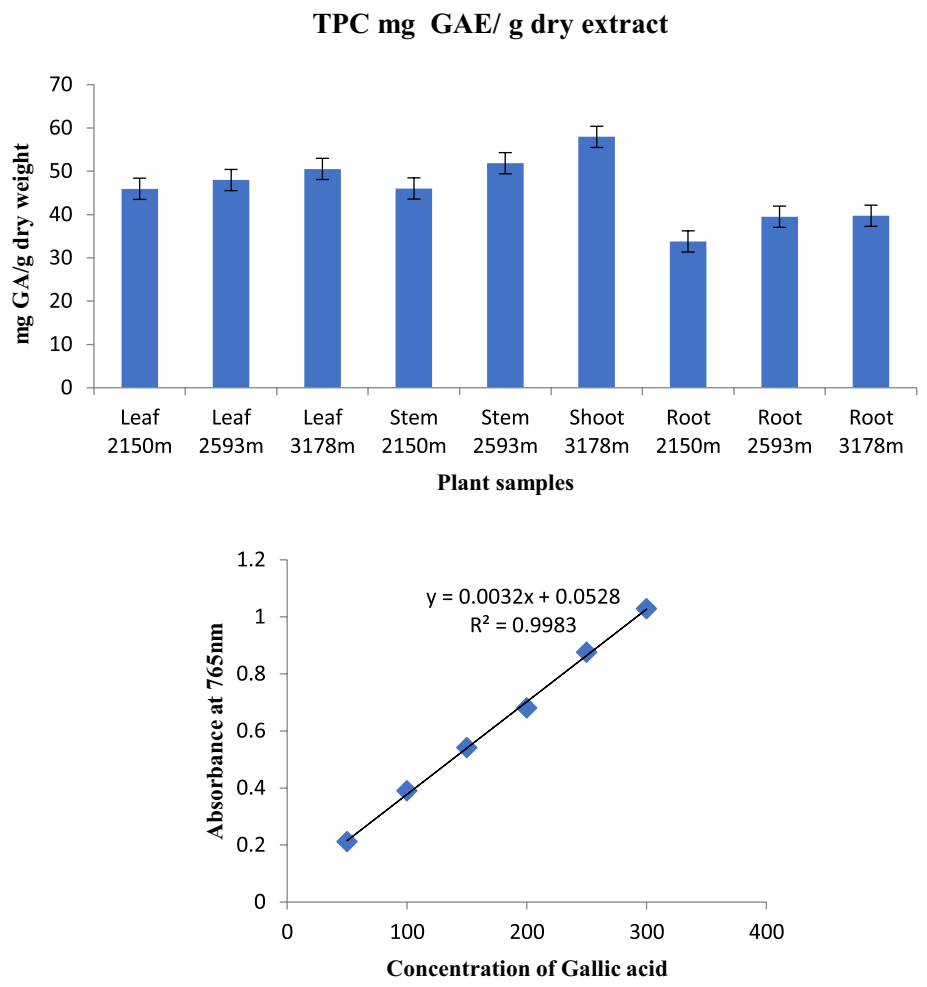


FIGURE 4
Total phenolic content in various plant components along various altitudinal gradients and a standard curve with gallic acid concentrations ranging from 0.5 to 4 g/mL.

activity (500 µg/mL) against the SS fungal strain was recorded in the aqueous extract of roots growing at 2,593 and 3,178 m (Table 4).

Cytotoxicity

The leaf, stem, and root extracts of plants growing at different altitudes were tested against A549, and it was observed that leaf extract does not have any cytotoxic activity even at higher doses. However, the stem and root extracts showed an inhibitory effect at a 0.25 mg/mL concentration, while complete cell cytotoxicity was found at the highest dose of (1 mg/mL) concentration. *Selinum carvifolia* leaf extracts from different altitudes showed mild cytotoxicity or inhibition of cell proliferation in higher concentrations, while root extracts of plants growing at the three altitudes (2,150, 2,593, and 3,178 m) inhibited cell proliferation even at a lower concentration of 0.125 mg/mL, and maximum cell growth inhibition could be seen at higher concentrations of 1 mg/mL. Plants growing at an altitude of 2,150 m displayed modest cytotoxicity at higher dosages, but plants growing at altitudes of 2,593 and 3,178 m exhibited cytotoxicity in A549 cells at lower concentrations, resulting in full cell death in a dose-dependent manner. The mild toxicity of leaf extracts from plants growing at the three altitudes (2,150, 2,593, and 3,178 m) and the stem of plants growing at an altitude of 3,178 m was detected, and the percentage of growth inhibition increased with the increasing concentration of test chemicals (as demonstrated in the graphs in Figures 8–10). A comparative graphical representation of the inhibitory activity of extracts from different plant parts (leaves, stems, and roots) of *S. carvifolia* collected from three altitudes (2,150, 2,593, and 3,178 m) against the A549 cell line using an MTT assay is shown in Figure 11.

Discussion

Phytochemical analysis is a significant laboratory or scientific process. This procedure identifies the fundamental elements of any plant portion, including the leaves, stems, and roots. Nobody is exactly

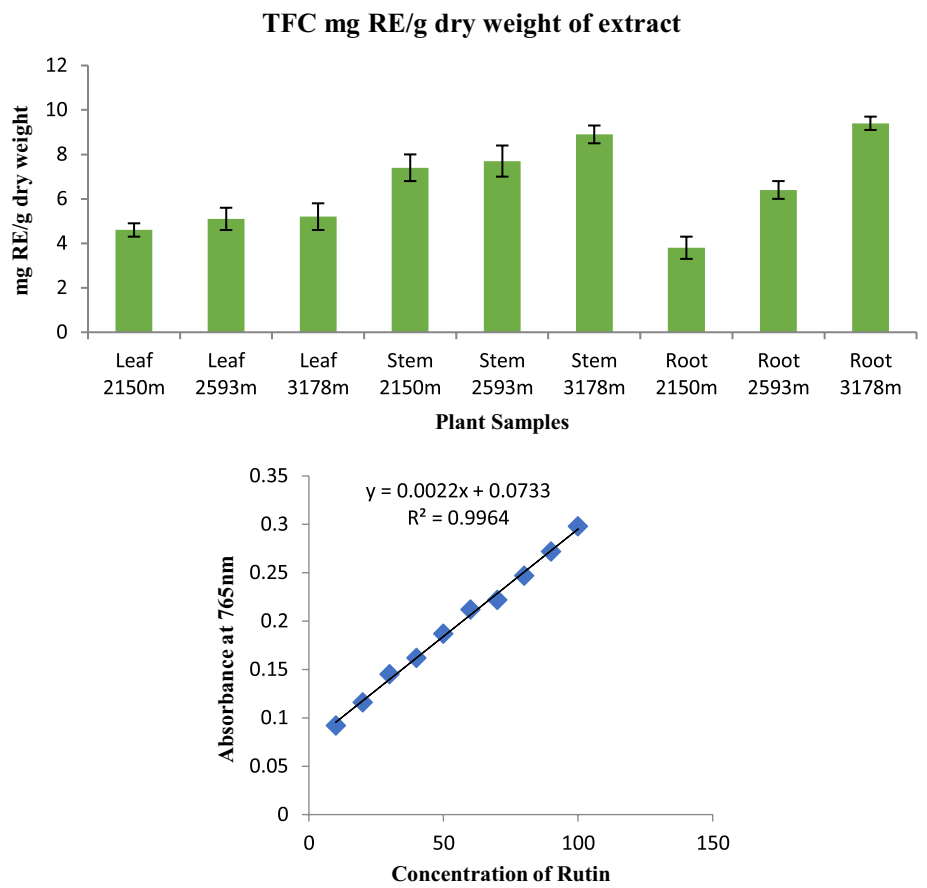


FIGURE 5
Total flavonoid content in different plant parts along various altitudinal gradients and a standard curve with rutin concentration ranging from 5 to 25 μg in the reaction mixture.

sure how many distinct types of medicinal plants are utilized globally today, but we do know that they play a significant role in both conventional and Western medicine. Thus, it is crucial to use a viable technique to test the phytochemicals present in the plant. It can be inferred from the HPLC fingerprints that this analytical method is a practical way to determine the presence of a wide range of substances contained in the methanolic extract of plant leaves, stems, and roots.

Various medicinal plants are used traditionally for the treatment of different diseases or symptoms. In order to correlate the traditional uses to modern pharmacological activities, it is important to analyze their major chemical constituents and to evaluate their biological activities. To gain a broad picture of a plants' phytochemical composition and its biological activities, it is essential to investigate different portions of the plant (leaf, stem, and root bark, fruits, flowers, and so on) to find the most promising source (64). Using reversed-phase HPLC with a UV detector at 230-400 nm, we investigated several components of *S. carvifolia* (root, stem, and leaf) and discovered that chloroform, hexane, and petroleum ether extracts contain substantial levels of phenolic compounds. The phenolic chemicals in the leaves, stems, and roots of *S. carvifolia* collected from three different altitudes were elucidated. Based on their retention lengths and standard calibration curves, gallic acid, rutin, chlorogenic acid, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid were identified and quantified from all the plant parts at each studied altitude. It was found that the concentration of phenolic compounds increased significantly along with the altitude in both aerial and underground parts. The leaves and stem of all three different altitudes contained more phenolic compounds than roots. The presence of physiologically and pharmacologically essential phenolic chemicals in quantifiable amounts as demonstrated in the study can be used by the pharmaceutical and phytopharmaceutical industries to build quality control profiles. Furthermore, this method used is accurate and repeatable, and it may be used to determine phenolic compounds in plant extracts.

To the best of our knowledge, this study is the first to isolate, characterize, biologically evaluate, quantify, and validate the chemical contents of *S. carvifolia* using HPLC. The results of this study reveal the significant *in vitro* cytotoxic potential of this plant toward cancerous cells, which suggests the anticancerous activity of plant extract can probably be attributed to the phenolic compounds, as determined by the MTT cytotoxicity assay. However, more studies are warranted to further elucidate the properties of the compounds present in the extracts and their mechanism of action. The presence of highly bioactive molecules creates new opportunities for therapeutic development. The created method is easy to use, sensitive, specific, and repeatable, and it can be expanded to assess the quality of

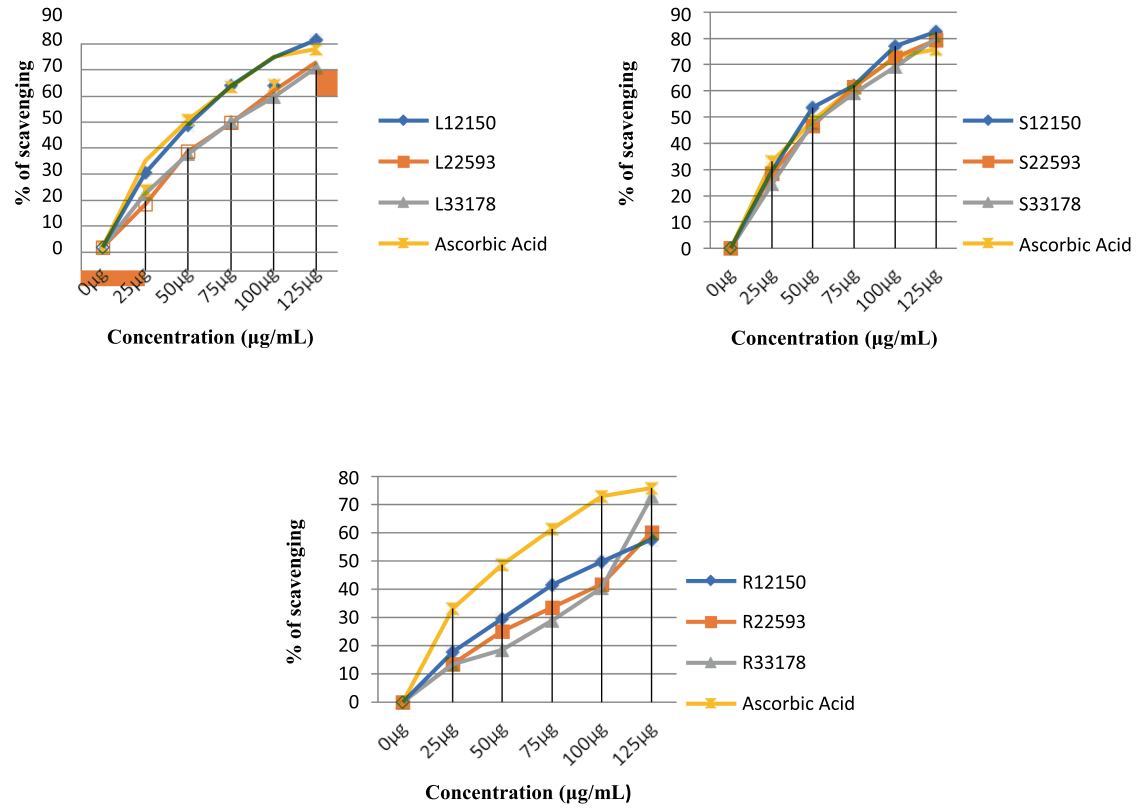


FIGURE 6
Measurement of DPPH radical scavenging activity in leaf, stem, and root extracts at various altitudes (Triple experiments were carried out). The data are presented as mean ± SD; $n = 3$, $p < 0.05$.

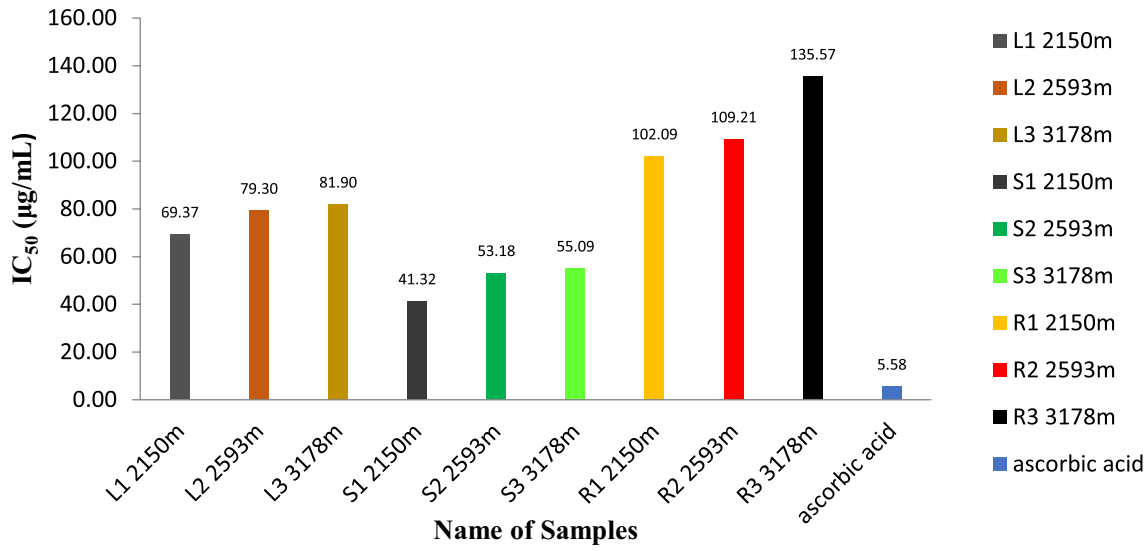


FIGURE 7
Determination of the IC50 value in leaf, stem, and root extracts at different altitudes.

S. carvifolia, according to the validation results. For quick study of its phytomolecules in diverse plants, herbal formulations, and plant products, the preliminary HPLC fingerprinting approach can be useful.

Susceptibility tests with MICs in the range of 500–1,500 g/mL are commonly used to classify the antimicrobial activity of plant extracts (65). In our findings, antibacterial activity of methanolic extract in the roots of *S. carvifolia* growing at an altitude of 2,150 m

TABLE 1 Minimum inhibitory concentration of ethanolic extract of roots (in μg/mL) of *Selinum carvifolia* against *Staphylococcus aureus* (SA-96), *Staphylococcus epidermidis* (SE), *Streptococcus mutans* (SM), *Klebsiella pneumoniae* (KP), *Escherichia coli* (EC), *Escherichia coli* (DH5α), *Salmonella typhimurium* (STM), and *Pseudomonas aeruginosa* (PA).

Altitudes	Samples strains							
	SA-96	SE	EC	KP	SM	DH5α	STM	PA
Root 2,150 m	250	--	--	--	--	--	--	--
Root 2,593 m	250	--	--	<500	--	--	--	<500
Root 3,178 m	500	500	--	--	--	--	--	--

TABLE 2 Minimum inhibitory concentration of aqueous extract of stems (in μg/mL) of *Selinum carvifolia* against *Staphylococcus aureus* (SA-96), *Staphylococcus epidermidis* (SE), *Streptococcus mutans* (SM), *Klebsiella pneumoniae* (KP), *Escherichia coli* (EC), *Escherichia coli* (DH5α), *Salmonella typhimurium* (STM), and *Pseudomonas aeruginosa* (PA).

Altitudes	Samples strains							
	SA-96	SE	EC	KP	SM	DH5α	STM	PA
Stem 2,150 m	--	--	-	-	--	--	--	--
Stem 2,593 m	--	--	--	--	--	--	--	--
Stem 3,178 m	--	--	--	--	--	--	--	<500

TABLE 3 Minimum fungicidal concentration of ethanolic extract of roots (in μg/mL) of *Selinum carvifolia* against *Candida albicans* (CA-3010), *Candida albicans* (CA-227), *Cryptococcus neoformans* (CN), and *Sporothrix schenckii* (SS).

Altitudes	Samples strains			
	CA-3010	CN	SS	CA-227
Root 2,150 m	--	--	--	--
Root 2,593 m	--	--	<500	--
Root 3,178 m	--	--	<500	<500

TABLE 4 Minimum fungicidal concentration of aqueous extract of roots (in μg/mL) of *Selinum carvifolia* against *Candida albicans* (CA-3010), *Candida albicans* (CA-227), *Cryptococcus neoformans* (CN), and *Sporothrix schenckii* (SS).

Altitudes	Samples strains			
	CA-3010	CN	SS	CA-227
Root 2,150 m	--	--	--	--
Root 2,593 m	--	--	<500	--
Root 3,178 m	--	--	<500	--

was reported to be active against only one strain among the selected bacterial strains, namely, SA-96, while the root parts collected from an altitude of 2,593 m showed activity against SA-96, KP, and PA strains. The methanolic extracts of roots collected from an altitude of 3,178 m showed moderate activity against SA-96 and SE strains, whereas the aqueous extracts of stems from the same altitude showed activity against the PA strain. The root methanolic extract from an altitude of 3,178 m showed antifungal activity against two out of four fungal strains, namely, SS and CA-227, while the roots obtained from an altitude of 2,593 m only showed activity against the SS strain. Antifungal activity was found in the aqueous extracts of roots from the altitudes of 2,593 and 3,178 m against the SS strain. This variation could be attributed to differences in the active compounds present in the extracts, extraction solvent, employed plant component, analysis method, environmental stress, and climatic and geographical factors (66).

The antioxidant properties of *S. carvifolia* were assessed by measuring total phenols, total flavonoids, and DPPH activity using the crude methanolic extract of various plant parts from different altitudes, and it was observed that there were variations in activity in the different altitudes, which could be due to several factors. Reducing sugar activity was found to be at a maximum in the methanolic extracts of the aerial parts as the altitude increased. The genotype, plant age, soil quality, geographical location, meteorological conditions, cultivation method, and abiotic stress may all play a role in this variation (67–69).

The cytotoxicity test was performed on the A549 cell line to determine whether the fractions were cytotoxic to cancer cell lines. The cytotoxic data of different fractions of the extracts of *S. carvifolia* revealed no toxicity in the extracts of the aerial parts, even at different altitudes from the stems collected at 3,178 m, but the root extracts showed mild inhibition at high doses. These results confirm the beneficial use and application of this plant against human pathogens and the associated human-and community-acquired infections (70, 71).

In the present investigation, it was interesting to observe that the total phenolic contents in nearly all the plant parts increased with altitude, thus also increasing the antibacterial, antifungal, and antioxidant properties along with the cancer cell cytotoxicity and reducing sugar activities. The secondary metabolite concentration increases with altitude as it plays an important role in the defense mechanisms of plants. The magnified UV radiation and low temperature produces a stressful environment, leading to the formation of free radicals that are countered by phenols, flavonoids, and terpenoids (72). Among the different plant parts tested, the bioactivity was found to be highest in the roots, which could be an indication about the importance of roots for the adaptation of plants to a changing environment, even at the lowest

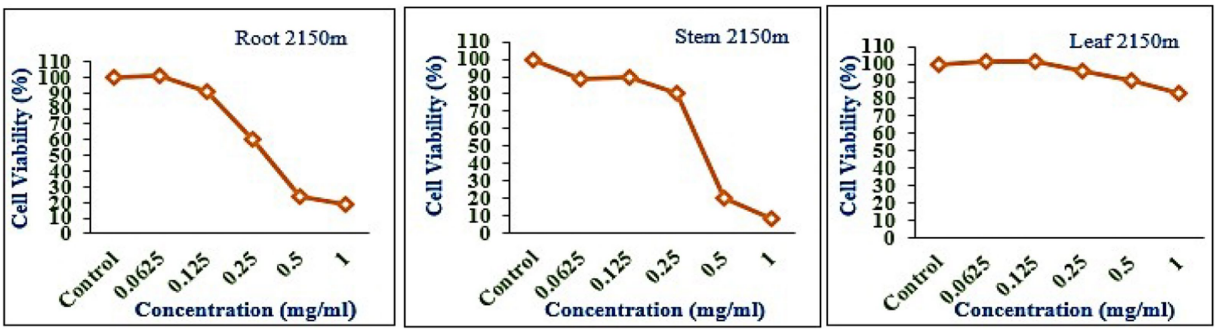


FIGURE 8
Inhibition of A549 cell proliferation in methanolic extract of different plant parts (root, stem, and leaf) of *Selinum carvifolia* at 2,150 m altitude.

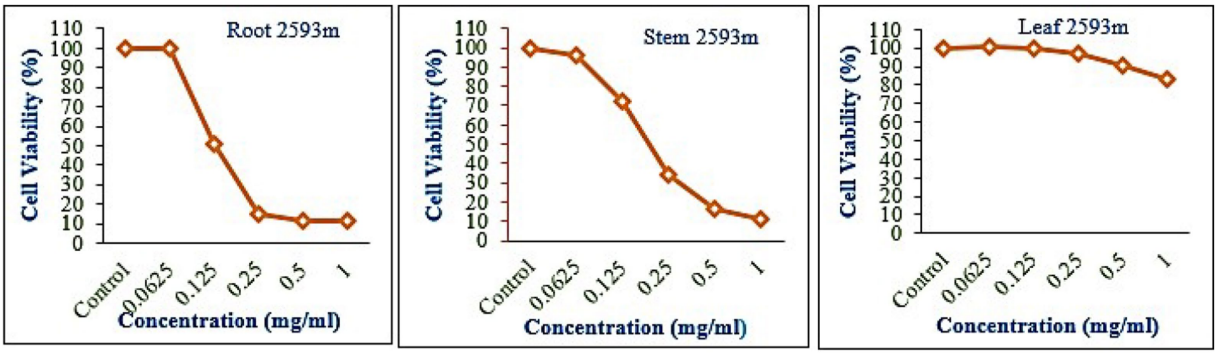


FIGURE 9
Inhibition of A549 cell proliferation in methanolic extract of different plant parts (root, stem, and leaf) of *Selinum carvifolia* at 2,593 m altitude.

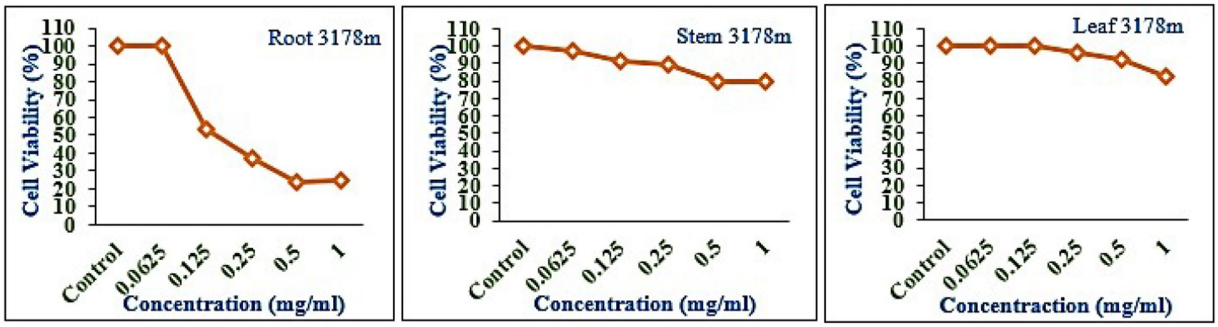


FIGURE 10
Inhibition of A549 cell proliferation in methanolic extract of different plant parts (root, stem, and leaf) of *Selinum carvifolia* at 31,78 m altitude.

level, as also supported by Souhir et al. (73). The cytotoxic activity of plant has also been shown to increase with the increase in altitude, with better activity observed against cancer cell lines (74).

Our findings are intriguing, and they may prompt additional investigation into the phytochemical, toxicological, and pharmacological characteristics of these extract products in order to promote their widespread use in antibacterial, antifungal, antioxidant, and anticancer activities. In the current study, the methanolic extracts gave higher yields of chemical constituents than initially expected or anticipated; the originality of this work is that positive results were achieved with a hydro-alcohol ratio and it will be useful to carry out other data analyses with MIC and other formulation studies, because hydro-alcohol is more suitable for clinical study than methanol or water extracts. In comparison to standard drugs, hydro-alcoholic extracts of *S. carvifolia* were found

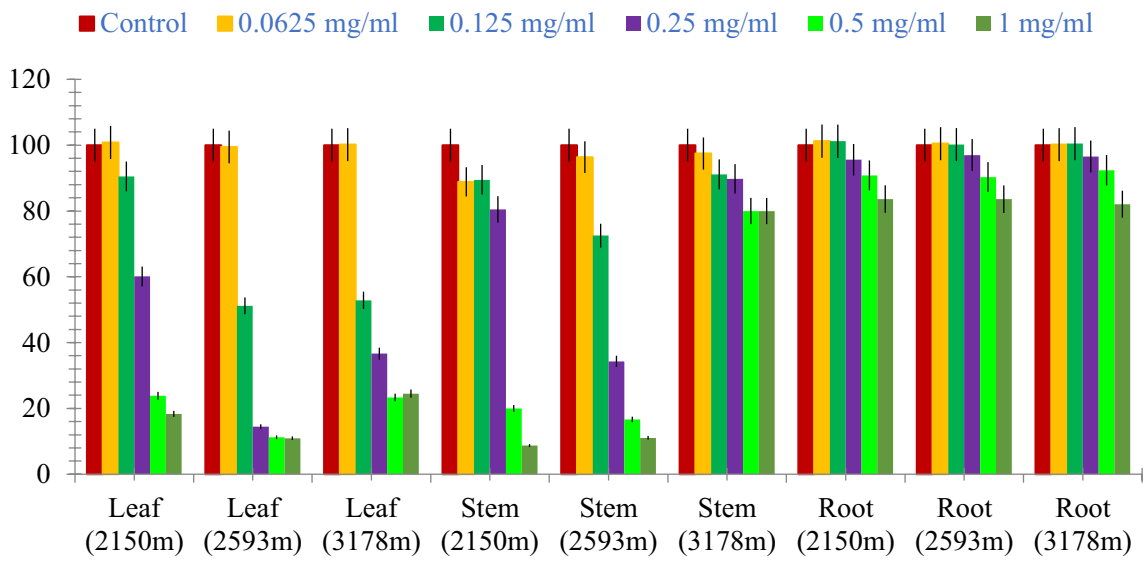


FIGURE 11
Comparative graphical representation of inhibitory activity of plant part extracts (leaf, stem, and roots) from three different altitudes (2,150, 2,593, and 3,178 m) against A549 cell proliferation using MTT assay. Data are represented as mean ± SE value of the three independent experiments.

to be active against the majority of clinically isolated microorganisms and fungi. The current study validated the claimed uses of the whole plant in traditional medicine to treat various infectious diseases caused by microbes. However, more research is needed to better assess the potential efficacy of crude extracts as antimicrobial agents. The altitudinal variation was also found in the plant extract, and according to our findings, the plant extract growing at the highest altitude produces the best results.

Conclusion

Biologically active compounds were identified and quantified in the leaf stems and roots of *Selinum carvifolia*. Eight phenolic compounds were analyzed through HPLC methods. For the first time, the presence of gallic acid, chlorogenic acid, syringic acid, vanillic acid, cinnamic acid, caffeic acid, and protocatechuic acid have been reported as non-volatile constituents of *S. carvifolia*. A significant difference in concentration was observed in phenolic compounds from lower to higher altitudes. This may be due to the fact that at higher altitudes, plants encounter greater environmental challenges than in lower altitudes. All the extracts showed antibacterial, antifungal, cytotoxicity, and antioxidant effects, notably the root extracts, which exhibited very good cytotoxicity activity and scavenging activity against DPPH radicals. *S. carvifolia,* which was researched in this study, is a rich source of physiologically active chemicals that can be utilized to treat and prevent a variety of ailments.

In the current study, the methanolic extract gave higher yields of chemical constituents than expected; the originality of this work is that some great results were obtained with the hydro-alcohol ratio, and it will be useful to carry out other data analyses with MIC and other formulation studies, because hydro-alcohol is more suitable for clinical study than methanol or water extracts. In comparison to standard drugs, the hydro-alcoholic extracts of *S. carvifolia* were found to be active against the majority of clinically isolated microorganisms and fungi. The current study validated the claimed uses of the whole plant in traditional medicine to treat various infectious diseases caused by microbes. However, more research is needed to better assess the potential efficacy of crude extracts as antimicrobial agents. The altitudinal variation was also found in the plant extract, and according to our findings, the plant extract growing at the highest altitude produced the best results. The current findings will be used to select plant species for further investigation in the search for new natural bioactive compounds.

The phytochemical composition, total phenol content, and antibacterial, antifungal, cytotoxicity, and DPPH activity of the methanolic extracts of *S. carvifolia* leaves, stems, and roots were investigated in this study. The antioxidant activity of different parts of *S. carvifolia* MeOH extract was tested against free radicals DPPH and reducing sugar activity. The results showed that root extract is a good antioxidant agent because it inhibits DPPH activity more effectively. The differences in the antioxidant activity of the different parts of *S. carvifolia* can be attributed to differences in the concentration of the identified phenolic compounds present and the higher levels of phenolics and flavonoids in the extract of the selected samples.

According to the current findings, this plant is a rich source of medicinally important phytoconstituents, and the observed antimicrobial, cytotoxic, and antioxidant potential could be attributed to these constituents. Although the parameters used in this study are not disease-specific, quantification of their properties can be used to guide the use of these plants in ROS-related complications such as diabetes. More research is needed to isolate and identify the responsible active compounds

and their mechanisms of action so that we can better understand their ability to control relevant diseases.

Data availability statement

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

Author contributions

The manuscript was structured and prepared by RS, LS, NS, SP, AS, HD, PV, SS, SK, SM, and SR under the guidance of GS. All authors contributed to the article and approved the submitted version.

Conflict of interest

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

Publisher's note

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

References

1. Dewal MB, Wani AS, Vidaillac C, Oupický D, Rybak MJ, Firestine SM. Thieno [2,3-d] pyrimidinedione derivatives as antibacterial agents. *Eur J Med Chem*. (2012) 51:145–53. doi: 10.1016/j.ejmech.2012.02.035

2. Nkomo LP. In vitro bioactivity of crude extracts of Lippia javanica on clinical isolates of Helicobacter pylori: Preliminary phytochemical screening. University of Fort Hare. Master' Thesis, University of Fort Hare, Faculty of Science and Agriculture (2010) p. 97.

3. Kalghatgi S, Spina CS, Costello JC, Liesa M, Morones-Ramirez JR, Slomovic S, et al. Bactericidal antibiotics induce mitochondrial dysfunction and oxidative damage in Mammalian cells. *Sci Transl Med*. (2013) 5:192ra85. doi: 10.1126/scitranslmed.3006055

4. Chanda S, Kaneria M, Nair R. Antibacterial activity of *Psoralea corylifolia* L. seed and aerial parts with various extraction methods. *Res J Microbiol*. (2011) 6:124.

5. Adeshina GO, Onaolapo JA, Ehinmidu JO, Odama LE. Phytochemical and antimicrobial studies of the ethyl acetate extract of Alchornea cordifolia leaf found in Abuja, Nigeria. *J Med Plant Res*. (2010) 4:649–58.

6. Rojas J, Buitrago A In: D MCT and M Rai, editors. *Essential oils and their products as antimicrobial agents: Progress and prospects. Chapter 13, 26p*: Therapeutic Medicinal Plants from Lab to Mark (2015)

7. Joubouhi C, Tamokou J-D-D, Ngnokam D, Voutquenne-Nazabadioko L, Kuiate J-R. Iridoids from Canthium subcordatum iso-butanol fraction with potent biological activities. *BMC Complement Altern Med*. (2017) 17:17. doi: 10.1186/s12906-016-1536-8

8. Nzogong RT, Ndjateu FST, Ekom SE, Fosso J-AM, Awouafack MD, Tene M, et al. Antimicrobial and antioxidant activities of triterpenoid and phenolic derivatives from two Cameroonian Melastomataceae plants: Dissotis senegambiensis and Amphiblemma monticola. *BMC Complement Altern Med*. (2018) 18:159. doi: 10.1186/s12906-018-2229-2

9. Davidson PM, Branen AL. Food antimicrobials—an introduction In: PM Davidson, JN Sofos and AL Branen, editors. *Antimicrobial in food. 3rd* ed. New York, NY: CRC Press (2005). 1–10.

10. Sieniawska E, Swatko-Ossor M, Sawicki R, Skalicka-Woźniak K, Ginalska G. Natural Terpenes Influence the Activity of Antibiotics against Isolated *Mycobacterium tuberculosis*. *Med Princ Pract*. (2017) 26:108–12. doi: 10.1159/000454680

11. Tagousop CN, Tamokou J-D-D, Kengne IC, Ngnokam D, Voutquenne-Nazabadioko L. Antimicrobial activities of saponins from Melanthera elliptica and their synergistic effects with antibiotics against pathogenic phenotypes. *Chem Cent J*. (2018) 12:97. doi: 10.1186/s13065-018-0466-6

12. Ahmad I, Aqil F. In vitro efficacy of bioactive extracts of 15 medicinal plants against ESβL-producing multidrug-resistant enteric bacteria. *Microbiol Res*. (2007) 162:264–75. doi: 10.1016/j.micres.2006.06.010

13. Rao K, Ch S, Banji D, Sandhya S, Mahesh V. A study on the nutraceuticals from the genus Rumex. *Hygeia J D Med*. (2011) 3:76–88.

14. Adnan M, Ali S, Sheikh K, Amber R. Review on antibacterial activity of Himalayan medicinal plants traditionally used to treat pneumonia and tuberculosis. *J Pharm Pharmacol*. (2019) 71:1599–625. doi: 10.1111/jphp.13156

15. Thakur A, Singh S, Dulta K, Singh N, Ali B, Hafeez A, et al. Nutritional evaluation, phytochemical makeup, antibacterial and antioxidant properties of wild plants utilized as food by the Gaddis-a tribal tribe in the Western Himalayas. *Front Agron*. (2022) 4:114. doi: 10.3389/fagro.2022.1010309

16. Mann A, Banso A, Clifford LC. An antifungal property of crude plant extracts from Anogeissus leiocarpus and Terminalia avicennioides. *Tanzan J Health Res*. (2008) 10:34–8.

17. Singh N, Pandey R, Chandraker SK, Pandey S, Malik S, Patel D. Use of wild edible plants can meet the needs of future generation In: *Agro-biodiversity and Agri-ecosystem Management*: Springer Nature Singapore (2022). 341–66.

18. Yahaya U, Abubakar S, Salisu A. Antifungal Activity of *Parkia biglobosa* Extract on Pathogenic Strain of *Candida albicans*. *J Appl Sci*. (2019) 19:235–40. doi: 10.3923/jas.2019.235.240

19. Webster D, Taschereau P, Belland RJ, Sand C, Rennie RP. Antifungal activity of medicinal plant extracts; preliminary screening studies. *J Ethnopharmacol*. (2008) 115:140–6. doi: 10.1016/j.jep.2007.09.014

20. Thangavelu T, Udayabhanu J, Thangavel S. Shanmugapriya K Antimicrobial activity of methanolic extracts of indigenous traditional Indian folk Medicinal Plant, *Gnaphalium polycaulon*. *Int J Green Pharm*. (2015) 9:39. doi: 10.4103/0973-8258.150921

21. Niño J, Narváez DM, Mosquera OM, Correa YM. Antibacterial, antifungal and cytotoxic activities of eight Asteraceae and two Rubiaceae plants from Colombian biodiversity. *Braz J Microbiol*. (2006) 37:566–70. doi: 10.1590/s151783822006000400030

22. Singh N, Mansoori A, Jiwani G, Solanke AU, Thakur TK, Kumar R, et al. Antioxidant and antimicrobial study of Schefflera vinosa leaves crude extracts against rice pathogens. *Arab J Chem*. (2021) 14:103243. doi: 10.1016/j.arabjc.2021.103243

23. Sahu N, Singh N, Arya KR, Reddy SS, Rai AK, Shukla V, et al. Assessment of the dual role of Lyonia ovalifolia (Wall.) Drude in inhibiting AGEs and enhancing GLUT4 translocation through LC-ESI-QTOF-MS/MS determination and in silico studies. *Front Pharmacol*. (2023) 14:1073327. doi: 10.3389/fphar.2023.1073327

24. Maoz M, Neeman I. Antimicrobial effects of aqueous plant extracts on the fungi Microsporum canis and Trichophyton rubrum and on three bacterial species. *Lett Appl Microbiol*. (1998) 26:61–3. doi: 10.1046/j.1472-765x.1998.00277.x

25. Sharma V, Jambaladinni K, Singh N, Mishra N, Kumar A, Kumar R. Understanding environmental associated abiotic stress response in plants under changing climate In: *Molecular Response and Genetic Engineering for Stress in Plants*: IOP Publishing (2022)

26. Singh N, Jiwani G, Rocha LS, Mazaheri R. Bioagents and Volatile Organic Compounds: An Emerging Control Measures for Rice Bacterial Diseases. In *Bacterial Diseases of Rice and Their Management*. Apple Academic Press. (2023) 255–74.

27. Nagavani V, Rao TR. Evaluation of antioxidant potential and qualitative analysis of major polyphenols by RP–HPLC in Nymphaea nouchali Burm flowers. *Int J Pharm*. (2010) 2:98–104.

28. Cartea ME, Francisco M, Lema M, Soengas P, Velasco P. Resistance of cabbage (*Brassica oleracea* capitata group) crops to Mamestra brassicae. *J Econ Entomol*. (2010) 103:1866–74. doi: 10.1603/EC09375

29. Krings U, Berger RG. Antioxidant activity of some roasted foods. *Food Chem*. (2001) 72:223–9. doi: 10.1016/S0308-8146(00)00226-0

30. Mundhe KS, Kale A, Gaikwad S, Deshpande NR, Kashalkar R. Evaluation of phenol, flavonoid contents and antioxidant activity of Polyalthia longifolia. *J Chem Pharm Res*. (2011) 3:764–9.

31. Singh A, Singh N, Singh S, Srivastava RP, Singh L, Verma PC, et al. The industrially important genus Kaempferia: An ethnopharmacological review. *Front Pharmacol*. (2023) 14:311.

32. Hakiman M, Maziah M. Non enzymatic and enzymatic antioxidant activities in aqueous extract of different Ficus deltoidea accessions. *J Med Plant Res*. (2009) 3:120–31.

33. Qusti SY, Abo-Khatwa AN, MAB L. Screening of antioxidant activity and phenolic content of some selected food items cited in the holy Quran. *EJBS*. (2010) 2:40–51.

34. Javanmardi J, Stushnoff C, Locke E, Vivanco JM. Antioxidant activity and total phenolic content of Iranian Ocimum accessions. *Food Chem*. (2003) 83:547–50. doi: 10.1016/S0308-8146(03)00151-1

35. Meng F, Zuo G, Hao X, Wang G, Xiao H, Zhang J, et al. Antifungal activity of the benzo [c] phenanthridine alkaloids from *Chelidonium majus* Linn. against resistant clinical yeast isolates. *J Ethnopharmacol*. (2009) 125:494–6. doi: 10.1016/j.jep.2009.07.029

36. Rajak BK, Rani P, Singh N, Singh DV. Sequence and structural similarities of ACCase protein of Phalaris minor and wheat: An insight to explain herbicide selectivity. *Front Plant Sci*. (2022) 2022, 13:1056474. doi: 10.3389/fpls.2022.1056474

37. EL-Hefny M, Salem MZ, Behiry SI, Ali HM. The potential antibacterial and antifungal activities of wood treated with *Withania somnifera* fruit extract and the phenolic, caffeine and flavonoid composition of the extract according to HPLC. *PRO*. (2020) 8:113. doi: 10.3390/pr8010113

38. Srivastava RP, Dixit P, Singh L, Nagpoore NK, Rawat V, Pandey S, et al. HPTLC fingerprinting, photopigments and nutrient analysis of Selinum tenuifolium along the altitudinal gradient. *J Liq Chromatogr Relat Technol*. (2021) 44:87–94. doi: 10.1080/10826076.2020.1848861

39. Chauhan RS, Nautiyal MC, Figueredo G, Chalard P. Volatile composition of underground parts of Selinum vaginatum and possible uses. *Chem Nat Compd*. (2012) 48:901–2. doi: 10.1007/s10600-012-0418-7

40. Srivastava RP, Dixit P, Singh L, Verma PC, Saxena G. Comparative morphological and anatomical studies of leaves, stem, and roots of Selinum vaginatum C.B. Clarke and Selinum tenuifolium Wall. *Flora*. (2018a) 248:54–60. doi: 10.1016/j.flora.2018.08.017

41. Bhattarai S, Chaudhary RP, Taylor RSL. Non-Medicinal uses of selected wild plants by the people of Mustang district. *Nepal J Nat Hist Mus*. (2009) 24:48–58. doi: 10.3126/jnhm.v24i1.2237

42. Gewali MB, Awale S. *Aspects of traditional medicine in Nepal*. Toyama, Japan: Institute of Natural Medicine, University of Toyama (2008).

43. Lama AD, Kim J, Martiskainen O, Klemola T, Salminen JP, Tyystjärvi E, et al. Impacts of simulated drought stress and artificial damage on concentrations of flavonoids in *Jatropha curcas* (L.), a biofuel shrub. *J Plant Res*. (2016) 129:1141–50. doi: 10.1007/s10265-016-0850-z

44. Srivastava RP, Dixit P, Singh L, Verma PC, Saxena G. Status of Selinum Spp. L. a Himalayan Medicinal Plant in India: A Review of Its Pharmacology, Phytochemistry and Traditional Uses. *Curr Pharma Bio*. (2018b) 19:1122–34. doi: 10.2174/1389201020666181227150829

45. Padalia RC, Verma RS, Chauhan A, Chanotiya CS, Yadav A. Variation in the volatile constituents of different plant parts of Ligusticopsis wallichiana from western Himalaya, India. *Nat Prod Commun*. (2012) 7:1077–8.

46. Joshi D, Melkani AB, Nailwal MK, Prasad R, Bisht LS. Terpenoid composition of essential oil from a new chemotype of Selinum wallichianum Raizada & Saxena. *Nat Prod Res*. (2018) 32:362–5. doi: 10.1080/14786419.2017.1356831

47. Chauhan RS, Nautiyal MC, Tava A, Mella M. Chemical composition of the volatile oil from the roots of Selinum tenuifolium Wall. *Helv Chim Acta*. (2012) 95:780–3. doi: 10.1002/hlca.201100480

48. Dev V, Oka M, Mathela CS, Murari ND, Stevens TH. The volatile constituents of roots of Selinum tenuifolium. *J Nat Prod*. (1984) 47:904–5. doi: 10.1021/np50035a035

49. Sood SS, Chopra MM, Jamwal RK. Essential oil of Selinum tenuifolium. *Indian Perfurmer*. (1978) 22:127.

50. Adhikari B, Devkota HP, Joshi KR, Watanabe T, Yahara S. Two new diacetylene glycosides: bhutkesoside A and B from the roots of Ligusticopsis wallichiana. *Nat Prod Res*. (2016) 30:1577–84. doi: 10.1080/14786419.2015.1118635

51. Abubakar AR, Haque M. (2020). Preparation of Medicinal Plants: Basic Extraction and Fractionation Procedures for Experimental Purposes. *J Pharm Bioallied Sci*. (2020) 12:1–10. doi: 10.4103/jpbs.JPBS_175_19

52. Belkacem N, Djaziri R, Lahfa F, El Haci IA, Boucherit Z. Phytochemical screening and in vitro antioxidant activity of various *Punica granatum* l. Peel extracts from Algeria: A comparative study. *Phytothérapie*. (2014) 12:372–9. doi: 10.1007/s10298-014-0850-x

53. Benariba N, Djaziri R, Bellakhdar W, Belkacem N, Kadiata M, Malaisse WJ, et al. (2013). Phytochemical screening and free radical scavenging activity of *Citrullus colocynthis* seeds extracts. *Asian Pac J Trop Biomed*. (2013) 3:35–40. doi: 10.1016/S2221-1691(13)60020-9

54. Kumar A, Singh N, Mansoori A, Jiwani G, Solanke AU, Kumar R. Evaluation of antioxidant and antimicrobial potential of Thespesia lampas root extracts. *J Exp Biol Agric Sci*. (2021) 9:87–99. doi: 10.18006/2021.9(1).87.99

55. Senhaji CB, ElH M, Hamdouch A, Heimeur N, Adil C, Ferji Z. Phytochemical screening, quantitative analysis and antioxidant activity of Asteriscus imbricatus and Pulicariam auritanica organic extracts. *Int Food Res J*. (2017) 24:2482–9.

56. Maisuthisakul P, Suttajit M, Pongsawatmanit R. Assessment of phenolic content and free radical-scavenging capacity of some Thai indigenous plants. *Food Chem*. (2007) 100:1409–18. doi: 10.1016/j.foodchem.2005.11.032

57. Krivorotova T, Sereikaite J. Determination of fructan exohydrolase activity in the crude extracts of plants. *Electron J Biotechnol*. (2014) 17:329–33. doi: 10.1016/j.ejbt.2014.09.005

58. Clinical and Laboratory Standards Institute (CLSI). *CLSI Document M07-A9. Methods for Dilution Antimicrobial Susceptibility Tests for Bacteria That Grow Aerobically: Approved Standard. 19th* ed. Wayne: CLSI (2012).

59. Clinical and Laboratory Standards Institute (CLSI). *Methods for Dilution Antimicrobial Susceptibility Test for Bacteria That Grow Aerobically. Approved Standard*, (2009). *29*: 1–65. Clinical and Laboratory Standards Institute (CLSI).

60. Mosman T. Rapid Colorimetric Assay for Cellular Growth and Survival, Application to Proliferation and Cytotoxicity Assays. *J Immunol Methods*. (1983) 65:55–63. doi: 10.1016/0022-1759(83)90303-4

61. Nemati F, Dehpouri AA, Eslami B, Mahdavi V, Mirzanejad S. Cytotoxic properties of some medicinal plant extracts from mazandaran, iran. *Iran Red Crescent Med J*. (2013) 15:e 8871. doi: 10.5812/ircmj.8871

62. Selvamani P, Sen DJ, Gupta JK. Pharmacognostical standardization of Commiphora berryi (Arn) Engl and phytochemical studies on its crude extracts. *Afr J Pharm Pharmacol*. (2009) 3:37–46.

63. Jain M, Kapadia R, Albert S, Mishra SH. Standardization of *Feronia limonia* L. leaves by HPLC, HPTLC, physicochemical and histological parameters. *Bol Latinoam Caribe Plantas Med Aromat*. (2011) 10:525–35.

64. Rahman MM, Islam MB, Biswas M, Khurshid Alam AH. In vitro antioxidant and free radical scavenging activity of different parts of *Tabebuia pallida* growing in Bangladesh. *BMC Res Notes*. (2015) 8:621. doi: 10.1186/s13104-015-1618-6

65. Aligiannis N, Kalpoutzakis E, Chinou IB, Mitakou S, Gikas E, Tsarbopoulos A. Composition and antimicrobial activity of the essential oils of five taxa of Sideritis from Greece. *J Agric Food Chem*. (2001) 49:811–5. doi: 10.1021/jf001018w

66. Iloki-Assanga SB, Lewis-Luján LM, Lara-Espinoza CL, Gil-Salido AA, Fernandez-Angulo D, Rubio-Pino JL, et al. Solvent effects on phytochemical constituent profiles and antioxidant activities, using four different extraction formulations for analysis of *Bucida buceras* L. and *Phoradendron californicum*. *BMC Res Notes*. (2015) 8:396. doi: 10.1186/s13104-015-1388-1

67. Akšić MF, Tosti T, Sredojević M, Milivojević J, Meland M, Natić M. Comparison of sugar profile between leaves and fruits of blueberry and strawberry cultivars grown inorganic and integrated production system. *Plan Theory*. (2019) 8:205. doi: 10.3390/plants8070205

68. Halford NG, Curtis TY, Muttucumaru N, Postles J, Mottram DS. Sugars in crop plants. *Ann Appl Biol*. (2011) 158:1–25. doi: 10.1111/j.1744-7348.2010.00443.x

69. Kumar A, Kumar R, Singh N, Mansoori A. Regulatory framework and policy decisions for genome-edited crops. In: A Bhattacharya, V Parkhi and B Char, editors. *CRISPR/Cas Genome Editing: Concept and Strategies in Plant Sciences*. (2020). 193–201.

70. Devkota HP, Adhikari B, Watanabe T, Yahara S. Nonvolatile chemical constituents from the leaves of Ligusticopsis wallichiana (DC.) Pimenov & Kljuykov and their free radical-scavenging activity. *J Anal Methods Chem*. (2018):1794650. doi: 10.1155/2018/1794650

71. Saraswat MK, Magruder JT, Crawford TC, Gardner JM, Duquaine D, Sussman MS, et al. Preoperative *staphylococcus aureus* screening and targeted decolonization in cardiac surgery. *Ann Thorac Surg*. (2017) 104:1349–56. doi: 10.1016/j.athoracsur.2017.03.018

72. Rana PS, Saklani P, Chandel C. Influence of altitude on Secondary Metabolites and antioxidant activity of Coleus forskohlii root extracts. *Res J Med Plant*. (2020) 14:43–52. doi: 10.3923/rjmp.2020.43.52

73. Souhir K, Dorra S, Ines BR, Mattew S, Neila T, Marie-Laure F, et al. Influence of climate variations on phenolic composition and antioxidant capacity of *Medicago minima* populations. *Sci Rep*. (2020):8293. doi: 10.1038/s41598-020-65160-4

74. Khalil N, El-Jalel L, Yousif M, Gonaid M. Altitude impact on the chemical profile and biological activities of Satureja thymbra L. essential oil. *BMC Complement Med Ther*. (2020) 20:186. doi: 10.1186/s12906-020-02982-9

Glossary

HPLC	High-pressure liquid chromatography
DNSA	3,5-dinitrosalicylic acid
MIC	Minimum inhibitory concentrations
DPPH	2,2-diphenyl-1-picrylhydrazyl
MTT	(3-(4,5-dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide)
A549	Adenocarcinomic human alveolar basal epithelial cells
NBRI	National Botanical Research Institute
GPS	Global positioning system
GBPIHED	Govind Ballabh Pant Institute of Himalayan Environment and Development
Hx	Hexane
Chl	Chloroform
PTE	Petroleum ether
TPC	Total phenolic content
TFC	Total flavonoid content
GA	Gallic acid
MTCC	Microbial type culture collection and gene bank
CLSI	Clinical and Laboratory Standards Institute
DMSO	Dimethyl sulfoxide
PDA	Potato
Dextrose	Agar
ATCC	American type culture collection
ELISA	Enzyme-linked immuno-sorbent assay
EDTA	Ethylene diamine tetra-acetic acid

frontiers | Frontiers in Nutrition

TYPE Original Research
PUBLISHED 26 September 2023
DOI 10.3389/fnut.2023.1228172

Check for updates

OPEN ACCESS

EDITED BY
Tarun Belwal,
Texas A&M University, United States

REVIEWED BY
Ajmer Singh Grewal,
Guru Gobind Singh College of Pharmacy, India
Charu Lata,
CSIR-National Institute of Science Communication and Policy Research, India

*CORRESPONDENCE
Nirupma Singh
✉ nirupmasingh@rediffmail.com

RECEIVED 24 May 2023
ACCEPTED 07 August 2023
PUBLISHED 26 September 2023

CITATION
Nagre K, Singh N, Ghoshal C, Tandon G, Iquebal MA, Nain T, Bana RS and Meena A (2023) Probing the potential of bioactive compounds of millets as an inhibitor for lifestyle diseases: molecular docking and simulation-based approach.
Front. Nutr. 10:1228172.
doi: 10.3389/fnut.2023.1228172

COPYRIGHT
© 2023 Nagre, Singh, Ghoshal, Tandon, Iquebal, Nain, Bana and Meena. This is an open-access article distributed under the terms of the Creative Commons Attribution License (CC BY). The use, distribution or reproduction in other forums is permitted, provided the original author(s) and the copyright owner(s) are credited and that the original publication in this journal is cited, in accordance with accepted academic practice. No use, distribution or reproduction is permitted which does not comply with these terms.

Probing the potential of bioactive compounds of millets as an inhibitor for lifestyle diseases: molecular docking and simulation-based approach

Kajal Nagre[1], Nirupma Singh[1]*, Chandrika Ghoshal[2], Gitanjali Tandon[3], Mir Asif Iquebal[3], Tarsem Nain[4], Ram Swaroop Bana[5] and Anita Meena[6]

[1]Division of Genetics, ICAR-Indian Agricultural Research Institute, Pusa Campus, New Delhi, India, [2]Division of Vegetable Science, ICAR-Indian Agricultural Research Institute, Pusa Campus, New Delhi, India, [3]Centre for Agricultural Bioinformatics, ICAR-Indian Agricultural Statistics Research Institute, Pusa Campus, New Delhi, India, [4]Department of Genetics, Maharshi Dayanand University, Rohtak, India, [5]Division of Agronomy, Indian Agricultural Research Institute, Pusa Campus, New Delhi, India, [6]ICAR-Central Institute for Arid Horticulture, Beechwal, Bikaner, India

Millets are becoming more popular as a healthy substitute for people with lifestyle disorders. They offer dietary fiber, polyphenols, fatty acids, minerals, vitamins, protein, and antioxidants. The nutritional importance of millets leads to the present *in-silico* study of selective bioactive compounds docked against the targets of lifestyle diseases, *viz*., diabetes, hypertension, and atherosclerosis using molecular docking and molecular simulations approach. Pharmacokinetic analysis was also carried out to analyse ADME properties and toxicity analysis, drug-likeliness, and finally target prediction for new targets for uncharacterized compounds or secondary targets for recognized molecules by Swiss Target Prediction was also done. The docking results revealed that the bioactive compound flavan-4-ol, among all the 50 compounds studied, best docked to all the four targets of lifestyle diseases, *viz*., Human dipeptidyl peptidase IV (−5.94 kcal mol^{-1} binding energy), Sodium-glucose cotransporter-2 (−6.49 kcal mol^{-1}) diabetes-related enzyme, the Human angiotensin-converting enzyme (−6.31 kcal mol^{-1}) which plays a significant role in hypertension, and Proprotein convertase subtilisin kexin type 9 (−4.67 kcal mol^{-1}) for atherosclerosis. Molecular dynamics simulation analysis substantiates that the flavan-4-ol forms a better stability complex with all the targets. ADMET profiles further strengthened the candidature of the flavan-4-ol bioactive compound to be considered for trial as an inhibitor of targets DPPIV, SGLT2, PCSK9, and hACE. We suggest that more research be conducted, taking Flavon-4-ol into account where it can be used as standard treatment for lifestyle diseases.

KEYWORDS

secondary metabolite, millet, molecular docking, MD simulations, drug likeness, admet

1. Introduction

Millions of individuals throughout the world suffer from the chronic condition of obesity and diabetes, both of which have significant social costs due to their high incidence rates. Insulin resistance, elevated levels of oxidative stress, and increased expression of inflammatory markers are all prominent symptoms of the complicated condition known as obesity, which results in increased body fat mass. A metabolic condition known as diabetes mellitus (DM) is characterized by decreased insulin secretion and dysfunction of pancreatic cells. Metabolic disorders such as diabetes mellitus (DM), hypertension, cardiovascular disease, and obesity all result from being overweight (1). Obesity is currently a global problem; it is a condition where having too much body fat raises the likelihood of developing health issues, which increase the chance of developing chronic illnesses including diabetes and heart disease. The prevalence of obesity and diabetes has increased dramatically over the past few decades as a result of the growing consumption of processed junk food. Diabetes, heart disease, stroke, gall bladder disease, fatty liver, rheumatoid arthritis, and joint diseases are only a few of the long-term health concerns associated with obesity. Foods rich in dietary fibers, beneficial bioactive compounds, and complex carbohydrates are in greater demand due to these health issues (2). The high amount of gluten in cereals makes it difficult to generate nutritious foods or nutraceuticals even though research is being done to biofortify wholegrain cereals like wheat and rice with phenolic acids that have antimutagenic, anti-glycemic, and antioxidative effects (3–5) There is an urgent need to locate new sources of nutraceuticals, natural foods, and other dietary supplements considering the growing lifestyle diseases and the expanding public knowledge of health care and nutrition.

Plant-based medicines are created from unprocessed plant extracts that are complex blends of several phytochemicals. These phytochemicals are used to treat both chronic and infectious disorders because of their unique and complicated biological effects. Even though there is a huge variety of bioactive secondary metabolites present in different plant species, only a small portion of them have undergone extensive research and have been shown as significant sources of bioactive substances. Also, the idea of treating diabetes, obesity, and related diseases with natural treatments has not received much attention. There are more than 5,000 naturally occurring flavonoids that have been identified in a variety of plants. Many studies showed the potential health advantages of natural flavonoids in the treatment of diabetes mellitus (DM) and obesity, and they reveal higher bioavailability and activity on numerous molecular targets. Flavonoids are divided into six main subgroups: flavanols (which include quercetin, kaempferol, and myricetin), flavanones (which include eriodictyol, hesperetin, and naringenin), flavonoids (which include daidzein, genistein, and glycitein), flavones (which include apigenin and luteolin), flavan (including cyanidin, peonidin, and petunidin). Flavonoids may be helpful in the treatment, prevention, and mitigation of a variety of viral illnesses as well as degenerative illnesses like cancer, diabetes, obesity, and other age-related illnesses (6, 7). According to accumulated epidemiological data, dietary flavan-3-ols have a significant effect in lowering the risk of Type II Diabetes Mellitus (8, 9).

Millets are the sixth most-grown cereals in the world, including pearl millet, foxtail millet, finger millet, and other minor millets. Due to their distinct characteristics of being a C_4 plant with high photosynthetic efficiency, high capacity for producing dry matter, and ability to grow under the most challenging agro-climatic conditions where other crops like sorghum and maize fail to yield, millets outperform all other cereals (10). The millets renamed as Nutri cereals contain alkaloids, flavonoids, terpenes, polyphenols, etc., compared to many other kinds of cereal, including barley, rice, maize, and wheat. Because of their distinctive and complex biological effects, bioactive substances are employed to treat both chronic and infectious diseases. The millet-identified bioactive constituents include gallic acid, protocatechuic acid, p-hydroxybenzoic acid, vanillic acid, syringic acid, ferulic acid, trans-coumaric acid, caffeic acid, sinapic acid, quercetin, and proanthocyanidins (condensed tannins) (11, 12). Even though plant species have a wide range of bioactive secondary metabolites, only a small percentage of them have undergone in-depth study. Recent *in-silico* studies have described the promising effects of bioactive compounds of millets in treating metabolic diseases like diabetes and obesity, hypertension, and cardiovascular disease (13–16).

The identification of bioactive compounds is greatly aided by *in silico* investigations, which also provide several benefits, including minimization of time and expense required for this process. It can be costly and time-consuming to synthesize and test several chemicals as is required by traditional experimental procedures. On the other hand, *in silico* techniques make use of computer simulations and computational models to screen and forecast the activity of hundreds or even millions of chemicals, reducing the number of prospective candidates for further experimental validation. By offering information on molecular interactions, lowering expenses, and improving the likelihood of finding possible therapeutic drugs, they supplement experimental approaches. In the present study, four proteins were used as target receptors Human Dipeptidyl Peptidase (DPPIV), Sodium-Glucose Cotransporter-2 (SGLT-2), Human Angiotensin-Converting Enzyme (hACE), and Proprotein Convertase Subtilisin Kexin type 9 (PCSK9). The aim of the study is to explore bioactive compounds with antidiabetic, antihypertension, and antiatherosclerosis properties in millets using *in-silico* approaches. The selected bioactive compounds were analyzed for pharmacokinetics and physiochemical properties, and MD simulations were carried out to evaluate the binding stability, conformation, and interactive ways between the ligands and target protein. Hence, we aim to investigate the pharmacological activities of the bioactive compounds from millets against diabetes mellitus, atherosclerosis, and hypertension through pharmacokinetics and pharmacological properties, and molecular modeling methods.

2. Materials and methods

2.1. Macromolecule as target preparation

The three-dimensional X-ray crystallographic structures for targets that play an extremely significant role in diabetes mellitus, atherosclerosis, and hypertension were retrieved from The Research

TABLE 1 Target with their Protein Data bank ID (PDBID) and amino acid chain used for molecular docking.

Sr. No.	PDB ID	Target or Macromolecule	Chain with Amino acid
1	7VSI	Sodium-glucose Cotransporter-2 (SGLT-2)	A chain 672AA
2	1J2E	Human Dipeptidyl Peptidase IV(DPPIV)	A chain 729 AA
3	2PMW	Proprotein convertase subtilisin Kexin type 9 (PCSK9)	A chain 126 AA
4	2XY9	Human Angiotensin-converting enzyme (ACE)	A chain 585 AA

Collaboratory for Structural Bioinformatics (RCSB)[1] and saved in PDB format (Table 1). The water molecules were deleted and polar hydrogen was added with correct partial charges in the target. The four targets chosen were Human Dipeptidyl Peptidase (DPPIV), Sodium-Glucose Cotransporter-2 (SGLT-2) for diabetes, Human Angiotensin-Converting Enzyme (hACE) for hypertension, and Proprotein Convertase Subtilisin Kexin type 9 (PCSK9) for atherosclerosis based on literature search because these proteins play important roles in lifestyle diseases and the search for their inhibitors from millets will help in dietary interventions (13–16).

2.2. Ligand selection

The bioactive compounds such as carotenoids, alkaloids, flavonoids, coumarin, and phenol were docked for antidiabetic, antihypertension, and antiatherosclerosis properties based on a literature search. Their structures were retrieved from ZINC DATABASE and PubChem in SDF format which were later converted into PDB format using OPEN BABLE. The 2D structure, PubChem ID, and type of bioactive compound details are given in Table 2. The FDA-approved drug empagliflozin was used as an antidiabetic positive control, while ramipril and atorvastatin were used as standards for antihypertension and antiatherosclerosis studies, respectively (17, 18).

2.3. Molecular docking

The process of "molecular docking" explores the potential links between molecules interacting under topographical restrictions or energy considerations to match the two molecules to the optimal interaction conformation (28). Molecular docking is the most modern, efficient, and cost-effective method for creating and testing pharmaceutical compounds (7). Computer-aided tools have become sophisticated drug discovery techniques that can be used to filter drugs from bioactive chemicals found in a variety of therapeutic plants (29). In the present study, in-silico molecular docking is done by using the software Autodock Vina and Autodock Tools. In this study, ligands were kept flexible while proteins were kept rigid. AutoDock was used to create grid boxes to prepare PDBQT files for proteins and ligands, among other intermediary stages. Docking is performed using 10 runs of the Lamarckian Genetic Algorithm. The grid was placed in a box whose coordinates were x, y, and z having 2,048,383 total grid points per map with a spacing of 1.0 Å. The dimensions of the grid box were set as $X = 126$, $Y = 126$, and $Z = 126$, and the center grid box was set with the coordinates as center $x = 48.648$, center $y = 59.931$, and center $z = 31.936$. We used the feature-rich molecular modeling application Discovery Studio Visualizer to view, share, and analyze data. The best docking outcomes were analyzed using Biovia Discovery Studio Visualizer, which is also useful for seeing and assessing predicted protein–ligand interactions.

2.4. Molecular dynamics simulations

MD simulation is important to practice the existing drug discovery developments as it assists in a better understanding of molecular structure-to-function relationships (30, 31). In this work, MD simulations have been carried out to evaluate the binding stability, conformation, and interactive ways between the ligands and target protein. This simulation was investigated for receptor-ligand complexes for 100 ns via GROMACS (Groningen Machine for Chemical Simulations) software version 2021 (31, 32). GROMACS is a non-commercial molecular simulation package that is useful in performing simulations of proteins, lipids, and nucleic acids (32). Firstly, the topology of the protein and the ligand was created using CHARMM36 force fields (33). For protein topology, the generation GROMACS program was used, while for ligand topology, CGenFF was used.[2] CGenFF (CHARMM General Force Field) program performs atom typing and assignment of parameters and charges by analogy in a completely automated fashion (33, 34). The complex was immersed in a dodecahedron box of simple point charge (SPC) water molecules. The solvated system was neutralized by adding counter-ions. Energy minimization of the solvated structures was done using the steepest descent and conjugate gradient algorithm till the maximum force reached below 100 KJ/mol/nm. To equilibrate, the system was then subjected to position–restrained dynamics simulation (NVT and NPT) at 300 K for 100 ns. Finally, this system was subjected to the MD production run for 100 ns at 300 K temperature and 1 bar pressure. For trajectory analysis, various parameters were computed using GROMACS. These included Root Mean Square Deviation (RMSD), Root Mean Square Fluctuation (RMSF), Potential Energy, SASA (solvent accessible surface area), and Molecular mechanics Poisson–Boltzmann surface area (MM/PBSA) calculation.

2.5. Pharmacokinetics properties

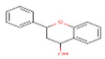
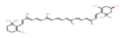
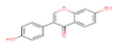
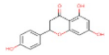
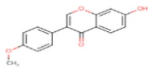
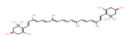
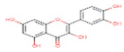
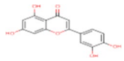
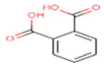
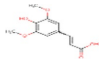
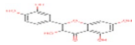
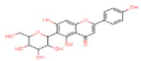
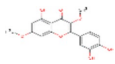
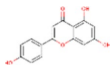
2.5.1. ADME prediction

Medicinal chemistry and pharmacokinetics of small compounds are demonstrated by ADME. Conducting drug metabolism and pharmacokinetics (DMPK) research, also known as ADME (Absorption, Distribution, Metabolism, and Excretion) investigations,

1 https://www.rcsb.org/

2 https://cgenff.µMaryland.edu/

TABLE 2 The 50 compounds studied with their 2D structure and PubChem ID.

Sr. No.	Compound	PubChem Id	2D Structure	Type of bioactive compound	Reference
1	Flavan-4-ol	439712		Flavonoid	(11)
2	β-cryptoxanthin	5281235		Carotenoid	(12)
3	Daidzein	5281708		Isoflavone	(19)
4	Naringenin	932		Flavone	(20)
5	Formononetin	5280378		Isoflavone	(19)
6	Violaxanthin	6384269		Carotenoid	(20)
7	Zeaxanthin	5280899		Carotenoid	(11)
8	Quercetin	5280343		Flavonoid	(20)
9	Luteolin	5280445		Flavones	(20)
10	Phthalic acid	1017		Benzene dicarboxylic acid	(21)
11	Trans- sinapic acid	637775		Hydroxycinnamic acids	(22)
12	Dihydroquerectin	471		Pentahydroxy flavone	(20)
13	p-coumaric acid	637542		Phenylpropanoid	(20)
14	Caffeic acid	689043		Polyphenols	(22)
15	3,4Dihydroxybenzoic acid	72		Benzoic acid	(20)
16	Isovitexin	162350		Glycosyl compound	(20)
17	Anthocyanin	56928084		Flavonoids	(20)
18	3, 7 Dimethylquerctin	5280417		Flavonoids	(23)
19	Apigenin	5280443		Flavonoids	(9)
20	Tricin	5281702		Flavonoids	(20)
21	Isorhamnetin-3-0-glucoside	5318645		Sesquiterpenoid	(24)

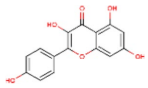
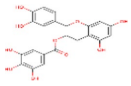
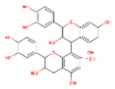
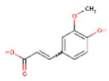
(Continued)

TABLE 2 (Continued)

Sr. No.	Compound	PubChem Id	2D Structure	Type of bioactive compound	Reference
22	Xylotriose	91873341		Oligosaccharide	(25)
23	Lucenin-1	44257923		Flavones	(23)
24	Orientin	5281675		Flavonoids	(19)
25	Myricetin	5281672		Flavonoids	(22)
26	Vitexin	5280441		Flavone glycoside	(22)
27	Syringic acid	10742		Benzenoids	(20)
28	Hexadecanoic acid	5282743		Saturated fatty acids	(21)
29	Chlorogenic acid	1794427		Hydrocinnamoyl derivatives	(20)
30	Xylotetraose	10230811		Tetrasaccharide	(25)
31	Gentisic acid	3469		Hydroxybenzoic acid	(22)
32	α tocopherol	86472		Phenols	(25)
33	Saponarin	441381		Flavonoid	(19)
34	Xylobiose	439538		Glycoside hydrolase	(26)
35	Vanillic acid	8468		Benzoic acid derivative	(22)
36	Epigallocatechin	72277		Flavonoid	(22)
37	Catechin	9064		Flavonoid	(19)
38	γ- tocopherol	92729		Phenols	(27)

(Continued)

TABLE 2 (Continued)

Sr. No.	Compound	PubChem Id	2D Structure	Type of bioactive compound	Reference
39	Gallic acid	370		Phenolic acid	(22)
40	Kaempherol	5280863		Flavonoids	(20)
41	Epicatechin	107905		Flavonoid	(22)
42	Procyanidin B1	11250133		Polyphenols	(22)
43	Cinnamic acid	444539		Unsaturated carboxylic acid	(20)
44	Violanthin	442665		Flavonoid	(22)
45	Proanthocyanidin	107876		Polyphenol	(22)
46	Ferulic acid	445858		Hydroxycinnamic acids	(20)
47	Atropine	174174		Alkaloid	(26)
48	p-Hydroxybenzoic acid	135		Phenolic acid	(22)
49	Scopolamine	3000322		Alkaloid	(26)
50	Proanthocyanidin	107876		Condensed tannins	(22)

is a crucial step in the discovery and development of new drugs (18). SWISSADME and admetSAR open-source tools are used for ADME analysis (35).

2.5.2. Toxicity analysis

The ProTox-II website studies the toxicity LD_{50} value and toxicity class. The LD_{50} is the fatal dose at which 50% of the tested population is fatal after ingesting the substance. By providing the SMILES from PubChem, the appropriate chemical can be studied in the online database known as pkCSM.[3] The website offers information such as whether a substance is Ames positive and thus mutagenic. It forecasts whether a specific substance is linked to skin sensitivity or not and suggests whether a particular drug will affect liver functions or not (18).

2.5.3. Drug-likeness properties

The Drug-likeness property is predicted by Lipinski's Rule. To determine whether a chemical compound has pharmacological or biological activities as an orally active drug in humans, Lipinski's Rule, a refinement of drug-likeness, is used.

2.5.4. Target prediction

The observed phenotypic effects are the result of the activity of bioactive small molecules, such as metabolites, being modulated by their binding to proteins or other macro-molecular targets. To

3 http://biosig.unimelb.edu.au/pkcsm/

understand the molecular processes behind the bioactivity of bioactive small compounds and to foretell any adverse effects or cross-reactivity, it is crucial to map their targets. We can computationally find new targets for uncharacterized compounds or secondary targets for recognized molecules. Swiss Target Prediction[4] is a web service that uses a collection of 2D and 3D similarity measurements with known ligands to precisely predict the targets of bioactive compounds (36).

3. Result

3.1. Molecular docking

The docking scores indicate how well the ligands fit into the active site of the target, and the more negative the value, the better the affinity of both ligand and target. The Autodock tools are used for ligand-target interactions and are used to analyze the outcome of docked compounds (37). In our study, 50 bioactive compounds docked against the four targets Human Dipeptidyl Peptidase (DPPIV), Sodium-Glucose Cotransporter-2 (SGLT-2) for diabetes, Human Angiotensin-Converting Enzyme (HACE) for hypertension, and Proprotein Convertase Subtilisin Kexin type 9 (PCSK9) for atherosclerosis. Out of 50 compounds, 20 compounds showed better binding energies with the chosen targets against their standard as shown in Figure 1.

3.1.1. Target I-DPPIV

Molecular docking facilitates the evaluation of the biological effects of small compounds by predicting binding affinity against the target protein. In the current docking studies of 50 bioactive compounds that docked against the DPPIV, the apigenin, zeaxanthin, flavan-4-ol, and violaxanthin showed better binding energy −6.29 kcal mol^{-1}, −6.20 kcal mol^{-1}, −5.94 kcal mol^{-1}, −5.29 kcal mol^{-1}, respectively than standard empagliflozin −4.65 kcal mol^{-1}. Of the inhibitors studied, the number of hydrogen bonds formed between the target and inhibitor depicts the stable complex formation; apigenin formed four hydrogen bonds with the target, and the rest of the compound formed one hydrogen bond. The amino acid involved in this bond formation were Pro475, Pro510, Lys512, Ile529, Phe559, Arg560, and Asn562, with 24.26 μM inhibition constant. Zeaxanthin hydrogen bond formation involved amino acids Ile102, Val121, Lys122, Phe240, and Ala707, with 28.44 μM inhibition constant, whereas flavan-4-ol formed bond with Leu90, Asn92, Phe95. Ile102, and 44.17 μM inhibition constant. Violaxanthin formed a bond with Val303, Ala465, and Lys466 amino acid residues and 132.40 μM inhibition constant as shown in Table 3.

3.1.2. Target II – SGLT-2

All the bioactive compounds that are docked against the macromolecule SGLT-2, from the flavan-4-ol, daidzein, luteolin, and naringenin showed more negative binding energies −6.49 kcal mol^{-1}, −5.97 kcal mol^{-1}, −5.95 kcal mol^{-1}, and − 5.85 kcal mol^{-1}, respectively than standard Empagliflozin with −4.67 kcal mol^{-1} value. The more the negative value, the more stability in the complex. The flavan-4-ol formed two hydrogen bonds with the amino acids Asn75, His80, Phe98, Glu99, Val157, Tyr290, and Gln457 with 17.57 μM inhibition constant. Daidzein formed three hydrogen bonds with amino acid Ser156, Val144, Ala446, Ala447, Gln451, Leu452, Tyr455, and Phe504 with 41.73 μM inhibition constant. The luteolin also formed three hydrogen bonds with amino acids Asn75, His80, Thr87, Val95, Phe98, and Glu99 with a 43.47 μM inhibition constant. Naringenin formed one hydrogen bond with Val444, Ala446, Ala447, Gln451, Leu452, Tyr 455, and Phe504 amino acids with 51.60 μM inhibition constant. All the values are depicted in Table 3.

3.1.3. Target III-hACE

In the docking analysis, violaxanthin, zeaxanthin, flavan-4-ol, and daidzein showed the property of hACE inhibitors. The more negative the binding energy value the more stable the complex formed. Compared to conventional Ramipril (−4.01 kcal mol^{-1}), the bioactive compound violaxanthin (−6.96 kcal mol^{-1}), zeaxanthin (−6.64 kcal mol^{-1}), flavan-4-ol (−6.31 kcal mol^{-1}), and daidzein (−5.77 kcal mol^{-1}) displayed higher binding energies with the target. The violaxanthin interaction with the target formed one hydrogen bond with the involvement of Leu194, Pro198, Lys478, Tyr481, Trp486, Pro500, Arg501, and Tyr619 amino acid residues and 7.89 μM inhibition constant. The zeaxanthin also formed one hydrogen bond with Glu134, Ile138, Leu194, Phe196, Pro198, Lys199, and Lys478 aa residues with 13.65 μM inhibition constant. Flavan-4-ol binds with the Trp59, Tyr62, Ala63, Asn66, Ala356, Trp357, Asp358, and Tyr360 and forms three conventional hydrogen bonds with the target with 23.70 μM inhibition constant. The Daidzein also interacted strongly with the target forming three hydrogen bonds with Asp121, Glu123, Arg124, Tyr135, Leu139, Ile204, Ala207, Ala208, Ser219, and Ser516 amino acid residues having inhibition constant of 5.77 μM as shown in Table 3.

3.1.4. Target IV – PCSK9

In the present study, the inhibitor phthalic acid showed more binding interaction with target PCKS9 than standard Atorvastatin −2.37 kcal mol^{-1} with −5.69 kcal mol^{-1} binding energy. The interaction was made by five hydrogen bonds with amino acid residues Arg97, Gln101, Arg104, and Arg105 with an inhibition constant of 67.61 μM. Whereas, naringenin formed three hydrogen bonds with Ala 62, Lys 83, Leu 135, and Lys 136 amino acid residues and − 4.89 kcal mol^{-1} binding energy, with 259.51 μM inhibition constant. Daidzein binds with the target and formed three hydrogen bonds with Glu84, Leu 88, Gly11, and Pro120 amino acids and had 290.97 μM inhibition constant. The flavan-4-ol with −4.67 kcal mol^{-1} binding energy binds with amino acid residues Ala62, Leu135, and Lys136 forming one hydrogen bond with 378.10 μM inhibition constant, as shown in Table 3.

In the *in-silico* study, stronger and more stable contact between the ligand and the target molecule is revealed by low-binding energy. The more negative the value of binding energy the more robust the complex. From the docking result, the top compounds flavan-4-ol, violaxanthin, zeaxanthin, apigenin, daidzein, luteolin, and naringenin, which are bioactive compounds, showed better binding against the positive control with the chosen targets. The flavan-4-ol is the only common molecule that docked against all four targets and established a stable complex. The binding interactions, hydrophobic interactions, and hydrogen bond formation between flavan-4-ol with the selected targets are shown in Figures 2–5.

4 http://www.swisstargetprediction.ch/

FIGURE 1
The top 20 compounds docked against the four selected targets Hamann Dipeptidyl Peptidase (DPPIV), Sodium-Glucose Cotransporter-2 (SGLT-2) Human Angiotensin-Converting Enzyme (ACE), and Proprotein Convertase Subtilisin Kexin type 9 (PCSK9) with their binding energies against their respective positive controls.

3.2. Molecular docking simulation studies

Although protein–ligand docking provides effective information, it only covers the static depiction of the binding conformations of the ligand in the active region of the receptor. Thus, the integration of Newton's equations of motion in the form of molecular dynamics provides an insight into atomic dynamics in the system throughout the defined timeline. For better understanding, an MD simulation of 100 ns was performed on all four complexes. MD trajectories analysis was used to determine the stability and fluctuation patterns of these complexes by using RMSD, RMSF (Root Mean Square Fluctuation), MM/PBSA, Radius of Gyration, and SASA (Solvent Accessible Surface Area) of the receptor atoms (Table 4).

3.2.1. Root means square deviation and root mean square fluctuation

RMSD aids in analyzing the change in the protein structure during simulations while RMSF measures the differences in the structural confirmation of the atoms. RMSD is a crucial measure for analyzing the equilibration in the stability of complex systems during the simulation process. To measure the structural conformational differences, the RMSD of the protein backbone atoms was plotted against time. During the simulation, minor fluctuations were observed for all the complexes; however, complex 2: DPPIV-Flavan-4-ol complex had minimum RMSD as shown in Figure 6A.

RMSF is another vital parameter to consider while simulating the stability and flexibility of complex systems. The RMSF was used to determine how amino acid residues of a target protein changed their behavior upon binding to a ligand. The RMSF values for the protein's C-alpha atoms were computed and showed against the residues. In all complexes, the amino acid residues exhibited very little variation during the simulation apart from compound 4: HACE2-flavan-4-ol complex which showed huge fluctuations and can be graphically viewed in Figure 6B.

3.2.2. Radiation of gyration

Furthermore, the complex system radius of gyration (RGY) was also computed. RGY measures the root mean square distance between the protein's atoms and the rotation axis. Being one of the critical parameters which designate the overall change in the compactness and dimensions of the protein structure throughout the simulation, elevated RGY values indicate that the protein is less compact and

TABLE 3 Molecular interaction profiling and docking score of the top four bioactive compounds against selected targets in comparison with positive controls.

Sr. No.	Compound	Binding Energy (kcal mol^{-1})	Inhibition Constant (uM)	Hydrogen bond	Amino acid involved
Target I – Human Dipeptidyl Peptidase (DPPIV)					
1	Apigenin	−6.29	24.26	4	Pro475, Pro510, Lys512, Ile529 Phe559, Arg560, Asn562
2	Zeaxanthin	−6.20	28.44	1	Ile102, Val121, Lys122, Phe240, Ala707
3	Flavan-4-ol	−5.94	44.17	1	Leu90, Asn92, Phe95. Ile102
4	Violaxanthin	−5.29	132.40	1	Val303, Ala465, Lys466
5	Standard Empagliflozin	−3.73	1.84	4	Val121, Lys 122, Gln 123, Trp124, Tyr 211, Asp 739
Target II – Sodium-Glucose Cotransporter-2 (SGLT-2)					
1	Flavan-4-ol	−6.49	17.57	2	Asn75, His80, Phe98, Glu99, Val157, Tyr290, Gln457
2	Daidzein	−5.97	41.73	3	Ser156, Val144, Ala446, Ala447, Gln451, Leu452, Tyr455, Phe504
3	Luteolin	−5.95	43.47	3	Asn75, His80, Thr87, Val95, Phe98, Glu99
4	Naringenin	−5.85	51.60	1	Val444, Ala446, Ala447, Gln451, Leu452, Tyr 455, Phe504
5	Standard Empagliflozin	−4.67	379.50	1	Tyr410, Glu421, Val425, Leu428, Trp429, Phe432
Target III – Human Angiotensin Converting Enzyme (hACE)					
1	Violaxanthin	−6.96	7.89	1	Leu194, Pro198, Lys478, Tyr481, Trp486, Pro500, Arg501, Tyr619
2	Zeaxanthin	−6.64	13.65	1	Glu134, Ile138, Leu194, Phe196, Pro198, Lys199, Lys478
3	Flavan-4-ol	−6.31	23.70	3	Trp59, Tyr62, Ala63, Asn66, Ala356, Trp357, Asp358, Tyr360
4	Daidzein	−5.77	5.77	3	Asp121, Glu123, Arg124, Tyr135, Leu139, Ile204, Ala207, Ala208, Ser219, Ser516
5	Standard Ramipril	−4.01	1.15	1	Trp59, Tyr62, Ile88, Trp357, Asp358, Tyr360
Target IV – Proprotein Convertase Subtilisin Kexin type 9 (PCSK9)					
1	Phthalic acid	−5.69	67.61	5	Arg97, Gln101, Arg104, Arg105
2	Naringenin	−4.89	259.51	3	Ala62, Lys 83, Leu 135, Lys 136
3	Daidzein	−4.82	290.97	3	Glu84, Leu 88, Gly11, Pro120
4	Flavan-4-ol	−4.67	378.10	1	Ala62, Leu135, Lys136
5	Standard Atorvastatin	−2.37	18.18	1	Ala62, Arg97, Leu135, Lys136, Pro138

flexible, whereas low values indicate that the protein is very compact and inflexible. RGY values of atoms of protein backbone were plotted against time to examine the changes in structural compactness. Protein and protein–ligand complexes showed a gradual decrease in the RGY value throughout the simulation, which revealed that the test molecules induced no major structural changes in the protein (Figure 6C). Energy parameters were of great help while studying the overall stability patterns of the protein in the system. After studying the change of potential energy patterns, it was observed that there were least fluctuations for all the four complexes. The potential energies for all the complexes are shown in Figure 6D.

3.2.3. Solvent accessible surface area

Solvent-accessible surface area (SASA) measures how much of a molecule's area is available to the solvent. It is used to measure the steric availability of an atom. SASA is a significant parameter for examining the degree of receptor exposure to the surrounding solvent molecules during simulation. In general, ligand binding may cause structural changes in the receptor, causing the region in contact with the solvent to alter. SASA values of protein were plotted against time to estimate the changes in surface area. SASA for all four complexes is provided in Table 4 and depicted in Figure 6E.

3.2.4. MM/PBSA – binding free energy analysis

MM/PBSAs are probably highly popular methods for predicting binding free energy because of their superior accuracy compared to most molecular docking scoring functions and lower processing requirements compared to free energy approaches. For biomolecular research on protein folding, protein–ligand binding, protein–protein interaction, and other topics, MM/PBSA has been extensively used. All complexes' binding free energy (ΔG bind) was determined using the MM/PBSA technique for the final 20 ns (80–100 ns) of the

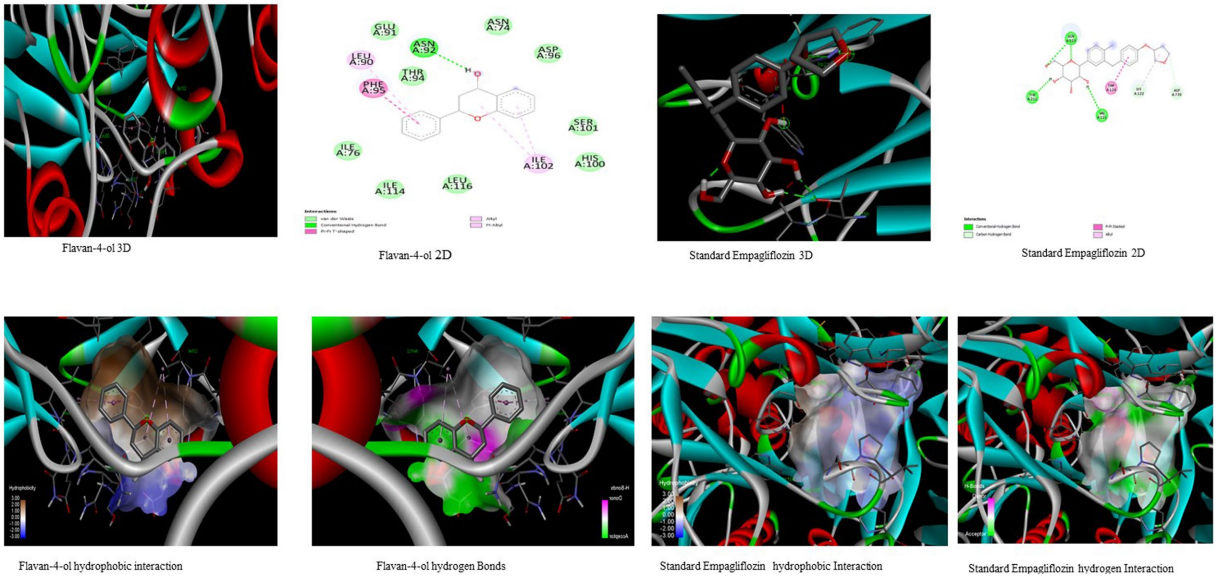


FIGURE 2
Visualization of binding interaction and hydrogen and hydrophobic bonds formed between the flavan-4-ol with Target I Human Dipeptidyl Peptidase (DPPIV) against their positive controls.

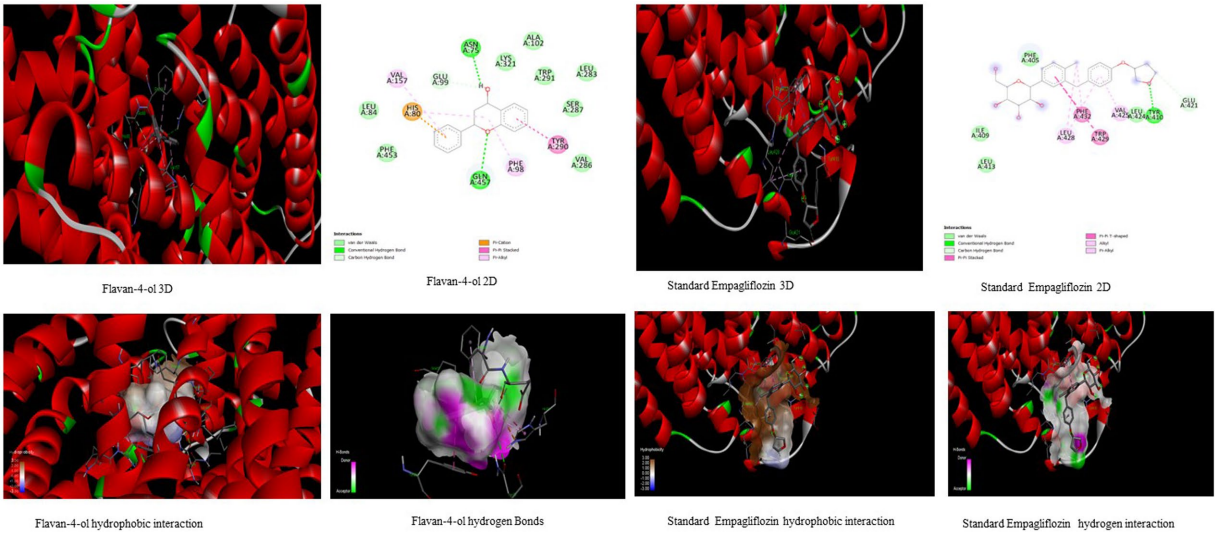


FIGURE 3
Visualization of binding interaction and hydrogen and hydrophobic bonds formed between the flavan-4-ol with Target II Sodium-Glucose Cotransporter2 (SGLT2) against their positive controls.

simulated trajectories with dt 1,000 frames. Low-negative free-binding energies indicate that the test ligands have a strong affinity for binding to the target protein. Binding energies for all the complexes are given in Table 5.

3.3. Pharmacokinetics studies

The SWISSADME, ProTox-II, and pkCSM are used to investigate the pharmacokinetics properties, druglike nature, and medicinal chemistry of substances such as absorption, distribution, metabolism, excretion, and toxicity (ADMET) profiling. In the present study, bioactive substances flavan-4-ol, apigenin, daidzein, luteolin, and phthalic acid were classified as class V, meaning they may be dangerous if ingested in amounts of 2,000 to 5,000 mg/kg, whereas naringenin was classified as class IV, meaning it would be harmful if ingested in amounts of 300 to 2,000 mg/kg. Violaxanthin was classified as class III, meaning it is harmful if consumed at $50 < LD_{50} \leq 300$ mg/kg. Zeaxanthin class II has a lethal dose of $5 < LD_{50} \leq 50$ mg/kg as shown in Table 6. Also, the compounds were found to be neither hepatotoxic nor carcinogenic.

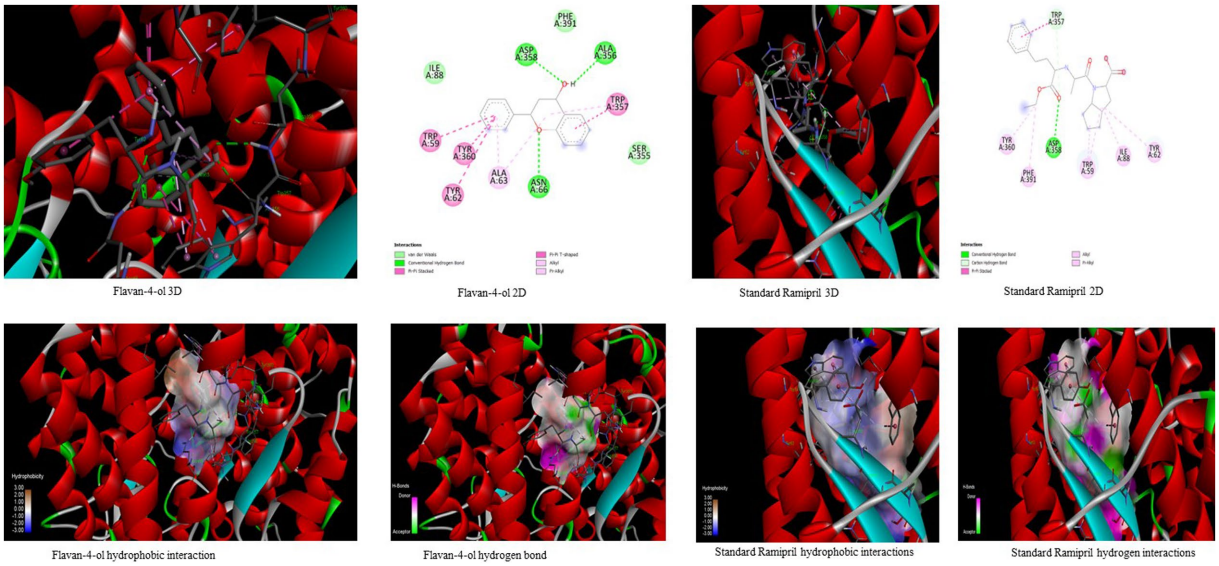


FIGURE 4
Visualization of binding interaction and hydrogen and hydrophobic bonds formed between the flavan-4-ol with Target III Human Angiotensin-Converting enzyme (ACE) against their positive controls.

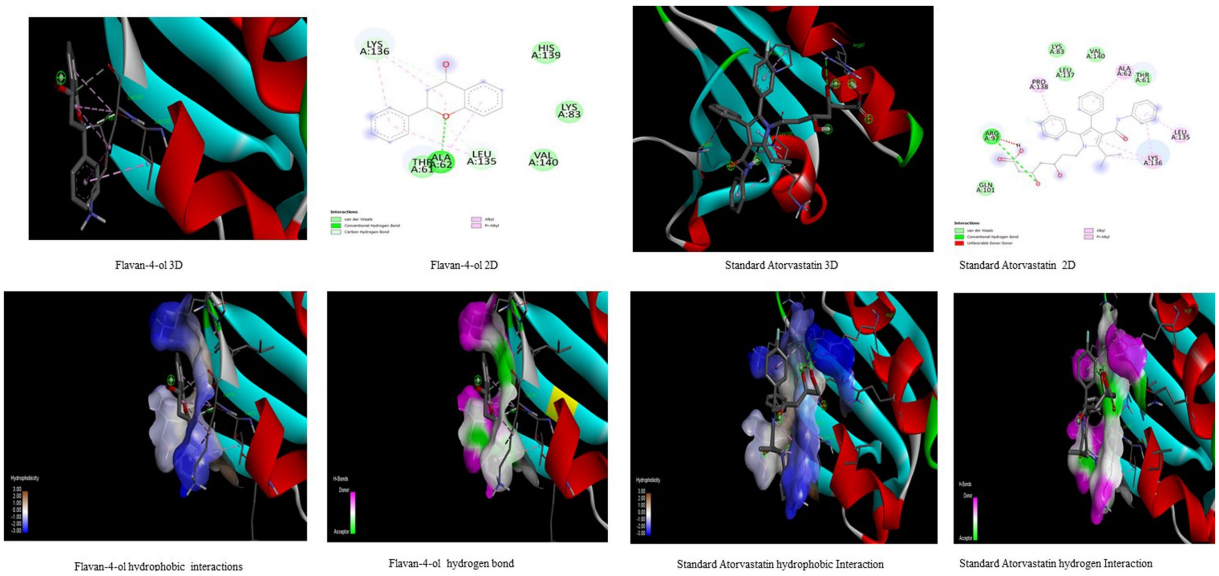


FIGURE 5
Visualization of binding interaction and hydrogen and hydrophobic bonds formed between the flavan-4-ol with Target IV Proprotein Convertase Subtilisin Kexin type 9 (PCSK9) against their positive controls.

The body's ability to absorb bioactive compounds is further influenced by their solubility and stability because of the severe pH values in the stomach and metabolism by gut microbes. A popular molecular descriptor in the investigation of drug transport characteristics, such as intestinal absorption, is Topological Polar Surface Area (TPSA). The TPSA value <140 Å shows good intestinal absorption and the TPSA value <70 Å shows good brain penetration. In the current study, the bioactive compounds luteolin, naringenin, daidzein, and apigenin have values <70 Å which means that they may act as good brain penetration compounds (Table 7).

The ability of bioactive compounds to be consumed in the human system is predicted by Lipinski's rule of five, which is one of the most often-used characteristics of bioactive constituents that resemble pharmacological properties. The characteristics of Lipinski were determined using the canonical SMILES that were taken from the PubChem database. The flavan-4-ol, daidzein, naringenin, quercetin, luteolin, and phthalic acid showed drug-likeness properties as they follow the Lipinski rule. Flavan-4-ol has a maximum tolerated dose of 0.194 log mg/kg/day for human consumption and an acute oral rat toxicity dose of 2.206 mol/kg (Table 6).

TABLE 4 Parameters for MD analysis.

S. No	Protein–ligand Complex	Average RMSD (nm)	Average RMSF (nm)	SASA (nm\S2\N)	Potential Energy (KJ mol^{-1})
1	PCSK9-Flavan-4-ol	0.48	0.0860	61.622	−296419.2922
2	DPP4- Flavan-4-ol	0.22	0.0918	333.168	−1712766.206
3	SGLT2-Flavan-4-ol	0.36	0.1635	245.842	−998130.8998
4	ACE2-Flavan-4-ol	0.93	1.4751	259.702	−2059289.941

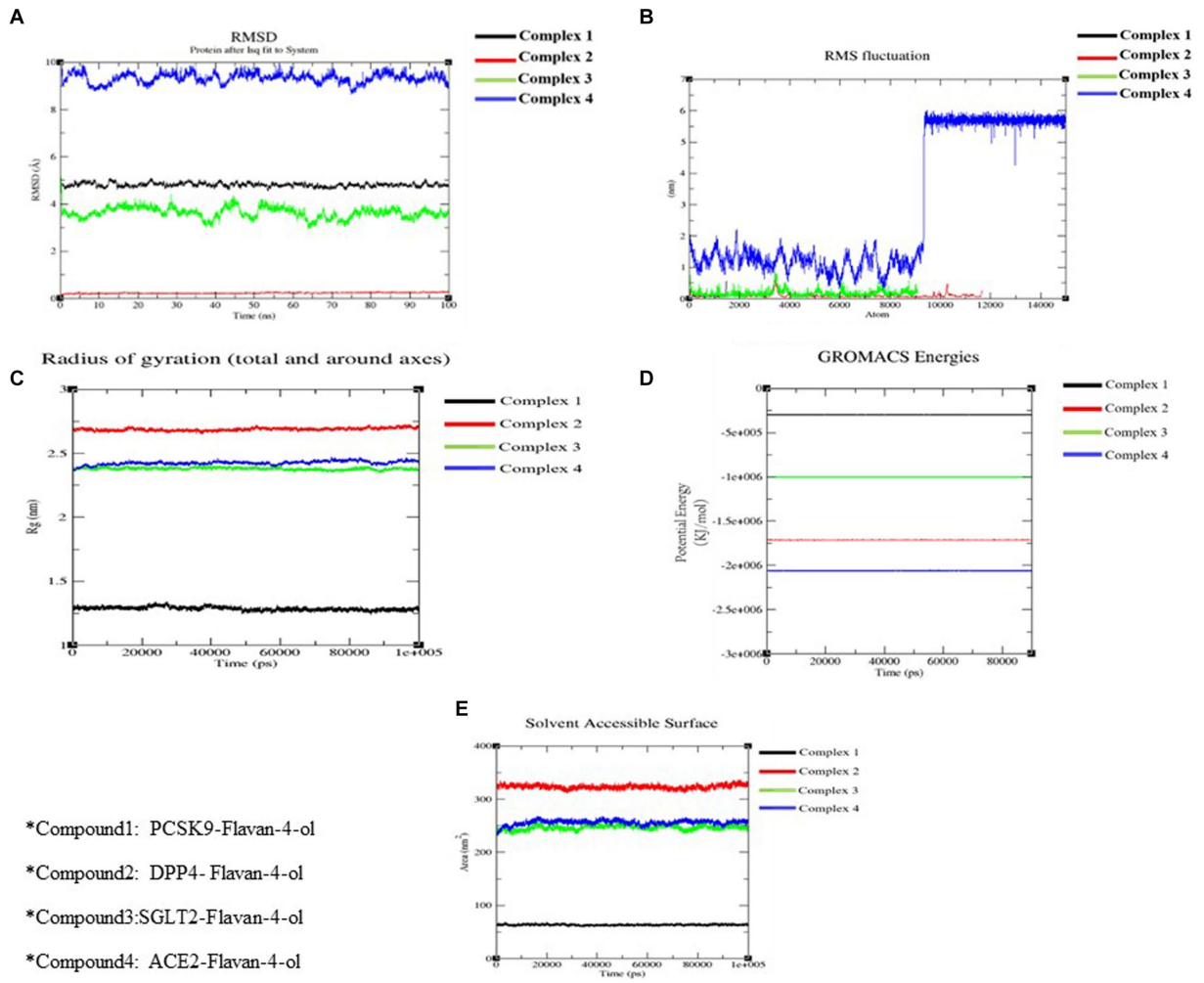


FIGURE 6
Molecular docking simulation analysis. **(A)** RMSD of all the four complexes. **(B)** RMSF of all the four complexes. **(C)** The radius of gyration of all the four complexes. **(D)** The potential energy of all the four complexes. **(E)** SASA of all the four complexes.

3.4. In-depth analysis of identified bioactive compound – Flavan-4-ol as a potential candidate with inhibitory capacity against lifestyle diseases

The bioavailability radar as shown in Figure 7 revealed that the colored zone, which considered features like flexibility, lipophilicity, saturation, size, polarity, and solubility, is the ideal physicochemical region for oral bioavailability (23). According to its physical characteristics, flavan-4-ol has a molecular formula of 226.27 g/mol. There are 17 heavy total atoms, with 12 aromatic heavy atoms. In the sp3 hybridization, 0.2 carbon atoms were present. There were one rotatable bond, two hydrogen bond acceptors, and one hydrogen bond donor. The topological polar surface area was determined to be 29.46 A° and the molar refractivity to be 66.24.

The log Po/w (log P) is 2.35, the log Po/w (Xlog P3) is 2.7, the log Po/w (Wlog P) is 2.6, the log Po/w (MlogP) is 2.54, the log Po/w (SILICOS-IT) is 2.98, and the consensus log Po/w is 2.63. Overall, log *p*-values indicate that the chemical has good lipophilic characteristics. A log S (ESOL) value of −3.4, which indicates that the chemical belongs to the moderately water-soluble class, was used to analyze the substance's water solubility.

The drug-likeness parameter is high as it is following Lipinski, Verber, and Egan rules with a bioavailability score of zero. Swiss ADME Synthetic Accessibility (SA) Score is based primarily on the assumption that the frequency of molecular fragments in 'really' obtainable molecules correlates with the ease of synthesis. The fragmental contribution to SA should be favorable for frequent chemical moieties and unfavorable for rare moieties. The synthetic accessibility score was found to be 3.07 which means it would not be tough to synthesize the molecule. There is no alert for PAINS, indicating the compound is quite specific in nature.

3.4.1. Pharmacokinetics properties

3.4.1.1. Absorption

The pharmacokinetic features of flavan-4-ol were investigated which showed Blood Brain Barrier (BBB+) with a computed probability value of 0.942; flavan-4-ol's permeability to the BBB is 0.559 log BB and its permeability to the central nervous system is −1.676 log PS, both of which indicate a low likelihood of CNS adverse effects. According to its P-glycoprotein I inhibitor and P-glycoprotein II inhibitor scores of 0.79 and 0.88, flavan-4-ol has a low probability of being a Pgp inhibitor. As a result, it is thought to be free of serious medication interactions. A score of one indicates an inhibitor, while a score of zero indicates a non-inhibitor. This output number displays the likelihood that it is a Pgp inhibitor. Its score of 0.60 indicates that it has the lowest likelihood of being a Pgp substrate. A Pgp substrate receives a score of one, while a non-substrate receives a score of zero. According to its score of 1.00, flavan-4-ol is projected to have a low intestine absorption rate in humans (Table 8).

TABLE 5 Binding free-energy calculations of selected complexes using MM-PBSA.

S. No	Protein–ligand Complex	ΔG bind (kJ mol^{-1})
1	PCSK9-Flavan-4-ol	-1,903.97
2	DPP4- Flavan-4-ol	−9,473.91
3	SGLT2-Flavan-4-ol	−3,247.02
4	ACE2-Flavan-4-ol	−8,699.47

3.4.1.2. Distribution

To analyze the distribution, the unbound fraction in plasma (Fu), the volume of distribution (VD), and the blood–brain barrier (BBB) permeability were taken into consideration (Tables 8B,C). The computed volume of distribution (VD) was 0.478 L/kg. The range between 0.04 to 20 L/kg is ideal for VD (38). The score of 0.559 indicates that the flavan-4-ol has a greater probability of blood–brain barrier penetration. The output value of 0.9424 is the likelihood of successfully crossing BBB. Calculations revealed that the plasma's unbound fraction (Fu) was 0.079. This suggests that more unbound plasma fractions available for pharmacological activity.

3.4.1.3. Metabolism

From the perspective of drug plasma concentration, this parameter is crucial. The database categorizes ligands as either category 0 (non-inhibitor) or category 1 (inhibitor), depending on whether they are likely to inhibit the enzyme or not. Similarly, to this, a score of one or zero represents the likelihood of being an enzyme substrate. The molecule is classified as a category 1 substrate while a category 0 non-substrate of the enzyme is indicated by the molecule (38). Due to the flavan-4-ol's assigned score of 0.7971, it is most likely not an inhibitor of CYP1A2. The likelihood of CYP2C19 inhibition is 0.82 although the likelihood of being a CYP2C19 substrate is extremely low. Additionally, there is no evidence of CYP2D6 substrate or inhibitor. The flavan-4-ol is considered to be neither CYP3A4 substrates nor inhibitors of CYP3A4 (Tables 8B,C).

TABLE 6 Prediction LD_{50} value, prediction toxicity class, and pkcsM toxicity of top compounds.

Sr. No.	Compound	Prediction LD_{50} (mg/kg)	Prediction toxicity class	Ames toxicity	Max. tolerate dose human (log mg/kg/day)	Acute oral rat Toxicity (mol/kg)	Chronic oral rat toxicity (Log mg/kg_BW/day)	Hepatotoxicity
1	Flavan-4-ol	2,500	5	No	0.194	2.206	1.761	No
2	Apigenin	2,500	5	No	0.328	2.45	2.298	No
3	Daidzein	2,430	5	No	0.187	2.164	1.187	No
4	Naringenin	2000	4	No	−0.176	1.791	1.944	No
5	Violaxanthin	55	3	No	−0.384	2.132	2.054	No
6	Zeaxanthin	10	2	No	−1.058	3.496	2.603	No
7	Luteolin	3,919	5	No	0.499	2.471	2.612	No
8	Phthalic acid	2,530	5	No	0.582	1.449	2.165	No
9	Standard Atorvastatin	5,000	5	No	0.193	2.877	4.839	No
10	Standard Empagliflozin	3,000	5	No	0.25	2.554	3.51	No
11	Standard Ramipril	10,000	6	No	0.163	2.108	2.046	No

TABLE 7 ADME (Absorption, Distribution, Metabolism, and Excretion) Prediction, and physiochemical properties of compounds.

Sr. No.	Compound	HA	HBD	HBA	MR	iLOGP	TPSA (Å)	Log S	Lipinski rule
1	Flavan-4-ol	17	1	2	66.24	2.35	29.46	−3.40	Yes
2	Apigenin	20	5	3	73.99	1.89	90.90	−3.94	Yes
3	Daidzein	19	2	4	71.97	1.77	70.67	−3.53	Yes
4	Naringenin	20	3	5	71.57	1.75	86.99	−3.49	Yes
5	Violaxanthin	4	2	4	185.80	7.22	65.52	−9.05	No
6	Zeaxanthin	42	2	2	186.76	7.23	40.46	−9.58	No
7	Luteolin	21	4	6	76.01	1.86	111.13	−3.71	Yes
8	Phthalic acid	12	2	4	40.36	0.60	74.60	−1.57	Yes
9	St. Atorvastatin	41	4	6	158.26	3.81	111.79	−5.99	Yes
10	St. Empagliflozin	31	4	7	113.41	3.40	108.61	−3.80	Yes
11	St. Ramipril	30	2	6	116.96	3.17	95.94	−2.75	Yes

**HA, No. of a heavy atom; HBD, No. of hydrogen bond donor; HBA, No. of hydrogen bond acceptor; MR, Molar refractiveness; TPSA, Topological polar surface area; iLOGP, lipophilicity; Log S, Water solubility; Lipinski rule, Drug likeness factor; St., standard.

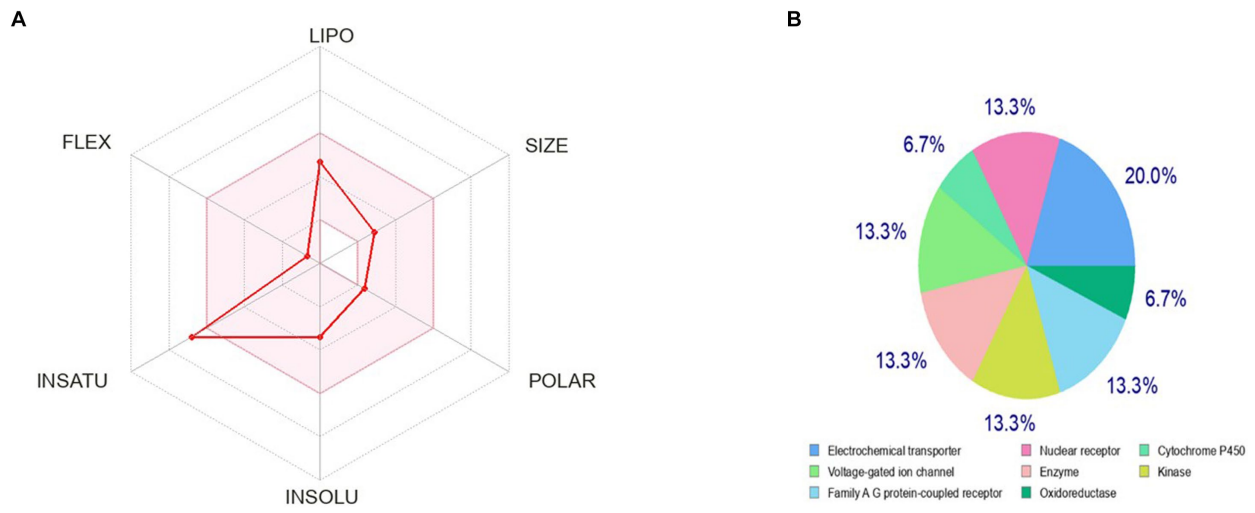


FIGURE 7
Pharmacokinetic properties and *in-silico* ADME modeling of Flavan-4-ol. **(A)** Oral bioavailability graph from the SwissADME database; the pink zone is the physicochemical space for oral bioavailability, and the red line defines oral bioavailability properties. **(B)** Target prediction.

3.4.1.4. Excretion

Flavan-4-ol has a low clearance rate of 0.176 mL/min/kg, which is calculated as the clearance (CL) (Table 8B). A drug's score will be >15 mL/min/kg if its clearance rate is high, 5–15 mL/min/kg if it is moderate, and 5 mL/min/kg if it has a low clearance rate (38). Renal organic cation transporters (ROCTs) are facilitated diffusion transporters that facilitate the vectorial transport of numerous physiological chemicals and xenobiotics or drugs in the kidney, liver, and placenta cells of mammals, assisting in their absorption and elimination. There was no inhibition of Renal Oct 2 by flavan-4-ol and it did not act as a substrate for Renal OCT2 (Table 8C).

3.4.1.5. Toxicity

The toxicity parameters include skin sensitization, hERG inhibition, human hepatotoxicity, AMES toxicity, carcinogenicity, rat oral acute toxicity, and hepatotoxicity. Flavan-4-ol was used in the ADMET test as a non-AMES substance and was not considered carcinogenic. There is a chance that flavan-4-ol will not cause skin sensitivity. The chemical is safe from negative reactions and has a decreased risk of carcinogenicity (0.09). A potassium ion channel that participates in the heart's normal repolarization activity is encoded by the human ether-à-go-go-related gene (hERG) (Table 8C). Long-term QT syndrome, which can result in arrhythmia and ultimately result in mortality, can be brought on by a drug's induction of hERG function blockage (39). With a predicted probability value of 0.89 for hERG inhibition (predictor I) and 0.88 for hERG inhibition (predictor II) for flavan-4-ol, they served as weak inhibitors and non-inhibitors of flavan-4-ol, respectively. The chance of the flavan-4-ol not easily degrading is 0.7280. All the pharmacokinetic parameters (absorption, distribution, metabolism, and excretion) are summarized in Table 8.

TABLE 8 Physicochemical and Pharmacokinetics Analysis of Flavan-4-ol.

(A) Physicochemical Properties of Flavan-4-ol (Swiss ADME)	
MW	226.27 g/mol
Heavy atoms	17
Aromatic heavy atoms	12
Fraction Csp3	0.2
Rotatable bonds	1
H-bond acceptors	2
H-bond donors	1
MR	66.24
TPSA	29.46 $Å^2$
iLOGP	2.35
XLOGP3	2.7
WLOGP	2.6
MLOGP	2.54
Silicos-IT Log P	2.98
Consensus Log P	2.63
ESOL Log S	−3.4
GI absorption	High
Lipinski violations	0
Ghose violations	0
Veber violations	0
Egan violations	0
Muegge violations	0.55
Bioavailability Score	0
PAINS alerts	0
Synthetic Accessibility	3.07

(B) ADMET Features of Flavan-4-ol (pkCSM open-source tool)	
Absorption	
Intestinal absorption (human)	94.242
P-glycoprotein I inhibitor	No
P-glycoprotein II inhibitor	No
Distribution	
VDss (human)	0.478 (log L/kg)
Fraction unbound (human)	0.079 (fu)
BBB permeability	0.559 (log BB)
CNS permeability	−1.676 (log PS)
Metabolism	
CYP2D6 substrate	No
CYP1A2 inhibitor	Yes
CYP2C19 inhibitor	Yes
CYP2C9 inhibitor	No
CYP2D6 inhibitor	No
CYP3A4 inhibitor	No

(Continued)

TABLE 8 (Continued)

(B) ADMET Features of Flavan-4-ol (pkCSM open-source tool)	
Excretion	
Total Clearance	0.176 (log ml/min/kg)
Renal OCT2 substrate	No
Toxicity	
AMES toxicity	No
Max. tolerated dose (human)	0.194 (log mg/kg/day)
hERG I inhibitor	No
hERG I inhibitor	No
Oral Rat Acute Toxicity (LD50)	2.206 (mol/kg)
Oral Rat Chronic Toxicity (LOAEL)	1.761 (log mg/kg_bw/day)
Hepatotoxicity	No
Skin Sensitisation	No

(C) ADMET features of Flavan-4-ol (admetSAR open-source tool)		Probability
Absorption		
Blood–Brain Barrier	BBB+	0.9424
Human Intestinal Absorption	HIA+	1.0000
Caco-2 Permeability	Caco2+	0.7438
P-glycoprotein Substrate	Non-substrate	0.6080
P-glycoprotein Inhibitor	Non-inhibitor	0.7945
	Non-inhibitor	0.8804
Aqueous solubility		−3.2626 LogS
Caco-2 Permeability		1.3393 Log Papp, cm/s
Distribution		
Subcellular localization	Mitochondria	0.6187
Metabolism		
CYP450 2C9 Substrate	Non-substrate	0.7794
CYP450 2D6 Substrate	Non-substrate	0.8637
CYP450 3A4 Substrate	Non-substrate	0.6939
CYP450 1A2 Inhibitor	Inhibitor	0.7971
CYP450 2C9 Inhibitor	Non-inhibitor	0.6064
CYP450 2D6 Inhibitor	Non-inhibitor	0.8971
CYP450 2C19 Inhibitor	Inhibitor	0.8289
CYP450 3A4 Inhibitor	Non-inhibitor	0.9369
CYP Inhibitory Promiscuity	Low CYP Inhibitory Promiscuity	0.6713
Excretion		
Renal Organic Cation Transporter	Non-inhibitor	0.8973
Toxicity		
Human Ether-a-go-go-Related Gene Inhibition	Weak inhibitor	0.8975
	Non-inhibitor	0.8841

(Continued)

TABLE 8 (Continued)

(C) ADMET features of Flavan-4-ol (admetSAR open-source tool)		Probability
AMES Toxicity	Non-AMES toxic	0.6569
Carcinogens	Non-carcinogens	0.9060
Biodegradation	Not readily biodegradable	0.7280
Acute Oral Toxicity	III	0.5338
Carcinogenicity (Three-class)	Non-required	0.4707
Rat Acute Toxicity	2.6087	LD50, mol/kg

3.5. Target prediction

The observed phenotypic effects are the result of the activity of bioactive small molecules, such as metabolites, being modulated by their binding to proteins or other macro-molecular targets. It is crucial to map their targets to understand the molecular processes behind the bioactivity of bioactive small compounds and foretell any adverse effects or cross-reactivity. We can computationally find new targets for uncharacterized compounds or secondary targets for recognized molecules. Swiss Target Prediction is a web service that uses a collection of 2D and 3D similarity measurements with known ligands to precisely predict the targets of bioactive compounds. Five distinct organisms can be used to perform predictions, and mapping assumptions by homology within and across species is possible for near paralogs and orthologs (4). The flavan-4-ol outcome of the closely related receptors was calculated using the UniProt ID, ChEMBL-ID, target class, likelihood, and known actives in 2D/3D. The results were 20% electrochemical transporter, 13.3% nuclear receptor, family A-G protein receptor, kinase, enzymes, voltage-gated ion channels, and 6.7% cytochrome P450. Flavan-4-ol forecasts these other proteins as a target as well, as shown in Figure 7.

4. Discussion and conclusion

Millets have shown positive health impacts, including antioxidant activity, anti-diabetic, anti-tumorigenic, anti-atherogenic, and antibacterial properties (27). Regular eating of whole grain millets and their products can reduce the risk of type II diabetes, gastrointestinal malignancies, cardiovascular disease, and a variety of other ailments (41). Most millets have a carbohydrate content between 60 and 70%, with the majority being non-starchy polysaccharides, contributing to millets' many health advantages (42). According to various epidemiological studies, eating millet enhances the immune system, detoxifies the body, lowers the risk of cancer, boosts energy, improves brain and muscular systems, and raises immunity in the respiratory system (3, 43). Typically, millets are eaten with the seed coat, which is high in phenolics, dietary fiber, minerals, and vitamins and is more beneficial to human health than other whole-grain cereals (26, 43).

The computational approaches investigated novel compounds from millet sources in terms of their interactions with DPPIV, SGLT-2, hACE, and PCSK9. A structure-based virtual screening method called molecular docking discovers active inhibitors based on predictions of the binding affinities and molecular interactions between ligand molecules (or inhibitors) and their corresponding target proteins or enzymes. The binding affinity of active inhibitors is typically evaluated using a flexible docking simulation methodology. For the protein–ligand complex created with low energy conformation, the most advantageous binding mechanism docking poses is examined (37). By docking, the optimal binding orientation of ligands for their corresponding target molecules is discovered.

DPPIV is an integral membrane aminopeptidase and member of the prolyl oligopeptidase that was initially identified as a T-cell differentiation antigen (CD26) and was reported on the diverse groups of epithelial cells, viz., kidney, liver intestine, prostate, lung, and placenta. DPPIV is a major glycemic mediator used to control type 2 diabetes mellitus which is associated with severe life-threatening coronary diseases such as stroke, heart failure, and many more cardiovascular adverse effects (8, 15). Inhibition of DPPIV by bioactive compounds provides proof as a tool for the treatment of type 2 diabetes mellitus (8). DPPIV inhibition reduces inflammation and immune system activation, which are frequent characteristics of diabetes and hypertension, indicating that these processes may play a significant part in DPPIV-mediated kidney damage. Sitagliptin and vildagliptin, two DPPIV inhibitors that are now available on the market, exhibit notable hypoglycemia effects. However, they can also cause rashes, upper respiratory tract infections, and hypersensitivity responses (44–46). Therefore, a trustworthy and tried method for finding new hypoglycemia medications is the discovery of DPP-IV inhibitors with novel structures, particularly among the secondary metabolites of plants. The results of our docking analysis indicate that among 50 bioactive compounds docked against DPPIV, apigenin, zeaxanthin, flavan-4-ol, and violaxanthin had better binding interactions when compared to the standard empagliflozin. Furthermore, the number of hydrogen bonds formed between the target and inhibitor was found to be a significant predictor of stable complex formation. For instance, apigenin formed four hydrogen bonds with the target protein while the other compounds formed one hydrogen bond. Our findings suggest that apigenin, zeaxanthin, flavan-4-ol, and violaxanthin hold promise as potential candidates for further study in drug discovery against DPPIV.

Type-2 diabetes is a metabolic disorder characterized by high levels of glucose in the blood. Renal glucose reabsorption is an important factor in maintaining elevated blood glucose levels. Sodium-glucose cotransporters, particularly SGLT-2, play a significant role in glucose reabsorption from the kidneys (25). SGLT-2 inhibitors have emerged as a new class of antihyperglycemic agents that help manage type-2 diabetes by inhibiting the SGLT-2 pathway of glucose reabsorption in the kidneys, leading to increased urinary excretion of excess glucose and lowering blood sugar levels (47). This mode of action is insulin-independent, which means that these inhibitors can be used alone or in combination with other antidiabetic agents to improve glycemic control. Furthermore, SGLT-2 inhibitors have been found to exert nephroprotective effects in patients with chronic kidney disease. As type-2 diabetes is a major health concern worldwide, the role of SGLT-2 inhibitors in managing this chronic condition cannot be overstated (47). The results of docking bioactive compounds against the macromolecule SGLT-2 showed that flavan-4-ol, daidzein, luteolin, and naringenin had more negative-binding energies than the standard Empagliflozin. This indicates that the tested compounds showed greater stability in complex with SGLT-2. Flavan-4-ol

demonstrated the highest stability, forming two hydrogen bonds with seven different amino acids in SGLT-2. Daidzein, luteolin, and naringenin also showed stability and formed three hydrogen bonds with amino acids in SGLT-2. These findings, obtained via a structure-based drug-design method, are critical in the development of drugs that can effectively target SGLT-2. Moreover, SGLT-2 inhibitors have been found to be a promising new type of anti-diabetic drug. The use of SGLT-2 inhibitors has been proven to be effective in reducing blood glucose and weight without increasing the risk of hypoglycemia. In addition, a meta-analysis demonstrated that SGLT-2 use has significant cardiovascular and renal protective effects.

Human angiotensin-converting enzyme (hACE) control leads to the management of hypertension which poses a serious risk of developing coronary disease, heart failure, stroke, and a variety of other cardiovascular diseases (48). It plays an integral role in the control of blood pressure through the integration of the Angiotensin 2 pathway synthesis. A high concentration of Angiotensin 2 affects the renal tubule to retain sodium and water which further results in hypertension (48). The maintenance of cardiovascular homeostasis depends on the renin-angiotensin system. ACE inhibition or angiotensin II receptor blockade is the mainstay of therapy for several cardiovascular disorders. Angiotensin-(1-7) levels in plasma and tissues may rise as a result of hACE inhibition, which prevents the conversion of angiotensin-(1-7) to angiotensin (49). Current clinical uses for hACE inhibitors include the management of hypertension, endothelial dysfunction, congestive heart failure, myocardial infarction, and renal illness (including diabetic nephropathy) (23). Our docking analysis identified five bioactive compounds as potent hACE inhibitors, with compounds violaxanthin, zeaxanthin, flavan-4-ol, and daidzein being the most effective. These compounds formed hydrogen bonds with key active site residues and exhibited higher binding energies and Ki values than Ramipril. Further *in vitro* and *in vivo* studies are warranted to confirm the efficacy of these compounds as potential therapeutic agents for hypertension and related cardiovascular diseases.

A serine protease called PCSK9 plays an integral role in the regulation of the cholesterol level of the body. It binds to hepatic-specific LDL (low-density lipoprotein) receptors and increases the intracellular degradation of the intricate LDL receptor, hence decreasing blood LDL clearance. Despite being synthesized to a lesser amount in other organs, PCSK9 is primarily released by the liver. In addition to its well-known role in the hepatic LDL receptor-mediated pathway, PCSK9 has also been linked to the claim that it may prevent vascular inflammation during atherogenesis (23). When LDL receptors are blocked, there is a rise in LDL concentration, which increases the risk of developing cardiovascular disease and stroke. The gain-of-function mutation of PCSK9 results in autosomal-dominated familial hypercholesteremia. Inhibition of PCSK9 is considerable promise for the management of hypercholesterolemia and its associated cardiovascular disease. The docking study identified phthalic acid, naringenin, daidzein, and flavan-4-ol as PCSK9 inhibitors with varying binding energies, the number of hydrogen bonds formed, and inhibition constants. The findings provide valuable insights for the design of more potent PCSK9 inhibitors for the treatment of hypercholesterolemia.

All the top four substances exhibit higher binding affinities than the standard. The binding energy was found through an *in-silico* analysis to indicate a stronger and more stable connection between the ligand and the target molecule. Bioactive substances including flavan-4-ol, violaxanthin, zeaxanthin, apigenin, daidzein, luteolin, and naringenin have stronger binding energies with the targeted molecules than the more often used atorvastatin, empagliflozin, and ramipril. The stronger the complex, the higher the negative binding energy value. Only the flavan-4-ol is the most prevalent molecule that docked against all four targets, and it formed a stable complex with all four targets.

Pharmacokinetics analysis supported the MD data. Also, the 100 ns MDs verified the examined compounds' affinity by demonstrating improved stability in the receptor-binding region. MM/PBSA binding Free Energy Analysis depicted low negative free binding energies indicating that the test ligands had a strong affinity for binding to the target protein (34). Among the four complexes, the PCSK9-flavan-4-ol had low binding energy −1,903.97 ΔG bind (kJ mol^{-1}) as shown in Table 5. They were then put through an MD simulation trajectory. The results of the RMSD analysis showed that the DPPIV- flavan-4-ol complex had minimum, and RMSF complex 4: HACE2-flavan-4-ol showed huge fluctuations; the analysis proceeds further for RGY and SASA during the whole 100 ns MD trajectory. Overall, the four complexes show fluctuations for more stability. The root means the square distance between a protein's atoms and its rotational axis is measured by RGY. It is one of the crucial variables that describe the overall change in the compactness and dimensions of the protein structure during the simulation. Low RGY values imply a protein that is extremely compact and inflexible, whereas elevated values denote a protein that is less compact and flexible. Protein's backbone RGY values were plotted over time to observe how the compactness of the structure changed over time. Throughout the simulation, the RGY value of the protein and protein–ligand complexes gradually decreased, indicating that the test compounds did not significantly alter the protein's structural composition (Figure 6).

The LD_{50} can measure acute toxicity, and the six classes of toxicity classifications are outlined by the Globally Harmonized System of Classification and Labelling of Chemicals (GHS) (50). In this study, flavan-4-ol, apigenin, daidzein, luteolin, and phthalic acid were classified as class V, meaning they may be dangerous if ingested in amounts of 2000 to 5,000 mg/kg, and naringenin was classified as class IV, meaning it would be harmful if ingested in amounts of 300 to 2000 mg/kg.

Bioactive compounds, such as violaxanthin and zeaxanthin, have been found to be neither hepatotoxic nor carcinogenic. They are more permeable than other compounds due to their solubility, stability, and metabolism by gut microbes. The TPSA value of luteolin, naringenin, daidzein, and apigenin has been used to measure their capacity to be orally active in the human system. Lipinski's rule of five has been used to predict the drug-likeness properties of bioactive compounds, except for zeaxanthin.

Bioactive compounds are important for discovering new drugs, but animal models are not reliable predictors of human toxicity. We found that the compound flavan-4-ol is best docked to all four targets of lifestyle diseases, and MD simulation analysis further strengthens our finding that the flavan-4-ol forms a better stability complex with all the targets. ADMET profiles substantiate the candidature of the flavan-4-ol bioactive compound to be considered for trial as an inhibitor of targets DPPIV, SGLT2, PCSK9, and hACE. We suggest that more research is conducted, taking

Flavon-4-ol into account when producing new medicines from millets. Multi-target therapeutic candidates can be created from it to suppress the biochemical pathway of diseases diabetes, hypertension, and atherosclerosis.

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

KN and NS: conception, design, methodology, statistical analysis, and writing manuscript. GT and MI: simulation analysis and manuscript writing. TN: visualization and review. RB, AM, and CG: writing—review and editing. All authors contributed to the article and approved the submitted version.

Conflict of interest

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

Publisher's note

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

References

1. Khalid M, Alqarni MH, Alsayari A, Foudah AI, Aljarba TM, Mukim M, et al. Anti-diabetic activity of bioactive compound extracted from spondias mangifera fruit: in-vitro and molecular docking approaches. *Plan Theory*. (2022) 11:562. doi: 10.3390/plants11040562

2. Shobana S, Malleshi NG. Preparation and functional properties of decorticated finger millet (Eleusine coracana). *J Food Eng*. (2007) 79:529–38. doi: 10.1016/j.jfoodeng.2006.01.076

3. Chandrasekara A, Shahidi F. Content of insoluble bound phenolics in millets and their contribution to antioxidant capacity. *J Agric Food Chem*. (2010) 58:6706–14. doi: 10.1021/jf100868b

4. Friedman M. Chemistry, biochemistry, and dietary role of potato polyphenols – a review. *J Agric Food Chem*. (1997) 45:1523–40. doi: 10.1021/jf960900s

5. Jones JM, Engleson J. Whole grains: benefits and challenges. *Annu Rev Food Sci Technol*. (2010) 1:19–40. doi: 10.1146/annurev.food.112408.132746

6. Kawser Hossain M, Abdal Dayem A, Han J, Yin Y, Kim K, KµMar Saha S, et al. Molecular mechanisms of the anti-obesity and anti-diabetic properties of flavonoids. *Int J Mol Sci*. (2016) 17:569. doi: 10.3390/ijms17040569

7. Konappa N, Udayashankar AC, Krishnamurthy S, Pradeep CK, Chowdappa S, Jogaiah S. GC–MS analysis of phytoconstituents from AmomµM nilgiricµM and molecular docking interactions of bioactive serverogenin acetate with target proteins. *Sci Rep*. (2020) 10:1–23. doi: 10.1038/s41598-020-73442-0

8. Burton-Freeman B, Brzeziński M, Park E, Sandhu A, Xiao D, Edirisinghe I. A selective role of dietary anthocyanins and flavan-3-ols in reducing the risk of type 2 diabetes mellitus: a review of recent evidence. *Nutrients*. (2019) 11:841. doi: 10.3390/nu11040841

9. Dias MC, Pinto DC, Silva AM. Plant flavonoids: chemical characteristics and biological activity. *Molecules*. (2021) 26:5377. doi: 10.3390/molecules26175377

10. Available at: http://www.aicpmip.res.in/aboutus.html

11. Gull A, Jan R, Nayik GA, Prasad K, KµMar P. A review of the significance of finger millet in nutrition, health, and value-added products. *MagnesiµM (mg)*. (2014) 130:120.

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

13. Jakubczyk A, Ćwiek P, Rybczyńska-Tkaczyk K, Gawlik-Dziki U, Złotek U. The influence of millet flour on antioxidant, anti-ACE, and anti-microbial activities of wheat wafers. *Foods*. (2020) 9:220. doi: 10.3390/foods9020220

14. Liu F, Shan S, Li H, Shi J, Hao R, Yang R, et al. Millet shell polyphenols prevent atherosclerosis by protecting the gut barrier and remodeling the gut microbiota in ApoE−/− mice. *Food Funct*. (2021) 12:7298–309. doi: 10.1039/d1fo00991e

15. Nishizawa N, Togawa T, Park KO, Sato D, Miyakoshi Y, Inagaki K, et al. Dietary Japanese millet protein ameliorates plasma levels of adiponectin, glucose, and lipids in type 2 diabetic mice. *Biosci Biotechnol Biochem*. (2009) 73:351–60. doi: 10.1271/bbb.80589

16. Tentolouris A, Vlachakis P, Tzeravini E, Eleftheriadou I, Tentolouris N. SGLT2 inhibitors: a review of their antidiabetic and cardioprotective effects. *Int J Environ Res Public Health*. (2019) 16:2965. doi: 10.3390/ijerph16162965

17. Available at: https://www.fda.gov/drugs/drug-safety-and-availability/fda-revises-labels-sglt2-inhibitors-diabetes-include-warnings-about-too-much-acid-blood-and-serious8

18. Available at: https://www.thermofisher.com/blog/connectedlab/the-role-of-adme-toxicology-studies-in-drug-discovery-development/

19. Lee JY, Jun DY, Yoon YH, Ko JY, Woo KS, Woo MH, et al. Anti-inflammatory effect of flavonoids kaempferol and biochanin a-enriched extract of barnyard millet (Echinochloa crus-galli var. frµMentacea) grains in LPS-stimulated RAW264. 7 cells. *J Life Sci*. (2014) 24:1157–67. doi: 10.5352/JLS.2014.24.11.1157

20. Li S, Dong X, Fan G, Yang Q, Shi J, Wei W, et al. Comprehensive profiling and inheritance patterns of metabolites in foxtail millet. *Front Plant Sci*. (2018) 9:1716. doi: 10.3389/fpls.2018.01716

21. Sharma S, Saxena DC, Riar CS. Analyzing the effect of germination on phenolics, dietary fibers, minerals, and γ-amino butyric acid contents of barnyard millet (Echinochloa frµMentaceae). *Food Biosci*. (2016) 13:60–8. doi: 10.1016/j.fbio.2015.12.007

22. Okwudili UH, Gyebi DK, Obiefuna JAI. Finger millet bioactive compounds, bioaccessibility, and potential health effects–a review. *Czech J Food Sci*. (2017) 35:7–17. doi: 10.17221/206/2016-CJFS

23. Menard J, Patchett AA. Angiotensin-converting enzyme inhibitors. *Adv Protein Chem*. (2001) 56:13–75. doi: 10.1016/S0065-3233(01)56002-7

24. Liberale L, Montecucco F, Camici GG, Dallegri F, Vecchie A, Carbone F, et al. Treatment with proprotein convertase subtilisin/kexin type 9 (PCSK9) inhibitors to reduce cardiovascular inflammation and outcomes. *Curr Med Chem*. (2017) 24:1403–16. doi: 10.2174/0929867324666170303123734

25. Mathanghi SK, Sudha K. Functional and phytochemical properties of finger millet (Eleusine coracana L.) for health. *Int J Pharm, Chem Biol Sci*. (2012) 2:431–8.

26. Begemann J, Ostovar S, Schwake-Anduschus C. Facing tropane alkaloid contamination in millet–analytical and processing aspects. *Qual Assur Safety Crops Foods*. (2021) 13:79–86. doi: 10.15586/qas.v13i2.888

27. Yang L, Allred KF, Geera B, Allred CD, Awika JM. SorghµM phenolics demonstrate estrogenic action and induce apoptosis in nonmalignant colonocytes. *Nutr Cancer*. (2012) 64:419–27. doi: 10.1080/01635581.2012.657333

28. Graw S, Chappell K, Washam CL, Gies A, Bird J, Robeson MS, et al. Multi-omics data integration considerations and study design for biological systems and disease. *Mol Omics*. (2021) 17:170–85. doi: 10.1039/D0MO00041H

29. Chaturvedi M, Nagre K, Yadav JP. In silico approach for identification of natural compounds as potential COVID-19 main protease (Mpro) inhibitors. *VirusDisease*. (2021) 32:325–9. doi: 10.1007/s13337-021-00701-7

30. Hess B, van der Spoel LE, Lindahl E. *GROMACS Groningen machine for chemical simulations, user manual version 4.5.4*. Groningen, The Netherlands: University of Groningen (2011).

31. Huang J, Rauscher S, Nawrocki G, Ran T, Feig M, De Groot BL, et al. CHARMM36m: an improved force field for folded and intrinsically disordered proteins. *Nat Methods*. (2017) 14:71–3. doi: 10.1038/nmeth.4067

32. Van Der Spoel D, Lindahl E, Hess B, Groenhof G, Mark AE, Berendsen HJ. GROMACS: fast, flexible, and free. *J Comput Chem*. (2005) 26:1701–18. doi: 10.1002/jcc.20291

33. Vanommeslaeghe K, Hatcher E, Acharya C, Kundu S, Zhong S, Shim J, et al. CHARMM general force field: a force field for drug-like molecules compatible with the

CHARMM all-atom additive biological force field. *J Comput Chem*. (2010) 31:671–90. doi: 10.1002/jcc.21367

34. Yu W, He X, Vanommeslaeghe K, MacKerell AD Jr. Extension of the CHARMM general force field to sulfonyl-containing compounds and its utility in biomolecular simulations. *J Comput Chem*. (2012) 33:2451–68. doi: 10.1002/jcc.23067

35. Mishra S, Dahima R. In vitro ADME studies of TUG-891, a GPR-120 inhibitor usingSWISS ADME predictor. *J Drug Delivery Therap*. (2019) 9:366–9.

36. Gfeller D, Grosdidier A, Wirth M, Daina A, Michielin O, Zoete V. SwissTargetPrediction: a web server for target prediction of bioactive small molecules. *Nucl Acids Res*. (2014) 42:W32–8. doi: 10.1093/nar/gku293

37. Rudrapal M, Gogoi N, Chetia D, Khan J, Banwas S, Alshehri B, et al. Repurposing of phytomedicine-derived bioactive compounds with promising anti-SARS-CoV-2 potential: molecular docking, MD simulation and drug-likeness/ADMET studies. *Saudi J Biol Sci*. (2022) 29:2432–46. doi: 10.1016/j.sjbs.2021.12.018

38. Ahmad I, Kuznetsov AE, Pirzada AS, Alsharif KF, Daglia M, Khan H. Computational pharmacology and computational chemistry of 4-hydroxyisoleucine: physicochemical, pharmacokinetic, and DFT-based approaches. *Front Chem*. (2023) 11:1145974. doi: 10.3389/fchem.2023.1145974

39. Quiñones M, Margalef M, Arola-Arnal A, Muguerza B, Miguel M, Aleixandre A. The blood pressure effect and related plasma levels of flavan-3-ols in spontaneously hypertensive rats. *Food Funct*. (2015) 6:3479–89. doi: 10.1039/C5FO00547G

40. McKeown NM. Whole grain intake is favourably associated with metabolic risk factors for type 2 diabetes and cardiovascular disease in the Framingham offspring study. *Am J Clin Nutr*. (2002) 76:390–8. doi: 10.1093/ajcn/76.2.390

41. Shivran AC (2016) Biofortification for nutrient-rich millets. *Biofortification of food crops*. Springer, New Delhi, 409–420

42. Manach C, Mazur A, Scalbert A. Polyphenols and prevention of cardiovascular diseases. *Curr Opin Lipidol*. (2005) 16:77–84. doi: 10.1097/00041433-200502000-00013

43. Antony U, Sripriya G, Chandra TS. Effect of fermentation on the primary nutrients in finger millet (Eleusine coracana). *J Agric Food Chem*. (1996) 44:2616–8. doi: 10.1021/jf950787q

44. Guasch L, Sala E, Ojeda MJ, Valls C, Bladé C, Mulero M, et al. Identification of novel hμMan dipeptidyl peptidase-IV inhibitors of natural origin (PartII): in silico prediction in antidiabetic extracts. *PLoS One*. (2012) 7:e44972. doi: 10.1371/journal.pone.0044972

45. Kasina SVSK, Baradhi KM. Dipeptidyl peptidase IV (DPP IV) inhibitors. [updated 2022 may 29]. *StatPearls [internet]*. Treasure Island (FL): StatPearls Publishing; (2022), Available at: https://www.ncbi.nlm.nih.gov/books/NBK542331/

46. Makrilakis K. The role of DPP-4 inhibitors in the treatment algorithm of type 2 diabetes mellitus: when to select, what to expect. *Int J Environ Res Public Health*. (2019) 16:2720. doi: 10.3390/ijerph16152720

47. Jabbour SA, Goldstein BJ. SodiμM-glucose co-transporter 2 inhibitors: blocking renal tubular reabsorption of glucose to improve glycaemic control in patients with diabetes. *Int J Clin Pract*. (2008) 62:1279–84. doi: 10.1111/j.1742-1241.2008.01829.x

48. Herman LL, Padala SA, Ahmed I, et al. Angiotensin converting enzyme inhibitors (ACEI) [updated 2022 Aug 5]. *StatPearls*. Treasure Island (FL): StatPearls Publishing; (2022), Available at: https://www.ncbi.nlm.nih.gov/books/NBK431051/

49. Roks AJ, van Geel PP, Pinto YM, Buikema H, Henning RH, de Zeeuw D, et al. Angiotensin-(1–7) is a modulator of the hμMan renin-angiotensin system. *Hypertension*. (1999) 34:296–301. doi: 10.1161/01.HYP.34.2.296

50. Ramya S, Soorya C, Pushpalatha GGL, Aruna D, Loganathan T, Balamurugan S, et al. Artificial intelligence and machine learning approach based in-silicoADME-Tox and pharmacokinetic profile of α-linolenic acid from Catharanthus roseus (L.) G.Don. *J Drug Delivery Therap*. (2022) 12:96–109. doi: 10.22270/jddt.v12i2-S.5274

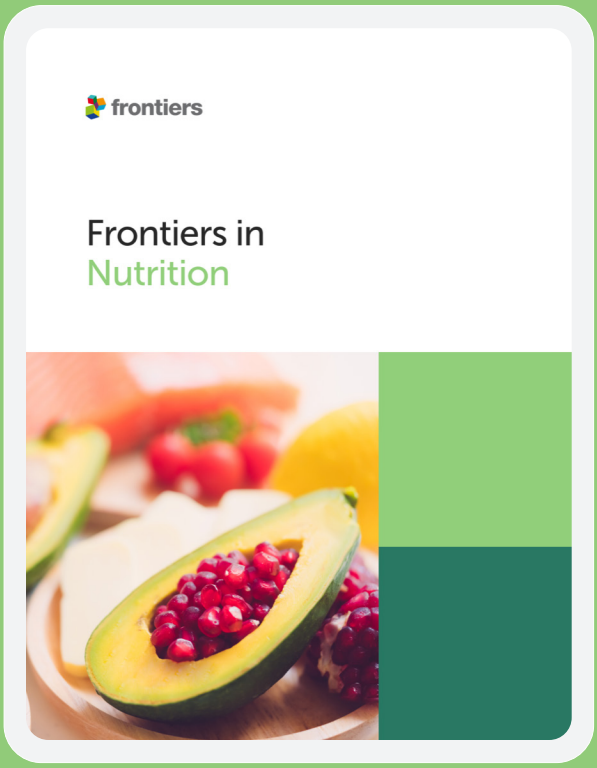
Frontiers in Nutrition

Explores what and how we eat in the context of health, sustainability and 21st century food science

A multidisciplinary journal that integrates research on dietary behavior, agronomy and 21st century food science with a focus on 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