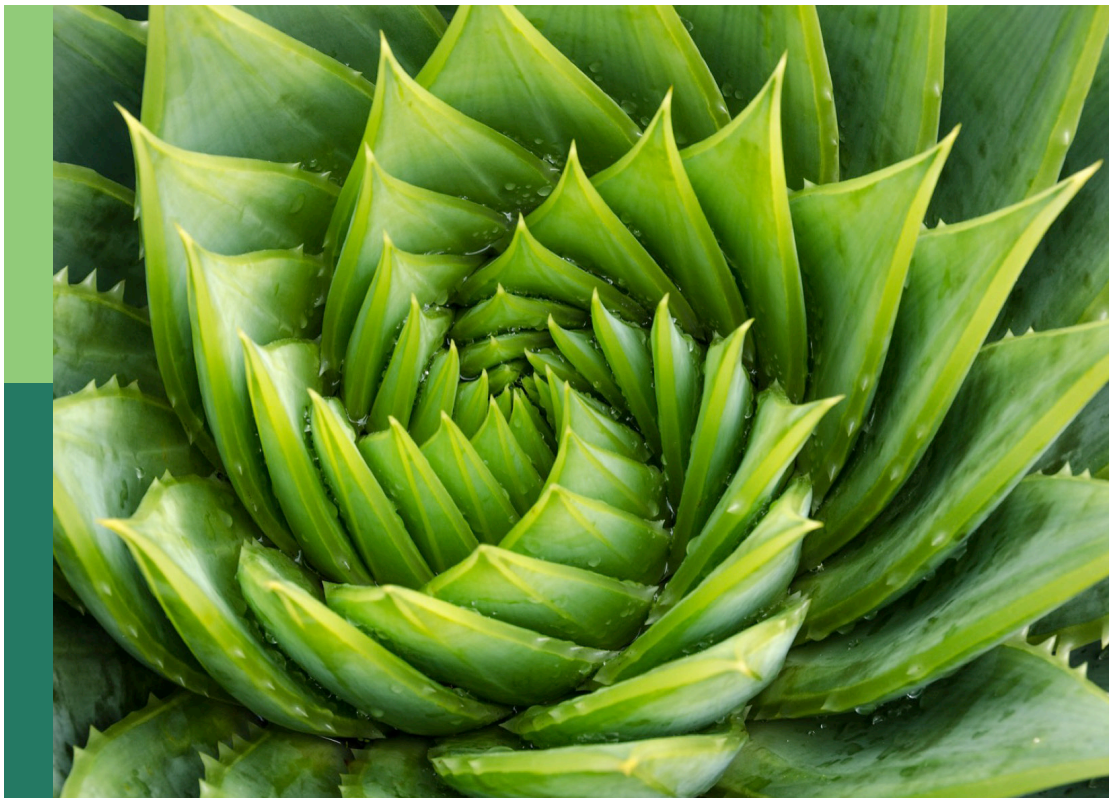
frontiers | Research Topics

Integrating advanced high-throughput technologies to improve plant resilience to environmental challenges

Edited by
Freddy Mora-Poblete, Sigfredo Fuentes and Parviz Heidari

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-2700-9
DOI 10.3389/978-2-8325-2700-9

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

Integrating advanced high-throughput technologies to improve plant resilience to environmental challenges

Topic editors

Freddy Mora-Poblete — University of Talca, Chile

Sigfredo Fuentes — The University of Melbourne, Australia

Parviz Heidari — Shahrood University of Technology, Iran

Citation

Mora-Poblete, F., Fuentes, S., Heidari, P., eds. (2023). *Integrating advanced high-throughput technologies to improve plant resilience to environmental challenges*. Lausanne: Frontiers Media SA. doi: 10.3389/978-2-8325-2700-9

Table of contents

frontiers | Frontiers in Plant Science

TYPE Editorial
PUBLISHED 31 May 2023
DOI 10.3389/fpls.2023.1218691

OPEN ACCESS

EDITED AND REVIEWED BY
Nunzio D'Agostino,
University of Naples Federico II, Italy

*CORRESPONDENCE
Freddy Mora-Poblete
morapoblete@gmail.com

RECEIVED 07 May 2023
ACCEPTED 23 May 2023
PUBLISHED 31 May 2023

CITATION
Mora-Poblete F, Heidari P and Fuentes S (2023) Editorial: Integrating advanced high-throughput technologies to improve plant resilience to environmental challenges. *Front. Plant Sci.* 14:1218691. doi: 10.3389/fpls.2023.1218691

Editorial: Integrating advanced high-throughput technologies to improve plant resilience to environmental challenges

Freddy Mora-Poblete[1]*, Parviz Heidari[2] and Sigfredo Fuentes[3,4]

[1]Laboratory of Forest Genetics and Biotechnology, Institute of Biological Sciences, University of Talca, Talca, Chile, [2]Faculty of Agriculture, Shahrood University of Technology, Shahrood, Iran, [3]Digital Agriculture, Food and Wine Sciences Group, School of Agriculture, Food and Ecosystem Sciences, Faculty of Science, The University of Melbourne, Melbourne, VIC, Australia, [4]Tecnologico de Monterrey, Escuela de Ingeniería y Ciencias, Monterrey, Mexico

KEYWORDS

abiotic stress resilience, high-throughput technologies, data integration, multilevel biological data, bioinformatics resources, multi-omics

Editorial on the Research Topic
Integrating advanced high-throughput technologies to improve plant resilience to environmental challenges

High-throughput technologies enable extensive omics datasets for genomics, transcriptomics, proteomics, phenomics, and metabolomics analysis. These advancements, accompanied by evolving bioinformatics tools, integrate omics-related data, providing critical information about plant molecular systems and their functions (Choi, 2019). These technologies significantly advance omics research in plants, investigating gene function, regulation, and adaptation. Furthermore, they contribute to the recovery of substantial plant diversity, vital for genetic improvement, food security, and conservation efforts (Kumar et al., 2021). By integrating multi-level biological data from genomics, transcriptomics, proteomics, and metabolomics, comprehensive investigations and insights into molecular aspects governing responses to abiotic stresses can be achieved.

This Research Topic integrates advanced high-throughput technologies, multi-omics, bioinformatics, systems biology, and artificial intelligence to explore plant stress and tolerance to environmental constraints. It includes nine original research articles, enhancing plant resilience to stressors like drought, cold, ultraviolet radiation, flooding, and low-nitrogen stress. The articles encompass important plant species: rice, potato, cabbage, sugarcane, poplar, Antarctic moss (*Pohlia nutans* and *Leptobryum pyriforme*), and an endangered plant species, *Myricaria laxiflora*. Additionally, a review examines recent progress in genome engineering and the role of CRISPR-Cas9-mediated genome editing in sustainable agriculture.

Various cutting-edge technologies are explored in this Research Topic to enhance plant resilience against environmental challenges. These include transcriptomics, proteomics, metabolomics, and phenomics. Dwivedi et al. conducted the first study to employ high-throughput phenomics parameters for selecting reproductive stage drought stress (RSDS)

tolerance in near-isogenic lines (NILs) of rice. Four traits were used: projected shoot area, water use, transpiration rate, and red-green-blue and near-infrared values. Results showed the potential role of quantitative trait loci (QTLs) in improving water use efficiency (WUE) and identified eleven NILs based on phenomics traits and performance under imposed drought in the field. In conclusion, high-throughput phenomics-based phenotyping enables efficient and rapid assessment of plant traits, facilitating the selection of stress-tolerant genotypes in challenging environments.

The integration of transcriptomics and proteomics is of paramount importance in advancing our understanding of complex biological systems. In this regard, Li et al. investigated the regulatory mechanisms of flooding stress by integrating transcriptomic and proteomic analyses in roots of *Myricaria laxiflora* (an endangered plant) during nine-time points under flooding and post-flooding recovery treatments. Genes related to auxin, cell wall, calcium signaling, and MAP kinase signaling were greatly down-regulated exclusively at the transcriptomic level during the early stages of flooding. Six groups of differentially expressed proteins were exclusively identified at the proteome level, showing specific expression patterns. These proteins were associated with glycolysis, major carbohydrate metabolism, redox reactions, and G-protein signaling. They played important roles in energy generation and redox homeostasis during flooding stress. Post-flooding recovery showed activation of genes related to ROS scavenging, mitochondrial metabolism, and development. In addition, the genes/proteins related to redox, hormones, and transcriptional factors also played vital roles in flooding stress of *M. laxiflora*. These findings enhance our understanding of flooding stress mechanisms and aid in preserving endangered plants in flood-prone areas. Chen et al. performed an integrated metabolomics and transcriptomics analysis to examine changes in metabolites and regulatory pathways in *Populus tomentosa* under low-N stress. Nitrogen is essential for plant growth and development, but little is known about how trees regulate their metabolism under N deficiency. RNA sequencing revealed differentially expressed genes related to carbohydrate metabolism and hormone pathways. Integrated metabolomics and transcriptomics analyses revealed a co-expression pattern of metabolites and genes. These findings provide insights into the metabolic and molecular mechanisms underlying N and carbon interactions in poplar.

Ponce et al. and Sajad et al. conducted transcriptomic studies to explore the genetic and molecular mechanisms underlying plant responses to different stress conditions such as drought and cold. Ponce et al. used RNA-sequencing study to investigate the drought response of two Andigenum varieties of potato with different drought tolerance. Comparative transcriptome analysis revealed significant early differences in the responses of the tolerant variety, suggesting the importance of rapid response. The tolerant variety exhibited more effective ABA synthesis and mobilization, feedback regulation, and differential expression of genes involved in cell wall reinforcement and remodeling. These findings provide insights into molecular bases of drought tolerance mechanisms and pave the way for improving potato resistance to drought stress. Sajad et al. conducted a genome-wide study on cabbage (*Brassica oleracea* var. capitata L.) and highlighted the crucial roles of the Hsp90 gene family in growth and development under cold conditions. This study discovered 12 BoHsp90 genes classified into five groups. Promoter evaluation revealed stress-related and hormone-responsive cis-elements in the genes. RNA-seq data analysis showed tissue-specific expression of BoHsp90-9 and BoHsp90-5, with six genes induced by cold stress, indicating their important role in cold acclimation and overall growth and development.

Integrative transcriptomics and metabolomics data analysis involves combining and analyzing data from transcriptomic studies (gene expression) and metabolomic studies (metabolite profiling) to gain a comprehensive understanding of biological processes. In this context, two studies investigated the molecular mechanisms underlying the adaptation of Antarctic moss species, *Pohlia nutans* and *Leptobryum pyriforme*, to UV-B radiation. Liu et al. studied the molecular mechanism of this adaptation and used transcriptomics and metabolomics to profile *Leptobryum pyriforme* moss under UV-B radiation. The results showed differentially expressed genes involved in UV-B signaling, flavonoid biosynthesis, ROS scavenging, and DNA repair. Flavonoids were the most changed metabolites, and the UVR8-mediated signaling, jasmonate signaling, flavonoid biosynthesis pathway, and DNA repair system contributed to adaptation. On the other hand, Liu et al. assembled the high-quality genome sequence of the Antarctic moss (*Pohlia nutans*) with 698.20 Mb and 22 chromosomes, revealing genomic features that aid in its adaptation to the extreme environment. The large size of the genome is due to a high proportion of repeat sequences and a recent whole-genome duplication event. The genome also shows evidence of massive gene duplications and expansions of gene families that facilitate neofunctionalization. The moss exhibits expanded gene families involved in phenylpropanoid biosynthesis, unsaturated fatty acid biosynthesis, and plant hormone signal transduction, likely supporting its Antarctic lifestyle. Additionally, the moss has developed adaptive strategies for UV-B radiation through the expansion and upregulation of genes encoding DNA photolyase, antioxidant enzymes, and flavonoid biosynthesis enzymes. The findings of these investigations provide insights into the adaptation of Antarctic moss to polar environments and assessing global climate change's impact on Antarctic land plants.

Malviya et al. and Xu et al. utilized advanced high-throughput technologies with the aim of advancing research on the breeding improvement of sugarcane. In the study by Malviya et al., advanced high-throughput sequencing techniques were likely employed to explore the diversity of root-associated microbiomes in sugarcane species. Results revealed a significant rhizospheric diversity across progenitors and close relatives. The rhizosphere microbial abundance of modern sugarcane progenitors was at the lower end of the spectrum, indicating the potential of introgression breeding to improve nutrient use and disease and stress tolerance of commercial sugarcane. On the other hand, Xu et al. examined the

effects of Indole-3-butyric acid (IBA) on sugarcane seedling growth and gene expression during root development. Results showed significant growth promotion and accumulation of plant hormones at 100 ppm IBA. Transcriptomic analysis revealed significant differentially expressed genes, mainly involved in metabolic, cellular, and single-organism processes. The study also analyzed the expression of genes related to plant hormones and signaling pathways. The findings provide new insights into the IBA response to sugarcane bud sprouting, which could aid in the breeding improvement of sugarcane.

Finally, one review article was included in this Research Topic. Ali et al. reviewed the recent progress in genome engineering and the role of CRISPR-Cas9-mediated genome editing in sustainable agriculture. Genome editing technology, particularly CRISPR-Cas9-based systems, offers significant potential for precise trait targeting, including enhancing resistance to microorganisms. Non-genetically modified plants can be developed using base editing systems. Disease-resistant crops have been produced using gene editing and are likely to gain greater public acceptance than conventionally genetically modified plants. Genome editing can enhance crop productivity to meet current and future nutritional requirements worldwide.

In summary, the Research Topic brought together recent findings and literature on utilizing advanced high-throughput technologies to enhance plant resilience against environmental challenges. These insights have shed light on how plants can withstand and adapt to environmental stress by applying such technologies (Veley et al., 2017; Gogolev et al., 2021).

Author contributions

FM-P wrote the first draft of the manuscript. All authors contributed to the conception of the Research Topic, manuscript revision, editing, and approved the submitted version.

Funding

FM-P was supported by the National Research and Development Agency (ANID, Chile) grant FONDECYT/Regular 1231681.

Conflict of interest

The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

Publisher's note

References

Choi, H. K. (2019). Translational genomics and multi-omics integrated approaches as a useful strategy for crop breeding. *Genes Genomics* 41, 133–146. doi: 10.1007/s13258-018-0751-8

Gogolev, Y. V., Ahmar, S., Akpinar, B. A., Budak, H., Kiryushkin, A. S., Gorshkov, V. Y., et al. (2021). Omics, epigenetics, and genome editing techniques for food and nutritional security. *Plants* 10 (7), 1423. doi: 10.3390/plants10071423

Kumar, A., Anju, T., Kumar, S., Chhapekar, S. S., Sreedharan, S., Singh, S., et al. (2021). Integrating omics and gene editing tools for rapid improvement of traditional food plants for diversified and sustainable food security. *Int. J. Mol. Sci.* 22 (15), 8093. doi: 10.3390/ijms22158093

Veley, K. M., Berry, J. C., Fentress, S. J., Schachtman, D. P., Baxter, I., and Bart, R. (2017). High-throughput profiling and analysis of plant responses over time to abiotic stress. *Plant Direct* 1 (4), e00023. doi: 10.1002/pld3.23

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 December 2021
doi: 10.3389/fpls.2021.788377

Transcriptomics Integrated With Metabolomics Reveal the Effects of Ultraviolet-B Radiation on Flavonoid Biosynthesis in Antarctic Moss

*Shenghao Liu[1,2], Shuo Fang[1], Chenlin Liu[1], Linlin Zhao[1,2], Bailin Cong[1] and Zhaohui Zhang[1,2]**

[1] Key Laboratory of Marine Ecology and Environment Science, First Institute of Oceanography, Natural Resources Ministry, Qingdao, China, [2] Marine Ecology and Environmental Science Laboratory, Pilot National Laboratory for Marine Science and Technology, Qingdao, China

Bryophytes are the dominant vegetation in the Antarctic continent. They have suffered more unpleasant ultraviolet radiation due to the Antarctic ozone layer destruction. However, it remains unclear about the molecular mechanism of Antarctic moss acclimation to UV-B light. Here, the transcriptomics and metabolomics approaches were conducted to uncover transcriptional and metabolic profiling of the Antarctic moss *Leptobryum pyriforme* under UV-B radiation. Totally, 67,290 unigenes with N_{50} length of 2,055 bp were assembled. Of them, 1,594 unigenes were significantly up-regulated and 3353 unigenes were markedly down-regulated under UV-B radiation. These differentially expressed genes (DEGs) involved in UV-B signaling, flavonoid biosynthesis, ROS scavenging, and DNA repair. In addition, a total of 531 metabolites were detected, while flavonoids and anthocyanins accounted for 10.36% of the total compounds. There were 49 upregulated metabolites and 41 downregulated metabolites under UV-B radiation. Flavonoids were the most significantly changed metabolites. qPCR analysis showed that UVR8-COP1-HY5 signaling pathway genes and photolyase genes (i.e., *LpUVR3*, *LpPHR1*, and *LpDPL*) were significantly up-regulated under UV-B light. In addition, the expression levels of JA signaling pathway-related genes (i.e., *OPR* and *JAZ*) and flavonoid biosynthesis-related genes were also significantly increased under UV-B radiation. The integrative data analysis showed that UVR8-mediated signaling, jasmonate signaling, flavonoid biosynthesis pathway and DNA repair system might contribute to *L. pyriforme* acclimating to UV-B radiation. Therefore, these findings present a novel knowledge for understanding the adaption of Antarctic moss to polar environments and provide a foundation for assessing the impact of global climate change on Antarctic land plants.

Keywords: abiotic stress, bryophytes, flavonoids, metabolome, transcriptome, ultraviolet-B radiation

OPEN ACCESS

Edited by:
Freddy Mora-Poblete, University of Talca, Chile

Reviewed by:
Alma Balestrazzi, University of Pavia, Italy
YeonKyeong Lee, Norwegian University of Life Sciences, Norway

***Correspondence:**
Zhaohui Zhang
zhangzhaohui@fio.org.cn

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

Received: *02 October 2021*
Accepted: *12 November 2021*
Published: *08 December 2021*

Citation:
Liu S, Fang S, Liu C, Zhao L, Cong B and Zhang Z (2021) Transcriptomics Integrated With Metabolomics Reveal the Effects of Ultraviolet-B Radiation on Flavonoid Biosynthesis in Antarctic Moss. Front. Plant Sci. 12:788377. doi: 10.3389/fpls.2021.788377

INTRODUCTION

Ultraviolet-B radiation (280–315 nm) is an inherent part of sunlight. Increased UV-B light have been observed on the earth's surface since the 1980s and 1990s due to depletion of stratospheric ozone layer, which results from increases of chlorofluorocarbons in the atmosphere (Hossaini et al., 2017; Neale et al., 2021). In addition to ozone effects, the UV-B light on the Earth's surface is also

affected by clouds, aerosols, reflectivity of the Earth's surface and solar activity (Bais et al., 2019). Since the mid-1990s, the emission of ozone depleting substances (ODS) in the atmosphere has gradually decreased and the concentration has been declining, due to the efficient implementation of the Montreal Protocol and its amendments (Mckenzie et al., 2011). Currently, UV-B radiation are only slightly stronger than in 1980 (increases less than ∼5%) at mid-latitudes, but increases are enormous at high and polar latitudes where ozone depletion is increasing markedly (Mckenzie et al., 2011). In future, changes of UV-B radiation at middle and low latitudes will possibly be dominated by concentration of aerosols, while levels of UV-B radiation at high latitudes will be affected by stratospheric ozone recovery, cloud layer and surface reflectivity (Bais et al., 2019). In Antarctica, due to the positively anticipated recovery of ozone layer, the erythemal UV-B irradiance will be reduced up to 40% in the spring of 2100 (Bais et al., 2019). However, according to NASA's report, the Antarctic ozone hole reached a peak of about 24×10^6 km^2 in early October 2020 and had extended to most areas of the Antarctic continent. The area of the Antarctic ozone hole was higher than the average in the past decade and the return of Antarctic ozone to pre-1980 levels could be substantially delayed (Hossaini et al., 2017).

The amount of UV-B light received at Earth's surface greatly influences the terrestrial and aquatic ecosystems. High UV-B levels, larger than 1 $\mu mol\ m^{-2}\ s^{-2}$ UV-B, is usually an environmental stress and cause the damage of their DNA, inducing cell to form pyrimidine dimer (Fuentes-León et al., 2020). Apart from DNA damage, high UV-B irradiance also causes membrane changes and protein crosslinking, and generates reactive oxygen species (ROS), as well as reduces the photosynthetic capacity and plant productivity (Núñez-Pons et al., 2018). However, low intensity of UV-B light can act as a growth signal mediating the plant photomorphogenesis (Kataria et al., 2014; Palma et al., 2020). Plants have evolved a variety of adaptive responses to UV-B light through morphological changes including cotyledon curling and hypocotyl growth inhibition (Dotto and Casati, 2017), and by biochemical changes including the accumulation of secondary metabolites such as flavonoids, anthocyanins, terpenoids, phenols, alkaloids, and beta-carotene (Ghasemi et al., 2019; Palma et al., 2020).

Plants are autotrophic sessile organisms and UVR8 (UV Resistance Locus 8) is the only characterized UV-B specific photoreceptor. Plants perceive UV-B light through the photoreceptor UVR8 and CONSTITUTIVE PHOTOMORPHOGENESIS1 (COP1) signaling pathway (Podolec et al., 2021). Upon UV-B light, UVR8 protein competitively binds to COP1 at the substrate binding site, inhibiting E3 ubiquitin ligase activity of COP1 and activating ELONGATED HYPOCOTYL5 (HY5) transcription factor (Tilbrook et al., 2016). The UVR8-COP1-HY5 form the main signaling components mediating the UV-B light signal transduction. In addition, UVR8 can also bind to a different set of transcription factors and directly inhibit their DNA binding. UVR8 can be inactivated by two WD40-repeat proteins RUP1 and RUP2, which provides negative feedback regulation (Podolec et al., 2021). The most current knowledge obtained from *Arabidopsis* showed that UVR8-mediated UV-B responses are accumulation of flavonols and anthocyanins, inhibition of hypocotyl growth, and changes in gene expression or protein accumulation (Yin and Ulm, 2017; Podolec et al., 2021). The UVR8-mediated signaling for inducing flavonoid accumulation was also conserved in the liverwort *Marchantia polymorpha* (Bowman et al., 2017; Kondou et al., 2019). Interestingly, UVR8-mediated signal transduction against UV-B radiation were also demonstrated in the green alga *Chlamydomonas reinhardtii* (Tilbrook et al., 2016). Although algae can generate purple phenolic pigments under abiotic stresses, UVB-induced flavonoid accumulation hasn't been identified in algae (Davies et al., 2020).

Plants in the terrestrial ecosystems of the Antarctic continent are undergoing higher UV-B radiation of 3.4–6.2 mW/cm^2 (Bao et al., 2018). These strong UV-B irradiation seriously restricts the growth and distribution of Antarctic terrestrial plants. Mosses and lichens are the dominant vegetation in the Antarctic ice free regions. The Antarctic field experiments demonstrate that chlorophyll contents were reduced in moss (*Bryum argenteum*) and lichen (*Umbilicaria aprina*) by continuing UV-B radiation, whereas the contents of UV-B absorbing compounds are increased (Singh and Singh, 2014). Several phenylpropanoids were found to function as protective barrier that increase resistance to UV radiation of moss *Ceratodon purpureus* in Antarctica (Clarke and Robinson, 2008). Meanwhile, flavonoids and carotenoids extracted from three Antarctic species [i.e., *Polytrichum juniperinum* Hedw, *Colobanthus quitensis* (Kunth) Bartl, and *Deschampsia antarctica* Desv] demonstrated the characteristics of UV-absorbing compounds, protecting cells and activating the DNA damage repair process (Pereira et al., 2009). It seems that these basal land plants can fight against the UV-B radiation through synthesizing antioxidants including UV-B-absorbing pigments, flavonoids and anthocyanins, functioning as the effective damage repair systems (Singh et al., 2011). However, anthocyanin compounds were not detected in the model plant *Physcomitrella patens* (Wolf et al., 2010). Currently, there are still limited about systematic and in-depth studies on the regulation of flavonoid biosynthesis and stress response in these basal land plants.

Here, the transcriptomics and metabolomics approaches were conducted to uncover transcriptional profiling of the Antarctic moss *Leptobryum pyriforme* under UV-B radiation. A total of 4947 differential expressed genes (DEGs) and 90 significantly changed metabolites (SCMs) were detected. Among them, flavonoids were one of the most significantly changed metabolites. An integrative data analysis demonstrated that UVR8-mediated signal signaling, Jasmonate signaling, flavonoid biosynthesis pathway and DNA repair system might play a critical role in the adaptation of the Antarctic moss *L. pyriforme* to

UV-B radiation. Interestingly, our evidence confirmed that anthocyanins compounds were present in the Antarctic moss *L. pyriforme*.

EXPERIMENTAL PROCEDURES

Plant Samples and Ultraviolet-B Radiation Treatments

The moss samples were collected from the Fildes Peninsula of Antarctica (S62°12.851′, W58°56.253′; **Figure 1**). The moss *L. pyriforme* was separated and purified from mixed growth samples. They were cultivated on a soil mixture of Pindstrup substrate (Pindstrup Mosebrug A/S, Ryomgaard, Denmark) and local soil (ratio 1:1) at 16°C, 50 μmol photons·m^{-2}·s^{-1} light with a 16-h-light/8-h-dark photoperiod. Seedlings were covered by transparent plastic film to keep moisture. Under this condition, the UV light is 0.09 mW/cm^2 that produced by the Philips T8 TLD36W/54-765 fluorescent tubes.

Two Philips TL20W/01RS narrowband UV-B tubes were used for UV-B treatment as described previously (Wolf et al., 2010). The average level of UV-B light was 0.30 mW/cm^2, which was measured by a UV-340A Ultraviolet Light Meter (Lutron Electronic Enterprise, Taiwan). Two Philips T8 TLD36W/54-765 fluorescent tubes were used to supplement the white light field and the photosynthetically active radiation was 13.5 μmol photons·m^{-2}·s^{-1} (1000 lux) which was measured by a LX-101A Light Meter (Lutron Electronic Enterprise). Two-month-old plants were firstly acclimated to a low-white light field (13.5 μmol photons·m^{-2}·s^{-1}) for 48 h. Seedlings were then treated with 0.30 mW/cm^2 UV-B light for 1.5 h, which were used for transcriptome sequencing. Seedlings treated with 0.30 mW/cm^2 UV-B light for 5 days were used for LC-MS/MS analysis. The plants without UV-B radiation were collected and used as control group. The plants collected from different flowerpots were divided into three biological replicates. The green gametophytes were collected and rapidly frozen in liquid nitrogen, and stored at −80°C.

FIGURE 1 | The field picture of Antarctic moss sampling site.

Measurement of Total Antioxidant Capacity, Proline, Total Chlorophyll and Flavonoid Contents

Mosses were treated with UV-B light as described above. After 6 h or 72 h of UV-B radiation, the moss gametophytes were cut off and ground in liquid nitrogen. The extraction and reaction buffer were obtained from commercial kits (Nanjing Jiancheng Bioengineering Institute, Nanjing, China). 0.1 g of sample powder was used for each biochemical determining according to the kit instructions. Briefly, total antioxidant capacity was measured by ferric-reducing antioxidant power (FRAP) method. Proline quantitation was achievable by reaction with ninhydrin. Total chlorophylls were dissolved in organic solvent and detected on spectrophotometer according to the Lambert-Beer's Law. Total flavonoids content was determined by ultraviolet-visible spectrophotometry at wavelengths of 325 nm. The experiments were repeated three times.

RNA Isolation and Transcriptome Sequencing

Transcriptome sequencing were conducted following the standard procedure (Zhang et al., 2019). Briefly, 2 g of moss gametophyte samples were ground into power in liquid nitrogen. Total RNA was extracted by using Trizol reagent (Invitrogen, United States). RNA integrity was detected by Agilent 2100 Bioanalyzer (Agilent Technologies, United States), while RNA quality was analyzed by 1% agarose gel electrophoresis. mRNA was isolated from total RNA by oligo (dT) magnetic beads (New England Biolabs, United States). The enriched mRNA was broken into short fragments. cDNA fragments were synthesized after reverse transcription reaction. cDNA libraries were produced from appropriately fragmented cDNA using NEBNext® Ultra™ RNA Library Prep Kit (New England Biolabs). The qualities of cDNA libraries of three UV-B groups and three control groups were detected by the Agilent Bioanalyzer 2100. Finally, the libraries were sequenced on an Illumina Hiseq 2500 platform.

Sequence Assembly, Functional Annotation and Differentially Expressed Genes Analysis

Transcriptome assembly was performed by Trinity software (Grabherr et al., 2011). Gene function was annotated using BLAST alignment against Non-Redundant Protein Sequences (NR) and Swiss-Prot databases. Gene Ontology (GO) enrichment was implemented by GOseqR packages. Kyoto encyclopedia of genes and genomes (KEGG) pathway enrichment was carried out by KOBAS software (Mao et al., 2005). The gene expression levels were estimated between UV-B treatments and control groups by the EdgeR package (Robinson et al., 2010). The gene expression levels were estimated with RPKM (Reads Per Kilobase per Million mapped reads). The $|\log_2(\text{Treat/Control})| > 1$ and the adjusted p-value < 0.005 were employed as the threshold to discriminate the DEGs.

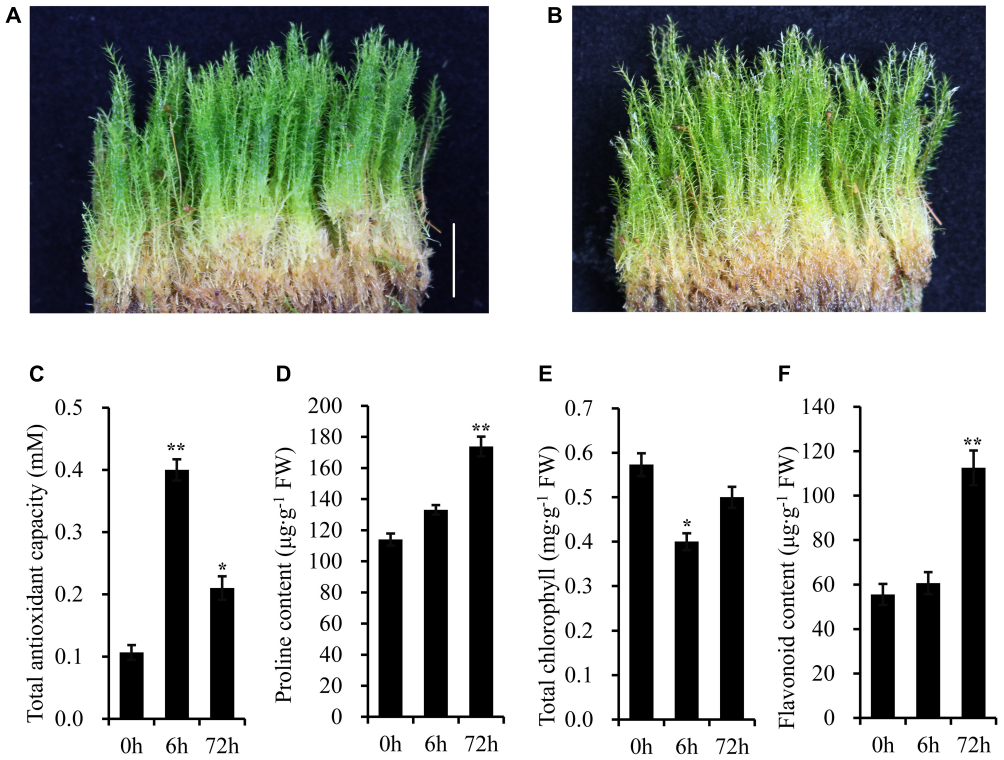


FIGURE 2 | The Antarctic moss *L. pyriforme* can tolerance strong UV-B radiation. **(A)** Photo of *L. pyriforme* under nornal condition. **(B)** Morphological changes of *L. pyriforme* under UV-B radiation. **(C–F)** The physiological parameters of *L. pyriforme* under UV-B radiation. Bar = 1.0 cm.

Phylogenetic Analysis

BLASTP alignment and HMMER program were used to identify the 2-oxoglutarate-dependent dioxygenase (2-OGD) family proteins, chalcone synthase (CHS), chalcone isomerase (CHI), flavonoid 3′-hydroxylase (F3′H), and flavonoid 3′,5′-hydroxylase (F3′,5′H) from the transcriptome data. Several representative 2-OGD, CHS, CHI, F3′H, and F3′,5′H from other land plants were retrieved from GenBank. Multiple sequence alignments were conducted using the ClustalW program. The phylogenetic tree was constructed by the neighbor-joining method using the Mega 6.0 (Tamura et al., 2013). The bootstrap values of each branch were calculated by 1000 bootstrap replicates.

Quantitative RT-PCR Analysis

To validate the expression levels of DGEs in transcriptome Sequencing, quantitative reverse transcription-polymerase chain reaction (RT-PCR) analysis were performed. Total RNA was isolated from moss gametophytes and 0.5 ng of total RNA were used to synthesize the first-strand cDNA using the *TransScript*® All-in-One First-Strand cDNA Synthesis SuperMix for qPCR with One-Step gDNA Removal Kits (Transgen, Beijing, China). The *Actin-1* gene of *L. pyriforme* was identified as the best reference gene to normalize the template. The gene specific primers were listed in **Supplementary Table 1**. Quantitative RT-PCR analysis was performed using *PerfectStart*® Green qPCR SuperMix Kits (Transgen). The cycling regime is 95°C for 5 min, followed by 40 cycles of amplification (95°C for 10 s, 57°C for 10 s, and 72°C 10 s) and run on a LightCycler96 qPCR instrument (Roche, Switzerland). Relative gene expression levels were calculated using the comparative Ct ($2^{-\Delta\Delta Ct}$) method (Livak and Schmittgen, 2001). The experiments were carried out using three biological replicates from three different experiments.

TABLE 1 | Statistical analysis of length distribution of transcripts and unigenes.

Nucleotide length	Transcripts	Unigenes
300–500 bp	48732	24416
500–1000 bp	48920	19427
1000–2000 bp	56761	11385
>2000bp	59278	12062
Total	213691	67290
Min length (bp)	301	301
Mean length (bp)	1559	1189
Median length (bp)	1122	650
Max length (bp)	19128	19128
N_{50} length (bp)	2373	2055
N_{90} length (bp)	707	464

Metabolomic Profiling Analysis

The moss gametophytes were collected and used for LC-MS/MS analysis. The sample extraction, metabolite identification, and quantification were conducted by Wuhan Metware

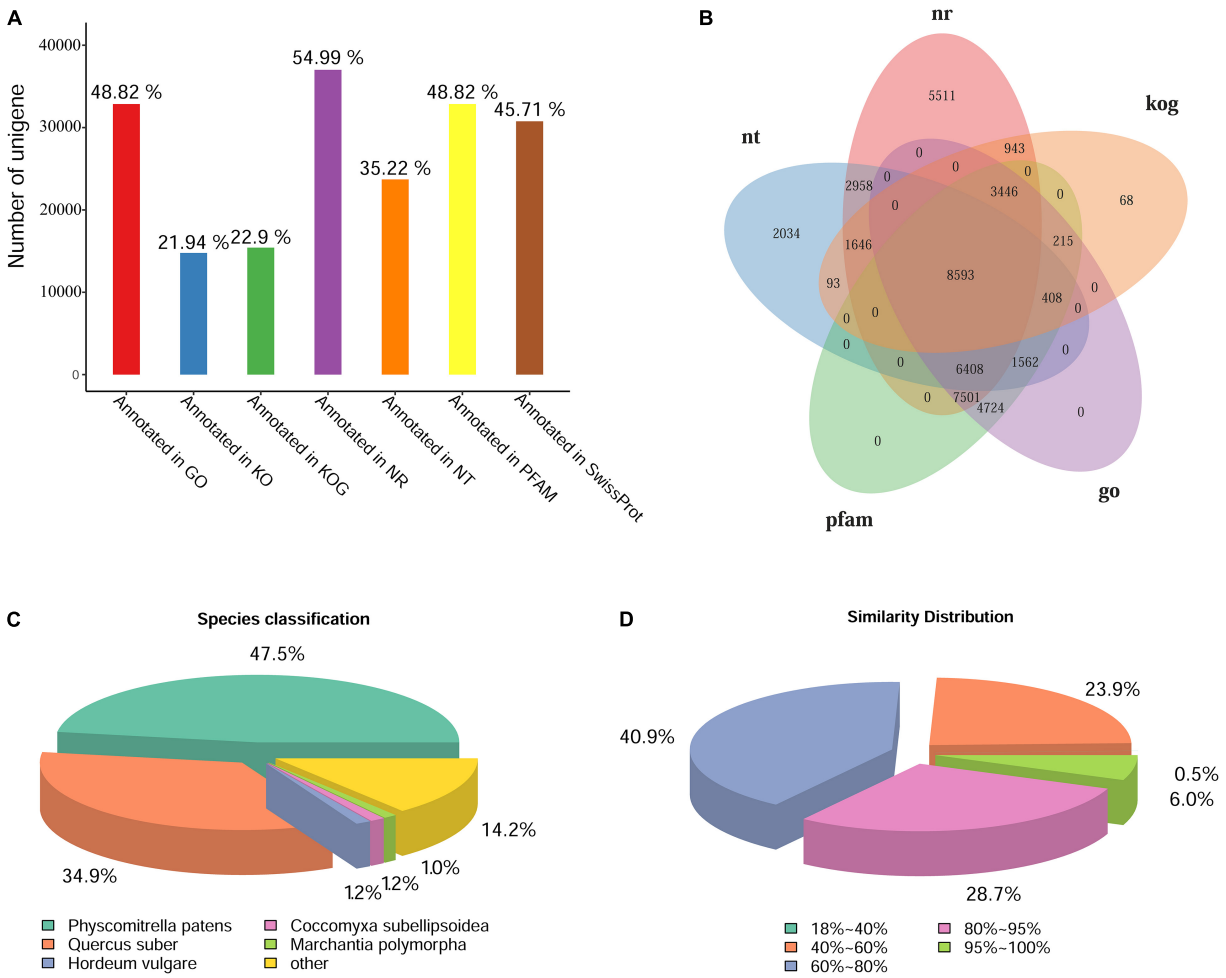


FIGURE 3 | The statistics of functional annotation of the Antarctic moss genes. **(A)** The statistics of gene annotation rate in seven databases. **(B)** Venn diagram of five database annotation results. **(C)** The statistics of species classification. **(D)** The statistics of similarity distribution. Bar = 0.5 cm.

Biotechnology Co., Ltd. following the standard procedures (Zhou et al., 2019; Li et al., 2021; Wang et al., 2021). Briefly, the freeze-dried seedlings were ground to powder using a grinder (MM 400, Retsch, Germany). Then, 0.1 g of the powder was extracted in 1.2 mL of 70% aqueous methanol at 4°C overnight. The extract mixtures were vortexed for three times during the incubation to increase the yield. The extracts were centrifuged at 10,000 g for 10 min, and the supernatant was filtered and used for LC-MS/MS analysis. The ultra-performance liquid chromatography (UPLC) (Shim-pack UFLC CBM30A, Shimadzu, Japan) and tandem mass spectrometry (MS/MS) (SCIEX QTRAP 6500, Applied Biosystems, United States) were used. The metabolites were determined by secondary spectral properties according to the public metabolite database and the self-built database with 5000 + compounds. Principal component analysis (PCA) and orthogonal projections to latent structure-discriminant analysis (OPLS-DA) were carried out for the identified metabolites. The OPLS-DA model were used to determine the relative importance of each metabolite using a parameter called "variable importance in project" (VIP). Metabolites with fold change $\geq$ 2 or fold change $\leq$ 0.5 and VIP $\geq$ 1 were considered as differentially accumulated metabolites.

Data Analysis

All experiments were performed in three biological replicates. All data were presented as the mean (± S.E.). The statistical significance of the differences between the UV-B groups and the control groups were calculated using Student's *t*-test ($^{*}P < 0.05$, $^{**}P < 0.01$).

RESULTS

Morphological and Physiological Changes Under Ultraviolet-B Radiation

Under normal conditions, the above-ground parts of gametophyte plants were green (**Figure 2A**). However, after 5 days of UV-B radiation, there was a bleaching

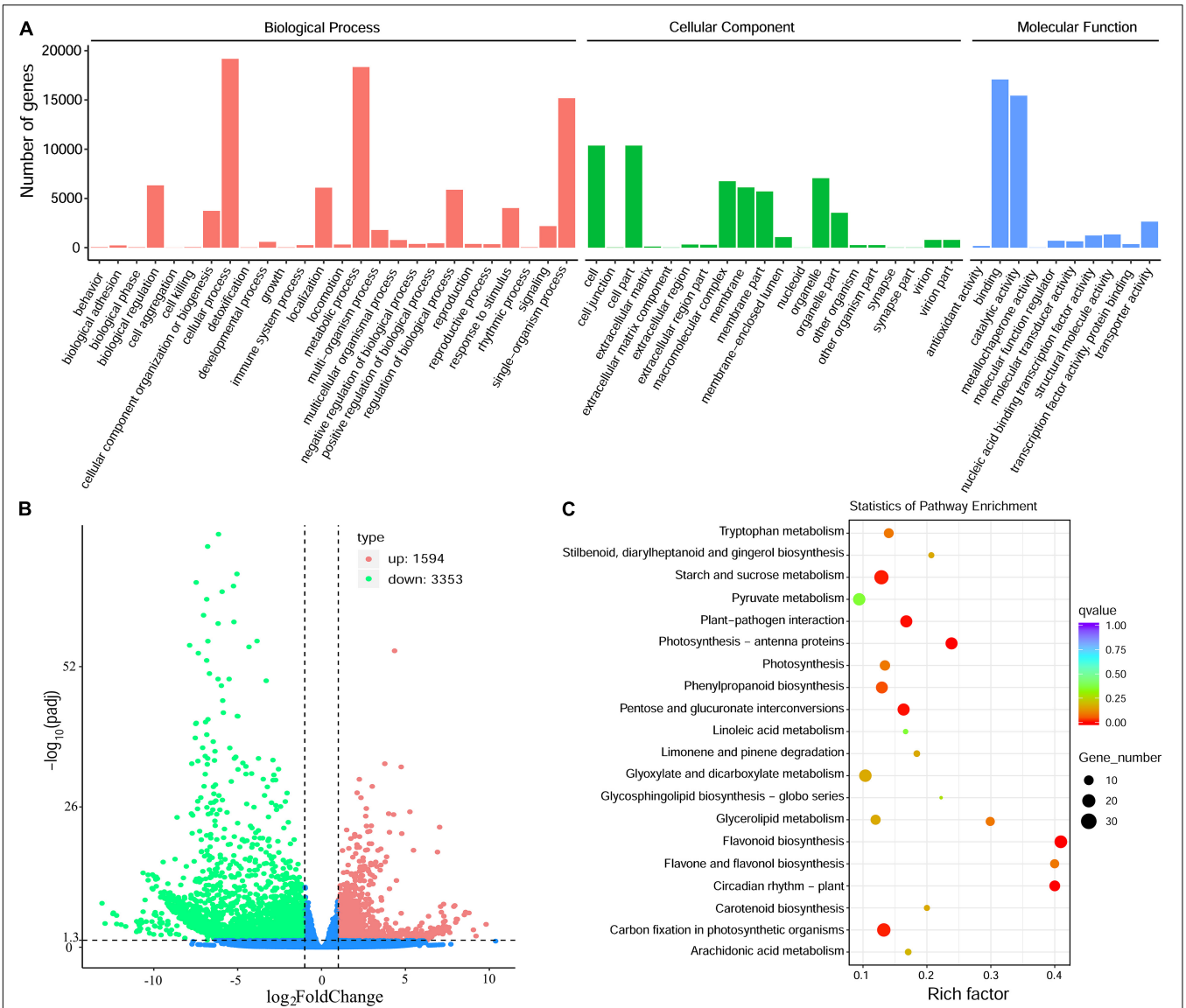


FIGURE 4 | Transcriptome sequencing of the Antarctic moss under UV-B light. **(A)** The classification of Gene Ontology of *L. pyriforme* unigenes. **(B)** The volcano plot showing the DEGs between UV-B radiation group and control group. The X-axis indicates fold change of gene expression (threshold, $|\log_2(\text{Treat/Control})| > 1$), while the Y-axis means the statistically significant level (threshold, q-value < 0.005). **(C)** KEGG pathway enrichment of DEGs. Rich factor represents the ratio of the number of DEGs to the total number of annotated genes in this pathway.

phenomenon appeared on the top of the plant (**Figure 2B**). The shoots of moss also turned yellow slightly after a long-term UV-B radiation. High level of UV-B radiation produces oxygen free radicals, and further leads to oxidative stress. Then, some physiological parameters were detected after UV-B radiation (**Figures 2C–F**). The total antioxidant capacity was significantly higher in UV-B irradiated moss than that of the control group. For example, the total antioxidant capacity was increased to 3.75-fold after 6 h of UV-B radiation. The proline content increased gradually and reached 1.52-fold at 72 h of UV-B radiation, while the total chlorophylls in UV-B irradiated moss were markedly lower (approximately 30.23% less) than that in control group plants after 6 h UV-B radiation. Finally, total flavonoids content was increased to 2.03-fold at 72h of UV-B radiation.

Statistical Analysis of Transcripts and Unigenes Obtained by Transcriptome Sequencing

The transcriptome sequencing was performed on an Illumina Hiseq 2500 platform (**Table 1**). To improve the quality and reliability of data, raw reads containing sequencing adapter or N (N represents that the base sequence is uncertain), and low-quality reads (reads with base number of $Q_{phred} \leq 20$ accounting

for more than 50% of the whole read) were removed from the original data. After filtering, the GC content distribution and the sequencing error rate were calculated, and clean reads are retrieved for subsequent analysis (**Supplementary Table 2**). Then, clean reads were assembled by Trinity software (Grabherr et al., 2011). The length of transcripts and cluster sequences (unigenes) were counted, respectively (**Table 1**). Firstly, a total of 213,691 transcripts ranging from 300 to 19,128 bp were obtained. The average length of transcripts was 1,559 bp, while the N_{50} and the N_{90} were 2,373 bp and 707 bp, respectively. Secondly, a total of 67,290 unigenes (non-redundant sequences) were assembled from the transcriptome sequencing; the statistics of length distribution showed that 24,416 unigenes were 300-500 bp, accounting for 36.28% of

TABLE 2 | Representative UV-B stress-related genes of the Antarctic moss *L. pyriforme.*

Gene_ID	Log_2Fold change (Treat/Control)	q-Value (p-adjusted)	Gene symbol	Functional annotation
UV-B signaling pathway and DNA repair system				
Cluster-30840.18892	1.25	4.36E-08	LpUVR8-1	ultraviolet-B receptor UVR8-like [*Physcomitrium patens*, XP_024386693.1]
Cluster-30840.21183	1.64	2.66E-05	LpUVR8-2	ultraviolet-B receptor UVR8-like [*Pistacia vera*, XP_031258370.1]
Cluster-30840.19346	5.70	0.00521	LpCOP1-1	E3 ubiquitin-protein ligase COP1 [*Selaginella moellendorffii*, XP_024539131.1]
Cluster-30840.14579	2.05	4.06E-08	LpCOP1-2	E3 ubiquitin-protein ligase COP1-like [*Physcomitrella patens*, XP_024371106.1]
Cluster-30840.13183	3.02	1.57E-30	LpCOP1-3	E3 ubiquitin-protein ligase COP1-like [*Physcomitrella patens*, XP_024379324.1]
Cluster-30840.1444	2.97	3.12E-23	LpCOP1-4	E3 ubiquitin-protein ligase COP1-like [*Physcomitrella patens*, XP_024371105.1]
Cluster-30840.4650	2.76	3.99E-22	LpHY5	transcription factor HY5 [*Carica papaya*, XP_021891410.1]
Cluster-30840.4633	2.26	1.12E-16	LpUVR3	(6-4) DNA photolyase [*Selaginella moellendorffii*, XP_024538612.1]
Cluster-9328.0	1.72	9.46E-06	LpPHR1	type II CPD DNA photolyase [*Pityrogramma austroamericana*, AAQ18175.1]
Cluster-20264.0	4.78	8.70E-22	LpDPL	Deoxyribodipyrimidine photolyase [*Talaromyces pinophilus*, KAF3397646.1]
Cluster-16194.0	1.86	0.00026	LpUVH1	DNA repair endonuclease UVH1-like [*Physcomitrella patens*, XP_024403241.1]
Cluster-3153.0	5.85	0.00011	LpRHP7	DNA repair protein RHP7 [*Quercus suber*, POE46986.1]
Cluster-8244.0	6.50	0.00024	LpFPG	Formamidopyrimidine-DNA glycosylase [Quercus suber, XP_023902507.1]
Flavonoid biosynthesis pathway				
Cluster-30840.15975	2.35	1.62E-11	LpCHS-1	chalcone synthase [*Pohlia nutans*, QBQ18373.1]
Cluster-30840.16944	2.03	1.26E-06	LpCHS-2	chalcone synthase [*Physcomitrium patens*, XP_024356740.1]
Cluster-30840.21027	1.51	0.028984	LpCHS-3	chalcone synthase [*Plagiochasma appendiculatum*, AIV42295.1]
Cluster-30840.24622	4.00	1.77E-25	LpCHS-4	chalcone synthase [*Selaginella moellendorffii*, XP_002965309.1]
Cluster-30840.17190	1.96	9.48E-11	LpCHI	chalcone isomerase [*Conocephalum conicum*, AOC83889.1]
Cluster-30840.18588	5.27	7.85E-26	LpF3H-1	flavanone 3-dioxygenase [*Physcomitrium patens*, XP_024379378.1]
Cluster-30840.4416	7.23	0.0023602	LpF3H-2	flavanone 3-dioxygenase [*Physcomitrium patens*, XP_024379378.1]
Cluster-30840.8906	1.53	2.58E-06	LpF3H-3	flavanone 3-dioxygenase [*Oryza brachyantha*, XP_006649297.2]
Cluster-30840.27891	2.11	0.00012	LpDMR6	DMR6-LIKE OXYGENASE [*Panicum virgatum*, XP_039773652.1]
Cluster-30840.26738	1.25	1.20E-18	LpF3′H-1	flavonoid 3′-monooxygenase [*Physcomitrium patens*, XP_024359127.1]
Cluster-28074.0	-2.78	0.00435	LpF3′H-2	flavonoid 3′-monooxygenase [*Vitis vinifera*, RVW52181.1]
Cluster-30840.24119	-3.95	1.84E-16	LpF3′H-3	flavonoid 3′-hydroxylase [*Raphanus sativus*, BAX90118.1]
Cluster-30840.2097	-2.37	0.00016	LpF3′,5′H-1	flavonoid 3′,5′-hydroxylase [*Pohlia nutans*, AHI15953.1]
Cluster-30840.23977	-4.88	2.59E-19	LpF3′,5′H-2	flavonoid 3′,5′-hydroxylase [*Handroanthus impetiginosus*, PIM99139.1]
Cluster-743.0	2.17	1.25E-07	LpANS-1	2-oxoglutarate-dependent dioxygenase ANS [*Physcomitrium patens*, XP_024373110.1]
Cluster-30840.16084	1.83	7.38E-09	LpANS-2	2-oxoglutarate-dependent dioxygenase ANS [*Selaginella moellendorffii*, XP_002974642.2]
Cluster-30840.19732	2.30	1.22E-07	LpANS-3	2-oxoglutarate-dependent dioxygenase ANS [*Physcomitrium patens*, XP_024374366.1]

(Continued)

TABLE 2 | (Continued)

Gene_ID	Log_2Fold change (Treat/Control)	q-Value (p-adjusted)	Gene symbol	Functional annotation
Jasmonic acid signaling pathway				
Cluster-30840.18763	1.57	5.20E-05	LpOPR-1	12-oxophytodienoate reductase 3 [*Rosa chinensis*, XP_024181065.1]
Cluster-30840.15749	4.35	1.17E-55	LpOPR-2	12-oxophytodienoate reductase 11 [*Physcomitrium patens*, XP_024365650.1]
Cluster-30840.2057	1.41	2.60E-05	LpOPR-3	12-oxophytodienoic acid reductase [*Chlorella sorokiniana*, PRW44569.1]
Cluster-20359.0	3.04	4.79E-16	LpJAZ-1	jasmonate ZIM domain protein 1 [*Calohypnum plumiforme*, QTY21848.1]
Cluster-30840.16608	1.76	0.01032	LpJAZ-2	jasmonate ZIM domain protein 3 [*Calohypnum plumiforme*, QTY21850.1]
Cluster-30840.17939	2.93	5.81E-16	LpJAZ-3	Jasmonate ZIM-domain protein 9 [*Saccharum* hybrid cultivar ROC22, AVF19699.1]
Cluster-30840.2623	4.29	2.59E-25	LpRUP2	WD repeat-containing protein RUP2 [*Physcomitrella patens*, XP_024394565.1]
Other plant abiotic stress resistance pathways				
Cluster-30840.28529	7.94	1.87E-08	LpGST-1	glutathione S-transferase [*Klebsormidium nitens*, GAQ91203.1]
Cluster-30840.14319	2.81	8.60E-15	LpGST-2	DHAR class glutathione S-transferase [*Physcomitrella patens*, AFZ39124.1]
Cluster-30840.28471	7.88	2.88E-08	LpMAP3K	Mitogen-activated protein kinase kinase kinase A [*Symbiodinium microadriaticum*, OLP91855.1]
Cluster-30840.1781	7.04	5.51E-23	LpADH	alcohol dehydrogenase [*Quercus suber*, XP_023883167.1]
Cluster-30840.1307	3.30	0.022598	LpALDH	aldehyde dehydrogenase [*Quercus suber*, XP_023882895.1]
Cluster-20579.1	3.36	1.13E-18	LpERF-1	ethylene-responsive transcription factor ERF038-like [*Physcomitrium patens*, XP_024392192.1]
Cluster-30840.29018	2.97	4.08E-09	LpERF-2	ethylene-responsive transcription factor ERF022 [*Physcomitrium patens*, XP_024369554.1]

the total sequences, whereas 19,427 unigenes were 500-1000 bp, accounting for 28.87% of the total sequences. The average length of unigenes was 1,189 bp, the N_{50} and the N_{90} were 2,055 bp and 464 bp, respectively.

Gene Functional Annotation and Gene Ontology Classification

The unigenes were annotated by BLAST against the databases of GO, KOG/COG, KEGG, Nr, Nt, Pfam, Swiss-Prot. The statistics of gene annotation rate were shown in **Figure 3A**. There were 37,006 unigenes annotated in Nr database, accounting for 54.99% of the total, while there were 32,857 unigenes annotated in GO or Pfam, accounting for 48.82% of the total. In addition, there were 46,422 unigenes annotated in at least one Database, accounting for 68.98% of the total. Five database annotation results were selected to draw Venn diagram, and a total of 8,593 genes were annotated into these five databases (**Figure 3B**). According to the annotation results of Nr database, the species distribution map and sequence similarity distribution map were drawn (**Figures 3C,D**). The species distribution map showed the similarity between the gene sequences of this species and its related species. Five species were identified from different taxonomic positions which bore the largest number of similar proteins (**Figure 3C**). Results showed that *L. pyriforme* was most closely related to the model plant *Physcomitrella patens*. In addition, sequence similarity distribution analysis showed that unigenes with 60%–80% similarity comparing with the protein in Nr database accounted for 40.9% of the total, and 80%–95% similarity accounted for 28.7% of the total, respectively (**Figure 3D**).

The annotated genes were classified according to three categories and further divided into 56 functional subgroups (**Figure 4A**). In biological process, these unigenes were predominantly distributed into cellular process (19,170 unigenes), metabolic process (18,337 unigenes), and single-organism process (15,172 unigenes). In cellular component, cell (10,365 unigenes) and cell part (10,365 unigenes) are the most highly enriched GO terms, followed by organelle (7,051 unigenes). In molecular function, the majority of unigenes were involved in binding (17,079 unigenes), catalytic activity (15,430 unigenes) and transporter activity (2,637 unigenes) (**Figure 4A**).

Functional Analysis of Differentially Expressed Genes in *Leptobryum pyriforme* Under Ultraviolet-B Radiation

A total of 4,947 DEGs were detected by differential expression analysis using the DEGseq R package (threshold, padj < 0.05 and $|\log_2\text{FoldChange}| > 1$). Of them, 1,594 unigenes were up-regulated and 3353 unigenes were down-regulated in the Antarctic moss *L. pyriforme* under UV-B radiation (**Figure 4B** and **Supplementary Table 3**). KEGG pathway enrichment was analyzed by rich factor, q-value, and the number of enriched

genes. Of them, Starch and sucrose metabolism, Photosynthesis-antenna proteins, Carbon fixation in photosynthetic organisms, Plant-pathogen interaction, Phenylpropanoid biosynthesis, Flavonoid biosynthesis, and Circadian rhythm were the highly enriched pathways (**Figure 4C**). The representative DEGs were shown in **Table 2**. According to their function, the DEGs were further classified into different classes involved in UV-B signaling, DNA repair, flavonoid biosynthesis, Jasmonate signaling, and ROS scavenging pathways.

Phylogenetic Analysis of Key Enzymes in Flavonoid Biosynthesis Pathway

The key enzymes involved in flavonoid biosynthesis including CHS, CHI, flavanone 3-hydroxylase (F3H), 2-oxoglutarate-dependent dioxygenase (2-ODD), flavonoid 3'-hydroxylase (F3'H), flavonoid 3',5'-hydroxylase (F3',5'H), were identified the Antarctic moss *L. pyriforme* transcriptome. The phylogenetic analysis of these enzymes showed that the CHS, CHI, 2-OGD were clustered together in each clade. In addition, F3'H and F3',5'H were clustered into one clade (**Figure 5**). Since flavone synthase, flavonol synthase, flavanone-3-hydroxylase, and anthocyanidin synthase were 2-OGD family proteins possessing the same conserved 2OG-FeII_Oxy domain and DIOX_N domain, they formed a large tree branch. Thus, it is difficult to distinguish 2-OGD family enzymes only through phylogenetic analysis. Thus, the catalytic activities of these enzymes can be further identified by using *in vitro* enzyme activity assay.

Metabolome Analysis of Antarctic Moss Under Ultraviolet-B Radiation

To uncover the potential mechanisms of the Antarctic moss adapted to UV-B stress, the metabolites were detected using the UPLC-MS/MS method. A total of 531 metabolites were detected which included 90 Amino acid and derivatives, 82 Organic acids and derivatives, 67 Lipids, 42 Nucleotide and derivates, 40 Phenylpropanoids, 34 Alkaloids, 22 Flavone, 19 Carbohydrates, 19 Terpenoids, 15 Vitamins and derivatives, 12 Alcohols, 10 Polyphenol, seven Flavanone, seven Flavonol, six Anthocyanins, six Flavonoid, five Indole derivatives, five Isoflavone, five Phenolamides, four Quinones, four Sterides, two Proanthocyanidins, and 28 Others (**Figure 6A** and **Supplementary Table 4**). Therefore, Flavonoid biosynthesis pathway products including Flavone, Flavanone, Flavonol, Flavonoid, Proanthocyanidins and Anthocyanins accounted for 10.36% of the total compounds. In addition, Naringenin, Kaempferol, Quercetin, Luteolin, Eriodictyol and Hesperetin were the key intermediate metabolites of flavonoid synthesis pathway, which were all detected in the Antarctic moss *L. pyriforme*. Furthermore, Anthocyanins as the downstream products of the flavonoid pathway, were also found, including Peonidin O-hexoside, Malvidin 3-O-glucoside, Pelargonidin, Cyanidin 3-O-rutinoside, Cyanidin 3-O-galactoside, and Peonidin 3-O-glucoside chloride (**Supplementary Table 4**).

The OPLS-DA model was conducted to understand the overall metabolic difference between the samples in each group and the degree of variation between the samples in the group. The R^2Y and Q^2Y values in OPLS-DA were greater than 0.90, demonstrating that the model was meaningful, and the differential metabolites could be screened according to Fold change and the VIP value (**Figure 6B**). In the present study, using thresholds of $|\log_2\text{Foldchange}| \geq 1$ and VIP (variable importance in project, VIP) ≥ 1, a total of 90 metabolites were identified to be significantly different in 531 metabolites after UV-B radiation. Of them, there were 49 upregulated metabolites and 41 downregulated metabolites (**Figure 6C** and **Supplementary Table 5**). Flavonoid biosynthesis pathway products accounted for 21.25% of the total differential metabolites. The top 20 SCMs in the UV-B radiation compared to the control group according to the order of $|\log_2\text{Fold change}|$ or VIP scores, were shown in **Figures 6D,E**, respectively. Among them, Syringetin that belonged to Flavonols, was the most significant metabolite with $\log_2$(Fold change) 15.48 and VIP score 5.26. Cyanidin 3-O-rutinoside, a kind of anthocyanins, was the second significantly changed metabolite with $\log_2$(Fold change) 14.68 and VIP score 5.11. The significantly changed metabolites were sorted according to the type of pathways in KEGG database. These metabolites were mainly participated in the secondary metabolic pathways, such as Anthocyanin biosynthesis, Alkaloids biosynthesis, Flavone and Flavonol biosynthesis, Flavonoid biosynthesis and other small molecules metabolism (**Figure 6F**). Meanwhile, either DEGs or SCMs in the flavonoid biosynthesis pathway were abundantly enriched under UV-B radiation. Therefore, the changes in content of these metabolites may contribute to resisting the destruction of ROS caused by UV-B radiation.

Ultraviolet-B Radiation Induces the Complex Network Responses

The recent discovery of the UV-B-specific photoreceptor UVR8 allows an in-depth evaluation of the role of downstream hormones. The DGE analysis of transcriptome sequencing showed that UVR8-COP1-HY5 signaling pathway genes were up-regulated under UV-B radiation (**Table 2**). Quantitative RT-PCR analysis confirmed that the gene expression levels of *LpUVR8-1*, *LpUVR8-2*, *LpCOP1-1*, *LpCOP1-2*, *LpHY5* were increased under UV-B radiation (**Figure 7**). Furthermore, we also found that several photolyases (i.e., *LpUVR3*, *LpPHR1* and *LpDPL*) participate in repairing the UV-induced DNA damage in a light-dependent manner, which were also up-regulated under UV-B radiation. Effects of UV-B on plants largely rely on the regulation and cross talking with hormonal pathways. In this study, JA signaling pathway-related genes (i.e., *LpOPR-1*, *LpOPR-2*, *LpJAZ-1*, *LpJAZ-2* and *LpJAZ-3*) were up-regulated under UV-B radiation (**Figure 7**). Jasmonates are critical signals involved in regulating the biosynthesis of secondary metabolites, especially in modulating anthocyanin accumulation. The differential expression analysis of transcriptome sequencing showed that several enzyme genes of flavonoid biosynthesis pathway, including *CHS*, *CHI*, *F3H*, *F3′H*, *F3′,5′H* and *ANS* were up-regulated under UV-B radiation (**Table 2**). Several flavonoid-related genes (*LpCHS-1*, *LpCHS-2*, *LpCHS-3*, *LpCHS-4*, *LpCHI*, *LpF3H-1*, *LpF3H-2*, *LpANS-1*, *LpANS-2*, *LpANS-3*)

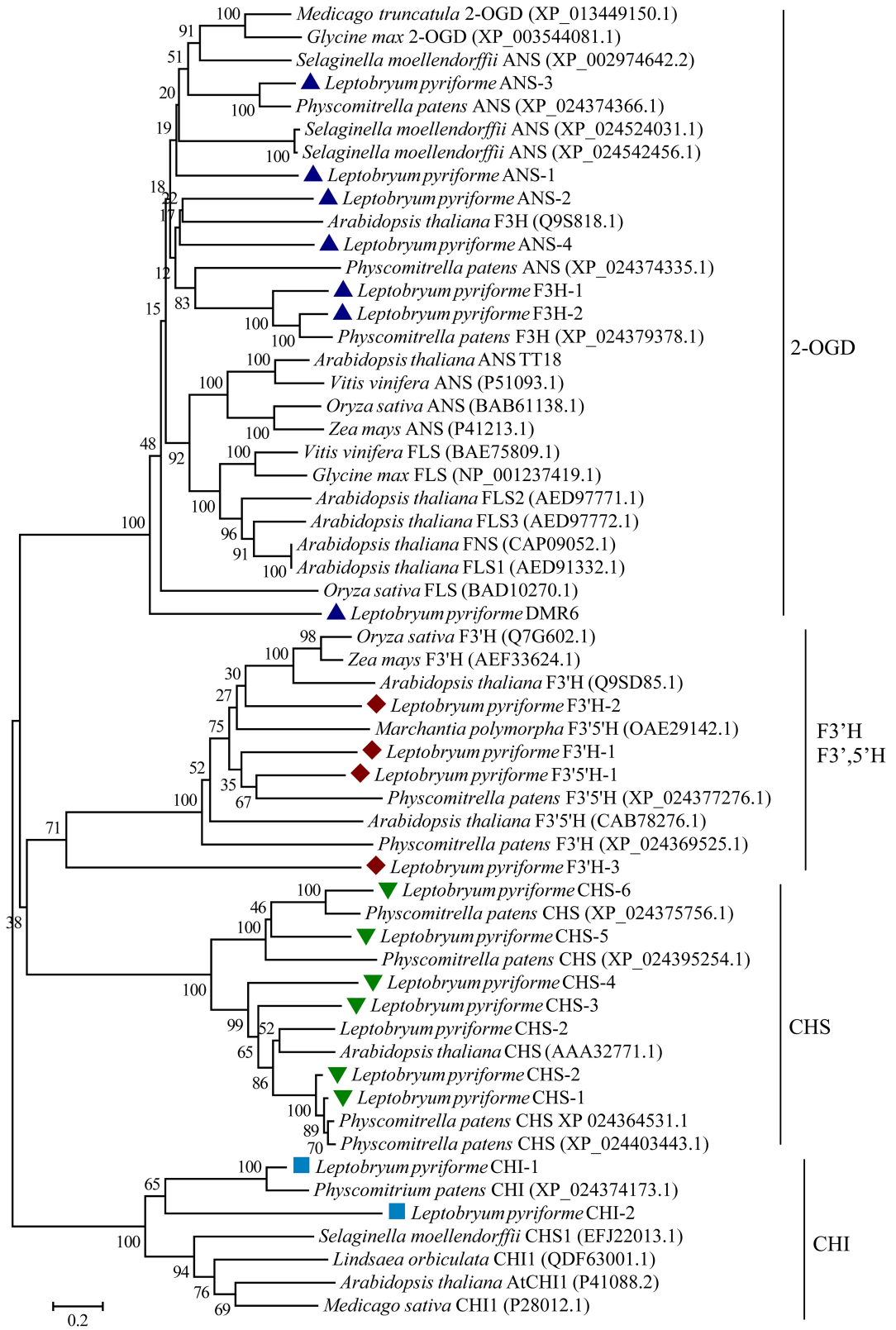


FIGURE 5 | Phylogenetic relationship of flavonoid biosynthesis-related enzymes in representative plants. 2-OGD, 2-oxoglutarate-dependent dioxygenase; ANS, anthocyanidin synthase; CHS, chalcone synthase; CHI, chalcone isomerase; F3H, flavanone 3-hydroxylase; FNS, flavone synthase; FLS, flavonol synthase; F 3′H, flavonoid 3′-hydroxylase; F3′,5′H, flavonoid 3′,5′-hydroxylase.

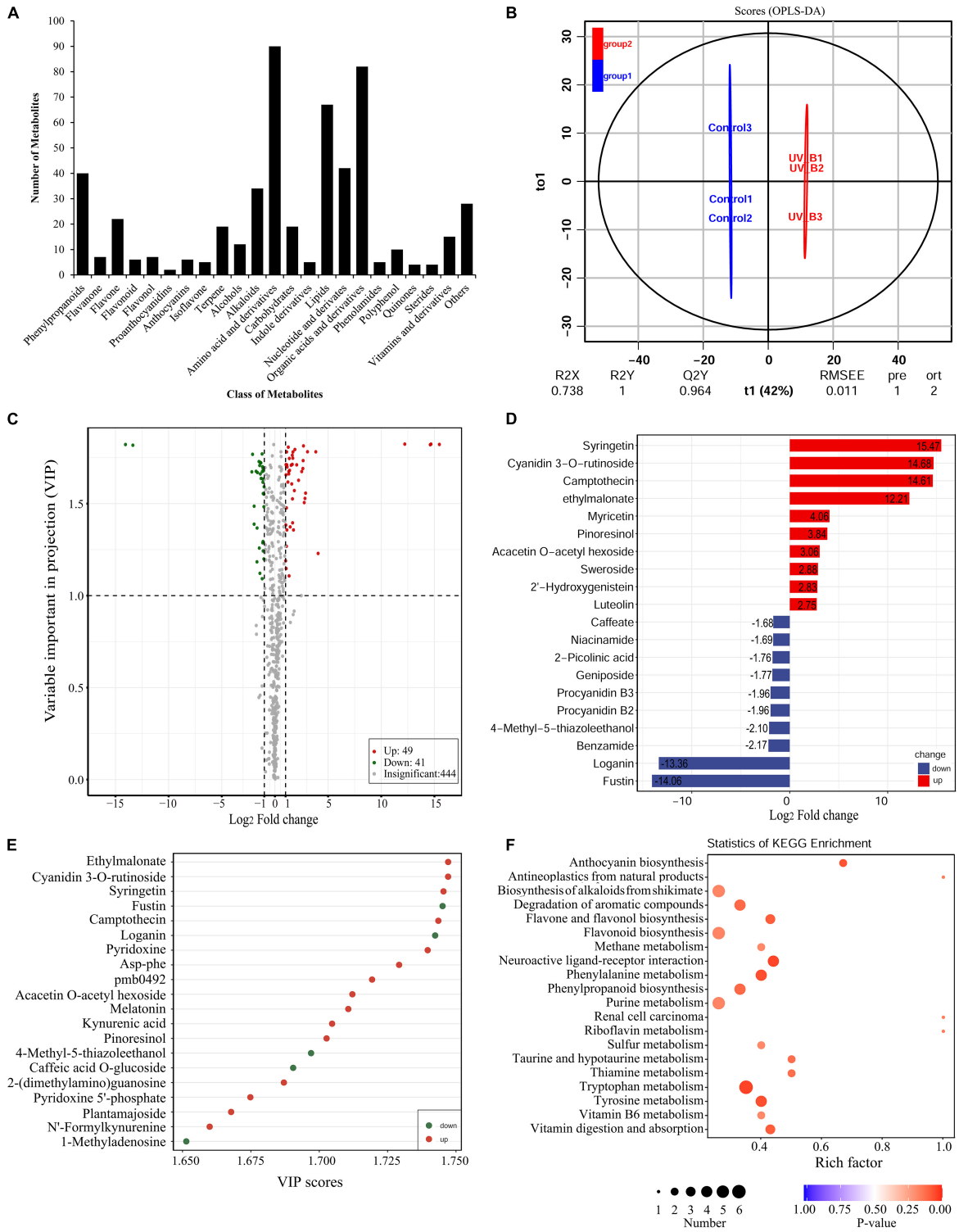


FIGURE 6 | Widely targeted metabolomic analysis of the Antarctic moss under UV-B radiation. **(A)** Statistical analysis of the classes of total metabolites. **(B)** The differences between UV-B radiation group and control group were calculated using OPLS-DA model. R^2X and R^2Y indicate the interpretation rate of X and Y matrix, respectively. Q^2Y represents the prediction ability of the model. A value closer to 1 means that the model is more stable and reliable. In addition, $Q^2Y > 0.5$ can be regarded as an effective model, and $Q^2Y > 0.9$ is an excellent model. **(C)** The volcano plot showing the contents of metabolites and the statistical significance. Each point represents a metabolite. Horizontal ordinate indicates the fold change of metabolites between two groups, while VIP value represents significant difference in statistical analysis. **(D)** The fold change of the top 20 significantly changed metabolites (SCMs) between two groups. **(E)** The VIP scores of the top 20 SCMs between two groups. **(F)** Statistics of KEGG enrichment for the SCMs.

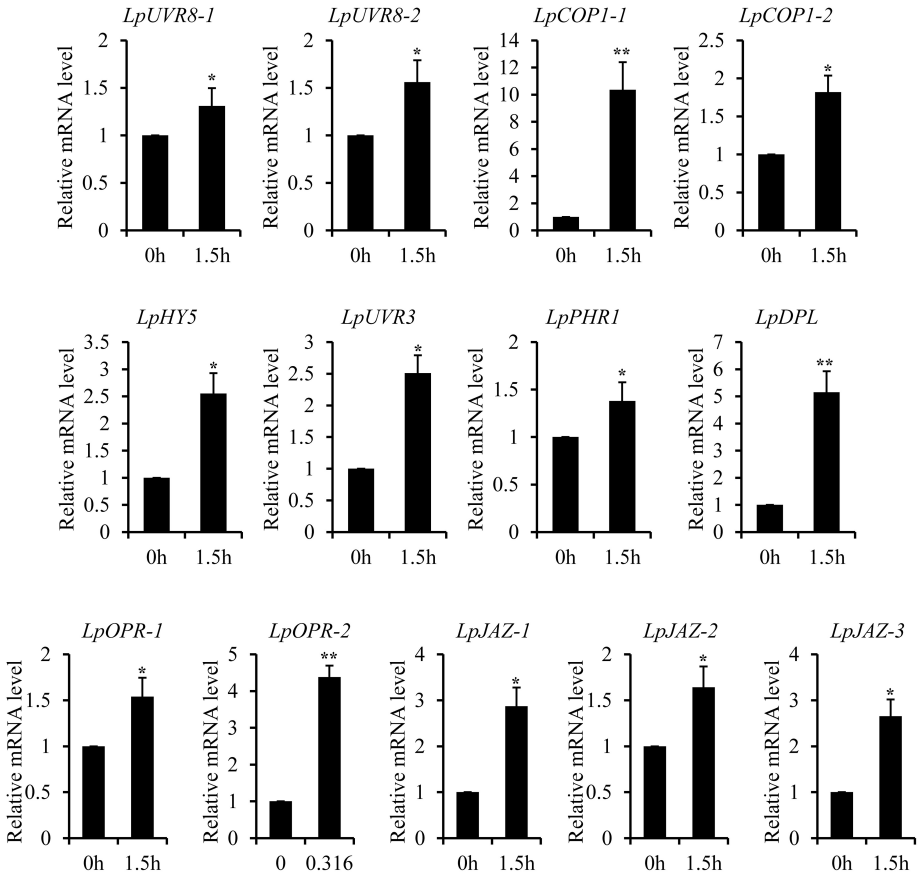


FIGURE 7 | Key genes of UVR8 and Jasmonate signaling pathway were up-regulated after UV-B treatment. The gene expression levels were analyzed by quantitative RT-PCR analysis; the Y-axis indicates the relative expression level; the X-axis indicates UV-B treatment time (h); The data were calculated from three biological replicates. Vertical bars are means ± SE. Significant difference (*$P < 0.05$, **$P < 0.01$).

were selected and the expression levels of these genes were all up-regulated after UV-B radiation which were analyzed by real-time PCR (**Figure 8**). These results were consistent with the metabolome analysis data that several Flavonoids were accumulated under UV-B radiation (**Supplementary Table 5**). Transcriptomics integrated with metabolomics showed that the contents of flavonones (Eriodictyol, 4′,5,7-Trihydroxyflavanone, Naringenin chalcone, and Naringenin) produced by the *CHS* and *CHI* genes were slightly upregulated with the fold changes of 3.14, 1.90, 1.89, and 1.72, respectively (**Figure 9** and **Supplementary Table 5**). Subsequently, flavones (Acacetin O-acetyl hexoside, Luteolin, Luteolin 7-O-glucoside, and Chrysoeriol O-hexosyl-O-pentoside), flavonols (Syringetin, Myricetin, Morin, and Kaempferol), and anthocyanins (Cyanidin 3-O-rutinoside, and Malvidin 3-O-glucoside), were generated under the action of 2-ODG family enzymes (FNS/FLS/ANS), which also showed significant upregulation under UV-B radiation (**Figure 9**). Therefore, we speculated that UV-B perception signaling, DNA repair system, Jasmonic acid signaling and Flavonoid biosynthesis pathways play the mutually coordinative roles in improving resistance against UV-B radiation in the Antarctic moss *L. pyriforme*.

DISCUSSION

In Antarctic, terrestrial ecosystems in particular experience harsh environments such as enhanced UV-B radiation, strong wind, less nutrient supply, less water availability, short growth season (Chown and Convey, 2016; Convey and Peck, 2019). Thus, Antarctic terrestrial plants are at the physiological limitations of survival. Consequently, bryophytes dominate the Antarctic land vegetation communities and only two vascular plants are present (Cannone et al., 2016; Bertini et al., 2021). Since 1974 the depletion of stratospheric ozone over the Antarctic has led to enhanced UV-B radiation (Rozema et al., 2005). It is therefore obvious that Antarctic terrestrial ecosystems have faced severe ozone depletion and enhanced UV-B for up to 47 years. However, only limited data on the impact of UV-B irradiance on Antarctic plants are available. In the present study, morphological and physiological analysis showed that the Antarctic moss *L. pyriforme* can endure strong UV-B radiation (**Figure 2**). Then, the transcriptomics and metabolomics profiling of the Antarctic moss *L. pyriforme* under UV-B radiation were detected. A total of 67,290 unigenes with the average length of 1,189 bp were generated from the transcriptome sequencing

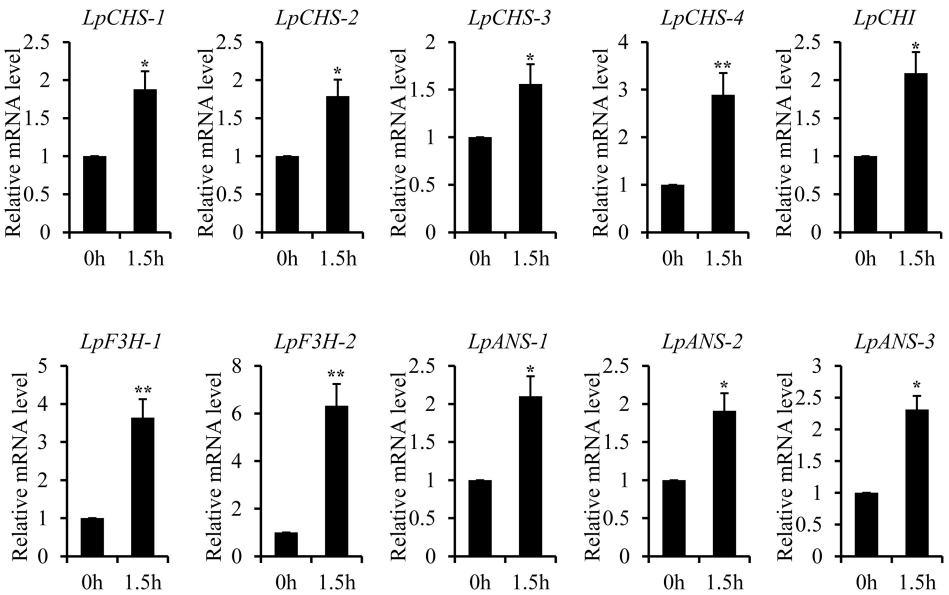


FIGURE 8 | Key enzyme genes of flavonoid synthesis pathway were up-regulated after UV-B treatment. The gene expression levels were analyzed by quantitative RT-PCR analysis; the Y-axis indicates the relative expression level; the X-axis indicates UV-B treatment time (h); The data were calculated from three biological replicates. Vertical bars are means ± SE. Significant difference (*$P < 0.05$, **$P < 0.01$).

(**Table 1**). The differential expression analysis was widely used in transcriptome sequencing to uncover stress-related genes (Wang et al., 2021). In this research, 1,594 genes were upregulated, and 3353 genes were downregulated in *L. pyriforme* under UV-B radiation (**Figure 4B**). These DEGs can be further classified into pathways of UV-B signaling (e.g., ultraviolet-B receptor UVR8, E3 ubiquitin-protein ligase COP1, and transcription factor HY5) and DNA repair system [e.g., (6–4) DNA photolyase, type II CPD DNA photolyase, deoxyribodipyrimidine photolyase, DNA repair endonuclease UVH1, and formamidopyrimidine DNA glycosylase], flavonoid biosynthesis (chalcone synthase, chalcone isomerase, lavanone 3-dioxygenase, flavonoid 3′-monooxygenase and flavonoid 3′,5′-hydroxylase), jasmonic acid signaling (12-oxophytodienoate reductase 3, jasmonate ZIM domain protein, and WD repeat-containing protein RUP2) (**Table 2**). The moss model plant *P. patens* was found to be more capable of surviving UV-B stress than Arabidopsis and approximately 400 differential expression genes were identified from moss in response to UV-B radiation (Wolf et al., 2010). UVR8 signaling well coordinate to regulate the expression of plant nuclear genes, leading to UV-B light-induced photomorphogenesis and environmental adaptation (Podolec et al., 2021). Our results demonstrated that UV-B radiation stimulated the gene expression of UVR8 signaling components and the accumulation of flavonoids in the Antarctic moss *L. pyriforme* (**Figures 6–8**). In addition, transcriptional profiling of another Antarctic moss *Pohlia nutans* showed that UV-B exposure enhanced the transcript abundance for UVR8 and flavonoid pathway genes (Li et al., 2019). These suggested that the UVR8-induced flavonoid production might act as core UV-B protection mechanism and have been already established in the moss species.

Metabolomics is one omics approach of qualitatively and quantitatively analyzing all metabolites to provide a functional screen of the cellular state (Jang et al., 2018). The HPLC-MS/MS-based plant metabolomics has been widely used to profile stress-responsive metabolites (Li and Song, 2019; Wang et al., 2021). In the present study, a widely targeted metabolomics approach based on the UPLC-MS/MS analytical platforms were used to analyze the metabolites of *L. pyriforme* under UV-B radiation. Results showed that 49 metabolites were up-regulated, and 41 metabolites were down-regulated in *L. pyriforme* under UV-B radiation (**Figure 6**). These significantly changed metabolites were classified into metabolic pathways of Anthocyanin biosynthesis, Alkaloids biosynthesis, Flavone and flavonol biosynthesis, Flavonoid biosynthesis and other small molecules metabolism (**Figure 6F**). Among the 10 metabolites with the largest increases, 6 metabolites belong to flavonoids (**Supplementary Table 5**). Among them, syringetin (a kind of flavonol) was the most significantly changed metabolite of with $\log_2$Fold change 15.47 and VIP score 5.26. Syringetin was considered as one of potent anti-photoaging agents due to its UV-absorbing and antioxidant properties (Jung et al., 2016). Cyanidin 3-O-rutinoside (a kind of anthocyanin) was the second significantly changed metabolite with $\log_2$Fold change 14.68 and VIP score 5.11 (**Supplementary Table 5**). Therefore, these results provide reliable evidence for the view that Antarctic bryophytes can form an effective protective mechanism by synthesizing UV-absorbing pigments (anthocyanins and carotenoids). Reports had demonstrated that the Antarctic mosses (*Ceratodon perpureus*, *Bryum pseudotriquetrum*, *Grimmia antarctici*, *Schistidium antarctici*) and liverwort (*M. polymorpha*) contain anthocyanins, while Antarctic algae do not contain anthocyanins

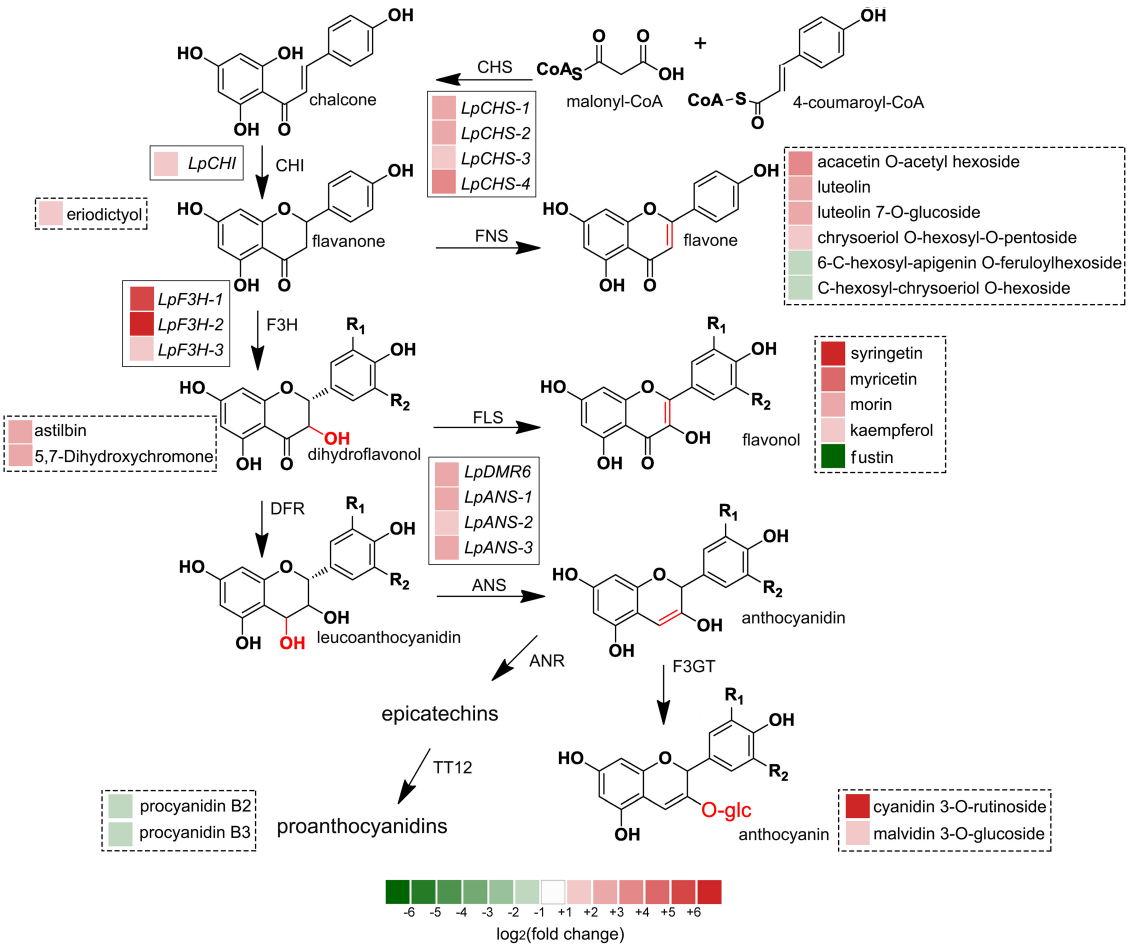


FIGURE 9 | Integrated transcriptome and metabolome analysis showed that flavonoid biosynthesis might contribute the resistance of Antarctic moss against UV-B radiation. The left block of each gene and metabolite indicated the $\log_2$(fold change) of this gene and metabolite between control and UV-B radiation. Genes labeled in solid line box and metabolites showed in dotted box.

(Singh et al., 2011). The accumulation of flavonoids will decrease the transmittance of UV-B light (Singh et al., 2012; Singh and Singh, 2014). Furthermore, integrated transcriptome and metabolome analysis revealed that flavonoid biosynthesis may dominate the resistance of the Antarctic moss *L. pyriforme* against UV-B radiation (**Figure 9**).

Besides flavonoids, plants also produce structurally diverse specialized metabolites, including bioactive alkaloids. Some of them were either similar to or even more active than standard antioxidants. In the present study, three alkaloids (i.e., camptothecin, L-dencichin, and melatonin) were markedly upregulated under UV-B radiation (**Supplementary Table 5**). Camptothecin generally produced by *Camptotheca acuminata* and functioned as a pentacyclic quinoline alkaloid with anti-cancer activity due to its ability to inhibit DNA topoisomerase (Zhao et al., 2017). The architecture of chromatin at a given promoter is critical for triggering the transcriptional readout (Bhadouriya et al., 2020). Here, we proposed that camptothecin and modulation of DNA topoisomerase are related with DNA stabilization against UV-B-induced denaturation while maintaining its metabolic activity. Melatonin is another alkaloid compound that is upregulated under UV-B radiation. Melatonin is a multifunctional signaling molecule, ubiquitously distributed in different parts of plants (Khan et al., 2020). Melatonin acted as a powerful growth regulator and antioxidant, which delayed leaf senescence, lessened photosynthesis inhibition, and improved redox homeostasis and the antioxidant system through a direct scavenging of ROS and reactive nitrogen species (RNS) under abiotic and biotic stress conditions (Debnath et al., 2019).

Flavonoids are widely distributed metabolites in land plants. Flavonoids are synthesized through the phenylpropanoid and acetate-malonate metabolic pathways (Buer et al., 2010). They are considered to have arisen during plant evolution from aquatic to terrestrial about 500 million years ago (Davies et al., 2020; Stiller et al., 2021). However, flavonoid biosynthetic pathway and its regulatory mechanism are less characterized for bryophytes than angiosperms (Davies et al., 2020). Through the comparison of genetic and molecular studies, it is found that bryophytes and angiosperms have both commonalities and significant differences in flavonoid biosynthesis and metabolic regulation

(Davies et al., 2020). The gene sequences of these enzymes have been identified from bryophytes, but *in vitro* enzymatic property analysis integrated with *in vivo* biological function analysis are recommended to identify the branch metabolic pathways. For example, at least 17 CHS genes were uncovered in *P. patens* genome, indicating that the gene family expansion and functional differentiation events occurred during the evolution of bryophytes (Koduri et al., 2010). In the present study, 6 CHS genes were identified from the Antarctic moss *L. pyriforme* transcriptome (**Figure 5**). The biosynthesis of flavones and flavonols requires chemical conversion of a common precursor, (2S)-flavanone, and is catalyzed by flavone synthase I (FNS I) and flavanone 3β-hydroxylases (F3Hs), respectively. Both enzymes as well as flavonol synthase (FLS) and anthocyanidin synthase (ANS) belong to the 2-OGD family proteins (Li et al., 2020; Ge et al., 2021). The functions of these enzymes in bryophytes are rarely reported. Previously, a liverwort FNSI gene was isolated from *Plagiochasma appendiculatum*, and its translating products showed efficient FNSI activity that convert naringenin to apigenin and 2-hydroxynaringenin (Han et al., 2014). The liverwort FNSIs evolved into a dual-function enzyme with both FNS I and F3H activities in both *P. patens* and *S. moellendorffii*, suggesting that they represent the functional transition forms between canonical FNSIs and F3Hs (Li et al., 2020). The sequence of key enzymes involved in flavonoid biosynthesis were also retrieved from the Antarctic moss *L. pyriforme* transcriptome (**Figure 5** and **Table 2**). However, more equivalent functional studies need to be carried out on the basal lant plants, such as mosses and liverworts.

Mosses are restricted to sparse ice-free areas of Antarctic frozen continent. They seem to have established an efficient DNA damage repair system through synthesizing antioxidants such as UV-B-absorbing pigments and anthocyanins (Singh et al., 2011; Waterman et al., 2018). However, the analysis approaches used in previous studies only utilized the simple instrument of spectrophotometer (i.e., UV-B-absorbing compounds at $AUC_{280-315nm}$, anthocyanins at A_{526nm}) (Waterman et al., 2018). Thus, it is urgent to carry out qualitative and quantitative analyses of individual flavonoids in bryophytes. In the present study, we analyzed the metabolome of the Antarctic moss *L. pyriforme* by widely targeted metabolomics technology (**Figure 6**). A total of 531 metabolites were detected and flavonoids accounted for 10.36% of the total compounds. Interestingly, Anthocyanin products were also detected, including Peonidin O-hexoside, Malvidin 3-O-glucoside, Pelargonidin, Cyanidin 3-O-rutinoside, Cyanidin 3-O-galactoside, and Peonidin 3-O-glucoside chloride (**Supplementary Table 4**). The anthocyanins cannot be detected in methanolic extracts of *P. patens* separated by HPLC method (Wolf et al., 2010). Thus, whether bryophytes can synthesize anthocyanins is still controversial and more research are recommended to draw a conclusion. In the Antarctic field experiments, the contents of total chlorophyll were decreased in moss (*B. argenteum*) and lichen (*U. aprina*) under an enhanced UV-B radiation, while the levels of carotenoids, phenolics, and UV-B-absorbing pigments were all increased (Singh and Singh, 2014). On the plants surface, the accumulation of flavonoids will decrease the transmittance of UV-B light and have antioxidant functions (Singh et al., 2011; Dias et al., 2020). Taken together, our results suggested that UV-B signaling and DNA repair system, flavonoid biosynthesis, Jasmonate signaling pathways contribute a critical role in Antarctic moss acclimating to ozone depletion and enhanced UV-B radiation conditions.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: National Center for Biotechnology Information (NCBI) BioProject database under accession number PRJNA767045.

AUTHOR CONTRIBUTIONS

SL and ZZ conceived the original research, designed the experiments, and wrote the manuscript. SL, SF, and LZ performed the experiments. SL, CL, LZ, and BC analyzed the data. ZZ supervised the experiments. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the National Natural Science Foundation of China (41976225) and Central Public-Interest Scientific Institution Basal Research Foundation of China (GY0219Q05).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.788377/full#supplementary-material

REFERENCES

Bais, A. F., Bernhard, G., Mckenzie, R. L., Aucamp, P. J., Young, P. J., Ilyas, M., et al. (2019). Ozone-climate interactions and effects on solar ultraviolet radiation. *Photochem. Photobiol. Sci.* 18, 602–640. doi: 10.1039/c8pp90059k

Bao, T., Zhu, R., Wang, P., Ye, W., Ma, D., and Xu, H. (2018). Potential effects of ultraviolet radiation reduction on tundra nitrous oxide and methane fluxes in maritime Antarctica. *Sci. Rep.* 8:3716. doi: 10.1038/s41598-018-21881-1

Bertini, L., Cozzolino, F., Proietti, S., Falconieri, G. S., Iacobucci, I., Salvia, R., et al. (2021). What Antarctic plants can tell us about climate changes: temperature as a driver for metabolic reprogramming. *Biomolecules* 11:biom11081094. doi: 10.3390/biom11081094

Bhadouriya, S. L., Mehrotra, S., Basantani, M. K., Loake, G. J., and Mehrotra, R. (2020). Role of chromatin architecture in plant stress responses: an update. *Front. Plant Sci.* 11:603380. doi: 10.3389/fpls.2020.603380

Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., et al. (2017). Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* 171, 287.e–304.e. doi: 10.1016/j.cell.2017.09.030

Buer, C. S., Imin, N., and Djordjevic, M. A. (2010). Flavonoids: new roles for old molecules. *J. Integr. Plant Biol.* 52, 98–111. doi: 10.1111/j.1744-7909.2010.00905.x

Cannone, N., Guglielmin, M., Convey, P., Worland, M. R., and Favero Longo, S. E. (2016). Vascular plant changes in extreme environments: effects of multiple drivers. *Clim. Change* 134, 651–665. doi: 10.1007/s10584-015-1551-7

Chown, S. L., and Convey, P. (2016). Antarctic entomology. *Annu. Rev. Entomol.* 61, 119–137. doi: 10.1146/annurev-ento-010715-023537

Clarke, L. J., and Robinson, S. A. (2008). Cell wall-bound ultraviolet-screening compounds explain the high ultraviolet tolerance of the Antarctic moss, *Ceratodon purpureus*. *New Phytol.* 179, 776–783. doi: 10.1111/j.1469-8137.2008.02499.x

Convey, P., and Peck, L. S. (2019). Antarctic environmental change and biological responses. *Sci. Adv.* 5, eaaz0888–eaaz0888. doi: 10.1126/sciadv.aaz0888

Davies, K. M., Jibran, R., Zhou, Y., Albert, N. W., Brummell, D. A., Jordan, B. R., et al. (2020). The evolution of flavonoid biosynthesis: a bryophyte perspective. *Front. Plant Sci.* 11:7. doi: 10.3389/fpls.2020.00007

Debnath, B., Islam, W., Li, M., Sun, Y., Lu, X., Mitra, S., et al. (2019). Melatonin mediates enhancement of stress tolerance in plants. *Int. J. Mol. Sci.* 20:ijms20051040. doi: 10.3390/ijms20051040

Dias, M. C., Pinto, D., Freitas, H., Santos, C., and Silva, A. M. S. (2020). The antioxidant system in *Olea europaea* to enhanced UV-B radiation also depends on flavonoids and secoiridoids. *Phytochemistry* 170:112199. doi: 10.1016/j.phytochem.2019.112199

Dotto, M., and Casati, P. (2017). Developmental reprogramming by UV-B radiation in plants. *Plant Sci.* 264, 96–101. doi: 10.1016/j.plantsci.2017.09.006

Fuentes-León, F., Peres De Oliveira, A., Quintero-Ruiz, N., Munford, V., Satoru Kajitani, G., Coimbra Brum, A., et al. (2020). DNA damage induced by late spring sunlight in Antarctica. *Photochem. Photobiol.* 96, 1215–1220. doi: 10.1111/php.13307

Ge, C., Tang, C., Zhu, Y. X., and Wang, G. F. (2021). Genome-wide identification of the maize 2OGD superfamily genes and their response to *Fusarium verticillioides* and *Fusarium graminearum*. *Gene* 764:145078. doi: 10.1016/j.gene.2020.145078

Ghasemi, S., Kumleh, H. H., and Kordrostami, M. (2019). Changes in the expression of some genes involved in the biosynthesis of secondary metabolites in *Cuminum cyminum* L. under UV stress. *Protoplasma* 256, 279–290. doi: 10.1007/s00709-018-1297-y

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Han, X. J., Wu, Y. F., Gao, S., Yu, H. N., Xu, R. X., Lou, H. X., et al. (2014). Functional characterization of a *Plagiochasma appendiculatum* flavone synthase I showing flavanone 2-hydroxylase activity. *FEBS Lett.* 588, 2307–2314. doi: 10.1016/j.febslet.2014.05.023

Hossaini, R., Chipperfield, M. P., Montzka, S. A., Leeson, A. A., Dhomse, S. S., and Pyle, J. A. (2017). The increasing threat to stratospheric ozone from dichloromethane. *Nat. Commun.* 8:15962. doi: 10.1038/ncomms15962

Jang, C., Chen, L., and Rabinowitz, J. D. (2018). Metabolomics and isotope tracing. *Cell* 173, 822–837. doi: 10.1016/j.cell.2018.03.055

Jung, H., Lee, E. H., Lee, T. H., and Cho, M. H. (2016). The methoxyflavonoid isosakuranetin suppresses uv-b-induced matrix metalloproteinase-1 expression and collagen degradation relevant for skin photoaging. *Int. J. Mol. Sci.* 17:ijms17091449. doi: 10.3390/ijms17091449

Kataria, S., Jajoo, A., and Guruprasad, K. N. (2014). Impact of increasing Ultraviolet-B (UV-B) radiation on photosynthetic processes. *J. Photochem. Photobiol. B* 137, 55–66. doi: 10.1016/j.jphotobiol.2014.02.004

Khan, A., Numan, M., Khan, A. L., Lee, I. J., Imran, M., Asaf, S., et al. (2020). Melatonin: awakening the defense mechanisms during plant oxidative stress. *Plants* 9:lants9040407. doi: 10.3390/plants9040407

Koduri, P. K. H., Gordon, G. S., Barker, E. I., Colpitts, C. C., Ashton, N. W., and Suh, D. Y. (2010). Genome-wide analysis of the chalcone synthase superfamily genes of *Physcomitrella patens*. *Plant Mol. Biol.* 72, 247–263. doi: 10.1007/s11103-009-9565-z

Kondou, Y., Miyagi, Y., Morito, T., Fujihira, K., Miyauchi, W., Moriyama, A., et al. (2019). Physiological function of photoreceptor UVR8 in UV-B tolerance in the liverwort *Marchantia polymorpha*. *Planta* 249, 1349–1364. doi: 10.1007/s00425-019-03090-w

Li, C., Liu, S., Zhang, W., Chen, K., and Zhang, P. (2019). Transcriptional profiling and physiological analysis reveal the critical roles of ROS-scavenging system in the Antarctic moss *Pohlia nutans* under Ultraviolet-B radiation. *Plant Physiol. Biochem.* 134, 113–122. doi: 10.1016/j.plaphy.2018.10.034

Li, D. D., Ni, R., Wang, P. P., Zhang, X. S., Wang, P. Y., Zhu, T. T., et al. (2020). Molecular basis for chemical evolution of flavones to flavonols and anthocyanins in land plants. *Plant Physiol.* 184, 1731–1743. doi: 10.1104/pp.20.01185

Li, P., Ruan, Z., Fei, Z., Yan, J., and Tang, G. (2021). Integrated transcriptome and metabolome analysis revealed that flavonoid biosynthesis may dominate the resistance of *Zanthoxylum bungeanum* against stem canker. *J. Agric. Food Chem.* 69, 6360–6378. doi: 10.1021/acs.jafc.1c00357

Li, Q., and Song, J. (2019). Analysis of widely targeted metabolites of the euhalophyte *Suaeda salsa* under saline conditions provides new insights into salt tolerance and nutritional value in halophytic species. *BMC Plant Biol.* 19:388. doi: 10.1186/s12870-019-2006-5

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Mao, X., Cai, T., Olyarchuk, J. G., and Wei, L. (2005). Automated genome annotation and pathway identification using the KEGG Orthology (KO) as a controlled vocabulary. *Bioinformatics* 21, 3787–3793. doi: 10.1093/bioinformatics/bti430

Mckenzie, R. L., Aucamp, P. J., Bais, A. F., Björn, L. O., Ilyas, M., and Madronich, S. (2011). Ozone depletion and climate change: impacts on UV radiation. *Photochem. Photobiol. Sci.* 10, 182–198. doi: 10.1039/c0pp90034f

Neale, R. E., Barnes, P. W., Robson, T. M., Neale, P. J., Williamson, C. E., Zepp, R. G., et al. (2021). Environmental effects of stratospheric ozone depletion, UV radiation, and interactions with climate change: UNEP Environmental Effects Assessment Panel, Update 2020. *Photochem. Photobiol. Sci.* 20, 1–67. doi: 10.1007/s43630-020-00001-x

Núñez-Pons, L., Avila, C., Romano, G., Verde, C., and Giordano, D. (2018). UV-protective compounds in marine organisms from the southern ocean. *Mar. Drugs* 16:md16090336. doi: 10.3390/md16090336

Palma, C. F. F., Castro-Alves, V., Morales, L. O., Rosenqvist, E., Ottosen, C. O., and Strid, Å (2020). Spectral composition of light affects sensitivity to UV-B and photoinhibition in Cucumber. *Front. Plant Sci.* 11:610011. doi: 10.3389/fpls.2020.610011

Pereira, B. K., Rosa, R. M., Da Silva, J., Guecheva, T. N., Oliveira, I. M., Ianistcki, M., et al. (2009). Protective effects of three extracts from Antarctic plants against ultraviolet radiation in several biological models. *J. Photochem. Photobiol. B* 96, 117–129. doi: 10.1016/j.jphotobiol.2009.04.011

Podolec, R., Demarsy, E., and Ulm, R. (2021). Perception and signaling of ultraviolet-B radiation in plants. *Annu. Rev. Plant Biol.* 72, 793–822. doi: 10.1146/annurev-arplant-050718-095946

Robinson, M. D., Mccarthy, D. J., and Smyth, G. K. (2010). edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Rozema, J., Boelen, P., and Blokker, P. (2005). Depletion of stratospheric ozone over the Antarctic and Arctic: responses of plants of polar terrestrial ecosystems to enhanced UV-B, an overview. *Environ. Pollut.* 137, 428–442. doi: 10.1016/j.envpol.2005.01.048

Singh, J., and Singh, R. P. (2014). Adverse effects of UV-B radiation on plants growing at schirmacher oasis, East Antarctica. *Toxicol. Int.* 21:101. doi: 10.4103/0971-6580.128815

Singh, J., Dubey, A. K., and Singh, R. P. (2011). Antarctic terrestrial ecosystem and role of pigments in enhanced UV-B radiations. *Rev. Environ. Sci. Bio/Technol.* 10, 63–77. doi: 10.1007/s11157-010-9226-3

Singh, J., Gautam, S., and Bhushan Pant, A. (2012). Effect of UV-B radiation on UV absorbing compounds and pigments of moss and lichen of *Schirmacher oasis* region, East Antarctica. *Cell Mol. Biol.* 58, 80–84. doi: 10.1170/T924

Stiller, A., Garrison, K., Gurdyumov, K., Kenner, J., Yasmin, F., Yates, P., et al. (2021). From fighting critters to saving lives: polyphenols in plant defense and human health. *Int. J. Mol. Sci.* 22:ijms22168995. doi: 10.3390/ijms22168995

Tamura, K., Stecher, G., Peterson, D., Filipski, A., and Kumar, S. (2013). MEGA6: molecular evolutionary genetics analysis version 6.0. *Mol. Biol. Evol.* 30, 2725–2729. doi: 10.1093/molbev/mst197

Tilbrook, K., Dubois, M., Crocco, C. D., Yin, R., Chappuis, R., Allorent, G., et al. (2016). UV-B perception and acclimation in *Chlamydomonas reinhardtii*. *Plant Cell* 28, 966–983. doi: 10.1105/tpc.15.00287

Wang, R., Shu, P., Zhang, C., Zhang, J., Chen, Y., Zhang, Y., et al. (2021). Integrative analyses of metabolome and genome-wide transcriptome reveal the regulatory network governing flavor formation in kiwifruit (*Actinidia chinensis*). *New Phytol.* 2021:17618. doi: 10.1111/nph.17618

Waterman, M. J., Bramley-Alves, J., Miller, R. E., Keller, P. A., and Robinson, S. A. (2018). Photoprotection enhanced by red cell wall pigments in three East Antarctic mosses. *Biol. Res.* 51, 1–13. doi: 10.1186/s40659-018-0196-1

Wolf, L., Rizzini, L., Stracke, R., Ulm, R., and Rensing, S. A. (2010). The molecular and physiological responses of *Physcomitrella patens* to ultraviolet-B radiation. *Plant Physiol.* 153, 1123–1134. doi: 10.1104/pp.110.154658

Yin, R., and Ulm, R. (2017). How plants cope with UV-B: from perception to response. *Curr. Opin. Plant Biol.* 37, 42–48. doi: 10.1016/j.pbi.2017.03.013

Zhang, W., Liu, S., Li, C., Zhang, P., and Zhang, P. (2019). Transcriptome sequencing of Antarctic moss under salt stress emphasizes the important roles of the ROS-scavenging system. *Gene* 696, 122–134. doi: 10.1016/j.gene.2019.02.037

Zhao, D., Hamilton, J. P., Pham, G. M., Crisovan, E., Wiegert-Rininger, K., Vaillancourt, B., et al. (2017). De novo genome assembly of *Camptotheca acuminata*, a natural source of the anti-cancer compound camptothecin. *Gigascience* 6, 1–7. doi: 10.1093/gigascience/gix065

Zhou, S., Kremling, K. A., Bandillo, N., Richter, A., Zhang, Y. K., Ahern, K. R., et al. (2019). Metabolome-scale genome-wide association studies reveal chemical diversity and genetic control of maize specialized metabolites. *Plant Cell* 31, 937–955. doi: 10.1105/tpc.18.00772

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

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 08 December 2021
doi: 10.3389/fpls.2021.769748

Metabolomics and Transcriptomics Integration of Early Response of *Populus tomentosa* to Reduced Nitrogen Availability

Min Chen[1,2†], Yiyi Yin[1†], Lichun Zhang[1], Xiaoqian Yang[1], Tiantian Fu[1], Xiaowei Huo[1] and Yanwei Wang[1]*

[1] National Engineering Laboratory for Tree Breeding, Key Laboratory of Genetics and Breeding in Forest Trees and Ornamental Plants, Ministry of Education, The Tree and Ornamental Plant Breeding and Biotechnology Laboratory of National Forestry and Grassland Administration, College of Biological Sciences and Biotechnology, Beijing Advanced Innovation Center for Tree Breeding by Molecular Design, Beijing Forestry University, Beijing, China, [2] School of Life Sciences, Tsinghua University, Beijing, China

Nitrogen (N) is one of the most crucial elements for plant growth and development. However, little is known about the metabolic regulation of trees under conditions of N deficiency. In this investigation, gas chromatography-mass spectrometry (GC-MS) was used to determine global changes in metabolites and regulatory pathways in *Populus tomentosa*. Thirty metabolites were found to be changed significantly under conditions of low-N stress. N deficiency resulted in increased levels of carbohydrates and decreases in amino acids and some alcohols, as well as some secondary metabolites. Furthermore, an RNA-sequencing (RNA-Seq) analysis was performed to characterize the transcriptomic profiles, and 1,662 differentially expressed genes were identified in *P. tomentosa*. Intriguingly, four pathways related to carbohydrate metabolism were enriched. Genes involved in the gibberellic acid and indole-3-acetic acid pathways were found to be responsive to low-N stress, and the contents of hormones were then validated by high-performance liquid chromatography/electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS). Coordinated metabolomics and transcriptomics analysis revealed a pattern of co-expression of five pairs of metabolites and unigenes. Overall, our investigation showed that metabolism directly related to N deficiency was depressed, while some components of energy metabolism were increased. These observations provided insights into the metabolic and molecular mechanisms underlying the interactions of N and carbon in poplar.

Keywords: metabolome, transcriptome, poplar, nitrogen deficiency, carbon

OPEN ACCESS

Edited by:
Freddy Mora-Poblete, University of Talca, Chile

Reviewed by:
Muhammad Imran, South China Agricultural University, China
Muhammad Atif Muneer, Fujian Agriculture and Forestry University, China

****Correspondence:***
Yanwei Wang
ywwang@bjfu.edu.cn

† These authors have contributed equally to this work

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

Received: *02 September 2021*
Accepted: *09 November 2021*
Published: *08 December 2021*

Citation:
Chen M, Yin Y, Zhang L, Yang X, Fu T, Huo X and Wang Y (2021) Metabolomics and Transcriptomics Integration of Early Response of Populus tomentosa to Reduced Nitrogen Availability. Front. Plant Sci. 12:769748. doi: 10.3389/fpls.2021.769748

INTRODUCTION

Plant growth is perturbed by various biotic and abiotic stresses (Rejeb et al., 2014). Among abiotic stresses, nitrogen (N) stress has a major effect on plant physiological activity. Many biological molecules, including nucleic acids, amino acids, proteins, chlorophyll, lipids, and a variety of other metabolites, contain N, which is required for their synthesis (Kusano et al., 2011). N is actively transported or taken up by the plant root system. Organisms utilize N from three sources. The first source is N_2 from the air, which can be assimilated by leguminous rhizobia (Zahran, 1999),

and the second is organic N in the soil, which can be taken up by plants in specific environments (Jones et al., 2005). However, the main sources of N, which are taken up by most higher plants *via* transporters, are ammonium (NH_4^+) and nitrate (NO_3^-) from the soil (Jackson et al., 2008). The absorption and utilization of N (NH_4^+ and NO_3^-) are highly regulated in plants (Patterson et al., 2010). A wide range of physiological activities of plants is disrupted by N deficiency, including photosynthesis, signal transduction, and the synthesis of phospholipids, endogenous hormones, and many secondary metabolites (Shao et al., 2020; de Bang et al., 2021; Mu and Chen, 2021).

Nitrogen (N) deficiency usually reduces amino acid synthesis (Albinsky et al., 2010), while metabolite profiling analyses of *Arabidopsis* and maize have shown that growth under low-N conditions causes increases in the levels of many amino acids (North et al., 2009; Broyart et al., 2010; Trachsel et al., 2013). This may be due to varying experimental conditions. For example, gamma-aminobutyric acid (GABA) was unaltered under conditions of nitrate deficiency in high-light conditions but was induced in low-light conditions in tomatoes (Urbanczyk-Wochniak and Fernie, 2005). Furthermore, N deficiency was shown to affect the biosynthesis of some carbohydrates, as the fundamental processes of carbon (C) and N metabolism are tightly coordinated (Cao et al., 2019). The C skeleton and energy provided by carbohydrates are required for photosynthesis and N uptake (Gutieìrrez et al., 2005; Zheng, 2009). It has been demonstrated that N deficiency suppresses the levels of carbohydrate and major soluble sugars, but stimulates the accumulation of starch (Scheible et al., 1997). A wide range of genes is involved in the low-N stress response in plants (Shi et al., 2016). Some N assimilation process-related genes, especially those involved in the ornithine urea cycle (OUC) and tricarboxylic acid (TCA) cycle, were identified in *Aureococcus anophagefferens* by transcriptomics analysis (Chen Q. et al., 2015). High levels of expression of genes involved in long-chain fatty acid and hydrocarbon biosynthesis were also found in *Botryococcus braunii* (Chlorophyta) under conditions of N deprivation (Fang et al., 2015). Moreover, transcription factors with driving roles in N × C interactions were shown to be associated with N stress in maize (Chen Q. et al., 2015).

It is well known that forest plantations of poplar species have large effects on C mitigation, the pulp industry, and biomass production (Studer et al., 2011). In recent years, some fast-growing tree species, such as *Populus* spp., have been widely planted worldwide (Rennenberg et al., 2010). Unlike annual plants, the yields of which are highly dependent on the addition of high-N fertilizer, perennial plants, such as trees, achieve their yields with minimal N input (20–50% less) because of their remobilization of resources, such as N in bark storage proteins (Karp and Shield, 2008). This difference exerts a large influence on life cycle studies of bioenergy chains, considering the energy consumption required to produce N fertilizer. Elucidating the genetic regulation underlying N use efficiency (NUE) (Liu et al., 2015), and identifying the important genes *via* genomic and other "-omics" approaches (Tuskan et al., 2018), will facilitate progress in genetically modified tree breeding for sustainable and efficient supply of biomass plants in the future. These investigations help maintain environmental and financial sustainability (Karp and Shield, 2008). Considering the increasingly recognized importance of forestry in ecological balance and the accelerating exhaustion of mineral resources for fertilizer, it is necessary to investigate the regulatory mechanisms and genes involved in the NUE of trees for tree improvement.

Populus tomentosa, also known as Chinese white poplar, is one of the fastest-growing poplar species; it is widely distributed in northern China and is of great economic and ecological importance (Du et al., 2012). *P. tomentosa* is considered a model system to explore and understand the morphological changes and molecular mechanisms of tree growth and development, as well as responses to the environment. Previously, our laboratory focused on systematically characterizing the molecular responses of *P. tomentosa* under conditions of N deficiency, including the global profiling of microRNAs (miRNAs) (Ren et al., 2015), the degradome (Chen M. et al., 2015), and long non-coding RNAs (lncRNAs) (Chen et al., 2016). It was reported that total C content, reactive oxygen species (ROS), ATP, peroxidase, superoxide dismutase (SOD), and glutamine synthetase (GS) were increased in two contrasting poplar clones *Nanlin 1388* and *Nanlin 895* (Wang X. et al., 2016). In addition, the results of transcriptomics analyses of N signaling, metabolism, and storage in poplar shoot growth and development have been reported. Transcriptomics studies showed that N starvation suppressed the expression of genes encoding most nitrate transporters (NRTs) and ammonium transporters (AMTs) in poplar leaves and genes involved in N assimilation in both roots and leaves (Luo et al., 2013). N starvation treatment was also shown to increase the fine root length and surface area, foliar starch concentration, and transcript abundance of several AMTs (AMT1;2) and NRTs (NRT1;2 and NRT2;4B) in the roots of slow-growing species (*P. popularis*) and a fast-growing species (*P. alba* × *P. glandulosa*) during acclimation to limiting N supply (Luo et al., 2013). Global transcriptomic reprogramming was shown to play a critical role underlying the physiological and morphological response of poplar leaves and roots to N starvation and excess (Luo et al., 2015). Similarly, global transcriptome reprogramming and activation of root growth were also revealed in poplar (*Populus tremula* × *Populus alba*) to low-N supply (Wei et al., 2013a). *PtaNAC1*-centered subnetwork was further revealed to be involved in increasing root biomass, which was helpful in the dynamic adjustment of poplar root architecture to low-N availability (Wei et al., 2013a,b). Further investigation demonstrated that *PtaNAC1* and *PtaRAP2.11* encoding transcription factors, F-box protein-encoding gene similar to *Hawaiian Skirt* (*PtaHWS*) had a markable influence on root development of poplar under low N (Dash et al., 2015).

Most studies to date have focused on morphological, physiological, and transcriptional changes in poplar, and few have examined the global metabolic changes in poplar combined with transcriptomics profiles under low-N conditions. To identify the metabolites and the corresponding regulatory pathways and genes involved in low-N signaling in trees, we identified genes and metabolites produced responsive to low-N stress through metabolomics and transcriptomics profile analyses. Intriguingly,

we detected alterations in metabolites and transcriptional reprogramming, providing insights into the physiological and metabolic changes involved in growth and development, and obtained information to improve NUE in plantations in both agriculture and forestry.

MATERIALS AND METHODS

Plant Materials and Treatment

Populus tomentosa clones (TC1521) were grown in culture on a half-strength Murashige–Skoog (MS) medium (Murashige and Skoog, 1962) (pH = 6.2) containing 20 g L^{-1} sucrose (Sigma-Aldrich, St. Louis, MO, United States) and 0.4 mg L^{-1} indole-3-butyric acid (IBA) (Sigma-Aldrich) at 25°C under a 16/8 h (day/night) photoperiod. Sixty-day-old plants were transferred into a hydroponic solution with sufficient N level for 5 days, which was changed for fresh solution every 2 days. The plants were then transferred to a solution with or without sufficient N as the control and treatment groups for 3 days as described previously (Ren et al., 2015). Briefly, plants were grown in modified half-strength mass spectrometry (MS) liquid medium (pH = 6.2) with 2 mM NH_4NO_3 (Sigma-Aldrich) and 1 mM KNO_3 (Sigma-Aldrich) as sufficient N conditions (KK) (control) or with.01 mM NH_4NO_3 and 1 mM KCl (Sigma-Aldrich) instead of KNO_3 for low-N treatment (DN). Whole *P. tomentosa* plants were harvested in the midmorning, immediately frozen in liquid N, and stored at −80°C.

Metabolite Extraction

Samples were taken from six 60-day-old *P. tomentosa* plants with or without low-N treatment and ground in liquid N, and 50 ± 2.5 mg materials were transferred to 1.5-mL tubes, followed by the addition of 1 mL of 100% methanol (precooled to −20°C) and 10 mL of phenylalanine (10 μg/mL) as an internal standard, and centrifuged for 10 s. The tubes were preheated, ultrasonicated for 15 min at 60°C, and then centrifuged for 10 min. The supernatants (0.4 mL) were then transferred to 0.2 mL of acetonitrile precooled to 0°C, and 0.4 mL ultrapure water was added to the new tubes and then centrifuged for 15 min. Then, aliquots of 200 μL of the supernatants were transferred to glass bottles and dried under a gentle stream of N_2 gas. Methoxyamine pyridine hydrochloride at a concentration of 20 mg/L (30 μL) was added to the bottles and shaken for 30 s. The oximation reaction proceeded at 37°C for 15 min. Finally, 30 μL of *N,O*-bistrifluoroacetamide (containing 1% trimethylchlorosilane) derivatization reagent was added and allowed to react for 1 h at 70°C. After these reactions, the samples were analyzed for their metabolite contents.

Gas Chromatography-Mass Spectrometry and Metabolite Profile Analysis

Metabolites were detected by gas chromatography-mass spectrometry (GC-MS) (7890A/5975C GC-MS system; Agilent Technologies, Santa Clara, CA, United States) at Shanghai Sensichip Infotech Co. Ltd. (Shanghai, China). The Restek capillary column was an HP-5 ms (30 m × .25 mm × .25 μm) (Agilent Technologies). The parameters were as follows: injection port temperature, 280°C; EI ion source temperature, 230°C; quadrupole rod temperature, 150°C; carrier gas, high-purity helium (99.99%); splitless injection; and sample size, 1 μL. The temperature program consisted of an initial temperature of 70°C for 2 min, 10°C/min up to 320°C, and was put on hold for 6 min. GC-MS was performed by the full-scan method with a range from 50 to 550 mass-to-charge. The XC/MS software was used for metabolomics data preprocessing in the R software package (R Development Core Team, Vienna, Austria) and then compiled to remove impurity peaks due to losses from the column and the sample preparation process. The results were then organized as a two-dimensional (2D) matrix, including observation values (samples), variables (retention time/mass-to-charge ratio), and peak strength. Finally, each sample was normalized relative to the total mass using the internal standard, and the normalized data were input into SIMCA-P (ver. 11.0) for principal component analysis (PCA) using the PLS-DA model with variable importance in projection (VIP) values > 1, combined with Student's *t*-test ($p \leq 0.05$) to identify the differentially expressed metabolites, and searched for metabolites in the National Institute of Standards and Technology[1] and Kyoto Encyclopedia of Genes and Genomes (KEGG)[2] database.

Metabolomics Data Analysis and Metabolic Pathway Construction

Before the data analysis, all data were standardized for mean-centering and unit-variance scaling using SIMCA-P with the default parameters (ver. 11.5[3]). Hierarchical clustering analysis (HCA) and PCA models were tested using all samples. Significant differences among metabolites between DN and KK were examined using the *t*-test ($p \leq 0.05$). A heatmap was built using Pearson's test and hierarchical clustering, performed with MATLAB 7.5 (MathWorks, Inc., Natick, MA, United States). Metabolic pathways were constructed with Metaboanalyst, and the *Arabidopsis* metabolic pathway database was used as a reference for the global algorithm. The enrichment pathways of metabolites were analyzed based on the KEGG database with a *p*-value ≤ 0.05 established as the false discovery rate (FDR) for multiple tests. The interactions among different metabolites were determined using KEGGSOAP[4] and metabolic pathway networks were constructed using Cytoscape[5].

Total RNA Extraction and Illumina Sequencing Analysis

Total RNA was isolated from three 60-day-old *P. tomentosa* plants with or without low-N treatment using TRIzol reagent (Invitrogen, Carlsbad, CA, United States). The quantity and quantity of total RNA were determined using 1% agarose gel

[1]http://www.nist.gov/index.html

[2]http://www.genome.jp/kegg/

[3]http://www.umetrics.com/simca

[4]http://www.bioconductor.org/packages/2.4/bioc/html/KEGGSOAP.html

[5]http://www.cytoscape.org

(Sigma-Aldrich) electrophoresis and an Agilent 2100 Bioanalyzer (Agilent Technologies). Complementary DNA (cDNA) libraries were constructed as described previously (Zhang et al., 2012). Briefly, total RNA was first treated with DNase I, and mRNAs were then enriched with oligo(dT) magnetic beads mixed with fragmentation buffer (Ambion, Austin, TX, United States). The fragmented messenger RNAs (mRNAs) were used to synthesize the random hexamer-primed cDNA, which was subjected to size selection and further PCR amplification. The quantity and quantity of cDNA libraries were determined with the Agilent 2100 Bioanalyzer and ABI StepOnePlus Real-Time PCR System (Applied Biosystems, Foster City, CA, United States). Finally, the cDNA libraries were sequenced using the Illumina HiSeq 2000 platform (Illumina, San Diego, CA, United States), and raw sequencing reads were processed to remove the dirty reads, i.e., reads with adapters, >5% unknown nucleotides, and low-quality reads, based on the National Institutes of Health Sequence Read Archive database (accession number: SRP063920). The obtained clean reads were applied to *de novo* assembly with Trinity[6], which contains three independent software modules: Inchworm, Chrysalis, and Butterfly (**Supplementary Figure S1**). Briefly, the programs in Trinity were applied sequentially to assemble the clean reads into unique full-length transcripts, map reads into contigs, and then assemble unigenes. Finally, unigene sequences were aligned by BLASTn to the National Center for Biotechnology Information (NCBI) non-redundant nucleotide database (NT)[7] with an *e*-value cutoff $< 10^{-5}$. Furthermore, unigenes were aligned with the NCBI non-redundant protein database (NR) (see footnote 7) and Swiss-Prot[8] protein database with an *e*-value cutoff $< 10^{-5}$. To predict and classify possible functions, unigenes were also searched against the Cluster of Orthologous Groups (COG) database[9] by BLASTx with an *e*-value cutoff $< 10^{-5}$. Unigenes not aligned to any database were further scanned by ESTScan to obtain the amino sequences of hypothetical proteins. By blasting against these databases, four parts of the analysis of unigenes were performed: SSR analysis, unigene expression annotation, single nucleotide polymorphisms (SNP) analysis, and unigene function annotation. Further, PCA analysis and unigene expression difference analysis were conducted based on unigene expression annotation. The fragments per kilobase per million reads method was used to assess the differential expression of unigenes, as described previously (Chen et al., 2013; Hou et al., 2016), and unigenes with FDR $\leq$ 0.001 and $|\log2\text{Ratio}| \geq 1$ were regarded as differentially expressed genes (DEGs) between the DN and KK groups.

Gene Ontology and Kyoto Encyclopedia of Genes and Genomes Analysis

Unigenes were aligned with Kyoto Encyclopedia of Genes and Genomes (KEGG)[2] to predict the metabolic pathways, with $p \leq 0.05$ and $q \leq 0.05$ taken to indicate significant enrichment (Hou et al., 2016; Wang X. et al., 2016), and the functions of gene products (using BLASTx[10]) with an *e*-value cutoff $< 10^{-5}$. Furthermore, Gene Ontology (GO) analysis was performed with NR annotation to annotate the functions of unigenes with the Blast2GO program[11], with $p \leq 0.05$ taken to indicate significance (Hou et al., 2016; Wang X. et al., 2016).

Hormone Quantification by High-Performance Liquid Chromatography/Electrospray Ionization Tandem Mass Spectrometry

Samples were taken from three 60-day-old *P. tomentosa* plants with or without low-N treatment and analyzed for the concentrations of indole-3-acetic acid (IAA), abscisic acid (ABA), and gibberellic acid (GA). Extraction and purification were performed as described previously (Pan et al., 2010). Briefly, the samples were ground into a powder with a mortar and pestle, and 50-mg samples were transferred to precooled 2-mL tubes and kept in liquid N. Then, 500 μL of 2-propanol/H_2O/concentrated HCl (2:1:0.002, vol/vol/vol) extraction solvent was added to each tube, and various volumes of internal standard solutions were added. The tubes were centrifuged at 100 rpm for 30 min on a shaker at 4°C. Then, 1 mL of dichloromethane was added to each sample and shaken for 30 min at 4°C. The tubes were further centrifuged at 13,000 × *g* for 5 min at 4°C. Then, 900 μL of solvent from the lower phase was transferred into screw-capped vials and concentrated using an N evaporator. The samples were redissolved in 100 μL of methanol. Then, 50 μL of sample solution was injected into the reverse-phase C18 Gemini high-performance liquid chromatography (HPLC) column for high-performance liquid chromatography/electrospray ionization tandem mass spectrometry (HPLC-ESI-MS/MS) analysis. Quantitative analysis of each plant hormone was performed as described previously (Pan et al., 2010).

Real-Time Quantitative Reverse Transcription PCR to Detect the Transcripts of Differentially Expressed Genes in Response to Low-N Treatment (LN)

Samples were taken from three 60-day-old *P. tomentosa* plants with or without low-N treatment for 0, 1, 3, or 5 days as described above. Total RNA was extracted using an RNA prep Pure Plant Kit (Tiangen, Beijing, China) and reverse transcribed using a FastQuant RT Kit (With gDNase) (Tiangen). To verify RNA sequencing (RNA-Seq) profiles in this investigation, Real-Time Quantitative Reverse Transcription PCR (qRT-PCR) was performed on an Applied Biosystems 7500 Fast Real-Time PCR System using SYBR Premix Ex Taq™ (Tli RNaseH Plus; Takara, Shiga, Japan) in accordance with the manufacturer's instructions. The primers of 14 randomly selected DEGs are listed in **Supplementary Table S1**. Reactions were performed in

[6]http://trinityrnaseq.sourceforge.net/

[7]http://www.ncbi.nlm.nih.gov

[8]http://www.expasy.ch/sprot

[9]http://www.ncbi.nlm.nih.gov/COG

[10]http://blast.ncbi.nlm.nih.gov/Blast.cgi

[11]http://www.blast2go.com/b2ghome

a volume of 10 μL containing 1 μL of cDNA, 5 μL of SYBR Green, 0.2 μL of forward primer, 0.2 μL of reverse primer, 0.2 μL of ROXII, and 3.4 μL of distilled water. The thermocycling conditions consisted of an initial denaturation step at 95°C for 10 min followed by 40 cycles of 95°C for 30 s, 95°C for 5 s, and 60°C for 30 s. All reactions were performed in triplicate for each gene. The $2^{-\Delta\Delta CT}$ relative quantification method was used to evaluate and calculate variations (Livak and Schmittgen, 2001), with 18S rRNA used as an internal reference. The correlations of gene expression between RNA-Seq and qRT-PCR were analyzed by Pearson's test with $p \leq 0.01$ taken to indicate significance.

RESULTS

Metabolite Profiling of *Populus tomentosa* Under N Deficiency

To determine the metabolomic regulatory mechanisms of the response of poplar to N deficiency, GC-MS was performed to analyze the changes in metabolites between *P. tomentosa* plants grown under conditions of low and sufficient N (DN and KK, respectively). The total ion chromatogram (TIC) is shown in **Supplementary Figure S2**. A total of 1,131 metabolites were finally identified by GC-MS (**Supplementary Table S2**), and partial least squares-discriminant analysis (PLS-DA) and PCA were performed to determine the accuracy and significance of differences between the KK and DN samples (**Supplementary Figure S3**). The KEGG enrichment analysis assigned the detected metabolites to 18 metabolic pathways, including amino acid- and sugar-related metabolism. Metabolites clustered in C fixation, starch and sucrose, and fructose and mannose metabolism were also enriched (**Supplementary Table S3**).

Among the 1,131 metabolites, 30 with significantly differential expression were identified by PLS-DA with VIP > 1 and $p \leq 0.05$ (**Figure 1**). Most of these metabolites (70%) were significantly suppressed under conditions of low-N stress. Marked reductions in the levels of metabolites with amino groups (such as valine, L-isoleucine, L-alanine, cadaverine, ethylamine, and ethanolamine) were assumed to result in the cessation of the de novo synthesis of the free amino acids because of N deficiency (**Table 1**). Furthermore, the levels of alcohols were reduced under low-nitrate stress, as shown by the *a*-hydroxycyclohexene, cyclohexanol, ethylene glycol, inositol, and xylitol contents compared with the controls. However, N deficiency led to the accumulation of some soluble sugars, including D-fructose, D-galactose, and D-glucose, in *P. tomentosa* plants.

Network Construction of Responsive Metabolites in *Populus tomentosa* Under N Deficiency

Based on the pathways of low-N-responsive metabolites, we further constructed metabolic networks to analyze their interactions (**Figure 2**). As shown in the metabolic networks, the induced sugars, including D-glucose, produced by the Calvin cycle influenced the synthesis of diverse downstream metabolites, including amino acids and putrescine belonging to the citrate cycle. Interestingly, in addition to the induced D-glucose, D-fructose, and D-galactose, other sugar-related metabolites, such as galactonic acid and oxalic acid, were enhanced under conditions of N deficiency in this study. The induced level of sugar may have been due to reduced carbohydrate metabolism during degradation and utilization by downstream metabolites of amino acids derived from N assimilation, in turn, due to low-N stress, consistent with the pattern of N assimilation-related gene regulation revealed by transcriptomics analysis, which also revealed the connection of C and N metabolism.

RNA-Seq of *Populus tomentosa* Under N Deficiency and Gene Function Annotations

To further determine the genes involved in regulating low-N signaling in poplar under conditions of N deficiency, RNA-Seq was conducted in DN and KK *P. tomentosa* plants. After discarding the contaminated raw data, 104,843,476 clean reads (52,903,032 reads from DN, 51,940,444 reads from KK) containing 943,591,2840 nt were obtained (**Supplementary Table S4**). Based on these clean reads, 154,857 and 146,823 contigs of the treatment and control, respectively, were further assembled to 71,801 (DN) and 82,908 (KK) unigenes with mean lengths of 582 and 508 nt, respectively (**Supplementary Table S5**).

The unigene sequences were then aligned to the NCBI NR protein database, Swiss-Prot, KEGG, and COG by BLASTx with an *e*-value cutoff $< 10^{-5}$, and to the NCBI NT nucleotide database by BLASTn with an *e*-value cutoff $< 10^{-5}$, to retrieve proteins with the highest sequence similarity to the given unigenes along with their protein functional annotations (**Supplementary Figure S1**). For function annotation analysis, we obtained 52,816, 54,584, 32,342, 29,222, 18,431, and 42,264 unigenes, which were annotated to the NR, NT, Swiss-Prot, KEGG, COG, and GO databases, respectively, and a final total of 59,125 unigenes were annotated (**Supplementary Tables S6, S7**).

Then, we searched the unigene sequences against the COG database to predict the possible functions and understand the global gene function distributions of the species. A total of 18,431 sequences from 59,125 unigenes were mapped to the COG database; among the 25 categories, the largest group of unigenes (n = 5,411, 29.36%) were annotated as "General function prediction only," followed by "Translation, ribosomal structure and biogenesis" (n = 3,384, 18.36%) and "Transcription" (n = 2,840, 15.41%), while genes functioning in "Nuclear structure" (n = 2, 0.01%) and "Extracellular structures" (n = 17, 0.09%) represented the smallest categories (**Figure 3**).

Global Transcriptomic Changes of *Populus tomentosa* Under N Deficiency

To identify the genes involved in the response of *P. tomentosa* to low-N stress, we further screened unigenes with significant changes based on the assembled data. A total of 1,561 downregulated and 1,101 upregulated DEGs were identified with a false discovery rate (FDR) ≤ 0.001 and $|\log2\text{Ratio}| \geq 1$ (**Supplementary Table S8a** and **Supplementary Figure S4**).

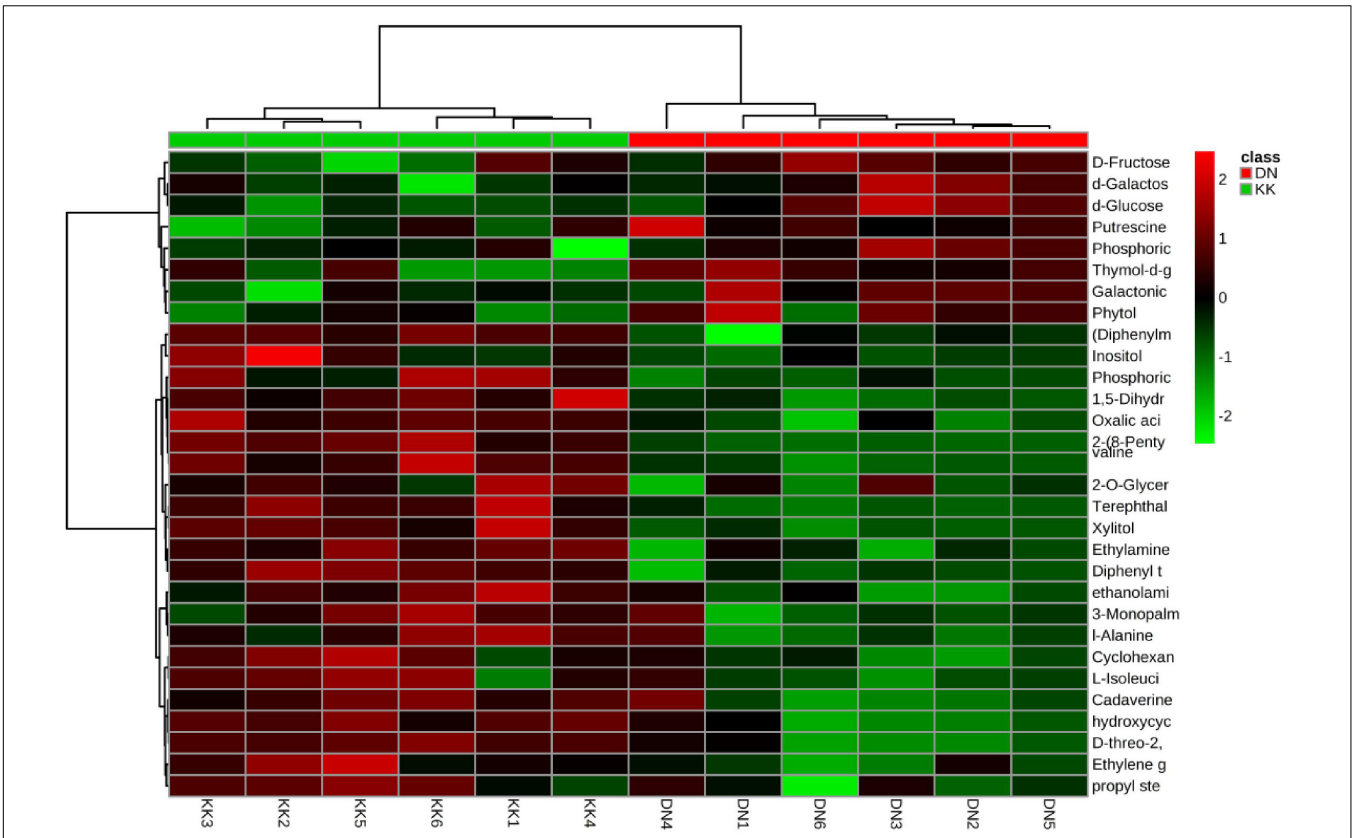


FIGURE 1 | Clustering analysis of differentially expressed metabolites in *Populus tomentosa* under nitrogen (N) deficiency. KK1, KK2, KK3, KK4, KK5, and KK6 represent six replicates of control samples, while DN1, DN2, DN3, DN4, DN5, and DN6 represent six replicates of low-N treatment samples.

To exclude statistical error, more rigorous criteria (fragments per kilobase per million reads of unigene in KK and DN > 1) were applied, and 1,017 DEGs were finally identified under conditions of N deficiency in *P. tomentosa* (**Supplementary Table S8b**). To understand the functions of these genes in the N stress response, we further aligned DEGs to the GO database. The results revealed 21, 15, and 12 classes of DEGs involved in biological process, cellular component, and molecular function, respectively (**Figure 4**). In the biological process, most DEGs fell into the cluster of "cellular process" (889, 1.50%), "metabolic process" (874, 1.48%), "single-organism process" (454,.77%), and "response to stimulus" (373,.63%). Three categories of DEGs, "organelle" (n = 931, 1.57%), "cell" (n = 1,072, 1.81%), and "cell part" (n = 1,072, 1.81%), were dominant in the cellular component, whereas "binding" (n = 604, 1.02%), "catalytic activity" (n = 464, 0.78%), and "structural molecule activity" (n = 342, 0.58%) represented the main DEG groups in molecular function (**Supplementary Table S9**).

To further understand the molecular pathways in which these *P. tomentosa* genes were involved, we scanned the DEGs into the KEGG database. As a result, 1,329 genes were finally annotated to 104 KEGG pathways, among which 16 pathways were significantly enriched at $p \leq 0.05$, including "genetic information processing," "cellular processes," "metabolism," "environmental information processing," and "organismal systems" under conditions of low-N stress in *P. tomentosa* (**Table 2**). Intriguingly, five pathways related to amino acid metabolism were significantly enriched and 20 DEGs were involved in N metabolism, indicating that *P. tomentosa* shows a marked response to low-N stress. On the other hand, four pathways related to carbohydrate metabolism were also enriched, i.e., amino sugar and nucleotide sugar metabolism, inositol phosphate metabolism, pentose and glucuronate interconversion, and propanoate metabolism, suggesting that carbohydrate regulation is involved in the response to low-N stress (**Table 2**). In addition, many DEGs could be clustered into several metabolic pathways, including N metabolism, auxin-related pathway, or function as some transporters and kinases, etc. (**Supplementary Table S10**).

Genes Related to N Acquisition, Allocation, and Assimilation in Response to Low-N Stress

Plants take up N from two sources in the soil, ammonium, and nitrate. Nitrate is reduced to nitrite (NO_2^-) and then ammonium, which is incorporated into glutamate and glutamine with a net supply of 2-oxoglutarate (provided by isocitrate dehydrogenase, IDH), catalyzed by the GS/glutamate

synthase (GOGAT) cycle. Other amino acids, such as aspartate and asparagine, are then generated by aspartate aminotransferase (*AspAT*) and asparagine synthetase (*AS*) (Suárez et al., 2002). In the nitrate assimilation process, N is incorporated into N-containing compounds, such as other amino acids, chlorophylls, and nucleic acids. Analysis of the transcriptome data revealed downregulation of N acquisition genes, such as putative nitrate reductase (NR; twofold downregulated) and GS (1.60-fold downregulated). The genes that participate in the biogenesis of various amino acids and other N-containing compounds from glutamine/glutamate were shown to be downregulated: putative IDH (11.60-fold downregulated), AspAT (1.17-fold downregulated), and AS (1.37-fold downregulated) (**Supplementary Table S10** and **Table 3**).

Consistent with these observations, several genes encoding putative amino acid synthesis-related proteins were suppressed under low-N conditions, such as proline-rich protein (*CL2231.Contig3_All*), aspartate aminotransferase (*CL8615.Contig2_All*), cationic amino acid transporter (*Unigene38715_All*), and GS (*Unigene27220_All*, *Unigene22953_All*) (**Supplementary Table S10**). These results indicated the decreased level of N assimilation and further utilization under conditions of N deficiency stress in *P. tomentosa*.

Interaction of C and N Metabolism Under Conditions of N Deficiency

C assimilation and N absorption control the status of plant growth and development. Extensive studies have shown the interdependence and interaction of C and N metabolism in photosynthesis, amino acid metabolism, lipid metabolism, and carbohydrate metabolism, including the TCA cycle, Calvin cycle, glycolysis, etc. (Sinha et al., 2015; Goel et al., 2016). Briefly, C metabolism is mostly dependent on photosynthesis, in which N participates in the biosynthesis of chlorophyll and other related proteins. Correspondingly, N assimilation demands energy and C skeletons from C metabolism.

Obviously, a low-N environment affects the functions of the chloroplasts and chlorophyll, and putative related proteins were downregulated in *P. tomentosa*, such as chlorophyll a–b binding protein of light-harvesting complex (LHC) II type I (2.73-fold downregulated), chlorophyll a–b binding protein CP26 (1.21-fold downregulated), photosystem I reaction center subunit psaK (1.37-fold downregulated), oxygen-evolving enhancer protein

TABLE 1 | Identification of changed metabolites in *Populus tomentosa* under N deficiency.

Metabolites	DN	KK	*p*-value	Fold change
Xylitol	91.93	488.77	0.01	−2.52
(Diphenylmethylene)(1-mesitylethyl) azane oxide	6.36	24.24	0.00	−1.93
Valine	80.97	202.64	0.00	−1.32
Terephthalic acid	19.14	43.26	0.00	−1.18
Diphenyl terephthalate	32.26	72.20	0.00	−1.16
2-(8-Pentyldodecahydropyrrolo[1,2-a] quinolin-3-yl) ethanol	61.45	129.38	0.00	−1.07
L-isoleucine	62.09	105.46	0.02	−0.76
L-alanine	594.97	962.70	0.02	−0.69
3-Monopalmitin ether	29.16	45.01	0.04	−0.63
1,5-Dihydroxy-6-methoxyxanthone	326.85	492.94	0.00	−0.59
Cadaverine	1343.61	2024.62	0.01	−0.59
Ethylamine	23.51	34.14	0.00	−0.54
Propyl stearate	81.68	114.46	0.05	−0.49
Cyclohexanol	45.99	64.14	0.01	−0.48
D-threo-2,5-hexodiulose	228.08	314.37	0.00	−0.46
A-hydroxycyclohexene	279.35	367.96	0.00	−0.40
Inositol	44.08	56.51	0.04	−0.36
Ethylene glycol	245.87	311.93	0.02	−0.34
Phosphoric acid	3664.68	4563.56	0.01	−0.32
2-*O*-glycerol-*a*-D-galactopyranoside	226.28	277.92	0.05	−0.30
Ethanolamine	232.47	278.27	0.01	−0.26
D-galactose	79.10	70.28	0.04	0.17
D-glucose	1299.78	1108.73	0.01	0.23
D-fructose	590.29	437.26	0.04	0.43
Putrescine	22.08	16.19	0.05	0.45
Phytol	137.39	96.67	0.03	0.51
Phosphoric acid propyl ester	48.54	27.87	0.04	0.80
Galactonic acid	24.84	14.24	0.02	0.80
Thymol-*a*-D-glucopyranoside	67.09	36.97	0.02	0.86
Oxalic acid	139.16	306.90	0.00	1.14

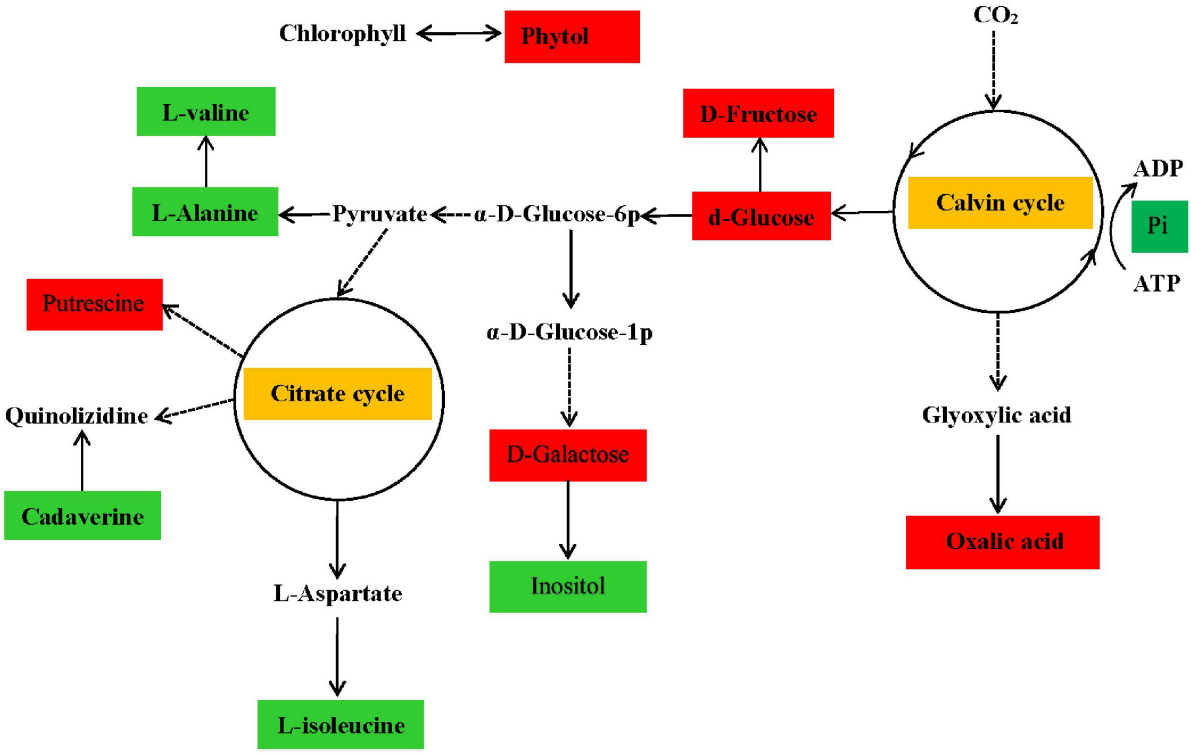


FIGURE 2 | Metabolic pathways showing the most primary metabolites in *P. tomentosa* under N deficiency. Major metabolites were enriched in the citrate and Calvin cycles. Solid arrows indicate direct transformation and dashed arrows indicate indirect or complicated connections. Metabolites indicated by red and green rectangles represent increased and decreased concentrations in the low-N treatment group, respectively.

2 of photosystem II (1.89-fold downregulated), ferredoxin–NR (1.05-fold downregulated), and chloroplast processing peptidase, indicating depressed photosynthesis capacity under low-N conditions (**Supplementary Table S10**).

On the other hand, our transcriptomics data revealed significant changes in the enzymes involved in the TCA cycle, glycolysis, and lipid degradation. Particularly, three enzymes involved in glycolysis, i.e., 6-phosphofructokinase 4, fructose-bisphosphate aldolase, and phosphoglycerate mutase, were decreased under low-N conditions. Moreover, five enzymes in the TCA cycle were altered in *P. tomentosa* under conditions of low-N stress, and genes encoding putative citrate synthase, succinyl-CoA synthetase, and succinate dehydrogenase were upregulated, while genes encoding putative aconitate hydratase, IDH (time-limiting enzyme), and PEPCase (flux regulator of TCA) were downregulated. Conversely, genes encoding five putative enzymes involved in lipid degradation were induced, i.e., stearoyl-CoA desaturase, 3-hydroxyacyl-CoA dehydrogenase, acetyl-CoA acyl-transferase, fatty acyl-CoA reductase, and long-chain-aldehyde dehydrogenase (**Supplementary Table S10**). Our data indicated decreased energy from glucose and increased energy and C skeletons from lipid, while the downregulation of putative phosphoenolpyruvate carboxylase (PEPCase) and IDH depressed the flux of oxaloacetate into the TCA cycle and lowered the production of ATP and nicotinamide adenine dinucleotide (NADH) in plants, respectively. Taken together, these results suggest reduced carbohydrate metabolism and energy production, particularly decreased energy generated from glucose in *P. tomentosa*, under conditions of low-N stress.

Hormone Signaling-Related Gene Profiles and Hormone Quantification

Auxin and ABA were reported to be closely related to N signaling (Sakakibara, 2003; Mockaitis and Estelle, 2008; Lu et al., 2015). The plant hormones IAA, ABA, and GA in *P. tomentosa* were quantified by HPLC-ESI-MS/MS in this study. The concentration of IAA was increased under low-N conditions (1.52) compared with normal conditions (1.16) (**Figure 5**). The accumulation of IAA was consistent with the increased expression of genes encoding auxin response genes. For example, putative *ARF6* (1.50-fold increase) and *IAA10* (2.23-fold increase) were markedly induced (**Table 4**).

On the other hand, the concentrations of ABA and GA were generally reduced, especially ABA, with more significant repression under conditions of low-N stress (**Figure 5**). ABA and GA are biosynthesized in the methylerythritol-4-phosphate (MEP) pathway. In the network, the conversion between isopentenyl pyrophosphate (IPP) and dimethylallyl diphosphate (DMAPP) is catalyzed by putative isopentenyl diphosphate isomerase (*IDI*; 12.09-fold). Next, isoprenyl pyrophosphate synthase (*IPPS*) catalyzes IPP to generate isopentenyl AMP

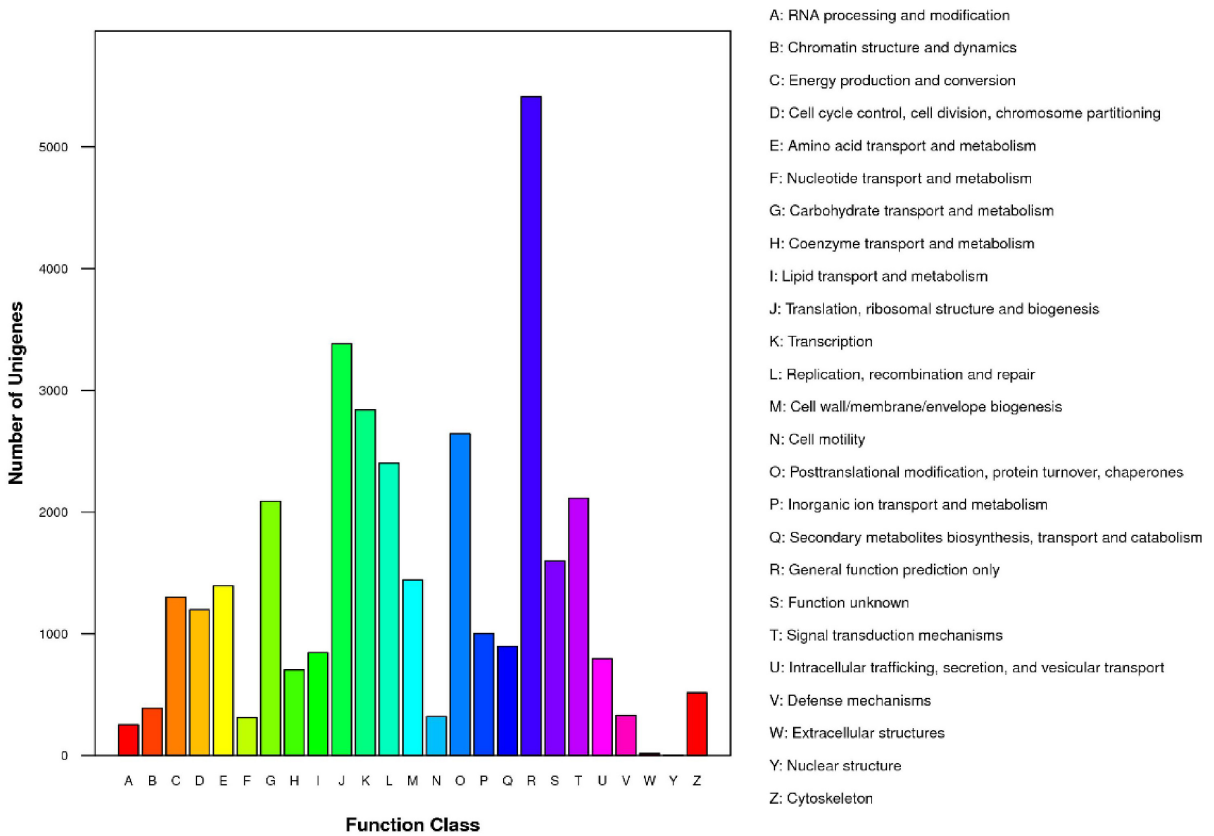


FIGURE 3 | Cluster of Orthologous Groups (COG) function categories of unigenes in *P. tomentosa*. COG function classes are shown on the *x*-axis and the full names of the functions are annotated on the right. The numbers of unigenes in one class are shown on the *y*-axis.

(iAMP), the precursor of cytokinin (CTK). Meanwhile, IPP and DMAPP can be transformed to geranylgeranyl diphosphate (GGDP), further generating GA and other carotenoids. In addition, GGDP can be reduced to phytyl diphosphate by putative geranyl-geranyl reductase (1.12-fold), thus providing phytol (0.507) for chlorophyll synthesis (**Table 4**). The concentrations of GA and ABA in this pathway showed small reductions, whereas another route of phytol utilization was increased under conditions of N deficiency.

Detection of Dynamic Expression of Differentially Expressed Genes by Real-Time Quantitative Reverse Transcription PCR

To validate the expression profiles of the identified DEGs and determine their possible dynamic responses to low-N stress at different treatment stages, the levels of expression of 14 DEGs were investigated by qRT-PCR after 0, 1, 3, and 5 days of low-N treatment (**Figure 6**). Except for *Unigene24651*, *Unigene3952*, *CL5328.Contig5*, and *CL4139.Contig1*, the expression changes of most (10) DEGs at 3 days under low-N treatment conditions quantified by qRT-PCR were consistent with the abundance determined by RNA-Seq analysis. The inconsistencies in the expression of the four genes quantified by these two methods may have been due to insufficient coverage sequencing depth to reflect the true distribution of these genes or to differences in the data normalization criteria for the two methods. Further expression correlation analyses of these ten DEGs demonstrated that the expression determined by RNA-Seq was positively correlated with that revealed by qRT-PCR (**Supplementary Figure S5**).

After further examining the expression changes by qRT-PCR, we found three different expression patterns (**Figure 6**). First, the expression abundance increased gradually after 1 and 3 days of low-N treatment, peaked at 3 days, and then declined steadily at 5 days for *Unigene 22453*, *Unigene19959*, *Unigene25149*, and *Unigene24651*. It is worth noting that *Unigene24651* was markedly repressed at 5 days relative to 0 days, in contrast to *Unigene 22453*, *Unigene19959*, and *Unigene25149*. Second, the expression levels of *Unigene3952*, *CL5328.Contig5*, *Unigene24078*, *Unigene3286*, *CL933.Contig 1*, and *CL9290.Contig2* peaked after 1 day of low-N treatment and decreased after 3 days. The expression levels of *Unigene3952*, *CL5328.Contig5,* and *Unigene24078* were increased, while those of *CL933.Contig 1* and *CL9290.Contig2* were reduced, after 5 days. Both of these patterns showed the dynamic fluctuation

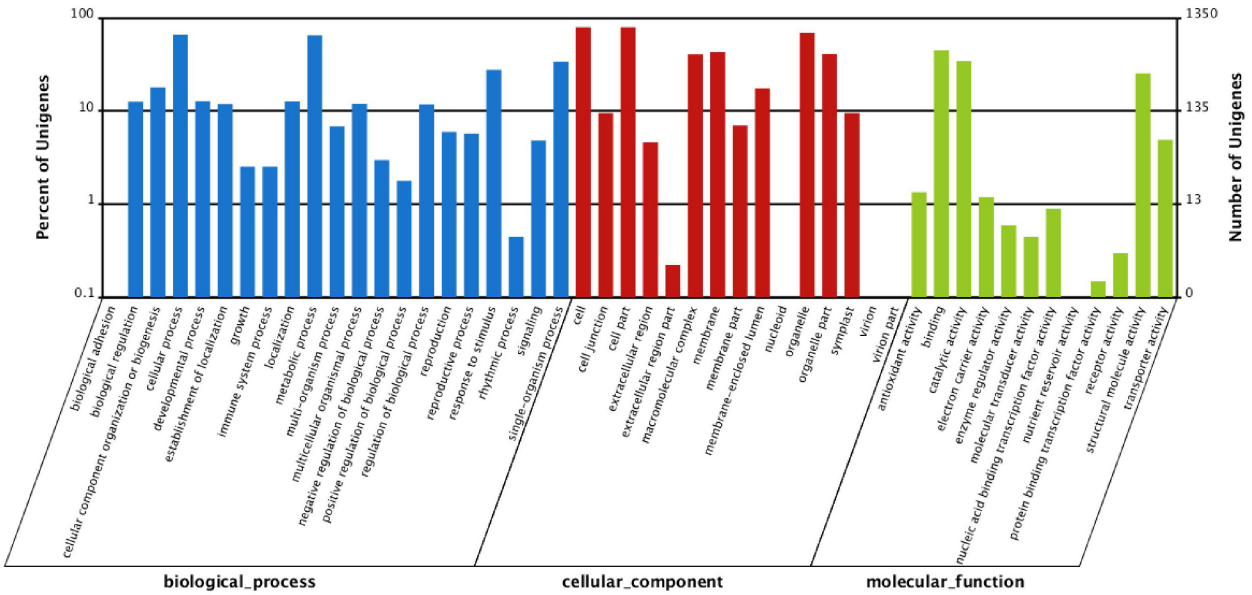


FIGURE 4 | The Gene Ontology (GO) classification of differentially expressed genes (DEGs; DN/KK) in *P. tomentosa* under conditions of N deficiency. The percentage of unigenes in one class is shown on the *y*-axis, while DEGs under DN treatment were classified into three classes (biological processes, cellular component, molecular function) on the *x*-axis.

TABLE 2 | Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway annotation of *Populus tomentosa*.

Pathway	DEGs (1329)	All genes (29,222)
Ribosome	471 (35.44%)	1661 (5.68%)
Phagosome	44 (3.31%)	450 (1.54%)
Endocytosis	64 (4.82%)	896 (3.07%)
Nitrogen metabolism	20 (1.5%)	190 (0.65%)
Arginine and proline metabolism	21 (1.58%)	252 (0.86%)
Biosynthesis of unsaturated fatty acids	11 (0.83%)	103 (0.35%)
Beta-alanine metabolism	13 (0.98%)	141 (0.48%)
Histidine metabolism	9 (0.68%)	87 (0.3%)
Amino sugar and nucleotide sugar metabolism	25 (1.88%)	354 (1.21%)
Inositol phosphate metabolism	17 (1.28%)	219 (0.75%)
Tryptophan metabolism	12 (0.9%)	139 (0.48%)
Pentose and glucuronate interconversions	23 (1.73%)	345 (1.18%)
Steroid biosynthesis	9 (0.68%)	104 (0.36%)
Lysine degradation	10 (0.75%)	121 (0.41%)
Propanoate metabolism	15 (1.13%)	207 (0.71%)
Phosphatidylinositol signaling system	15 (1.13%)	207 (0.71%)

of miRNA expression in *P. tomentosa* in response to low-N stress at different stages. Third, *Unigene18608*, *CL6853.Contig2*, and *CL4139.Contig1* were all induced, while *Unigene13209* was repressed, at all stages of low-N stress.

Correlation of Gene Expression and Metabolic Changes Under Low-N Stress

Based on the global metabolome, we found an array of primary metabolites with differential changes under DN treatment, meanwhile, the transcriptome data showed accompanied changes of genes, which were involved in biosynthesis or degradation of these key metabolites. Thereafter we compared the expression of inducement or suppression of genes with these corresponding metabolites. Coordinated metabolomics and transcriptomics analyses revealed the relative expression of five pairs of metabolites and unigenes (**Table 5**). Our results revealed five genes encoding enzymes involved in the biogenesis or degradation of associated carbohydrates, and genes and metabolites showed a co-expression pattern. The data indicated that the accumulation of D-glucose and galactonic acid could be attributed to their increased biosynthesis, catalyzed by dTDP-D-glucose 4,6-dehydratase and β-galactosidase, respectively, while the accumulation of D-fructose and D-galactose may have been the result of a reduction of fructose-2,6-bisphosphatase and galactose mutarotase expression.

DISCUSSION

Due to their ability to perform high-throughput genome-wide characterization, transcriptomics analyses have been extensively applied to systemically investigate the genes responsible for the development and adaptation of plants to different environmental stresses (Vogel et al., 2016; Wang Y. N. et al., 2016; Woo et al., 2016). It is well known that N, one of the most important nutrients, functions as a component of various key cell molecules, such as proteins (amino acids), nucleic acids, chlorophyll, and secondary metabolites. Consequently, many studies have focused on the mechanisms underlying the roles of N in plants, and DEGs in response to N deficiency. Several genes involved in N assimilation and utilization have been identified in a variety

TABLE 3 | Differentially expressed genes (DEGs) identified by GO and KEGG involved in N metabolism in *Populus tomentosa*.

Gene ID	log_2Ratio (DN/KK)	Regulation (DN/KK)	Annotation
CL8615.Contig2_All	−1.61	Down	Aminomethyltransferase, mitochondrial
Unigene23815_All	−1.34	Down	Asparagine synthase
CL6264.Contig2_All	−1.07	Down	Carbonic anhydrase, chloroplastic
Unigene30284_All	2.15	Up	Cytochrome b5
Unigene29769_All	2.77	Up	Cytochrome b-c1 complex subunit Rieske-3, mitochondrial
Unigene35581_All	−11.65	Down	Cytochrome c1–1, heme protein, mitochondrial
Unigene31830_All	12.10	Up	Cytochrome c1–2, heme protein, mitochondrial
Unigene24651_All	−1.05	Down	Ferredoxin–nitrite reductase, chloroplastic
CL4868.Contig1_All	1.01	Up	Glutamate dehydrogenase 1
Unigene30445_All	3.65	Up	Glutamate dehydrogenase 3
Unigene31623_All	3.92	Up	Glutamate dehydrogenase 3
Unigene22953_All	−1.60	Down	Glutamine synthetase, chloroplastic
Unigene27220_All	−1.37	Down	Glutamine synthetase, chloroplastic
Unigene38713_All	−12.18	Down	NADP-specific glutamate dehydrogenase
Unigene18435_All	−3.30	Down	Nitrate reductase
CL4201.Contig1_All	3.05	Up	Nitrate reductase (NADH)
CL933.Contig1_All	−2.55	Down	Nitrate reductase (NADH)
CL933.Contig2_All	−2.14	Down	Nitrate reductase (NADH)
CL933.Contig3_All	−2.22	Down	Nitrate reductase (NADH)
Unigene34301_All	−11.88	Down	Nitrate reductase (NADH)
Unigene38634_All	−12.26	Down	Nitrate reductase (NADH)

DN, low N treated; KK, control. NADP, nicotinamide adenine dinucleotide phosphate; NADH, nicotinamide adenine dinucleotide.

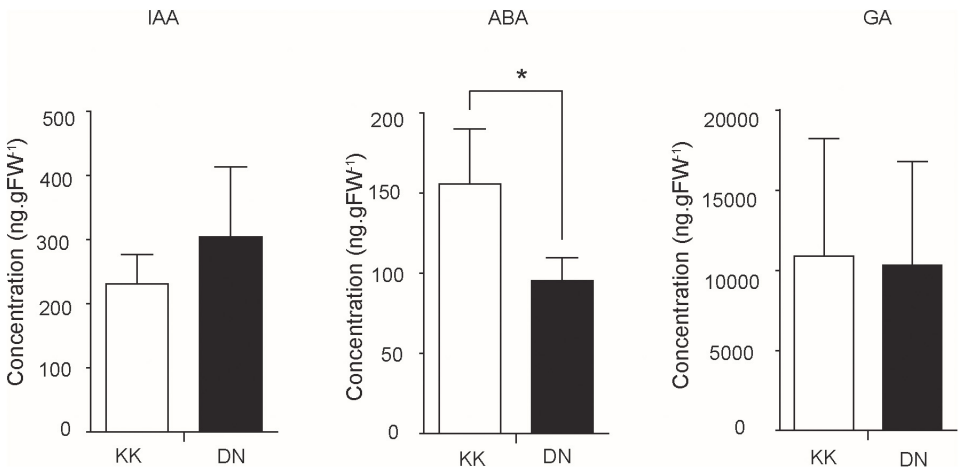


FIGURE 5 | Quantification of the concentrations of gibberellic acid (GA), abscisic acid (ABA), and indole-3-acetic acid (IAA) in *Populus tomentosa* plants between control (KK) and low-N treatment (DN) groups. The *y*-axis represents the hormone concentration, while the *x*-axis represents the KK and DN samples. *denotes significant difference ($p \leq 0.05$).

of plants, including cucumber (Zhao et al., 2015), rice (Yang et al., 2015), maize (Chen Q. et al., 2015), *Medicago truncatula* (Bonneau et al., 2013), and *Arabidopsis* (Günther et al., 2012).

Global Changes in the Transcriptomic Network of *Populus tomentosa* Under N Deficiency

The results of the present study showed that N deficiency leads to the differential expression of many genes, depresses N assimilation, reduces amino acid biosynthesis, and reduces photosynthetic and energy production capacity. Similar to previous studies, a large number of DEGs were identified in *P. tomentosa* in response to low-N conditions, including several genes directly related to N metabolism, encoding putative nitrate reductase (*NR*), glutamate dehydrogenase (*GDH*), glutamine synthetase (*GS*), among others (**Supplementary Table S10**). Generally, plants can directly use ammonium, but not nitrate (Jackson et al., 2008). Nitrate first must be reduced by NR (Bowsher et al., 1988), which is then reduced to ammonium

TABLE 4 | Hormone-related DEGs in *Populus tomentosa* under N deficiency.

Gene ID	Log_2Ratio (DN/KK)	Regulation (DN/KK)	Annotation
Unigene21573_All	1.50	Up	Auxin response factor 6
CL5591.Contig2_All	2.23	Up	Auxin-responsive protein IAA10
Unigene40069_All	−11.57	Down	IAA-amino acid hydrolase ILR1-like 1
Unigene9971_All	−1.43	Down	IAA-amino acid hydrolase 11 (ILL11)
Unigene36745_All	−1.77	Down	UDP-glycosyltransferase 76C3
Unigene28931_All	12.09	Up	Isopentenyl diphosphate isomerase
Unigene39732_All	−11.54	Down	Similar to geranylgeranyl hydrogenase
Unigene5624_All	1.12	Up	Geranylgeranyl reductase
Unigene40194_All	−12.20	Down	RAB proteins geranylgeranyl transferase component A
Unigene39275_All	−11.71	Down	RAB proteins geranylgeranyl transferase component A

RAB proteins, Ras -Associated binding protein.

by the concerted action of NR (Foyer and Mullineaux, 1998). With the action of two enzymes, GOGAT and GS, ammonium can form the amino acids that combine with the organic acids produced by photosynthesis.

This study showed that five transcripts of putative NR and two genes encoding putative GS were downregulated under conditions of low-N stress because of N source limitations (**Supplementary Table S10**). However, the expression levels of *NR* and *GS* in *Arabidopsis* and rice were elevated under low-N conditions, possibly indicating differences in the mechanisms regulating N uptake and accumulation among *P. tomentosa*, *Arabidopsis*, and rice. A previous investigation revealed direct coupling of N assimilation and photosynthesis in chloroplasts, in a mechanism designated as nitrate photo assimilation (Bot et al., 2009). Several studies have revealed universal physiological and metabolic changes involved in the responses of plants to N deficiency, including significant reductions of chlorophyll content, growth, and photosynthesis, protein, starch, etc. (Mesnard and Ratcliffe, 2005). We detected four decreased transcripts of putative chlorophyll a–b binding proteins, indicating downregulated photosynthesis under low-N conditions. Similarly, previous investigations showed that genes involved in photosynthesis were downregulated in the leaves of *Arabidopsis* (Liu et al., 2017) and roots of rice (Cai et al., 2012).

Responses of Hormone and Hormone-Related Genes to N Deficiency

Among plant hormones, auxin and ABA have been reported to be closely related to N signaling (Sakakibara, 2003; Lu et al., 2015; Yu et al., 2016). Our results demonstrated the interaction of GA signaling with N deprivation. Transcripts of putative GA-insensitive (*GAI*, a DELLA domain-containing GRAS family transcription factor) were induced (1.57-fold upregulation) and putative repressor of GA1-3-like 1 (*RGL1*) was suppressed (1.78-fold downregulation) under conditions of low-N stress. In cucumbers, suboptimal root zone temperatures were shown to suppress GA biogenesis and plant growth, while exogenous GA restored seedling biomass and enhanced N uptake (Bai et al., 2016). These results show the interdependent promotional role between GA biogenesis and N acquisition.

In *Arabidopsis*, N deprivation induced an increase in auxin content and expression of the auxin synthetic gene, *TAR2*, and the expression of auxin influx carriers (AUX/LAX family, such as *AUX1*, *LAX1*, *LAX2*, *LAX3*) and efflux transporter factors (PIN and PGP, e.g., *PIN1*, *PIN2*, and *PIN4*) could be regulated by the nitrogen/carbon (N/C) ratio (Gutiérrez et al., 2007). In addition, auxin signaling is related to N availability, and *ARF8* could be induced by nitrate deficiency. miR160 has been shown to suppress *ARF10*, *ARF16*, and *ARF17*, while miR167 targets *ARF6* and *ARF8* (Rhoades et al., 2002). In maize, low N supply increased the levels of root auxin and NO, and then enhanced root elongation (Mi et al., 2008). Auxin was shown to trigger the accumulation of miR393, leading to the repression of *TIR1*/*AFBs*, depressed auxin perception and response, and ultimately, homeostasis (Chen et al., 2011). In addition, by binding to auxin and AUX/IAA, TIR1 could direct the ubiquitination and degradation of AUX/IAA proteins. On the other hand, binding to AUX/IAA promotes the activation of ARFs and the depression of other early genes in the auxin response pathway (Mockaitis and Estelle, 2008). Studies of *Arabidopsis* have revealed that the coordination of miRNA nodes directs the nitrate and auxin signaling involved in lateral root formation. Both miR393 and its target, *AFB3*, are known to be induced by nitrate and glutamine/glutamate, which affect the uptake of auxin (Vidal et al., 2010). However, miR167 is depressed under conditions of nitrate treatment, thus causing the accumulation of *ARF8* and the downstream gene product, *GH3*, leading to a homeostatic level of auxin and modulation of the root architecture (Yang et al., 2006). Previous studies have validated the interaction of pto-miR160a and its target mRNA, *pto-ARF16*, which are involved in tree growth (Tian et al., 2016). Our previous investigation demonstrated the downregulation of pto-miR160f, pto-miR167a/b, pto-miR167c–f, pto-miR167g, and pto-miR390a–d and the upregulation of pto-miR393a/b under conditions of low-N stress in *P. tomentosa*, and thus the association of nitrate with auxin signaling (Ren et al., 2015). Furthermore, our transcriptomics data showed the corresponding induction of putative *pto-IAA10* (*CL5591.Contig2_All*, 2.2-fold upregulation) and *pto-ARF6* (*Unigene21573_All*, 1.5-fold upregulation), indicating that the nitrate-induced auxin signaling pathway plays a role in root architecture plasticity.

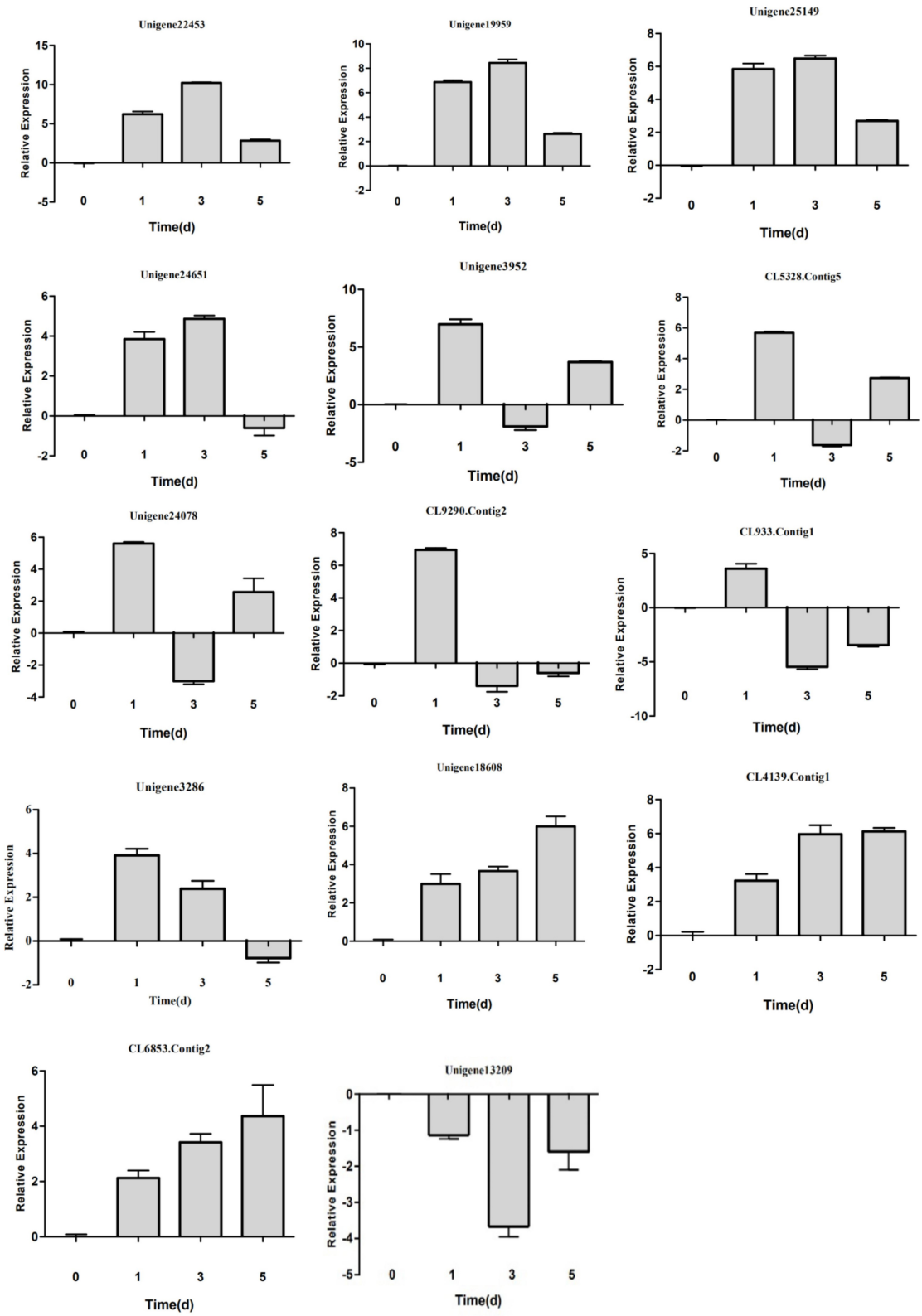


FIGURE 6 | Detection of DEG expression in *Populus tomentosa* under control (KK) and low-N treatment (DN) conditions by real-time quantitative reverse transcription-polymerase chain reaction (qRT-PCR). The *x*-axis represents the treatment time (0, 1, 3, and 5 days), while the *y*-axis represents relative gene expression level.

TABLE 5 | Coordinated changes in gene expression that led to alternation of some key primary metabolites in *Populus tomentosa* under N deficiency.

Metabolite	Log_2 (DN/KK)	Gene annotation	Gene ID	Function	GO molecular function
D-glucose	0.23	dTDP-D-glucose 4,6-dehydratase	Unigene29785_All (4.38), Unigene33572_All (2.54)	Biosynthesis	Oxidoreductase activity
Galactonic acid	0.80	Beta-galactosidase	CL1096.Contig2_All (0.95), CL1096.Contig3_All (0.95), CL1096.Contig4_All (0.95), CL1096.Contig5_All (1.19), CL1096.Contig6_All (0.95), CL1096.Contig7_All (0.12), CL1096.Contig8_All (0.69), CL1096.Contig9_All (0.69),	Biosynthesis	Beta-galactosidase activity
D-fructose	0.43	Fructose-2,6-bisphosphatase	CL4139.Contig1_All (−1.15), CL5072.Contig1_All (−0.17), CL5072.Contig2_All (0.03), CL5072.Contig3_All (−0.11),	Degradation	Phosphoglycerate mutase activity
Inositol	−0.36	Inositol oxygenase 2	Unigene29428_All (1.07), CL1004.Contig2_All (0.01),	Degradation	Inositol oxygenase activity
D-galactose	0.17	Galactose mutarotase and related enzymes	CL4080.Contig1_All (−1.34), CL1004.Contig3_All (−0.63), CL1004.Contig4_All (−1.19) Unigene34048_All (0.10), Unigene34101_All (−0.48), Unigene38082_All (−0.55)	Degradation	Aldose 1-epimerase activity

Previous studies provided evidence that the biosynthesis of CTK is dependent on N status to induce IPT genes, and that nitrate and glutamine act as inducers in *Arabidopsis* (Takei et al., 2004) and rice (Kamada-Nobusada et al., 2013). Similar to ABA, reduced CTK in a low-N environment was reported to enhance lateral root formation (Kiba et al., 2010). In addition, cytokinin-*N*-glucosyltransferase was shown to be involved in CTK homeostasis in *Arabidopsis*. Our study detected the downregulation of putative cytokinin-*N*-glucosyltransferase (1.77-fold downregulation), consistent with the results seen in *Arabidopsis* under conditions of N starvation (Bi et al., 2007) and rice (Cai et al., 2012). By catalyzing *N*-glucosylation of CTKs, CTKs then regulated CTK-response genes, such as *CKX3*, *AHK2*, *AHK3*, *ARR1*, and *LOG2* (Wang et al., 2013), thus controlling the response to N deficiency in *P. tomentosa*. In addition, ethylene response regulator was increased in *Botrytis cinerea*-infected *Solanum lycopersicum* under nitrate-limiting conditions, suggesting that N status may be related to pathogen susceptibility (Vega et al., 2015). As a consequence, diverse plant hormones should be involved in the adaptation of *P. tomentosa* to conditions of low-N stress.

Global Changes of Metabolic Networks in Response to N Deficiency

To investigate the effects of low-N stress on the metabolic regulation of *P. tomentosa* and validate gene expression at the level of the transcriptome, GC-MS was further applied for identification and robust quantification of primary metabolites in plant samples, including sugars, sugar alcohols, amino acids, organic acids, and polyamines. These results should enable a comprehensive evaluation of metabolic changes in *P. tomentosa* exposed to low-N stress, and highlight fields of interest for future studies of the NUE of poplar. The contents of amino acids were changed markedly under conditions of low N supply, including valine (1.32-fold downregulation), L-alanine (0.69-fold downregulation), and L-isoleucine (0.76-fold downregulation), similar to the gene expression pattern of the transcriptome. These changes were somewhat different from those seen in previous studies related to N starvation. In tomatoes, the L-alanine level was reported to decrease to less than half that recorded at the start of the experiment, while L-isoleucine and valine did not change markedly (Urbanczyk-Wochniak and Fernie, 2005). Tschoep et al. (2009) found that a low-N growth regime caused shoot growth reduction and increased the levels of many amino acids. The levels of amino acids, including alanine, serine, glycine, valine, leucine, isoleucine, phenylalanine, and tyrosine, derived from pyruvate and phosphoenolpyruvate increased in *Synechocystis* under conditions of N deficiency (Tschoep et al., 2009). This may have been related to differences between the species or experimental conditions. Urbanczyk-Wochniak and Fernie (2005) reported that different environmental factors could cause different changes in amino acids, although the variation of amino acids was not linear under low-N conditions, with some amino acids showing a marginal decline initially but an increased level at the end of the experiment. The results of this investigation were also somewhat consistent with previous studies of the poplar transcriptome, which indicated that poplar species could slow down N assimilation under conditions of limited N supply (Luo et al., 2013).

In this study, the downregulation of genes participating in the glyoxylate cycle [isocitrate lyase, malate synthase,

and phosphoenolpyruvate carboxykinase (ATP: oxaloacetate carboxy-lyase (transphosphorylation), EC 4.1.1.49] may have had adverse effects on the growth of *P. tomentosa*, indicating a depressed state of C metabolism providing less carbon (by synthesizing a variety of carbohydrates) and energy (providing NADPH and then ATP) for plant growth and development. On the other hand, N and C metabolism could be mutually affected due to the requirement for C skeletons for the production of N metabolites, such as amino acids. In dicots, nitrate levels could affect carbohydrate metabolism (Scheible et al., 1997). Phloem transport and C export are suppressed under low-N conditions (Nunes-Nesi et al., 2010). The significantly altered amino acid production seen in this study was related to pyruvate recycling. The accumulation of foliar starch was also found in *Arabidopsis* and maize exposed to low-N conditions due to the reduced demand of C skeletons for N compounds, including amino acids and proteins (Ikram et al., 2012; Schlüter et al., 2012). Concerning C metabolism, the expression of some sugars, i.e., D-fructose (.433-fold), D-galactose (.171-fold), and D-glucose (.229-fold), was induced under conditions of N deficiency in this investigation. Similarly, it was reported that glucose and fructose accumulated under low-N conditions in *Synechocystis* (Asayama et al., 2004). In contrast, fructose and galactose did not vary markedly under high-light conditions in tomatoes, whereas they accumulated under low-light conditions (Urbanczyk-Wochniak and Fernie, 2005). Glucose and fructose are signaling molecules and energy sources in plants under conditions of abiotic stress (Bonneau et al., 2013). Although foliar starch, glucose, and fructose accumulate under low-N conditions, N deficiency inhibits photosynthetic capacity and growth. Poplar roots actively forage for nutrients under low-N conditions, while the acquired N appears insufficient for the biosynthesis of photosynthetic enzymes and metabolic precursors, leading to decreased photosynthetic capacity (Sardans and Peñuelas, 2012). In this study, the level of phytol related to chlorophyll synthesis increased significantly, which would directly affect the chlorophyll content in photosynthesis. Moreover, there was a marked accumulation of oxalic acid, which is involved in in the photosynthetic C3 cycle and photorespiration. Accumulation of phosphoric acid was also observed; this is involved in cyclic photophosphorylation. Interestingly, in contrast to most previous studies, four amines (putrescine, ethanolamine, ethylamine, and cadaverine) showed significant changes in this study. As components of N assimilation, the synthesis of ethylamine, ethanolamine, and putrescine decreased. Consistently, previous studies have shown that polyamines play an important role in modulating the sensitivity/tolerance to N stress (Syed et al., 2011). In addition to carbohydrate metabolic processes, N deficiency also affects the metabolism of other nutrients, such as phosphate and sulfur transport, indicating alteration of the N:C/P/S ratios and potential interactions (**Supplementary Table S7**). Specifically, several sugar-related pathways are known to interconnect with N-responsive pathways, and C/N balance has particular significance for plant growth. Obviously, crosstalk between N and C/P/S metabolism is involved in nutrient utilization and signal transduction, and a limitation of one element could cause imbalances in the nutrient network and affect the uptake, accumulation, and utilization of other nutrients (Schachtman and Shin, 2007; Rennenberg et al., 2010).

CONCLUSION

In summary, this study revealed the global changes of transcripts and metabolites occurring in response to N deprivation in *P. tomentosa*. Specifically, 2,662 DEGs, 30 significantly changed metabolites, and three altered hormones were detected in our study. Combining transcriptomic and metabolic profiles, our investigation revealed a general depression of molecular and physiological metabolism under low N stress, including N absorption and assimilation, photosynthesis, glycolysis, and the TCA cycle. This extends our understanding of the relationships between responsive genes and downstream metabolic responses and provides a basis for the integrated and comprehensive analysis of molecular responses to N deficiency, which may improve understanding of the molecular and metabolic mechanisms underlying the phenotype of *P. tomentosa*, and the interaction of C and N under conditions of N deficiency.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. This data can be found here: National Center for Biotechnology Information (NCBI) BioProject database under accession number PRJNA296440.

AUTHOR CONTRIBUTIONS

YW designed the experiment and edited the manuscript. MC and YY performed the research and wrote the manuscript. XY, LZ, TF, and XH contributed to the analytical tools and reagents. All the authors read and approved the final manuscript.

FUNDING

This work was supported by the National Natural Science Foundation of China (Nos. 32071504, 31670671).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.769748/full#supplementary-material

REFERENCES

Albinsky, D., Kusano, M., Higuchi, M., Hayashi, N., Kobayashi, M., Fukushima, A., et al. (2010). Metabolomic screening applied to rice FOX *Arabidopsis* lines leads to the identification of a gene-changing nitrogen metabolism. *Mol. Plant* 3, 125–142. doi: 10.1093/mp/ssp069

Asayama, M., Imamura, S., Yoshihara, S., Miyazaki, A., Yoshida, N., Sazuka, T., et al. (2004). SigC, the Group 2 sigma factor of RNA polymerase, contributes to the late-stage gene expression and nitrogen promoter recognition in the *Cyanobacterium Synechocystis* sp. Strain PCC 6803. *Biosci. Biotech. Biochem.* 68, 477–487. doi: 10.1271/bbb.68.477

Bai, L., Deng, H., Zhang, X., Yu, X., and Li, Y. (2016). Gibberellin is involved in inhibition of cucumber growth and nitrogen uptake at suboptimal root-zone temperatures. *PLoS One* 11:e0156188. doi: 10.1371/journal.pone.0156188

Bi, Y.-M., Wang, R.-L., Zhu, T., and Rothstein, S. J. (2007). Global transcription profiling reveals differential responses to chronic nitrogen stress and putative nitrogen regulatory components in *Arabidopsis*. *BMC Genomics* 8:281. doi: 10.1186/1471-2164-8-281

Bonneau, L., Huguet, S., Wipf, D., Pauly, N., and Truong, H. N. (2013). Combined phosphate and nitrogen limitation generates a nutrient stress transcriptome favorable for arbuscular mycorrhizal symbiosis in *Medicago truncatula*. *New Phytol.* 199, 188–202. doi: 10.1111/nph.12234

Bot, J. L., Bénard, C., Robin, C., Bourgaud, F., and Adamowicz, S. (2009). The 'trade-off' between synthesis of primary and secondary compounds in young tomato leaves is altered by nitrate nutrition: experimental evidence and model consistency. *J. Exp. Bot.* 60, 4301–4314. doi: 10.1093/jxb/erp271

Bowsher, C. G., Emes, M. J., Cammack, R., and Hucklesby, D. P. (1988). Purification and properties of nitrite reductase from roots of pea (*Pisum sativum* cv. Meteor). *Planta* 175, 334–340. doi: 10.1007/BF00396338

Broyart, C., Fontaine, J. X., Molinié, R., Cailleu, D., Tercé-Laforgue, T., Dubois, F., et al. (2010). Metabolic profiling of maize mutants deficient for two glutamine synthetase isoenzymes using 1H-NMR-based metabolomics. *Phytochem. Anal.* 21, 102–109. doi: 10.1002/pca.1177

Cai, H., Lu, Y., Xie, W., Zhu, T., and Lian, X. (2012). Transcriptome response to nitrogen starvation in rice. *J. Biosci.* 37, 731–747. doi: 10.1007/s12038-012-9242-2

Cao, Y. W., Qu, R. J., Miao, Y. J., Tang, X. Q., Zhou, Y., Wang, L., et al. (2019). Untargeted liquid chromatography coupled with mass spectrometry reveals metabolic changes in nitrogen-deficient Isatis indigotica Fortune. *Phytochemistry* 166, 112058. doi: 10.1016/j.phytochem.2019.112058

Chen, M., Bao, H., Wu, Q., and Wang, Y. (2015). Transcriptome-wide identification of mirna targets under nitrogen deficiency in Populus tomentosa using degradome sequencing. *Int. J. Mol. Sci.* 16, 13937–13958. doi: 10.3390/ijms160613937

Chen, M., Wang, C., Bao, H., Chen, H., and Wang, Y. (2016). Genome-wide identification and characterization of novel lncRNAs in Populus under nitrogen deficiency. *Mol. Genet. Genomics* 291, 1663–1680. doi: 10.1007/s00438-016-1210-3

Chen, Q., Liu, Z., Wang, B., Wang, X., Lai, J., and Tian, F. (2015). Transcriptome sequencing reveals the roles of transcription factors in modulating genotype by nitrogen interaction in maize. *Plant Cell Rep.* 34, 1761–1771. doi: 10.1007/s00299-015-1822-9

Chen, X., Zhu, W., Azam, S., Li, H., Zhu, F., Li, H., et al. (2013). Deep sequencing analysis of the transcriptomes of peanut aerial and subterranean young pods identifies candidate genes related to early embryo abortion. *Plant Biotechnol. J.* 11, 115–127. doi: 10.1111/pbi.12018

Chen, Z.-H., Bao, M.-L., Sun, Y.-Z., Yang, Y.-J., Xu, X.-H., Wang, J.-H., et al. (2011). Regulation of auxin response by miR393-targeted transport inhibitor response protein1 is involved in normal development in *Arabidopsis*. *Plant Mol. Biol.* 77, 619–629. doi: 10.1007/s11103-011-9838-1

Dash, M., Yordanov, Y. S., Georgieva, T., Kumari, S., Wei, H., and Busov, V. (2015). A systems biology approach identifies new regulators of poplar root development under low nitrogen. *Plant J.* 84, 335–346. doi: 10.1111/tpj.13002

de Bang, T. C., Husted, S., Laursen, K. H., Persson, D. P., and Schjoerring, J. K. (2021). The molecular-physiological functions of mineral macronutrients and their consequences for deficiency symptoms in plants. *New Phytol.* 229, 2446–2469. doi: 10.1111/nph.17074

Du, Q., Wang, B., Wei, Z., Zhang, D., and Li, B. (2012). Genetic diversity and population structure of Chinese White poplar (*Populus tomentosa*) revealed by SSR markers. *J. Hered.* 103, 853–862. doi: 10.1093/jhered/ess061

Fang, L., Sun, D., Xu, Z., He, J., Qi, S., Chen, X., et al. (2015). Transcriptomic analysis of a moderately growing subisolate *Botryococcus braunii* 779 (Chlorophyta) in response to nitrogen deprivation. *Biotechnol. Biofuels* 8:130. doi: 10.1186/s13068-015-0307-y

Foyer, C. H., and Mullineaux, P. M. (1998). The presence of dehydroascorbate and dehydroascorbate reductase in plant tissues. *FEBS Lett.* 425, 528–529.

Goel, P., Bhuria, M., Kaushal, M., and Singh, A. K. (2016). Carbon: nitrogen interaction regulates expression of genes involved in N-uptake and assimilation in *Brassica juncea* L. *PLoS One* 11:e0163061. doi: 10.1371/journal.pone.0163061

Günther, T., Lampei, C., and Schmid, K. J. (2012). Mutational bias and gene conversion affect the intraspecific nitrogen stoichiometry of the *Arabidopsis thaliana* transcriptome. *Mol. Biol. Evol.* 30, 561–568. doi: 10.1093/molbev/mss249

Gutiérrez, R. A., Lejay, L. V., Dean, A., Chiaromonte, F., Shasha, D. E., and Coruzzi, G. M. (2007). Qualitative network models and genome-wide expression data define carbon/nitrogen-responsive molecular machines in *Arabidopsis*. *Genome Biol.* 8, R7. doi: 10.1186/gb-2007-8-1-r7

Gutieìrrez, R. A., Shasha, D. E., and Coruzzi, G. M. (2005). Systems biology for the virtual plant. *Plant Physiol.* 138, 550–554. doi: 10.1104/pp.104.900150

Hou, J., Wu, Q., Zuo, T., Guo, L., Chang, J., Chen, J., et al. (2016). Genome-wide transcriptomic profiles reveal multiple regulatory responses of poplar to *Lonsdalea quercina* infection. *Trees* 30, 1389–1402.

Ikram, S., Bedu, M., Daniel-Vedele, F., Chaillou, S., and Chardon, F. (2012). Natural variation of *Arabidopsis* response to nitrogen availability. *J. Exp. Bot.* 63, 91–105.

Jackson, L. E., Burger, M., and Cavagnaro, T. R. (2008). Roots, nitrogen transformations, and ecosystem services. *Annu. Rev. Plant Biol.* 59, 341–363. doi: 10.1146/annurev.arplant.59.032607.092932

Jones, D. L., Healey, J. R., Willett, V. B., Farrar, J. F., and Hodge, A. (2005). Dissolved organic nitrogen uptake by plants—an important N uptake pathway? *Soil Biol. Biochem.* 37, 413–423. doi: 10.1016/j.soilbio.2004.08.008

Kamada-Nobusada, T., Makita, N., Kojima, M., and Sakakibara, H. (2013). Nitrogen-dependent regulation of de novo cytokinin biosynthesis in rice: the role of glutamine metabolism as an additional signal. *Plant Cell Physiol.* 54, 1881–1893. doi: 10.1093/pcp/pct127

Karp, A., and Shield, I. (2008). Bioenergy from plants and the sustainable yield challenge. *New Phytol.* 179, 15–32. doi: 10.1111/j.1469-8137.2008.02432.x

Kiba, T., Kudo, T., Kojima, M., and Sakakibara, H. (2010). Hormonal control of nitrogen acquisition: roles of auxin, abscisic acid, and cytokinin. *J. Exp. Bot.* 62, 1399–1409. doi: 10.1093/jxb/erq410

Kusano, M., Fukushima, A., Redestig, H., and Saito, K. (2011). Metabolomic approaches toward understanding nitrogen metabolism in plants. *J. Exp. Bot.* 62, 1439–1453.

Liu, Q., Chen, X., Wu, K., and Fu, X. (2015). Nitrogen signaling and use efficiency in plants: what's new? *Curr. Opin. Plant Biol.* 27, 192–198. doi: 10.1016/j.pbi.2015.08.002

Liu, W., Sun, Q., Wang, K., Du, Q., and Li, W.-X. (2017). Nitrogen limitation adaptation (NLA) is involved in source-to-sink remobilization of nitrate by mediating the degradation of NRT1.7 in *Arabidopsis*. *New Phytol.* 214, 734–744. doi: 10.1111/nph.14396

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the 2(-Delta Delta C(T)) Method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Lu, Y., Sasaki, Y., Li, X., Mori, I. C., Matsuura, T., Hirayama, T., et al. (2015). ABI1 regulates carbon/nitrogen-nutrient signal transduction independent of ABA biosynthesis and canonical ABA signalling pathways in *Arabidopsis*. *J. Exp. Bot.* 66, 2763–2771. doi: 10.1093/jxb/erv086

Luo, J., Li, H., Liu, T., Polle, A., Peng, C., and Luo, Z.-B. (2013). Nitrogen metabolism of two contrasting poplar species during acclimation to limiting nitrogen availability. *J. Exp. Bot.* 64, 4207–4224. doi: 10.1093/jxb/ert234

Luo, J., Zhou, J., Li, H., Shi, W., Polle, A., Lu, M., et al. (2015). Global poplar root and leaf transcriptomes reveal links between growth and stress responses under nitrogen starvation and excess. *Tree Physiol.* 35, 1283–1302. doi: 10.1093/treephys/tpv091

Mesnard, F., and Ratcliffe, R. G. (2005). NMR analysis of plant nitrogen metabolism. *Photosynth. Res.* 83, 163–180. doi: 10.1007/s11120-004-2081-8

Mi, G., Chen, F., and Zhang, F. (2008). Multiple signaling pathways controls nitrogen-mediated root elongation in maize. *Plant Signal. Behav.* 3, 1030–1032. doi: 10.4161/psb.6800

Mockaitis, K., and Estelle, M. (2008). Auxin receptors and plant development: a new signaling paradigm. *Annu. Rev. Cell Dev. Biol.* 24, 55–80. doi: 10.1146/annurev.cellbio.23.090506.123214

Mu, X., and Chen, Y. (2021). The physiological response of photosynthesis to nitrogen deficiency. *Plant Physiol. Biochem.* 158, 76–82. doi: 10.1016/j.plaphy.2020.11.019

Murashige, T., and Skoog, F. (1962). A revised medium for rapid growth and bio assays with tobacco tissue cultures. *Physiol. Plant.* 15, 473–497. doi: 10.1111/j.1399-3054.1962.tb08052.x

North, K. A., Ehlting, B., Koprivova, A., Rennenberg, H., and Kopriva, S. (2009). Natural variation in *Arabidopsis* adaptation to growth at low nitrogen conditions. *Plant Physiol. Biochem.* 47, 912–918. doi: 10.1016/j.plaphy.2009.06.009

Nunes-Nesi, A., Fernie, A. R., and Stitt, M. (2010). Metabolic and signaling aspects underpinning the regulation of plant carbon nitrogen interactions. *Mol. Plant* 3, 973–996. doi: 10.1093/mp/ssq049

Pan, C., Valente, J. J., LoBrutto, R., Pickett, J. S., and Motto, M. (2010). Combined application of high resolution and tandem mass spectrometers to characterize methionine oxidation in a parathyroid hormone formulation. *J. Pharm. Sci.* 99, 1169–1179. doi: 10.1002/jps.21901

Patterson, K., Cakmak, T., Cooper, A., Lager, I., Rasmusson, A. G., and Escobar, M. A. (2010). Distinct signalling pathways and transcriptome response signatures differentiate ammonium- and nitrate-supplied plants. *Plant Cell Environ.* 33, 1486–1501. doi: 10.1111/j.1365-3040.2010.02158.x

Rejeb, I. B., Pastor, V., and Mauch-Mani, B. (2014). Plant responses to simultaneous biotic and abiotic stress: molecular mechanisms. *Plants (Basel)* 3, 458–475. doi: 10.3390/plants3040458

Ren, Y., Sun, F., Hou, J., Chen, L., Zhang, Y., Kang, X., et al. (2015). Differential profiling analysis of miRNAs reveals a regulatory role in low N stress response of Populus. *Funct. Integr. Genomics* 15, 93–105. doi: 10.1007/s10142-014-0408-x

Rennenberg, H., Wildhagen, H., and Ehlting, B. (2010). Nitrogen nutrition of poplar trees. *Plant Biol. (Stuttg)* 12, 275–291. doi: 10.1111/j.1438-8677.2009.00309.x

Rhoades, M. W., Reinhart, B. J., Lim, L. P., Burge, C. B., Bartel, B., and Bartel, D. P. (2002). Prediction of plant microRNA targets. *Cell* 110, 513–520.

Sakakibara, H. (2003). Nitrate-specific and cytokinin-mediated nitrogen signaling pathways in plants. *J. Plant Res.* 116, 253–257. doi: 10.1007/s10265-003-0097-3

Sardans, J., and Peñuelas, J. (2012). The role of plants in the effects of global change on nutrient availability and stoichiometry in the plant-soil system. *Plant Physiol.* 160, 1741–1761. doi: 10.1104/pp.112.208785

Schachtman, D. P., and Shin, R. (2007). Nutrient sensing and signaling: NPKS. *Annu. Rev. Plant Biol.* 58, 47–69. doi: 10.1146/annurev.arplant.58.032806.103750

Scheible, W. R., Gonzalez-Fontes, A., Lauerer, M., Muller-Rober, B., Caboche, M., and Stitt, M. (1997). Nitrate acts as a signal to induce organic acid metabolism and repress starch metabolism in tobacco. *Plant Cell* 9, 783–798. doi: 10.1105/tpc.9.5.783

Schlüter, U., Mascher, M., Colmsee, C., Scholz, U., Bräutigam, A., Fahnenstich, H., et al. (2012). Maize source leaf adaptation to nitrogen deficiency affects not only nitrogen and carbon metabolism but also control of phosphate homeostasis. *Plant Physiol.* 160, 1384–1406. doi: 10.1104/pp.112.204420

Shao, C. H., Qiu, C. F., Qian, Y. F., and Liu, G. R. (2020). Nitrate deficiency decreased photosynthesis and oxidation-reduction processes, but increased cellular transport, lignin biosynthesis and flavonoid metabolism revealed by RNA-Seq in Oryza sativa leaves. *PLoS One* 15:e0235975. doi: 10.1371/journal.pone.0235975

Shi, H. W., Wang, L. Y., Li, X. X., Liu, X. M., Hao, T. Y., He, X. J., et al. (2016). Genome-wide transcriptome profiling of nitrogen fixation in *Paenibacillus* sp. WLY78. *BMC Microbiol.* 16:25. doi: 10.1186/s12866-016-0642-6

Sinha, S. K., Rani, M., Bansal, N., Gayatri, Venkatesh, K., and Mandal, P. K. (2015). Nitrate starvation induced changes in root system architecture, carbon:nitrogen metabolism, and mirna expression in nitrogen-responsive wheat genotypes. *Appl. Biochem. Biotechnol.* 177, 1299–1312. doi: 10.1007/s12010-015-1815-8

Studer, M. H., DeMartini, J. D., Davis, M. F., Sykes, R. W., Davison, B., Keller, M., et al. (2011). Lignin content in natural *Populus* variants affects sugar release. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6300–6305. doi: 10.1073/pnas.1009252108

Suárez, M. F., Avila, C., Gallardo, F., Cantón, F. R., García-Gutiérrez, A., Claros, M. G., et al. (2002). Molecular and enzymatic analysis of ammonium assimilation in woody plants. *J. Exp. Bot.* 53, 891–904. doi: 10.1093/jexbot/53.370.891

Syed, D. N., Afaq, F., Maddodi, N., Johnson, J. J., Sarfaraz, S., Ahmad, A., et al. (2011). Inhibition of human melanoma cell growth by the dietary flavonoid fisetin is associated with disruption of Wnt/β-catenin signaling and decreased Mitf levels. *J. Invest. Dermatol.* 131, 1291–1299. doi: 10.1038/jid.2011.6

Takei, K., Ueda, N., Aoki, K., Kuromori, T., Hirayama, T., Shinozaki, K., et al. (2004). AtIPT3 is a Key Determinant of nitrate-dependent cytokinin biosynthesis in *Arabidopsis*. *Plant Cell Physiol.* 45, 1053–1062. doi: 10.1093/pcp/pch119

Tian, J., Chen, J., Li, B., and Zhang, D. (2016). Association genetics in Populus reveals the interactions between Pto-miR160a and its target Pto-ARF16. *Mol. Genet. Genomics* 291, 1069–1082. doi: 10.1007/s00438-015-1165-9

Trachsel, S., Kaeppler, S. M., Brown, K. M., and Lynch, J. P. (2013). Maize root growth angles become steeper under low N conditions. *Field Crops Res.* 140, 18–31. doi: 10.1016/j.fcr.2012.09.010

Tschoep, H., Gibon, Y., Carillo, P., Armengaud, P., Szecowka, M., Nunes-Nesi, A., et al. (2009). Adjustment of growth and central metabolism to a mild but sustained nitrogen-limitation in *Arabidopsis*. *Plant Cell Environ.* 32, 300–318. doi: 10.1111/j.1365-3040.2008.01921.x

Tuskan, G. A., Groover, A. T., Schmutz, J., DiFazio, S. P., Myburg, A., Grattapaglia, D., et al. (2018). Hardwood tree genomics: unlocking woody plant biology. *Front. Plant Sci.* 9:1799. doi: 10.3389/fpls.2018.01799

Urbanczyk-Wochniak, E., and Fernie, A. R. (2005). Metabolic profiling reveals altered nitrogen nutrient regimes have diverse effects on the metabolism of hydroponically-grown tomato (*Solanum lycopersicum*) plants. *J. Exp. Bot.* 56, 309–321. doi: 10.1093/jxb/eri059

Vega, A., Canessa, P., Hoppe, G., Retamal, I., Moyano, T. C., Canales, J., et al. (2015). Transcriptome analysis reveals regulatory networks underlying differential susceptibility to Botrytis cinerea in response to nitrogen availability in *Solanum lycopersicum*. *Front. Plant Sci.* 6:911. doi: 10.3389/fpls.2015.00911

Vidal, E. A., Araus, V., Lu, C., Parry, G., Green, P. J., Coruzzi, G. M., et al. (2010). Nitrate-responsive miR393/AFB3 regulatory module controls root system architecture in *Arabidopsis thaliana*. *Proc. Natl. Acad. Sci. U.S.A.* 107, 4477–4482. doi: 10.1073/pnas.0909571107

Vogel, C., Bodenhausen, N., Gruissem, W., and Vorholt, J. A. (2016). The *Arabidopsis* leaf transcriptome reveals distinct but also overlapping responses to colonization by phyllosphere commensals and pathogen infection with impact on plant health. *New Phytol.* 212, 192–207. doi: 10.1111/nph.14036

Wang, J., Ma, X.-M., Kojima, M., Sakakibara, H., and Hou, B.-K. (2013). Glucosyltransferase UGT76C1 finely modulates cytokinin responses via cytokinin N-glucosylation in *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 65, 9–16. doi: 10.1016/j.plaphy.2013.01.012

Wang, X., Li, X., Zhang, S., Korpelainen, H., and Li, C. (2016). Physiological and transcriptional responses of two contrasting *Populus* clones to nitrogen stress. *Tree Physio.l* 36, 628–642. doi: 10.1093/treephys/tpw019

Wang, Y. N., Tang, L., Hou, Y., Wang, P., Yang, H., and Wei, C. L. (2016). Differential transcriptome analysis of leaves of tea plant (*Camellia sinensis*) provides comprehensive insights into the defense responses to Ectropis oblique attack using RNA-Seq. *Funct. Integr. Genomics* 16, 383–398. doi: 10.1007/s10142-016-0491-2

Wei, H., Yordanov, Y., Georgieva, T., Li, X., and Busov, V. (2013a). Nitrogen deprivation promotes *Populus* root growth through global transcriptome reprogramming and activation of hierarchical genetic networks. *New Phytol.* 200, 483–497. doi: 10.1111/nph.12375

Wei, H., Yordanov, Y., Kumari, S., Georgieva, T., and Busov, V. (2013b). Genetic networks involved in poplar root response to low nitrogen. *Plant Signal. Behav.* 8, e27211. doi: 10.4161/psb.27211

Woo, H., Koo, H., Kim, J., Jeong, H., Yang, J. O., Lee, I., et al. (2016). Programming of plant leaf senescence with temporal and inter-organellar coordination of

transcriptome in *Arabidopsis*. *Plant Physiol.* 171, 452–467. doi: 10.1104/pp.15.01929

Yang, J. H., Han, S. J., Yoon, E. K., and Lee, W. S. (2006). Evidence of an auxin signal pathway, microRNA167-ARF8-GH3, and its response to exogenous auxin in cultured rice cells. *Nucleic Acids Res.* 34, 1892–1899. doi: 10.1093/nar/gkl118

Yang, W., Yoon, J., Choi, H., Fan, Y., Chen, R., and An, G. (2015). Transcriptome analysis of nitrogen-starvation-responsive genes in rice. *BMC Plant Biol.* 15:31. doi: 10.1186/s12870-015-0425-5

Yu, J., Han, J., Wang, R., and Li, X. (2016). Down-regulation of nitrogen/carbon metabolism coupled with coordinative hormone modulation contributes to developmental inhibition of the maize ear under nitrogen limitation. *Planta* 244, 111–124. doi: 10.1007/s00425-016-2499-1

Zahran, H. H. (1999). Rhizobium-legume symbiosis and nitrogen fixation under severe conditions and in an arid climate. *Microbiol. Mol. Biol. Rev.* 63, 968–989. doi: 10.1128/MMBR.63.4.968-989.1999

Zhang, J., Liang, S., Duan, J., Wang, J., Chen, S., Cheng, Z., et al. (2012). De novo assembly and characterisation of the transcriptome during seed development, and generation of genic-SSR markers in Peanut (*Arachis hypogaea* L.). *BMC Genomics* 13:90. doi: 10.1186/1471-2164-13-90

Zhao, W., Yang, X., Yu, H., Jiang, W., Sun, N., Liu, X., et al. (2015). RNA-Seq-based transcriptome profiling of early nitrogen deficiency response in cucumber seedlings provides new insight into the putative nitrogen regulatory network. *Plant Cell Physiol.* 56, 455–467. doi: 10.1093/pcp/pcu172

Zheng, Z. L. (2009). Carbon and nitrogen nutrient balance signaling in plants. *Plant Signal. Behav.* 4, 584–591. doi: 10.4161/psb.4.7.8540

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

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 05 January 2022
doi: 10.3389/fpls.2021.752730

Drought Tolerant Near Isogenic Lines of Pusa 44 Pyramided With *qDTY2.1* and *qDTY3.1*, Show Accelerated Recovery Response in a High Throughput Phenomics Based Phenotyping

Priyanka Dwivedi[1], Naleeni Ramawat[2], Dhandapani Raju[3,4], Gaurav Dhawan[1], S. Gopala Krishnan[1], Viswanathan Chinnusamy[3,4], Prolay Kumar Bhowmick[1], K. K. Vinod[1], Madan Pal[4], Mariappan Nagarajan[5], Ranjith Kumar Ellur[1], Haritha Bollinedi[1] and Ashok K. Singh[1]*

[1] Division of Genetics, ICAR-Indian Agricultural Research Institute (ICAR-IARI), New Delhi, India, [2] Amity Institute of Organic Agriculture, Amity University, Noida, India, [3] Nanaji Deshmukh Plant Phenomics Centre, ICAR-IARI, New Delhi, India, [4] Division of Plant Physiology, ICAR-IARI, New Delhi, India, [5] Rice Breeding and Genetics Research Centre, ICAR-IARI, Aduthurai, India

OPEN ACCESS

Edited by:
Parviz Heidari,
Shahrood University of Technology, Iran

Reviewed by:
Raju Bheemanahalli,
Mississippi State University, United States
Alireza Pour-Aboughadareh,
Seed and Plant Improvement Institute, Iran

***Correspondence:**
S. Gopala Krishnan
gopal_icar@yahoo.co.in

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

Received: 03 August 2021
Accepted: 15 November 2021
Published: 05 January 2022

Citation:
Dwivedi P, Ramawat N, Raju D, Dhawan G, Gopala Krishnan S, Chinnusamy V, Bhowmick PK, Vinod KK, Pal M, Nagarajan M, Ellur RK, Bollinedi H and Singh AK (2022) Drought Tolerant Near Isogenic Lines of Pusa 44 Pyramided With qDTY2.1 and qDTY3.1, Show Accelerated Recovery Response in a High Throughput Phenomics Based Phenotyping. Front. Plant Sci. 12:752730. doi: 10.3389/fpls.2021.752730

Reproductive stage drought stress (RSDS) is a major challenge in rice production worldwide. Cultivar development with drought tolerance has been slow due to the lack of precise high throughput phenotyping tools to quantify drought stress-induced effects. Most of the available techniques are based on destructive sampling and do not assess the progress of the plant's response to drought. In this study, we have used state-of-the-art image-based phenotyping in a phenomics platform that offers a controlled environment, non-invasive phenotyping, high accuracy, speed, and continuity. In rice, several quantitative trait loci (QTLs) which govern grain yield under drought determine RSDS tolerance. Among these, *qDTY2.1* and *qDTY3.1* were used for marker-assisted breeding. A set of 35 near-isogenic lines (NILs), introgressed with these QTLs in the popular variety, Pusa 44 were used to assess the efficiency of image-based phenotyping for RSDS tolerance. NILs offered the most reliable contrast since they differed from Pusa 44 only for the QTLs. Four traits, namely, the projected shoot area (PSA), water use (WU), transpiration rate (TR), and red-green-blue (RGB) and near-infrared (NIR) values were used. Differential temporal responses could be seen under drought, but not under unstressed conditions. NILs showed significant level of RSDS tolerance as compared to Pusa 44. Among the traits, PSA showed strong association with yield (80%) as well as with two drought tolerances indices, stress susceptibility index (SSI) and tolerance index (TOL), establishing its ability in identifying the best drought tolerant NILs. The results revealed that the introgression of QTLs helped minimize the mean WU per unit of biomass per day, suggesting the potential role of these QTLs in improving WU-efficiency (WUE). We identified 11 NILs based on phenomics traits as well as performance under

imposed drought in the field. The study emphasizes the use of phenomics traits as selection criteria for RSDS tolerance at an early stage, and is the first report of using phenomics parameters in RSDS selection in rice.

Keywords: phenomics, controlled environment, image-based phenotyping, drought tolerance, rice

INTRODUCTION

Abiotic stresses such as drought are detrimental to the growth, development, productivity, and grain quality in rice (Kumar et al., 2012; Raman et al., 2012), because more than 80% of its growth period is water-dependent. This renders drought as an extremely hazardous abiotic stress affecting rice production globally (Kumar, 2018). The threat due to drought becomes more pertinent in the wake of global climate change, owing to its frequent occurrence around the globe. In India, particularly in states such as Odisha, Chhattisgarh, eastern Uttar Pradesh, Bihar, and Jharkhand, recent episodes of severe drought had resulted in significant yield losses in rice (Kumar A. et al., 2014; Kumar S. et al., 2014). Some of these states suffered yield losses of up to 40%.[1] Drought-related yield loss in rice during 2015–2016 alone is estimated to be 1.17 million tonnes (DACFW, 2017).

Drought is a complex phenomenon that combines interaction of various climatic, geographical, and edaphic factors resulting in deprivation of water for crop growth. Drought often occurs unpredicted after a failed rainfall and is commonly associated with high temperatures. Managing drought stress through agronomic practice after its onset is virtually unfeasible in rainfed ecosystems. Therefore, it becomes crucial to opt for preventive strategies against any drought occurrence during a crop cycle. Among the available options, developing rice varieties with inbuilt drought tolerance is the most strategic in terms of sustainability. Rice is naturally sensitive to drought stress, but the degree of damage depends on the affected crop stage, stress duration and intensity. Among the crop stages, drought incidence at the reproductive stage is most damaging in terms of economic returns. Therefore, reproductive stage drought stress (RSDS) tolerance is one of the most desirable attributes for cultivar development in rice. Drought tolerance is a complex quantitative trait with variable phenotypic impacts at different developmental stages (Oladosu et al., 2019). In the past, many quantitative trait loci (QTLs) have been mapped in rice, notably associated with RSDS tolerance and targeting grain yield under drought. However, utilizing these QTLs for the development of improved varieties requires a robust crop phenotyping strategy, which is efficient and time-saving (Tuberosa, 2012; Deery et al., 2014; Underwood et al., 2017).

Robust phenotyping has been implicated as a major impediment in evaluating breeding lines for drought tolerance in rice (Deery et al., 2016). The common methods in use often resulted in poor concordance of crop response between artificially managed and naturally occurring drought stress, leading to dubious conclusions (Courtois et al., 2000; Lafitte et al., 2006; Bernier et al., 2008; Venuprasad et al., 2009). In the traditional phenotyping methods, the poor precision is due to the limited use of actual crop response parameters, that are relatively simple, inaccurate, and intermittent. Besides, traditional drought phenotyping is laborious, time-consuming, hectic, economically ineffective, and plant destructive (Furbank and Tester, 2011; Chen et al., 2014). Notwithstanding, imposed drought treatments under natural field conditions had always been under the threat of unforeseen rains (Hoover et al., 2018). Although lately, facilities such as rain-out shelters have improved the screening system by preventing rain interference (Singh et al., 2016), they seldom offered improved fidelity in the response data. Rainout shelters are huge structures that require large land areas and are laborious to operate and maintain. Furthermore, the scaling up of rainout shelters to handle large populations becomes practically unfeasible after establishment.

Until the recent development of phenomics platforms, high throughput phenotyping has been seldom used in agricultural research (Lippman and Zamir, 2007). Assembled in an environment-controlled screen house, phenomics platforms provide image-based monitoring of crop ontogenesis on a continuous time scale. With its non-destructive phenotyping setup equipped to generate high-resolution data (Ubbens and Stavness, 2017), phenomics platforms help in screening for abiotic stress responses with the capabilities of handling large populations and automated data recording. The multidimensional imaging in the phenomics facilities uses visible, infrared (IR), near-IR (NIR), and hyperspectral bands, providing opportunities for temporal assessment of several physiological traits (Peñuelas and Filella, 1998; Tester and Langridge, 2010). For instance, the stomatal response to drought stress is an important trait for screening tolerance that is measured using thermal imaging and canopy temperature (Jones et al., 2009; Rischbeck et al., 2017). Likewise, several stress responses can be assessed through continuous crop monitoring.

Phenotyping in a phenomics platform is an automated process, in which each plant travels through a battery of imaging devices at pre-programmed intervals. During this process, several plant health parameters are monitored continuously to ensure ideal cultural environment for the test plants. Therefore, the phenotyping ensures accuracy and precision, while saving time and labor of handling large populations. The data generated are highly reliable than those generated from rainout shelters and field-based platforms. Contemporarily, the number of studies employing high throughput phenotyping is on the rise for studying quantitative inheritance of various traits in rice. Yang et al. (2015) could identify nine QTLs associated with leaf traits from a genome-wide association study (GWAS) involving 533 rice accessions. Similarly, Guo et al. (2018) employed 51 image-based traits to evaluate drought tolerance in 507

[1] www.irri.org/climate-change-ready-rice

rice accessions. Dynamic quantification of RSDS response among 40 rice accessions grown under two cultural conditions, pot and field, using high throughput phenotyping, revealed that four traits, viz., greenness plant area ratio (GPAR), total plant area/bounding rectangle area ratio (TBR), total plant area/convex hull area ratio (TAR) and perimeter area ratio (PAR) were capable of differentiating drought resistant and susceptible genotypes (Duan et al., 2018). Kim et al. (2020) demonstrated the efficiency of red-green-blue (RGB) images in predicting plant area, color, and compactness; NIR images for assessing plant water content; IR images for assessing plant temperature and fluorescence; and gravimetric platform (DroughtSpotter®) for measurement of water use efficiency (WUE), plant water loss rate, and transpiration rate at various crop stages. They used the data to assess the drought response between a drought-tolerant rice mutant, *osphyb,* and its wild type (WT), by estimating photosynthetic efficiency. Although, high throughput phenotyping has been used to study drought and high-temperature responses in crops such as wheat (Shirdelmoghanloo et al., 2016), tomato (Danzi et al., 2019), and *Brassica* (Chen et al., 2019), the use of image-based screening for assessing stress tolerance in rice is still in its infancy. Therefore, detailed experimentations are needed for understanding the versatility of this method in breeding for tolerance to abiotic stresses such as drought.

Near isogenic lines (NILs) form an ideal set of experimental material for comparative phenotypic evaluation and assessment of the precision achieved in phenomics platforms. Unlike the mutants, NILs are easy to generate and provide relatively accurate genomic comparison. Moreover, several NILs can be simultaneously tested to generate comprehensive data suited for a reliable analysis. Marker-assisted backcross breeding (MABB) and genetic engineering are used to generate NILs in crops. MABB has been proved to be one of the best ways to incorporate target traits in cultivars (Dhawan et al., 2021; Singh et al., 2021). In recent years, several QTLs and meta-QTLs have been identified in rice which are associated with different levels of tolerance to drought. Further, several improved lines carrying these QTLs such as *qDTY1.1*, *qDTY2.1*, *qDTY2.2*, *qDTY3.1*, *qDTY4.1*, *qDTY3.2*, *qDTY9.1*, *qDTY10.1,* and *qDTY12.1* (Vinod et al., 2019; Dhawan et al., 2021; Dwivedi et al., 2021; Oo et al., 2021a) have been developed in rice.

In the present study, a set of 35 NILs developed in the background of the popular rice cultivar, Pusa 44 introgressed with two QTLs, *qDTY2.1* and *qDTY3.1* (Dwivedi et al., 2021) were evaluated in the field as well as in a phenomics facility, under induced RSDS. The objective was to identify the most promising NILs showing resilience under RSDS using different image-based parameters such as NIR and visual (VIS) imaging. We hypothesized that a superior drought tolerant NIL would be the one that performs better than Pusa 44 during drought stress while showing a less significant reduction in the projected shoot area (PSA), a slower increase in NIR intensity as well as a maximum reduction in water use (WU) and transpiration rate (TR). Agromorphologic data were also recorded manually at the time of harvest from all the test genotypes for comparative evaluation.

MATERIALS AND METHODS

Plant Materials

A subset consisting of 35 promising NILs was identified from a population of 143 di-QTL introgression lines carrying two RSDS tolerant QTLs, *qDTY2.1* and *qDTY3.1,* in the background of Pusa 44. The subset was selected based on agronomic performance and grain quality parameters. Additionally, four genotypes were used as checks, which included, Pusa 44 – the recipient parent (RP), IR81896-B-B-142 – the donor parent (DP), and two RSDS tolerant checks *viz.,* IR81896-B-B-195 (RC1) and IR87728-59-B-B (RC2). Of these, Pusa 44 is a popular rice variety of north-western India suitable to the rice-wheat cropping system, with good yield, grain quality, and ideal agronomic characteristics. The donor, IR81896-B-B-142, is a line developed from the cross Apo/Swarna*2 carrying both the QTLs: *qDTY2.1* and *qDTY3.1* (Venuprasad et al., 2009). Among the checks, RC1 was developed from the same cross as that of the DP, while the second check, RC2, was derived from the cross, Aday sel/IR64. RC2 carries the QTLs *qDTY2.2* and *qDTY4.1* (Sandhu et al., 2018). The plant materials were grown with standard cultural practices for rice under transplanted conditions. Both the checks were proven to be field tolerant to RSDS from multiple experiments and locations (Singh et al., 2016; Dwivedi et al., 2021).

Field Phenotyping for Drought Stress Response

The first field experiment was set up using 35 NILs grown in an augmented randomized block design (ARBD) with six blocks along with four checks. For each genotype, 120 plants were grown per treatment in four rows using a spacing of 20 × 15 cm. The NILs were subjected to two treatments, unstressed and stressed. In the unstressed treatment, a normal irrigation regime was maintained throughout the crop duration. Under stressed treatment, irrigation was discontinued after the initiation of the reproductive phase. However, lifesaving irrigation was provided when the soil water potential reached −70 KPa. Both the treatments received the same agronomic management, except for the imposition of drought stress (**Figure 1**). At the physiological maturity stage, five uniform looking plants were selected from each genotype per treatment for recording phenotypic data such as days to 50 percent flowering (DFF), plant height (PH), number of productive tillers (NPT), panicle length (PL), biomass (BM), spikelet fertility (SF), grain yield per plant (YP), hulling percentage (H%), milling percentage (M%), yield per hectare (PY), the weight of 1,000 grains (TW), and elongation ratio (ER). The experiment was carried out at the research farm of the Division of Genetics, ICAR-Indian Agricultural Research Institute (ICAR-IARI), New Delhi. In the next season, nine selected best NILs were field evaluated in a randomized complete block design (RCBD) along with the same four checks.

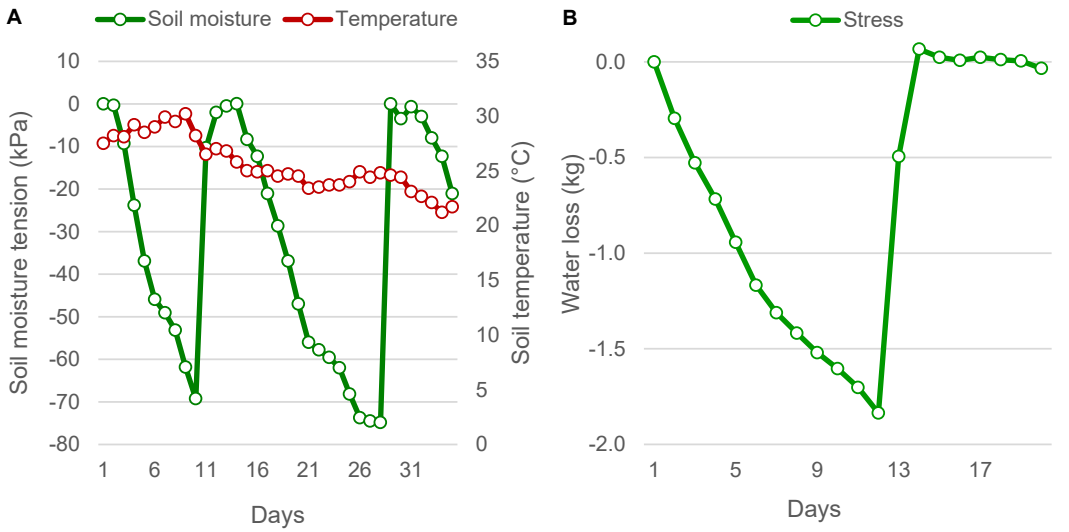


FIGURE 1 | Soil moisture status under stress in the field and pot culture experiments. **(A)** In the field, soil moisture depletion was determined by the increasing soil moisture tension, measured using tensiometers. The soil moisture tension reached below -70 kPa twice during the reproductive stage, which forced lifesaving irrigations. Average ambient air temperature during the period did not vary much. It took an average of 9.5 days to reach the maximum stress from field capacity. **(B)** The pot culture under phenomics also was subjected to similar stress, the moisture depletion was measured as the falling weight of the pots. When the pots lost 1,840 g of water on the eleventh day, the plants experienced severe stress as that seen in the field, which was followed by lifesaving irrigation on the twelfth day of imposing stress.

Controlled Environmental Phenotyping

A pot culture experiment was conducted using 35 NILs along with their parents, Pusa 44 and IR 81896-B-B-142, in a controlled environment facility at the Nanaji Deshmukh Plant Phenomics Centre (NDPPC) at ICAR-IARI, New Delhi. NILs were initially sown in a raised bed nursery and transplanted after 30 days into the field. At the active tillering stage (∼ 30 days after transplanting), the plants were shifted to pots filled with 15 Kg of soil. Care was taken to keep the soil core intact without damaging the roots. After planting, the pots were flood irrigated with an equal volume of water (∼3 liters) and the weight on the pots was equated to 18.5 kg using soil. Pots were allowed for 12 h to get the soil fully saturated. The next day, all the pots were weighed again to confirm the uniformity in pot weight. The pot weight was measured using an automated weighing station. The potted plants were then shifted to the climate-controlled greenhouses at the NDPPC. A total of 156 pots were maintained, with two plants per treatment for each genotype. All the pots were tagged according to the cars they were loaded to. The plants were irrigated twice a day at 6 AM and 6 PM up to saturation pot weight and maintained well-watered with full saturation of 25% w/v soil moisture content (SMC). First, imaging of the plants was done 80 days after sowing. Thereafter, irrigation for one set of pots (stressed treatment) was restricted till the soil moisture dropped to 50% of saturation coinciding ∼ 12 % SMC. At this stage, all the plants showed severe drought symptoms and the SMC was estimated through the gravimetric method. Lifesaving irrigation was given at this stage to reach the saturation pot weight, raising the SMC to 25%. During the interval between the withdrawal of irrigation and the saturation, daily pot weight data was collected for measuring the progress of water depletion in the drought-stressed pots. From this data, the evapotranspiration (ET) was computed using the pot weight difference between the consecutive days. A set of pots without plants (mocks) were used as the internal control for measuring evaporation (E), using which whole plant transpiration per pot (T) was calculated and expressed in grams per pot. WU was also measured from the water consumption data and recorded in ml/g/day. The potted rice plants were maintained in a climate control led greenhouse, providing a photo regimen of 14 h light/10 h dark, a temperature regimen of 32/25^{o}C day/night, and relative humidity of ∼ 60%.

Imaging of the genotypes was carried out for five-time snaps with five days' intervals using RGB and NIR sensors. These stages, represented as I, II, III, IV, and V, indicated the number of five-day intervals. Stages I, II, and III denoted the phases under drought stress, after withholding the irrigation, while the stages IV and V denoted the recovery phase after providing the life-saving irrigation. The RGB images were captured using a Prosilica GT6400 (Allied Vision, Stadtroda, Germany) series visual camera with an RGB spectral range of 400-700 nm installed in Scanalyzer3D imaging platform (Lemna Tec GmbH, Aachen, Germany). Images were captured from one top view and three side views (SV0, SV120, and SV240) having a pixel matrix of 6576 × 4384. The raw images were acquired with the basic setup of the camera with an exposure set to 30,000, gain set to 20, gamma set to 125, the red-white balance of 110, and blue-white-red balance of 195, and provided with a constant fluorescent light source. The raw images were processed using the LemnaGrid® software (LemnaTec GmbH, Aachen, Germany) using standard image processing tools. RGB images captured from all the three side views were used to estimate the PSA

of the plant and expressed in cm^2 (Neilson et al., 2015; Parent et al., 2015). The estimated PSA was used for normalizing whole plant transpiration to calculate the transpiration rate (TR) and expressed in g per cm^2. To capture the leaf architectural changes specific to drought stress response, the stage-wise PSA was normalized with tiller number and expressed as PSA per tiller. The RGB data was also used to estimate a projected plant area (PPA), which was used for computing the growth rate. The rate of growth was calculated by dividing the PPA by the time interval in days between the two measurements. The hydration status of the plant was measured using NIR imaging. A Goldeye® G032 NIR camera (AVT Allied Vision, Exton, Pennsylvania, United States) with a pixels' resolution of 636 × 508 was used to acquire NIR images at 0°, 120°, and 240° from the side-view. To delineate the region of interest for the stress response, the NIR gray images were matched with corresponding RGB images. However, to compensate for the differences in the plant positions and in the image resolution between NIR and RGB images, a local matching method using 130 matching points was used. The matched images were cropped to a uniform size, and the plant area calculated from the RGB image was extracted to confirm the average water content. The average NIR pixel intensities were obtained at the water absorption wavelength of 1,450 nm, which was read within a range of zero to 255. Therefore, plants with a high-water content showed a low NIR intensity. The mean gray value was used for easy assessment of plant water status. At the time of maturity, the plants of individual pots were harvested and final phenotypic data was recorded, for traits namely PH, BM, TN, and YP.

Calculation of Drought Indices

To assess the response for grain yield under drought, nine stress indices were calculated for both pot and field experiments. These indices were computed for each genotype using two parameters, yield under unstressed (Y_p) and yield under stressed conditions (Y_s). The average grain yield of all genotypes under unstress and stress conditions were represented by $\bar{Y}_p$ and $\bar{Y}_s$, respectively. Indices were used for initially ranking the NILs for each index based on the key drought tolerance indicator attached to it. Details of the indices and their key indicators for drought tolerance are given in **Table 1**.

Data Processing and Statistical Analyses

The data were analyzed for variability and trend over the experimental period. Additionally, Pearson's correlations were worked out between drought indices and phenomics parameters. All the basic data operations and elemental statistical analyses were carried out using the Data Analysis Addin in Microsoft Excel® v. 2019 (Microsoft Corporation, Redmond, United States). The agromorphologic data from the field and pot experiments were subjected to analysis of variance, carried out using STAR package 2.0.1 (IRRI, 2014). For the grain yield, statistical comparisons were done under stressed and unstressed treatments.

RESULTS

Intensity of Drought

In both field and pot experiments, the intensity of drought reached critical levels producing significant drought responses in plants (**Figure 1**). In the field conditions, soil moisture level dropped twice to critical level falling to a maximum tension of −74.8 kPa, within a 35 day-window during the reproductive phase. The moisture depletion took 10 days to reach the critical level of −70 kPa from saturation during both episodes. During the 2016 trial, despite two peak drought incidences, there was only one lifesaving irrigation provided, due to a rain of 37.8 mm received on the tenth day. This rain did not affect the experiment, because by the time the soil moisture tension had reached −69.3 kPa. However, during the 2017 season, two lifesaving irrigations were needed, as there was no rain during the reproductive stage (**Supplementary Table 1**). Following the lifesaving irrigation, the moisture level was restored to saturation. The ambient average temperature during this period did not show much variation but had shown a dropping trend as the days advanced. The temperature was maximum on the 9^{th} day with an average of 30.2°C. Under the pot experiment, soil moisture drop was induced only once, which showed a similar pattern of depletion as observed in the field. However, it took 12 days to reach the 50% depletion level. At this point, the average SMC ranged between 10.01 and 13.80 %, under gravimetric determination. There was no variation in the ambient temperature under pot culture.

Variability of Near Isogenic Lines' Response Under Imposed Drought

The ANOVA revealed significant variations for yield and related agronomic traits under both stress and unstress conditions in the field as well as pot-based evaluation (**Table 2**). The variation among the NILs was higher for most of the traits under stress conditions than those recorded under unstressed treatment. The ratio of stress and unstress variances indicated higher variation for NPT, H%, M%, ER, and SF. However, the variation among NILs was found to be higher in unstress conditions for a few traits like DFF, PH, and TW. The variance for BM and PY were relatively similar under both conditions in the field. Whereas, under the pot culture significant variation could be observed only for for NP and YP among the NILs and checks under stress. But, significant variation could be noticed only for YP under unstressed treatment in this experiment. The ratio of variances also revealed inconspicuous differences.

Image-Based Phenotyping of Near Isogenic Lines Under Stressed and Unstressed Conditions

Under controlled phenotyping, four major traits (PSA, WU, TR, and NIR intensity) estimated from RGB and NIR images signified the effect of drought at various test stages. The box and whisker plots for traits, indicated relatively lesser variability for all the traits under unstressed conditions than under stress (**Figure 2**). Compared with the response under unstressed pots, an apparent

TABLE 1 | Drought tolerance indices used in the study.

Drought tolerance index	Notation	Formula	Unit	Range	Tolerance indicators	References
Relative stress tolerance	TOL	$100(Y_p–Y_s)/Y_p$	%	0–100%	A low value indicates tolerance	Rosielle and Hamblin (1981)
Mean productivity	MP	$0.5(Y_p+Y_s)$	g	NIL	A high value indicates a better yield	Rosielle and Hamblin (1981)
Geometric mean productivity	GMP	$(Y_p*Y_s)^{0.5}$	g	NIL	Similar to MP	Fernandez (1992)
Stress index	SI	$1–(\tilde{Y}_s/\tilde{Y}_p)$	–	0.0 – 1.0	A low value indicates mild stress or high mean tolerance	Fischer and Maurer (1978)
Stress susceptibility index	SSI	$[(1–(Y_s/Y_p)]/SI$	–	0.0 – 1.0	A low value indicates relatively high tolerance over the average stress	Fischer and Maurer (1978)
Stress tolerance index	STI	$(Y_p*Y_s)/(\tilde{Y}_p)^2$	–	0.0 – 1.0	A high value indicates better tolerance	Fernandez (1992)
Drought yield index	YI	$Y_s/\tilde{Y}_s$	–	NIL	A high value indicates better tolerance	Lin et al. (1986)
Yield stability index	YSI	Y_s/Y_p	–	0 – 1.0	A high value indicates better tolerance	Bouslama and Schapaugh (1984)
Relative tolerance	RT	$(Y_p–Y_s)/Y_s*100$	%	0 – 100	A low value indicates tolerance	Oo et al. (2021b)

TABLE 2 | The ANOVA for agronomic traits showing components of variance under stressed and unstressed treatments in the near-isogenic lines (NILs) in the field trial.

Trait	Variance under stress (Vs)			Variance under unstress (Vus)			Vs/Vus for NILs
	Checks	NILs	Residual	Checks	NILs	Residual	
DF_f	144.0*	2.2*	54.4*	121.0*	13.3*	42.71*	0.17
PH_f	1354.0*	9.0*	750.5*	1253.0*	20.0*	1261.9*	0.45
NP_f	0.2*	101.1*	6.9*	2.6*	1.6*	1.1*	62.62
PL_f	0.4*	1.0*	0.3*	0.04*	1.5*	7.6*	0.68
BM_f	24.6*	83.0*	161.1*	161.8*	90.3*	39.4*	0.92
SF_f	8.8*	64.5*	301.0*	8.6*	44.1*	385.4*	1.46
YP_f	9.6*	10.5*	69.5*	3.2*	19.5*	46.2*	0.54
$H\%_f$	6.0*	223.8*	161.6*	1.7*	1.3*	1.2*	178.18
$M\%_f$	3.1*	163.3*	241.3*	3.2*	1.7*	0.1*	96.74
TW_f	15.7*	1.8*	9.2*	0.3*	3.5*	5.9*	0.52
PY_f	1,058,244*	1,158,970*	4,108,996*	777,060*	1,158,038*	29,022,097*	1.00
ER_f	0.01*	0.01*	0.02*	0.00*	0.001*	0.01*	8.00
PH_p	419.1*	18.6^{ns}	1537.9*	619.5*	19.2^{ns}	685.6*	0.97
$W1_p$	79.3^{ns}	26.4^{ns}	7.7^{ns}	93.9^{ns}	50.0^{ns}	590.6^{ns}	0.53
$W2_p$	73.4*	9.9^{ns}	21.8^{ns}	72.9^{ns}	7.9^{ns}	117.5^{ns}	1.25
BM_p	195.4^{ns}	46.3^{ns}	67.4^{ns}	275.5^{ns}	61.9^{ns}	1234.9^{ns}	0.75
NP_p	6.8*	2.1*	0.3^{ns}	0.3^{ns}	1.1^{ns}	11.3^{ns}	1.79
YP_p	9.5*	3.8*	20.0*	3.0*	3.6*	19.8	1.05

*DF, days to 50% flowering; PH, plant height in cm; NP, number of panicle-bearing tillers; PL, length of panicle in cm; BM, biomass per plant in g; W1, the weight of plant part above the pot in g; W2, the weight of plant part between pot and soil level in g; SF, spikelet fertility in %; YP, grain yield per plant in g; H, hulling per cent; M, milling percentage; TW, weight of 1000 grains in g; PY, Plot yield in kg per ha; ER, elongation ratio; *significant at 5% level. The suffixes f and p indicate field and pot culture, respectively; ns, non-significant.*

drop was observed for stage-wise PSA, WU, and TR in the stressed pots. The drop occurred till stage III, beyond which the traits showed improvement, thereby producing a positive trend following irrigation. In the case of NIR intensity, under drought, the mean gray values increased with progressive stress, which declined at the recovery phase. Analyzing the mean performance of treatment and genotypes at different stages of the trial, it was noticed that the measured traits were statistically on par before the stress initialization (stage I). As the trial progressed, the trait means began to show significant statistical differences among the plants that were subjected to stress (**Supplementary Table 2**). For traits such as WU, TR, and NIR, apparent treatment effects could be detected beginning from stage II. The treatment difference for PSA was found remarkable beginning from stage III. Further, the differences in treatment effect could be seen even at the recovery stages after irrigation among the traits such as PSA,

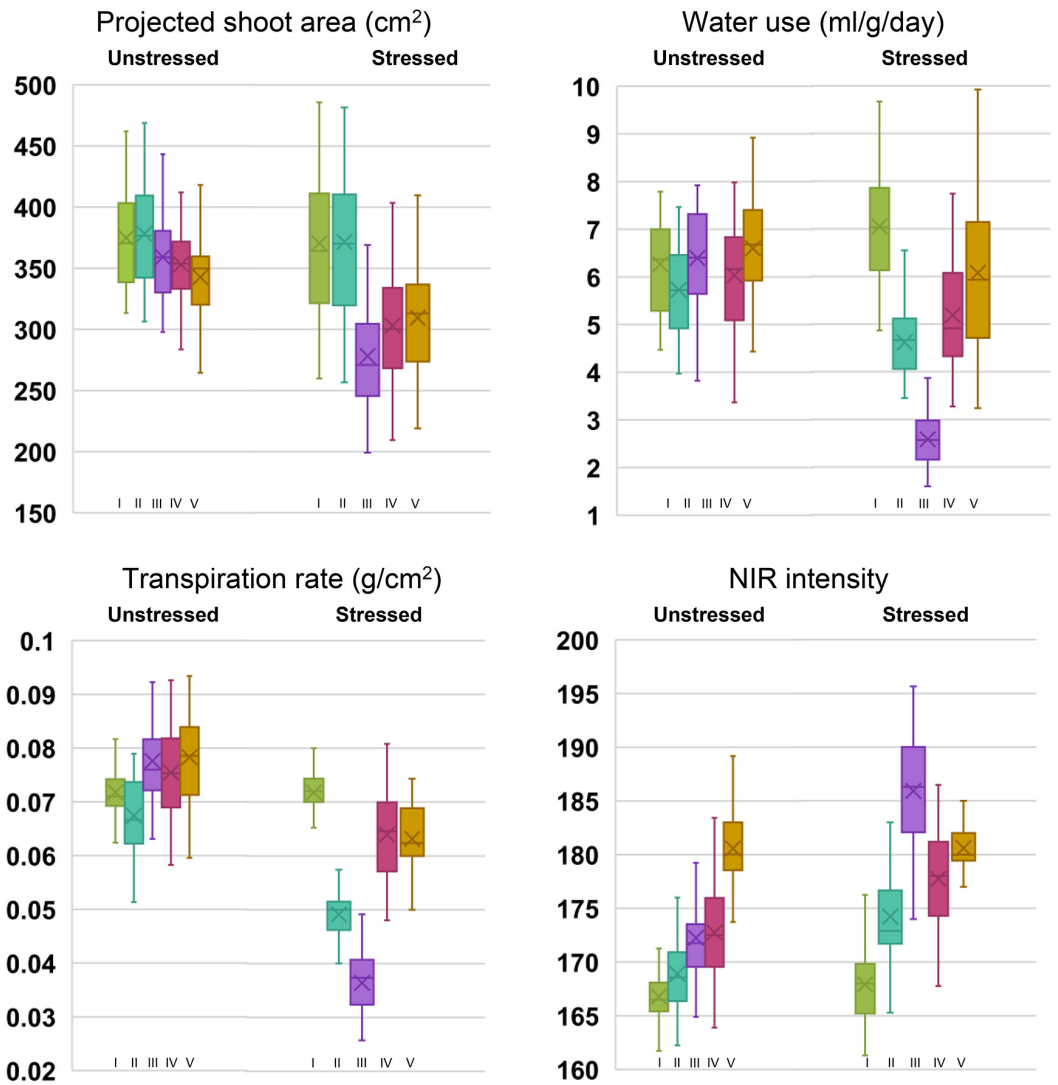


FIGURE 2 | Box plots of phenomics parameters, projected shoot area (PSA), water use (WU), transpiration rate (TR), and near-infrared values (NIR) among the genotypes under stressed and unstressed treatments for five stages of drought exposure. Stages I to V corresponds to the initial phase without drought (0 days), 5 days after withholding irrigation in drought set, peak drought phase at 10 days after drought imposition when soil moisture content reached 12%, two days after irrigation, and 5 days into the recovery phase, respectively.

WU, and TR. The NIR values, however, showed a quick reversal during the recovery phase (stage V) by indicating non-significant differences among the treatment means. The genotypic variability was also found to vary with the stress treatment. We could observe significant variations among the parents and the NILs for PSA beginning from stage III. Whereas, for WU and TR, the genotypic differences became apparent only beyond stage IV. Interestingly, the differences in NIR values could be noticed only during stage IV. The average treatment values for PSA, WU, and TR were higher under irrigated treatment than the stressed one. Among the parental lines, under both stressed and unstressed conditions, higher PSA, WU, and NIR intensity were observed for the donor parent, IR81896-B-B-142, over the recurrent parent, Pusa 44. The NILs, however, showed significant improvement over Pusa 44 for these traits. In the case of NIR intensity, the mean values steadily increased from stage I (166.81) to stage V (180.7) under unstressed pots. When subjected to stress, the intensity initially increased from 167.56 (stage I) to 185.62 (stage III) and recovered back to 180.54 at stage V. The rate of transpiration decreased significantly from 0.072 to 0.036 and recovered up to 0.063 g/cm^2 under stress conditions. The coefficient of variation (CV) between the stages for each genotypic class showed a conspicuous increase in the CV between stressed and unstressed treatments. Prominent CV difference was observed for traits such as WU, TR, and NIR.

Comparing the individual genotypic response, under drought stress, all the NILs were found to perform better than Pusa 44, while several of them outperformed the donor parent, IR

81896-B-B-142 (**Supplementary Table 3**). Although there was a reduction in parameter estimates of PSA, WU, and TR as the drought treatment progressed through stages I to III, NILs along with the donor parent generally showed a rapid recovery than Pusa 44. For PSA, the percentage reduction among the NILs from stage I to stage III was ranging between 9.1 and 33.7%. The WU and TR (**Supplementary Tables 4, 5**) were significantly reduced from stage I to stage III and ranged between 44 to 76.1% and 25.9 to 67.4%, respectively. Similarly, NIR intensity values were found to increase during drought stress and percent increase ranged between 3.5 and 14.6 % (**Supplementary Table 6**). Out of 35 NILs, four lines (P1823-12-42, P1823-12-82, P1823-12-118, and P1823-12-143) were found to be superior with a lesser reduction in PSA; lesser increase in NIR intensity with the highest reduction in WU and TR during progressive drought stress (Stage I to III). A few additional NILs (P1823-12-21, P1823-12-44, P1823-12-48, P1823-12-49, P1823-12-50, P1823-12-63, P1823-12-77, P1823-12-79, P1823-12-80, and P1823-12-81) were also found good as they showed a lesser reduction in PSA, TR, and WU along with a lesser increase in NIR intensity.

We could identify NILs with rapid recovery, another key indicator of drought tolerance. These lines showed a rapid gain in PSA, WU and TR while having a rapid reduction in NIR values. The post irrigation PSA increase at stage III among the NILs ranged between 0.4 and 23.7 %, as against the 20.6% of the donor parent, IR81896-B-B-142. PSA increment in Pusa 44 was 11.2%. Among the NILs, P1823-12-65, P1823-12-114, P1823-12-77, P1823-12-48, P1823-12-64, P1823-12-50, P1823-12-32, P1823-12-141, P1823-12-42, P1823-12-23, P1823-12-49, P1823-12-123, P1823-12-98, P1823-12-127, P1823-12-63, and P1823-12-104 were identified to perform better than Pusa 44. The values for WU and TR were also found to increase from stage I to stage III in all NILs between 31.8 to 71.2 % and 18.1 to 55.8 %, respectively. Where, higher WU and TR values were observed in the donor parent (63.8% and 50.8%, respectively) than in Pusa 44 (59.6% and 41.8 %). The NILs such as, P1823-12-23, P1823-12-32, P1823-12-42, P1823-12-48, P1823-12-64, P1823-12-65, P1823-12-84, P1823-12-104, P1823-12-118, P1823-12-124, and P1823-12-130 were found to have better WU as well as TR than Pusa 44. The reduction in NIR intensity values during the post-stress recovery period ranged between 0.2 and 6.7% among the genotypes. The highest reduction was observed in IR81896-B-B-142 (6.0%) than in Pusa 44 (4.2%). The NILs which indicated better recovery through a rapid decline in NIR intensity included P1823-12-32, P1823-12-48, P1823-12-65, P1823-12-80, P1823-12-96, P1823-12-104, P1823-12-124, P1823-12-132, and P1823-12-141, which was better than Pusa 44.

Temporal Responses of Parents and Near Isogenic Lines

Graphical comparison of the genotypic response pattern over the five stages of phenotyping, indicated significant contrast between the parents for traits such as PSA, WU, TR, and NIR under unstressed conditions (**Figure 3**). However, the response curves were largely flat indicating inconspicuous differences between the stages. In contrast, under drought, all the genotypes showed a common response pattern with a declining phase during progressive drought and a recovery phase after irrigation as described earlier. Interestingly, the NILs were found almost similar to Pusa 44 under unstressed conditions, while they exhibited a significant shift away from Pusa 44 under drought. Particularly for NIR intensity, NILs showed a lower value when compared with Pusa 44, indicating less water use to maintain higher leaf hydration status.

Differential responses were recorded for some of the NILs in terms of their values at various stages. Out of the 35 NILs evaluated, 14 NILs were showing desirable trends for all the four parameters both under stress and non-stress conditions. While some of the lines showed a flatter curve during drought, indicating their endurance, while others showed a trend parallel to the donor parent. These criteria were used for selecting the best-performing ones. Some NILs outperformed both the parents under drought conditions and while remaining similar to Pusa 44 during unstressed conditions. Based on these patterns, three NILs, namely, P1823-12-21, P1823-12-32, and P1823-12-82 which showed tangible differences in the trend were selected for further comparison (**Figure 4**).

Under non-stress, all the three NILs showed a near similar trend for all the traits, with a clear advantage over Pusa 44. Under stress, P1823-12-21 showed a nearly flat trend for PSA along with low WU and NIR, and an intermediate TR between the parents. While under stress, this NIL showed a very similar pattern for PSA, WU, and TR, like that of both the parents. The NIR values of P1823-12-21 were lower than the parents. On the contrary, P1823-12-32 did not show differences in PSA values at all stages under stress from the parents. This NIL also showed no significant variation in WU and TR, except for a low NIR intensity in the initial stages of stress indicating better maintenance of systemic hydration status. The third NIL, P1823-12-82 showed significantly improved PSA and WU under stress with low NIR intensity and an intermediate TR that was apparent during the recovery phase. Accordingly, P1823-12-82 was considered to have the best drought response followed by P1823-12-32 and P1823-12-21. All the NILs showed trends similar to that of the three selected NILs for different parameters. The line graphs for all other NILs are presented in **Supplementary Figure 1**.

Agronomic Performance of the Near Isogenic Lines Under Controlled Environment Screen

Following the controlled drought exposure under the phenomics platform, the agronomic traits of the genotypes were compared for the agronomic performance (**Table 3**). The results indicated significant differences among the stressed and unstressed treatments as well as within the genotypes. Considering the CV, stressed treatment showed a larger variation for NT and YP relative to unstressed treatment. However, PH and PB did not show explicit variation at the maturity stage. Further, the variation for TN was also not much remarkable as that of YP. Therefore, among the traits, grain yield was the trait largely affected by the transient drought treatment provided at the

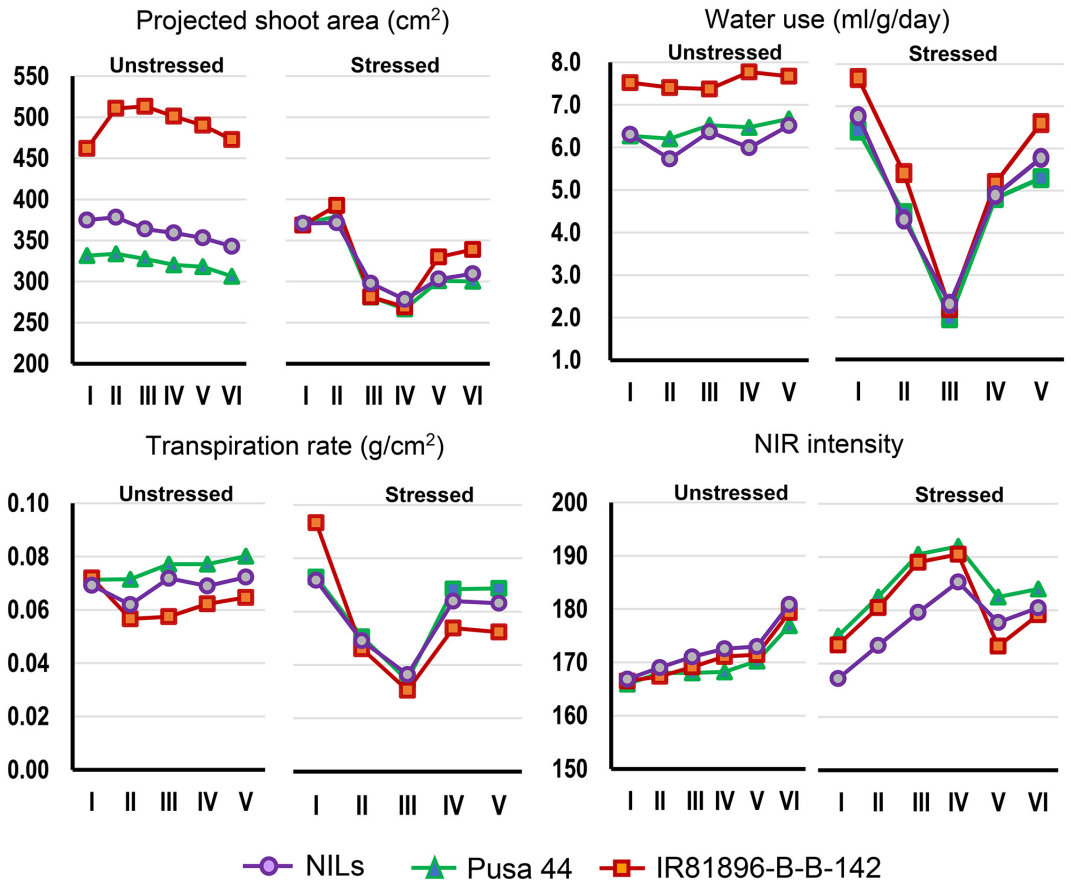


FIGURE 3 | Graphical comparison of the temporal response of genotypes under stressed and unstressed treatments for the four phenomics parameters. The overall trend indicated differential response under stress which was largely absent under unstressed conditions.

reproductive stage. Comparing the relative yield reduction (RDY) between the NILs and parents could bring out the effect of drought on the genotypes. The RDY values for the NILs ranged between 0.0 and 0.73 with an average of 0.34. The recurrent parent, Pusa 44 showed an RDY of 0.71, while IR 81896-B-B-142 had that of 0.49. This indicated that most of the NILs performed better than the parent, under stress. Seven NILs, P1823-12-23, P1823-12-48, P1823-12-63, P1823-12-64, P1823-12-65, P1823-12-114, and P1823-12-141 possessed higher RDY values than the tolerant parents IR81896-B-B-142. The rest of the 28 NILs were significantly better than both the parents under stress. One of the NILs, P1823-12-89 produced equal yield under both stress and unstressed conditions bringing the RDY to zero. Similarly, there were two other NILs, P1823-12-104 and P1823-12-122, which had RDY less than 10%. Among the 35 NILs tested, 34 of them have RDY values above that of Pusa 44, while 28 exceeded the donor parent. Also, the relative tolerance (RT) was compared for the 28 NILs that performed better than both parents, and the RT values were found to range from 0 to 87. RT of tolerant (donor) parent IR81896-B-B-142 was 100, while that of Pusa 44 was 256. Based on the pattern of drought response that emerged from the controlled phenotyping, the NILs' performance was further compared to their actual field performance to establish any correspondence.

Drought Indices

The pattern of drought tolerance indices divulged a better picture of the comparative evaluation of the drought tolerant NILs. Since these indices had different tolerance indicators emphasizing various drought responses, we have employed all of them together for evaluation. Seven drought indices calculated using yield per plant under control and stress conditions suggested that almost all the NILs performed better than recipient parent Pusa 44 in the field (**Table 4**). No indices were calculated for the pot experiment, as the drought imposed was not enough to dissect the tolerance response for grain yield, since only one episode of the drought was induced in this evaluation. The lower values of stress susceptibility index (SSI) and higher values of stress tolerance index (STI) indicated the better performance of the line under stress. All the NILs except five were showing lower values of SSI while 32 NILs had higher values for STI in comparison to Pusa 44 having an SSI was 1.56 and STI was 0.39. Corresponding SSI and STI values for IR81896-B-B-142 were 0.98 and 0.56, respectively.

A similar pattern was seen for all other indices too, with NILs having higher mean productivity (MP) and yield index (YI) than

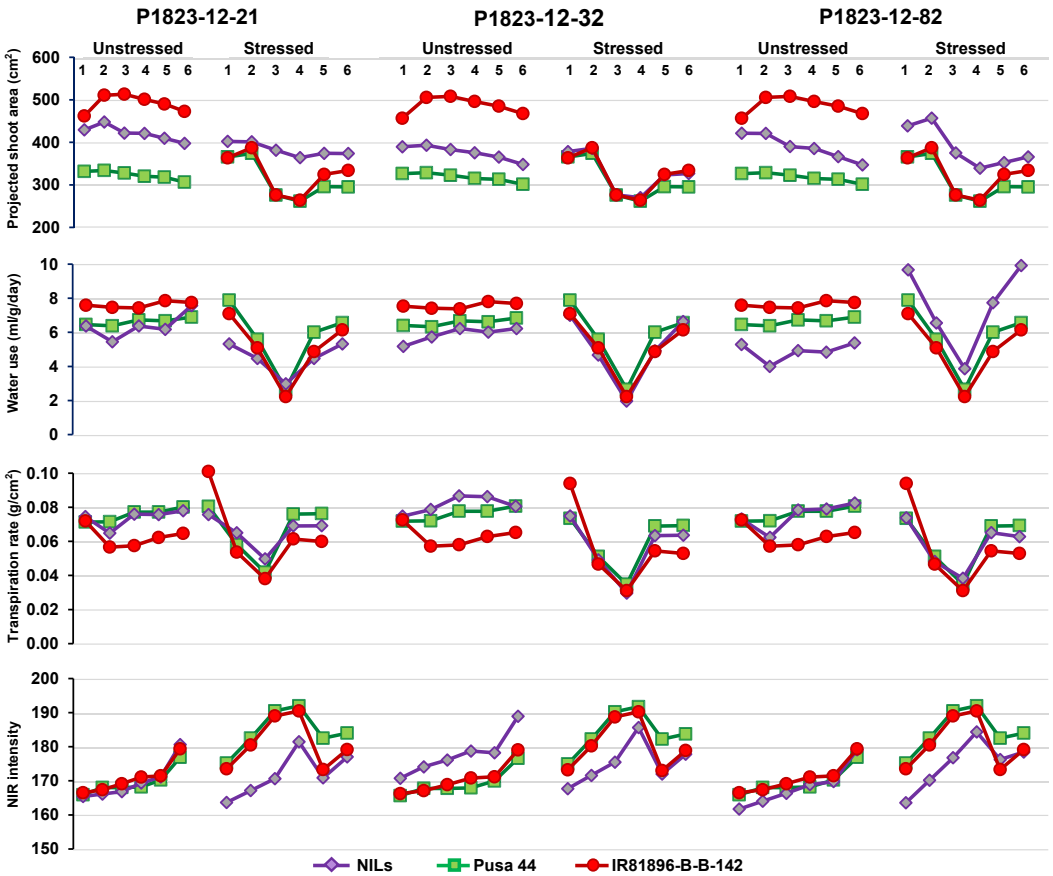


FIGURE 4 | Graphical trends of three variable near-isogenic lines (NILs) (P1823-12-21, P1823-12-32, and P1823-12-82) under control and stress conditions for phenomics parameters.

Pusa 44. Besides, there were 27 NILs with higher MP and 24 with higher YI than the donor parent. Among the NILs, SSI ranged between 0.22 and 1.67, while STI had a range of 0.22 to 1.13. The corresponding averages were 0.975 and 0.686, respectively. Similarly, the average MP of NILs was 20.65g with a range of 12.39g to 26.83g. The MP of the parents was 16.55g for Pusa 44 and 18.77g for IR81896-B-B-142. Relative stress tolerance (TOL) indicates the drought impact on yield, irrespective of its magnitude. In this study, the TOL for Pusa 44 was 11.35 while that of the tolerant (donor) parent was 7.18. Lower values of TOL indicates less reduction of yield under stress as compared to yield under normal unstressed condition. Among the NILs, 29 lines had lower TOL values than Pusa 44, while 16 among these outperformed IR81896-B-B-142 with lower TOL values. The lowest TOL of 1.45 was recorded for P1823-12-114, followed by P1823-12-127 with 2.10 and P1823-12-130 with 2.44. Based on agronomic performance, grain quality, and drought indices, the number of selected NILs was narrowed down and the final best selections came out to be 14 NILs.

Correlation Analyses

Correlation analyses were carried out to explore the correspondence between yield, drought indices, and phenomics parameters. When all the 35 NILs were used for computing the relations between grain yield and phenomics parameters, we could strike only a few significant correlations. This could be due to difference among the NILs for phenomics traits. However, when NILs were shortlisted based on the distinct drought tolerance responses, significant correlations emerged from the selected NILs. Initially 14 NILs were shortlisted based on the phenomics data which was subsequently reduced to nine based on both and field and phenomics data. Significant correlations could be noticed between grain yields from field and pot experiments (**Table 5**).

Correlations between the TOL and grain yield and TOL and SSI values showed the expected pattern. While grain yield and STI indicated a strong negative correlation with TOL, a positive association was found between SSI and TOL. The stage-wise correlations worked out between the field-based yield and phenomics parameters also showed significance. Positive correlations for PSA could be observed with yield in both pot and field experiments, particularly at stages after stress. This was also reflected in the significant negative correlations of PSA with tolerance indices such as TOL and SSI. Although not apparent, positive significant correlations with grain yield were also exhibited by WU at various stages. The correlation for NIR

TABLE 3 | The average performance of NILs and their relative tolerance under stress in the pot experiment under drought stress and unstressed conditions.

Genotypes	Unstressed				Stressed				RDY	RT
	PH	PB	NT	YP	PH	PB	NT	YP		
P1823-12-4	84.0^{cd}	58.6^{c}	13.0^{b}	13.0^{bc}	86.0^{ab}	61.1^{a}	12.0^{b}	7.5^{d}	0.42	73.3
P1823-12-21	76.0^{k}	50.2^{f-h}	10.0^{e}	7.5^{m}	74.0^{ij}	47.0^{e-h}	9.0^{e}	6.0^{ef}	0.20	25.0
P1823-12-23	82.0^{ef}	53.4^{de}	10.0^{e}	10.0^{hi}	74.0^{ij}	47.0^{e-h}	9.0^{e}	4.0^{ij}	0.60	150.0
P1823-12-32	82.0^{ef}	47.8^{h-l}	11.0^{d}	8.0^{lm}	79.0^{de}	40.7^{kl}	10.0^{d}	6.5^{e}	0.19	23.1
P1823-12-36	80.0^{gh}	77.2^{a}	14.0^{a}	10.5^{gh}	70.0^{mn}	51.3^{c}	13.0^{a}	6.5^{e}	0.38	61.5
P1823-12-42	83.0^{de}	40.9st	11.0^{d}	7.5^{m}	71.0^{lm}	37.9^{mn}	10.0^{d}	5.5^{fg}	0.27	36.4
P1823-12-44	76.0^{k}	42.7^{p-s}	11.0^{d}	8.0^{lm}	65.0^{q}	37.3^{mn}	10.0^{d}	6.0^{ef}	0.25	33.3
P1823-12-48	84.0^{cd}	48.7^{g-j}	12.0^{c}	10.5^{gh}	74.0^{ij}	32.6^{q}	10.0^{d}	4.0^{ij}	0.62	163.0
P1823-12-49	85.0^{bc}	49.7^{f-j}	12.0^{c}	14.0^{a}	75.0^{hi}	48.7^{def}	13.0^{a}	8.5^{bc}	0.39	64.7
P1823-12-50	79.0^{hi}	42.7^{p-s}	11.0^{d}	13.5^{ab}	77.0^{fg}	32.1^{q}	9.0^{e}	9.0^{ab}	0.33	50.0
P1823-12-63	84.0^{cd}	43.1^{o-s}	10.0^{e}	11.5^{ef}	71.0^{lm}	45.7^{ghi}	11.0^{c}	4.5^{hi}	0.61	156.0
P1823-12-64	85.0^{bc}	47.2^{i-m}	13.0^{b}	9.0^{jk}	71.0^{lm}	40.3^{l}	12.0^{b}	3.0^{k}	0.67	200.0
P1823-12-65	85.0^{bc}	45.3^{l-p}	13.0^{b}	11.0^{fg}	76.0^{gh}	39.4^{lm}	12.0^{b}	3.0^{k}	0.73	267.0
P1823-12-77	82.0^{ef}	49.8^{f-i}	12.0^{c}	8.5^{kl}	81.0^{c}	40.9^{kl}	12.0^{b}	6.0^{ef}	0.29	41.7
P1823-12-79	85.0^{bc}	45.5^{l-o}	11.0^{d}	9.0^{jk}	80.0^{cd}	44.8^{hij}	13.0^{a}	8.0^{cd}	0.11	12.5
P1823-12-80	75.0^{kl}	32.1^{u}	10.0^{e}	7.5^{m}	77.0^{fg}	44.1^{ij}	12.0^{b}	4.0^{ij}	0.47	87.5
P1823-12-81	86.0^{b}	53.9^{de}	13.0^{b}	9.5^{ij}	85.0^{b}	54.5^{b}	13.0^{a}	7.5^{d}	0.21	26.7
P1823-12-82	81.0^{fg}	55.8^{d}	11.0^{d}	11.5^{ef}	79.0^{de}	33.6^{pq}	10.0^{d}	8.5^{bc}	0.26	35.3
P1823-12-84	83.0^{de}	54.3^{de}	12.0^{c}	9.0^{jk}	73.0^{jk}	45.5^{ghi}	10.0^{d}	5.0^{gh}	0.44	80.0
P1823-12-89	71.0^{m}	48.6^{g-j}	12.0^{c}	9.5^{ij}	69.0^{no}	49.3^{cde}	13.0^{a}	9.5^{a}	0.00	0.0
P1823-12-96	84.0^{cd}	45.2^{m-p}	12.0^{c}	11.0^{fg}	73.0^{jk}	50.5^{cd}	13.0^{a}	8.0^{cd}	0.27	37.5
P1823-12-98	78.0^{ij}	38.7^{t}	12.0^{c}	12.5^{cd}	78.0^{ef}	38.6^{lm}	12.0^{b}	8.5^{bc}	0.32	47.1
P1823-12-104	70.0^{mn}	42.6^{q-s}	12.0^{c}	10.0^{hi}	76.0^{gh}	46.0^{ghi}	13.0^{a}	9.0^{ab}	0.10	11.1
P1823-12-114	81.0^{fg}	44.5^{n-r}	12.0^{c}	12.0^{de}	80.0^{cd}	37.7^{mn}	11.0^{c}	5.5^{fg}	0.54	118.0
P1823-12-118	86.0^{b}	45.0^{m-q}	10.0^{e}	8.5^{kl}	73.0^{jk}	54.9^{b}	10.0^{d}	5.5^{fg}	0.35	54.5
P1823-12-120	78.0^{ij}	64.9^{b}	14.0^{a}	12.0^{de}	71.0^{lm}	46.6^{fgh}	12.0^{b}	9.5^{a}	0.21	26.3
P1823-12-122	80.0^{gh}	44.1^{n-r}	11.0^{d}	9.0^{jk}	80.0^{cd}	42.6^{jk}	9.0^{e}	8.5^{bc}	0.06	5.9
P1823-12-123	81.0^{fg}	42.3^{rs}	11.0^{d}	11.5^{ef}	77.0^{fg}	36.0^{no}	9.0^{e}	9.0^{ab}	0.22	27.8
P1823-12-124	76.0^{k}	52.1^{ef}	12.0^{c}	12.5^{cd}	71.0^{lm}	39.2^{lm}	10.0^{d}	8.0^{cd}	0.36	56.3
P1823-12-127	78.0^{ij}	42.6^{q-s}	12.0^{c}	8.0^{lm}	75.0^{hi}	44.7^{hij}	10.0^{d}	4.5^{hi}	0.44	77.8
P1823-12-130	75.0^{kl}	48.2^{g-k}	12.0^{c}	8.5^{kl}	68.0^{op}	40.3^{l}	10.0^{d}	6.5^{e}	0.24	30.8
P1823-12-132	86.0^{b}	42.9^{o-s}	11.0^{d}	8.5^{kl}	80.0^{cd}	48.8^{def}	9.0^{e}	6.5^{e}	0.24	30.8
P1823-12-134	80.0^{gh}	48.8^{g-j}	11.0^{d}	7.5^{m}	74.0^{ij}	42.8^{jk}	11.0^{c}	6.0^{ef}	0.20	25.0
P1823-12-141	69.0^{n}	50.7^{fg}	12.0^{c}	8.5^{kl}	72.0^{kl}	34.0^{opq}	12.0^{b}	3.6^{jk}	0.58	136.0
P1823-12-143	74.0^{l}	45.1^{m-q}	13.0^{b}	9.0^{jk}	81.0^{c}	34.9^{op}	12.0^{b}	7.5^{d}	0.17	20.0
Pusa 44	83.0^{de}	45.7^{k-n}	13.0^{b}	8.0^{lm}	78.0^{ef}	35.8^{nop}	11.0^{c}	2.3^{l}	0.71	256.0
IR81896-B-B-142	97.0^{a}	47.1^{j-m}	11.0^{d}	8.5^{kl}	87.0^{a}	47.2^{efg}	12.0^{b}	4.3^{hi}	0.49	100.0
SED	0.9	1.3	0.17	0.3	0.8	1.1	0.2	0.3	0.03	11.0
CD (0.05)	1.8	2.6	0.35	0.6	1.7	2.3	0.5	0.7	0.06	22.3
CV%	6.7	15.9	9.2	19.0	6.6	15.7	12.8	32.1		

PH, plant height in cm; PB, plant biomass in g; NT, number of tilers; YPP, yield per plant in g; means suffixed with same letters are statistically not different by Tukey's honestly significant different test.

and TR was negative for grain yield but was not pronounced as that of PSA and WU. Similarly, strong positive correlations could be observed between TR and two drought indices, TOL and SSI. Since TOL and SSI are correlated positively and strongly (.90), both provided equal opportunity in assessing the drought-tolerant status of the NILs. Correlations of agronomic parameters under field evaluation showed no correspondence due to the different performance of NILs under both unstressed and stressed treatments (**Supplementary Table 7**).

DISCUSSION

Drought is a complex stress to breed against, since its nefarious effects are inconsistent across genotypes and environments. Drought occurs when plants suffer from inadequate uptake of water to meet the water loss through transpiration. Although there can be several factors that incite drought, the most common ones in the agricultural systems are scarcity of irrigation water and irregularity in precipitation coupled with

TABLE 4 | Drought indices and yield performance of Pusa 44 *qDTY* NILs in field experiment.

Entry	Yp	Ys	TOL	MP	GMP	SSI	STI	YI	YSI
Pusa44	22.2^{wx}	10.9^{s}	11.4^{d}	16.6^{r}	15.6^{r}	1.6^{ab}	0.4^{pq}	0.7^{t}	0.5^{tu}
IR81896-B-B-142	22.4^{v-x}	15.2^{mno}	7.2^{hij}	18.8^{no}	18.4^{o-q}	1.0^{fgh}	0.6^{klm}	0.9^{mno}	0.7^{k-n}
P1823-12-4	25.2^{nop}	21.1^{ab}	4.0^{nop}	23.2^{fgh}	23.1^{f-h}	0.5^{o}	0.9^{d}	1.3^{ab}	0.8^{b}
P1823-12-21	19.5^{z}	15.7^{lmn}	3.7^{op}	17.6^{pqr}	17.5^{p-r}	0.6^{no}	0.5^{mno}	1.0^{lm}	0.8^{bc}
P1823-12-23	29.0^{e-i}	21.7^{a}	7.3^{hi}	25.3^{bc}	25.1^{bc}	0.8^{kl}	1.0^{b}	1.3^{a}	0.8^{e-h}
P1823-12-32	29.6^{c-f}	14.0^{opq}	15.6^{ab}	21.8^{ij}	20.4^{ij}	1.6^{a}	0.7^{gh}	0.9^{opq}	0.5^{u}
P1823-12-36	28.7^{e-j}	18.6^{d-g}	10.0^{def}	23.6^{ef}	23.1^{ef}	1.1^{fg}	0.9^{d}	1.1^{e-h}	0.7^{mno}
P1823-12-42	29.6^{cde}	19.9^{bcd}	9.7^{f}	24.8^{cd}	24.3^{cd}	1.0^{fgh}	1.0^{b}	1.2^{bcd}	0.7^{lmn}
P1823-12-44	25.8^{mn}	16.4^{j-m}	9.4^{f}	21.1^{jk}	20.6^{g-i}	1.1^{ef}	0.7^{g}	1.0^{kl}	0.6^{nop}
P1823-12-48	24.0^{o-u}	16.6^{l-m}	7.5^{gh}	20.3^{klm}	19.9^{k-m}	1.0^{hi}	0.7^{ghij}	1.0^{jkl}	0.7^{j-m}
P1823-12-49	27.6^{i-l}	21.2^{ab}	6.4^{h-k}	24.4^{cde}	24.2^{c-e}	0.7^{k-n}	1.0^{bc}	1.3^{a}	0.8^{c-g}
P1823-12-50	22.9^{s-w}	17.2^{h-k}	5.7^{j-m}	20.1^{k-n}	19.8^{k-m}	0.8^{kl}	0.7^{gij}	1.0^{ij}	0.8^{e-h}
P1823-12-63	20.8^{yz}	15.5^{mn}	5.3^{k-n}	18.2^{opq}	18.0^{p-r}	0.8^{jkl}	0.5^{l-n}	0.9^{lmn}	0.7^{f-i}
P1823-12-64	30.7^{cd}	14.0^{opq}	16.7^{a}	22.3^{ghi}	20.7^{g-i}	1.7^{a}	0.7^{g}	0.8^{o-r}	0.5^{u}
P1823-12-65	29.6^{c-g}	20.8^{abc}	8.8^{fg}	25.2^{bc}	24.8^{bc}	0.9^{hij}	1.0^{b}	1.3^{abc}	0.7^{i-l}
P1823-12-77	25.3^{no}	11.5^{rs}	13.8^{c}	18.4^{nop}	17.1^{p-r}	1.7^{a}	0.5^{no}	0.7^{t}	0.5^{u}
P1823-12-79	33.6^{a}	18.7^{d-g}	15.0^{bc}	26.2^{ab}	25.1^{ab}	1.4^{cd}	1.0^{b}	1.1^{d-h}	0.6^{rs}
P1823-12-80	24.3^{o-s}	14.5^{nop}	9.8^{ef}	19.4^{l-n}	18.7^{no}	1.2^{de}	0.6^{jkl}	0.9^{nop}	0.6^{pqr}
P1823-12-81	29.5^{c-h}	18.3^{e-h}	11.2^{de}	23.9^{def}	23.3^{d-f}	1.2^{e}	0.9^{cd}	1.1^{f-i}	0.6^{opq}
P1823-12-82	24.5^{n-r}	19.3^{def}	5.2^{k-n}	21.9^{ij}	21.8^{ij}	0.7^{lmn}	0.8^{ef}	1.2^{def}	0.8^{cde}
P1823-12-84	16.5^{z}	11.6^{rs}	4.9^{l-o}	14.01^{s}	13.8^{s}	0.9^{hij}	0.3^{r}	0.7^{s}	0.7^{i-l}
P1823-12-89	27.8^{ijk}	13.2^{pq}	14.6^{bc}	20.5^{kl}	19.1^{k-m}	1.6^{a}	0.6^{ijk}	0.8^{qrs}	0.5^{u}
P1823-12-96	21.4^{xy}	17.1^{h-l}	4.3^{no}	19.2^{m-o}	19.1^{k-m}	0.6^{mno}	0.6^{ijk}	1.0^{ijk}	0.8^{bcd}
P1823-12-98	24.2^{o-t}	19.6^{cde}	4.6^{mno}	21.9^{ij}	21.8^{ij}	0.6^{no}	0.8^{ef}	1.2^{de}	0.8^{bc}
P1823-12-104	17.5^{z}	12.7^{qr}	4.8^{l-o}	15.1^{s}	14.9^{r}	0.8^{ijk}	0.4^{qr}	0.8^{rs}	0.7^{h-k}
P1823-12-114	20.0^{yz}	18.5^{d-h}	1.5^{q}	19.2^{m-o}	19.2^{k-m}	0.2^{p}	0.6^{h-k}	1.1^{e-h}	0.9^{a}
P1823-12-118	31.0^{bc}	16.1^{klm}	14.9^{bc}	23.5^{ef}	22.3^{ef}	1.5^{bc}	0.8^{de}	1.0^{kl}	0.5^{st}
P1823-12-120	32.2^{ab}	21.5^{a}	10.7^{de}	26.83^{a}	26.3^{a}	1.0^{fgh}	1.1^{a}	1.3^{a}	0.7^{lmn}
P1823-12-122	22.6^{u-x}	17.7^{g-j}	4.9^{l-o}	20.1^{k-m}	20.0^{jk}	0.7^{lmn}	0.7^{ghi}	1.1^{hij}	0.8^{c-f}
P1823-12-123	23.8^{p-v}	14^{opq}	9.8^{ef}	18.9^{no}	18.3^{o-q}	1.3^{d}	0.6^{klm}	0.8^{o-r}	0.6^{qr}
P1823-12-124	25.1^{n-q}	19.1^{d-g}	6.0^{i-l}	22.1^{h-j}	21.9^{h-j}	0.7^{klm}	0.8^{ef}	1.2^{d-g}	0.8^{d-h}
P1823-12-127	20.0^{yz}	17.9^{f-i}	2.1^{q}	19.0^{no}	18.9^{no}	0.3^{p}	0.6^{ijkl}	1.1^{ghi}	0.9^{ab}
P1823-12-130	22.2^{wx}	19.8^{bcd}	2.4^{pq}	21.0^{jk}	21.0^{jk}	0.3^{p}	0.7^{fg}	1.2^{cde}	0.9^{ab}
P1823-12-132	26.8^{klm}	19.6^{cde}	7.2^{hi}	23.2^{fg}	23.0^{fg}	0.8^{jk}	0.9^{d}	1.2^{de}	0.7^{g-j}
P1823-12-134	20.4^{yz}	14.0^{opq}	6.4^{hijk}	17.2^{qr}	16.9^{qr}	1.0^{ghi}	0.5^{nop}	0.9^{opq}	0.7^{j-m}
P1823-12-141	16.8^{z}	8.0^{t}	8.8^{fg}	12.4^{t}	11.6^{s}	1.6^{ab}	0.2^{s}	0.5^{u}	0.5^{tu}
P1823-12-143	20.5^{yz}	13.4^{pq}	7.2^{hij}	17.0^{r}	16.6^{qr}	1.1^{fg}	0.5^{op}	0.9^{p-s}	0.7^{m-o}
SED	0.73	0.56	0.66	0.56	0.56	0.07	0.04	0.03	0.02
CD(0.05)	1.48	1.13	1.34	1.13	1.13	0.13	0.07	0.07	0.04

Y_{us}, yield per plant (g) under unstress; Y_s, yield under stress; TOL, relative stress tolerance; MP, mean productivity; GMP, geometric mean productivity; SSI, stress susceptibility index; SI, stress index; STI, stress tolerance index; YI, yield index; means suffixed with same letters are statistically not different by Tukey's honestly significant different test.

high temperature. Unlike other stresses, drought's impact on the plant community is wide and ruinous. When drought strikes, every plant in the community perceives drought in varying magnitude, effects of which can be seen commonly in agricultural fields as dynamic phenotypic expressions. Therefore, phenotyping for drought response is a cumbersome task that requires a large area and population, with timebound activities and destructive sampling, resulting in compromised precision. Besides, various intervening factors such as unpredicted rains and poor infrastructure, contribute more to unexplained variance. To address this problem, controlled environment facilities such as rainout shelters are often advocated in drought-related studies, wherein the interference of unpredicted rains can be avoided. However, rainout shelters are huge and require a lot of manual interventions to operate, still do not offer the expected precision in phenotyping, and involve destructive sampling. Although expensive, phenomics platforms are highly improvised facilities equipped with precision instruments to image and measure plants as they grow in a controlled environment glasshouse. Unlike field-based screening systems where plants grow in a static position, in the phenomics platform individual plants are dynamically positioned in specialized pots which can be

TABLE 5 | Correlations between grain yield, phenomics parameters, and major tolerance indices among the selected NILs.

Stages	Yield parameters		35 NILs				14 NILs					9 Selected NILs			
			PSA	WU	NIR	TR	PSA	WU	NIR	TR	TOL	PSA	WU	NIR	TR
I	PY_f	**PY_p**	−0.06	0.02	−0.16	−0.18	0.13	0.17	−0.15	−0.10	0.40*	0.82*	0.63*	−0.52	0.16
II		0.40*	−0.08	−0.06	−0.22	−0.23	0.32	0.18	−0.22	−0.04		0.89*	0.50	−0.52	−0.65*
III			0.10	0.17	−0.22	0.00	0.55*	0.57*	−0.22	−0.42		0.79*	0.41	−0.35	−0.26
IV			0.05	0.17	−0.17	−0.20	0.48*	0.40*	−0.04	0.02		0.80*	0.47	−0.36	−0.48
V			−0.07	0.10	0.00	−0.13	0.46*	0.46*	−0.25	−0.22		0.77*	0.50	−0.32	−0.50
I	YP_f	**YP_p**	0.06	−0.07	−0.09	−0.13	0.54*	−0.10	−0.05	−0.26	−0.51*	0.57	0.39	−0.46	0.21
II		0.52*	0.06	−0.20	−0.08	−0.17	0.20	−0.08	−0.34	−0.06		0.67*	0.36	−0.38	−0.70*
III			0.27	0.08	−0.06	0.03	0.56*	0.37	−0.44	−0.32		0.55	0.43	−0.18	−0.08
IV			0.21	0.00	0.03	−0.03	0.44	0.16	−0.37	−0.22		0.54	0.45	−0.28	−0.39
V			0.12	−0.03	0.02	−0.25	0.39	0.39	−0.28	−0.39		0.60*	0.50	−0.37	−0.51
I	YP_p		0.06	−0.07	−0.09	−0.13	0.55*	0.49*	−0.32	−0.11	–	0.43	0.20	−0.32	0.24
II			0.06	−0.20	−0.08	−0.17	0.56*	0.18	−0.17	−0.40		0.44	−0.11	−0.39	−0.65*
III			0.27	0.08	−0.06	0.03	0.71*	0.38	−0.07	−0.03		0.67*	0.17	−0.32	0.11
IV			0.21	0.00	0.03	−0.03	0.67*	0.38	0.04	−0.10		0.66*	0.08	−0.19	0.33
V			0.12	−0.03	0.02	−0.25	0.68*	0.46*	0.11	−0.04		0.61*	0.27	−0.41	−0.29
I	TOL		0.04	0.13	0.11	0.04	−0.46*	−0.09	0.28	0.03	–	−0.32	−0.09	0.29	0.01
II			0.11	0.44*	0.21	0.07	−0.45	0.15	0.28	0.18		−0.31	0.17	0.50	0.23
III			−0.14	0.21	0.23	0.12	−0.62*	−0.06	0.24	0.13		−0.53	−0.04	0.53	0.11
IV			−0.11	0.39*	0.14	0.22	−0.60*	0.07	0.14	0.42		−0.48	0.14	0.41	0.33
V			0.08	0.30*	−0.02	0.12	−0.54*	−0.11	0.09	0.32		−0.35	−0.04	0.39	0.21
I	SSI		0.08	−0.11	−0.05	−0.22	−0.32	−0.11	0.10	−0.32	0.90*	−0.33	−0.11	0.09	−0.35
II			0.15	0.35*	−0.02	0.09	−0.33	0.07	0.07	0.22		−0.32	0.16	0.17	0.29
III			−0.03	0.07	0.00	0.16	−0.49*	−0.02	0.03	0.21		−0.40	0.27	0.17	0.44
IV			−0.03	0.07	−0.05	0.21	−0.48*	0.08	−0.03	0.49*		−0.38	0.40	0.08	0.65*
V			0.13	0.02	−0.06	0.13	−0.47*	−0.08	0.16	0.42		−0.30	0.23	0.32	0.41

*YP, yield per plant in g; PY, plot yield in kg/ha; TOL, relative stress tolerance from the field; SSI, stress susceptibility index; STI; PSA, projected shoot area; WU, water use; NIR, near-infrared intensity; TR, transpiration rate; stages I-V are the imaging stage points from the phenomics platform. Stage I, 1 DAI (days after irrigation); stage II, 5DAI; stage III, 9DAI with peak drought stress, irrigated on the next day; stage IV, 2DAR (days after recovery); stage V, 5DAR; * significant at 5% level. The suffixes f and p indicate field and pot culture, respectively.*

moved and randomized to provide uniform growing conditions. Besides, the plants travel and are continuously phenotyped, thereby providing unparalleled precision, high throughput, and non-invasive sampling.

Adapted to a semi-aquatic environment, the rice crop consumes about 2,500 liters of water to produce one kilogram of rough rice (Bouman, 2009). Reducing the water supply to rice crops from its normal levels has a penalty attached because moisture stress can cause yield loss up to 70% under severe deprivation (Kumaraswamy and Shetty, 2016). However, genotypic differences do exist in the level of yield loss, which can be translated into the development of drought-tolerant rice varieties. A tolerant genotype can endure drought by yielding grains, where a sensitive cultivar normally fails. Recently, there are several QTLs identified in rice related to grain yield under drought stress, named with the prefix qDTY (Vinod et al., 2019). Currently, several of these QTLs are being integrated into mainstream cultivars for augmenting them with drought tolerance. In this study, we have used a set of NILs developed from Pusa 44, a mega-scale cultivar of northern India, harboring a combination of two qDTYs, *qDTY2.1,* and *qDTY3.1* (Dwivedi et al., 2021; Oo et al., 2021a).

There are already reports of the usage of image-based parameters for assessment of plant health under different stress conditions like drought, heat, salinity, etc. in various crops (Jones et al., 2002; Roy et al., 2011; Yonemaru and Morita, 2012; Hairmansis et al., 2014; Campbell et al., 2015; Humplík et al., 2015). However, the use of these parameters for assessing drought tolerance has been seldom experimented in rice. In this study, we have used 35 improved NILs of Pusa 44 to assess the feasibility of deploying a phenomics platform for drought tolerance assessment. NILs being genetically near-uniform, provide a greater opportunity to examine subtle response variations conditioned by the integrated QTLs/ genes in comparison with the recurrent parent. Therefore, we have attempted to compare the image-based crop response parameters from the phenomics with field performance, to identify best performing NILs. This has provided us the opportunity to assess the critical relationship between the phenomics parameters and grain yield under drought.

Comparing the parental response to induced drought, we found that the donor line, IR81896-B-B-142 had higher values of PSA and WU over Pusa 44. IR81896-B-B-142 possessed

long and broad leaves than Pusa 44 with a tall plant stature. Pusa 44 had higher NIR intensity values and TR indicating its inclination to lose internal water due to the elevated TR. Primarily, these observations indicated the tendency of Pusa 44 to fail under drought. As the drought progressed, there was a decline observed among all the traits, except for NIR intensity which showed a reverse trend. This was because NIR increased with reducing internal water, while PSA, WU, and TR showed reduction as the drought progressed. However, the pattern of PSA decline was found to be different in some of the NILs, which showed little variation in the PSA over the drought period. Coincidently, those NILs that showed a lesser decline in PSA also showed better grain yield. It should be inferred that plants that showed lesser shrinkage have a better mechanism to conserve internal water status, thereby being able to produce better yield (Pandey and Shukla, 2015). This made us rely upon PSA as a selection parameter, because it showed a consistent association with grain yield among the NILs that produced higher yield under drought in both field and pot culture experiments. PSA is a measure of biomass obtained from RGB images taken from three dimensions, top and two side views (Honsdorf et al., 2014). Under the drought treatment, the trend of PSA decline showed a sharp decrease until stage III which corresponded to the onset of drought. This is corroborated from the PSA pattern under unstressed conditions which remained almost the same between the imaging stages with only statistically insignificant differences. Reduction in PSA indicated that the plant biomass was shrinking due to the loss of water. However, this decline reflected only a transient architectural change rather than a physical biomass loss. Visual imaging captures drought-induced leaf rolling and wilting in the plants and predicts lowering PSA. This is the reason why the PSA started to revert during stages IV and V following the irrigation. Therefore, in this study, PSA decline recorded a temporal response to drought, indicating that image-based phenotyping was capable of capturing even minute variations in plant architecture. If a prolonged drought is provided, the PSA can accurately reflect actual biomass loss. A similar response for PSA has been reported earlier corresponding to progressive drought stress (Mir et al., 2019; Kim et al., 2020). Leaf rolling under water stress is an adaptive mechanism by plants to minimize water loss from leaves that helps in survival. During this process, stomatal closure occurs to reduce transpiration loss of water (Kim et al., 2020). This helps the plants to use internal water judiciously and to maintain higher relative water content and to decrease leaf drying. Maintenance of water homeostasis in plants requires lowering water loss while maintaining a steady uptake from the soil, although at a reduced pace.

Other parameters such as NIR intensity, WU, and TR could also be used for further evaluation of lines. There are earlier reports of the use of these parameters to assess drought tolerance in crops including rice (Gupta et al., 2012; Deshmukh et al., 2014; Kumar et al., 2015; Malinowska et al., 2017; Kim et al., 2020). Plant responses in favor of tolerance could be perceived from reduced TR and NIR values. While TR translates directly to the water loss from the leaves, NIR indicated the hydration status of the plants. Hydrated plants generally show lower gray values for the NIR intensity, which increases along with drying up as drought progresses. Near-IR spectrometry is a recognized tool for determining leaf water content in plants (Zhang et al., 2012; Jin et al., 2017). Water use which indicates the internal water replenishment status also showed a positive trend with grain yield among the high yielding NILs. The TR also showed a decreasing trend as the drought progressed, concordantly to the internal water depletion. With an increase in stress, mean values for PSA began to decrease at a lesser rate than changes in WU and TR, indicating that changes in PSA occur after the water loss. Therefore, a deep decline in any of these traits indicated a sensitive reaction, as the tolerant genotypes are expected to have a slower change. Further, tolerant genotypes showed quicker recovery than the sensitive ones, reverting to the original values following irrigation. In this study, we have seen that NILs responded rapidly for re-watering and entered into the recovery phase than the recurrent parent. However, we could observe an elevation in the NIR values as the recovery progressed perhaps due to the plants entering their maturity.

Although all the NILs possessed the same combinations of QTLs, *qDTY2.1* and *qDTY3.1*, there were variations among them for the degree of drought tolerance. However, these variations remained latent under field evaluation, since the final yield was the major trait used to evaluate the tolerance. Whereas in the phenomics evaluation, instead of the outcome, dynamic observation of the crop response was made under drought, which could unfurl subtle variations in the plant response. This was the reason for the absence of a meaningful correlation between phenomics variables and grain yield when the data from all the 35 NILs were used for computations. However, when the NILs were shortlisted based on the desired response pattern based on phenomics variables, the correlations became apparent and robust. Phenotypic deviations among the NILs introgressed with the same QTLs are not an uncommon feature in maker-assisted introgression breeding because of the transmission of untargeted donor loci. Although such segments may not produce perceptible phenotypic deviation from the recurrent parent, they can indulge in interactions with the background genome resulting in varying expression levels of the introgressed QTLs. Background interactions from the small introgressed segments have been reported to influence target gene/QTL expression under varying environments (Li et al., 2020).

We could establish that the NILs used in this study showed improved adaptation to drought than the recipient parent, Pusa 44, but with varying degrees of advantage. The phenomics parameters could, however, dissect these variations into a specific pattern with PSA, WU, TR, and NIR striking a significant correlation with grain yield. Among the varying patterns, the most desirable combination was a flatter curve for the PSA as seen in P1823-12-21, higher WU as found in P1823-12-82 coupled with low TR and NIR values. This combination can be used as selection criteria for identifying superior drought tolerant NILs from the phenomics platform.

Furthermore, considering the individual traits, PSA and NIR could be directly used for selection. Based on this criterion, a final set of nine NILs such as P1823-12-23, P1823-12-32, P1823-12-42, P1823-12-48, P1823-12-64, P1823-12-65, P1823-12-104, P1823-12-124, and P1823-12-141 were selected. These outstanding NILs performing better under both phenomics-based screening and field screening could be used for varietal development.

The pattern of stage-wise correlation of phenomics parameters with yield provided another clue. It was found that a high positive correlation with grain yield got stronger at stage III of the phenomics screening, particularly for the PSA. Stage III had the peak drought, following which the plant has entered into the recovery phase after lifesaving irrigation. Therefore, it must be understood that the best correlation with the grain yield under drought was obtained by the phenomics parameters recorded at the peak drought stage. The strength of correlation was found to go weak at the recovery phase. This gave us the indication of the criticality of the stage at which comparison is to be made for selection purposes. At stage I, wherein no statistically validated differences were available between stressed and unstressed treatments, the association of the PSA with yield was weak, which may be due to the lack of apparent drought response at this stage. As the drought started to show its effect on PSA beginning from stage II, the correlation got stronger and reached its maximum at stage III when the drought was at its peak. At the recovery phase, the strength of association began to revert to a lower magnitude. This implied that the drought response such as leaf rolling and wilting could predict the drought tolerance response among the NILs. However, in this study, we did not have enough data proving this hypothesis but it is worth examining in a large NIL set with several plants. From the observations herein, we suggest that the phenomics parameters at peak drought stage (stage III in this case) could be considered critical for the selection of drought-tolerant lines.

Several drought indices were used in previous works targeted to identify drought-tolerant lines/ genotypes in many crops (Babu et al., 2011; Dadbakhsh et al., 2011; Farshadfar and Elyasi, 2012; Naghavi et al., 2013; Ali and El-Sadek, 2016; Bennani et al., 2017; Mau et al., 2019), since a simultaneous selection for yield and tolerance indices could be a good selection criterion (Raman et al., 2012; Garg and Bhattacharya, 2017; Oo et al., 2021a). As different indices were having separate response emphasis, it would be prudent to use multiple indices for a comprehensive evaluation of drought tolerance (Oo et al., 2021a). Because of this reason, we have used multiple drought indices in the present study and found that indices such as TOL and SSI showed a good correlation with image-based parameters. This provided a further implication of the use of tolerant indices along with phenomics parameters and yield. No previous study comparing the drought indices and phenomics parameters has been reported earlier. Based on the field performance, we could select a final set of nine NILs. These NILs can be further evaluated for national testing for cultivar release and can be used as parental lines for drought tolerance breeding.

CONCLUSION

A set of 35 NILs of Pusa 44 was simultaneously evaluated for drought response in a field evaluation as well as in a phenomics platform under induced drought. The field tolerance of the lines was confirmed from two years of evaluation and compared with the image-based phenomics traits such as PSA, WU, NIR, and TR. From the significant association of these traits, particularly PSA at peak drought, we found that phenomics-based traits can be directly used as selection criteria for identifying drought-tolerant lines. In addition, the phenomics evaluation provided an early-stage non-invasive method that can be carried on continuously and scaled to a large population, and automated. This opens a novel vista in the evaluation and selection of rice breeding lines based on various physiological parameters. Additionally, if scaled up, this method could be labor-saving, cost-effective, and accurate as well as could overcome the limitations of conventional drought screening methods such as using rainout shelters. Rapid and high throughput phenotyping is the need of the hour, for scaling up the breeding programs to genomic selection pipelines as well as in accelerating genetic gain. Besides, accelerated approaches provide promise to generate climate-resilient cultivars to ensure future food security, especially under the changing climate and water limitations.

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

SG and VC: conceptualization. SG, VC, MP, and DR: methodology. KV, PB, RE, and SG: software. PD and GD: validation. PD, GD, KV, and SG: formal analysis. PD, PB, HB, SG, KV, VC, and DR: investigation. SG, AS, MP, MN, and DR: resources. PD, KV, and SG: data curation, visualization, and writing – original draft preparation. KV and SG: writing – review and editing. SG, AS, NR, VC, and DR: supervision. SG, MP, and AS: project administration. SG, MP, and VC: funding acquisition. All authors have read and agreed to the published version of the manuscript.

FUNDING

This research was funded by the ICAR-funded network project on National Innovations on Climate Resilient Agriculture (NICRA), the National Agricultural Science Fund (NASF) project on "Phenomics of Moisture Deficit Stress Tolerance and Nitrogen Use Efficiency in Rice and Wheat – Phase II," and the National Agricultural Higher Education Project (NAHEP) – Centre for Advanced Agricultural Science and Technology (CAAST)

project on "Genomics Assisted Breeding for Crop Improvement," of the World Bank and the Indian Council of Agricultural Research, New Delhi.

ACKNOWLEDGMENTS

The study is part of the Ph.D. thesis research of PD and PD wishes to acknowledge ICAR-Indian Agricultural Research Institute, New Delhi, and Indian Council of Agriculture Research (ICAR) for providing all the necessary facilities.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2021.752730/full#supplementary-material

REFERENCES

Ali, M. B., and El-Sadek, A. (2016). Evaluation of drought tolerance indices for wheat (*Triticum aestivum* L.) under irrigated and rainfed conditions. *Commun. Biometry Crop Sci.* 11, 77–89.

Babu, C. M., Hemlatha, T., and Naik, B. R. (2011). Comparison of remote sensing based indices for drought monitoring in Anantapur. *Int. J. Appl. Res.* 2, 449–456.

Bennani, S., Nsarellah, N., Jlibene, M., Tadesse, W., Birouk, A., and Ouabbou, H. (2017). Efficiency of drought indices under different severities for bread wheat selection. *Aust. J. Crop Sci.* 11, 395–405.

Bernier, J., Kumar, A., Serraj, R., Spaner, D., and Atlin, G. (2008). Review: breeding upland rice for drought resistance. *J. Sci. Food Agric.* 88, 927–939. doi: 10.1002/jsfa.3153

Bouman, B. (2009). How much water does rice use? *Rice Today* 8, 28–29.

Bouslama, M., and Schapaugh, W. T. (1984). Stress tolerance in soybean, Part I: evaluation of three screening techniques for heat and drought tolerance. *Crop Sci. J.* 24, 933–937. doi: 10.2135/cropsci1984.0011183X002400050026x

Campbell, M. T., Knecht, A. C., Berger, B., Brien, C. J., Wang, D., and Walia, H. (2015). Integrating image-based phenomics and association analysis to dissect the genetic architecture of temporal salinity responses in rice. *Plant Physiol.* 168, 1476–1489. doi: 10.1104/pp.15.00450

Chen, D., Neumann, K., Friedel, S., Kilian, B., Chen, M., Altmann, T., et al. (2014). Dissecting the phenotypic components of crop plant growth and drought responses based on high-throughput image analysis. *Plant Cell* 26, 4636–4655. doi: 10.1105/tpc.114.129601

Chen, S., Guo, Y., Sirault, X., Stefanova, K., Saradadevi, R., Turner, N. C., et al. (2019). Nondestructive phenomic tools for the prediction of heat and drought tolerance at anthesis in *Brassica* species. *Plant Phenomics* 2019, 3264872. doi: 10.34133/2019/3264872

Courtois, B., McLaren, G., Sinha, P. K., Prasad, K., Yadav, R., and Shen, L. (2000). Mapping QTLs associated with drought avoidance in upland rice. *Mol. Breed.* 6, 55–66. doi: 10.1023/A:1009652326121

DACFW (2017). Annual Report 2016-17. Department of Agriculture, Cooperation and Farmers Welfare. Government of India. 188p. Available online at: www.agricoop.nic.in (accessed November 5, 2021).

Dadbakhsh, A., Yazdansepas, A., and Ahmadizadeh, M. (2011). Study drought stress on yield of wheat (*Triticum aestivum* L.) genotypes by drought tolerance indices. *Adv. Environ. Biol.* 5, 1804–1810.

Danzi, D., Briglia, N., Petrozza, A., Summerer, S., Povero, G., Stivaletta, A., et al. (2019). Can high throughput phenotyping help food security in the mediterranean area? *Front. Plant Sci.* 10:15. doi: 10.3389/fpls.2019.00015

Deery, D., Jimenez-Berni, J., Jones, H., Sirault, H., and Furbank, R. (2014). Proximal remote sensing buggies and potential applications for field-based phenotyping. *Agronomy* 4, 349–379. doi: 10.3390/agronomy4030349

Deery, D. M., Rebetzke, G. J., Jimenez-Berni, J. A., James, R. A., Condon, A. G., Bovill, W. D., et al. (2016). Methodology for high-throughput field phenotyping of canopy temperature using airborne thermography. *Front. Plant Sci.* 7:1808. doi: 10.3389/fpls.2016.01808

Deshmukh, R., Sonah, H., Patil, G., Chen, W., Prince, S., Mutava, R., et al. (2014). Integrating omic approaches for abiotic stress tolerance in soybean. *Front. Plant Sci.* 5:244. doi: 10.3389/fpls.2014.00244

Dhawan, G., Kumar, A., Dwivedi, P., Gopala Krishnan, S., Pal, M., Vinod, K. K., et al. (2021). Introgression of qDTY1.1governing reproductive stage drought tolerance into an elite basmati rice variety "Pusa Basmati 1" through marker assisted backcross breeding. *Agronomy* 11:202. doi: 10.3390/agronomy11020202

Duan, L., Han, J., Guo, Z., Tu, H., Yang, P., Zhang, D., et al. (2018). Novel digital features discriminate between drought resistant and drought sensitive rice under controlled and field conditions. *Front. Plant Sci.* 9:492. doi: 10.3389/fpls.2018.00492

Dwivedi, P., Ramawat, N., Dhawan, G., Gopala Krishnan, S., Vinod, K. K., Singh, M. P., et al. (2021). Arought tolerant near isogenic lines (NILs) of Pusa 44 developed through marker assisted introgression of *qDTY2.1* and *qDTY3.1* enhances yield under reproductive stage drought stress. *Agriculture* 11:64. doi: 10.3390/agriculture11010064

Farshadfar, E., and Elyasi, P. (2012). Screening quantitative indicators of drought tolerance in bread wheat (*T. aestivum*) landraces. *Eur. J. Exp. Biol.* 2, 577–584.

Fernandez, G. C. J. (1992). "Effective selection criteria for assessing plant stress tolerance," in *Proceeding of 4th International Symposium on Adaptation of Food Crop Temperature and Water Stress*, ed. C. G. Kuo (Tainan: Asian Vegetable and Research and Development Center), 257–270.

Fischer, R. A., and Maurer, R. (1978). Drought resistance in spring wheat cultivars. I. Grain yield responses. *Aust. J. Agric. Res.* 29, 892–912. doi: 10.1071/AR9780897

Furbank, R. T., and Tester, M. (2011). Phenomics—technologies to relieve the phenotyping bottleneck. *Trends Plant Sci.* 16, 635–644. doi: 10.1016/j.tplants.2011.09.005

Garg, H. S., and Bhattacharya, C. (2017). Drought tolerance indices for screening some of rice genotypes. *IJABR* 7, 671–674.

Guo, Z., Yang, W., Chang, Y., Ma, X., Tu, H., Xiong, F., et al. (2018). Genome-wide association studies of image traits reveal genetic architecture of drought resistance in rice. *Mol. Plant* 11, 789–805. doi: 10.1016/j.molp.2018.03.018

Gupta, P. K., Balyan, H. S., Gahlaut, V., and Kulwal, P. L. (2012). Phenotyping, genetic dissection, and breeding for drought and heat tolerance incommon wheat: status and prospects. *Plant Breed. Rev.* 36, 85–168.

Hairmansis, A., Berger, B., Tester, M., and Roy, S. J. (2014). Image-based phenotyping for non-destructive screening of different salinity tolerancetraits in rice. *Rice* 7, 1–10. doi: 10.1186/s12284-014-0016-3

Honsdorf, N., March, T. J., Berger, B., Tester, M., and Pillen, K. (2014). High-throughput phenotyping to detect drought tolerance QTL in wild barley introgression lines. *PLoS One* 9:e97047. doi: 10.1371/journal.pone.0097047

Hoover, D. L., Wilcox, K. R., and Young, K. E. (2018). Experimental droughts with rainout shelters: a methodological review. *Ecosphere* 9:e02088. doi: 10.1002/ecs2.2088

Humplík, J. F., Lazár, D., Husičková, A., and Spíchal, L. (2015). Automated phe- notyping of plant shoots using imaging methods for analysis ofplant stress responses - a review. *Plant Methods* 11:29. doi: 10.1186/s13007-015-0072-8

IRRI (2014). *STAR- Statstical Tool for Agricultural Research Version 2.0.1. Biometrics and Breeding Informatics, PBGB Division*. Los Baños: International Rice Research Institute.

Jin, X., Shi, C., Yu, C. Y., Yamada, T., and Sacks, E. J. (2017). Determination of leaf water content by visible and near-infrared spectrometry and multivariate calibration in Miscanthus. *Front. Plant Sci.* 8:721. doi: 10.3389/fpls.2017.00721

Jones, H., Serraj, R., Loveys, B. R., Xiong, L., Wheaton, A., and Price, A. H. (2009). Thermal infrared imaging of crop canopies for the remote diagnosis and quantification of plant responses to water stress in the field. *Funct. Plant Biol.* 36, 978–989. doi: 10.1071/FP09123

Jones, H. G., Stoll, M., Santos, T., de Sousa, C., Chaves, M. M., and Grant, O. M. (2002). Use of infrared thermography for monitoring stomatal closure in the field: application to grapevine. *J. Exp. Bot.* 53, 2249–2260. doi: 10.1093/jxb/erf083

Kim, S. L., Kim, N., Lee, H., Lee, E., Cheon, K. S., Kim, M., et al. (2020). High-throughput phenotyping platform for analyzing drought tolerance in rice. *Planta* 252:38. doi: 10.1007/s00425-020-03436-9

Kumar, A., Dixit, S., Ram, T., Yadaw, R. B., Mishra, K. K., and Mandal, N. P. (2014). Breeding high-yielding drought-tolerant rice genetic variations and conventional and molecular approaches. *J. Exp. Bot.* 65, 6265–6278. doi: 10.1093/jxb/eru363

Kumar, S., Dwivedi, S. K., Singh, S. S., Jha, S. K., Lekshmy, S., Elanchezhian, R., et al. (2014). Identification of drought tolerant rice genotypes by analysing drought tolerance indices and morpho-physiological traits. *SABRAO J. Breed. Genet.* 46, 217–230.

Kumar, A., Verulkar, S. B., Mandal, N. P., Variar, M., Shukla, V. D., Dwivedi, J. L., et al. (2012). High-yielding, drought-tolerant, stable rice genotypes for the shallow rainfed lowland drought-prone ecosystem. *Field Crops Res.* 133, 37–47. doi: 10.1016/j.fcr.2012.03.007

Kumar, J., Kumar, S., and Pratap, A. (2015). "Plant phenomics: an overview," in *Phenomics in Crop Plants: Trends, Options and Limitations*, eds J. Kumar, A. Pratap, and S. Kumar (Berlin: Springer International), 1–10. doi: 10.1007/978-81-322-2226-2_1

Kumar, R. (2018). Development of drought resistance in rice. *Int. J. Curr. Microbiol. Appl. Sci.* 7, 1439–1456. doi: 10.20546/ijcmas.2018.705.171

Kumaraswamy, S., and Shetty, P. (2016). Critical abiotic factors affecting implementation of technological innovations in rice and wheat production: a review. *Agric. Rev.* 37, 268–278.

Lafitte, R. H., Ismail, A. M., and Bennett, J. (2006). Abiotic stress tolerance in tropical rice: progress and the future. *Oryza* 43, 171–186.

Li, D., Quan, C., Song, Z., Li, X., Yu, G., Li, C., et al. (2020). High-throughput plant phenotyping platform (HT3P) as a novel tool for estimating agronomic traits from the lab to the field. *Front. Bioeng. Biotechnol.* 8:623705. doi: 10.3389/fbioe.2020.623705

Lin, C. S., Binns, M. R., and Lefkovitch, L. P. (1986). Stability analysis: where do we stand? *Crop Sci.* 26, 894–900. doi: 10.2135/CROPSCI1986.0011183X002600050012X

Lippman, Z. B., and Zamir, D. (2007). Heterosis: revisiting the magic. *Trend Genet.* 23, 60–66. doi: 10.1016/j.tig.2006.12.006

Malinowska, M., Donnison, I. S., and Robson, P. R. H. (2017). Phenomics analysis of drought responses in *Miscanthus* collected from differentgeographical locations. *GCB Bioenergy* 9, 78–91. doi: 10.1111/gcbb.12350

Mau, Y. S., Ndiwa, A. S. S., Oematan, S. S., and Markus, J. E. R. (2019). Drought tolerance indices for selection of drought tolerant, high yielding upland rice genotypes. *AJCS* 13, 170–178. doi: 10.21475/ajcs.19.13.01.p1778

Mir, R. R., Reynolds, M., Pinto, F., Khan, M. A., and Bhat, M. A. (2019). High-throughput phenotyping for crop improvement in the genomics era. *Plant Sci.* 282, 60–72. doi: 10.1016/j.plantsci.2019.01.007

Naghavi, M. R., Pour-Aboughadareh, A. R., and Khalili, M. (2013). Evaluation of drought tolerance indices for screening some of corn (*Zea mays* L.) cultivars under environmental conditions. *Not. Sci. Biol.* 5, 388–393. doi: 10.15835/nsb539049

Neilson, E. H., Edwards, A. M., Blomstedt, C. K., Berger, B., Møller, B. L., and Gleadow, R. M. (2015). Utilization of a high-throughput shoot imaging system to examine the dynamic phenotypic responses of a c4 cereal crop plant to nitrogen and water deficiency over time. *J. Exp. Bot.* 66:1817. doi: 10.1093/jxb/eru526

Oladosu, Y., Rafii, M. Y., Samuel, C., Fatai, A., Magaji, U., Kareem, I., et al. (2019). Drought resistance in rice from conventional to molecular breeding: a review. *Int. J. Mol. Sci.* 20:3519. doi: 10.3390/ijms20143519

Oo, K. S., Krishnan, S. G., Vinod, K. K., Dhawan, G., Dwivedi, P., Kumar, P., et al. (2021a). Molecular breeding for improving productivity of *Oryza sativa* L. cv. Pusa 44 under reproductive stage drought stress through introgression of a major QTL, *qDTY12.1*. *Genes* 12:967.

Oo, K. S., Dhawan, G., Pankaj, Prakash, V., Krishnan, S. G., Bhowmick, P. K., et al. (2021b). Development and evaluation of Pusa 44 backcross derived lines possessing qDTY2.1, qDTY3.1 and their combinations indicate differential yield response under severe reproductive stage drought stress. *Indian J. Genet. Plant Breed.* 81, 221–235. doi: 10.31742/IJGPB.81.2.5

Pandey, V., and Shukla, A. (2015). Acclimation and tolerance strategies of rice under drought stress. *Rice Sci.* 22, 147–161. doi: 10.1016/S1672-6308(14)60289-4

Parent, B., Shahinnia, F., Maphosa, L., Berger, B., Rabie, H., Chalmers, K., et al. (2015). Combining field performance with controlled environment plant imaging to identify the genetic control of growth and transpiration underlying yield response to water-deficit stress in wheat. *J. Exp. Bot.* 66, 5481–5492. doi: 10.1093/jxb/erv320

Peñuelas, J., and Filella, L. (1998). Visible and near-infrared reflectance techniques for diagnosing plant physiological status. *Trends in Plant Sci.* 3, 151–156. doi: 10.1016/S1360-1385(98)01213-8

Raman, A., Verulkar, S., Mandal, N., Variar, M., Shukla, V. D., Dwivedi, J. L., et al. (2012). Drought yield index to select high yielding rice lines under different drought stress severities. *Rice* 5:31. doi: 10.1186/1939-8433-5-31

Rischbeck, P., Cardellach, P., Mistele, B., and Schmidhalter, U. (2017). Thermal phenotyping of stomatal sensitivity in spring barley. *J. Agron. Crop Sci.* 203, 483–493. doi: 10.1111/jac.12223

Rosielle, A. A., and Hamblin, J. (1981). Theoretical aspects of selection for yield in stress and non-stress environment. *Crop Sci.* 21, 943–946. doi: 10.2135/cropsci1981.0011183X002100060033x

Roy, S. J., Tucker, E. J., and Tester, M. (2011). Genetic analysis of abiotic stresstolerance in crops. *Curr. Opin. Plant Biol.* 14, 232–239. doi: 10.1016/j.pbi.2011.03.002

Sandhu, N., Dixit, S., Swamy, B. P. M., Vikram, P., Venkateshwarlu, C., Catolos, M., Kumar, A., et al. (2018). Positive interactions of major-effect QTLs with genetic background that enhances rice yield under drought. *Sci. Rep.* 8, 1626. doi: 10.1038/s41598-018-20116-7

Shirdelmoghanloo, H., Taylor, J. D., Lohraseb, I., Rabie, H., Brien, C., Timmins, A., et al. (2016). A QTL on the short arm of wheat (*Triticum aestivum* L.) chromosome 3B affects the stability of grain weight in plants exposed to a brief heat shock early in grain filling. *BMC Plant Biol.* 16:100. doi: 10.1186/s12870-016-0784-6

Singh, R., Singh, Y., Xalaxo, S., Verulkar, S., Yadav, N., Singh, S., et al. (2016). From QTL to variety-harnessing the benefits of QTLs for drought, flood and salt tolerance in mega rice varieties of India through a multi-institutional network. *Plant Sci. J.* 242, 278–287. doi: 10.1016/j.plantsci.2015.08.008

Singh, V. J., Vinod, K. K., Krishnan, S. G., and Singh, A. K. (2021). "Rice adaptation to climate change: opportunities and priorities in molecular breeding," in *Molecular Breeding for Rice Abiotic Stress Tolerance and Nutritional Quality,* eds M. A. Hossain, L. Hassan, K. M. Ifterkharuddaula, A. Kumar, and R. Henry (Wiley-Blackwell), 1–25. doi: 10.1002/9781119633174.ch1

Tester, M., and Langridge, P. (2010). Breeding technologies to increase crop production in a changing world. *Science* 327, 818–822. doi: 10.1126/science.1183700

Tuberosa, R. (2012). Phenotyping for drought tolerance of crops in the genomics era. *Front. Physiol.* 3:347. doi: 10.3389/fphys.2012.00347

Ubbens, J. R., and Stavness, I. (2017). Deep plant phenomics: a deep learning platform for complex plant phenotyping tasks. *Front. Plant Sci.* 8:1190. doi: 10.3389/fpls.2017.01190

Underwood, J., Wendel, A., Schofield, B., McMurray, L., and Kimber, R. (2017). Efficient in-field plant phenomics for row-crops with an autonomous ground vehicle. *J. Field Robot.* 34, 1061–1083. doi: 10.1002/rob.21728

Venuprasad, R., Dalid, C. O., Del Valle, M., Zhao, D., Espiritu, M., Cruz, M. T. S., et al. (2009). Identification and characterization of large-effect quantitative trait loci for grain yield under lowland drought stress in rice using bulk-segregant analysis. *Theor. Appl. Genet.* 120, 177–190. doi: 10.1093/jhered/esm109

Vinod, K. K., Krishnan, S. G., Thribhuvan, R., and Singh, A. K. (2019). "Genetics of drought tolerance, candidate genes and their utilization in rice improvement," in *Genomics Assisted Breeding of Crops for Abiotic Stress Tolerance, Sustainable Development and Biodiversity 21*, Vol. 2, ed. V. R. Rajpal (Cham: Springer), doi: 10.1007/978-3-319-99573-1_9

Yang, W., Guo, Z., Huang, C., Wang, K., Jiang, N., Feng, H., et al. (2015). Genome-wide association study of rice (*Oryza sativa* L.) leaf traits with a high-throughput leaf scorer. *J. Exp. Bot.* 66, 5605–5615. doi: 10.1093/jxb/erv100

Yonemaru, J. I., and Morita, S. (2012). Image analysis of grain shape to eval-uate the effects of high temperatures on grain filling of rice, *Oryza sativa* L. *Field Crop Res.* 137, 268–271. doi: 10.1016/j.fcr.2012.08.003

Zhang, Q., Li, Q., and Zhang, G. (2012). Rapid determination of leaf water content using VIS/NIR spectroscopy analysis with wavelength selection. *J. Spectrosc.* 27, 276795. doi: 10.1155/2012/276795

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

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 24 February 2022
doi: 10.3389/fpls.2022.829337

High-Throughput Sequencing-Based Analysis of Rhizosphere and Diazotrophic Bacterial Diversity Among Wild Progenitor and Closely Related Species of Sugarcane (*Saccharum* spp. Inter-Specific Hybrids)

Mukesh Kumar Malviya[1†], *Chang-Ning Li*[1†], *Prakash Lakshmanan*[1,2,3], *Manoj Kumar Solanki*[4], *Zhen Wang*[5], *Anjali Chandrol Solanki*[6], *Qian Nong*[1], *Krishan K. Verma*[1], *Rajesh Kumar Singh*[1], *Pratiksha Singh*[1], *Anjney Sharma*[1], *Dao-Jun Guo*[1,7], *Eldessoky S. Dessoky*[8], *Xiu-Peng Song*[1*] and *Yang-Rui Li*[1,7*]

[1] Key Laboratory of Sugarcane Biotechnology and Genetic Improvement (Guangxi), Ministry of Agriculture and Rural Affairs, Sugarcane Research Center, Chinese Academy of Agricultural Sciences, Guangxi Key Laboratory of Sugarcane Genetic Improvement, Sugarcane Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, China, [2] Interdisciplinary Research Center for Agriculture Green Development in Yangtze River Basin, College of Resources and Environment, Southwest University, Chongqing, China, [3] Queensland Alliance for Agriculture and Food Innovation, University of Queensland, St. Lucia, QLD, Australia, [4] Plant Cytogenetics and Molecular Biology Group, Faculty of Natural Sciences, Institute of Biology, Biotechnology and Environmental Protection, University of Silesia in Katowice, Katowice, Poland, [5] Guangxi Key Laboratory of Agricultural Resources Chemistry and Biotechnology, College of Biology and Pharmacy, Yulin Normal University, Yulin, China, [6] Department of Agriculture Science, Mansarovar Global University, Bhopal, India, [7] College of Agriculture, Guangxi University, Nanning, China, [8] Department of Biology, College of Science, Taif University, Taif, Saudi Arabia

Considering the significant role of genetic background in plant-microbe interactions and that most crop rhizospheric microbial research was focused on cultivars, understanding the diversity of root-associated microbiomes in wild progenitors and closely related crossable species may help to breed better cultivars. This study is aimed to fill a critical knowledge gap on rhizosphere and diazotroph bacterial diversity in the wild progenitors of sugarcane, the essential sugar and the second largest bioenergy crop globally. Using a high-throughput sequencing (HTS) platform, we studied the rhizosphere and diazotroph bacterial community of *Saccharum officinarum* L. cv. Badila (BRS), *Saccharum barberi* (*S. barberi*) Jesw. cv Pansahi (PRS), *Saccharum robustum* [*S. robustum;* (RRS), *Saccharum spontaneum* (*S. spontaneum*); SRS], and *Saccharum sinense* (*S. sinense*) Roxb. cv Uba (URS) by sequencing their 16S rRNA and *nifH* genes. HTS results revealed that a total of 6,202 bacteria-specific operational taxonomic units (OTUs) were identified, that were distributed as 107 bacterial groups. Out of that, 31 rhizobacterial families are commonly spread in all five species. With respect to *nifH* gene, *S. barberi* and *S. spontaneum* recorded the highest and lowest number of OTUs, respectively. These results were validated by quantitative PCR analysis of both genes. A total of 1,099 OTUs were

OPEN ACCESS

Edited by:
Freddy Mora-Poblete, University of Talca, Chile

Reviewed by:
Peifang Zhao, Yunnan Academy of Agricultural Sciences, China
Jay Prakash Verma, Banaras Hindu University, India

***Correspondence:**
Xiu-Peng Song
xiupengsong@gxaas.net
Yang-Rui Li
liyr@gxaas.net
orcid.org/0000-0002-7559-9244

[†]These authors have contributed equally to this work

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

Received: *05 December 2021*
Accepted: *04 January 2022*
Published: *24 February 2022*

Citation:
Malviya MK, Li C-N, Lakshmanan P, Solanki MK, Wang Z, Solanki AC, Nong Q, Verma KK, Singh RK, Singh P, Sharma A, Guo D-J, Dessoky ES, Song X-P and Li Y-R (2022) High-Throughput Sequencing-Based Analysis of Rhizosphere and Diazotrophic Bacterial Diversity Among Wild Progenitor and Closely Related Species of Sugarcane (Saccharum spp. Inter-Specific Hybrids). Front. Plant Sci. 13:829337. doi: 10.3389/fpls.2022.829337

identified for diazotrophs with a core microbiome of 9 families distributed among all the sugarcane species. The core microbiomes were spread across 20 genera. The increased microbial diversity in the rhizosphere was mainly due to soil physiochemical properties. Most of the genera of rhizobacteria and diazotrophs showed a positive correlation, and few genera negatively correlated with the soil properties. The results showed that sizeable rhizospheric diversity exists across progenitors and close relatives. Still, incidentally, the rhizosphere microbial abundance of progenitors of modern sugarcane was at the lower end of the spectrum, indicating the prospect of *Saccharum* species introgression breeding may further improve nutrient use and disease and stress tolerance of commercial sugarcane. The considerable variation for rhizosphere microbiome seen in *Saccharum* species also provides a knowledge base and an experimental system for studying the evolution of rhizobacteria-host plant association during crop domestication.

Keywords: rhizosphere soil, sugarcane, *nifH*, 16S rRNA, diazotroph diversity

INTRODUCTION

Sugarcane is an important agricultural crop grown in nearly 110 countries worldwide. China is the third-largest producer of sugarcane (a collective term for *Saccharum* species, but more commonly applied to cultivated *Saccharum officinarum* (*S. officinarum*) L. and *Saccharum* spp. inter-specific hybrids). It is a major crop in southern China, accounting for ≈90% of Chinese sugar production (Li and Yang, 2015). Over the years, sugarcane has been developed as a multi-purpose agro-industrial crop as it provides the raw material for different industries, such as food, thermal, energy/fuel, and paper (Goldemberg et al., 2008; Fischer et al., 2012). Sugarcane is mainly grown as a monoculture for extended periods resulting in yield decline, which is attributed to degraded soil, imbalanced soil biology, and build-up of pests and diseases (Shoko et al., 2007). Restoration of soil biology and soil fertility is now emerging as a priority for improving soil health, reducing the yield gap, and sustaining profitable green agriculture (Brackin et al., 2013; Schultz et al., 2017).

The rhizosphere is rich in microbial diversity and abundance. These functionally diverse microbial communities include saprophytes, epiphytes, pathogens, and also plant growth-promoting microorganisms (Avis et al., 2008). Bacteria are the most abundant rhizospheric microbiota, covering up to 15% of the total root surface (van Loon, 2007). About 2–5% of rhizobacteria are known to promote plant growth (Antoun and Prévost, 2005). Many plant growth-promoting rhizobacteria (PGPRs) are capable of nitrogen fixation, solubilization of inorganic molecules, such as phosphate, production of plant growth regulators/hormones, siderophores, and compounds that control phytopathogens (Cawoy et al., 2011; Jangu and Sindhu, 2011; Velineni and Brahmaprakash, 2011). Thus, PGPRs and other rhizosphere bacteria are now well-recognized as an essential component of sustainable agriculture systems. The most commonly found rhizospheric bacterial genera are *Bacillus*, *Pseudomonas*, *Rhizobia*, *Arthrobacter*, *Agrobacterium*, *Micrococcus*, *Cellulomonas*, *Azotobacter*, *Alcaligenes*, *Mycobacterium*, and *Flavbacter* (Teixeira et al., 2010; Prashar et al., 2014). There are several PGPRs and diazotrophs genera, such as *Bacillus*, *Paenibacillus*, *Pseudomonas*, *Enterobacter*, *Arthrobacter*, *Azotobacter*, *Burkholderia*, and *Azospirillum*, that are associated with sugarcane rhizosphere (Ahmad et al., 2016; Lamizadeh et al., 2016; Li et al., 2017; Malviya et al., 2019; Pereira et al., 2020; Rosa et al., 2020). The majority of PGPRs are not culturable. Thus studying those using traditional laboratory methods is challenging and time-consuming (Prashar et al., 2014; Wei et al., 2018). Rapid progress in molecular biology, especially the advent of cost-effective, high-throughput DNA sequencing technologies and the associated data analytics, has improved the understanding of rhizosphere microflora by culture-independent studies (Reuter et al., 2015; Wei et al., 2018). The next-generation sequencing approaches provide an efficient and comprehensive system to identify microbial species in the rhizosphere irrespective of microbial abundance (Uroz et al., 2013). As a result, through the sequencing of the 16S rRNA gene, remarkable progress in the taxonomic characterization of highly diversified rhizospheric bacteria has been achieved (Dong et al., 2017; Gong et al., 2019). Further, modern molecular techniques permit an in-depth analysis of soil bacterial communities' compositional and functional dynamics in changing soil environmental conditions, a recurring feature of agricultural soil (Dini-Andreote et al., 2010; Wei et al., 2018). Recently rhizosphere microbiomes community structure of three endangered Plants (Xu et al., 2020) and blueberry varieties (Wang et al., 2019) has been studied by high-throughput sequencing (HTS).

Sugarcane rhizosphere microbes need strong attention to understand their diversity and role in crop improvement. Several novel PGPRs from the sugarcane microbiome have been identified and improved crop production (Pisa et al., 2011; Schultz et al., 2017; Armanhi et al., 2018). Most of the studies on rhizosphere microbes and endophytes are limited to modern commercial sugarcane cultivars, and little is known about their occurrence and abundance in the wild sugarcane progenitor species. They are primarily involved in nitrogen fixation and plant hormone production, thus positively affecting sugarcane growth (Goldemberg et al., 2008; Mehnaz, 2011). The present

study focuses on wild sugarcane progenitor species because these species are regarded as high fibrous plants with significant geographic distributions and can survive a broad range of abiotic stresses, such as droughts, saline, floods, and freezing conditions (Santchurn et al., 2019). Rhizobacteria play a significant role in nitrogen fixation in sugarcane crops (Li et al., 2017; Malviya et al., 2019). However, much remains to be learned about these diazotrophic rhizobacteria, a key driver of soil health and fertility. There are several reports of identification and characterization of PGPRs from sugarcane (Inui-Kishi et al., 2012; Lamizadeh et al., 2016) and other crops (Kumar et al., 2014; Gaikwad and Sapre, 2015; Tsegaye et al., 2019) and some of them are being used for crop productivity improvement (Kashyap et al., 2017; Swarnalakshmi et al., 2020). Not surprisingly, most sugarcane studies are conducted with modern cultivated sugarcane hybrid varieties. There is no previous attempt to understand the diversity of rhizobacteria in their wild progenitors and closely related species, such as *S. officinarum*, *Saccharum spontaneum* (*S. spontaneum*), *Saccharum robustum* (*S. robustum*), *Saccharum barberi* (*S. barberi*), and *Saccharum sinense* (*S. sinense*). Rhizospheric microorganisms interact with plants for their survival and nutritional requirements (Berendsen et al., 2012; Finkel et al., 2017). Many of them are beneficial to plants for nutrient uptake and to cope with pathogens and abiotic stresses (Pisa et al., 2011; Dagnaw et al., 2015; Pereira et al., 2020).

Thus, the current study is aimed to understand the role of rhizosphere bacterial communities and identify new species of nitrogen-fixing bacteria using high-throughput 16S rRNA and *nifH* gene sequencing using the Illumina platform. This study reports novel and valuable findings on the diversity of bacterial communities in five *Saccharum* progenitor species, namely, *S. officinarum* L. cv Badila (BRS), *S. barberi* Jesw. cv Pansahi (PRS), *S. robustum* (RRS), *S. spontaneum* (SRS), and *S. sinense* Roxb. cv Uba (URS) and provides a knowledge base to study the influence of sugarcane genotype on rhizosphere bacteria in this necessary sugar and energy crop.

MATERIALS AND METHODS

Soil Sampling

Rhizospheric soil of five sugarcane species, BRS, PRS, RRS, SRS, and URS, was maintained in the sugarcane field, germplasm of Sugarcane Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, Guangxi, China (22°49′ N, 108°18′ E, 800–1731 masl). Climate conditions of the site were as follows: natural night (10 h), temperature (minimum 22°C to maximum 35°C), and air humidity was 75–80%. The rhizosphere soil sample of all five sugarcane species was collected at the maturing stage of the first ratoon crop. Five soil samples were separately collected from each species by shaking roots in a sterile bag for 5 min. These soil samples were pooled to prepare a composite soil sample (≈5 g) for each species. Soil DNA from three sub-samples from the composite sample of each species were extracted separately. A schematic diagram represents the overall experimental design (**Figure 1**).

Soil Physiochemical Parameters

As described by Solanki et al. (2019), soil chemical analysis activities were carried out. Soil water content was measured by the weighing method. Soil pH was analyzed by a pH meter, and soil organic carbon was measured by the dichromate oxidation method (Walkley and Black, 1934). Total nitrogen (N) was estimated through the semi-micro-Kjeldahl method (Bremmer and Mulvaney, 1982). The FeSO4/Zn reduction method was used for the estimation of nitrate-nitrogen (NO_3^-N) and ammonium nitrogen (NH_4^+N) (Carter, 1993). Total phosphorus (P) was measured via the sodium carbonate fusion method (Carter, 1993). Available P was estimated by the sample extraction method (Bao, 2002). Total potassium (K) estimation was done by the photometry method (Bao, 2002). The ammonium acetate extraction-flame photometry method was applied to detect the available K in the soil (Bao, 2002). All analysis was performed in three biological replicates.

High-Throughput Sequencing

A culture-independent method was applied in this study to identify the bacterial composition of test species (Ranjard et al., 2000). Total microbial DNA was extracted using cetyltrimethylammonium bromide/sodium dodecyl sulfate (CTAB/SDS) isolation method with minor modifications (Barbier et al., 2019). The purity of extracted DNA was assayed with a NanoDrop One spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Three DNA samples were pooled as one and used for DNA sequencing. Bacterial 16S rRNA was amplified with the universal primers 341F (ACTCCTACGGGAGGCAGCAG) and 806R (GGACTACHVGGGTWTCTAAT), which target the V3–V4 regions (Dong et al., 2018). The *nifH* gene was amplified with primers Pol-F (TGCGAYCCSAARGCBGACTC) and Pol-R (ATSGCCATCATYTCRCCGGA), as previously reported (Poly et al., 2001). PCR amplification of identified 16S rRNA and *nifH* genes was performed, and each PCR contains 25 μl reaction that included 12.5 μl ready to use PCR mix (Tiangen Biotech, Beijing, China), 1.0 μl of each primer (10 μM), 2.5 μl of DNA template (10 ng/ml), and 9.0 μl PCR grade water. PCR amplification consisted of a 3 min denaturation at 95°C; 25 cycles of 30 s at 95°C, 30 s at 55°C, and 30 s at 72°C; and 5 min at 72°C for 16S rRNA. For *nifH* gene amplification, thermal cycling consisted of initial denaturation at 98°C for 1 min, followed by 30 cycles of denaturation at 98°C for 10 s, annealing at 50°C for 30 s, and elongation at 72°C for 60 s, finally, 72°C for 5 min, following the protocols reported previously by Zhang et al. (2015). Visualization and quantification of PCR products were conducted by mixing an equal volume of 1× loading buffer containing SYB green dye with the PCR products and electrophoresed on 2% agarose gel. The 400–450 bp DNA fragments were isolated and used for further experiments. Equimolar amounts of PCR products from all samples were pooled, and the mixture was purified using Qiagen Gel Extraction Kit (Qiagen, Germany). Sequencing libraries of purified amplicons were generated using TruSeq® DNA (Illumina, San Diego, CA, USA). PCR-Free Sample Preparation Kit (Illumina, San Diego, CA, USA) in accordance with the

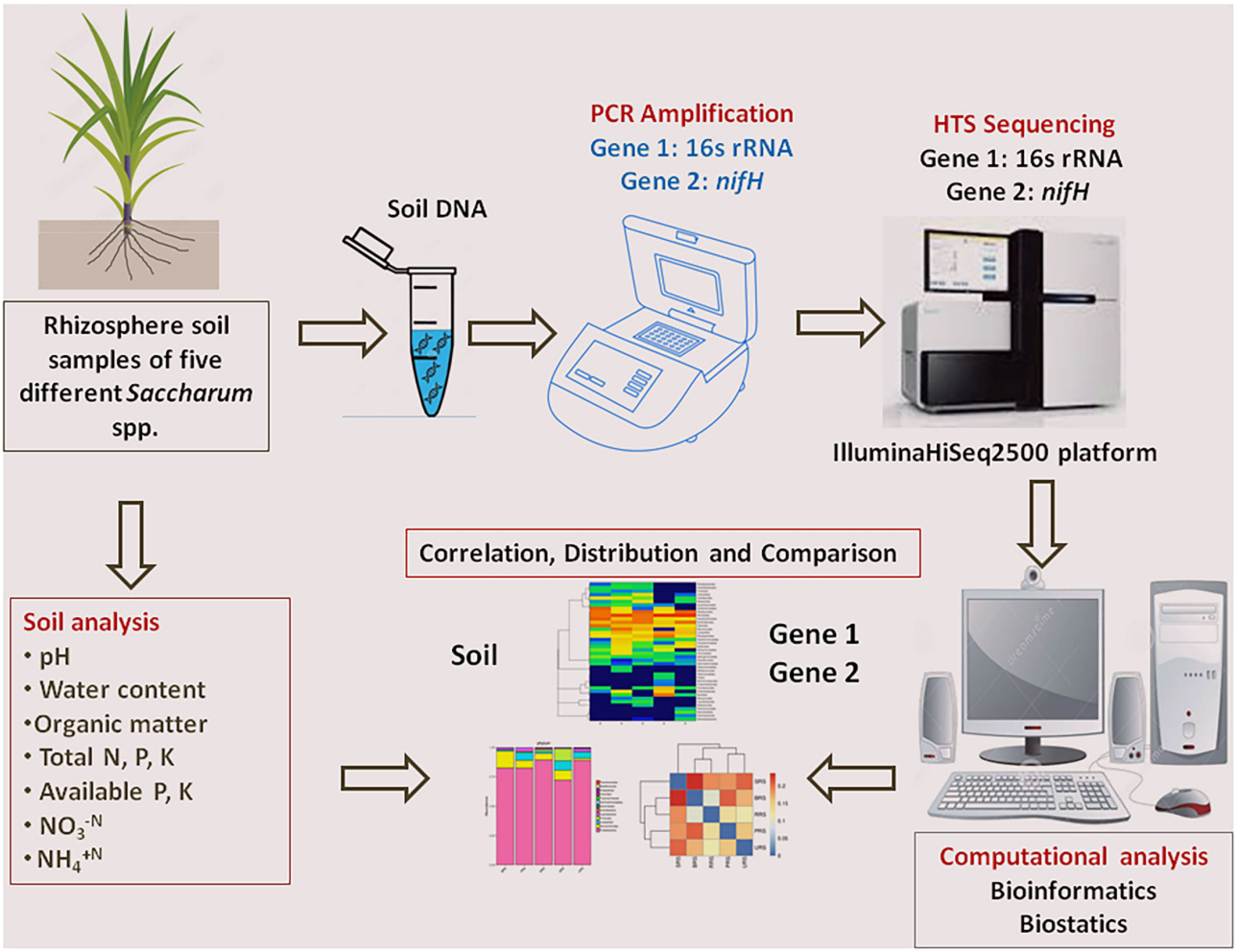


FIGURE 1 | Schematic diagram of experimental design and analysis that used in the current study. HTS, High-throughput sequencing.

manufacturer's protocol, and index codes were added. The library quality and concentration were assessed on the Qubit@ 2.0 Fluorometer (Thermo Scientific, Waltham, MA, USA) and Agilent Bioanalyzer 2100 system. To perform sequencing, the qualified libraries were fed into the IlluminaHiSeq2500 platform, and 250 bp paired-end reads were generated.

HTS Data Analysis

Based on their unique barcode, trimming of barcode, and primer sequencing were done. Reads were assembled using FLASH (V.1.2.7) (http://ccb.jhu.edu/software/FLASH/) (Magoè and Salzberg, 2011) to generate raw tags. To obtain high-quality clean tags from raw tags, we performed quality filtration using QIIME (V1.7.0) (http://qiime.org/index.html) (Caporaso et al., 2010). Removal of chimera sequences was done by comparing the tags with the reference database Unite Database (https://unite.ut.ee/) using the UCHIME algorithm (http://www.drive5.com/usearch/manual/uchime_algo.html) (Edgar et al., 2011). The above step is critical to obtain effective tags. Operational taxonomic unit (OTU) identification was done with UPARSE software (v.7.0.1001) (http://drive5.com/uparse/) (Edgar, 2013). Based on ≥95% of sequence similarity, all the effective tags were clustered into OTUs. For each OTU cluster, a representative sequence was screened to perform taxonomic annotation.

Operational Taxonomic Units were taxonomically annotated following a Basic Local Alignment Search Tool (BLAST) analysis against the Unite Database of each identified representative bacterial sequence done in QIIME software. Multiple sequence alignment was conducted with MUSCLE software (V.3.8.3) (http://www.drive5.com/muscle/) (Edgar, 2004), and phylogenetic relationship of different OTUs was established to understand the diversity of microbial species in various samples (groups). Alpha diversity analysis is carried out to find the complexity of species diversity for each sample using 6 indices, which include observed species, Chao1, Shannon, Simpson, abundance-based coverage estimator (ACE), and Good's coverage. Indices calculation for all the samples was done using QIIME and visualized in R software (V. 2.15.3). Community richness was identified with two selected indices using Chao—the Chao1 richness estimator (http://www.mothur.org/wiki/Chao).

The Shannon index (http://www.mothur.org/wiki/Shannon) and Simpson index (http://www.mothur.org/wiki/Simpson) indices were used for the identification of community diversity in all the samples. To characterize the sequencing depth and coverage, the Good's coverage (http://www.mothur.org/wiki/Coverage) was used. Beta diversity analysis was performed to evaluate the differences in bacterial species among all the samples. QIIME software was used to calculate beta diversity using weighted and unweighted UniFrac distances. The raw data of 16S rRNA (accession no. PRJNA678588) and *nifH* gene (accession no. PRJNA681283) were submitted to the NCBI Sequence Read Archive.

Quantitative (q) PCR Analysis

Quantitative PCR was performed to determine the relative abundance of the 16S rRNA and *nifH* gene in each individual rhizospheric soil. The bacterial gene copy numbers were determined with the primers 341F/518R (GC-341 F 5-CCTACGGGAGGCAGCAG-3) and 518 R (5-ATTACCGCGGCTGCTGG-3) (Moore et al., 2011) and the *nifH* gene with the primers *nifH*-F (AAAGGYGGWATCGGYAARTCCACCAC) and *nifH*-R (TTGTTSGCSGCRTACATSGCCAT CAT) (Zhang et al., 2016). Briefly, each 20 μl PCR reaction contained 1.0 μl of DNA template (2.5 ng/μl), 1.0 μl of each primer (10 μmol/l), 7.0 μl of molecular-grade water, and 10 μl iQ™ SYBR Green SuperMix (Bio-Rad Laboratories, Hercules, CA, USA). Conditions of the qPCR for 16S rRNA are as follows: initial denaturation at 95°C for 5 min; 40 cycles of denaturation at 95°C for 15 s, and annealing at 56°C for 30 s. Condition of the qPCR for *nifH* is as follows: 95°C for 30 s followed by 40 cycles of 95°C for 5 s, 60°C for 40 s. The standard curve of DNA and copy number was constructed by the standard formula: $y = -3.406 \times 37.05$ [y: Ct value; × LOG10 (copy number)]. All samples were done in three replicates.

Statistical Analysis

Using the FactoMineR and ggplot2 packages in R software, results were visualized. Later, Principal Coordinate Analysis (PCoA) was performed to obtain principal coordinates and to visualize them. To compare microbial diversity in different samples, the UniFrac method was used to generate weighted or unweighted UniFrac values among samples that were transformed to give uncorrelated/orthogonal axes. Visualization of PCoA results was done using weighted gene co-expression network analysis (WGCNA) package, stat, and ggplot2 packages in R software. QIIME software was used for hierarchical clustering by unweighted pair-group method with arithmetic mean method (UPGMA) (Sokal, 1958) to infer the distance matrix using average linkage. Heatmaps and Venn plots were generated using the package "ggplots" of software R (v3.0.3). The experiments were conducted in replicates, and data were analyzed using standard ANOVA followed by the Duncan's Multiple Range Test (DMRT) through Origin 2017SR2 software (Northampton, MA, USA). Circos plots were drawn by Circos Table Viewer v0.63-9 software (Krzywinski et al., 2009) to calculate Spearman's rank correlation coefficient between soil variables and bacterial taxa by using PAST3 software (Hammer et al., 2001), and heatmap is generated by using ClustVis online tool (Metsalu and Vilo, 2015) and Morpheus (https://software.broadinstitute.org/morpheus/).

RESULTS

This study was aimed to understand the rhizosphere soil microbiota of five different critical ancestral sugarcane species using 16S rRNA and *nifH* gene sequencing to understand their bacterial community diversity, especially that of diazotrophs. This is a major knowledge gap as the significance of rhizosphere biology has a strong genetic basis. It is now increasingly recognized as a target for crop productivity improvement in all major crops, such as sugarcane.

Soil Physiochemical Parameters

The physiochemical properties of all the rhizosphere soil samples are presented in **Figure 2**. Higher water content values were found in *S. sinense* (19.1 ± 0.38%), and the low values were found in *S. spontaneum* (14% ± 0.33). The pH range was recorded (*S. sinense*) 7.1 ± 0.14 to 5.7 ± 0.11 (*S. officinarum*). Organic matter (OM) was observed higher in *S. sinense* (16.30 ± 0.33) and lowest in *S. spontaneum* (8.29 ± 0.17). The higher amounts of total N were recorded in *S. barberi* (0.78 ± 0.02), and low quantity resulted in *S. spontaneum* (0.48 ± 0.11). Total K, available P, and available K were recorded higher in *S. officinarum* and low in *S. spontaneum*. The higher value of available NO_3^-N was observed in *S. officinarum* (24.70 ± 0.49) and the lowest value in *S. sinense* (4.43 ± 0.09). A higher concentration of available NH_4^+N was recorded in *S. sinense* (22.20 ± 0.44) and low in *S. spontaneum* (5.79 ± 0.12).

Sequencing Data

A total of 233,455 effective sequences with an average length of 413 bp were obtained for 16S rRNA genes from different sugarcane species samples. Sequencing of *nifH* genes from five different species samples produced a total of 182,185 sequences with 357 average bp length. These raw reads of 16S rRNA and nifH genes were filtered using QIIME quality filters, followed by OTU identification, clustering, and analysis, respectively (**Supplementary Tables S1, S2**).

Operational Taxonomic Units

The *S. spontaneum* sample, compared to those from all the other species, showed less species evenness at a low OTU rank. We analyzed common and unique OTUs based on 16S rRNA and *nifH* gene sequences for each sample (**Supplementary Figures S1A,B**). In the 16S rRNA sequence data, 6202 OTUs were identified from all samples collectively, of which 519 OTUs were common across all species. The relative frequency of OTUs in the studied species was as follows: *S. sinense* > *S. robustum* > *S. barberi* > *S. officinarum* > *S. spontaneum* (**Figure 3A**). A total of 1,099 OTUs were identified for *nifH* gene from the sequence data of species combined. Among them, 14 were common OTUs found across all the samples. The occurrence of OTUs in the plant samples was as follows: *S. barberi*

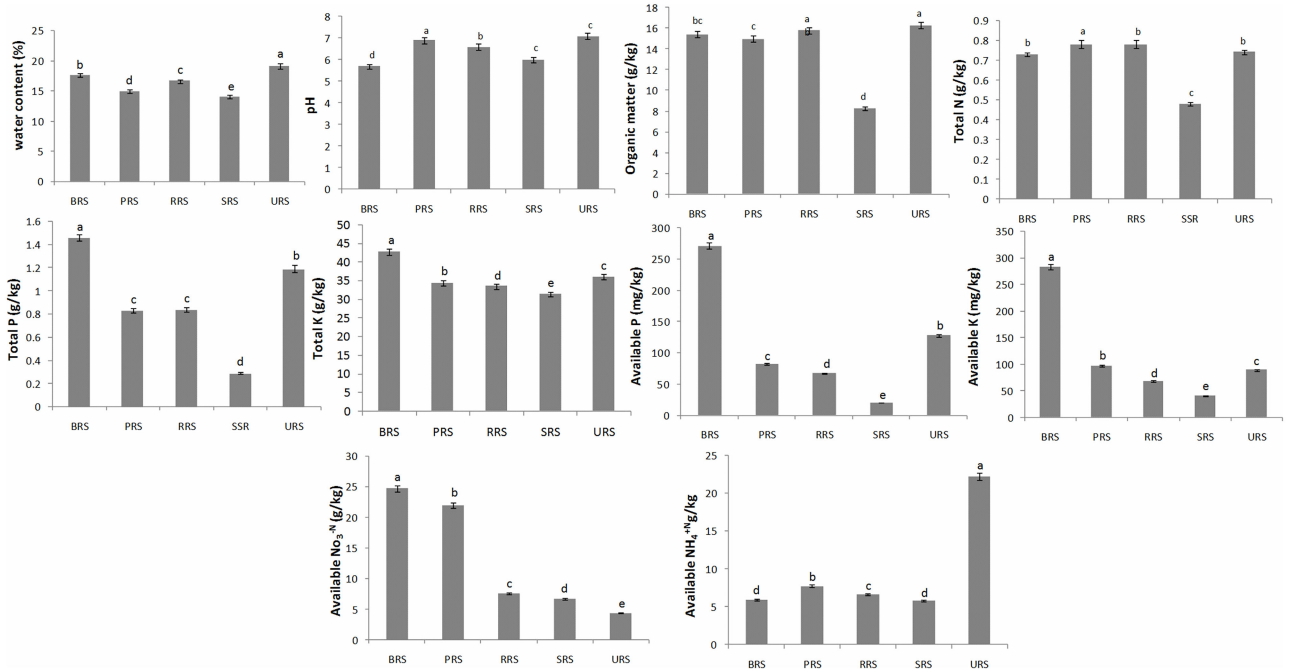


FIGURE 2 | Soil physiochemical properties of rhizospheric soil samples of sugarcane species. Values followed by the same lowercase letters in the same bar are not significantly different according to DMRT 5%. *S. officinarum* L. cv Badila (BRS), *S. barberi* Jesw. cv Pansahi (PRS), *S. robustum* (RRS), *S. spontaneum* (SRS), and *S. sinense* Roxb. cv Uba (URS). DMRT, Duncan's Multiple Range test.

> *S. robustum* > *S. officinarum* > *S. sinense* > *S. spontaneum*. Thus, *S. barberi* and *S. spontaneum* recorded the highest and lowest number of OTUs (**Figure 3B**).

Principal Component Analysis (PCA)

To understand the rhizobacterial community composition, PCA plots were generated based on 16S rRNA and *nifH* gene OTUs data (**Supplementary Figures S2A,B**). The OTUs of 16S rRNA samples show that *S. officinarum* and *S. robustum* are not identical, *S. sinense* and *S. robustum* have more similarity, and *S. spontaneum* is distinct from all the other species studied. PCA of *nifH* gene samples showed a close identity between *S. barberi* and *S. spontaneum,* whereas *S. sinense, S. officinarum,* and *S. robustum* remained distinct.

Diversity Index and Microbial Composition

Alpha diversity refers to the diversity within a particular sample individually, and it is usually represented by the microbial species (i.e., species richness) enumerated in the test samples. Alpha diversity analysis was done using Shannon, Simpson, and Chao indices Rarefaction curves for both 16S rRNA and *nifH* sequence data. **Supplementary Figure S3A** consists of plots displaying Shannon, Simpson, and Chao indices, built using 16S rRNA samples. The Shannon index was increased as both the species richness and the evenness in the community was increased. Among the 16S rRNA data of all the samples analyzed, the Shannon and Chao indices of *S. sinense* sample were higher than the other four samples, with a slight variation found among them. Whereas, Simpson indices was increased as the diversity was increased. In the Simpson index plot, *S. officinarum* samples showed the highest value, implying more species diversity than the other four samples. Shannon and Chao indices of *nifH* gene data of all the samples showed more species diversity in *S. barberi* sample than others. Like 16S rRNA data, *S. officinarum* sample showed the highest value in Simpson index based on *nifH* gene data (**Supplementary Figure S3B**).

The relative abundance of the microbial communalities was differed among the five sugarcane species analyzed. The abundant phyla identified in all sugarcane species using 16S rRNA data were Proteobacteria, Firmicutes, Actinobacteria, Acidobacter, Bacteroidetes, Chloroflexi, Gemmatimonadetes, Planctomycetes, and Nitrospirae (**Figure 4A**). Firmicutes were the highest phyla present in *S. officinarum* compared to other samples. Gemmatimonadetes were the highest in *S. barberi*, Acidobacter was more elevated in *S. spontaneum*, and Bacteroidetes was most increased in *S. sinense*. Based on *nifH* gene data, we found Proteobacteria and Verrucomicrobia (**Figure 4B**). Firmicutes were abundant in *S. spontaneum* samples compared to others. Many unclassified phyla were also represented abundantly in *S. spontaneum* followed by *S. sinense* and *S. barberi* samples.

Genus distribution using 16S rRNA sequence data is given in **Figure 5A**. *Bacillus* was the most abundant genus in *S. officinarum, S. robustum,* and *S. sinense,* while *Pseudomonas* became the number one genus in *S. barberi* and *S. spontaneum*. *Bacillus*, *Pseudomonas*, *Pseudarthobacter*, *Massilia*, *Lysobacter*, *Nitrospira*, *Gemmatimonas*, and

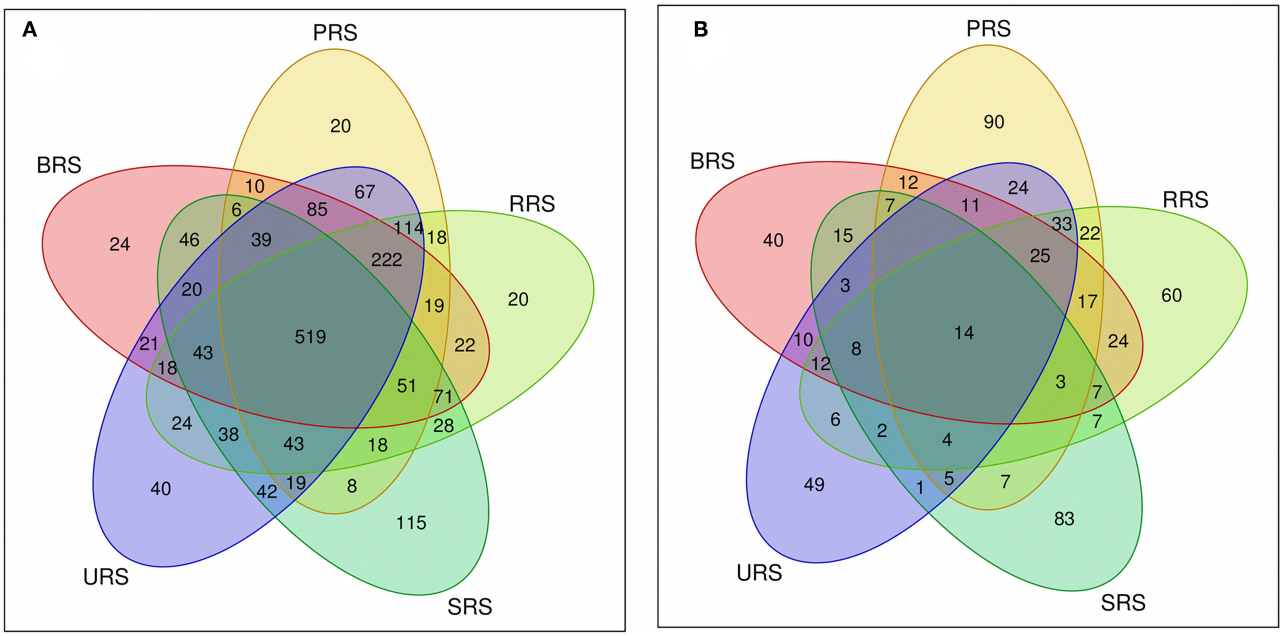


FIGURE 3 | Venn diagram showing the OTUs distribution between the sugarcane species soil samples **(A)** 16S rRNA and **(B)** *nifH* gene. *S. officinarum* L. cv Badila (BRS), *S. barberi* Jesw. cv Pansahi (PRS), *S. robustum* (RRS), *S. spontaneum* (SRS), and *S. sinense* Roxb. cv Uba (URS). OTUs, operational taxonomic units.

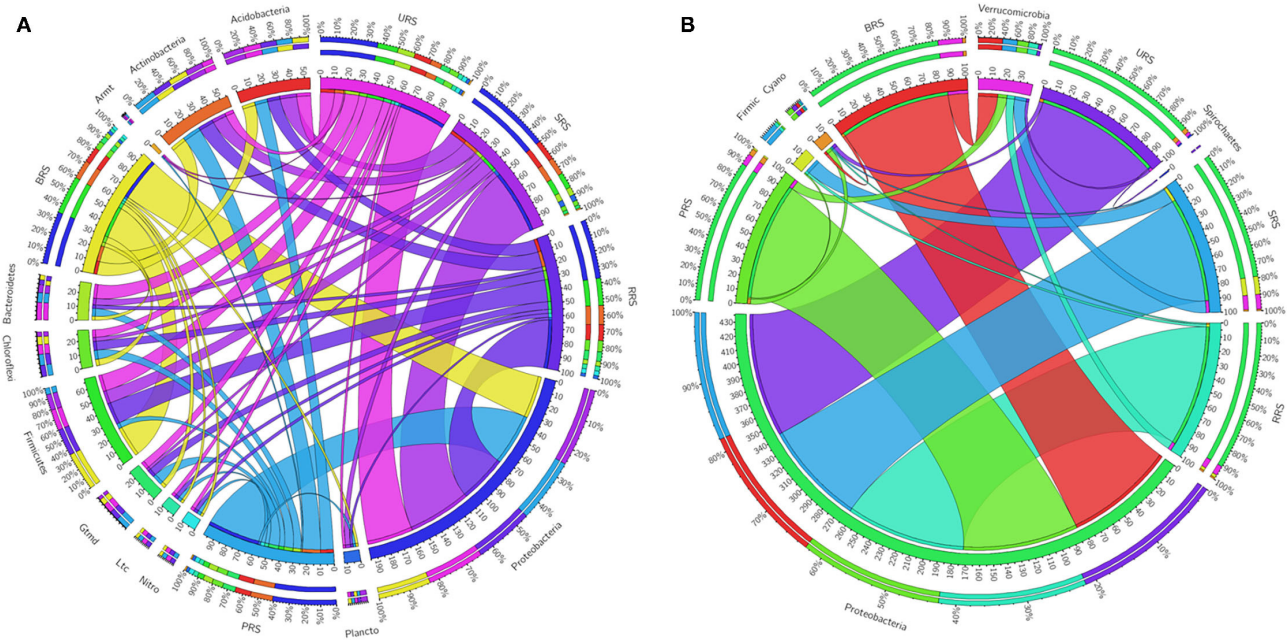


FIGURE 4 | Circular representation of the proportional structure of bacterial communities at the phylum level based on 16s rRNA **(A)** and *nifH* gene **(B)** associated with the sugarcane rhizosphere of different species. Values within the inner circle indicate the number of reads of a phylum within the normalized dataset. *S. officinarum* L. cv Badila (BRS), *S. barberi* Jesw. cv Pansahi (PRS), *S. robustum* (RRS), *S. spontaneum* (SRS), and *S. sinense* Roxb. cv Uba (URS).

Streptomyces, and *Rhizobium* were the most abundant in *S. officinarum*. However, *S. barberi* soil sample was dominated by *Bacillus*, *Pseudomonas*, *Pseudarthobacter*, *Lysobacter*, *Gemmatimonas*, and *Sphingomonas*. *S. robustum* contained *Bacillus*, *Pseudomonas*, *Pseudarthobacter*, *Massilia*, *Nitrospira*, *Gemmatimonas*, *Streptomyces*, *Paenibacillus*, and

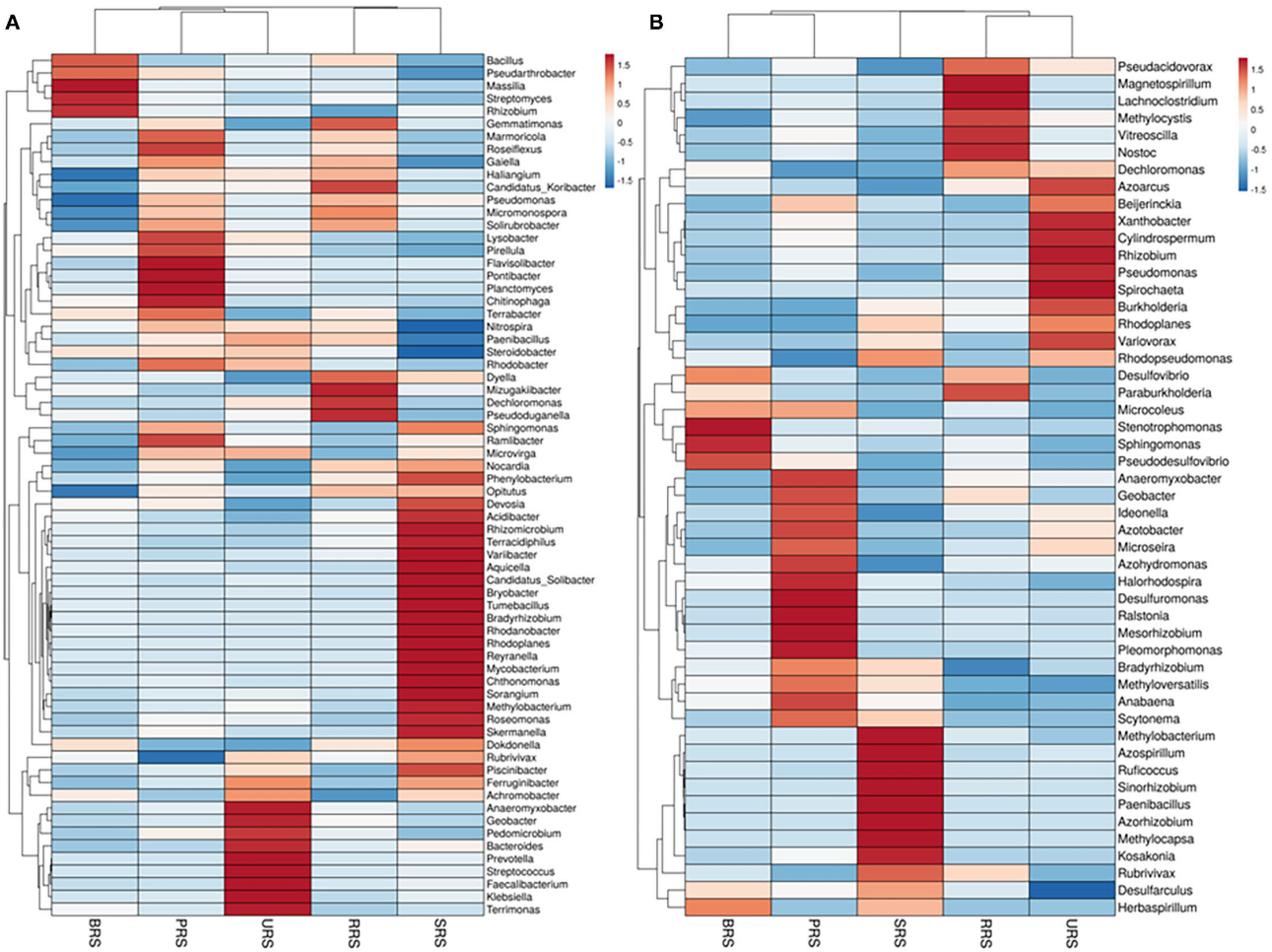


FIGURE 5 | Heat map showing the abundances of bacterial communities at genus level in the soil samples of sugarcane species **(A)** 16S rRNA and **(B)** *nifH* gene. *S. officinarum* L. cv Badila (BRS), *S. barberi* Jesw. cv Pansahi (PRS), *S. robustum* (RRS), *S. spontaneum* (SRS), and *S. sinense* Roxb. cv Uba (URS).

Dechloromonas. *S. spontaneum* sample contained *Pseudomonas*, *Pseudarthobacter*, *Massilia*, *Tumebacillus, Remibacter, Sphingomonas*, and *Skermanelia. S. sinense* sample was dominated by *Bacillus*, *Pseudomonas*, *Pseudarthobacter*, *Massilia*, *Lysobacter*, *Nitrospira*, *Faecalibacterium,* and *Streptococus*.

Genus distribution using *nifH* gene data is presented in **Figure 5B**. *Bradyrhizobium*, *Dechloromonas*, *Desulfovibrio*, and *Stenotrophomonas* were abundant in *S. officinarum*, while *Bradyrhizobium*, *Desulfovibrio*, *Xanthobacter,* and *Anaeromyxobacter* were the leading genera in *S. barberi*. *Bradyrhizobium*, *Dechloromonas*, *Desulfovibrio*, and *Anaeromyxobacter* were the dominant groups in *S. robustum*. *S. spontaneum* contained *Bradyrhizobium*, *Azospirillum*, *Methanobacterium*, and *Paenibacilllus* species. *Bradyrhizobium*, *Dechloromonas*, *Xanthobacter* and *Anaeromyxobacter* were present in *S. sinense*. Genus *Burkolderia* was found in *S. barberi*, *S. spontaneum*, and *S. sinense* while *Beijerinckia* was recorded in *S. barberi, S. spontaneum*, and *S. sinense*. Genus *Idenella* was commonly present in all the samples except *S. spontaneum,* and *Kosakonia* was normally present in *S. barberi and S. spontaneum*.

Star analysis was conducted using the top 10 genera of each sample. The top 10 bacterial genera of 16S rRNA data-based analysis were *Pseudarthobacter*, *Pseudomonas*, *Bacillus*, *Massilia*, *Gemmatimonas*, *Nitrospora*, *Haliangium*, *Ramibacter*, *Tumebacillus,* and *Lysobacter*. From this analysis, the *Bacillus* genus was the leading one in *S. officinarum* and *S. robustum* samples, while *Pseudomonas* was the dominant one in *S. barberi*, *S. robustum*, *S. spontaneum,* and *S. sinense. Tumebacillus* genus was found only in *S. spontaneum* (**Supplementary Figure S4A**). Similarly, star analysis for *nifH* gene samples identified the presence of *Bacillus* in all the samples. *Desulfovibrio* was found in significant numbers *in S. officinarum, S. robustum,* and *S. barberi*. *Xanthobacter* was abundant in *S. sinense. Anaeromyxobacter* was found in all the samples except that of *S. spontaneum. Pseudoacidovorax* genus was found only in *S. robustum*, *S. barberi,* and *S. sinense* samples. *Azospirillum* and *Methylobacterium* were unique to *S. spontaneum* (**Supplementary Figure S4B**). The relative abundance of the top 20 diazotrophs at the genus level present in all sugarcane species is shown in **Figure 6**. Diazotrophs

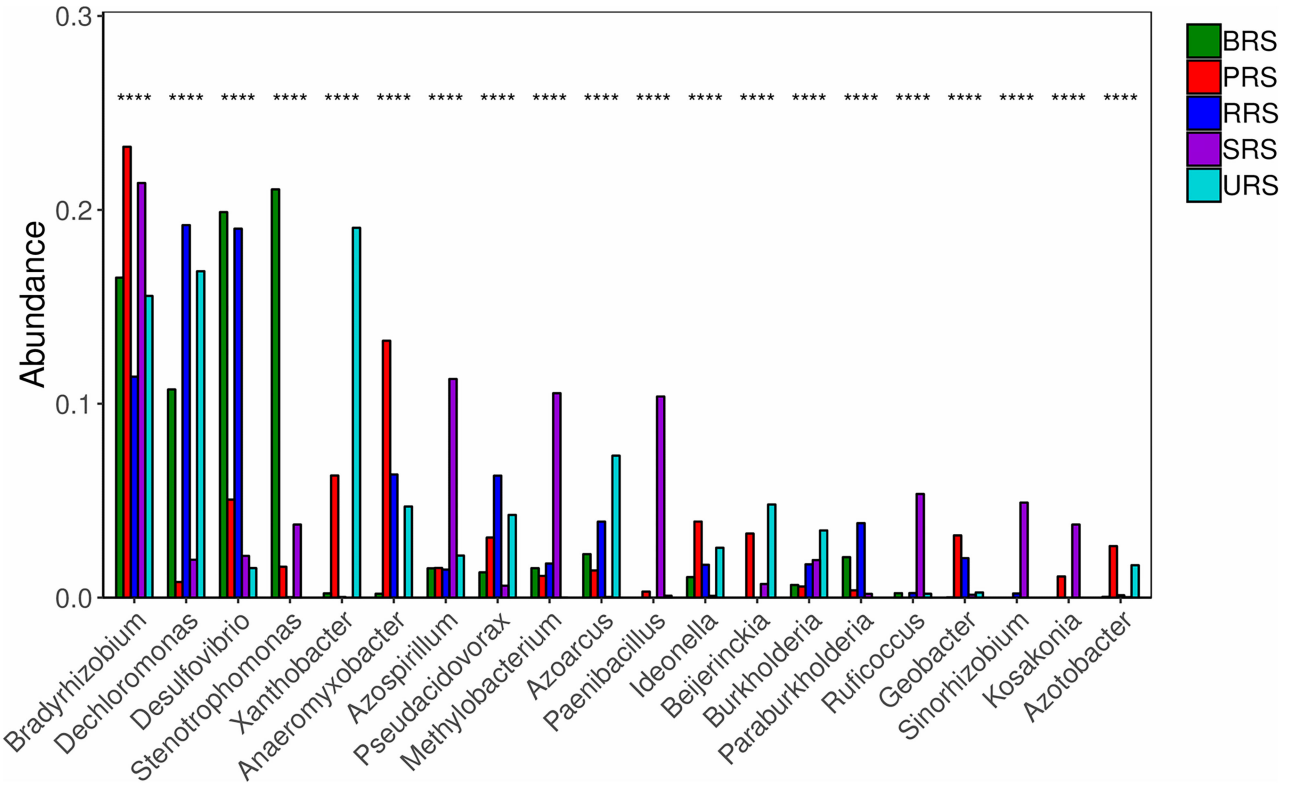


FIGURE 6 | The relative abundances of the top 20 diazotrophs at the genus level in the soil of sugarcane species. *S. officinarum* L. cv Badila (BRS), *S. barberi* Jesw. cv Pansahi (PRS), *S. robustum* (RRS), *S. spontaneum* (SRS), and *S. sinense* Roxb. cv Uba (URS). Note: Select species with $p < 0.05$ to draw a column chart, each column is the relative abundance of different species, where **** means $p < 0.0001$.

belonging to the genus *Bradyrhizobium* were present in all the samples tested.

Beta Diversity

Beta diversity measures species diversity among different samples collected from similar or different environments. We performed the PCoA and UniFrac-based cluster analysis to understand the beta diversity of 16S rRNA and *nifH* gene in the sugarcane species studied (**Supplementary Figures S5, S6**). Based on 16S rRNA data, *S. sinense,* along with *S. barberi*, and *S. officinarum,* along with *S. robustum*, formed independent clusters. Moreover, *S. spontaneum* was segregated away from all others, displaying higher beta diversity (**Supplementary Figure S5A**). Similarly based on *nifH* data, *S. spontaneum* was segregated away from all others samples (**Supplementary Figure S6A**). Hierarchical clustering based on the UniFrac cluster analysis showed similar results for both 16S rRNA and *nifH* gene data, containing identical sequences showing 0 (blue color) distances. *S. sinense* to *S. spontaneum* showed more distance (red color = 0.3), indicating dissimilarity in sequences of *S. sinense* to *S. spontaneum* samples (**Supplementary Figures S5b**, **S6b**).

The distribution of dominant genera based on their relative abundance performed with the Bray-Curtis algorithm showed *Bacillus* being the dominant genus in all the samples except those from *S. spontaneum* (**Figure 7A**). *Pseudomonas* was the second dominant genus, while *Pseudarthobacter* became the third leading one. *S. spontaneum* had most of the unidentified genera in our analysis. Similarly, the distribution of dominant genera in *nifH* samples showed the presence of *Dechloromonas* spp. in all samples except those from *S. barberi*. *Xanthobacter* and *Bradyrhizobium* were found to be prevalent in *S. officinarum* and *S. sinense,* respectively (**Figure 7B**).

qPCR

To quantify bacterial abundance in all five species, we performed qPCR analysis using the 16S rRNA and *nifH* gene-specific primer (**Figure 8**). Occurrences of gene copy number, indicating bacterial abundance, in all species were as follows: *S. sinense* > *S. robustum* > *S. barberi* > *S. officinarum* > *S. spontaneum* (**Figure 8A**). The *nifH* gene copy number results showed that N-fixing bacterial population was highest in *S. barberi* followed by *S. robustum* > *S. officinarum* > *S. sinense* with the least in *S. spontaneum* (**Figure 8B**). Thus, *S. sinense* had the most considerable rhizospheric bacterial abundance, while N-fixing bacteria were most prolific in the rhizosphere of *S. barberi*.

Spearman's Rank Correlation

Spearman rank correlation analysis was calculated based on 16S rRNA and *nifH* gene abundance (top 50), diversity indices, gene copy, and soil variables; and values were illuminated in a heat map (**Figure 9**). Spearman's rank correlation coefficients were calculated to assess the association between soil physiochemical

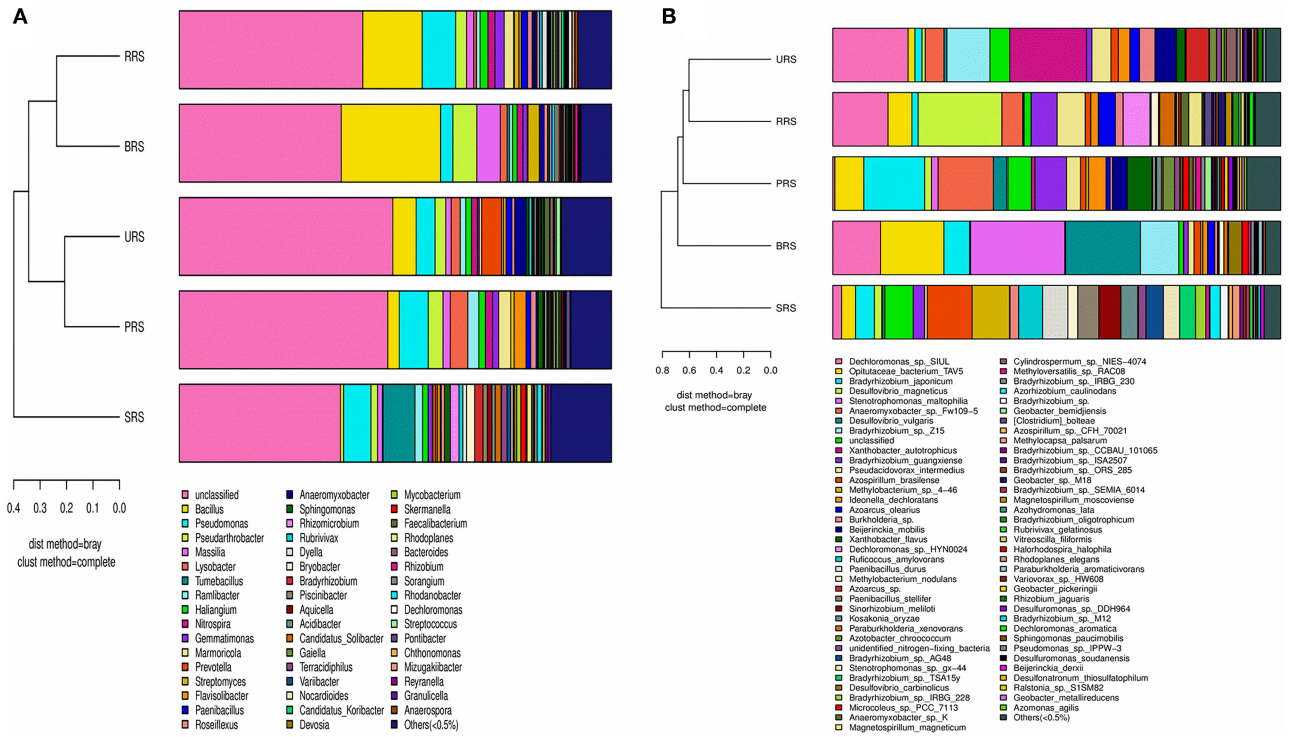


FIGURE 7 | A combination of horizontal multi-sample similarity trees and histograms 16S rRNA **(A)** and *nifH* genes **(B)**. On the left is hierarchical clustering between samples based on community composition (Bray-Curtis algorithm), and on the right is a column chart of the sample's community structure. *S. officinarum* L. cv Badila (BRS), *S. barberi* Jesw. cv Pansahi (PRS), *S. robustum* (RRS), *S. spontaneum* (SRS), and *S. sinense* Roxb. cv Uba (URS).

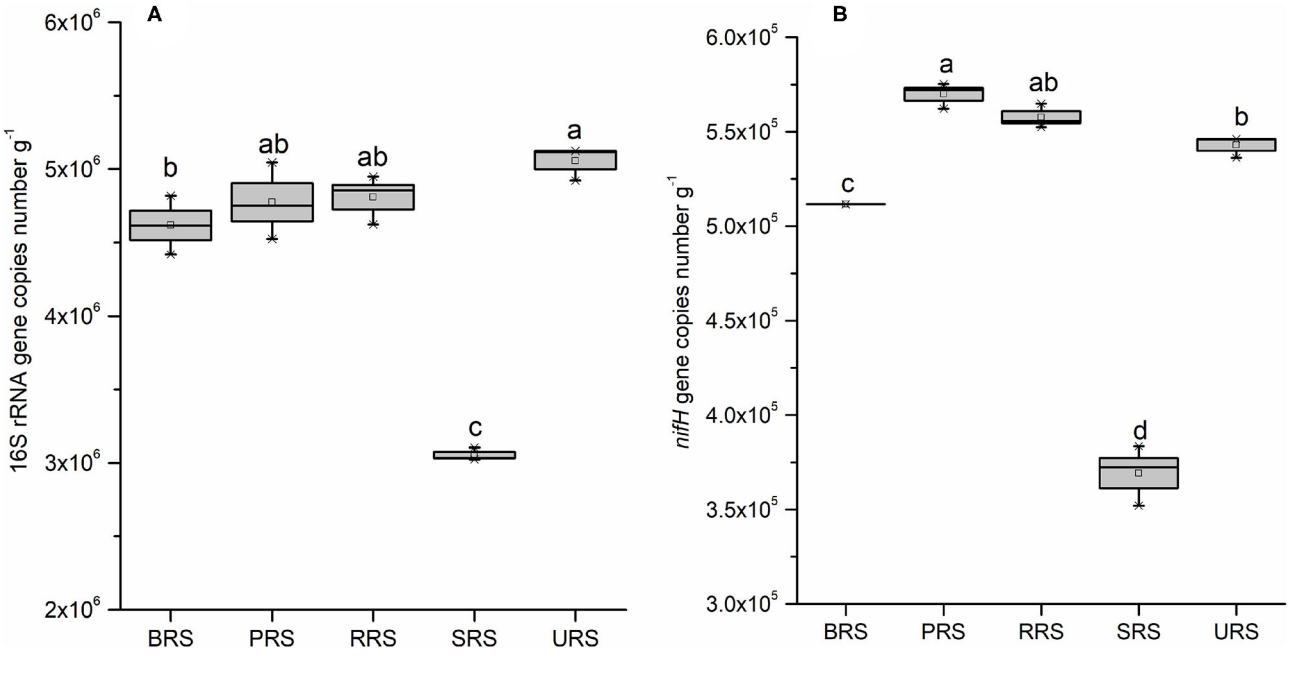


FIGURE 8 | Box plot estimated values of gene copies of 16S rRNA **(A)**, and *nifH* **(B)** in the soil samples of sugarcane species. Within each box, Small squares denote mean values; boxes extend from the 10th to the 90th percentile of each group's distribution of values. Values followed by the same lowercase letters in the same column are not significantly different according to DMRT 5%. DMRT, Duncan's Multiple Range Test.

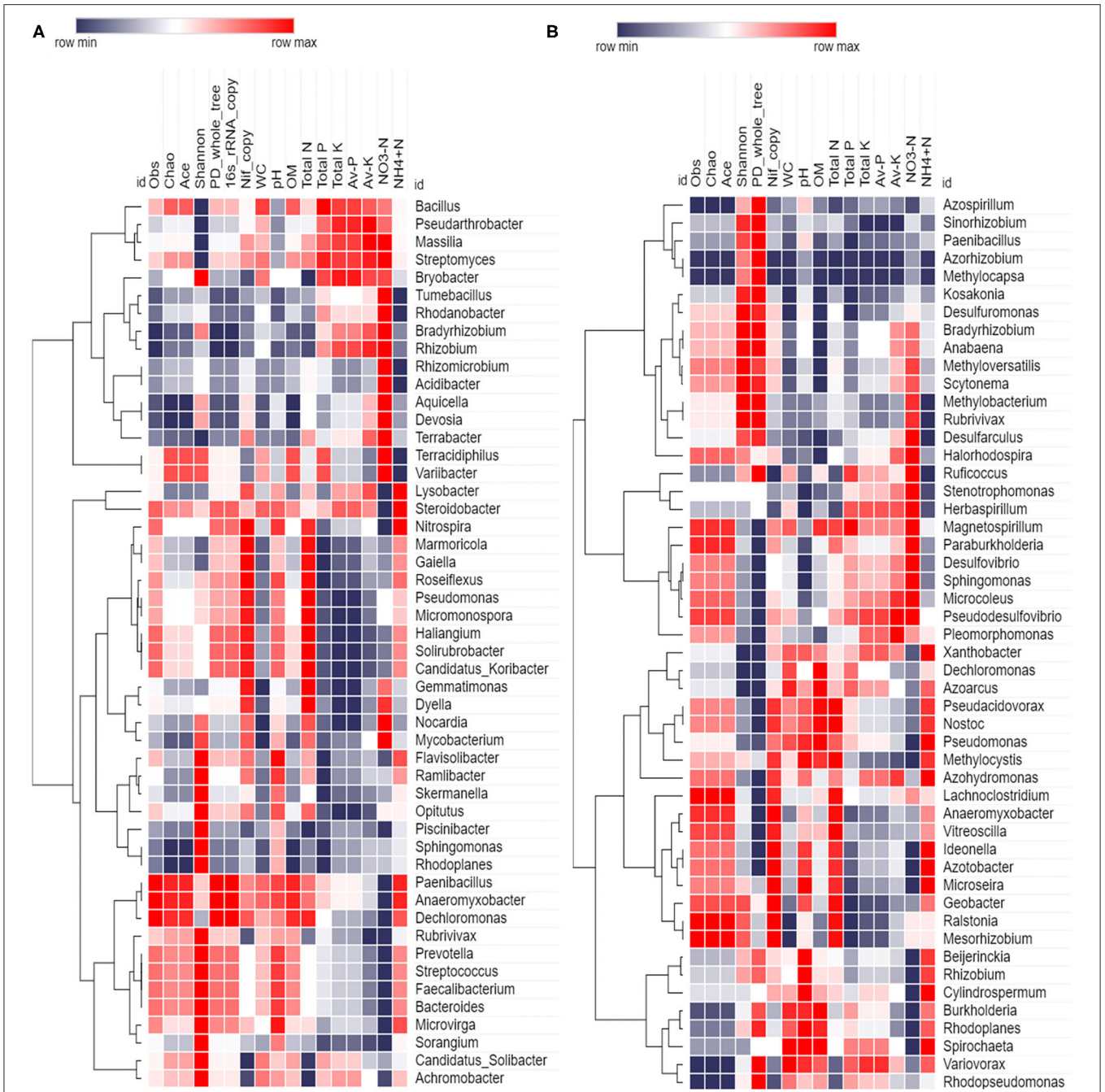


FIGURE 9 | Spearman correlation heatmap based on abundance (top 50 genera), diversity indices, gene copy, and soil variables **(A)** 16S rRNA **(B)** *nifH* gene.

properties and dominant (16S rRNA and *nifH*) bacterial genera. The relationships were estimated between these taxa and physiochemical properties to understand the role of the microbial community shift. Related heat maps show that soil physiochemical factors significantly affected the relative abundance of bacterial taxa (phyla and genera). 16S rRNA bacterial abundance shows a strong positive correlation ($r = 0.8$, $p < 0.1$) between pH and abundance of bacteria, such as *Nitrospira*, *Flavisolibacter*, and *Paenibacillus*. OM had a positive correlation ($p < 0.05$) with *Paenibacillus*, *Anaeromyxobactor,* and *Microvirga*. Total N had a positive correlation ($r = 0.9$, $p < 0.05$) with *Marmoricola* and *Gaiella*. Total P had positive correlation ($p < 0.05$) with *Bacillus* ($r = 0.9$) and *Streptomyces* ($r = 0.6$). Total K had positive correlation ($r = 0.9$, $p < 0.01$) with *Pseudarthobacter* and *Streptomyces*. Available P had positive correlation with *Pseudarthobacter* ($r = 0.9$, $p < 0.05$). While, available K had

positive correlation with *Pseudarthobacter* ($r = 1$, $p < 0.05$) and *Massilia* ($r = 0.9$, $p < 0.05$). Moreover, available NO_3^-N showed a positive correlation with *Massilia* ($r = 0.9$, $p < 0.05$), *Terrabacter* ($r = 0.9$, $p < 0.5$), and *Streptomyces* ($r = 0.8$, $p < 0.1$). Available NH_4^+N had a positive correlation with *Lysobacter* ($r = 0.8$, $p < 0.1$), *Nitrospira* ($r = 0.9$, $p < 0.05$), *Paenibacillus* ($r = 0.9$, $p < 0.05$), *Anaeromyxobacter* ($r = 0.9$, $p < 0.05$), and *Steroidobacter* ($r = 0.9$, $p < 0.05$). *Paenibacillus* and *Anaeromyxobacter* strongly correlated with obs, Chao, PD_whole tree, and OM. However, *Nocardia* had a significant negative correlation with total K, while *Tumebacillus* showed a negative correlation with pH. In the case of *nifH* gene, diversity indices (observed species, Ace, Chao) had a strong positive correlation with *Anaeromyxobacter* ($r = 0.9$, $p < 0.05$), *Lachnoclostridium* ($r = 0.9$, $p < 0.05$), and *Vitreoscilla* ($r = 0.8$, $p < 0.1$). The qRT-PCR data (*nifH* gene copy number) had a strong positive correlation ($p < 0.5$) with *Anaeromyxobacter*, *Ideonella*, *Geobacter*, *Azotobacter*, and *Vitreoscilla*. WC had positive correlation with *Azoarcus* ($r = 0.9$, $p < 0.05$), and soil pH showed a strong positive correlation with *Ideonella* ($r = 0.8$, $p < 0.1$), *Azotobacter* ($r = 0.8$, $p < 0.1$), *Rhizobium* ($r = 0.9$, $p < 0.05$), and *Pseudomonas* ($r = 0.8$, $p < 0.1$). Similarly, soil OM had a strong positive correlation with *Azoarus* ($r = 1$, $p < 0.05$), *Pseudomonas* ($r = 0.9$, $p < 0.05$), and *Nostoc* ($r = 0.8$, $p < 0.1$). Soil total N revealed a strong positive correlation with *Vitreoscilla* ($r = 0.9$, $p < 0.05$), and available K showed with *Pseudosulfovibrio* ($r = 0.9$, $p < 0.05$). While, soil available NO_3^-N was correlated positively to *Desulfovibrio* ($r = 0.9$, $p < 0.05$), *Microcoleus* ($r = 0.97$, $p < 0.05$), *Halorhodospira* ($r = 0.8$, $p < 0.1$), *Sphingnomonas* ($r = 0.9$, $p < 0.05$), and *Pseudosulfovibrio* ($r = 0.9$, $p < 0.05$). Moreover, soil available NH_4^+N showed strong correlation with *Xanthobacter* ($r = 0.9$, $p < 0.05$), *Ideonella* ($r = 0.9$, $p < 0.05$), *Azotobacter* ($r = 0.9$, $p < 0.05$), *Pseudomonas* ($r = 0.9$, $p < 0.05$), and *Azohydromonas* ($r = 0.08$, $p < 0.1$). However, pH was correlated significantly negative with the *Stenotrophomonas* ($r = -0.9$, $p < 0.05$), and OM was correlated strongly negative with *Desulfarculus* ($r = -0.9$, $p < 0.05$) and *Anabaena* ($r = -0.9$, $p < 0.1$). Similarly, soil available NH_4^+N showed a negative correlation with *Methylobacter* ($r = -0.9$, $p < 0.05$) and *Desulfarculus* ($r = -0.9$, $p < 0.05$).

DISCUSSION

The present study focuses on wild sugarcane progenitors because these species are distributed in different geographic distributions and service in various biotic and abiotic stresses (Santchurn et al., 2019). To understand the PGPRs and nitrogen-fixing diazotrophs, we need to study the sugarcane rhizosphere microbiome community. In most crops, such as sugarcane, a large proportion of PGPRs are yet to be identified. Bacteria are the most abundant of all the rhizospheric microbiota, and many are known to promote plant growth (Antoun and Prévost, 2005; van Loon, 2007; Dagnaw et al., 2015). The present study made an effort to understand the diverse and complex bacterial communities present in the rhizospheric soils of progenitors and closely related species of modern sugarcane hybrids as they play critical roles in plant nutrition, biotic and abiotic stress tolerance, and growth and development (Baldani et al., 2002; Li et al., 2017).

16S rRNA sequence data revealed 6,202 OTUs assigned to different bacterial species colonizing the rhizosphere of different sugarcane species studied. Analysis of these OTUs showed that *S. sinense* rhizosphere has the most significant number of rhizobacterial communities compared to other related species studied here. These results were also confirmed by qPCR analysis. qPCR analysis results demonstrated that 16S rRNA gene copy numbers were observed highest in *S. sinense* and lowest in *S. spontaneum*. The highest *nifH* gene numbers were observed in *S. robustum* and lowest in *S. spontaneum*. Besides this, alpha diversity analysis also predicted the highest bacterial diversity in *S. sinense* samples. However, other sugarcane species also showed considerable species diversity. Phylum level distribution studies identified the dominance of Proteobacteria, Firmicutes, Actinobacteria, Bacteroidetes, Chloroflexi, Gemmatimonadetes, Planctomycetes, and Nitrospirae. Many of them were reported to be present in the sugarcane rhizosphere previously (Gao et al., 2019). Out of the total OTUs identified, 1,099 were from diazotrophs based on *nifH* gene data. The nifH gene sequence analysis helped to identify the top twenty genera (*Bradyrhizobium*, *Dechloromonas*, *Desulfovibrio*, *Stenotrophomonas*, Xanthobacter *Anthobacter*, *Anaeromyxobacter*, *Azospirillum*, *Pseudoacidovorax*, *Methylobacterium*, *Azoarcus*, *Paenibacillus*, *Ideonella*, *Beijerinckia*, *Paraburkholderia*, *Burkholderia*, *Ruficoccus*, *Geobacter*, *Sinorhizobium*, *Kosakonia,* and *Azotobacter*). Out of these 20 genera, most of the genera were found to fix nitrogen in sugarcane and other plants. Genus *Bradyrhizobium* is known to fix nitrogen in sorghum and sugarcane (de Matos et al., 2019; Hara et al., 2019). *Azotobacter* genus consists of seven different species and they are involved in atmospheric nitrogen fixation in different crops (Jiménez et al., 2011). Genera *Azotobacter* and *Beijerinckia* were studied for diazotrophic attributes in the early 1960s in rice and other cereal crops (Dobereiner, 1961; Pankievicz et al., 2019). Contrary to *Azotobacter*, the *Beijerinckia* genus is restricted mainly to the tropics, and its nitrogen fixation ability has been reported in various plants (Nassar et al., 2020). *Kosakonia* spp. fixes nitrogen on cucumber roots (Sun et al., 2018). Roots of switchgrass are inhabited with nitrogen-fixing bacteria that belong to the genera *Dechloromonas*, *Desulfovibrio*, *Azoarcus*, *Ideonella*, *Paraburkholderia,* and *Burkholderia* (Bahulikar et al., 2019). It is hard to identify and classify most bacteria in culture because of their morphological similarities. But, culture-independent methods, such as 16S rRNA sequencing, are highly efficient, cost-effective and provide accurate identification and classification of rhizobacteria. Overall, the dominant genera identified in this study are known to fix atmospheric nitrogen, facilitating plant growth.

In the present study, we observed a few taxa, such as *Proteobacteria*, *Acidobacteria*, *Actinobacteria*, *Firmicute*, *Sphingomonas*, *Bradyrhizobium*, and *Gemmatimonas,* were dominant with *Bacillus*, *Pseudomonas*, *Bradyrhizobium*, *Burkholderia*, and *Rhizobium* as leading genera. In our previous studies, we observed the occurrence of *Proteobacteria*, *Acidobacteria*, *Actinobacteria*, *Firmicute*, *Bacteroidetes,*

Sphingomonas, *Bradyrhizobium*, *Bryobacter*, and *Gemmatimonas* in the sugarcane rhizosphere (Solanki et al., 2018, 2020). Some genera, such as *Bacillus*, *Pseudomonas,* and *Burkholderia*, known for their plant-growth-promoting and nitrogen-fixing properties, were found to be enriched in the sugarcane rhizosphere (Li et al., 2017; Malviya et al., 2019; Singh et al., 2020). Community composition analysis of 16S rRNA sequence data helped to track phylum and genus level distribution of rhizobacteria among different sugarcane species studied. Interestingly, we observed the presence of a few genera, namely *Streptococcus*, *Rhodanobacter*, *Anaeromyxobacter,* and *Prevotella,* shared among all the species studied here. The members of these genera were found commonly in soil and can colonize crop plants. From the previous reports, it appears that *Rhodanobacter*, *Anaeromyxobacter,* and *Prevotella* species were isolated from different soil and plant sources, and they were found to have *nifH* gene and nitrogen fixation abilities (Igai et al., 2016; Espenberg et al., 2018). Thus, we believe that more characterization of these bacteria colonizing the sugarcane rhizosphere will be beneficial for developing bio-based crop products to improve sugarcane crop productivity. Diversity among these nitrogen-fixing bacteria was revealed by alpha and beta diversity analyses. The top genera with the highest abundance were found to be *Bacillus*, *Desulfovibrio*, *Xanthobacter*, *Anaeromyxobacter*, *Pseudoacidovorax*, *Azospirillum,* and *Methylobacterium*. Among these, *Bacillus* is a common bacterial diazotroph in sugarcane (Reis et al., 2007). *Desulfovibrio*, *Anaeromyxobacter*, *Azospirillum,* and *Xanthobacter* are nitrogen fixers in rice (Sessitsch et al., 2012; Bao et al., 2014; Yoneyam et al., 2017; Rosenblueth et al., 2018). Correlation results showed that among all bacteria, genera *Bacillus*, *Pseudomonas*, *Massilia*, *Pseudarthobacter,* and *Streptomyces* correlate positively with various soil parameters. Genera *Bacillus*, *Pseudomonas,* and Streptomyces are well-known for plant growth and plant disease suppression activities in sugarcane and other crops (Wang et al., 2019; Amna et al., 2020; Chandra et al., 2020; Jiao et al., 2021). Genus *Massilia* is a significant group of rhizosphere and root colonizing bacteria of many plant species (Ofek et al., 2012). Genus *Pseudarthobacter* plays an essential role in biodegradation (Gupta et al., 2020; Chen et al., 2021). Genera *Azotobacter and Ideonella* revealed a positive correlation with NH_4^+N, and several previous reports also demonstrated that these bacteria could fix atmospheric nitrogen efficiently and enhance the crop yield (Bhromsiri and Bhromsiri, 2010; Aasfar et al., 2021). Diazotroph *Anaeromyxobacter* correlated positively with NH_4^+N and recently reported that *Anaeromyxobacter* enhances the nitrogen-fixing activity of paddy soils (Masuda et al., 2020). *Desulfovibrio* had a positive correlation with available NO_3^-N. *Desulfovibrio* is known for its sulfur-reducing property and plays a particular and significant role in the fixation of biological nitrogen (Mistry et al., 2022). On the contrary, soil physiochemical, such as OM, showed a negative correlation with genera *Desulfarculus* and *Anabaena,* while NH_4^+N showed a negative correlation with genera *Methylobacter* and *Desulfarculus*. It may be due to a decrease of soil nutrients in the sugarcane rhizosphere that increases microbial completion (Jones et al., 2018; Solanki et al., 2020).

CONCLUSIONS

The sugarcane species studied here showed a significant number of N-fixing rhizobacteria, which strengthens the contention that exploring rhizosphere bacteria may help to develop a sustainable low resource input sugarcane crop production system, particularly, for meeting its N requirement from atmospheric nitrogen fixation. Substantial genetic variation for rhizobacteria, such as diazotroph communities, exists among different progenitors and closely related species of modern cultivated sugarcane hybrids. However, considering the vast natural habitats of these wild species spanning tropics and subtropics, similar studies using accessions sourced from different locations and environmental conditions will significantly advance our understanding of sugarcane rhizobiome. Future research should also focus on *Saccharum* species introgression breeding and isolation and practical application of beneficial PGPRs, especially diazotrophs. Filling the significant knowledge gap on microbiota and sugarcane interactions is critical for exploiting these beneficial microbes for sustainable sugarcane agriculture.

DATA AVAILABILITY STATEMENT

The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found in the article/**Supplementary Material**.

AUTHOR CONTRIBUTIONS

MM and C-NL conceived and designed the experiments and performed the experiments. MM, C-NL, and MS analyzed the data. ZW, ACS, KV, AS, RS, PS, and ED contributed to reagents/materials/analysis tools. Y-RL contributed to project administration. C-NL, QN, X-PS, and Y-RL contributed to funding acquisition. MM, MS, and PL contributed in writing-original draft and revision. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by grants from the Guangxi Innovation Teams of Modern Agriculture Technology (No. gjnytxgxcxtd-2021-03-01 to Y-RL), National Natural Science Foundation of China (Nos. 31701489 to QN, 31801288 to C-NL, and 31901594 to X-PS), the Natural Science Foundation of Guangxi Province (Nos. 2019GXNSFDA185004 and 2021GXNSFAA196041 to C-NL), and Guangxi Academy of Agricultural Sciences Fund (Nos. Guinongke 2021YT09 and Guinongke 2021JM01 to C-NL).

ACKNOWLEDGMENTS

The authors extend their appreciation to National Natural Science Foundation of China, China Postdoctoral Foundation, the Youth program of the National Natural Science Foundation of China, Guangxi Innovation Term of Modern Agriculture

Technology, and Guangxi Academy of Agricultural Sciences, Nanning, Guangxi, China, for providing the necessary funding and facilities for this study. The authors also extend their appreciation to Taif University Researchers Supporting Project number (TURSP-2020/85).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.829337/full#supplementary-material

REFERENCES

Aasfar, A., Bargaz, A., Yaakoubi, K., Hilali, A., Bennis, I., Zeroual, Y., et al. (2021). Nitrogen fixing azotobacter species as potential soil biological enhancers for crop nutrition and yield stability. *Front. Microbiol.* 12, 628379. doi: 10.3389/fmicb.2021.628379

Ahmad, F., Ahmad, I., Aqil, F., Ahmed Wani, A., and Sousche, Y. S. (2016). Plant growth promoting potential of free-living diazotrophs and other rhizobacteria isolated from Northern Indian soil. *Biotechnol. J.* 1, 1112–1123. doi: 10.1002/biot.200600132

Amna, Xia, Y., Farooq, M. A., Javed, M. T., Kamran, M. A., Mukhtar, T., Ali, J., et al. (2020). Multi-stress tolerant PGPR Bacillus xiamenensis PM14 activating sugarcane (*Saccharum officinarum* L.) red rot disease resistance. *Plant Physiol. Biochem.* 151, 640–649. doi: 10.1016/j.plaphy.2020.04.016

Antoun, H., and Prévost, D. (2005). "Ecology of plant growth promoting rhizobacteria," in *PGPR: Biocontrol and Biofertilization*, ed Z. A. Siddiqui (Netherlands: Springer), 1–38. doi: 10.1007/1-4020-4152-7_1

Armanhi, J. S., de Souza, R. S., Damasceno, N. D., de Araújo, L. M., Imperial, J., and Arruda, P. A. (2018). Community-based culture collection for targeting novel plant growth-promoting bacteria from the sugarcane microbiome. *Front. Plant Sci.* 8, 2191. doi: 10.3389/fpls.2017.02191

Avis, T. J., Gravel, V., Antoun, H., and Tweddell, R. J. (2008). Multifaceted beneficial effects of rhizosphere microorganisms on plant health and productivity. *Soil Biol. Biochem.* 40, 1733–1740. doi: 10.1016/j.soilbio.2008.02.013

Bahulikar, R. A., Chaluvadi, S. R., Torres-Jerez, I., Mosali, J., Bennetzen, J. L., and Udvardi, M. (2019). Nitrogen fertilization reduces nitrogen fixation activity of diverse diazotrophs in switchgrass roots. *Phytobiomes J.* 5, 80–87. doi: 10.1094/PBIOMES-09-19-0050-FI

Baldani, J. I., Reism, V. M., Baldani, V. L., and Döbereiner, J. (2002). A brief story of nitrogen fixation in sugarcane-reasons for success in Brazil. *Funct. Plant Biol.* 29, 417–423. doi: 10.1071/PP01083

Bao, S. D. (2002). *Soil Agricultural Chemical Analysis, 3rd Edn*. Beijing: China Agricultural Press.

Bao, Z., Okubo, T., Kubota, K., Kasahara, Y., Tsurumaru, H., Anda, M., et al. (2014). Metaproteomic identification of diazotrophic methanotrophs and their localization in root tissues of field-grown rice plants. *Appl. Environ. Microbiol.* 80, 5043–5052. doi: 10.1128/AEM.00969-14

Barbier, F. F., Chabikwa, T. G., Ahsan, M. U., Cook, S. E., Powell, R., Tanurdzic, M., et al. (2019). A phenol/chloroform-free method to extract nucleic acids from recalcitrant, woody tropical species for gene expression and sequencing. *Plant Methods* 15, 62. doi: 10.1186/s13007-019-0447-3

Berendsen, R. L., Pieterse, C. M., and Bakker, P. A. (2012). The rhizosphere microbiome and plant health. *Trends Plant Sci.* 17, 478–486. doi: 10.1016/j.tplants.2012.04.001

Bhromsiri, C., and Bhromsiri, A. (2010). Isolation, screening of growth-promoting activities and diversity of rhizobacteria from vetiver grass and rice plants. *Thai J. Agric. Sci.* 43, 217–230.

Brackin, R., Robinson, N., Lakshmanan, P., and Schmidt, S. (2013). Microbial function in adjacent subtropical forest and agricultural soil. *Soil Biol. Biochem.* 57, 68–77. doi: 10.1016/j.soilbio.2012.07.015

Bremmer, J. M., and Mulvaney, C. S. (1982). "Total nitrogen," in methods of soil analysis—Part 2," in *Chemical and Microbiological Properties*, ed. C. A. Bluck (Madison: American Society of Agronomy), 595–624.

Caporaso, J., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F. D., Costello, E. K., et al. (2010). QIIME allows analysis of high-throughput community sequencing data. *Nat. Methods* 7, 335–336. doi: 10.1038/nmeth.f.303

Carter, M. R. (1993). *Soil Sampling and Methods of Analysis*. Boca Raton, FL: Lewis Publishers.

Cawoy, H., Bettiol, W., Fickers, P., and Ongena, M. (2011). Bacillus-based biological control of plant diseases. *Pesticides Modern World-Pesticides Use Manage.* 21, 273–302. doi: 10.5772/17184

Chandra, H., Kumari, P., Bisht, R., Prasad, R., and Yadav, S. (2020). Plant growth promoting *Pseudomonas aeruginosa* from Valeriana wallichii displays antagonistic potential against three phytopathogenic fungi. *Mol. Biol. Rep.* 47, 6015–6026. doi: 10.1007/s11033-020-05676-0

Chen, F. Y., Chen, Y., Chen, C., Feng, L., Dong, Y., Chen, J., et al. (2021). High-efficiency degradation of phthalic acid esters (PAEs) by *Pseudarthro bacter* defluvii E5: performance, degradative pathway, and key genes. *Sci. Total Environ.* 794, 148719. doi: 10.1016/j.scitotenv.2021.148719

Dagnaw, F., Assefa, F., Gebrekidan, H., and Argaw, A. (2015). Characterization of plant growth promoting bacteria from sugarcane (*Saccharum officinarum* L.) rhizosphere of Wonji-shoa sugar estate and farmers landraces of Ethiopia. *Biotechnol.* 14, 58. doi: 10.3923/biotech.2015.58.64

de Matos, G. F., Zilli, J. E., de Araújo, J. L. S., Parma, M. M., Melo, I. S., Radl, V., et al. (2019). *Bradyrhizobium sacchari* sp. nov., a legume nodulating bacterium isolated from sugarcane roots. *Arch. Microbiol.* 199, 1251–1258. doi 10.1007/s00203-017-1398-6

Dini-Andreote, F., Andreote, F. D., Costa, R., Taketani, R. G., van Elsas, J. D., and Araújo, W. L. (2010). Bacterial soil community in a Brazilian sugarcane field. *Plant Soil* 336, 337–349. doi: 10.1007/s11104-010-0486-z

Dobereiner, J. (1961). Nitrogen-fixing bacteria of the genus Beijerinckia Derx in the rhizosphere of sugar cane. *Plant Soil.* 15, 211–216. doi: 10.1007/BF01400455

Dong, L., Xu, J., Zhang, L., Yang, J., Liao, B., Li, X., et al. (2017). High-throughput sequencing technology reveals that continuous cropping of American ginseng results in changes in the microbial community in arable soil. *Chinese Med.* 12, 18. doi: 10.1186/s13020-017-0139-8

Dong, M., Yang, Z., Cheng, G., Peng, L., Xu, Q., and Xu, J. (2018). Diversity of the bacterial microbiome in the roots of four saccharum species: S. *spontaneum, S. robustum, S. barberi, and S. officinarum. Front. Microbiol.* 9, 267. doi: 10.3389/fmicb.2018.00267

Edgar, R. C. (2004). MUSCLE: a multiple sequence alignment method with reduced time and space complexity. *BMC Bioinform.* 5, 113. doi: 10.1186/1471-2105-5-113

Edgar, R. C. (2013). UPARSE: highly accurate OTU sequences from microbial amplicon reads. *Nat. Methods* 10, 996. doi: 10.1038/nmeth.2604

Edgar, R. C., Haas, B. J., Clemente, J. C., Quince, C., and Knight, R. (2011). UCHIME improves sensitivity and speed of chimera detection. *Bioinform* 27, 2194–2200. doi: 10.1093/bioinformatics/btr381

Espenberg, M., Truu, M., Mander, Ü., Kasak, K., Nõlvak, H., Ligi, T., et al. (2018). Differences in microbial community structure and nitrogen cycling in natural and drained tropical peat land soils. *Sci. Rep.* 8, 4742. doi: 10.1038/s41598-018-23032-y

Finkel, O. M., Castrillo, G., Herrera Paredes, S., Salas González, I., and Dangl, J. L. (2017). Understanding and exploiting plant beneficial microbes. *Curr. Opin. Plant Biol.* 38, 155–163. doi: 10.1016/j.pbi.2017.04.018

Fischer, D., Pfitzner, B., Schmid, M., Simoes-Araujo, J. L., Reis, V. M., Pereira, W., et al. (2012). Molecular characterization of the diazotrophic bacterial community in uninoculated and inoculated field-grown sugarcane (Saccharum sp.). *Plant Soil* 356, 83–99. doi: 10.1007/s11104-011-0812-0

Gaikwad, S., and Sapre, V. (2015). Structural and functional diversity of rhizobacterial strains isolated from rhizospheric zone of different plants of Sholapur- Maharashtra Region, India. *Int. J. Curr. Microbiol. Appl. Sci.* 4, 263–273.

Gao, X., Wu, Z., Liu, R., Wu, J., Zeng, Q., and Qi, Y. (2019). Rhizosphere bacterial community characteristics over different years of sugarcane ratooning in consecutive monoculture. *BioMed. Res. Int.* 2019:4943150. doi: 10.1155/2019/4943150

Goldemberg, J., Coelho, S. T., and Guardabassi, P. (2008). The sustainability of ethanol production from sugarcane. *Energy Policy* 36, 2086–2097. doi: 10.1016/j.enpol.2008.02.028

Gong, B., Cao, H., Peng, C., Perčulija, V., Tong, G., and Fang, H. (2019). High-throughput sequencing and analysis of microbial communities in the mangrove swamps along the coast of Beibu Gulf in Guangxi, China. *Sci. Rep*. 9, 9377. doi: 10.1038/s41598-019-45804-w

Gupta, N., Skinner, K. A., Summers, Z. M., Edirisinghe, J. N., Weisenhorn, P. B., Faria, J. P., et al. (2020). Draft genome sequence of Pseudarthrobacter sp. strain ATCC 49442 (Formerly Micrococcus luteus), a pyridine-degrading bacterium. *Microbiol. Resour. Announc*. 9, e00299-20 doi: 10.1128/MRA.00299-20

Hammer, Ø., Harper, D. A. T., and Ryan, P. D. (2001). Past: paleontological statistics software package for education and data analysis. *Palaeontol. Electron*. 4, 178. Available online at: http://palaeo-electronica.org/2001_1/past/issue1_01.htm

Hara, S., Morikawa, T., Wasai, S., Kasahara, Y., Koshiba, T., Yamazaki, K., et al. (2019). Identification of nitrogen-fixing bradyrhizobium associated with roots of field-grown sorghum by metagenome and proteome analyses. *Front. Microbiol*. 10, 407. doi: 10.3389/fmicb.2019.00407

Igai, K., Itakura, M., Nishijima, S., Tsurumaru, H., Suda, W., Tsutaya, T., et al. (2016). Nitrogen fixation and nifH diversity in human gut microbiota. *Sci. Rep*. 6, 31942. doi: 10.1038/srep31942

Inui-Kishi, R. N., Kishi, L. K., Picchi, S. C., Barbosa, J. C., Lemos, M. T. O., Marcondes, J., et al. (2012). Phosphorus solubilizing and IAA production activities in plant growth promoting rhizobacteria from Brazilian soils under sugarcane cultivation. *J. Eng. Appl. Sci*. 27, 1819–6608.

Jangu, O. P., and Sindhu, S. S. (2011). Differential response of inoculation with indole acetic acid producing Pseudomonas sp. in green gram *(Vignaradiata L.)* and black gram *(Vignamungo L.). Microbiol. J*. 1, 159–173. doi: 10.3923/mj.2011.159.173

Jiao, X., Takishita, Y., Zhou, G., and Smith, D. L. (2021). Plant associated rhizobacteria for biocontrol and plant growth enhancement. *Front. Plant Sci*. 17, 634796. doi: 10.3389/fpls.2021.634796

Jiménez, D. J., Montaña, J. S., and Martínez, M. M. (2011). Characterization of free nitrogen fixing bacteria of the genus azotobacterin organic vegetable-grown colombian soils. *Braz. J. Microbiol*. 42, 1517–8382. doi: 10.1590/S1517-83822011000300003

Jones, D. L., Magthab, E. A., Gleeson, D. B., Hill, P. W., Sánchez-Rodríguez, A. R., Roberts, P., et al. (2018). Microbial competition for nitrogen and carbon is as intense in the subsoil as in the topsoil. *Soil Biol. Biochem*. 117, 72–82. doi: 10.1016/j.soilbio.2017.10.024

Kashyap, A. S., Pandey, V. K., Kannojia, P., Singh, U., and Sharma, P. K. (2017). "Role of plant growth-promoting rhizobacteria for improving crop productivity in sustainable agriculture," in *Plant-Microbe Interactions in Agro-Ecological Perspectives*, eds D. Singh, H. Singh, and R. Prabha (Singapore: Springer), pp. 673–693. doi: 10.1007/978-981-10-6593-4_28

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

Kumar, A., Maurya, B. R., and Raghuwanshi, R. (2014). Isolation and characterization of PGPR and their effect on growth, yield and nutrient content in wheat (Triticum aestivum L.). *Biocatal. Agric. Biotechnol*. 3, 121–128. doi: 10.1016/j.bcab.2014.08.003

Lamizadeh, E., Enayatizamir, N., and Motamedi, H. (2016). Isolation and identification of plant growth-promoting rhizobacteria (pgpr) from the rhizosphere of sugarcane in saline and non-saline soil. *Int. J. Curr. Microbiol. Appl. Sci*. 5, 1072–1083. doi: 10.20546/ijcmas.2016.510.113

Li, H. B., Singh, R. K., Singh, P., Song, Q. Q., Xing, Y. X., Yang, L. T., et al. (2017). Genetic diversity of nitrogen-fixing and plant growth promoting Pseudomonas species isolated from sugarcane rhizosphere. *Front. Microbiol*. 8, 1268. doi: 10.3389/fmicb.2017.01268

Li, Y. R., and Yang, L. T. (2015). Sugarcane agriculture and sugar industry in China. *Sugar Tech*. 17, 1–8. doi: 10.1007/s12355-014-0342-1

Magoè, T., and Salzberg, S. L. (2011). FLASH: fast length adjustment of short reads to improve genome assemblies. *Bioinform* 27, 2957–2963. doi: 10.1093/bioinformatics/btr507

Malviya, M. K., Solanki, M. K., Li, C. N., Htun, R., Singh, R. K., Singh, P., et al. (2019). Beneficial linkages of endophytic *Burkholderia anthina* MYSP113 towards sugarcane growth promotion. *Sugar Tech*. 21, 737–748. doi: 10.1007/s12355-019-00703-2

Masuda, Y., Yamanaka, H., Xu, Z. X., Shiratori, Y., Aono, T., Amachi, S., et al. (2020). Diazotrophic anaeromyxobacter isolates from soils. *Appl. Environ. Microbiol*. 86, e00956-20. doi: 10.1128/AEM.00956-20

Mehnaz, S. (2011). Plant growth-promoting bacteria associated with sugarcane. in bacteria in agrobiology. *Crop Ecosyst*. 11, 165–187. doi: 10.1007/978-3-642-18357-7_7

Metsalu, T., and Vilo, J. (2015). ClustVis: a web tool for visualizing clustering of multivariate data using Principal Component Analysis and heatmap. *Nucleic Acids Res*. 43, 566–570. doi: 10.1093/nar/gkv468

Mistry, H., Thakor, R., and Bariy, H. (2022). "Isolation and identification of associative symbiotic n2 fixing microbes: Desulfovibrio," in *Practical Handbook on Agricultural Microbiology,* eds N. Amaresan, P. Patel, and D. Amin (New York, NY: Springer Protocols Handbooks), 77–83. doi: 10.1007/978-1-0716-1724-3_10

Moore, T. A., Xing, Y., Lazenby, B., Lynch, M. D. J., Schiff, S., Robertson, W. D., et al. (2011). Prevalence of anaerobic ammonium-oxidizing bacteria in contaminated groundwater. *Environ. Sci. Technol*. 45, 7217–7225. doi: 10.1021/es201243t

Nassar, R., Seleem, E. A., Caruso, G., Sekara, A., and Abdelhamid, M. T. (2020). The Nitrogen-fixing bacteria—effective enhancers of growth and chemical composition of Egyptian henbane under varied mineral n nutrition. *Agronomy* 10, 921. doi: 10.3390/agronomy10070921

Ofek, M., Hadar, Y., and Minz, D. (2012). Ecology of root colonizing Massilia (Oxalobacteraceae). *PLoS One* 7, e40117. doi: 10.1371/journal.pone.0040117

Pankievicz, V. C. S., Irving, T. B., Maia, L. G. S., and An,é, J. M. (2019). Are we there yet? The long walk towards the development of efficient symbiotic associations between nitrogen-fixing bacteria and non-leguminous crops. *BMC Biol*. 17, 99. doi: 10.1186/s12915-019-0710-0

Pereira, L. B., Andrade, G. S., Meneghin, S. P., Vicentini, R., and Ottoboni, L. M. M. (2020). Prospecting plant growth promoting bacteria isolated from the rhizosphere of sugarcane under drought stress. *Curr. Microbiol*. 76, 1345–1354. doi: 10.1007/s00284-019-01749-x

Pisa, G., Magnani, G. S., Weber, H., Souza, E. M., Faoro, H., Monteiro, R. A., et al. (2011). Diversity of 16S rRNA genes from bacteria of sugarcane rhizosphere soil. *Braz. J. Med. Biol. Res*. 4, 1215–1221. doi: 10.1590/S0100-879X2011007500148

Poly, F., Monrozier, L. J., and Bally, R. (2001). Improvement in the RFLP procedure for studying the diversity of nifH genes in communities of nitrogen fixers in soil. *Res. Microbiol*. 152, 95–103. doi: 10.1016/S0923-2508(00)01172-4

Prashar, P., Kapoor, N., and Sachdeva, S. (2014). Rhizosphere: its structure, bacterial diversity and significance. *Rev. Environ. Sci. Biotechnol*. 13, 63–77. doi: 10.1007/s11157-013-9317-z

Ranjard, L., Poly, F., and Nazaret, S. (2000). Monitoring complex bacterial communities using culture-independent molecular techniques: application to soil environment. *Res. Microbiol*. 151, 167–177. doi: 10.1016/S0923-2508(00)00136-4

Reis, V., Lee, S., and Kennedy, C. (2007). "Biological nitrogen fixation in sugarcane," in *Associative and Endophytic Nitrogen-fixing Bacteria and Cyanobacterial Associations*, vol 5, eds C. Elmerich, and W. E. Newton (Dordrecht: Springer), pp. 213–232.

Reuter, J. A., Spacek, D. V., and Snyder, M. P. (2015). High-throughput sequencing technologies. *Mol. Cell*. 58, 586–597. doi: 10.1016/j.molcel.2015.05.004

Rosa, P. A. L., Mortinho, E. S., Jalal, A., Galindo, F. S., Buzetti, S., Fernandes, G. C., et al. (2020). Inoculation with growth-promoting bacteria associated with the reduction of phosphate fertilization in sugarcane. *Front. Environ. Sci*. 8, 32. doi: 10.3389/fenvs.2020.00032

Rosenblueth, M., Ormeño-Orrillo, E., López-López, A., Rogel, M. A., Reyes-Hernández, B. J., Martínez-Romero, J. C., et al. (2018). Nitrogen fixation in cereals. *Front. Microbiol*. 9, 1794. doi: 10.3389/fmicb.2018.01794

Santchurn, G., Badaloo, M. G. H., Zhou, M., and Labuschagne, M. T. (2019). Contribution of sugarcane crop wild relatives in the creation of improved varieties in Mauritius. *Plant Genet. Resour*. 17, 151–163. doi: 10.1017/S1479262118000552

Schultz, N., Pereira, W., Silva, P., d, A., Baldani, J. I., Boddey, R. M., et al. (2017). Yield of sugarcane varieties and their sugar quality grown in different soil types

and inoculated with a diazotrophic bacteria consortium. *Plant Produc. Sci.* 20, 366–374. doi: 10.1080/1343943X.2017.1374869

Sessitsch, A., Hardoim, P., Döring, J., Weilharter, A., Krause, A., and Woyke, T. (2012). Functional characteristics of an endophyte community colonizing rice roots as revealed by metagenomic analysis. *Mol. Plant-Microbe Interact.* 25, 28–36. doi: 10.1094/MPMI-08-11-0204

Shoko, M. D., Tagwira, F., and Zhou, M. (2007). The potential of reducing nitrogen fertilizers in a soybean-sugarcane production system in Zimbabwe. *Afr. J. Agric. Res.* 2, 475–480. Available online at: https://academicjournals.org/journal/AJAR/article-full-text-pdf/D6D871A33883 (accessed July 26, 2007).

Singh, R. K., Singh, P., Li, H. B., Song, Q. Q., Guo, D. J., Solanki, M. K., et al. (2020). Diversity of nitrogen-fixing rhizobacteria associated with sugarcane: a comprehensive study of plant-microbe interactions for growth enhancement in Saccharum spp. *BMC plant Biol.* 20, 220. doi: 10.1186/s12870-020-02400-9

Sokal, M. A. (1958). Statistical method for evaluating systematic relationships. *Univ. Kansas Sci. Bull.* 38, 1409–1438.

Solanki, M. K., Wang, F. Y., Li, C. N., Wang, Z., Lan, T. J., Singh, R. K., et al. (2019). Impact of sugarcane–legume intercropping on diazotrophic microbiome. *Sugar Tech.* 22, 52–64. doi: 10.1007/s12355-019-00755-4

Solanki, M. K., Wang, F. Y., Wang, Z., Li, C. N., Lan, T. J., Singh, R. K., et al. (2018). Rhizospheric and endospheric diazotrophs mediated soil fertility intensification in sugarcane-legume intercropping systems. *J. Soil Sedi.* 19, 1911–1927. doi: 10.1007/s11368-018-2156-3

Solanki, M. K., Wang, Z., Wang, F. Y., Li, C. N., Lal, C. G., Singh, R. K., et al. (2020). Assessment of diazotrophic proteobacteria in sugarcane rhizosphere when intercropped with legumes (Peanut and Soybean) in the field. *Front. Microbiol.* 11, 1814. doi: 10.3389/fmicb.2020.01814

Sun, S., Chen, Y., Cheng, J., Li, Q., Zhang, Z., and Lan, Z. (2018). Isolation, characterization, genomic sequencing, and GFP-marked insertional mutagenesis of a high-performance nitrogen-fixing bacterium, Kosakonia radicincitans GXGL-4A and visualization of bacterial colonization on cucumber roots. *Folia Microbiol.* 63, 789–802. doi: 10.1007/s12223-018-0608-1

Swarnalakshmi, K., Yadav, V., Tyagi, D., Dhar, D. W., Kannepalli, A., and Kumar, S. (2020). Significance of plant growth promoting rhizobacteria in grain legumes: growth promotion and crop production. *Plants* 9, 1596. doi: 10.3390/plants9111596

Teixeira, L. C., Peixoto, R. S., Cury, J. C., Sul, W. J., Pellizari, V. H., Tiedje, J., et al. (2010). Bacterial diversity in rhizosphere soil from Antarctic vascular plants of Admiralty Bay, maritime Antarctica. *ISME J.* 4, 989–1001. doi: 10.1038/ismej.2010.35

Tsegaye, Z., Gizaw, B., Tefera, G., Feleke, A., Chaniyalew, S., Alemu, T., et al. (2019). Isolation and biochemical characterization of Plant Growth Promoting (PGP) bacteria colonizing the rhizosphere of Tef crop during the seedling stage. *J. Plant Sci. Phytopathol.* 3, 013–027. doi: 10.29328/journal.jpsp.1001027

Uroz, S., Ioannidis, P., Lengelle, J., Cébron, A., Morin, E., Buée, M., et al. (2013). Functional assays and metagenomic analyses reveals differences between the microbial communities inhabiting the soil horizons of a Norway spruce plantation. *PLoS One* 8:e55929. doi: 10.1371/journal.pone.0055929

van Loon, L. C. (2007). Plant responses to plant growth promoting bacteria. *Euro. J. Plant Pathol.* 119, 243–254. doi: 10.1007/s10658-007-9165-1

Velineni, S., and Brahmaprakash, G. P. (2011). Survival and phosphate solubilizing ability of Bacillus megaterium in liquid inoculants under high temperature and desiccation stress. *J. Agric. Sci. Technol.* 13, 795–802.

Walkley, A., and Black, I. A. (1934). An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Sci.* 37, 29–38. doi: 10.1097/00010694-193401000-00003

Wang, Z., Solanki, M. K., Yu, Z. X., Yang, L. T., An, Q. L., Dong, D. F., et al. (2019). Draft genome analysis offers insights into the mechanism by which Streptomyces chartreusis wzs021 increases drought tolerance in sugarcane. *Front. Microbiol.* 9, 3262. doi: 10.3389/fmicb.2018.03262

Wei, Y. J., Wu, Y., Yan, Y. Z., Zou, W., Xue, J., Ma, W. R., et al. (2018). High-throughput sequencing of microbial community diversity in soil, grapes, leaves, grape juice and wine of grapevine from China. *PloS one* 13:e0193097. doi: 10.1371/journal.pone.0193097

Xu, D., Yu, X., Yang, J., Zhao, X., and Bao, Y. (2020). High-throughput sequencing reveals the diversity and community structure in rhizosphere soils of three endangered plants in western ordos, China. *Curr. Microbiol.* 77, 2713–2723. doi: 10.1007/s00284-020-02054-8

Yoneyam, T., Terakado-Tonooka, J., and Minamisawa, K. (2017). Exploration of bacterial N2-fixation systems in association with soil-grown sugarcane, sweet potato, and paddy rice: a review and synthesis. *Soil Sci. Plant Nutri.* 63, 578–590. doi: 10.1080/00380768.2017.1407625

Zhang, B., Penton, C. R., Xue, C., Wang, Q., Zheng, T., and Tiedje, J. M. (2015). Evaluation of the Ion Torrent Personal Genome Machine for gene-targeted studies using amplicons of the nitrogenase gene nifH. *Appl. Environ. Microbiol.* 81, 4536–4545. doi: 10.1128/AEM.00111-15

Zhang, M. Y., Xu, Z. H., Teng, Y., Christie, P., Wang, J., Ren, W. J., et al. (2016). Non-target effects of repeated chlorothalonil application on soil nitrogen cycling: the key functional gene study. *Sci. Tot. Environ.* 543, 636–643. doi: 10.1016/j.scitotenv.2015.11.053

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

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

frontiers in Plant Science

REVIEW
published: 17 March 2022
doi: 10.3389/fpls.2022.860281

Genome Engineering Technology for Durable Disease Resistance: Recent Progress and Future Outlooks for Sustainable Agriculture

Qurban Ali[1,2], Chenjie Yu[2], Amjad Hussain[3], Mohsin Ali[3], Sunny Ahmar[4], Muhammad Aamir Sohail[3], Muhammad Riaz[5], Muhammad Furqan Ashraf[6], Dyaaaldin Abdalmegeed[2,7], Xiukang Wang[8], Muhammad Imran[9], Hakim Manghwar[10] and Lei Zhou[1]**

[1]State Key Laboratory for Managing Biotic and Chemical Threats to the Quality and Safety of Agro-products, Institute of Agro-product Safety and Nutrition, Zhejiang Academy of Agricultural Sciences, Hangzhou, China, [2]Key Laboratory of Monitoring and Management of Crop Disease and Pest Insects, College of Plant Protection, Ministry of Education, Nanjing Agricultural University, Nanjing, China, [3]College of Plant Science and Technology, Huazhong Agricultural University, Wuhan, China, [4]Institute of Biology, Biotechnology, and Environmental Protection, Faculty of Natural Sciences, University of Silesia, Katowice, Poland, [5]State Key Laboratory for Conservation and Utilization of Subtropical Agro-bioresources, Root Biology Center, College of Natural Resources and Environment, South China Agricultural University, Guangzhou, China, [6]State Key Laboratory of Subtropical Silviculture, Zhejiang A&F University, Hangzhou, China, [7]Department of Botany and Microbiology, Faculty of Science, Tanta University, Tanta, Egypt, [8]College of Life Sciences, Yan'an University, Yan'an, China, [9]Key Laboratory for Conservation and Utilization of Subtropical Agro-Bioresources, College of Agriculture, South China Agriculture University, Guangzhou, China, [10]Lushan Botanical Garden, Chinese Academy of Sciences, Jiujiang, China*

OPEN ACCESS

Edited by:
Sigfredo Fuentes, The University of Melbourne, Australia

Reviewed by:
Wajid Zaman, University of Chinese Academy of Sciences, China
Rajesh Yarra, University of Florida, United States
Giseli Valentini, Agricultural Research Service (USDA), United States

***Correspondence:**
Lei Zhou
zhoul@zaas.ac.cn
Hakim Manghwar
hakim@lsbg.cn

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

Received: *22 January 2022*
Accepted: *22 February 2022*
Published: *17 March 2022*

Citation:
Ali Q, Yu C, Hussain A, Ali M, Ahmar S, Sohail MA, Riaz M, Ashraf MF, Abdalmegeed D, Wang X, Imran M, Manghwar H and Zhou L (2022) Genome Engineering Technology for Durable Disease Resistance: Recent Progress and Future Outlooks for Sustainable Agriculture. Front. Plant Sci. 13:860281. doi: 10.3389/fpls.2022.860281

Crop production worldwide is under pressure from multiple factors, including reductions in available arable land and sources of water, along with the emergence of new pathogens and development of resistance in pre-existing pathogens. In addition, the ever-growing world population has increased the demand for food, which is predicted to increase by more than 100% by 2050. To meet these needs, different techniques have been deployed to produce new cultivars with novel heritable mutations. Although traditional breeding continues to play a vital role in crop improvement, it typically involves long and laborious artificial planting over multiple generations. Recently, the application of innovative genome engineering techniques, particularly CRISPR-Cas9-based systems, has opened up new avenues that offer the prospects of sustainable farming in the modern agricultural industry. In addition, the emergence of novel editing systems has enabled the development of transgene-free non-genetically modified plants, which represent a suitable option for improving desired traits in a range of crop plants. To date, a number of disease-resistant crops have been produced using gene-editing tools, which can make a significant contribution to overcoming disease-related problems. Not only does this directly minimize yield losses but also reduces the reliance on pesticide application, thereby enhancing crop productivity that can meet the globally increasing demand for food. In this review, we describe recent progress in genome engineering techniques, particularly CRISPR-Cas9 systems, in development of disease-resistant crop plants. In addition, we describe the role of CRISPR-Cas9-mediated genome editing in sustainable agriculture.

Keywords: plant pathogen, genome editing, CRISPR-Cas system, pesticide, disease resistance

INTRODUCTION

Phytopathogens are one of the most common causes of plant diseases and pose a threat to global agricultural prosperity, as well as the safety of agro-based products. Plant diseases, caused by phytopathogenic bacteria, fungi, nematodes, viruses, invertebrate pests, and weeds account for approximately 20–40% of losses in agricultural crop yields worldwide (Ali et al., 2021, 2022). In the past few years, advances in crop breeding have provided a number of new technologies in the food and agriculture industry. Crops not only provide food for human consumption but also provide fuel and animal feed. The world population is expected to reach 9.6 billion by 2050, with a rise in global food demand by 100 to 110% compared with that in 2005 (Savary et al., 2019). However, current reductions in the extent of cultivable arable land and increasing water deficiencies highlight the urgency for innovate genome editing technologies in crop breeding for sustainable agriculture production. Moreover, given the emergence of new pathogens and development of resistance in existing pathogens, plant breeders, pathologists, and horticulturists need to develop different approaches to produce new cultivars with novel heritable mutations. Although traditional breeding practices have for long played a vital role in crop improvement, these typically involve prolonged laborious artificial planting over multiple generations.

Compared with conventional breeding, genetic engineering, which entails the use of biotechnology for direct editing of the genetic material of organisms (Christou, 2013), has numerous benefits. First, it can facilitate the insertion, deletion, modification, disruption, or fine-tuning of particular genes of interest and causes minimal, if any, undesirable alterations in the remaining crop genome (Hussain et al., 2021). Moreover, crops with desired traits can be obtained within fewer generations. Second, genetic engineering requires the exchange of genetic material between species. Consequently, the initial genetic material that can be used in this phase is not restricted to a single organism (Dong and Ronald, 2019). Third, in the process of genetic modification, plant transformation can introduce new genes into vegetatively propagated crops, including cassava (*Manihot esculenta*), potato (*Solanum tuberosum*), and banana (*Musa* sp.). Collectively, genetic engineering, thus, represents a potentially effective approach for enhancing the resistance to plant pathogens (Dong and Ronald, 2019).

Numerous aspects of crop genetic engineering are dependent on either traditional transgenic techniques or new genome-editing technologies. Using traditional transgenic methods, genes encoding proteins associated with required agronomic characteristics are inserted into random positions within the genome *via* transformation processes (Lorence and Verpoorte, 2004). These methods typically generate variants containing foreign DNA. In contrast, genome editing facilitates the modification of endogenous plant DNA at specific targets *via* deletion/insertion and replacement of the requisite DNA fragments (Barrangou and Doudna, 2016). In certain parts of the world, including the United States (USDA, 2018), Argentina (Agriculture, Livestock, Fisheries and Food Secretariat, Argentina, 2015), and Brazil, genome-edited plants that do not contain foreign DNA are exempt from regulatory measures applicable to genetically modified plants (CTNBio, 2018) and, accordingly, have a status equivalent to that of crop plants developed using traditional breeding techniques (Orozco, 2018). Despite these differences in regulatory practices, however, both traditional transgenic and new genome editing strategies represent important crop enhancement methods.

During the course of evolution, plants have developed multi-layer protective mechanisms against microbial pathogens (Chisholm et al., 2006; Jones and Dangl, 2006). For example, pre-formed physiological barriers and their enhancement prevent possible pathogens from entering the cell (Uma et al., 2011). Moreover, plants can mount appropriate defensive responses triggered by the perception of physical pathogen contact mediated *via* plasma membrane-bound and intracellular immune receptors that recognize pathogen-derived elicitors or by indirect alteration of host targets (Zhang and Coaker, 2017; Kourelis and Van Der Hoorn, 2018).

Furthermore, plant-derived antimicrobial peptides and other compounds can inhibit pathogens by directly detoxifying or inhibiting virulence factor activity (Ahuja et al., 2012). Plants also initiate RNA silencing or RNA interference (RNAi) processes that detect invasive viral pathogens and cut targeted viral RNA (Rosa et al., 2018). However, pathogens have in turn evolved effective counter-strategies that enable them to circumvent host plant defensive responses. For example, numerous fungal and bacterial pathogens have been found to release cell wall-degrading enzymes (Kubicek et al., 2014), whereas when within the host cytoplasm, certain pathogen-derived effectors (Franceschetti et al., 2017) can inhibit host defenses or promote susceptibility. Moreover, it has been established that almost all plant viruses have developed RNAi inhibitors to counter RNAi-based host defense responses, as in the case of viruses that can also hijack the host RNAi system to silence host genes, thereby enhancing viral pathogenicity (Wang et al., 2012).

These host–microbe interactions, thus, provide important clues for disease resistance-targeted genetic engineering (Dangl et al., 2013). For instance, genes encoding proteins that break down mycotoxins (Karlovsky, 2011), inhibit enzymic cell wall degradation, or species of nucleic acid that can isolate inhibitors of the RNA virus (Wang et al., 2012) can be inserted into plants to reduce microbial virulence. Furthermore, plants can be engineered to synthesize and secrete antimicrobial compounds that specifically inhibit pathogen colonization (Dong and Ronald, 2019), whereas by targeting viral RNA for degradation, plant RNAi mechanisms can be manipulated to confer high viral immunity (Rosa et al., 2018). In addition, to enhance the robustness and widen the spectrum of disease tolerance, natural or edited immune receptors that recognize different pathogen strains can be inserted individually or in combination (Fuchs, 2017), and basic defense hub regulatory genes can be reprogrammed for the fine-tuning of defense responses (Pieterse and Van Loon, 2004). Similarly, genetic engineering can be used to generate host bait proteins that trap pathogens, thereby altering pathogen identification specificity (Malnoy et al., 2016).

For an additional detailed summary of the aspects of gene-edited disease-resistant plants, please refer to the review article previously published by Cook et al. (2015). In this review, we cover recent developments in the engineering of plant resistance to microbial pathogens based on the molecular mechanisms underlying plant–pathogen interactions and describe recent biotechnological advances. In the following sections, we also provide an overview of the breakthroughs in plant genetic engineering aimed at enhancing disease resistance and highlight some of the techniques that have proved promising in field trials.

ZFN, TALEN, AND CRISPR-Cas GENOME-EDITING TECHNIQUES

Genome editing entails site-specific genome targeting, which is used to modify the genomic DNA of plant or animal cells with high precision and efficiency. Here, we compare three of the most widely used genome editing technologies. Among these, zinc finger nuclease (ZFN)-based modification, which is based on the use of programmable nucleases, is considered the first breakthrough in the field of genome engineering (Chandrasegaran and Carroll, 2016; **Figure 1**). The transcription activator-like effector nuclease (TALEN) editing using the bacterial transcription activator-like effector (TALE) is known to expand the potential utility of genome editing (**Figure 1**). However, the most recently developed clustered regularly spaced short palindrome repeats (CRISPR)-Cas system has attracted the most remarkable attention from researchers worldwide, on the account of its simplicity, ease of use, high efficiency, and ability to allow transgene production (Mali et al., 2013). In different organisms, including plants, application of the CRISPR-Cas9 system has rapidly surpassed the ZFN and TALEN systems (Mali et al., 2013). Unlike ZFN and TALEN systems, CRISPR-Cas9, a system that uses protein motifs for target recognition, relies on DNA/RNA recognition to generate double-strand breaks (DSBs). CRISPR-Cas9 can be considered superior to ZFNs and TALENs in the following respects: (i) simplicity of target design, (ii) efficiency of Cas9 protein and gRNA, (iii) the ease with which simultaneous targeted mutations can be generated in multiple genes (Ma et al., 2015; Malzahn et al., 2017), and (iv) vector design, given difficulties in the usability and access to developed bioinformatics tools (Ahmad et al., 2020). However, despite the multiple advantages of the CRISPR-Cas technology and significant developments to date, the technology still warrants further improvements.

Since its introduction, in recent years, the CRISPR system has been undergoing continual modifications, such as CRISPR-Cas12a (Shimatani et al., 2017) and base editing tools (Bharat et al., 2019; Butt et al., 2020), for easier use and adaptability in response to different constraints. The developed SpCas9 variant can target the expanded NGN protospacer adjacent motif (PAM), and the enzyme has been optimized to produce a near-PAMless SpCas9 variant, referred to as SpRY (NRN > NYN PAM), with SpRY nuclease and base editing variants being able to target almost all PAMs (Walton et al., 2020). Currently, the low efficiency of the homologous recombination pathway (knock-in/gene replacement) and the similar low efficiency of the transformation of homologous donor sequences in plant cells have contributed to the complexity and poor efficiency of knock-in mutations (Lu et al., 2020). Consequently, an efficient gene knock-in procedure in plants based on CRISPR-Cas-mediated homologous recombination is still required. Furthermore, the current CRISPR-Cas9 system has minimal effects with respect to the control of both RNA and DNA viruses, thereby highlighting the need to develop a useful and acceptable CRISPR system to overcome these current limitations. In this regard, recent findings have indicated that Cas13 proteins (Cas13a, Cas13b, and Cas13c) have considerable potential applicability as robust RNA regulators of RNA viruses (Abudayyeh et al., 2017; Zhang et al., 2018a). For example, CRISPR-Cas13a has been shown to confer RNA virus resistance in both monocot and dicot plants (Zhang et al., 2019a). Targeted-site gene editing has also been performed to design eIF4E resistance alleles that play important roles in virus resistance (Bastet et al., 2018, 2019).

Recently, a new genome editing technology, "prime editing," has been developed, which can be employed to perform different types of editing, such as particular base-to-base transfers, including all transformation (C → T, G → A), (A → G and T → C), and transverse (C → A, C → G, G → C, G → T, A → C, A → T, T → A, and T → G) mutations, along with small-scale insertion/deletions, without causing DNA double-strand breakage. Given that the prime editing system has sufficient versatility to complete specific forms of genome editing, these new developments have considerable potential. Moreover, it can be modified for different purposes (including crop production, resistance to abiotic and biotic stresses, and crop plant quality improvement; Yasmin et al., 2020; Ayaz et al., 2021a). Moreover, the Cas12b (C2c1) system has recently been successfully applied for editing genomes in cotton plants to enhance the resistance to high temperatures, thereby paving the way for the development of varieties that can be cultivated under heat stress conditions (Wang et al., 2020). Consequently, by facilitating sustainable agricultural crop production, application of the CRISPR system is predicted to make a significant contribution to overcome food scarcity and ensure global food security.

RESISTANCE GENES DEPLOYED FOR BROAD-SPECTRUM DURABLE RESISTANCE

During the early 1940s, innovative genetic studies examining the plant–pathogen interaction between flax and the flax rust fungus *Melampsora lini* were conducted (Flor, 1956), which contributed to the development of the "gene-for-gene" theory, proposing that in plant–pathogen interactions, a host plant resistance (*R*) locus matches an avirulence factor (Avr) in the pathogen (Flor, 1971). The theory maintains that as long as the *R* gene and the homologous Avr occur simultaneously, interactions between the plant and pathogen are incompatible,

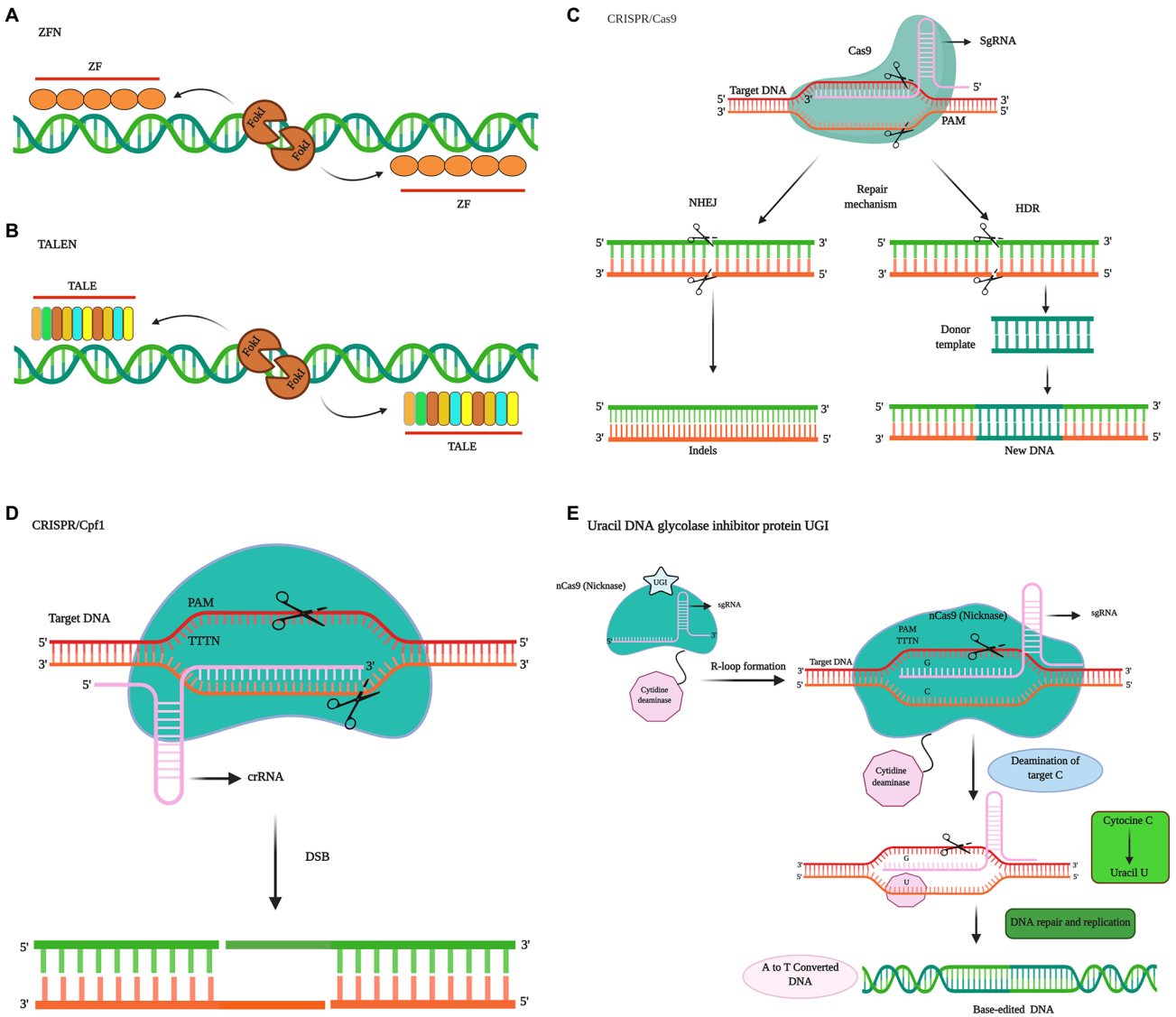


FIGURE 1 | Comparison of different widely used genome editing systems in plants. Site-specific techniques, such as those based on zinc finger nucleases (ZFNs; **A**) and protein-dependent DNA cleavage systems, such as those based on transcription activator-like effector nucleases (TALENs; **B**). The two sequence-specific zinc finger proteins flank the target sequence, and *FokI* nuclease follows the C terminus of each protein. Zinc finger proteins are used to direct *FokI* dimers to target the specific DNA sites to be cleaved. TALENs consist of two sequence-specific TALEN proteins arranged on the target sequence, with *FokI* nuclease followed by the C terminus of each protein. The TALEN protein directs *FokI* dimers to the target specific DNA sites to be cleaved. **(C)** Illustration of a cluster of the regularly spaced short palindrome repeats (CRISPR)/CRISPR-associated nuclease 9 (Cas9) system. CRISPR-Cas induces double-strand breaks (DSB) in the DNA strand. CRISPR RNA (crRNA) directs the Cas9 protein. Trans-activating CRISPR RNA (tracrRNA) is used to stabilize the structure and activate Cas9 to cut the target DNA. Single guide RNA [sgRNA (Gulabi)] recognizes the target gene, and then the Cas9 protein (green) cuts the two strands of the target DNA at the RuvC and His Asn-His (HNH) domains. The cutting of DNA by Cas9 is dependent on the presence of spacer adjacent motif (PAM) sequences. A 20-base sgRNA defines the binding specificity. In DNA, there are two mechanism of double-strand break (DSB) repair, namely, homology-directed repair (HDR), which is activated in the presence of a template and leads to knock-in or gene replacement, and non-homologous end joining (NHEJ), which is not sufficiently precise and leads to permanent gene knockout. **(D)** The Cpf1 system. Cpf1 is a two-component RNA programmable DNA nuclease related to CRISPR. The PAM of Cpf1 is TTTN. Cpf1 cuts the target DNA and introduces a DSB, a 5-nt staggered cut at the 5′ end of the T-rich PAM. CRISPR-Cas9 nickase introduces breaks only in the strand complementary to sgRNA. The double nickase with two sgRNAs introduces a staggered DSB into the DNA, after which HDR repairs the DSB. Uracil DNA glycosylase inhibitor (UGI) protein prevents removal of uracil from the DNA and subsequent repair pathways and contributes to increasing the frequency of mutations.

and the host has complete pathogen resistance (Flor, 1971). At the beginning of the 20th century, the efficacy of *R* gene-mediated resistance was first established in wheat (*Triticum* sp.) by the British scientist Rowland Biffen (Biffen, 1905). Subsequently, a number of different *R* genes have been identified, emulated, and introduced into several different varieties related to the same species, then introduced into other species of the same genera (Song et al., 1995), and ultimately across genera

(Tai et al., 1999). For example, rice (*Oryza sativa*) plants expressing an *R* gene (*Rxo1*) derived from maize (*Zea mays*) were found to show resistance to bacterial streak in a laboratory environment (Hussain et al., 2019).

Similarly, multi-year trials have revealed that tomatoes (*Solanum lycopersicum*) expressing the pepper *Bs2 R* gene maintained durable field resistance to bacterial spot (*Xanthomonas* sp.) disease (Horvath et al., 2015; Kunwar et al., 2018), and under field conditions, transgenic wheat expressing different alleles of the wheat resistance locus *Pm3* showed race-specific resistance against stem rust (*Puccinia graminis* f. sp. *tritici*; Brunner et al., 2012). Furthermore, the overexpression of *R* genes, such as *Rpi-vnt1.1* or *RB* from wild potato, has been shown to confer durable resistance to potato late blight (*Phytophthora infestans*) in commercial potatoes (Jones et al., 2014; Ayaz et al., 2021b). Notably, to date, genetically modified transgenic potato overexpressing the *Rpi-vnt1.1* gene, which has enhanced resistance to potato late blight, is the only crop that has been approved for commercial use (Dong and Ronald, 2019).

Given that pathogens have the capacity to evade detection based on host *R* genes (Jones and Dangl, 2006), the disease resistance conferred by a single *R* gene typically lacks durability under field conditions, as pathogens can evolve alternative virulent forms *via* Avr gene mutation. Consequently, to obtain broad-spectrum disease resistance and thereby ensure long-lasting field resistance, multiple *R* genes are generally introduced simultaneously, a procedure referred to as gene stacking (Fuchs, 2017; Mundt, 2018). Resistance based on the stacking of *R* genes is anticipated to be both broad-spectrum and durable, given that pathogen strains are unlikely to overwhelm the resistance conferred by multiple *R* genes.

A well-established methodology for stacking *R* genes at pre-existing *R* loci is cross-breeding (**Figure 2**), using which, breeders can identify and select the desired progeny with requisite *R* genes, based on marker-assisted selection (Das and Rao, 2015). An example is the bacterial blight pathogen of rice (*X. oryzae* pv. *oryzae*), a lethal and devastating disease in some areas of Africa and Asia (Niño-Liu et al., 2006), for which cross-breeding has been performed to introduce three stacked *R* genes (*Xa21*, *Xa5*, and *Xa13*) and was subsequently established to contribute to resistance against this disease. These three stacked genes have also been cloned and introduced into Jalmagna (a deep-water rice cultivar; Pradhan et al., 2015), which was found to show significant durable resistance against eight *X. oryzae* isolates under field conditions (Pradhan et al., 2015).

However, despite the impressive results obtained using stacked resistance genes, the selection process can be excessively time-consuming and laborious when selection is based on the identification of a large number of loci. As a more viable alternative to gene stacking, researchers can assemble several *R* gene cassettes into a plasmid and then introduce the entire *R* gene cluster into specific genetic site *via* plant transformation (Que et al., 2010; Solangi et al., 2021). In this way, all *R* genes are inherited as a single genetic locus, thereby minimizing the time needed for selection, as illustrated by the molecular stacking of three broad-spectrum potato late blight *R* genes (*Rpi-blb3*, *Rpi-sto1*, and *Rpi-vnt1.1*) *via* *Agrobacterium*-mediated transformation of a susceptible potato cultivar (Zhu et al., 2012). Under controlled greenhouse conditions, this tri-transgenic potato cultivar simultaneously expressing the aforementioned three resistance genes was found to be characterized by wide-spectrum resistance, which is equivalent to the development of strain-specific resistance conferred by each of the three *Rpi* genes (Zhu et al., 2012). Similarly, *Agrobacterium*-mediated transformation has been used to simultaneously introduce *Rpi-vnt1.1* and *Rpi-sto1* into three distinct varieties of potato *via* a single DNA fragment insertion (Jo et al., 2014), which was found to confer broad-spectrum resistance to late blight. Notably, apart from the *R* genes, no external DNA, for example, in the inserted DNA fragment or a selectable marker gene, were inserted, which is considered to be advantageous with respect to transgene regulation (Jo et al., 2014). In the case of both the double- and triple-gene-stacked potatoes described here, resistance was verified under field conditions. Moreover, it has been established that the appropriate spatio-temporal deployment of potato cultivars containing *R* genes against late blight can minimize fungicide usage by more than 80% (Haverkort et al., 2016).

Similar to the aforementioned potato cultivars, a recent study conducted on African highland potato varieties in Uganda showed that the molecular stacking of three *R* genes (*Rpi-vnt1.1*, *RB*, and *Rpi-blb2*) was observed to confer durable field resistance to the potato late blight pathogen (Ghislain et al., 2019). In addition, the yield of these modified potato cultivars was established to be three times higher than the national average. These findings accordingly serve to emphasize that the gene stacking approach not only confers durable field resistance but has no detrimental impact with respect to crop yields (Ghislain et al., 2019). Moreover, these studies highlight the simplicity and efficacy of the molecular stacking used to confer broad-spectrum disease resistance in important vegetatively propagated crop organisms, for which more conventional breeding-based stacking is not practical (**Figure 2**). However, despite the notable benefits of molecular stacking, the performance of the stacked genes is highly dependent on the respective vectors (Que et al., 2010). Selection of appropriate vectors will facilitate the insertion of exogenous DNA sequences into plant genomes at specified targets and also enable the introduction of multiple *R* gene cassettes in the vicinity of an established *R* gene cluster (Ainley et al., 2013). In this regard, the latest innovations in genome engineering have contributed to the development of the targeted insertion of DNA segments with desired features, which can be used to incorporate diverse traits in complex crop species (Ramu et al., 2016; Voytas, 2017).

The field of genome engineering is undergoing continuous evolution and is currently in a phase of heightened activity and frequent groundbreaking developments. We can accordingly anticipate a continuous stream of new innovations that will contribute to enhance the efficiency of targeted insertion and reduce the size of inserted DNA fragments, with applications in numerous plant species. Moreover, further developments and breakthroughs in specific gene insertions will provide new opportunities to stack larger numbers of *R* genes and engineer broad-spectrum viral resistance by altering a single locus, which not only offers convenience with respect to breeding but also confers durable disease resistance.

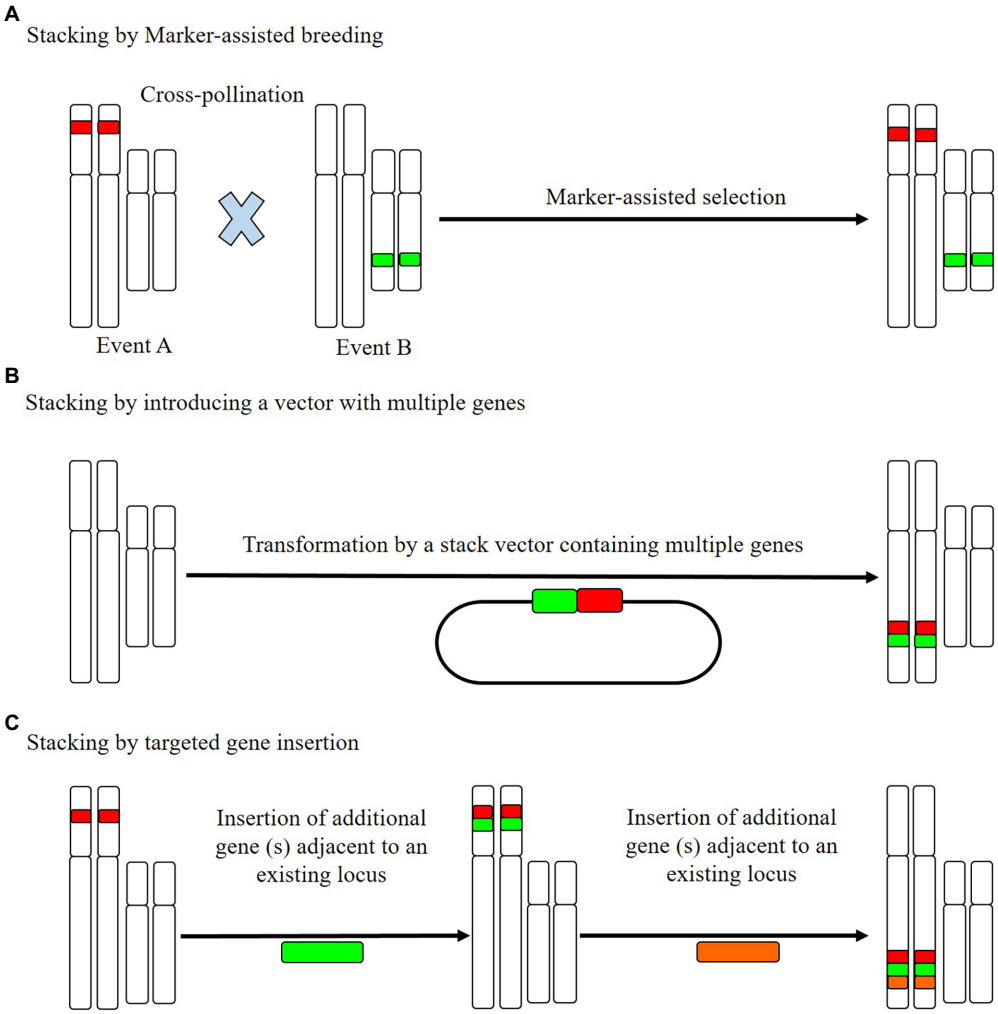


FIGURE 2 | The resistance (*R*) gene stacking method. **(A)** Cross-pollinated individuals, based on established trait loci, stack marker-assisted breeding and combine the combined trait loci for marker-assisted selection of the offspring. **(B)** A single transgene stacking event can be completed by integrating multiple genes into a single stack vector and introducing these simultaneously. **(C)** Targeted insertion stacking aims to insert new genes near established loci. To stack a large number of genes, this process can be repeated. The stacked genes in B and C are genetically related, such that they can be readily inserted at a single site.

CRISPR-Cas-INDUCED MUTATIONS AND ENHANCED BIOTIC RESISTANCE BASED ON TRANSGENE-FREE METHODS

Although sequence-specific nucleases, such as CRISPR-Cas, play an important role in transgene manipulations (for example, when the T-DNA in *Agrobacterium tumefaciens* provides CRISPR-Cas9 and sgRNA), the induced mutations may be genetically independent of the site of construct integration. This means that even if the CRISPR process involves an intermediate transgenic state, it is possible to produce genetically modified-free crops by merely isolating the mutation site from the insertion site using CRISPR-Cas techniques. In this regard, PCR can be used to characterize individual plants and determine the sequences of entire genomes, which is beneficial with respect to assessing mutations at the target site and also in identifying potential off-target mutations (Kim and Kim, 2016). Although in the United States, CRISPR-modified crops have been evaluated *via* product-based legislation and are not covered by genetically modified organism principles (Waltz, 2018), the current situation in Europe tends be more complex, in that regulation in European countries is dependent on process-based legislation (Sprink et al., 2016). However, given that it is practically impossible to trace back to the initial induction of a small mutation (whether it be introduced naturally; chemically by ethyl methanesulfonate; *via* X-ray radiation, oligonucleotide-directed mutagenesis, or TALENs, or physically by CRISPR-Cas), the attempts to regulate CRISPR-mediated mutations in the European Union would appear to be unrealistic (Urnov et al., 2018). In this context, transfection of plant protoplasts with the ribonucleoprotein complex of any genetically modified intermediate consisting of Cas9 protein and sgRNA is considered

TABLE 1 | The application of diverse techniques for the improvement of different agronomic and disease resistance traits in crops species.

Crop species	Technology	Target genes	Result/Trait improvement	Reference
Triticum aestivum	TALEN	*TaMLO*	Powdery mildew resistance	Wang et al., 2014
Manihot esculenta	CRISPR-Cas9	*MenCBP-1/2*	Cassava brown streak virus resistance	Gomez et al., 2019
Hordeum vulgare	CRISPR-Cas9	*HvMorc1*	*Blumeria graminis*/*Fusarium graminearum* resistance	Kumar et al., 2018
Gossypium hirsutum	CRISPR-Cas9	*Gh14-3-3d*	*Verticillium dahlia* resistance	Zhang et al., 2018b
Nicotiana benthamiana	CRISPR-Cas9	*IR*, *C1*	Cotton leaf curl multan virus resistance	Yin et al., 2019a
Brassica napus	CRISPR-Cas9	*BnCRT1a*	*Verticillium longisporum* resistance	Pröbsting et al., 2020
Triticum aestivum	CRISPR-Cas9	*TaNFXL1*	*Fusarium graminearum* resistance	Brauer et al., 2020
Citrullus lanatus	CRISPR-Cas9	*Clpsk1*	*Fusarium oxysporum* resistance	Zhang et al., 2020a
Oryza sativa	CRISPR-Cas9	*OsCul3a*	*Xanthomonas oryzae/Magnaporthe oryzae* resistance	Gao et al., 2020
Oryza sativa	CRISPR-Cas9	*OsPi21*, *OsXa13*	*Magnaporthe oryzae/Xanthomonas oryzae* resistance	Li et al., 2019a
Solanum tuberosum	TALEN	*ALS*	Herbicide resistance	Butler et al., 2016
Zea mays	ZFN	*PAT*	Herbicide resistance	Schornack et al., 2006
Gossypium hirsutum	CRISPR-Cas9	*ALARP*	Cotton fiber development	Sander and Joung, 2014
Manihot esculenta	CRISPR-Cas9	*eIF4E isoforms nCBP-1 & nCBP-2*	*Cassava brown streak virus*	Gomez et al., 2019
Oryza sativa	CRISPR-Cas9	*SWEET11*, *SWEET13* and *SWEET14/promoter*	*Xanthomonas oryzae* pv. *Oryzae* resistance	Oliva et al., 2019
Oryza sativa	CRISPR-Cpf1	*OsEPFL9*	Regulation of stomatal density	Yin et al., 2019b
Citrus	CRISPR-Cas9	*Citrus CsLOB1*	*Xanthomonas citri* subsp. *Citri* resistance	Jia et al., 2017
Solanum lycopersicum	CRISPR-Cas9	*Solyc08g075770*	*Fusarium oxysporum* f. sp. *lycopersici,*	Prihatna et al., 2018
Cucumis sativus	CRISPR-Cas9	*elf4E/cds*	Resistance to CVYV, ZYMV, and PRSMV	Chandrasekaran et al., 2016
Glycine max	CRISPR-Cas9	*GmF3H1/2*, *FNSII-1*	Soybean mosaic virus	Zhang et al., 2020b
Malus domestica	CRISPR-Cas9	*PDS*, *TFL1*	Albino phenotype, early flowering	Charrier et al., 2019
Vitis vinifera	CRISPR-Cas9	*VvPDS, MLO-7*	Albino phenotype	Nakajima et al., 2017
Musa spp.	CRISPR-Cas9	*ORF1, 2, 3* and *IR of BSV*	Resistance against Banana streak virus	Tripathi et al., 2019
Arabidopsis thaliana	CRISPR-Cas9	*eIF4E*	Transgene free resistant against Clover yellow vein virus	Bastet et al., 2019
Citrullus lanatus	CRISPR-Cas9	*Clpsk1*	Resistance to *Fusarium oxysporum f.* sp. *niveum*	Zhang et al., 2020a
Oryza sativa	CRISPR-Cas9	*EBEs of OsSWEET14*	Resistance to *Xanthomonas oryzae* pv. *oryzae*	Zafar et al., 2020
Capsicum annuum	CRISPR-Cas9	*CaERF28*	Anthracnose disease resistance	Mishra et al., 2021

a theoretical solution (Metje-Sprink et al., 2019). Successful attempts to circumvent problems associated with the genetically modified status of edited plant using "DNA-free" systems have previously been reported for a number of species, including *Chlamydomonas reinhardtii* (Baek et al., 2016), *Zea mays* (Svitashev et al., 2016), petunia (Subburaj et al., 2016), wheat (Liang et al., 2017), apple (*Malus domestica*) and grapevine (*Vitis vinifera*; Malnoy et al., 2016), lettuce (*Lactuca sativa*), rice, tobacco (*Nicotiana tabacum*), and *Arabidopsis thaliana*, and can contribute to further reductions in possible off-target mutations (**Table 1**).

Genes encoding certain susceptibility factors are also considered potentially useful targets for manipulation, as resistance can already be increased based on a simple knockout (Zaidi et al., 2018). For example, in some plant species, it has been established that plants harboring recessive MLO alleles are characterized by resistance to powdery mildew, as demonstrated by the modification of MLO in tomatoes (Nekrasov et al., 2017). TALENs and genome editing provide a basis for the so-called new breeding technology involving CRISPR-Cas editing, which has been used to mutate all six wheat MLO alleles, an in the targeted inactivation of existing alleles and genomes in a number of polyploidy crops, resulting in enhanced resistance to the *Blumeria graminis* f. sp. *tritici* fungal pathogen (Wang et al., 2014; **Table 1**). In rice, resistance to *X. oryzae* pv. *oryzae* has been obtained by mutating *SWEET* resistance genes or by engineering genes their promoters with a bacterial pathogen (Oliva et al., 2019; Xu et al., 2019).

Enhanced tolerance can also be achieved by deleting, rewriting, or inserting *cis* elements in the promotors of susceptibility or resistance genes. The *cis* elements targeted by effectors in susceptibility genes can be eliminated by taking advantage of the fact that the gene remains intact and can still perform its normal plant growth functions. Furthermore, TALENs have been used to modify *cis* elements in the promoter of the *OsSWEET14* gene in rice, which targets the rice blast *AvrXa7* gene. Although this reduction in *cis*-factor activity leads to fewer extreme disease symptoms, it is achieved *via* a TALEN-based method, which promote bacteria (Li et al., 2012). Although rewriting *cis* elements *via* homology-directed repair (HDR) using repair templates to modify the promoter is considered more difficult, it does, nevertheless, ensure that indels do not interfere with the spatial distribution of *cis* elements in promoters, which is necessary for the maintenance of appropriate host plant gene regulation. Moreover, the insertion of new *cis* elements into the promoters of defense-related genes can contribute to enhance the expression in response to new stress signals/pathogens.

However, a double mutation in the BIK1 protein (G230A/D231A) has been found to result in an flg22-induced defense signaling dominant-negative effect, thereby indicating that these two amino acid residues are functionally essential (Zhang et al.,

2010). Any of the hyper-phosphorylated amino acids may be substituted to dampen effector-mediated host susceptibility. The typical role of RIN4 in suppressing the immune response involves FLS2-activated serine 141 phosphorylation, and threonine 166 phosphorylation is mediated by the AvrB effector (Chung et al., 2014). Although AvrB has been observed to reduce pathogen-induced callus deposition in wild-type plants, this inhibitory effect is eliminated in mutant plants that express non-phosphorylated RIN4 (T166A). These findings, thus, emphasize the necessity of taking multiple factors into consideration. Generally, however, to enhance plant resistance, it should be feasible to substitute effectors with amino acids. Targeted-base editing of catalytically dead Cas9, such as the cytosine deaminase domain, has the potential to become a powerful tool, enabling molecular biologists to share unique amino acids that disable effectors and target interactions. These potential developments in CRISPR-Cas-based resistance engineering are also summarized in **Figure 3**. In all these cases, it is worth noting that resistance is race-specific against the pathogen that has deployed the corresponding effector. Some host genome modifications (mutation stacking) can also contribute to the development of new pathogen-resistant crop varieties and more efficient forms of plant immunity.

TRANSGENIC TECHNIQUES THAT USE CRISPR-Cas SYSTEMS TO ENHANCE DISEASE RESISTANCE

Genome editing also enables the transfer of heterologous genes involved in resistance pathway mechanisms. The distinction between this approach and T-DNA transfer methods is that the space between identified highly expressed protection genes can be precisely selected using CRISPR-Cas-HDR in response to the pressure exerted by pathogen strains. This alternative approach will predictably improve regulatory accessibility and increase expression of the genes conferred by transfer resistance. By using this strategy, it will also be possible to reduce the potential negative effects on other host genes in the vicinity of the site of integration (Schenke and Cai, 2020). Given that the flanking regions are known, string integration can typically detect positive events in such genomic regions without the necessity of using selectable markers, and thus, it is possible to perform PCR-based locus amplification and sequencing. The first generation of genetically modified organisms has often been criticized, notably on account of certain possible side effects attributable to the random selection of genetically modified genes and because antibiotic resistance genes were necessary for the detection of positive integration events (Gepts, 2002). These drawbacks can now be overcome, although it remains important to control the introduction of foreign DNA, a time-consuming and cost-intensive procedure, which normally has low consumer acceptance (Waltz, 2018).

CRISPR-Cas that can be used to transfer HDR-mediated resistance genes into the actively expressed chromatin regions of susceptible hosts to confer stress tolerance can contribute to the breeding of resistant varieties. This feature acquisition technique can be simplified using CRISPR-Cas genome editing. For example, it is possible to jointly regulate *R* genes closely linked in a head-to-head configuration to ensure that they can function in tandem to impart resistance. The expression of *R* genes in such regulatory modules will reduce the probability of imbalance and further minimize fitness costs (Karasov et al., 2017). In this regard, *R* genes, such as nucleotide-binding site leucine-rich repeat sequence (NLR) genes, typically function in pairs, including NLR accessory molecules and sensors that can recognize pathogenic effectors, although they do not control autoimmunity prevention mechanisms in the case of pathogens (Wu et al., 2017). By transferring *R* genes to crops, in which they can be expressed in response to biotic stress, pathogen resistance can be enhanced, although the number of appropriate *R* genes is limited and the identity of the interacting auxiliary NLR will need to be verified (Wang et al., 2019).

Furthermore, it has been established that dominant *R* gene-mediated drug resistance is not as stable as is recessive drug resistance and is typically confined to a minority of race-specific isolates that can readily be resolved by higher mutation rates (Kourelis and Van Der Hoorn, 2018). Therefore, achieving stronger resistance can be achieved by stacking/pyramiding different traits, and in this regard, CRISPR-Cas-induced HDR should be further encouraged (Pandolfi et al., 2017). Furthermore, pattern recognition receptor (PRR) transfer has been demonstrated to confer broader and more robust resistance in *Arabidopsis* and tomatoes (Lacombe et al., 2010; Saqib et al., 2020), banana and rice (Tripathi et al., 2014), and wheat (Manghwar et al., 2021). This is feasible owing to the maintenance of comparable immune signaling systems in monocots and dicots (Holton et al., 2015). The benefit of these approaches is that the modifications are tailored to specific pathogens, with defense responses being initiated when plants are threatened by the corresponding pathogens, thereby reducing potential losses of yield (Karasov et al., 2017). It is well established that by generating receptive chromatin regions, stress can induce a priming response, thereby promoting the more rapid expression of protective genes during secondary challenges by same stress source (Manghwar et al., 2022; Manghwar and Hussain, 2022). These genomic regions can accordingly have beneficial effects with respect to the expression of heterologous *R* genes or PRRs.

BASE EDITING FOR CONFERRING DISEASE RESISTANCE IN CROPS

Base editing technology has evolved as a revolutionary technique for genome editing in plants that is both efficient and effective. Despite being developed only recently, several papers have already been published describing the use of these techniques to enhance agronomic features and disease resistance in a range of agriculturally important crops. Base editing has shown considerable potential for trait development in rice, and this technique might also be applied in other monocot crops, such as maize, sugarcane, barley, wheat, and sorghum (Yarra and Sahoo, 2021). Base editing involves the use of a catalytically deficient CRISPR-Cas nuclease coupled to a nucleotide deaminase and, in some cases, DNA repair proteins. Using this approach, single-nucleotide variations can be engineered at specified loci

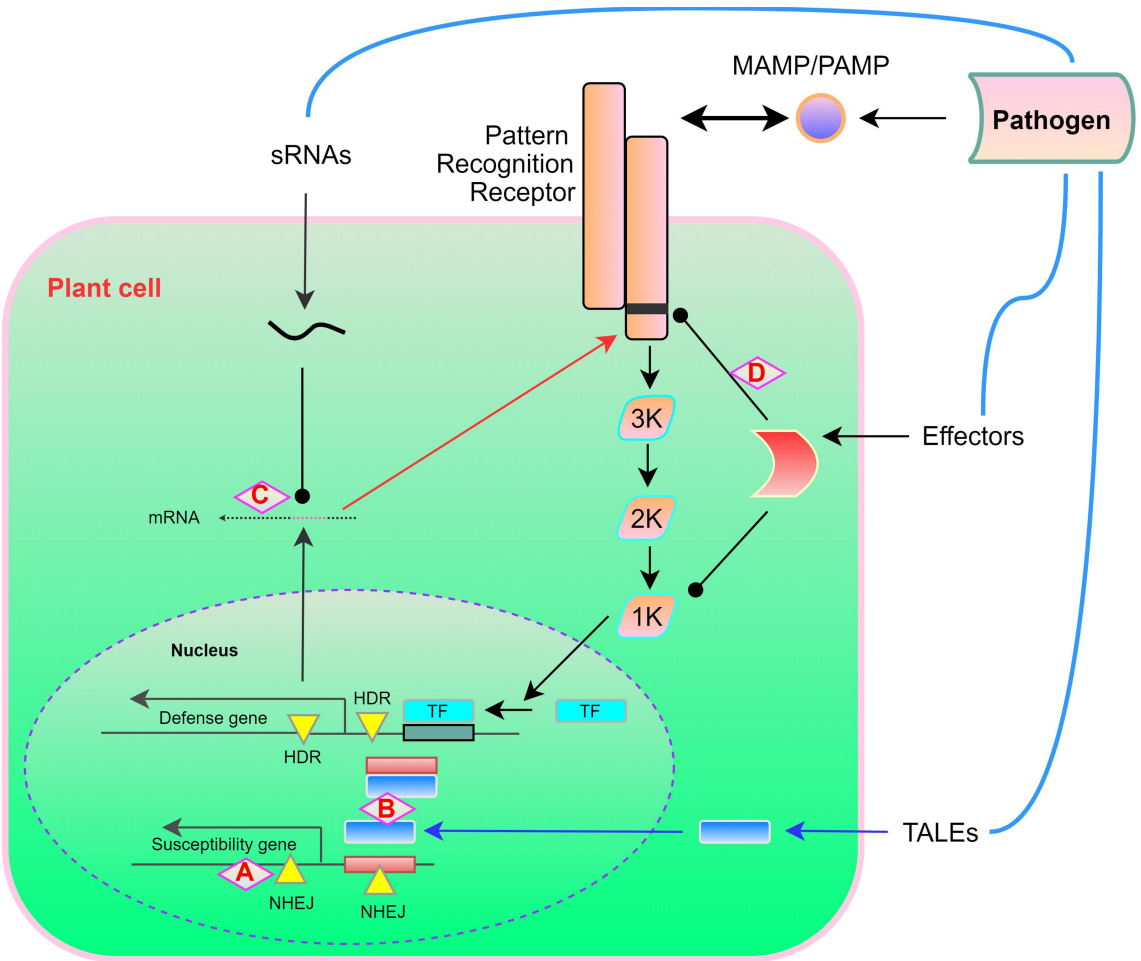


FIGURE 3 | A potential target for applying CRISPR-Cas for engineering of "non-transgenic" disease resistance in plants. In response to pathogen recognition, the plant-induced defense response involves a MAP kinase phosphorylation cascade, which leads to the activation of transcription factors and the expression of defense-related genes. **(A)** Open reading frame disruption by non-homologous end joining (NHEJ) causes frameshift mutations in susceptibility genes. **(B)** NHEJ eliminates *cis* elements to prevent transcription activator-like effector (TALE) activation of susceptibility genes and homology-directed repair (HDR) introduces *cis* elements, which promotes the TALE-triggered activation of defense genes. **(C)** Rewriting of transboundary RNAi *via* HDR-mediated pathogen siRNA targeting. **(D)** HDR-mediated coding sequence rewriting replaces the effectors produced by pathogens for amino acid residues required for protein recognition, thereby preventing, for example, cleavage or modification. The figures are adopted with modification from those presented by Schenke and Cai (2020).

in the DNA (nuclear or organellar) or RNA of both dividing and non-dividing cells. Base editors (BEs) can be divided into two categories, namely, DNA BEs, which can be used to introduce specific point mutations in DNA, and RNA Bes, which can alter single ribonucleotides. The different types of DNA BEs that are currently accessible include cytosine BEs (CBEs), adenine BEs (ABEs), C-to-G BEs (CGBEs), dual-base editing, and organellar BEs, among which, CBE and ABE, given their simplicity and efficacy in precise base editing, have been widely employed in numerous organisms, including plants, for both gene functional annotation and gene correction (Veillet et al., 2019b; Yarra and Sahoo, 2021).

Cytidine Base Editors

Cytidine base editors, the first-developed type of DNA base editor, are used to facilitate the conversion from C-G to T-A (Komor et al., 2016), and in two seminal investigations, CBEs with diverse topologies incorporating a Cas9 nickase (nCas9, for example, with a D10A mutation) coupled to cytidine deaminase and uracil glycosylase inhibitor (UGI) were described (Nishida et al., 2016; **Figure 4A**). Similar to the canonical CRISPR–Cas systems, CBEs are guided to target genomic region by a sgRNA. After binding to the target DNA, an sgRNA–CBE complex forms a single-stranded DNA R-loop (Jiang et al., 2016), and CBE cytidine deaminase, which catalyzes the hydrolytic deamination of an exposed cytosine, gains access to this non-target single-stranded DNA. C-to-T base editing is mediated *via* deamination and subsequent cellular mismatch repair. This process tends to be hampered by uracil (U) base excision repair (BER), which either regenerates the original base pair or results in indels (Porto et al., 2020). However, the action of UGI subverts BER and increases the likelihood of C-to-T

editing. Although antibiotics are commonly used to select transformants, it may be difficult to detect base-edited cells in a population. To circumvent this constraint, a surrogate reporter system based on the repair of a faulty hygromycin-resistance gene has been constructed in plants (Xu et al., 2020b). The editing of cytidine bases for enhancing crop agronomic traits and disease resistance has already found application in a range of plant species, including rice (Hua et al., 2019; Ren et al., 2021; Yarra and Sahoo, 2021); *Arabidopsis* (Chen et al., 2017; Li et al., 2019b); wheat (Zong et al., 2017; Zhang et al., 2019b); maize (Zong et al., 2017); potato and tomato (Veillet et al., 2019a; Hunziker et al., 2020); watermelon, cotton, soya bean, apple, and pear (Zhao et al., 2018; Cai et al., 2020; Qin et al., 2020; Malabarba et al., 2021); and strawberry and rapeseed (Xing et al., 2020; Cheng et al., 2021).

Adenine Base Editors

On the basis of the characterization of CBEs, it was assumed that a combination of adenine deaminase and nCas9 would give rise to ABEs, which could be used to convert an A-T base pair to a G-C base pair. However, none of the naturally occurring adenine deaminases have been found to work with DNA (Gaudelli et al., 2017; Molla et al., 2021; **Figure 4B**). However, Gaudelli et al. developed a single-stranded DNA-specific transfer RNA (tRNA) adenosine deaminase (TadA) variant using directed evolution and protein engineering (Gaudelli et al., 2017), into which mutations were introduced to generate an engineered version (TadA*). Given that TadA catalyzes deamination as a dimer, a heterodimeric protein containing a non-catalytic wild-type TadA monomer and a designed catalytic monomer (TadA*) was developed (Gaudelli et al., 2017; Rees and Liu, 2018). The fusion of this heterodimer (TadA–TadA*) with nCas9 was accordingly found to yield ABEs that can efficiently convert A to G in high-purity mammalian cells (Molla and Yang, 2019). In contrast to uracil excision repair, cellular inosine excision repair is comparatively weak and causes little interference with A-T to G-C conversions (Yarra and Sahoo, 2021). Consequently, no other glycosylase inhibitor protein is necessary in the development of ABEs (Molla et al., 2021). To date, ABE systems have been used to modify growth traits and disease resistance in *A. thaliana* and *Brassica napus* (Kang et al., 2018), rice (Molla et al., 2020; Yarra and Sahoo, 2021), and wheat and *Nicotiana benthamiana* (Li et al., 2018; Wang et al., 2021). Collectively, the base editing tools developed thus far have shown considerable potential for enhancing the efficiency of single-base editing, with applications for the modification of a broad range of agronomic and disease resistance traits in crops.

PRIME EDITING FOR PLANT DISEASE RESISTANCE

A pioneering genome editing approach that overcomes the issue of transversion editing was developed prior to the introduction of transversion base editors (**Figure 5**). The "prime editor" technique, which can be used to introduce 12 different base changes in human cells, comprises nCas9 (H840A) coupled to the reverse transcriptase of the Moloney murine leukemia virus (M-MLV RT), as well as a prime editing guide RNA (pegRNA) with a reverse transcriptase template and a primer-binding site at the 3′ end of the sgRNA. The reverse transcriptase template contains the genetic material for the desired mutation, whereas the primer-binding site joins the nCas9 (H840A)-nicked ssDNA strand (Azameti and Dauda, 2021). Following the priming of reverse transcription, the genetic information from the reverse transcriptase template is incorporated into the genome, and in response to cutting of the ssDNA approximately 3 base pairs upstream of the PAM sequence on the non-target strand, a 3-base pair extension of the pegRNA, which contains both a primer-binding site and a reverse transcription sequence (RT sequence), an appropriate polymorphism can be introduced at the target location (Anzalone et al., 2019). Although it induces base replacements and introduces a few indels at a relatively wide range of locations (+1 to +33), the primary editor is not limited by its PAM sequence (Azameti and Dauda, 2021).

To date, maize, wheat, and rice are among the crops for which prime editing systems have been developed and deployed (Hua et al., 2020; Lin et al., 2020; Xu et al., 2020a), and recently, prime editing technology has been used to generate point mutations, insertions, and deletions in the protoplasts of rice and wheat plants, achieving a regeneration of prime-edited rice plants of up to 21.8% (Lin et al., 2020). In addition, an HPT-ATG reporter has been developed in rice to generate and assess the activity of a plant prime editor 2 (pPE2) system (Xu et al., 2020a), with the authors obtaining up to 31.3% edited transgenic T_0 plants. Using the prime editor to target *OsALS-1* and *OsALS-2*, Xu et al. (2020c) found that the predicted G-to-T and specific C-to-T substitution editing efficiencies were 1.1% (1/87) and 1.1% (1/88), respectively. In a further example of the application of this technology, Hua et al. (2020) developed the prime editor Sp-PE3 and investigated its efficacy in rice calli, achieving an editing efficiency of up to 17.1% at the targeted locations, among which, the rice endogenous acetolactate synthase (ALS) gene was edited to generate the desired G-A base transition in four of the 44 (9.1%) assessed transgenic lines, with no insertions or deletions. Interestingly, using the same editor system, no mutations were detected at the ABERRANT PANICLE ORGANIZATION 1 (APO1) locus, thereby implying that Sp-PE3 can be used to facilitate precise base substitution with varying degrees of efficacy, depending on the targeted location. Prime editors have also been used to generate mutant maize lines with double (*W542L/S621I*) mutations (Jiang et al., 2020), with 43.75% (7 of 16) of the lines transformed with *pZ1WS* also containing an edited *S621I*, and a single line found to contain homozygous mutations in both *ZmALS1* and *ZmALS2*. Attempts have also been made to engineer herbicide resistance in rice by targeting ACETOLACTATE SYNTHASE (OsALS), with a verified editing efficiency of 0.26–2% at the targeted location (Butt et al., 2020; Hussain et al., 2021), and reported development of resistant rice as a consequence of base substitutions. Furthermore, Veillet et al. (2020) used CRISPR-mediated plant prime editing to successfully target potato *StALS* genes, with an editing efficacy

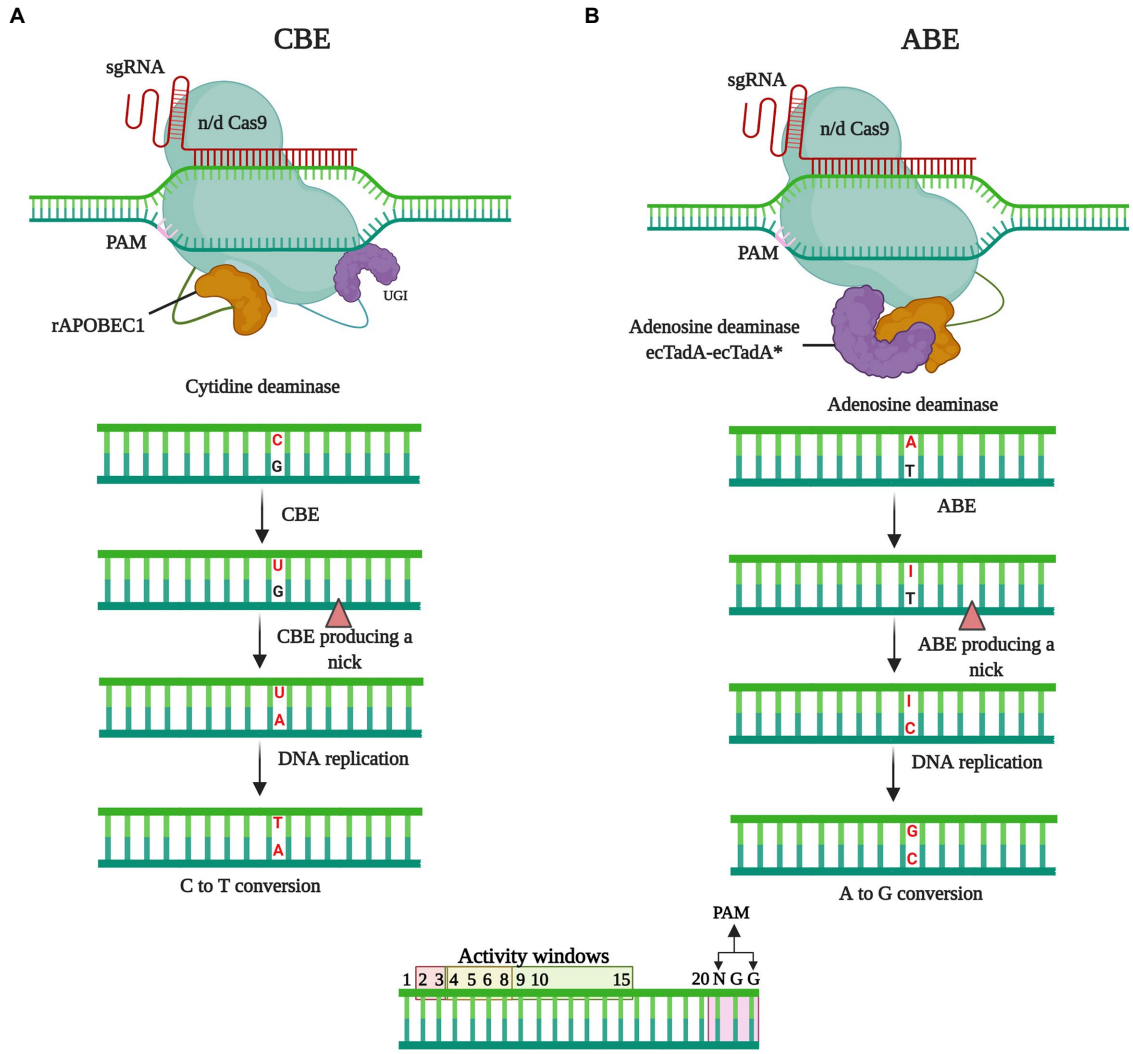


FIGURE 4 | The mechanisms of base editing. In the absence of double-strand breaks (DSBs), base editing facilitates the introduction of precise point mutations at specified target sites in the genome *via* nucleotide substitution. **(A)** Cytidine deaminase base editing (CBE) is performed in conjunction with the use of an APOBEC1 cytidine deaminase, which converts C to U. Subsequently the U-G mismatch is resolved *via* cellular mismatch repair or base editing mismatch repair machinery that leads to the formation of T-A at the target locus. **(B)** The adenine base editing (ABE) leads A-T to G-C substitution. After recruiting to the targeted genomic locus, the ABE delaminates targeted A base to I (inosine) leading to I-T base pairing. The cellular mismatch repair mechanism or DNA replication resolves the I-T, forming G-C base pairing.

of 92%. Nevertheless, despite the use of orthogonal approaches, such as reverse transcriptase orthologs with differing catalytic activity, the efficacy of prime editor-based editing in plants remains limited (Lin et al., 2020; Xu et al., 2020a).

RNAi APPROACHES FOR CONFERRING DISEASE RESISTANCE IN PLANTS

Plants have evolved inherent RNA interference (RNAi) mechanism, which they deploy to minimize the presence of unwanted elements and develop virus resistance (Kamthan et al., 2015; Rosa et al., 2018). RNAi functions by suppressing gene expression or neutralizing specific mRNA molecules, and the dominant nature of functional RNAi hairpins may be seen into plants. As an example of RNAi-based genome editing in plants, a gene encoding P450, an enzyme responsible for herbicide resistance, has been targeted using an RNAi expression element inserted into a CRISPR-Cas9 construct, with transgene-free genome-edited resistant plants being obtained, as determined by herbicide-based isolation phenotyping rather than PCR to identify and separate the transgene-free plants (Majumdar et al., 2017). Numerous recent studies have similarly used RNAi approaches to confer disease resistance in plants (e.g., Kuo and Falk, 2020; Liu et al., 2021).

Similar resistance may also be conferred using non-coding RNA, with miRNA expressed *via* RNA interference with polymerase II (RNAi) typically being the underlying mechanism

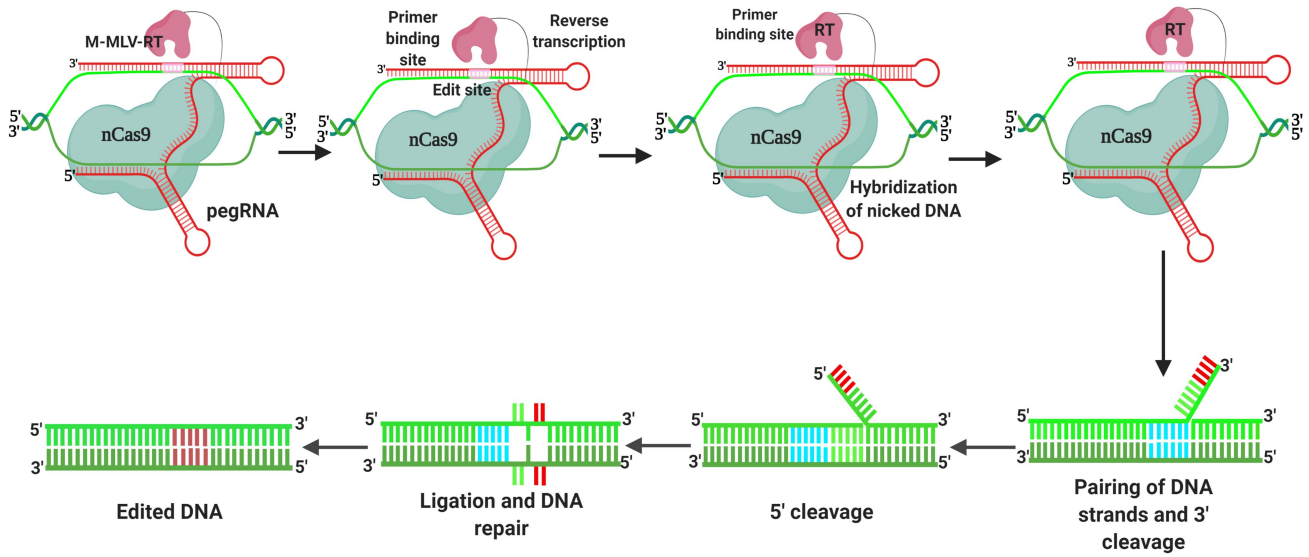


FIGURE 5 | Prime editing is a recently designed base editing technique that facilitates the accurate substitution, insertion, and deletion of sequences. The fusion of Nickase Cas9 (nCas9) and reverse transcriptase (RT) is the most important element in prime editing. The intended modifications are encoded by the prime editing gRNA, which directs the nCas9-RT complex to the target gene sequence. The DNA is cut by the prime editor, which is then hybridized to the primer-binding site, resulting in reverse transcription. DNA ligation and repair are followed by base pairing of 30 or 50 flaps, resulting in DNA editing.

in eukaryotic cells. Small RNAs are mobile elements that can be interchanged among plants and pathogens (Hua et al., 2018). It has been shown that pathogen-derived sRNAs target host mRNAs encoding genes associated with RNAi degradation and defense (Weiberg et al., 2013), whereas in response to pathogen attack, plants initiate the differential expression of endogenous sRNAs (Shen et al., 2014) and may even target *in vivo* factor to pathogens (Zhang et al., 2016). Thus, given that RNAi is a natural phenomenon, in order to target pathogens, it is only necessary to modify the underlying mechanisms by converting small artificial miRNAs homologous to viral genomes to small molecular miRNAs. For example, a common target is the viral coat protein or viral replication (Dong and Ronald, 2019). Moreover, stable constitutive expression of CRISPR-Cas systems can be employed to target viral genomes with adequate sgRNA, thereby generating a new immune system (Zhang et al., 2018a).

Both small sgRNA and virus-specific proteins can be modified using CRISPR-Cas technology. The mechanisms are sufficient to degrade pathogen genomes or intermediate replication in response to the viral infection of host cells. Notably, this miRNA or sgRNA should be carefully engineered such that it does not inadvertently attack the host or human nucleic acids. Given that the principal mechanisms of resistance (degradation of nucleic acids derived from important pathogens) are broadly comparable, the application of either system could be used to generate transgenic crops *via* host gene-induced silencing (HIGS) and CRISPR-Cas methods. In this regard, although counter-defensive measures (inhibitor proteins) against the host RNAi mechanisms have been identified in at least some plant viruses (Voinnet, 2005), which could nullify HIGS, these would not detrimentally influence the bacterial CRISPR-Cas system in the event of co-infection events. However, as demonstrated by the *Geminivirus* infection of cassava, it should be taken into consideration that CRISPR-Cas9 can also influence virus evolution (Mehta et al., 2019; Rybicki, 2019).

It is assumed that it would be feasible to engineer specific mutations in host defense factor coding sequences based on HDR. However, given that a sufficient number of repair templates is necessary to facilitate repair, the efficiency is naturally lower than that achieved based on non-homologous end joining (NHEJ). Given the degenerative nature of the genetic code, base triplets may be substituted without altering start sequence amino acids in the open reading frame. For example, by targeting plant defense genes, pathogen-derived siRNAs may succeed in hijacking the host RNAi machinery, thereby preventing the degradation of mRNAs encoding (trans-kingdom RNAi). The potential objectives of this approach have been established with respect to the interaction between tomato and *Botrytis cinerea*, which affects the defense signal transduction sequence MAPKKK4 (Weiberg et al., 2013), and in the interaction between *Arabidopsis* and *B. cinerea*, targeting *AtWRKY7*, *AtPMR6*, and *AtFEI2* (Wang et al., 2017). The advantage of this approach is that there is no implementation cost. Specific amino acids required for pathogen effector recognition and/or cleavage may also be altered in targeted plant proteins. For example, HopB1, which cleaves BAK1 (Li et al., 2016), can be used to cleave signaling components PBS1, PBL1, PBL2, PBL6, and BIK1 co-receptors or AvrPphB among *Pseudomonas syringae* effector proteases (Zhang et al., 2010), as even a single amino acid change in the Aβ cleavage motif prevents AvrPphB cleavage. This RNA surveillance system also functions as a powerful antiviral defense mechanism in plants.

POTENTIAL UTILITY OF CRISPR-Cas-MEDIATED GENOME EDITING IN SUSTAINABLE AGRICULTURE

Over the past 100 years, agricultural productivity has grown substantially as a consequence of continual technical advances (Voytas and Gao, 2014). In recent times, the advent of molecular genetic tools has ushered in a new era of genomic breeding (**Figure 6**), based on genetic engineering and molecular breeding (Wallace et al., 2018). Both transgenic and non-transgenic methodologies have, for a number of years, contributed to groundbreaking developments in agriculture and horticulture. However, despite the fact that transgenic crops continue to be the focus of crop improvement (Green, 2018; Nakka et al., 2019), public acceptance of these modified plants is typically limited (Voytas and Gao, 2014). To circumvent the wariness associated with the introduction of foreign genetic material, numerous genome editing approaches are currently being assessed for their utility in the development non-transgenic crops, some of which have recently been marketed (Metje-Sprink et al., 2020). In this context, the majority of the studies conducted to date have tended to focus on "proof of concept" or improving the precision and delivery of the site directed nucleases SDN (Metje-Sprink et al., 2020). Although genome editing has been used to generate a number of crops, there have yet been very few agricultural trials of these crops. Among those genome-edited crops that have reached the field trial stage are herbicide-resistant canola and flax plants. For example, in the United States, CIBUS conducted one of the first field trials for herbicide-resistant canola (Hussain et al., 2021), which, given that it is not genetically modified. Bayer Crop Science has developed a genome-edited flax with glyphosate tolerance, which, in 2019, was successfully cultivated over an extensive area of approximately 50 million acres (Hussain et al., 2021). The cultivation of such genome-edited crops, particularly those characterized by multiple resistance, can represent a key strategy for simultaneously combating numerous stresses and enhancing crop yields, while also preserving soil moisture and texture (Zhang et al., 2020c), and it is hoped that trials of a larger number of these crops are inaugurated in the coming years.

Advances in genome editing, particularly the introduction of CRISPR-Cas systems, have opened up new avenues for crop development (Han and Kim, 2019; Zhang et al., 2019c). In this regard. The editing of plant genomes has a number of notable benefits. For example, wheat plants engineered for the triple knockout of *TaMLO* not only showed resistance to powdery mildew but also showed resistance to chlorophyll (Wang et al., 2014), whereas the triple mutants obtained based on non-conservative EMS-induced *TaMLO* target mutations showed evident pleiotropic effects (Acevedo-Garcia

FIGURE 6 | An overview of genome engineering technology and the potential use of the CRISPR-Cas system in different agriculture science disciplines.

et al., 2017). Recently, a CRISPR system has been used to alter the coding and promoter regions of the citrus canker susceptibility gene *LOB1*, thereby conferring citrus canker resistance (Jia et al., 2017; Peng et al., 2017), and the fruits of these modified plants have been marketed as a sustainably cultivated commercial product for human consumption (Waltz, 2016). Furthermore, CRISPR-Cas9 has been used to facilitate the knockout of *Ms1* in wheat lines for hybrid seed production (Okada et al., 2019). Moreover, genome editing enables the direct insertion of exogenous genes into plants to impart biotic or abiotic stress resistance (Fartyal et al., 2018). For example, phosphinothricin acetyltransferase (PAT) and bialaphos resistance (BAR) genes, initially obtained from species of *Streptomyces*, have been introduced into plants to confer glufosinate herbicide resistance (Schütte et al., 2017), whereas Han and Kim have used CRISPR-Cas9 to induce loss-of-function mutations of 5-oxoprolinase (OXP) and phosphoribosyl anthranilate isomerase (PAI) in plants, resulting in resistance to sulfamethoxazole and 6-methylanthranilate, respectively (Han and Kim, 2019).

Genome editing, particularly programmed base editing and prime editing, is of crucial importance as it can be used to introduce heritable targeted changes that give rise to transgene-free crops (Tian et al., 2018; Butt et al., 2020; Liu et al., 2020) that are genetically indistinguishable from plants developed based on classical mutagenesis approaches (Huang et al., 2016). Furthermore, base editing facilitates the substitution of multiple amino acids in specific genes, which can contribute to enhance the resistance spectrum of crops, such as resistance to multiple diseases (Zhang et al., 2019b). Moreover, in contrast to traditionally genetically modified plants, non-transgenic plants generated using CRISPR-Cas systems are exempt from regulatory approval in different countries (Lassoued et al., 2019). Similarly, given that genome-edited crops produced without the introduction of foreign DNA do not require risk evaluation (Ishii and Araki, 2017), developers can bring new crops to the market years sooner and at considerably less cost than is the case with genetically modified crops (Waltz, 2016). Consequently, by contributing to reductions in crop development time and expense, the use of genome editing has the potential to accelerate the commercialization of crops compared with mutagenesis and conventional genetic modification.

CONCLUSION

At present, the viability of the entire agriculture sector is under threat on three fronts. The evolution of new pathogens, along with the development of resistance in pre-existing pathogens, are a source of significant direct losses in crop production, whereas the depletion of environmental resources (reductions in the area of cultivable arable land and water resources) is contributing to indirect reductions in crop yield. In addition, the ever-expanding human population continues to drive the demand for sufficient food supplies to meet nutritional needs. Phytopathogens are arguably the most important causal agents of plant diseases. To date, a range of techniques have been employed to manage crop diseases, including traditional and transgenic breeding. However, although these methods have primarily remained effective, they are typically laborious and time-consuming. Recent developments in genome editing technology, particularly CRISPR-Cas9-based systems, do, nevertheless, offer considerable potential for the improvement of crop plants *via* precise trait targeting. Such genome editing can be used to introduce specific targeted modifications in terms of both gain- and loss-of-function. Moreover, these technologies enable the rapid classification of new immune receptor genes (such as guided molecular evolution and Ren-Seq) and have contributed to a significant expansion in the pool of deployable genes for enhancing resistance to a range of microorganisms. In addition, it is predicted that the development of molecular stacking and targeted gene insertion will play an increasingly important role in the generation of broad-spectrum resistance to both viral and non-viral pathogens. At the current stage of technological development, CRISPR-Cas9 systems have emerged as the most efficient and suitable alternative genome editing-based solutions for the development of disease-resistant crops, which will predictably contribute to higher crop productivity with simple and effective disease management. In addition, the emergence of new base editing systems has facilitated the development of transgene-free non-genetically modified plants, which are likely to be indistinguishable from the same plants altered using transgenic or conventional crop breeding methods. To date, several disease-resistant crops have been produced using gene editing, which will undoubtedly gain greater public acceptance than that gained by conventionally genetically modified plants. Accordingly, we firmly believe that, used responsibly, genome editing in the agricultural industry stands to make significant contributions to the enhancement of crop productivity that can benefit both producers and consumers, and it goes a long way in meeting current and future increases in human nutritional requirements.

AUTHOR CONTRIBUTIONS

QA, HM, and LZ planned and designed this review manuscript. QA wrote this review paper. CY, MA, MI, and MS help to draw the figures. AH, SA, XW, and MR helped to improve the manuscript writing. QA, LZ, DA, and HM contributed to the critically revising of the manuscript. All the authors have reviewed, edited, and approved the manuscript before submission.

FUNDING

This work was supported by the China Postdoctoral Science Foundation (No: 2014M561669). The high-talent introduction and continuous training fund supported by Zhejiang Academy of Agricultural Sciences (No: 10300000021LL05).

REFERENCES

Abudayyeh, O. O., Gootenberg, J. S., Essletzbichler, P., Han, S., Joung, J., Belanto, J. J., et al. (2017). RNA targeting with CRISPR-Cas13. *Nature* 550, 280–284. doi: 10.1038/nature24049

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

Ahmad, S., Wei, X., Sheng, Z., Hu, P., and Tang, S. (2020). CRISPR/Cas9 for development of disease resistance in plants: recent progress, limitations and future prospects. *Brief. Funct. Genomics* 19, 26–39. doi: 10.1093/bfgp/elz041

Ahuja, I., Kissen, R., and Bones, A. M. (2012). Phytoalexins in defense against pathogens. *Trends Plant Sci.* 17, 73–90. doi: 10.1016/j.tplants.2011.11.002

Ainley, W. M., Sastry-Dent, L., Welter, M. E., Murray, M. G., Zeitler, B., Amora, R., et al. (2013). Trait stacking *via* targeted genome editing. *Plant Biotechnol. J.* 11, 1126–1134. doi: 10.1111/pbi.12107

Ali, Q., Ahmar, S., Sohail, M. A., Kamran, M., Ali, M., Saleem, M. H., et al. (2021). Research advances and applications of biosensing technology for the diagnosis of pathogens in sustainable agriculture. *Environ. Sci. Pollut. Res.* 28, 9002–9019. doi: 10.1007/s11356-021-12419-6

Ali, Q., Zheng, H., Rao, M. J., Ali, M., Hussain, A., Saleem, M. H., et al. (2022). Advances, limitations, and prospects of biosensing technology for detecting phytopathogenic bacteria. *Chemosphere* 296:133773. doi: 10.1016/j.chemosphere.2022.133773

Anzalone, A. V., Randolph, P. B., Davis, J. R., Sousa, A. A., Koblan, L. W., Levy, J. M., et al. (2019). Search-and-replace genome editing without double-strand breaks or donor DNA. *Nature* 576, 149–157. doi: 10.1038/s41586-019-1711-4

Ayaz, A., Huang, H., Zheng, M., Zaman, W., Li, D., Saqib, S., et al. (2021a). Molecular cloning and functional analysis of GmLACS2-3 reveals its involvement in cutin and suberin biosynthesis along with abiotic stress tolerance. *Int. J. Mol. Sci.* 22:9175. doi: 10.3390/ijms22179175

Ayaz, A., Saqib, S., Huang, H., Zaman, W., Lü, S., and Zhao, H. (2021b). Genome-wide comparative analysis of long-chain acyl-CoA synthetases (LACSs) gene family: a focus on identification, evolution and expression profiling related to lipid synthesis. *Plant Physiol. Biochem.* 161, 1–11. doi: 10.1016/j.plaphy.2021.01.042

Azameti, M. K., and Dauda, W. P. (2021). Base editing in plants: applications, challenges, and future prospects. *Front. Plant Sci.* 12:664997. doi: 10.3389/fpls.2021.664997

Baek, K., Kim, D. H., Jeong, J., Sim, S. J., Melis, A., Kim, J.-S., et al. (2016). DNA-free two-gene knockout in *Chlamydomonas reinhardtii via* CRISPR-Cas9 ribonucleoproteins. *Sci. Rep.* 6:30620. doi: 10.1038/srep30620

Barrangou, R., and Doudna, J. A. (2016). Applications of CRISPR technologies in research and beyond. *Nat. Biotechnol.* 34, 933–941. doi: 10.1038/nbt.3659

Bastet, A., Lederer, B., Giovinazzo, N., Arnoux, X., German-Retana, S., Reinbold, C., et al. (2018). Trans-species synthetic gene design allows resistance pyramiding and broad-spectrum engineering of virus resistance in plants. *Plant Biotechnol. J.* 16, 1569–1581. doi: 10.1111/pbi.12896

Bastet, A., Zafirov, D., Giovinazzo, N., Guyon-Debast, A., Nogué, F., Robaglia, C., et al. (2019). Mimicking natural polymorphism in eIF4E by CRISPR-Cas9 base editing is associated with resistance to potyviruses. *Plant Biotechnol. J.* 17, 1736–1750. doi: 10.1111/pbi.13096

Bharat, S. S., Li, S., Li, J., Yan, L., and Xia, L. (2019). Base editing in plants: current status and challenges. *Crop J.* 8, 384–395. doi: 10.1016/j.cj.2019.10.002

Biffen, R. H. (1905). Mendel's laws of inheritance and wheat breeding. *J. Agric. Sci.* 1, 4–48. doi: 10.1017/S0021859600000137

Brauer, E. K., Balcerzak, M., Rocheleau, H., Leung, W., Schernthaner, J., Subramaniam, R., et al. (2020). Genome editing of a deoxynivalenol-induced transcription factor confers resistance to *Fusarium graminearum* in wheat. *Mol. Plant-Microbe Interact.* 33, 553–560. doi: 10.1094/MPMI-11-19-0332-R

Brunner, S., Stirnweis, D., Diaz Quijano, C., Buesing, G., Herren, G., Parlange, F., et al. (2012). Transgenic Pm3 multilines of wheat show increased powdery mildew resistance in the field. *Plant Biotechnol. J.* 10, 398–409. doi: 10.1111/j.1467-7652.2011.00670.x

Butler, N. M., Baltes, N. J., Voytas, D. F., and Douches, D. S. (2016). Geminivirus-mediated genome editing in potato (*Solanum tuberosum* L.) using sequence-specific nucleases. *Front. Plant Sci.* 7:1045. doi: 10.3389/fpls.2016.01045

Butt, H., Rao, G. S., Sedeek, K., Aman, R., Kamel, R., and Mahfouz, M. (2020). Engineering herbicide resistance via prime editing in rice. *Plant Biotechnol. J.* 18, 2370–2372. doi: 10.1111/pbi.13399

Cai, Y., Chen, L., Zhang, Y., Yuan, S., Su, Q., Sun, S., et al. (2020). Target base editing in soybean using a modified CRISPR/Cas9 system. *Plant Biotechnol. J.* 18:1996, –1998. doi: 10.1111/pbi.13386

Chandrasegaran, S., and Carroll, D. (2016). Origins of programmable nucleases for genome engineering. *J. Mol. Biol.* 428, 963–989. doi: 10.1016/j.jmb.2015.10.014

Chandrasekaran, J., Brumin, M., Wolf, D., Leibman, D., Klap, C., Pearlsman, M., et al. (2016). Development of broad virus resistance in non-transgenic cucumber using CRISPR/Cas9 technology. *Mol. Plant Pathol.* 17, 1140–1153. doi: 10.1111/mpp.12375

Charrier, A., Vergne, E., Dousset, N., Richer, A., Petiteau, A., and Chevreau, E. (2019). Efficient targeted mutagenesis in apple and first time edition of pear using the CRISPR-Cas9 system. *Front. Plant Sci.* 10:40. doi: 10.3389/fpls.2019.00040

Chen, Y., Wang, Z., Ni, H., Xu, Y., Chen, Q., and Jiang, L. (2017). CRISPR/Cas9-mediated base-editing system efficiently generates gain-of-function mutations in Arabidopsis. *Sci. China Life Sci.* 60, 520–523. doi: 10.1007/s11427-017-9021-5

Cheng, H., Hao, M., Ding, B., Mei, D., Wang, W., Wang, H., et al. (2021). Base editing with high efficiency in allotetraploid oilseed rape by A3A-PBE system. *Plant Biotechnol. J.* 19:87, –97. doi: 10.1111/pbi.13444

Chisholm, S. T., Coaker, G., Day, B., and Staskawicz, B. J. (2006). Review host-microbe interactions: shaping the evolution of the plant immune response. *Cell* 124, 803–814. doi: 10.1016/j.cell.2006.02.008

Christou, P. (2013). Plant genetic engineering and agricultural biotechnology 1983–2013. *Trends Biotechnol.* 31, 125–127. doi: 10.1016/j.tibtech.2013.01.006

Chung, E.-H., El-Kasmi, F., He, Y., Loehr, A., and Dangl, J. L. (2014). A plant phosphoswitch platform repeatedly targeted by type III effector proteins regulates the output of both tiers of plant immune receptors. *Cell Host Microbe* 16, 484–494. doi: 10.1016/j.chom.2014.09.004

Cook, D. E., Mesarich, C. H., and Thomma, B. P. H. J. (2015). Understanding plant immunity as a surveillance system to detect invasion. *Annu. Rev. Phyto. pathol.* 53, 541–563.

CTNBio (2018). Normative Resolution No. 16 of January 15, 2018.

Dangl, J. L., Horvath, D. M., and Staskawicz, B. J. (2013). Pivoting the plant immune system from dissection to deployment. *Science* 341, 746–751. doi: 10.1126/science.1236011

Das, G., and Rao, G. J. N. (2015). Molecular marker assisted gene stacking for biotic and abiotic stress resistance genes in an elite rice cultivar. *Front. Plant Sci.* 6:698. doi: 10.3389/fpls.2015.00698

Dong, O. X., and Ronald, P. C. (2019). Genetic engineering for disease resistance in plants: recent progress and future perspectives. *Plant Physiol.* 180, 26–38. doi: 10.1104/pp.18.01224

Fartyal, D., Agarwal, A., James, D., Borphukan, B., Ram, B., Sheri, V., et al. (2018). Co-expression of P173S mutant rice EPSPS and igrA genes results in higher glyphosate tolerance in transgenic rice. *Front. Plant Sci.* 9:144. doi: 10.3389/fpls.2018.00144

Flor, H. H. (1956). The complementary genic systems in flax and flax rust. *Adv. Genet.* 8, 29–54.

Flor, H. H. (1971). Current status of the gene-fob-gene concept 3531. *Annu. Rev. Phytopathol.* 9, 275–296. doi: 10.1146/annurev.py.09.090171.001423

Franceschetti, M., Maqbool, A., Jiménez-Dalmaroni, M. J., Pennington, H. G., Kamoun, S., and Banfield, M. J. (2017). Effectors of filamentous plant pathogens: commonalities amid diversity. *Microbiol. Mol. Biol. Rev.* 81:e00066-16. doi: 10.1128/MMBR.00066-16

Fuchs, M. (2017). Pyramiding resistance-conferring gene sequences in crops. *Curr. Opin. Virol.* 26, 36–42. doi: 10.1016/j.coviro.2017.07.004

Gao, Z., Liu, Q., Zhang, Y., Chen, D., Zhan, X., Deng, C., et al. (2020). OsCUL3a-associated molecular switches have functions in cell metabolism, cell death, and disease resistance. *J. Agric. Food Chem.* 68, 5471–5482. doi: 10.1021/acs.jafc.9b07426

Gaudelli, N. M., Komor, A. C., Rees, H. A., Packer, M. S., Badran, A. H., Bryson, D. I., et al. (2017). Programmable base editing of T to G C in

genomic DNA without DNA cleavage. *Nature* 551, 464–471. doi: 10.1038/nature24644

Gepts, P. (2002). A comparison between crop domestication, classical plant breeding, and genetic engineering. *Crop Sci.* 42, 1780–1790. doi: 10.2135/cropsci2002.1780

Ghislain, M., Byarugaba, A. A., Magembe, E., Njoroge, A., Rivera, C., Román, M. L., et al. (2019). Stacking three late blight resistance genes from wild species directly into African highland potato varieties confers complete field resistance to local blight races. *Plant Biotechnol. J.* 17, 1119–1129. doi: 10.1111/pbi.13042

Gomez, M. A., Lin, Z. D., Moll, T., Chauhan, R. D., Hayden, L., Renninger, K., et al. (2019). Simultaneous CRISPR/Cas9-mediated editing of cassava eIF 4E isoforms nCBP-1 and nCBP-2 reduces cassava brown streak disease symptom severity and incidence. *Plant Biotechnol. J.* 17, 421–434. doi: 10.1111/pbi.12987

Green, J. M. (2018). The rise and future of glyphosate and glyphosate-resistant crops. *Pest Manag. Sci.* 74, 1035–1039. doi: 10.1002/ps.4462

Han, Y.-J., and Kim, J.-I. (2019). Application of CRISPR/Cas9-mediated gene editing for the development of herbicide-resistant plants. *Plant Biotechnol. Rep.* 13, 447–457. doi: 10.1007/s11816-019-00575-8

Haverkort, A. J., Boonekamp, P. M., Hutten, R., Jacobsen, E., Lotz, L. A. P., Kessel, G. J. T., et al. (2016). Durable late blight resistance in potato through dynamic varieties obtained by Cisgenesis: scientific and societal advances in the DuRPh project. *Potato Res.* 59, 35–66. doi: 10.1007/s11540-015-9312-6

Holton, N., Nekrasov, V., Ronald, P. C., and Zipfel, C. (2015). The phylogenetically-related pattern recognition receptors EFR and XA21 recruit similar immune signaling components in monocots and dicots. *PLoS Pathog.* 11:e1004602. doi: 10.1371/journal.ppat.1004602

Horvath, D. M., Pauly, M. H., Hutton, S. F., Vallad, G. E., Scott, J. W., Jones, J. B., et al. (2015). The pepper Bs2 gene confers effective field resistance to bacterial leaf spot and yield enhancement in Florida tomatoes. *Acta Hortic.* 1069, 47–51. doi: 10.17660/ActaHortic.2015.1069.5

Hua, K., Jiang, Y., Tao, X., and Zhu, J. (2020). Precision genome engineering in rice using prime editing system. *Plant Biotechnol. J.* 18:2167, –2169. doi: 10.1111/pbi.13395

Hua, K., Tao, X., and Zhu, J. (2019). Expanding the base editing scope in rice by using Cas9 variants. *Plant Biotechnol. J.* 17, 499–504. doi: 10.1111/pbi.12993

Hua, C., Zhao, J.-H., and Guo, H.-S. (2018). Trans-kingdom RNA silencing in plant–fungal pathogen interactions. *Mol. Plant* 11, 235–244. doi: 10.1016/j.molp.2017.12.001

Huang, S., Weigel, D., Beachy, R. N., and Li, J. (2016). A proposed regulatory framework for genome-edited crops. *Nat. Genet.* 48, 109–111. doi: 10.1038/ng.3484

Hunziker, J., Nishida, K., Kondo, A., Kishimoto, S., Ariizumi, T., and Ezura, H. (2020). Multiple gene substitution by target-AID base-editing technology in tomato. *Sci. Rep.* 10:20471. doi: 10.1038/s41598-020-77379-2

Hussain, A., Amna, Kamran, M. A., Javed, M. T., Hayat, K., Farooq, M. A., et al. (2019). Individual and combinatorial application of Kocuria rhizophila and citric acid on phytoextraction of multi-metal contaminated soils by *Glycine* max L. *Environ. Exp. Bot.* 159, 23–33. doi: 10.1016/j.envexpbot.2018.12.006

Hussain, A., Ding, X., Alariqi, M., Manghwar, H., Hui, F., Li, Y., et al. (2021). Herbicide resistance: another hot agronomic trait for plant genome editing. *Plan. Theory* 10:621. doi: 10.3390/plants10040621

Ishii, T., and Araki, M. (2017). A future scenario of the global regulatory landscape regarding genome-edited crops. *GM Crops Food* 8, 44–56. doi: 10.1080/21645698.2016.1261787

Jia, H., Zhang, Y., Orbović, V., Xu, J., White, F. F., Jones, J. B., et al. (2017). Genome editing of the disease susceptibility gene CsLOB1 in citrus confers resistance to citrus canker. *Plant Biotechnol. J.* 15, 817–823. doi: 10.1111/pbi.12677

Jiang, Y.-Y., Chai, Y.-P., Lu, M.-H., Han, X.-L., Lin, Q., Zhang, Y., et al. (2020). Prime editing efficiently generates W542L and S621I double mutations in two ALS genes in maize. *Genome Biol.* 21:257. doi: 10.1186/s13059-020-02170-5

Jiang, F., Taylor, D. W., Chen, J. S., Kornfeld, J. E., Zhou, K., Thompson, A. J., et al. (2016). Structures of a CRISPR-Cas9 R-loop complex primed for DNA cleavage. *Science* 351, 867–871. doi: 10.1126/science.aad8282

Jo, K.-R., Kim, C.-J., Kim, S.-J., Kim, T.-Y., Bergervoet, M., Jongsma, M. A., et al. (2014). Development of late blight resistant potatoes by cisgene stacking. *BMC Biotechnol.* 14:50. doi: 10.1186/1472-6750-14-50

Jones, J. D. G., and Dangl, J. L. (2006). The plant immune system. *Nature* 444, 323–329. doi: 10.1038/nature05286

Jones, J. D. G., Perkins, S., Foster, S., Tomlinson, L., Verweij, W., Jupe, F., et al. (2014). Elevating crop disease resistance with cloned genes. *Philos. Trans. R. Soc. B Biol. Sci.* 369, 20130087–20130087. doi: 10.1098/rstb.2013.0087

Kamthan, A., Chaudhuri, A., Kamthan, M., and Datta, A. (2015). Small RNAs in plants: recent development and application for crop improvement. *Front. Plant Sci.* 6:208. doi: 10.3389/fpls.2015.00208

Kang, B.-C., Yun, J.-Y., Kim, S.-T., Shin, Y., Ryu, J., Choi, M., et al. (2018). Precision genome engineering through adenine base editing in plants. *Nat. Plants* 4, 427–431. doi: 10.1038/s41477-018-0178-x

Karasov, T. L., Chae, E., Herman, J. J., and Bergelson, J. (2017). Mechanisms to mitigate the trade-off between growth and defense. *Plant Cell* 29, 666–680. doi: 10.1105/tpc.16.00931

Karlovsky, P. (2011). Biological detoxification of the mycotoxin deoxynivalenol and its use in genetically engineered crops and feed additives. *Appl. Microbiol. Biotechnol.* 91, 491–504. doi: 10.1007/s00253-011-3401-5

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

Komor, A. C., Kim, Y. B., Packer, M. S., Zuris, J. A., and Liu, D. R. (2016). Programmable editing of a target base in genomic DNA without double-stranded DNA cleavage. *Nature* 533, 420–424. doi: 10.1038/nature17946

Kourelis, J., and Van Der Hoorn, R. A. L. (2018). Defended to the nines: 25 years of resistance gene cloning identifies nine mechanisms for R protein function. *Plant Cell* 30, 285–299. doi: 10.1105/tpc.17.00579

Kubicek, C. P., Starr, T. L., and Glass, N. L. (2014). Plant cell wall–degrading enzymes and their secretion in plant-pathogenic fungi. *Annu. Rev. Phytopathol.* 52, 427–451. doi: 10.1146/annurev-phyto-102313-045831

Kumar, N., Galli, M., Ordon, J., Stuttmann, J., Kogel, K., and Imani, J. (2018). Further analysis of barley MORC 1 using a highly efficient RNA-guided Cas9 gene-editing system. *Plant Biotechnol. J.* 16, 1892–1903. doi: 10.1111/pbi.12924

Kunwar, S., Iriarte, F., Fan, Q., Evaristo da Silva, E., Ritchie, L., Nguyen, N. S., et al. (2018). Transgenic expression of EFR and Bs2 genes for field management of bacterial wilt and bacterial spot of tomato. *Phytopathology* 108, 1402–1411. doi: 10.1094/PHYTO-12-17-0424-R

Kuo, Y. W., and Falk, B. W. (2020). RNA interference approaches for plant disease control. *BioTechniques* 69, 469–477. doi: 10.2144/btn-2020-0098

Lacombe, S., Rougon-Cardoso, A., Sherwood, E., Peeters, N., Dahlbeck, D., Van Esse, H. P., et al. (2010). Interfamily transfer of a plant pattern-recognition receptor confers broad-spectrum bacterial resistance. *Nat. Biotechnol.* 28, 365–369. doi: 10.1038/nbt.1613

Lassoued, R., Macall, D. M., Hesseln, H., Phillips, P. W. B., and Smyth, S. J. (2019). Benefits of genome-edited crops: expert opinion. *Transgenic Res.* 28, 247–256. doi: 10.1007/s11248-019-00118-5

Li, L., Kim, P., Yu, L., Cai, G., Chen, S., Alfano, J. R., et al. (2016). Activation-dependent destruction of a co-receptor by a *Pseudomonas syringae* effector dampens plant immunity. *Cell Host Microbe* 20, 504–514. doi: 10.1016/j.chom.2016.09.007

Li, T., Liu, B., Spalding, M. H., Weeks, D. P., and Yang, B. (2012). High-efficiency TALEN-based gene editing produces disease-resistant rice. *Nat. Biotechnol.* 30, 390–392. doi: 10.1038/nbt.2199

Li, S., Shen, L., Hu, P., Liu, Q., Zhu, X., Qian, Q., et al. (2019a). Developing disease-resistant thermosensitive male sterile rice by multiplex gene editing. *J. Integr. Plant Biol.* 61, 1201–1205. doi: 10.1111/jipb.12774

Li, Z., Xiong, X., Wang, F., Liang, J., and Li, J. (2019b). Gene disruption through base editing-induced messenger RNA missplicing in plants. *New Phytol.* 222, 1139–1148. doi: 10.1111/nph.15647

Li, C., Zong, Y., Wang, Y., Jin, S., Zhang, D., Song, Q., et al. (2018). Expanded base editing in rice and wheat using a Cas9-adenosine deaminase fusion. *Genome Biol.* 19:59. doi: 10.1186/s13059-018-1443-z

Liang, Z., Chen, K., Li, T., Zhang, Y., Wang, Y., Zhao, Q., et al. (2017). Efficient DNA-free genome editing of bread wheat using CRISPR/Cas9 ribonucleoprotein complexes. *Nat. Commun.* 8:14261. doi: 10.1038/ncomms14261

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

Liu, X., Qin, R., Li, J., Liao, S., Shan, T., Xu, R., et al. (2020). A CRISPR-Cas9-mediated domain-specific base-editing screen enables functional assessment of ACCase variants in rice. *Plant Biotechnol. J.* 18:1845–1847. doi: 10.1111/pbi.13348

Liu, P., Zhang, X., Zhang, F., Xu, M., Ye, Z., Wang, K., et al. (2021). A virus-derived siRNA activates plant immunity by interfering with ROS scavenging. *Mol. Plant* 14, 1088–1103. doi: 10.1016/j.molp.2021.03.022

Lorence, A., and Verpoorte, R. (2004). Gene transfer and expression in plants. *Methods Mol. Biol.* 267, 329–350. doi: 10.1385/1-59259-774-2:329

Lu, Y., Tian, Y., Shen, R., Yao, Q., Wang, M., Chen, M., et al. (2020). Targeted, efficient sequence insertion and replacement in rice. *Nat. Biotechnol.* 38, 1402–1407. doi: 10.1038/s41587-020-0581-5

Ma, X., Zhang, Q., Zhu, Q., Liu, W., Chen, Y., Qiu, R., et al. (2015). A robust CRISPR/Cas9 system for convenient, high-efficiency multiplex genome editing in monocot and dicot plants. *Mol. Plant* 8, 1274–1284. doi: 10.1016/j.molp.2015.04.007

Majumdar, R., Rajasekaran, K., and Cary, J. W. (2017). RNA interference (RNAi) as a potential tool for control of mycotoxin contamination in crop plants: concepts and considerations. *Front. Plant Sci.* 8:200. doi: 10.3389/fpls.2017.00200

Malabarba, J., Chevreau, E., Dousset, N., Veillet, F., Moizan, J., and Vergne, E. (2021). New strategies to overcome present CRISPR/Cas9 limitations in apple and pear: efficient dechimerization and base editing. *Int. J. Mol. Sci.* 22:319. doi: 10.3390/ijms22010319

Mali, P., Yang, L., Esvelt, K. M., Aach, J., Guell, M., DiCarlo, J. E., et al. (2013). RNA-guided human genome engineering via Cas9. *Science* 339, 823–826. doi: 10.1126/science.1232033

Malnoy, M., Viola, R., Jung, M.-H., Koo, O.-J., Kim, S., Kim, J.-S., et al. (2016). DNA-free genetically edited grapevine and apple protoplast using CRISPR/Cas9 Ribonucleoproteins. *Front. Plant Sci.* 7:1904. doi: 10.3389/fpls.2016.01904

Malzahn, A., Lowder, L., and Qi, Y. (2017). Plant genome editing with TALEN and CRISPR. *Cell Biosci.* 7:21. doi: 10.1186/s13578-017-0148-4

Manghwar, H., and Hussain, A. (2022). Mechanism of tobacco osmotin gene in plant responses to biotic and abiotic stress tolerance: a brief history. *Biocell* 46, 623–632. doi: 10.32604/biocell.2022.017316

Manghwar, H., Hussain, A., Ali, Q., and Liu, F. (2022). Brassinosteroids (BRs) role in plant development and coping with different stresses. *Int. J. Mol. Sci.* 23:1012. doi: 10.3390/ijms23031012

Manghwar, H., Hussain, A., Ali, Q., Saleem, M. H., Abualreesh, M. H., Alatawi, A., et al. (2021). Disease severity, resistance analysis, and expression profiling of pathogenesis-related protein genes after the inoculation of *Fusarium equiseti* in wheat. *Agronomy* 11:2124. doi: 10.3390/agronomy11112124

Mehta, D., Stürchler, A., Anjanappa, R. B., Zaidi, S. S.-A., Hirsch-Hoffmann, M., Gruissem, W., et al. (2019). Linking CRISPR-Cas9 interference in cassava to the evolution of editing-resistant geminiviruses. *Genome Biol.* 20:80. doi: 10.1186/s13059-019-1678-3

Metje-Sprink, J., Menz, J., Modrzejewski, D., and Sprink, T. (2019). DNA-free genome editing: past, present and future. *Front. Plant Sci.* 9:1957. doi: 10.3389/fpls.2018.01957

Metje-Sprink, J., Sprink, T., and Hartung, F. (2020). Genome-edited plants in the field. *Curr. Opin. Biotechnol.* 61, 1–6. doi: 10.1016/j.copbio.2019.08.007

Mishra, R., Nath, J., Bijayalaxmi, M., Raj, M., and Joshi, K. (2021). A single transcript CRISPR/Cas9 mediated mutagenesis of CaERF28 confers anthracnose resistance in chilli pepper (*Capsicum annuum* L.). *Planta* 254, 5–17. doi: 10.1007/s00425-021-03660-x

Molla, K. A., Shih, J., and Yang, Y. (2020). Single-nucleotide editing for zebra3 and wsl5 phenotypes in rice using CRISPR/Cas9-mediated adenine base editors. *aBIOTECH* 1, 106–118. doi: 10.1007/s42994-020-00018-x

Molla, K. A., Sretenovic, S., Bansal, K. C., and Qi, Y. (2021). Precise plant genome editing using base editors and prime editors. *Nat. Plants* 7, 1166–1187. doi: 10.1038/s41477-021-00991-1

Molla, K. A., and Yang, Y. (2019). CRISPR/Cas-mediated base editing: technical considerations and practical applications. *Trends Biotechnol.* 37, 1121–1142. doi: 10.1016/j.tibtech.2019.03.008

Mundt, C. C. (2018). Pyramiding for resistance durability: theory and practice. *Phytopathology* 108, 792–802. doi: 10.1094/PHYTO-12-17-0426-RVW

Nakajima, I., Ban, Y., Azuma, A., Onoue, N., Moriguchi, T., Yamamoto, T., et al. (2017). CRISPR/Cas9-mediated targeted mutagenesis in grape. *PLoS One* 12:e0177966. doi: 10.1371/journal.pone.0177966

Nakka, S., Jugulam, M., Peterson, D., and Asif, M. (2019). Herbicide resistance: development of wheat production systems and current status of resistant weeds in wheat cropping systems. *Crop J.* 7, 750–760. doi: 10.1016/j.cj.2019.09.004

Nekrasov, V., Wang, C., Win, J., Lanz, C., Weigel, D., and Kamoun, S. (2017). Rapid generation of a transgene-free powdery mildew resistant tomato by genome deletion. *Sci. Rep.* 7:482. doi: 10.1038/s41598-017-00578-x

Niño-Liu, D. O., Ronald, P. C., and Bogdanove, A. J. (2006). *Xanthomonas oryzae* pathovars: model pathogens of a model crop. *Mol. Plant Pathol.* 7, 303–324. doi: 10.1111/j.1364-3703.2006.00344.x

Nishida, K., Arazoe, T., Yachie, N., Banno, S., Kakimoto, M., Tabata, M., et al. (2016). Targeted nucleotide editing using hybrid prokaryotic and vertebrate adaptive immune systems. *Science* 353:aaf8729. doi: 10.1126/science.aaf8729

Okada, A., Arndell, T., Borisjuk, N., Sharma, N., Watson-Haigh, N. S., Tucker, E. J., et al. (2019). CRISPR/Cas9-mediated knockout of Ms1 enables the rapid generation of male-sterile hexaploid wheat lines for use in hybrid seed production. *Plant Biotechnol. J.* 17, 1905–1913. doi: 10.1111/pbi.13106

Oliva, R., Ji, C., Atienza-Grande, G., Huguet-Tapia, J. C., Perez-Quintero, A., Li, T., et al. (2019). Broad-spectrum resistance to bacterial blight in rice using genome editing. *Nat. Biotechnol.* 37, 1344–1350. doi: 10.1038/s41587-019-0267-z

Orozco, P. (2018). Argentina and Brazil merge law and science to regulate new breeding techniques. Cornell Alliance Science.

Pandolfi, V., Ribamar Costa Ferreira Neto, J., Daniel da Silva, M., Lindinalva Barbosa Amorim, L., Carolina Wanderley-Nogueira, A., Lane de Oliveira Silva, R., et al. (2017). Resistance (R) genes: applications and prospects for plant biotechnology and breeding. *Curr. Protein Pept. Sci.* 18, 323–334. doi: 10.2174/1389203717666160724195248

Peng, A., Chen, S., Lei, T., Xu, L., He, Y., Wu, L., et al. (2017). Engineering canker-resistant plants through CRISPR/Cas9-targeted editing of the susceptibility gene CsLOB1 promoter in citrus. *Plant Biotechnol. J.* 15, 1509–1519. doi: 10.1111/pbi.12733

Pieterse, C. M. J., and Van Loon, L. C. (2004). NPR1: the spider in the web of induced resistance signaling pathways. *Curr. Opin. Plant Biol.* 7, 456–464. doi: 10.1016/j.pbi.2004.05.006

Porto, E. M., Komor, A. C., Slaymaker, I. M., and Yeo, G. W. (2020). Base editing: advances and therapeutic opportunities. *Nat. Rev. Drug Discov.* 19, 839–859. doi: 10.1038/s41573-020-0084-6

Pradhan, S. K., Nayak, D. K., Mohanty, S., Behera, L., Barik, S. R., Pandit, E., et al. (2015). Pyramiding of three bacterial blight resistance genes for broad-spectrum resistance in deepwater rice variety, Jalmagna. *Rice* 8:51. doi: 10.1186/s12284-015-0051-8

Prihatna, C., Barbetti, M. J., and Barker, S. J. (2018). A novel tomato fusarium wilt tolerance gene. *Front. Microbiol.* 9:1226. doi: 10.3389/fmicb.2018.01226

Pröbsting, M., Schenke, D., Hossain, R., Häder, C., Thurau, T., Wighardt, L., et al. (2020). Loss-of-function of CRT1a (Calreticulin) reduces plant susceptibility to *Verticillium longisporum* in both *Arabidopsis thaliana* and oilseed rape (*Brassica napus*). *Plant Biotechnol. J.* 18, 2328–2344. doi: 10.1111/pbi.13394

Qin, L., Li, J., Wang, Q., Xu, Z., Sun, L., Alariqi, M., et al. (2020). High-efficient and precise base editing of C• G to T• A in the allotetraploid cotton (*Gossypium hirsutum*) genome using a modified CRISPR/Cas9 system. *Plant Biotechnol. J.* 18, 45–56. doi: 10.1111/pbi.13168

Que, Q., Chilton, M.-D. M., de Fontes, C. M., He, C., Nuccio, M., Zhu, T., et al. (2010). Trait stacking in transgenic crops: challenges and opportunities. *GM Crops* 1, 220–229. doi: 10.4161/gmcr.1.4.13439

Ramu, V. S., Paramanantham, A., Ramegowda, V., Mohan-Raju, B., Udayakumar, M., and Senthil-Kumar, M. (2016). Transcriptome analysis of sunflower genotypes with contrasting oxidative stress tolerance reveals individual-and combined-biotic and abiotic stress tolerance mechanisms. *PLoS One* 11:e0157522. doi: 10.1371/journal.pone.0157522

Rees, H. A., and Liu, D. R. (2018). Base editing: precision chemistry on the genome and transcriptome of living cells. *Nat. Rev. Genet.* 19, 770–788. doi: 10.1038/s41576-018-0059-1

Ren, Q., Sretenovic, S., Liu, S., Tang, X., Huang, L., He, Y., et al. (2021). PAM-less plant genome editing using a CRISPR–SpRY toolbox. *Nat. Plants* 7, 25–33. doi: 10.1038/s41477-020-00827-4

Rosa, C., Kuo, Y.-W., Wuriyanghan, H., and Falk, B. W. (2018). RNA interference mechanisms and applications in plant pathology. *Annu. Rev. Phytopathol.* 56, 581–610. doi: 10.1146/annurev-phyto-080417-050044

Rybicki, E. P. (2019). CRISPR–Cas9 strikes out in cassava. *Nat. Biotechnol.* 37, 727–728. doi: 10.1038/s41587-019-0169-0

Sander, J. D., and Joung, J. K. (2014). CRISPR-Cas systems for editing, regulating and targeting genomes. *Nat. Biotechnol.* 32, 347–355. doi: 10.1038/nbt.2842

Saqib, S., Uddin, S., Zaman, W., Ullah, F., Ayaz, A., Asghar, M., et al. (2020). Characterization and phytostimulatory activity of bacteria isolated from tomato (*Lycopersicon esculentum* Mill.) rhizosphere. *Microb. Pathog.* 140:103966. doi: 10.1016/j.micpath.2020.103966

Savary, S., Willocquet, L., Pethybridge, S. J., Esker, P., McRoberts, N., and Nelson, A. (2019). The global burden of pathogens and pests on major food crops. *Nat. Ecol. Evol.* 3, 430–439. doi: 10.1038/s41559-018-0793-y

Schenke, D., and Cai, D. (2020). Applications of CRISPR/Cas to improve crop disease resistance: Beyond inactivation of susceptibility factors. *iScience* 23:101478. doi: 10.1016/j.isci.2020.101478

Schornack, S., Meyer, A., Römer, P., Jordan, T., and Lahaye, T. (2006). Gene-for-gene-mediated recognition of nuclear-targeted AvrBs3-like bacterial effector proteins. *J. Plant Physiol.* 163, 256–272. doi: 10.1016/j.jplph.2005.12.001

Schütte, G., Eckerstorfer, M., Rastelli, V., Reichenbecher, W., Restrepo-Vassalli, S., Ruohonen-Lehto, M., et al. (2017). Herbicide resistance and biodiversity: agronomic and environmental aspects of genetically modified herbicide-resistant plants. *Environ. Sci. Eur.* 29:5. doi: 10.1186/s12302-016-0100-y

Shen, D., Suhrkamp, I., Wang, Y., Liu, S., Menkhaus, J., Verreet, J., et al. (2014). Identification and characterization of micro RNA s in oilseed rape (*Brassica napus*) responsive to infection with the pathogenic fungus *Verticillium longisporum* using Brassica AA (*Brassica rapa*) and CC (*Brassica oleracea*) as reference genomes. *New Phytol.* 204, 577–594. doi: 10.1111/nph.12934

Shimatani, Z., Kashojiya, S., Takayama, M., Terada, R., Arazoe, T., Ishii, H., et al. (2017). Targeted base editing in rice and tomato using a CRISPR-Cas9 cytidine deaminase fusion. *Nat. Biotechnol.* 35, 441–443. doi: 10.1038/nbt.3833

Solangi, Z. A., Yao, Y., Zhang Yani, T. M., Rattar, Q. A., Tahir, S., Mustafa, G., et al. (2021). QTL mapping for seed weight and seed yield-related traits in oil-seed crop rapeseed (*Brassica* sp.). *Int. J. Agric. Res.* 9, 51–60. doi: 10.33500/ijaar.2021.09.005

Song, W.-Y., Wang, G.-L., Chen, L.-L., Kim, H.-S., Pi, L.-Y., Holsten, T., et al. (1995). A receptor kinase-Like protein encoded by the rice disease resistance gene, Xa21. *Science* 270, 1804–1806. doi: 10.1126/science.270.5243.1804

Sprink, T., Eriksson, D., Schiemann, J., and Hartung, F. (2016). Regulatory hurdles for genome editing: process-vs. product-based approaches in different regulatory contexts. *Plant Cell Rep.* 35, 1493–1506. doi: 10.1007/s00299-016-1990-2

Subburaj, S., Chung, S. J., Lee, C., Ryu, S.-M., Kim, D. H., Kim, J.-S., et al. (2016). Site-directed mutagenesis in petunia×hybrida protoplast system using direct delivery of purified recombinant Cas9 ribonucleoproteins. *Plant Cell Rep.* 35, 1535–1544. doi: 10.1007/s00299-016-1937-7

Svitashev, S., Schwartz, C., Lenderts, B., Young, J. K., and Cigan, A. M. (2016). Genome editing in maize directed by CRISPR–Cas9 ribonucleoprotein complexes. *Nat. Commun.* 7:13274. doi: 10.1038/ncomms13274

Tai, T. H., Dahlbeck, D., Clark, E. T., Gajiwala, P., Pasion, R., Whalen, M. C., et al. (1999). Expression of the Bs2 pepper gene confers resistance to bacterial spot disease in tomato. *Proc. Natl. Acad. Sci.* 96, 14153–14158. doi: 10.1073/pnas.96.24.14153

Tian, S., Jiang, L., Cui, X., Zhang, J., Guo, S., Li, M., et al. (2018). Engineering herbicide-resistant watermelon variety through CRISPR/Cas9-mediated base-editing. *Plant Cell Rep.* 37, 1353–1356. doi: 10.1007/s00299-018-2299-0

Tripathi, J. N., Lorenzen, J., Bahar, O., Ronald, P., and Tripathi, L. (2014). Transgenic expression of the rice Xa21 pattern-recognition receptor in banana (M usa sp.) confers resistance to *Xanthomonas campestris* pv. Musacearum. *Plant Biotechnol. J.* 12, 663–673. doi: 10.1111/pbi.12170

Tripathi, J. N., Ntui, V. O., Ron, M., Muiruri, S. K., Britt, A., and Tripathi, L. (2019). CRISPR/Cas9 editing of endogenous banana streak virus in the B genome of *Musa* spp. overcomes a major challenge in banana breeding. *Commun. Biol.* 2:46. doi: 10.1038/s42003-019-0288-7

Uma, B., Rani, T. S., and Podile, A. R. (2011). Warriors at the gate that never sleep: non-host resistance in plants. *J. Plant Physiol.* 168, 2141–2152. doi: 10.1016/j.jplph.2011.09.005

Urnov, F. D., Ronald, P. C., and Carroll, D. (2018). A call for science-based review of the European court's decision on gene-edited crops. *Nat. Biotechnol.* 36, 800–802. doi: 10.1038/nbt.4252

USDA (2018). Secretary Perdue Issues USDA Statement on Plant Breeding Innovation. USDA press release no. 0070. 18. Available at: https://www.usda.gov/media/press-releases/2018/03/28/secretary-perdue-issues-usda-statement-plant-breeding-innovation (Accessed March 12, 2019).

Veillet, F., Chauvin, L., Kermarrec, M.-P., Sevestre, F., Merrer, M., Terret, Z., et al. (2019a). The *Solanum tuberosum* GBSSI gene: a target for assessing gene and base editing in tetraploid potato. *Plant Cell Rep.* 38, 1065–1080. doi: 10.1007/s00299-019-02426-w

Veillet, F., Kermarrec, M. P., Chauvin, L., Guyon-Debast,, A., Chauvin, J. E., Gallois, J., et al. (2020). Prime editing is achievable in the tetraploid potato, but needs improvement. *BioRxiv.* doi: 10.1101/2020.06.18.159111

Veillet, F., Perrot, L., Chauvin, L., Kermarrec, M.-P., Guyon-Debast, A., Chauvin, J.-E., et al. (2019b). Transgene-free genome editing in tomato and potato plants using agrobacterium-mediated delivery of a CRISPR/Cas9 cytidine base editor. *Int. J. Mol. Sci.* 20:402. doi: 10.3390/ijms20020402

Voinnet, O. (2005). Induction and suppression of RNA silencing: insights from viral infections. *Nat. Rev. Genet.* 6, 206–220. doi: 10.1038/nrg1555

Voytas, D. (2017). "Optimizing gene targeting in plants," in *In Vitro Cellular and Developmental Biology-Animal* (New York, NY, USA: Springer), S23–S23.

Voytas, D. F., and Gao, C. (2014). Precision genome engineering and agriculture: opportunities and regulatory challenges. *PLoS Biol.* 12:e1001877. doi: 10.1371/journal.pbio.1001877

Wallace, J. G., Rodgers-Melnick, E., and Buckler, E. S. (2018). On the road to breeding 4.0: unraveling the good, the bad, and the boring of crop quantitative genomics. *Annu. Rev. Genet.* 52, 421–444. doi: 10.1146/annurev-genet-120116-024846

Walton, R. T., Christie, K. A., Whittaker, M. N., and Kleinstiver, B. P. (2020). Unconstrained genome targeting with near-PAMless engineered CRISPR-Cas9 variants. *Science* 368, 290–296. doi: 10.1126/science.aba8853

Waltz, E. (2016). Gene-edited CRISPR mushroom escapes US regulation. *Nature* 532:293. doi: 10.1038/nature.2016.19754

Waltz, E. (2018). With a free pass, CRISPR-edited plants reach market in record time.

Wang, Q., Alariqi, M., Wang, F., Li, B., Ding, X., Rui, H., et al. (2020). The application of a heat-inducible CRISPR/Cas12b (C2c1) genome editing system in tetraploid cotton (*G. hirsutum*) plants. *Plant Biotechnol. J.* 18, 2436–2443. doi: 10.1111/pbi.13417

Wang, Z., Liu, X., Xie, X., Deng, L., Zheng, H., Pan, H., et al. (2021). ABE8e with polycistronic tRNA-gRNA expression cassette significantly improves adenine base editing efficiency in *Nicotiana benthamiana*. *Int. J. Mol. Sci.* 22:5663. doi: 10.3390/ijms222413627

Wang, M.-B., Masuta, C., Smith, N. A., and Shimura, H. (2012). RNA silencing and plant viral diseases. *Mol. Plant-Microbe Interact.* 25, 1275–1285. doi: 10.1094/MPMI-04-12-0093-CR

Wang, Y., Shan, Q., Cheng, X., Zhang, Y., Liu, J., Gao, C., et al. (2014). Simultaneous editing of three homoeoalleles in hexaploid bread wheat confers heritable resistance to powdery mildew. *Nat. Biotechnol.* 32, 947–951. doi: 10.1038/nbt.2969

Wang, M., Weiberg, A., Dellota, E. Jr., Yamane, D., and Jin, H. (2017). Botrytis small RNA Bc-siR37 suppresses plant defense genes by cross-kingdom RNAi. *RNA Biol.* 14, 421–428. doi: 10.1080/15476286.2017.1291112

Wang, L., Zhao, L., Zhang, X., Zhang, Q., Jia, Y., Wang, G., et al. (2019). Large-scale identification and functional analysis of NLR genes in blast resistance in the Tetep rice genome sequence. *Proc. Natl. Acad. Sci.* 116, 18479–18487. doi: 10.1073/pnas.1910229116

Weiberg, A., Wang, M., Lin, F.-M., Zhao, H., Zhang, Z., Kaloshian, I., et al. (2013). Fungal small RNAs suppress plant immunity by hijacking host RNA interference pathways. *Science* 342, 118–123. doi: 10.1126/science.1239705

Wu, C.-H., Abd-El-Haliem, A., Bozkurt, T. O., Belhaj, K., Terauchi, R., Vossen, J. H., et al. (2017). NLR network mediates immunity to diverse plant pathogens. *Proc. Natl. Acad. Sci.* 114, 8113–8118. doi: 10.1073/pnas.1702041114

Xing, S., Chen, K., Zhu, H., Zhang, R., Zhang, H., Li, B., et al. (2020). Fine-tuning sugar content in strawberry. *Genome Biol.* 21:230. doi: 10.1186/s13059-020-02146-5

Xu, R., Li, J., Liu, X., Shan, T., Qin, R., and Wei, P. (2020a). Development of plant prime-editing systems for precise genome editing. *Plant Commun.* 1:100043. doi: 10.1016/j.xplc.2020.100043

Xu, Z., Xu, X., Gong, Q., Li, Z., Li, Y., Wang, S., et al. (2019). Engineering broad-Spectrum bacterial blight resistance by simultaneously disrupting variable TALE-binding elements of multiple susceptibility genes in Rice. *Mol. Plant* 12, 1434–1446. doi: 10.1016/j.molp.2019.08.006

Xu, W., Yang, Y., Liu, Y., Kang, G., Wang, F., Li, L., et al. (2020b). Discriminated sgRNAs-based SurroGate system greatly enhances the screening efficiency of plant base-edited cells. *Mol. Plant* 13, 169–180. doi: 10.1016/j.molp.2019.10.007

Xu, W., Zhang, C., Yang, Y., Zhao, S., Kang, G., He, X., et al. (2020c). Versatile nucleotides substitution in plant using an improved prime editing system. *Mol. Plant* 13, 675–678. doi: 10.1016/j.molp.2020.03.012

Yarra, R., and Sahoo, L. (2021). Base editing in rice: current progress, advances, limitations, and future perspectives. *Plant Cell Rep.* 40, 595–604. doi: 10.1007/s00299-020-02656-3

Yasmin, H., Naeem, S., Bakhtawar, M., Jabeen, Z., Nosheen, A., Naz, R., et al. (2020). *Halotolerant rhizobacteria Pseudomonas pseudoalcaligenes* and *Bacillus subtilis* mediate systemic tolerance in hydroponically grown soybean (*Glycine max* L.) against salinity stress. *PLoS One* 15:e0231348. doi: 10.1371/journal.pone.0231348

Yin, X., Anand, A., Quick, P., and Bandyopadhyay, A. (2019b). Editing a stomatal developmental gene in rice with CRISPR/Cpf1. *Methods Mol. Biol.* 1917, 257–268. doi: 10.1007/978-1-4939-8991-1_19

Yin, K., Han, T., Xie, K., Zhao, J., Song, J., and Liu, Y. (2019a). Engineer complete resistance to cotton leaf curl multan virus by the CRISPR/Cas9 system in *Nicotiana benthamiana*. *Phytopathol. Res.* 1, 1–9. doi: 10.1186/s42483-019-0017-7

Zafar, K., Sedeek, K. E. M., Rao, G. S., Khan, M. Z., Amin, I., Kamel, R., et al. (2020). Genome editing technologies for Rice Improvement: Progress, prospects, and safety concerns. *Front. Genome Ed.* 2:5. doi: 10.3389/fgeed.2020.00005

Zaidi, S. S.-A., Mukhtar, M. S., and Mansoor, S. (2018). Genome editing: targeting susceptibility genes for plant disease resistance. *Trends Biotechnol.* 36, 898–906. doi: 10.1016/j.tibtech.2018.04.005

Zhang, M., and Coaker, G. (2017). Harnessing effector-triggered immunity for durable disease resistance. *Phytopathology* 107, 912–919. doi: 10.1094/PHYTO-03-17-0086-RVW

Zhang, P., Du, H., Wang, J., Pu, Y., Yang, C., Yan, R., et al. (2020b). Multiplex CRISPR/Cas9-mediated metabolic engineering increases soya bean isoflavone content and resistance to soya bean mosaic virus. *Plant Biotechnol. J.* 18, 1384–1395. doi: 10.1111/pbi.13302

Zhang, Z., Ge, X., Luo, X., Wang, P., Fan, Q., Hu, G., et al. (2018b). Simultaneous editing of two copies of Gh14-3-3d confers enhanced transgene-clean plant defense against *Verticillium dahliae* in allotetraploid upland cotton. *Front. Plant Sci.* 9:842. doi: 10.3389/fpls.2018.00842

Zhang, J., Li, W., Xiang, T., Liu, Z., Laluk, K., Ding, X., et al. (2010). Receptor-like cytoplasmic kinases integrate signaling from multiple plant immune receptors and are targeted by a *Pseudomonas syringae* effector. *Cell Host Microbe* 7, 290–301. doi: 10.1016/j.chom.2010.03.007

Zhang, R., Liu, J., Chai, Z., Chen, S., Bai, Y., Zong, Y., et al. (2019b). Generation of herbicide tolerance traits and a new selectable marker in wheat using base editing. *Nat. Plants* 5, 480–485. doi: 10.1038/s41477-019-0405-0

Zhang, A., Liu, Y., Wang, F., Li, T., Chen, Z., Kong, D., et al. (2019a). Enhanced rice salinity tolerance via CRISPR/Cas9-targeted mutagenesis of the OsRR22 gene. *Mol. Breed.* 39:47. doi: 10.1007/s11032-019-0954-y

Zhang, M., Liu, Q., Yang, X., Xu, J., Liu, G., Yao, X., et al. (2020a). CRISPR/Cas9-mediated mutagenesis of Clpsk1 in watermelon to confer resistance to *Fusarium oxysporum* f. sp. niveum. *Plant Cell Rep.* 39, 589–595. doi: 10.1007/s00299-020-02516-0

Zhang, Y., Malzahn, A. A., Sretenovic, S., and Qi, Y. (2019c). The emerging and uncultivated potential of CRISPR technology in plant science. *Nat. Plants* 5, 778–794. doi: 10.1038/s41477-019-0461-5

Zhang, Y., Pribil, M., Palmgren, M., and Gao, C. (2020c). A CRISPR way for accelerating improvement of food crops. *Nat. Food* 1, 200–205. doi: 10.1038/s43016-020-0051-8

Zhang, T., Zhao, Y.-L., Zhao, J.-H., Wang, S., Jin, Y., Chen, Z.-Q., et al. (2016). Cotton plants export microRNAs to inhibit virulence gene expression in a fungal pathogen. *Nat. Plants* 2:16153. doi: 10.1038/nplants.2016.153

Zhang, T., Zheng, Q., Yi, X., An, H., Zhao, Y., Ma, S., et al. (2018a). Establishing RNA virus resistance in plants by harnessing CRISPR immune system. *Plant Biotechnol. J.* 16, 1415–1423. doi: 10.1111/pbi.12881

Zhao, H., Shar, A. G., Li, S., Chen, Y., Shi, J., Zhang, X., et al. (2018). Effect of straw return mode on soil aggregation and aggregate carbon content in an annual maize-wheat double cropping system. *Soil Tillage Res.* 175, 178–186. doi: 10.1016/j.still.2017.09.012

Zhu, S., Li, Y., Vossen, J. H., Visser, R. G. F., and Jacobsen, E. (2012). Functional stacking of three resistance genes against *Phytophthora infestans* in potato. *Transgenic Res.* 21, 89–99. doi: 10.1007/s11248-011-9510-1

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

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

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

frontiers in Plant Science

ORIGINAL RESEARCH
published: 17 March 2022
doi: 10.3389/fpls.2022.852886

Transcriptome Analysis Reveals a Gene Expression Pattern That Contributes to Sugarcane Bud Propagation Induced by Indole-3-Butyric Acid

Lin Xu[1†], Zhi-Nian Deng[1†], Kai-Chao Wu[1†], Mukesh Kumar Malviya[1], Manoj Kumar Solanki[2], Krishan K. Verma[1], Tian Pang[1], Yi-Jie Li[1], Xiao-Yan Liu[1], Brijendra Kumar Kashyap[3], Eldessoky S. Dessoky[4,5], Wei-Zan Wang[1] and Hai-Rong Huang[1]**

[1]Key Laboratory of Sugarcane Biotechnology and Genetic Improvement (Guangxi), Ministry of Agriculture and Rural Area, Sugarcane Research Center, Chinese Academy of Agricultural Sciences, Guangxi Key Laboratory of Sugarcane Genetic Improvement, Sugarcane Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, China, [2]Plant Cytogenetics and Molecular Biology Group, Institute of Biology, Biotechnology and Environmental Protection, Faculty of Natural Sciences, University of Silesia in Katowice, Katowice, Poland, [3]Department of Biotechnology Engineering, Institute of Engineering and Technology, Bundelkhand University, Jhansi, India, [4]Department of Plant Genetic Transformation, Agriculture Genetic Engineering Research Institute, Agriculture Research Center, Giza, Egypt, [5]Department of Biology, College of Science, Taif University, Taif, Saudi Arabia

Sugarcane is a cash crop that plays an integral part in the sugar industry. The Sustainable Sugarcane Initiative (SSI) has been adopted globally, ensuring enough and aiming for more yield, helping increase disease-free sugarcane cultivation. Single-bud seeds could be the best approach for sugarcane cultivation. Indole-3-butyric acid (IBA) is a rooting agent utilized significantly in seedling propagation. Greenhouse experiment results discovered the significant growth promotion in sugarcane seedlings and accumulation of plant hormones at 100 ppm IBA. Next, we performed transcriptomic analysis of sugarcane buds using RNA sequencing and compared their gene expression during root development due to affect of IBA (100 ppm). A total of 113,475 unigenes were annotated with an average length of 836 bp (N50 = 1,536). The comparative RNA-seq study between the control (CK) and IBA-treated (T) buds showed significant differentially expressed unigenes (494 upregulated and 2086 downregulated). The IBA influenced major biological processes including metabolic process, the cellular process, and single-organism process. For cellular component category, cell, cell part, organelle, membrane, and organelle part were mainly affected. In addition, catalytic activity and binding were primarily affected in the molecular function categories. Furthermore, the expression of genes related to plant hormones and signaling pathways was analyzed by qRT-PCR, which was consistent with the RNA-seq expression profile. This study provides new insights into the IBA response to the bud sprouting in sugarcane based on RNA sequencing, and generated information could help further research on breeding improvement of sugarcane.

Keywords: IBA, single-bud seed, sugarcane, root, transcriptome

OPEN ACCESS

Edited by:
Parviz Heidari,
Shahrood University of Technology, Iran

Reviewed by:
Jen-Tsung Chen,
National University of Kaohsiung, Taiwan
Abdullah,
Quaid-i-Azam University, Pakistan

***Correspondence:**
Hai-Rong Huang
hhrong15937@126.com
Wei-Zan Wang
wwz003411@126.com

[†]These authors have contributed equally to this work

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

Received: *11 January 2022*
Accepted: *14 February 2022*
Published: *17 March 2022*

Citation:
Xu L, Deng Z-N, Wu K-C, Malviya MK, Solanki MK, Verma KK, Pang T, Li Y-J, Liu X-Y, Kashyap BK, Dessoky ES, Wang W-Z and Huang H-R (2022) Transcriptome Analysis Reveals a Gene Expression Pattern That Contributes to Sugarcane Bud Propagation Induced by Indole-3-Butyric Acid. Front. Plant Sci. 13:852886. doi: 10.3389/fpls.2022.852886

INTRODUCTION

Sugarcane (*Saccharum officinarum*) is a globally grown commercial crop that belongs to the family Gramineae and is grown extensively all over China, which is the third-largest sugarcane producer in the world after Brazil and India with a production of around 11.6 MT of white sugar (Zhang and Govindaraju, 2018). Guangxi province contributes the highest (7.21 MT), which is one of the main sugarcane-growing areas in China other than Guangdong and Yunnan (Chen et al., 2021). The Chinese sugar industry played a significant part in several bio-product processing units like ethanol, energy (biomass power, bioethanol), yeast production, paper, chemicals, cane juice, and other associated byproducts (Lu and yang, 2019; Khaire et al., 2021). The sugar industry has a significant annual output that makes it very important for the Chinese economy, and it contributes significantly to the development of rural infrastructures like roads, education, medicine, and other facilities, consecutively playing a vital role in the livelihood of many sugarcane farmers and workers directly employed in sugar mills.

Sugarcane productivity is influenced by seed quality, land treatment, monoculture cropping system, fertilization timing, climate change, etc. Still, the most crucial factor is the planting distance, which decides the volume of nutrients absorbed by the plants (Singh et al., 2019). The investment in the area of the sugarcane plantation is vast, and since the economy rests on it, sometimes the conventional methods might incur a loss in its production. This problem can be handled by adopting the Sustainable Sugarcane Initiative (SSI), which combines different technologies to enhance the yield of sugarcane by enabling the crop to utilize plant nutrients (Loganandhan et al., 2013). One of the methods used is the bud chip technology, also called single-bud planting (Mohanty and Nayak, 2021). It uses the axillary buds from the sugarcane stems, where a root primordial along with the small quantity of tissue attached to the bud is used to regenerate the sugarcane plant. This technology saves more than 70% of the stalk material and reduces the cost of cultivation. The advantages of using bud chips are (1) high budding rate, (2) quicker stem development, (3) harvesting time decreased, (4) increased yield, and (5) free from disease and infection (Mohanty and Nayak, 2021).

Considering the socio-economic state of China, which is dependent on the sugar industry, it is essential for scientists working on a micro level, to the farmers, working at a macro level through business enthusiasts, educationists, marketing firms, and industrialists, to come together and help toward the increase of its production (Morris et al., 2017). However, the sugarcane plant is an ideal model plant species extensively researched by plant biologists and agricultural specialists on physiology, biochemistry, and molecular biology (Li et al., 2016a). One of the most critical aspects of the plant is the development and germination of its roots and leaves. For that, the rooting system should be proficient and well established. Phytohormone, like auxin, plays a vital role in establishing the root system (Jing and Strader, 2019). There have been many studies where scientists have used indole-3-butyric acid (IBA) and indole 3-acetic acid (IAA) to study *in vitro* the role of these hormones in inducing the roots (Ludwig-Muller et al., 2005; Rout, 2006; Pacurar et al., 2014). IAA and IBA have been seen to induce adventitious root formation in mung bean (Li et al., 2016a), lotus (Libao et al., 2018), tea (Wei et al., 2014), sugarcanes (Li et al., 2020), strawberry (Hossain and Gony, 2020), carrot taproot (Khadr et al., 2020), Arabidopsis (Yang et al., 2021), and *Euryodendron excelsum* HT (Xiong et al., 2022). Studies have shown that genes expressed predominantly on IBA treatment are involved in cell replication and weakening (Brinker et al., 2004). A protein phosphatase 2A gene was upregulated during the IBA-induced adventitious root formation on Arabidopsis stem segments (Fattorini et al., 2017). Transcriptome analysis has previously been performed on sugarcane internodes (Kasirajan et al., 2018; Wang et al., 2019), root (Zeng et al., 2015; Dharshini et al., 2019; Malviya et al., 2020), leaves (Li et al., 2016a; Liu et al., 2018), mature and furled immature leaves (Manechini et al., 2021), and stalk (Zhu et al., 2021). To the best of our knowledge, this is the first-of-its-kind attempt to study the transcriptome analysis of the single bud of sugarcane in response to IBA. The present study focuses on applying IBA on single-bud seeds to find out the critical differential expressed genes and metabolic pathways on sugarcane rooting and provides a theoretical basis for enhancing sugarcane growth through phytohormone.

MATERIALS AND METHODS

Plant Material and Pot Experiment

The sugarcane variety Guitang 55 (GT55) was obtained from the germplasm of Sugarcane Research Institute, Guangxi Academy of Agricultural Sciences, Nanning, Guangxi, China. The GT55 variety is the least susceptible to bacterial infection. Therefore, we chose GT55 as the experimental material to reduce experimental error. Single bud of sugarcane sets (length 3 cm) with 3–5th buds was taken from cane stem of variety GT55 by counting ground eyes as first buds. Single buds of GT55 sugarcane sets (length 3 cm) were cut and immersed in the IBA at different concentrations (50, 100, and 200 ppm) for 10 min. The control set was immersed into the sterilized distilled water for 10 min as control. All of the single sugarcane buds were transplanted into pots containing 3.5 kg of field soil with moderate fertility (each kg soil containing 21 g organic matter, 130 mg available N, 85 mg available P, 74 mg available K, 0.81 total N, 3.10 g total P, 7.10 g total K and pH was 6.21). Soil analysis was performed according to Solanki et al. (2019) in the Soil testing laboratory of Guangxi University, Nanning, Guangxi. The plants were grown in a greenhouse with temperatures maintained between 25 to 30°C and relative humidity between 50 to 80%. As per the requirement, one time autoclaved tap water was used to irrigate sugarcane plants. After 35 days, top visible dewlap leaf samples were detached from 5 plants for each treatment and frozen in liquid nitrogen for phytohormone analysis. Five plants samples were separated from each treatment for plant growth parameter analysis.

Next, we performed an experiment for transcriptome analysis of IBA (100 mg/L) treated single buds of GT55 sugarcane (T) and the sterilized distilled water called control (CK). The sugarcane seedlings were transplanted into pots and watered same in the greenhouse. After 2 days, all the sugarcane buds were taken out and washed. After the sugarcane rooting of the single bud was cut two sides of the cross section and removed, the single eye sets were cut in half lengthwise. The half lengthwise were wrapped with tin paper, stored in liquid nitrogen for 40 min, and delivered by dry ice. The Schematic diagram of the overall experimental design is present in **Figure 1**.

Quantification of Plant Hormones

To quantify the concentrations of the hormones, leaf extract was extracted according to Li et al. (2016a). Briefly, leaf sample (1 g) was crushed in iced mortar and pestle with 5 ml of 80% methanol (v/v) containing 1 mM butylated hydroxytoluene, and the homogenate was incubated at 4°C for 12 h, then centrifuged at 10,000 × *g* for 20 min. The supernatant was collected and then eluted through a Sep-Pak C18 cartridge (Waters, Milford, MA, United States) to get rid of the pigment. The solution was lyophilized as powder and liquefied with 2 ml of phosphate-buffered saline. Next, concentrations of abscisic acid (ABA), indole-3-acetic acid (IAA), and gibberellin (GA3) were quantified by an enzyme-linked immunosorbent assay described by Yang et al. (2001). A 96-well microtiter plate pre-coated with coating buffer containing synthetic ovalbumin conjugates for IAA, GA3, and ABA was used according to manufacturer protocol (China Agricultural University, Beijing, China). In brief, 50μLsample diluted in assay buffer was added to each well, followed by 50 μl of diluted (1:2,000) antibodies in assay buffer. The plates were incubated for 45 min at 37°C and then washed four times with scrubbing buffer, and anti-mouse IgG was added to each well and incubated for 30 min at 37°C, and washed. Next, 100 μl of a 1.5 mg mL^{-1} ortho-phenylenediamine substrate solution and 0.04% (v/v) of 30% H_2O_2 in substrate buffer were added. The enzyme reaction was carried out in the dark at 37°C for 15 min and then stopped by adding 50 μL of 2 M $H_2SO_4$to each well. The absorbance was recorded at 490 nm. Concentrations of IAA, GA3, and ABA were calculated by standard curve data. All the experiments were performed in five biological replicates.

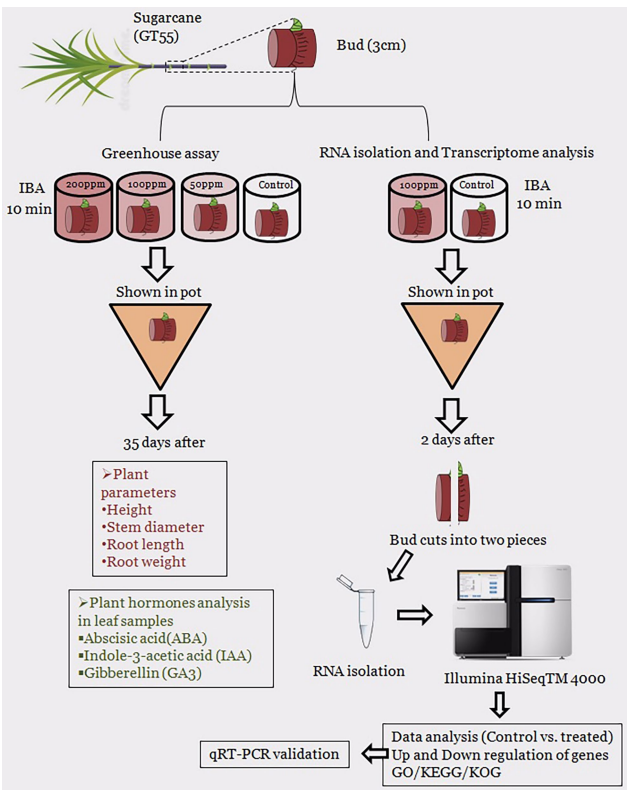


FIGURE 1 | Schematic diagram of an overall research experiment that used in the current study.

RNA Isolation, cDNA Library Preparation, and Sequencing

Total RNA was extracted according to the manufacturer's procedure using a TRIzol reagent kit (Invitrogen, Carlsbad, CA, United States). The quality of the RNA was determined using an Agilent 2,100 Bioanalyzer (Agilent Technologies, Palo Alto, CA, United States) and RNAse-free agarose gel electrophoresis. Following total RNA extraction, eukaryotic mRNA was enriched using Oligo (dT) beads, whereas prokaryotic mRNA was enriched using the Ribo-ZeroTM Magnetic Kit to remove rRNA (Epicentre, Madison, WI, United States). The enriched mRNA was then fragmented into short fragments with fragmentation buffer before being reverse transcribed into cDNA using random primers. DNA polymerase I, RNase H, dNTP, and buffer were used to make second-strand cDNA. The cDNA fragments were then purified using a QiaQuick PCR extraction kit (Qiagen, Venlo, Netherlands), end-repaired, poly (A) added, and ligated to Illumina sequencing adapters. The ligation products were size selected by agarose gel electrophoresis; PCR amplified and sequenced using Illumina HiSeqTM 4,000 by Gene Denovo Biotechnology Co. assembly and analysis.

The raw reads were processed under the data quality control protocol to ensure the data quality before further analysis. To filter low-quality data, fastp (Chen et al., 2018) was used to perform quality control on the offline raw reads and get clean reads. It removes low-quality reads (the number of bases with a quality value $Q \leq 20$ accounts for more than 50% of the entire read). After the data is filtered, the base composition and quality distribution were analyzed to display the data quality visually. The more balanced the base composition, the higher the quality and the more accurate the subsequent analysis. The reads were assembled using the Trinity software (Grabherr et al., 2011). The software combines reads of a certain length of the overlap to form contigs, which are longer fragments without N. These give rise to unigenes, which are clusters of genes that is unique for a particular function. This software also sorts all unigenes from longest to shortest and accumulates the lengths in turn. The quality of the assembly result can be evaluated from the N50 value. When the accumulated fragment length reaches 50% of the total fragment

length (the length of all Unigenes), the length and quantity of that fragment are the Unigene N50 length and quantity. The longer the Unigene N50, the smaller the quantity, the better the assembly quality. To access the integrity of the assembly, BUSCO (Simao et al., 2015) Software was used. Based on the results of each sample unigene, PCA analysis and calculation of the Pearson correlation coefficient between the control (CK) and IBA-treated (T) samples were done using the R suite to help exclude outliers.

Functional Annotation and Classification of DEGs

The unigenes were further annotated. First, blastx was used to align the Unigene sequence according to their sequence similarity. To rule out any discrepancy, four protein databases were used. Swiss-Prot and NCBI NR database are two well-known protein databases, among which Swiss-Prot is strictly selected to eliminate any redundancy, NR (Non-Redundant Protein Sequence Database) included in NCBI, is equivalent to a cross-reference based on nucleic acid sequence, linking nucleic acid data and protein data, COG/KOG, a database for orthologous classification of gene products, and KEGG. This database systematically analyzes the metabolic pathways and functions of gene products in cells, with value of $e < 0.00001$, to obtain the protein function annotation information of the Unigene (**Table 1**).

To further check the processes where the expressed unigenes play a role, the Gene Ontology (Biological Processes, Cellular Components, and Molecular Functions) analysis was done using Blast2GO (Conesa et al., 2005) and GO (Ashburner et al., 2000) software. Gene Ontology (referred to as GO) is an internationally standardized gene function classification system that provides a set of dynamically updated standard vocabulary (controlled vocabulary) to comprehensively describe the attributes of genes and gene products in organisms. On the one hand, GO function analysis provides annotations for the GO function classification of differentially expressed genes; on the other hand, it gives a significant enrichment analysis of GO functions of differentially expressed genes. Furthermore, COG and KEGG pathway annotations were performed using Blastall software against COG and KEGG databases. Principal component analysis (PCA) and heatmap analysis were performed with R package models.[1]

Identification of Differentially Expressed Genes

The differential gene expression analysis was done using the software edgeR v14 (Robinson et al., 2010), where the reads were normalized, followed by the calculation of the probability of hypothesis test (value of p) according to the model and Finally, the multiple hypothesis test correction is carried out to obtain the FDR value (False discovery rate). Genes were considered to be significantly differentially expressed when FDR value < 0.05 and $|\log2FC| > 1$. Differential gene expression patterns are hierarchically clustered, and heat maps are used to present clustering results. Genes with similar expression patterns may have common functions or participate in common metabolic and signaling pathways.

[1]http://www.r-project.org/

TABLE 1 | Summary for the BLASTx results of sugarcane transcriptome data set.

Annotation database	Annotation number
Total Unigenes	113,475
Nr	67,143
KEGG	58,470
COG	32,246
Swiss-Prot	40,453
Annotation genes	69,255
Without annotation gene	44,220

Quantitative Real-Time PCR Assay

For expression validation of differentially expressed genes, special primers were designed for 8 differentially expressed genes were selected by using Premier 5.0 software. Sequencing PCR products checked the specificity of each primer set, and efficiencies of the different primer sets were similar. Details of all primers are described in **Supplementary Table S1**. Total RNA used as the template for reverse transcriptase reactions was initially treated with DNase I Amplification Grade enzyme (Invitrogen), an aliquot of treated RNA was used in qPCR to rule out DNA contamination. cDNA synthesis was done using SuperScript First-Strand Synthesis System for RT-PCR (Invitrogen) and random hexamers and oligo(dT) primers. qPCR reaction was performed in a volume of 20 µl containing 10 µl 2 × All-in-OneTM (GeneCopoeia, Los Angeles, United States) qPCR Mix, 2 µl cDNA, 1 µl of each 4 µM primer, and 6 µl RNase-free sterile water. To normalize the relative expression of selected genes, the GAPDH gene was used as a reference. qPCR was performed using an iQ5 Real-Time PCR Detection System (Bio-Rad, Hercules, United States). PCR reactions were performed at 95°C for 10 min, followed by 40 cycles of 95°C for 10 s, 60°C for 20 s, and 72°C for 20 s. Melting curve analysis was conducted for each reaction to confirm the specificity of the reaction, and all the cDNA samples were analyzed in triplicate. Relative expression levels of candidate genes were calculated using the $2^{-\Delta\Delta Ct}$ method (Livak and Schmittgen, 2001).

Statistical Analysis

Statistical processing of plant phenotype data, hormone content, and qRT-PCR data was performed in Excel. The experiments were conducted in replicates, and data were analyzed using standard analysis of variance (ANOVA) followed by Duncan's multiple range test (DMRT) through Origin 2017SR2 software (Northampton, MA, United States).

RESULT

Affect of IBA on Plant Growth Parameters and Plant Hormones

Affect of different concentrations of IBA on sugarcane bud sprouting was identified based on the greenhouse experiment.

In comparison to control, IBA-treated plants were showed significant enrichment in the plant height, stem diameter, root length, and root weight (**Figure 2**). IBA-treated plants (100 and 200 ppm) height was enhanced significantly ($p < 0.05$) over the control. In the case of stem diameter, 100 ppm IBA showed relatively higher growth ($p < 0.05$) than the control plants. For root length, significant ($p < 0.05$) enhancement recovered with 100 ppm IBA over the control plants (**Figures 2, 3**). Similarly, root weight increased significantly in 100 ppm IBA-treated plants, and 200 ppm treated plants. Plant hormones analysis in response to IBA, ABA, IAA, and GA3 were determined. Compared to the control, the leaves of the 100 and 200 ppm IBA-treated plants had significant concentrations of all three hormones (**Figure 4**). The ABA concentration was enhanced by about 44% over to the control in IBA-treated (100 and 200 ppm) plant leaves. For IAA concentration, the maximum quantity resulted in 100 and 200 ppm treated plant leaves showing up to 27.1 and 28.8% enhancement over control, respectively. Moreover, a higher concentration of GA3 also recovered with 100, and 200 ppm treated plants than the control (**Figure 4**).

Transcriptome Profile of RNA Samples of Sugarcane Buds

To obtain genes involved in IBA-induced sugarcane buds, we generate two cDNA libraries from the control and IBA-treated samples (**Supplementary Table S2**). An average of 66.758 million clean reads with Q20 values from 99.79 to 99.86% per sample were generated. Data pre-processing statistics and quality control are presented in **Supplementary Table S3**. These data were then deposited in the National Center for Biotechnology Information (NCBI) with the accession number PRJNA766098. After trimming low-quality sequences, the software Trinity performed *de novo* assembly with combined cleaned reads from both libraries. As a result, the sequencing data were assembled into 113,475 unigenes with lengths ranging from 200 to 16,723 bp (mean length = 836 bp).

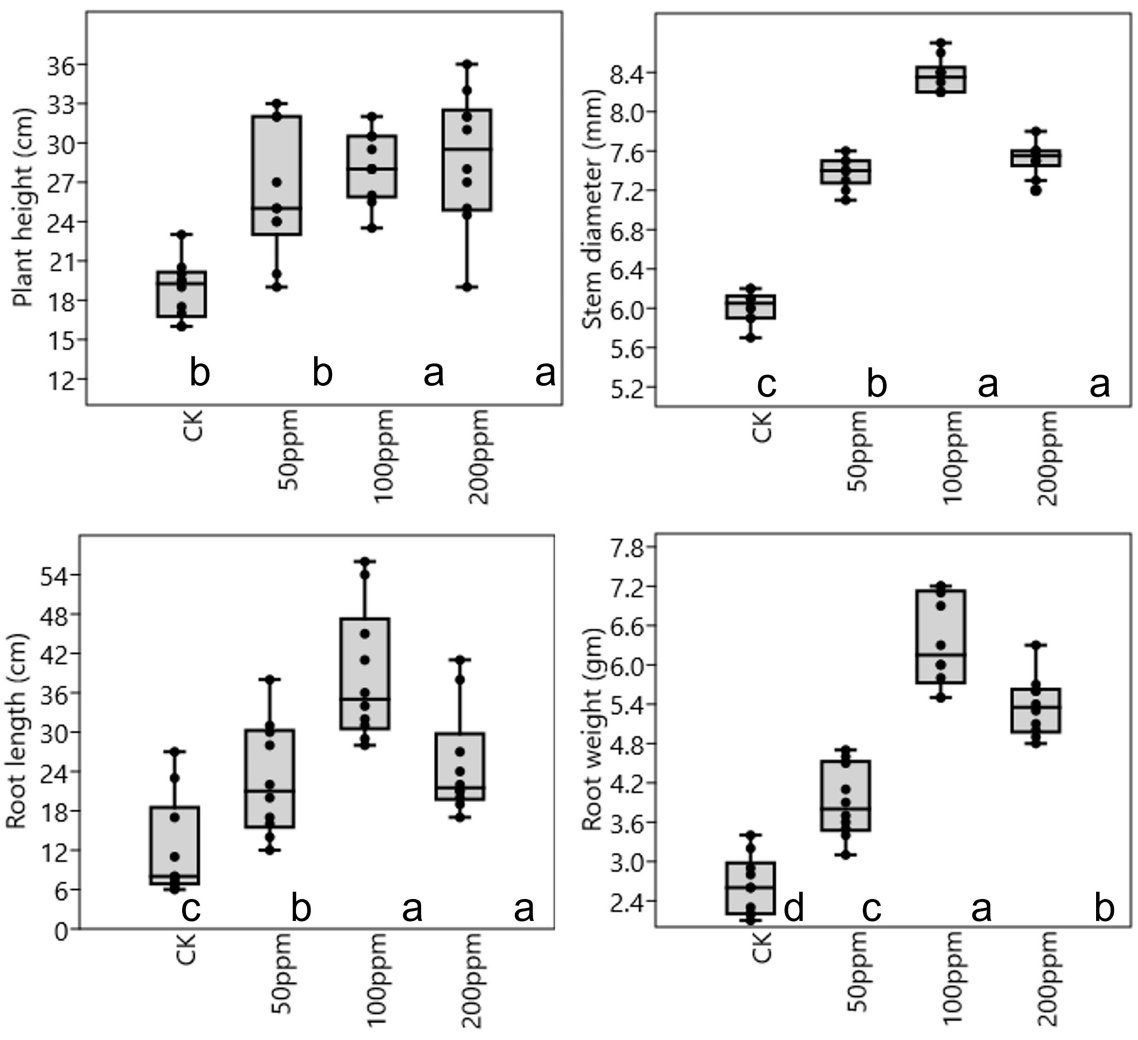


FIGURE 2 | Plant physiological parameter in response to different concentrations of IBA. Bars with different letters indicate significant differences at the $p < 0.05$ according to DMRT.

Total length of the unigens was 94.94 Mb (94938,961 bp). GC percentage of the assembled transcripts was 50.64% and N50 length was 1,536. The genetic testing data also represents the percent sequenced gene among all samples varied from 58 to 88% (**Supplementary Table S4**). The unigene lengths distribution and frequency percentage are shown in **Supplementary Figure S1**. Distribution data showed the quality of assembly for further analysis. To investigate the difference between CK and T samples, the R statistical package was used to assess gene expression levels and determine DEGs. All genes expressed by different samples of CK and T are shown in a violin plot and curve plots (**Figures 5A,B**). Additionally, a principal component analysis (PCA) graph was made for both groups, and the higher difference was found between them specially, in the T samples (**Supplementary Figure S2**) and Pearson comparison clustering showed the similarity among the groups (**Figure 5C**), both groups were clustered together with replicates.

FIGURE 3 | Response of different concentrations of IBA on sugarcane root.

Functional Annotation of DEGs

A total of 113,475 unigenes were annotated by blasting their sequences against four public databases. Significant annotation matches were found for 67,143 unigenes (59%) in the Non-Redundant Protein Sequence Database (Nr database), 58,470 (52%) unigenes in the Kyoto Encyclopedia of Genes and Genomes (KEGG) database, 40,453 (36%) unigenes in the Swiss-Prot database, 32,246 unigenes (28%) in the COG/KOG database. A total of 27,539 unigenes were found common among all databases. Four major databases are annotated in the Venn diagram (**Figure 6A**), and maximum unique genes were annotated with Nr database, followed by Swiss-Prot, KEGG, and COG. Interestingly, KEGG-Nr and KEGG-Swiss-Prot shared the maximum common unigenes. The NR database queries revealed that 67,143 unigenes annotated sequences were aligned with different plant species to known nucleotide sequences similarity with other plants. Annotation results showed 27, 15, 10, 8, 4, and 3%, matching with *Sorghum bicolor*, *Zea mays*, other species, *Quercus suber*, *Oryza sativa Japonica* Group, *Arabidopsis thaliana*, and *Searia italica*, respectively (**Figure 6B**). These results support the genomic similarity of sugarcane tissues with the other grass plants, such as *Zea mays* and *Oryza sativa*.

The annotation of GO terms discovered that unigenes 139,152 (50%) were assigned to biological processes, 97,637 (35%) to molecular functions, and 40,001 (14%) to cellular components (**Supplementary Figure S3**). Most annotated unigenes in biological processes were involved in "metabolic process" 23,453 (14%), "cellular process" 23,286 (14%), and "single-organism process" 19,215 (14%). In the cellular component category, most annotated unigenes were annotated as "cell" 22,875 (23%), "cell part" 22,820 (23%) and "organelle" 19,686 (20%). In the molecular functions, most annotated unigenes were categorized as "catalytic activity" 17,911 (45%) and "binding" 16,876 (42%). For KEGG analysis, 14,854 unigenes were classified in 141 pathways based

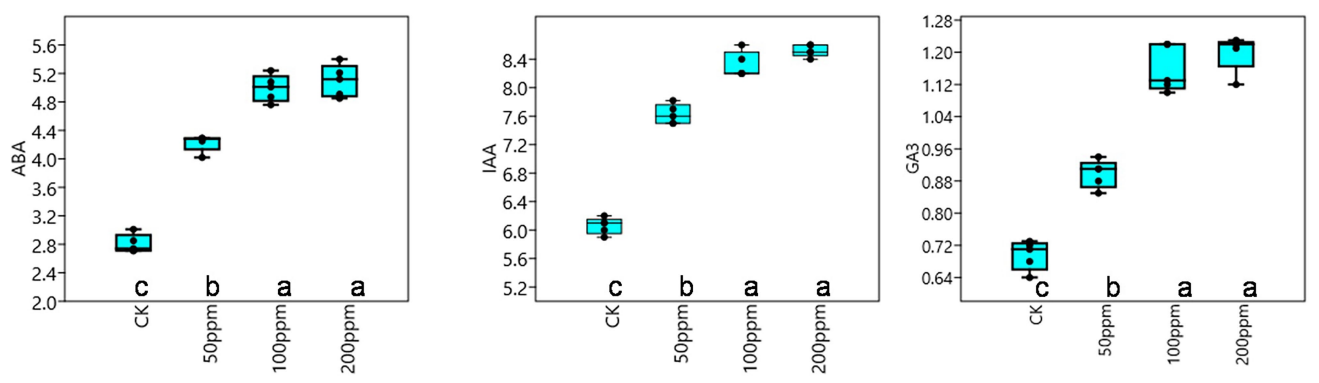


FIGURE 4 | Response of IBA on different endogenous hormones, ABA (μg g^{-1} FW), IAA (nMol g^{-1} FW), and GA_3 (nMol g^{-1} FW). Bars with different letters indicate significant differences at the $p < 0.05$ according to DMRT.

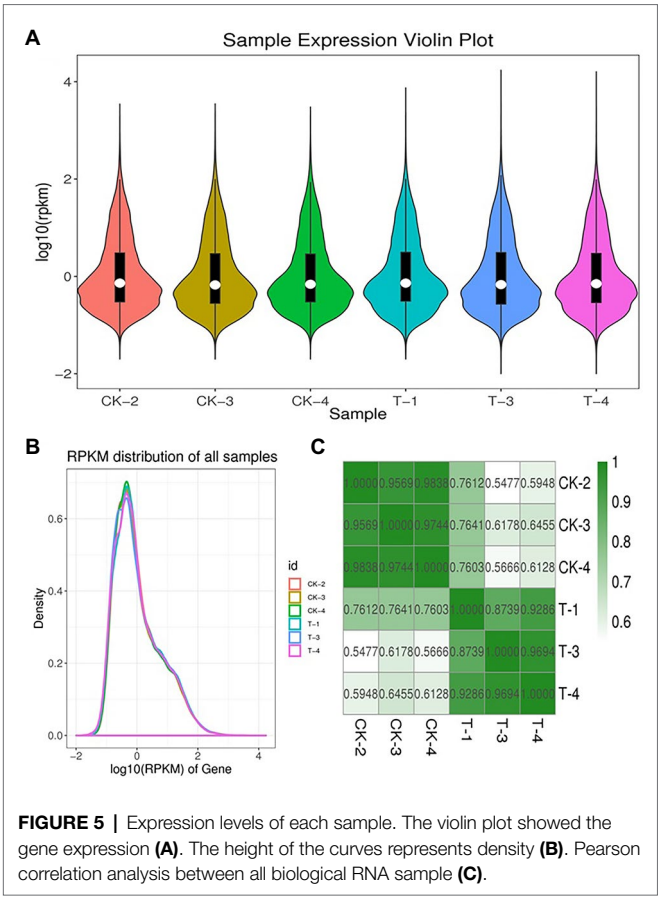


FIGURE 5 | Expression levels of each sample. The violin plot showed the gene expression **(A)**. The height of the curves represents density **(B)**. Pearson correlation analysis between all biological RNA sample **(C)**.

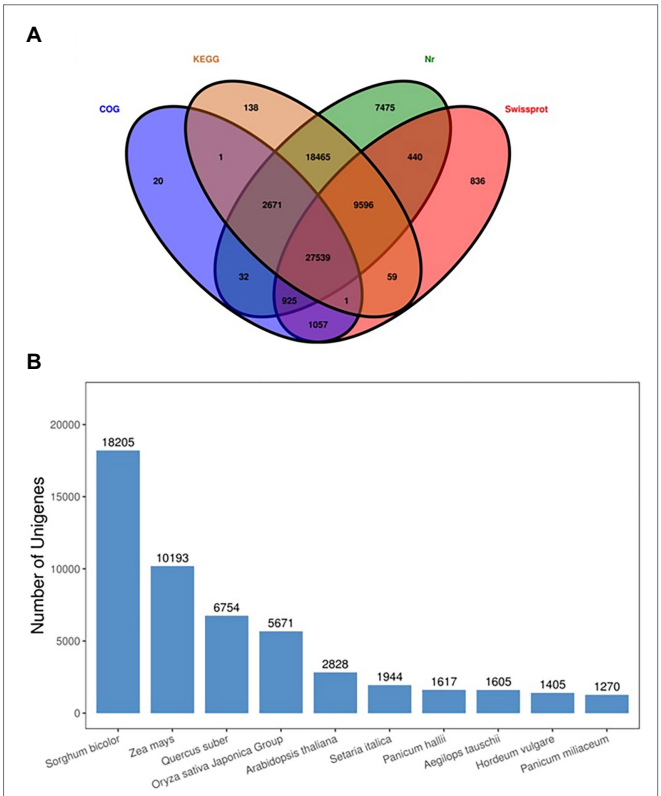


FIGURE 6 | Venn diagram annotated of the four major databases **(A)**, species distribution of the top BLAST hits against the NR database **(B)**.

on metabolism, genetic information processing, environmental information processing, cellular processes, and organismal systems pathway (**Supplementary Figure S4**). A total of 7,323 (49%) unigenes were annotated into "Metabolic pathways" of Global and overview maps, "translation" of genetic information processing, "signal transduction" of environmental information processing, "transport and catabolism" of cellular processes, and "Environmental adaptation" of organismal system. Besides, we compared unigenes with the KOG classification, and 38,945 unigenes were found to be aligned to the database and distributed into 25 categories (**Supplementary Figure S5**). Among them, the KOG category "general function prediction only" represented the largest group, followed by "Signal transduction mechanisms" and "posttranslational modification, protein turnover, chaperones."

Functional Enrichment of DGEs (CK vs. T)

The DEGs were define by using value of $p < 0.05$ or $|\log2FC| \geq 1$. Among the total 2,580 DEGs, 494 upregulated and 2,086 downregulated unigenes were identified between CK and T (**Figures 7A,B**). Moreover, the hierarchical cluster (H-cluster) analysis of all DEGs was performed to show the expression pattern of DEGs in different samples (**Supplementary Figure S6**). GO, and KEGG analyses were performed to determine biological processes and functions enriched in DEGs in response to IBA (CK vs. T). Three major Go categories, biological process, cellular component, and molecular function classification, revealed that 52 pathways (24 biological processes, 17 cellular component, and 11 molecular functions) interact in CK vs. T (**Figure 8A**). In response to IBA, major influenced pathways in the biological process category are metabolic process, cellular process, and single-organism process. For the cellular component category, cell, cell part, organelle, membrane, and organelle part are mainly affected. In addition, catalytic activity and binding were primarily affected in the molecular function category (**Figure 8A**). **Figure 8B** showed the up and downregulation of top 20 GO IDs, and biological process genes influence more than molecular and cellular components. **Supplementary Table S7** contains the list of gene functions and their *p*-values that revealed a significant enrichment of biological processes and molecular functions. Moreover, the KEGG pathway enrichment analysis of DEGs revealed that Global and overview maps, Carbohydrate metabolism, Translation Folding, sorting and degradation, Signal transduction, and Environmental adaptation were the most affected categories (**Figure 9A**), and the top 20 enriched pathways represented in **Figure 9B**, results showed the up and downregulation of genes in the pathways. In case of CK vs. T, KEGG pathways raveled significant enrichment in ko04626 Plant–pathogen interaction ($p < 0.001$), ko04016 MAPK signaling pathway—plant ($p < 0.001$, ko04075 Plant hormone signal transduction ($p < 0.001$), ko00564 Glycerophospholipid metabolism ($p < 0.01$), ko00908 Zeatin biosynthesis ($p < 0.01$, ko00940 Phenylpropanoid biosynthesis

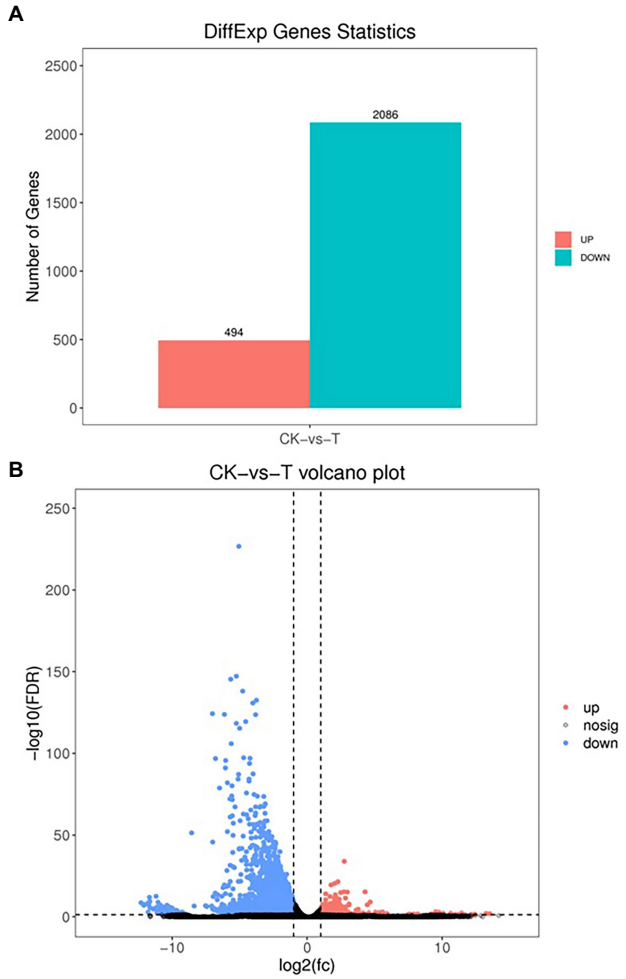


FIGURE 7 | Differential expression analysis of RNA-seq **(A)**, the volcano indicates the differentially expressed genes (DEGs; **B**). Each dot in the figure signifies a particular DEG. The red dot shows upregulated DEGs. The blue dot indicates downregulated DEGs, and the dark grey dot is a non-significant differential gene.

($p < 0.01$), ko01110 Biosynthesis of secondary metabolites ($p < 0.05$), ko00904 Diterpenoid biosynthesis ($p < 0.05$) and ko00941 Flavonoid biosynthesis ($p < 0.05$; **Figure 9C**; **Supplementary Figure S6**).

DEGs Enriched in Plant Hormone Signal Transduction Pathway

GO annotation results provide the information about the IBA treatment to enrich the essential biological process (**Supplementary Table S8**). GO results showed 101 genes were expressed in GO term GO:0009755 hormone-mediated signaling pathway. Next, we selected a few GO pathways to determine the response of IBA on plant hormone-related genes (**Supplementary Table S9**). In the cytokinin metabolic process, 6 genes were significantly ($p = 0.004$) expressed in the CK vs. T, and 3 (IPT5, CKX4, and LOGL1) were upregulated, and 3 (CKX5, LOG, and LOGL10) were downregulated. The two genes ACC1 and ACS7 of the 1-aminocyclopropane-1-carboxylate metabolic process ($p = 0.005$) were downregulated in the CK vs. T. For the response to salicylic acid genes, 30 genes were significantly ($p = 0.005$) enriched. Among them, 6 genes (AATP1, PGIP1, PGIP1, DIR1, ALD1, and NAC079) were upregulated, and 24 were downregulated. A total of 14 genes were expressed in the ethylene metabolic process ($p = 0.06$); among them, only ALD1 gene was upregulated. For the jasmonic acid metabolic process, total 24 genes were significantly ($p = 0.006$) enriched; among them, 20 genes were downregulated, while only 4 genes (At3g11180, At3g11180, SSL5, and NAC079) were upregulated. However, three genes of L-phenylalanine metabolic process were non-significantly downregulated in the biological process (**Figure 10**; **Supplementary Table S9**).

DEGs Enriched in Other Functions of Plant

In CK vs. T, the DEGs were categorized into more than 50 different terms that were dominated by biological processes pathways (**Supplementary Table S9**). In primary biological process related GO term includes different pathways (**Figure 11**). The maximum significant genes were found with GO:0050896 response to stimulus (596), and then followed by GO:0065007 biological regulation (477), GO:0050789 regulation of biological process (433), GO:0007154 cell communication (228), GO:0009719 response to endogenous stimulus (202), GO:0007165 signal transduction (199), GO:0023052 signaling (199), GO:0044700 single-organism signaling (199), GO:0001101 response to acid chemical (186), GO:1901698 response to nitrogen compound 69, and GO:0010243 response to organonitrogen compound (65; **Figure 11**; **Supplementary Table S10**). In CK vs. T, molecular function pathways classified higher with GO:0016740 transferase activity (307), GO:0016772 transferase activity, transferring phosphorus-containing groups (184), GO:0016301 kinase activity (141), GO:0016773 phosphotransferase activity, alcohol group as acceptor (120), and GO:0004672 protein kinase activity (117; **Supplementary Figure S7**). In cellular component related GO term includes GO:0030054 cell junction (125), GO:0005911 cell–cell junction (124), GO:0071944 cell periphery (96), GO:0030312 external encapsulating structure (85), and GO:0005618 cell wall (34; **Supplementary Figure S8**).

In GO term biological process, we selected a few unigenes to see their differential pattern compared to the control (**Figure 12**; **Supplementary Table S11**). For GO:0010243 response to organonitrogen compound, 65 genes were expressed significantly ($p < 0.0001$), among them, 58 were downregulated, while 7 genes (ERF1B, RLK5, CHIA (*Sorghum bicolor*), HAR1, PUB25, PTI5, and CHIA) were upregulated. In the case of GO:0071229 cellular response to acid chemical, 46 genes were downregulated and 6 genes (AATP1, PGIP1, PGIP1, DIR1, ALD1, and NAC079) were upregulated. A total of 68 genes were significantly ($p = 0.045$) enriched with the GO:0071554 cell wall organization or biogenesis. Among them, 23 genes were upregulated, and 45 genes were downregulated. For GO:0009845 seed germination, 12 differentially expressed genes were found, and among them, only one gene RR9 was upregulated (**Figure 12**; **Supplementary Table S11**).

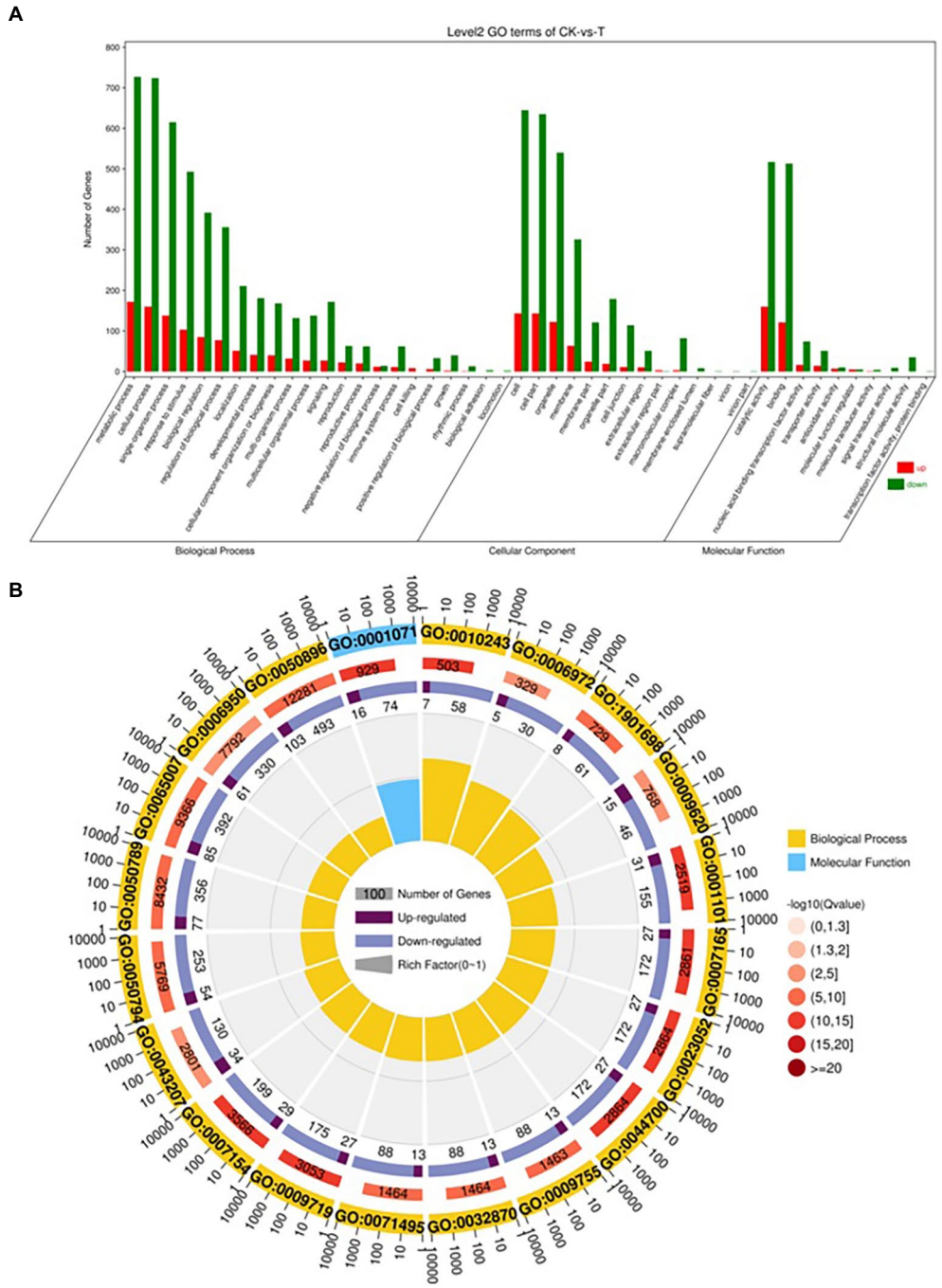
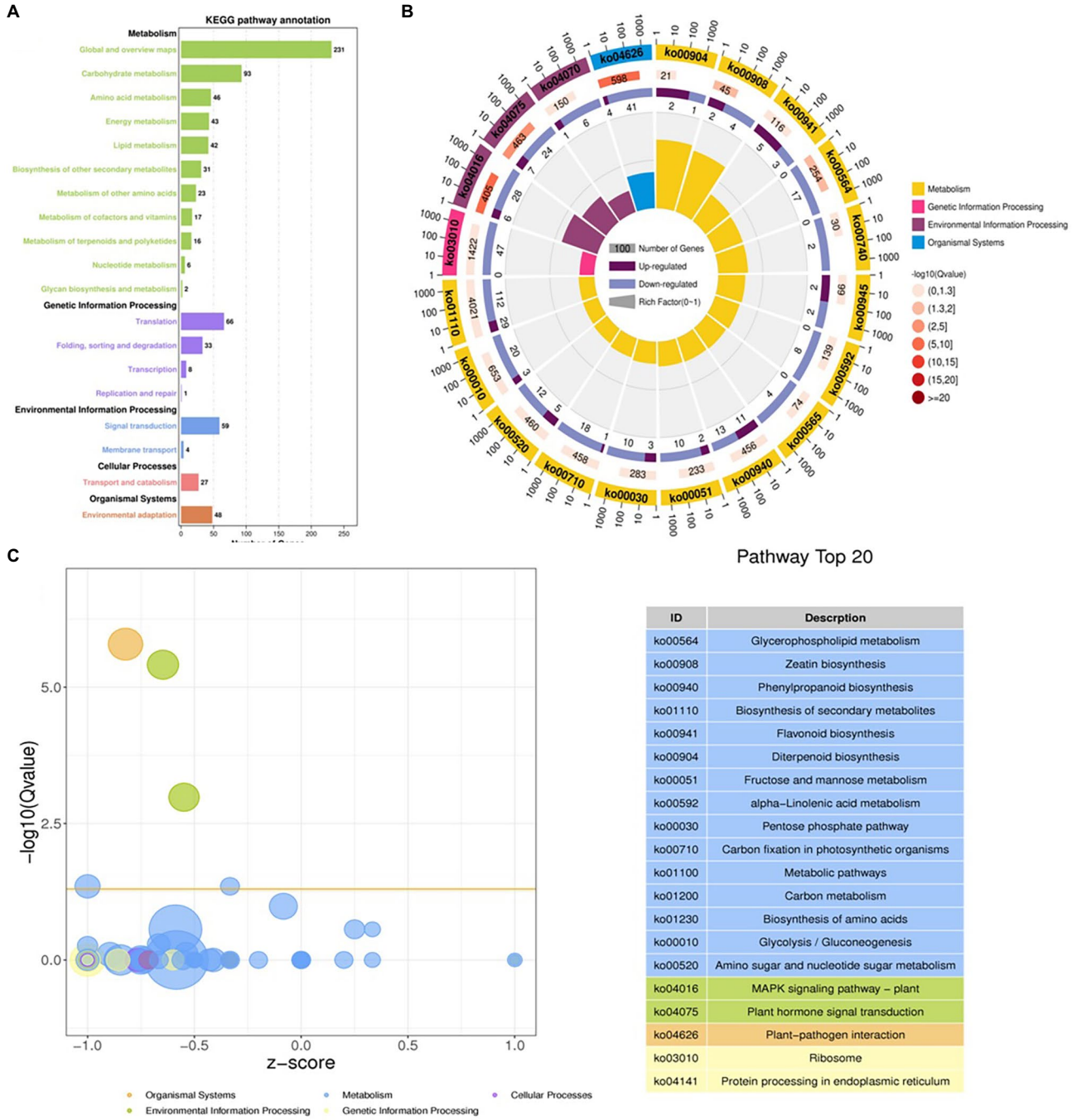


FIGURE 8 | Gene classification was based on Gene Ontology (GO) analysis for differentially expressed genes (DEGs). Different classes are shown for biological processes, cellular components, and molecular functions **(A)**. GO Rich Circle Map **(B)**: (First lap: The pathway of the top 20 riches, the coordinates of the number of genes outside the circle.) Different colors represent different classes; Second lap: the number of pathways in the background gene and the Q value. The longer the gene, the smaller the Q value, the redder the color; the third circle: the upper reduction gene scale bar chart, dark purple represents the increase of the gene ratio, the light purple represents the reduction of the gene ratio; the specific values are shown below; the fourth circle: the number of differential genes in each pathway divided by all quantities), the background grid line, each grid represents 0.1).

Validation of Gene Expression Using qRT-PCR

Next, to gene function, eight differentially expressed unigenes were selected to test our RNA-seq results with real-time RT-PCR. These unigenes are related to plant hormone signal transduction, organonitrogen metabolism as well as other enzymatic processes (**Table 2**). Based on RNA-seq results, Unigene0054023 was showed higher log fold change as compared to other genes.



ID	Descrption
ko00564	Glycerophospholipid metabolism
ko00908	Zeatin biosynthesis
ko00940	Phenylpropanoid biosynthesis
ko01110	Biosynthesis of secondary metabolites
ko00941	Flavonoid biosynthesis
ko00904	Diterpenoid biosynthesis
ko00051	Fructose and mannose metabolism
ko00592	alpha–Linolenic acid metabolism
ko00030	Pentose phosphate pathway
ko00710	Carbon fixation in photosynthetic organisms
ko01100	Metabolic pathways
ko01200	Carbon metabolism
ko01230	Biosynthesis of amino acids
ko00010	Glycolysis / Gluconeogenesis
ko00520	Amino sugar and nucleotide sugar metabolism
ko04016	MAPK signaling pathway – plant
ko04075	Plant hormone signal transduction
ko04626	Plant–pathogen interaction
ko03010	Ribosome
ko04141	Protein processing in endoplasmic reticulum

FIGURE 9 | Distributions of the KEGG pathways. Putative proteins were mapped to the reference canonical pathways in the KEGG database. The bar chart shows the number of sequences in different pathway categories **(A)**, KEGG Rich Circle Map **(B)**, Bubble plot for visualizing top 20 KEGG Annotation **(C)**. The z-score predicts existence of a bias in gene regulation.

The qRT-PCR results indicated that most of the genes were upregulated, but significant ($p < 0.05$) upregulation resulted only with 3 genes [Unigene0063754 Auxin-responsive protein SAUR32 (*Zea mays*), Unigene0066169 Cytokinin oxidase 2 (*Saccharum officinarum*), and Unigene0111066 two-component response regulator ORR9 (*Sorghum bicolor*)] as compared to control (**Figure 13**). Others genes ERF1B, MKK9, IPT5, PTI5, and GA20ox1B were showed a positive trend similar to RNA-seq analysis (**Figure 13**; **Table 2**).

DISCUSSION

Bud sprouting is a critical phenomenon of the sugarcane plant that allows the development of adventitious roots (AR) during

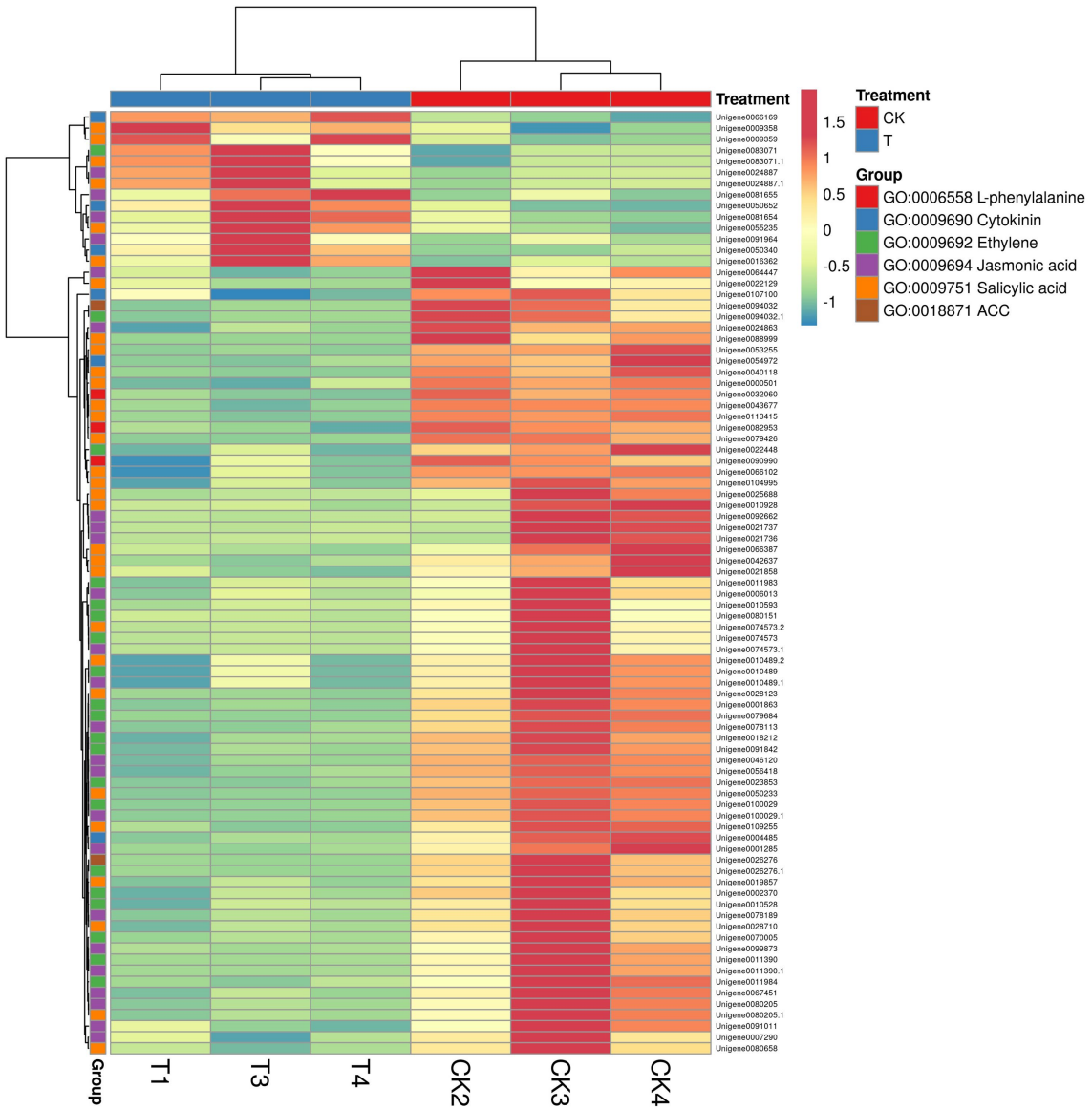


FIGURE 10 | The heat map representation of genes involved in phytohormone metabolism. The expression values are RNA-seq FPKM values.

rooting. Phytohormones like IBA and Auxin are the major drivers in AR developments that regulate biological, cellular, and molecular functions (Vidoz et al., 2010; An et al., 2020; Khadr et al., 2020; Li et al., 2020). During plant rooting, IAA played a significant role at the stage of root primordia formation (Li et al., 2020; Xiong et al., 2022). Next to Auxin, IBA is a well-adopted plant growth regulator (PGR) that is frequently used for the clonal propagation of different crops. It is a need to understand the response of PGR like IBA to sugarcane bud sprouting to induce AR during *in vitro* propagation. On the other hand, inadequate cellular and molecular mechanisms information are accessible to answer how IBA regulates the AR of the bud of a long perennial grass-like sugarcane. In the present study, IBA-treated (50, 100, 200 ppm) sugarcane buds responded to a significant ($p < 0.05$) growth promotion in plant physiological parameters like root length, plant height, stem diameter with the 100 ppm concentration. Subsequently, three plant growth hormones like IAA, ABA, and GA_3 accumulated significantly ($p < 0.05$) higher in 100 ppm concentration than control. These results are supported by several previous reports that discussed the efficacy of IBA in inducing the plant growth parameters, such as length, root weight, and leaves diameter, in different plants like strawberry and carrot plants (Hossain and Gony, 2020; Khadr et al., 2020). Auxin influence many plant growth regulating functions during developmental processes, such as cell division, elongation, and differentiation (Swarup et al., 2002; Yang et al., 2021; Xiong et al., 2022). Likewise, IBA is also known as a growth booster for the plant that regulates plant functions regards to roots, such as elongation, root hair development, lateral root formation,

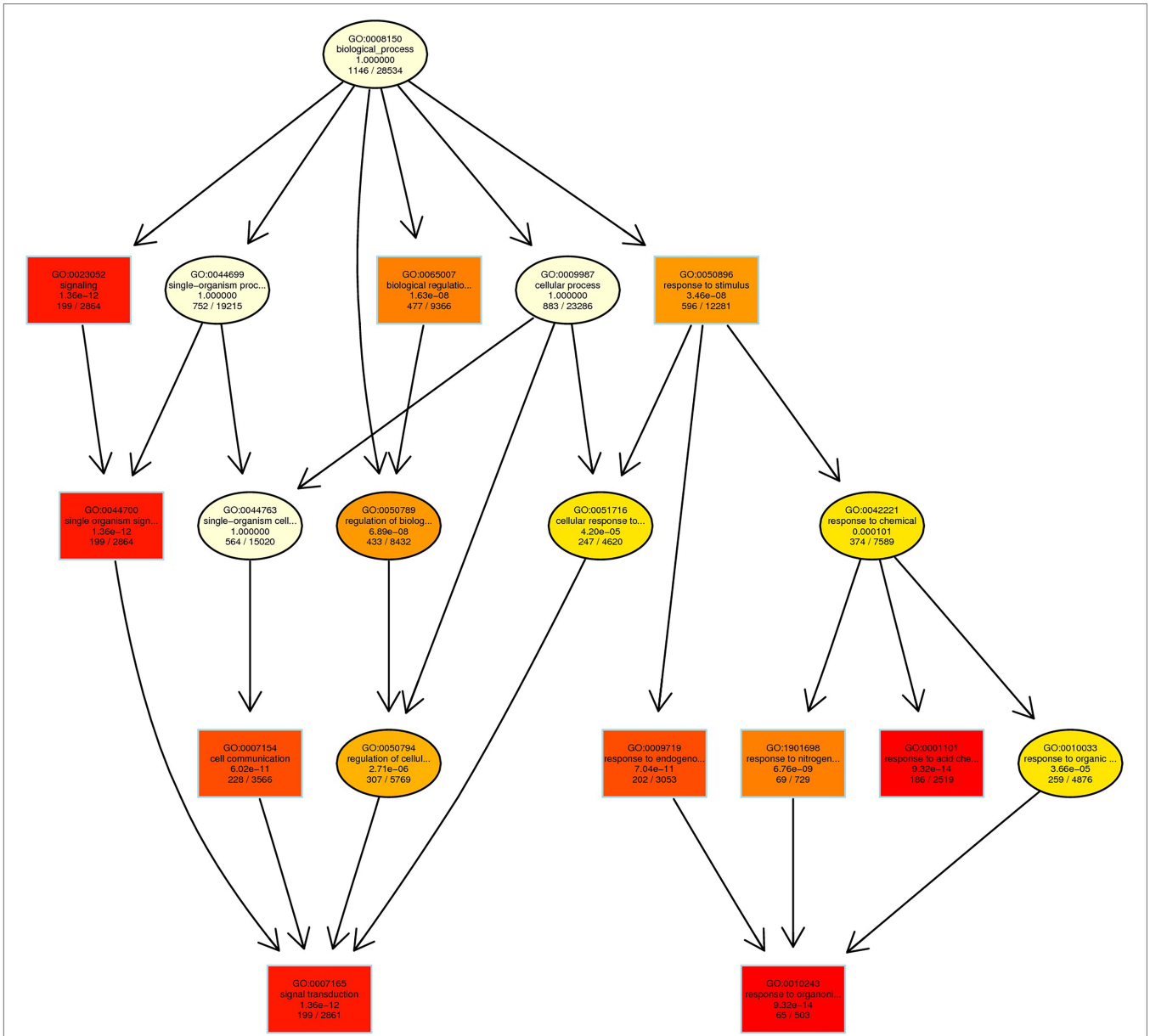


FIGURE 11 | Biological process-based enriched GO term in CK vs. T. Red to yellow color exhibiting the high to low enrichment DEGs in each GO term.

secondary metabolism, and nutrient transport (Muday et al., 2012; Damodaran and Stra, 2019; Wu et al., 2021). Recently, Meng et al. (2019) reported IBA treatment significantly enhanced the plant hormones at 3 days compared to control in apple rootstocks. Phytohormone-related genes, particularly those associated with auxin, ethylene, cytokinin, jasmonic acid, and salicylic acid, are vital pillars of plant growth promotion, especially in regulating adventitious AR formation (Wei et al., 2019; Li et al., 2020).

In addition to physiological data about hormones and plant growth promotion, we further analyzed the data at the molecular level by transcriptomic analysis experiment by using two kinds of samples, 100 ppm IBA-treated buds (T) and control buds (CK). In recent times, modern biotechnological, bioinformatics tools and big-data platforms have helped understand the functions of genes and gene families up to depth (Hoang et al., 2017; Abdullah et al., 2021; Faraji et al., 2021; Heidari et al., 2021; Manechini et al., 2021). In this study, RNA-seq results revealed that The IBA treatment significantly influenced DEGs regulation. An average of 66.758 million clean reads were assembled into 113,475 unigenes, and the total length of the unigenes was 94.94 Mb (94,938,961 bases). The GC percentage of the assembled transcripts was 50.64%. The N50 length was 1,536. Our results are aligned with the few previous research that generates the 59.4 mb unigenes length from basal tissue of sugarcane micro shoots treated with auxin (Li et al., 2020). In the case of CK

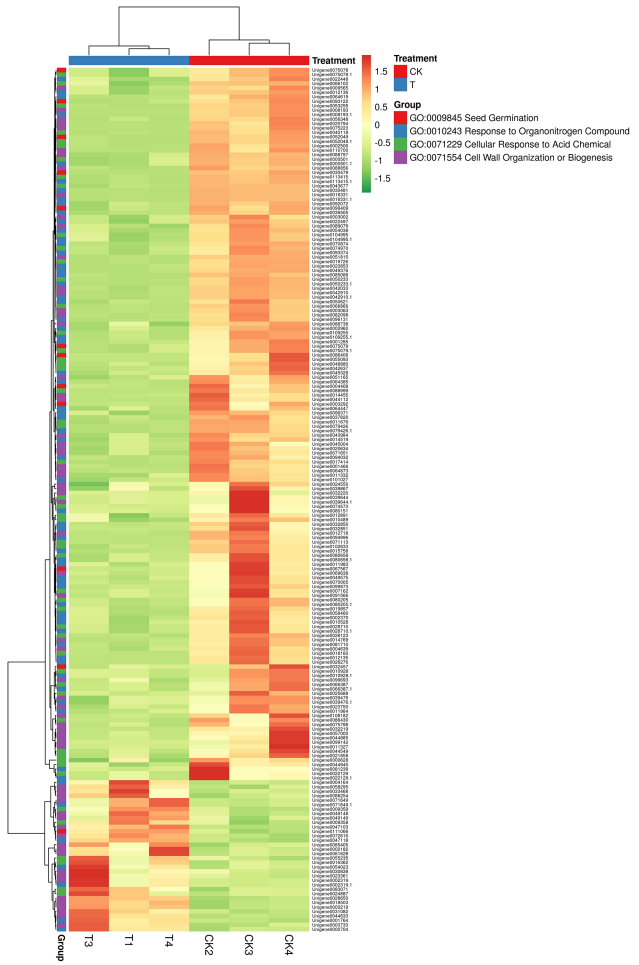


FIGURE 12 | The heat map representation of genes involved in different biological functions. The expression values are RNA-seq FPKM values.

vs. T, 86% unigenes were assembled, which is longer than reported previously in studies using the same technology (Kim et al., 2012; Yue et al., 2016; Quan et al., 2017; Li et al., 2020). Thus, the present sequencing depth contains high-quality bases for further investigation. To our knowledge, this study attempts the first time to investigate the sugarcane bud transcriptome dataset in response to IBA treatment. Additionally, the types and quantities of the genes expressed in CK and IBA-treated single-bud cuttings, along with their functions, classifications, and metabolic pathways regulating sugarcane bud sprouting. A total of 2,580 DEGs were identified, and this number was higher than similar previous studies that discussed the transcriptome of sugarcane tissues (Wei et al., 2014; Li et al., 2016a, 2020). Furthermore, *in silico* analysis showed that among 2,580 DEGs, 494 DEGs were upregulated, and 2,086 DEGs were downregulated with the cutoff FDR value < 0.05 and $|\log2FC| > 1$. The response IBA in plants is not dominated by a single factor but by a balance of biosynthesis, metabolism, and function through various signaling pathways. Our results showed that IBA influenced the biological process dominantly, and metabolic process, cellular process, and single-organism process had a significant influence of IBA. In the case of hormone-mediated signaling pathway and IBA, upregulation of genes IPT5, CKX4, and LOGL1 and downregulated genes CKX5, LOG, and LOGL10 in the cytokinin metabolic process. Several previous studies discussed these genes in plant development in different plants (Kurakawa et al., 2007; Ramirez-Carvajal et al., 2009; Böttcher et al., 2015; Wang et al., 2020; Shibasaki et al., 2021), and Cytokinin has been known to be a negative regulator of adventitious rooting *via* its negative regulation of auxin (Ramirez-Carvajal et al., 2009; Li et al., 2016b). The two downregulated genes, ACC1 and ACS7 of the 1-aminocyclopropane-1-carboxylate metabolic process, need to understand more in the case of sugarcane. These two genes played a significant role in ethylene biosynthesis.

TABLE 2 | RNA-seq results of selected genes for qRT-PCR validation.

Gene ID	Symbol	Description	CK_mean_ rpkm	T_mean_ rpkm	log2(fc)	p-value	FDR
Unigene0000704	ERF1B	Ethylene-responsive transcription factor 1B (*Sorghum bicolor*)	3.99	10.91	1.45	0.0000	0.0014
Unigene0014585	MKK9	Mitogen-activated protein kinase kinase 9 (*Sorghum bicolor*)	21.96	45.63	1.06	0.0008	0.0154
Unigene0050652	IPT5	Adenylate isopentenyltransferase 5, chloroplastic (*Sorghum bicolor*)	8.31	17.82	1.10	0.0001	0.0037
Unigene0054023	PTI5	AP2-EREBP transcription factor, partial (*Zea mays*)	3.71	16.62	2.17	0.0029	0.0467
Unigene0063754	SAUR32	Auxin-responsive protein SAUR32 (*Zea mays*)	9.95	21.61	1.12	0.0000	0.0000
Unigene0066169	CKX4	Cytokinin oxidase 2 (*Saccharum officinarum*)	1.72	5.43	1.66	0.0000	0.0000
Unigene0097937	GA20ox1B	Gibberellin 20 oxidase 1 (*Saccharum* hybrid cultivar ROC22)	87.22	192.26	1.14	0.0000	0.0000
Unigene0111066	RR9	Two-component response regulator ORR9 (*Sorghum bicolor*)	10.58	23.57	1.15	0.0000	0.0001

Rpkm, reads per kilobase of transcript.

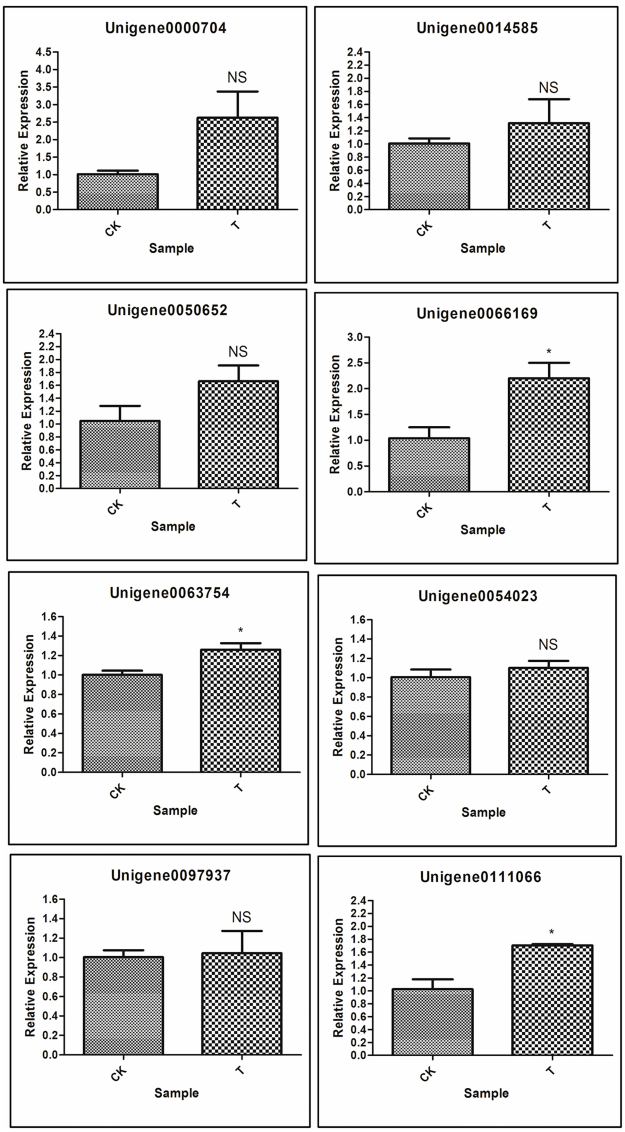


FIGURE 13 | Validation of selected genes involved in rooting by qRT-PCR; Significant ($p < 0.05$) differences between CK and treatment are indicated by an asterisk.

Twenty-four genes were downregulated for the salicylic acid pathway, while 6 genes like AATP1, PGIP1, PGIP1, DIR1, ALD1, and NAC079 were upregulated. Upregulated genes may play a significant role in defense against stress-inducing factors. However, downregulation of a higher number of genes may limit the defense responses that may enhance the plant growth (Karasov et al., 2017; Smith, 2019), although it is not the absolute truth (Campos et al., 2016). A gene ALD1 was upregulated in the ethylene metabolic process among the 14 genes. These results indicated that more genes downstream of metabolic and regulatory networks might be activated under long-term stress, which would likely eventually affect sugarcane growth (Cunha et al., 2017). Similarly, another stress relating pathway (jasmonic acid) had 20 downregulated genes were found, while only four genes (At3g11180, At3g11180, SSL5, and NAC079) were upregulated. Jasmonic acid pathways played a decisive role in the defense management of plants against biotic and abiotic stresses (Antico et al., 2012; Du et al., 2013; Sanches et al., 2017; Kim et al., 2021). As per the experiment, we used a healthy sugarcane bud and healthy soil with proper moisture for the bud sprouting that limited the stress-inducing factors. The present study provides helpful information by the DEGs results of different stress regulating plant hormone pathways. It is known that sugarcane can tolerate early phase stress deficits without significantly affecting future yields (da Silva et al., 2013).

Next, to plant hormones, different biological processes were examined. Higher downregulated DEGs of GO:0010243 and ERF1B, RLK5, CHIA (*Sorghum bicolor*), HAR1, PUB25, PTI5 upregulated DEGs. These DEGs may play a significant role in the nitrogen metabolism and secondary metabolism of the plant. Zhang et al. (2020) reported the ERF1B gene played a considerable role in root development. Jinn et al. (2000) concluded that RLK5 played a significant function in floral organ abscission. CHIA and PTI5 genes are involved in plant defense to biotic and abiotic stresses (Gu et al., 2002; Su et al., 2014), HAR1 gene in nodulation (Okuma et al., 2020), and PUB25 gene in root meristem development (Wendrich et al., 2017). Next, GO:0071229 cellular response to acid chemical, AATP1, PGIP1, PGIP1, DIR1, ALD1, and NAC079 genes were upregulated, and their functions are cell development and plant defense (Song et al., 2004; Champigny et al., 2013; Yang et al., 2014; Kalunke et al., 2015; Conrath et al., 2017). A total of 23 genes were upregulated in cell wall organization or biogenesis, and one gene RR9 was upregulated with the seed germination pathway. Dissimilarity in the expression levels of specific genes encoding the cell wall components involved in cell wall modifications and the maintenance of cell wall integrity, as well as cytoskeleton-related proteins were detected in T and CK, and, perhaps, some of them were related to adventitious root induction in bud sprouting. Upregulated genes of the cell wall are associated with the cell developmental and environmental signals, including lateral root formation and root development (Villacorta-Martín et al., 2015; Pizarro and Díaz-Sala, 2022). We further validate the RNA-seq data of eight DEGs through the qRT-PCR (**Figure 11**; **Table 2**). qRT-PCR validation concluded that RNA-seq analysis data showed more accurate expression than the qRT-PCR, and based on these conclusions, we can avoid the use of qRT-PCR for gene quantification. However, systematic quantification by using two different methods could help to optimize information for the readers and new researchers.

CONCLUSION

The comprehensive study shows that use of IBA on the single-bud seeds of sugarcane holds a lot of potential in sugarcane industry and showing the changes at the molecular level. The transcriptomic data generated in the present study

for sugarcane provides insights into mechanisms of IBA-induced adventitious rooting and opens up economic avenues for agriculturists and farmers. Future works should characterize the functional role of these identified individual DEGs and their regulatory networks. The expression patterns of genes encoding phytohormones, cell wall development, and other function in response to IBA in sugarcane bud sprouting were concluded by the RNA-seq data. Identified DEGs and their functions may be associated with plant growth stimulation and specific modifications of the cell wall. Multiple gene profiling by cutting-edge technologies discover competitive information about bud sprouting and adventitious root formation, and these molecular maps will help to understand the expressional signature of different DEGs under different kinds of stress regulating factors, and that may help to improve the crop production against the climate resilience. Moreover, a depth study will help to stabilize the networking and functions of DEGs associated with IBA treatment and plant growth. Future activities to better understand the root growth stimulation of sugarcane under different stress like saline, drought, root inhibiting biotic pathogen, and heavy metal treatments might spotlight essential genes functions to compare with relative plant species and link with the plant evolutionary biology. Big-data information would significantly contribute for the long perennial grasses to conduct experiments and compare with the database that is not present abundantly for the grass crops like sugarcane.

DATA AVAILABILITY STATEMENT

The raw data have been deposited to the NCBI, Sequence Read Archive (SRA) database with accession number PRJNA766098. The test links are as follows: https://dataview.ncbi.nlm.nih.gov/object/PRJNA766098?reviewer=cpfmjtoasijgvpa9dq0q3ur0td.

AUTHOR CONTRIBUTIONS

LX, Z-ND, K-CW, W-ZW, and H-RH conceived and designed the experiments. LX, H-RH, MM, MS, and KV drafted the manuscript. TP, Y-JL, and X-YL contributed to reagents, materials, and analysis tools. MS, BK, and ED analyzed of the sequence data and performed the statistical analysis. All authors contributed to the article and approved the submitted version.

FUNDING

This research was supported by the National key research and development program of China (2020YFD1000600), Government Booting Local Program (ZY20158007), Guangxi Natural Science Foundation (2020GXNSFBA159024), National Natural Science Foundation of China (32060468), and Guangxi Academy of Agricultural Sciences, Nanning, Guangxi, China, for providing the necessary facilities for this study.

ACKNOWLEDGMENTS

National key research and development program of China, Guangxi Natural Science Foundation, National Natural Science Foundation of China, and Guangxi Academy of Agricultural Sciences, Nanning, Guangxi, China, for providing the necessary funding and facilities for this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.852886/full#supplementary-material

REFERENCES

Abdullah, Faraji, S., Heidari, P., and Poczai, P. (2021). The *BAHD* gene family in cacao (*Theobroma cacao*, Malvaceae): genome-wide identification and expression analysis. *Front. Ecol. Evol.* 9:707708. doi: 10.3389/fevo.2021.707708

An, H., Zhang, J., Xu, F., Jiang, S., and Zhang, X. (2020). Transcriptomic profiling and discovery of key genes involved in adventitious root formation from green cuttings of highbush blueberry (*Vacciniumcorymbosum* L.). *BMC Plant Biol.* 20:182. doi: 10.1186/s12870-020-02398-0

Antico, C. J., Colon, C., Banks, T., and Ramonell, K. M. (2012). Insights into the role of jasmonic acid-mediated defenses against necrotrophic and biotrophic fungal pathogens. *Front. Biol.* 7, 48–56. doi: 10.1007/s11515-011-1171-1

Ashburner, M., Ball, C. A., Blake, J. A., Botstein, D., Butler, H., Cherry, J. M., et al. (2000). Gene ontology: tool for the unification of biology. The gene ontology consortium. *Nat. Genet.* 25, 25–29. doi: 10.1038/75556

Böttcher, C., Burbidge, C. A., Boss, P. K., and Davies, C. (2015). Changes in transcription of cytokinin metabolism and signaling genes in grape (*Vitis vinifera* L.) berries are associated with the ripening-related increase in isopentenyl adenine. *BMC Plant Biol.* 15:223. doi: 10.1186/s12870-015-0611-5

Brinker, M., Van Zyl, L., Liu, W., Craig, D., Sederoff, R. R., Clapham, D. H., et al. (2004). Microarray analyses of gene expression during adventitious root development in *Pinus contorta*. *Plant Physiol.* 135, 1526–1539. doi: 10.1104/pp.103.032235

Campos, M. L., Yoshida, Y., Major, I. T., de Oliveira Ferreira, D., Weraduwage, S. M., Froehlich, J. E., et al. (2016). Rewiring of jasmonate and phytochrome B signaling uncouples plant growth-defense tradeoffs. *Nat. Commun.* 7:12570. doi: 10.1038/ncomms12570

Champigny, M. J., Isaacs, M., Carella, P., Faubert, J., Fobert, P. R., and Cameron, R. K. (2013). Long distance movement of DIR1 and investigation of the role of DIR1-like during systemic acquired resistance in *Arabidopsis*. *Front. Plant Sci.* 4:230. doi: 10.3389/fpls.2013.00230

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34, i884–i890. doi: 10.1093/bioinformatics/bty560

Chen, D., Zhou, W., Yang, J., Ao, J., Huang, Y., Shen, D., et al. (2021). Effects of seaweed extracts on the growth, physiological activity, cane yield and sucrose content of sugarcane in China. *Front. Plant Sci.* 12:659130. doi: 10.3389/fpls.2021.659130

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

Conrath, U., Linke, C., Jeblick, W., Geigenberger, P., Quick, W. P., and Neuhaus, H. E. (2017). Enhanced resistance to *Phytophthora infestans and Alternaria solani* in leaves and tubers, respectively, of potato plants with decreased activity of the plastidic ATP/ADP transporter. *Planta* 217, 75–83. doi: 10.1007/s00425-003-0974-y

Cunha, C. P., Roberto, G. G., Vicentini, R., Lembke, C. G., Souza, G. M., Ribeiro, R. V., et al. (2017). Ethylene-induced transcriptional and hormonal responses at the onset of sugarcane ripening. *Sci. Rep.* 7:43364. doi: 10.1038/srep43364

da Silva, V., da Silva, B. B., Albuquerque, W. G., Borges, C. J. R., de Sousa, I. F., and Neto, J. D. (2013). Crop coefficient, water requirements, yield and water use efficiency of sugarcane growth in Brazil. *Agric. Water Manag.* 128, 102–109. doi: 10.1016/j.agwat.2013.06.007

Damodaran, S., and Stra, L. C. (2019). Indole 3-butyric acid metabolism and transport in *Arabidopsi thaliana*. *Front. Plant Sci.* 10:851. doi: 10.3389/fpls.2019.00851

Dharshini, S., Hoang, N. V., Mahadevaiah, C., Sarath Padmanabhan, T. S., Alagarasan, G., Suresha, G. S., et al. (2019). Root transcriptome analysis of *Saccharum spontaneum* uncovers key genes and pathways in response to low-temperature stress. *Environ. Experi. Bot.* 117:103935. doi: 10.1016/j.envexpbot.2019.103935

Du, H., Liu, H., and Xiong, L. (2013). Endogenous auxin and jasmonic acid levels are differentially modulated by abiotic stresses in rice. *Front. Plant Sci.* 4:397. doi: 10.3389/fpls.2013.00397

Faraji, S., Mehmood, F., Malik, H. M. T., Ahmed, I., Heidari, P., and Poczai, P. (2021). The GASA gene family in cacao (*Theobroma cacao*, Malvaceae): genome wide identification and expression analysis. *Agronomy* 11:1425. doi: 10.3390/agronomy11071425

Fattorini, L., Veloccia, A., Della Rovere, F., D'angeli, S., Falasca, G., and Altamura, M. M. (2017). Indole-3-butyric acid promotes adventitious rooting in *Arabidopsis thaliana* thin cell layers by conversion into indole-3-acetic acid and stimulation of anthranilate synthase activity. *BMC Plant Biol.* 17:121. doi: 10.1186/s12870-017-1071-x

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Gu, Y. Q., Wildermuth, M. C., Chakravarthy, S., Loh, Y. T., Yang, C., He, X., et al. (2002). Tomato transcription factors pti4, pti5, and pti6 activate defense responses when expressed in *Arabidopsis*. *Plant Cell* 14, 817–831. doi: 10.1105/tpc.000794

Heidari, P., Sahar, F., and Peter, P. (2021). Magnesium transporter gene family: genome-wide identification and characterization in *Theobroma cacao*, *Corchorus capsularis* and *Gossypium hirsutum* of family Malvaceae. *Agronomy* 11:1651. doi: 10.3390/agronomy11081651

Hoang, N. V., Furtado, A., Mason, P. J., Marquardt, A., Kasirajan, L., Thirugnanasambandam, P. P., et al. (2017). A survey of the complex transcriptome from the highly polyploid sugarcane genome using full-length isoform sequencing and de novo assembly from short read sequencing. *BMC Genomics* 18:395. doi: 10.1186/s12864-017-3757-8

Hossain, M. M., and Gony, O. (2020). Influence of indole butyric acid on root induction in daughter plants of strawberry. *J. Appl. Hortic.* 22, 209–214. doi: 10.37855/jah.2020.v22i03.37

Jing, H., and Strader, L. C. (2019). Interplay of auxin and cytokinin in lateral root development. *Int. J. Mol. Sci.* 20:486. doi: 10.3390/ijms20030486

Jinn, T. L., Stone, J. M., and Walker, J. C. (2000). HAESA, an *Arabidopsis leucine*-rich repeat receptor kinase, controls floral organ abscission. *Genes Dev.* 14, 108–117. doi: 10.1101/gad.14.1.108

Kalunke, R. M., Tundo, S., Benedetti, M., Cervone, F., De Lorenzo, G., and D'Ovidio, R. (2015). An update on polygalacturonase-inhibiting protein (PGIP), a leucine-rich repeat protein that protects crop plants against pathogens. *Front. Plant Sci.* 6:146. doi: 10.3389/fpls.2015.00146

Karasov, T. L., Chae, E., Herman, J. J., and Bergelson, J. (2017). Mechanisms to mitigate the trade-off between growth and defense. *Plant Cell* 29, 666–680. doi: 10.1105/tpc.16.00931

Kasirajan, L., Hoang, N. V., Furtado, A., Botha, F. C., and Henry, R. J. (2018). Transcriptome analysis highlights key differentially expressed genes involved in cellulose and lignin biosynthesis of sugarcane genotypes varying in fiber content. *Sci. Rep.* 8:11612. doi: 10.1038/s41598-018-30033-4

Khadr, A., Wang, G. L., Wang, Y. H., Zhang, R. R., Wang, X. R., Xu, Z. S., et al. (2020). Effects of auxin (indole-3-butyric acid) on growth characteristics, lignification, and expression profiles of genes involved in lignin biosynthesis in carrot taproot. *PeerJ* 8:e10492. doi: 10.7717/peerj.10492

Khaire, K. C., Moholkar, V. S., and Goyal, A. (2021). Bioconversion of sugarcane tops to bioethanol and other value added products: an overview. *Mater. Sci. Energy Technol.* 4, 54–68. doi: 10.1016/J.MSET.2020.12.004

Kim, H., Seomun, S., Yoon, Y., and Jang, G. (2021). Jasmonic acid in plant abiotic stress tolerance and interaction with abscisic acid. *Agronomy* 11:1886. doi: 10.3390/agronomy11091886

Kim, S. B., Yu, J. G., Lee, G. H., and Park, Y. D. (2012). Characterization of Brassica rapa sadenosyl-L-methionine synthetase gene including its roles in biosynthesis pathway. *Hortic. Environ. Biotechnol.* 53, 57–65. doi: 10.1007/s13580-012-0084-5

Kurakawa, T., Ueda, N., Maekawa, M., Kobayashi, K., Kojima, M., Nagato, Y., et al. (2007). Direct control of shoot meristem activity by a cytokinin-activating enzyme. *Nature* 445, 652–655. doi: 10.1038/nature05504

Li, A., Lakshmanan, P., He, W., Tan, H., Liu, L., Liu, H., et al. (2020). Transcriptome profiling provides molecular insights into Auxin-induced adventitious root formation in sugarcane (*Saccharum* spp. interspecific hybrids) microshoots. *Plants* 9:931. doi: 10.3390/plants9080931

Li, C., Nong, Q., Solanki, M. K., Liang, Q., Xie, J., Liu, X., et al. (2016a). Differential expression profiles and pathways of genes in sugarcane leaf at elongation stage in response to drought stress. *Sci. Rep.* 6:25698. doi: 10.1038/SREP25698

Li, S. W., Shi, R. F., Leng, Y., and Zhou, Y. (2016b). Transcriptomic analysis reveals the gene expression profile that specifically responds to IBA during adventitious rooting in mung bean seedlings. *BMC Genomics* 17:43. doi: 10.1186/s12864-016-2372W-4

Libao, C., Runzhi, J., Jianjun, Y., Xiaoyong, X., Haitao, Z., and Shuyan, L. (2018). Erratum to: Transcriptome profiling reveals an IAA-regulated response to adventitious root formation in lotus seedling. *Z. Naturforsch. C* 73:409. doi: 10.1515/znc-2018-0133

Livak, K. J., and Schmittgen, T. D. (2001). Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta Ct}$ method. *Methods* 25, 402–408. doi: 10.1006/meth.2001.1262

Liu, X., Zhang, R., Ou, H., Gui, Y., Wei, J., Zhou, H., et al. (2018). Comprehensive transcriptome analysis reveals genes in response to water deficit in the leaves of *Saccharum narenga* (Nees ex Steud.). *BMC Plant Biol.* 18:250. doi: 10.1186/s12870-018-1428-9

Loganandhan, N., Gujja, B., Goud, V. V., and Natarajan, U. (2013). Sustainable sugarcane initiative (SSI): a methodology of 'more with less'. *Sugar Tech.* 15, 98–102. doi: 10.1007/s12355-012-0180-y

Lu, Y.-H., and Yang, Y.-Q. (2019). "Sugarcane biofuels production in China," in *Sugarcane Biofuels*. eds. M. T. Khan and I. A. Khan (Switzerland: Springer), 139–155.

Ludwig-Muller, J., Vertocnik, A., and Town, C. D. (2005). Analysis of indole-3-butyric acid-induced adventitious root formation on Arabidopsis stem segments. *J. Exp. Bot.* 56, 2095–2105. doi: 10.1093/jxb/eri208

Malviya, M. K., Li, C. N., Solanki, M. K., Singh, R. K., Htun, R., Singh, P., et al. (2020). Comparative analysis of sugarcane root transcriptome in response to the plant growth-promoting *Burkholderia anthina* MYSP113. *PLoS One* 15:e0231206. doi: 10.1371/journal.pone.0231206

Manechini, J. R. V., de Santos, P. H. S., Romanel, E., de Brito, M. S., Scarpari, M. S., Jackson, S., et al. (2021). Transcriptomic analysis of changes in gene expression during flowering induction in sugarcane under controlled photoperiodic conditions. *Front. Plant Sci.* 12:635784. doi: 10.3389/fpls.2021.635784

Meng, Y., Xing, L., Li, K., Wei, Y., Wang, H., Mao, J., et al. (2019). Genome-wide identification, characterization and expression analysis of novel long non-coding RNAs that mediate IBA-induced adventitious root formation in apple root stocks. *Plant Growth Regul.* 87, 287–302. doi: 10.1007/s10725-018-0470-9

Mohanty, M., and Nayak, P. (2021). Bud chip method of sugarcane planting: a review. *J. Pharm. Innov.* 10, 150–153. doi: 10.22271/tpi.2021.v10.i3c.5753

Morris, E., Amy, A., and Noé, H. (2017). The impact of falling sugar prices on growth and rural livelihoods. Inter-American Development Bank. Country Office in Belize. IV. Title. V. Series. IDB-TN-1237.

Muday, G. K., Rahman, A., and Binder, B. M. (2012). Auxin and ethylene: collaborators or competitors? *Trends Plant Sci.* 17, 181–195. doi: 10.1016/j.tplants.2012.02.001

Okuma, N., Soyano, T., Suzaki, T., and Kawaguchi, M. (2020). *MIR2111-5* locus and shoot-accumulated mature miR2111 systemically enhance nodulation depending on HAR1 in *Lotus japonicus*. *Nat. Commun.* 11:5192. doi: 10.1038/s41467-020-19037-9

Pacurar, D. I., Perrone, I., and Bellini, C. (2014). Auxin is a central player in the hormone cross-talks that control adventitious rooting. *Physiol. Plant.* 151, 83–96. doi: 10.1111/ppl.12171

Pizarro, A., and Díaz-Sala, C. (2022). Expression levels of genes encoding proteins involved in the cell wall–plasma membrane–cytoskeleton continuum are associated with the maturation-related adventitious rooting competence of pine stem cuttings. *Front. Plant Sci.* 12:783783. doi: 10.3389/fpls.2021.783783

Quan, J., Meng, S., Guo, E., Zhang, S., Zhao, Z., and Yang, X. (2017). De novo sequencing and comparative transcriptome analysis of adventitious root development induced by exogenous indole-3-butyric acid in cuttings of tetraploid black locust. *BMC Genomics* 18:179. doi: 10.1186/s12864-017-3554-4

Ramirez-Carvajal, G. A., Morse, A. M., Dervinis, C., and Davis, J. M. (2009). The cytokinin type-B response regulator PtRR13 is a negative regulator of adventitious root development in *Populus*. *Plant Physiol.* 150, 759–771. doi: 10.1104/pp.109.137505

Robinson, M. D., Mccarthy, D. J., and Smyth, G. K. (2010). edgeR: a bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Rout, G. R. (2006). Effect of auxins on adventitious root development from single node cuttings of Camellia sinensis (L.) Kuntze and associated biochemical changes. *Plant Growth Regul.* 48, 111–117.

Sanches, P., Santos, F., Peñaflor, M., and Bento, J. (2017). Direct and indirect resistance of sugarcane to *Diatraea saccharalis* induced by jasmonic acid. *Bull. Entomol. Res.* 107, 828–838. doi: 10.1017/S0007485317000372

Shibasaki, K., Takebayashi, A., Makita, N., Kojima, M., Takebayashi, Y., Kawai, M., et al. (2021). Nitrogen nutrition promotes rhizome bud outgrowth via regulation of cytokinin biosynthesis genes and an *Oryza longistaminata* ortholog of FINE CULM 1. *Front. Plant Sci.* 12:788. doi: 10.3389/fpls.2021.670101

Simao, F. A., Waterhouse, R. M., Ioannidis, P., Kriventseva, E. V., and Zdobnov, E. M. (2015). BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics* 31, 3210–3212. doi: 10.1093/bioinformatics/btv351

Singh, P., Singh, S. N., Tiwari, A. K., Pathak, S. K., Singh, A. K., Srivastava, S., et al. (2019). Integration of sugarcane production technologies for enhanced cane and sugar productivity targeting to increase farmers' income: strategies and prospects. *3 Biotech* 9:48. doi: 10.1007/S13205-019-1568-0

Smith, L. M. (2019). Salicylic acid, senescence, and Heterosis. *Plant Physiol.* 180, 3–4. doi: 10.1104/pp.19.00260

Solanki, M. K., Wang, F. Y., Wang, Z., Li, C. N., Lan, T. J., Singh, R. K., et al. (2019). Rhizospheric and endospheric diazotrophs mediated soil fertility intensification in sugarcane-legume intercropping systems. *J. Soils Sediments* 19, 1911–1927. doi: 10.1007/s11368-018-2156-3

Song, J. T., Lu, H., McDowell, J. M., and Greenberg, J. T. (2004). A key role for ALD1 in activation of local and systemic defenses in *Arabidopsis*. *Plant J.* 40, 200–212. doi: 10.1111/j.1365-313X.2004.02200.x

Su, Y., Xu, L., Fu, Z., Yang, Y., Guo, J., Wang, S., et al. (2014). ScChi, encoding an acidic class III chitinase of sugarcane, confers positive responses to biotic and abiotic stresses in sugarcane. *Int. J. Mol. Sci.* 15, 2738–2760. doi: 10.3390/ijms15022738

Swarup, R., Parry, G., Graham, N., Allen, T., and Bennett, M. (2002). Auxin cross-talk: integration of signaling pathways to control plant development. *Plant Mol. Biol.* 49, 411–426. doi: 10.1007/978-94-010-0377-3_12

Vidoz, M. L., Loreti, E., Mensuali, A., Alpi, A., and Perata, P. (2010). Hormonal interplay during adventitious root formation in flooded tomato plants. *The Plant J.* 63, 551–562. doi: 10.1111/j.1365-313X.2010.04262.x

Villacorta-Martín, C., Sánchez-García, A. B., Villanova, J., Cano, A., van de Rhee, M., de Haan, J., et al. (2015). Gene expression profiling during adventitious root formation in carnation stem cuttings. *BMC Genomics* 16:789. doi: 10.1186/s12864-015-2003-5

Wang, J. G., Zhao, T. T., Wang, W. Z., Feng, C. L., Feng, X. Y., Xiong, G. R., et al. (2019). Culm transcriptome sequencing of Badila (*Saccharum officinarum* L.) and analysis of major genes involved in sucrose accumulation. *Plant Physiol. Biochem.* 144, 455–465. doi: 10.1016/j.plaphy.2019.10.016

Wang, X., Ding, J., Lin, S., Liu, D., Gu, T., Wu, H., et al. (2020). Evolution and roles of cytokinin genes in angiosperms 2: do ancient CKX s play housekeeping roles while non-ancient CKX s play regulatory roles? *Horti. Res.* 7, 1–15. doi: 10.1038/s41438-020-0246-z

Wei, K., Ruan, L., Wang, L., and Cheng, H. (2019). Auxin-induced adventitious root formation in nodal cuttings of *Camellia sinensis*. *Int. J. Mol. Sci.* 20:4817. doi: 10.3390/ijms20194817

Wei, K., Wang, L. Y., Wu, L. Y., Zhang, C. C., Li, H. L., Tan, L. Q., et al. (2014). Transcriptome analysis of indole-3-butyric acid-induced adventitious root formation in nodal cuttings of *Camellia sinensis* (L.). *PLoS One* 9:e107201. doi: 10.1371/journal.pone.0107201

Wendrich, J. R., Möller, B. K., Li, S., Saiga, S., Sozzani, R., Benfey, P. N., et al. (2017). Framework for gradual progression of cell ontogeny in the *Arabidopsis* root meristem. *Proc. Natl. Acad. Sci. U.S.A.* 114, E8922–E8929. doi: 10.1073/pnas.1707400114

Wu, K., Duan, X., Zhu, Z., Sang, Z., Zhang, Y., Li, H., et al. (2021). Transcriptomic analysis reveals the positive role of abscisic acid in endodormancy maintenance of leaf buds of *Magnolia wufengensis*. *Front. Plant Sci.* 12:742504. doi: 10.3389/fpls.2021.742504

Xiong, Y., Chen, S., Wei, Z., Chen, X., Guo, B., Zhang, T., et al. (2022). Transcriptomic analyses provide insight into adventitious root formation of Euryodendron excelsum HT Chang during ex vitro rooting. *Plant Cell Tissue Organ Cult.* 148, 649–666. doi: 10.1007/s11240-021-02226-9

Yang, Z. T., Wang, M. J., Sun, L., Lu, S. J., Bi, D. L., Sun, L., et al. (2014). The membrane-associated transcription factor NAC089 controls ER-stress-induced programmed cell death in plants. *PLoS Genet.* 10:e1004243. doi: 10.1371/journal.pgen.1004243

Yang, L., You, J., Li, J., Wang, Y., and Chan, Z. (2021). Melatonin promotes *Arabidopsis* primary root growth in an IAA dependent manner. *J. Exp. Bot.* 72, 5599–5611. doi: 10.1093/jxb/erab196

Yang, J. C., Zhang, J. H., Wang, Z. Q., Zhu, Q. S., and Wang, W. (2001). Hormonal changes in the grains of rice subjected to water stress during grain filling. *Plant Physiol.* 127, 315–323.

Yue, R., Lu, C., Qi, J., Han, X., Yan, S., Guo, S., et al. (2016). Transcriptome analysis of cadmium-treated roots in maize (*Zea mays* L.). *Front. Plant Sci.* 7:1298. doi: 10.3389/fpls.2016.01298

Zeng, Q., Ling, Q., Fan, L., Li, Y., Hu, F., Chen, J., et al. (2015). Transcriptome profiling of sugarcane roots in response to low potassium stress. *PLoS One* 10:e0126306. doi: 10.1371/journal.pone.0126306

Zhang, M., and Govindaraju, M. (2018). "Sugarcane production in China," in *Sugarcane - Technology and Research*. ed. A. B. de Oliveira.

Zhang, J., Liu, J., Jiang, C., Nan, T. G., Kang, L. P., Zhou, L., et al. (2020). Expression analysis of transcription factor ERF gene family of *Panax ginseng*. *Zhongguo Zhong Yao Za Zhi* 45, 2515–2522. doi: 10.19540/j.cnki.cjcmm.20200329.101

Zhu, K., Yang, L. T., Li, C.-H., Lakshmanan, P., Xing, Y. X., Xing, Y. X., et al. (2021). A transcriptomic analysis of sugarcane response to *Leifsonia xyli* subsp. *xyli* infection. *PLoS One* 16:e0245613. doi: 10.1371/journal.pone.0245613

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 10 June 2022
doi: 10.3389/fpls.2022.924490

Integrated Transcriptomic and Proteomic Analyses Uncover the Regulatory Mechanisms of *Myricaria laxiflora* Under Flooding Stress

Linbao Li[1,2], Guiyun Huang[1,2], Weibo Xiang[2], Haofei Zhu[1,2], Haibo Zhang[1,2], Jun Zhang[1,2], Zehong Ding[3], Jihong Liu[4] and Di Wu[1,2]**

[1] Rare Plants Research Institute of Yangtze River, China Three Gorges Corporation, Yichang, China, [2] National Engineering Research Center of Eco-Environment Protection for Yangtze River Economic Belt, Beijing, China, [3] Sanya Research Institute of Chinese Academy of Tropical Agricultural Sciences, Sanya, China, [4] College of Horticulture and Forestry, Huazhong Agricultural University, Wuhan, China

Flooding is one of the major environmental stresses that severely influence plant survival and development. However, the regulatory mechanisms underlying flooding stress remain largely unknown in *Myricaria laxiflora*, an endangered plant mainly distributed in the flood zone of the Yangtze River, China. In this work, transcriptome and proteome were performed in parallel in roots of *M. laxiflora* during nine time-points under the flooding and post-flooding recovery treatments. Overall, highly dynamic and stage-specific expression profiles of genes/proteins were observed during flooding and post-flooding recovery treatment. Genes related to auxin, cell wall, calcium signaling, and MAP kinase signaling were greatly down-regulated exclusively at the transcriptomic level during the early stages of flooding. Glycolysis and major CHO metabolism genes, which were regulated at the transcriptomic and/or proteomic levels with low expression correlations, mainly functioned during the late stages of flooding. Genes involved in reactive oxygen species (ROS) scavenging, mitochondrial metabolism, and development were also regulated exclusively at the transcriptomic level, but their expression levels were highly up-regulated upon post-flooding recovery. Moreover, the comprehensive expression profiles of genes/proteins related to redox, hormones, and transcriptional factors were also investigated. Finally, the regulatory networks of *M. laxiflora* in response to flooding and post-flooding recovery were discussed. The findings deepen our understanding of the molecular mechanisms of flooding stress and shed light on the genes and pathways for the preservation of *M. laxiflora* and other endangered plants in the flood zone.

Keywords: ***Myricaria laxiflora*, post-flooding recovery, transcriptome and proteome, flooding stress, transcriptional and post-transcriptional regulation**

OPEN ACCESS

Edited by:
Freddy Mora-Poblete, University of Talca, Chile

Reviewed by:
Krishan K. Verma, Guangxi Academy of Agricultural Science, China
Muhammad Ahmad Hassan, Anhui Agricultural University, China

***Correspondence:**
Di Wu
763394826@qq.com
Jihong Liu
liujihong@mail.hzau.edu.cn

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

Received: *20 April 2022*
Accepted: *23 May 2022*
Published: *10 June 2022*

Citation:
Li L, Huang G, Xiang W, Zhu H, Zhang H, Zhang J, Ding Z, Liu J and Wu D (2022) Integrated Transcriptomic and Proteomic Analyses Uncover the Regulatory Mechanisms of Myricaria laxiflora Under Flooding Stress. Front. Plant Sci. 13:924490. doi: 10.3389/fpls.2022.924490

INTRODUCTION

Flooding (including submergence and waterlogging) is a major problem that seriously influences plant growth and development worldwide, leading to significant loss of yield and even death of the plant (necrosis) (Jia et al., 2021). The frequency of flooding is increasing on natural and agricultural species by the building of dams and roads and divergence of rivers (Dat et al., 2004; Boulange et al., 2021). Under flooding environments, the access of plants to light

and oxygen is drastically limited, accompanied by outbursts of reactive oxygen species (ROS). Meantime, the plant growth is suppressed, and the energy/nutrient supply is markedly inhibited (Yin and Komatsu, 2017).

Hormones have been demonstrated to play significant roles in flooding stress (Jia et al., 2021). Among them, ethylene is a primary signaling molecule and it is accumulated in plants to adapt to flooding stress (Kuroha et al., 2018). Abscisic acid (ABA), auxin, gibberellic acid (GA), cytokinin, and salicylic acid (SA) are other hormones that prevent plants against flooding stress by regulating adventitious root formation or by controlling carbohydrate consumption (Agullo-Anton et al., 2014; Du et al., 2015; Jia et al., 2021). Moreover, ethylene can interact with a hormonal cascade of ABA, GA, and auxin to promote adventitious root growth upon flooding in rice, bitterzoet, and tomato (Steffens et al., 2006; Vidoz et al., 2010; Dawood et al., 2016), supporting a complex regulation of hormones during flooding stress.

In addition to hormones, ROS plays vital roles in plant responses to flooding stress (Sasidharan et al., 2018). As an oxidative stress indicator, hydrogen peroxide is required for ethylene-induced epidermal impairments to promote aerenchyma formation under flooding conditions (Steffens and Sauter, 2009; Steffens et al., 2011). The activities of several ROS-scavenging enzymes, including catalase (CAT), peroxidase (POD), glutathione reductase (GR), and superoxide dismutase (SOD), are increased during the flooding stress (Ye et al., 2015; Wang et al., 2019). Meantime, ROS accumulation triggers the expression of downstream fermentation genes required for hypoxia (insufficient oxygen supply) acclimation and survival (Baxter-Burrell et al., 2002).

Transcription factors are also involved in the flooding responses and interact with hormones. For instance, *AtMYB2* and *AtbZIP50* are both induced by exogenous ABA, and their expression levels are increased under hypoxia conditions (Klok et al., 2002; Hsu et al., 2011; De Ollas et al., 2021). Two ethylene-responsive factors, *ERF1* and *ERF2*, are induced during the post-flooding recovery stage (Tsai et al., 2014; Yin et al., 2019). Overexpression of a NAC transcription factor (*SHYG*) enhances flooding-induced leaf movement and cell expansion in abaxial cells of the basal petiole region (Rauf et al., 2013). Two WRKY transcription factors, *WRKY33* and *WRKY12*, interact with each other to up-regulate *RAP2.2* for *Arabidopsis* adaptation to submergence stress (a state that the plant partially or fully immerses in water) (Tang et al., 2021). Meantime, WRKY-mediated transcriptional regulation triggers plant immunity upon flooding (Hsu et al., 2013).

Over the past decades, significant progress has been achieved regarding the plant response to flooding stress, and thousands of genes and many signaling pathways have been identified. A time-course transcriptomic analysis revealed that jasmonic acid (JA) participated in submergence-mediated internode elongation and improved flooding tolerance in rice (Minami et al., 2018). Meantime, the expression of genes referred to cell wall modification, trehalose biosynthesis, GA biosynthesis, and transcription factors was significantly changed (Minami et al., 2018). Analysis of enrichment pathways exhibited the involvement of antioxidant process, ROS signaling, photosynthesis, carbohydrate metabolism, stress response, hormone biosynthesis, and signal transduction in flooding stress (Wu et al., 2017; Wang et al., 2019; Qiao et al., 2020). Genes related to alcohol fermentation, ethylene biosynthesis, pathogen defense, and cell wall were altered at both transcriptional and proteomic levels, while ROS scavengers and chaperons were changed only at the translational level (Komatsu et al., 2009). Proteins involved in fermentation, ROS removal, glycolysis, and defense response were also affected under flooding treatments (Hashiguchi et al., 2009; Nanjo et al., 2012). These studies provided effective bases for exploring the mechanism of flooding responses at the transcriptomic and/or proteomic levels.

Myricaria laxiflora (Tamaricaceae) is an endangered plant mainly distributed in the flood zone of the Yangtze River, China (Li et al., 2021). In the native habitats, *M. laxiflora* is completely submerged during the long-lasting flooding periods in summer and rapidly sprouts in the autumn and winter after the flooding (Wu et al., 1998; Wang et al., 2003). Under summer flooding conditions, the primary and secondary branch number, aboveground biomass, and total biomass were greatly reduced in *M. laxiflora* seedlings (Chen and Xie, 2009). Meanwhile, the total soluble sugar, fructose, and starch contents also decreased in the branches and leaves (Zhou et al., 2021). Now summer flooding is considered as a major ecological process to influence the growth and development of *M. laxiflora*. In recent years, the construction of the Three Gorges Dam (TGD) significantly disrupted the *M. laxiflora* habitats, and only a few natural populations are remained (Bao et al., 2010). *M. laxiflora* has been regarded as an ideal plant to study the ecological restoration of Yangtze River (Chen and Xie, 2009; Chen et al., 2019). To date, many studies have been performed to investigate the influences of flooding on the seedling survival and physiology responses of *M. laxiflora* (Chen and Xie, 2009; Guan et al., 2020; Zhou et al., 2020); however, the molecular mechanisms underlying the flooding stress remain elusive.

In this work, we performed an integrated transcriptomic and proteomic analysis to investigate the dynamic changes in gene and protein expression of *M. laxiflora* during nine time-points under the flooding and post-flooding recovery treatments. These results shed light on the molecular mechanism of flooding stress in *M. laxiflora* and other endangered plants in the flood zone.

MATERIALS AND METHODS

Plant Material and Experiment Design

The cuttings of a *M. laxiflora* plant were collected from Yanzhiba (Yichang, China) to prepare sufficient amount of experimental materials with the same genetic background. These materials were cultivated in the experimental farm of the Rare Plants Research Institute of Yangtze River in March 2020.

For the flooding treatment, 1-year-old seedlings with uniform growth were transferred to an experimental pond located in the TGD. The water in the experimental pond is connected to the Yangtze River, as it is a good strategy to simulate the natural flooding on *M. laxiflora*. The daily temperature ranged from

19 to 27°C. The flooding treatment started at 11:00 AM. The root samples were collected at 0, 6, 12, 18, 24, 30, 36, and 48 h (designated as R0, R6, R12, R18, R24, R30, R36, and R48) after the flooding treatment, respectively. Then, the seedlings were transferred to the normal condition for 12 h recovery (RR12), at which the root samples were also collected. The samples were immediately frozen in liquid nitrogen and stored at −80°C until use.

Library Construction and Transcriptome Analysis

Total RNA extraction, library construction, and RNA-seq sequencing were performed by the Shanghai OE Biotech Company (Shanghai, China). Total RNA was isolated by the mirVana miRNA Isolation Kit (Ambion) according to the manufacturer's protocol. RNA integrity was examined by the Agilent 2100 Bioanalyzer (Agilent Technologies, CA, United States). The RNA-seq libraries were constructed by the TruSeq Stranded mRNA LTSample Prep Kit (Illumina, CA, United States) and then sequenced on the Illumina HiSeq 2500 platform in PE150 model. Each sample was performed in triple replicates.

The clean reads were obtained by Trimmomatic (Bolger et al., 2014) to remove adaptors and low-quality reads. The transcripts were *de novo* assembled by Trinity (Grabherr et al., 2011) in paired-end method. The longest transcript was selected as a unigene according to the similarity and length for further analysis. Gene expression was calculated by using the FPKM (Fragments Per Kilobase per Million mapped reads) method.

Proteomic Analysis

The samples used for RNA-seq sequencing were also subjected to proteome analysis by the Shanghai OE Biotech Company (Shanghai, China) using the data-independent acquisition (DIA) approach, which is regarded as a novel mass spectrometric method that shows higher precision and better reproducibility across replicates (Bruderer et al., 2015). Briefly, total protein was acquired from each sample, and the concentration was measured by the bicin-chonininc acid method. The quality was examined by SDS-PAGE with G250. After trypsin digestion, the peptides were desalted using a SOLA^{TM} SPE 96 Column and then eluted with 60% methanol three times.

The LC–MS/MS assays were performed as previously described (Wang et al., 2020). Briefly, the peptide mixture was loaded onto an Agilent ZORBAX Extend-C18 reversed-phase column (5 μm, 15 cm × 2.1 mm) on an 1100 HPLC System (Agilent, CA, United States). Buffer A (2% ACN) and buffer B (90% ACN) were applied for the reverse gradient. The solvent gradient was set as follows: 0–10 min, 2% B; 10–10.01 min, 2–5% B; 10.01–37 min, 5–20% B; 37–48 min, 20–40% B; 48–48.01 min, 40–90% B; 48.01–58 min, 90% B; 58–58.01 min, 90–2% B; 58.01–63 min, 2% B. The flow rate was 250 μl/min. In total, 10 fractions were collected and then dried by vacuum centrifugation. For DIA mass spectrum scanning, dry gas was operated at 3.0 L/min at 180°C and the acquisition of MS data was performed in the mass range of 100–1700 m/z. The ion mobility was 0.7–1.3, and the collision energy was 20–59 eV.

The LC–MS/MS raw files were searched against the protein database derived from our RNA-seq transcripts by Spectronaut Pulsar software. The false discovery rate of proteins and peptides was set to 0.01. For any set of samples, the proteins with expression value in more than 50% samples were kept, and the missing value was filled with the mean in the same group.

Transcriptomic and Proteomic Integrative Analysis

As previously described (Ding et al., 2020), the abundances of genes and proteins were min-max normalized between −1 and 1 among the time-points during flooding treatment. The expression patterns were identified by the standard procedure of WGCNA software (Langfelder and Horvath, 2008) and then visualized by the R package pheatmap. For further analysis, genes and proteins that were not assigned to any of the groups were excluded.

To interpret the biological functions influenced by flooding treatment, the genes were functionally annotated and classified into Gene Ontology (GO) categories by Blast2GO (Conesa et al., 2005), or assigned into hierarchical categories by the MapMan classification system (Thimm et al., 2004). The significantly enriched categories were determined by the Fisher's exact test (Ding et al., 2020).

Statistical Analysis

Differentially expressed genes (DEGs) were identified by DESeq2 (Love et al., 2014) setting the false discovery rate < 0.05 and $|\log_2 FC \text{ (fold-change)}| > 1$. Differentially expressed proteins (DEPs) were identified by the Student's *t*-test setting a fold-change > 2 (or <0.5) and *P*-value < 0.05. Each sample was performed in triple replicates. The correlations between the DEPs and their corresponding genes were calculated by the Pearson's correlation test (Ding et al., 2019).

RESULTS

Differentially Expressed Genes and Differentially Expressed Proteins Identification

Significantly altered phenotypes were observed in *M. laxiflora* as the flooding treatment time progressed: the color of the leaves was changed from bright green to yellow–green and small aerial roots began to grow. To investigate the dynamic expression changes of genes and proteins, root samples were collected at 0, 6, 12, 18, 24, 30, 36, and 48 h after flooding treatment and at 12 h after post-flooding recovery treatment, and subjected to transcriptome and proteome analyses, respectively.

Differentially expressed genes and DEPs were identified between a time point and R0 (0 h, the control) to signify the genes and proteins up-regulated or down-regulated for a certain period of flooding (**Figures 1A,B**). The number of DEGs and DEPs fluctuated slightly and then increased as the flooding stress time prolonged, although there was a time-shift (e.g., R24 vs. R36) between the transcriptomic and proteomic levels.

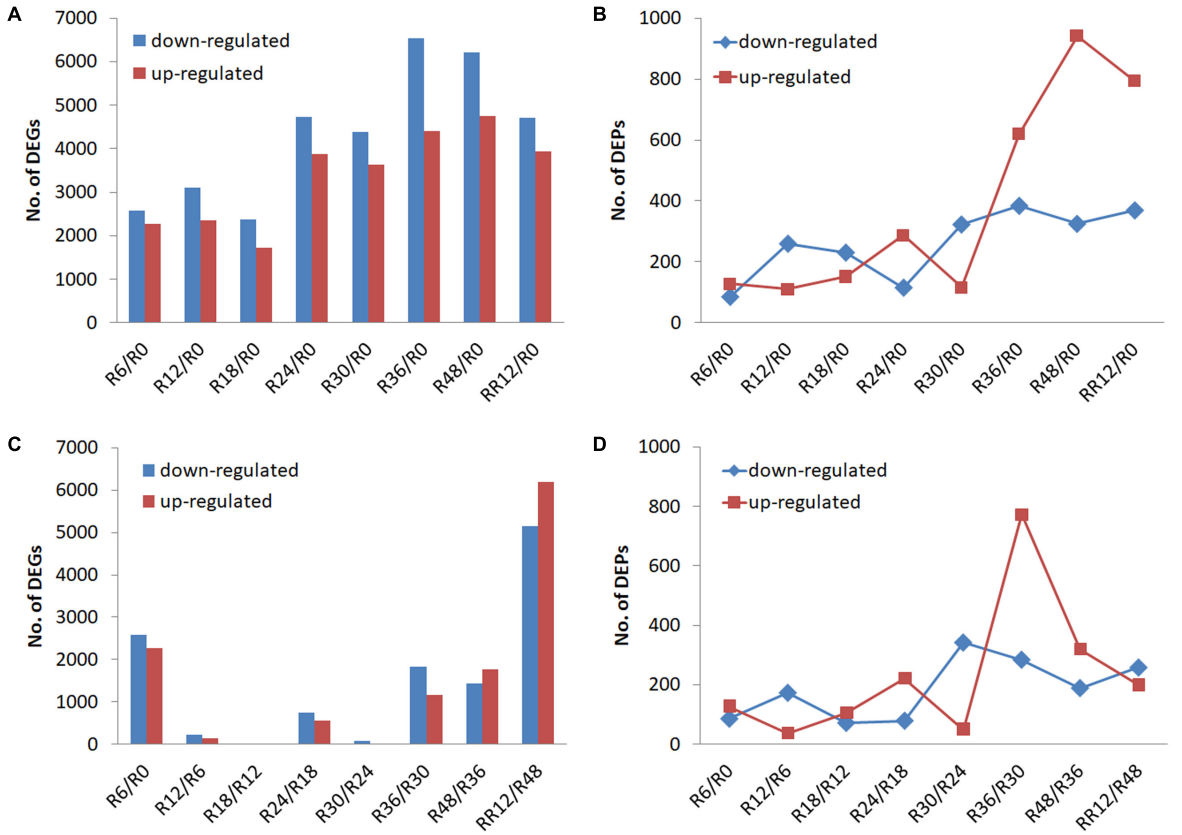


FIGURE 1 | Transcriptome and proteome profiling of *M. laxiflora* at different time-points during flooding and post-flooding recovery treatment. **(A,B)** Up-regulated and down-regulated DEGs and DEPs, respectively, identified for a certain period of flooding compared with the control (R0). **(C,D)** Up-regulated and down-regulated DEGs and DEPs, respectively, identified between a time point and the preceding time point. R0, R6, R12, R18, R24, R30, R36, and R48 represent the samples collected at 0, 6, 12, 18, 24, 30, 36, and 48 h after the flooding treatment, respectively, while RR12 indicates the samples collected at 12 h after post-flooding recovery treatment.

Differentially expressed genes and DEPs were also identified between a time point and the preceding time point to signify significant transitions in gene and protein expression during flooding stress. The highest number of DEGs was detected from R48 to RR12 (RR12/R48), followed by R6/R0, while less number of DEGs was found at R12/R6, R18/R12, and R30/R24 (**Figure 1C**). On the contrary, the highest number of up-regulated and down-regulated DEPs was detected at R36/R30 and R30/R24, respectively (**Figure 1D**).

Subsequently, these DEGs and DEPs were merged, and their co-expression patterns were detected at the transcriptomic and/or proteomic levels to dissect their over-represented functions, respectively.

Differentially Expressed Genes Exclusively Identified at the Transcriptome Level

In total, eight groups (gM1–gM8) of DEGs were determined at the transcriptome level but not at the proteome level (**Figure 2A** and **Supplementary Table 1**).

Genes from gM1 to gM3 showed the highest expression at R0 and their levels were dramatically depressed upon flooding treatment. The genes from gM1 were significantly enriched in auxin metabolism, cell wall, RNA regulation of transcription, secondary metabolism, calcium signaling, and MAP kinase signaling (**Figure 2D**). Many genes included in this group were involved in cellulose synthesis (*CESA6*, *CSLC4*, *CSLC6*, *CSLD3*, *CSLD5*), cell wall degradation (*XYL4*, *ASD1*, *BXL1*, *BXL2*), and lignin biosynthesis (*4CL2*, *4CL3*, *COMT1*, *CCOAMT*, *CCR1*, *CAD5*, *CAD9*). In addition, many auxin-responsive genes (*SAUR39*, *SAUR55*, *DFL2*, *TCP14*) and auxin efflux carriers (*PIN1*, *PIN4*, *PIN7*, *LAX3*) were also included. These results strongly suggested a coordinated contribution of auxin, cell wall, and lignin biosynthesis genes to growth inhibition under flooding treatment. Accordingly, several genes referred to calcium signaling (*CAM8*, *CBL10*, *CML37*, *CML41*, *CPK1*, *CPK13*), MAP kinase signaling (*MAPKKK5*, *MAPKKK13*, *MKK3–MKK7*, *MPK15*, *MPK16*), and abiotic stress (*DIL9*, *RD22*, and *ERD3* for drought; *HSFA2* and *HSP90.1* for heat) were identified, supporting a major role of these genes in rapid response to flooding stress. However, no enrichment was found for gM2 and gM3.

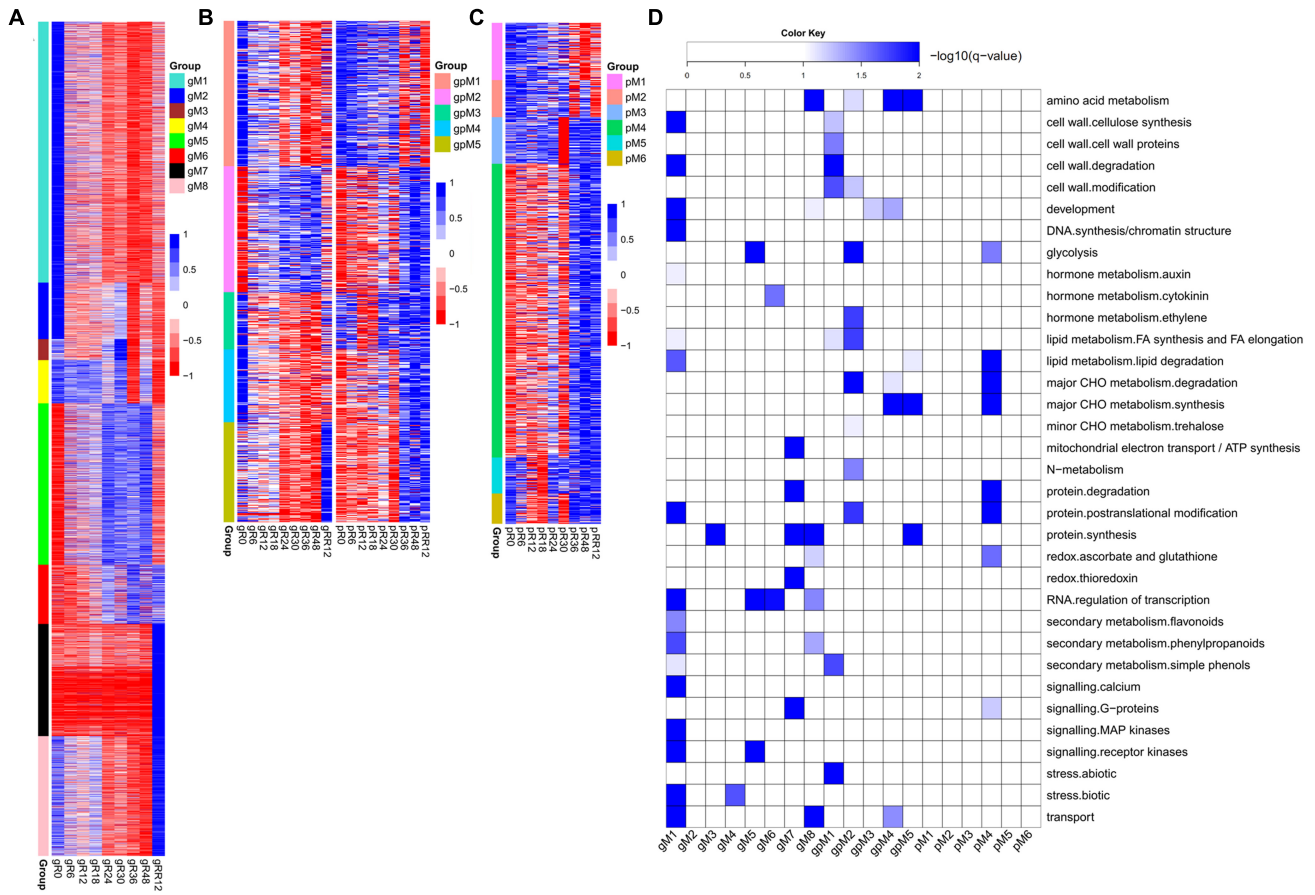


FIGURE 2 | Dynamic and stage-specific expression changes of genes and proteins upon flooding stress. **(A)** Profiles of DEGs identified exclusively at the transcriptomic level. A total of eight groups (gM1–gM8) were determined according to their expression patterns. **(B)** Profiles of DEGs/DEPs identified at both the transcriptomic and proteomic levels. Five groups (gpM1–gpM5) were found. **(C)** Profiles of DEPs identified exclusively at the proteomic level. Six groups (pM1–pM6) were found. **(D)** Functional category enrichment of DEGs/DEPs corresponding to those identified in **(A–C)**. The samples/groups are prefixed as follows: g, transcriptomic level; p, proteomic level; gp, both transcriptomic and proteomic levels.

Genes from gM4 maintained a high expression level from R0 to R18 and then fluctuated until RR12. These genes were significantly enriched in stress response, of which *CSC1*, *CTL1*, and *RBOHE* were involved in drought, heat, and defense responses.

Genes from gM5 to gM6 were highly expressed from R24 to R48, although their levels were quite different at RR12. The enriched categories of gM5 were glycolysis, RNA regulation of transcription, and receptor kinase signaling. Many glycolysis-related genes (*GAPDH1*, *ENO1*, *PFK2*, *PFK3*, *PGK*, *PGM2*, *TPI*) were found in this group. A few receptor kinases related to abiotic stress (*CRK29* and *RLK7*) and root hair elongation (*MRH1*) were also found. The genes from gM6 were enriched in cytokinin metabolism and RNA regulation of transcription. A cytokinin receptor (*HK3*) and two cytokinin synthesis-degradation genes (*IPT3* and *CKX1*) were included in this group. The results indicated a significant role of these genes in the late stages (e.g., R24–R48) of flooding stress.

Genes from gM7 to gM8 were expressed highest at RR12. The enriched categories of gM7 were mitochondrial electron transport/ATP synthesis, protein synthesis and degradation, thioredoxin (TRX) redox, and G-protein signaling. Several genes referred to mitochondrial electron transport/ATP synthesis (*ATP3*, *ATP5*, *COB*, *PHB1*, *PHB7*) and redox reaction (*CAT2*, *CSD2*, *GPX7*, *GR1*, *PER1*, *TRX*) were included in this group. The genes from gM8 were enriched in development, redox, RNA regulation of transcription, and secondary metabolism. Many genes involved in meristem development (*AP2* and *NAC2*), dormancy (*EMB2750* and *EMB2746*), plant growth (*CPL4*, *AGL8*, *EMF2*), and lignin biosynthesis (*C4H*, *4CL*, *COMT1*, *CCoAOMT1*, *CCR1*) were included. In addition, a few genes involved in ascorbate biosynthesis (*GME*) and ROS scavenging (*APX1*, *FSD3*, *CSD1*) were also found in this group. These results indicated a demand of energy for plant growth at the recovery stage after flooding.

Together, these findings revealed stage-specific functions of genes exclusively regulated at the transcriptome level. For example, auxin, cell wall, calcium signaling, and MAP kinase signaling were associated with the early stage of flooding stress; glycolysis, cytokinin, and RNA regulation of transcription

were related to the late stage; while mitochondrial electron transport/ATP synthesis, redox, and lignin biosynthesis were associated with the recovery stage after flooding.

Differentially Expressed Genes/Differentially Expressed Proteins Identified at the Transcriptome and Proteome Levels

Five groups of genes (gpM1–gpM5) were determined to be differentially expressed at both the transcriptome and proteome levels (**Figure 2B** and **Supplementary Table 1**).

The genes from gpM1 (R = 0.72, P = 0.028) to gpM2 (R = 0.57, P = 0.113) exhibited similar expression trends between the transcriptome and proteome levels. The expression levels of gpM1 were gradually deceased from R0 to RR12, and the genes of this group were significantly enriched in cell wall, FA synthesis and elongation, simple phenols, and abiotic stress (**Figure 2D**). Many cell wall related genes, including *CSLA2*, *IRX1*, and *IRX6* for cellulose synthesis, *FLA8* and *FLA17* for cell wall proteins, *GH9B5* and *BXL2* for cell wall degradation, and *EXLA1*, *EXLB1*, and *EXPA4* for cell wall modification were included. Similarly, two simple phenol genes (*LAC12* and *LAC17*) involved in oxidation reduction and lignin catabolic processes, and a few genes referred to salt (*OSM34*), cold (*RCI3* and *ESK1*), and heat response (*HSP18.2* and *HSP17.3*) were also included.

In contrast to gpM1, the expression of genes in gpM2 was gradually increased from R0 to RR12. These genes were significantly enriched in glycolysis, ethylene, major CHO metabolism, N-metabolism, and protein posttranslational modification (**Figure 2D**). Many genes related to glycolysis (*GAPDH3*, *PFK2*, *PFK3*, *PFP1*, *PFP2*, *PK1*, *PK2*) and major CHO degradation (*SEX1*, *HXK1*, *SUS3*, *SUS4*) were found in this group. In addition, a few genes referred to trehalose biosynthesis (*TPS6* and *TPS7*), N-metabolism (*GLT1* and *NIA1*), ethylene metabolism (*ERS1* and *ACO1*), and abiotic stress (*TAP46*, *SnRK2*, and *OST1*) were also included. These results strongly suggested a request for energy generation *via* carbohydrate degradation and glycolysis pathways to defend against the flooding stress.

The genes from gpM4 exhibited an opposite expression trend (R = −0.72, P = 0.029) between the transcriptome and proteome profiles. The enriched categories were amino acid (AA) metabolism, development, major CHO metabolism, and transport. Several genes involved in starch synthesis (*APL1*, *SBE2.2*, and *SS4*) and degradation (*GWD3* and *PHS1*), cell proliferation and growth (*ANT*, *LUG*, *RAPTOR1*), and AA metabolism (*FAH*, *ADT1*, *URE*, *PGDH2*) were found in this group. Interestingly, three transporters (*CUE1*, *PHT1;1*, *PDR2*) involved in phosphate transport were included.

However, no correlation was detected between the transcriptome and proteome levels for the genes from gpM3 to gpM5. The genes from gpM3 were enriched in development, and a few genes referred to cell division (*SWA1* and *ELO1*), apical meristem development (*NAC2*), and leaf senescence (*DET1* and *YLS8*) were included. The enriched categories of gpM5 were AA metabolism, lipid degradation, synthesis of major CHO metabolism, and protein synthesis. Several genes involved in AA biosynthesis (*MS1*, *HMT2*, *P5CS1*) and starch synthesis (*APL3*, *ISA1*, *GBSS1*) were found in this group.

Collectively, these results suggested a crucial contribution of transcriptional regulation (e.g., in cell wall and abiotic stress) and post-transcriptional regulation (e.g., in AA metabolism, development, and major CHO metabolism) upon flooding stress.

Differentially Expressed Proteins Exclusively Identified at the Proteome Level

Six groups (pM1–pM6) of DEPs were determined at the proteome level but not at the transcriptome level (**Figure 2C** and **Supplementary Table 1**).

The expression of proteins in pM4 was gradually increased from R0 to RR12. These proteins were enriched in glycolysis, major CHO metabolism, redox, and G-proteins signaling (**Figure 2D**). Many proteins referred to glycolysis (ENO3, PPC2, PPC3, PFK3, PK), major CHO synthesis (APL2, SBE2.2, SS1, SS3, SPS1F, SPS2F), and degradation (DPE1, DPE2, ISA3, SEX4, AMY3, BMY3, PHS2) were included. In addition, a few proteins related to redox (APX3, MDAR6, OXP1, GPX8, TO1, TY1) were also found in this group.

The proteins from pM1 to pM3 were highly expressed from R0 to R24, while those from pM5 to pM6 were highly expressed from R24 to RR12. However, no functional enrichment was significantly determined for these groups.

Together, these findings revealed stage-specific expression patterns of genes exclusively regulated at the protein level. The key genes and pathways determined by our integrated transcriptome and proteome analysis were summarized in **Figure 3**.

Response of Genes Involved in Sucrose-Starch Metabolism and Glycolysis

Diverse expression patterns were observed for the sucrose-starch metabolism genes at the transcriptome and/or proteome levels (**Figures 4A,B** and **Supplementary Table 2**).

Many genes involved in sucrose degradation (*INV1*, *FK1*) and starch degradation (*BMY1a*, *BMY1b*, *BMY8a*, *BMY8b*, *SEX4a*, *PHS2a*) were down-regulated at R6 by flooding at the transcriptome level. Several starch biosynthesis genes (including *APL1*, *APL3*, *SS4*, *SBE3*, *SBE4*, *GBSS1a*, *GBSS1b*) were also down-regulated, but they showed ∼6 h delay compared with those degradation genes, indicating a quicker response of starch degradation genes than the biosynthesis genes in response to flooding.

A few sucrose degradation genes were up-regulated at R6–R48 by flooding (*SUS3a*, *SUS3b*, *INV2*, *FK2*) or at RR12 by recovery treatment (*INV3*, *INV4*, *FK3*, *FK4*) exclusively at the transcriptome level. By comparison, many starch degradation genes (*AMY3b*, *BMY3*, *DPE1*, *DPE2*, *SEX4b*, *ISA3*, *PHS2b*) were up-regulated at R36–RR12 exclusively at the proteome level, suggesting different regulation mechanisms of sucrose-starch genes under flooding treatment. In addition, more than a dozen genes

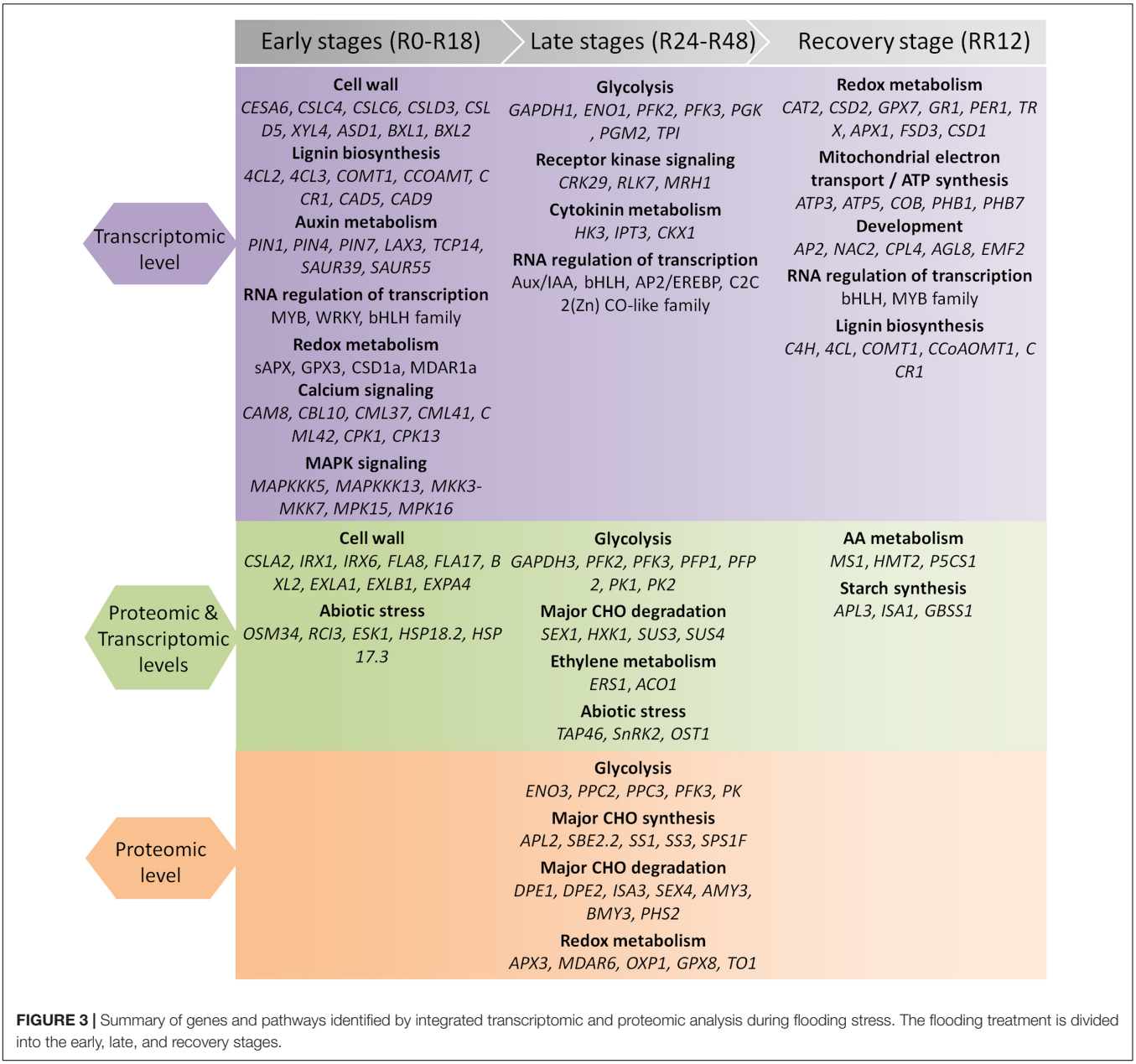


FIGURE 3 | Summary of genes and pathways identified by integrated transcriptomic and proteomic analysis during flooding stress. The flooding treatment is divided into the early, late, and recovery stages.

related to sucrose degradation (*SUS3c*, *SUS4*, *HK1*), starch biosynthesis (*APL1*, *APL3*, *SS4*, *SBE3*, *SBE4*, *GBSS1*), and starch degradation (*SEX1a*, *SEX1b*, *PHS1a*, *PHS1b*, *AMY3a*, *ISA1*) were regulated at both the transcriptome and proteome levels with low expression correlations, supporting a major contribution of post-transcriptional regulation in sucrose-starch metabolism under flooding.

Similar expression patterns were found for the glycolysis genes (**Figures 4A,C**), since most of them were up-regulated at distinct stages during flooding at the transcriptome, proteome, or both levels. Notably, 21 genes covering the whole glycolysis pathway were up-regulated from R6 to R48 at the transcriptome level, suggesting an emergency of energy demand under flooding treatment.

Collectively, these results clearly suggested stage-specific responses of sucrose-starch and glycolysis genes under flooding treatment mainly *via* post-transcriptional regulation.

Response of Genes Associated With Redox Metabolism

In total, 81 redox genes were differentially expressed in response to flooding, with the four most abundant types were TRX, SOD, glutathione peroxidase (GPX), and ascorbate peroxidase (APX). These genes were mainly grouped into gM1, gM5, gM7, gM8, and pM4 (**Figure 5** and **Supplementary Table 3**).

Many genes involved in ascorbate biosynthesis (*VTC2*, *GALDH*), hormone response (*HB2a*, *GPX3*, *GRX480a*), abiotic

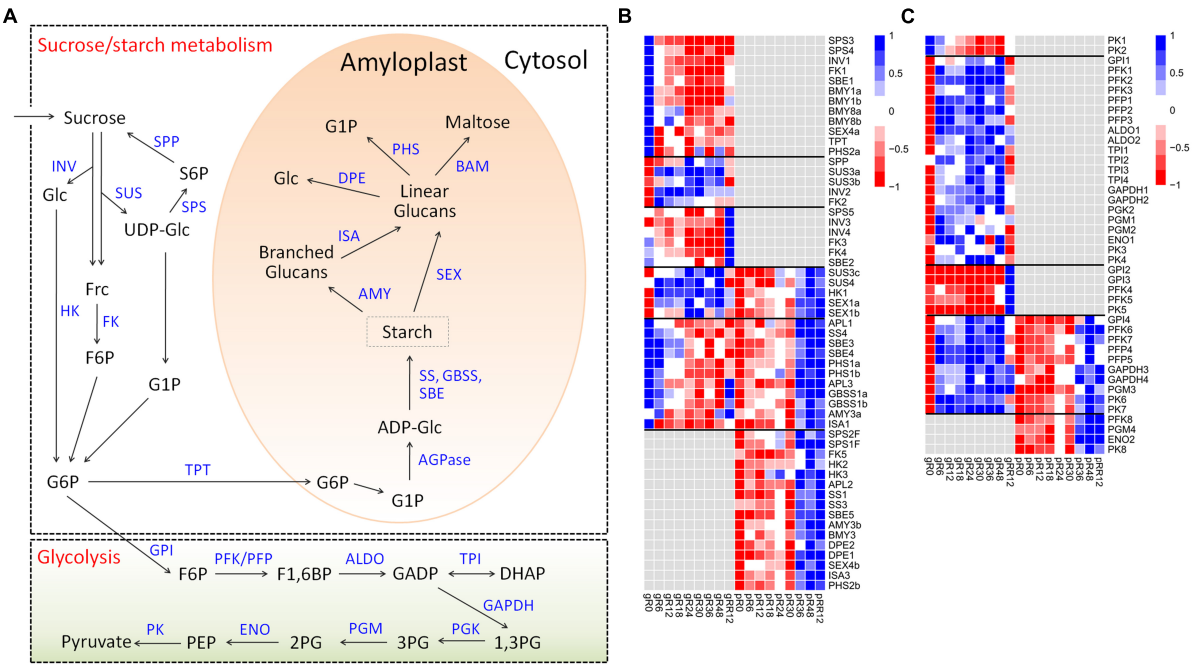


FIGURE 4 | Expression profiles of DEGs and DEPs related to sucrose-starch metabolism and glycolysis during flooding stress. **(A)** Summary of pathways of sucrose-starch metabolism and glycolysis. **(B)** Heatmap of DEGs and DEPs related to sucrose-starch metabolism. **(C)** Heatmap of DEGs and DEPs involved in glycolysis metabolism. The samples are prefixed as follows: g, transcriptomic level; p, proteomic level.

stress (*sAPX*, *GPX6a*, *FSD2*, *CSD1a*), and ROS removal (*MDAR1a*, *MDAR1b*) were found in gM1, and their expression levels were significantly depressed by flooding stress. In comparison, several genes related to cold stress (*HB1a*, *HB1b*) and lateral root development (*RLFa*, *RLFb*) were included in gM5, and they were up-regulated under flooding treatment. The genes in gM7 and gM8 were significantly up-regulated at RR12, indicating their major roles at the recovery stage. A few genes involved in temperature stress (*BAS1*, *GRXS17*) and superoxide detoxification (*CAT2a*, *CSD2a*) were found in gM7, while those related to H_2O_2 scavenger (*APX1a*), ascorbate biosynthesis (*GME*, *GLDH*, *VTC4*), and abiotic stress (*CSD1b*) were found in gM8. In addition, several genes involved in cold stress (*MDAR6*) and oxidative damage (*APX3*, *GPX8*) were found in pM4. These redox genes were regulated only at the transcriptome or proteome level.

On the contrary, a few genes functioned in drought (*CDSP32*), oxidative stress (*GPX6b* and *GPX6c*), and ROS-scavenging (*tAPX* and *GRX480b*) were regulated both at the transcriptome and proteome levels.

Together, these results suggested that redox genes were coordinately regulated in a stage-specific manner at the transcriptome and/or proteome levels under flooding treatment.

Response of Genes Involved in Hormone Metabolism

Diverse stage-specific expression patterns were observed for the genes involved in hormone biosynthesis and signaling transduction (**Figure 6** and **Supplementary Table 4**).

Four auxin influx carriers (*LAX3*, *PIN1*, *PIN4*, *PIN7*) and three auxin-responsive genes (*DFL2*, *SAUR39*, *TCP14*) were highly expressed at R0, and their expression levels were greatly depressed by flooding treatment (**Supplementary Table 4**). Two auxin-responsive genes (*ARF19*, *PID*) were highly expressed from R0 to R18, while two auxin receptors (*TIR1*, *AFB3*) were highly expressed from R24 to R48 at the transcriptome level. *PIN3* showed similar expression patterns to *TIR1* and *AFB3* at the transcriptome level, but it was also regulated at the proteome level. An auxin transport (*BIG*) was up-regulated from R36 to RR12 exclusively at the proteome level.

Likely, the brassinosteroid (BR) biosynthesis and signaling genes were highly expressed at R0 (*SMT1*, *DWF4*, *BIN2*, *BEH4a*, *BEH4b*), R6–R36 (*DWF7* and *BZR1*), and RR12 (*CPD*, *TRIP1*, *DET2*, and *DWF1*) at the transcriptome level, respectively. In addition, *DWF5* was highly expressed at R48–RR12 exclusively at the proteome level.

Four ABA biosynthesis genes (*NCED3b*, *NCED4*, *NCED3a*, *AAO3*) and a few ABA-responsive genes (*ABI1*, *CCD8*, *HVA22J*) were highly expressed at R0 and their expression levels were down-regulated by flooding stress. Interestingly, *NCED3a* and *AAO3* were up-regulated from R48 to RR12 while *NCED3b* and *NCED4* were not changed, indicating their different roles at the recovery stage. In addition, *AAO3* and *ABA1* were also up-regulated from R36 to RR12 at the proteome level.

Seven genes involved in cytokinin biosynthesis-degradation (*IPT1a*, *IPT1b*, *IPT3a*, *CKX5*) and signal transduction (*HK5a*, *WOLa*, *WOLb*) were highly expressed at R0, and their expression was greatly depressed by flooding treatment. By contrast, another

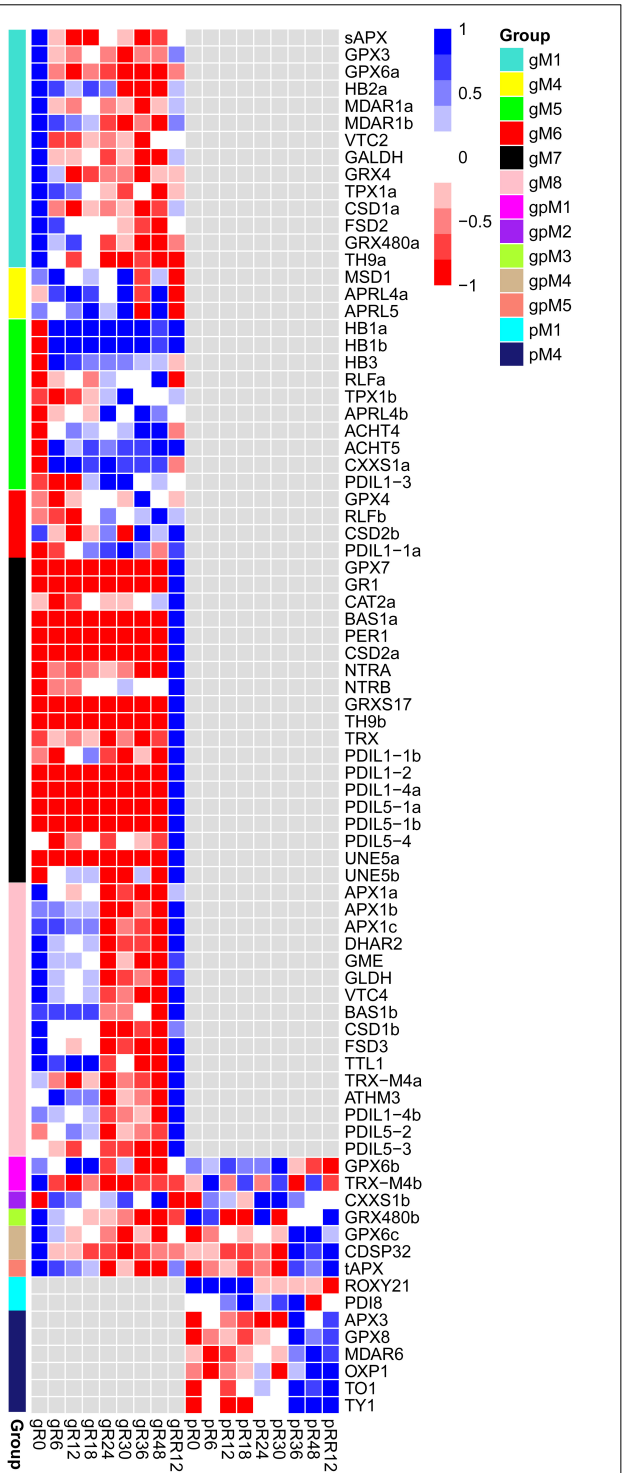


FIGURE 5 | Expression profiles of DEGs and DEPs involved in redox metabolism during flooding stress. Each row represents a gene with its name at the right.

six genes involved in cytokinin biosynthesis-degradation (*IPT3b*, *CKX1*) and signal transduction (*HK1*, *HK3a*, *HK3b*, *HK3c*) were highly induced from R24 to RR12.

Five genes involved in ethylene biosynthesis (*EFEa*, *EFEb*) and signal transduction (*ERF4*, *ETR2b*, *EIN4*) were highly expressed at RR12 at the transcriptome level. On the contrary, another six genes related to ethylene biosynthesis (*ACO1a*, *ACO1b*) and signal transduction (*ERF3b*, *ETR1*, *ETR2a*, *ERS1*) were highly expressed from R6 to R48. Moreover, *ERS1* and *ACO1b*, together with three ethylene biosynthesis genes (*EFEc*, *EFEd*, *ACS10*), showed high expression levels from R36 to RR12 at the proteome level.

GA20OX1 and *GA3OX1* were two vital genes involved in GA biosynthesis. These two genes and a GA receptor (*GID1B*) were specially induced at RR12 at the transcriptome level.

Four JA biosynthesis genes (*OPR3*, *AOS*, *LOX2*, *AOC3*) were highly expressed at R0–R12, while *OPR1* was highly expressed from R6 to R18 at the transcriptome level. *AOC4*, *LOX3*, and *LOX5* were highly expressed from R0 to R12 at the transcriptome level and from R36 to RR12 at the proteome level.

Collectively, these results revealed diverse stage-specific expression patterns of hormone genes during flooding treatment: ABA and JA mainly functioned at the early stage, GA mainly functioned at the recovery stage, while auxin, BR, CK, and ethylene functioned across all stages. These results also suggested a major role of hormone genes in flooding stress *via* post-transcriptional regulation.

Response of Transcriptional Factors

In total, 349 differentially expressed transcriptional factors (TFs) were found, with most of them regulated exclusively at the transcriptome level (**Figure 7** and **Supplementary Table 5**). The TF families with the most abundant numbers were MYB (65) and bHLH (57), and then followed by WRKY (31), AP2/EREBP (30), HB (22), and C2H2 (21). Notably, the TFs from each family trended to be enriched in a few expression patterns (e.g., gM1 and gM5), indicating the stage-specific roles of TFs under flooding treatment. Since the TFs were overall enriched in gM1, gM5, and gM8, these three groups were further inspected to interpret their molecular functions, respectively.

In gM1, many stress-responsive TFs were found from diverse gene families, including AP2/EREBP (*CBF1*, *CBF2*, *DREB2A*), bZIP (*GBF3a-c*), C2H2 (*AZF1*, *RHL41*), and MYB (*MYB65*) for low temperature, AP2/EREBP (*DREB2A*, *ERF53*) and MYB (*MYB102a*) for drought, and AP2/EREBP (*HRD*), C2H2 (*AZF1*), and WRKY (*WRKY33a*) for salt. Hormones play an important role in stress response, and accordingly, AP2/EREBP (*CBF1*, *CBF2*) and bZIP (*GBF3a-c*) for ABA, ARF (*ARF4*, *ARF16a*) for auxin, ARR (*ARR3a-b*, *ARR4*, *ARR10*, *ARR11*) for cytokinin, and HB (*ATH1*) for GA responses were found. Meantime, bHLH (*bHLH32*, *LRL1*) for root hair development, bHLH (*EGL3*) and MYB (*MYB5a*, *MYB23*) for trichome development, and MYB (*FLP*) and bHLH (*FMA*) for stomata development were observed. Interestingly, a few TF families involved in nutrition response were also included, including bHLH (*PYEa-b*) for iron response, C2H2 (*STOP1*) for aluminum toxicity, and WRKY (*WRKY6*) and bHLH (*bHLH32*) for phosphorus response.

In gM5, the members from AP2/EREBP (*RAV1*) and C2C2(Zn) CO-like (*BBX29*) related to low temperature,

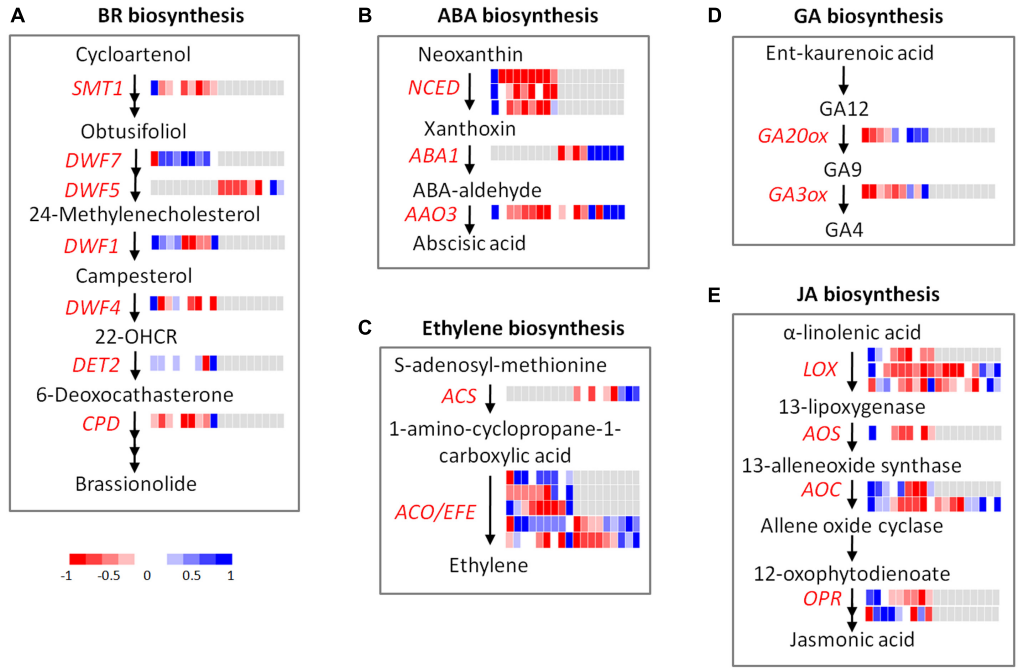


FIGURE 6 | Expression profiles of DEGs and DEPs involved in hormone metabolism during flooding stress. **(A)** BR biosynthesis, **(B)** ABA biosynthesis, **(C)** ethylene biosynthesis, **(D)** GA biosynthesis, **(E)** JA biosynthesis. Each row represents a gene, and the cells (from left to right) indicate the normalized expression value of R0, R6, R12, R18, R24, R30, R36, R48, and RR12 at the transcriptomic and proteomic levels, respectively.

C2C2(Zn) DOF (*DAG1*) involved in seed germination, and MYB (*MYB5b*) responsible for trichome development were found.

In gM8, the members from bHLH (*TT8a*) and MYB (*MYB5c*) for trichome development, HB (*HB6b*) and MYB (*MYB102b-c* and *MYB41*) for abiotic stress and ABA response, bHLH (*FIT1*) for iron uptake, and MYB (*TT2*) and bHLH (*TT8a*) for anthocyanin biosynthesis were included.

These results suggested stage-specific roles of TFs under flooding treatment: e.g., TFs related to abiotic stress, hormone, development, and nutrition were depressed at the early-stage (gM1), those involved in cold and seed germination were highly expressed at the mid-stage (gM5), while those referred to development, abiotic stress, and anthocyanin biosynthesis were induced at the recovery-stage (gM8).

DISCUSSION

The Genes Regulated Both at the Transcriptomic and Proteomic Levels

Transcriptome and proteome methods have widely been applied to explore DEGs and DEPs in response to flooding stress (Nanjo et al., 2012; Minami et al., 2018). However, very few studies were performed by integrative analysis of these two approaches, a promising way to systematically investigate complex physiological processes (Ding et al., 2020). To better understand the molecular mechanisms underlying flooding response, transcriptome and proteome were carried out in parallel to identify the critical genes and proteins under a naturally simulated flooding stress in *M. laxiflora* in this work. In total, 16,893 DEGs and 1,900 DEPs were identified under flooding. Of which, only 955 were commonly found at the transcriptome and proteome levels (**Supplementary Table 1**), suggesting that the proteins with significant expression changes did not always show a corresponding change at the transcriptome level.

Many genes and pathways have been demonstrated to be involved in flooding responses (Nanjo et al., 2012; Minami et al., 2018); however, their regulatory roles remain elusive at both the transcriptome and proteome levels. Carbohydrate metabolism genes were significantly depressed by flooding at the transcriptome level (Chen et al., 2016). This phenomenon was also observed in this work. Besides, a few sucrose degradation genes were up-regulated by flooding (*SUS3a*, *SUS3b*, *INV2*, *FK2*) or by post-flooding recovery treatment (*INV3*, *INV4*, *FK3*, *FK4*) exclusively at the transcriptome level (**Figure 4B**). In addition to the starch degradation genes (*AMY3b*, *BMY3*, *DPE1*, *DPE2*, *SEX4b*, *ISA3*, *PHS2b*) up-regulated by flooding exclusively at the proteome level, many starch biosynthesis (*APL1*, *APL3*, *SS4*, *SBE3*, *SBE4*, *GBSS1*) and degradation (*SEX1a*, *SEX1b*, *PHS1a*, *PHS1b*, *AMY3a*, *ISA1*) genes were regulated at both the transcriptome and proteome levels with low expression correlations (**Figure 4B**). These results supported a complex regulation of carbohydrate metabolism genes under flooding conditions (Wang and Komatsu, 2018).

Due to the energy shortage instigated by reduced carbohydrate metabolism, glycolysis is activated to produce energy for plant survival under flooding conditions (Lin et al., 2019). Consistently,

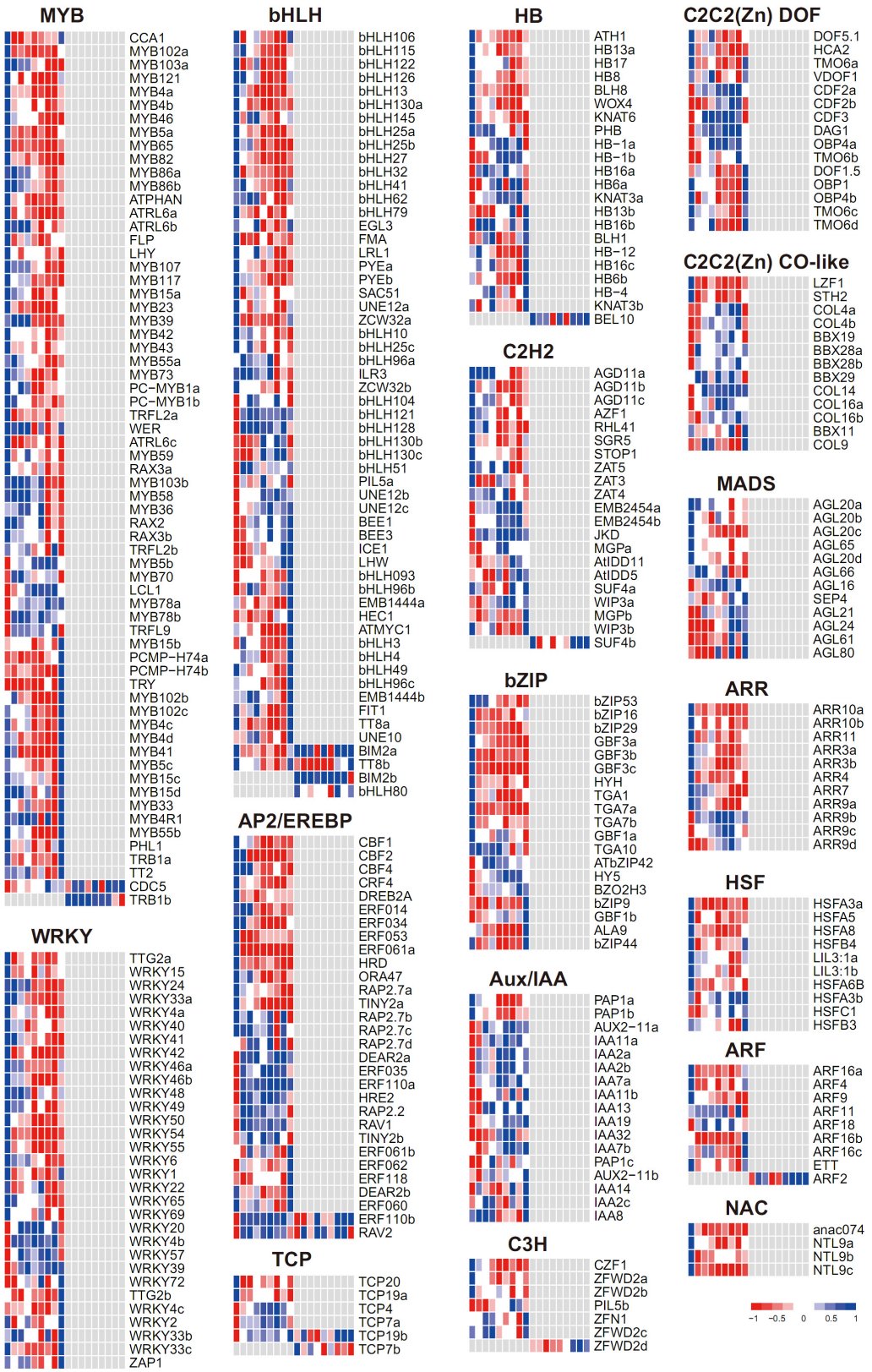


FIGURE 7 | Expression profiles of transcriptional factors during flooding stress. Each row represents a TF with its name at the right. The cells (from left to right) indicate the normalized expression value of R0, R6, R12, R18, R24, R30, R36, R48, and RR12 at the transcriptomic and proteomic levels, respectively.

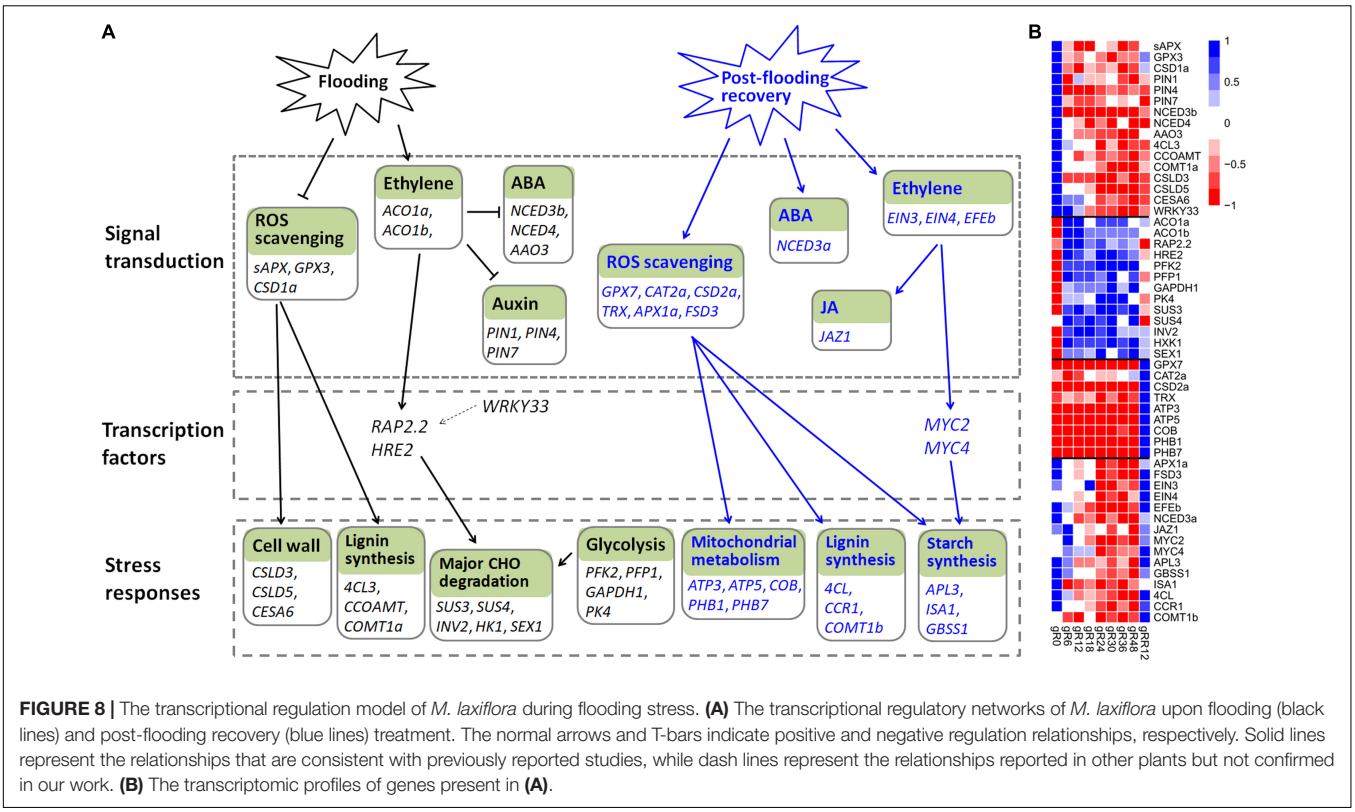


FIGURE 8 | The transcriptional regulation model of *M. laxiflora* during flooding stress. **(A)** The transcriptional regulatory networks of *M. laxiflora* upon flooding (black lines) and post-flooding recovery (blue lines) treatment. The normal arrows and T-bars indicate positive and negative regulation relationships, respectively. Solid lines represent the relationships that are consistent with previously reported studies, while dash lines represent the relationships reported in other plants but not confirmed in our work. **(B)** The transcriptomic profiles of genes present in **(A)**.

diverse regulatory patterns were observed for the glycolysis genes at transcriptomic and/or proteomic levels (**Figure 4C**), in accord with the dynamic changes of starch-sucrose genes. Proteins related to AA metabolism, transport, and development were identified in soybean in response to flooding stress (Yin et al., 2014). Similarly, many genes relevant to AA metabolism (*FAH*, *ADT1*, *URE*, *PGDH2*), phosphate transport (*CUE1*, *PHT1;1*, *PDR2*), and apical meristem development (*NAC2*, *DET1*, *YLS8*) were identified in this work. Moreover, these genes displayed negatively correlations or no correlations between the transcriptomic and proteomic levels (**Supplementary Table 1**).

On the contrary, many genes related to cell wall (including *CSLA2* and *IRX1* for cellulose synthesis, *GH9B5* and *BXL2* for cell wall degradation, *EXLA1*, *EXLB1*, and *EXPA4* for cell wall modification) and abiotic stress (*OSM34*, *RCI3*, *ESK1*, *HSP18.2*), which also played an essential role in flooding response (Komatsu et al., 2009; Minami et al., 2018), showed a consistent expression change at the transcriptome and proteome levels, indicating a major role of transcriptional regulation in flooding response.

These results deepen our understanding of the regulatory mechanisms of diverse gene pathways and suggested a major contribution of both transcriptional and post-transcriptional regulation to flooding stress.

The Regulatory Network of *Myricaria laxiflora* Under Flooding

Ethylene is a primary signal to trigger plant-adaptive responses to flooding (Kuroha et al., 2018). It accumulates quickly under flooding conditions, and the activities of ethylene biosynthetic enzymes (such as ACO and ACS) are induced in the meantime (Van Veen et al., 2013; Sasidharan and Voesenek, 2015). The elevated ethylene causes a massive depression of ABA, mediated by the down-regulation of 9-*cis*-epoxycarotenoid dioxygenase and the up-regulation of ABA-8-hydroxylase (Van Veen et al., 2013). In lines with previous studies, two ACO genes (*ACO1a* and *ACO1b*) were induced by flooding, while three genes (*NCED3b*, *NCED4*, *AAO3*) involved in ABA biosynthesis were greatly down-regulated in this work. Meantime, ethylene inhibited auxin transport in roots and hampered the normal growth of adventitious roots (Vidoz et al., 2010), and accordingly, the expression levels of three auxin transporters (*PIN1*, *PIN4*, *PIN7*) were depressed (**Figure 8** and **Supplementary Table 6**).

Reactive oxygen species is another important signal in hypoxia sensing and ethylene-mediated root development (Steffens et al., 2011). Consistently, a few ROS scavenging genes (*sAPX*, *GPX3*, *CSD1a*) were depressed by flooding stress. Several genes related to calcium signaling (*CAM8*, *CBL42*, *CML37*, *CPK1*) and MAPK signaling (*MAPKKK5*, *MKK3*, *MPK15*), which are key signals involved in flooding and abiotic stress (Zeng et al., 2017; Wang and Komatsu, 2018), were also dramatically depressed.

ERF transcription factors play an essential role in flooding and hypoxic stress (Van Dongen and Licausi, 2015). *RAP2.2* and *HRE2*, which were involved in hypoxia-responsive gene regulation and flooding survival (Sasidharan and Voesenek, 2015), were significantly up-regulated by flooding in this work. *RAP2.2* also induced the expression of genes encoding sugar metabolism (Hinz et al., 2010). However, the transcription

factor *WRKY33*, which up-regulated *RAP2.2* for *Arabidopsis* adaptation to submergence stress (Tang et al., 2021), was down-regulated by flooding, indicating a different regulatory network of *M. laxiflora* upon flooding.

Flooding-induced reduction in oxygen and CO_2 inhibits photosynthesis and aerobic respiration and causes a strong decline in energy and carbohydrate availability (Yin and Komatsu, 2017). Consistent with the severely reduction in cell wall formation (Bailey-Serres and Voesenek, 2010), a few cell wall (*CSLD3*, *CSLD5*, *CESA6*) and lignin biosynthesis (*4CL3*, *COMT1a*, *CCOAMT*) genes were down-regulated by flooding. To cope with energy supply, the sucrose-starch and glycolysis pathways are activated (Bailey-Serres and Voesenek, 2008). Accordingly, several genes involved in glycolysis (*PFK2*, *PFP1*, *GAPDH1*, *PK4*) and major CHO metabolism (*SUS3*, *SUS4*, *INV2*, *HK1*, *SEX1*) were greatly induced (**Figure 8**). Collectively, these results provide a complex regulatory network of *M. laxiflora* under flooding stress.

The Regulatory Network of *Myricaria laxiflora* Upon Post-flooding Recovery

Flooding responses have been extensively studied in plants, however, less is known about the signaling networks at the post-flooding recovery stage (Yeung et al., 2018).

Reactive oxygen species acts as a primary signaling in plant post-flooding recovery due to the combined effects of reoxygenation, reillumination, dehydration, and senescence stresses (Yeung et al., 2019). ROS generated during post-flooding disrupts the photosynthetic apparatus, and consequently, hampers the photosynthetic recovery and carbohydrate replenishment (Sasidharan et al., 2018). Therefore, the ROS scavengers are subsequently activated during recovery (Yeung et al., 2018). In this work, many genes related to ROS scavenging (*GPX7*, *CAT2a*, *CSD2a*, *TRX*, *APX1a*, *FSD3*) were up-regulated at post-flooding recovery stage. This efficient ROS scavenging may result in faster starch replenishment, correlating with higher biomass and survival ratios (Qin et al., 2013; Yeung et al., 2018). Accordingly, many genes related to starch biosynthesis (*APL3*, *ISA1*, *GBSS1*), development (*AP2*, *NAC2*, *EMB2746*, *AGL8*), and lignin biosynthesis (*4CL*, *CCR1*, *COMT1b*) were up-regulated upon post-flooding recovery (**Figure 8** and **Supplementary Table 6**). ROS accumulation caused by reoxygenation and reillumination is also associated with the high-energy demands of reactivated mitochondrial metabolism (Sasidharan et al., 2018), therefore, a few genes referred to mitochondrial electron transport/ATP synthesis (*ATP3*, *ATP5*, *COB*, *PHB1*, *PHB7*) were up-regulated by post-flooding recovery treatment (**Figure 8**). These results suggested that the quickly growth of *M. laxiflora* during post-flooding recovery might due to its highly efficient ROS scavenging system (Yeung et al., 2018).

In addition to ROS, hormones and TFs also confer post-flooding recovery responses (Yeung et al., 2019). Several ethylene biosynthesis genes (*ACS2*, *ACS8*, *ACO2*) and ethylene-responsive factors (*ERF1* and *ERF2*) were induced upon reoxygenation (Tsai et al., 2014). Over-expression of *MYC2* enhanced stress tolerance to reoxygenation by JA signaling during submergence recovery (Yuan et al., 2017). Furthermore, *MYC2* physically interacted with the transcription factors *EIN3* and *EIL1* to regulate ethylene-mediated post-flooding responses (Song et al., 2014). In this work, an ethylene-forming enzyme (*EFEb*) and two ethylene receptors (*EIN3* and *EIN4*) were greatly induced by post-flooding recovery treatment. However, none of the ethylene-responsive factors (*ERF1*, *ERF3*, and *ERF4*) showed a similar expression pattern, indicating distinct regulatory pathways of ethylene in post-flooding recovery between *M. laxiflora* and *Arabidopsis* (Song et al., 2014).

MYC2 and *MYC4* function redundantly in JA signaling (Song et al., 2014). These two genes, together with their interaction component *JAZ1* involved in JA signaling (Zhou et al., 2016), were significantly up-regulated upon recovery. *MYC2* also co-expressed with several starch synthesis genes (*APL3*, *ISA1*, and *GBSS1*) during recovery, suggesting a possible involvement of *MYC2* in carbohydrate metabolism *via* regulating the expression of starch-metabolic genes (Bian et al., 2021). In addition, *NCED3a*, a key gene involved in ABA biosynthesis, was significantly induced upon recovery, in accord with the crucial roles of ABA in post-flooding recovery (Yeung et al., 2018). Together, these results suggested a complex regulatory network of ROS, JA, ethylene, and ABA signaling in *M. laxiflora* upon post-flooding recovery.

CONCLUSION

In summary, the molecular mechanisms underlying flooding stress were investigated by integrated transcriptomic and proteomic approaches in *M. laxiflora* in nine time-points. Our results not only uncovered the highly dynamic and stage-specific expression changes of genes/proteins during flooding and post-flooding recovery, but also highlighted the genes and pathways that functioned specially at different stages. The genes involved in auxin, cell wall, calcium signaling, and MAP kinase signaling were depressed during the early stages of flooding at the transcriptomic level, the genes referred to glycolysis and major CHO metabolism mainly functioned during the late stages of flooding at the transcriptomic and/or proteomic levels, while those genes related to ROS scavenging, mitochondrial metabolism, and development were activated upon post-flooding recovery at the transcriptomic level. In addition, the genes/proteins related to redox, hormones, and TFs also played vital roles in flooding stress of *M. laxiflora*. The findings will improve our understanding of molecular mechanisms of flooding stress, and provide useful information for the preservation of *M. laxiflora* and other endangered plants in the flood zone.

DATA AVAILABILITY STATEMENT

The original contributions presented in the study are publicly available. The accession numbers are here: NCBI, PRJNA840865 and iProX, IPX0004477000.

AUTHOR CONTRIBUTIONS

LL, DW, and JL conceived and designed the research and revised the manuscript. GH, WX, HFZ, HBZ, and JZ performed the experiments. LL and ZD analyzed the data and drafted the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This work was supported by the Ecological Environmental Protection Fund of China Three Gorges Corporation (WWKY-2020-0267). The funder was not involved in the study design, collection, analysis, interpretation of data, the writing of this manuscript, or the decision to submit it for publication.

ACKNOWLEDGMENTS

We thank the Shanghai OE Biotech Company (Shanghai, China) for the technical support of this study.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.924490/full#supplementary-material

REFERENCES

Agullo-Anton, M. A., Ferrandez-Ayela, A., Fernandez-Garcia, N., Nicolas, C., Albacete, A., Perez-Alfocea, F., et al. (2014). Early steps of adventitious rooting: morphology, hormonal profiling and carbohydrate turnover in carnation stem cuttings. *Physiol. Plant* 150, 446–462. doi: 10.1111/ppl.12114

Bailey-Serres, J., and Voesenek, L. A. (2008). Flooding stress: acclimations and genetic diversity. *Annu. Rev. Plant Biol.* 59, 313–339. doi: 10.1146/annurev.arplant.59.032607.092752

Bailey-Serres, J., and Voesenek, L. A. (2010). Life in the balance: a signaling network controlling survival of flooding. *Curr. Opin. Plant. Biol.* 13, 489–494. doi: 10.1016/j.pbi.2010.08.002

Bao, D., Lu, Z., Jiang, M., Xu, S., Yao, Q., Liu, Q., et al. (2010). Population structure and dynamics of remanent Myricaria laxiflora downstream from the Three Gorges dam. *J. Wuhan Bot. Res.* 28, 711–716.

Baxter-Burrell, A., Yang, Z., Springer, P. S., and Bailey-Serres, J. (2002). RopGAP4-dependent Rop GTPase rheostat control of Arabidopsis oxygen deprivation tolerance. *Science* 296, 2026–2028. doi: 10.1126/science.1071505

Bian, S., Tian, T., Ding, Y., Yan, N., Wang, C., Fang, N., et al. (2021). bHLH transcription factor NtMYC2a regulates carbohydrate metabolism during the pollen development of tobacco (Nicotiana tabacum L. cv. TN90). *Plants* 11:17. doi: 10.3390/plants11010017

Bolger, A. M., Lohse, M., and Usadel, B. (2014). Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30, 2114–2120. doi: 10.1093/bioinformatics/btu170

Boulange, J., Hanasaki, N., Yamazaki, D., and Pokhrel, Y. (2021). Role of dams in reducing global flood exposure under climate change. *Nat. Commun.* 12:417. doi: 10.1038/s41467-020-20704-0

Bruderer, R., Bernhardt, O. M., Gandhi, T., Miladinovic, S. M., Cheng, L. Y., Messner, S., et al. (2015). Extending the limits of quantitative proteome profiling with data-independent acquisition and application to acetaminophen-treated three-dimensional liver microtissues. *Mol. Cell Proteomics* 14, 1400–1410. doi: 10.1074/mcp.M114.044305

Chen, F., Guan, S., Ma, Y., Xie, Z., Lv, K., Huang, Y., et al. (2019). Impact of regulated water level fluctuations on the sexual reproduction of remnant Myricaria laxiflora populations. *Glob. Ecol. Conserv.* 18:e00628.

Chen, F.-Q., and Xie, Z.-Q. (2009). Survival and growth responses of Myricaria laxiflora seedlings to summer flooding. *Aquat. Bot.* 90, 333–338.

Chen, W., Yao, Q., Patil, G. B., Agarwal, G., Deshmukh, R. K., Lin, L., et al. (2016). Identification and comparative analysis of differential gene expression in soybean leaf tissue under drought and flooding stress revealed by RNA-Seq. *Front. Plant Sci.* 7:1044. eCollection 2016 doi: 10.3389/fpls.2016.01044

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

Dat, J. F., Capelli, N., Folzer, H., Bourgeade, P., and Badot, P. M. (2004). Sensing and signalling during plant flooding. *Plant Physiol. Biochem.* 42, 273–282. doi: 10.1016/j.plaphy.2004.02.003

Dawood, T., Yang, X., Visser, E. J., Te Beek, T. A., Kensche, P. R., Cristescu, S. M., et al. (2016). A co-opted hormonal cascade activates dormant adventitious root primordia upon flooding in Solanum dulcamara. *Plant Physiol.* 170, 2351–2364. doi: 10.1104/pp.15.00773

De Ollas, C., Gonzalez-Guzman, M., Pitarch, Z., Matus, J. T., Candela, H., Rambla, J. L., et al. (2021). Identification of ABA-mediated genetic and metabolic responses to soil flooding in tomato (Solanum lycopersicum L. Mill). *Front. Plant Sci.* 12:613059. doi: 10.3389/fpls.2021.613059.

Ding, Z., Fu, L., Tie, W., Yan, Y., Wu, C., Dai, J., et al. (2020). Highly dynamic, coordinated, and stage-specific profiles are revealed by a multi-omics integrative analysis during tuberous root development in cassava. *J. Exp. Bot.* 71, 7003–7017. doi: 10.1093/jxb/eraa369

Ding, Z., Fu, L., Tie, W., Yan, Y., Wu, C., Hu, W., et al. (2019). Extensive post-transcriptional regulation revealed by transcriptomic and proteomic integrative analysis in cassava under drought. *J. Agric. Food Chem.* 67, 3521–3534. doi: 10.1021/acs.jafc.9b00014

Du, H., Chang, Y., Huang, F., and Xiong, L. (2015). GID1 modulates stomatal response and submergence tolerance involving abscisic acid and gibberellic acid signaling in rice. *J. Integr. Plant Biol.* 57, 954–968. doi: 10.1111/jipb.12313

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., et al. (2011). Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat. Biotechnol.* 29, 644–652. doi: 10.1038/nbt.1883

Guan, S., Chen, F., Zhou, J., Xie, Z., and Huang, Y. (2020). Spatiotemporal photosynthetic physiology responses of remnant Myricaria laxiflora populations to regulated water level fluctuations. *Conserv. Physiol.* 8:coaa020. doi: 10.1093/conphys/coaa020

Hashiguchi, A., Sakata, K., and Komatsu, S. (2009). Proteome analysis of early-stage soybean seedlings under flooding stress. *J. Proteome. Res.* 8, 2058–2069. doi: 10.1021/pr801051m

Hinz, M., Wilson, I. W., Yang, J., Buerstenbinder, K., Llewellyn, D., Dennis, E. S., et al. (2010). Arabidopsis RAP2.2: an ethylene response transcription factor that is important for hypoxia survival. *Plant Physiol.* 153, 757–772. doi: 10.1104/pp.110.155077

Hsu, F. C., Chou, M. Y., Chou, S. J., Li, Y. R., Peng, H. P., and Shih, M. C. (2013). Submergence confers immunity mediated by the WRKY22 transcription factor in Arabidopsis. *Plant Cell* 25, 2699–2713. doi: 10.1105/tpc.113.114447

Hsu, F. C., Chou, M. Y., Peng, H. P., Chou, S. J., and Shih, M. C. (2011). Insights into hypoxic systemic responses based on analyses of transcriptional regulation in Arabidopsis. *PLoS One* 6:e28888. doi: 10.1371/journal.pone.0028888

Jia, W., Ma, M., Chen, J., and Wu, S. (2021). Plant morphological, physiological and anatomical adaption to flooding stress and the underlying molecular mechanisms. *Int. J. Mol. Sci.* 22:1088. doi: 10.3390/ijms22031088

Klok, E. J., Wilson, I. W., Wilson, D., Chapman, S. C., Ewing, R. M., Somerville, S. C., et al. (2002). Expression profile analysis of the low-oxygen response

in Arabidopsis root cultures. *Plant Cell* 14, 2481–2494. doi: 10.1105/tpc.004747

Komatsu, S., Yamamoto, R., Nanjo, Y., Mikami, Y., Yunokawa, H., and Sakata, K. (2009). A comprehensive analysis of the soybean genes and proteins expressed under flooding stress using transcriptome and proteome techniques. *J. Proteome. Res.* 8, 4766–4778. doi: 10.1021/pr900460x

Kuroha, T., Nagai, K., Gamuyao, R., Wang, D. R., Furuta, T., Nakamori, M., et al. (2018). Ethylene-gibberellin signaling underlies adaptation of rice to periodic flooding. *Science* 361, 181–186. doi: 10.1126/science.aat1577

Langfelder, P., and Horvath, S. (2008). WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics* 9:559. doi: 10.1186/1471-2105-9-559

Li, L., Wu, D., Zhen, Q., Zhang, J., Qiu, L., Huang, G., et al. (2021). Morphological structures and histochemistry of roots and shoots in Myricaria laxiflora (Tamaricaceae). *Open Life Sci.* 16, 455–463. doi: 10.1515/biol-2021-0049

Lin, Y., Li, W., Zhang, Y., Xia, C., Liu, Y., Wang, C., et al. (2019). Identification of genes/proteins related to submergence tolerance by transcriptome and proteome analyses in soybean. *Sci. Rep.* 9:14688. doi: 10.1038/s41598-019-50757-1

Love, M. I., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome. Biol.* 15:550. doi: 10.1186/s13059-014-0550-8

Minami, A., Yano, K., Gamuyao, R., Nagai, K., Kuroha, T., Ayano, M., et al. (2018). Time-course transcriptomics analysis reveals key responses of submerged deepwater rice to flooding. *Plant Physiol.* 176, 3081–3102. doi: 10.1104/pp.17.00858

Nanjo, Y., Skultety, L., Uvackova, L., Klubicova, K., Hajduch, M., and Komatsu, S. (2012). Mass spectrometry-based analysis of proteomic changes in the root tips of flooded soybean seedlings. *J. Proteome Res.* 11, 372–385. doi: 10.1021/pr200701y

Qiao, D., Zhang, Y., Xiong, X., Li, M., Cai, K., Luo, H., et al. (2020). Transcriptome analysis on responses of orchardgrass (Dactylis glomerata L.) leaves to a short term flooding. *Hereditas* 157:20. doi: 10.1186/s41065-020-00134-0

Qin, X., Li, F., Chen, X., and Xie, Y. (2013). Growth responses and non-structural carbohydrates in three wetland macrophyte species following submergence and de-submergence. *Acta Physiol. Plant* 35, 2069–2074.

Rauf, M., Arif, M., Fisahn, J., Xue, G. P., Balazadeh, S., and Mueller-Roeber, B. (2013). NAC transcription factor speedy hyponastic growth regulates flooding-induced leaf movement in Arabidopsis. *Plant Cell* 25, 4941–4955. doi: 10.1105/tpc.113.117861

Sasidharan, R., Hartman, S., Liu, Z., Martopawiro, S., Sajeev, N., Van Veen, H., et al. (2018). Signal dynamics and interactions during flooding stress. *Plant Physiol.* 176, 1106–1117. doi: 10.1104/pp.17.01232

Sasidharan, R., and Voesenek, L. A. (2015). Ethylene-mediated acclimations to flooding stress. *Plant Physiol.* 169, 3–12. doi: 10.1104/pp.15.00387

Song, S., Huang, H., Gao, H., Wang, J., Wu, D., Liu, X., et al. (2014). Interaction between MYC2 and ETHYLENE INSENSITIVE3 modulates antagonism between jasmonate and ethylene signaling in Arabidopsis. *Plant Cell* 26, 263–279. doi: 10.1105/tpc.113.120394

Steffens, B., Geske, T., and Sauter, M. (2011). Aerenchyma formation in the rice stem and its promotion by H2O2. *New Phytol.* 190, 369–378. doi: 10.1111/j.1469-8137.2010.03496.x

Steffens, B., and Sauter, M. (2009). Epidermal cell death in rice is confined to cells with a distinct molecular identity and is mediated by ethylene and H2O2 through an autoamplified signal pathway. *Plant Cell* 21, 184–196. doi: 10.1105/tpc.108.061887

Steffens, B., Wang, J., and Sauter, M. (2006). Interactions between ethylene, gibberellin and abscisic acid regulate emergence and growth rate of adventitious roots in deepwater rice. *Planta* 223, 604–612. doi: 10.1007/s00425-005-0111-1

Tang, H., Bi, H., Liu, B., Lou, S., Song, Y., Tong, S., et al. (2021). WRKY33 interacts with WRKY12 protein to up-regulate RAP2.2 during submergence induced hypoxia response in Arabidopsis thaliana. *New Phytol.* 229, 106–125. doi: 10.1111/nph.17020

Thimm, O., Blasing, O., Gibon, Y., Nagel, A., Meyer, S., Kruger, P., et al. (2004). MAPMAN: a user-driven tool to display genomics data sets onto diagrams of metabolic pathways and other biological processes. *Plant J.* 37, 914–939. doi: 10.1111/j.1365-313x.2004.02016.x

Tsai, K. J., Chou, S. J., and Shih, M. C. (2014). Ethylene plays an essential role in the recovery of Arabidopsis during post-anaerobiosis reoxygenation. *Plant Cell Environ.* 37, 2391–2405. doi: 10.1111/pce.12292

Van Dongen, J. T., and Licausi, F. (2015). Oxygen sensing and signaling. *Annu. Rev. Plant Biol.* 66, 345–367.

Van Veen, H., Mustroph, A., Barding, G. A., Vergeer-Van Eijk, M., Welschen-Evertman, R. A., Pedersen, O., et al. (2013). Two Rumex species from contrasting hydrological niches regulate flooding tolerance through distinct mechanisms. *Plant Cell* 25, 4691–4707. doi: 10.1105/tpc.113.119016

Vidoz, M. L., Loreti, E., Mensuali, A., Alpi, A., and Perata, P. (2010). Hormonal interplay during adventitious root formation in flooded tomato plants. *Plant J.* 63, 551–562. doi: 10.1111/j.1365-313X.2010.04262.x

Wang, J., Sun, H., Sheng, J., Jin, S., Zhou, F., Hu, Z., et al. (2019). Transcriptome, physiological and biochemical analysis of Triarrhena sacchariflora in response to flooding stress. *BMC Genet* 20:88. doi: 10.1186/s12863-019-0790-4

Wang, X., and Komatsu, S. (2018). Proteomic approaches to uncover the flooding and drought stress response mechanisms in soybean. *J. Proteomics* 172, 201–215. doi: 10.1016/j.jprot.2017.11.006

Wang, Y., Sang, Z., Xu, S., Xu, Q., Zeng, X., Jabu, D., et al. (2020). Comparative proteomics analysis of Tibetan hull-less barley under osmotic stress via data-independent acquisition mass spectrometry. *Gigascience* 9:giaa019. doi: 10.1093/gigascience/giaa019

Wang, Y., Wu, J., Tao, Y., Li, Z., and Huang, H. (2003). Natural distribution and ex situ conservation of endemic species Myricaria laxiflora in water-level-fluctuation Zone within Three-Gorges reservoir area of Changjiang River. *Wuhan Bot. Res.* 21, 415–422.

Wu, J., Zhao, H. B., Yu, D., and Xu, X. (2017). Transcriptome profiling of the floating-leaved aquatic plant *Nymphoides peltata* in response to flooding stress. *BMC Genomics* 18:119. doi: 10.1186/s12864-017-3515-y

Wu, J., Zhao, Z., Jin, Y., and Shen, Z. (1998). Investigation and study on the endemic plant Myricaria laxiflora in the Three-Gorge reservoir area. *Wuhan Bot. Res.* 16, 111–116.

Ye, T., Shi, H., Wang, Y., and Chan, Z. (2015). Contrasting changes caused by drought and submergence stresses in Bermudagrass (Cynodon dactylon). *Front. Plant Sci.* 6:951. doi: 10.3389/fpls.2015.00951

Yeung, E., Bailey-Serres, J., and Sasidharan, R. (2019). After the deluge: plant revival post-flooding. *Trends Plant Sci.* 24, 443–454. doi: 10.1016/j.tplants.2019.02.007

Yeung, E., Van Veen, H., Vashisht, D., Sobral Paiva, A. L., Hummel, M., Rankenberg, T., et al. (2018). A stress recovery signaling network for enhanced flooding tolerance in Arabidopsis thaliana. *Proc. Natl. Acad. Sci. U. S. A.* 115, E6085–E6094. doi: 10.1073/pnas.1803841115

Yin, D., Sun, D., Han, Z., Ni, D., Norris, A., and Jiang, C. Z. (2019). PhERF2, an ethylene-responsive element binding factor, plays an essential role in waterlogging tolerance of petunia. *Hortic. Res.* 6:83. doi: 10.1038/s41438-019-0165-z

Yin, X., and Komatsu, S. (2017). Comprehensive analysis of response and tolerant mechanisms in early-stage soybean at initial-flooding stress. *J. Proteomics* 169, 225–232. doi: 10.1016/j.jprot.2017.01.014

Yin, X., Sakata, K., Nanjo, Y., and Komatsu, S. (2014). Analysis of initial changes in the proteins of soybean root tip under flooding stress using gel-free and gel-based proteomic techniques. *J. Proteomics* 106, 1–16. doi: 10.1016/j.jprot.2014.04.004

Yuan, L. B., Dai, Y. S., Xie, L. J., Yu, L. J., Zhou, Y., Lai, Y. X., et al. (2017). Jasmonate regulates plant responses to postsubmergence reoxygenation through transcriptional activation of antioxidant synthesis. *Plant Physiol.* 173, 1864–1880. doi: 10.1104/pp.16.01803

Zeng, C., Ding, Z., Zhou, F., Zhou, Y., Yang, R., Yang, Z., et al. (2017). The discrepant and similar responses of genome-wide transcriptional profiles between drought and cold stresses in cassava. *Int. J. Mol. Sci.* 18:2668. doi: 10.3390/ijms18122668

Zhou, J., Chen, F., Lv, K., Guan, S., and Huang, Y. (2020). Endodormancy induction and photosynthetic physiology of Myricaria laxiflora remnant populations under chronic summer submersion. *Flora* 271:151682.

Zhou, J., Chen, F., Lv, K., Wu, Y., and Huang, Y. (2021). Summer dormancy in an endangered riparian shrub Myricaria laxiflora: Changes in branches, leaves, and nonstructural carbohydrates. *Glob. Ecol. Conserv.* 31:e01809.

Zhou, M., Sun, Z., Li, J., Wang, D., Tang, Y., and Wu, Y. (2016). Identification of JAZ1-MYC2 complex in Lotus corniculatus. *J. Plant Growth Regul.* 35, 440–448.

Conflict of Interest: LL, GH, HFZ, HBZ, JZ, and DW were employed by the China Three Gorges Corporation.

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

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 Plant Science

ORIGINAL RESEARCH
published: 17 June 2022
doi: 10.3389/fpls.2022.920138

The Antarctic Moss *Pohlia nutans* Genome Provides Insights Into the Evolution of Bryophytes and the Adaptation to Extreme Terrestrial Habitats

Shenghao Liu[1,2], Shuo Fang[1], Bailin Cong[1,2], Tingting Li[1], Dan Yi[1], Zhaohui Zhang[1], Linlin Zhao[1,2] and Pengying Zhang[3]**

[1]Key Laboratory of Marine Eco-Environmental Science and Technology, First Institute of Oceanography, Ministry of Natural Resources, Qingdao, China, [2]School of Advanced Manufacturing, Fuzhou University, Jinjiang, China, [3]National Glycoengineering Research Center, School of Life Sciences and Shandong University, Qingdao, China

OPEN ACCESS

Edited by:
Freddy Mora-Poblete,
University of Talca, Chile

Reviewed by:
Marely Cuba-Díaz,
Universidad de Concepción, Campus Los Ángeles, Chile
Gustavo E. Zuñiga,
University of Santiago, Chile
Hyoungseok Lee,
Korea Polar Research Institute, South Korea

***Correspondence:**
Linlin Zhao
zhaolinlin@fio.org.cn
Pengying Zhang
zhangpy80@sdu.edu.cn

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

Received: 14 April 2022
Accepted: 19 May 2022
Published: 17 June 2022

Citation:
Liu S, Fang S, Cong B, Li T, Yi D, Zhang Z, Zhao L and Zhang P (2022) The Antarctic Moss Pohlia nutans Genome Provides Insights Into the Evolution of Bryophytes and the Adaptation to Extreme Terrestrial Habitats. Front. Plant Sci. 13:920138. doi: 10.3389/fpls.2022.920138

The Antarctic continent has extreme natural environment and fragile ecosystem. Mosses are one of the dominant floras in the Antarctic continent. However, their genomic features and adaptation processes to extreme environments remain poorly understood. Here, we assembled the high-quality genome sequence of the Antarctic moss (*Pohlia nutans*) with 698.20 Mb and 22 chromosomes. We found that the high proportion of repeat sequences and a recent whole-genome duplication (WGD) contribute to the large size genome of *P. nutans* when compared to other bryophytes. The genome of *P. nutans* harbors the signatures of massive segmental gene duplications and large expansions of gene families, likely facilitating neofunctionalization. Genomic characteristics that may support the Antarctic lifestyle of this moss comprise expanded gene families involved in phenylpropanoid biosynthesis, unsaturated fatty acid biosynthesis, and plant hormone signal transduction. Additional contributions include the significant expansion and upregulation of several genes encoding DNA photolyase, antioxidant enzymes, flavonoid biosynthesis enzymes, possibly reflecting diverse adaptive strategies. Notably, integrated multi-omic analyses elucidate flavonoid biosynthesis may function as the reactive oxygen species detoxification under UV-B radiation. Our studies provide insight into the unique features of the Antarctic moss genome and their molecular responses to extreme terrestrial environments.

Keywords: whole-genome sequencing, gene-family expansions, Antarctic bryophyte, metabolomic profiling, flavonoid biosynthesis, UV-B radiation

INTRODUCTION

The features of the Antarctic terrestrial environment are characterized by high ultraviolet radiation, freezing, and extreme dryness (Perera-Castro et al., 2020; Cassaro et al., 2021). Notably, the strong ultraviolet radiation in Antarctica is a typical consequence of global climate change and human activities. Enhanced UV-B radiation (280–315 nm) has been widely concerned

since the 1980s due to ozone depletion, which results from the breakdown of chlorofluorocarbons (CFCs) in the atmospheric stratosphere (Hossaini et al., 2017; Neale et al., 2021). Due to the Montreal Protocol and its amendments prohibiting the release of harmful CFCs, the ozone depletion seems to have slowed down. However, ozone hole run up to 24.8×10^6 km^2 in 2021 and enlarged to most regions of the Antarctic continent. The Antarctic terrestrial plants are undergoing the higher UV-B light of 3.4–6.2 mW/cm^2 (Bao et al., 2018). They may battle the UV-B radiation by producing antioxidants such as flavonoids and other UV-B-absorbing pigments, modulating reactive oxygen species (ROS)-scavenging enzyme activities (Pereira et al., 2009; Singh et al., 2011; Singh and Singh, 2014; Wang et al., 2020). The field experiments demonstrated that continuing UV-B radiation reduced the chlorophyll contents in the Antarctic moss (*Bryum argenteum*) and lichen (*Umbilicaria aprina*), and increased the contents of UV-B absorbing compounds (Singh and Singh, 2014). The Multi-omics analysis indicated that UVR8-mediated signaling, flavonoid biosynthesis, and DNA repair machinery might facilitate the adaptation of Antarctic moss to UV-B radiation (Li et al., 2019; Liu et al., 2021). The moderate moss *Physcomitrella patens* was more qualified of surviving UV-B radiation than Arabidopsis, and several flavonoid biosynthesis genes were also markedly upregulated in response to UV-B radiation (Wolf et al., 2010). However, compared with other land plants, the molecular mechanism of bryophytes acclimation to strong UV-B radiation was far less documented.

Antarctica is considered to be the coldest continent on Earth. The daily values of air temperature over 0°C are only achieved for the few short summer months or weeks (Convey et al., 2018). Field experiments showed that the moss surface temperatures were over +4°C for 43.4% of the peak Antarctic summer, while surface temperatures exceeded 14°C for an average of just 2.5% of the time (Perera-Castro et al., 2020). Antarctic mosses also underwent a higher frequency of air freeze-thaw cycles during the austral summer. However, the regional climate in maritime Antarctica such as the South Shetlands Islands is relatively milder, with temperature ranges from −5°C to 13°C in the summer daytime (Perera-Castro et al., 2020). Furthermore, Antarctic mosses usually have surface temperatures well above air temperature, over 15°C at midday in summer, relying on the water content of moss tundra. Antarctic mosses propagating in this cold habitat largely depend on the capacity to maximize photosynthesis for short summer and reduce respiratory carbon losses (Perera-Castro et al., 2020).

Most regions of the East Antarctica are increasing dryness due to global climate change. In recent decades, the more positive Southern Annular Mode, driven by ozone depletion and greenhouse gas emissions, maintains freezing temperature in East Antarctica (Robinson et al., 2018). Subsequently, the freezing temperature causes the drying climate in the East Antarctica. Vegetation distributions are markedly influenced by local availability of ice-free lands and water. For instance, in Windmill Islands, East Antarctica, moss communities are the well-developed and extensive; but the health of moss-beds is declining due to regional dryness (Malenovský et al., 2015). Drought stress usually causes photoinhibition and photodamage of photosynthetic apparatus with growth retardation and yield reduction.

Global climate change and human activities are having a significant impact on the Antarctic terrestrial ecosystem (Convey and Peck, 2019; Cannone et al., 2022). The Antarctic Peninsula is now experiencing some of the most rapid warming on earth (Sato et al., 2021). The ice-free areas of Antarctica is expected to increase by approximately 25% by the end of 21st century due to climate change, while most of this expansion will occur in the Antarctic Peninsula (Lee et al., 2017). Warming had caused the special phenomena of flowering plant spread and snow algae outbreak (Gray et al., 2020; Cannone et al., 2022). Similarly, a significant increase was found in biological activity over the past 50 years, measured through moss growth or accumulation rates (Amesbury et al., 2017). Currently, Antarctic plants appears to be promising models for studying the adaptation mechanism to various abiotic stresses (Convey and Peck, 2019; Perera-Castro et al., 2020; Bertini et al., 2021; Liu et al., 2021) and monitoring the regional climate changes (Amesbury et al., 2017; Lee et al., 2017; Robinson et al., 2018; Cannone et al., 2022) as well as assessing the impact of human activities (Malenovský et al., 2015; Hughes et al., 2020). Particularly, the genomic features were widely clarified in the Antarctic bacteria (Benaud et al., 2021), algae (Zhang et al., 2020b), and fishes (Kim et al., 2019), but the underpinnings remain unclear in the adaptation of mosses to extreme environments.

Mosses and lichens are the dominant floras in the coastal regions of Antarctica (Cannone et al., 2016; Bertini et al., 2021). The moss *Pohlia nutans* is abundant in the Fildes Peninsula and Victoria Land of Antarctica where water supply is plentiful (Skotnicki et al., 2002; Liu et al., 2019; Wang et al., 2020). They usually occur as small and isolated colonies with short shoots (1–2 cm length). Similar to other mosses in these regions, they appear to reproduce asexually through dispersal of vegetative propagules-small fragments of colonies (Skotnicki et al., 2002). Protonema proliferation plays an essential role in the asexual processes of differentiation and regeneration (Zhao et al., 2019). Since bryophytes commonly possess most key innovations of land plant evolution, the Antarctic psychrophilic mosses represent an emerging model system for studying responses and sensitivities to environmental changes (Convey and Peck, 2019). Here, the whole-genome sequencing, transcriptome and metabolome profiling, as well as comparative genomic analysis will enlarge our understanding of early land plant evolution and the adaptations of these basal land plants to the polar terrestrial environments.

MATERIALS AND METHODS

Plant Materials and Stress Treatments

The moss *P. nutans* isolate NO.L were gathered from the Fildes Peninsula of Antarctica (S62°13.260′, W58°57.291′), in March 2014. They were cultivated under conditions according to previous reports (Wang et al., 2020; Liu et al., 2021). For cold stress, mosses were placed under 0°C. For mock drought stress, mosses were treated with 20% PEG6000. Two Philips

TL20W/01RS narrowband UV-B tubes were used for UV-B light source as described previously (Wolf et al., 2010; Liu et al., 2021). Mosses without treatments or treated with sterile water were used as control group. The green parts of gametophytes were collected and used for genome and transcriptome sequencing, and UPLC-MS/MS analysis.

Estimation of Genome Size

Genomic DNA was prepared from the gametophytes of *P. nutans* using a modified cetyltrimethylammonium bromide (CTAB) method (Allen et al., 2006). DNA quality was detected using a Nanodrop 2000 Spectrophotometer (Thermo Fisher Scientific, United States) and an Agilent BioAnalyzer (Agilent, United States). DNA samples were broken into fragments with a length of 350 bp by an ultrasonic disruptor. DNA library was constructed and subjected to Paired-end sequencing using the Illumina Hiseq sequencing platform. Finally, clean reads were used to estimate the *P. nutans* genome size using *K*-mer frequency (*K*-mer = 17 bp; Zhang et al., 2020b).

Library Construction and Genome Sequencing

The paired-end genomic libraries were constructed using the Illumina TruSeq DNA PCR-Free Prep kit and sequenced with an Illumina HiSeq × 10 platform; the long inserts of SMRT Cell libraries were constructed and sequenced with a PacBio Sequel II instrument (Pacific Biosciences, CA, United States); the chromatin interaction mapping (Hi-C) libraries were constructed with 350–500 bp insertions and sequenced on an Illumina HiSeq platform × 10 platform. The quality of the Hi-C sequencing was assessed using the HiCUP pipeline (Wingett et al., 2015). Raw reads were trimmed to discard adaptor sequences, potential contaminants, and others.

Genome Assembly and Quality Assessment

We utilized the PacBio SMRT-sequencing and Hi-C-based scaffolding approaches with further polishing using Illumina short reads to assemble a high-quality of the *P. nutans* genome. In brief, 31.16 Gb of PacBio single-molecule long reads (average length, 13.73 kb) from SMRT sequencing were assembled into contigs using hifiasm (Cheng et al., 2021). The Illumina clean short reads obtained previously were aligned to the PacBio assembly using BWA (Li and Durbin, 2009) and Minimap2 (Li, 2018). The repetitive polishing was conducted using Pilon (v.1.22; Walker et al., 2014). The assembled contigs were submitted to BLAST search against the NCBI non-redundant (NR) nucleotide database to discard organellar DNA and prokaryotic sequences. Then, we combined the assembled contigs from PacBio sequencing and Illumina clean short reads into scaffolds using SSPACE (v.3.0) tool (Boetzer et al., 2010). Finally, 63.26 Gb of Hi-C sequencing clean reads were employed to cluster, orientate, and link the assembled scaffolds into 22 pseudo-chromosomes using HiCUP pipeline (Wingett et al., 2015). The completeness of genome assembly was evaluated by alignment with the plantae database of Benchmarking Universal Single-Copy Orthologs (v.3; Manni et al., 2021). Gene region completeness was assessed using transcriptome data of the Antarctic moss *P. nutans*.

Repeat Annotation

Repeat sequences annotation was carried out using a combined method of Repeatmasker, Proteinmask, and *de novo* that were described in previous publication (Zhang et al., 2020a). For the *de novo* approach, LTR_FINDER (Xu and Wang, 2007), PILER (Edgar and Myers, 2005), and RepeatModeler (Flynn et al., 2020) were used for TEs prediction and build the *de novo* repeat sequence libraries. Then, RepeatMasker was conducted to search for repeats in the *P. nutans* genome refer to the *de novo* repeat libraries as reference libraries. For the homology-based approach, TEs were identified using RepeatMasker and RepeatProteinMask with the integrate Repbase database of known repeat sequences and the *de novo* repeat sequence libraries (Zhang et al., 2020a). The final non-redundant repeat sequences were obtained by integrated together overlapping TEs from both *de novo* and homology-based predictions. The transposable elements (TEs) were calculated and summarized in the *P. nutans* genome.

Insertion Time of LTR Retrotransposon

Full-length LTR-RTs were further investigated from the *P. nutans* genome assembly using LTRharvest (Ellinghaus et al., 2008) and LTRretriever (Ou and Jiang, 2018). Default parameters were used except for -minlenltr 100 -maxlenltr 7000 -mintsd 4 -maxtsd 6 -motif TGCA -seqids yes. *Copia* and *Gyspy* in all non-redundant LRT-RTs file were used for tree construction. Full-length LTRs were aligned with MAFFT (Katoh and Standley, 2013). The insertion time of LRT-RTs was calculated by LTRretriever.

Whole-Genome Duplication Incidents

To find the potential whole-genome duplications (WGD), the *P. nutans* genome was self-aligned or aligned with *P. patens* to identify duplicated gene pairs. All *Ks* (rate of synonymous substitutions) distributions were calculated using wgd command-line tool (v.3.0; Zwaenepoel and Van De Peer, 2019). All *Ks* values ≤0.1 were removed to avoid the incorporation of allelic and/or splice variants (Lang et al., 2018). A rough dating procedure of the WGD event were conducted based on the observed sequence divergence using the *Ks* of 0.11 and a referenced substitution rate (r) of 9.4×10^{-9} per synonymous site per year in *P. patens* (Rensing et al., 2008). The time gene insertion (T) was calculated using the formula $T = Ks/2r$.

Gene Prediction and Functional Annotation

Protein-coding genes were predicted from the *P. nutans* genome using three strategies: (1) *de novo* prediction, (2) homology-based prediction, and (3) RNA-sequencing annotation. For *de novo* prediction, AUGUSTUS (v.2.5.5) and GlimmerHMM (v3.0.1) were used to detect genes (Majoros et al., 2004; Stanke et al., 2006). For homology-based prediction, proteins of five well-annotated species (*P. patens*, *Amborella trichopoda*, *Ananas comosus*, *Abrus precatorius*) were aligned to the *P. nutans* genome using TBLASTN, with a threshold E-value $\leq 1 \times 10^{-5}$. The resultant alignments were

analyzed using Genewise (v.2.4.1; Birney et al., 2004). The *de novo* assembly and homologues results were incorporated in MAKER to generate a consensus dataset (Holt and Yandell, 2011). To supplement the dataset, we aligned the RNA-sequencing data to the genome using TopHat (v2.1.1), and the alignments were used as input for Cufflinks (v.2.2.1; Trapnell et al., 2009, 2012). We merged the MAKER consensus and transcripts to generate the final dataset. Gene functional annotation was conducted according to best hit of BLASTP (*E*-value ≤1E-05) against SwissProt, Non-Redundant, TrEMBL, and KEGG databases.

Evolution Analysis

To build the dataset for gene-family clustering, we obtained the protein-coding genes from the genomes of 14 green plants, containing five bryophytes (the Antarctic moss *P. nutans*, the moss model plant *P. patens*, the liverwort model plant *Marchantia polymorpha*, the hornwort model plant *Anthoceros angustus*, and the moss *Ceratodon purpureus*). We only used the longest transcript of each gene. Orthogroups or gene families were detected after conducting an all-against-all BLASTP search using OrthoMCL (v.2.0, threshold *E*-value of 1.0×10^{-5}; Li et al., 2003). A venn diagram was present by comparison of gene families identified from different plants. The syntenic blocks were identified using MCScanX or jcvi software (threshold *E*-value $\leq 1.0 \times 10^{-5}$; Wang et al., 2012).

We constructed species phylogenetic tree using single-copy orthologs. Since the *P. nutans* genome underwent a recent WGD events with few single-copy orthologs, we therefore used half of the paired chromosomes of *P. nutans* for species evolution analyses. The single-copy orthologs were aligned using MAFFT and finally joined into the super alignment matrix (Katoh and Standley, 2013). Then, the maximum likelihood phylogenetic analyses were conducted by RAxML (v.7.2.3; Stamatakis, 2006). Divergence times were calculated using MCMCTree in PAML v4.9 (Yang, 2007). Calibration points on the tree branches were inferred from Timetree.[1] The time-calibrated phylogeny was mainly congruent with previous phylogenetic analyses (Zhang et al., 2020a). Gene-family sizes were estimated using the OrthoMCL program. The gene family was filtered by threshold of variance/mean < 10 and missing rate < 25%. The gene family expansion or contraction were calculated using CAFÉ (v2.1) based on the maximum likelihood model (*q*-value ≤0.05; De Bie et al., 2006). The expanded gene families were analyzed by KEGG pathway enrichment.

Transcription Factor Annotation

We utilized the plant transcription factors (TFs) prediction program iTAK (v.1.7) to identify TFs (Zheng et al., 2016).

Identification and Phylogenetic Analysis of Gene Families

Hmmsearch program was used to identify the L-phenylalanine ammonia-lyase (PAL), chalcone synthase (CHS), 2-oxoglutarate-dependent dioxygenase (2-OGD), DNA photolyase, and other family proteins from five bryophyte genomes with default parameters of --cut_tc. Sequence alignments were conducted using MAFFT (Katoh and Standley, 2013). Phylogenetic trees were constructed with IQ-TREE software (v1.6.12) from amino-acids sequences (with maximum-likelihood method, 1,000 replicates, and LG + F + R7 model; Nguyen et al., 2015).

Transcriptome Sequencing

Mosses were treated with UV-B light, Cold (0°C), and drought (20% PEG6000). RNA isolation and transcriptome sequencing were conducted following the procedure described previously (Li et al., 2021). The gene expression levels were estimated with RPKM (Reads Per Kilobase per Million mapped reads). The $|\log_2(\text{Treat/Control})| > 1$ and FDR (*q*-value, Corrected *p*-value) < 0.05 were used as the threshold to discriminate the differentially expressed genes (DEGs).

Metabolomics Profiling and Analyses

Mosses were treated with UV-B light for 3 days. The moss gametophytes from UV-B radiation groups (i.e., UV-B) and control groups (i.e., CG) were cut and used for metabolite analysis. For cold stress, mosses were placed under 0°C for 60 h. The metabolite extraction, qualitative, and quantitative analysis were performed in accordance with previously described methods (Li et al., 2021; Wang et al., 2021b).

Statistical Data Analysis

Data were statistically compared between different groups and the statistical significance were calculated using Student's *t* test ($^*p < 0.05$, $^{**}p < 0.01$). For multivariate data analysis of metabolome, principal component analysis (PCA), hierarchical clustering analysis (HCA), and orthogonal projections to latent structure-discriminant analysis (OPLS-DA) were performed according to earlier publications (Nyamundanda et al., 2010; Li et al., 2021). Metabolites with $\log_2 |\text{fold change}| \geq 1$ and VIP ≥ 1 were considered as differentially changed metabolites between two groups.

RESULTS

Genome Assembly and Annotation

The moss gametophytes were used for genome and transcriptome sequencing (**Figure 1A**). To obtain a high-quality genome of *P. nutans*, we used a combination of Illumina short-read and PacBio Sequel II HiFi-read sequencing approaches. The estimated genome size was 708 Mb, and the proportion of repeat sequences 62.02%, based on the distributions of *K*-mer frequency. We employed a hierarchical assembly approach using 31.16 Gb (44.01-fold coverage) of PacBio HiFi reads, 83.00 Gb (131.36-fold coverage) of Illumina short-reads, and 63.26 Gb (90.19-fold coverage) of Hi-C sequencing data. In the process of Hi-C-assisted genome assembly, the original contigs were sorted according to the interaction map and 98.16% of the contig length could be mapped to the chromosome. Finally, we obtained an optimized assembly of 698.20 Mb with 22 chromosomes

[1]http://www.timetree.org/

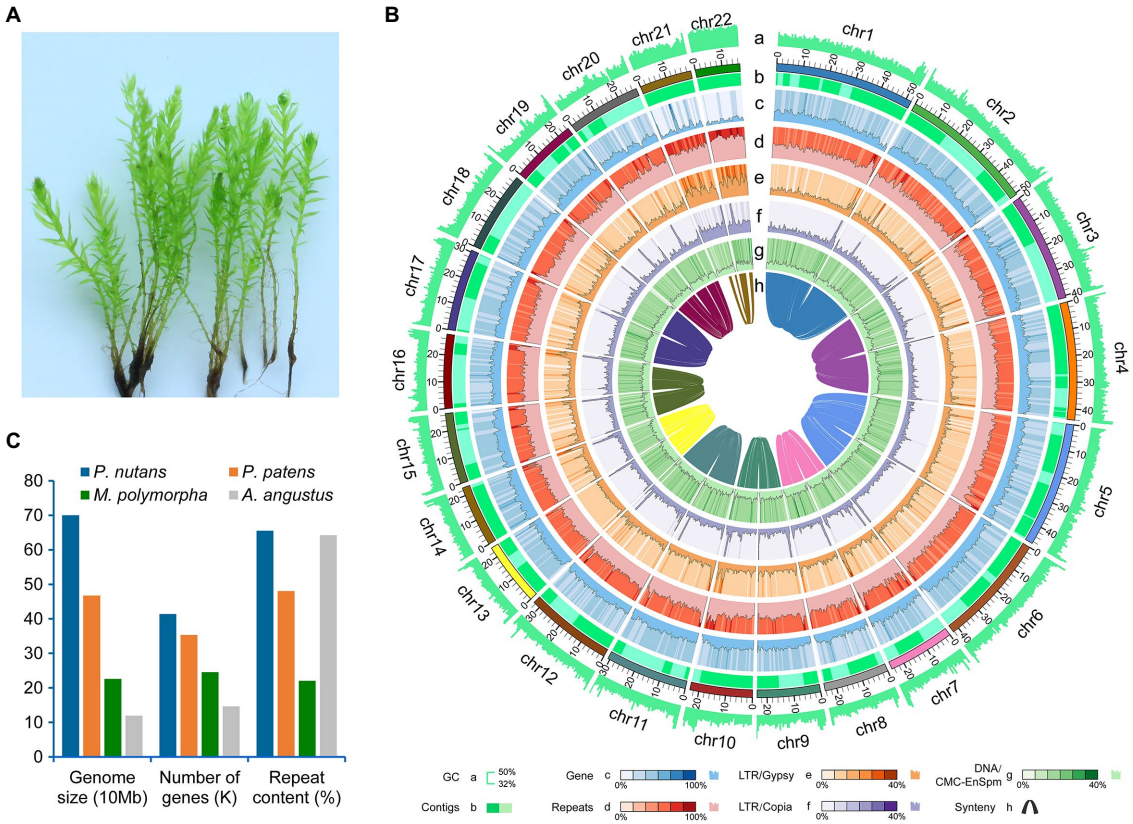


FIGURE 1 | Genomic structure of the Antarctic moss *Pohlia nutans*. **(A)** Photo of *P. nutans* under normal condition. **(B)** The landscape of genome assembly and annotation of *P. nutans*. Tracks from outside to inside correspond to: a, GC content; b, contigs density; c, gene density; d, repeat density; e, LTR/*Gypsy* content; f, LTR/*Copia* content; g, DNA/CMC-EnSpm; and h, syntonic relationship. **(C)** The genome size, gene content, and repeat content of four bryophyte genomes (the Antarctic moss *P. nutans*, the moss model plant *Physcomitrella patens*, the liverwort model plant *Marchantia polymorpha*, the hornwort model plant *Anthoceros angustus*).

TABLE 1 | Assembly and annotation characteristics of the *Pohlia nutans* genome.

Genome assembly features	
Genome size (Mb)	698.20
Max scaffold length (bp)	52,153,166
Scaffold N50 (bp)	30,709,574
Total contig length (bp)	752,686,138
Max contig length (bp)	26,973,851
Contig N50 (bp)	13,886,535
GC ratio	41.53%
Genome annotation features	
Number of protein coding genes	40,905
Gene density (genes/100 kb)	5.90
Average gene or CDS length (bp)	5,021/1,353
Average exons per gene	6.28
Average exon length (bp)	378.83
Average intron length (bp)	500.57
Repeat content	65.53%

and 313 scaffolds. The contig N50 length was 796.64 kb and the scaffold N50 length was 1.09 Mb (**Figure 1B**; **Table 1**).

A total of 40,905 protein-coding genes were identified by combining three annotation strategies, with an average coding-sequence length of 1,353.45 bp and an average of 6.28 exons per gene (**Table 1**). *Pohlia nutans* possessed the largest genome size and maximum gene number among the published bryophyte genomes (**Figure 1C**). About 93.95% of the protein-coding genes had their best homologs on plant sequences from NCBI non-redundant database, and 94.71% were functionally annotated through Swissprot, TrEMBL, Pfam, GO and KEGG databases. The genome assembly captured 86.0% of the Benchmarking Universal Single-Copy Orthologs (BUSCO) plantae dataset with 83.9% complete gene models plus 2.1% fragmented gene models (**Supplementary Table 1**). To assess assembly accuracy, we mapped sequencing reads to the assembled genome, with mapping rates of 97.67 and 98.86% for the paired-end and long reads, respectively (**Supplementary Table 2**). In addition, the Illumina paired-end reads from transcriptome data were used to evaluate assembly accuracy, receiving a 96.75% mapping rate (**Supplementary Figures 1A,B**). Besides protein-coding genes, non-coding RNA sequences were annotated in the *P. nutans* genome, including 3,740 transfer RNAs, 55 micro RNAs, 545 ribosomal RNAs, 999 small nuclear RNAs (**Supplementary Table 3**). These data showed that the integrity and correctness of the *P. nutans* genome assembly were high, guaranteeing the reliability of following genomic analyses.

We annotated repeat sequences after the genome assembly using multiple methods. Repeat sequences comprised 65.53% of the *P. nutans* assembled genome. The transposable elements (TEs) occupied 63.79% of the *P. nutans* genome assembly (**Supplementary Table 4**). Among the TEs annotation, long terminal repeats (LTRs) accounted for 28.77% of the genome assembly, while DNA transposons occupied 23.68%. The number of LTR-RTs in *P. nutans* genome was about 3.18-fold greater than that in *P. patens* genome (**Supplementary Figure 2A**). The superfamilies *Copia* and *Gypsy* were two main types of LTR-RTs, and they each accounted for 9.67 and 8.95%, respectively, of the *P. nutans* genome assembly. Particularly, *P. nutans* and hornworts (*A. angustus*) held a relative higher proportion of *Copia* type of LTR-RTs, whereas *P. patens* and liverworts (*M. polymorpha*) possessed more *Gypsy* type of LTR-RTs (**Supplementary Figure 2B**). *Pohlia nutans* underwent a most recent LTR-RT (*Copia* and *Gypsy*) amplification, followed by *P. patens*, *A. angustus*, and *M. polymorpha* in order (**Figure 2A**). Nevertheless, *Gypsy* comprised merely 1.64% of *A. angustus* genome sequences that it did not show in the plot. The density decreased between 1.0 to 6.0 Ma, which reflected element deterioration that was hard to discriminate these elements. These results indicated that LTR-RT amplification were likely contributed to the *P. nutans* genome size expansion.

Land plants normally harbor genomic signatures of whole-genome duplications (WGDs; that is, polyploidy). In order to investigate the potential WGD event, we detected the distributions of synonymous substitutions per synonymous site (*Ks*) values for paralogs to identify the amplification of duplicate genes from WGDs. This method was under the assumption that synonymous substitutions between duplicate genes accrue at a relatively constant rate. Consequently, *Ks* distribution of the detected pairs formed one peaks ($Ks = 0.11$) demonstrating that the *P. nutans* genome presented the evidence of having undergone a recent WGD incident. Particularly, the WGD events of the *P. nutans* genome occurred more recently about at 5.85 Ma (**Figure 2B**). Thus, we suggested that WGD incident was another driving force for genome size expansion of *P. nutans* even more than TEs amplification.

Comparative Genome Analysis

To evaluate the gene conservation or loss after polyploidization or speciation, we conducted the alignment of the homology of pairs of genes in the chromosomal fragments "synteny blocks." We investigated syntenic blocks within the *P. nutans* genome or between *P. nutans* and other moss genomes using all-versus-all BLASTP alignments. The synteny analysis within the *P. nutans* genome detected a total of 1,289 syntenic blocks with 11,030 gene pairs. Generally, two adjacent chromosomes in *P. nutans* genome showed strong syntenic relationships (**Supplementary Figure 2C**). The data supported the fact that the *P. nutans* genome undergo a recent WGD event. The intergenomic synteny analysis between the genomes of *P. nutans* and the moss *C. purpureus* identified a total of 1,565 syntenic blocks with 16,852 gene pairs. Almost each of the *C. purpureus* chromosomes was highly syntenic with a pair of the *P. nutans* chromosomes (**Figure 2C**). Several exceptions were that chr7 or chr10 in *C. purpureus* exhibited syntenies with the adjacent chr7 and chr8 in the *P. nutans*, respectively; chr1 or chr11 in *C. purpureus* corresponds to the adjacent chr1 and chr2 in *P. nutans*, respectively. Meanwhile, the intergenomic synteny analysis between the genomes *P. nutans* and *P. patens* showed that there were relatively less syntenic blocks (i.e., 1,299) and gene pairs (i.e., 10,493; **Figure 2C**). Thus, the intergenomic syntenies between the genomes *P. nutans* and *C. purpureus* were stronger than those between the genomes of *P. nutans* and *P. patens*. In addition, the syntenic blocks between the genomes of *P. nutans* and *C. purpureus* were not congruent with those between the genomes *P. nutans* and *P. patens* (**Supplementary Figure 2D**).

For sequence similarity-based clustering of homologues, we conducted the analyses with the predicted proteomes among *P. nutans* and other 13 green plants. These representative green plants included bryophytes (*P. patens*, *C. purpureus*, *M. polymorpha*, and *A. angustus*), algae (*Chlamydomonas reinhardtii*, *Chlorella variabilis*, and *Klebsormidium nitens*), fern (*Selaginella Moellendorfii*), monocots (*Oryza sativa* and *Vitis vinifera*), and eudicots (*Arabidopsis thaliana*, *A. trichopoda*, and *Clematoclethra variabilis*). We identified 37,125 genes distributed among 12,650 gene families in *P. nutans*. A total of 5,357 gene families were shared by four species *P. nutans*, *P. patens*, *A. angustus*, and *K. nitens* genomes, while a total of 3,523 gene families collectively presented in above four species and *C. reinhardtii* (**Supplementary Figures 3A,B**). In addition, 2,184 gene families were shared by 14 represent green plants with different evolutionary status, whereas 1,047 gene families including 2,790 genes, which appeared to be unique to *P. nutans* (**Figure 2D**). These lineage-specific family genes were markedly enriched in various biosynthetic categories (e.g., phenylpropanoid biosynthesis, flavonoid biosynthesis, carotenoid biosynthesis, and other biosynthesis), likely facilitating neofunctionalization (**Supplementary Figure 3C**). *Pohlia nutans* genome possessed the largest number of protein-coding genes, but few amount of single-copy orthologs (**Supplementary Figure 3D**). We therefore use half of the paired chromosomes of *P. nutans* for species evolutionary analyses. Finally, the phylogenetic position of *P. nutans* was determined using 51 single-copy orthologs (**Figure 3A**). *Pohlia nutans* was clustered within bryophyte lineages, sister to the temperate *C. purpureus* species. Our molecular dating analyses indicated that the speciation of bryophyte lineages occurred at 511.1 Ma and the divergent event between *P. nutans* and *C. purpureus* occurred at 98.0 Ma (**Figure 3A**).

We assessed the gene family expansion or contraction by using the homolog matrix of orthogroups to calculate ancestral and lineage-specific genes on the phylogenetic tree. The genome of *P. nutans* covered 7,807 markedly expanded gene families and 635 notably contracted gene families ($p < 0.05$), respectively (**Figure 3A**). Particularly, *P. nutans* harbored the largest number of expanded gene families among 14 green plants. Furthermore, KEGG pathway enrichment analyses revealed that the expanded gene families participated in biosynthesis of metabolites (i.e., phenylpropanoid biosynthesis, tropane, piperidine, and pyridine alkaloid biosynthesis, alpha-Linolenic acid metabolism, and fatty acid elongation), environmental adaptation (i.e., plant hormone signal transduction, and plant-pathogen interaction; **Figure 3B**).

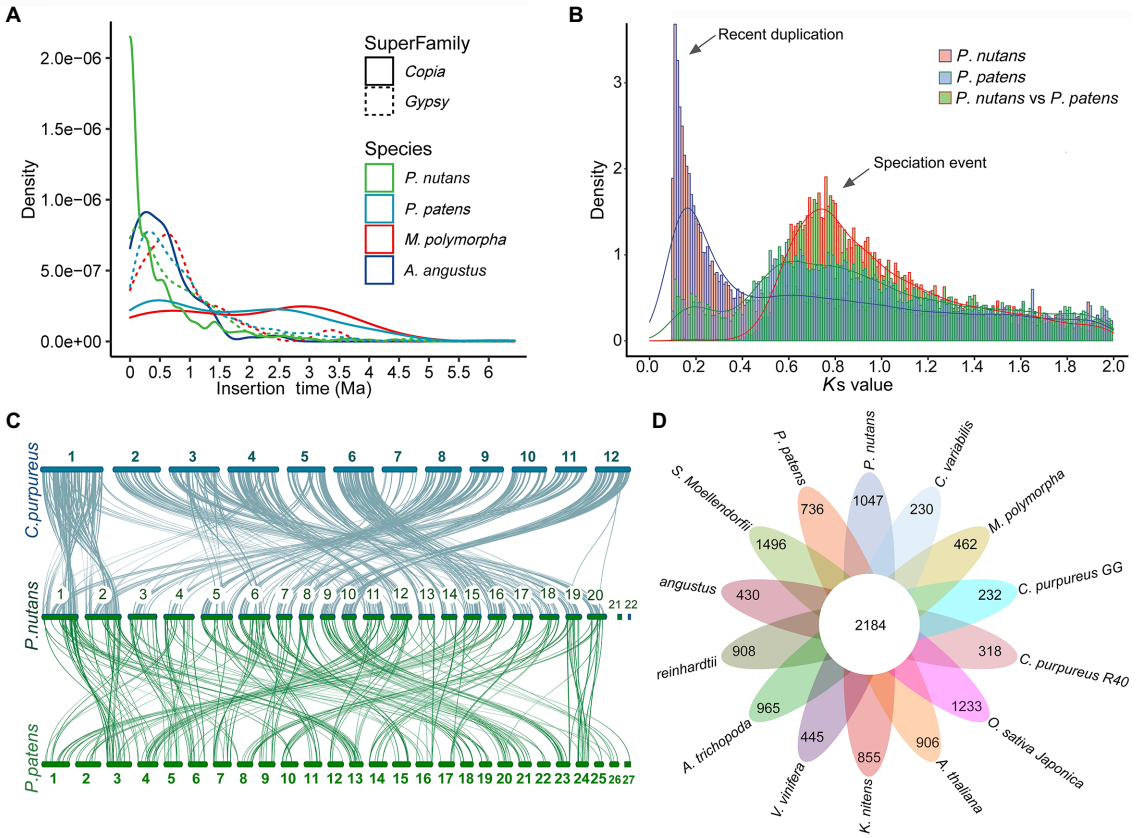


FIGURE 2 | Evolutionary and comparative genome analyses of the Antarctic moss *Pohlia nutans*. **(A)** LTR-RT (*Copia* and *Gypsy*) amplification analyses of four bryophytes (the Antarctic moss *P. nutans*, the moss model plant *Physcomitrella patens*, the liverwort model plant *Marchantia polymorpha*, the hornwort model plant *Anthoceros angustus*). **(B)** Whole-genome duplication and speciation events identified between *P. nutans* and *P. patens*. **(C)** The intergenomic synteny analysis between the genomes *P. nutans*, *Ceratodon purpureus*, and *P. patens*. **(D)** Comparison of the number of gene families identified by OrthoMCL. Several representative plants were selected for evolutionary analysis such as bryophytes (*P. nutans*, *P. patens*, *C. purpureus*, *M. polymorpha*, and *A. angustus*), algae (*Chlamydomonas reinhardtii*, *Chlorella variabilis*, and *Klebsormidium nitens*), fern (*Selaginella Moellendorfii*), monocots (*Oryza sativa* and *Vitis vinifera*), eudicots (*Arabidopsis thaliana*, *A. trichopoda*, and *Clematoclethra variabilis*). The Venn diagram shows the shared and unique gene families in green plants of different evolutionary status.

The *P. nutans* genome comprised of 1,328 transcription factors (TFs) forming 63 families, which was substantially larger than these in other four bryophyte genomes (**Figure 3C**). In addition, the *P. nutans* genome also harbored more stress-related TFs like AP2, bHLH, bZIP, MYB, NAC, Trihelix, and WRKY (**Figure 3C**).

Evolutionary Innovations of Stress-Related Genes in Extreme Terrestrial Habitats

Although the Antarctic Peninsula is warming and the East Antarctica is likely to become dryness, it is obviously that strong ultraviolet radiation is the most significant adverse result of climate change and human activities. Therefore, we will focus on the molecular mechanism of *P. nutans* in response to strong UV-B radiation. We found that the gene family of DNA photolyase was highly expended in *P. nutans* when compared with other four bryophytes. There were 16 of DNA photolyase homologs in the *P. nutans* genome, nine in the *C. purpureus* genome, eight in the *P. patens* genome, eight in the *M. polymorpha* genome, and six in the *A. angustus* genome (**Figure 4A**). In addition, we performed the transcriptome sequencing of *P. nutans* under UV-B radiation, cold stress, and drought stress. Six DNA photolyase genes were significantly upregulated under UV-B radiation and cold stress (**Figure 4E**). We also found that gene families encoding antioxidant and detoxicant enzymes were expanded in *P. nutans*, such as glutaredoxin, glutathione S-transferase, and protein detoxification of MATE family (**Figures 4C,D**). They were also differentially expressed under UV-B radiation, cold and drought stresses (**Figures 4F–H**). Thus, the DNA photolyases and antioxidant enzymes could provide the Antarctic moss *P. nutans* with DNA repair machinery and ROS-scavenging system against environmental stresses.

Integrated Transcriptomic and Metabolomic Analysis Highlight the Role of Flavonoid Biosynthesis Under UV-B Light

We found the gene family encoding phenylalanine ammonia-lyase (PAL), catalyzing the first committed step of the phenylpropanoid pathway, was highly expanded in *P. nutans* genome. Phylogenetic analysis showed that most of PAL homologs from mosses were clustered together (**Supplementary Figure 4A**).

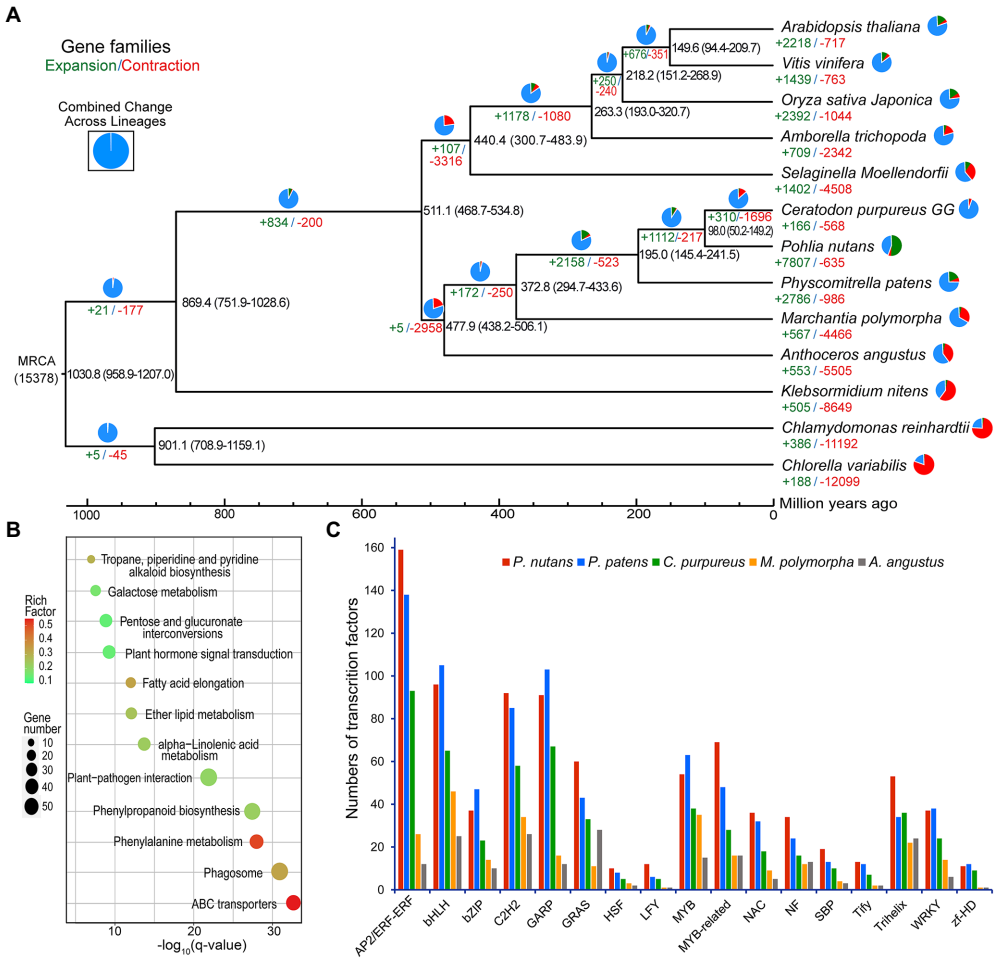


FIGURE 3 | Evolutionary analysis and gene-family expansion and contraction among several green plants. **(A)** The number of significantly expanded and contracted gene families was shown on a time-calibrated phylogenetic tree. The numbers of expanded (green) and contracted (red) gene families were identified by CAFE software and shown above the branches. The estimated age of speciation was indicated on the branches with blue numbers (million years ago, Ma). **(B)** Scatterplot of KEGG enrichment of the expanded gene families in *Pohlia nutans*. **(C)** Numbers of selected stress-related transcription factor genes in five bryophyte species.

Two PAL homologs from hornwort and four moss PAL homologs were close to the tree root, which seemed to be more ancestral. In addition, we performed the transcriptome sequencing of *P. nutans* under UV-B radiation (**Figure 5A**). We found that 14 PAL genes were significantly upregulated in *P. nutans* under UV-B radiation (**Supplementary Figure 4B**). The flavonoid biosynthesis enzymes were expanded in the *P. nutans* genome, including chalcone synthase (CHS) and 2-oxoglutarate-dependent dioxygenase (2-OGD; **Figure 5B**). 2-OGD family proteins like flavonol synthase, flavone synthase, flavanone-3-hydroxylase, and anthocyanidin synthase harbored the conserved domains of 2OG-FeII_Oxy and DIOX_N. Transcriptome sequencing showed that *CHS* and *2-OGD* were markedly upregulated under UV-B radiation. In higher plants, the transcription factor R2R3-MYB is responsible for regulating the transcription of flavonoid biosynthesis enzymes (Albert et al., 2018). Transcriptome sequencing also showed that six *R2R3-MYB* were markedly upregulated in *P. nutans* under UV-B radiation (**Figure 5C**).

To further reveal the responses of *P. nutans* to UV-B radiation, we used the UPLC-MS/MS method to detect the metabolites. A total of 415 metabolites were identified (**Supplementary Table 5**). Flavonoids accounted for 9.88% of the total compounds. In addition, anthocyanins were also identified in *P. nutans*, including cyanidin 3-O-(6″-Malonylglucoside), malvidin 3-O-galactoside, and peonidin O-hexoside. PCA score plot, heatmap cluster, and the supervised OPLS-DA analyses indicated that the data were stable and reliable, and significant differences were detected in metabolic phenotypes after UV-B radiation (**Supplementary Figures 5A–C**). Using thresholds of $|\log_2\text{Foldchange}| \geq 1$ and VIP (variable importance in project, VIP) ≥ 1, a total of 97 significantly changed metabolites (SCMs) were detected under UV-B radiation. Of them, 80 metabolites were significantly upregulated and 17 metabolites were markedly downregulated (**Figure 5D**; **Supplementary Table 6**). The top 20 SCMs in UV-B group compared to control group according to the values of $\log_2$(Fold change) are shown in **Figure 5E**.

Flavonoids accounted for 37.11% of the total SCMs (**Supplementary Figure 5D**). Gallic acid was the most

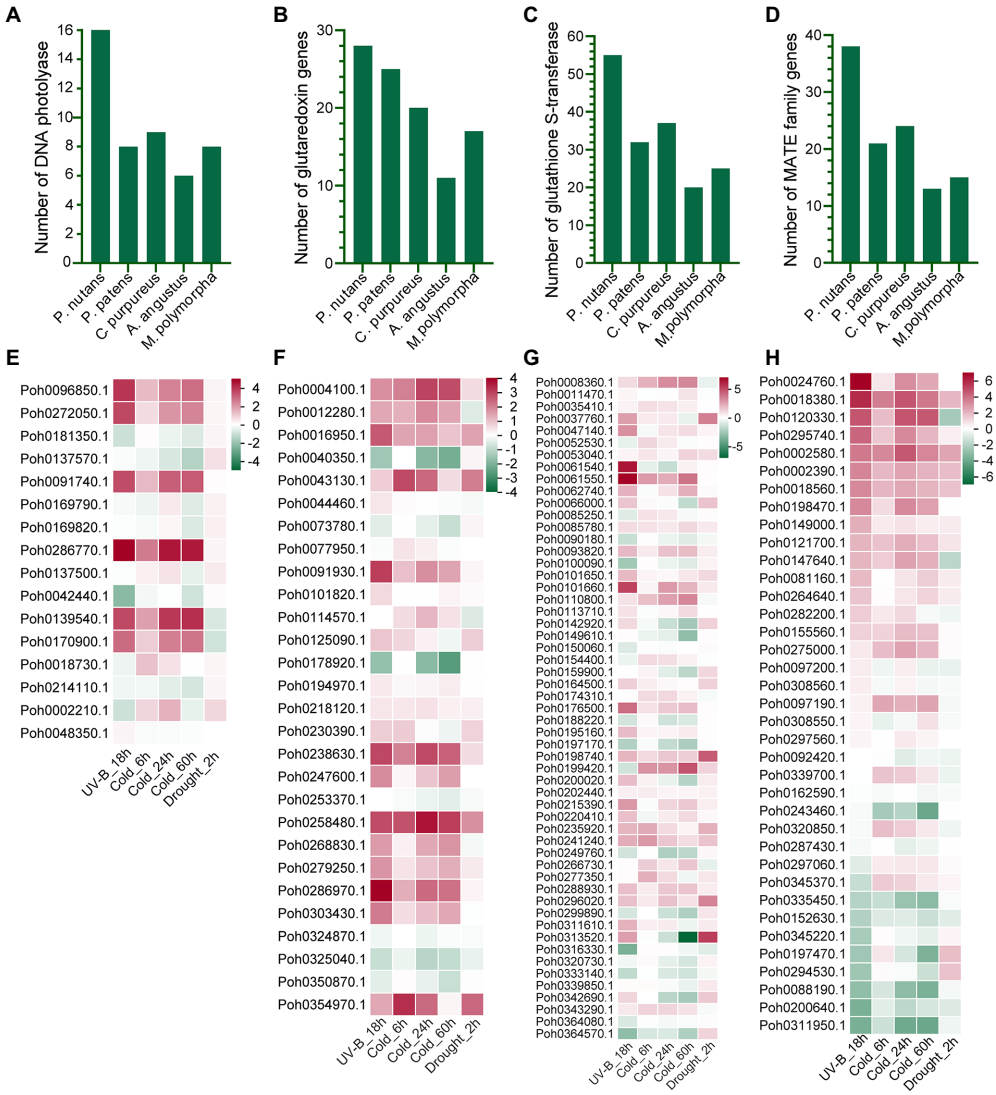


FIGURE 4 | The gene families related to oxidative stress were expanded and upregulated under various abiotic stresses. Gene families of DNA photolyase **(A)**, glutaredoxin **(B)**, glutathione S-transferase **(C)**, and protein detoxification with of MATE domain **(D)** were expanded in the *P. nutans* genome. Genes encoding DNA photolyase **(E)**, glutaredoxin **(F)**, glutathione S-transferase **(G)**, and protein detoxification with MATE domain **(H)** were markedly upregulated under UV-B radiation, cold, and drought stresses detected by transcriptome sequencing.

accumulated metabolite with $\log_2$(Fold change) 14.66 and VIP score 1.50. Caffeoyl-p-coumaroyltartaric acid, a kind of phenolic acids, was the second significantly changed metabolite with $\log_2$(Fold change) 12.11 and VIP score 1.50. In addition, cyanidin 3-O-(6″-malonylglucoside), a kind of anthocyanins, was also markedly accumulated metabolite with $\log_2$(Fold change) 11.02 and VIP score 1.5, exhibiting significant antioxidant activity (Rahman et al., 2021). KEGG pathway analysis showed that the SCMs were mainly involved in flavonoid biosynthesis pathways (**Figure 5F**). Notably, DEGs and SCMs were both abundantly enriched in flavonoid biosynthesis pathway under UV-B radiation. These metabolites might facilitate mosses in resisting the extra ROS damages. Therefore, the gene-family expansions, upregulated mRNA levels, and accumulated metabolites in flavonoid biosynthesis pathway might represent one of molecular adaptations to life in the Antarctic strong UV-B environments.

DISCUSSION

The Antarctic Moss Genome Provides New Resource for Evolutionary and Functional Studies

Antarctic extremely environmental pressures including higher UV-B radiation, freezing, and extreme dryness, lower nutrient supply and strong wind severely influence the reproduction and distribution of terrestrial plants (Convey and Peck, 2019; Zhang et al., 2020b). Mosses and lichens dominate the Antarctic

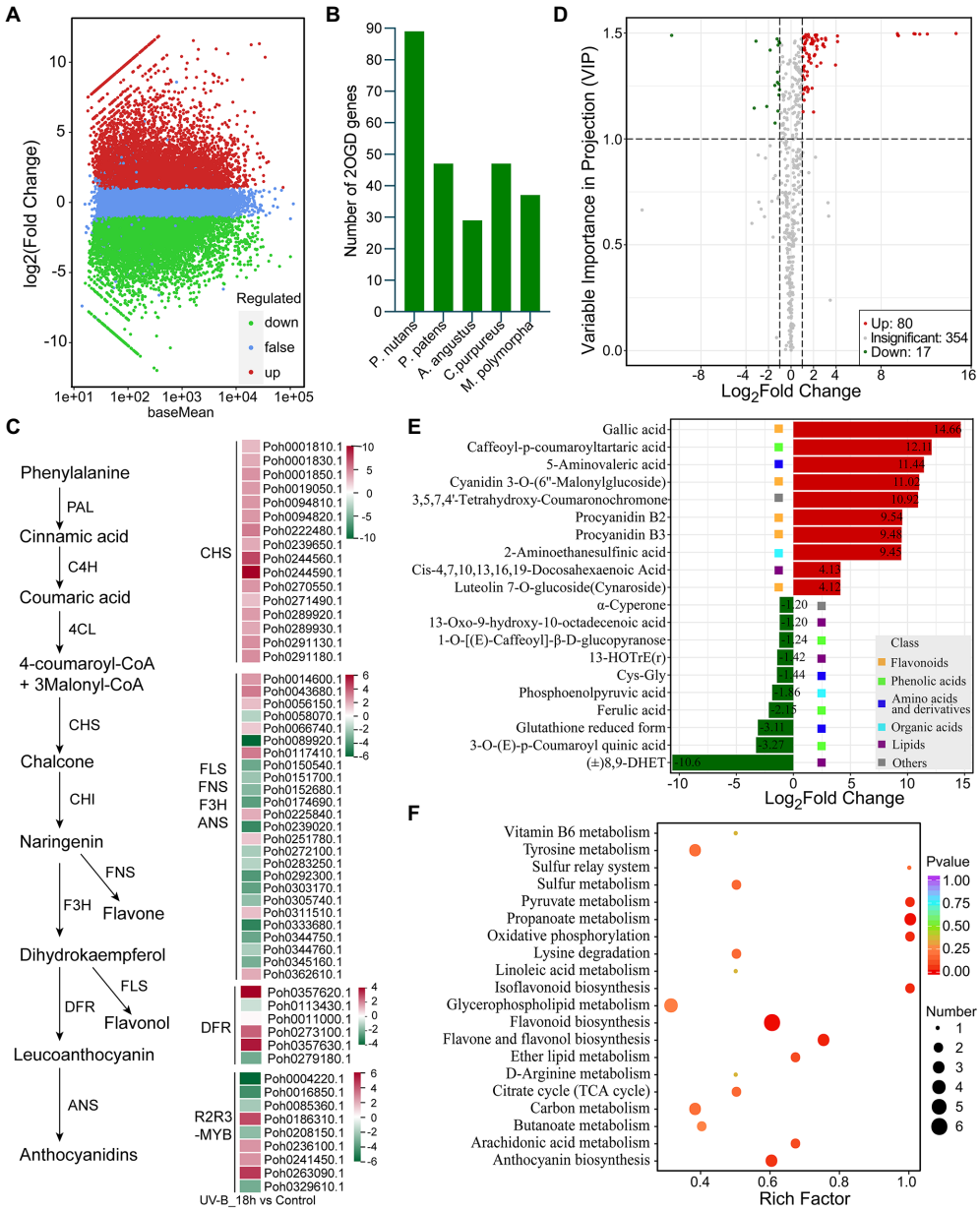


FIGURE 5 | Integrated multi-omic analyses reveal the role of flavonoid pathway under UV-B light. **(A)** The volcano plot showing the DEGs between UV-B radiation group and control group. The *Y*-axis indicates fold change of gene expression (threshold, $|\log_2(\text{Treat/Control})| > 1$), while the *X*-axis means the statistically significant level (threshold, *q*-value < 0.05). **(B)** The gene family of 2-oxoglutarate-dependent dioxygenase (2-OGD) was expanded in the *Pohlia nutans* genome. **(C)** Flavonoid biosynthesis pathway genes were markedly upregulated under UV-B radiation detected by transcriptome sequencing. **(D)** The volcano plot showing the contents of metabolites and the statistical significance. Each point represents a metabolite. Horizontal ordinate indicates the fold change of metabolites between two groups, while VIP value represents significant difference in statistical analysis. **(E)** The fold change of the top 20 significantly changed metabolites (SCMs) between two groups. **(F)** Statistics of KEGG enrichment for the SCMs.

flora and are restricted to sparse ice-free areas (Skotnicki et al., 2002). They are normally at the survival limitations. These native terrestrial plants have adapted to the extreme conditions over many millions of years (Convey and Peck, 2019). They are now threatened by global climate changes and the direct impacts of human activities (Malenovský et al., 2015; Amesbury et al., 2017; Robinson et al., 2018), as well as invasions of non-native species (Hughes et al., 2020). Here, we complete a high-quality genome sequence of the Antarctic moss *P. nutans* with 699.88 Mb with 22 chromosomes and 362 scaffolds (**Figure 1**). The distribution of *K*-mer frequency was widely used for the estimation of genome size (Zhang et al., 2020b). The optimized genome assembly was largely congruent with the *K*-mer analysis. The genomes of representative species such as mosses (*P. patens* and *Syntrichia caninervis*), liverworts (*M. polymorpha*), and hornworts

(*A. angustus*) have been sequenced and offer new perspectives for understanding the molecular adaptations to terrestrial life (Bowman et al., 2017; Lang et al., 2018; Zhang et al., 2020a; Silva et al., 2021).

Compared with the genome size of other bryophytes, the *P. nutans* genome was substantially larger with a higher proportion of LTR-RTs and more protein encoding genes (**Figure 1C**; **Table 1**). Repeat sequences comprised 65.53% of the *P. nutans* assembled autosomal genome (**Table 1**; **Supplementary Table 4**), similar to that of *A. angustus* (64.21%; Zhang et al., 2020a) and the Antarctic ice algae ICE-L (63.78%; Zhang et al., 2020b), but a remarkable higher than that of *M. polymorpha* (22%; Bowman et al., 2017) and *P. patens* (48%; Rensing et al., 2008). TEs comprised 63.79% of the *P. nutans* genome assembly, whereas they accounted for 57% of the *P. patens* genome (Lang et al., 2018). Notably, LTR-RTs in *P. nutans* genome was the most abundant among the four bryophyte genomes (**Figure 2A**; **Supplementary Figure 2A**). We therefore proposed that the amplification of LTR-RTs in *P. nutans* likely enlarges its genome size and may play an essential role in the adaptations to extreme environments.

High proportion of repetitive sequences usually causes a large barrier to reliable assembly of genomes from short-read sequences (Uauy, 2017). Consequently, the *P. nutans* genome assembly captured 86.0% of the BUSCO plantae dataset with 83.9% complete gene models plus 2.1% fragmented gene models (Manni et al., 2021), compared with 89.64, 93.51, and 92.15% captured in *A. angustus* (Zhang et al., 2020a), *P. patens* (Rensing et al., 2008; Lang et al., 2018), and *M. polymorpha* (Bowman et al., 2017), respectively (**Supplementary Table 1**). Similarly, in Antarctic green alga *Chlamydomonas* sp. ICE-L, the genome assembly captured 83.7% of the BUSCO datasets (Zhang et al., 2020b). In Antarctic green alga *Chlamydomonas* sp. UWO241, the genome assembly contains 16,325 protein-coding genes, capturing 85% of the BUSCO datasets (Zhang et al., 2021). We speculate that organisms living in extreme conditions generated multiple new metabolic pathways and functional genes, but that several metabolic pathways are missing, resulting in a lower BUSCO score.

Bryophytes comprise of mosses, hornworts, and liverworts that are the three early diverging extant land plant lineages (Bowman et al., 2017). Given that they emerged from the early lineage in the divergence of land plants, they offer the critical clues to investigate the early land plant evolution (Zhang et al., 2020a). Thus, this high-quality assembled genome will provide an important resource for evolutionary and functional studies in mosses, particularly those associated with UV-B radiation, cold, and drought resistance studies.

The Recent WGD Was a Major Fountain for Gene-Family Expansion and Function Diversification

Whole-genome duplication incident is an important driving force for speciation, with approximately 15% of angiosperm and 31% of fern speciation events involving this process (Wood et al., 2009; Hong et al., 2021). In moss model plant *P. patens*, *Ks*-based analysis indicated two WGD incidents dating back to 27–35 and 40–48 Ma, respectively (Lang et al., 2018). Like that of *P. patens*, the *P. nutans* genome presented the evidence of having undergone a WGD event, which occurred markedly earlier than that of *P. patens* (**Figure 2B**). In *P. patens* genome, the *Ks* values of the structure-based peaks are between 0.75–0.90 (older WGD1) and 0.50–0.65 (younger WGD2; Lang et al., 2018). *Ks* frequency analysis between *P. nutans* and *P. patens* identified a notable *Ks* peak at about 0.75 of duplicated syntenic pairs, indicating that the older WGD1 event was ancestrally shared by the two genomes and subsequently the speciation event was occurred (**Figure 2B**). The syntenic blocks within the *P. nutans* genome showed that most of two adjacent chromosomes had strong syntenic relationships (**Figure 1B**; **Supplementary Figure 2C**). In haploid *P. nutans*, chromosome number $n = 11$, 22, 33 are the most common and have been recorded from many countries, including Estonia ($n = 11$; Fetisova and Visotskaya, 1970), Australia ($n = 22$; Ramsay and Spence, 1996), and British and Irish (Smith and Newton, 1967). Given the syntenic relationships (**Supplementary Figure 2C**), one of the most likely assumptions for the extant chromosome number of the Antarctic moss *P. nutans* was the duplication of 11 ancestral chromosomes by WGD event, followed by chromosomal rearrangement, and DNA segment exchange, fusion and loss, as well as other fragmentation events. Finally, the genome of Antarctic moss *P. nutans* formed 22 chromosomes. This genome also showed high synteny with the previously released genomes of *P. patens* and *C. purpureus* (**Figure 2C**). Interestingly, almost each chromosome in *C. purpureus* was syntenic with two adjacent chromosomes in *P. nutans*. Thus, we inferred that the *P. nutans* genome was forming haploidy after undergoing a recent WGD event and still retained a large number of duplicated blocks within its genome.

Whole-genome duplication also represents one of the primary mechanisms for gene duplication, along with three principal evolutionary patterns such as tandem duplication, segmental duplication, and transposition events (Hong et al., 2021). In addition, gene duplicates comprise 8–20% of the genes in eukaryotic genomes, which facilitates the establishment of new gene functions and the generation of evolutionary novelty (Moore and Purugganan, 2003; Van De Peer et al., 2009). Segmental and tandem duplications are two predominate causes for gene family expansion in plants. We identified 28,529 segmental duplication gene pairs in *P. nutans* and accounted for 76.29% of the total duplication gene pairs. Segmental duplications multiple genes are generated through the chromosome rearrangements following WGD (Hong et al., 2021). Thus, the high proportion of segmental duplication gene pairs was largely congruent with the above analysis that *P. nutans* had undergone a recent WGD event (**Figure 2B**). Tandem duplications are defined as multiple genes of one family emerging within the same intergenic region or in neighboring intergenic regions. In absence of WGD in *A. angustus*, the tandem gene duplications mainly contribute to the expansion of specialized gene family and the adaptive evolution of the hornwort during the colonization of terrestrial ecosystems (Zhang et al., 2020a).

Some central gene-family expansions continuously occurred after the origin of embryophytes, especially in the origin of

land plants and the evolution of bryophytes (Leebens-Mack et al., 2019). WGD provides a major fountain for large expansions of gene families. Through WGD events, the duplication of all genes could offer an extraordinary opportunity for the occurrence of evolutionary novelties and functional diversifications (Van De Peer et al., 2009; Hong et al., 2021). These mechanisms may support adaptive evolution and have likely contributed plant survival in terrestrial habitats. Accordingly, we identified 7,807 of significantly expanded and 635 of markedly contracted gene families in the *P. nutans* genome (**Figure 3A**). The origin of land plants and the evolution of bryophytes accompany the largest number of gene-family expansions (i.e., transition from streptophyte algae to bryophytes; Leebens-Mack et al., 2019). Notably, in the *P. nutans* genome, the number of expanded gene families was remarkable large, being even 2.80-fold than that in the *P. patens* genome (**Figure 3A**). TFs-encoding genes were among the most highly retained following WGD incident, which exhibited in the present comparison of five bryophyte genomes. Since *P. nutans* and *P. patens* genomes both underwent at least one WGDs that they possessed a large number of TFs, whereas *A. angustus* and *M. polymorpha* genomes did not experience WGDs holding far less TFs. Particularly, several TFs like AP2, bHLH, bZIP, and WRKY emerged in the latest ancestor of Viridiplantae, whereas GRAS and NAC family genes arose in early streptophytes after divergence from the chlorophytes (Leebens-Mack et al., 2019). We found that they all emerged in these five bryophyte genomes (**Figure 3C**).

The gene families encoding DNA photolyase, antioxidant enzymes, PAL, CHS, and 2-OGD were significantly upregulated under various abiotic stresses, possibly representing diverse adaptive processes (**Figures 4, 5**). Photolyase is responsible for repairing the UV-induced DNA damage in a light-dependent manner. We found that the gene family of DNA photolyase was highly expended in *P. nutans* when compared with other four bryophytes (**Figure 4A**). Previously, we reported a CPD photolyase gene *PnPHR1* isolated from Antarctic moss *P. nutans*. PnPHR1 can repair photoproducts of cyclobutane pyrimidine dimers (CPD) and enhance the plant resistance to UV-B radiation by scavenging extra ROS (Wang et al., 2021a). PnPHR1 also increases the survival rate of *Escherichia coli* strain after UV-B radiation. Therefore, the successful colonization of *P. nutans* in Antarctic extreme environments can be attributed to the large expansion of gene families with functions associated with DNA repairing machine, ROS scavenging system, unsaturated fatty acid biosynthesis, and flavonoid biosynthesis.

Flavonoid Metabolites Play an Important Role in UV-B Tolerance of *Pohlia nutans*

Antarctic terrestrial organisms have undergone some of the most extremely environmental pressures including higher UV irradiation, cold, and extreme dryness. The production of a diverse repertoire of specialized metabolites is one of the features of land plants in response to environmental stresses (Yang et al., 2018; de Vries et al., 2021). The phenylpropanoid pathway is the sources for lots of metabolites that act in warding off environmental stressors (de Vries et al., 2021). The phenylpropanoid and flavonoid pathways likely facilitate the earliest steps of plants on land, protecting pioneer plants against the threatens of the terrestrial stresses such as drought and increased UV radiation (Jiao et al., 2020; de Vries et al., 2021). While some of flavonoids are the well-known UV screening compounds, various other phenylpropanoid-derived metabolites are equally potent UV protectants (Zeng et al., 2020; Nichelmann and Pescheck, 2021). We found that PAL gene family was markedly expanded in *P. nutans* when compared with other four bryophytes (**Supplementary Figure 4**). Of them, 14 PAL genes were markedly upregulated in *P. nutans* under UV-B radiation. Some phenylpropanoids undertake a screening barrier increasing resistance to UV-B light in the Antarctic moss *C. purpureus* (Clarke and Robinson, 2008). Flavonoids are produced through the phenylpropanoid and acetate-malonate metabolic pathways (Buer et al., 2010). They are deemed to have emerged during plant landing and adapting to terrestrial habitats about 500 million years ago (Davies et al., 2020; Stiller et al., 2021). Bryophytes and angiosperms likely have both commonalities and significant differences in flavonoid biosynthesis and metabolic regulation (Davies et al., 2020). We found that enzymes of flavonoid pathway were expanded in *P. nutans*, such as CHS and 2-OGD (**Figure 5B**). In addition, some of these genes were upregulated under UV-B radiation (**Figure 5C**). Previously, we characterized a type I flavone synthase (i.e., PnFNSI) from *P. nutans*. PnFNSI could catalyze the conversion of naringenin to apigenin and increase plant tolerance to UV-B radiation and drought stress (Wang et al., 2020). Similarly, flavonoids and carotenoids extracted from three Antarctic moss species demonstrate the features of UV-absorbing compounds and activate the DNA damage repair process (Pereira et al., 2009).

Gene-family expansions or contractions contribute to the production of a diverse repertoire of specialized metabolites in response to environmental stresses (Leebens-Mack et al., 2019; de Vries et al., 2021). However, analytical approaches in earlier publications largely use the simple equipment of spectrophotometer (i.e., UV-B-absorbing compounds at $AUC_{280–315\,nm}$, anthocyanins at $A_{526\,nm}$; Waterman et al., 2018). The HPLC-MS/MS-based plant metabolomics provides a qualitative and quantitative solution and has been extensively used to profile diverse repertoire of specialized metabolites (Li and Song, 2019; Wang et al., 2021b). Here, a widely targeted metabolomics built on the UPLC-MS/MS platforms found that 80 metabolites were markedly upregulated, and 17 metabolites were notably downregulated in *P. nutans* after UV-B stress (**Figure 5D**). KEGG enrichment indicated that these SCMs were mainly involved in flavonoid biosynthesis, flavone and flavonol biosynthesis, anthocyanin biosynthesis, and isoflavonoid biosynthesis and other small molecule metabolites (**Figure 5F**). Among them, gallic acid (a kind of flavonol) was the most significantly changed metabolite of with $\log_2$Fold change 14.66. Particularly, abiotic stress usually causes the generation of ROS and oxidative damage. Plants generally maintains a delicate balance of ROS through efficient ROS scavenging system. Gallic acid serves a natural defense mechanism by exhibiting excellent antioxidant and improving leaf membrane stability (Choubey et al., 2018; Saidi et al., 2021). In addition, cyanidin

3-O-(6″-malonylglucoside), a kind of anthocyanins, was also significantly accumulated with $\log_2$(Fold change) 11.02 (**Supplementary Table 6**), exhibiting predominant antioxidant effects against various oxidative stress (Rahman et al., 2021). In contrast, anthocyanins were not detected in methanolic extracts of *P. patens* separated by HPLC method (Wolf et al., 2010). Similarly, several SCMs like procyanidin and luteolin were detected in *P. nutans*, and functioned as an efficient antioxidant (Gong et al., 2020). Notably, several Antarctic mosses (e.g., *Bryum pseudotriquetrum*, *C. perpureus*, and *Schistidium antarctici*) and liverwort (*M. polymorpha*) possess anthocyanins, but Antarctic algae do not have anthocyanins (Singh et al., 2011). Flavonoids will reduce the transmittance of UV-B light and enabled plants to first colonize land (Singh et al., 2012; Singh and Singh, 2014). Therefore, transcriptomics integrated with metabolomics offer evidence for the opinion that flavonoids act as efficient antioxidants might dominate the tolerance of *P. nutans* against UV-B radiation.

CONCLUSION

Mosses are the basal land plants thriving in Antarctic ice-free continent and have evolved to survive, grow, and propagate in this harsh environment. We found that the Antarctic moss *P. nutans* genome harbors the signatures of a recent WGD incident with a high proportion of repeat sequences. In particular, the massive segmental gene duplications and remarkable expansions of gene families provide the primary driving force for the evolution of novel gene function and the production of specialized metabolite repertoire. Notably, the integration of multi-omics data elucidates the underlining mechanism of Antarctic moss *P. nutans* adaptation to extreme environments, which can be largely attributed to DNA repairing, ROS scavenging, and flavonoid biosynthesis. Collectively, the high-quality genome of *P. nutans* provides insight into the unique features of early land plant evolution and the molecular mechanism of moss surviving under rapidly changed environments in Antarctica.

DATA AVAILABILITY STATEMENT

The plant material is vouchered and available on request from Dr. Linlin Zhao at the First Institute of Oceanography, Natural Resources Ministry of China. The whole genome sequence data presented in the study are deposited in the National Genomics Data Center (NGDC, https://ngdc.cncb.ac.cn), BioProject number PRJCA008231. Of them, the PacBio HiFi reads of *P. nutans* were deposited in Genome Sequence Archive (GSA), accession number CRA006048; the Hi-C sequencing reads were deposited in GSA under the accession number CRA006049; the assembly and annotation data were deposited in Genome Warehouse (GWH), accession number GWHBHNB00000000. In addition, the transcriptome sequencing data of *P. nutans* were also deposited in GSA, accession number CRA006053, CRA006553, and CRA006556.

AUTHOR CONTRIBUTIONS

LZ and PZ conceived and designed the project and revised the manuscript. SL prepared the samples and performed the analyses of the genome, transcriptome sequence, with assistance from TL. SF performed the comparative genomics. BC, DY, and ZZ performed the metabolomics profiling and analysis. SL wrote the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This research was funded by the National Natural Science Foundation of China, grant/award number: 41976225; the Central Public-Interest Scientific Institution Basal Research Foundation of China, grant/award number: GY0219Q05; and the Development Project of Shandong Province, grant/award number: 2019GSF107064.

ACKNOWLEDGMENTS

We thank Qi Liu and Bing Peng in Wuhan Onemore-tech Co., Ltd. for their assistance with genome sequencing and analysis.

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.920138/full#supplementary-material

Supplementary Table 1 | Comparison of BUSCO assessment of genome annotation among four bryophytes.

Supplementary Table 2 | Statistics of sequencing reads mapping to the Pohlia nutans genome assmebly.

Supplementary Table 3 | Identification of non-coding RNA genes in the Pohlia nutans genome.

Supplementary Table 4 | Summary of transposable elements (TEs) identified in the Pohlia nutans genome.

Supplementary Table 5 | All detected metabolites in Pohlia nutans under UV-B Radiation.

Supplementary Table 6 | Significantly changed metabolites in Pohlia nutans under UV-B radiation.

Supplementary Figure 1 | Mapping analysis of transcriptome sequencing data. **(A)** The density of Illumina paired-end reads from transcriptome data in 22 chromosomes of *Pohlia nutans*. **(B)** Number of mapped reads from transcriptome sequencing in 22 chromosomes of *P. nutans*. Transcriptome sequencing data from cold stress were used for this assessment. Fix layout: justify the text.

Supplementary Figure 2 | LTR-RT evolution and chromosome syntenic analyses. **(A)** Number of intact LTR-RTs of *Pohlia nutans*, *Physcomitrella patens*, *Marchantia polymorpha*, and *Anthoceros angustus*. **(B)** Phylogenetic relationships of *Copia* and *Gypsy* retrotransposons across four bryophytes. **(C)** Dotplot representation of the introgenomic synteny analysis within the *P. nutans* genome.

(D) Compare the consistency of syntenic blocks between two groups of *P. nutans* vs. *C. purpureus*, and *P. nutans* vs. *P. patens*.

Supplementary Figure 3 | Venn diagram of gene families and clustering statistics. **(A)** The Venn diagram shows the shared and unique gene families in *Pohlia nutans*, *Physcomitrella patens*, *Anthoceros angustus*, and *Klebsormidium nitens* genomes. **(B)** The Venn diagram shows the shared and unique gene families in *P. nutans*, *P. patens*, *A. angustus*, *K. nitens*, and *Chlamydomonas reinhardtii* genomes. **(C)** Scatterplot of KEGG enrichment of the unique family genes in *P. nutans*. **(D)** Clustering statistics of gene family among the green plants of different evolutionary status.

Supplementary Figure 4 | Phenylalanine ammonia-lyase (PAL) family was expanded and upregulated under various abiotic stresses. **(A)** Phylogenetic analysis of PAL families in five bryophyte species (*Pohlia nutans*, *Physcomitrella patens*, *Marchantia polymorpha*, and *Anthoceros angustus*). The branches of *P. nutans* were colored in red, and the gene ID of *P. nutans* were shown. **(B)** Heatmap showed that PAL genes were markedly upregulated under cold stress, and UV-B radiation detected by transcriptome sequencing.

Supplementary Figure 5 | Statistical analysis of data quality and differential metabolites. **(A)** Principal component analysis (PCA) for mass spectrum data of UV-B radiation group and control group, and mix (quality control sample). The *x* axis and *y* axis indicate the PC1 and PC2, respectively. Three groups (i.e., CG, UV-B, and mix) were clearly separated, and three biological replicates of each group were clustered together, showing that the experiments were reproducible and reliable. **(B)** Heatmap shows that there were significant differences in the classes and quantities of metabolites between UV-B radiation group and control group. The green color indicates the decrease of differentially expressed metabolites, while the red color indicates the increase of differentially expressed metabolites. **(C)** Score scatter plot of the latent structure discriminant analysis (OPLS-DA) model for the differences between UV-B radiation group and control group. R^2X and R^2Y indicate the interpretation rate of *X* and *Y* matrix, respectively. Q^2Y represents the prediction ability of the model, while a value closer to 1.0 means that the model is more stable and reliable. In addition, $Q^2Y > 0.5$ can be regarded as an effective model, and $Q^2Y > 0.9$ is an excellent model. The data indicated that UV-B radiation strongly changed the metabolite profiles of the Antarctic moss *Pohlia nutans*. **(D)** Statistics of significantly changed metabolites in *P. nutans* under UV-B Radiation.

REFERENCES

Albert, N. W., Thrimawithana, A. H., Mcghie, T. K., Clayton, W. A., Deroles, S. C., Schwinn, K. E., et al. (2018). Genetic analysis of the liverwort *Marchantia polymorpha* reveals that R2R3MYB activation of flavonoid production in response to abiotic stress is an ancient character in land plants. *New Phytol.* 218, 554–566. doi: 10.1111/nph.15002

Allen, G. C., Flores-Vergara, M. A., Krasynanski, S., Kumar, S., and Thompson, W. F. (2006). A modified protocol for rapid DNA isolation from plant tissues using cetyltrimethylammonium bromide. *Nat. Protoc.* 1, 2320–2325. doi: 10.1038/nprot.2006.384

Amesbury, M. J., Roland, T. P., Royles, J., Hodgson, D. A., Convey, P., Griffiths, H., et al. (2017). Widespread biological response to rapid warming on the Antarctic peninsula. *Curr. Biol.* 27, 1616.e2–1622.e2. doi: 10.1016/j.cub.2017.04.034

Bao, T., Zhu, R., Wang, P., Ye, W., Ma, D., and Xu, H. (2018). Potential effects of ultraviolet radiation reduction on tundra nitrous oxide and methane fluxes in maritime Antarctica. *Sci. Rep.* 8:3716. doi: 10.1038/s41598-018-21881-1

Benaud, N., Edwards, R. J., Amos, T. G., D'agostino, P. M., Gutiérrez-Chávez, C., Montgomery, K., et al. (2021). Antarctic desert soil bacteria exhibit high novel natural product potential, evaluated through long-read genome sequencing and comparative genomics. *Environ. Microbiol.* 23, 3646–3664. doi: 10.1111/1462-2920.15300

Bertini, L., Cozzolino, F., Proietti, S., Falconieri, G. S., Iacobucci, I., Salvia, R., et al. (2021). What Antarctic plants can tell us about climate changes: temperature as a driver for metabolic reprogramming. *Biomol. Ther.* 11:1094. doi: 10.3390/biom11081094

Birney, E., Clamp, M., and Durbin, R. (2004). Genewise and genomewise. *Genome Res.* 14, 988–995. doi: 10.1101/gr.1865504

Boetzer, M., Henkel, C. V., Jansen, H. J., Butler, D., and Pirovano, W. (2010). Scaffolding pre-assembled contigs using SSPACE. *Bioinformatics* 27, 578–579. doi: 10.1093/bioinformatics/btq683

Bowman, J. L., Kohchi, T., Yamato, K. T., Jenkins, J., Shu, S., Ishizaki, K., et al. (2017). Insights into land plant evolution garnered from the *Marchantia polymorpha* genome. *Cell* 171, 287.e15–304.e15. doi: 10.1016/j.cell.2017.09.030

Buer, C. S., Imin, N., and Djordjevic, M. A. (2010). Flavonoids: new roles for old molecules. *J. Integr. Plant Biol.* 52, 98–111. doi: 10.1111/j.1744-7909.2010.00905.x

Cannone, N., Guglielmin, M., Convey, P., Worland, M. R., and Favero Longo, S. E. (2016). Vascular plant changes in extreme environments: effects of multiple drivers. *Clim. Chang.* 134, 651–665. doi: 10.1007/s10584-015-1551-7

Cannone, N., Malfasi, F., Favero-Longo, S. E., Convey, P., and Guglielmin, M. (2022). Acceleration of climate warming and plant dynamics in Antarctica. *Curr. Biol.* 32, 1599.e2–1606.e2. doi: 10.1016/j.cub.2022.01.074

Cassaro, A., Pacelli, C., Aureli, L., Catanzaro, I., Leo, P., and Onofri, S. (2021). Antarctica as a reservoir of planetary analogue environments. *Extremophiles* 25, 437–458. doi: 10.1007/s00792-021-01245-w

Cheng, H., Concepcion, G. T., Feng, X., Zhang, H., and Li, H. (2021). Haplotype-resolved de novo assembly using phased assembly graphs with hifiasm. *Nat. Methods* 18, 170–175. doi: 10.1038/s41592-020-01056-5

Choubey, S., Goyal, S., Varughese, L. R., Kumar, V., Sharma, A. K., and Beniwal, V. (2018). Probing gallic acid for its broad spectrum applications. *Mini-Rev. Med. Chem.* 18, 1283–1293. doi: 10.2174/1389557518666180330114010

Clarke, L. J., and Robinson, S. A. (2008). Cell wall-bound ultraviolet-screening compounds explain the high ultraviolet tolerance of the Antarctic moss, *Ceratodon purpureus*. *New Phytol.* 179, 776–783. doi: 10.1111/j.1469-8137.2008.02499.x

Convey, P., Coulson, S. J., Worland, M. R., and Sjöblom, A. (2018). The importance of understanding annual and shorter-term temperature patterns and variation in the surface levels of polar soils for terrestrial biota. *Polar Biol.* 41, 1587–1605. doi: 10.1007/s00300-018-2299-0

Convey, P., and Peck, L. S. (2019). Antarctic environmental change and biological responses. Science. *Advances* 5:eaaz0888. doi: 10.1126/sciadv.aaz0888

Davies, K. M., Jibran, R., Zhou, Y., Albert, N. W., Brummell, D. A., Jordan, B. R., et al. (2020). The evolution of flavonoid biosynthesis: a bryophyte perspective. *Front. Plant Sci.* 11:7. doi: 10.3389/fpls.2020.00007

De Bie, T., Cristianini, N., Demuth, J. P., and Hahn, M. W. (2006). CAFE: a computational tool for the study of gene family evolution. *Bioinformatics* 22, 1269–1271. doi: 10.1093/bioinformatics/btl097

de Vries, S., Fürst-Jansen, J. M. R., Irisarri, I., Dhabalia Ashok, A., Ischebeck, T., Feussner, K., et al. (2021). The evolution of the phenylpropanoid pathway entailed pronounced radiations and divergences of enzyme families. *Plant J.* 107, 975–1002. doi: 10.1111/tpj.15387

Edgar, R. C., and Myers, E. W. (2005). PILER: identification and classification of genomic repeats. *Bioinformatics* 21, i152–i158. doi: 10.1093/bioinformatics/bti1003

Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for de novo detection of LTR retrotransposons. *BMC Bioinformatics* 9:18. doi: 10.1186/1471-2105-9-18

Fetisova, L., and Visotskaya, E. (1970). Chromosome numbers in the mosses from Estonia. *Bot. Zh.* 55, 1150–1152.

Flynn, J. M., Hubley, R., Goubert, C., Rosen, J., Clark, A. G., Feschotte, C., et al. (2020). RepeatModeler2 for automated genomic discovery of transposable element families. *Proc. Natl. Acad. Sci.* 117, 9451–9457. doi: 10.1073/pnas.1921046117

Gong, X., Xu, L., Fang, X., Zhao, X., Du, Y., Wu, H., et al. (2020). Protective effects of grape seed procyanidin on isoflurane-induced cognitive impairment in mice. *Pharm. Biol.* 58, 200–207. doi: 10.1080/13880209.2020.1730913

Gray, A., Krolikowski, M., Fretwell, P., Convey, P., Peck, L. S., Mendelova, M., et al. (2020). Remote sensing reveals Antarctic green snow algae as important terrestrial carbon sink. *Nat. Commun.* 11:2527. doi: 10.1038/s41467-020-16018-w

Holt, C., and Yandell, M. (2011). MAKER2: an annotation pipeline and genome-database management tool for second-generation genome projects. *BMC Bioinformatics* 12:491. doi: 10.1186/1471-2105-12-491

Hong, S., Lim, Y. P., Kwon, S.-Y., Shin, A.-Y., and Kim, Y.-M. (2021). Genome-wide comparative analysis of flowering-time genes; insights on the gene family expansion and evolutionary perspective. *Front. Plant Sci.* 12:702243. doi: 10.3389/fpls.2021.702243

Hossaini, R., Chipperfield, M. P., Montzka, S. A., Leeson, A. A., Dhomse, S. S., and Pyle, J. A. (2017). The increasing threat to stratospheric ozone from dichloromethane. *Nat. Commun.* 8:15962. doi: 10.1038/ncomms15962

Hughes, K. A., Pescott, O. L., Peyton, J., Adriaens, T., Cottier-Cook, E. J., Key, G., et al. (2020). Invasive non-native species likely to threaten biodiversity and ecosystems in the Antarctic peninsula region. *Glob. Chang. Biol.* 26, 2702–2716. doi: 10.1111/gcb.14938

Jiao, C., Sørensen, I., Sun, X., Sun, H., Behar, H., Alseekh, S., et al. (2020). The *Penium margaritaceum* genome: hallmarks of the origins of land plants. *Cell* 181, 1097.e12–1111.e12. doi: 10.1016/j.cell.2020.04.019

Katoh, K., and Standley, D. M. (2013). MAFFT multiple sequence alignment software version 7: improvements in performance and usability. *Mol. Biol. Evol.* 30, 772–780. doi: 10.1093/molbev/mst010

Kim, B. M., Amores, A., Kang, S., Ahn, D. H., Kim, J. H., Kim, I. C., et al. (2019). Antarctic blackfin icefish genome reveals adaptations to extreme environments. *Nat. Ecol. Evol.* 3, 469–478. doi: 10.1038/s41559-019-0812-7

Lang, D., Ullrich, K. K., Murat, F., Fuchs, J., Jenkins, J., Haas, F. B., et al. (2018). The *Physcomitrella patens* chromosome-scale assembly reveals moss genome structure and evolution. *Plant J.* 93, 515–533. doi: 10.1111/tpj.13801

Lee, J. R., Raymond, B., Bracegirdle, T. J., Chades, I., Fuller, R. A., Shaw, J. D., et al. (2017). Climate change drives expansion of Antarctic ice-free habitat. *Nature* 547, 49–54. doi: 10.1038/nature22996

Leebens-Mack, J. H., Barker, M. S., Carpenter, E. J., Deyholos, M. K., Gitzendanner, M. A., Graham, S. W., et al. (2019). One thousand plant transcriptomes and the phylogenomics of green plants. *Nature* 574, 679–685. doi: 10.1038/s41586-019-1693-2

Li, H. (2018). Minimap2: pairwise alignment for nucleotide sequences. *Bioinformatics* 34, 3094–3100. doi: 10.1093/bioinformatics/bty191

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics* 25, 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, C., Liu, S., Zhang, W., Chen, K., and Zhang, P. (2019). Transcriptional profiling and physiological analysis reveal the critical roles of ROS-scavenging system in the Antarctic moss *Pohlia nutans* under ultraviolet-B radiation. *Plant Physiol. Biochem.* 134, 113–122. doi: 10.1016/j.plaphy.2018.10.034

Li, Q., and Song, J. (2019). Analysis of widely targeted metabolites of the euhalophyte *Suaeda salsa* under saline conditions provides new insights into salt tolerance and nutritional value in halophytic species. *BMC Plant Biol.* 19:388. doi: 10.1186/s12870-019-2006-5

Li, L., Stoeckert, C. J. Jr., and Roos, D. S. (2003). OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 13, 2178–2189. doi: 10.1101/gr.1224503

Li, W., Wen, L., Chen, Z., Zhang, Z., Pang, X., Deng, Z., et al. (2021). Study on metabolic variation in whole grains of four proso millet varieties reveals metabolites important for antioxidant properties and quality traits. *Food Chem.* 357:129791. doi: 10.1016/j.foodchem.2021.129791

Liu, S., Fang, S., Liu, C., Zhao, L., Cong, B., and Zhang, Z. (2021). Transcriptomics integrated with metabolomics reveal the effects of ultraviolet-B radiation on flavonoid biosynthesis in Antarctic moss. *Front. Plant Sci.* 12:788377. doi: 10.3389/fpls.2021.788377

Liu, S., Zhang, P., Li, C., and Xia, G. (2019). The moss jasmonate ZIM-domain protein PnJAZ1 confers salinity tolerance via crosstalk with the abscisic acid signalling pathway. *Plant Sci.* 280, 1–11. doi: 10.1016/j.plantsci.2018.11.004

Majoros, W. H., Pertea, M., and Salzberg, S. L. (2004). TigrScan and GlimmerHMM: two open source ab initio eukaryotic gene-finders. *Bioinformatics* 20, 2878–2879. doi: 10.1093/bioinformatics/bth315

Malenovský, Z., Turnbull, J. D., Lucieer, A., and Robinson, S. A. (2015). Antarctic moss stress assessment based on chlorophyll content and leaf density retrieved from imaging spectroscopy data. *New Phytol.* 208, 608–624. doi: 10.1111/nph.13524

Manni, M., Berkeley, M. R., Seppey, M., Simão, F. A., and Zdobnov, E. M. (2021). BUSCO update: novel and streamlined workflows along with broader and deeper phylogenetic coverage for scoring of eukaryotic, prokaryotic, and viral genomes. *Mol. Biol. Evol.* 38, 4647–4654. doi: 10.1093/molbev/msab199

Moore, R. C., and Purugganan, M. D. (2003). The early stages of duplicate gene evolution. *Proc. Natl. Acad. Sci.* 100, 15682–15687. doi: 10.1073/pnas.2535513100

Neale, R. E., Barnes, P. W., Robson, T. M., Neale, P. J., Williamson, C. E., Zepp, R. G., et al. (2021). Environmental effects of stratospheric ozone depletion, UV radiation, and interactions with climate change: UNEP environmental effects assessment panel, update 2020. *Photochem. Photobiol. Sci.* 20, 1–67. doi: 10.1007/s43630-020-00001-x

Nguyen, L. T., Schmidt, H. A., Von Haeseler, A., and Minh, B. Q. (2015). IQ-TREE: a fast and effective stochastic algorithm for estimating maximum-likelihood phylogenies. *Mol. Biol. Evol.* 32, 268–274. doi: 10.1093/molbev/msu300

Nichelmann, L., and Pescheck, F. (2021). Solar UV-B effects on composition and UV screening efficiency of foliar phenolics in *Arabidopsis thaliana* are augmented by temperature. *Physiol. Plant.* 173, 762–774. doi: 10.1111/ppl.13554

Nyamundanda, G., Brennan, L., and Gormley, I. C. (2010). Probabilistic principal component analysis for metabolomic data. *BMC Bioinformatics* 11:571. doi: 10.1186/1471-2105-11-571

Ou, S., and Jiang, N. (2018). LTR_retriever: a highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol.* 176, 1410–1422. doi: 10.1104/pp.17.01310

Pereira, B. K., Rosa, R. M., Da Silva, J., Guecheva, T. N., Oliveira, I. M., Ianistcki, M., et al. (2009). Protective effects of three extracts from Antarctic plants against ultraviolet radiation in several biological models. *J. Photochem. Photobiol. B* 96, 117–129. doi: 10.1016/j.jphotobiol.2009.04.011

Perera-Castro, A. V., Waterman, M. J., Turnbull, J. D., Ashcroft, M. B., Mckinley, E., Watling, J. R., et al. (2020). It is hot in the sun: Antarctic mosses have high temperature optima for photosynthesis despite cold climate. *Front. Plant Sci.* 11:1178. doi: 10.3389/fpls.2020.01178

Rahman, S., Mathew, S., Nair, P., Ramadan, W. S., and Vazhappilly, C. G. (2021). Health benefits of cyanidin-3-glucoside as a potent modulator of Nrf2-mediated oxidative stress. *Inflammopharmacology* 29, 907–923. doi: 10.1007/s10787-021-00799-7

Ramsay, H., and Spence, J. (1996). Chromosome data on Australasian Bryaceae. *J. Hattori Bot. Lab.* 80, 251–270.

Rensing, S. A., Lang, D., Zimmer, A. D., Terry, A., Salamov, A., Shapiro, H., et al. (2008). The *Physcomitrella* genome reveals evolutionary insights into the conquest of land by plants. *Science* 319, 64–69. doi: 10.1126/science.1150646

Robinson, S. A., King, D. H., Bramley-Alves, J., Waterman, M. J., Ashcroft, M. B., Wasley, J., et al. (2018). Rapid change in East Antarctic terrestrial vegetation in response to regional drying. *Nat. Clim. Chang.* 8, 879–884. doi: 10.1038/s41558-018-0280-0

Saidi, I., Guesmi, F., Kharbech, O., Hfaiedh, N., and Djebali, W. (2021). Gallic acid improves the antioxidant ability against cadmium toxicity: impact on leaf lipid composition of sunflower (*Helianthus annuus*) seedlings. *Ecotoxicol. Environ. Saf.* 210:111906. doi: 10.1016/j.ecoenv.2021.111906

Sato, K., Inoue, J., Simmonds, I., and Rudeva, I. (2021). Antarctic peninsula warm winters influenced by Tasman Sea temperatures. *Nat. Commun.* 12, 1497–1499. doi: 10.1038/s41467-021-21773-5

Silva, A. T., Gao, B., Fisher, K. M., Mishler, B. D., Ekwealor, J. T. B., Stark, L. R., et al. (2021). To dry perchance to live: insights from the genome of the desiccation-tolerant biocrust moss *Syntrichia caninervis*. *Plant J.* 105, 1339–1356. doi: 10.1111/tpj.15116

Singh, J., Dubey, A. K., and Singh, R. P. (2011). Antarctic terrestrial ecosystem and role of pigments in enhanced UV-B radiations. *Rev. Environ. Sci. Biotechnol.* 10, 63–77. doi: 10.1007/s11157-010-9226-3

Singh, J., Gautam, S., and Bhushan Pant, A. (2012). Effect of UV-B radiation on UV absorbing compounds and pigments of moss and lichen of Schirmacher oasis region, East Antarctica. *Cell. Mol. Biol.* 58, 80–84. doi: 10.1170/T924

Singh, J., and Singh, R. P. (2014). Adverse effects of UV-B radiation on plants growing at schirmacher oasis, East Antarctica. *Toxicol. Int.* 21, 101–106. doi: 10.4103/0971-6580.128815

Skotnicki, M., Bargagli, R., and Ninham, J. (2002). Genetic diversity in the moss *Pohlia nutans* on geothermal ground of mount Rittmann, Victoria land, Antarctica. *Polar Biol.* 25, 771–777. doi: 10.1007/s00300-002-0418-3

Smith, A., and Newton, M. (1967). Chromosome studies on some British and Irish mosses. II. *Trans. Br. Bryol. Soc.* 5, 245–270. doi: 10.1179/00681387804804304

Stamatakis, A. (2006). RAxML-VI-HPC: maximum likelihood-based phylogenetic analyses with thousands of taxa and mixed models. *Bioinformatics* 22, 2688–2690. doi: 10.1093/bioinformatics/btl446

Stanke, M., Keller, O., Gunduz, I., Hayes, A., Waack, S., and Morgenstern, B. (2006). AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 34, W435–W439. doi: 10.1093/nar/gkl200

Stiller, A., Garrison, K., Gurdyumov, K., Kenner, J., Yasmin, F., Yates, P., et al. (2021). From fighting critters to saving lives: polyphenols in plant defense and human health. *Int. J. Mol. Sci.* 22:8995. doi: 10.3390/ijms22168995

Trapnell, C., Pachter, L., and Salzberg, S. L. (2009). TopHat: discovering splice junctions with RNA-Seq. *Bioinformatics* 25, 1105–1111. doi: 10.1093/bioinformatics/btp120

Trapnell, C., Roberts, A., Goff, L., Pertea, G., Kim, D., Kelley, D. R., et al. (2012). Differential gene and transcript expression analysis of RNA-seq experiments with TopHat and cufflinks. *Nat. Protoc.* 7, 562–578. doi: 10.1038/nprot.2012.016

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

Van De Peer, Y., Maere, S., and Meyer, A. (2009). The evolutionary significance of ancient genome duplications. *Nat. Rev. Genet.* 10, 725–732. doi: 10.1038/nrg2600

Walker, B. J., Abeel, T., Shea, T., Priest, M., Abouelliel, A., Sakthikumar, S., et al. (2014). Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One* 9:e112963. doi: 10.1371/journal.pone.0112963

Wang, H., Liu, S., Wang, T., Liu, H., Xu, X., Chen, K., et al. (2020). The moss flavone synthase I positively regulates the tolerance of plants to drought stress and UV-B radiation. *Plant Sci.* 298:110591. doi: 10.1016/j.plantsci.2020.110591

Wang, H., Liu, H., Yu, Q., Fan, F., Liu, S., Feng, G., et al. (2021a). A CPD photolyase gene PnPHR1 from Antarctic moss *Pohlia nutans* is involved in the resistance to UV-B radiation and salinity stress. *Plant Physiol. Biochem.* 167, 235–244. doi: 10.1016/j.plaphy.2021.08.005

Wang, R., Shu, P., Zhang, C., Zhang, J., Chen, Y., Zhang, Y., et al. (2021b). Integrative analyses of metabolome and genome-wide transcriptome reveal the regulatory network governing flavor formation in kiwifruit (*Actinidia chinensis*). *New Phytol.* 233, 373–389. doi: 10.1111/nph.17618

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40:e49. doi: 10.1093/nar/gkr1293

Waterman, M. J., Bramley-Alves, J., Miller, R. E., Keller, P. A., and Robinson, S. A. (2018). Photoprotection enhanced by red cell wall pigments in three East Antarctic mosses. *Biol. Res.* 51, 49–13. doi: 10.1186/s40659-018-0196-1

Wingett, S., Ewels, P., Furlan-Magaril, M., Nagano, T., Schoenfelder, S., Fraser, P., et al. (2015). HiCUP: pipeline for mapping and processing Hi-C data. *F1000Res.* 4:1310. doi: 10.12688/f1000research.7334.1

Wolf, L., Rizzini, L., Stracke, R., Ulm, R., and Rensing, S. A. (2010). The molecular and physiological responses of *Physcomitrella patens* to ultraviolet-B radiation. *Plant Physiol.* 153, 1123–1134. doi: 10.1104/pp.110.154658

Wood, T. E., Takebayashi, N., Barker, M. S., Mayrose, I., Greenspoon, P. B., and Rieseberg, L. H. (2009). The frequency of polyploid speciation in vascular plants. *Proc. Natl. Acad. Sci. U. S. A.* 106, 13875–13879. doi: 10.1073/pnas.0811575106

Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* 35, W265–W268. doi: 10.1093/nar/gkm286

Yang, Z. (2007). PAML 4: phylogenetic analysis by maximum likelihood. *Mol. Biol. Evol.* 24, 1586–1591. doi: 10.1093/molbev/msm088

Yang, L., Wen, K. S., Ruan, X., Zhao, Y. X., Wei, F., and Wang, Q. (2018). Response of plant secondary metabolites to environmental factors. *Molecules* 23:762. doi: 10.3390/molecules23040762

Zeng, X., Yuan, H., Dong, X., Peng, M., Jing, X., Xu, Q., et al. (2020). Genome-wide dissection of co-selected UV-B responsive pathways in the UV-B adaptation of qingke. *Mol. Plant* 13, 112–127. doi: 10.1016/j.molp.2019.10.009

Zhang, J., Fu, X.-X., Li, R.-Q., Zhao, X., Liu, Y., Li, M.-H., et al. (2020a). The hornwort genome and early land plant evolution. *Nat. Plants* 6, 107–118. doi: 10.1038/s41477-019-0588-4

Zhang, X., Cvetkovska, M., Morgan-Kiss, R., Hüner, N. P. A., and Smith, D. R. (2021). Draft genome sequence of the Antarctic green alga Chlamydomonas sp. UWO241. *iScience* 24:102084. doi: 10.1016/j.isci.2021.102084

Zhang, Z., Qu, C., Zhang, K., He, Y., Zhao, X., Yang, L., et al. (2020b). Adaptation to extreme Antarctic environments revealed by the genome of a sea ice green alga. *Curr. Biol.* 30, 3330.e7–3341.e7. doi: 10.1016/j.cub.2020.06.029

Zhao, W., Li, Z., Hu, Y., Wang, M., Zheng, S., Li, Q., et al. (2019). Development of a method for protonema proliferation of peat moss (*Sphagnum squarrosum*) through regeneration analysis. *New Phytol.* 221, 1160–1171. doi: 10.1111/nph.15394

Zheng, Y., Jiao, C., Sun, H., Rosli, H. G., Pombo, M. A., Zhang, P., et al. (2016). iTAK: a program for genome-wide prediction and classification of plant transcription factors, transcriptional regulators, and protein kinases. *Mol. Plant* 9, 1667–1670. doi: 10.1016/j.molp.2016.09.014

Zwaenepoel, A., and Van De Peer, Y. (2019). Wgd-simple command line tools for the analysis of ancient whole-genome duplications. *Bioinformatics* 35, 2153–2155. doi: 10.1093/bioinformatics/bty915

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

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

frontiers | Frontiers in Plant Science

ORIGINAL RESEARCH
published: 22 June 2022
doi: 10.3389/fpls.2022.908511

Genome-Wide Study of *Hsp90* Gene Family in Cabbage (*Brassica oleracea* var. *capitata* L.) and Their Imperative Roles in Response to Cold Stress

*Shoukat Sajad[1], Shuhan Jiang[1], Muhammad Anwar[2], Qian Dai[1], Yuxia Luo[1], Muhammad A. Hassan[3], Charles Tetteh[4] and Jianghua Song[1]**

[1]College of Horticulture, Vegetable Genetics and Breeding Laboratory, Anhui Agricultural University, Hefei, China, [2]Guangdong Technology Research Center for Marine Algal Bioengineering, Guangdong Key Laboratory of Plant Epigenetics, College of Life Sciences and Oceanography, Shenzhen University, Shenzhen, China, [3]School of Agronomy, Anhui Agricultural University, Hefei, China, [4]Department of Plant Pathology, College of Plant Protection, Anhui Agricultural University, Hefei, China

OPEN ACCESS

Edited by:
Parviz Heidari, Shahrood University of Technology, Iran

Reviewed by:
Abdullah, Quaid-i-Azam University, Pakistan
Mostafa Ahmadizadeh, University of Hormozgan, Iran

***Correspondence:**
Jianghua Song
jhsong@ahau.edu.cn

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

Received: *30 March 2022*
Accepted: *24 May 2022*
Published: *22 June 2022*

Citation:
Sajad S, Jiang S, Anwar M, Dai Q, Luo Y, Hassan MA, Tetteh C and Song J (2022) Genome-Wide Study of Hsp90 Gene Family in Cabbage (Brassica oleracea var. capitata L.) and Their Imperative Roles in Response to Cold Stress. Front. Plant Sci. 13:908511. doi: 10.3389/fpls.2022.908511

Heat shock protein 90 (Hsp90) plays an important role in plant developmental regulation and defensive reactions. Several plant species have been examined for the *Hsp90* family gene. However, the *Hsp90* gene family in cabbage has not been well investigated to date. In this study, we have been discovered 12 *BoHsp90* genes in cabbage (*Brassica oleracea* var. *capitata* L.). These *B. oleracea Hsp90* genes were classified into five groups based on phylogenetic analysis. Among the five groups, group one contains five *Hsp90* genes, including *BoHsp90-1*, *BoHsp90-2*, *BoHsp90-6*, *BoHsp90-10*, and *BoHsp90-12*. Group two contains three *Hsp90* genes, including *BoHsp90-3*, *BoHsp90-4*, and *BoHsp90*. Group three only includes one *Hsp90* gene, including *BoHsp90-9*. Group four were consisting of three *Hsp90* genes including *BoHsp90-5*, *BoHsp90-7*, and *BoHsp90-8*, and there is no *Hsp90* gene from *B. oleracea* in the fifth group. Synteny analysis showed that a total of 12 *BoHsp90* genes have a collinearity relationship with 5 *Arabidopsis* genes and 10 *Brassica rapa* genes. The promoter evaluation revealed that the promoters of *B. oleracea Hsp90* genes included environmental stress-related and hormone-responsive *cis-elements*. RNA-seq data analysis indicates that tissue-specific expression of *BoHsp90-9* and *BoHsp90-5* were highly expressed in stems, leaves, silique, and flowers. Furthermore, the expression pattern of *B. oleracea BoHsp90* exhibited that *BoHsp90-2, BoHsp90-3, BoHsp90-7, BoHsp90-9, BoHsp90-10, and BoHsp90-11* were induced under cold stress, which indicates these *Hsp90* genes perform a vital role in cold acclimation and supports in the continual of normal growth and development process. The cabbage *Hsp90* gene family was found to be differentially expressed in response to cold stress, suggesting that these genes play an important role in cabbage growth and development under cold conditions.

Keywords: cabbage, genome wide, *Hsp90*, expression pattern, phylogenetic analysis

INTRODUCTION

Crop plants during their growth cycle often confronted with various biotic and abiotic stresses (Food and Agriculture Organization [FAO], 2020). Recent climatic changes continuously trigger extreme temperature events and disrupt optimal plant growth (IPCC, 2021; Masson-Delmotte et al., 2021). Increased global warming instigated the risk of cold damage to crop plants (Gu et al., 2008; Moreno and Orellana, 2011; Hassan et al., 2021). Plants contain regulatory mechanisms that enable them to withstand harsh environmental conditions, called acclimatization (Liu et al., 2019). To acclimatize the extreme temperature (low and high) conditions, plants exhibit a number of protein expressions either to promote biosynthesis of new proteins or to protect existing proteins; cold acclimation process carried through expressions of highly conserved stress proteins called heat shock proteins (Hsps; Queitsch et al., 2002). Various types of Hsps are particularly found in all plant species (Xiong and Ishitani, 2006). Based on molecular mass, Hsps were categorized into five groups, i.e., the sHsp family, the chaperonin (Hsp60/GroEL) family, the 70-kDa heat shock protein (Hsp70/DnaK) family, the Hsp90 family, and the Hsp100/ClpB family (Pratt and Toft, 2003; Wang et al., 2004; Al-Whaibi, 2011).

The Hsp90 protein family is a renowned and extremely conserved class of molecular chaperones in eukaryotic cytoplasm (Chen et al., 2006)." For instance, in *Arabidopsis*, there are 7 Hsp90 proteins, of which AtHsp90-1, AtHsp90-2, AtHsp90-3 and AtHsp90-4 are present in the cytoplasm, while AtHsp90-5, AtHsp90-6 and AtHsp90-7 are present in the mitochondria, chloroplast, and endoplasmic reticulum, respectively (Krishna and Gloor, 2001; Yamada et al., 2007). It is a part of the ATPase superfamily (GHKL: gyrase, histidine kinase, and MutL; Dutta and Inouye, 2000); It functions as an ATP-regulated dimer and chaperone protein with three highly conserved domains: a 25 kDa C-terminal substrate-binding domain, a 12 kDa intermediate ATP-binding domain, and a 12 kDa N-terminal ATP-binding domain (Prodromou et al., 1997; Pearl and Prodromou, 2000)." The protein channel that binds to ATP in Hsp90 is frequently closed (Sung et al., 2016) which results in the poor activity of N-terminal ATPase (Raman and Suguna, 2015). The cytoplasm of eukaryotic cells contains the intermediate and N-terminal domains of Hsp90; the charged portions are varied in length between species (Pearl and Prodromou, 2000). It is commonly known that a protein's function is determined by its ability to fold into a three-dimensional structure. The Hsp90 proteins together with other molecular chaperones provide a course that facilitates protein folding (Wang et al., 2004). The *Hsp90* gene family suppress protein aggregation and enable the ubiquitination of inactive proteins (Picard, 2002); further, it belongs to a group of molecular chaperones that are involved in the creation of the spatial structure of kinase substrates, DNA repair and substrate activation, early stress signaling, and transcription factor spatial structure maintenance (Jackson et al., 2004; Lachowiec et al., 2015)." The expression of stressor *Hsp90* gene is upregulated when plants are stressed; it interacts and repairs the deformed proteins with support of non-proteinaceous compounds (Mukhopadhyay et al., 2003).

This protein (Hsp90) has been found in a number of plant species such as 7 in *Arabidopsis* (Krishna and Gloor, 2001), 7 in *Solanum lycopersicum* and 7 in *Asteridae* and tobacco (Liu et al., 2014; Song et al., 2019) and pepper (Jing et al., 2020). Nine and ten *Hsp90* genes were identified in rice (Hu et al., 2009) and Populus (Zhang et al., 2013), respectively. Recent studies have shown that Hsp90 has a vital role in plant' acclimation to various biotic and abiotic stresses (Di Donato and Geisler, 2019). The *HSP* gene family were considered to be involved or induced in cold environments (Krishna et al., 1995); it has the potential to protect plant biological systems from cold damage by preventing freezing-induced protein denaturation (Yan et al., 2006). Many studies reported that the low-temperature stress activates the gene expressions of both chaperonins and Hsps, particularly Hsp90, Hsp70, and the *smHsps* (Wang et al., 2004). In *Brassica napus*, the expression of Hsp90 was upregulated under cold stress (Krishna et al., 1995; Gill et al., 2015). Likewise, rice also exhibited the up-regulated expression of *OsHsp90-2* and *OsHsp90-4* under drought, cold, and heat stress (Hu et al., 2009). Liu et al. (2022) reported that in rubber trees (*Hevea brasiliensis* Müll. Arg.) the expression of the *HbHsp90.1* gene was significantly up-regulated at 4°C for 6 h of cold stress treatment. Similarly, Song et al. (2019) observed an identical trend in the tobacco plant, where *NtHsp90-4*, *NtHsp90-5* and *NtHsp90-9* showed high levels of transcription at 4°C for 6 to 12 h duration under cold stress. Increased Hsp90 protein accumulation was observed in winter wheat under cold stress (Vítámvás et al., 2012).

Brassica oleracea var. *capitata* (Cabbage) is a popular vegetable crop cultivated all over the world and contains a variety of nutrients and health-promoting compounds (Lv et al., 2020). Temperature required for optimal cabbage growth is 17°C, while its normal growth and development decreased when temperature falls below 10°C; further it can withstand up to −5°C for short phase of cold exposure, prolonged exposure leads to cell death (Food and Agricultural Organization of the United Nations (FAO), 2007). This investigation executed a genome-wide analysis of the *Brassica* Genome Network database using Hsp90 protein sequences of *Arabidopsis thaliana*. We analyzed the gene structure, phylogenetic tree, chromosomal location, homologous gene pairs, conserved motifs, collinearity analysis, sectioning pressure and cis-elements of the cabbage *Hsp90* gene promoter by utilizing bioinformatics approaches. Moreover, this study explored the relevant roles of cabbage *Hsp90* gene, provides a vibrant approach for molecular breeding and develop a better understanding for its tolerance responses under cold conditions.

MATERIALS AND METHODS

Hsp90 Gene Family Identification in *Brassica oleracea*

To extract the *Hsp90* gene from the *B. oleracea* genome, "the *Arabidopsis Hsp90* gene was used as a query in the online

database Phytozome 11.0 (Goodstein et al., 2012).[1] The genome, CDS, protein and 1.5 kb upstream promoter region of the *Hsp90* gene were extracted and used as a query against the NCBI database[2] to identify conserved domains and discard *Hsp90* domains missing genes. To ensure the presence of Hsp90 specific domains, all non-redundant protein sequences were checked using manual and Pfam databases, and 12 *BoHsp90* genes were finally found.

Physical Properties and Position on Chromosomes

To predict the protein physical properties of *B. oleracea* Hsp90 (*BoHsp90*), "the ExPASy tool is utilized.[3] The phytozome was used to obtain information about gene start-end position, the number of amino acids, and the chromosome location of the *BoHsp90* family. Plot phenograms[4] indicate the location of each *BoHsp90* gene on its chromosome. Using standard parameters in the ExPASy program, the protein sequence of the Bo*Hsp90* gene was determined[5]" (Gasteiger et al., 2003)."

Structural Analysis and Phylogenetic Relationship and Motif Analysis

Using the CDS and genome sequences of the related *BoHsp90* genes, construct gene structures in the Gene Structure Display Server to find exons-introns within genes (Hu et al., 2015)[6]. Multiple alignments of Hsp90 protein sequences of *B. oleracea*, *Brassica rapa, B. napus*, cucumber, tomato, potato, soybean, *Vitis vinifera*, *Zea mays, Arabidopsis,* rice, tobacco, cotton, *Medicago truncatula*, *Elaeis guineensis*, *Nelumbo nucifera*, *and P. bretschneideri* were performed by MEGA 7.0 with default parameters. Then, based on the alignment results, a phylogenetic tree was constructed using the maximum likelihood method with the following parameters: partial deletion, and 1,000 bootstrap tests. MEME is used to carry out the conserved sections of the *BoHsp90* gene (Bailey et al., 2006)[7] and the Pfam database was used to screen their genomic assemblies (Punta et al., 2012).[8]

Synteny Analysis of *Hsp90* Genes in *Brassica oleracea*, *Arabidopsis*, and *Brassica rapa*

By using a software MCScanX, a syntenic study of the *Hsp90* gene in *B. oleracea*, *Arabidopsis*, and *B. rapa* was performed to better understand the syntenic interactions between *A. thaliana*, *B. oleracea*, and *B. rapa* (Wang et al., 2012). It was used to check the *Hsp90* genes synteny in, *Arabidopsis*, *B. oleracea* and *B. rapa*.

Cis-Regulatory Element Prediction for *Brassica oleracea Hsp90* Gene Promoters

The promoter sequence (1.5 kb DNA sequence in the start codon stream) of each *Hsp90* gene was derived from the *Brassica* genome. "Using the PlantCARE database, the cis-acting regulatory element then examines the promoter of each *B. oleracea Hsp90* gene[9] (Plantcare; Lescot et al., 2002). The anticipated cis-acting regulatory elements were categorized based on their regulatory functions."

Gene Duplication, Ka/Ks Calculation and Positive Selection Analysis

Segmental and Tandem duplication occurred when two closely related *BoHsp90* genes were discovered on the same or different chromosomes."Additionally, the gene duplication events *BoHsp90* genes and also synonymous substitution rate (Ka) and the non-synonymous substitution rate (Ks) were calculated using an online tool and computed their ratios.[10] The divergence time (T) for *B. oleracea* was calculated by using T = ks/2r where R 1.5×10^{-8}; Koch et al., 2000)." The ratio of Ka/Ks evaluates the selection pressure of *BoHsp90* repeated genes. When the ratio of Ka/Ks is>1, <1, or = 1, it is considered a positive, negative, or neutral selection. "The amino acids in the BoHsp90 protein that are under selective pressure are then clarified using the Selection 2.2 tool.[11] The maximum likelihood test from the Bayesian inference method is used to calculate the offset ratio between codons in an aligned sequence (Yang, 1997)." Different kinds the selection type appears in the selection result, and evaluate by scale color.

Transcriptomic Data Analysis

Gene expression patterns of cabbage in different tissues (roots, stems, flowers, leaves, scales, bulbs and callus) were retrieved from the NCBI database[12] based on FPKM expression data with accession number GSE42891 (Yu et al., 2014). Finally, an expression heatmap was created using the TBtools software.[13]

Plant Materials and Stress Treatments

The cabbage (cultivar Yingchuan) seeds were planted in plug trays in moistened mixed soil (peat moss) were provided by School of Horticulture, Anhui Agricultural University, Hefei, China. Seedlings were germinated and cultured for 5 to 6 weeks in a glass chamber at 22°C on a 16/8 h. (light/dark) cycle. Then, they are treated at low temperatures (−2°C, 0°C, and 4°C) for 6, 12 and 24 h (Ahmed et al., 2015). Plants that had not been treated were cultivated normally. At 6, 12, and 24 h following treatment, all true leaves were sampled and quickly put into liquid nitrogen and stored at −80°C for RNA extraction and synthesis of cDNA.

[1]https://phytozome-next.jgi.doe.gov/
[2]www.ncbi.nlm.nih.gov/Structure/
[3]http://web.expasy.org/protparam
[4]http://visualization.ritchielab.org/phenograms/plot
[5]http://www.expasy.org/tools
[6]http://gsds.cbi.pku.edu.cn/
[7]https://www.meme-suite.org/meme/
[8]https://pfam.sanger.ac.uk/
[9]http://bioinformatics.psb.ugent.be/
[10]https://services.cbu.uib.no/tools/kaks
[11]http://selecton.tau.ac.il
[12]https://www.ncbi.nlm.nih.gov/
[13]http://www.tbtools.com/

RNA Isolation, cDNA Synthesis and qRT-PCR Analysis to Validate the *BoHsp90* Genes Expression

The Novaprotein reagent or RNA isolation kit for RNA extraction from cabbage leaves and quantification using a NanoDrop spectrophotometer (Thermo Scientific). "The superscript III Reverse Transcriptase (Novaprotein) according to the manufacturer's guidelines, and the cDNA was dilute to 100 ng/l with ddH2O for further analysis. Quantitative RT-PCR was performed with the CFX96 Real Time System (Bio-Rad) with 96 well plates were used, and each well contained a total reaction mixture (20 µl) consisting of 10 µl of Power SYBR Master mix, 1 µl primer mix each F/R primers +6 µl ddH2O) and 2 µl cDNA." The transcript level was normalized with a housekeeping gene *BolActin*. The 3 biological replicates were used for each *BoHsp90s*. The primers sequence for each gene were designed by using Primer5 software as shown in **Supplementary Table S2**.

RESULTS

Identification of the *Hsp90* Gene Family in *Brassica oleracea*

Using the *A. thaliana* Hsp90 protein sequence (BAB09283.1) as a query sequence, by using the online database phytozome[14] to find *Hsp90* genes in the *Brassica* genome. "The blast outcomes of *A. thaliana* in *Brassica* gave 18 hits in phytozome." The signature *BoHsp90* domain for each protein was confirmed using SMART Pfam[15] (Punta et al., 2012). Additionally, NCBI web-based conserved domain search tool[16] (Marchler-Bauer et al., 2015) was used to verify the conserved domain, and the protein sequences lacking the *BoHsp90* domain were removed. During an extensive search of the *Hsp90* gene, 12 putative genes were found. The annotation of known *Brassica BoHsp90* family members were carried out, based on their chromosomal wise arrangements (*BoHsp901-12*). The protein sequences of *BoHsp90* genes extended from 44 (*BoHsp90-12*) to 1819 (*BoHsp90-11*) amino (aa) residue with molecular weight (MW) range of 46.93–201.60 kDa. All 12 Hsp90 members have an average length of 773.4 aa and an average MW of 87.92 KDa furthermore, 41.66% (5 out of 12 genes) of *BoHsp90* genes had high pI values (more than 5), while 58.33% (7 out of 12 genes) had pI values below 5. The predicted subcellular localization includes nuclear, endoplasmic reticulum and cytoplasmic. These results revealed that the maximum number of genes are found in the cytoplasm and percentage found in the nuclear (**Supplementary Table S1**).

Phylogenetic Relationship of the *BoHsp90* Genes in *Brassica oleracea*

To study their evolutionary relationship, the predicted protein sequence of BoHsp90 was employed to create a rootless phylogenetic tree. The phylogenetic study of the *Hsp90* genes distributed into five groups (G1, G2, G3, G4, and G5) containing *B. oleracea*, *B. rapa*, *B. napus*, tomato, cucumber, potato, *V. vinifera*, *Sorghum bicolor*, cotton, tobacco, *Arabidopsis*, rice, *Z. mays*, soybean, *M. truncatula, E. guineensis*, Chinese white pear (*Pyrus bretschneideri*) and *N. nucifera*. Among them G1 is the largest group with a total of 25 members, having *B. oleracea* (4 genes), B. *napa* (4 genes), *B. rapa* (3 genes), rice (3 genes), cucumbers (2 genes), potatoes (2 genes), *A. thaliana* (2 genes), 1 *V. vinifera* (1 gene), *E. guineensis* (1 gene), *N. nucifera* (1 gene) and one from sorghum. The second largest group, G4, comprises of 22 members, having *B. oleracea* (3 genes), *B. rapa* (3 genes), *B. napus* (1 gene), *A. thaliana* (3 genes), *soybeans* (2 genes), tomato (2 genes), rice (2 genes), *Z. mays* (2 genes), potato (1 gene), tobacco (1 gene), cotton (1 gene), and one *P. bretschneideri*. The G2 group contain 15 members, among which 3 in *B. oleracea*, 3 in *B. rapa*, one in *A. thaliana*, 2 in *Z. mays*, 2 in rice, 2 in tomato, one in cucumber, and one in soybean. While G3 comprised of 10 members as 2 in cotton, 2 in soybean, one in *B. oleracea*, *B. rapa*, *A. thaliana*, cucumber, tomato, and *M. truncatula*. G5 composed of 9 members and considered as a small group, having 3 in cotton, 2 in soybean, 2 in *M. truncatula*, one in cucumber and rice member (**Figure 1**).

Structural Analysis of the *BoHsp90* Genes

The arrangement of the *BoHsp90* genes was plotted using the online tool GSDS based on their coding sequence. Within each *BoHsp90* gene, "(**Figure 2**) displays the comparative lengths of introns and the conservation of corresponding exon sequences." All these genes have between 1 and 19 introns. However, half of the *BoHsp90* genes contain 1 to 14 introns, while another half of *BoHsp90* genes contain different numbers of introns likewise *BoHsp90-2* has 19 introns, *BoHsp90–10* contains 17 introns, 12 in *BoHsp90-11*, 9 in *BoHsp90-6*, *BoHsp90*-12 has 8 introns and 3 introns in the *BoHsp90–9*. The position and length of introns also vary widely. In order to better understand the structural properties of BoHsp90 protein, 10 conserved motifs in the protein were identified using the online program MEME motif search tool and the distribution of these conserved motifs in BoHsp90 protein was explored." The findings revealed that the majority of the genes comprised 10 conserved motifs. There are similar motifs in closely related genes, indicating similar functions in the *BoHsp90* gene family. The *BoHsp90-3* and *BoHsp90*-12 contain six motifs, *BoHsp90-7* has eight and 9 motifs were found in *BoHsp90-9*. According to these findings, structural motif alignment varies between members of the *BoHsp90* family genes but is similar within closely related genes (**Supplementary Table S5**).

Chromosomal Distribution, Collinearity Analysis of the *Hsp90* Gene Family in *Brassica oleracea*

The chromosomal location map shown in (**Figure 3**) was created to specify the chromosomal orientation and deletion site of the *BoHsp90* gene. The findings revealed that a total of 12 genes were linked to four chromosomes in cabbage. Of these, 83% of the *BoHsp90* gene (10 loci) are present on three chromosomes in cabbage. Chromosomes 3, 2, and the more repressed *Hsp90* members each have three and four genes, while chromosome

[14]http://phytozome.jgi.doe.gov/pz/portal.html
[15]http://pfam.xfam.org/
[16]https://www.ncbi.nlm.nih.gov/Structure/cdd/wrpsb.cgi

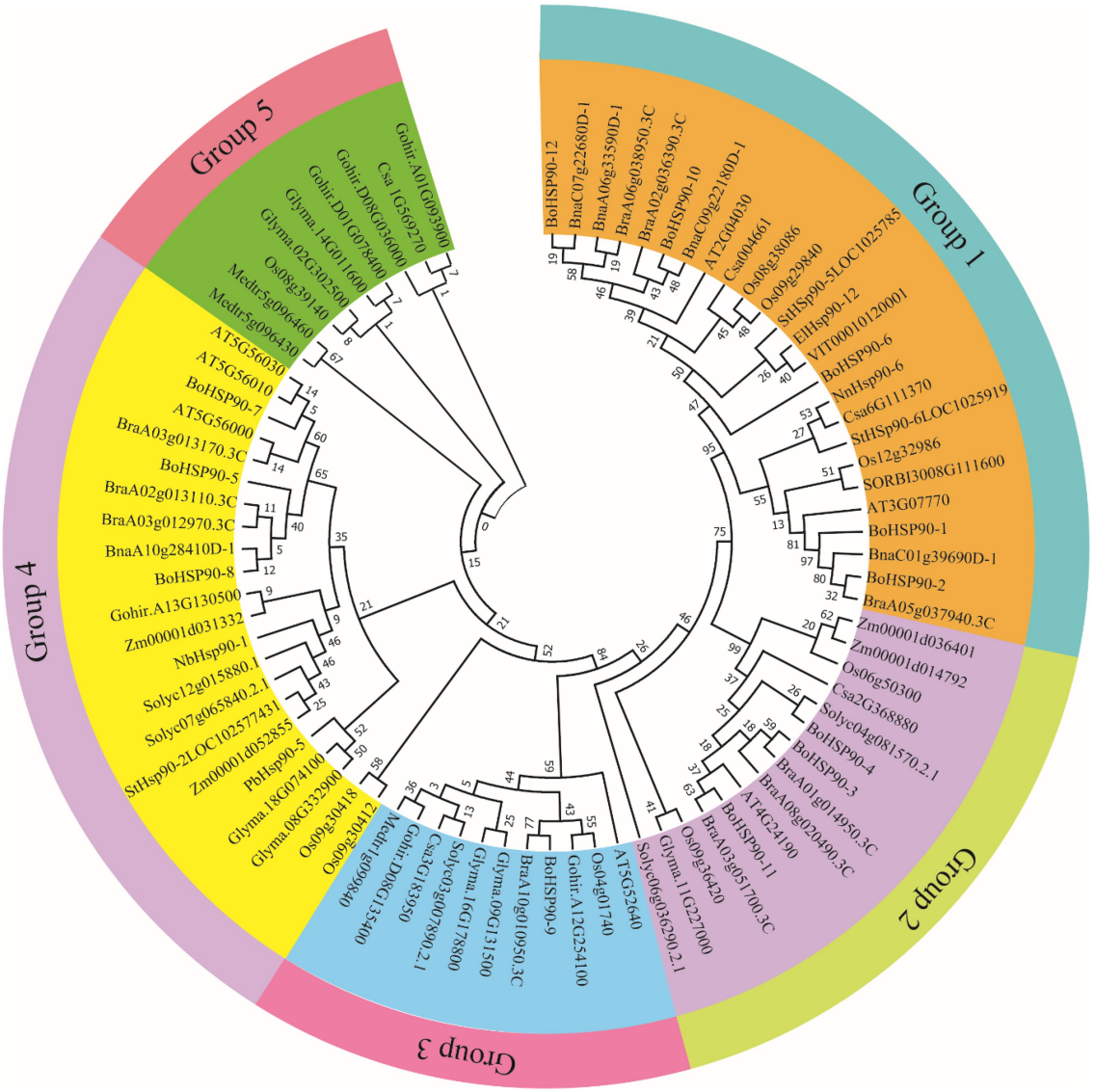


FIGURE 1 | Phylogenetic tree of the Hsp90 proteins from *Brassica oleracea*, *Brassica rapa*, *Brassica napus*, tomato, cucumber, potato, *Vitis vinifera*, *Sorghium bicolor*, cotton, tobacco, *Arabidopsis*, rice, *Zea mays*, soybean, *Medicago truncatula, Elaeis guineensis*, Chinese white pear (*Pyrus bretschneideri*) and *Nelumbo nucifera*"

4 contains two genes. Furthermore, we also explored the collinearity relationships between the Cabbage *BoHsp90* genes and related genes from 2 representative species including *A. thaliana* and *B. rapa*. A total of 12 cabbage *Hsp90* genes have collinearity relationships with 5 *A. thaliana* genes. 10 *B. rapa* genes.

Duplication, Ka/Ks and Evolutionary Analysis of the *BoHsp90* Genes in *Brassica oleracea*

The Ka/Ks ratio can be used to practice the selection process history of coding sequences "(Li et al., 1981). To explore the bias of duplicated *BoHsp90* members, the Ka, Ks and Ka/Ks ratios of Hsp90 paralogous members were determined." The Ks values were used to estimate each gene's fraction. "All of the paralogous Hsp90 pairs had a Ks value ranging from 0.1 to 0.3, which was made up of cabbage duplication events. In addition, the third pair of *BoHspP90* genes obtained duplications from 0.28 to 13.05 Mya (**Table 1**). The Ka/Ks ratio can also be used to determine selection pressure. The value of Ka/Ks1, greater than or equal to one, shows that the paralogous pair was subjected to, positive, purifying and neutral selection, or rare selection, respectively (Juretic et al., 2005)." Consequently, all the values were smaller than one, the representative that the *BoHsp90* genes changed largely under the influence of purifying selection (**Table 1**). "Consistent evolution of amino acid positions is imperative for conservation of function and protein structure." Certain amino acid identifications may be useful in elucidating selection pressures and assessing the

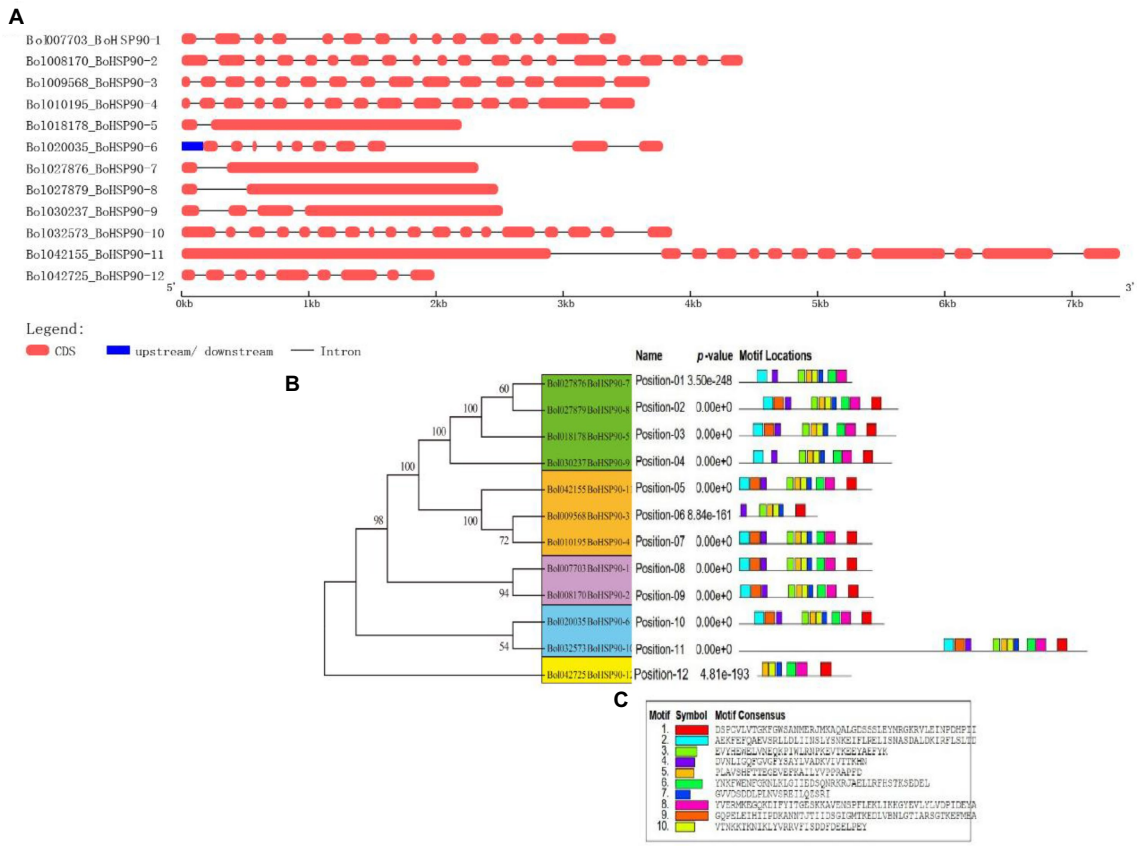


FIGURE 2 | Gene structure and conserved motif analysis of *BoHsp90* genes. **(A)** Gene structure of the *BoHsp90* genes in cabbage. **(B)** Motif composition of *BoHsp90*. **(C)** A conserved motif in *BoHsp90* gene was detected with MEME. Ten different motifs are represented by variously colored boxes.

relevance of specific amino acid interactions in Hsp90 protein interactions. The MEC model's positive selection test for Hsp90 protein estimates the selection pressure of a given coding region. Our research shows that 38.76% (271 out of 699) amino acids are purifying selected, while the rest are affected by the neutral process (**Figure 4**). The results clarify the evolution all gene pairs are affected by purification selection.

Analysis of the Promoter Region of *Brassica oleracea Hsp90* genes

Each *BoHsp90* gene was isolated from the *Brassica* plant genome, and the cis-regulatory region (1,500 bp upstream) was assessed using PlantCARE to better understand its possible functions and regulatory processes. The 12 *BoHsp90* promoter region of contain 50% of light-responsive elements, responsive to hormones 29%, environmental responsive elements 7%, development-related elements contain 11% and site binding elements 2% (**Figure 5A**). "There are various transcription factor binding sites that govern hormones such as gibberellic acid (GA), abscisic acid (ABA), jasmonic acid, auxin, silicic acid, and ethylene. In addition to the previously mentioned regions (**Figure 5B**) or by stress response (defense, heat, low-temperature stress, etc). The *BoHsp90-3* promoter region contains a maximum of 11 ABRE-related elements and 3 for jasmonic acid-responsive elements (**Supplementary Table S3**).

Tissue-Specific Expression Profile Analysis of *Brassica oleracea* the *Hsp90* Genes

The expression profile of the *BoHsp90* gene was generated using RNA-seq data from seven tissues (root, stem, leaf, shoot, flower, callus, and silique). The expression pattern was derived from publicly accessible Illumina RNA-Seq data for Cabbage at NCBI (Liu et al., 2014).[17] The expression of the *BoHsp90* gene was found to be higher in stem and silique tissues than in other tissues (**Figure 6**). The 12 predicted *BoHsp90* genes were found to be active in at least one of the six tissues. The expression of *BoHsp90–9* and *BoHsp90–5* was found to be high in the stem, silique, leaves, and flowers. Two genes were upregulated in stem, silique, leaves, and flowers (*BoHsp90-9* and *BoHsp90-5*). "Target genes involved in plant growth and development have been shown to be affected by genes that are highly expressed in plant tissues or organs. In this study, we found development-related *BoHsp90* genes in the cabbage genome at the transcription level, and two genes (*BoHsp90–5*, and *9*) were selected in stem, leaves, silique, and flowers that were upregulated, downregulated, or not different in other tissues according to normalized FPKM values. (**Supplementary Table S4**).

[17]www.ncbi.nlm.nih.gov/gds/?term=GSE42891

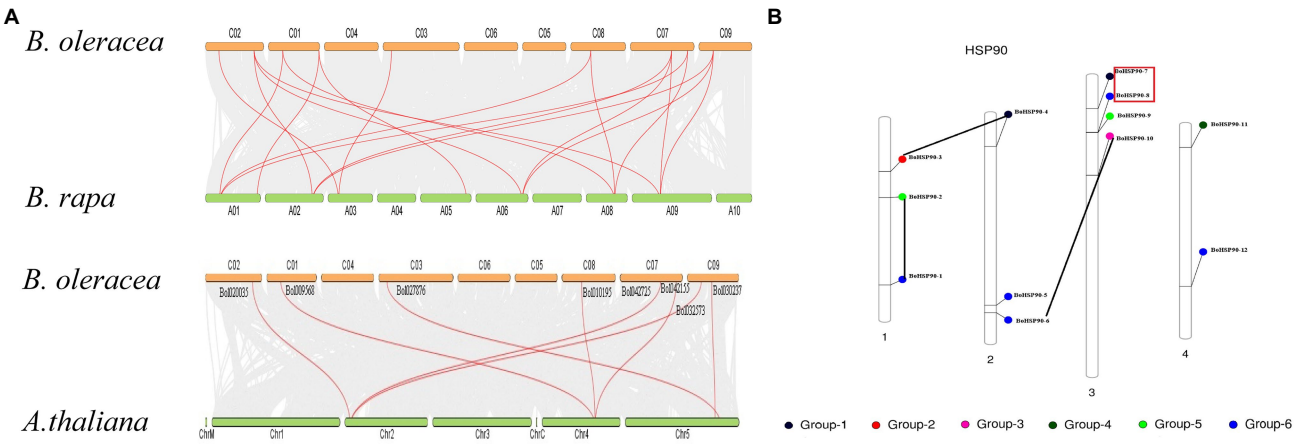


FIGURE 3 | Chromosomal location and duplication of Cabbage (*B. oleracea) Hsp90* genes. **(A)** Red boxes indicate tandem duplications, and black lines indicate segmental duplications. **(B)** Collinearity relationship or Synteny analysis of *Hsp90* genes in *B. oleracea*, *B. rapa*, and *Arabidopsis*.

TABLE 1 | Duplicated *BoHsp90* genes in cabbage.

#	Paralogous pair	Ka	Ks	KaKs	Time (Mya)
1	*Bol027876(BoHsp90-7)-Bol027879(BoHsp90-8)*	0.007	0.186	0.042	6.21
2	*Bol009568(BoHsp90-3)-Bol010195(BoHsp90-4)*	0.023	0.309	0.076	10.32
3	*Bol020035(BoHsp90-6)-Bol032573(BoHsp90-10)*	0.110	0.391	0.281	13.05

Expression Profiles Analysis of *Brassica oleracea Hsp90* Genes Under Cold Stress

To confirm the expression patterns determined by the qRT-PCR was employed to analyze the expression patterns of 12 different *BoHsp90* genes under cold stress conditions. Different expression profiles of 12 *BoHsp90* genes in response to cold stress were identified. (**Figure 7**). The result showed that most of the genes showed high expression levels at 4°C for 12 h intervals. Besides the expression level of *BoHsp90-2, BoHsp90-3, BoHsp90-7, BoHsp90-9, BoHsp90-2, and BoHsp90-11* were highly upregulated at 4°C after 12 h exposure of cold stress. we have further noticed that expression patterns of genes, including, *BoHsp90-2*, *BoHsp90-3*, *BoHsp90-4*, *BoHsp90-5*, *BoHsp90-7*, *BoHsp90-8*, *BoHSP90-10*, and *BoHSP90-12* were gradually downregulated as time passed. Similarly, *BoHSP90-1*, *BoHsp90-4*, *BoHsp90-6*, *BoHsp90-9* were highly upregulated at −2 and 0°C for 24 h intervals, but these genes showed downregulation trend at 6 h. The findings from this study exhibited that the *B.oleracea Hsp90* genes family, which are responded to cold stress promptly and with a long response time.

DISCUSSION

The unpredictability of global temperatures is followed by an increase in the frequency of cold spells, which is harmful for crop growth and effected the yields, and understanding the response of plants to low-temperature stress is very important for plant growth (Ding et al., 2019). When the temperature fluctuates (rises and falls) in comparison to normal, the physiological and biochemical reactions in the organism are inhibited, and cold stress inhibits the majority of protein and mRNA transcription. However, temperature stress (high and low) typically alters gene expression in plants, and an important type of highly conserved protein known as HSPs is rapidly synthesized. The Hsp90 is a chaperone protein that is highly conserved in prokaryotes and all eukaryotes and is involved in the cell cycle, signaling, and other key biological processes (Pearl and Prodromou, 2006). HSPs were first identified in the salivary gland chromosomes of Drosophila larvae (Ritossa, 1962). The function of *Hsp90* genes has been extensively studied in animal and fungal systems to date, but not widely studied in plants. However, with the continuous advancement and wider application of genome sequencing technologies, more and more plant genome data for Hsp90 have been released recently (Kersey, 2019). The *Hsp90* gene was found to be involved not only in stress signal transduction, receptor folding, transcription factors and kinases, and physiological processes in plants (Chory and Wu, 2001; Mach, 2009; Cha et al., 2013; Han et al., 2015) but also in assisting cell survival in stressful situations (Jackson et al., 2004). The

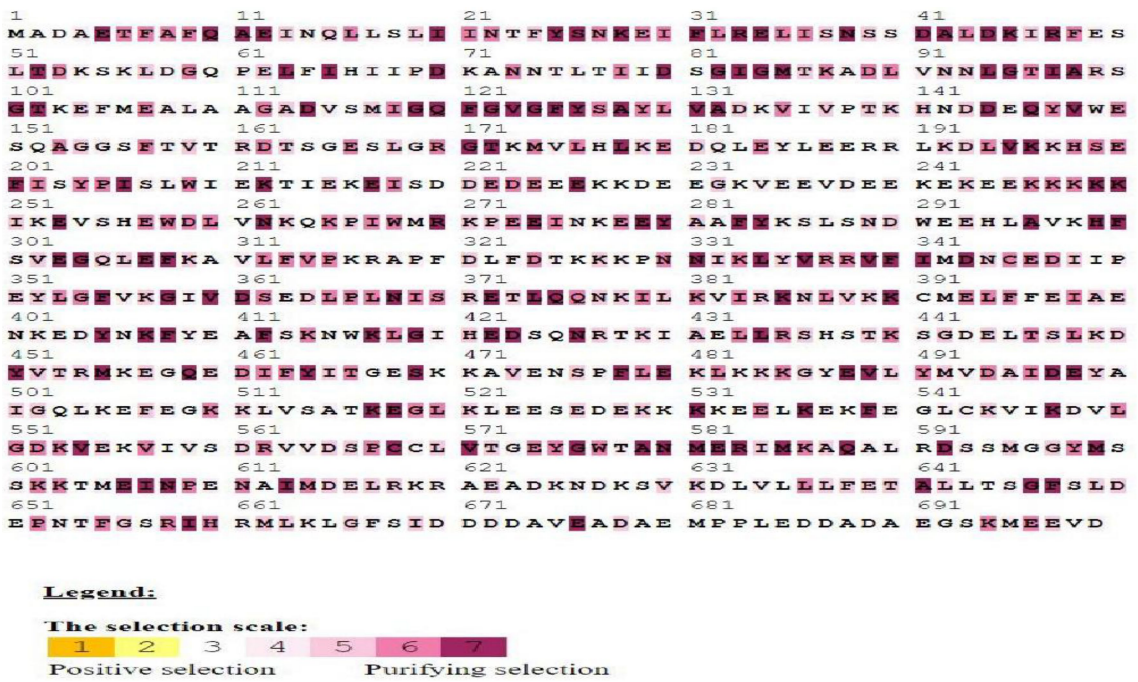


FIGURE 4 | Selection pressures among *Hsp90* gene sequences using mechanistic-empirical combination (MEC) model of selection online tool. Gray and white highlights represent neutral selection, and purple highlight represents purifying or negative selection on codons."

Hsp90 gene family has been identified in two model plants, *Arabidopsis* and rice (Hu et al., 2009). Subsequently, the *Hsp90* gene family was discovered in some major vegetable crops, such as tomato (Liu et al., 2014) and pepper (Jing et al., 2020). Cabbage is a very important vegetable crop cultivated in many areas worldwide. *B. oleracea* seedlings must be vernalized at 0–10°C and then licked, blossomed, and bore fruit under long sunshine and moderate temperature. Due to its characteristics of different maturity, suitable for different seasons, wide adaptability, and strong stress resistance, *B. oleracea* occupies an important position in the annual production of vegetables (Zhu et al., 2017). But the low temperature was one of the most important limiting factors for the wintering cultivation of *B. oleracea* in open fields; during winter *B. oleracea* often suffers from low temperatures and even sub-zero temperatures. The low-temperature conditions make its yield and quality decline. However, *Hsp90* gene family has not been characterized in cabbage (*B. oleracea*). Therefore, we investigated the whole genome of the *BoHsp90* gene family in cabbage, including phylogeny, chromosomal location, gene structure, conserved motifs, and expression profiles. A total of 12 *Hsp90* genes in cabbage were identified in this study, which are greater than other plant species such as 7 in *Arabidopsis* (Krishna and Gloor, 2001), 9 in rice (Hu et al., 2009), 10 in poplar (Zhang et al., 2013), 7 in tomato (Liu et al., 2014), and 7 in pepper (Jing et al., 2020). This could be the one of the main reasons for the difference in cold tolerance ability among different plant species. In this study, using the *Arabidopsis* Hsp90 protein sequence, 12 *Hsp90* genes were retrieved from the *Brassica* plant genome. These Hsp90 family proteins are vital in acclimatizing unfavorable environmental conditions through regulation and maintenance of physiological mechanism of cabbage plant (Zhang et al., 2021). The various physicochemical properties of the Hsp90 family proteins imply that there is a great deal of diversity among members, which will help in further research into the functions of *Hsp90* genes. Different features of the Hsp90 protein family, such as molecular weight, isoelectric point, and the number of exon-introns, were investigated in cabbage in this work. "Furthermore, all Hsp90 proteins in *B. oleracea* were acidic nature, which was similar with findings from *A. thaliana*, tomato, and other plants (Liu et al., 2014). These results suggested that Hsp90 proteins were conserved across plant species. The amino acid sequences of Hsp90 family proteins provided phylogenetic information based on their basic structures. In different *BoHsp90* gene sequences, we found different numbers of introns (1–19 introns). The fewer introns a plant has, the better its capability to respond to various environmental stimuli and developmental processes (Jin et al., 2014). Long-term evolution has resulted in a high number of introns in *BoHsp90s*. Song et al. (2019) revealed that conserved motif analysis of *Hsp90* genes in tobacco, *Nitab4.5_0001622g0050* and *Nitab4.5_0003328g0120* consist of fewer motifs, hinting that it had lost sequencing parts during evolution. Referring to previous investigations, the cabbage *Hsp90* genes' conserved motif analysis, *Bol009568. BoHsp90-3* and *Bol42725.BoHsp90-12* shown fewer motifs,

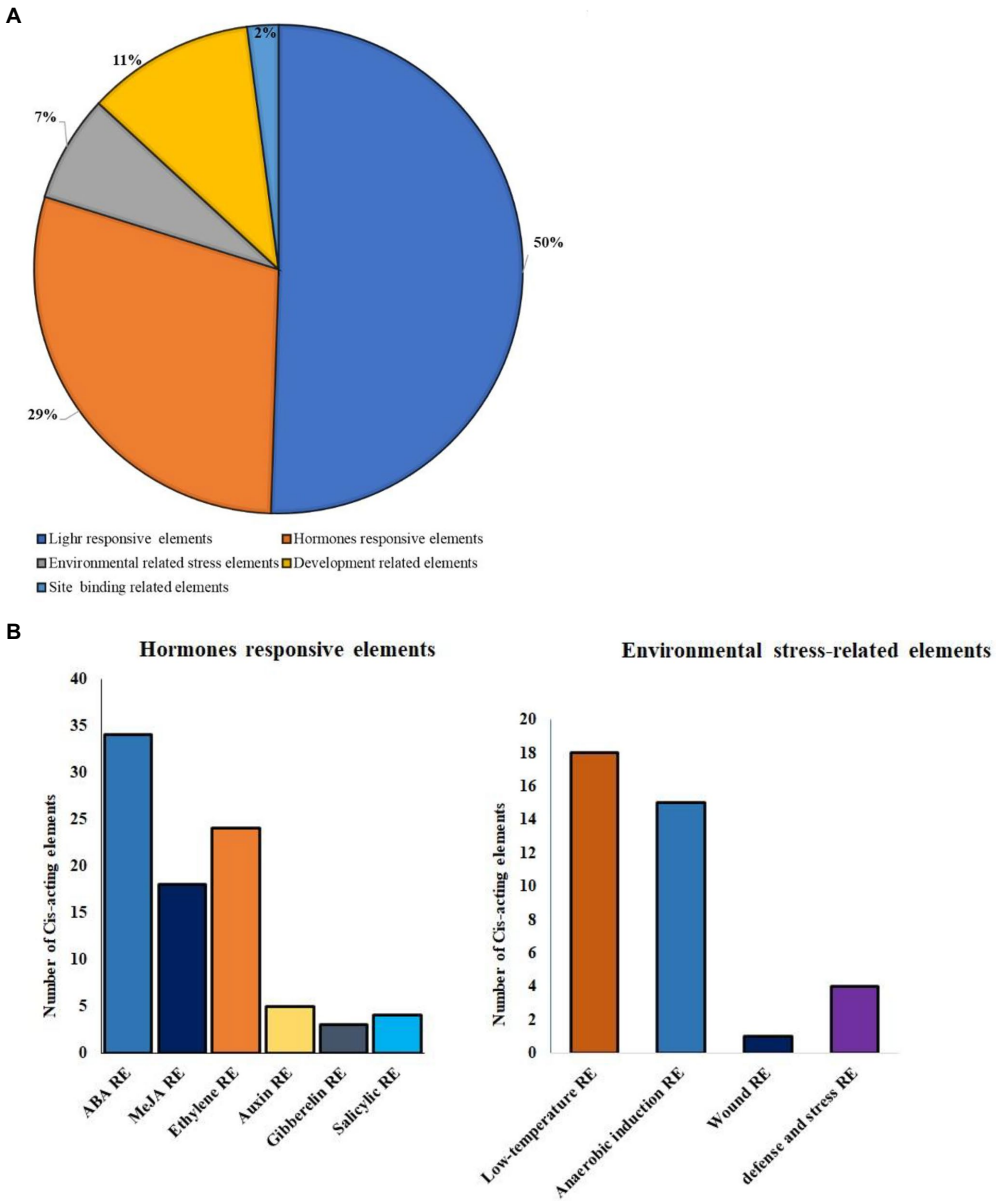


FIGURE 5 | Distribution of cis-elements in the promoter of *B. oleracea Hsp90* genes. **(A)** The proportion of cis-elements in the promoters of *B. oleracea Hsp90* genes are are shown in a pie chart. **(B)** Different hormones responsive elements and environmental stress-related elements in *12 BoHsp90* genes cis-elements regions.

which might be due to misplaced sequence sections during evolution. In addition, *BoHsp90-3* and *BoHsp90-12* contain 6 motifs, *BoHsp90-7* has 8 motifs, and 9 motifs were found in *BoHsp90-9*. These findings suggested that the structural motifs are aligned differently in various members of the *BoHsp90* family genes but are similar within closely related genes.

Gene duplication is a crucial mechanism in the evolution of gene families (Yang et al., 2008). In cabbage, at least three pairs of duplicated genes were discovered, representing that gene replication may have occurred throughout the evolution of the *Hsp90* genes. Phylogenetic analysis is frequently used to learn more about species' evolutionary changes and to discover the orthologs and paralogs within species (Moreira and Philippe, 2000). An unrooted phylogenetic tree was constructed using protein sequences from *B. oleracea*, *B. rapa*, *B. napus*, tomato, cucumber, potato, *V. vinifera, Sorghium bicolor,* cotton, tobacco, *Arabidopsis*, rice, *Z. mays*, soya bean, *E. guineensis*, Chinese white pear (*Pyrus bretschneideri*) and *N. nucifera*. Hsp90*s* can be split into five groups based on evolutionary study. Protein function is determined by its structure (Skolnick and Fetrow, 2000).

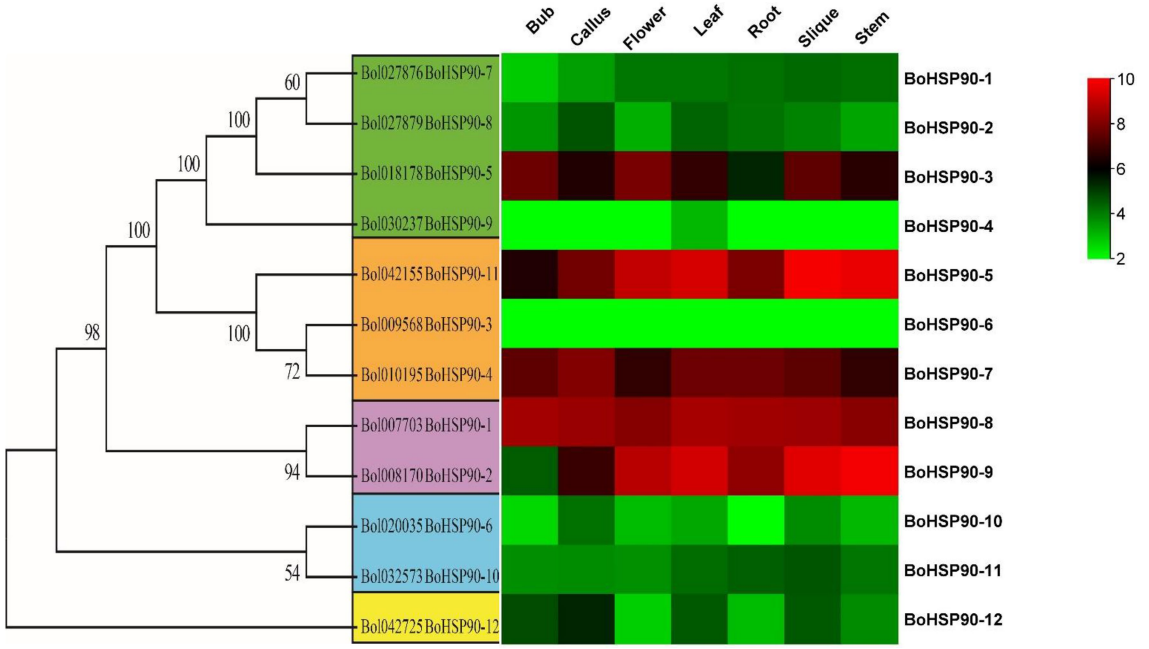


FIGURE 6 | The heat map of 12 *BoHsp90* genes was constructed using Tbtools software. The FPKM values were log2 transformed and mean-centered by genes using the scale 10 for minimum expression (red color) and 2 (green color) for maximum to make a heat map visible for each cabbage tissue.

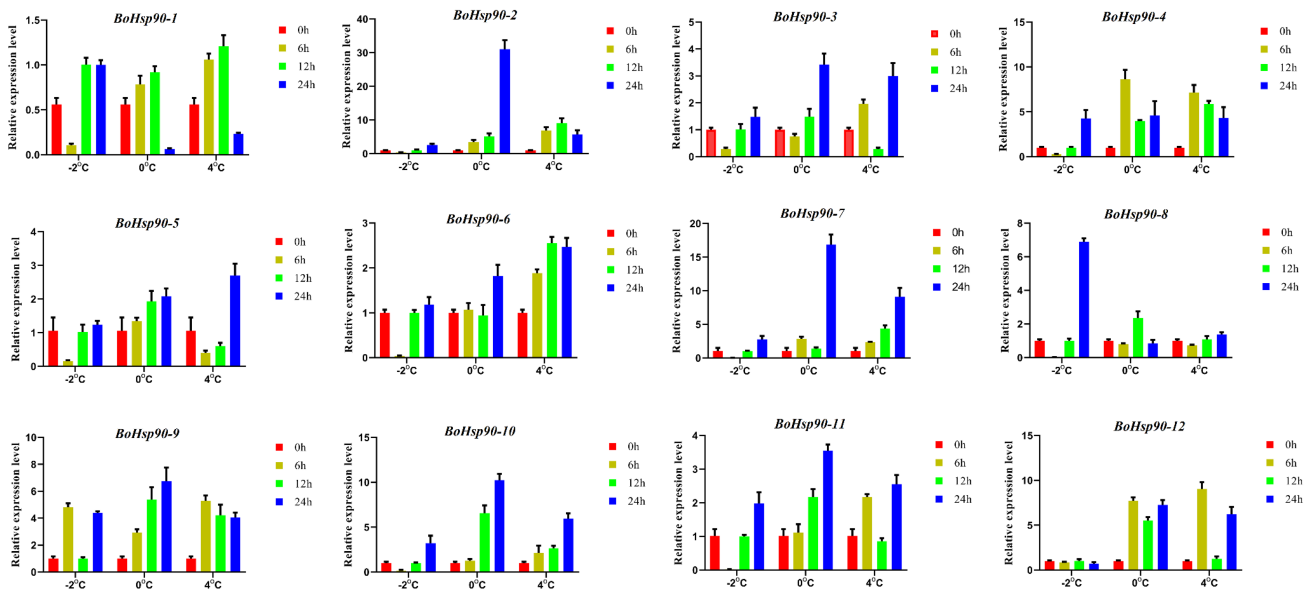


FIGURE 7 | Relative expression levels of 12 *Hsp90* genes in *B. oleracea* leaves by qRT-PCR after 0 h (control), 6 h, 12 h and 24 h after treatment of cold stress at −2°C, 0°C, and 4°C. The expression levels of *Hsp90* genes were normalized to *BoActin* using as an internal control. Error bars indicate the standard error (SE) between three replicates.

According to the collinearity relationships between the *B. orleacea BoHsp90* family genes and 2 representative plant species including A. thalian and *B. rapa*, only three pair of paralogous pair (*Bol027876BoHsp90-7-Bol027879. BoHsp90-8, Bol01019568.BoHsp90-3-Bol010195Hsp90.4 and Bol020035.BoHsp90-6-Bol032573.BolHsp90-10*) were perceived and segmental duplication were not found, representing that enlargement of *Hsp90* gene in *B. oleracea* was mainly caused

by tandem duplication genes. Only four *B. oleracea Hsp90* genes *Bol7703-BoHsp90-1, Bol008170.BolHsp90-2, Bol018178. BoHsp90-5 and Bol027879.BoHsp90-8* was conservative and not collinearity relation with *Hsp90* genes in *B. rapa* and *Arabidopsis*. However, eight *B. oleracea Hsp90* genes have collinearity relationships with 5 genes of *A. thaliana* and 10 genes of *B. rapa*. These findings suggested that the *Hsp90* gene family expands differently in each species, as this pattern was also noticed in other plant gene families (Huang et al., 2016). Furthermore, promoter analysis revealed that the promoter of the *B. oleracea Hsp90* gene contained various regulatory elements related to light-responsive elements, responsive to hormones, environmental responsive elements, development-related elements, and site binding elements, among others, implying that these regulatory elements actively engage in a variety of biotic and abiotic stress responses in plants and also indicating that *Hsp90* genes are important in cell-cycle control, signal transduction, genomic silencing, protein degradation, and protein trafficking.

Recent studies have shown that Hsp90 has a vital role in plant acclimation to various biotic and abiotic stresses (Di Donato and Geisler, 2019). The *Hsp* gene family was considered to be involved or induced under cold environments (Krishna et al., 1995); it has the potential to protect plant biological systems from cold damage by preventing freezing-induced protein denaturation (Yan et al., 2006). In this study, the publicly online tissue specific expression analysis showed that the *BoHsp90* gene was found to be higher in stem and silique tissues than other tissues. Similarly, the expression of *BoHsp90–9* and *BoHsp90–5* was found to be high in the stem, silique, leaves, and flowers. While upregulated in stem, silique, leaves and flowers. These findings indicated that the functions of *B. oleracea Hsp90* genes varied across tissues. Many studies indicated that *Hsp90* gene family induced under low-temperature stress (Krishna et al., 1995; Di Donato and Geisler, 2019).

Gene expression patterns were highly correlated with their functions. The tissue-specific expression patterns at a given developmental stage are crucial for determining the gene functions in which they are involved. The expression pattern of different genes was studied in our work under diverse stresses and distinct cabbage tissues to better understand their possible significance during cold stress and development responses. The majority of the genes investigated exhibited expression in developmental tissues, indicating their function in these tissues (**Figure 7**). However, the *Hsp90* gene family in cabbage has not been investigated to date. Thus, to understand the functions of *Hsp90* genes under cold stress, we accomplish qRT-PCR of 12 different genes under cold stress. These results indicated that expressions of *Hsp90* genes family, *BoHsp90-1/6/8/9/11*, *BoHsp90-1*, and *BoHsp90-9* were highly upregulated at 4°C, 0°C, and -2°C respectively, while remaining were down-regulated at respective low temperatures for 6 h duration. In the same manner, gene expressions of *BoHsp90-2/3/5/6/7/9/10/11/12*, *BoHsp90-1/4/9/10/11/12*, and *BoHsp90-1* were significantly upregulated at 4°C, 0°C, and -2°C, respectively; while *BoHsp90-1/4/8*, *BoHsp90-2/3/6/7/8*, *BoHsp90-2/3/4/5/6/7/8/9/10/11/12* were downregulated at 4°C, 0°C, and -2°C, respectively, for 12 h duration. Lastly for 24 h duration, gene expressions of *BoHsp90-3/5/6/7/10/11/12*, *BoHsp90-1/4/6/9*, and *BoHsp90-1/3/4/6/9/11/12* were upregulated at 4°C, 0°C, and -2°C, respectively; while *BoHsp90-1/2/4/8/9*, *BoHsp90-2/3/5/7/8/10/11/12*, and *BoHsp90-2/5/7/8/10* were downregulated at 4°C, 0°C, and -2°C, respectively. Previous studies also reported that Hsp90 protein accumulation was found to be altered in cold-treated winter wheat followed cold treatment (Vítámvás et al., 2012). Overall, the cabbage *Hsp90* gene family was found to be differentially expressed under cold stress, suggesting that these genes play an important role in cabbage growth and development under cold conditions. We found the evolution of *Hsp90* genes in cabbage using comparative genomics analysis, and we screened putative regulatory genes associated with cold stress. The biological activities of these genes will be examined in the subsequent study, which will give an essential theoretical foundation for increasing cabbage quality.

CONCLUSION

In this study, *Hsp90* family genes were identified and characterized in the *Brassica oleracea*. A total of 12 *Hsp90* genes of *Brassica* were identified and characterized using phylogenetic tree, gene structure, physicochemical characteristics, conserved motif analysis, chromosomal location, synteny, selection pressure, homologous gene pairs, and cis-elements in the promoters, which revealed a rich evolutionary history for this *Brassica* family. The expression pattern analysis of *B. oleracea BoHsp90* exhibited that *BoHsp90-2, BoHsp90-3, BoHsp90-7, BoHsp90-9, BoHsp90-10,* and *BoHsp90-11* were induced under cold stress, which indicates these *Hsp90* genes perform a vital role in cold acclimation and supports in the continual of normal growth and development process. Our findings provide a scientific foundation for more comprehensive understanding of the cabbage *Hsp90* gene family, and it will also contribute to the development of new high yielding and stress tolerant cultivars of *B. oleracea*.

DATA AVAILABILITY STATEMENT

Publicly available datasets were analyzed in this study. This data can be found here: https://www.ncbi.nlm.nih.gov/gds/?term=GSE42891.

AUTHOR CONTRIBUTIONS

SS and SJ conceived and designed the work. MA, QD, YL, MH, and CT performed the experiments and conduct the bioinformatics analysis. SS and JS wrote and revised the manuscript. All authors contributed to the article and approved the submitted version.

FUNDING

This study was financially supported by the National Natural Science Foundation of China (31872108, 31272170) and the Key Research and Development Projects in Anhui Province (No. 202104b11020006).

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.908511/full#supplementary-material

REFERENCES

Ahmed, N. U., Jung, H.-J., Park, J.-I., Cho, Y.-G., Hur, Y., and Nou, I.-S. (2015). Identification and expression analysis of cold and freezing stress responsive genes of *Brassica oleracea*. *Gene* 554, 215–223. doi: 10.1016/j.gene.2014.10.050

Al-Whaibi, M. H. (2011). Plant heat-shock proteins: a mini review. *J. King Saud Univ.-Sci.* 23, 139–150. doi: 10.1016/j.jksus.2010.06.022

Bailey, T. L., Williams, N., Misleh, C., and Li, W. W. (2006). MEME: discovering and analyzing DNA and protein sequence motifs. *Nucleic Acids Res.* 34, W369–W373. doi: 10.1093/nar/gkl198

Cha, J.-Y., Ahn, G., Kim, J. Y., Kang, S. B., Kim, M. R., Su'udi, M., et al. (2013). Structural and functional differences of cytosolic 90-kDa heat-shock proteins (Hsp90s) in *Arabidopsis thaliana*. *Plant Physiol. Biochem.* 70, 368–373. doi: 10.1016/j.plaphy.2013.05.039

Chen, B., Zhong, D., and Monteiro, A. (2006). Comparative genomics and evolution of the *HSP90* family of genes across all kingdoms of organisms. *BMC Genomics* 7, 1–19. doi: 10.1186/1471-2164-7-156

Chory, J., and Wu, D. (2001). Weaving the complex web of signal transduction. *Plant Physiol.* 125, 77–80. doi: 10.1104/pp.125.1.77

Di Donato, M., and Geisler, M. (2019). HSP 90 and co-chaperones: a multitaskers' view on plant hormone biology. *FEBS Lett.* 593, 1415–1430. doi: 10.1002/1873-3468.13499

Ding, Y., Shi, Y., and Yang, S. (2019). Advances and challenges in uncovering cold tolerance regulatory mechanisms in plants. *New Phytol.* 222, 1690–1704. doi: 10.1111/nph.15696

Dutta, R., and Inouye, M. (2000). GHKL, an emergent ATPase/kinase superfamily. *Trends Biochem. Sci.* 25, 24–28. doi: 10.1016/S0968-0004(99)01503-0

Food and Agriculture Organization [FAO] (2020). *Responding to the Impact of the COVID-19 Outbreak on Food Value Chains Through Efficient Logistics*. Rome: Food and Agriculture Organization.

Food and Agricultural Organization of the United Nations (FAO) (2007). *FAOSTAT*. Rome: FAO.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R. D., and Bairoch, A. (2003). ExPASy: the proteomics server for in-depth protein knowledge and analysis. *Nucleic Acids Res.* 31, 3784–3788. doi: 10.1093/nar/gkg563

Gill, R. A., Ali, B., Islam, F., Farooq, M. A., Gill, M. B., Mwamba, T. M., et al. (2015). Physiological and molecular analyses of black and yellow seeded *Brassica napus* regulated by 5-aminolivulinic acid under chromium stress. *Plant Physiol. Biochem.* 94, 130–143. doi: 10.1016/j.plaphy.2015.06.001

Goodstein, D. M., Shu, S., Howson, R., Neupane, R., Hayes, R. D., Fazo, J., et al. (2012). Phytozome: a comparative platform for green plant genomics. *Nucleic Acids Res.* 40, D1178–D1186. doi: 10.1093/nar/gkr944

Gu, L., Hanson, P. J., Post, W., Mac Kaiser, D. P., Yang, B., Nemani, R., et al. (2008). The 2007 eastern US spring freeze: increased cold damage in a warming world? *Bioscience* 58, 253–262. doi: 10.1641/B580311

Han, Y., Chen, Y., Yin, S., Zhang, M., and Wang, W. (2015). Over-expression of TaEXPB23, a wheat expansin gene, improves oxidative stress tolerance in transgenic tobacco plants. *J. Plant Physiol.* 173, 62–71. doi: 10.1016/j.jplph.2014.09.007

Hassan, M. A., Xiang, C., Farooq, M., Muhammad, N., Yan, Z., Hui, X., et al. (2021). Cold stress in wheat: plant acclimation responses and management strategies. *Front. Plant Sci.* 12:676884. doi: 10.3389/FPLS.2021.676884

Hu, W., Hu, G., and Han, B. (2009). Genome-wide survey and expression profiling of heat shock proteins and heat shock factors revealed overlapped and stress specific response under abiotic stresses in rice. *Plant Sci.* 176, 583–590. doi: 10.1016/j.plantsci.2009.01.016

Hu, B., Jin, J., Guo, A.-Y., Zhang, H., Luo, J., and Gao, G. (2015). GSDS 2.0: an upgraded gene feature visualization server. *Bioinformatics* 31, 1296–1297. doi: 10.1093/bioinformatics/btu817

Huang, Z., Duan, W., Song, X., Tang, J., Wu, P., Zhang, B., et al. (2016). Retention, molecular evolution, and expression divergence of the auxin/indole acetic acid and auxin response factor gene families in *Brassica rapa* shed light on their evolution patterns in plants. *Genome Biol. Evol.* 8, 302–316. doi: 10.1093/gbe/evv259

IPCC (2021). *Climate Change 2021: The Physical Science Basis*. Cambridge: IPCC.

Jackson, S. E., Queitsch, C., and Toft, D. (2004). Hsp90: from structure to phenotype. *Nat. Struct. Mol. Biol.* 11, 1152–1155. doi: 10.1038/nsmb1204-1152

Jin, Z., Chandrasekaran, U., and Liu, A. (2014). Genome-wide analysis of the Dof transcription factors in castor bean (*Ricinus communis* L.). *Genes Genomics* 36, 527–537. doi: 10.1007/s13258-014-0189-6

Jing, W., Fangjun, T., Chengliang, L., Xilu, Z., Lijun, O., Juntawong, N., et al. (2020). Genome-wide identification and analysis of *HSP90* gene family in pepper. *Acta Horticult. Sin.* 47, 665–674. doi: 10.16420/j.issn.0513-353x.2019-0470

Juretic, N., Hoen, D. R., Huynh, M. L., Harrison, P. M., and Bureau, T. E. (2005). The evolutionary fate of MULE-mediated duplications of host gene fragments in rice. *Genome Res.* 15, 1292–1297. doi: 10.1101/gr.4064205

Kersey, P. J. (2019). Plant genome sequences: past, present, future. *Curr. Opin. Plant Biol.* 48, 1–8. doi: 10.1016/j.pbi.2018.11.001

Koch, M. A., Haubold, B., and Mitchell-Olds, T. (2000). Comparative evolutionary analysis of chalcone synthase and alcohol dehydrogenase loci in *Arabidopsis*, Arabis, and related genera (Brassicaceae). *Mol. Biol. Evol.* 17, 1483–1498. doi: 10.1093/oxfordjournals.molbev.a026248

Krishna, P., and Gloor, G. (2001). The Hsp90 family of proteins in *Arabidopsis thaliana*. *Cell Stress Chaperones* 6, 238–246. doi: 10.1379/1466-1268(2001)006<0238:THFOPI>2.0.CO;2

Krishna, P., Sacco, M., Cherutti, J. F., and Hill, S. (1995). Cold-induced accumulation of *hsp90* transcripts in *Brassica napus*. *Plant Physiol.* 107, 915–923. doi: 10.1104/pp.107.3.915

Lachowiec, J., Lemus, T., Borenstein, E., and Queitsch, C. (2015). Hsp90 promotes kinase evolution. *Mol. Biol. Evol.* 32, 91–99. doi: 10.1093/molbev/msu270

Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van De Peer, Y., et al. (2002). PlantCARE, a database of plant cis-acting regulatory elements and a portal to tools for in silico analysis of promoter sequences. *Nucleic Acids Res.* 30, 325–327. doi: 10.1093/nar/30.1.325

Li, W.-H., Gojobori, T., and Nei, M. (1981). Pseudogenes as a paradigm of neutral evolution. *Nature* 292, 237–239. doi: 10.1038/292237a0

Liu, Y., Dang, P., Liu, L., and He, C. (2019). Cold acclimation by the CBF-COR pathway in a changing climate: Lessons from Arabidopsis thaliana. *Plant Cell Rep* 38, 511–519. doi: 10.1007/s00299-019-02376-3

Liu, Y., Wan, H., Yang, Y., Wei, Y., Li, Z., Ye, Q., et al. (2014). Genome-wide identification and analysis of heat shock protein 90 in tomato. *Yi chuan=Hereditas* 36, 1043–1052. doi: 10.3724/SP.J.1005.2014.1043

Liu, M., Wang, L., Ke, Y., Xian, X., Wang, J., Wang, M., et al. (2022). Identification of *HbHSP90* gene family and characterization *HbHSP90*. 1 as a candidate gene for stress response in rubber tree. *Gene*:827, 146475. doi: 10.1016/j.gene.2022.146475

Lv, H., Wang, Y., Han, F., Ji, J., Fang, Z., Zhuang, M., et al. (2020). A high-quality reference genome for cabbage obtained with SMRT reveals novel genomic features and evolutionary characteristics. *Sci. Rep.* 10, 1–9. doi: 10.1038/s41598-020-69389-x

Mach, J. (2009). Alternative splicing produces a JAZ protein that is not broken down in response to jasmonic acid. *Plant Cell* 21:14. doi: 10.1105/tpc.108.210111

Marchler-Bauer, A., Derbyshire, M. K., Gonzales, N. R., Lu, S., Chitsaz, F., Geer, L. Y., et al. (2015). CDD: NCBI's conserved domain database. *Nucleic Acids Res.* 43, D222–D226. doi: 10.1093/nar/gku1221

Masson-Delmotte, V., Zhai, P., Pirani, A., Connors, S., Péan, C., Berger, S., et al. (2021). "Climate change 2021: the physical science basis," in *Working Group I contribution to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change.* ambridge: Cambridge University Press.

Moreira, D., and Philippe, H. (2000). Molecular phylogeny: pitfalls and progress. *Int. Microbiol.* 3, 9–16. PMID: 10963328

Moreno, A. A., and Orellana, A. (2011). The physiological role of the unfolded protein response in plants. *Biol. Res.* 44, 75–80. doi: 10.4067/S0716-97602011000100010

Mukhopadhyay, I., Nazir, A., Saxena, D., and Chowdhuri, D. K. (2003). Heat shock response: hsp70 in environmental monitoring. *J. Biochem. Mol. Toxicol.* 17, 249–254. doi: 10.1002/jbt.10086

Pearl, L. H., and Prodromou, C. (2000). Structure and in vivo function of Hsp90. *Curr. Opin. Struct. Biol.* 10, 46–51. doi: 10.1016/S0959-440X(99)00047-0

Pearl, L. H., and Prodromou, C. (2006). Structure and mechanism of the Hsp90 molecular chaperone machinery. *Annu. Rev. Biochem.* 75, 271–294. doi: 10.1146/annurev.biochem.75.103004.142738

Picard, D. (2002). Heat-shock protein 90, a chaperone for folding and regulation. *Cell. Mol. Life Sci. CMLS* 59, 1640–1648. doi: 10.1007/pl00012491

Pratt, W. B., and Toft, D. O. (2003). Regulation of signaling protein function and trafficking by the Hsp90/hsp70-based chaperone machinery. *Exp. Biol. Med.* 228, 111–133. doi: 10.1177/153537020322800201

Prodromou, C., Roe, S. M., O'brien, R., Ladbury, J. E., Piper, P. W., and Pearl, L. H. (1997). Identification and structural characterization of the ATP/ADP-binding site in the Hsp90 molecular chaperone. *Cell* 90, 65–75. doi: 10.1016/S0092-8674(00)80314-1

Punta, M., Coggill, P. C., Eberhardt, R. Y., Mistry, J., Tate, J., Boursnell, C., et al. (2012). The Pfam protein families database. *Nucleic Acids Res.* 40, D290–D301. doi: 10.1093/nar/gkr1065

Queitsch, C., Sangster, T. A., and Lindquist, S. (2002). Hsp90 as a capacitor of phenotypic variation. *Nature* 417, 618–624. doi: 10.1038/nature749

Raman, S., and Suguna, K. (2015). Functional characterization of heat-shock protein 90 from Oryza sativa and crystal structure of its N-terminal domain. *Acta Crystallograph. Sect. F: Struct. Biol. Commun.* 71, 688–696. doi: 10.1107/S2053230X15006639

Ritossa, F. (1962). A new puffing pattern induced by temperature shock and DNP in Drosophila. *Experientia* 18, 571–573. doi: 10.1007/BF02172188

Skolnick, J., and Fetrow, J. S. (2000). From genes to protein structure and function: novel applications of computational approaches in the genomic era. *Trends Biotechnol.* 18, 34–39. doi: 10.1016/S0167-7799(99)01398-0

Song, Z., Pan, F., Yang, C., Jia, H., Jiang, H., He, F., et al. (2019). Genome-wide identification and expression analysis of *HSP90* gene family in *Nicotiana tabacum*. *BMC Genet.* 20, 1–12. doi: 10.1186/s12863-019-0738-8

Sung, N., Lee, J., Kim, J.-H., Chang, C., Joachimiak, A., Lee, S., et al. (2016). Mitochondrial Hsp90 is a ligand-activated molecular chaperone coupling ATP binding to dimer closure through a coiled-coil intermediate. *Pro. Nati. Acad. Sci.* 113, 2952–2957. doi: 10.1073/pnas.1516167113

Vítámvás, P., Prášil, I. T., Kosová, K., Planchon, S., and Renaut, J. (2012). Analysis of proteome and frost tolerance in chromosome 5A and 5B reciprocal substitution lines between two winter wheats during long-term cold acclimation. *Proteomics* 12, 68–85. doi: 10.1002/pmic.201000779

Wang, Y., Tang, H., Debarry, J. D., Tan, X., Li, J., Wang, X., et al. (2012). MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Res.* 40:e49. doi: 10.1093/nar/gkr1293

Wang, W., Vinocur, B., Shoseyov, O., and Altman, A. (2004). Role of plant heat-shock proteins and molecular chaperones in the abiotic stress response. *Trends Plant Sci.* 9, 244–252. doi: 10.1016/j.tplants.2004.03.006

Xiong, L., and Ishitani, M. (2006). "Stress signal transduction: components, pathways and network integration." in *Abiotic Stress Tolerance in Plants*. eds. M. Way and K. J. Green. (Berlin: Springer), 3–29.

Yamada, K., Fukao, Y., Hayashi, M., Fukazawa, M., Suzuki, I., and Nishimura, M. (2007). Cytosolic HSP90 regulates the heat shock response that is responsible for heat acclimation in *Arabidopsis thaliana*. *J. Biol. Chem.* 282, 37794–37804. doi: 10.1074/jbc.M707168200

Yan, S.-P., Zhang, Q.-Y., Tang, Z.-C., Su, W.-A., and Sun, W.-N. (2006). Comparative proteomic analysis provides new insights into chilling stress responses in rice. *Mol. Cell. Proteom.* 5, 484–496. doi: 10.1074/mcp.M500251-MCP200

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* 13, 555–556. doi: 10.1093/bioinformatics/13.5.555

Yang, Z., Zhou, Y., Wang, X., Gu, S., Yu, J., Liang, G., et al. (2008). Genomewide comparative phylogenetic and molecular evolutionary analysis of tubby-like protein family in *Arabidopsis*, rice, and poplar. *Genomics* 92, 246–253. doi: 10.1016/j.ygeno.2008.06.001

Yu, J., Tehrim, S., Zhang, F., Tong, C., Huang, J., Cheng, X., et al. (2014). Genome-wide comparative analysis of NBS-encoding genes between Brassica species and Arabidopsis thaliana. *BMC Genom.* 15, 1–18. doi: 10.1186/1471-2164-15-3

Zhang, K., He, S., Sui, Y., Gao, Q., Jia, S., Lu, X., et al. (2021). Genome-wide characterization of *HSP90* gene family in cucumber and their potential roles in response to abiotic and biotic stresses. *Front. Genet.* 12, 95. doi: 10.3389/fgene.2021.584886

Zhang, J., Li, J., Liu, B., Zhang, L., Chen, J., and Lu, M. (2013). Genome-wide analysis of the Populus *Hsp90* gene family reveals differential expression patterns, localization, and heat stress responses. *BMC Genomics* 14, S1–S14. doi: 10.1186/1471-2164-14-S8-S1

Zhu, X., Bo, T., Chen, J., Tai, X., and Ren, Y. (2017). Research progress on mapping of genes associated with main agronomic traits of cabbage. *Acta Horticult. Sin.* 44, 1729–1737.

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

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

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 27 September 2022
DOI 10.3389/fpls.2022.1003907

OPEN ACCESS

EDITED BY
Freddy Mora-Poblete,
University of Talca, Chile

REVIEWED BY
Samuel Henrique Kamphorst,
Universidade Estadual do Norte Fluminense Darcy Ribeiro, Brazil
Rodrigo Iván Contreras-Soto,
Universidad de O'Higgins, Chile

*CORRESPONDENCE
Lindsey Compton
l.j.compton@bham.ac.uk
Gisella Orjeda
morjedaf@unmsm.edu.pe

†These authors have contributed equally to this work and share senior authorship

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

RECEIVED 26 July 2022
ACCEPTED 06 September 2022
PUBLISHED 27 September 2022

CITATION
Ponce OP, Torres Y, Prashar A, Buell R, Lozano R, Orjeda G and Compton L (2022) Transcriptome profiling shows a rapid variety-specific response in two Andigenum potato varieties under drought stress.
Front. Plant Sci. 13:1003907.
doi: 10.3389/fpls.2022.1003907

Transcriptome profiling shows a rapid variety-specific response in two Andigenum potato varieties under drought stress

Olga Patricia Ponce[1], Yerisf Torres[2,3], Ankush Prashar[4], Robin Buell[5], Roberto Lozano[3,6], Gisella Orjeda[7*†] and Lindsey Compton[1*†]

[1]School of Biosciences, University of Birmingham, Birmingham, United Kingdom, [2]Department of Plant Science, Wageningen University, Wageningen, Netherlands, [3]Unidad de genómica, Laboratorios de Investigación y Desarrollo (LID), Universidad Peruana Cayetano Heredia, Lima, Peru, [4]School of Natural and Environmental Sciences, Newcastle University, Newcastle upon Tyne, United Kingdom, [5]Department of Crop & Soil Sciences, Institute for Plant Breeding, Genetics & Genomics, Center for Applied Genetic Technology, University of Georgia, Athens, GA, United States, [6]Digital Science and Technology Department, Joyn Bio LLC, Boston, MA, United States, [7]Faculty of Biological Sciences, Universidad Nacional Mayor de San Marcos, Lima, Peru

Potato is a drought-sensitive crop whose global sustainable production is threatened by alterations in water availability. Whilst ancestral *Solanum tuberosum* Andigenum landraces retain wild drought tolerance mechanisms, their molecular bases remain poorly understood. In this study, an aeroponic growth system was established to investigate stress responses in leaf and root of two Andigenum varieties with contrasting drought tolerance. Comparative transcriptome analysis revealed widespread differences in the response of the two varieties at early and late time points of exposure to drought stress and in the recovery after rewatering. Major differences in the response of the two varieties occurred at the early time point, suggesting the speed of response is crucial. In the leaves and roots of the tolerant variety, we observed rapid upregulation of ABA-related genes, which did not occur until later in the susceptible variety and indicated not only more effective ABA synthesis and mobilization, but more effective feedback regulation to limit detrimental effects of too much ABA. Roots of both varieties showed differential expression of genes involved in cell wall reinforcement and remodeling to maintain cell wall strength, hydration and growth under drought stress, including genes involved in lignification and wall expansion, though the response was stronger in the tolerant variety. Such changes in leaf and root may help to limit water losses in the tolerant variety, while limiting the reduction in photosynthetic rate. These findings provide insights into molecular bases of drought tolerance mechanisms and pave the way for their reintroduction into modern cultivars with improved resistance to drought stress and yield stability under drought conditions.

KEYWORDS

drought, potato, abiotic stress, rehydration, transcriptome

1 Introduction

As the human population is projected to approach 9 billion within four decades, the growing food gap necessitates at least a 50% increase in crop-based production to ensure food security (Grafton et al., 2015). Potato (*Solanum tuberosum* L.) is the fourth most important food crop after maize, wheat and rice, with an estimated annual tuber production of 370 million tonnes (FAO, 2021). The potato crop has highly desirable characteristics, including rich and balanced nutrition, high yields and adaptability to diverse cultivation environments, maintaining stable yields in marginal soil with limited labour inputs (Lutaladio and Castaldi, 2009; Scott, 2011; Zaheer and Akhtar, 2016). However, for potato as for other crops, drought presents a serious threat to food security as the climate changes, as it decreases crop growth and yield more than any other abiotic or biotic stress (George et al., 2018).

Modern potato varieties are particularly susceptible to periodic water shortages. Their shallow rooted systems have a weak soil penetration ability and poor nutrient uptake capacity, thus requiring consistent irrigation and making them especially susceptible to periodic droughts (Iwama, 2008). Drought conditions shorten the potato growth cycle, hamper growth and reduce the final tuber number (Deblonde and Ledent, 2001; Eiasu et al., 2007; Kumar et al., 2007). Low soil water potential can also reduce tuber quality by lowering dry matter concentration and increasing abundance of reducing sugars, resulting in "sugar ends" formation (Haverkort and Verhagen, 2008). A key goal is therefore the adaptation of existing potato varieties to drought conditions, including progressively more frequent and intense agricultural droughts due to the combined effects of growing evapo-transpiration demand and below-normal precipitation regimes (Monneveux et al., 2013). Without such adaptation, global yield losses are projected to range between 18% and 32%, particularly at lower latitudes (Hijmans, 2003).

Plants, including potato, have a broad range of adaptive strategies for responding to drought stress at morphological, physiological and molecular levels, enabling either drought escape or tolerance of lower water potential (Dahal et al., 2020). For example, drought tolerance is often associated with the accumulation of solutes such as sugar alcohols and proline that can decrease the leaf water potential and facilitate uptake of water (Dahal et al., 2020). At the molecular level, drought induces the expression of many genes involved in tolerance to the stress. One of the most well characterised responses to drought, and other forms of abiotic stress, involves the phytohormone abscisic acid (ABA). Synthesis of ABA induces transcriptional reprogramming leading to a variety of outcomes, including accumulation of osmo-protectants and stomatal closure (Sah et al., 2016). Its endogenous content is predominantly regulated through the oxidative cleavage of β-carotene in plastids. In this pathway, the NCED family of enzymes catalyse the rate-limiting cleavage of violaxanthin and neoxanthin *cis*-isomers to produce xanthoxin, which is exported to the cytosol and converted to ABA in a 2-step enzymatic process (Sah et al., 2016; Ali et al., 2020). While ABA is a crucial signaling molecule in the response to drought, there are many drought-responsive genes that do not respond to exogenous application of ABA, showing that there are also ABA-independent mechanisms (Dahal et al., 2020).

Several studies have taken omics-based approaches to understanding differences in responses to drought in leaf or root tissues between potato varieties with varying degrees of tolerance (Vasquez-Robinet et al., 2008; Evers et al., 2010; Zhang et al., 2014; Sprenger et al., 2016; Moon et al., 2018; Pieczynski et al., 2018; Chen et al., 2020). Of particular interest are the tetraploid Andean potato varieties, *S. tuberosum* subsp. *andigena*, that are well adapted to harsh climatic conditions (Vasquez-Robinet et al., 2008). These varieties may provide an important primary gene pool for improving the stress responses of the more widely grown potato *S. tuberosum* subsp. *tuberosum* (Sukhotu and Hosaka, 2006). Andigena landraces can more effectively maintain photosynthesis levels under prolonged drought stress compared with Tuberosum (Vasquez-Robinet et al., 2008). Moreover, the more stress tolerant Andigena landraces show key differences in leaves relating to resistance, including lower reactive oxygen species (ROS) accumulation, higher mitochondrial activity and more active chloroplast defence responses (Vasquez-Robinet et al., 2008). In this study, RNA-sequencing was performed at early and late stages of exposure to drought stress and after rewatering, in leaf and root tissues of two Andigenum varieties with contrasting drought tolerance phenotypes. The objective was to identify key genes and molecular pathways associated with tolerance to drought to help inform the breeding of new *S. tuberosum* cultivars with improved yield and quality under drought stress.

2 Methods

2.1 Plant material and stress treatment

Two CIP potato varieties of the subspecies *Andigenum* (*Solanum tuberosum* subsp. *andigena*) were employed in this study, namely "Negrita" (CIP accession number: 703671) and "Wila Huaka Lajra" (CIP accession number: 703248), which are tolerant and susceptible to drought, respectively. The two genotypes were chosen based on data from the PapaSalud project presented by Barandalla et al. (2010) at the XXIV congress of the Latin American Potato Association in Cusco, Peru. The PapaSalud project evaluated 77 native potato accessions for their level of resistance against different pests and diseases, their nutritional properties and adaptive potential for different environments to identify appropriate genotypes for sustainable agriculture. Their tolerance to abiotic stresses was

also evaluated and data on drought tolerance was provided by the Neiker Institute (E. Ritter personal communication; Supplementary Table 1). Additional information on the two varieties covering a broad range of phenotypic traits is provided in Supplementary Table 2.

Plants were grown as described by Torres et al., 2013 in an aeroponic system installed in the Estación experimental Santa Ana (INIA - Huancayo) in Huancayo, Peru, where the temperature oscillation was between 6°C and 18°C. The aeroponic system was established based on Otazu (2010) (Supplementary Figure 1A). Briefly, it consisted of several wooden tables/boxes within a net house. The tables were covered with black plastic to avoid transmission of light underneath. The table top and plastic were perforated and the holes filled with Styrofoam leaving a space where *in vitro* plants were inserted so that the roots were in complete darkness under the table. Nutrients were prepared and delivered directly to the root system using a pump connected to a timer, allowing us to control the amount and timing of water and nutrients.

After 3 months of normal irrigation when plants were at the tuber initiation stage, the two potato varieties were exposed to hydric stress to simulate a drought condition by removal of water from the aeroponic system. Photosynthetic rate was measured using the CI-340 Handheld Photosynthesis System (CI-340) at different time points during stress and rewatering (Torres et al., 2013). Based on the patterns of photosynthetic activity described by Vasquez-Robinet et al. (2008), the time taken for the initial photosynthetic rate to decrease by 25% and 60% after removal of water from the system, defined as the early and late responses to drought, was determined by Torres et al., 2013. They also determined the time when plants recovered 80% of their initial photosynthetic rate after re-irrigation (Torres et al., 2013). The different responses were defined in the tolerant variety for the early response at 40 minutes after drought induction (T1), for the late response at 120 minutes after drought induction (T2), and for the recovery phase (T3) at 20 minutes after rewatering which took place at 190 min (Supplementary Figure 1B). The photosynthetic rate of the susceptible variety decreased more rapidly compared with the tolerant variety. While in the tolerant variety, the photosynthetic rate was reduced to 60% at 120 minutes, the same reduction in the susceptible variety occurred at 100 minutes. Moreover, after rewatering, the susceptible variety only recovered only ~50% of its initial photosynthetic rate (Torres et al., 2013). Leaves and roots from the two varieties were collected at the three stress time points (T1, T2, T3) and control samples (T0) were collected from normal irrigation conditions (i.e. before the stress), with samples from 3 individual plants providing biological replication in each condition (Supplementary Figure 1B). In total, there were 48 samples (4 'treatments' (3 stress time points, one control before stress) x 2 varieties x 2 tissues x 3 biological replicates).

2.2 RNA extraction and sequencing

From each leaf or root sample, total RNA was extracted from 1-2g of material using Tri®Reagent (Sigma) followed by treatment with DNAase using the DNA-free™ (Ambion) kit. Sample purity and concentration was determined by the OD260/OD280 and OD260/OD230 ratios using NanoDrop™ and sample integrity was verified with agarose gel electrophoresis. mRNA library construction and sequencing of all 48 samples was carried out at Michigan State University using the Illumina Hi-Seq™ 2000 platform in 1 x 50 nt single end mode.

2.3 Sample QC and read alignment

Read quality was assessed using FastQC v0.11.9. TrimGalore v0.6.5 (Krueger, 2021) was used to trim Illumina adapters and remove reads with a Phred score< 28 or length below 20 nt. Quality-filtered reads were aligned to the potato genome v6.1 (Pham et al., 2020) downloaded from SpudDB (http://spuddb.uga.edu/), using STAR v2.7.2b with default parameters (Dobin et al., 2013). Before mapping, index files were built for the potato reference genome with the option –runMode genomeGenerate using the gff3 genome annotation file. For mapping, the option –quantMode TranscriptomeSAM was used to align the reads to the genome. The number of counts per gene excluding ambiguous reads was obtained using HTSeq v0.11.0 (Anders et al., 2015) with the parameters: –stranded=no, –mode=union and –nonunique=none. Subsequent analysis was carried out using the R statistical software v3.6.3 and RStudio v.1.2.1335.

2.4 Differential expression analysis

Gene count normalization, sample quality assessment and differential expression analysis were performed using DESeq2 v1.26.0 (Love et al., 2014). Genes with raw counts below 10 across all samples were removed prior to downstream analysis. Outlier assessment was carried out after DESeq2 count normalization and variance-stabilizing transformation. Principal component analysis was performed on the top 1,000 most variable genes. To identify significant differentially expressed genes (DEGs) between individual stress time-points compared with the control, or between two varieties at any given time point, we used a threshold of below 0.05 for the default Benjamini-Hochberg adjusted P-values reported by DESeq2 and a threshold of >= 1 for the absolute shrunken log2-fold change (log2FC). This approach provides an overall false discovery rate (FDR) of 5%.

2.5 Functional enrichment analysis

The gene ontology (GO) terms associated with the annotated genes were downloaded from SpudDB (http://spuddb.uga.edu/GO). GO enrichment analysis of the DEGs was performed with the gost() function of the gprofiler2 R package v0.2.0, which uses the hypergeometric test to determine the significance of functional terms (Raudvere et al., 2019). Enriched GO terms were defined where the Bonferroni-corrected P-value from the hypergeometric test was below 0.05.

2.6 Identification of drought and rewatering responsive genes in the tolerant variety

We selected genes that were either up- or down- regulated in response to drought stress and then partially or fully recovered their expression after rewatering ("drought-responsive" genes) in three steps. First, we selected significantly up- or down- regulated genes in the early response (T1) compared with the control condition, with Benjamini-Hochberg adjusted P-value below 0.05 and no threshold for log 2-fold change. Second, we selected up (or down) regulated genes whose expression level was maintained or further increased (or decreased) at T2. Third, we selected genes whose expression changed in an opposite direction from T2 to T3 (i.e. recovered partially or completely) with a log 2-fold change difference of more than 0.5.

3 Results

To characterize drought-induced transcriptional changes in *S. tuberosum* subsp. *andigenum* we performed gene expression profiling using RNA-seq of leaf and root tissues of two varieties with contrasting drought tolerance phenotypes. RNA-seq reads from all 48 samples mapped well to the potato reference genome, with 82-90% of reads mapping uniquely (Supplementary Table 3). The responses of the stress tolerant variety Negrita and the susceptible variety Wila Huaka Lajra to water stress were compared at early (T1, 40-minutes) and late (T2, 120-minutes) time points and in a recovery time point following rewatering (T3). A false discovery rate of 5% and an absolute log 2-fold change >= 1 was used to identify DEGs in each variety that were up- or down- regulated in each time compared to the non-stressed control, or genes that were differentially expressed between varieties at a given time point.

3.1 Overview of drought-induced transcriptional changes in leaf and root

Principal component analysis (PCA) showed clear separation between tolerant and susceptible varieties in both leaf and root, showing widespread differences in gene expression between varieties (Figures 1A, B). In both varieties and tissues, samples were distributed according to the duration of stress treatment, with T1 samples most similar to control samples and T2 and T3 samples increasingly distant from control (Figures 1A, B). This suggested there is a progressive response to drought stress and the observed morphological recovery in response to rewatering was not matched by a full recovery of gene expression in either variety. Also evident from the PCA plot is the relatively tight clustering in both tissues of the control and early stress (T1) samples in the susceptible compared with the tolerant variety (Figures 1A, B). This indicated a stronger, faster response to drought stress in the tolerant variety, particularly in leaf.

The progressive response to drought can also be observed by looking at the differentially expressed genes in each time point compared with the non-stressed control (Figures 1, 2; Supplementary Figure 2). Overall, there was more widespread upregulation of gene expression in response to drought stress compared with downregulation (Figures 1C-F, 3A). For both varieties and both tissues, there was a progressive increase in the number of DEGs in each time point (Figures 1C-F) and a corresponding increase in the number of DEGs unique to any particular time point from T1 through to T3 (Figures 2A-D). In each variety, most genes differentially expressed at the early time point continued to be differentially expressed at the later time point(s), suggesting a sustained response to stress in both tissues (Figures 2A-D). Only a very small proportion of genes were differentially expressed across all three time points (Figures 2A-D).

Leaves of the tolerant variety responded more rapidly to drought stress than susceptible leaves, with 389 genes upregulated compared to only 31 in the susceptible, and 102 genes downregulated compared to 7 in the susceptible at the early time point (Figures 1C, D, 3B). The same but less pronounced pattern was also observed in root (Figures 1E, F, 3B). This suggested the speed of response to stress could be a crucial difference between the two varieties. The rapid response in the leaves of the tolerant variety could also be seen in the volcano plots (Figures 1G, H; Supplementary Figure 2), which showed more upregulated genes in the tolerant variety and more genes with a larger log2 fold change, including some genes with fold changes greater than 5.

While the early response to drought stress was largely unique to the tolerant variety, particularly in leaf, the response of the two varieties became more similar across the stress time points (Figures 3B-D). In the late stress response (T2), the tolerant variety continued to upregulate more genes than the susceptible variety in both tissues (Figures 1C, E, 3C). The opposite pattern was observed for downregulated genes in leaves, with the susceptible variety downregulating more genes than the tolerant variety at T2, with this trend continuing into the recovery phase, T3 (Figures 1D, 3C, D). The delayed response

10.3389/fpls.2022.1003907

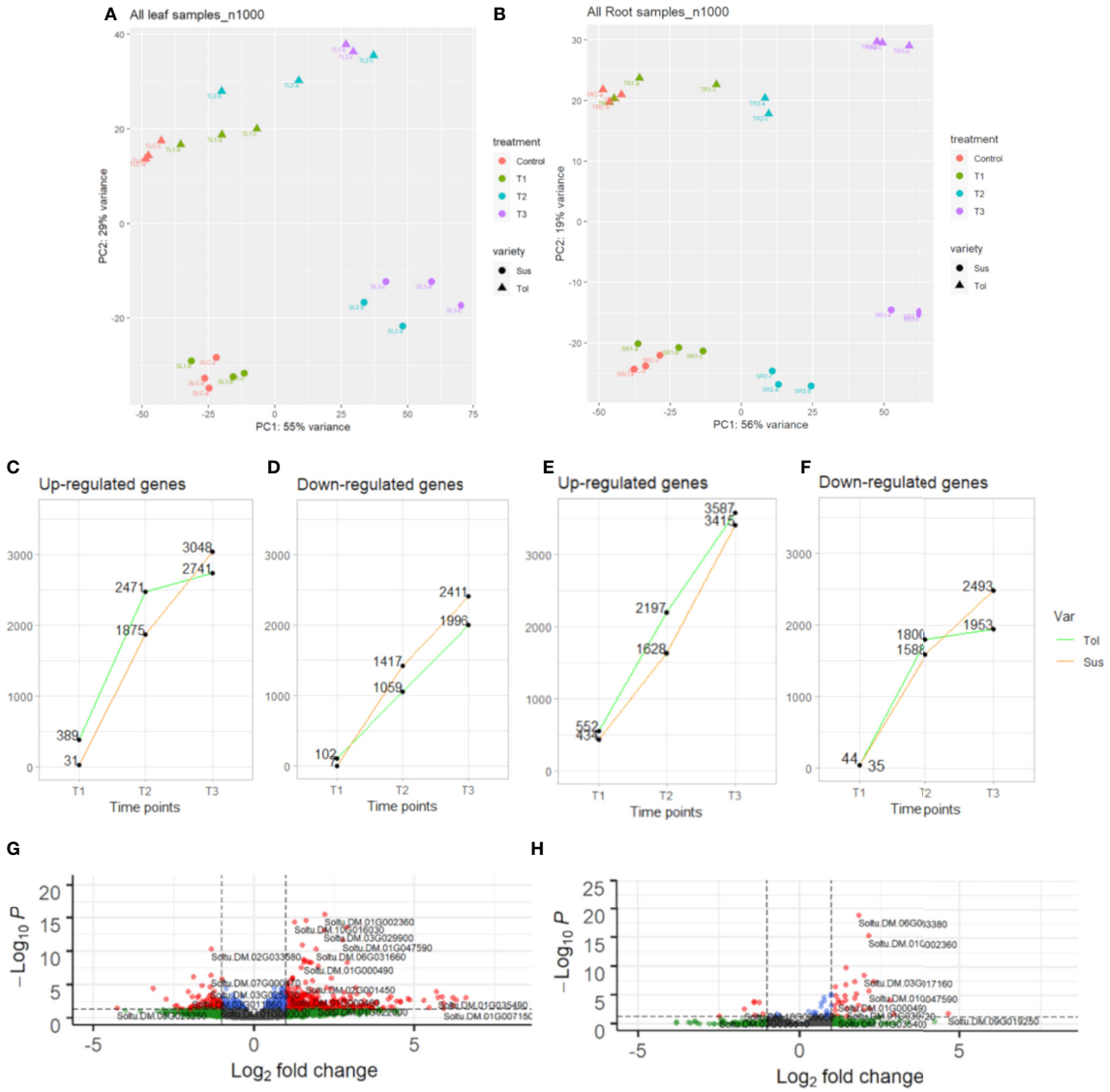


FIGURE 1
Transcriptomic overview of progressive drought response in two potato varieties. Principal component analysis (PCA) of leaf **(A)** and root **(B)** in tolerant (Tol) and susceptible (Sus) varieties in the non-stressed control and across three drought-stressed time points (early T1, late T2 and recovery after rewatering T3). The number of DEGs in each time point compared with the non-stressed control that are up-regulated in leaf **(C)** or root **(E)** or downregulated in leaf **(D)** or root **(F)**. Volcano plot showing differential expression of genes in T1 in leaf **(G)** or root **(H)** compared with the control. Vertical dashed lines indicate absolute log2FC ≥2. Horizontal dashed lines indicate padj. equal to 5%. Genes passing neither threshold are shown in grey, while non-significant genes passing the FC threshold are shown in green. In blue are genes with a small but significant fold change and in red are genes passing both thresholds.

in the leaves and roots of the susceptible variety can also be observed as an increase in the proportion of genes that were uniquely differentially expressed in T3 compared with the tolerant variety (Figure 2). Meanwhile, there was some indication that the tolerant variety could recover its gene expression more effectively in response to rewatering compared with the susceptible variety. This can be seen by a drop-off in the rate of gene upregulation in T3 in leaf tissue of the tolerant variety (Figure 1C), while the susceptible variety had more than double the number of upregulated genes unique to the T3 phase (1,437 genes) compared with the tolerant variety (698 genes) (Figures 2A, B).

Similarly, in roots, the susceptible variety downregulated more genes in the recovery phase, T3, compared with the tolerant variety (Figures 1F, 3D). A higher proportion of these downregulated genes in the susceptible variety (46.8%) were uniquely downregulated in the recovery phase compared with the tolerant (34.6%) (Figures 2C, D). Meanwhile, in the tolerant

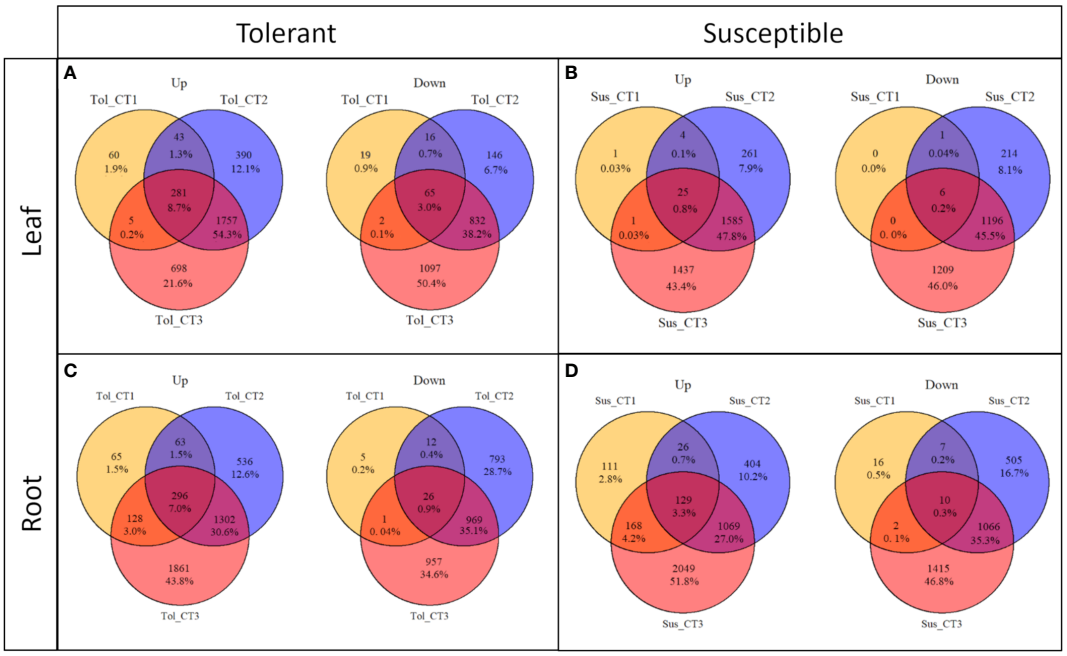


FIGURE 2
Commonalities and differences in DEGs across drought treatment time points in two potato varieties. Up and down-regulated DEGs in leaf **(A, B)** and root **(C, D)** in tolerant (Tol) and susceptible (Sus) varieties in three drought-stressed time points (early T1, late T2 and recovery after rewatering T3) compared with the non-stressed control condition.

variety, there was a drop-off in the rate of gene downregulation in T3 compared with T2 (Figure 1F).

3.2 Common, variety-specific and tissue-specific responses to drought stress

We identified DEGs that were up- or down- regulated in each variety compared with the non-stressed control condition and performed a gene ontology enrichment analysis to identify the functional roles of DEGs responding to drought and rewatering and the variety-specific responses that may be related to the tolerance phenotype (Supplementary Figure 3).

3.2.1 Early responses to drought stress

Consistent with the higher number of DEGs observed in the tolerant leaf or root in the early response to drought, there were more enriched GO terms unique to the tolerant variety, particularly in leaf. Most of the enriched terms exclusive to the tolerant variety in leaves were processes related to osmotic or abiotic stress, including the responses to ABA, regulation of stomatal movement and responses to chitin. Notably, these terms were also enriched for upregulated genes in the roots of both varieties but were not enriched in susceptible leaf (Supplementary Figure 3A). Meanwhile, genes downregulated in tolerant leaves were enriched for seven GO terms relating to DNA replication, negative regulation of transcription factor activity and cell division, likely reflecting a generalized shut down in growth occurring in leaf, not observed in the susceptible variety until T2 (Supplementary Figures 3A, B).

Within the response to ABA (GO:0009737), there were 39 upregulated DEGs in leaves of the tolerant variety compared with only 5 in susceptible leaves (Figure 4). These included 5 genes encoding PP2C proteins, two ABA transporters (Soltu.DM.11G011430-AtABCG25, Soltu.DM.05G023720-AtABCG40), 3 ABI five binding proteins (Soltu.DM.04G000490, Soltu.DM.02G030840, Soltu.DM.05G000860), and 2 AtRD26 proteins (Soltu.DM.12G029330, Soltu.DM.07G024710), which were not upregulated in the susceptible leaf in T1. Two of these PP2Cs were among the 20 most upregulated DEGs in only the tolerant leaf (Supplementary Table 4B). Most of the genes that responded to ABA early in the tolerant variety only began to be upregulated in the susceptible leaves during the late response to drought (T2) (Figure 4). However, there were some commonalities in the early response of leaf in both varieties. In leaf, though not in root, the upregulated DEGs in both varieties were enriched for the process of cell wall modification (GO:0009827, Supplementary Figure 3A). Six genes encoding plant invertase/pectin methylesterase inhibitor superfamily proteins were upregulated to similar levels in both varieties, though only three were significant DEGs in the susceptible. However, by T2, seven genes coding for these proteins were DE in both varieties with a high log2FC (Supplementary Table 5).

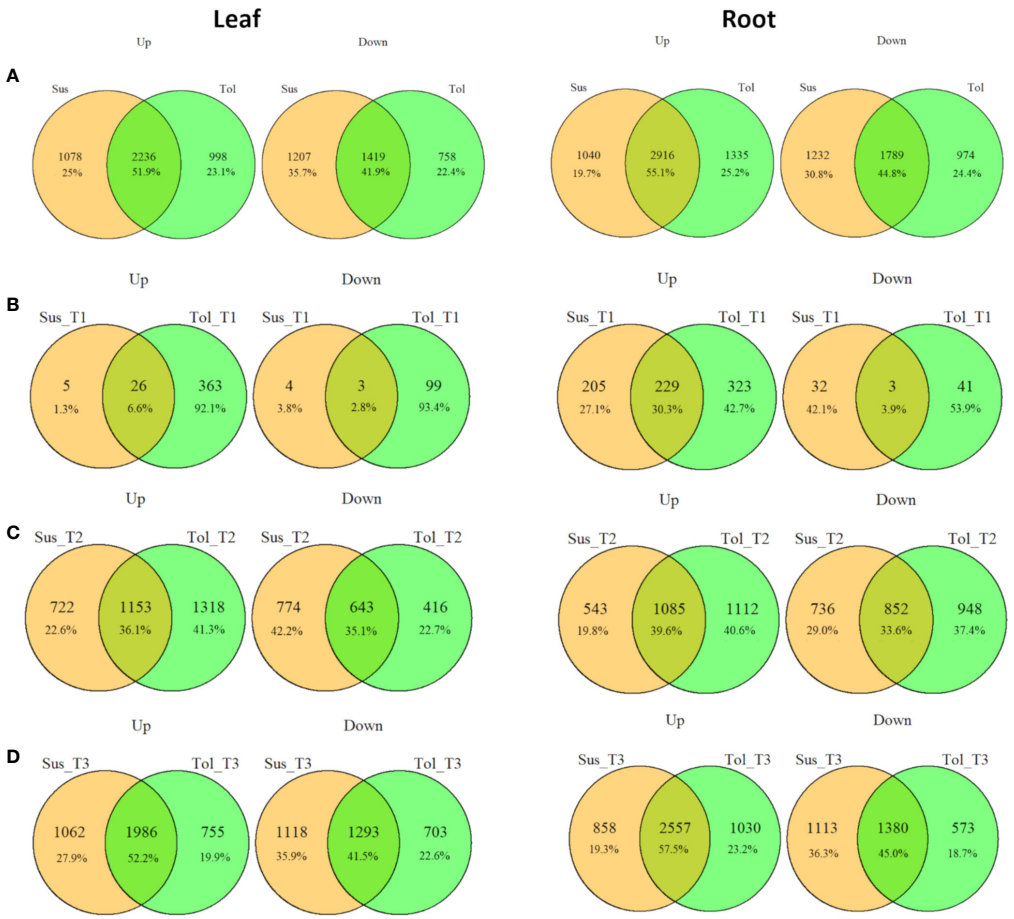


FIGURE 3
Commonalities and differences in drought-responsive DEGs between tolerant and susceptible varieties. Overlap between up- and down-regulated DEGs in leaf and root in tolerant (Tol) and susceptible (Sus) varieties in at least one drought-stressed time point (T1-T3) **(A)**, or in each of the early T1 **(B)**, late T2 **(C)** or recovery after rewatering T3 **(D)** time points compared with the non-stressed control condition.

While there was very little similarity between the two varieties for gene enrichment in leaves, in root there were 21 common enriched GO terms for upregulated genes (Supplementary Figure 3A). Although the response to ABA term was enriched in both varieties (GO:0009737), it also contained many DEGs specific for the tolerant root, which barely overlapped with DEGs specific to the tolerant leaf (Supplementary Table 6). These included 3 ABCG11 transporters (for cutin transport), 2 ABCG40 transporters (for ABA transport), 1 raffinose synthase family protein, and most highly upregulated were 1 galactinol synthase (Soltu.DM.02G006360) and 2 MYB domain proteins (Soltu.DM.05G023310, Soltu.DM.12G001820) (Supplementary Table 6). These 2 last DEGs were among the top 20 most upregulated genes in only the tolerant root (Supplementary Table 4).

Among the 21 GO terms enriched in the roots of both varieties were those related to salicylic acid, lignin and L-phenylalanine catabolic processes, the cellular response to hypoxia and response to oxidative stress (Supplementary Figure 3A). In the lignin catabolic process (GO:0046274) and L-phenylalanine catabolic process (GO:0006559), both varieties upregulated six DEGs encoding phenylalanine ammonia lyases (PALs), which are involved in the first step of production of lignin by converting phenylalanine into cinnamic acid (Figure 5). However, there were also unique components of the tolerant response. Among the top 20 most upregulated genes in only the tolerant variety was a laccase gene (Soltu.DM.04G028320) and a peroxidase superfamily protein homologous to Per52 in *A. thaliana* (Soltu.DM.06G032730) (Supplementary Table 4), both involved in the polymerization of lignin monomers. The laccase Soltu.DM.04G028320 was also one of the genes with the largest difference in expression between tolerant and susceptible varieties (Supplementary Table 4).

While genes involved in the response to oxidative stress were upregulated in the roots of both varieties, including the 6

10.3389/fpls.2022.1003907

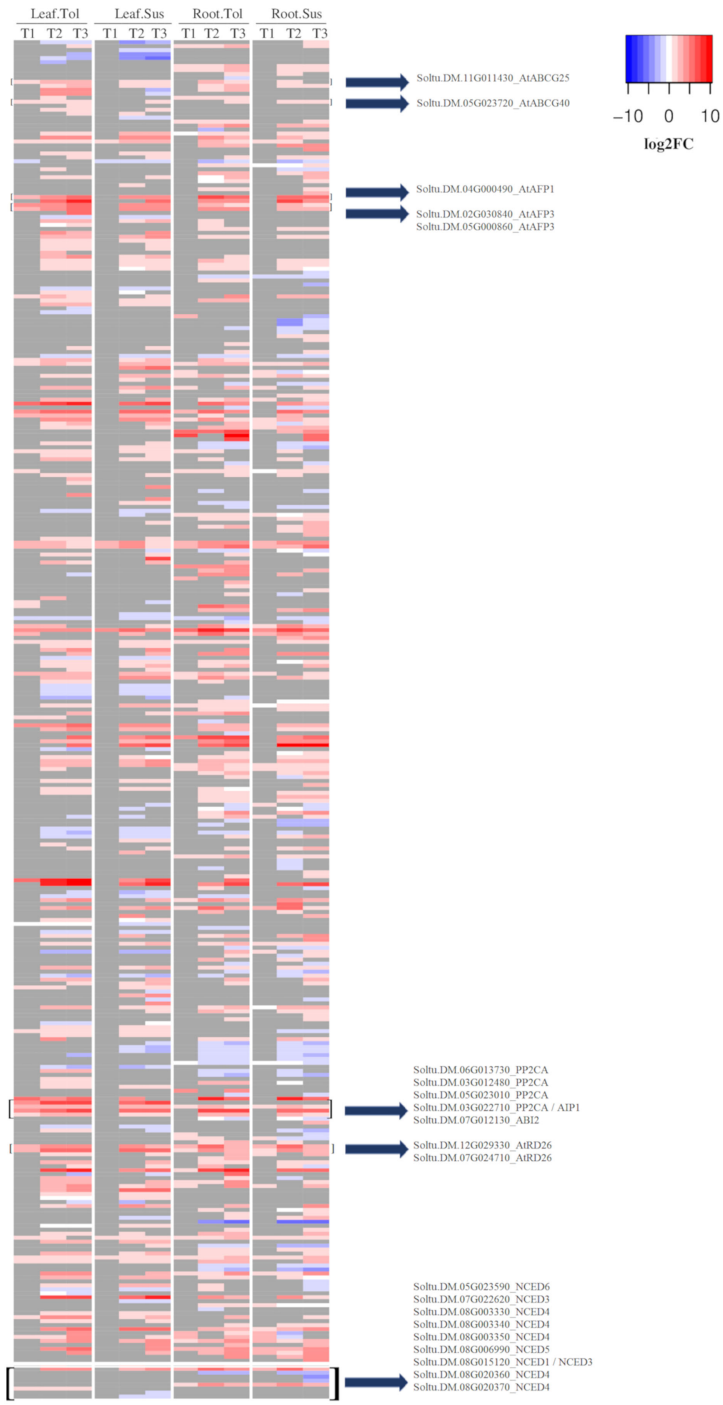


FIGURE 4
Heatmap of the log2 fold change for DEGs related to the ABA response. Shown is the log2FC compared to the non-stressed control for all DEGs annotated with GO:0009737 and all *NCED* genes. In grey are genes with no significant change in the respective time point (early T1, late T2, and recovery T3), tissue, or variety.

10.3389/fpls.2022.1003907

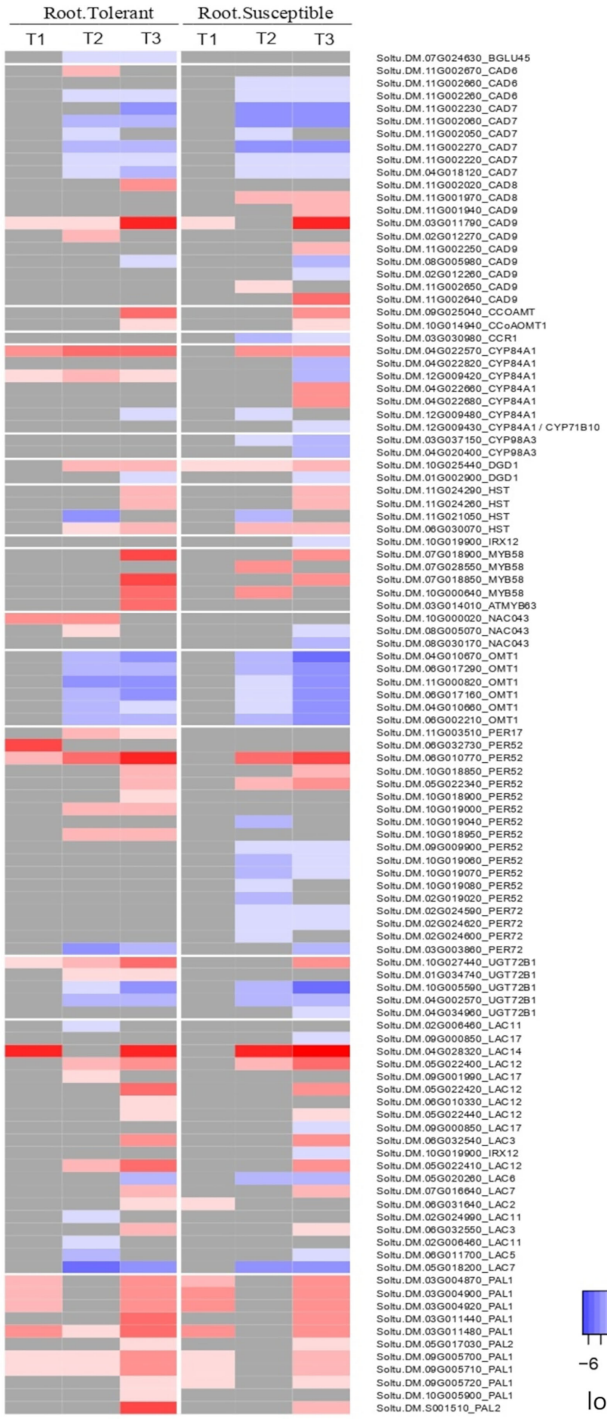


FIGURE 5
Heatmap of the log2 fold change of DEGs related to lignin biosynthesis in roots. Shown is the log2FC compared to the non-stressed control for all DEGs annotated with lignin biosynthetic process (GO:0009809), lignin catabolic process (GO:0046274) or as laccase genes in the potato genome v6.1. In grey are genes with no significant change in the respective time point (early T1, late T2, and recovery T3), tissue, or variety.

PAL genes, there were more upregulated genes in the tolerant than in the susceptible variety, including 16 DEGs upregulated only in the tolerant. These included the upregulation of one galactinol synthase (Soltu.DM.02G006360), three peroxidase superfamily genes in addition to *PER52*, and 3 serine-type endopeptidase inhibitors (Soltu.DM.03G003070, Soltu.DM.06G018620, and Soltu.DM.06G018610), 2 of which were among the 20 most upregulated genes in only the tolerant root (Supplementary Table 7). There were also 10 enriched GO terms for upregulated genes observed only in the tolerant but not susceptible roots. These included phenylpropanoid and camalexin biosynthetic processes, and with a higher number of DEGs were GO terms related to the response to salt stress and the defense response to fungus (Supplementary Figure 3A).

Since the response to ABA was one of the main differences observed between the two varieties, particularly in leaf but also in root, the expression of the potato 9-*cis*-epoxycarotenoid dioxygenase (*NCED*) genes were observed. In roots, Soltu.DM.07G022620 *(AtNCED3*) and Soltu.DM.08G006990 *(AtNCED5*) were upregulated in both varieties across all three time points. At*NCED3* was also upregulated in leaves of both varieties at the later stress time points. Soltu.DM.08G015120, homologous to At*NCED3* and At*NCED1*, was not upregulated in roots or in the susceptible variety at any time point but was upregulated at all three time points in tolerant leaves (Figure 4; Supplementary Table 8).

Returning to the top 20 most highly upregulated genes in the tolerant variety (but not differentially expressed in the susceptible), we observed cytochrome P450 superfamily proteins in both leaf (Soltu. DM.08G010200, also expressed at a significantly higher level than in the susceptible variety) and root (Soltu.DM.04G027470, Soltu.DM.01G005280) (Supplementary Tables 4A, B). At least 8 of the top 20 upregulated genes in the tolerant leaf were involved in processes related directly or indirectly to the ABA response, including transport of organic compounds (sugars, amino acids, lipids), and chaperones controlling protein folding. In the tolerant root, highly upregulated genes included an expansin-like B1 involved in cell wall modification, three genes involved in terpene biosynthesis, and two genes encoding UDP-glycosyl-transferases. The top 20 downregulated genes included three genes encoding NAD(P)-binding Rossmann-fold superfamily proteins, two in leaf (Soltu.DM.10G029820, Soltu.DM.10G025190) and one in root (Soltu.DM.09G001780) (Supplementary Tables 4A, B). In leaf, two genes encoding bifunctional inhibitor/lipid-transfer protein/seed storage 2S albumin superfamily proteins were also strongly downregulated in the early response and throughout the duration of stress treatments.

3.2.2 Late responses to drought stress

In the late response to drought, the responses became more similar between the two varieties and tissues. This was reflected in seven shared enriched GO terms for upregulated genes, including the responses to ABA (GO:0009737), hypoxia (GO:0071456), salt stress (GO:0009651), heat (GO:0009408) and water deprivation (GO:0009414), with a large number of DEGs in each term (Supplementary Figure 3B). In leaf, most of these GO terms (other than response to hypoxia and salt stress) were already enriched in the tolerant variety at the early time point. A similar pattern was also observed for downregulated genes related to growth, with genes involved in DNA replication (GO:1902975) and cell population proliferation (GO:0008283) reducing their expression in leaf at T1 in the tolerant variety, but not until T2 in the susceptible variety. Similarly, for upregulated genes in roots, there were GO terms enriched in the tolerant variety at the early stage of stress that were not enriched in the susceptible response until T2, including the response to salt stress (GO:0009651) and the ABA activated signaling pathway (GO:0009738) (Supplementary Figure 3B).

A total of 19 GO terms were enriched in tolerant but not susceptible leaves at the late time point, of which 4 were also enriched at the early time point (regulation of stomatal movement, regulation of transcription, response to cold and response to chitin). The remaining 15 enriched terms included regulation of the jasmonic acid signaling pathway, salicylic acid biosynthetic process, and more terms related to the defense response against biotic stresses, including wounding (GO:0009611), fungi (GO:0050832), bacteria (GO:0042742) and oomycetes (GO:0002239) (Supplementary Figure 3B). Similarly, tolerant roots were enriched for responses to fungus (GO:0050832) and the response to wounding (GO:0009611), though susceptible roots were not. Meanwhile, unique to the late response of the susceptible leaves were upregulated DEGs enriched for processes involved in protein refolding.

In both leaves and roots of the tolerant variety, the defense response to fungus process (GO:0050832) was enriched among upregulated genes but was not enriched until the recovery phase in the susceptible variety. These genes included 15 WRKY genes; in leaf, 3 were significantly upregulated in both varieties and 8 uniquely in the tolerant, while in root, 4 were significantly upregulated in both varieties and 8 uniquely in the tolerant, including Soltu.DM.12G007400 *(AtWRKY51)*, which was also one of the top 20 most upregulated genes in the late response of root (Supplementary Table 9). Another WRKY gene, Soltu.DM.08G028850 (*WRKY53*), was also in the top 20 DEGs uniquely upregulated in tolerant leaves. Considering all annotated WRKY genes in the potato genome v6.1, the tolerant variety clearly upregulated more WRKY genes in the late response, especially in leaf, four of which also showed an early upregulation in T1 only in the tolerant variety (Figure 6). Many of these genes were not upregulated in the susceptible variety until the recovery phase, T3. Within the response to fungus process (GO:0050832), there were also 4 genes whose products interact with calcium, including two calmodulin genes (Soltu.DM.10G026220, Soltu.DM.10G026210) upregulated in all but the susceptible roots during stress (Supplementary Table 9).

In the late response of roots, the susceptible variety downregulated more genes than the tolerant variety (Figures 1D; 3C), which was reflected in a large number of GO terms uniquely enriched in the susceptible variety (Supplementary Figure 3B). These included three terms relating to biosynthesis of lignin (GO:0009809), melatonin (GO:0030187) and aromatic compounds (GO:0019438). Similarly, downregulated genes in the leaves of the susceptible but not the tolerant variety were enriched for lignan (GO:0009807) and cutin (GO:0010143) biosynthetic processes.

Both varieties showed evidence for cell wall remodeling during the late response. In tolerant and susceptible roots, downregulated genes were enriched for the xyloglucan metabolic process (GO:0010411) and cell wall biogenesis (GO:0042546), both involving a large number of xyloglucan endotransglucosylase/hydrolase genes (Supplementary Table 10). Only the tolerant variety was enriched for genes involved in cell wall modification (GO:0042545), involving downregulation of 10 genes encoding pectin methylesterase inhibitors, only 4 of which were downregulated in the susceptible. Moreover, among the top 20 genes downregulated only in the tolerant roots were four cell wall related genes, including one xyloglucan endotransglucosylase/hydrolase (Soltu.DM.12G025120), one plant invertase/pectin methylesterase inhibitor superfamily (Soltu.DM.02G001870) and one pectin lyase-like superfamily gene (Soltu.DM.04G014020) (Supplementary Table 4A).

3.3 Drought and rewatering responsive genes in the tolerant variety

To identify drought-responsive genes in the tolerant variety, we identified genes whose expression significantly increased in the early response, was maintained or continued to be increased in the late response and returned partially or fully towards non-stressed control levels in response to rewatering. Genes that decreased and then recovered their expression were identified similarly. We identified 45 and 176 such genes in leaf and root, respectively. From the 45 genes in leaf, 39 upregulated and 6 genes downregulated their expression in the early response to drought, while in root, these numbers were 131 and 45 genes, respectively (Supplementary Table 11).

In leaf, the drought-responsive genes were involved in various processes including cell wall modification, responses against pathogens, starch breakdown, transport and calcium binding, among others (Supplementary Table 11A). From the 45 genes, 25 showed a stronger response to drought than the susceptible variety (a difference of more than 0.80 log2FC at T1), including 21 upregulated genes and 4 downregulated genes. Notably, Soltu.DM.06G031870, a HSP20-like chaperones superfamily protein and Soltu.DM.08G010200, a cytochrome P450 family 71 subfamily B polypeptide, were particularly highly upregulated at T1, with log2FC values of > 5 and > 4 respectively. Both were in the top 20 upregulated genes in the tolerant variety and both were significantly more upregulated in T1 compared with susceptible variety, which barely changed its expression compared with the control. The susceptible variety only upregulated these genes from T2 onwards, and even then, Soltu.DM.08G010200 was expressed at a significantly lower level than in the tolerant variety, indicating a weaker, delayed response to drought stress.

Among the 25 genes with a quick response to drought and rewatering in the tolerant variety were 12 genes that were not differentially expressed in the susceptible variety in either the early or late response to drought. These included genes related to the response to pathogens (Soltu.DM.12G00530), calcium-binding (Soltu.DM.01G032110) and transcription regulation (Soltu.DM.09G019660). Two of the upregulated genes from control to T1 (Soltu.DM.01G028100, a beta glucosidase and Soltu.DM.01G032110, an EF hand calcium-binding protein family gene) had a higher expression in the tolerant than in the susceptible leaves during T1, and one downregulated gene (Soltu.DM.08G002160, an FAD-dependent oxidoreductase family protein) had a lower expression in the tolerant than the susceptible at T1.

In root, the drought-responsive genes in the tolerant variety included transcription factors, genes that respond to pathogens and genes involved in transport and signaling cascades. Among 176 genes, 42 had more extreme changes compared to the susceptible variety (a difference of more than 0.80 log2FC compared with the susceptible at T1), including 38 upregulated genes and 4 downregulated genes (Supplementary Table 11B). Comparing the expression of these 42 genes between the two varieties at each time point, 7 upregulated and 2 downregulated genes had a significant difference in expression between tolerant and susceptible varieties in T1 (Supplementary Table 11B). Among the 7 genes were 3 basic chitinases (Soltu.DM.07G005400, Soltu.DM.07G005390, Soltu.DM.02G022960), which were not differentially expressed in the susceptible at any time, and one beta-1,3-glucanase (Soltu.DM.02G033060), which was not differentially expressed in the susceptible variety until the recovery phase.

Among the most highly upregulated genes from control to T1 in the tolerant variety were Soltu.DM.05G002810, encoding an alpha/beta-hydrolases superfamily protein and Soltu.DM.09G019250, encoding an EID1-like protein, whose expression also significantly changed in the susceptible across the 3-time points, but with a lower log2FC. Of the 42 genes shown in Supplementary Table 11B, 18 genes were not differentially expressed in the susceptible root from control to T1 or T2, and 15 were not differentially expressed at any time point. These 15 genes that were responding to drought stress only in the tolerant roots included, among the upregulated genes, 2 NAC-domain containing proteins (Soltu.DM.07G014750 and Soltu.DM.10G000020), the three basic chitinases (Soltu.DM.07G005400, Soltu.DM.07G00539, and Soltu.DM.02G022960), and one nitrate transporter

10.3389/fpls.2022.1003907

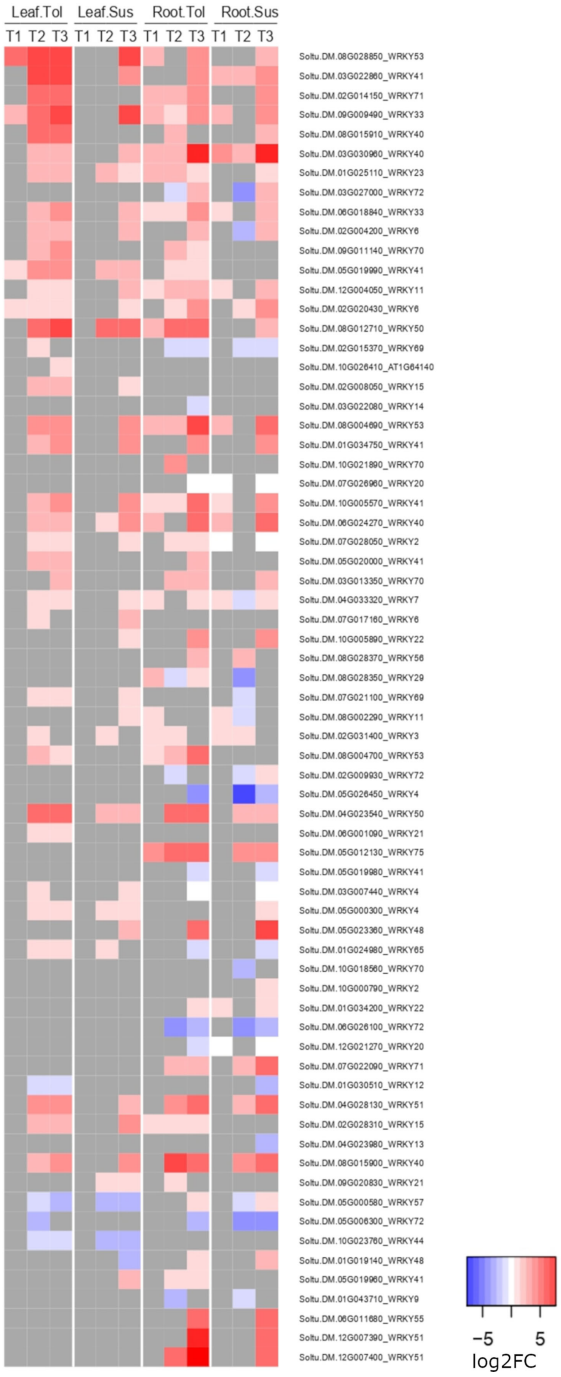


FIGURE 6
Heatmap of the log2 fold change of WRKY genes. Shown is the log2FC expression of WRKY genes differentially expressed in either tissue in two potato varieties at any time point compared to the non-stressed control.

(Soltu.DM.06G030890); the downregulated genes included a flavanone 3-hydroxylase (Soltu.DM.02G023850).

4 Discussion

The increased frequency and severity of abiotic stress conditions caused by climate change, particularly drought, creates a need to identify key genes and molecular pathways that enable potato plants to adapt to or tolerate the stress (Hijmans, 2003; George et al., 2018). To address this need, the transcriptomic response in leaf and root of two Andean potato varieties with contrasting drought tolerance phenotypes was analysed and key genes and pathways associated with tolerance were identified in the early and late responses to drought.

4.1 Faster response of the tolerant variety to drought stress

Both tissues showed a faster change in gene transcription in response to drought in the tolerant compared with the susceptible variety. This was observed in the higher number of DEGs in leaves and in higher fold changes compared with the non-stressed control condition in roots of the tolerant compared with the susceptible variety in the early response. This suggested rapid recognition of the lack of water by the tolerant root and transmission of signals to enable a fast response in leaf. The delayed response of the susceptible leaves was confirmed by observing the different biological processes related to the response to drought stress that began to be enriched in the late response, many of which were already enriched in the tolerant variety since the early time point. One such response was the widespread downregulation of genes involved in DNA replication and cell division, which suggested a generalised shut down/arrest of cell growth. This behaviour is an important mechanism enabling plants to conserve energy under stress, and this reduction under water deficit has been observed to occur independently from changes in photosynthesis (Granier and Tardieu, 1999; Skirycz and Inzé, 2010).

The recovery phase after rewatering produced only a partial recovery of genome-wide gene expression in both varieties. Nevertheless, there were many 'drought-responsive' genes in the tolerant variety that increased (or decreased) their expression in response to the stress and whose expression returned towards non-stressed control levels in response to rewatering, and the response of these genes was much reduced in the susceptible variety. In the tolerant root, these genes included several NAC domain-containing proteins whose expression increased under drought but decreased after rewatering. NACs are a large transcription factor family that are involved in diverse biological processes in plants. These include plant development, cell division, senescence, cell wall formation, plant immunity and responses to abiotic stress (Singh et al., 2021). Several NAC proteins respond to hydric stress and regulate genes in the ABA-dependent pathway (Shen et al., 2017; Jiang et al., 2019). In rice, the expression of *OsNAC2* was associated with the increase in ABA by activating the expression of the *NCED3* gene (Jiang et al., 2019). A NAC gene in potato, *StNAC053*, became highly expressed under ABA and drought treatment. When overexpressed in *Arabidopsis* transgenic lines, this gene enabled plants to better tolerate drought compared with the wild type (Wang et al., 2021).

Nitrate transporters also responded quickly to the availability of water in tolerant compared with susceptible roots. Nitrate excretion transporter1 (*NAXT1*), mainly expressed in the cortex of mature roots (Segonzac et al., 2007), is responsible for nitrate (NO_3^-) efflux from the root into the external medium and is stimulated by cytoplasmic acidic pH (Aslam et al., 1995; Segonzac et al., 2007). Under drought, it was observed that some nitrate transporters were involved in ABA transport and stomatal closure (Kanno et al., 2012). In both the tolerant leaf and root, several genes encoding 2-oxoglutarate (2OG) and Fe (II)-dependent oxygenase superfamily proteins were upregulated rapidly under drought and downregulated with rewatering but did not respond strongly in the susceptible variety. This superfamily protein is involved in diverse processes in plants and subsequently affects responses to biotic or abiotic stresses; these include DNA repair, histone demethylation, biosynthesis or catabolism of enzymes, such as gibberellin, ethylene, auxin, and salicylic acid, and metabolism of secondary metabolites like flavonoids, and coumarin (Farrow and Facchini, 2014).

4.2 Rapid drought-induction of ABA related genes in tolerant leaves and roots

One of the most obvious differences between the tolerant and susceptible varieties was the more widespread and rapid upregulation of ABA-related genes in the early time point, in both tolerant leaves and roots compared with the susceptible variety. Tolerant leaves upregulated a large number of DEGs (33 genes) that were not upregulated in the susceptible variety in the early the response, while tolerant roots upregulated a different set of 19 genes that were also not upregulated in the early response of the susceptible variety. In tolerant leaves, the upregulated genes included 5 PP2C genes, 2 ABA transporters including the ATP-binding cassette family G25 (ABCG25), and 3 ABI five binding proteins (AFPs). Two of these PP2Cs were among the 20 most upregulated DEGs in the tolerant variety. Similarly in roots, the upregulated genes included 5 ABA related transporters. Most of these genes that responded to ABA early in the tolerant variety only began to be upregulated in the

susceptible leaves or roots during the late response to drought, showing the delayed response of the susceptible variety, which has also been observed in *Brassica rapa* varieties in response to drought (Guo et al., 2014).

ABA is a phytohormone that regulates different physiological processes under drought, inducing transcriptional reprogramming leading to a variety of outcomes, including stomatal closure and synthesis of osmoprotectants (Sah et al., 2016). ABCG25 is located in the plasma membrane of vascular tissues and functions as an ABA exporter to allow mobilization of this phytohormone towards the guard cells (Ma et al., 2018). AFP proteins, including PP2C proteins, are negative regulators of the ABA response (Lynch et al., 2017). While upregulation of ABCG25 could indicate more effect of the ABA hormone in tolerant leaves, the upregulation of AFP and PP2C genes would suggest that limiting its effects is also important. In the absence of ABA, PP2C interacts with and inactivates SnRK2, but in the presence of ABA, PP2C is inactivated by interaction with PYR/PYL/RCAR receptors, allowing the release of SnRK2. Free SnRK2 phosphorylates and activates itself and other downstream factors, including SLAC1 and SLAH3 transporters involved in ion import into guard cells, and transcription factors, such as AREB/ABF, to mediate stomatal closure and decrease water transpiration (Ali et al., 2020). Therefore, a higher level of PP2C proteins could maintain more inactive SnRK2 to limit or reduce the downstream effects of the ABA signal in the tolerant variety.

It has been reported that higher levels of ABA can be detrimental to plants in various ways, including accelerating senescence and increasing disease susceptibility (Gietler et al., 2020). In addition, plants under drought stress still need an adequate amount of CO_2 to maintain photosynthesis (Jung et al., 2020), which would also be important for crop yield. Therefore, although the ABA mediated response may be important in response to drought, so too is its effective regulation to make sure that levels are properly modulated so as not to confer a threat. This modulation can occur by regulating the production of ABA or by regulating the response to ABA through the action of PP2C proteins (Jung et al., 2020). Interestingly, the expression of PP2C can in turn be induced by ABA, specifically by the action of the transcription factors AREB/ABF that are activated in the ABA signaling pathway. This may be considered as an important form of negative feedback regulation within the ABA response pathway (Jung et al., 2020). Previous work in potato has also found an increase in PP2C gene expression under drought stress in the leaves of tolerant potato plants (Chen et al., 2020) and in the stolon tissue (Gong et al., 2015). Interestingly, in the potato plants evaluated here, the rate of photosynthesis decreased more rapidly in the susceptible than in the tolerant variety. The tighter regulation of ABA levels through upregulation of PP2C proteins likely allowed the tolerant variety to better maintain its rate of photosynthesis under prolonged stress compared with the susceptible variety.

On the other hand, the early enrichment of genes that respond to ABA in the tolerant variety indicated that there might be more production of this phytohormone in this variety, though ABA levels were not measured in this study. This could be related to the expression of NCED genes. In leaf, an early upregulation of the Soltu.DM.08G015120 gene homologous to *NCED1/NCED3* in *Arabidopsis* was observed in only the tolerant variety, and maintained across all three time points, while the susceptible variety did not upregulate this gene. Since *NCED3* catalyses the rate-limiting step in the ABA biosynthetic pathway, this difference might lead to earlier accumulation of ABA in the tolerant compared with the susceptible leaves. The role of *NCED3* in drought tolerance was previously reported in *Arabidopsis*, where its antisense suppression produced a drought-sensitive phenotype (Iuchi et al., 2001).

Interestingly, the roots of both varieties upregulated *NCED3* and *NCED5* genes throughout drought stress. Previous work also reported that *NCED5* contributes together with *NCED3* to the synthesis of ABA in response to water deficit (Frey et al., 2012). Therefore, in our varieties, functional enrichment of the processes related to ABA was associated with NCED gene expression, which would enable more ABA to be produced at the earlier time point in the roots of both varieties, and potentially more ABA to be produced in tolerant compared with susceptible leaves. Moreover, in tolerant but not susceptible leaves, a beta-glucosidase gene, Soltu.DM.01G028100, was rapidly upregulated in response to drought and recovered its expression after rewatering, but did not significantly respond to stress in the susceptible variety. BGlu enzymes have been associated with the activation of ABA by hydrolyzing ABA glucose ester (ABA-GE) to convert it to free active ABA. In *Oryza sativa*, Os4Glu9, Os4Glu10, Os4BGlu11, Os4BGlu12, and Os4BGlu13 belong to a group of beta glucosidase enzymes that act on ABA-GE (Kongdin et al., 2021). The coding sequence of Soltu.DM.01G028100 was similar to *GLU12*, suggesting that the tolerant variety may not only respond to drought by synthesising 'new' ABA, but also by activating pre-existing ABA in the leaf. Overall, these results suggest an earlier, stronger ABA-mediated response in the tolerant compared with the susceptible variety, coupled with more effective feedback regulation.

4.3 Conserved lignification of root tissue is enhanced in the tolerant variety

During the early response to drought, there was a conserved upregulation of several genes involved in lignification in both tolerant and susceptible roots, including six phenylalanine ammonia lyases (PALs), which catalyse the first step in lignin biosynthesis. This is consistent with the widely reported role of lignin in enhancing tolerance to drought and other abiotic stresses in many plant species (Moura et al., 2010; Chun et al., 2019; Karlova et al., 2021). Lignin provides rigidity to the cell wall and forms a hydrophobic barrier around the xylem to reduce water loss through leakage, thus facilitating more effective

water transport through the plant (Karlova et al., 2021). However, only the tolerant variety strongly upregulated other genes involved in lignin polymerization. These included a laccase homologous to *A. thaliana LAC14* among the top 20 most strongly upregulated genes, and four peroxidase superfamily proteins homologous to *PER52* in *A. thaliana*, one of which was in the top 20 and all of which are involved in the polymerisation of monolignols to produce the final lignin polymer (Chun et al., 2019). These results suggest that the tolerant variety responds to drought stress by more strongly inducing the expression of lignin biosynthetic genes in order to reinforce the plant cell wall and reduce water loss. Similarly, Sprenger et al. (2016) found a constitutively higher expression of genes for lignin biosynthesis in drought tolerant compared with susceptible plants, while Moon et al. (2018) reported upregulation of Laccase 14 in tolerant potato varieties at 6, 12, 24 and 48 hours after drought stress.

4.4 ROS-related damage limitation in the tolerant variety

It is well known that ROS accumulation is one of the first responses to stress in plants. Although at high concentrations it can produce severe damage to cellular structures including proteins, lipids, and nucleic acids, at lower concentrations it functions as a stress signal that allows plants to respond to adverse conditions (Halliwell, 2006; Petrov et al., 2015). Although the roots of both varieties responded early to oxidative stress, roots of the tolerant variety upregulated double the number of genes relating to this process in the early response to drought. In particular, only the tolerant variety highly upregulated two genes homologous to *MYB78* in *Arabidopsis* (Soltu.DM.05G023310 and Soltu.DM.12G001820), which may lead to a rapid accumulation of ROS and downstream signaling, as observed in *Brassica napus* where expression of *BnaMYB78* led to ROS accumulation (Chen et al., 2016). On the other hand, in roots of the tolerant variety we also observed more widespread upregulation of genes that may protect against ROS-induced damage, which is consistent with other studies that have shown greater accumulation of sugars, proline, and molecular chaperones in Andean potato varieties under drought stress (Vasquez-Robinet et al., 2008).

Upregulated genes in the tolerant variety included one galactinol synthase and three serine-type endopeptidase inhibitors. These three genes showed a log2FC of more than four in the early response of the tolerant variety, but a negligible change in the susceptible variety. Serine-type endopeptidase inhibitors (SPIs) are enzymes that regulate the action of proteases to avoid excessive protein degradation that could lead to cellular damage (Clemente et al., 2019). The expression of protease inhibitors was highly induced under abiotic stress in *Arabidopsis* and their overexpression conferred resistance against drought, salt, cold, and oxidative stresses (Zhang et al., 2008). In addition, *Arabidopsis* transgenic lines overexpressing an SPI gene had less oxidative damage than the wild type under drought, showing less lipid peroxidation and more antioxidant activities (Malefo et al., 2020). Upregulation of these genes in the tolerant variety could therefore play a key role in avoiding and/or limiting cellular damage. This is consistent with the upregulation of genes involved in protein refolding that was only observed in the susceptible leaf, suggesting this variety had suffered greater protein damage by the late drought stress time point.

Galactinol synthase is involved in the synthesis of raffinose; this enzyme converts UDP-galactose into galactinol, which in turn is converted into raffinose by the raffinose synthase enzyme (Taji et al., 2002). In addition to upregulating a galactinol synthase, only the tolerant variety upregulated a raffinose synthase gene (Soltu.DM.02G033230) in the early drought response. Both genes were also upregulated in the susceptible variety, but not until the later stage of drought stress and to a lesser degree. Moreover, two UDP-glycosyltransferases were among the top 20 most highly upregulated genes in the early response of only the tolerant variety, while related genes in *Arabidopsis* have a known role in cold, salt and drought tolerance (Li et al., 2017). Raffinose family oligosaccharides (RFOs) accumulate during seed development and play an important role in desiccation tolerance of the seed (Taji et al., 2002). In *Arabidopsis* rosettes, over-expression of galactinol synthase results in increased galactinol and raffinose levels and enhanced dehydration tolerance (Taji et al., 2002). The accumulation of galactinol and raffinose in plants can protect against ROS-related damage under stress (ElSayed et al., 2014). The high expression of these antioxidant proteins in only the tolerant variety could be alleviating the oxidative damage produced under drought. Variety-specific accumulation of raffinose and galactinol is supported by other studies that show conserved accumulation in some potato varieties such as Alegria, Milva, Desiree and Saturna (Sprenger et al., 2016), but no accumulation in other Andean varieties, Sullu and SS2613 (Evers et al., 2010).

4.5 Biotic stress related response of the tolerant variety is conserved across tissues

In the late response to drought, many genes involved in the response against pathogens, including fungi, bacteria and oomycetes, changed their expression in both tissues. Crosstalk between the responses to abiotic and biotic stresses in plants involves processes that respond to hormones, such as ABA, salicylic acid, or jasmonic acid, as well as ROS generation as a signal of stress (Fujita et al., 2006). Overexpression of transcription factors, such as MYB, NAC, HSF, and WRKY

are also involved in this crosstalk (Fujita et al., 2006; Bai et al., 2018). Such crosstalk was clearly observed between these two types of stress, in the early response to drought in roots and increasingly so in the late response, particularly in leaves. Among the genes involved in the crosstalk, the tolerant variety upregulated more WRKY genes in both tissues compared with the susceptible variety. In the potato genome, 129 genes were annotated as a putative WRKY, whose expression responded to different types of stress, such as heat, salt, and drought, and to salicylic acid treatment (Zhang et al., 2017).

Specifically, Soltu.DM.08G028850, annotated as *AtWRKY53* in *Arabidopsis*, was one of the most highly upregulated genes in the tolerant leaves, but was not upregulated in the susceptible leaf until the recovery phase. The same was observed for Soltu.DM.12G007400 *(AtWRKY51)* in the late response of root. Members of the WRKY protein family are involved in the regulation of the ABA pathway, and their overexpression promotes drought tolerance in tomato, tobacco, and rice (Bai et al., 2018). It was reported that expression of *AtWRKY53* was modulated under biotic stress, induced by SA but repressed by JA, and was involved in plant senescence (Zentgraf and Doll, 2019). In contrast to the result observed here, the upregulation of this specific *AtWRKY53* under drought was correlated with reduced drought tolerance (Sun and Yu, 2015), where its overexpression decreased the hydrogen peroxide levels and stomatal closure in *Arabidopsis* lines that did not survive after drought and rewatering. In contrast, the lower reduction in the rate of photosynthesis observed in our tolerant variety could be correlated with reduced stomatal closure.

Other upregulated genes in the tolerant variety relating to biotic stress responses included genes encoding basic chitinases and genes relating to calcium signaling. Under both biotic and abiotic stress, the fluctuation of calcium functions as a signal, activating stress-responsive calcium sensors like calmodulins (CaMs) or calcineurin B-like proteins (CBLs), calcium-dependent protein kinases (CPK) and calcium/calmodulin-dependent protein kinases (CCaMKs) (Ku et al., 2018). The tolerant variety showed stronger upregulation of genes involved in calcium signaling in both tissues, including two calmodulin genes upregulated in all but the susceptible roots (Soltu.DM.10G026220, Soltu.DM.10G026210), and a BCL-2-associated athanogene 6 (*BAG6*) upregulated in all but the susceptible leaves. BAG proteins including Bag6 mediate the response to multiple kinds of stress in *Arabidopsis*, including the response to salt stress (Arif et al., 2021).

Three basic chitinases were only upregulated in the tolerant root during the early and late responses to drought stress and largely recovered after rewatering, while the susceptible variety did not upregulate chitinases at all. Chitinases are enzymes that participate in the first line of the plant defence during PAMP-triggered immunity by degrading chitin, a major component of the fungal cell wall. However, chitinases are not only induced under pathogen attack, but also under salt, cold, and drought stress (Takenaka et al., 2009) and play a role in plant growth and development. In potato, a class I chitinase was observed within a group of genes conferring drought tolerance identified by a yeast functional screening approach (Kappachery et al., 2013). In clover leaves, chitinases and β-1,3 glucanases increase their expression under drought during the early stage of stress, and were significantly correlated with an increase in proline, with a possible role in detoxification of accumulated ammonia under drought (Lee et al., 2008). Here also, the tolerant variety upregulated a beta-1,3-glucanase (Soltu.DM.02G033060), which then recovered its expression after rewatering, while the susceptible variety did not upregulate this gene until the recovery phase. β-1,3 glucanases hydrolyse glycosidic bonds in the glucans of the fungal cell wall, to protect against fungal pathogens (Oide et al., 2013). The upregulation of such genes involved in the conserved pathways between biotic and abiotic stress responses may contribute to the improved response to drought stress in the tolerant variety.

4.6 Cell wall remodeling in response to drought stress

Both tolerant and susceptible varieties altered the expression of cell-wall related genes in the early and late responses to drought stress, though the response was stronger and faster in the tolerant variety. In tolerant leaves, more plant invertase/pectin methylesterase inhibitor superfamily proteins (INV/PMEI-SP) were upregulated in the early response compared with the susceptible variety. This family includes pectin methylesterase inhibitors (PMEIs) and invertase inhibitor (INVI) proteins that regulate the PME and INV enzymes, respectively (Coculo and Lionetti, 2022). Since the tolerant variety showed stronger upregulation of PMEIs in leaf, this may translate into more inhibition of the action of PMEs. Under drought, one of the most important mechanisms generated by the plant is the regulation of stomatal aperture/closure. In *Arabidopsis*, the activity of PME upon methylesterified pectin was important for proper regulation of stomatal aperture under heat and drought stress (Wu et al., 2107). *Arabidopsis* mutants not expressing *PME34*, whose activation depended on ABA, showed enhanced stomatal aperture and a lethal phenotype to heat stress (Wu et al., 2017). In pepper, the overexpression of CaPMEI1 increased tolerance to drought in *Arabidopsis* plants (An et al., 2008). Therefore, the regulation of PME by PMEI is an important factor influencing stomatal opening during drought stress.

In contrast to early responses in leaf, in the late response the tolerant roots showed more widespread and stronger downregulation of INV/PMEI-SP genes than the susceptible roots, including one gene in the top 20 most downregulated genes (Soltu.DM.02G001870). PME demethylesterifies oligomers of the pectin backbone, then these blockwise

demethylesterified pectins may bind to each other by crosslinking with calcium ions to form a rigid structure called the "egg-box" in which calcium ions interact with molecules of water to keep the cell wall hydrated (Wormit and Usadel, 2018). Therefore, downregulation of PMEIs may facilitate formation of the egg box structure and maintenance of root cell wall hydration. However, different experiments have shown contradictory results in *Arabidopsis* regarding the relationship between root growth and PMEI activity. While in some cases root growth was promoted by overexpression of PMEI (An et al., 2008), in others it was promoted by inhibition of PMEI (Wormit and Usadel, 2018). In rice, the high expression of PMEI provoked a negative effect on plant growth, producing dwarfed plants (Nguyen et al., 2017). In transgenic potato expressing a Petunia PME, whose activity was pronounced in leaf and tubers, more plant growth at the early stage but no difference in growth after 35 days was observed (Pilling et al., 2000). In the case of invertase inhibitors, their expression was favourable against drought in maize (Chen et al., 2019), in contrast to observations in cucumber where overexpression of vacuolar invertase reduced drought tolerance (Chen et al., 2022). Therefore, further investigation is needed to understand the species and tissue-specific effects of PMEIs, PMEs and invertase inhibitors in the abiotic stress response.

Among the most strongly upregulated genes in root during the early response to drought was an expansin like-B1 (*EXLB1*). Expansins are a class of non-enzymatic cell wall proteins that play a role in the regulation of cell growth by disrupting hydrogen bonds, facilitating cell wall loosening and expansion (Marowa et al., 2016). Strong overexpression of an expansin like-B1 gene was also observed in the stolon of potato variety Ningshu under drought stress (Gong et al., 2015). In *Brassica rapa*, *BrEXLB1* was preferentially expressed in root, and under drought stress its expression was highly elevated, contributing to enhanced root growth and drought tolerance (Muthusamy et al., 2020). In maize, Exp1, Exp5 and ExpB8 levels were increased in roots under low water potential, allowing continued root elongation under stress (Wu et al., 2001). Here, while both varieties upregulated expansin-like B1 during the early response to drought stress, the response was much stronger in the tolerant than in the susceptible variety, with log2 fold changes of > 6 and > 2 respectively. Continued strong upregulation of this gene in the tolerant variety throughout the late and recovery responses to drought stress may have contributed to enhanced root growth compared with the susceptible variety. Consistent with this idea, the root hair specific gene Soltu.DM.01G006590 homologous to *Arabidopsis* AT5G22410 and encoding the cell-wall localised peroxidase PER60/RHS18 was strongly downregulated in the late response of only the tolerant variety (log2FC > 5). Its overexpression in *Arabidopsis* mutants reduced root hair length by 16% compared with wild type, thus its downregulation in the tolerant variety may further facilitate cell wall expansion and growth under drought stress.

A common response in the late response of root between both varieties was the downregulation of xyloglucan endotransglucosylase/hydrolase genes (XTHs). However, though the response was common, more XTHs were downregulated in the tolerant variety and to a greater extent than in the susceptible variety, including one gene in the top 20 most downregulated genes (Soltu.DM.12G025120). XTHs have the capacity to cleave and re-ligate the xyloglucan fragments and their increased expression has been correlated with improved drought tolerance (Le Gall et al., 2015). In maize, the expression and the activity of XTHs in the root differed depending on the evaluated region. The expression of XTH in the apical zone was downregulated, while it was upregulated in the subapical zones (Iurlaro et al., 2016). Therefore, it may be important to differentiate the expression of these enzymes in the different root zones to better understand the responses of the tolerant and susceptible varieties.

4.7 Conclusions

There are commonalities and differences in the transcriptomic response between potato varieties that differ in their tolerance to drought stress, many of which involve genes related to the plant cell wall. Strikingly, leaves and roots of the tolerant variety show more widespread and stronger upregulation of genes relating to the ABA response in the early stages of stress compared with the susceptible variety, indicating the speed of response may be crucial. Similarly, there is a general early shut down in growth in the tolerant variety that is not seen until the late response in the susceptible variety. The tolerant roots upregulate many more genes involved in the response to oxidative stress than susceptible roots, facilitating maintenance of protein integrity and early accumulation of metabolites including galactinol and raffinose that may enhance desiccation tolerance. In addition, the tolerant roots showed stronger upregulation of genes involved in lignin biosynthesis, which likely strengthens the cell walls and maintains water transport/minimizes water loss under drought stress. In the late response to stress, the tolerant variety upregulates many genes involved in the response to various biotic stresses, including WRKY family proteins, chitinases and glucanases that may modulate hormone signaling and facilitate effective detoxification of cells under drought stress.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI, PRJNA874012.

Author contributions

GO conceived and designed the experiments. YT and RL carried out the experiments and RB generated the RNA-seq data. OP and LC led the analysis of the data, with contributions from GO and AP. OP and LC wrote the manuscript with contributions from all authors.

Please see here for full authorship criteria.

Funding

This work was supported by the Catedra-CONCYTEC in Genómica funcional y AgroNegocios (Grant no. 242-2011-CONCYTEC-OAJ) and FINCyT (grant no. 099-FINCyT-EQUIP-2009)/(076-FINCyT-PIN-2008). OP received a fellowship from FONDECYT–CONCYTEC, Lima, Peru (grant no. 126-2017-FONDECYT).

Acknowledgments

The authors acknowledge the FONDECYT – CONCYTEC-Peru for the attribution of a PhD grant to OP (grant no. 126-2017-FONDECYT).

Conflict of interest

RL is employed by Joyn Bio LLC.

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1003907/full#supplementary-material

References

Ali, S., Hayat, K., Iqbal, A., and Xie, L. (2020). Implications of abscisic acid in the drought stress tolerance of plants. *Agronomy* 10 (9), 1–28. doi: 10.3390/agronomy10091323

Anders, S., Pyl, P. T., and Huber, W. (2015). HTSeq – a Python framework to work with high-throughput sequencing data. *Bioinformatics* 32 (2), 166–169. doi: 10.1093/bioinformatics/btu638

An, S. H., Sohn, K. H., Choi, H. W., Hwang, I. S., Lee, S. C., and Hwang, B. K. (2008). Pepper pectin methylesterase inhibitor protein CaPMEI1 is required for antifungal activity, basal disease resistance and abiotic stress tolerance. *Planta* 228 (1), 61–78. doi: 10.1007/s00425-008-0719-z

Arif, M., Li, Z., Luo, Q., Li, L., Shen, Y., and Men, S. (2021). The BAG2 and BAG6 genes are involved in multiple abiotic stress tolerances in arabidopsis thaliana. *Int. J. Mol. Sci.* 22 (11), 5856. doi: 10.3390/ijms22115856

Aslam, M., Travis, R. L., and Huffaker, R. C. (1995). Effect of pH and calcium on short-term NO3- fluxes in roots of barley seedlings. *Plant Physiol.* 108 (2), 727–734. doi: 10.1104/pp.108.2.727

Bai, Y., Sunarti, S., Kissoudis, C., Visser, R. G. F., and van der Linden, C. G. (2018). The role of tomato WRKY genes in plant responses to combined abiotic and biotic stresses. *Front. Plant Sci.* 13 (9). doi: 10.3389/fpls.2018.00801

Barandalla, L., Aragonés, A., López, R., Lucca, F., Ruiz de Galarreta, J., and Ritter, E. (2010). "Evaluacion y analisis molecular de tolerancias a estreses abioticos en papa," in *Memoria del XXIV congreso de la asociación latinoamericana de la papa (ALAP) 2010 y I simposium internacional de recursos genéticos de la papa*, Instituto Nacional de Innovación Agraria - INIA, Peru. 417–418.

Chen, L., Liu, X., Huang, X., Luo, W., Long, Y., Greiner, S., et al. (2019). Functional characterization of a drought-responsive invertase inhibitor from maize (*Zea mays* l.). *Int. J. Mol. Sci.* 20 (17), 4081. doi: 10.3390/ijms20174081

Chen, Y., Li, C., Yi, J., Yang, Y., Lei, C., and Gong, M. (2020). Transcriptome response to drought, rehydration and re-dehydration in potato. *Int. J. Mol. Sci.* 21 (1), 159. doi: 10.3390/ijms21010159

Chen, B., Niu, F., Liu, W. Z., Yang, B., Zhang, J., Ma, J., et al. (2016). Identification, cloning and characterization of R2R3-MYB gene family in canola (*Brassica napus* l.) identify a novel member modulating ROS accumulation and hypersensitive-like cell death. *DNA Res.* 23 (2), 101–114. doi: 10.1093/dnares/dsv040

Chen, L., Zheng, F., Feng, Z., Li, Y., Ma, M., Wang, G., et al. (2022). A vacuolar invertase CsVI2 regulates sucrose metabolism and increases drought tolerance in *Cucumis sativus* l. *Int. J. Mol. Sci.* 23 (1), 176. doi: 10.3390/ijms23010176

Chun, H. J., Baek, D., Cho, H. M., Lee, S. H., Jin, B. J., Yun, D. J., et al. (2019). Lignin biosynthesis genes play critical roles in the adaptation of arabidopsis plants to high-salt stress. *Plant Signal Behav.* 14 (8), 1625697. doi: 10.1080/15592324.2019.1625697

Clemente, M., Corigliano, M. G., Pariani, S. A., Sánchez-López, E. F., Sander, V. A., and Ramos-Duarte, V. A. (2019). Plant serine protease inhibitors: Biotechnology application in agriculture and molecular farming. *Int. J. Mol. Sci.* 20 (6), 1345. doi: 10.3390/ijms20061345

Coculo, D., and Lionetti, V. (2022). The plant Invertase/Pectin methylesterase inhibitor superfamily. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.863892

Dahal, K., Li, X. Q., Tai, H., Creelman, A., and Bizimungu, B. (2020). Improving potato stress tolerance and tuber yield under a climate change scenario – a current overview. *Front. Plant Sci.* 10 (563). doi: 10.3389/fpls.2019.00563

Deblonde, P. M. K., and Ledent, J. F. (2001). Effects of moderate drought conditions on green leaf number, stem height, leaf length and tuber yield of potato cultivars. *Eur. J. Agron.* 14 (1), 31–41. doi: 10.1016/S1161-0301(00)00081-2

Dobin, A., Davis, C. A., Schlesinger, F., Drenkow, J., Zaleski, C., Jha, S., et al. (2013). STAR: ultrafast universal RNA-seq aligner. *Bioinformatics* 29, 15–21. doi: 10.1093/bioinformatics/bts635

Eiasu, B. K., Soundy, P., and Hammes, P. S. (2007). Response of potato (*Solarium tuberosum*) tuber yield components to gel-polymer soil amendments and irrigation regimes. *New Zeal. J. Crop Hortic. Sci.* 35 (1), 25–31. doi: 10.1080/01140670709510164

ElSayed, A. I., Rafudeen, M. S., and Golldack, D. (2014). Physiological aspects of raffinose family oligosaccharides in plants: Protection against abiotic stress. *Plant Biol. (Stuttg).* 16 (1), 1–8. doi: 10.1111/plb.12053

Evers, D., Lefèvre, I., Legay, S., Lamoureux, D., Hausman, J. F., Rosales, R. O., et al. (2010). Identification of drought-responsive compounds in potatothrough a

combined transcriptomic and targeted metabolite approach. *J. Exp. Bot.* 61 (9), 2327–2343. doi: 10.1093/jxb/erq060

FAO (2021) FAO statistical database. FAOSTAT, (2021). Available at: http://www.fao.org/faostat/en/#data/QC (Accessed 4th October 2021).

Farrow, S. C., and Facchini, P. J. (2014). Functional diversity of 2-oxoglutarate/Fe(II)-dependent dioxygenases in plant metabolism. *Front. Plant Sci.* 9 (5). doi: 10.3389/fpls.2014.00524

Frey, A., Effroy, D., Lefebvre, V., Seo, M., Perreau, F., Berger, A., et al. (2012). Epoxycarotenoid cleavage by NCED5 fine-tunes ABA accumulation and affects seed dormancy and drought tolerance with other NCED family members. *Plant J.* 70 (3), 501–512. doi: 10.1111/j.1365-313X.2011.04887.x

Fujita, M., Fujita, Y., Noutoshi, Y., Takahashi, F., Narusaka, Y., Yamaguchi-Shinozaki, K., et al. (2006). Crosstalk between abiotic and biotic stress responses: a current view from the points of convergence in the stress signaling networks. *Curr. Opin. Plant Biol.* 9 (4), 436–442. doi: 10.1016/j.pbi.2006.05.014

George, T. S., Taylor, M. A., Dood, I. C., and White, P. (2018). Climate change and consequences for potato production: a review of tolerance to emerging abiotic stress. *Potato. Res.* 60, 239–268. doi: 10.1007/s11540-018-9366-3

Gietler, M., Fidler, J., Labudda, M., and Nykiel, M. (2020). Abscisic acid–enemy or savior in the response of cereals to abiotic and biotic stresses? *Int. J. Mol. Sci.* 21 (13), 1–29. doi: 10.3390/ijms21134607

Gong, L., Zhang, H., Gan, X., Zhang, L., Chen, Y., Nie, F., et al. (2015). Transcriptome profiling of the potato (*Solanum tuberosum* l.) plant under drought stress and water-stimulus conditions. *PloS One* 10 (5), e0128041. doi: 10.1371/journal.pone.0128041

Grafton, R. Q., Williams, J., and Jiang, Q. (2015). Food and water gaps to 2050: preliminary results from the global food and water system (GFWS) platform. *Food Sec.* 7, 209–220. doi: 10.1007/s12571-015-0439-8

Granier, C., and Tardieu, F. (1999). Water deficit and spatial pattern of leaf development. variability in responses can be simulated using a simple model of leaf development. *Plant Physiol.* 119 (2), 609–619. doi: 10.1104/pp.119.2.609

Guo, Y. M., Samans, B., Chen, S., Kibret, K. B., Hatzig, S., Turner, N. C., et al. (2014). Drought-tolerant *Brassica rapa* shows rapid expression of gene networks for general stress responses and programmed cell death under simulated drought stress. *Plant Mol. Biol. Rep.* 35, 416–430. doi: 10.1007/s11105-017-1032-4

Halliwell, B. (2006). Redox biology is a fundamental theme of aerobic life. *Plant Physiol.* 141 (2), 312–322. doi: 10.1104/pp.106.077073.312

Haverkort, A. J., and Verhagen, A. (2008). Climate change and its repercussions for the potato supply chain. *Potato. Res.* 51, 223–237. doi: 10.1007/s11540-008-9107-0

Hijmans, R. J. (2003). The effect of climate change on global potato production. *Am. J. Potato. Res.* 80, 271–279. doi: 10.1007/BF02855363

Iuchi, S., Kobayashi, M., Taji, T., Naramoto, M., Seki, M., Kato, T., et al. (2001). Regulation of drought tolerance by gene manipulation of 9-cis-epoxycarotenoid dioxygenase, a key enzyme in abscisic acid biosynthesis in arabidopsis. *Plant J.* 27 (4), 325–333. doi: 10.1046/j.1365-313x.2001.01096.x

Iurlaro, A., De Caroli, M., Sabella, E., De Pascali, M., Rampino, P., De Bellis, L., et al. (2016). Drought and heat differentially affect XTH expression and XET activity and action in 3-day-old seedlings of durum wheat cultivars with different stress susceptibility. *Front. Plant Sci.* 10 (7). doi: 10.3389/fpls.2016.01686

Iwama, K. (2008). Physiology of the potato: New insights into root system and repercussions for crop management. *Potato. Res.* 51, 333–353. doi: 10.1007/s11540-008-9120-3

Jiang, D., Zhou, L., Chen, W., Ye, N., Xia, J., and Zhuang, C. (2019). Overexpression of a microRNA-targeted NAC transcription factor improves drought and salt tolerance in rice *via* ABA-mediated pathways. *Rice.* 12 (1), 76. doi: 10.1186/s12284-019-0334-6

Jung, C., Nguyen, N. H., and Cheong, J. J. (2020). Transcriptional regulation of protein phosphatase 2C genes to modulate abscisic acid signaling. *Int. J. Mol. Sci.* 21 (24), 9517. doi: 10.3390/ijms21249517

Kanno, Y., Hanada, A., Chiba, Y., Ichikawa, T., Nakazawa, M., Matsui, M., et al. (2012). Identification of an abscisic acid transporter by functional screening using the receptor complex as a sensor. *Proc. Natl. Acad. Sci. U. S. A.* 109 (24), 9653–9658. doi: 10.1073/pnas.1203567109

Kappachery, S., Yu, J. W., Baniekal-Hiremath, G., and Park, S. W. (2013). Rapid identification of potential drought tolerance genes from *Solanum tuberosum* by using a yeast functional screening method. *C. R. Biol.* 336 (11–12), 530–545. doi: 10.1016/j.crvi.2013.09.006

Karlova, R., Boer, D., Hayes, S., and Testerink, C. (2021). Root plasticity under abiotic stress. *Plant Physiol.* 187 (3), 1057–1070. doi: 10.1093/plphys/kiab392

Kongdin, M., Mahong, B., Lee, S. K., Shim, S. H., Jeon, J. S., and Ketudat Cairns, J. R. (2021). Action of multiple rice β-glucosidases on abscisic acid glucose ester. *Int. J. Mol. Sci.* 22 (14), 7593. doi: 10.3390/ijms22147593

Krueger, F. (2021) *Trim galore!*. Available at: http://www.bioinformatics.babraham.ac.uk/projects/trim_galore/

Kumar, S., Asrey, A., and Mandal, G. (2007). Effect of differential irrigation regimes on potato (*Solanum tuberosum*) yield and post-harvest attributes. *Indian J. Agric. Sci.* 77 (6), 366–368. https://krishi.icar.gov.in/jspui/handle/123456789/25584

Ku, Y. S., Sintaha, M., Cheung, M. Y., and Lam, H. M. (2018). Plant hormone signaling crosstalks between biotic and abiotic stress responses. *Int. J. Mol. Sci.* 19 (10), 3206. doi: 10.3390/ijms19103206

Lee, B. R., Jung, W. J., Lee, B. H., Avice, J. C., Ourry, A., and Kim, T. H. (2008). Kinetics of drought-induced pathogenesis-related proteins and its physiological significance in white clover leaves. *Physiol. Plant* 132 (3), 329–337. doi: 10.1111/j.1399-3054.2007.01014.x

Le Gall, H., Philippe, F., Domon, J. M., Gillet, F., Pelloux, J., and Rayon, C. (2015). Cell wall metabolism in response to abiotic stress. *Plants (Basel).* 4 (1), 112–166. doi: 10.3390/plants4010112

Li, P., Li, Y. J., Zhang, F. J., Zhang, G. Z., Jiang, X. Y., Yu, H. M., et al. (2017). The arabidopsis UDP-glycosyltransferases UGT79B2 and UGT79B3, contribute to cold, salt and drought stress tolerance *via* modulating anthocyanin accumulation. *Plant J.* 89 (1), 85–103. doi: 10.1111/tpj.13324

Love, M., Huber, W., and Anders, S. (2014). Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* 15 (12), 550. doi: 10.1186/s13059-014-0550-8

Lutaladio, N. B., and Castaldi, L. (2009). Potato: The hidden treasure. *J. Food Compos. Anal.* 22 (6), 491–493. doi: 10.1016/j.jfca.2009.05.002

Lynch, T. J., Erickson, B. J., Miller, D. R., and Finkelstein, R. R. (2017). ABI5-binding proteins (AFPs) alter transcription of ABA-induced genes *via* a variety of interactions with chromatin modifiers. *Plant Mol. Biol.* 93 (4–5), 403–418. doi: 10.1007/s11103-016-0569-1

Ma, Y., Cao, J., He, J., Chen, Q., Li, X., and Yang, Y. (2018). Molecular mechanism for the regulation of ABA homeostasis during plant development and stress responses. *Int. J. Mol. Sci.* 19 (11), 1–14. doi: 10.3390/ijms19113643

Malefo, M. B., Mathibela, E. O., Crampton, B. G., and Makgopa, M. E. (2020). Investigating the role of bowman-birk serine protease inhibitor in arabidopsis plants under drought stress. *Plant Physiol. Biochem.* 149, 286–293. doi: 10.1016/j.plaphy.2020.02.007

Marowa, P., Ding, A., and Kong, Y. (2016). Expansins: roles in plant growth and potential applications in crop improvement. *Plant Cell Rep.* 35 (5), 949–965. doi: 10.1007/s00299-016-1948-4

Monneveux, P., Ramírez, D. A., and Pino, M. T. (2013). Drought tolerance in potato (*S. tuberosum* l.). can we learn from drought tolerance research in cereals? *Plant Sci.* 205–206, 76–86. doi: 10.1016/j.plantsci.2013.01.011

Moon, K. B., Ahn, D. J., Park, J. S., Jung, W. Y., Cho, H. S., Kim, H. R., et al. (2018). Transcriptome profiling and characterization of drought-tolerant potato plant (*Solanum tuberosum* l.). *Mol. Cells* 41 (11), 979–992. doi: 10.14348/molcells.2018.031

Moura, J. C., Bonine, C. A., de Oliveira Fernandes, V. J., Dornelas, M. C., and Mazzafera, P. (2010). Abiotic and biotic stresses and changes in the lignin content and composition in plants. *J. Integr. Plant Biol.* 52 (4), 360–376. doi: 10.1111/j.1744-7909.2010.00892.x

Muthusamy, M., Kim, J. Y., Yoon, E. K., Kim, J. A., and Lee, S. I. (2020). BrEXLB1, a *Brassica rapa* expansin-like b1 gene is associated with root development, drought stress response, and seed germination. *Genes* 11 (4), 404. doi: 10.3390/genes11040404

Nguyen, H. P., Jeong, H. Y., Jeon, S. H., Kim, D., and Lee, C. (2017). Rice pectin methylesterase inhibitor28 (OsPMEI28) encodes a functional PMEI and its overexpression results in a dwarf phenotype through increased pectin methylesterification levels. *J. Plant Physiol.* 208, 17–25. doi: 10.1016/j.jplph.2016.11.006

Oide, S., Bejai, S., Staal, J., Guan, N., Kaliff, M., and Dixelius, C. (2013). A novel role of PR2 in abscisic acid (ABA) mediated, pathogen-induced callose deposition in Arabidopsis thaliana. *New Phytol.* 200 (4), 1187–1199. doi: 10.1111/nph.12436

Otazu, V. (2010). *Manual on quality seed potato production using aeroponics* (Lima (Peru: International Potato Center (CIP), 42, ISBN 978-92-9060-392-4.

Petrov, V., Hille, J., Mueller-Roeber, B., and Gechev, T. S. (2015). ROS-mediated abiotic stress-induced programmed cell death in plants. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00069

Pham, G. M., Hamilton, J. P., Wood, J. C., Burke, J. T., Zhao, H., Vaillancourt, B., et al. (2020). Construction of a chromosome-scale long-read reference genome assembly for potato. *GigaScience* 9 (9), giaa100. doi: 10.1093/gigascience/giaa100

Pieczynski, M., Wyrzykowska, A., Milanowska, K., Boguszewska-Mankowska, D., Zagdanska, B., Karlowski, W., et al. (2018). Genomewide identification of genes involved in the potato response to drought indicates functional evolutionary conservation with arabidopsis plants. *Plant Biotech. J.* 16 (2), 603–614. doi: 10.1111/pbi.12800

Pilling, J., Willmitzer, L., and Fisahn, J. (2000). Expression of a *Petunia inflata* pectin methyl esterase in *Solanum tuberosum* l. enhances stem elongation and modifies cation distribution. *Planta* 210 (3), 391–399. doi: 10.1007/PL00008147

Raudvere, U., Kolberg, L., Kuzmin, I., Arak, T., Adler, P., Peterson, H., et al. (2019). g:Profiler: A web server for functional enrichment analysis and conversions of gene list, (2019 Update). *Nucleic Acids Res.* 47 (W1), W191–W198. doi: 10.1093/nar/gkz369

Sah, S. K., Reddy, K. R., and Li, J. (2016). Abscisic acid and abiotic stress tolerance in crop plants. *Front. Plant Sci.* 7. doi: 10.3389/fpls.2016.00571

Scott, G. (2011). Plants, people, and the conservation of biodiversity of potatoes in Peru. *Nat. Conserv.* 9 (1), 21–38. doi: 10.4322/natcon.2011.003

Segonzac, C., Boyer, J. C., Ipotesi, E., Szponarski, W., Tillard, P., Touraine, B., et al. (2007). Nitrate efflux at the root plasma membrane: Identification of an *Arabidopsis* excretion transporter. *Plant Cell.* 19 (11), 3760–3777. doi: 10.1105/tpc.106.048173

Shen, J., Lv, B., Luo, L., He, J., Mao, C., Xi, D., et al. (2017). The NAC-type transcription factor OsNAC2 regulates ABA-dependent genes and abiotic stress tolerance in rice. *Sci. Rep.* 7, 4–0641. doi: 10.1038/srep40641

Singh, S., Koyama, H., Bhati, K. K., and Alok, A. (2021). The biotechnological importance of the plant-specific NAC transcription factor family in crop improvement. *J. Plant Res.* 134 (3), 475–495. doi: 10.1007/s10265-021-01270-y

Skirycz, A., and Inzé, D. (2010). More from less: Plant growth under limited water. *Curr. Opin. Biotechnol.* 21 (2), 197–203. doi: 10.1016/j.copbio.2010.03.002

Sprenger, H., Kurowsky, C., Horn, R., Erban, A., Seddig, S., Rudack, K., et al. (2016). The drought response of potato reference cultivars with contrasting tolerance. *Plant Cell Environ.* 39 (11), 2370–2389. doi: 10.1111/pce.12780

Sukhotu, T., and Hosaka, K. (2006). Origin and evolution of andigena potatoes revealed by chloroplast and nuclear DNA markers. *Genome* 49 (6), 636–647. doi: 10.1139/g06-014

Sun, Y., and Yu, D. (2015). Activated expression of AtWRKY53 negatively regulates drought tolerance by mediating stomatal movement. *Plant Cell Rep.* 34 (8), 1295–1306. doi: 10.1007/s00299-015-1787-8

Taji, T., Ohsumi, C., Iuchi, S., Seki, M., Kasuga, M., Kobayashi, M., et al. (2002). Important roles of drought- and cold-inducible genes for galactinol synthase in stress tolerance in *Arabidopsis thaliana*. *Plant J.* 29 (4), 417–426. doi: 10.1046/j.0960-7412.2001.01227.x

Takenaka, Y., Nakano, S., Tamoi, M., Sakuda, S., and Fukamizo, T. (2009). Chitinase gene expression in response to environmental stresses in *Arabidopsis thaliana*: Chitinase inhibitor allosamidin enhances stress tolerance. *Biosci. Biotechnol. Biochem.* 73 (5), 1066–1071. doi: 10.1271/bbb.80837

Torres, Y., Lozano, R., Merino, C., and Orjeda, G. (2013). Identificación de genes relacionados a sequía en papas nativas empleando RNA-seq. *Rev. Peru. Biol.* 20 (3), 211–214. doi: 10.15381/rpb.v20i3.5208

Vasquez-Robinet, C., Mane, S. P., Ulanov, A. V., Watkinson, J. I., Stromberg, V. K., De Koeyer, D., et al. (2008). Physiological and molecular adaptations to drought in Andean potato genotypes. *J. Exp. Bot.* 59 (8), 2109–2123. doi: 10.1093/jxb/ern073

Wang, Q., Guo, C., Li, Z., Sun, J., Deng, Z., Wen, L., et al. (2021). Potato NAC transcription factor StNAC053 enhances salt and drought tolerance in transgenic. *Arabidopsis. Int. J. Mol. Sci.* 22 (5), 2568. doi: 10.3390/ijms22052568

Wormit, A., and Usadel, B. (2018). The multifaceted role of pectin methylesterase inhibitors (PMEIs). *Int. J. Mol. Sci.* 19 (10), 2878. doi: 10.3390/ijms19102878

Wu, H. C., Huang, Y. C., Stracovsky, L., and Jinn, T. L. (2017). Pectin methylesterase is required for guard cell function in response to heat. *Plant Signal Behav.* 12 (6), e1338227. doi: 10.1080/15592324.2017.1338227

Wu, Y., Meeley, R. B., and Cosgrove, D. J. (2001). Analysis and expression of the α-expansin and β-expansin gene families in maize. *Plant Physiol.* 126 (1), 222–232. doi: 10.1104/pp.126.1.222

Zaheer, K., and Akhtar, M. H. (2016). Potato production, useage, and nutrition – a review. *Crit. Rev. Food Sci. Nutr.* 56 (5), 711–721. doi: 10.1080/10408398.2012.724479

Zentgraf, U., and Doll, J. (2019). Arabidopsis WRKY53, a node of multi-layer regulation in the network of senescence. *Plant* 8 (12), 578. doi: 10.3390/plants8120578

Zhang, N., Liu, B., Ma, C., Zhang, G., Chang, J., Si, H., et al. (2014). Transcriptome characterization and sequencing-based identification of drought-responsive genes in potato. *Mol. Biol. Rep.* 41 (1), 505–517. doi: 10.1007/s11033-013-2886-7

Zhang, X., Liu, S., and Takano, T. (2008). Two cysteine proteinase inhibitors from *Arabidopsis thaliana*, AtCYSa and AtCYSb, increasing the salt, drought, oxidation and cold tolerance. *Plant Mol. Biol.* 68 (1–2), 131–143. doi: 10.1007/s11103-008-9357-x

Zhang, C., Wang, D., Yang, C., Kong, N., Shi, Z., Zhao, P., et al. (2017). Genome-wide identification of the potato WRKY transcription factor family. *PloS One* 12 (7), e0181573. doi: 10.1371/journal.pone.0181573

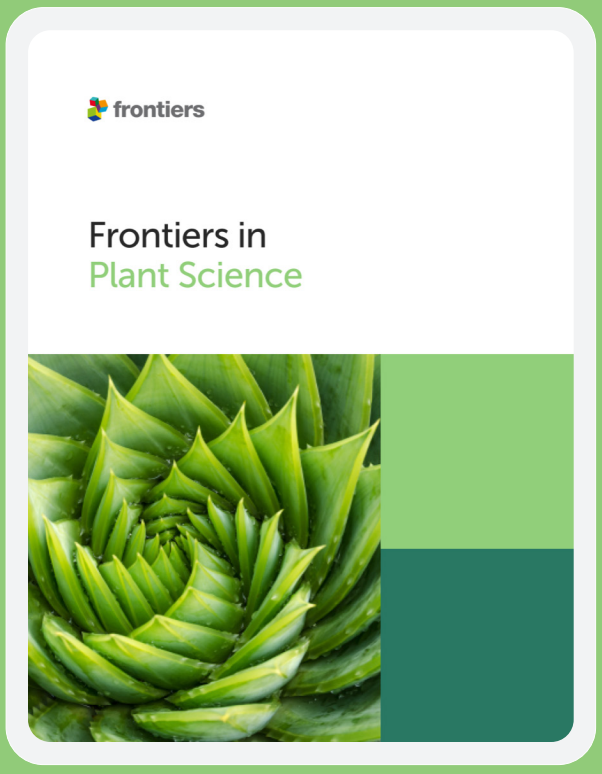
Frontiers in
Plant Science

Cultivates the science of plant biology and its applications

The most cited plant science journal, which advances our understanding of plant biology for sustainable food security, functional ecosystems and human health.

Discover the latest Research Topics

See more →



Frontiers

Avenue du Tribunal-Fédéral 34
1005 Lausanne, Switzerland
frontiersin.org

Contact us

+41 (0)21 510 17 00
frontiersin.org/about/contact