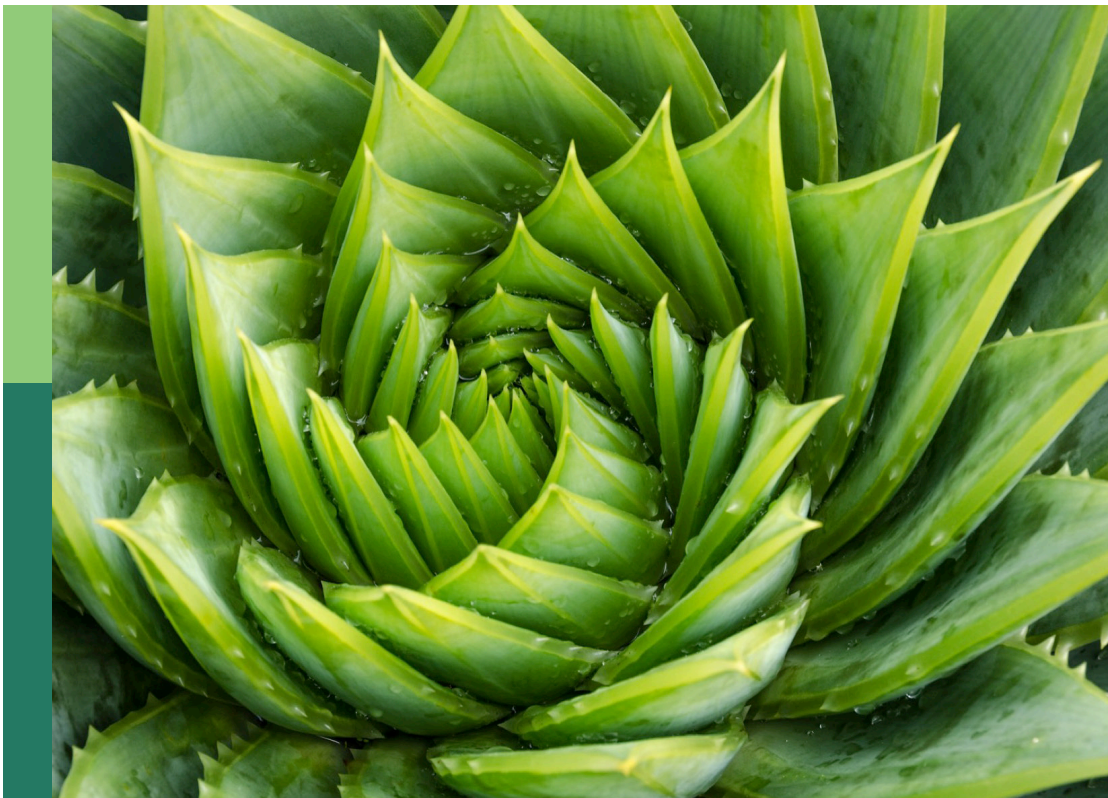
frontiers | Research Topics

Identification of genes involved in environmental adaption and yield traits in forage and other ecologically important plants

Edited by
Qiang Guo, Shangang Jia, Qing Ma, Guo-Qiang Wu, Chi Zhang, Xuemin Wang and Jiejun Xi

Published in
Frontiers in Plant Science



ISSN 1664-8714
ISBN 978-2-8325-3861-6
DOI 10.3389/978-2-8325-3861-6

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

Identification of genes involved in environmental adaption and yield traits in forage and other ecologically important plants

Topic editors

Qiang Guo — Beijing Academy of Agricultural and Forestry Sciences, China
Shangang Jia — China Agricultural University, China
Qing Ma — Lanzhou University, China
Guo-Qiang Wu — Lanzhou University of Technology, China
Chi Zhang — University of Nebraska-Lincoln, United States
Xuemin Wang — Institute of Animal Sciences, Chinese Academy of Agricultural Sciences, China
Jiejun Xi — Northwest A&F University, China

Citation

Guo, Q., Jia, S., Ma, Q., Wu, G.-Q., Zhang, C., Wang, X., Xi, J., eds. (2023). *Identification of genes involved in environmental adaption and yield traits in forage and other ecologically important plants*. Lausanne: Frontiers Media SA.
doi: 10.3389/978-2-8325-3861-6

Table of contents

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 28 October 2022
DOI 10.3389/fpls.2022.1037272

OPEN ACCESS

EDITED BY
Qiang Guo,
Beijing Academy of Agricultural and Forestry Sciences, China

REVIEWED BY
Jiyu Zhang,
Lanzhou University, China
Shan Xiao Wang,
Yangzhou University, China

*CORRESPONDENCE
Tiejun Zhang
tiejunzhang@126.com
Junmei Kang
kangjunmei@caas.cn

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 05 September 2022
ACCEPTED 04 October 2022
PUBLISHED 28 October 2022

CITATION
He X, Zhang F, He F, Shen Y, Yu L-X, Zhang T and Kang J (2022) Accuracy of genomic selection for alfalfa biomass yield in two full-sib populations. *Front. Plant Sci.* 13:1037272. doi: 10.3389/fpls.2022.1037272

Accuracy of genomic selection for alfalfa biomass yield in two full-sib populations

Xiaofan He[1], Fan Zhang[2], Fei He[2], Yuhua Shen[3], Long-Xi Yu[4], Tiejun Zhang[1*] and Junmei Kang[2*]

[1]School of Grassland Science, Beijing Forestry University, Beijing, China, [2]Institute of Animal Science, Chinese Academy of Agricultural Sciences, Beijing, China, [3]College of Chemistry and Life Sciences, Chifeng University, Chifeng, China, [4]Plant Germplasm Introduction and Testing Research, United States Department of Agriculture-Agricultural Research Service, Prosser, WA, United States

Alfalfa (*Medicago sativa*) is one of the most important leguminous forages, widely planted in temperate and subtropical regions. As a homozygous tetraploid, its complex genetic background limits genetic improvement of biomass yield attributes through conventional breeding methods. Genomic selection (GS) could improve breeding efficiency by using high-density molecular markers that cover the whole genome to assess genomic breeding values. In this study, two full-sib F_1 populations, consisting of 149 and 392 individual plants (P149 and P392), were constructed using parents with differences in yield traits, and the yield traits of the F_1 populations were measured for several years in multiple environments. Comparisons of individual yields were greatly affected by environments, and the best linear unbiased prediction (BLUP) could accurately represent the original yield data. The two hybrid F_1 populations were genotyped using GBS and RAD-seq techniques, respectively, and 47,367 and 161,170 SNP markers were identified. To develop yield prediction models for a single location and across locations, genotypic and phenotypic data from alfalfa yields in multiple environments were combined with various prediction models. The prediction accuracies of the F_1 population, including 149 individuals, were 0.11 to 0.70, and those of the F_1 population, consisting of 392 individuals, were 0.14 to 0.67. The BayesC and RF models had the highest average prediction accuracy of 0.60 for two F_1 populations. The accuracy of the prediction models for P392 was higher than that of P149. By analyzing multiple prediction models, moderate prediction accuracies are obtained, although accuracies will likely decline across multiple locations. Our study provided evidence that GS can accelerate the improvement of alfalfa yield traits.

KEYWORDS

***Medicago sativa*, genomic selection, full-sib populations, biomass yield, single nucleotide polymorphism**

Introduction

Alfalfa (*Medicago sativa*) is a perennial leguminous forage with a high yield, high forage value, and wide adaptability and is known as the "Queen of Forages". Yield is one of the essential traits of alfalfa and is the primary goal of breeding. Yield is a quantitative trait controlled by multiple genes and influenced by the environment (Annicchiarico, 2015). Alfalfa is a homozygous tetraploid, has a complex genetic background, and suffers from severe inbreeding depression and non-additive inheritance, thus limiting the process of alfalfa breeding (Bingham et al., 1994; Annicchiarico et al., 2010). In the past decades, yield improvement has been achieved by increasing the tolerance of alfalfa to stress, so this method of yield improvement only works when stressed (Lamb et al., 2006). Conventional yield breeding has stagnated in recent years, and the narrow heritability for alfalfa yield ranges from 0.15 to 0.30. Low breeding values, long selection cycles, high costs, and low efficiency limit the improvement of alfalfa yield, which requires breeders to use new methods instead of conventional phenotypic selection to improve breeding efficiency (Annicchiarico, 2015).

Genomic selection (GS) is a new approach to select candidate individuals at an early stage or in an off-season nursery or greenhouse based on the prediction of breeding values, which can accelerate genetic gain by capturing genetic information through high-density SNP markers (Lande and Thompson, 1990). The GS study is divided into two main parts, marker development and model prediction. With the development of sequencing technology, there have been various strategies for developing SNP markers. Li et al. identified millions of SNPs in alfalfa by transcriptome sequencing, some of which have been used to develop an Illumina Infinium SNP array containing approximately 10,000 SNP markers (Li et al., 2014). With the rapid development of reduced-representation genome sequencing (RRGS) technology, low-cost RRGS results in SNP markers sufficient to cover the entire genome, of which genotyping-by-sequencing (GBS) and restriction site-associated DNA sequence (RAD-seq) are the two most common methods for RRGS. Based on this, GBS can sequence multiple individuals of a species with known genome sequences and perform differential analyses of individuals or populations (Wallace and Mitchell, 2017). The RAD-seq identifies specific restriction endonuclease sites. Hybridizing RAD tags to DNA microarrays allows simultaneous screening of thousands of polymorphic markers and can be used for bulk genotyping of populations (Baird et al., 2008). These sequencing methods offer the opportunity to obtain high-density and high-quality SNP molecular markers. And next-generation sequencing technology largely increases the efficiency of the determination of the exact genotype of an alfalfa individual (Stritzler et al., 2022). The second step is the prediction of the model, which involves phenotypic and genotypic data from the training population to examine the effects of all SNP markers and develop prediction models to predict GS. A test population is also required to validate the accuracy of the prediction model. Individual genetic strengths were estimated directly as the best linear unbiased prediction (BLUP) (Russell et al., 1997). The genomic best linear unbiased prediction (gBLUP) models for individual-based prediction, the ridge-regression best linear unbiased prediction (rrBLUP), and Bayesian models for SNP benefit value-based prediction, support vector machines (SVM) and random forest (RF), two machine learning methods, are the main approaches currently used to perform GS (Corinna and Vladimir, 1995; Meuwissen et al., 2001; Leo, 2001; VanRaden, 2008; Park and Casella, 2008; Endelman, 2011; Habier et al., 2011). These methods have to be selected according to different populations and different heritabilities of traits. Its prediction accuracy finally reveals the best model. Based on the constructed GS prediction equations, researchers estimate the genomic breeding value (GEBV) of individuals and select candidates early in the nursery or greenhouse, thus improving breeding efficiency.

GS breeding methods have been used for many plants, such as barley, maize, and wheat (Mendes and de Souza, 2016; Schmidt et al., 2016; Juliana et al., 2022). As alfalfa is a perennial forage, the collection of multiple harvest yield data for the purpose of evaluating breeding values is required throughout the breeding process. Furthermore, GS can significantly reduce breeding years to 3–4 years of phenotypic selection if breeding values can be assessed based on genotypes at the seedling stage. Li et al. (2015) used the GS breeding method to make single-location, cross-location, and cross-generation predictions in a comprehensive breeding population of alfalfa. Furthermore, their single-location prediction accuracy was 0.43 to 0.66. The accuracy of cross-generation prediction could also reach 0.40. The GS selection efficiency was higher than the phenotypic selection, which indicates that the GS breeding method could improve the alfalfa breeding efficiency (Li et al., 2015). Creating new germplasm with excellent integrative traits by artificial hybridization is a standard method for alfalfa breeding. GS modeling studies have not been reported for this breeding population. Furthermore, in GS modeling studies, population type, population size, and the number of SNP markers dramatically influence the accuracy of GS prediction, and different GS prediction models can affect prediction accuracy (Paolo et al., 2015). Recent developments in genotyping and prediction models have made it possible to investigate GS in alfalfa breeding populations.

The present study evaluated two full-sib populations of 149 and 392 progeny lines segregating for yield-related traits in three field trials over seven years. The two F_1 populations were then genotyped using GBS and RAD-seq to obtain SNP markers covering the entire genome. The number of SNP markers from the two library construction methods was compared, and eight GS prediction models were established to analyze the effects of population size and SNP markers on the accuracy of genome prediction.

Materials and methods

Plant materials and experimental design

Population development and field management were provided in a previous study (Zhang et al., 2019; Zhang et al., 2020). Due to its excellent forage yield, large leaves, high stem/leaf ratio, and good persistence of the yield, the same maternal parent, P2, was derived from cultivar Zhongmu No. 1 (CF0032020). Two separate P1 paternal parents (P1-1 and P1-2) with low yield but early flowering time were chosen from a local cultivar, Cangzhou (CF000735, Figure 1). The first population, which was designated as P149, included 149 F_1 individuals, whereas the second population, which was designated as P392, included 392 F_1 individuals (Figure 2). From 2014 to 2019, in Tongzhou and Langfang, eight environments for P149 were sampled for phenotyping. From 2016 to 2020, phenotypic data of P392 from eight environments were also recorded at Langfang and Changping. Harvest dates varied from year to year and depended on the location (Table 1).

Phenotypic data analysis

Every year, one to four harvests were carried out at each location (Table 1). Fresh yield of each plant were evaluate, and random samples of individual were harvested for biomass, weighed while still wet, and dried for five days at 60°C in a forced-air facility. The average dry matter biomass yield was calculated based on the average dry-matter content of each plant.

BLUP was performed to evaluate the phenotypic results collected over the years at various locations using PROC MIXED (Russell et al., 1997).

The following was the random-effects model utilized for BLUP:

$$Y_{ijkh} = m + l_k + r_{i(k)} + g_j + y_h + gl_{jk} + gy_{jh} + gly_{jkh} + e_{ijkh}$$

where Y_{ijkh} represents the Y for the jth genotype in the ith replication of the kth location in the hth year; m represents the grand mean; $r_{i(k)}$ represents the effect of the ith replication nested in the kth location; y_h represents the effect of the hth year; g_j represents the genetic effect of the jth genotype; gl_{jk} represents the interaction effect of the jth genotype and kth location. The interaction effect of the jth genotype and the hth year is denoted by the symbol gy_{jh}; the interaction effect of the jth genotype, the kth location, and the hth year is denoted by the symbol gly_{jkh}; and the residual is denoted by the symbol e_{ijkh}. The grand mean was the only element that was not determined by random effects.

The BLUP estimate was carried out using the lme4 package in R (Bates et al., 2015). Estimated BLUPs were used as observed phenotypic values of total biomass output to make a more accurate GS prediction model.

DNA isolation, RAD, and GBS library construction

DNA extraction and the development of RAD for P149 and GBS for P392 libraries have been described in previous studies (Zhang et al., 2019; Zhang et al., 2020). The RAD sequences and the raw GBS data were uploaded to the NCBI Sequence Read Archive with the BioProject IDs PRJNA503672 and PRJNA522887.

Sequence analysis and SNP genotyping

The genome of *M. sativa* 'Xinjiangdaye' was used as a reference for GBS sequencing analysis of P392 (Chen S. et al., 2020). TASSEL-GBS was used to call SNP with the default

FIGURE 1
The phenotypic difference between the parents P1-1, P1-2 and P2 for flowering time.

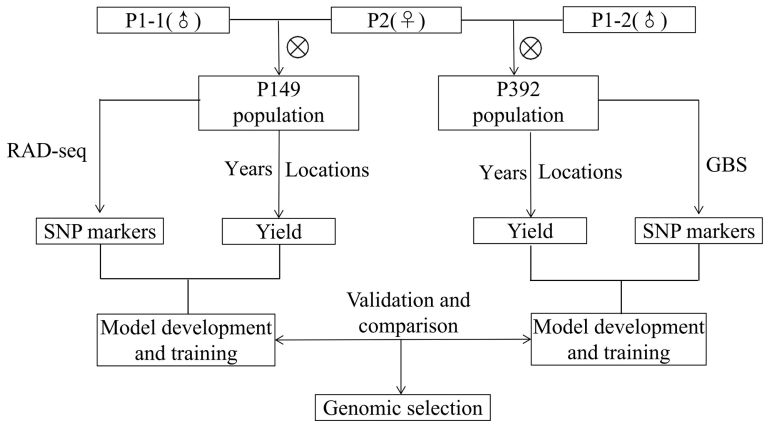


FIGURE 2
Schematic diagram of training, validation and comparison for genomic selection of alfalfa in full-sib populations.

parameter (Glaubitz et al., 2014). SNPs with a 50% missing rate and a MAF of 0.05 were filtered using VCFtools (Danecek et al., 2011).

The sequence data from the RAD-seq for P149 was initially filtered for quality using the Trimmomatic program with default parameters (Bolger et al., 2014). The reads obtained from paired-end sequencing were used in BWA-MEM to map *M. sativa* 'Xinjiangdaye' genome, using the program's default mapping settings (Li and Durbin, 2010). It was achieved with SAMtools by converting the SAM data into BAM files and then sorting the BAM files with the default settings (Li et al., 2009). Both Picard Tools and Genome Analysis ToolKit were used in this study. Picard Tools (http://broadinstitute.github.io/picard/) was utilized to mark duplicate reads, while Genome Analysis ToolKit was utilized to repair indels that can be mistaken for SNPs (Van der Auwera et al., 2013). To identify SNPs, the SAMtools mpileup and VarScan programs were utilized (Koboldt et al., 2012). Furthermore, the SNP data were filtered using VCFtools to have a missing rate of less than 10%, a minor allele frequency greater than 0.05, and a mean read depth greater than 20.

The SNP data from the P149 and P392 populations were uploaded to figshare (http://figshare.com) with the following dios: https://doi.org/10.6084/m9.figshare.21172051.v1 and https://doi.org/10.6084/m9.figshare.21162283.v2.

TABLE 1 Biomass yield harvested by the two populations.

Population	Location	Year	Harvests	Max	Min	Ave.	SE	CV	H^2
P149	Tongzhou (TZ), Beijing	2014	2	137.07	4.17	56.19	1.66	0.52	0.83
		2015	4	429.17	12.22	104.77	3.78	0.89	0.66
	Langfang (LF), Hebei	2014	3	650.00	9.48	278.87	5.95	0.53	0.44
		2015	4	270.19	2.80	108.53	2.48	0.49	0.91
		2016	4	690.00	14.00	226.65	5.31	0.58	0.91
		2017	1	448.03	22.83	217.12	8.47	0.48	0.02
		2018	1	644.33	99.60	337.46	11.49	0.42	0.05
		2019	4	489.68	1.75	111.94	4.07	0.90	0.94
P392	Langfang (LF), Hebei	2016	3	710.00	14.00	203.02	3.50	0.59	0.97
		2017	1	625.23	25.75	253.99	7.04	0.55	0.83
		2018	3	953.70	6.63	285.15	4.88	0.59	0.80
		2019	2	728.40	20.00	252.88	4.54	0.50	0.96
	Changping (CP), Beijing	2017	4	514.59	5.18	108.56	1.98	0.72	0.81
		2018	3	605.10	22.70	208.64	2.90	0.48	0.95
		2019	4	691.00	25.00	197.27	2.94	0.59	0.89
		2020	4	560.60	2.23	153.31	2.53	0.65	0.97

Correlation analysis

The relationships between the harvest phenotype value in different years and the BLUP estimate in different locations were determined by correlation analysis.

Model development and validation for genomic selection prediction

A variety of statistical models have been devised for genomic selection (Heffner et al., 2009; Lorenz et al., 2011). The present study evaluated gBLUP, rrBLUP, four Bayesian models, SVM, and RF. The accuracy of the predictions was determined by calculating the Pearson correlation between the predicted and observed phenotypes. To achieve this, 90% of the individuals were assigned to a training set, while the remaining 10% were assigned to a validation set. The genomic prediction accuracy was estimated as the Pearson correlation (r) between estimated GEBVs and estimated BLUPs of phenotypic values. This cross-validation process was carried out 5000 times, and the accuracy ratings were averaged.

gBLUP involves the construction of a covariate matrix by utilizing individual kinships, followed by predicting the phenotype through utilizing the values of individual species (VanRaden, 2008). The following constitutes the statistical model:

$$y = Xb + Zu + e,$$

where y represents the association matrix of the phenotypic vector, X represents the association matrix of the fixed effect coefficient, and b represents the fixed effect; Z represents the association matrix of random additive genetic effects; u represents a random additive genetic effect, also known as an individual genomic breeding value; and e represents a residual effect.

rrBLUP operates under the presumption of a linear mixed additive model, in which each marker is given an effect as a solution to the following equation:

$$y = m + Zu + e,$$

where y is the observed value of the phenotype, m represents the intercept, Z represents the marker matrix, and u represents a vector of estimated marker effects (Endelman, 2011).

Models based on Bayesian theory provide prior densities for the marker effects that induce the various forms of shrinkage. The answer is produced using a Gibbs sampling strategy to take samples from the resultant posterior density to solve the problem, as detailed in (Geman and Geman, 1984; George and Edward, 2012). In this study, we evaluated the phenotypic prediction capabilities of four different Bayesian models: (i) BayesA (Meuwissen et al., 2001), (ii) BayesB (Habier et al., 2011), (iii) BayesC (Habier et al., 2011), and (iv) Bayesian Ridge regression (Park and Casella, 2008). Using the following settings, the Bayesian models were examined using the BGLR package of the R programming language (Perez and de Los Campos, 2014). The number of iterations equals 5000, while the burnin threshold is set at 1500.

RF uses the bootstrap resampling method to extract multiple samples from the original sample, model a decision tree for each bootstrap sample, combine the predictions of multiple decision trees, and obtain the final prediction result through voting (Leo, 2001). SVM is a model for binary classification, and its learning algorithm is an optimization algorithm for solving convex quadratic programming (Corinna and Vladimir, 1995).

Results

Phenotypic analysis

In the two breeding populations, the coefficient of the variation of annual yield ranged from 0.42 to 0.90 (Table 1), indicating that the yields of the breeding individuals were quite different. In the P149 population, the broad-sense heritability (H^2) of LF_2017 and LF_2018 was only 0.02 and 0.05, indicating that the yield in these two years was greatly affected by the environment (Table 1).

The yield was evaluated by employing the BLUP. Correlation analysis was performed between different years of the two breeding populations using the mean value of the annual yield data and BLUP (Figures 3, 4). The correlations of the yields in different years ranged from 0.19 to 0.73, indicating that the comparisons of the individual yields in different years were quite different. For the yields of P149 and P392, the correlation coefficients between BLUP and each year ranged from 0.63 to 0.91. Overall, the correlation coefficients between BLUP and each year were higher than the values between different years, the only exception being in 2015 and 2016, when the P149 population gave the higher correlation coefficients, 0.66 rather than 0.64.

Between different locations, a correlation analysis was performed using the mean and BLUP values, and the coefficients for P149 and P392 were 0.47 and 0.59, respectively, indicating a significant difference in yields between different locations (Figure 5). In the P392 population, the correlation coefficients between two locations and the overall BLUP were 0.92 and 0.85, indicating that the overall BLUP could represent the value of each location (Figure 5B). In the P149 population, the correlation coefficients were 0.61 for TZ_BLUP and TZ-LF_BLUP and 0.97 for LF_BLUP and TZ-LF_BLUP (Figure 5A), which were consistently greater than 0.60.

Using the BLUP values, the H^2 of the P149 population at the TZ and LF locations was 0.48 and 0.57, respectively, while the H^2

10.3389/fpls.2022.1037272

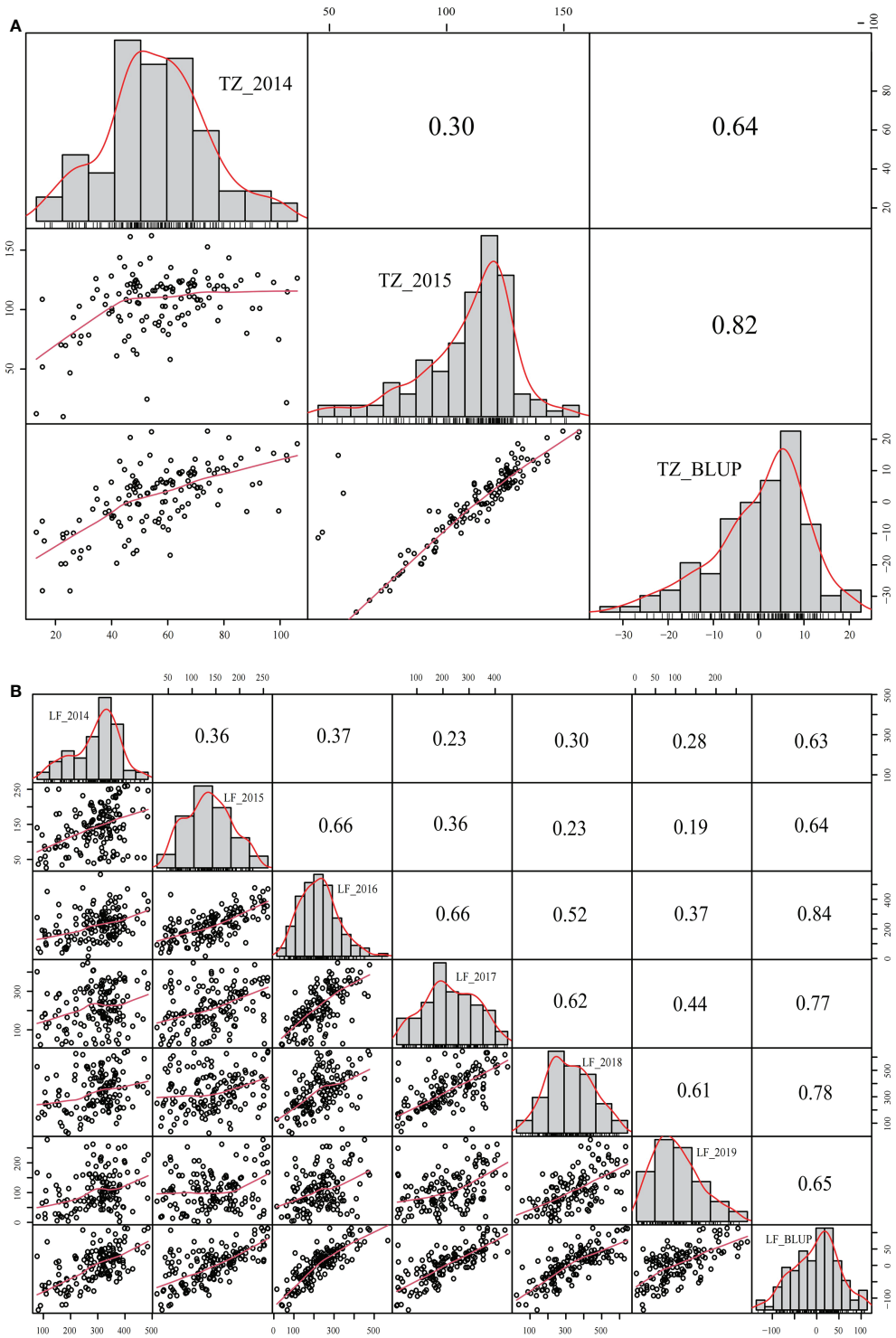


FIGURE 3
Correlation analysis of different years in the P149 population. **(A)** BLUP was calculated based on years in Tongzhou, Beijing. **(B)** BLUP was calculated based on years in Langfang, Hebei. BLUP, best linear unbiased prediction.

10.3389/fpls.2022.1037272

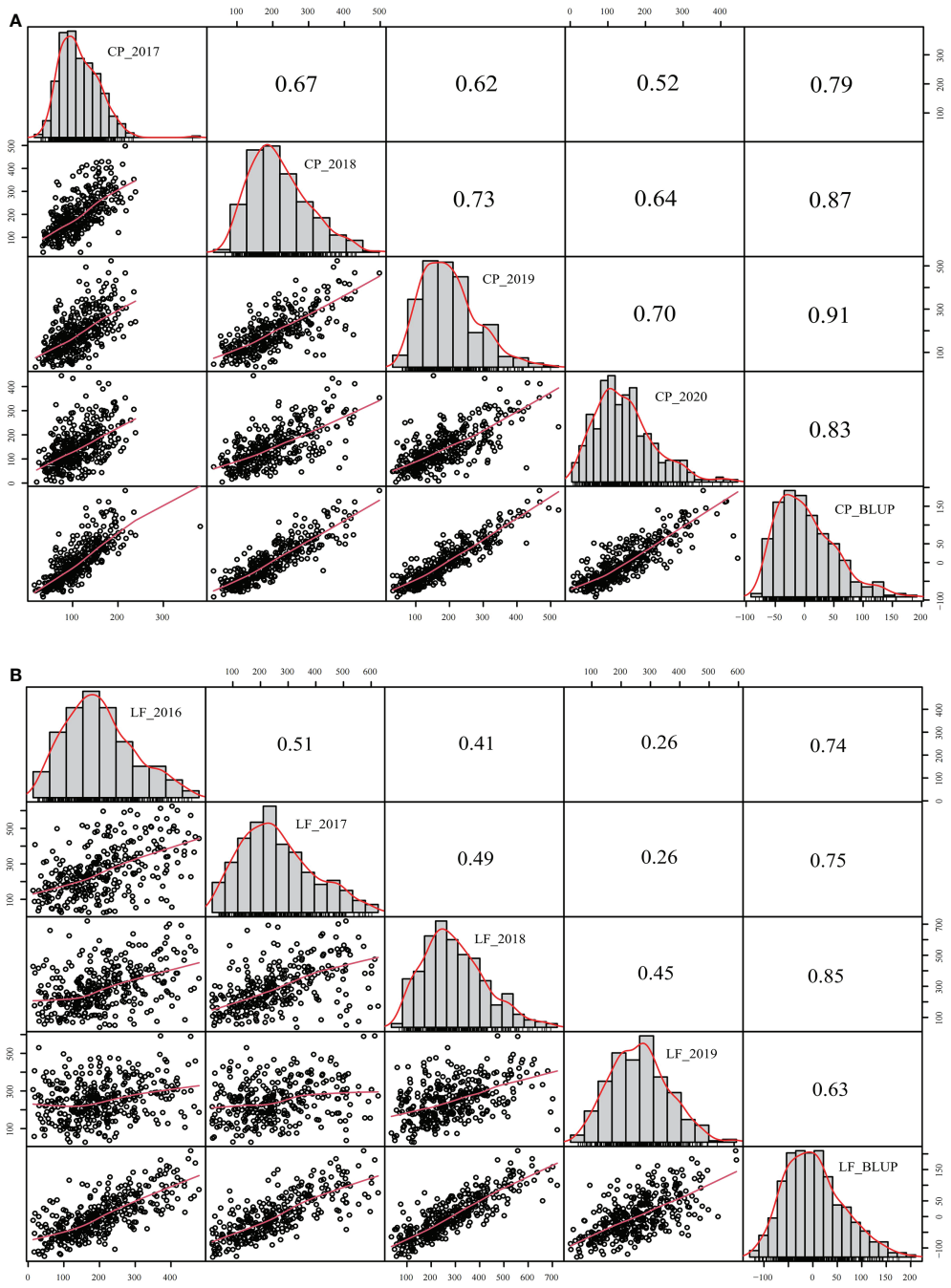


FIGURE 4
Correlation analysis of different years in the P392 population. **(A)** BLUP was calculated based on years in Changping, Beijing. **(B)** BLUP was calculated based on years in Langfang, Hebei. BLUP, best linear unbiased prediction.

of the combined location was only 0.19, which was much lower than the H^2 of the individual locations (Table 2). These results showed significant genetic diversity in the biomass produced by hybrid populations during each harvest and in each environment. On the other hand, the accuracy of the BLUP value and the means of the original data in the two populations was between 0.90 and 0.98, and the correlation coefficients of the combined location were higher than the values of the individual locations, indicating that BLUP could accurately represent the original yield data (Table 2 and Figure 6).

10.3389/fpls.2022.1037272

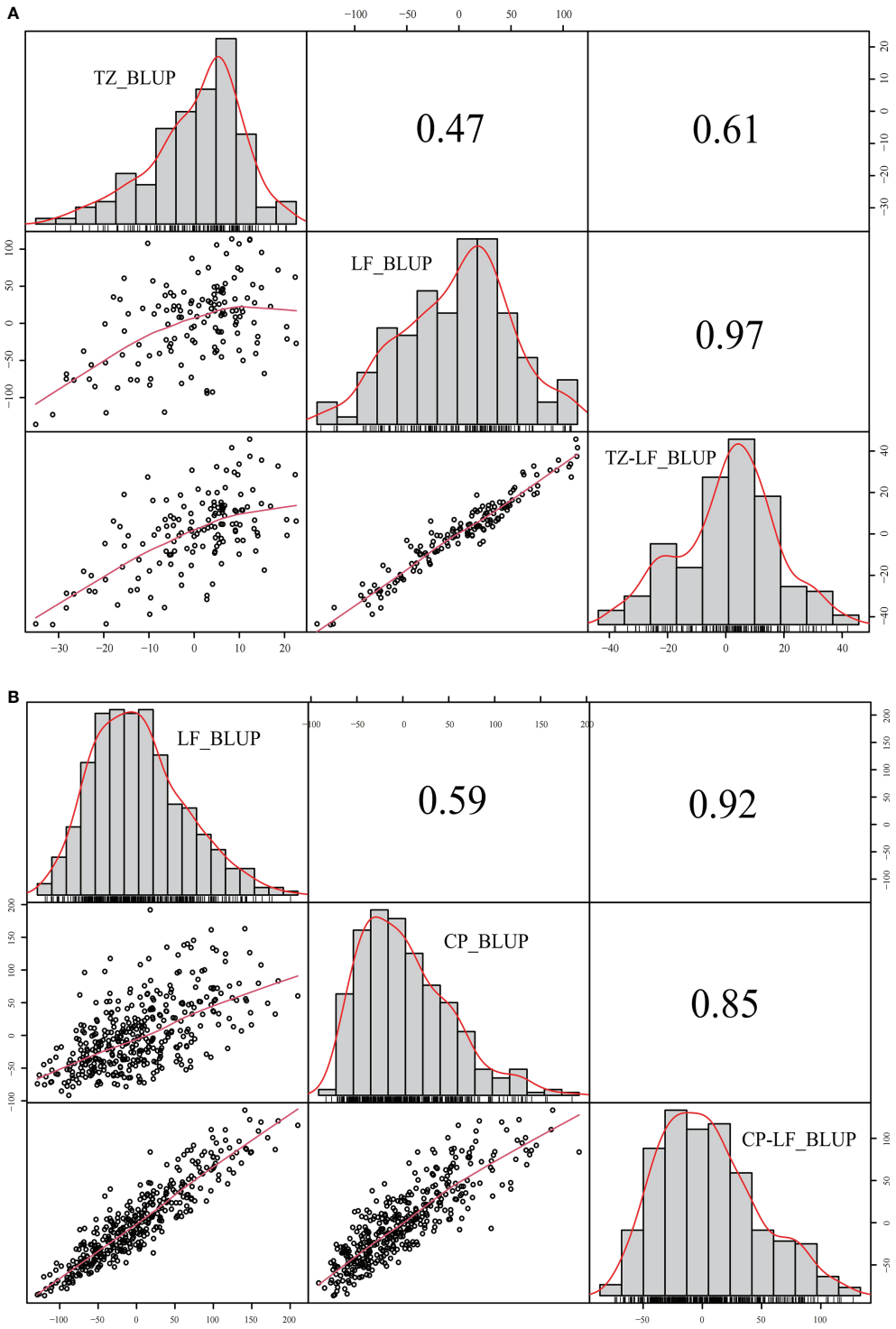


FIGURE 5
Correlation analysis of different locations in the P149 and P392 populations. **(A)** BLUP was calculated based on locations in the P149 population. **(B)** BLUP was calculated based on locations in the P392 population. Abbreviation: BLUP, best linear unbiased prediction.

TABLE 2 The result of BLUP in LF and TZ. .

Population	Location†	min	max	mean	cor‡	H^2
P149	TZ	-35.05	22.58	0.00	0.90	0.48
	LF	-136.14	113.80	0.00	0.94	0.57
	TZ-LF	-43.80	45.72	0.00	0.95	0.19
P392	LF	-129.11	209.97	0.00	0.98	0.57
	CP	-91.63	191.90	0.00	0.98	0.69
	LF-CP	-86.42	133.38	0.00	0.98	0.48

†TZ, Tongzhou, Beijing; LF, Langfang, Hebei; TZ-LF is the total biomass yield at both TZ and LF locations; CP, Changping, Beijing; LF-CP is the total biomass yield at both LF and CP locations. ‡Correlation analysis was performed between the means of phenotype value and BLUP estimate in the two populations.

Genotyping-by-sequencing genotype data

After quality filtering and processing, a total of 114,943,310,784 sequencing reads were obtained in P149, with an average of 771,431,616 per genotype, ranging from 197,919,225 to 2,813,293,425. A total of 98,802,178,615 sequencing reads were obtained in the P392 population. The average number of reads per genotype was 252,046,374, ranging from 2,637,626 to 779,056,669.

To analyze the sequence reads obtained from two breeding populations coupled with eight chromosomes of *M. sativa* 'Xinjiangdaye', we utilized a variety of UNEAK pipelines. After thoroughly examining the readings, we discovered 161,170 SNP markers for the breeding population of P149. After removing markers with missing values greater than 50% of the time, 47,367 SNP markers were retained in the P392 breeding population.

Furthermore, since RAD-seq and GBS sequencing were used in the two populations, we further analyzed the percentage of SNP makers captured jointly and found that the common SNP makers of the two populations were extremely low. A total of 116 shared SNP loci were obtained after the first attempt, which only accounted for 0.07% of the total number of SNPs sequenced by RAD-seq.

Validation of the P149 population

In P149, the total biomass output was calculated using the BLUP based on six harvests of two years in Tongzhou and

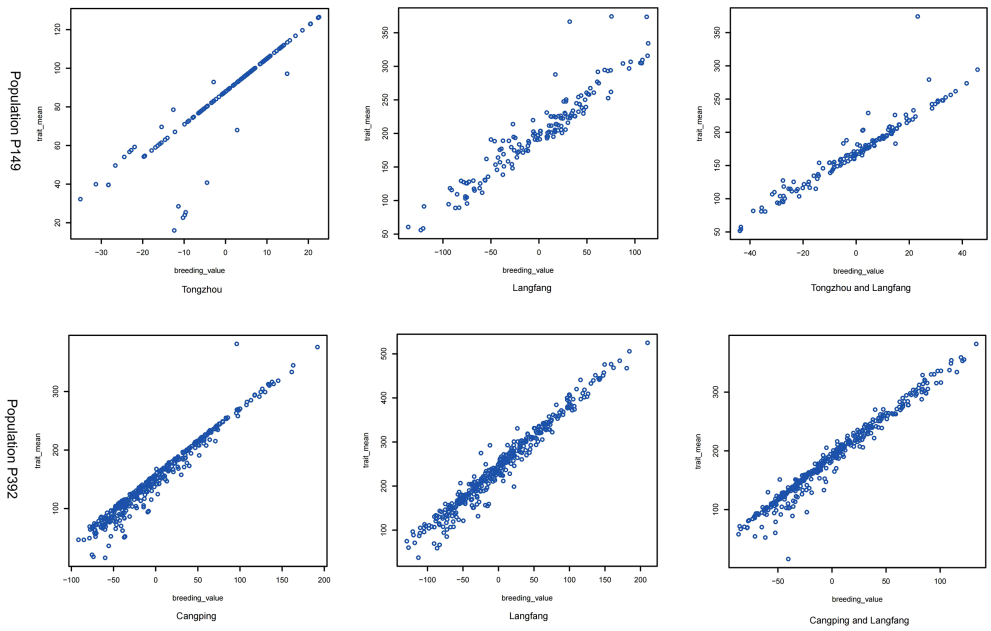


FIGURE 6
The correlation of BLUP (trait_mean, harvest phenotype value; breeding_value, BLUP phenotype value) in the P149 and P392 populations at a single location and across locations.

seventeen harvests of six years in Langfang. Furthermore, the BLUP values were used as phenotypic data for the GS prediction model. Using the dataset of 161170 SNPs, each one had fewer than 50% of its genotypic values missing.

We used eight GS models to estimate the accuracy of genomic prediction of single location and across locations, and the accuracy of these models ranged from 0.11 to 0.70. The BayesC model had the highest prediction accuracy of 0.70 in Tongzhou, the BayesB model had the highest prediction accuracy of 0.58 in Langfang, and the BayesB model had the highest prediction accuracy of 0.55 for total biomass at both sites (Table 3 and Figure 7A). Among these eight models, the gBLUP model predicted the lowest average accuracy of 0.22, while the BayesB and BayesC models predicted the highest average accuracy of 0.43.

It should be noted that the accuracy of the predictions of the TZ-LF and TZ-TZ-LF models for the four Bayesian models was very low in P149, and to explore the reason for this, we subsequently performed the validation of the GS model following the same method using biomass data for two years (2014 and 2015) from Tongzhou and Langfang. The model prediction accuracies ranged from 0.17 to 0.70, with single-location predictions from TZ, LF, and TZ-LF having the highest accuracy of the BayesC model at 0.70, 0.57, and 0.65, respectively. The gBLUP model had the lowest prediction accuracy of 0.19, and the BayesC model had the highest prediction accuracy of 0.60 among the eight models (Table 4).

Validation within the P392 population

For the P392 population, the total biomass yield was calculated by applying BLUP, using harvest yield data for four years of nine harvests at Langfang and four years of 15 harvests at Changping, and the BLUP values were used for the GS prediction model. When using the data set consisting of 47,367 SNPs, each had fewer than 50% of their genotypic values missing.

The highest accuracy of the total biomass yield was predicted by the RF model in Langfang, two locations, and the rrBLUP model in Changping, with 0.56, 0.67, and 0.61, respectively. The

TABLE 3 The genomic prediction accuracy of total biomass yield in the P149 population validated within and across locations.

		Validation†			
Model	**Estimation**	**TZ**	**LF**	**TZ-LF**	**Mean**
rrBLUP	TZ	0.47	0.33	0.34	0.39
	LF	0.38	0.40	0.41	
	TZ-LF	0.39	0.40	0.41	
gBLUP	TZ	0.18	0.18	0.19	0.22
	LF	0.19	0.25	0.26	
	TZ-LF	0.19	0.25	0.25	
BRR	TZ	0.69	0.15	0.20	0.40
	LF	0.36	0.46	0.49	
	TZ-LF	0.42	0.40	0.42	
Bayes A	TZ	0.68	0.11	0.17	0.43
	LF	0.34	0.55	0.58	
	TZ-LF	0.40	0.49	0.53	
Bayes B	TZ	0.69	0.11	0.17	0.43
	LF	0.31	0.58	0.60	
	TZ-LF	0.37	0.52	0.55	
Bayes C	TZ	0.70	0.20	0.23	0.43
	LF	0.33	0.52	0.55	
	TZ-LF	0.39	0.46	0.49	
SVM	TZ	0.44	0.29	0.32	0.36
	LF	0.34	0.38	0.39	
	TZ-LF	0.35	0.38	0.38	
RF	TZ	0.44	0.27	0.29	0.36
	LF	0.42	0.36	0.38	
	TZ-LF	0.41	0.32	0.38	

†TZ, Tongzhou, Beijing; LF, Langfang, Hebei; TZ-LF is the total biomass yield at both TZ and LF locations.

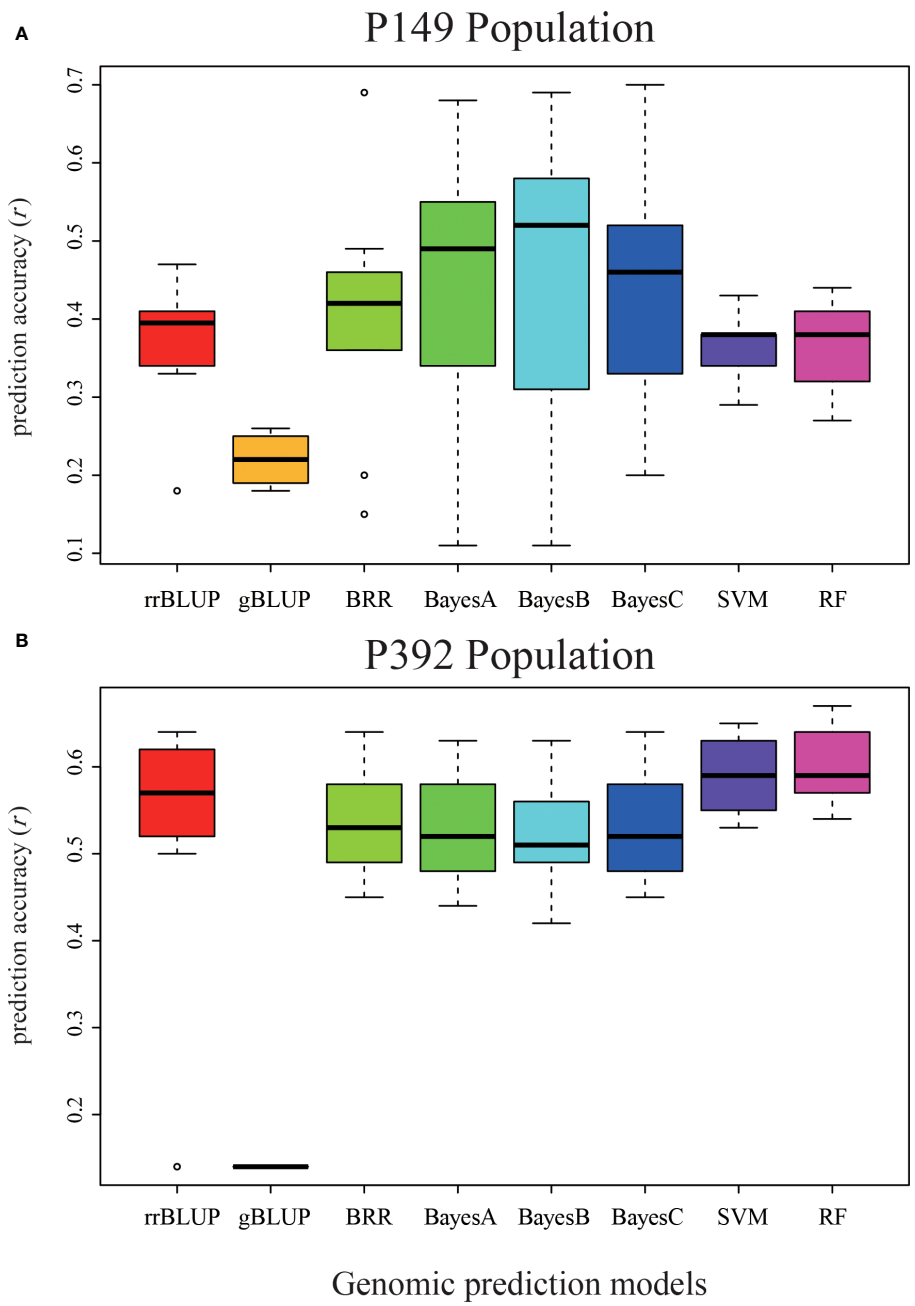


FIGURE 7
Validation of eight different models in the P149 and P392 populations. **(A)** The genomic prediction accuracy of total biomass yield in the P149 population was validated. **(B)** The genomic prediction accuracy of total biomass yield in the P392 population. rrBLUP, ridge-regression best linear unbiased prediction; gBLUP, genomic best linear unbiased prediction; BRR, Bayesian Ridge regression; BayesA, BayesianA; BayesB, BayesianB; BayesC, BayesianC; SVM, support vector machines; RF, random forest.

overall prediction accuracy of the eight models ranged from 0.14 to 0.67 (Table 5). The gBLUP model predicted the lowest average accuracy of 0.14, and the RF model predicted the highest average accuracy of 0.60 (Table 5 and Figure 7B). The highest average accuracy of 0.60 was achieved by the RF model.

Validation based on the number of markers

In the P149 population, we randomly selected different SNP markers for the GS analysis of the Tongzhou yield data using the

TABLE 4 The genomic prediction accuracy for two years (2014 and 2015) of biomass yield in the P149 population validated within and across locations.

		Validation†			
Model	**Estimation†**	**TZ**	**LF**	**TZ-LF**	**Mean**
rrBLUP	TZ	0.47	0.42	0.41	0.45
	LF	0.50	0.44	0.43	
	TZ-LF	0.49	0.43	0.43	
gBLUP	TZ	0.18	0.17	0.18	0.19
	LF	0.20	0.19	0.21	
	TZ-LF	0.19	0.19	0.20	
BRR	TZ	0.69	0.53	0.62	0.58
	LF	0.54	0.54	0.63	
	TZ-LF	0.55	0.53	0.62	
Bayes A	TZ	0.68	0.54	0.63	0.58
	LF	0.50	0.55	0.64	
	TZ-LF	0.53	0.55	0.64	
Bayes B	TZ	0.69	0.52	0.62	0.58
	LF	0.51	0.53	0.62	
	TZ-LF	0.54	0.53	0.63	
Bayes C	TZ	0.70	0.57	0.65	0.60
	LF	0.52	0.57	0.65	
	TZ-LF	0.54	0.57	0.65	
SVM	TZ	0.44	0.37	0.38	0.39
	LF	0.42	0.37	0.37	
	TZ-LF	0.44	0.38	0.37	
RF	TZ	0.44	0.36	0.36	0.41
	LF	0.53	0.35	0.35	
	TZ-LF	0.47	0.42	0.38	

†TZ, Tongzhou, Beijing; LF, Langfang, Hebei; TZ-LF is the total biomass yield at both TZ and LF locations.

rrBLUP model, with five replications for each number of SNP markers. When the number of SNP markers is 1000, the prediction accuracy is at least 0.45, and when the number of SNP markers is more than 5000, the prediction accuracy is 0.47 in all cases (Figure 8).

Discussion

Genotyping-by-sequencing and restriction-site-associated DNA sequencing

GBS and RAD-Seq are two of the most common simplified genome sequencing technologies that have been used to detect SNP loci in the genomes of many plants. The GBS and RAD-seq methods employ methylation-sensitive enzymes to trim genomic sequences and decrease redundancy. However, the sensitivity enzymes and sequencing methodologies used by the two systems are distinct (Zhang et al., 2019). In this study, we performed RAD-Seq and GBS on the P149 and P392 populations, respectively, to explore the effect of the number of SNP markers on the prediction accuracy of the GS model.

GBS simplifies the library building step and is a cost-effective method for SNP typing, but the number of SNPs obtained is significantly lower than RAD-Seq (Davey et al., 2011). In the P149 population, we identified 161,170 SNP loci in tetraploid alfalfa that could be used for genotyping using RAD-Seq, and in the P392 population, we obtained 47,367 SNP loci that could be used for genotyping using GBS, which was only 29.4% of the RAD-Seq method. It showed that the RAD-Seq sequencing method obtained significantly more SNPs than the GBS method.

Due to the covariance between the genomes of *M. truncatula* and *M. sativa*, many previous studies have used the *M. truncatula* genome as a reference for genotype data analysis, although *M. truncatula* is self-pollinated and diploid (Chen H. et al., 2020). The most significant challenge we encounter with alfalfa (as we do with other allogamous polysomic polyploids) is the presence of heterozygosity. Multiple readings are necessary to get an accurate determination of whether a genotype is heterozygous or homozygous. The availability of the whole genome sequence of *M.*

10.3389/fpls.2022.1037272

TABLE 5 The genomic prediction accuracy of total biomass yield in the P392 population validated within and across locations.

Model	Estimation†	Validation†			
		LF	CP	LF-CP	Mean
rrBLUP	LF	0.52	0.55	0.62	0.58
	CP	0.50	0.61	0.62	
	LF-CP	0.53	0.59	0.64	
gBLUP	LF	0.14	0.14	0.14	0.14
	CP	0.14	0.14	0.14	
	LF-CP	0.14	0.14	0.14	
BRR	LF	0.46	0.45	0.53	0.53
	CP	0.53	0.58	0.64	
	LF-CP	0.49	0.53	0.58	
Bayes A	LF	0.46	0.44	0.52	0.53
	CP	0.51	0.58	0.63	
	LF-CP	0.48	0.53	0.58	
Bayes B	LF	0.45	0.42	0.51	0.52
	CP	0.52	0.56	0.63	
	LF-CP	0.49	0.50	0.57	
Bayes C	LF	0.45	0.45	0.52	0.53
	CP	0.52	0.58	0.64	
	LF-CP	0.48	0.53	0.58	
SVM	LF	0.54	0.58	0.63	0.59
	CP	0.53	0.60	0.65	
	LF-CP	0.55	0.59	0.65	
RF	LF	0.56	0.59	0.64	0.60
	CP	0.57	0.59	0.67	
	LF-CP	0.54	0.59	0.67	

†LF, Langfang, Hebei; CP, Changping, Beijing; LF-CP is the total biomass yield at both TZ and LF locations.

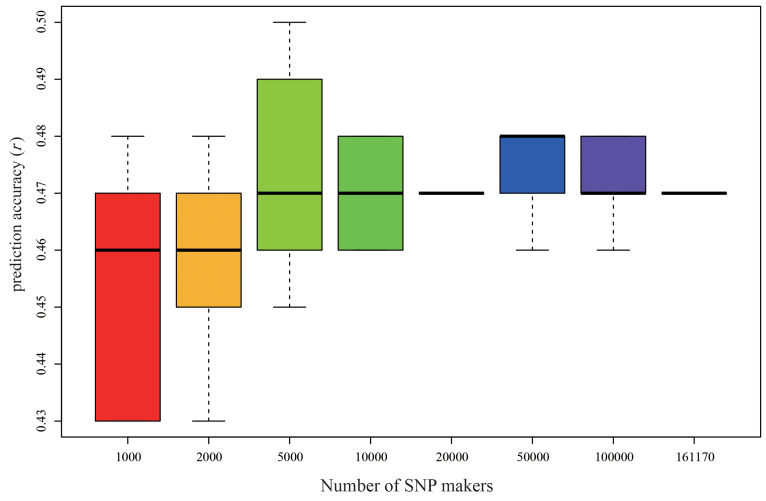


FIGURE 8
Genomic prediction accuracy of biomass yield at Tongzhou in P149 using different levels of marker datasets by rrBLUP. rrBLUP, ridge-regression best linear unbiased prediction.

sativa 'Xinjiangdaye' (Chen S. et al., 2020) allows us to compare the simplified genome sequence with the entire alfalfa genome, eliminating the dilemma of using *M. truncatula*. In this study, we aligned the simplified genome sequences with the reference genome of alfalfa and obtained 161,170 and 47,367 SNP markers that increased by 42% and 65% compared to the previous analysis using the *M. truncatula* genome (Chen S. et al., 2020). Increasing the number of SNP loci can significantly increase the power of association mapping and genomic prediction.

Genomic prediction accuracy based on the number of markers

As the number of SNPs increased, the accuracy of the GS model prediction also increased, but the accuracy of the predictions eventually reached a plateau. Predictions using 35,000 GBS SNP markers with up to 80% missing data were more accurate than predictions using 2000 diversity array technology (DArT) markers with 2% missing data in barley (*Triticum aestivum* L.) (Poland et al., 2012). For increasing marker counts and avoiding bias caused by DArT markers, GBS was mainly considered for increasing the accuracy of genomic prediction (Lu et al., 2013), thus improving the accuracy of genomic prediction. In alfalfa, the average accuracy of GS increased as the number of markers increased until the amount of missing data per marker exceeded the limitation of 70 to 80%, as a higher imputation error may be caused by higher missing values (Li et al., 2015). In this study, the number of SNP markers increased from 1000 to 161,170, but the prediction accuracy did not improve significantly and fluctuated around 0.47 when the number of SNP markers exceeded 5000.

Since simplified sequencing can only cover about 3% of the genomic regions, GBS and RAD-seq can capture a few genomic regions together, and thus the number of common SNPs that can be detected is also small. In this study, after several attempts, we found that the common SNP loci only represented 0.07% of the total SNPs sequenced by RAD-seq, which was close to the theoretical value of 0.09%. Therefore, it was impossible to combine the data from the two sequencing results for the integration validation of the yield model in two populations, P149 and P392, at a later stage. In order to improve the applicability of the yield model, representative loci with comprehensive coverage and uniformity should be selected to design genotype detection chips and establish a yield prediction model based on a uniform genotype data matrix. Then, when genotype analysis is done on the population to be tested, the yield model created in the first step can be used to predict the yield of the population.

Genomic prediction accuracy based on the model developed

GS breeding methods have been used for many plants, such as winter barley, with prediction accuracies of 0.80 for malt quality, 0.72 for maize yield, and 0.94 for oil palm fruit yield (Cros et al., 2015; Mendes and de Souza, 2016; Schmidt et al., 2016). In this study, two half-sib populations were harvested at three different locations, and multi-year yield data were collected. We obtained 0.70 accuracy using the BayesC model in the P149 population, indicating that sufficient accuracy can be predicted for complex traits such as alfalfa yield in half-sib alfalfa populations.

Different GS models resulted in different accuracies in marker assumptions and treatment effects. This study used the gBLUP model, which predicts on an individual basis, and the rrBLUP, BRR, BayseA, BayesB, and BayesC models, which predict on an SNP benefit value basis. The results showed that gBLUP was less effective than the other models, which indicates that predicting yield using individuals is not ideal for alfalfa. Moreover, prediction based on SNP effect values could lead to better prediction accuracy.

The worst prediction accuracy of the gBLUP model may be related to the fact that the gBLUP model is based on individual prediction, with a heritability of yield traits of approximately 19% and approximately 80% of phenotypic variation that cannot be predicted by individual effects, resulting in lower prediction accuracy. Furthermore, the covariance used in this study was replaced by the principal component analysis (PCA) results using TASSEL to calculate genotype data. However, the covariance between individuals for a specific trait is determined by the association at the causal loci rather than the association across the entire genome (Endelman and Jannink, 2012). This is because causal loci are more likely to be inherited. Recent research has shown that even for complicated characteristics with very high marker density, the connection at the whole genome level can only poorly approximate the causative locus relationship for individuals that are very distantly related (Hill and Weir, 2011; de Los Campos et al., 2013). This is why the prediction accuracy of the gBLUP model is lower than that of other methods. Because of the low prediction accuracy of the gBLUP model, we will not discuss this model afterwards.

Genomic prediction accuracy based on location

For the GS models used in the P149 population, we established four Bayesian models, among which the TZ-LF and TZ-TZ-LF models predicted very low accuracy. However, the prediction accuracy of the TZ single-location was all above 0.67. This is probably because Tongzhou had only two years of data for six harvests. Because of the disparity in the volume of the data, a GS model was developed that failed to capture the majority of biomass yield genes across ages or growth seasons, and as a result, it had an inaccurate forecast. We then used two years of data from Tongzhou and Langfang (2014 and 2015). When the GS model was rebuilt, the accuracy of its predictions

increased from 0.11–0.33 to 0.42–0.57, and the accuracy of both the local and cross-location predictions increased. This showed us that the stubble numbers of the modeling and validation populations should not be too different when conducting GS breeding and that the quality of biomass data can significantly impact the accuracy of the GS model prediction.

In addition, population size also affects the prediction accuracy of GS models. Generally, the larger the breeding population, the higher the accuracy of GS prediction and the higher the cost. We constructed two populations including different individuals to investigate the effect of the population size on the accuracy of the GS prediction. The seven models showed high prediction accuracy for the P392 population (0.42–0.67). In conclusion, when developing the alfalfa GS model, we should establish as large a breeding population as we can afford.

Implications for alfalfa breeding

In this study, the prediction accuracy of large breeding populations was higher than that of small breeding populations. The prediction accuracy of total biomass was high, and high accuracy was also observed in cross-location prediction. Between different models the prediction accuracy varied greatly. It is also surprising that when the number of SNP makers increases, the prediction accuracy does not increase significantly, which will be our next research direction. This study shows that GS has a high potential to accelerate alfalfa breeding.

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

TZ and JK designed and supervised the project. XH performed bioinformatic analysis and drafted the manuscript. FZ and FH performed experiments and participated in the bioinformatic analysis. YS collected the samples, maintained the growth of the experiment materials, and performed DNA extraction. L-XY revised the manuscript and participated in the interpretation of the data. TZ and JK contributed equally as corresponding authors. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the Breeding and Industrialization Demonstration of New High-quality Alfalfa Varieties (No. 2022JBGS0020), National Natural Science Foundation of China (No. 31772656), and China Forage and Grass Research System (CARS-34).

Conflict of interest

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

Publisher's note

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

References

Annicchiarico, P. (2015). Alfalfa forage yield and leaf/stem ratio: Narrow-sense heritability, genetic correlation, and parent selection procedures. *Euphytica* 205, 409–420. doi: 10.1007/s10681-015-1399-y

Annicchiarico, P., Scotti, C., Carelli, M., and Pecetti, L. (2010). Questions and avenues for lucerne improvement. *Czech J. Genet. Plant Breed.* 46, 1–13. doi: 10.17221/90/2009-CJGPB

Baird, N. A., Etter, P. D., Atwood, T. S., Currey, M. C., Shiver, A. L., Lewis, Z.A., et al. (2008). Rapid SNP discovery and genetic mapping using sequenced RAD markers. *PloS One* 3, e3376. doi: 10.1371/journal.pone.0003376

Bates, D., Maechler, M., Bolker, B. M., and Walker, S. C. (2015). Fitting linear mixed-effects models using lme4. *J. Stat. Software* 67, 1–48. doi: 10.18637/jss.v067.i01

Bingham, E. T., Groose, R. W., Woodfield, D. R., and Kidwell, K. K. (1994). Complementary gene interactions in alfalfa are greater in autotetraploids than diploids. *Crop Sci.* 34, 823–829. doi: 10.2135/cropsci1994.0011183X003400040001x

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

Chen, S., Huilong, D., Zhuo, C., Hongwei, L., Fugui, Z., and Chen, H.. (2020). The chromosome-level genome sequence of the autotetraploid alfalfa and resequencing of core germplasms provide genomic resources for alfalfa research. *Mol. Plant* 13, 1250–1261. doi: 10.1016/j.molp.2020.07.003

Chen, H., Zeng, Y., Yang, Y., Huang, L., Tang, B., Zhang, H., et al. (2020). Allele-aware chromosome-level genome assembly and efficient transgene-free genome editing for the autotetraploid cultivated alfalfa. *Nat. Commun.* 11, 2494. doi: 10.1038/s41467-020-16338-x

Corinna, C., and Vladimir, V. (1995). Support-vector networks. *Mach. Learn.* 20, 273–297. doi: 10.1007/BF00994018

Cros, D., Denis, M., Sanchez, L., Cochard, B., Flori, A., Durand-Gasselin, T., et al. (2015). Genomic selection prediction accuracy in a perennial crop: Case study

of oil palm (*Elaeis guineensis* jacq.). *Theor. Appl. Genet.* 128, 397–410. doi: 10.1007/s00122-014-2439-z

Danecek, P., Auton, A., Abecasis, G., Albers, C. A., Banks, E., DePristo, C.A., et al. (2011). The variant call format and VCFtools. *Bioinformatics* 27, 2156–2158. doi: 10.1093/bioinformatics/btr330

Davey, J. W., Hohenlohe, P. A., Etter, P. D., Boone, J. Q., Catchen, J. M., Blaxter, M.L., et al. (2011). Genome-wide genetic marker discovery and genotyping using next-generation sequencing. *Nat. Rev. Genet.* 12, 499–510. doi: 10.1038/nrg3012

de Los Campos, G., Vazquez, A. I., Fernando, R., Klimentidis, Y. C., and Sorensen, D. (2013). Prediction of complex human traits using the genomic best linear unbiased predictor. *PloS Genet.* 9, e1003608. doi: 10.1371/journal.pgen.1003608

Endelman, J. B. (2011). Ridge regression and other kernels for genomic selection with r package rrBLUP. *Plant Genome* 4, 250–255. doi: 10.3835/plantgenome2011.08.0024

Endelman, J. B., and Jannink, J. (2012). Shrinkage estimation of the realized relationship matrix. *G3-Genes Genomes Genet.* 2, 1405–1413. doi: 10.1534/g3.112.004259

Geman, S., and Geman, D. (1984). Stochastic relaxation, gibbs distributions, and the bayesian restoration of images. *IEEE Trans. On Pattern Anal. Mach. Intell.* 6, 721–741. doi: 10.1109/tpami.1984.4767596

George, C., and Edward, I. G. (2012). Explaining the gibbs sampler. *Am. Statistician* 46, 167–174. doi: 10.2307/2685208

Glaubitz, J. C., Casstevens, T. M., Lu, F., Harriman, J., Elshire, R. J., Sun, Q., et al. (2014). TASSEL-GBS: A high capacity genotyping by sequencing analysis pipeline. *PloS One* 9, e90346. doi: 10.1371/journal.pone.0090346

Habier, D., Fernando, R. L., Kizilkaya, K., and Garrick, D. J. (2011). Extension of the bayesian alphabet for genomic selection. *BMC Bioinf.* 12, 186. doi: 10.1186/1471-2105-12-186

Heffner, E. L., Sorrells, M. E., and Jannink, J. (2009). Genomic selection for crop improvement. *Crop Sci.* 49, 1–12. doi: 10.2135/cropsci2008.08.0512

Hill, W. G., and Weir, B. S. (2011). Variation in actual relationship as a consequence of mendelian sampling and linkage. *Genet. Res.* 93, 47–64. doi: 10.1017/S0016672310000480

Juliana, P., He, X., Poland, J., Roy, K. K., Malaker, P. K., Mishra, V.K., et al. (2022). Genomic selection for spot blotch in bread wheat breeding panels, full-sibs and half-sibs and index-based selection for spot blotch, heading and plant height. *Theor. Appl. Genet.* 135, 1965–1983. doi: 10.1007/s00122-022-04087-y

Koboldt, D. C., Zhang, Q., Larson, D. E., Shen, D., McLellan, M. D., Lin, L., et al. (2012). VarScan 2: Somatic mutation and copy number alteration discovery in cancer by exome sequencing. *Genome Res.* 22, 568–576. doi: 10.1101/gr.129684.111

Lamb, J., Sheaffer, C. C., Rhodes, L. H., Sulc, R. M., Undersander, D. J., and Brummer, E.C.. (2006). Five decades of alfalfa cultivar improvement: Impact on forage yield, persistence, and nutritive value. *Crop Sci.* 46, 902–909. doi: 10.2135/cropsci2005.08-0236

Lande, R., and Thompson, R. (1990). Efficiency of marker-assisted selection in the improvement of quantitative traits. *Genetics* 124, 743–756. doi: 10.1093/genetics/124.3.743

Leo, B. (2001). Random forests. *Mach. Learn.* 45, 5–32. doi: 10.1023/A:1010933404324

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

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence Alignment/Map format and SAMtools. *Bioinformatics* 25, 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, X., Han, Y., Wei, Y., Acharya, A., Farmer, A. D., Ho, J., et al. (2014). Development of an alfalfa SNP array and its use to evaluate patterns of population structure and linkage disequilibrium. *PloS One* 9, e84329. doi: 10.1371/journal.pone.0084329

Li, X., Wei, Y., Acharya, A., Hansen, J. L., Crawford, J. L., Viands, D.R., et al. (2015). Genomic prediction of biomass yield in two selection cycles of a tetraploid alfalfa breeding population. *Plant Genome* 8, 2. doi: 10.3835/plantgenome2014.12.0090

Lorenz, A. J., Chao, S., Asoro, F. G., Heffner, E. L., Hayashi, T., Iwata, H., et al. (2011). "Genomic selection in plant breeding: Knowledge and prospects" in: *Advances in Agronomy*. 110, 77–123. doi: 10.1016/B978-0-12-385531-2.00002-5

Lu, F., Lipka, A. E., Glaubitz, J., Elshire, R., Cherney, J. H., Casler, M.D., et al. (2013). Switchgrass genomic diversity, ploidy, and evolution: Novel insights from a network-based SNP discovery protocol. *PloS Genet.* 9, e1003215. doi: 10.1371/journal.pgen.1003215

Mendes, M. P., and de Souza, (2016). Genomewide prediction of tropical maize single-crosses. *Euphytica* 209, 651–663. doi: 10.1007/s10681-016-1642-1

Meuwissen, T. H., Hayes, B. J., and Goddard, M. E. (2001). Prediction of total genetic value using genome-wide dense marker maps. *Genetics* 157, 1819–1829. doi: 10.1093/genetics/157.4.1819

Paolo, A., Brent, B., E., C. B., Bernadette, J., and Athole, H. M. (2015). Achievements and challenges in improving temperate perennial forage legumes. *Crit. Rev. Plant Sci.* 34, 327–380. doi: 10.1080/07352689.2014.898462

Park, T., and Casella, G. (2008). The bayesian lasso. *J. Am. Stat. Assoc.* 103, 681–686. doi: 10.1198/016214508000000337

Perez, P., and de Los Campos, G. (2014). Genome-wide regression and prediction with the BGLR statistical package. *Genetics* 198, 463–483. doi: 10.1534/genetics.114.164442

Poland, J., Endelman, J., Dawson, J., Rutkoski, J., Wu, S., Manes, Y., et al. (2012). Genomic selection in wheat breeding using genotyping-by-sequencing. *Plant Genome* 5, 103–113. doi: 10.3835/plantgenome2012.06.0006

Russell, W., Walter, T. F., and Olga, C. B. (1997). Recovering information in augmented designs, using SAS PROC GLM and PROC mixed. *Agron. J.* 89, 856–859. doi: 10.2134/agronj1997.00021962008900060002x

Schmidt, M., Kollers, S., Maasberg-Prelle, A., Grosser, J., Schinkel, B., Tomerius, A., et al. (2016). Prediction of malting quality traits in barley based on genome-wide marker data to assess the potential of genomic selection. *Theor. Appl. Genet.* 129, 203–213. doi: 10.1007/s00122-015-2639-1

Stritzler, M., Pascuan, C., Bottero, E., Gómez, C., Frare, R., Puebla, A., et al. (2022). Rapid and cloning-free screening of edited alfalfa *via* next-generation sequencing. *Plant Cell Tissue Organ Culture* 151, 1–6. doi: 10.1007/s11240-022-02358-6

Van der Auwera, G. A., Carneiro, M. O., Hartl, C., Poplin, R., Del, A. G., Levy-Moonshine, A., et al. (2013). From FastQ data to high confidence variant calls: The genome analysis toolkit best practices pipeline. *Curr. Protoc. Bioinf.* 43, 11.10.1-11.10.33. doi: 10.1002/0471250953.bi1110s43

VanRaden, P. M. (2008). Efficient methods to compute genomic predictions. *J. Dairy Sci.* 91, 4414–4423. doi: 10.3168/jds.2007-0980

Wallace, J. G., and Mitchell, S. E. (2017). Genotyping-by-Sequencing. *Curr. Protoc. Plant Biol.* 2, 64–77. doi: 10.1002/cppb.20042

Zhang, F., Kang, J., Long, R., Yu, L., Sun, Y., Wang, Z., et al. (2020). Construction of high-density genetic linkage map and mapping quantitative trait loci (QTL) for flowering time in autotetraploid alfalfa (*Medicago sativa* l.) using genotyping by sequencing. *Plant Genome* 13, e20045. doi: 10.1002/tpg2.20045

Zhang, F., Kang, J., Long, R., Yu, L., Wang, Z., Zhao, Z., et al. (2019). High-density linkage map construction and mapping QTL for yield and yield components in autotetraploid alfalfa using RAD-seq. *BMC Plant Biol.* 19, 165. doi: 10.1186/s12870-019-1770-6

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 December 2022
DOI 10.3389/fpls.2022.1085497

OPEN ACCESS

EDITED BY
Shangang Jia,
China Agricultural University, China

REVIEWED BY
Yuefei Xu,
Northwest A&F University, China
Jun Li,
Inner Mongolia University, China
Xunzhong Zhang,
Virginia Tech, United States

*CORRESPONDENCE
Liebao Han
hanliebao@bjfu.edu.cn
Yuehui Chao
chaoyuehui@bjfu.edu.cn

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 31 October 2022
ACCEPTED 25 November 2022
PUBLISHED 08 December 2022

CITATION
Li S, Xie H, Zhou L, Dong D, Liu Y, Jia C, Han L, Chao Y and Chen Y (2022) Overexpression of *MsSAG113* gene promotes leaf senescence in alfalfa *via* participating in the hormone regulatory network. *Front. Plant Sci.* 13:1085497. doi: 10.3389/fpls.2022.1085497

Overexpression of *MsSAG113* gene promotes leaf senescence in alfalfa *via* participating in the hormone regulatory network

Shuwen Li[1], Hong Xie[1], Lingfang Zhou[1], Di Dong[1], Yaling Liu[2], Chenyan Jia[2], Liebao Han[1]*, Yuehui Chao[1]* and Yinglong Chen[3]

[1]School of Grassland Science, Beijing Forestry University, Beijing, China, [2]Inner Mongolia M-Grass Ecology And Environment (Group) Co., Ltd, Hohhot, China, [3]The University of Western Australia (UWA) Institute of Agriculture, and University of Western Australia School of Agriculture and Environment, The University of Western Australia, Perth, WA, Australia

Introduction: Alfalfa (Medicago sativa) is a kind of high quality leguminous forage species, which was widely cultivated in the world. Leaf senescence is an essential process in plant development and life cycle. Here, we reported the isolation and functional analysis of an alfalfa SENESCENCE-ASSOCIATED GENE113 (MsSAG113), which belongs to the PP2C family and mainly plays a role in promoting plant senescence.

Methods: In the study, Agrobacterium-mediated, gene expression analysis, next generation sequencing, DNA pull-down, yeast single hybridization and transient expression were used to identify the function of MsSAG113 gene.

Results: The MsSAG113 gene was isolated from alfalfa, and the transgenic plants were obtained by Agrobacterium-mediated method. Compared with the wildtype, transgenic plants showed premature senescence in leaves, especially when cultivated under dark conditions. Meanwhile, application of exogenous hormones ABA, SA, MeJA, obviously acclerated leaf senescence of transgenic plants. Furthermore, the detached leaves from transgenic plants turned yellow earlier with lower chlorophyll content. Transcriptome analysis identified a total of 1,392 differentially expressed genes (DEGs), involving 13 transcription factor families. Of which, 234 genes were related to phytohormone synthesis, metabolism and transduction. Pull-down assay and yeast one-hybrid assay confirmed that alfalfa zinc finger CCCH domain-containing protein 39 (MsC3H-39) could directly bind the upstream of MsSAG113 gene. In conclusion, the MsSAG113 gene plays a crucial role in promoting leaf senescence in alfalfa via participating in the hormone regulatory network.

Discussion: This provides an essential basis for further analysis on the regulatory network involving senescence-associated genes in alfalfa.

KEYWORDS

alfalfa, *MsC3H-39*, *MsSAG113*, RNA-seq, senescence

Introduction

Leaf senescence is the late stage in plant development, during which leaves change from pale yellow to yellow. During senescence, nutrients from old leaves are diverted to other parts of the plant to initiate new tissue, such as new leaves or developing organs and seeds (Jing et al., 2002). As a result, plant leaf cells undergo a series of changes in morphology, structure, and metabolism leading to leaf abscission, a process known as senescence syndrome (Bleecker and Patterson, 1997). Although senescence is an age-dependent process, it is often triggered by a range of biological and abiotic stresses (Gan and Amasino, 1997). External factors, such as environmental stress and long-term darkness may also induce leaf senescence, which is considered as a self-protection mechanism to cope with the stress to prevent crop yield or quality reduction (Chao et al., 2018).

Recently, the study of plant leaf senescence has been favored by a large number of researchers, and the related molecular mechanisms have also made good progress. In senescent leaves, the senescence-associated genes (*SAG*) are expressed, which are considered as a marker of cellular senescence (Wistrom and Villeponteau, 1992). Senescence associated genes have been cloned and identified in many plants. *OsSAG12-2* codes for a functional protease that negatively regulates stress-induced cell death in rice (Singh et al., 2016). A gene encoding an acyl hydrolase is involved in leaf senescence in *Arabidopsis thaliana*, chemically induced overexpression of *SAG101* causing premature senescence in both attached and isolated leaves of transgenic *Arabidopsis* plants (He and Gan, 2002). The senescence associated gene *AhSAG* under aluminum stress promoted senescence in transgenic plants and further induced or promoted the occurrence of programmed cell death (PCD) (Zhan et al., 2013). Growing isolated leaves in the dark can induce leaf yellowing and reduce chlorophyll content. Therefore, isolated leaves are used as models to study leaf senescence (Liebsch and Keech, 2016).

In addition, plant hormones play an important role in plant growth, development and metabolism. The senescence process is also regulated by various plant hormones such as salicylic acid (SA), abscisic acid (ABA), ethylene, jasmonic acid (JA), cytokinins, and auxin (He et al., 2002). Studies have shown that the ABA receptor gene *PLY9* has a role in aging old leaves and inhibiting the growth of young tissues (Zhao et al., 2016). There are many reports on the complex mechanisms of gene interactions regulating leaf senescence. *RhHB1* is a gene strongly over-expressed in senescent petals, which can be induced expression by ABA or ethylene. It was shown that ABA or ethylene treatment could induce increased *RhHB1* expression, while gibberellin GA_3 treatment delayed this process (Lü et al., 2014). Leaf senescence is usually accompanied by a decrease in cytokinin content, but the increase of cytokinin content can alleviate leaf senescence (Balibrea Lara et al., 2004). In 2015, Surya et al. found that the expression of *converted structural genes* (*IPT*) encoding the cytokinin biosynthetic enzyme isoprenyl transferase delayed leaf senescence of the transgenic rapeseed plants under controlled environments (Kant et al., 2015).

Protein phosphatases of the PP2C type are monomeric enzymes present in prokaryotes and eukaryotes. PP2C is the largest family of protein phosphatases in plant species. Plant PP2C family proteins play important roles in transcriptional regulatory pathways such as stress response, ABA synthesis, and developmental signaling (Schweighofer et al., 2004). The *AHG1* gene of the PP2C protein family has special functions during seed development and germination (Nishimura et al., 2007). Two phosphatases 2C genes, *ABI1* and *ABI2*, negatively feedback the ABA signaling pathway (Merlot et al., 2001). *SAG113* belongs to the PP2C family and was founded to localize to the cis-Golgi as a negative regulator of ABA signaling, specifically inhibiting stomatal closure, further leading to senescence and eventual desiccation (Zhang et al., 2012).

Alfalfa (*Medicago sativa*)is an important leguminous herbage speceis, which is rich in protein, trace elements and vitamins. It has been widely used in the world because of its important ecological and economic values (Hrbáčková et al., 2020). Therefore, it is important to study the senescence of alfalfa leaves to improve its application value.

Although senescence associated genes have been widely reported in many different plants, only a few studies involved in alfalfa. Therefore, in this study the role of a senescence associated gene, *MsSAG113* was explored in the senescence regulatory network, aiming to lay a foundation for improving alfalfa quality by genetic engineering methods.

Materials and methods

Plant materials and growth conditions

Seeds of alfalfa (*Medicago sativa*) (cv. Zhongmu No. 1) were vernalized at 4°C for 4 d and cultivated under 16 h light (25°C) and 8 h dark (23°C) until root development. The seedlings were transferred to pots filled with mixed peat, vermiculite and perlite, and grown under 16 h light (25°C) and 8 h dark (23°C). It is difficult to control the developmental senescence, thus many senescence-related studies have been performed on isolated leaves in dark cultures or attached leaves in dark cultures alone (Pruzinská et al., 2005). For phenotype analysis of detached leaves, alfalfa leaves of 3-month-old wild type (WT) and *MsSAG113* transgenic plants were collected and immersed in 2-(N-Morpholino) ethanesulfonic acid hydrate (MES) buffer (3mM, pH 5.8) in the dark or with different plant hormone treatments.

Subcellular localization

The SAG113-GFP-F/R primers (Supplementary Table S1) were used for the construction of subcellular localization vector *35S::MsSAG113:GFP* and the *35S::GFP* vector was used as a control. The recombinant plasmid was transferred into EHA105 *Agrobacterium tummefaciens* and injected into tobacco plants and cultured for subcellular localization. The position of green fluorescent protein (GFP) signal in the cells was observed with a confocal microscope (Leica TCS SP8, Leica Microsystems).

Exogenous hormone treatments and gene expression

Two-month-old alfalfa plants were selected and sprayed with hormones involving 50 μmol/L ABA, 0.5 mmol/L SA, 10 μmol/L MeJA, 10 μmol/L 6BAP, respectively. The leaves were collected at 0 min, 10 min, 20 min, 30 min, 1 h, 3 h, 6 h, 9 h, and 12 h for expression analysis. Meanwhile, leaves at different developmental stage (1 month-old, 2 month-old and 3 month-old) were harvested for *MsC3H-39* gene expression analysis. Total RNA from the leaf samples of different treatments was extracted by a Plant RNA Kit (Omega Bio-tek, Inc., USA), and cDNA was synthesize using PrimeScript™ RT reagent Kit (TaKaRa, Japan) for real-time quantitative RT-PCR (qRT-PCR) with primer pair of SAG113-RT-F/R and C3H-39-RT-F/R (Supplementary Table S1). Expression of *MsSAG113* and *MsC3H-39* was calculated by the $2^{-\Delta\Delta CT}$ method with three biological replicates (Schmittgen and Livak, 2008).

Transformation of plants

The sequence of *MsSAG113* (GenBank Access No.: KT592510) was obtained from NCBI (National Center for Biotechnology Information) and analyzed in MEGA version 6.0. Two primers for SAG113-F/R (Supplementary Table S1) were used to clone the coding domain sequence (CDS) from alfalfa. Primer 3302Y-SAG113-F/R (Supplementary Table S1) was used to obtain plant expression vector *35S::SAG113*. The expression vector was then transformed into alfalfa mediated by *Agrobacterium transformation* (Hoekema et al., 1983). Three different lines (S1, S6 and S7) with high expression levels were selected for further experiments.

Measurement of endogenous hormones and chlorophyll contents

Two-month-old transgenic and WT leaves with similar state were selected and endogenous hormones were detected by enzyme-linked immunoassay (ELISA) as described previously (Engvall and Perlmann, 1971). Chloroform was used to extract the chlorophyll content as previously reported (Matile et al., 1988). The absorbance of the extract was measured by spectrophotometer, and the content of each pigment in the extract was calculated. All samples were repeated three times.

Next generation sequencing and analysis

Mature leaves of three independent lines were selected for transcriptiome sequencing and DEGs were identified by DESeq R package (version 1.18.0) (Robinson et al., 2009) with parameters: adjusted p-value <0.05 and $|\log_2 FC| \geq 1$. Gene Ontology (GO) analysis of DEGs between samples was performed using GOseq R package (version: Release 2.12) (Young et al., 2010). KOBAS software (version: 2.0) was used for functional analysis of DEGs in the KEGG pathway (Mao et al., 2005).

DNA pull-down

The potential *cis* elements of upstream of *MsSAG113* were investigated *via* the PlantCARE website (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/). Based on the predicted result, one pair of primers SAG113-bition-F/R labeled with biotin (Supplementary Table S1) was synthesized and used for DNA pull-down. The alfalfa nuclear proteins were extracted by a nuclear protein extraction kit (Solarbio, China) and the extraction was incubated with the target sequence of *MsSAG113* containing biotin. Compound was pulled down with streptavidin magnetic beads (Beyotime, China) according to the manual. The compound products were washed five times, separated by SDS-PAGE, and identified by mass spectrometry.

Yeast one-hybrid

Two pairs of primers, pAbAi-SAG113pro-F/R and pGADT7-C3H-39-F/R (Supplementary Table S1) were used for construction of vectors for yeast one-hybrid. Recombinant plasmids pAbAi-$MsSAG113_{pro}$ and pGADT7-MsC3H-39 were co-transformed into yeast Y1H strain. Different concentrations of aureobasidin (AbA) (TaKaRa, Japan) were used to screen a suitable concentration for yeast one-hybrid, and transformed yeast cells were grown on SD/-Leu/AbA medium for further investigation.

Transient expression of *MsC3H-39*

The *MsC3H* CDS was cloned with MsC3H-39-F/R primers (Supplementary Table S1) and ligased into 3302Y vector,

resulting in plant expression vector. The recombinant plasmids were transiently transferred into 2 month-old alfalfa leaves with empty vector 3302Y as controls and the leaf phenotype was observed. After 48h of dark treatment, total RNA was isolated from each leaf sample, and the expression levels of *MsSAG113* was analyzed with qRT-PCR as described above.

Results

Location and expression pattern of *MsSAG113*

Based on the Cell-PLoc package and UniProt analysis, *MsSAG113* was predicted to be a nuclear-localized protein. The expression of *MsSAG113*-GFP fusion protein in tobacco leaves was captured showing strong GFP signal in the nucleus and cytoplasm (Figure 1), which indicated subcellular localization of *MsSAG113* protein was in nucleus and cytoplasm.

All hormones used in this study had significant effects on the expression of *MsSAG113*,which was increased rapidly at first and then decreased with the passage of time in ABA,SA and MeJA treated plants. The expression of *MsSAG113* reached its peak at 20 min after spraying with ABA and MeJA hormones (Figures 2A–C). It indicated that the expression of *MsSAG113* gene was regulated by ABA, MeJA or SA in the short term. The expression of *MsSAG113* in plants treated with 6BAP showed a down-regulated trend within 6 h and then recovered rapidly at 9 h (Figure 2D). These results indicated that *MsSAG113* was involved in short-term regulatory of plant hormone network.

The role of *MsSAG113* in premature leaf senescence

Through the analysis of transgenic plants S1, S6 and S7, we found that under normal condition, transgenic plants exhibited early senescence phenotype and leaves turned yellow much earlier than those of WT plants. All three transgenic samples showed significantly higher expression levels of *MsSAG113* than those in WT plants (Figure 3B). Meanwhile, the chlorophyll contents in transgenic plants were drastically lower than those in WT (Figure 3C). These results indicated that *MsSAG113* promoted early senescence and chlorophyll degradation in alfalfa.

Dark accelerates senescence of detached leaves of transgenic alfalfa

The detached leaves from transgenic plants turned yellow much earlier than WT under continuous dark for 7 days (Figure 4A), and chlorophyll contents were consistent with the senescence phenotype. The chlorophyll content of transgenic lines S1, S6 and S7 was 19.6%, 21.1% and 9.3% of that of wild-type lines, respectively (Figure 4B).

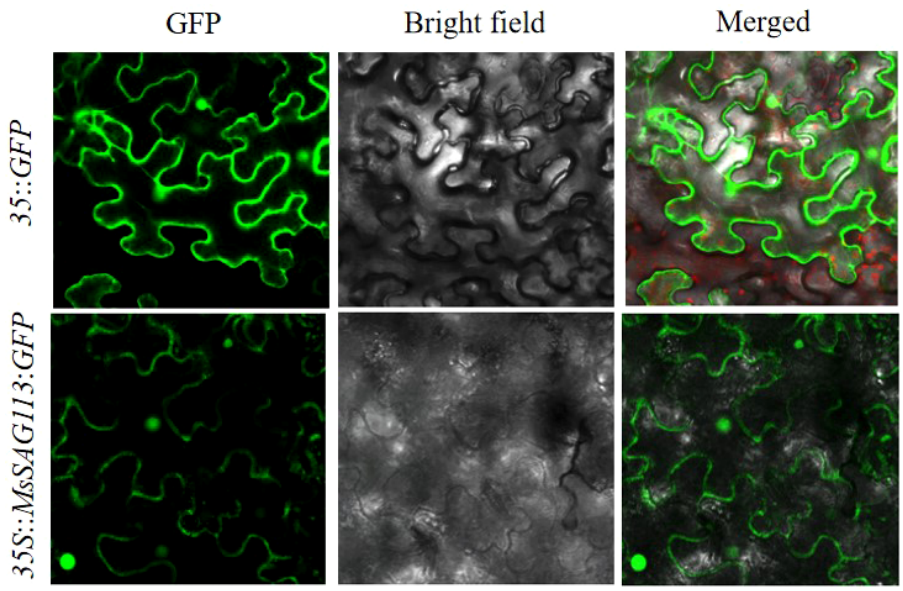


FIGURE 1
Subcellular localization of *MsSAG113* in the nucleus and cytoplasm detected by GFP signal. Scale bars =25 μm.

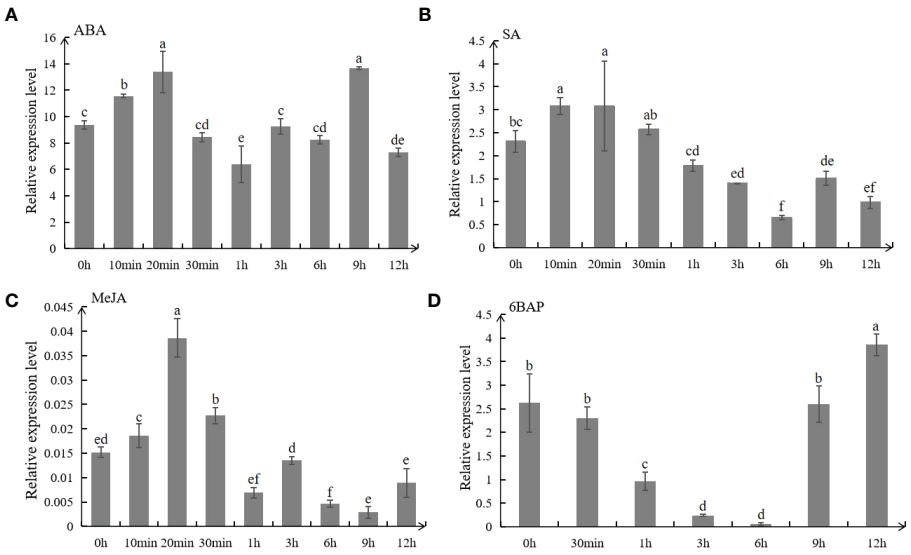


FIGURE 2
Expression analysis of *MsSAG113* in alfalfa by qRT-PCR. *MsSAG113* under different hormones treatments: 50 μM ABA **(A)**, 0.5 mM SA **(B)**, 10 μM MeJA **(C)**, and 10 μM 6BAP **(D)**. The values are means ± SD (n=3). In graphs, bars with different lowercase letters indicate significant difference ($p<0.05$).

Detection of endogenous hormones in overexpressed plants

The contents of plant hormones in WT and transgenic plants were detected. Results showed that MeJA and ZR contents of transgenic plants were lower than those of WT (Figures 5A, B). In addition, ABA contents of the three *MsSAG113* transgenic lines were all lower than those in WT. Compared with WT, ABA content in S1, S6 and S7 plants was decreased by 53.3%, 16.7% and 41.2%, respectively (Figure 5C).

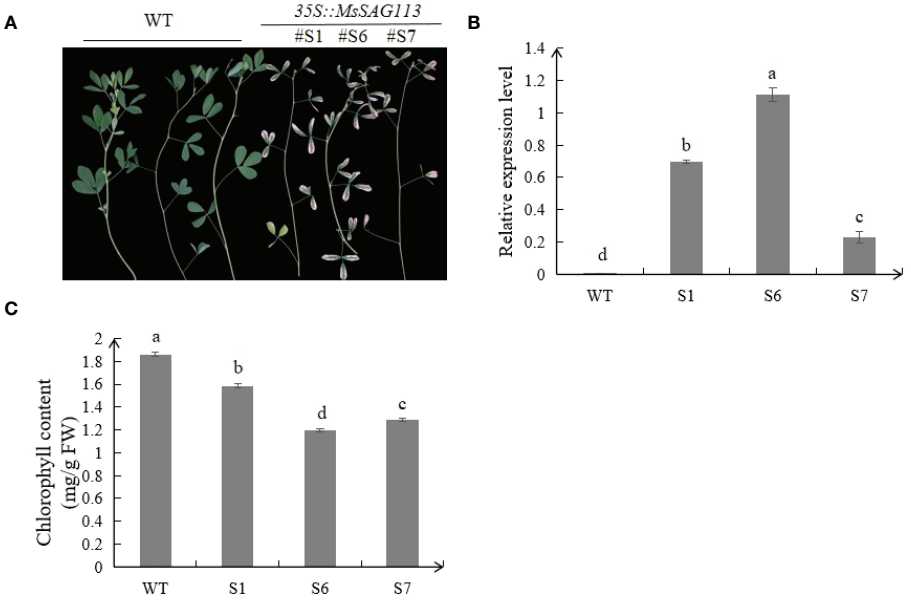


FIGURE 3
Phenotypes analysis of transgenic and WT alfalfa. Phenotype of three transgenic lines and WT alfalfa of 2 month-old **(A)**. Relative expression levels of *MsSAG113* in transgenic and WT plants **(B)**. Chlorophyll contents of three transgenic and WT plants **(C)**. The values are means ± SD (n = 3). In graphs, bars with different lowercase letters indicate significant difference at $p<0.05$.

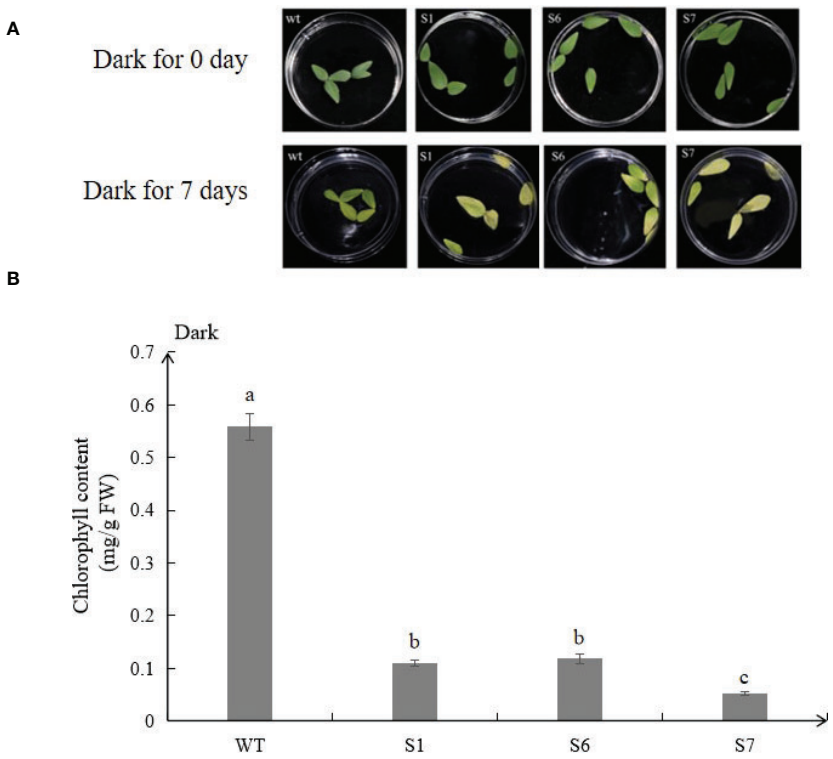


FIGURE 4
MsSAG113 is involved in dark-induced leaf senescence. Phenotypic analysis of 2 month-old wild type and transgenic isolated leaves after dark treatment **(A)**. Chlorophyll contents of leaves under dark treatment **(B)**. The values are means ± SD (n = 3). In graphs, bars with different lowercase letters indicate significant difference at $p<0.05$.

Analysis of detached leaves of transgenic *MsSAG113* treated with exogenous hormones

Detached leaves of S1, S6, and S7 treated with ABA, SA or MeJA showed obvious senescence on day 5, 8 or 12, respectively (Figure 6). Meanwhile, leaf senescence in S1, S6, and S7 leaves were also more pronounced than that in WT (Figures 6A–C). After 12 days of treatment with 10 μM 6BAP, the leaves of S1, S6 and S7 were significantly greener than those without *MsSAG113* gene in plant senescence, indicating that *MsSAG113* gene is involved in hormone network.

Under the treatments of ABA, SA and MeJA, the chlorophyll contents of transgenic plants, especially in S6 and S7 were lower than those from WT, which were consistent with the senescence phenotypes (Figures 7A–C). Under 6BAP treatments, the chlorophyll contents of WT leaves were higher than those in S6 and S7 (Figure 7D). However, samples of transgenic S1 showed no significant difference in chorophyll contents with WT under different hormone treatments except MeJA.

DEGs and TFs were identified by transcriptome analysis

Transcriptome analysis identified a total of 1,392 DEGs in the transgenic and WT plants, of which 898 were up-regulated and 494 were down-regulated (Figure 8A; Supplementary Table S2). There were 42 differentially expressed TFs from 13 families identified, of which MYB (13), bHLH (7), Homeobox (5), ZBTB (4) and zf-CCCH (2) families were among the top 5 (Figure 8B; Supplementary Table S3). Meanwhile, 234 DEGs related to plant hormones were identified mainly involving ABA, Auxin, JA and SA, among which 145 were up-regulated and 77 were down-regulated (Figure 8C; Supplementary Table S4). These results showed that over-expression of *MsSAG113* gene affected multiple senescence associated genes in alfalfa.

TF MsC3H-39 enhanced MsSAG113 expression

By DNA pull-down, 54 target proteins were collected, including alfalfa zinc finger protein MsC3H-39, which was

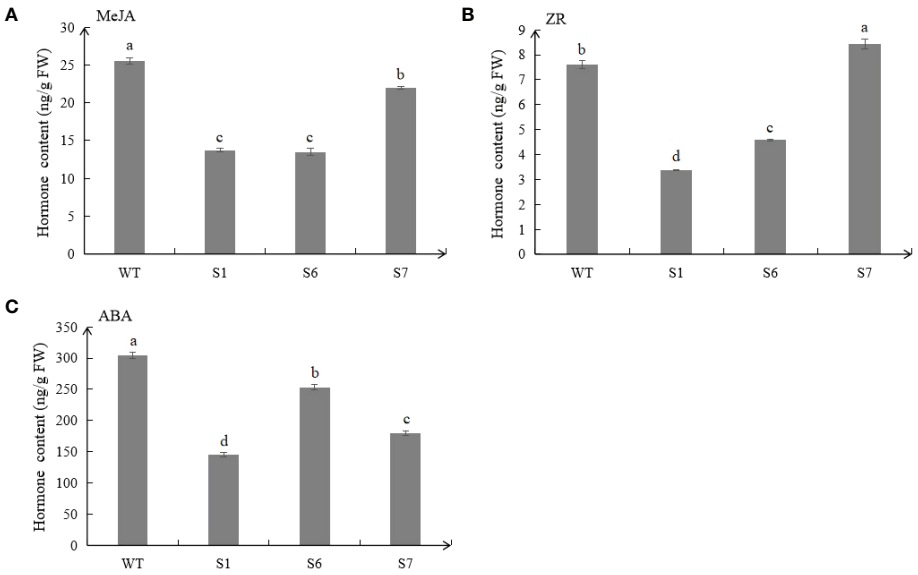


FIGURE 5
Endogenous hormone contents in transgenic and WT plants. Different hormones, including MeJA **(A)**, ZR **(B)** and ABA **(C)** contents in transgenic and WT plants.The values are means ± SD (n = 3). In graphs, bars with different lowercase letters indicate significant difference at $p<0.05$.

also found in transcriptome analysis. To verify whether the TF can recognize and change *MsSAG113* expression, we performed yeast one-hybrid and qRT-PCR analysis of alfalfa with transient expression of *MsC3H-39* gene. Yeast one-hybrid assay showed that only those yeast clones containing *MsSAG113* promoter and *MsC3H-39* gene could survive on selection medium (SD/-leu/AbA_{300}) (Figure 9A). These results suggest that MsC3H-39 transcription factor can bind directly to the promoter of *MsSAG113*.

The expression of *MsC3H-39* in alfalfa leaves was the highest in senescence leaves indicating the involvement of *MsC3H-39* in plant senescence. Meanwhile, transiently expression of *MsC3H-39* result in apparent senescence symptoms compared with empty vector 3302Y (Figures 9C, D). Transient expression of *MsC3H-39* gene and qRT-PCR analysis of *MsSAG113* showed that *MsC3H-39* significantly increased the expression of *MsSAG113*, suggesting that MsC3H-39 transcription factor positively regulated the expression of *MsSAG113* by binding the upstream of the gene (Figures 9B–D).

Discussion and conclusion

The regulatory network of senescence associated genes is constantly updated, and many *SAGs* were found being involved in gene expression regulation, signal transduction, macromolecular degradation and other senescence processes (He et al., 2001; Gepstein et al., 2003; Buchanan-Wollaston et al., 2005; Guo and Gan, 2012). *SAG113* belongs to the PP2C superfamily and plays an indispensable role in plant senescence. Senescence can limit crop yield and normal plant growth and development (Gan and Amasino, 1997). In this study, we explored the senescence regulation mechanism of *SAG113* gene in alfalfa, which could provide a reference for studying the growth and development mechanism of other legumes and improving crop yield.

Expression analysis showed that *MsSAG113* gene was affected by ABA, SA, MeJA and 6BAP hormone treatments. These hormones are considered to be the main internal factors that control the aging process of leaves (Balibrea Lara et al., 2004; Liebsch and Keech, 2016). It was reported that application of ABA-treated daffodils promoted senescence, and the increase in leaf ABA content was visually consistent with leaf senescence (Hunter et al., 2004). In *Arabidopsis thaliana* mutants deficient in SA signaling, the expression patterns of many genes were altered, and the mutants delayed yellowing and reduced necrosis (Morris et al., 2000). Exogenous application of JA led to premature senescence in leaves of wild-type *Arabidopsis thaliana*, but such effect absented in JA mutant plants (He et al., 2002). In our study, exogenous ABA, SA or MeJA accerlated early senescence of detached leaves with *MsSAG113* overexpression (Figure 7). These results indicated that ABA, SA

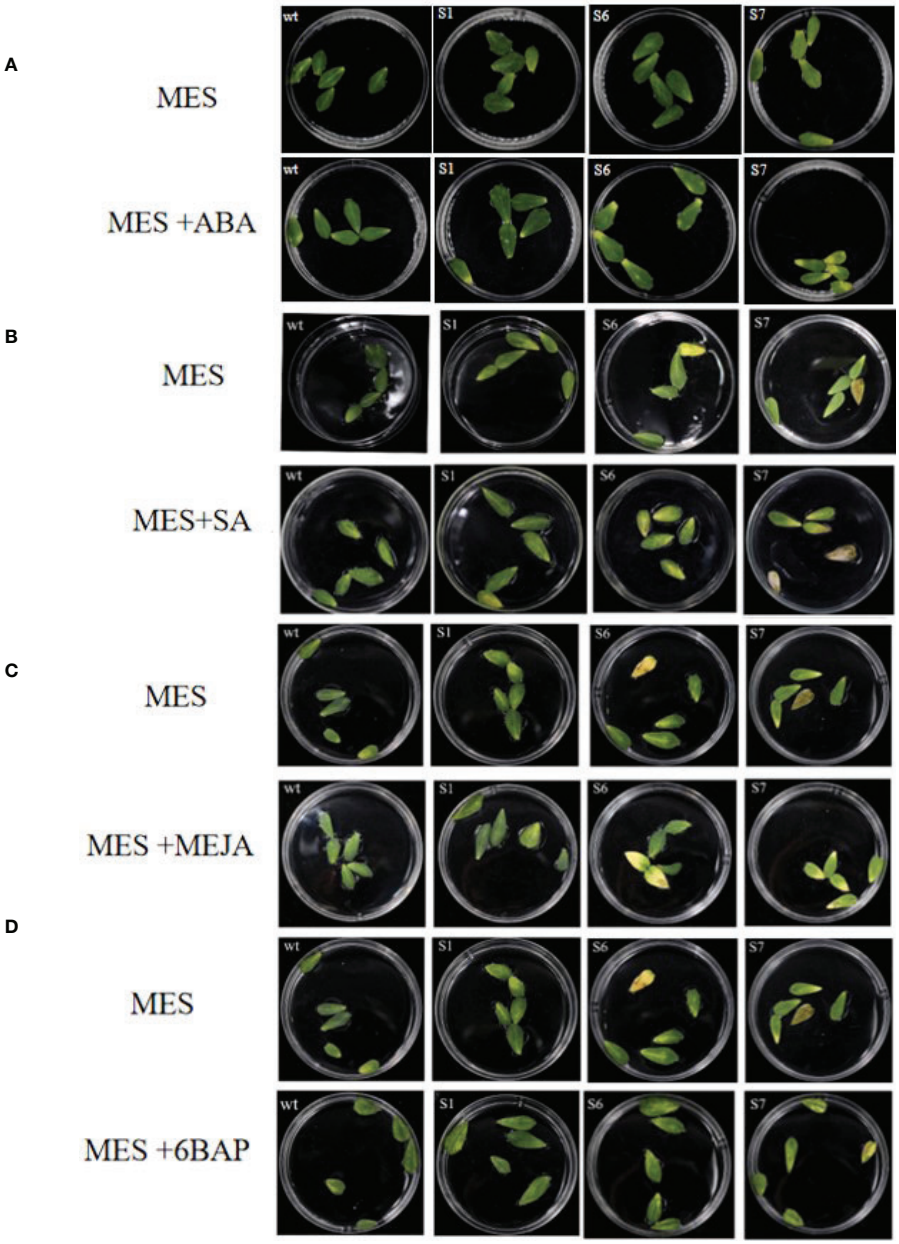


FIGURE 6
Detached leaves in MES with different hormones. Transgenic and WT detached leaves with or without ABA treatments for 5 days **(A)**. Transgenic and WT detached leaves with or without SA treatments for 8 days **(B)**. Transgenic and WT detached leaves with or without MeJA treatments for 12 days **(C)**. Transgenic and WT detached leaves with or without 6BAP treatments for 12 days **(D)**.

and MeJA might participate in plant senescence regulation of *MsSAG113* gene. Abscisic acid promoted aging, while 6BAP generally inhibited aging (Smart, 1994). Early physiological studies have shown that exogenous cytokinin treatment of most monocotyledonous and dicotyledonous plant species could delay leaf senescence (Richmond et al., 1957). Our study achieved similar results. Exogenous 6BAP hormone treatment of isolated leaves indicated that leaf senescence was delayed, and the delay in transgenic leaf senescence was more obvious (Figure 7). Therefore, *MsSAG113* gene, as a senescence regulator, may have a role in regulating leaf senescence by participating in a hormonal regulatory network.

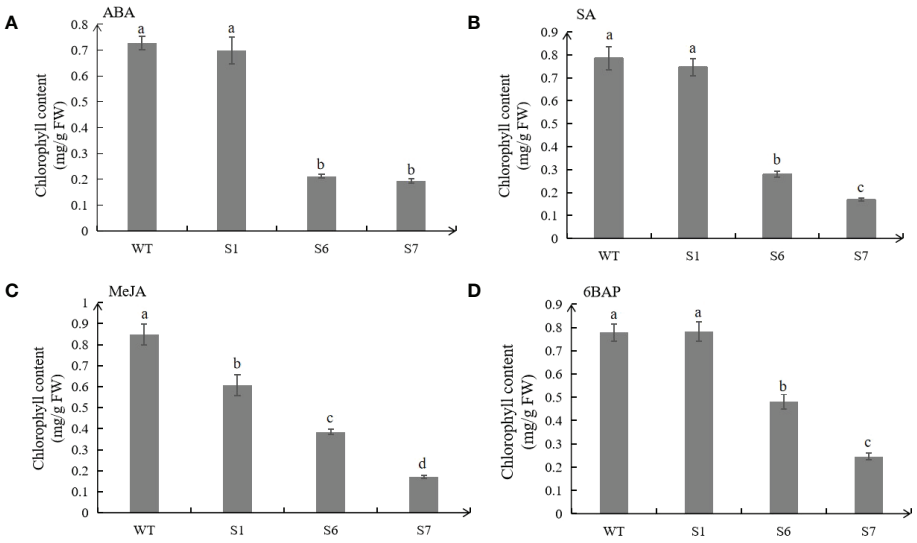


FIGURE 7
Chlorophyll contents in leaves of wild type (WT) and transgenic plants under hormone treatments. Chlorophyll contents of WT and transgenic leaves treated with 50 μM ABA for 5 days **(A)**. Chlorophyll contents of WT and transgenic leaves treated with 0.5 mM SA for 8 days **(B)**. Chlorophyll contents of WT and transgenic leaves treated with 10 μM MeJA for 10 days **(C)**. **C**hlorophyll contents of WT and transgenic leaves treated with 10 μmol/L 6BAP for 12 days **(D)**. The values are means ± SD (n = 3). In graphs, bars with different lowercase letters indicate significant difference at p<0.05.

PP2C enzymes play an important role in signal transduction. Six groups of PP2C family protein phosphatases are considered to be negative regulators of ABA (Schweighofer et al., 2004). Previous studies showed that SAG113 is a negative regulator of ABA signaling pathway, which can inhibit stomatal closing in leaves and specifically involved in the control of water loss during leaf senescence (Wang et al., 2011; Zhang and Gan, 2012). Overexpression of *FsPP2C1* in *Arabidopsis* reduced the degree of seed dormancy and was insensitive to ABA, so *FsPP2C1* had a negative regulatory effect on ABA (Gonzalez-Garcia et al., 2003). This precisely corroborates our results that ABA contents were reduced in plants overexpressing *MsSAG113*. Therefore, it is reasonable to speculate that the increased transcription level of *SAG113* may be involved in negatively regulating ABA signaling, leading to the inhibition of ABA synthesis.

Compared with the wild type, the leaves of *MsSAG113* transgenic alfalfa had obvious yellowing, and the chlorophyll content was significantly reduced (Figure 3). These results are consistent with previous research. During senescence, chlorophyll was degraded, which was visually manifested as yellowing of leaves (Curty and Engel, 1996). Overexpression of *MYBR1* could delay leaf senescence in *Arabidopsis*. Loss-of-function *MYBR1* plants exhibited faster chlorophyll loss and senescence (Jaradat et al., 2013). Darkness was often thought to be a cause of senescence, in which the synthesis of chlorophyll was blocked during the reductive phase (Weaver and Amasino, 2001; Eckhardt et al., 2004). Our results showed that after 7 days of dark treatment, the color of the isolated leaves in wild type was greener than that of the transgenic plants, with about 10 times difference in chlorophyll content (Figures 4, 5). These results suggest that *MsSAG113* gene be involved in the metabolic process of chlorophyll to regulate leaf senescence in alfalfa.

Our study showed that application of exogenous ABA, SA and MeJA significantly increased chlorophyll content in WT compared to transgenic plants. Previous studies on other plants observed the similar results (He et al., 2002; Lü et al., 2014; Zhao et al., 2016). However, we found that in isolated leaves treated with exogenous 6BAP, the chlorophyll content in S1 was increased compared with WT, while S6 and S7 were significantly reduced. Studies have shown that at the onset of senescence, under the activation of the *SAG12* promoter, feedback regulation of cytokinins occurs, which can be inhibited in delaying leaf senescence (Kant et al., 2015). The *SAG12* and *SAG13* promoters retarded the degradation of chlorophyll in isolated tomato senescent leaves under dark induction, while only *PSAG13::IPT* plants showed delayed chlorophyll degradation in isolated leaves and florets of broccoli (Chen et al., 2001; Swartzberg et al., 2006). Therefore, we inferred that under exogenous 6BAP treatment, the *SAG113*

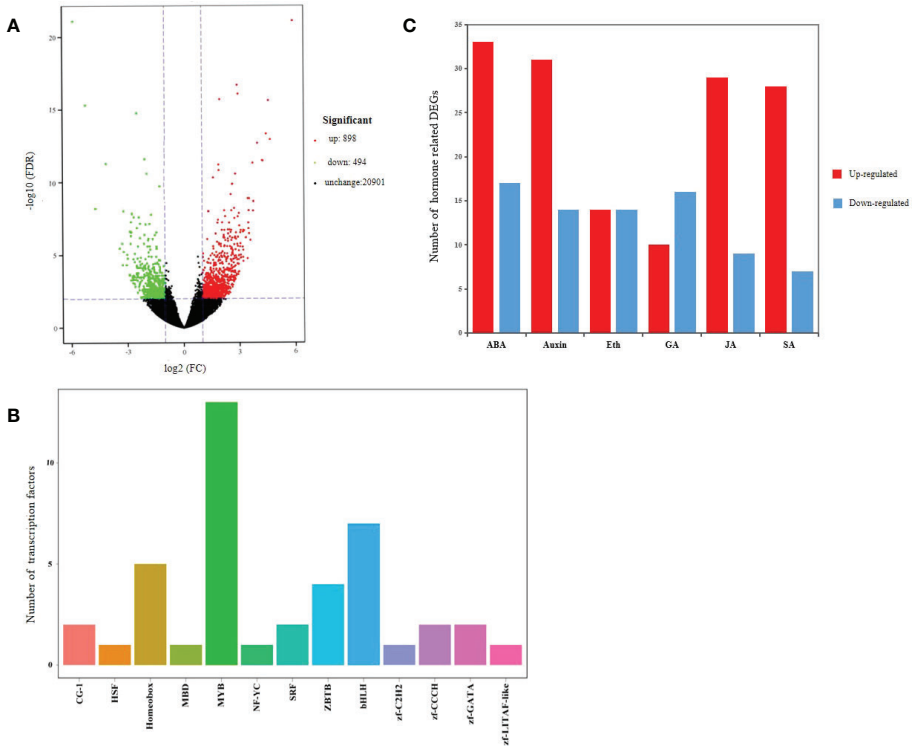


FIGURE 8
Transcriptome analysis of transgenic and WT alfalfa. Volcano plot of DEGs **(A)**. In graphs, each plot represents one gene with three colors, including red (up-regulated), green (down-regulated) and blank (unchanged). The X-axis represents the value of $\log_2$ (Fold Change) in the two samples, and the Y-axis indicates the negative value of $\log_{10}$ (FDR). Differentially expressed transcription factor genes identified with *MsSAG113* overexpression **(B)**. In graphs, the X-axis represents transcription factor family, and the Y-axis represents the number of transcription factor genes. Hormone related DEGs **(C)**. Eth indicates Ethylene. Red indicates up-regulated, and blue indicates down-regulated.

promoter was activated in S6 and S7 plants, resulting in feedback regulation of cytokinins, affecting chlorophyll synthesis, and thus reducing chlorophyll content of leaves. In addition, there was no significant difference in chlorophyll content between S1 plants and WT plants treated with exogenous hormones ABA, SA and 6BAP, which may be due to the different expression levels of *MsSAG113* in different transgenic lines and a phenotype in the transgenic plant is affected by gene expression levels and environmental factors.

RNA-Seq has become one of the important techniques for analyzing plant leaf senescence. Transcription factors and plant hormones are central to the senescence regulatory networks. In Arabidopsis, NAC, WRKY, AP2-EREBP and bHLH family transcription factors are regulated (Balazadeh et al., 2008). In this study, a total of 1,164 aging-related genes and 42 TFs from 13 families were identified, mainly including MYB, bHLH, Homeobox, ZBTB and zf-CCCH. The studies of plant senescence in *Arabidopsis thaliana* showed that many hormones were involved in the regulation of senescence in different ways (van der Graaff et al., 2006). By transcriptome analysis, we identified a total of 234 genes involved in hormone biosynthesis, metabolism, signal transduction and response. These results provided further evidence that *MsSAG113* gene plays important roles in regulating senescence in alfalfa and participating in hormone related regulatory network.

The identification of upstream regulators revealed the mechanism of *MsSAG113* gene in regulating senescence. This study screened the upstream regulatory protein of MsC3H-39 by DNA pull-down and verified it by yeast one-hybrid and transient expression analysis. MsC3H-39 protein belongs to CCCH-like of zinc finger proteins, which can be classified into C2H2, C2C2, C2HC, C2C2C2C2, C2HCC2C2 and CCCH types according to the number and order of zinc ion-bound Cys and His residues in the secondary structure of fingers (Sanchez-Garcia and Rabbitts, 1994; Klug and Schwabe, 1995; Mackay and Crossley, 1998). CCCH-like zinc lipoproteins have been shown to be involved in senescence regulation in previous

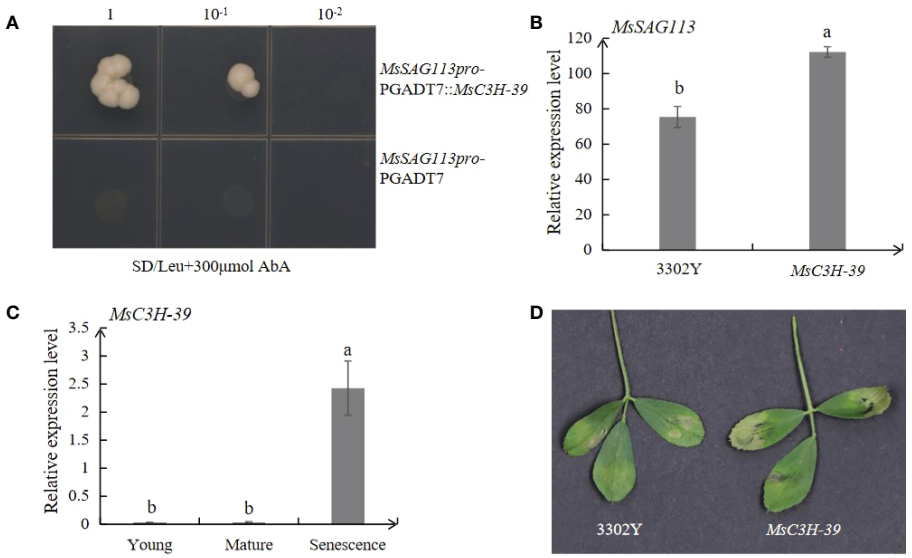


FIGURE 9
Interaction and expression analysis of *MsSAG113* with TF MsC3H-39. **(A)** Interaction analysis of upstream of *MsSAG113* and MsC3H-39 transcription factors. The number indicates dilution times of transformed yeast cells. **(B)** Relative expression levels of *MsSAG113* gene in alfalfa with transcient expression of empty vector or *MsC3H-39*. **(C)** Expression of *MsC3H-39* in young, mature and senescent leaves of alfalfa. **(D)** 3302Y empty vector and *MsC3H-39* transiently express the senescent phenotype in alfalfa leaves. The values are means ± SD (n = 3). In graphs, bars with different lowercase letters indicate significant difference at p<0.05.

reports. As a CCCH-type protein, OsDOS can delay leaf senescence, partially through the JA pathway (Jan et al., 2013). Two CCCH zinc finger proteins, AtC3H49 and AtC3H20 can delay leaf senescence and participate in ABA and JA responses in *Arabidopsis* (Lee et al., 2012). KHZ1 and KHZ2, two proteins of CCCH, positively regulated leaf senescence in *Arabidopsis thaliana* (Yan et al., 2017). The present study identified *MsSAG113* as a target gene of *MsC3H-39* by DNA pull-down assay and confirmed by yeast one-hybrid and qRT-PCR analysis. Meanwhile, we found that expression levels of *MsC3H-39* were the highest in senescence leaves, much higher than those in young and mature samples. Furthermore, MsC3H-39 can cause an earlier senescence phenotype and the senescence associated gene, *MsSAG113* can be induced by MsC3H-39 (Figures 9A–D). These data suggest that MsC3H-39 directly regulates the expression of *MsSAG113* and promotes leaf senescence.

In summary, *MsSAG113* gene, as a senescence regulator, can be induced by hormones and regulates plant senescence in alfalfa by participating in the hormone regulatory network. MsC3H-39 directly recognizes and binds the upstream of *MsSAG113* and enhances its expression (Figure 10). Meanwhile, A total of 42 TFs associated with plant senescence and 234 genes involved in hormone regulation was identified by transcriptome analysis, which further indicated the important role of *MsSAG113* gene in leaf senescence and hormone regulation network in alfalfa. This

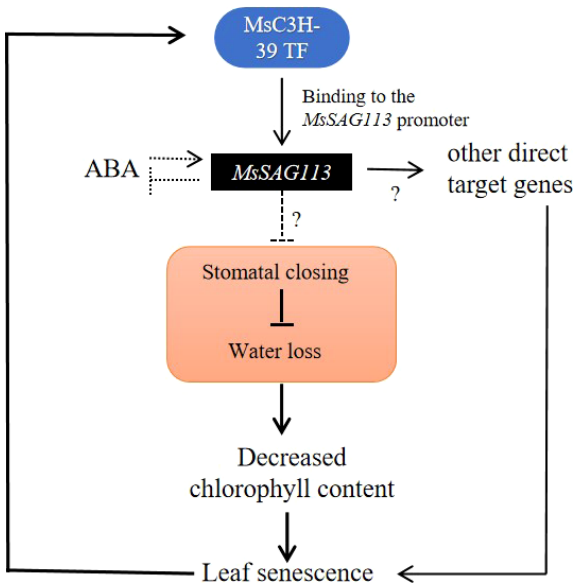


FIGURE 10
Potiential regulatory of *MsSAG113* in senescence in alfalfa. MsC3H-39 TF can enhance the expression of *MsSAG113* gene by recognizing and binding its promoter. ABA induces *MsSAG113* expression and *MsSAG113* negatively regulates ABA accumulation. *MsSAG113* may negatively regulate stomatal closure, resulting in accelerated water loss, a decrease in chlorophyll content and triggering plant senescence in alfalfa.

study lays a foundation for studying the mechanism of alfalfa senescence and provides new insights for improving the growth and production in pasture legumes.

Data availability statement

The data presented in the study are deposited in the NCBI repository. The accession numbers are PRJNA822614 and OP880214.

Author contributions

YHC, LH, and SL conceived and designed the experiment. SL and LZ conducted the experiments. DD and HX analyzed the data. CJ and YL contributed analysis tools. YHC provided financial support. SL wrote the manuscript. YHC and YLC edited and revised the manuscript. All authors made significant contributions to the manuscript and approved the final version for publication.

Funding

This work was supported by the Inner Mongolia Science and Technology Plan "Jiebangguashuai" Project (Breeding and Industrialization Demonstration of New Alfalfa Varieties with High Quality) and the Fundamental Research Funds for the Central Universities (Grant No. 2021ZY84).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1085497/full#supplementary-material

SUPPLEMENTARY TABLE 1
List of primers.

SUPPLEMENTARY TABLE 2
Differentially expressed genes.

SUPPLEMENTARY TABLE 3
Involved in hormone regulation genes in DEGs.

SUPPLEMENTARY TABLE 4
Transcriptional regulators in DEGs.

References

Balazadeh, S., Riano-Pachon, D. M., and Mueller-Roeber, B. (2008). Transcription factors regulating leaf senescence in arabidopsis thaliana. *Plant Biol.* 10, 63–75. doi: 10.1111/j.1438-8677.2008.00088.x

Balibrea Lara, M. E., Gonzalez Garcia, M., Fatima, T., Ehneß, R., Lee, T. K., Proels, R., et al. (2004). Extracellular invertase is an essential component of cytokinin-mediated delay of senescence. *Plant Cell* 16, 1276–1287. doi: 10.1105/tpc.018929

Bleecker, A. B., and Patterson, S. E. (1997). Last exit: Senescence, abscission, and meristem arrest in arabidopsis. *Plant Cell* 9, 1169–1179. doi: 10.1105/tpc.9.7.1169

Buchanan-Wollaston, V., Page, T., Harrison, E., Breeze, E., Lim, P. O., Nam, H. G., et al. (2005). Comparative transcriptome analysis reveals significant differences in gene expression and signalling pathways between developmental and dark/starvation-induced senescence in arabidopsis. *Plant J.* 42, 567–585. doi: 10.1111/j.1365-313X.2005.02399.x

Chao, Y., Xie, L., Yuan, J., Guo, T., Li, Y., Liu, F., et al. (2018). Transcriptome analysis of leaf senescence in red clover (Trifolium pratense l.). *Physiol. Mol. Biol. Plants* 24, 753–765. doi: 10.1007/s12298-018-0562-z

Chen, L. F. O., Hwang, J. Y., Charng, Y. Y., Sun, C. W., and And Yang, S. F. (2001). Transformation of broccoli (Brassica oleracea var. italica)with isopentenyltransferase gene *via* agrobacterium tumefaciens for post-harvest yellowing retardation. *Mol. Breed.* 7, 243–257. doi: 10.1023/A:1011357320259

Curty, C., and Engel, N. (1996). Detection, isolation and structure elucidation of a chlorophyll a catabolite from autumnal senescent leaves of cercidiphyllum japonicum. *Phytochemistry* 42, 1531–1536. doi: 10.1016/0031-9422(96)00155-0

Eckhardt, U., Grimm, B., and Hortensteiner, S. (2004). Recent advances in chlorophyll biosynthesis and breakdown in higher plants. *Plant Mol. Biol.* 56, 1–14. doi: 10.1007/s11103-004-2331-3

Engvall, E., and Perlmann, P. (1971). Enzyme-linked immunosorbent assay (ELISA) quantitative assay of immunoglobulin G. *Immunochemistry* 8, 871–874. doi: 10.1016/0019-2791(71)90454-X

Gan, S., and Amasino, R. M. (1997). Making sense of senescence. *Plant Physiol.* 113, 313–319. doi: 10.1104/pp.113.2.313

Gepstein, S., Sabehi, G., Carp, M., Hajouj, T., Nesher, M. F. O., Yariv, I., et al. (2003). Large-Scale identification of leaf senescence-associated genes. *Plant J.* 36, 629–642. doi: 10.1046/j.1365-313X.2003.01908.x

Gonzalez-Garcia, M. P., Rodriguez, D., Nicolas, C., Rodriguez, P. L., Nicolas, G., and Lorenzo, O. (2003). Negative regulation of abscisic acid signaling by the fagus sylvatica FsPP2C1 plays a role in seed dormancy regulation and promotion of seed germination. *Plant Physiol.* 133, 135–144. doi: 10.1104/pp.103.025569

Guo, Y., and Gan, S. (2012). Convergence and divergence in gene expression profiles induced by leaf senescence and 27 senescence-promoting hormonal, pathological and environmental stress treatments. *Plant Cell Environ.* 35, 644–655. doi: 10.1111/j.1365-3040.2011.02442.x

He, Y., Fukushige, H., Hildebrand, D. F., and Gan, S. (2002). Evidence supporting a role of jasmonic acid in arabidopsis leaf senescence. *Plant Physiol.* 128, 876–884. doi: 10.1104/pp.010843

He, Y., and Gan, S. (2002). A gene encoding an acyl hydrolase is involved in leaf senescence in arabidopsis. *Plant Cell* 14, 805–815. doi: 10.1105/tpc.010422

He, Y. Y., Tang, W. W., Swain, J. D. J. D., Green, A. L. A. L., Jack, T. P. T. P., and Gan, S. S. (2001). Networking senescence-regulating pathways by using arabidopsis enhancer trap lines1. *Plant Physiol. (Bethesda)* 126, 707–716. doi: 10.1104/pp.126.2.707

Hoekema, A., Hirsch, P. R., and Al, H. P. J. J. (1983). 20-a binary plant vector strategy based on separation of vir- and T-region of the agrobacterium tumefaciens Ti-plasmid. *Nature* 303, 179–180. doi: 10.1038/303179a0

Hrbáčková, M., Dvořák, P., Takáč, T., Tichá, M., Luptovčiak, I., Šamajová, O., et al. (2020). Biotechnological perspectives of omics and genetic engineering methods in alfalfa. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00592

Hunter, D. A., Ferrante, A., Vernieri, P., and Reid, M. S. (2004). Role of abscisic acid in perianth senescence of daffodil (Narcissus pseudonarcissus " Dutch master "). *Physiol. Plant.* 121, 313–321. doi: 10.1111/j.1399-3054.2004.00311.x

Jan, A., Maruyama, K., Todaka, D., Kidokoro, S., Abo, M., Yoshimura, E., et al. (2013). OsTZF1, a CCCH-tandem zinc finger protein, confers delayed senescence and stress tolerance in rice by regulating stress-related genes. *Plant Physiol.* 161, 1202–1216. doi: 10.1104/pp.112.205385

Jaradat, M. R., Feurtado, J. A., Huang, D., Lu, Y., and Cutler, A. J. (2013). Multiple roles of the transcription factor AtMYBR1/AtMYB44 in ABA signaling, stress responses, and leaf senescence. *BMC Plant Biol.* 13, 192. doi: 10.1186/1471-2229-13-192

Jing, H. C., Sturre, M. J., Hille, J., and Dijkwel, P. P. (2002). Arabidopsis onset of leaf death mutants identify a regulatory pathway controlling leaf senescence. *Plant J.* 32, 51–63. doi: 10.1046/j.1365-313x.2002.01400.x

Kant, S., Burch, D., Badenhorst, P., Palanisamy, R., Mason, J., and Spangenberg, G. (2015). Regulated expression of a cytokinin biosynthesis gene IPT delays leaf senescence and improves yield under rainfed and irrigated conditions in canola (Brassica napus l.). *PloS One* 10, e116349. doi: 10.1371/journal.pone.0116349

Klug, A., and Schwabe, J. W. (1995). Protein motifs 5. zinc fingers. *FASEB J.* 9, 597–604. doi: 10.1096/fasebj.9.8.7768350

Lee, S., Jung, H. J., Kang, H., and Kim, S. Y. (2012). Arabidopsis zinc finger proteins AtC3H49/AtTZF3 and AtC3H20/AtTZF2 are involved in ABA and JA responses. *Plant Cell Physiol.* 53, 673–686. doi: 10.1093/pcp/pcs023

Liebsch, D., and Keech, O. (2016). Dark-induced leaf senescence. new insights into a complex light-dependent regulatory pathway. *New Phytol.* 212, 563–570. doi: 10.1111/nph.14217

Lü, P., Zhang, C., Liu, J., Liu, X., Jiang, G., Jiang, X., et al. (2014). RhHB1 mediates the antagonism of gibberellins to ABA and ethylene during rose (Rosa hybrida) petal senescence. *Plant J.* 78, 578–590. doi: 10.1111/tpj.12494

Mackay, J. P., and Crossley, M. (1998). Zinc fingers are sticking together. *Trends Biochem. Sci.* 23, 1–4. doi: 10.1016/S0968-0004(97)01168-7

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

Matile, P., Ginsburg, S., Schellenberg, M., and Thomas, H. (1988). Catabolites of chlorophyll in senescing barley leaves are localized in the vacuoles of mesophyll cells. *Proc. Natl. Acad. Sci. United States America* 85, 9529–9532. doi: 10.1073/pnas.85.24.9529

Merlot, S., Gosti, F., Guerrier, D., Vavasseur, A., and Giraudat, J. (2001). The ABI1 and ABI2 protein phosphatases 2C act in a negative feedback regulatory loop of the abscisic acid signalling pathway. *Plant J.* 25, 295–303. doi: 10.1046/j.1365-313x.2001.00965.x

Morris, K., Mackerness, S. A. H., Page, T., John, C. F., Murphy, A. M., Carr, J. P., et al. (2000). Salicylic acid has a role in regulating gene expression during leaf senescence. *Plant J. For Cell Mol. Biol.* 23, 677–685. doi: 10.1046/j.1365-313x.2000.00836.x

Nishimura, N., Yoshida, T., Kitahata, N., Asami, T., Shinozaki, K., and Hirayama, T. (2007). ABA-hypersensitive Germination1 encodes a protein phosphatase 2C, an essential component of abscisic acid signaling in arabidopsis seed. *Plant J.* 50, 935–949. doi: 10.1111/j.1365-313X.2007.03107.x

Pruzinská, A., Tanner, G., Aubry, S., Anders, I., Moser, S., Müller, T., et al. (2005). Chlorophyll breakdown in senescent arabidopsis leaves. characterization of chlorophyll catabolites and of chlorophyll catabolic enzymes involved in the degreening reaction. *Plant Physiol.* 139, 52–63. doi: 10.1104/pp.105.065870

Richmond,, Amos, E., and Lang, A. (1957). Effect of kinetin on protein content and survival of detached xanthium leaves. *Science* 125, 650–651. doi: 10.1126/science.125.3249.650

Robinson, M. D., McCarthy, D. J., and Smyth, G. K. (2009). EdgeR: A bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26, 139–140. doi: 10.1093/bioinformatics/btp616

Sanchez-Garcia, I., and Rabbitts, T. H. (1994). The LIM domain a new structural motif found in zinc-finger-like proteins. *Trends Genet.* 10, 315–320. doi: 10.1016/0168-9525(94)90034-5

Schmittgen, T. D., and Livak, K. J. (2008). Analyzing real-time PCR data by the comparative CT method. *Nat. Protoc.* 3, 1101–1108. doi: 10.1038/nprot.2008.73

Schweighofer, A., Hirt, H., and Meskiene, I. (2004). Plant PP2C phosphatases: Emerging functions in stress signaling. *Trends Plant Sci.* 9, 236–243. doi: 10.1016/j.tplants.2004.03.007

Singh, S., Singh, A., and Nandi, A. K. (2016). The rice OsSAG12-2 gene codes for a functional protease that negatively regulates stress-induced cell death. *J. Biosci.* 41, 445–453. doi: 10.1007/s12038-016-9626-9

Smart, C. M. (1994). Tansley review no. 64. gene expression during leaf senescence. *New Phytol.* 126, 419–448. doi: 10.1111/j.1469-8137.1994.tb04243.x

Swartzberg, D., Dai, N., Gan, S., Amasino, R., and Granot, D. (2006). Effects of cytokinin production under two SAG promoters on senescence and development of tomato plants. *Plant Biol.* 8, 579–586. doi: 10.1055/s-2006-924240

van der Graaff, E. E., Schwacke, R. R., Schneider, A. A., Desimone, M. M., Flügge, U. U., and Kunze, R. R. (2006). Transcription analysis of arabidopsis membrane transporters and hormone pathways during developmental and induced leaf senescence1. *Plant Physiol. (Bethesda)* 141, 776–792. doi: 10.1104/pp.106.079293

Wang, R. S., Pandey, S., Li, S., Gookin, T. E., Zhao, Z., Albert, R., et al. (2011). Common and unique elements of the ABA-regulated transcriptome of arabidopsis guard cells. *BMC Genomics* 12, 216. doi: 10.1186/1471-2164-12-216

Weaver, L. M., and Amasino, R. M. (2001). Senescence is induced in individually darkened arabidopsis leaves, but inhibited in whole darkened plants. *Plant Physiol.* 127, 876–886. doi: 10.1104/pp.010312

Wistrom, C., and Villeponteau, B. (1992). Cloning and expression of SAG: A novel marker of cellular senescence. *Exp. Cell Res.* 199, 355. doi: 10.1016/0014-4827(92)90445-E

Yan, Z., Jia, J., Yan, X., Shi, H., and Han, Y. (2017). Arabidopsis KHZ1 and KHZ2, two novel non-tandem CCCH zinc-finger and K-homolog domain proteins, have redundant roles in the regulation of flowering and senescence. *Plant Mol. Biol.* 95, 549–565. doi: 10.1007/s11103-017-0667-8

Young, M. D., Wakefield, M. J., Smyth, G. K., and Oshlack, A. (2010). Gene ontology analysis for RNA-seq: Accounting for selection bias. *Genome Biol.* 11, R14. doi: 10.1186/gb-2010-11-2-r14

Zhang, K., and Gan, S. (2012). An abscisic acid-AtNAP transcription factor-SAG113 protein phosphatase 2C regulatory chain for controlling dehydration in senescing arabidopsis leaves. *Plant Physiol.* 158, 961–969. doi: 10.1104/pp.111.190876

Zhang, K., Xia, X., Zhang, Y., and Gan, S. (2012). An ABA-regulated and golgi-localized protein phosphatase controls water loss during leaf senescence in arabidopsis. *Plant J.* 69, 667–678. doi: 10.1111/j.1365-313X.2011.04821.x

Zhan, J., He, H., Wang, T., Wang, A., Li, C., and He, L. (2013). Aluminum-induced programmed cell death promoted by AhSAG, a senescence-associated gene in arachis hypoganea l. *Plant Sci.* 210, 108–117. doi: 10.1016/j.plantsci.2013.05.012

Zhao, Y., Chan, Z., Gao, J., Xing, L., Cao, M., Yu, C., et al. (2016). ABA receptor PYL9 promotes drought resistance and leaf senescence. *Proc. Natl. Acad. Sci.* 113, 1949–1954. doi: 10.1073/pnas.1522840113

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 December 2022
DOI 10.3389/fpls.2022.1098605

OPEN ACCESS

EDITED BY
Qiang Guo,
Beijing Academy of Agricultural and Forestry Sciences, China

REVIEWED BY
Jiyu Zhang,
Lanzhou University, China
Peizhi Yang,
Northwest A&F University, China

*CORRESPONDENCE
Zhuo Yu
yuzhuo58@sina.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 15 November 2022
ACCEPTED 05 December 2022
PUBLISHED 20 December 2022

CITATION
Lu Q, Yu X, Wang H, Yu Z, Zhang X and Zhao Y (2022) Quantitative trait locus mapping for important yield traits of a sorghum-sudangrass hybrid using a high-density single nucleotide polymorphism map.
Front. Plant Sci. 13:1098605.
doi: 10.3389/fpls.2022.1098605

Quantitative trait locus mapping for important yield traits of a sorghum-sudangrass hybrid using a high-density single nucleotide polymorphism map

Qianqian Lu†, Xiaoxia Yu†, Huiting Wang, Zhuo Yu*, Xia Zhang and Yaqi Zhao

Agricultural College, Inner Mongolia Agricultural University, Hohhot, Inner Mongolia, China

The sorghum-sudangrass hybrid is a vital gramineous herbage.The F2 population was obtained to clarify genetic regularities among the traits of sorghum-sudangrass hybrids by bagging and selfing in the F1 generation using 'scattered ear sorghum' and 'red hull sudangrass.' This hybrid combines the characteristics of the strong resistance of parents, high yield, and good palatability and has clear heterosis. A thorough understanding of the genetic mechanisms of yield traits in sorghum-sudangrass hybrids is essential in improving their yield. Therefore, we conducted quantitative trait locus (QTL) mapping for plant height, stem diameter, tiller number, leaf number, leaf length, leaf width, and fresh weight of each plant in three different environments, using a high-density genetic linkage map based on single nucleotide polymorphism markers previously constructed by our team. A total of 55 QTLs were detected, uniformly distributed over the 10 linkage groups (LGs), with logarithm of odds values ranging between 2.5 and 7.1, which could explain the 4.9–52.44% phenotypic variation. Furthermore, 17 yield-related relatively high-frequency QTL (RHF-QTL) loci were repeatedly detected in at least two environments, with an explanatory phenotypic variation of 4.9–30.97%. No RHF-QTLs were associated with the tiller number. The genes within the confidence interval of RHF-QTL were annotated, and seven candidate genes related to yield traits were screened. Three QTL sites overlapping or adjacent to previous studies were detected by comparative analysis. We also found that QTL was enriched and that qLL-10-1 and qFW-10-4 were located at the same location of 25.81 cM on LG10. The results of this study provide a foundation for QTL fine mapping, candidate gene cloning, and molecular marker-assisted breeding of sorghum-sudangrass hybrids.

KEYWORDS

sorghum-sudangrass hybrid, ultra-high-density genetic map, F2 population, yield traits, quantitative trait locus mapping, candidate gene prediction

1 Introduction

The sorghum-sudangrass hybrid (*Sorghum bicolor* × *Sorghum sudanense* [Piper] Stapf) is an annual warm-season forage that provides water and soil conservation benefits. It combines the advantages of high grass and large leaf yield, improved tillering, good regeneration, and high nutritional value from both parents, providing substantial heterosis. Cattle, sheep, geese, ducks, poultry, and other livestock utilize this hybrid as feed. The sorghum-sudangrass hybrid presents a developed root system, good soil capacity, and strong ecological adaptability and further presents broad development and utilization prospects in environmental protection and animal husbandry, especially the development of the breeding industry in agricultural areas (Saballos et al., 2008; Han et al., 2015; Mahmoudzadeh and Oad, 2018; Peng et al., 2020). In recent years, with the rapid development of animal husbandry and frequent extreme weather worldwide, increasing demand for high-quality forage has indicated a shortage of high-quality green and fresh forage. This is the main reason for limiting the quality and efficiency of animal husbandry in the northwestern Himalayan region of India (Mir et al., 2020; Kumar et al., 2022), the southern part of the Korean Peninsula and Jeju Island (Peng et al., 2020), the northern and southeastern United States (Wells et al., 2014; Nave et al., 2020), northern Europe (Liatukienė et al., 2008), Canada (Bélanger et al., 2006), and parts of China (Fang, 2018). Therefore, increasing forage yield is crucial in meeting sufficient livestock nutrition and accelerating production. Recently, the sorghum-sudangrass hybrid has been widely cultivated and applied due to its high grass yield, strong resistance, rich nutrient content, and wide adaptability, which will ultimately play a positive role in sustainable animal husbandry development and alleviate the contradiction between supply and demand of herbage.

Forage grass is a crucial material base of animal husbandry and the main source of livestock nutrition (Li et al., 2015; Zhang et al., 2021). Plant height (PH), stem thickness (ST), tiller number (TN), leaf-related traits (leaf number (LN), leaf length (LL), and leaf width (LW), and fresh weight (FW) represent the most important yield traits. An understanding of the genetic basis of these traits carries considerable importance regarding increasing the output of sorghum-sudangrass hybrids (Ma et al., 2018). In the past decades, gene identification related to crop target traits was mainly achieved *via* hybrid population breeding (Girma et al., 2019). Researchers from the United States, India, Japan, Australia, and China have cultivated several forage varieties, such as Jianlibao (Jumbo), Leyi (Everlush), Mengong Green series forage, Wancao No. 2 (WC-2), and Jicao No. 1 (JC-1) (Han et al., 2015). However, conventional breeding results from the perennial selection of crop hybrid populations for important phenotypic traits, which are time-consuming and inefficient regarding targeted modifications for complex traits (Huang and Yan, 2019; Islam et al., 2021). With molecular biology's rapid development, molecular marker-assisted breeding has received increasing attention in forage breeding. Understanding the quantitative trait locus (QTL) function associated with important yield traits is an effective way to improve crop yield.

QTL mapping has been widely used to study yield traits in sorghum-sudangrass hybrids. Shi et al. (2017) constructed genetic linkage maps based on amplified fragment length polymorphism (AFLP) molecular markers and detected 13, 11, 10, 2, and 9 QTLs for stem diameter (SD), LL, TN, LW, and LN, respectively. Wang et al. (2021) detected four QTLs associated with PH and one QTL related to TN based on a genetic map containing 133 SSR markers. Using a genetic linkage map containing 181 SSR markers, Yu et al. (2018) further detected two QTLs related to LN, LL, LW, and TN each.; four QTLs with SD; and one QTL associated with PH. Liu et al. (2015) constructed a genetic linkage map with 124 SSR markers and identified one QTL each for PH, SD and TN, and three QTLs each for LN and FW. However, genetic maps constructed using traditional molecular markers, such as simple-sequence repeats (SSR) and AFLP, have low density and few QTL sites, which complicates meeting the needs of further QTL fine mapping (West et al., 2006).

With the gradual increase in available plant genome sequences, the combination of next-generation sequencing technology and reference genomes provides a new research strategy for QTL mapping and molecular marker-assisted breeding of important plant traits (Hart et al., 2001). The genome size of sorghum is about 730 Mb. Since the sudangrass and sorghum-sudangrass hybrid genome sequencing was incomplete, we used the sorghum genome as a reference genome to develop single nucleotide polymorphism (SNP) markers for sorghum-sudangrass hybrid. Compared to AFLP and SSR markers, SNP markers are abundant and polymorphic, allowing high-density genetic map construction (Wen et al., 2020). A single genetic linkage map has been constructed using SNP markers in sorghum-sudangrass hybrids (Jin et al., 2021). SNP markers can be developed on a large scale through different high-throughput sequencing technologies, such as genotyping-by-sequencing (GBS) (Elshire et al., 2011), restriction site-related DNA tag sequencing (RAD-seq) (Xu et al., 2018), specific-length amplified fragment sequencing (SLAF-seq) (Sun et al., 2013), and whole-genome resequencing (WGRS) (Li et al., 2009). WGRS is a sequence-based genotyping method that is not limited by a restriction site and can quickly obtain a considerable number of recombinant breakpoints and marker density. According to research needs, sequencing coverage can be adjusted to the whole or part of the population, which vastly improves the accuracy of QTL mapping (Hart et al., 2001). WGRS has been successfully applied to map construction and QTL mapping of related traits of a variety of important crops, such as rice (Smulders et al., 2019), sorghum

(Zhang et al., 2021), and peanuts (Jiang et al., 2021). However, WGRS has not been used in QTL mapping for traits related to the sorghum-sudangrass hybrid, and it is rarely reported in gene prediction.

Therefore, the aim of this study was to conduct QTL mapping analysis for seven traits (i.e., PH, SD, TN, FW, and the leaf-related traits LN, LL, and LW), and the genes in the stably detected QTL intervals were annotated based on the high-density SNP map of sorghum-sudangrass hybrids constructed in previous research by our group. These results will serve as a foundation for further studies such as QTL fine mapping, functional key gene verification, and marker-assisted breeding.

2 Materials and methods

2.1 Plant material

Sorghum-sudangrass hybrid F_2 populations and the parental material were obtained by bagging and selfing in the F_1 generation using 'scattered ear sorghum' and 'red hull sudangrass' as parents, which were provided by the College of Agronomy, Inner Mongolia Agricultural University. Each was separated by their tillers per seedlings, cloned after tissue propagation, and stored in a group culture room. The female sorghum is the main cultivar in northeast China, while the male sorghum is widely cultivated in Inner Mongolia. There are substantial differences in agronomic traits such as PH, SD, TN, FW, and leaf-related traits (LN, LL, and LW) between the two varieties (Figure 1).

2.2 Site conditions

Three field experiments were performed in Khorqin Grassland Station of Tongliao, Inner Mongolia (T), and a teaching farm in the East Zone of Inner Mongolia Ancient Agricultural University in Hohhot (H), Inner Mongolia, in 2021 and 2022. These were recorded as 2021-H, 2022-H, and 2022-T. The randomized complete block design (RCBD) was used in the field experiment and was repeated three times at each field test site. Three plants were used for each replication. The F_2

FIGURE 1
Phenotype and field growth of parents and part of F2 population of sorghum-sudangrass hybrid. **(A)** Scattered ear sorghum (♀); **(B)** Red hull sudangrass (♂); **(C)** Male parent growing in the field; **(D)** Female parent growing in the field; **(E)** F_2 population growing in the field.

population and parents were transplanted into the field at the end of April each year, using a plant and row spacing of 20 and 45 cm, respectively. The soil type of the test site was meadow and sandy loam soil, with moderate soil fertility and good irrigation conditions. Timely tillage, weeding, and insect and disease control management were performed throughout the growing season.

2.3 Trait measurement and data analysis

At the full flowering stage, each trait index of 150 F_2 isolated individual plants and their parents were determined. The PH was measured from the plant root to the spike tip, and the LL and LW of the second leaf were measured. The SD was measured from the middle of the stem near the soil using vernier calipers, and the LN and TN of each plant were recorded. Finally, the whole plant shoot was weighed to determine the FW (Liu et al., 2015).

Excel 2010 (Microsoft Corp., Redmond, WA, USA) and SPSS 22.0 software (SPSS, Chicago, Illinois, USA) were used to calculate the phenotypic data of different traits and perform analysis of variance (ANOVA). The OriginPro software (OriginLab, Northampton, MA, USA) was used to construct the characteristics of frequency distribution and correlation analysis.

2.4 QTL mapping analysis

Based on SNP markers developed using WGRS, our research group previously constructed a high-density genetic linkage map of sorghum-sudangrass hybrids, which contains 5656 SNP markers distributed in 10 linkage groups (LGs), with 'scattered ear sorghum' and 'red hull sudangrass' as parents and 150 individual F_2 isolates as the test material (Lu et al., 2022). The total length of the covered genome was 2192.84 cM, and the average distance between markers was 0.39 cM. Based on the high-density genetic linkage map, the composite interval mapping (CIM) method of WinQTL software (https://brcwebportal.cos.ncsu.edu/qtlcart/WQTLCart.htm) was used to locate and analyze the phenotypic values of PH, SD, TN, LN, LL, LW, FW and their average values (AV) in three environments (2021-H, 2022-H, and 2022-T). The Logarithm of the odds (LOD) threshold for QTL was verified by 1000 arrangements with $P < 0.05$. The final genetic map was constructed using Mapchart 2.2 (https://www.mapchart.net). The QTL locus positions were annotated. The QTLs detected in more than two environments were defined as relatively high-frequency QTLs (RHF-QTLs) (Guo et al., 2020). QTL loci with two or more traits that overlap or were in close proximity were defined as QTL clusters, and QTLs that could explain more than 10% of phenotypic variation in different environments were considered major QTLs (Yang et al., 2022). The QTLs were named as follows: "q + trait abbreviation + linkage group + QTL serial number," for example, in "qPH-7-1," "q" is the QTL abbreviation, "PH" represents plant height, "7" represents the 7th linkage group, and "1" represents the first QTL on linkage group 7 (McCouch and Xiao, 1998; Ma et al., 2018).

2.5 Candidate gene annotation and prediction

Samples from the previous study (Lu et al., 2022) that used 'scattered ear sorghum' and 'red hull sudangrass' for high-throughput sequencing of the whole genome were re-sequenced with the Illumina HiSeqTM PE150 platform (Illumina, San Diego, CA, USA). The parental genotypes were sequenced separately at a sequencing depth of 29.71× and 28.77×, and 227.2 Mb and 234.42 Mb of data were collected, respectively. The specific steps of candidate gene prediction are as follows: (1) Detect SNP sites (different and homozygous sites) between parents, which were aligned to the reference genome sorghum *Sorghum bicolor* (sorghum) (https://phytozome-next.jgi.doe.gov/info/Sbicolor_v3_1_1) using Burrows-Wheeler Aligner (BWA) (http://bio-bwa.sourceforge.net/). Duplicate parts (rmDup) were removed using SAMTOOLS (https://www.htslib.org). The Bcftools command "merge" was used to combine 'scattered ear sorghum' and 'red hull sudangrass' genotypes, and SNP loci (different and homozygous) were retained. (2) On the basis of the results of step 1, in if the E value was $\leq 1e^{-10}$, BLAST (https://blast.ncbi.nlm.nih.gov/) was used to map markers on both sides of the QTL confidence interval to the physical location of the sorghum genome, to determine the variation sites in the target interval. (3) Based on the functional annotation of sorghum homologous genes, candidate genes for yield traits of sorghum were identified from the mutant loci of the second step.

3 Results

3.1 Phenotype and correlation analysis of yield-related traits

Phenotypic values of seven traits of the F_2 population and their parents (scattered ear sorghum (P1) and red hull sudangrass (P2)) are shown in Table 1 and Supplementary Table 1. Each parent trait showed significant differences in different environments. The SD (AV: 14.95 vs. 12.98) and LW (AV: 5.4 vs. 4.6) of P1 were higher than those of P2, while the opposite was true for the other five traits (PH, TN, LN, LL, and FW). The genetic variation of each trait in the F_2 population was high, and the coefficient of variation was between 7.5 and 57.74 (Table 1). The results of ANOVA showed that genotype had a significant influence on all traits ($P < 0.001$ or $P < 0.01$), and

10.3389/fpls.2022.1098605

TABLE 1 Phenotypic analysis for seven quality traits of sorghum-sudangrass hybrids.

Trait [a]	Environment [b]	Parent [c]		F_2 population						
		P1	P2	Max	Min	AV [d]	SD [e]	CV(%) [f]	Skewness	Kurtosis
PH	2021-H	227.5 ± 0.49**	340.2 ± 0.45	418	148	308.19	42.36	13.74	-0.51	0.96
	2022-H	189.6 ± 0.33**	330.8 ± 0.48	454	233	355.27	43.59	12.27	-0.36	-0.01
	2022-T	184.1 ± 0.36**	328.2 ± 0.44	395	208	285.93	35.66	12.47	0.18	0.20
	AV	200.69 ± 0.24**	333.1 ± 0.31	377	252	316.46	23.81	7.5	-0.09	-0.18
SD	2021-H	14.86 ± 0.36**	13.35 ± 0.33	16	8.16	12.11	1.59	13.12	0.01	-0.51
	2022-H	15.02 ± 0.35**	12.37 ± 0.46	19	9.02	13.51	1.92	14.21	0.02	-0.17
	2022-T	14.97 ± 0.28**	13.21 ± 0.45	20.02	7.92	12.23	2.43	19.86	0.51	-0.01
	AV	14.95 ± 0.16**	12.98 ± 0.13	15.67	9.01	12.62	1.35	10.69	0.05	-0.27
TN	2021-H	2 ± 0.33**	5 ± 0.42	8	0	3.29	1.89	57.74	0.58	-0.04
	2022-H	3 ± 0.33**	6 ± 0.31	8	0	3.32	1.74	52.40	0.23	-0.36
	2022-T	3 ± 0.33**	6 ± 0.33	8	0	3.31	1.65	49.84	0.32	0.04
	AV	3 ± 0.22**	6 ± 0.15	7	0	3.31	1.06	32.02	1.13	0.94
LN	2021-H	9 ± 0.22**	6 ± 0.25	12	6	9.34	1.14	12.21	-0.26	-0.25
	2022-H	7 ± 0.25**	12 ± 0.34	12	5	8.36	1.29	15.43	0.005	-0.27
	2022-T	8 ± 0.31**	10 ± 0.31	12	6	9.27	1.26	13.59	-0.14	-0.10
	AV	8 ± 0.15**	11 ± 0.15	11	7	8.99	0.83	9.23	-0.09	-0.10
LL	2021-H	63.4 ± 0.42**	69.5 ± 0.47	94	41.9	69.81	11.16	15.98	-0.11	-0.47
	2022-H	58.9 ± 0.52**	64.1 ± 0.32	101.6	43.5	69.54	10.83	15.57	0.13	-0.24
	2022-T	53.5 ± 0.43**	61.0 ± 0.52	91	38	65.44	11.97	18.29	-0.08	-0.64
	AV	58.6 ± 0.20**	64.9 ± 0.27	92.3	49.1	68.26	7.39	10.82	-0.09	0
LW	2021-H	5.7 ± 0.34**	4.27 ± 0.14	7.2	1.8	4.71	0.84	17.83	0.10	0.99
	2022-H	5.0 ± 0.51**	3.39 ± 0.12	6.5	3.3	4.77	0.65	13.62	0.07	-0.28
	2022-T	5.4 ± 0.27**	4.0 ± 0.16	7	2.6	4.72	0.89	18.85	-0.21	-0.14
	AV	5.4 ± 0.24**	3.9 ± 0.04	6.5	3.1	4.73	0.54	11.41	-0.10	0.49
FW	2021-H	148.8 ± 0.49**	233.3 ± 0.61	331.4	41.5	142.1	52.11	36.67	0.68	0.99
	2022-H	168.5 ± 0.35**	233.1 ± 0.64	374	102.2	211.3	48.81	23.10	0.77	0.98
	2022-T	343.9 ± 0.40**	282.9 ± 0.53	712	140	367.3	122.43	33.32	0.52	0.56
	AV	220.4 ± 0.25**	279.8 ± 0.35	392.4	140.1	240.2	52.38	21.80	0.37	-0.19

[a]PH, plant height; SD, stem diameter; TN, tiller number; LN, leaf number; LL, leaf length; LW, leaf width; FW, fresh weight. The units for PH, LL, LW are cm, SD is mm, FW is g.
[b]The populations planted in Hohhot in 2021 (2021-H); Hohhot in 2022 (2022-H); and Tongliao in 2022 (2022-T).
[c]P1, Scattered ear sorghum (♀); P2, Red hull sudangrass (♂).
[d]AV, Mean of the F_2 population.
[e]SD, Standard deviation.
[f]CV, coefficient of variation (SD/AV*100%).
** Indicates significance at the 0.01 level.

environmental factors had a significant influence on all traits except TN and LW (P <0.001) (Table 2). All traits had super parental separation in the 2021-H, 2022-H, and 2022-T environments and the average environment. The skewness and kurtosis of each trait were between 0.51–0.77 and 0.64–0.99, respectively. The absolute value was <1, which meets the normal distribution characteristics and is suitable for the next QTL localization study (Figure 2).

TABLE 2 Analysis of variance (ANOVA) for seven traits in three environments.

Trait	Factor	Sum of squares	Mean square	F
PH	Environment	1128193.65	564096.83	301.55***
	Genotype	71589.12	35794.56	19.14***
	Error	2508515.41	1870.63	
SD	Environment	540.98	270.49	29.23***
	Genotype	665.81	332.91	35.97***
	Error	12410.22	9.25	
TN	Environment	0.16	0.08	0.014
	Genotype	74.62	37.31	6.53**
	Error	7660.81	5.71	
LL	Environment	5389.44	2699.22	11.21***
	Genotype	3893.14	1946.57	8.09***
	Error	322805.89	240.72	
LW	Environment	0.87	0.44	0.33
	Genotype	14.01	7.0	5.33**
	Error	1761.1	1.31	
LN	Environment	269.85	134.93	23.12***
	Genotype	92.80	46.40	7.95***
	Error	7825.97	5.84	
FW	Environment	11981607.16	5990803.58	579.85***
	Genotype	98111.79	49055.89	4.75***
	Error	13854676.43	10331.60	

*** and ** Indicates significance at the 0.001 and 0.01 level, respectively.

Seven yield correlations were evaluated, and the results are shown in Figure 3. In the 2021-H, 2022-H, and 2022-T environments as well as the average environment, the Pearson correlation coefficient (r) in TN was the least significantly correlated with FW. All other traits were positively correlated with FW. In 2022-H, compared Compared with the other environments, the phenotypic difference in 2022-H changed substantially, indicating that the planting environment changed considerably. In the three environments and the average environment, the Pearson correlation coefficients (r) among leaf-related traits (LL, LW and LN) were significant, LN was significantly correlated with SD and PH, and LN was negatively correlated with SD. This was consistent with the research results of Jin et al. (2021). Consequently, we speculate that the genes controlling TN and SD traits may restrict each other. Therefore, leaf-related traits (LL, LW, and LN), FW, SD, and PH or LN may be in the same QTL cluster (i.e., with close or overlapping confidence intervals) and play a pleiotropic role in phenotypic control.

3.2 QTL mapping

The phenotypic data of PH, SD, TN, LL, LW, LN, and FW in the three environments and the average environment were analyzed by QTL mapping. A total of 266 QTLS QTLs were detected, which were distributed on 10 chromosomes, and the phenotypic variation explained ranged from 0 to 52.42% (Supplementary Table 2). The results revealed that 55 major QTLs and RHF-QTLs were detected, which were relatively evenly distributed in different intervals of 10 linkage groups. This included four PH-related QTLs, nine SD-related QTLs, four TN-related QTLs, and seven LL-related QTLs. There were 12 QTLs for LW, 9 for LN, and 10 for FW (Table 3 and Figure 4). The LOD values ranged from 2.5–7.1, which could explain the 4.9–52.44% phenotypic variation. There were 50 major QTLs with genetic contribution >10%, and 17 RHF-QTLs were detected in at least two environments, which were relatively evenly distributed among 10 linkage groups. These traits were reported for the first time based on the high-density genetic map, which could provide the foundation for improved yield and the

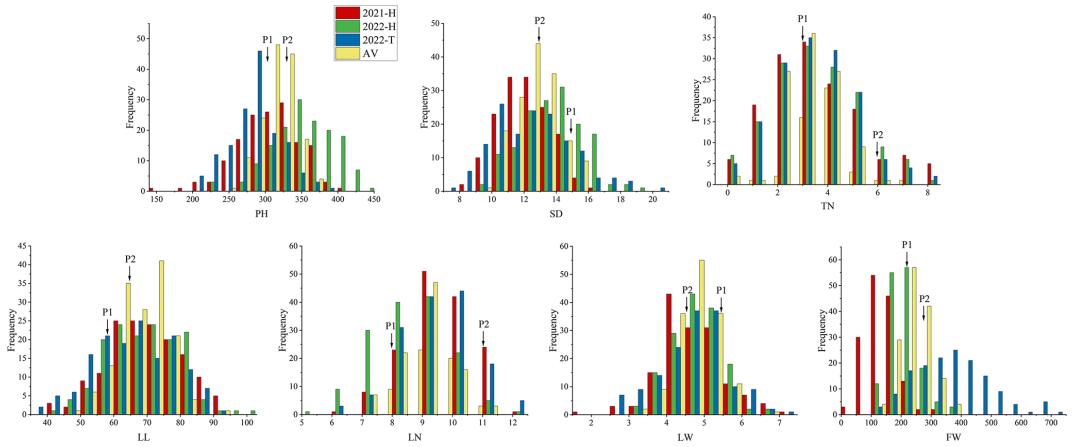


FIGURE 2
Frequency distribution maps of each yield-related trait in the F_2 population in three environments. The traits are plant height (PH), stem diameter (SD), tiller number (TN), leaf length (LL), leaf width (LW), leaf number (LN) and fresh weight (FW). Red, green, blue, and yellow labels indicate 2021-H, 2022-H, 2022-T, and AV, respectively. 2021-H, 2022-H, 2022-T and AV represent Hohhot (2021), Hohhot (2022), Tongliao (2022) and the average environment, respectively.

improvement of important agronomic traits of the sorghum-sudangrass hybrid. The specific positioning is indicated in Table 3.

For PH, four QTLs were detected on LG7 (Table 3 and Figure 4), with QTL ranging from 2.6–7.1, and the phenotypic variation rate was between 9.2% (2021-H, qPH-7-1) and 18.0% (2022-T, qPH-7-3). The qPH-7-1 locus was detected in 2021-H, 2022-T, and the average environment, and the phenotypic variation rates were >10%, indicating that the PH trait of this locus was less affected by the environment. It was the dominant QTL locus controlling the trait of the plant.

Nine QTLs associated with SD were identified in four linkage groups: LG7, LG8, LG9, and LG10. Four QTLs, namely qSD-8-1, qSD-8-2, qSD-8-3, and qSD-10-1, were detected in at least two environments, explaining the 6–14% phenotypic variation with negative additive effects, indicating a large influence by P2. The other five QTLs were detected in only one environment. Although they were notably affected by the environment, they explained between 12.15% and 25.65% of the phenotypic variation, and were the main effect were associated with SD. These QTLs will be located in future studies.

Four QTLs related to TN traits were identified. These were distributed on LG2, LG7, and LG9, and the phenotypic contribution rate was more than 10%. Only the additive effect of qTN-2-1 was negative, indicating that P2 had a promoting effect on the QTL, while the additive effect of qTN-7-1, qTN-9-1, and qTN-9-2 were positive, mainly promoting the effect of P1. No RFH-QTL was associated with TN, indicating that the environment strongly influenced TN.

For leaf-related traits (LL, LW, and LN), 7, 12 and 9 QTLs were detected for LL, LW, and LN, respectively (Table 3 and Figure 4). These explained 5.96% (AV, qLN-7-4) to 30.97% (2021-H, QLN-7-3) of the phenotypic variation. The additive effect values of LL and LN were mostly negative, and only two and one were positive, indicating that P2 increased the effect value of QTL in LL and LN. On the contrary, LW was considerably affected by P1, which was consistent with the results of the phenotypic traits analysis, indicating that LW was mainly controlled by P1.

Ten QTLs were associated with FW on LG1, LG6, LG9, and LG10, with the highest value of a single QTL being 5.7, explaining 4.9% (AV, qFW-10-3) to 52.44% (AV, qFW-6-2) of the phenotypic variation. All QTLs showed positive additive effects, and P2 increased the QTL effect. The loci qFW-10-1 (167.7–169.3 cM), qFW-10-2 (0.7–3.4 cM), qFW-10-3 (0.7–3.4 cM), and qFW-10-4 (23–26.4 cM) (related to FW) were located on LG10 under three environments and the average environment. However, due to the influence of the environment, the four QTLs were different from the same main locus, and the genetic differences were substantial.

3.3 RHF-QTL and QTL enrichment

Fifty-five QTLs associated with sorghum-sudangrass yield traits were identified in three environments and the average environment, which were relatively evenly distributed across 10 LGs. RHF-QTL of qph-7-1, qPH-7-2, qSD-8-1, qSD-8-2, qSD-8-3, qSD-10-1, qLL-10-1, qLW-4-1, qLW-4-3, qLN-7-2, qLN-7-3, qLN-7-4, qFW-1-1, qFW-6-1, qFW- 9-1, qFW-10-3, and qFW-10-4 were repeatedly detected in at least two environments, and no stable QTL association with TN was detected (Table 4).

10.3389/fpls.2022.1098605

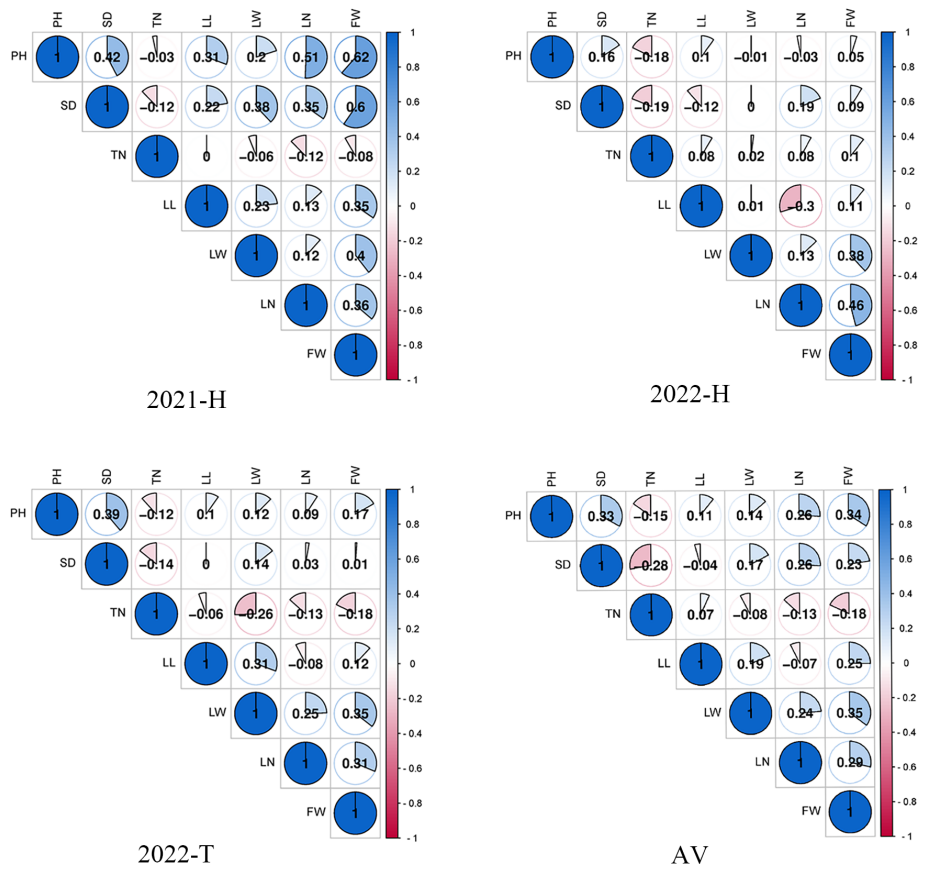


FIGURE 3
Pearson correlation coefficient (r) between traits in different environments.

TABLE 3 Quantitative trait loci (QTLs) identified for seven traits across three different environments (major and relatively high-frequency QTL [RHF-QTL]).

Traits [a]	QTLs	Treatments	Chrom [b]	Position	Marker interval	LODs [c]	Additive effects [d]	R^2 (%) [e]
PH	qPH-7-1	2021-H	7	4.71	4.4-5.3	3.3	15.94	9.2
		2022-T			4.4-5.5	5.2	16.45	10.25
		AV			4.2-5.4	3.4	9.8	11.66
	qPH-7-2	2022-H	7	8.11	7-9.4	3.1	16.00	7.76
		2022-T			7-9.4	4.4	16.73	13.56
		AV			7-10.5	3.0	9.83	12.45
	qPH-7-3	2022-T	7	10.81	10.5-11.8	4.1	14.76	18
	qPH-7-4	AV	7	0.71	0-3	3.4	8.82	13
SD	qSD-7-1	2022-T	7	263.01	261.3-263.4	3.1	-0.79	12.15
	qSD-8-1	2021-H	8	49.31	48.3-50.5	3.2	-0.59	9.7
		AV			48.7-50	4.2	-0.75	14

(Continued)

TABLE 3 Continued

Traits [a]	QTLs	Treatments	Chrom [b]	Position	Marker interval	LODs [c]	Additive effects [d]	R^2 (%) [e]
	qSD-8-2	2021-H	8	40.51	39.9-41.2	2.6	-0.51	9.5
		AV			40.2-41.2	3.7	-0.41	13.05
	qSD-8-3	2021-H	8	41.91	41.2-42.6	2.6	-0.47	9.04
		AV			41.2-43.1	3.4	-0.45	11.77
	qSD-8-4	AV	8	51.31	51-51.6	3.6	-0.64	15.29
	qSD-8-5	AV	8	52.71	52.5-54.4	5.0	-0.68	19.16
	qSD-9-1	2021-H	9	171.21	170.5-171.4	7.1	-0.92	25.62
	qSD-9-2	2021-H	9	177.91	177.3-178.3	3.9	-0.62	14.52
	qSD-10-1	2022-T	10	226.91	225.4-228.7	4.7	-0.48	12.19
		AV		227.71	225.9-230.2	3.2	-0.12	6.0
TN	qTN-2-1	2022-T	2	188.51	187.7-191.2	3.35	-0.62	11.15
	qTN-7-1	2022-T	7	119.41	118.5-121.8	4.1	0.74	17.56
	qTN-9-1	2022-H	9	82.81	82.2-84.8	3.4	0.68	14.7
	qTN-9-2	2022-H	9	90.51	84.8-94.5	3.1	0.78	15.86
LL	qLL-3-1	2021-H	3	163.51	162.4-163.8	3.1	-2.67	10.1
	qLL-4-1	2021-H	4	5.41	5-6.1	3.6	-2.49	10.32
	qLL-7-1	2021-H	7	13.81	12.8-14.8	3.1	6.57	10.21
	qLL-7-2	2021-H	7	148.01	144.3-150.3	3.2	-3.69	11.36
	qLL-7-3	2021-H	7	174.41	172.9-176.5	4.4	-7.03	21.65
	qLL-8-1	2022-H	8	180.21	179.1-181.6	5	2.82	17.06
	qLL-10-1	2021-H	10	25.81	24.2-26.6	4	-2.8	11.61
		AV			23-26.6	3.5	-1.41	9.11
LW	qLW-4-1	2021-H	4	149.11	145.6-151.8	3	-0.3	8.00
		2022-H			147.8-149.8	2.8	-0.19	9.83
	qLW-4-2	2021-H	4	155.21	153.2-156.4	3.2	-0.33	10.71
	qLW-4-3	2022-T	4	145.71	143.2-149.4	2.9	0.05	7.61
		AV			143.5-146.6	3.2	-0.07	8.50
	qLW-4-4	2022-T	4	170.61	168.7-173.7	3.0	-0.51	25.95
	qLW-5-1	2021-H	5	73.71	71.3-74.4	3.9	0.25	12.60
	qLW-5-2	2021-H	5	75.01	74.4-76.4	4.2	0.28	13.70
	qLW-5-3	2021-H	5	81.81	81-83.8	5.1	0.22	14.58
	qLW-5-4	2021-H	5	86.81	86.4-90	3.7	0.23	12.04
	qLW-5-5	2021-H	5	90.71	90-91.8	3.48	0.28	10.79
	qLW-8-1	AV	8	138.31	136.3-139.1	4.2	-0.16	13.48
	qLW-8-2	AV	8	226.31	225.4-227.2	4.6	-0.12	15.22
	qLW-9-1	2011-H	9	172.21	171-171.6	3.5	-0.31	10.65
LN	qLN-1-1	AV	1	6.81	6.2-8.4	3.3	0.33	14.54
	qLN-2-1	AV	2	43.41	42.4-45.1	3.1	-0.15	10.57

(Continued)

TABLE 3 Continued

Traits [a]	QTLs	Treatments	Chrom [b]	Position	Marker interval	LODs [c]	Additive effects [d]	R^2 (%) [e]
	qLN-2-2	AV	2	157.21	155.9-157.9	3.1	-0.16	9.39
	qLN-2-3	AV	2	159.21	157.9-159.7	3.5	-0.14	10.1
	qLN-5-1	2011-T	5	20.41	19.8-22.3	3	-0.54	14.37
	qLN-7-1	2021-H	7	198.01	197.8-198.9	4.2	-0.35	24.67
	qLN-7-2	2021-H	7	203.01	202-203.5	4.1	-0.43	25.33
		AV			203-204.1	2.6	-0.3	16.38
	qLN-7-3	2021-H	7	210.81	208.3-212.4	5.4	-0.66	30.97
		AV		208.81	204.1-211.7	2.6	-0.37	12.75
	qLN-7-4	2021-H	7	96.41	94.5-96.6	3.9	-0.58	14.84
		AV			94.4-98.4	2.6	-0.32	5.96
FW	qFW-1-1	2022-T	1	18.41	18.1-20.1	5	26.48	14.61
		AV		19.41		3.1	25.1	16.54
	qFW-3-1	AV	3	154.51	152.6-155.8	3.3	31.69	20.01
	qFW-6-1	2022-T	6	138.01	135.6-140.9	3.1	51.24	21.39
		AV		137.01	134.9-139.8	3.6	20.63	21.22
	qFW-6-2	AV	6	62.51	60.1-65.9	3.3	23.18	52.44
	qFW-6-3	AV	6	125.31	123.3-126.5	3.3	12.49	11.40
	qFW-9-1	2022-T	9	150.71	150.1-151.8	5.6	1.58	5.05
		AV			150.1-151.9	4.1	3.69	5.6
	qFW-10-1	2021-H	10	168.41	167.7-169.3	5.2	23.49	8.1
	qFW-10-2	2022-H	10	1.17	0.7-3.4	5.1	19.83	21.8
	qFW-10-3	2022-H	10	9.11	7.2-10.4	5.7	22.15	22.94
		AV			8.5-10.1	2.5	21.02	4.9
	qFW-10-4	2022-T	10	25.81	23-26.4	2.6	-41.02	8.26
		AV			23-19.6	3.4	-20.11	12.75

[a]The traits are PH, plant height; SD, stem diameter; TN, tiller number; LN, leaf number; LL, leaf length; LW, leaf width; FW, fresh weight 2021-H, 2022-H, 2022-T and AV represent the populations planted in Hohhot in 2021 (2021-H); Hohhot in 2022 (2022-H); Tongliao in 2022 (2022-T). and Mean of the F2 population respectively.
[b]Chrom, chomosome.
[c]LOD, logarithm of odds.
[d]Additive effect, positive effect was contributed by P1, negative effect was contributed by P2.
[e]R^2, Phenotypic variation.

Among these RHF-QTL, 6 showed a positive additive effect, and 11 showed a negative additive effect. The phenotypic variation was between 4.9% and 30.97%. qPH-7-1 and qPH-7-2, which control PH, were detected in 2021-H, 2022-T, and the average environment, with LOD values between 3.0–5.2, the highest phenotypic variation explained was 11.66% and 12.45% (>10%). Therefore, it is suggested that qPH-7-1 and qPH-7-2 may be the key QTLs controlling PH. We also found that the loci qLL-10-1 (23–26.6 cM, 25.81 cM) and qFW-10-4 (23–29.6 cM, 25.81 cM), related to the control of LL, overlapped on LG10 with QTL enrichment, which will be further explored in future research. In addition, compared to previous studies, QTL loci related to LW, LN, and FW overlapped or were close to the loci related to previous studies, which further confirmed the mapping accuracy (Table 5).

3.4 Candidate gene prediction

A total of 132 annotation genes were screened using gene mining of the RHF-QTL mapping intervals (Supplementary Table 3). According to the functional comparison of sorghum

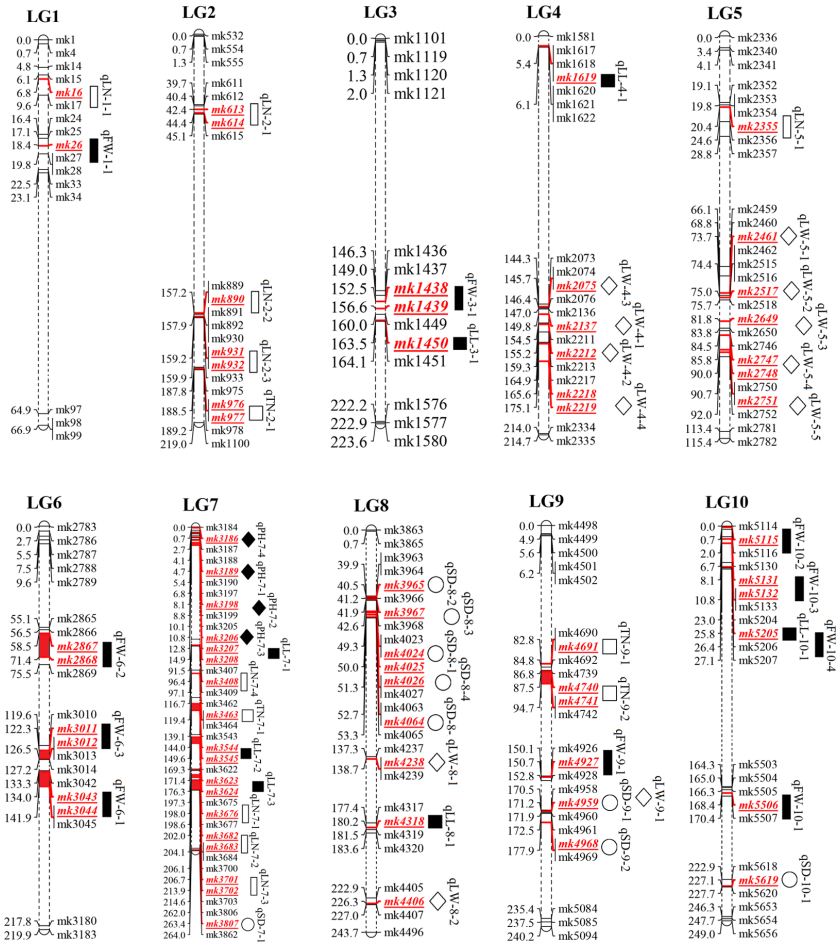


FIGURE 4
QTL mapping of yield traits in F_2 population of sorghum-sudangrass represented QTLs for plant height (PH), stem diameter (SD), tiller number (TN), leaf length (LL), leaf width (LW), leaf number (LN) and fresh weight (FW), respectively.

homologous genes, seven candidate genes that may affect the yield traits of sorghum-sudangrass were screened. The homologous genes are listed in Table 6. Among them, the homologous sorghum gene of *gene23531* was LOC8071161, encoding LRR receptor-like serine/threonine-protein kinase; the homologous gene of *gene26589* in sorghum was LOC8068853, encoding psbP domain-containing protein; the *gene15585* sorghum for LOC8056062 homologous gene, encoding galacturonosyltransferase protein; the sorghum homologous gene of *gene15584* was LOC8059564, encoding magnesium transporter NIPA2 protein; the homologous gene of *gene23381* in sorghum was LOC8080898, encoding DNA polymerase IA, chloroplastic protein; the homologous gene of *gene24523* in sorghum was LOC8054823, encoding RNA polymerase II transcription subunit 15a protein; and the homologous gene of gene31524 in sorghum was LOC8061987, encoding cytochrome P450 711A1 protein.

4 Discussion

The rapid development of molecular biology and the wide application of high-throughput sequencing technology have promoted new breeding strategies to increase crop yield and improve important yield-related traits. Several genetic linkage maps have been created, and some progress has been made in the QTL mapping of related traits. However, most of the existing genetic maps were constructed using restriction fragment length polymorphism (RFLP), sequence-related amplified polymorphism (SRAP), and simple-sequence repeats (SSR), which present few markers and large QTL confidence intervals, limiting their use in QTL fine mapping and marker-assisted breeding (Liu et al., 2015; Shi et al., 2017; Yu et al., 2018; Wang et al., 2021). Compared with the above molecular markers, SNP markers are widely used in map construction and QTL mapping of a variety of crops due to

TABLE 4 Relatively high-frequency quantitative trait loci (RHF-QTLs) detected in multiple environments (at least two) statistical analysis.

Traits	QTLs	Treatments	Marker interval	LODs [a]		Additive effects		R^2 (%)	
				Max	Min	Max	Min	Max	Min
PH	qPH-7-1	2021-H、2022-T、AV	4.2-5.5	5.2	3.4	16.45	9.8	11.66	9.2
	qPH-7-2	2022-H、2022-T、AV	7-10.5	4.4	3.0	16.73	9.83	12.45	7.76
SD	qSD-8-1	2021-H、AV	48.3-50.5	3.2	4.2	-0.59	-0.75	14	9.7
	qSD-8-2	2021-H、AV	39.9-41.2	3.7	2.6	-0.41	-0.51	13.05	9.5
	qSD-8-3	2021-H、AV	41.2-42.6	3.4	2.6	-0.47	-0.45	11.77	9.04
	qSD-10-1	2022-T、AV	225.4-230.2	3.2	4.7	-0.12	-0.48	6.0	12.19
LL	qLL-10-1	2021-H、AV	23-26.6	4	3.5	-1.41	-2.8	11.61	9.11
LW	qLW-4-1	2021-H、2022-H	145.6-151.8	3	2.8	-0.19	-0.3	9.83	8.00
	qLW-4-3	2022-T、AV	143.2-149.4	3.2	2.9	0.05	-0.07	7.61	8.50
LN	qLN-7-2	2021-H、AV	202-204.1	4.1	2.6	-0.3	-0.43	25.33	16.38
	qLN-7-3	2021-H、AV	204.1-212.4	5.4	2.6	-0.37	-0.66	30.97	12.75
	qLN-7-4	2021-H、AV	94.4-98.4	3.9	2.6	-0.32	-0.58	14.84	5.96
FW	qFW-1-1	2022-T、AV	18.1-20.1	5	3.1	26.48	25.1	16.54	14.61
	qFW-6-1	2022-T、AV	134.9-140.9	3.6	3.1	51.24	20.63	21.39	21.22
	qFW-9-1	2022-T、AV	150.1-151.8	5.6	4.1	3.69	1.58	5.6	5.05
	qFW-10-3	2022-H、AV	7.2-10.4	5.7	2.5	22.15	21.02	22.94	4.9
	qFW-10-4	2022-T、AV	23-26.4	3.4	2.6	-20.11	-41.02	12.75	8.26

[a]LOD, logarithm of odds.

their high density, uniform and extensive distribution on chromosomes, and high genetic stability (Smulders et al., 2019; Jiang et al., 2021; Zhang et al., 2021). Our research group previously constructed a high-density genetic linkage map of the sorghum-sudangrass hybrid, which contained 5656 SNP markers, covering a total genome length of 2192.84 cM, and the average distance between markers was 0.39 cM. (Lu et al., 2022) Compared with the previously constructed maps, the marker density increased, which effectively improved the accuracy of QTL mapping and provided a possibility for further screening of key genes and fine mapping of QTLs.

With the rapid development of sequencing technology and cost reductions, an increasing number of crop genomes have been exploited and applied in related studies (Li et al., 2021). Methods to detect QTL loci physically similar to or overlapping with the reference genome by combining specific crop genome sequences (at the chromosomal level) and high-density genetic maps have been widely used in the rapid identification of QTL loci and potential candidate genes (Luo et al., 2020). Since the sequencing of sudangrass has not been completed, Jin et al. (2021) reported that it is feasible to use the sorghum genome as the reference genome of the sorghum-sudangrass hybrid. In this study, 55 QTLs related to yield traits were detected in three

TABLE 5 Quantitative trait loci (QTLs) mapped on the same chromosome or adjacent marker regions in the current study and previous studies.

Chromosomes	The closest markers	QTLs in this study	QTLs detected in previous studies	
			Related traits [a]	Reference
4	qlw2	qLW-4-4(AV)	LW	Shi et al. (2017)[19]
5	LN2-1	qLN-5-1(AV)	LN	Yu et al. (2018)[21]
6	QFBMS6.1	qFW-6-1(2022-T,AV)	FW	Wang et al. (2014)[39]
	qFW6			Jin et al. (2021)[25]
	qTW6			Kajiya-K et al. (2020)[40]

[a]Related traits, LW, leaf width; LN, leaf number; FW, fresh weight.

TABLE 6 Annotated genes in interval of relatively high-frequency quantitative trait loci (RHF-QTLs).

Trait	QTL	Candidate genes	Homologous genesin *Sorghum bicolor* L.	Functional annotation
PH	qPH-7-2	*gene23531*	*LOC8071161*	LRR receptor-like serine/threonine-protein kinase
SD	qSD-8-1	*gene26589*	*LOC8068853*	psbP domain-containing protein
LW	qLW-4-1	*gene15585*	*LOC8056062*	galacturonosyltransferase 8
		gene15584	*LOC8059564*	probable magnesium transporter NIPA2
LN	qLN-7-2, qLN-7-3	*gene23381*	*LOC8080898*	DNA polymerase IA, chloroplastic
	qLN-7-3	*gene24523*	*LOC8054823*	RNA polymerase II transcription subunit 15a
FW	qFW-10-3	*gene31524*	*LOC8061987*	cytochrome P450 711A1

environments based on a high-density genetic linkage map (Table 2). Markers located in the same or similar positions on the same chromosome as those in previous studies are listed in Table 4. qFW-6-1 is a stable QTL associated with FW and was detected in two of the three environments. In contrast to the previous studies, qFW-6-1 (2402139–50582068 bp) was close to QFBMS6.1 (47686626–50991177 bp) (Wang et al., 2014), qFW6 (45, 156, 899–55, 463, 230 bp) (Jin et al., 2021), and qTW6 (49894350–51216671 bp) (Kajiya-K et al., 2020), which explained 21.22–21.39% of the phenotypic variation, and is an important QTL associated with FW. The marker qLW-4-4 (170.61 cM) was close to the P9m58-453-P9m58-208 marker (171.1 cM) (Shi et al., 2017) (with a negligible difference of 0.49 cM between the two sites), and the phenotypic variation was as high as 25.95%, which was the main effect site of LW. Furthermore, qLN-5-1 is located at 20.41 cM (19.8–22.3 cM) on LG 5, close to marker LN2-1 (22.9 cM) (Yu et al., 2018), with a phenotypic variation of 14.37%. It is a major QTL associated with LN. In addition, most QTLs located on the same LG were detected in the new marker interval due to the differences in marker types, population types, size, and material planting environments used to construct genetic maps (Guo et al., 2020). For PH, two stable QTLs were identified, namely qPH-7-1 and qPH-7-2, which were located at 4.71 cM and 8.11 cM on LG 7. The qPH-7-1 and qPH-7-2 LOD values were 3.4–5.2 and 3.0–4.4, respectively, and the additive effect between alleles was positive, which explained 7.76–11.66% of the phenotypic variation, indicating a stable QTL locus associated with PH. Compared with previous studies, qPH-7-1 (4.2–5.4 cM) and qPH-7-2 (7–10.5 cM) were co-localized with qPH7 (110.3–112.92 cM) loci on chromosome 7 (Zou et al., 2012). However, both qPH-7-1 and qPH-7-2 were inconsistent with qPH7 markers, suggesting that these may be novel QTLs for PH. In subsequent studies, these QTLs will be the focus of our attention. Additionally, compared with previous studies, it was found that several QTLs related to yield traits were detected at different positions on the same chromosome. For example, the QTLs controlling SD was detected on LG7, LG8, LG9, and LG10 (Wang et al., 2014; Liu et al., 2015; Kong et al., 2020; Jin et al., 2021); that controlling LL on LG3, LG4, LG7, LG8, LG9, and LG10 (Wang et al., 2014; Shi et al., 2017; Yu et al., 2018); and the one related to TN on LG2 (Li et al., 2015; Rayaprolu et al., 2021).

Yield traits such as PH, SD, TN, leaf-related traits, and FW are important quantitative traits, which are influenced considerably by the environment. The results of QTL localization are different in different environments, and the accuracy of the localization results can be guaranteed by setting up multi-year and multi-environmental tests to verify the QTL detected in multiple environments (Wang et al., 2009; Hou et al., 2015). Generally, QTLs that can be located in multiple environments (at least two environments) or all environments and have similar effects are defined as stable QTLs; otherwise, they are considered to be greatly influenced by the environments that they interact with (Xie et al., 2008). Guo et al. (2020) detected 38 QTLs related to protein and 68 related to starch in three environments and the average environment, respectively, and 26 stable QTLs were detected in more than two environments. Liu et al. (2022) identified 183 QTLs related to cotton fiber and yield traits in six environments, 62 QTLs for fiber and 10 QTLs for yield stability were identified in multiple environments. Ma et al. (2018) mapped maize leaf-related traits in three environments and eight stable QTLs in two or three environments, explaining 4.38–19.99% of the phenotypic variation. Additionally, Yang et al. (2022) identified 105 QTLs related to cotton in three environments, a total of 25 stable QTLs were detected in more than two environments. In this study, 55 QTLs related to yield traits of the sorghum-sudangrass hybrid were mapped in 3 environments and the average environment, and 17 RHF-QTLs were repeatedly detected in at least two environments (Table 3). Among them, there were two stable sites related to PH, four in SD, one in LL, two in LW, three in LN, and five in FW. These could explain 4.9–30.97% of the phenotypic variation. qPH-7-1 and qPH-7-2, which control plant height, could be detected in 2021-H, 2022-T, and the average environments, explaining 7.76–13.56% of the phenotypic variation (>10%), which may be

stable QTL controlling PH formation. Meanwhile, we also found that a pair of stable QTLs controlling different traits were co-localized at the same position and chromosome. The qLL-10-1 (23–26.6 cM, 25.81 cM) and qFW-10-4 (23–29.6 cM, 25.81 cM) markers were completely or only partially overlapped in LG10.

The LRR receptor-like serine/threonine-protein kinase encoded by the *gene23531* has high homology with the brassinolide insensitivity (BRI) gene and jointly regulates brassinosteroid signals. Ou et al. (2015) investigated dwarfing and non-dwarfing rootstock pear varieties based on RNA-sequence, and the results showed that the LRR receptor-like serine/threonine-protein kinase, a key gene controlling PH growth, was significantly up-regulated in the dwarfing varieties (Li et al., 2002). The psbP domain-containing protein encoded by *gene26589* participates in plant photosynthesis and plays an important role in the assembly of plant PS II and maintaining conformation stability (Bricker et al., 2013). Many studies have shown that the suppression of PsbP will cause the decline of the oxygen evolution ability of plants, a change in the direction of the electron transfer chain, and a lack of PS II oxidation and reduction function (Yi et al., 2007; Ido et al., 2009). The *gene15585*, coding the galacturonosyltransferase 8 (GT8) family proteins, is divided into two branches. The former mainly includes subclades of galacturonosyltransferase (GAUT) and galacturonosyltransferase-like TL (GA) genes, which play an important role in the synthesis of cell walls (Cheng et al., 2018). Kong et al. (2011) found that GAUT1 was involved in the synthesis process of pectin. In a tomato study, de Godoy et al. (2013) found that GAUT4 gene silencing treatment significantly reduced the pectin content. Therefore, it was speculated that *gene15585* might be key in leaf growth. The magnesium transporter NIPA2, encoded by the *gene15584*, is involved in plant photosynthesis and plays an important role in leaf growth and aging delay. Hermans and Verbruggen (2005) and Horlitz and Klaff (2000) also found that magnesium transporter NIPA2 is an important Mg^{2+} transporter, which could accelerate the continuous transport of Mg^{2+} to green tissues such as leaves, promoting the synthesis of green pigment and carbon assimilation and accelerating plant growth. It was speculated that the gene might be related to leaf photosynthesis. Chloroplasts are semi-autonomous organelles that contain their own DNA and can self-replicate. The *gene23381* and *gene24523* encode chloroplast DNA polymerase and RNA polymerase, respectively, which are key enzymes in plant regulation of chloroplast DNA synthesis and transcription in plants. RNA polymerase participates in the synthesis of various mRNA species, transcribes tRNA genes, regulates rRNA synthesis, and plays an important role in maintaining the growth and development of plant leaves (Börner et al., 2015). Zoschke et al. (2007) measured chloroplasts in *Arabidopsis* seeds and young, and old leaves at the transcriptional level and found that the transcriptional activity was relatively stable in all three. In this study, *gene23381* and *gene24523* were annotated in qLN-5-1, a QTL related to LN traits, suggesting that these genes may be important in controlling the growth and development of leaves of the sorghum-sudangrass hybrid. The *gene31524* encodes plant cytochrome P450. Renault et al. (2014) found that it has high catalytic activity, participates in a variety of metabolic reactions in plants, and plays an important role in signal transduction, pigment synthesis, light, electron transport, and biological defense. Therefore, it is speculated that *gene31524* may be the key gene affecting the growth and metabolism of the sorghum-sudangrass hybrid.

Under different environmental conditions, QTL loci detected at the same or adjacent loci on the same chromosome are called "QTL hotspots" or "QTL clusters," which are the result of the regionalized distribution of QTLs related to different traits (Yang et al., 2022). They are also the preferred regions for fine mapping and candidate gene identification (Xie et al., 2008). It can introduce genes related to crop quality, yield, and resistance into recipient crops simultaneously, control the correlation between traits in stable "QTL hotspots" or "QTL clusters" of different traits, and regulate pleiotropy of various traits through different metabolic pathways (Huang and Yan, 2019; Waheed et al., 2021). Numerous studies have shown that the phenomenon of "QTL hotspots" or "QTL clusters" are prevalent in a variety of crops, such as crested wheatgrass (Yang et al., 2022), wheat (Cui et al., 2016), and sorghum (Mace et al., 2012). In this study, we found that qLL-10-1 and qFW-10-4 were located at the same position of 25.81 cM on LG10 among the seven stable QTL controlling yield traits of sorghum-sudangrass hybrid. In addition, FW was significantly positively correlated with LL, in agreement with previous studies.

5 Conclusions

In this study, 55 major QTLs related to PH, SD, TN, LL, LW, LN, and FW were identified based on the high-density SNP map of the sorghum-sudangrass hybrid, among which 17 relatively RHF-QTL were detected in at least two environments. A stable QTL cluster containing QTLs controlling LL and FW (including at least one RHF-QTL) was detected, and three QTLs overlapping or located adjacent to the previously studied sites were identified. The genes in the RHF-QTL intervals were annotated, and seven candidate genes that might be related to PH, SD, LW, LN, and FW were screened. The results of this study will promote the fine mapping of QTL for yield traits of sorghum-sudangrass hybrids, cloning of key genes, and marker-assisted breeding.

Data availability statement

The data presented in the study are deposited in the figshare repository, accession number https://doi.org/10.6084/m9.figshare.21717146.v1.

Author contributions

ZY and QL conceived and designed the study. ZY, XY and QL performed the experiments. QL and XY wrote the article. HW, XZ, and YZ assisted in the performance of the experiments. ZY is the corresponding authors at the request of the Institute. All authors have read and approved the final manuscript.

Funding

This work was supported by grants from the National Natural Science Foundation of China (32060389), and the Fundamental Research Funds for the Universities (RZ2200001156). The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.

Acknowledgments

We thank the National Natural Science Fund of China, graduate student scientific research innovation projects in Inner Mongolia, and the Fundamental Research Funds for the Universities for funding received; HW and XZ for their assistance in the trait measurements; YZ for maintenance of the field plantings of the populations; XY and ZY for revising the manuscript; and Editage (editage.com) for English language editing services.

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1098605/full#supplementary-material

SUPPLEMENTARY TABLE 4
Phenotype data of Parental. 10.6084/m9.figshare.21667115

References

Bélanger, G., Castonguay, Y., Bertrand, A., Dhont, C., Rochette, P., Couture, L., et al. (2006). Winter damage to perennial forage crops in eastern Canada: causes, mitigation, and prediction. *Can. J. Plant Sci.* 86, 33–47. doi: 10.4141/P04-171

Börner, T., Aleynikova, A. Y., Zubo, Y. O., and Kusnetsov, V. V. (2015). Chloroplast RNA polymerases: role in chloroplast biogenesis. *Biochim. Biophys. Acta* 1847 (9), 761–769. doi: 10.1016/j.bbabio.2015.02.004

Bricker, T. ,. M., Roose, J. ,. L., Zhang, P., and Frankel, L. ,. K. (2013). The PsbP family of proteins. *Photosynth. Res.* 116, 235–250. doi: 10.1007/s11120-013-9820-7

Cheng, L., Ni, X., Zheng, M., Sun, L., Wang, X., and Tong, Z. (2018). Expressional characterization of galacturonosyltransferase-like gene family in eucalyptus grandis implies a role in abiotic stress responses. *Tree Genets Genomes* 14 (6), 81–91. doi: 10.1007/s11295-018-1294-5

Cui, F., Fan, X., Chen, M., Zhang, N., Zhao, C., Zhang, W., et al. (2016). QTL detection for wheat kernel size and quality and the responses of these traits to low nitrogen stress. *Theor. Appl. Genet.* 129, 469–484. doi: 10.1007/s00122-015-2641-7

de Godoy, F., Bermúdez, L., Lira, B. S., de Souza, A. P., Elbl, P., Demarco, D., et al. (2013). Galacturonosyltransferase 4 silencing alters pectin composition and carbon partitioning in tomato. *J. Exp. Bot.* 64 (8), 2449–2466. doi: 10.1093/jxb/ert106

Elshire, R. J., Glaubitz, J. C., Sun, Q., Poland, J. A., Kawamoto, K., Buckler, E. S., et al. (2011). A robust, simple genotyping-by-sequencing (GBS) approach for high diversity species. *PloS One* 6, e19379. doi: 10.1371/journal.pone.0019379

Fang, J. Y. (2018). Welcome the era when grass husbandry becomes half of china's modern agriculture. *Sci. Technol. Bull.* 63, 1615–1618. doi: 10.1360/N972018-00560

Girma, G., Nida, H., Seyoum, A., Mekonen, M., Nega, A., Lule, D., et al (2019). A large-scale genome-wide association analyses of ethopian sorghum landrace collection reveal loci associated with important traits. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.00691

Guo, Y., Zhang, G., Guo, B., Qu, C., Zhang, M., Kong, F., et al. (2020). QTL mapping for quality traits using a high-density genetic map of wheat. *PloS One* 15, e0230601. doi: 10.1371/journal.pone.0230601.g002

Han, L. Y., Li, J., Na, R. S., Yu, Z., and Zhou, H. (2015). Effect of two additives on the fermentation, *in vitro* digestibility and aerobic security of sorghum–sudangrass hybrid silages. *Grass Forage Sci.* 70, 185–194. doi: 10.1111/gfs.12092

Hart, G. E., Schertz, K. F., Peng, Y., and Syed, N. H. (2001). Genetic mapping of sorghum bicolor (L.) moench QTLs that control variation in tillering and other morphological characters. *Theor. Appl. Genet.* 103, 1232–1242. doi: 10.1007/s001220100582

Hermans, C., and Verbruggen, N. (2005). Physiological characterization of mg deficiency in arabidopsis thaliana. *J. Exp. Bot.* 56 (418), 2153–2161. doi: 10.1093/jxb/eri215

Horlitz, M., and Klaff, P. (2000). Gene-specific trans-regulatory functions of magnesium for chloroplast mRNA stability in higher plants. *J. Bio Chem.* 275 (45), 35638–35645. doi: 10.1074/jbc.M005622200

Hou, X., Liu, Y., Xiao, Q., Wei, B., Zhang, X., Gu, Y., et al. (2015). Genetic analysis for canopy architecture in an F2:3 population derived from two-type foundation parents across multi-environments. *Euphytica* 205, 421–440. doi: 10.1007/s10681-015-1401-8

Huang, L., and Yan, X. (2019). Construction of a genetic linkage map in pyropia yezoensis (Bangiales, rhodophyta) and QTL analysis of several economic traits of blades. *PloS One* 14, e0209128. doi: 10.1371/journal.pone.0209128

Ido, K., Ifuku, K., Yamamoto, Y., Ishihara, S., Murakami, A., Takabe, K., et al. (2009). Knockdown of the PsbP protein does not prevent assembly of the dimeric PSII core complex but impairs accumulation of photosystem II supercomplexes in tobacco. *Biochim. Biophys. Acta* 1787 (7), 873–881. doi: 10.1016/j.bbabio.2009.03.004

Islam, M. S., Pan, Y. B., Lomax, L., and Grisham, M. P. (2021). Identification of quantitative trait loci (QTL) controlling fibre content of sugarcane (Saccharum hybrids spp.). *Plant Breed.* 140, 360–366. doi: 10.1111/pbr.12912

Jiang, Y., Luo, H., Yu, B., Ding, Y., Kang, Y., Huang, L., et al. (2021). High-density genetic linkage map construction using whole-genome resequencing for mapping QTLs of resistance to aspergillus flavus infection in peanut. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.745408

Jin, P., Wang, L., Zhao, W., Zheng, J., Wang, Y. H., Liu, Y., et al. (2021). Construction of high density genetic map and QTL mapping in sorghum × sudangrass. *Euphytica* 217, 1–11. doi: 10.1007/s10681-021-02895-9

Kajiya-K, H., Takanashi, H., Fujimoto, M., Ishimori, M., Ohnishi, N., Wacera, W. F., et al. (2020). RAD-seq-based high-density linkage map construction and QTL mapping of biomass-related traits in sorghum using the Japanese landrace takakibi NOG. *Plant Cell Physiol.* 61, 1262–1272. doi: 10.1093/pcp/pcaa056

Kong, W., Nabukalu, P., Cox, T. S., Goff, V. H., Pierce, G. J., Lemke, C., et al. (2020). Transmission genetics of a sorghum bicolor × s. halepense backcross populations. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.00467

Kong, Y. ,. Z., Zhou, G. ,. K., Yin, Y. ,. B., Xu, Y., Pattathil, S., and Hahn, M. ,. G. (2011). Molecular analysis of a family of arabidopsis genes related to galacturonosyltransferases. *Plant Physiol.* 155 (4), 1791–1805. doi: 10.1104/pp.110.163220

Kumar, R., Joshi, R., Kumar, R., Srivatsan, V., Satyakam., Chawla, A., et al. (2022). Nutritional quality evaluation and proteome profile of forage species of Western himala ya. *Grassl. Sci.* 68, 214–225. doi: 10.1111/GRS.12357

Liatukienė, A., Liatukas, Ž., and Ruzgas, V. (2008). Winterhardiness as the key factor for selecting accessions of medicago sativa l. high-yielding germplasm. *Biologija* 54, 129–133. doi: 10.2478/v10054-008-0027-3

Li, F., Han, Z., Qiao, W., Wang, J., Song, Y., Cui, Y., et al. (2021). High-quality genomes and high-density genetic map facilitate the identification of genes from a weedy rice. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.775051

Li, R. Q., Li, Y. R., Fang, X. D., Yang, H. M., Wang, J., Kristiansen, K., et al. (2009). SNP detection for massively parallel whole-genome resequencing. *Genome Res.* 19, 1124–1132. doi: 10.1101/gr.088013.108

Liu, Y. L., Wang, L. H., Li, J. Q., Zhan, Q. W., Zhang, Q., Li, J. F., et al. (2015). QTL mapping of forage yield and forage yield component traits in sorghum bicolor × s. sudanense. *Genet. Mol. Res.* 14, 3854–3861. doi: 10.4238/2015.April.22.14

Liu, X., Yang, L., Wang, J., Wang, Y., Guo, Z., Li, Q., et al. (2022). Analyzing quantitative trait loci for fiber quality and yield-related traits from a recombinant inbred line population with *gossypium hirsutum* race palmeri as one parent. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.817748

Li, J. Q., Wang, L. H., Zhan, Q. W., Liu, Y. L., Zhang, Q., Li, J. F., et al. (2015). Mapping quantitative trait loci for five forage quality traits in a sorghum-sudangrass hybrid. *Genet. Mol. Res.* 14, 13266–13273. doi: 10.4238/2015.October.26.23

Li, J., Wen, J., Lease, K. ,. A., Doke, J. ,. T., Tax, F. ,. E., and Walker, J. ,. C. (2002). BAK1, an arabidopsis LRR receptor-like protein kinase, interacts with BRI1 and modulates brassinosteroid signaling. *Cell* 110 (2), 213–222. doi: 10.1016/s0092-8674(02)00812-7

Luo, X., Xu, L., Wang, Y., Dong, J., Chen, Y., Tang, M., et al. (2020). An ultra-high-density genetic map provides insights into genome synteny, recombination landscape and taproot skin colour in radish (*Raphanus sativus* l.). *Plant Biotechnol. J.* 18, 274–286. doi: 10.1111/pbi.13195

Lu, Q. Q., Yu, X. X., Yu, Z., Wang, H. T., Zhang, X., and Zhao, X. Q. (2022). Construction of ultra-high-density genetic linkage map of a sorghum-sudangrass hybrid using whole genome resequencing. *PloS One* 17 (11), e0278153. doi: 10.1371/journal.pone.0278153

Mace, E. S., Singh, V., Van, Oosterom., E. J., Hammer, G. L., Hunt, C. H., and Jordan, D. R. (2012). QTL for nodal root angle in sorghum (Sorghum bicolor l. moench) co-locate with QTL for traits associated with drought adaptation. *Theor. Appl. Genet.* 124, 97–109. doi: 10.1007/s00122-011-1690-9

Ma, L., Guan, Z., Zhang, Z., Zhang, X., Zhang, Y., Zou, C., et al. (2018). Identification of quantitative trait loci for leaf-related traits in an IBM Syn10 DH maize population across three environments. *Plant Breed.* 137, 127–138. doi: 10.1111/pbr.12566

Mahmoudzadeh, V. M., and Oad, R. (2018). Sorghum-sudangrass water productivity under subsurface drip irrigation. *Irrig Drain.* 67, 702–712. doi: 10.1002/ird.2278

McCouch, S. R., and Xiao, J. (1998). "From malthus to molecular mapping: prospects for the utilization of genome analysis to enhance the world food supply," in *Molecular dissection of complex traits*. Ed. A. H. Paterson (Boca Raton: FL: CRC Press), pp.267–pp.278.

Mir, N. H., Ahmad, S., and Bhat, S. S. (2020). Forage yield of orchard grass (Dactylis glomerata l.) under different fertility levels on a karewa upland of Kashmir himalaya. *IJCS* 8, 2444–2448. doi: 10.22271/chemi.2020.v8.i3ai.9575

Nave, R. L., Quinby, M. P., Griffith, A. P., Corbin, M. D., and Bates, G. E. (2020). Forage mass, nutritive value, and economic viability of cowpea overseeded in tall fescue and sorghum-sudangrass swards. *Crop Forage Turfgrass Manage.* 6, e20003. doi: 10.1002/cft2.20003

Ou, C., Jiang, S., Wang, F., Tang, C., and Hao, N. (2015). An RNA-seq analysis of the pear (Pyrus communis l.) transcriptome, with a focus on genes associated with dwarf. *Plant Gene* 4, 69–77. doi: 10.1016/j.plgene.2015.08.003

Peng, J. L., Kim, M., Kim, K., and Sung, K. (2020). Climatic suitability mapping and driving factors detection for whole crop maize and sorghum–sudangrass hybrid production in the south area of the Korean peninsula and jeju island. *Grassl. Sci.* 66, 207–214. doi: 10.1111/grs.12270

Rayaprolu, L., Selvanayagam, S., Rao, D. M., Gupta, R., Das, R. R., Rathore, A., et al. (2021). Genome-wide association study for major biofuel traits in sorghum using minicore collection. *Protein Pept. Lett.* 28, 909–928. doi: 10.2174/0929866528666210215141243

Renault, H., Bassard, J. E., Hamberger, B., and Werck-Reichhart, D. (2014). Cytochrome P450-mediated metabolic engineering: current progress and future challenges. *Curr. Opin. Plant Biol.* 19, 27–34. doi: 10.1016/j.pbi.2014.03.004

Saballos, A., Vermerris, W., Rivera, L., and Ejeta, G. (2008). Allelic association, chemical characterization and saccharification properties of brown midrib mutants of sorghum (Sorghum bicolor (L.) moench). *Bioenerg. Res.* 1, 193–204. doi: 10.1007/s12155-008-9025-7

Shi, Y., Yu, X. X., Yu, Z., Yang, D. S., Zhang, M. F., Li, Z. Z., et al. (2017). QTL mapping analysis of agronomic traits such as stem and leaf of sorghum-sudangrass hybrid. *Chinese J. Grassland* 9, 8–14. doi: 10.16742/j.zgcdxb.2017-04-02

Smulders, M. J. M., Arens, P., Bourke, P. M., Debener, T., Linde, M., Riek, J., et al. (2019). In the name of the rose: a roadmap for rose research in the genome era. *Hortic. Res.* 6, 948–964. doi: 10.1038/s41438-019-0156-0

Sun, X. W., Liu, D. Y., Zhang, X. F., Li, W. B., Liu, H., Hong, W. G., et al. (2013). SLAF-seq: an efficient method of large-scale denovo SNP discovery and genotyping using high-throughput sequencing. *PloS One* 8, e58700. doi: 10.1371/journal.pone.0058700

Waheed, R., Ignacio, J. C., Arbelaez, J. D., Juanillas, V. M., Asif, M., Henry, A., et al. (2021). Drought response QTLs in a super basmati ×Azucena population by high-density GBS-based SNP linkage mapping. *Plant Breed.* 140, 758–774. doi: 10.1111/pbr.12961

Wang, Y., Cheng, L., Sun, Y., Zhou, Z., Zhu, L., Xu, Z., et al. (2009). Genetic background effect on QTL expression of heading date and plant height and their interaction with environment in reciprocal introgression lines of rice. *Acta Agronom. Sin.* 35, 1386–1394. doi: 10.1016/S1875-2780(08)60095-4

Wang, L. H., Lu, Y. F., Lu, Y. J., Liu, Y. L., and Li, J. Q. (2021). QTL mapping of the agronomic traits using an RIL population of "T×623A×Sa". *J. Shanxi Agric. Sci.* 49 (12), 1491–1496. doi: 10.3969/j.issn.1002-2481.2021.12.18

Wang, X., Mace, E., Hunt, C., Cruickshank, A., Henzell, R., Parkes, H., et al. (2014). Two distinct classes of QTL determine rust resistance in sorghum. *BMC Plant Biol.* 14, 366. doi: 10.1186/s12870-014-0366-4

Wells, M. S., Martinson, K. L., Undersander, D. J., and Sheaffer, C. C. (2014). A survey investigating alfalfa winter injury in Minnesota and Wisconsin from the winter of 2012-2013. *Forage Grazinglands* 12, 1–7. doi: 10.2134/FG-2013-0051-RS

Wen, Y., Fang, Y., Hu, P., Tan, Y., Wang, Y., Hou, L., et al. (2020). Construction of a high-density genetic map based on SLAF markers and QTL analysis of leaf size in rice. *Front. Plant Sci.* 11. doi: 10.3389/fpls.2020.01143

West, M. A., Van, L. H., Kozik, A., Kliebenstein, D. J., Doerge, R. W., Clair, D. A. S., et al. (2006). High-density haplotyping with microarray-based expression and single feature polymorphism markers in arabidopsis. *Genome Res.* 16, 787–795. doi: 10.1101/gr.5011206

Xie, X., Jin, F., Song, M. H., Su, J. P., Huang, H. G., Kim, Y. G., et al. (2008). Fine mapping of a yield-enhancing QTL cluster associated with transgressive variation

in an oryza sativa x o. rufipogon cross. *Theor. Appl. Genet.* 116, 613–622. doi: 10.1007/s00122-007-0695-x

Xu, L. Y., Wang, L. Y., Wei, K., Tan, L. Q., Su, J. J., and Cheng, H. (2018). High-density SNP linkage map construction and QTL mapping for flavonoid-related traits in a tea plant (Camellia sinensis) using 2b-RAD sequencing. *BMC Genomics* 19, 1–11. doi: 10.1186/s12864-018-5291-8

Yang, P., Sun, X., Liu, X., Wang, W., Hao, Y., Chen, L., et al. (2022). Identification of candidate genes for lint percentage and fiber quality through QTL mapping and transcriptome analysis in an allotetraploid interspecific cotton CSSLs population. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.882051

Yang, D., Yu, X., Yu, Z., Li, J., Li, J., Lu, Q., et al. (2022). Quantitative trait locus mapping of five significant quality traits such as crude protein content of tetraploid hybrid crested wheatgrass in multiple environments. *Grassl. Sci.* 68, 134–144. doi: 10.1111/grs.12349

Yi, X., Hargett, S. ,. R., Liu, H., Frankel, L. ,. K., and Bricker, T. ,. M. (2007). The PsbP protein is required for photosystem II complex assembly/stability and photoautotrophy in arabidopsis thaliana. *J. Biol. Chem.* 282 (34), 24833–24841. doi: 10.1074/jbc.M705011200

Yu, X. X., Nan, Z. B., Yu, Z., Yang, D. S., Xie, R., Shi, Y., et al. (2018). Construction of genetic linkage map base on SSR markers and identification of QTLs for main traits of sorghum × sudangrass. *Mol. Plant Breed.* 16, 5347–5358. doi: 10.13271/j.mpb.016.005347

Zhang, J., Li, W., Lv, P., Yang, X., Xu, W., Ni, X., et al. (2021). Whole-genome resequencing and transcriptome analysis provide insights on aphid-resistant quantitative trait loci/genes in sorghum bicolor. *Plant Breed.* 140, 618–629. doi: 10.1111/pbr.12946

Zoschke, R., Liere, K., and Börner, T. (2007). From seedling to mature plant: Arabidopsis plastidial genome copy number, RNA accumulation and transcription are differentially regulated during leaf development. *Plant J.* 50 (4), 710–722. doi: 10.1111/j.1365-313X.2007.03084

Zou, G., Zhai, G., Feng, Q., Yan, S., Wang, A., Zhao, Q., et al. (2012). Identification of QTLs for eight agronomically important traits using an ultra-high-density map based on SNPs generated from high-throughput sequencing in sorghum under contrasting photoperiods. *J. Exp. Bot.* 63, 5451–5462. doi: 10.1093/jxb/ers205

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 20 January 2023
DOI 10.3389/fpls.2022.1097076

OPEN ACCESS

EDITED BY
Qiang Guo,
Beijing Academy of Agricultural and Forestry Sciences, China

REVIEWED BY
Yigong Zhang,
Xinjiang University, China
Pei Wang,
Southwest Minzu University, China

*CORRESPONDENCE
Guangxin Cui
cuiguangxin@caas.cn
Hongshan Yang
yanghongshan@caas.cn

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 13 November 2022
ACCEPTED 28 December 2022
PUBLISHED 20 January 2023

CITATION
Duan H, Tiika RJ, Tian F, Lu Y, Zhang Q, Hu Y, Cui G and Yang H (2023) Metabolomics analysis unveils important changes involved in the salt tolerance of *Salicornia europaea*. *Front. Plant Sci.* 13:1097076. doi: 10.3389/fpls.2022.1097076

Metabolomics analysis unveils important changes involved in the salt tolerance of *Salicornia europaea*

Huirong Duan[1], Richard John Tiika[1,2], Fuping Tian[1], Yuan Lu[1], Qian Zhang[1], Yu Hu[1], Guangxin Cui[1]* and Hongshan Yang[1]*

[1]Lanzhou Institute of Husbandry and Pharmaceutical Science, Chinese Academy of Agricultural Sciences, Lanzhou, China, [2]College of Forestry, Gansu Agricultural University, Lanzhou, China

Salicornia europaea is one of the world's salt-tolerant plant species and is recognized as a model plant for studying the metabolism and molecular mechanisms of halophytes under salinity. To investigate the metabolic responses to salinity stress in *S. europaea*, this study performed a widely targeted metabolomic analysis after analyzing the physiological characteristics of plants exposed to various NaCl treatments. *S. europaea* exhibited excellent salt tolerance and could withstand extremely high NaCl concentrations, while lower NaCl conditions (50 and 100 mM) significantly promoted growth by increasing tissue succulence and maintaining a relatively stable K^+ concentration. A total of 552 metabolites were detected, 500 of which were differently accumulated, mainly consisting of lipids, organic acids, saccharides, alcohols, amino acids, flavonoids, phenolic acids, and alkaloids. Sucrose, glucose, p-proline, quercetin and its derivatives, and kaempferol derivatives represented core metabolites that are responsive to salinity stress. Glycolysis, flavone and flavonol biosynthesis, and phenylpropanoid biosynthesis were considered as the most important pathways responsible for salt stress response by increasing the osmotic tolerance and antioxidant activities. The high accumulation of some saccharides, flavonoids, and phenolic acids under 50 mM NaCl compared with 300 mM NaCl might contribute to the improved salt tolerance under the 50 mM NaCl treatment. Furthermore, quercetin, quercetin derivatives, and kaempferol derivatives showed varied change patterns in the roots and shoots, while coumaric, caffeic, and ferulic acids increased significantly in the roots, implying that the coping strategies in the shoots and roots varied under salinity stress. These findings lay the foundation for further analysis of the mechanism underlying the response of *S. europaea* to salinity.

KEYWORDS

***Salicornia europaea*, salt tolerance, metabolomics, flavonoids, phenolic acids**

1 Introduction

Land salinization is one of the most serious abiotic stresses affecting modern agriculture worldwide, which limits the crop productivity and spatial distribution of plants (Morton et al., 2019; Gong et al., 2020). Nearly 10% of the land surface and 50% of irrigated land in the world are affected by salinized soils (Zhang et al., 2020a). The rising problem of global soil salinization and the serious losses in crop production call for a better understanding of the key mechanisms of salt tolerance in crops (Shabala and Mackay, 2011; Duan et al., 2015). An effective way of obtaining such information comes from studying halophytes. Halophytes possess remarkable abilities to tolerate and even benefit from the saline environment that may kill most other plant species (Flower et al., 2010; Shabala, 2013). Over the past few decades, although many important mechanisms for salt tolerance in halophytes have been intensively studied, breeding of crop plants with salt tolerance has not been very successful to date (Lu et al., 2020; Wang et al., 2020). Salt tolerance of plants is an intricate multigenic trait exhibiting heterosis, dominance, and additive effects and is also physiological omnifariously under the regulation of multiple tissue- and age-specific components (Shabala and Cuin, 2008). Therefore, salt tolerance of halophytes will be reflected by numerous subtraits, and comprehensive and systematic analyses focusing on salt tolerance should be conducted extensively.

The changes of metabolites reflect how an organism systemically and directly adapts to environmental changes (Weckwerth, 2003; Bundy et al., 2009). As a developing and promising methodology, metabolomics can analyze the metabolite changes qualitatively and quantitatively in organisms, thus revealing the relationship between plant species and environment and also showing the connection among phenotypes, metabolic networks, metabolic regulation, function, and plant growth (Schauer and Fernie, 2006; Töpfer et al., 2015). Plant metabolism is affected under salt stress, and plants need to adjust their metabolic levels to maintain basic metabolism and reach new homeostasis (Shulaev et al., 2008; Arbona et al., 2013). Consequently, metabolomics is the suitable tool for investigating plants' responses to salt stress. The metabolic responses of some plants to salt stress have already been examined, for example, *Suaeda salsa*, *Atriplex halimus*, and *Hordeum vulgare*, they provided new insights into the physiological mechanism of salt tolerance and nutritional value in halophytic species (Nemat Alla et al., 2012; Shelden et al., 2016; Li and Song, 2019).

Salicornia europaea, belonging to Amaranthaceae, is an annual euhalophyte widely distributed in coastal and inland salt marshes (Nie et al., 2015; Lv et al., 2020). As one of the most salt-tolerant plant species worldwide, the growth of *S. europaea* can be stimulated by 200 to 400 mM NaCl, and it can even withstand more than 1,000 mM NaCl (Lv et al., 2015). Thus, *S. europaea* is recognized as a model plant for studying the metabolism and molecular mechanisms of halophytes under salinity (Feng et al., 2015). A series of studies have been conducted in *S. europaea* under different salt treatments, which indicated that metabolic responses were caused under high salinity, including glycine betaine, proline, trigonelline, D-(+)-glucose, 2-propenyl (sinigrin), fructose-1-phosphate, and so on in the seedlings of *S. europaea* (Guy et al., 1984; Momonoki et al., 1994; Moghaieb et al., 2004; Wang et al., 2021). However, comprehensive studies on the metabolomics of *S. europaea* in a saline environment are limited, and how *S. europaea* adapts to the saline habitat by regulating the change of metabolites needs to be further deciphered.

In the present study, the effects on the growth and ion accumulation of *S. europaea* under salt treatments were compared. Moreover, the changes in the metabolic profiles of *S. europaea* under salinity were analyzed by widely targeted metabolomics. Our results comprehensively depict the characteristics of *S. europaea* metabolic response under salinity and provide a vigorous foundation for further understanding of the salt tolerance mechanisms and evaluating the nutritional value in *S. europaea*.

2 Materials and methods

2.1 Plant material and growth conditions

Seeds of *S. europaea* were collected from Liangcao Village of Jingtai County (37°21′2″, 104°5′28″), Gansu Province, China. Healthy seeds were rinsed with distilled water for three to five times after sterilization with 2% NaClO for 5 min. Thereafter, the seeds were germinated on a filter paper under 26°C in the dark. Uniform seedlings with a 2 to 3 cm radicle were transplanted into plastic pots (5 cm × 5 cm × 5 cm; one seedling/pot) containing sterilized sand. These seedlings were grown in an artificial climate chamber with a day/night temperature of 26°C/23°C. The daily photoperiod was 16/8 h (light/dark), with a light flux density of 600 μmol/m^2·s, and the relative humidity was 65%. The seedlings were divided into 10 groups and were watered with modified 1/2 strength Hoagland nutrient solutions (Tiika et al., 2021) supplemented with different NaCl concentrations (0, 50, 100, 200, 300, 400, 500, 600, 700, and 800 mM), respectively.

2.2 Growth index measurements

The seedlings were harvested 30 days after the imposition of different NaCl concentrations (0–400 mM) and then rinsed in distilled water. Four independent biological replicates of shoot and root samples were collected for this analysis. The fresh weight (FW) of separated root and shoot tissue, root length, and shoot height were measured. Thereafter, the samples were dried in an oven at 105°C for 10 mins before adjusting to 80°C for 72 h, and the dry weight (DW) of root and shoot tissue was determined. The degree of succulence was calculated by the following equation: succulence degree = FW of shoots / DW of shoots (Qi et al., 2009).

2.3 Measurement for Na^+ and K^+ concentrations in plants

Uniform seedlings were collected for the measurement of Na^+ and K^+ concentrations after treatment with additional NaCl for 30 days. The roots of the seedlings were rinsed twice (once for 4 min) in ice-cold 20 mM $CaCl_2$ to exchange cell wall-bound Na^+, and then the shoots were washed in deionized water to remove surface salts (Ma et al., 2017). Four independent biological replicates of shoot and root samples were harvested. Then, the tissues were dried at 80°C for 72 h to obtain dry weights. Na^+ and K^+ were extracted from the dried root

and shoot tissues in 100 mM acetic acid at 90°C for 2 h, and the ions were assayed using atomic absorption spectrophotometry (2655-00, Cole-Parmer Instrument Co., Vernon Hills, USA).

2.4 Sample preparation and extraction

Seedlings treated with 0 mM (CK), 50 mM (L), and 300 mM (H) NaCl for 30 days were rinsed in distilled water and then collected for widely targeted metabolic profiling analyses. Three independent biological replicates of shoot and root samples were collected under each NaCl condition. The root and shoot were freeze-dried and then crushed to powder using a mixer mill (MM 400, Retsch) with a zirconia bead at 30 Hz for 1.5 min (Li and Song, 2019). Furthermore, 100 mg powder was dissolved with 0.6 ml of 70% aqueous methanol, vortexing for 30 s every 30 min for six times, and was extracted overnight at 4°C (Sun et al., 2020). Following centrifugation at 10,000 *g* for 10 min, the extracts were absorbed and filtrated with a 0.22 µm pore size Nylon Syringe Filter (SCAA-104, anpel, Shanghai, China) before the UPLC–MS/MS analysis. Quality control (QC) samples were prepared by mixing sample extracts.

2.5 Data acquisition and quality control

The sample extracts were analyzed using an UPLC–ESI–MS/MS system (UPLC, Shim-pack UFLC SHIMADZU CBM30A system; MS, Applied Biosystems 4500 Q TRAP) from Metvare Biotechnology Co., Ltd., Wuhan, China. The analytical conditions were set according to the report of Li and Song (2019). To monitor the technical reproducibility, one QC sample was inserted into each of the 10 detected samples during the stability evaluation of the analysis conditions (Yi et al., 2021). Qualitative annotations of the metabolites were obtained from the Metware database together with the public databases. Quantifications of metabolites were conducted based on the information of retention time and peak pattern of metabolites and normalized by the R program (www.r-project.org/) (Xiao et al., 2021). Then, hierarchical cluster analysis was performed on the accumulation patterns of metabolites and visualized with a heat map.

2.6 Analysis of different metabolites

Multivariate statistical analysis combined with orthogonal partial least squares discriminant analysis and principal component analysis was conducted to analyze and verify the differences and reliability of metabolites in the samples. Different metabolites were screened by the thresholds with variable importance in projection (VIP) ≥ 1 and fold change (FC) ≥ 2 or ≤ 0.5. Then, the different metabolites were mapped to the Kyoto Encyclopedia of Genes and Genomes (KEGG) database and used in the significant enrichment analysis and major enriched pathways. To further analyze the change trend of relative content of metabolites, the relative contents of different metabolites were normalized and centralized, and then K-means clustering analysis was performed.

2.7 Statistical analysis

The physiological values were performed by ANOVA using SPSS statistical software (Version 16.0, SPSS Inc., Chicago, IL, USA). Duncan's multiple range tests were used to detect differences among means at a significance level of $p < 0.05$.

3 Results

3.1 Growth performance of *S. europaea* under salt stress

There were obvious differences in growth performances of *S. europaea* under different NaCl conditions (Figure 1). Severe salt stresses (≥ 300 mM NaCl) significantly reduced the growth of *S. europaea*, but *S. europaea* was still alive even under 800 mM NaCl. The fresh weights of shoots increased evidently and peaked at 50 and 100 mM NaCl, which were approximately 3.7- and 3.4-fold higher, respectively, than that under control condition (0 mM NaCl), and then reduced significantly (Figure 2A). The fresh weights of roots increased quickly with NaCl increasing and peaked at 100 mM NaCl, then decreased. The dry weights of shoots and roots exhibited similar

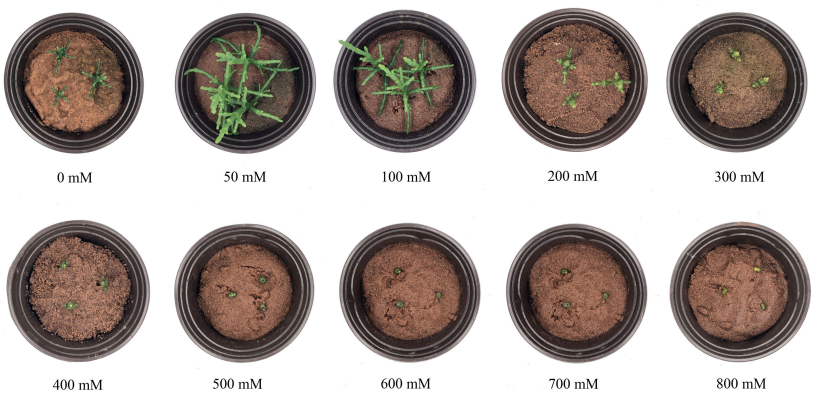


FIGURE 1
Growth performance of *Salicornia europaea* seedlings exposed to different levels of salt stress for 30 days.

10.3389/fpls.2022.1097076

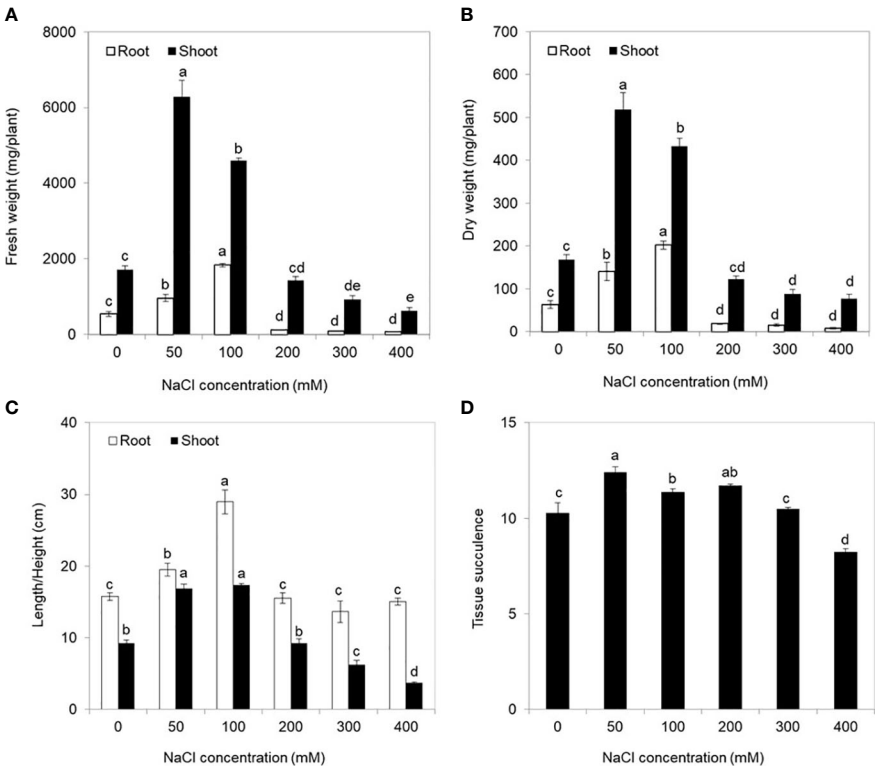


FIGURE 2
Physiological indices of *Salicornia europaea* seedlings exposed to different levels of salt stress for 30 days. **(A)** Fresh weight, **(B)** dry weight, **(C)** length/height, and **(D)** tissue succulence. Values are the means ± standard errors (SE) (n = 4), and bars indicate SE. Significant differences between columns in the roots and shoots are indicated by different lowercase letters (Duncan's test, $p < 0.05$).

change patterns (Figure 2B). Compared with the control, 50 and 100 mM NaCl promoted shoot height and root length significantly, and relative higher NaCl conditions (200–400 mM) reduced the shoot height evidently (Figure 2C). Besides, the tissue succulence of *S. europaea* increased firstly, and peaked at 50 mM NaCl, then decreased gradually with increasing NaCl concentrations (Figure 2D).

3.2 Ion accumulation of *S. europaea* under salt stress

Na^+ and K^+ were primarily distributed in the shoots after NaCl treatments as well as in the absence of salt (0 mM NaCl) (Figure 3). With NaCl application, the Na^+ concentration in roots and shoots was significantly higher than that in the control, and no significant difference was found among the 50–200 mM NaCl conditions and or between 300 and 400 mM NaCl conditions. Conversely, with the increase of salt stress, the K^+ concentration in both the shoots and roots exhibited declining levels and then remained stable from the 200 mM NaCl condition to the 400 mM NaCl condition.

3.3 Metabolite statistics of *S. europaea*

To investigate the metabolic changes of *S. europaea* adapted to different salt conditions, we carried out three different NaCl concentrations of *S. europaea* for extensive metabolic analysis using the UPLC-MS-based metabolomics approach, including 0, 50 mM NaCl (positively promoted concentration) and 300 mM NaCl (inflection point of inhibition concentration). As shown in Figures 4A, B, the three replicates of shoot and root from each treatment exhibited similar PC scores and were clustered together by Pearson's correlation analysis, revealing small variations among replicates. The root samples from different NaCl concentrations presented more discrete distribution than the shoot samples. After identification and analysis, a total of 552 metabolites were detected, and assigned to different classes, including 96 phenolic acids, 95 lipids, 81 flavonoids, 58 amino acids and their derivatives, 45 organic acids, 44 nucleotides and their derivatives, 30 alkaloids, 15 terpenoids, 12 lignans and coumarins, 4 tannins, 3 quinones, 3 steroids, and 66 others (Figure 4C; Supplementary Table S1).

3.4 Analysis and enrichment of the different metabolites of *S. europaea*

A total of 500 different metabolites were screened among the different NaCl conditions based on a VIP value ≥ 1 and FC ≥ 2 or ≤ 0.5. Based on the K-means analysis result, 12 classes that exhibited distinct clustering of different metabolite variations were separated, and class 4 contained the most different metabolites with a total of 68 (Figure 4D). In order to evaluate the significant variations of metabolites among different tissues and NaCl concentrations, comparisons were conducted among different groups of CK-R

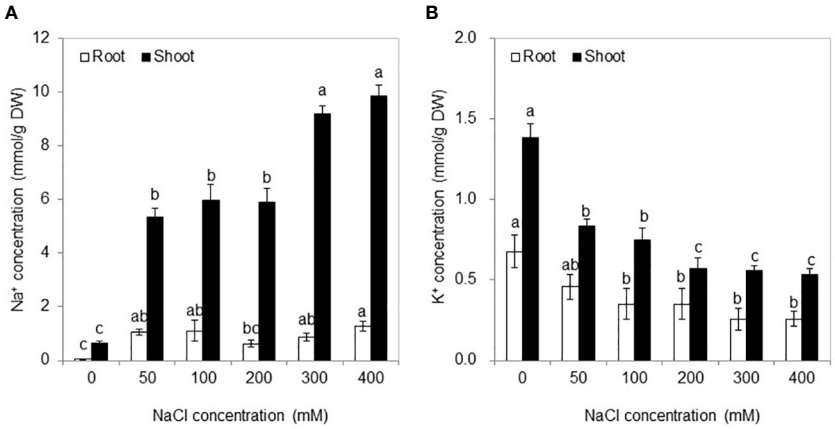


FIGURE 3
Na^+ **(A)** and K^+ **(B)** concentrations in the roots and shoots under different NaCl concentrations for 30 days. Values are means ± standard errors (SE) (n = 4), and bars indicate SE. Significant differences between columns in the roots and shoots are indicated by different lowercase letters (Duncan's test, $p < 0.05$).

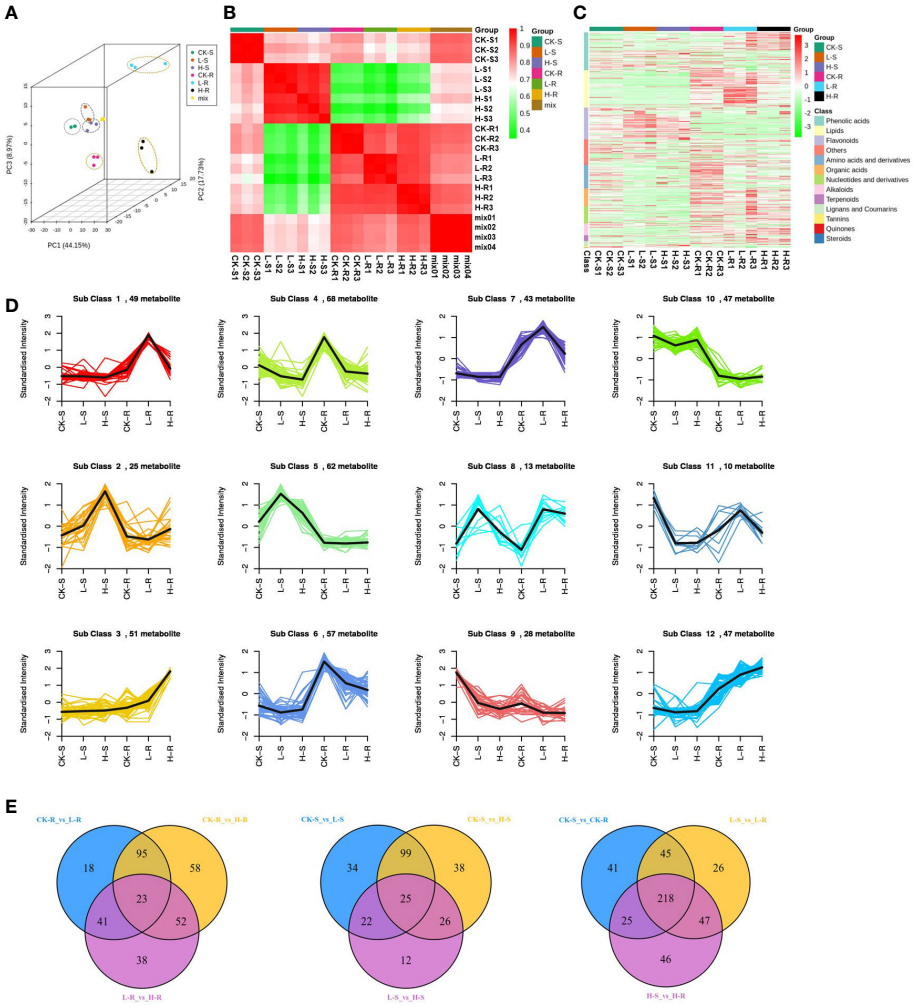


FIGURE 4
Dynamic metabolome of *Salicornia europaea* seedlings under different NaCl conditions. **(A)** Principal component analysis 3D plot of metabolomics data from three different NaCl treatments of the roots (R) and shoots (S). CK, 0 mM NaCl; L, 50 mM NaCl; H, 300 mM NaCl. "Mix" means the balanced mixture of all the samples (quality control). **(B)** Pearson correlation analysis of all the samples. "Mix" means the balanced mixture of all the samples (quality control). **(C)** Heat map of hierarchical clustering analysis of all detected metabolites in the roots and shoots among three different NaCl concentrations. Each sample is represented by a column, and each metabolite is represented by a row. **(D)** K-means cluster analysis of metabolites according to their variation tendencies. **(E)** Venn diagrams of different metabolites in multiple pairwise comparisons.

versus (*vs.*) L-R, L-R *vs.* H-R, CK-S *vs.* L-S, L-S *vs.* H-S, CK-R *vs.* CK-S, L-R *vs.* L-S, and H-R *vs.* H-S (Table 1). In the same plant tissue, the highest number of different metabolites was found in CK-R *vs.* H-R, including 112 upregulated and 116 downregulated metabolites, suggesting that the 300 mM NaCl condition significantly influenced metabolite accumulation in the roots. Under the same treatment, more abundant different metabolites accumulated in the roots of *S. europaea*, except for flavonoids which were more accumulated in the shoots. After searching the top 20 different metabolites in CK-S *vs.* CK-R, L-S *vs.* L-R, and H-S *vs.* H-R, lipids, flavonoids, and phenolic acids were the top three types of metabolites that accumulated differently in the shoots and roots. Additionally, venn diagrams were constructed to analyze the common different metabolites among all the groups compared. A total of 23 different metabolites were common in CK-R *vs.* L-R, CK-R *vs.* H-R, and L-R *vs.* H-R; 25 different metabolites were common in CK-S *vs.* L-S, CK-S *vs.* H-S, and L-S *vs.* H-S; and 218 different metabolites were common in CK-S *vs.* CK-R, L-S *vs.* L-R, and H-S *vs.* H-R (Figure 4E).

The different metabolites were mapped to the KEGG pathways to investigate the involved biological processes (Supplementary Table S2). The top four KEGG pathways with the highest number of different metabolites in most pairwise comparisons were involved in "metabolic pathways", "biosynthesis of secondary metabolites", "biosynthesis of amino acids", and "ABC transporters", except for the H-S *vs.* H-R group ("biosynthesis of amino acids" was not on the top four lists of the KEGG pathways). In the roots, compared with CK, the most enriched KEGG pathways (p-value ≤ 0.01) common in L and H contained "biosynthesis of amino acids" and "biosynthesis of secondary metabolites". In the shoots, compared with CK, the most enriched KEGG pathways (p-value ≤ 0.01) common in L and H contained "butanoate metabolism". In the comparisons of S *vs.* R, only "phenylpropanoid biosynthesis" was significantly enriched (p-value ≤ 0.01) in H-S *vs.* H-R.

3.5 Profiles of primary metabolites under salinity

Lipids were identified as one of the most different primary metabolites in *S. europaea* seedlings under salt stress, and 60 and 26 different lipids were accumulated in the roots and shoots, respectively (Supplementary Table S3). With the increase of NaCl concentrations, 40 different lipids presented increasing accumulation levels in L-R (for example, lysoPE 15:1 was accumulated 81.9-fold higher in L-R than CK-R) and then rapidly decreased in H-R. The majority of these lipids were compounds of lysophosphatidylethanolamine (LPE) and lysophosphatidylcholine (LPC). Sixteen different lipids showed decreasing trends, and four different lipids exhibited an increasing accumulation. In the shoots, except for lysoPC 20:0 and lysoPC 17:1, the accumulation of 24 different lipids was reduced under the salinity treatments compared with the control, especially MAG(18:4) isomer3 and 2-linoleoylglycerol-diglucoside, which were not detected under the salinity treatments.

Amino acids and their derivatives were the second most abundant different primary metabolites in *S. europaea* seedlings under salt stress. In total, 35 and 32 different amino acids and their derivatives were detected in the roots and shoots, respectively (Supplementary Table S3). The different amino acids and their derivatives in the roots showed varied change patterns with increasing NaCl concentrations, including six members that increased continuously (the reduced form of glutathione was only detected in H-R), 10 members that evidently declined in L-R and then slightly increased in H-R, and 19 members that decreased significantly (alanylleucine was only detected in CK-R). To better understand the variation in individual amino acids, KEGG enrichment analysis was performed, and it was found that "biosynthesis of amino acids", "lysine biosynthesis", and "aminoacyl-tRNA biosynthesis" were the most enriched pathways in CK-R *vs.* L-R and CK-R *vs.* H-R. The amino acids related to these pathways declined drastically in both roots and shoots as NaCl increased (Figure 5A). In contrast, some amino acids such as L-proline and L-cysteine exhibited upward trends as the salinity increased. The amounts of L-proline and L-cysteine were increased by 1.46- and 2.11-fold under L-R compared with CK-R, respectively, and increased by 4.7- and 33.6-fold under H-R compared with CK-R, respectively. In the shoots, except for L-proline and L-cysteine which increased, other amino acids and their derivatives declined in L-S. Eleven members then increased slightly in H-S (only the contents of three members were higher in H-S than CK-S) and 20 members decreased continuously.

A total of 16 and 21 "saccharides and alcohols" were differently accumulated in the roots and shoots of *S. europaea* seedlings, respectively (Supplementary Table S3). Compared with CK, seven members showed a drastically high accumulation in L-R, H-R, and L-S, including D-glucose, sedoheptulose, isomaltulose, melibiose, D-(+)-sucrose, galactinol, and turanose. In particular, the content of turanose was 67.8- and 58.1-fold higher in L-R and H-R, respectively, than in CK-R. Compared with CK, D-(-)-arabinose accumulated significantly in L-R and L-S. In addition, the KEGG pathway of "galactose metabolism" was the most enriched in CK-S *vs.* L-S (Supplementary Table S2), and six metabolites were related to this pathway, including D-glucose, inositol, D-fructose 6-phosphate, melibiose, D-(+)-sucrose, and galactinol. Except for D-fructose 6-phosphate, the other metabolites were drastically elevated in L-S compared with CK-S (Figure 5B).

In total, 31 and 31 organic acids were differently accumulated in the roots and shoots of *S. europaea* seedlings, respectively (Supplementary Table S3). In the roots, four different change patterns emerged, including 12 members with increasing levels in L-R that then decreased in H-R (3-methyl-2-oxobutanoic acid was not detected in CK-R), 12 members that declined continuously, five members that gradually increased (2-furanoic acid was not detected in CK-R), and two members that declined in L-R and then slightly increased in H-R. In the shoots, the different organic acids also showed four different change patterns, in which 23 members declined gradually with increasing salinity. In particular, the KEGG pathway of "butanoate metabolism" was the most enriched (p-value ≤ 0.01) in CK-S *vs.* L-S, and the KEGG pathways of "butanoate metabolism" and "carbon metabolism" were the most enriched (p-value ≤ 0.01) in CK-S *vs.* H-S. The two pathways covered γ-aminobutyric acid, succinic acid, fumaric acid, L-(-)-malic acid, 3-hydroxypropanoic acid, citric acid, and gluconic acid. All these organic acids were decreased dramatically with increasing salinity compared with CK (Figure 5C).

10.3389/fpls.2022.1097076

TABLE 1 Comparison of different metabolites in the shoots and roots of *Salicornia europaea* under salt stress.

Group class	CK-R vs L-R			L-R vs H-R			CK-R vs H-R			CK-S vs L-S			L-S vs H-S			CK-S vs H-S			CK-S vs CK-R			L-S vs L-R			H-S vs H-R		
	Upregulated	Downregulated		Upregulated	Downregulated		Upregulated	Downregulated		Upregulated	Downregulated		Upregulated	Downregulated		Upregulated	Downregulated		Upregulated	Downregulated		Upregulated	Downregulated		Upregulated	Downregulated	
Phenolic acids	19	15		13	6		31	16		17	8		3	12		16	12		32	26		31	24		37	25	
Lipids	33	6		2	41		5	17		0	23		5	2		1	16		64	9		77	7		63	6	
Flavonoids	3	8		23	3		16	8		24	4		5	8		15	4		5	40		4	47		4	48	
Amino acids and their derivatives	2	22		8	2		8	24		0	24		7	4		3	23		23	6		25	8		13	6	
Organic acids	6	7		1	14		7	14		2	20		2	7		3	27		16	5		20	1		26	1	
Nucleotide and their derivatives	5	5		5	3		15	6		4	12		3	2		5	15		16	8		21	4		25	4	
Alkaloids	6	5		7	0		7	5		2	4		3	1		3	6		14	5		12	3		16	4	
Terpenoids	1	0		2	4		4	3		0	1		0	0		0	4		7	3		5	3		9	3	
Lignans and coumarins	2	1		4	1		5	1		0	0		1	0		3	1		3	3		2	2		4	3	
Tannins	0	2		1	0		1	2		1	2		0	1		1	1		0	2		0	3		0	1	
Quinones	0	1		0	0		0	1		0	1		0	1		0	0		0	1		0	0		0	2	
Steroids	1	0		0	2		1	0		0	1		0	0		0	1		2	0		2	0		2	0	
Others	13	14		2	10		12	19		16	14		2	16		7	21		20	19		18	17		18	16	
Total	91	86	177	68	86	154	112	116	228	66	114	180	31	54	85	57	131	188	202	127	329	217	119	336	217	119	336

3.6 Profiles of flavonoids under salinity treatments

Flavonoids are important secondary metabolites in plants adapted to salinity. A total of 34 and 39 different flavonoids were accumulated in the roots and shoots, respectively, under the three salinity conditions. Most of these flavonoids were involved in the pathway of "flavone and flavonol biosynthesis" (ko00944) based on the KEGG enrichment analysis. The different flavonoids are shown in the heat map in Supplementary Figure S1.

In the roots, the content of kaempferol decreased significantly upon salt application, becoming almost undetectable in L-R and H-R. However, some derivatives of kaempferol showed increasing accumulation levels and peaked in H-R, for example, kaempferol-3-O-(6″-malonyl)-galactoside and kaempferol-3-O-(6″-malonyl)-glucoside (Figure 6). In contrast, quercetin and some of its derivatives increased continuously under salinity and peaked at H-R, including quercetin-3-O-(6″-O-acetyl)-galactoside and quercetin-3-O-(6″-O-malonyl)-galactoside. Four derivatives of isorhamnetin exhibited a similar change trend with a significant decrease in L-R and an increase in H-R, especially isorhamnetin-O-glucoside-O-malonyl-O-glucoside, which was undetectable in L-R but highly accumulated in H-R (FC value = 13.6).

In the shoots, naringenin chalcone and nepetin showed significantly increased levels under increasing salinity and peaked at L-S, with FC values of 6.0 and 12.2 compared with CK-S, respectively. It was found that kaempferol declined evidently with the application of salt stress; however, six of the nine derivatives of kaempferol continued to accumulate and peaked at L-S (Figure 6). Moreover, quercetin and eight of 11 of its derivatives increased continuously under increasing salinity and peaked at L-S, and quercetin-3-O-rhamnoside (quercitrin) had the highest FC value (13.1) at L-S. Three anthocyanins were differentially accumulated in the shoots under different salt conditions. Among them, delphinidin-3-O-glucoside (mirtillin) and delphinidin-3-O-rutinoside increased constantly and peaked at H-S, while delphinidin-3-O-arabinoside decreased evidently.

3.7 Profiles of phenolic acids under salinity treatments

Phenolic acids were important different secondary metabolites in *S. europaea* seedlings under salt stresses. In total, 62 and 40 different phenolic acids were accumulated in the roots and shoots of *S. europaea* seedlings under salt stress, respectively. Most of these metabolites belonged to the derivatives of hydroxybenzoic acid and hydroxycinnamic acid and were involved in "phenylpropanoid biosynthesis" (ko00940). The accumulation patterns of the representative members are illustrated in the heat map in Supplementary Figure S2.

In the roots, some different phenolic acids presented significantly increasing trends with increasing salinity, including p-coumaraldehyde, p-coumaryl alcohol, coniferaldehyde, sinapyl alcohol, sinapinaldehyde, and 3-O-p-coumaroyl quinic acid (Figure 7A). In particular, the FC values of p-coumaraldehyde and p-coumaryl alcohol at H-R were 32.02- and 68.31-fold higher than at CK-R, respectively. A total of

10.3389/fpls.2022.1097076

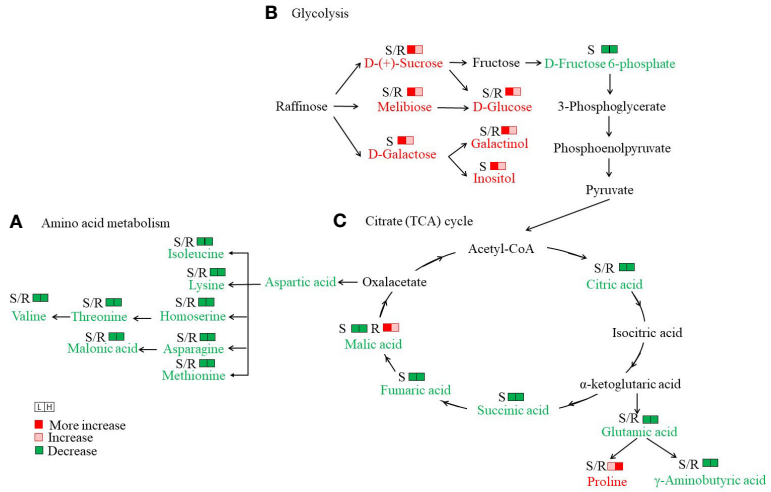


FIGURE 5
Analysis of important metabolic networks in the seedlings of *Salicornia europaea* under different NaCl conditions. **(A)** Amino acid metabolism, **(B)** Glycolysis, and **(C)** Citrate (TCA) cycle. S/R, shoot or root. The metabolites with a red color indicate an increase under both L (50 mM NaCl) and H (300 mM NaCl) conditions compared with CK (0 mM NaCl), while the metabolites with a green color indicate decreases under both L and H conditions compared with CK.

two hydroxybenzoic acids (protocatechuic acid and syringic acid) and five hydroxycinnamic acids (p-coumaric acid, caffeic acid, ferulic acid, sinapic acid, and chlorogenic acid) were detected with different change trends under varied NaCl conditions. Among them, protocatechuic acid, syringic acid, and sinapic acid showed continuously increasing levels with increasing salt concentrations. The contents of p-coumaric acid and ferulic acid under salinity treatments were significantly higher than CK-R, and peak values were found in L-R (FC values of 4.7 and 4.3, respectively). However, chlorogenic acid decreased quickly with increasing salinity.

In the shoots, we found the contents of some different phenolic acids increased under salinity and then peaked in L-S, including salidroside, coniferin, syringin, anthranilate O-hexosyl-O-hexoside, and so on, and their FC values were 2.3, 3.6, 5.0, and 9.6 than that in CK-S, respectively (Figure 7B). Differently, some different phenolic acids exhibited continuously increased levels with increasing salt conditions, including protocatechuic aldehyde, 2,5-dihydroxybenzoic acid, protocatechuic acid, salicylic acid, tyrosol, and protocatechuic acid among others.

3.8 Profiles of alkaloids under salinity

In total, 17 and 11 different alkaloids were accumulated in the roots and shoots of *S. europaea* seedlings under salt stress. The change patterns of these alkaloids under different NaCl conditions are visualized in the heat map shown in Figure 8. In the roots, six different alkaloids significantly decreased after NaCl application compared with CK-R, while their contents in L-R and H-R did not differ greatly. In contrast, 11 different alkaloids showed increasing accumulation under salinity compared with CK-R. Among these alkaloids, some obtained maximum values in L-R, such as betanin, 3-O-acetylhamayne, and grossamide, while some members accumulated the highest levels in H-R, for example, the content of "(5-8)-dihydroxy-1-(3,4-dihydroxyphenyl)-N2,N3-bis(4-hydroxyphenethyl)-1,2-dihydronaphthalene-2,3-dicarboxamide" and "3-{(2,3-trans)-2-(4-hydroxy-3-methoxyphenyl)-3-hydroxymethyl-2,3-dihydrobenzo[b][1,4]dioxin-6-yl}-N-(4-hydroxyphenethyl) acrylamide" was 55.3- and 53.3-fold higher in H-R compared with CK-R, respectively. In the shoots, there were two different change

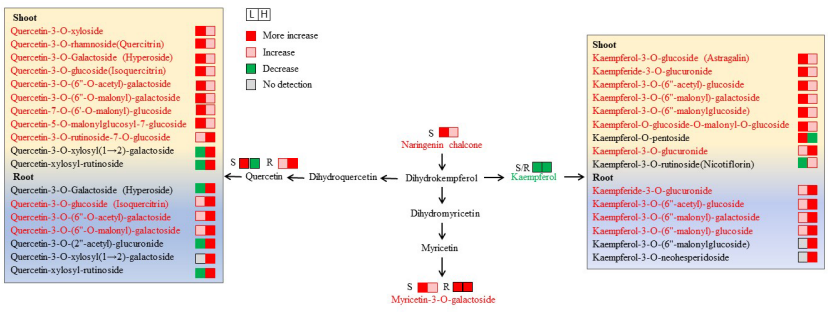


FIGURE 6
Changes in metabolites mapped to the metabolic pathway of flavone and flavonol biosynthesis. The metabolites with a red color indicate an increase under both L (50 mM NaCl) and H (300 mM NaCl) condition compared with CK (0 mM NaCl), while the metabolites with a green color indicate a decrease under both L and H condition compared with CK.

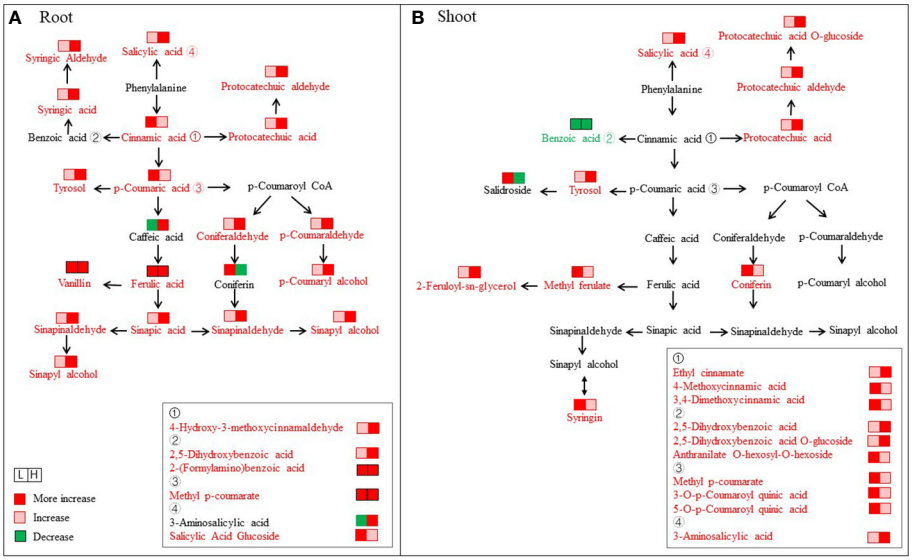


FIGURE 7
Changes in metabolites mapped to the metabolic pathway of phenylpropanoid biosynthesis. **(A)** Root. **(B)** Shoot. The metabolites shown in red color indicate increases under both L (50 mM NaCl) and H (300 mM NaCl) conditions compared with CK (0 mM NaCl), while the metabolites with a green color indicate a decrease under both L and H conditions compared with CK.

trends of the different alkaloids, except for betanin and lithospermoside. Seven members decreased significantly under increasing NaCl conditions, the contents of which differed little between L-S and H-S, such as synephrine and lumichrome. Three members showed increasing trends with increasing NaCl concentrations, especially "(5-8)-dihydroxy-1-(3,4-dihydroxy phenyl)-N2,N3-bis(4-hydroxyphenethyl)-1,2-dihydronaphthalene-2,3-dicarboxamide" (the FC value in H-S was 28.2).

4 Discussion

4.1 *S. europaea* is suitable for studying metabolic change responses to salt stress

Salinity is a major abiotic stress affecting the productivity of important economic crops worldwide (Negrão et al., 2017). Under salt stress, the comprehensive metabolic responses of halophytes are altered. This metabolic reprogramming offers an opportunity to identify important salt-responsive compounds (Wang et al., 2019b). In the present study, under the application of different NaCl concentrations, *S. europaea* exhibited excellent performance accompanied by increased tissue succulence and relatively stable K^+ levels, providing a model halophyte for salt tolerance research. Stress defense and active growth suppression are two complementary strategies for plants in adverse environments (Zhang et al., 2020b). *S. europaea* was able to survive under a very high NaCl concentration (800 mM) accompanied by evident growth inhibition (Figure 1), indicating that *S. europaea* positively regulated survival under high salinity by actively suppressing growth and resetting the balance between salt resistance and growth. The mechanism of this beneficial survival strategy has yet to be defined. In addition, in an intriguing twist, the results showed that lower NaCl concentrations were beneficial for *S. europaea* to promote growth, while higher NaCl concentrations directly inhibited growth. However, the diversity of metabolites and their functional characterization among low- and high-NaCl treatments have not yet been compared in *S. europaea*. In the present investigation, an abundant and diverse set of metabolites was identified in the roots and shoots of *S. europaea* exposed to various NaCl conditions, including primary metabolites and secondary metabolites.

4.2 Lipid composition is involved in the salt resistance of *S. europaea*

Salt stress often impairs the plasma membrane, and thus changes in membrane composition and properties are crucial in imparting salt tolerance among several higher plants (Mansour et al., 2015). Lipids are important in maintaining the stability and fluidity of the membrane, and regulatory roles are affected by changes in lipid composition (Gogna et al., 2020). Previous studies reported that the maintenance of membrane fluidity was regulated by the levels of unsaturated fatty acids (Kerkeb et al., 2001; Turk et al., 2004). In the roots of *S. europaea* subjected to salt stress, two unsaturated fatty acids namely 9,12,13-trihyroxy-10,15-octadecadienoic acid and 9,10,13-trihyroxy-11-octadecadienoic acid increased significantly. The major compounds of LPE and LPC were evidently accumulated in L-R. LPE and LPC are involved in the acyl modification of fatty acid unsaturation, and LPE may protect membrane integrity by inhibiting membrane lipid degradation (Ryu et al., 1997). Hence, the increased content of important unsaturated fatty acids and the regulation of LPE and LPC might be involved in the salt resistance of *S. europaea*. In addition, different variations were detected in lipids in the roots and shoots exposed to salinity; unlike lysoPC 20:0 and lysoPC 17:1, which increased significantly, the other different lipids decreased in the shoots under salt stress, implying that

major changes in lipid composition might not be necessary to maintain plasma membrane fluidity in the shoots. It was concluded that lipid composition varied in different tissues, which enabled *S. europaea* to adapt so successfully to external NaCl treatments.

4.3 Salinity stress disturbs another type of primary metabolism in *S. europaea*

Plant metabolism is significantly perturbed under abiotic stress (Sanchez et al., 2008; Kumari et al., 2015). In the present study, with the exception of lipids, many different primary metabolites were found as belonging to the classes of "amino acids and their derivatives", "saccharides and alcohols", and organic acids. It is well known that amino acids and their derivatives regulate plant defense against abiotic stress by maintaining the osmotic balance and the stability of the cell membrane (Khan et al., 2019). The previous studies of Shelden et al. (2016) and Li and Song (2019) revealed a significant increase in certain amino acids (*i.e.*, proline) under salt stress in barley and *S. salsa*. In accordance with these previous findings, in the present study, it was found that L-proline and L-cysteine accumulated continuously with increasing NaCl conditions both in the roots and shoots of *S. europaea*. As an osmotic regulator substance, proline enhances cell tolerance and reduces cell damage from various abiotic stresses (Zhao et al., 2021). Thus, proline may be an important osmotic regulator in *S. europaea* subjected to salt stress. In contrast, the levels of some amino acids in both the roots and shoots decreased significantly under NaCl treatments compared with the control (Figure 5A), and the KEGG pathways of "lysine biosynthesis" and "aminoacyl-tRNA biosynthesis" were highly enriched compared with the other pathways. Studies have shown that amino acids can function as the precursors of secondary metabolites that can protect plants from various stresses (Zhang et al., 2018). Thus, it was concluded that these decreased amino acids might be degraded or participate in the synthesis of other metabolites for regulating amino acid metabolism and deploying defense responses to salinity stress.

Saccharides can function as compatible osmolytes for osmotic adjustment and detoxification (Li and Song, 2019). As a primary adaptive strategy in response to stress, the biosynthesis and accumulation of some soluble sugars are often induced under salinity (Jia et al., 2019). Wang et al. (2020) stated that some soluble sugars including glucosamine, maltose, and D-(+)-sucrose increased significantly in the leaves of *S. salsa* under salt stress. A high accumulation of fructose and glucose in the halophyte *Limonium albuferae* was found under high-salinity conditions, and they were considered as major osmolytes in *Limonium* genus (González-Orenga et al., 2019). Similarly, our data showed that salt treatment upregulated the sucrose and glucose contents, indicating that they might be the main osmotic adjustment substances involved in *S. europaea*'s resistance to salt stress. In addition, it was found that the saccharide metabolites associated with glycolysis increased significantly under salinity under L-S compared with CK (Figure 5B). As an important process of energy production, the increase of glycolysis metabolites could be linked to the enhanced production of reactive oxygen species (ROS) (Guo et al., 2022). Thus, an elevated accumulation of glycolysis metabolites might be coupled with the generation of ROS in *S. europaea* especially.

Organic acids are osmotic regulator substances in plants and play important roles in the physiological process of salt resistance (Sanchez et al., 2008). In the present study, it was found that salinity induced the increase of certain kinds of organic acids in both the roots and shoots of *S. europaea*, especially 3-methyl-2-oxobutanoic acid, trans, trans-muconic acid, and 2,3-dihydroxybenzoic acid (Supplementary Table S3). Hence, these organic acids might accumulate more contents to decrease the osmotic potential and help *S. europaea* deal with salt stress. In addition, the citrate (TCA) cycle is an important pathway to release more energy and accelerate the physiological metabolic reaction against stress (Zhang et al., 2016). A detailed analysis showed that some organic acids that were TCA cycle intermediates decreased to a marked degree in the shoots of *S. europaea*, and their contents were lower in H than in L (Figure 5C). Similar observations have been reported in the halophytes *Thellungiella halophila* and *Limonium latifolium* in response to salt stress, suggesting that many changes in organic solute composition are predominantly controlled by constitutive developmental programs (Gong et al., 2005; Gagneul et al., 2007). The decrease of these metabolites might be accompanied by the suppression of energy production, implying that *S. europaea* adopted an energy conservation strategy from plant growth to the induction of protective mechanisms under salt stress; meanwhile, the inhibited TCA cycle flux might partially counteract ROS overproduction associated with glycolysis as a response of *S. europaea* to salt (Guo et al., 2022).

4.4 Flavonoids and phenolic acids positively dominate the salt resistance of *S. europaea*

When plants are exposed to a saline environment, ROS overgeneration disrupts the balance between ROS production and scavenging, inducing the oxidative damage of cells (Hao et al., 2021). Halophytes evolve a multifaceted antioxidant defense network to reduce ROS overgeneration under abiotic stress (Şirin and Aslım, 2019; Li et al., 2022). The phenylpropanoid pathway involves the biogenesis of diverse phenolic polymers, such as flavonoids and phenolic acids, which play important roles of ROS scavenging in plant defense against abiotic stresses (Liu et al., 2020; Li et al., 2021a). In the present study, KEGG analyses revealed that the pathways of "phenylpropanoid biosynthesis" and "flavone and flavonol biosynthesis" were dramatically enriched. Our results were consistent with previous reports (Li and Song, 2019; Wang et al., 2019a). For example, in peanut response to salt stress, "phenylpropanoid biosynthesis" and "flavonoids biosynthesis" were also enriched, and a dramatic accumulation of flavonoids and phenolic compounds was found, revealing that the abundances of these metabolites were regulated by salinity or salinity recovery specifically (Cui et al., 2018).

Flavonoids are well-known secondary metabolites produced when plants encounter abiotic stress, including flavone, flavonol, and flavonoid (Górniak et al., 2019; Wang et al., 2019a). Flavonoids occur in plants as glycosides and commonly function by increasing the water solubility to protect their reactive groups from oxidation by free radicals (Agati and Tattini, 2010). In addition, glycosylated flavonoids can be further acylated by aliphatic or aromatic acetyl

groups (Zhang et al., 2018). In the current study, glycosylated flavonoids, malonylated flavonoids, and flavonoids with acetyl modifications were found in the seedlings of *S. europaea* (Supplementary Figure S1). In particular, it was found that these flavonoids displayed distinct variation patterns under L and H and in both the shoots and roots (Figure 6), implying that flavonoids with different substituent groups might play different roles in different tissues of *S. europaea* in response to salinity. Furthermore, flavonoids with different chemical structures possess different levels of antioxidant activity. Researchers have reported that ortho-dihydroxylated B-ring flavonoids increase in plants under stress and function as important potential ROS scavengers, including quercetin and its derivatives (Hernández et al., 2009). Quercetin is an anti-oxidative flavonoid with a wide distribution throughout plants and serves as a promising candidate for cancer prevention (Qiu et al., 2010; Li et al., 2019). Kaempferol-3-O-glucoside is a bioactive flavonoid with antioxidant, anti-inflammatory, and protective characteristics (Kashyap et al., 2017). The significantly increased kaempferol derivatives, quercetin, and quercetin derivatives detected in the present study might scavenge the ROS in the shoots and roots to protect the plant from oxidative stress caused by increased salinity. Interestingly, most of these compounds were highly accumulated under low NaCl conditions in the shoots, while they were highly accumulated under high NaCl conditions in the roots. The results indicated that coping strategies in the shoots and roots under increasing salinity varied by modulating flavonoid accumulation with different substituent groups. However, the exact roles that these compounds play in salinity stress response remain to be investigated.

As another major category with antioxidant properties, phenolic acids are generally distributed in plant cell walls (Pradeep and Sreerama, 2018; Li et al., 2021b). In the present study, it was found that some of the phenolic acids were significantly higher in the seedlings of *S. europaea* both in the shoots and roots under the NaCl treatments, and the contents of these phenolic acids increased consistently with increased NaCl concentrations, including protocatechuic acid, protocatechuic aldehyde, salicylic acid, and tyrosol (Figure 7, S2). Among them, protocatechuic acid is a natural phenolic acid with biologically active components such as anti-aging, anti-atherosclerotic, and antibacterial activities (Wang et al., 2020). Similarly, as a plant endogenous signal molecule, salicylic acid possesses the ability to improve the metabolic activities and reduce the inhibition effects of plant growth under salinity (Bai et al., 2009). Changes in the level of salicylic acid were observed in the salt-tolerant wild soybean compared with the common wild soybean (Jiao et al., 2018). Besides, *p*-coumaric acid is a chain-breaking antioxidant by radical scavenging activity related to its electron- or hydrogen-donating ability and to the capacity to delocalize/stabilize the resulting phenoxyl radical within its structure (Pei et al., 2016). It was also found that the content of *p*-coumaric acid increased significantly in the roots of *S. europaea* under salt stress. We speculated that the increased contents of these phenolic acids were relevant to the antioxidant content and antioxidant activity, protecting *S. europaea* against oxidative stress due to salinity. Interestingly, the relative contents of some phenolic acids in the shoots and roots were significantly different due to different NaCl treatments (Figure 7). Generally, for the phenolic acids participating in the process of phenylpropanoid biosynthesis, most were highly accumulated in the roots and increased with increasing NaCl concentration. Only a small number of phenolic acids increased significantly downstream of the pathway, including syringin, coniferin, and tyrosol. Plant root systems are exposed directly to saline environments and must regulate metabolic processes firstly to maintain life activities (Li et al., 2017). It was speculated that these upregulated phenolic acids in the process of phenylpropanoid biosynthesis reflected the different defense effects of the roots compared with the shoots. Furthermore, suberin is a complex

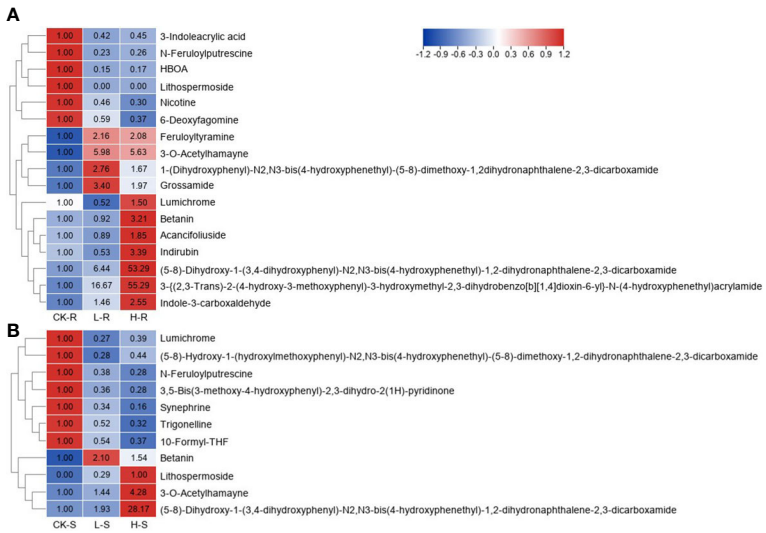


FIGURE 8
Heat map of different alkaloids among three different NaCl conditions. CK, 0 mM NaCl; L, 50 mM NaCl; H, 300 mM NaCl. **(A)** Heat map in the root (R) tissue. **(B)** Heat map in the shoot (S) tissue. Different metabolites were selected based on projection (VIP) >1 and fold change (FC) ≥2 or ≤0.5 in any of the comparison groups in S and R. The FC values are displayed in the heat map which uses the data of CK-R or CK-S as the calibrator, except for lithospermoside (no detection in the CK-S, and thus H-S is used as the calibrator).

polyester of aliphatics and phenolics deposited in the cell walls that creates a hydrophobic barrier to control the movements of water and solutes through cell walls, thus playing important roles in the salt adaptation of halophytes (Cui et al., 2021). Coumaric, caffeic, and ferulic acids are revolved in the suberin biosynthetic pathway, and they are linked to fatty alcohols by BAHD-type acyltransferases to produce alkyl hydroxycinnamates found in suberin-associated waxes (Vishwanath et al., 2015). Thus, the significant accumulation of coumaric, caffeic, and ferulic acids in the roots might contribute to suberin biosynthesis in *S. europaea*, thereby further helping to reduce the uncontrolled transport of water and dissolved ions (*e.g.*, excess Na^+) through the apoplast and transcellularly.

As a class of basic organic compounds in nature, alkaloids exhibit strong antioxidant and anti-tumor activities (Li and Song, 2019). In the present study, some alkaloids were significantly upregulated in the seedlings of *S. europaea*, including betanin, 3-O-acetylhamayne, and "(5-8)-dihydroxy-1-(3,4-dihydroxyphenyl)-N2,N3-bis(4-hydroxyphenethyl)-1,2-dihydronaphthalene-2,3-dicarboxamide" (Figure 8). The results implied that alkaloids might also play key roles in *S. europaea* under salt stress, which needs to be further validated.

5 Conclusion

The present study analyzed the physiological changes of *S. europaea* under different NaCl concentrations and unraveled the metabolic differences using widely targeted metabolomic analysis. *S. europaea* exhibited excellent performance under a variety of NaCl conditions, and moderate NaCl concentrations could promote the growth of the plant. The current investigation revealed that the biosynthesis of some key primary metabolites, including lipids, amino acids, soluble sugars and organic acids, was significantly enhanced under NaCl treatment. The significant changes in the contents of secondary metabolites corresponded well to the changes of saccharides, flavonoids, and phenolics acids, which were highly associated with the glycolysis pathway, phenylpropanoid biosynthesis, and flavonoid biosynthesis. These results provide new perspectives on important metabolites and their crucial contributions to the salt tolerance of *S. europaea*.

Data availability statement

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

Author contributions

GC and HY designed and supervised the project. HD performed bioinformatic analysis and drafted the manuscript. RT and FT performed the experiments and participated in the bioinformatic analysis. YL and QZ collected the samples, maintained the growth of the experiment materials, and measured the physiological values. YH revised the manuscript and participated in the interpretation of the data. GC and HY contributed equally as corresponding authors. All authors contributed to the article and approved the submitted version.

Funding

This research was funded by the Natural Science Foundation of China (31700338), Central Public-interest Scientific Institution Basal Research Fund (1610322022016), Science and Technology Innovation Program of Lanzhou Institute of Husbandry and Pharmaceutical Science, Chinese Academy of Agricultural Sciences (CAAS-LMY-04), Lanzhou Science and Technology Plan Project (2021-1-172 and 2019-RC-59), and Science and Technology Project of Gansu Province (LCJ20210023 and 21JR7RA036).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2022.1097076/full#supplementary-material

SUPPLEMENTARY FIGURE 1
Heat map of different metabolites of flavonoids among three different NaCl conditions. CK, 0 mM NaCl; L, 50 mM NaCl; H, 300 mM NaCl. **(A)** Heat map in the root (R) tissue. **(B)** Heat map in the shoot (S) tissue. Different metabolites were selected based on projection (VIP) >1 and fold change (FC) ≥2 or ≤0.5 in any of the comparison groups in S and R. The FC values are displayed in the heat map which uses the data of CK-R or CK-S as the calibrator, except for kaempferide-3-O-glucuronide and kaempferol-3-O-(6″-malonylglucoside) (no detection in the CK-R, and thus H-R is used as the calibrator).

SUPPLEMENTARY FIGURE 2
Heat map of different phenolic acids among three different NaCl conditions. CK, 0 mM NaCl; L, 50 mM NaCl; H, 300 mM NaCl. **(A)** Heat map in the root (R) tissue. **(B)** Heat map in the shoot (S) tissue. Different metabolites were selected based on projection (VIP) >1 and fold change (FC) ≥2 or ≤0.5 in any of the comparison groups in S and R. The FC values are displayed in the heat map which uses the data of CK-R or CK-S as the calibrator, except for terephthalic acid, pyrocatechol (no detection in the CK-R, and thus H-R is used as the calibrator), and methyl p-coumarate (H-S is used as the calibrator).

SUPPLEMENTARY TABLE 1
Metabolites detected in *Salicornia europaea* seedlings.

SUPPLEMENTARY TABLE 2
Kyoto Encyclopedia of Genes and Genomes pathways of different metabolites among different comparison groups.

SUPPLEMENTARY TABLE 3
Classification of primary metabolites detected in the roots and shoots of *Salicornia europaea* seedlings under salt stress.

References

Agati, G., and Tattini, M. (2010). Multiple functional roles of flavonoids in photoprotection. *New Phytol.* 186, 786–793. doi: 10.1111/j.1469-8137.2010.03269.x

Arbona, V., Manzi, M., Ollas, C. D., and Gómez-Cadenas, A. (2013). Metabolomics as a tool to investigate abiotic stress tolerance in plants. *Int. J. Mol. Sci.* 14, 4885–4911. doi: 10.3390/ijms14034885

Bai, T., Li, C. Y., Ma, F. W., Shu, H. R., and Han, M. Y. (2009). Exogenous salicylic acid alleviates growth inhibition and oxidative stress induced by hypoxia stress in *Malus robusta* rehd. *Plant Growth Regul.* 28, 358–366. doi: 10.1007/s00344-009-9104-9

Bundy, J. G., Davey, M. P., and Viant, M. R. (2009). Environmental metabolomics: a critical review and future perspectives. *Metabolomics* 5, 3–21. doi: 10.1007/s11306-008-0152-0

Cui, B., Liu, R. R., Flowers, T. J., and Song, J. (2021). Casparian bands and suberin lamellae: Key targets for breeding salt tolerant crops? *Environ. Exp. Bot.* 191, 104600. doi: 10.1016/j.envexpbot.2021.104600

Cui, F., Sui, N., Duan, G. Y., Liu, Y. Y., Han, Y., Liu, S. S., et al. (2018). Identification of metabolites and transcripts involved in salt stress and recovery in peanut. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00217

Duan, H. R., Ma, Q., Zhang, J. L., Hu, J., Bao, A. K., Wei, L., et al. (2015). The inward-rectifying k^+ channel SsAKT1 is a candidate involved in k^+ uptake in the halophyte *Suaeda salsa* under saline condition. *Plant Soil* 395, 173–187. doi: 10.1007/s11104-015-2539-9

Feng, J. J., Wang, J. H., Fan, P. X., Jia, W. T., Nie, L. L., Jiang, P., et al. (2015). High-throughput deep sequencing reveals that microRNAs play important roles in salt tolerance of euhalophyte *salicornia europaea*. *BMC Plant Biol.* 15, 63. doi: 10.1186/s12870-015-0451-3

Flowers, T. J., Galal, H. K., and Bromham, L. (2010). Evolution of halophytes: Multiple origins of salt tolerance in land plants. *Funct. Plant Biol.* 37, 604–612. doi: 10.1071/FP09269

Gagneul, D., Aïnouche, A., Duhazé, C., Lugan, R., Lahrer, F. R., and Bouchereau, A. (2007). A reassessment of the function of the so-called compatible solutes in the halophytic plumbginaceae *Limonium latifolium*. *Plant Physiol.* 144, 1598–1611. doi: 10.1104/pp.107.099820

Gogna, M., Choudhary, A., Mishra, G., Kapoor, R., and Bhatla, S. C. (2020). Changes in lipid composition in response to salt stress and its possible interaction with intracellular na^+-k^+ ratio in sunflower (*Helianthus annuus* l.). *Environ. Exp. Bot.* 178, 104147. doi: 10.1016/j.envexpbot.2020.104147

Gong, Q. Q., Li, P. H., Ma, S. S., Rupassara, S. I., and Bohnert, H. J. (2005). Salinity stress adaptation competence in the exptremophile *Thellungiella halophila* in comparison with its relative *Arabidopsis thaliana*. *Plant J.* 44, 826–839. doi: 10.1111/j.1365-313X.2005.02587.x

Gong, Z. Z., Xiong, L. M., Shi, H. Z., Yang, S. H., Herrera-Estrella, L. R., Xu, G. H., et al. (2020). Plant abiotic stress response and nutrient use efficiency. *Sci. China Life Sci.* 63, 635–674. doi: 10.1007/s11427-020-1683-x

González-Orenga, S., Ferrer-Gallego, P. P., Laguna, E., López-Gresa, M. P., Donat-Torres, M. P., Verdeguer, M., et al. (2019). Insights on salt tolerance of two endemic *Limonium* species from Spain. *Metabolites* 9, 294. doi: 10.3390/metabo9120294

Górniak, I., Bartoszewski, R., and Króliczewski, J. (2019). Comprehensive review of antimicrobial activities of plant flavonoids. *Phytochem. Rev.* 18, 241–272. doi: 10.1007/s11101-018-9591-z

Guo, Q., Han, J. W., Li, C., Hou, X. C., Zhao, C. Q., Wang, Q. H., et al. (2022). Defining key metabolic roles in osmotic adjustment and ROS homeostasis in the recretohalophyte *Karelinia caspia* under salt stress. *Physiol. Plantarum* 174, e13663. doi: 10.1111/ppl.13663

Guy, R. D., Warne, P. G., and Reid, D. M. (1984). Glycinebetaine content of halophytes: improved analysis by liquid chromatography and interpretations of results. *Physiol. Plantarum* 61, 195–202. doi: 10.1111/j.1399-3054.1984.tb05896.x

Hao, S. H., Wang, Y. R., Yan, Y. X., Liu, Y. H., Wang, J. Y., and Chen, S. (2021). A review on plant responses to salt stress and their mechanisms of salt resistance. *Horticulturae* 7, 132. doi: 10.3390/horticulturae7060132

Hernández, I., Alegre, L., Van Breusegem, F., and Munné-Bosch, S. (2009). How relevant are flavonoids as antioxidants in plants? *Trends Plant Sci.* 14, 125–132. doi: 10.1016/j.tplants.2008.12.003

Jiao, Y., Bai, Z. Z., Xu, J. Y., Zhao, M. L., Khan, Y., Hu, Y. J., et al. (2018). Metabolomics and its physiological regulation process reveal the salt-tolerant mechanism in *Glycine soja* seedling roots. *Plant Physiol. Bioch.* 126, 187–196. doi: 10.1016/j.plaphy.2018.03.002

Jia, X. M., Zhu, Y. F., Hu, Y., Zhang, R., Cheng, L., Zhu, Z. L., et al. (2019). Integrated physiologic, proteomic, and metabolomic analyses of *Malus halliana* adaptation to saline-alkali stress. *Hortic. Res.* 6, 91. doi: 10.1038/s41438-019-0172-0

Kashyap, D., Sharma, A., Tuli, H. S., Sak, K., Punia, S., and Mukherjee, T. K. (2017). Kaempferol-a dietary anticancer molecule with multiple mechanisms of action: recent trends and advancements. *J. Funct. Foods* 30, 203–219. doi: 10.1016/j.jff.2017.01.022

Kerkeb, L., Donaire, J. P., Venema, K., and Rodríguez-Rosales, M. P. (2001). Tolerance to NaCl induces changes in plasma membrane lipid composition, fluidity and h^+-ATPase activity of tomato calli. *Physiol. Plantarum* 113, 217–224. doi: 10.1034/j.1399-3054.2001.1130209.x

Khan, N., Bano, A., Rahman, M. A., Rathinasabapathi, B., and Babar, M. A. (2019). UPLC-HRMS-based untargeted metabolic profiling reveals changes in chickpea (*Cicer arietinum*) metabolome following long-term drought stress. *Plant Cell Environ.* 42, 115–132. doi: 10.1111/pce.13195

Kumari, A., Das, P., Parida, A. K., and Agarwal, P. K. (2015). Proteomics, metabolomics, and ionomics perspectives of salinity tolerance in halophytes. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00537

Li, M. X., Guo, R., Jiao, Y., Jin, X. F., Zhang, H. Y., and Shi, L. X. (2017). Comparison of salt tolerance in soja based on metabolomics of seedling roots. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01101

Li, C., Mur, L. A., Wang, Q. H., Hou, X. C., Zhao, C. Q., Chen, Z. M., et al. (2022). ROS scavenging and ion homeostasis is required for the adaptation of halophyte *Karelinia caspia* to high salinity. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.979956

Li, P. Q., Ruan, Z., Fei, Z. X., Yan, J. J., and Tang, G. H. (2021a). 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

Liu, Y. Q., Liu, J. G., Wang, R., Sun, H., Li, M., Strappe, P., et al. (2020). Analysis of secondary metabolites induced by yellowing process for understanding rice yellowing mechanism. *Food Chem.* 342, 128204. doi: 10.1016/j.foodchem.2020.128204

Li, W., Wen, L. C., Chen, Z. T., Zhang, Z. L., Pang, X. L., Deng, Z. C., et al. (2021b). 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

Lu, C., Feng, Z., Yuan, F., Han, G., Guo, J., Chen, M., et al. (2020). The SNARE protein LbSYP61 participates in salt secretion in *Limonium bicolor*. *Environ. Exp. Bot.* 176, 104076. doi: 10.1016/j.envexpbot.2020.104076

Lv, S. L., Jiang, P., Nie, L., Chen, X., Tai, F., Wang, D. L., et al. (2015). H^+-pyrophosphatase from *Salicornia europaea* confers tolerance to simultaneously occurring salt stress and nitrogen deficiency in arabidopsis and wheat. *Plant Cell Environ.* 38, 2433–2449. doi: 10.1111/pce.12557

Lv, S. L., Tai, F., Guo, J., Jiang, P., Lin, K. Q., Wang, D. L., et al. (2020). Phosphatidylserine synthase from *Salicornia europaea* is involved in plant salt tolerance by regulating plasma membrane stability. *Plant Cell. Physiol.* 62, 66–79. doi: 10.1093/pcp/pcaa141

Ma, Q., Hu, J., Zhou, X. R., Yuan, H. J., Kumar, T., Luan, S., et al. (2017). ZxAKT1 is essential for k^+ uptake and K^+/Na^+ homeostasis in the succulent xerophyte *Zygophyllum xanthoxylum*. *Plant J.* 90, 48–60. doi: 10.1111/tpj.13465

Mansour, M. M. F., Salama, K. H. A., and Allam, H. Y. H. (2015). Role of the plasma membrane in saline conditions: lipids and proteins. *Bot. Rev.* 81, 416–451. doi: 10.1007/s12229-015-9156-4

Moghaieb, R. E. A., Saneoka, H., and Fujita, K. (2004). Effect of salinity on osmotic adjustment, glycinebetaine accumulation and the betaine aldehyde dehydrogenase gene expression in two halophytic plants, *Salicornia europaea* and *Suaeda maritima*. *Plant Sci.* 166, 1345–1349. doi: 10.1016/j.plantsci.2004.01.016

Momonoki, S. Y., Kato, S., and Kamimura, H. (1994). Studies on the mechanism of salt tolerance in *Salicornia europaea jpn. J. Crop Sci.* 63, 650–656. doi: 10.1626/jcs.63.518

Morton, M. J. L., Awlia, M., Al-Tamimi, N., Saade, S., Pailles, Y., Negrão, S., et al. (2019). Salt stress under the scalpel - dissecting the genetics of salt tolerance. *Plant J.* 97, 148–163. doi: 10.1111/tpj.14189

Negrão, S., Schmöckel, S. M., and Tester, M. (2017). Evaluating physiological responses of plants to salinity stress. *Ann. Bot.* 119, 1–11. doi: 10.1093/aob/mcw191

Nemat Alla, M. M., Khedr, A. A., Serag, M. M., Abu-Alnaga, A. Z., and Nada, R. M. (2012). Regulation of metabolomics in atriplex halimus growth under salt and drought stress. *Plant Growth Regul.* 67, 281–304. doi: 10.1007/s10725-012-9687-1

Nie, L. L., Feng, J. J., Fan, P. X., Chen, X. Y., Guo, J., Lv, S. L., et al. (2015). Comparative proteomics of root plasma membrane proteins reveals the involvement of calcium signalling in NaCl-facilitated nitrate uptake in *Salicornia europaea*. *J. Exp. Bot.* 66, 4497–4510. doi: 10.1093/jxb/erv216

Pei, K., Ou, J., Huang, J., and Ou, B. S. (2016). p-Coumaric acid and its conjugates: dietary sources, pharmacokinetic properties and biological activities. *J. Sci. Food. Agric.* 96, 2952–2962. doi: 10.1002/jsfa.7578

Pradeep, P. M., and Sreerama, Y. N. (2018). Phenolic antioxidants of foxtail and little millet cultivars and their inhibitory effects on α -amylase and α -glucosidase activities. *Food Chem.* 247, 46–55. doi: 10.1016/j.foodchem.2017.11.103

Qi, C. H., Chen, M., Song, J., and Wang, B. S. (2009). Increase in aquaporin activity is involved in leaf succulence of the euhalophyte *Suaeda salsa*, under salinity. *Plant Sci.* 176, 200–205. doi: 10.1016/j.plantsci.2008.09.019

Qiu, Y., Liu, Q., and Beta, T. (2010). Antioxidant properties of commercial wild rice and analysis of soluble and insoluble phenolic acids. *Food Chem.* 121, 140–147. doi: 10.1016/j.foodchem.2009.12.021

Ryu, S. B., Karlsson, B. H., Özgen, M., and Palta, J. P. (1997). Inhibition of phospholipase d by lysophosphatidylethanolamine, a lipid-derived senescence retardant. *Proc. Natl. Acad. Sci. U. S. A.* 94, 12717–12721. doi: 10.1073/pnas.94.23.12717

Şirin, S., and Aslım, B. (2019). Determination of antioxidant capacity, phenolic acid composition and antiproliferative effect associated with phenylalanine ammonia lyase (PAL) activity in some plants naturally growing under salt stress. *Medicinal Chem. Res.* 28, 229–238. doi: 10.1007/s00044-018-2278-6

Sanchez, D. H., Siahpoosh, M. R., Roessner, U., Udvardi, M., and Kopka, J. (2008). Plant metabolomics reveals conserved and divergent metabolic responses to salinity. *Physiol. Plant* 132, 209–219. doi: 10.1111/j.1399-3054.2007.00993.x

Schauer, N., and Fernie, A. R. (2006). Plant metabolomics: towards biological function and mechanism. *Trends Plant Sci.* 11, 508–516. doi: 10.1016/j.tplants.2006.08.007

Shabala, S. (2013). Learning from halophytes: physiological basis and strategies to improve abiotic stress tolerance in crops. *Ann. Bot.* 112, 1209–1221. doi: 10.1093/aob/mct205

Shabala, S., and Cuin, T. A. (2008). Potassium transport and plant salt tolerance. *Physiol. Plantarum* 133, 651–669. doi: 10.1111/j.1399-3054.2007.01008.x

Shabala, S., and Mackay, A. (2011). Ion transport in halophytes. *Adv. Bot. Res.* 57, 151–199. doi: 10.1016/B978-0-12-387692-8.00005-9

Shelden, M. C., Dias, D. A., Jayasinghe, N. S., Bacic, A., and Roessner, U. (2016). Root spatial metabolite profiling of two genotypes of barley (*Hordeum vulgare* l.) reveals differences in response to short-term salt stress. *J. Exp. Bot.* 67, 3731–3745. doi: 10.1093/jxb/erw059

Shulaev, V., Cortes, D., Miller, G., and Mittler, R. (2008). Metabolomics for plant stress response. *Physiol. Plantarum* 132, 199–208. doi: 10.1111/j.1399-3054.2007.01025.x

Sun, X., Li, L., Pei, J., Liu, C., and Huang, L. F. (2020). Metabolome and transcriptome profiling reveals quality variation and underlying regulation of three ecotypes for *Cistanche deserticola*. *Plant Mol. Biol.* 102, 253–269. doi: 10.1007/s11103-019-00944-5

Tiika, R. J., Wei, J., Cui, G. X., Ma, Y. J., Yang, H. S., and Duan, H. R. (2021). Transcriptome-wide characterization and functional analysis of xyloglucan endo-transglycosylase/hydrolase (XTH) gene family of *Salicornia europaea* l. under salinity and drought stress. *BMC Plant Biol.* 21, 491. doi: 10.1186/s12870-021-03269-y

Töpfer, N., Kleessen, S., and Nikoloski, Z. (2015). Integration of metabolomics data into metabolic networks. *Front. Plant Sci.* 6. doi: 10.3389/fpls.2015.00049

Turk, M., Méjanelle, L., Sentjurc, M., Grimalt, J. O., Gunde-Cimerman, N., and Plemenitas, A. (2004). Salt-induced changes in lipid composition and membrane fluidity of halophilic yeast-like melanized fungi. *Extremophiles* 8, 53–61. doi: 10.1007/s00792-003-0360-5

Vishwanath, S. J., Delude, C., Domergue, F., and Rowland, O. (2015). Suberin: biosynthesis, regulation, and polymer assembly of a protective extracellular barrier. *Plant Cell Rep.* 34, 573–586. doi: 10.1007/s00299-014-1727-z

Wang, X., Bai, J., Wang, W., Zhang, G., Yin, S., and Wang, D. (2021). A comparative metabolomics analysis of the halophyte *Suaeda salsa* and *Salicornia europaea*. *Environ. Geochem. Health* 43, 1109–1122. doi: 10.1007/s10653-020-00569-4

Wang, F., Chen, L., Chen, H. P., Chen, S. W., and Liu, Y. P. (2019a). Analysis of flavonoid metabolites in citrus peels (*Citrus reticulata* "dahongpao") using UPLC-ESI-MS/MS. *Molecules* 24, 2680. doi: 10.3390/molecules24152680

Wang, W. Y., Liu, Y. Q., Duan, H. R., Yin, X. X., Cui, Y. N., Chai, W. W., et al. (2020). SsHKT1;1 is coordinated with SsSOS1 and SsNHX1 to regulate na^+ homeostasis in *Suaeda salsa* under saline conditions. *Plant Soil* 449, 117–131. doi: 10.1007/s11104-020-04463-x

Wang, Y. L., Zeng, X. Q., Xu, Q. J., Mei, X., Yuan, H. J., Jiabu, D. Z., et al. (2019b). Metabolite profiling in two contrasting *Tibetan hulless* barley cultivars revealed the core salt-responsive metabolome and key salt-tolerance biomarkers. *AOB Plants* 11, plz021. doi: 10.1093/aobpla/plz021

Weckwerth, W. (2003). Metabolomics in systems biology. *Annu. Rev. Plant Biol.* 54, 669–689. doi: 10.1146/annurev.arplant.54.031902.135014

Li, W. F. L., Yang, H. Y., Li, C. M., Tan, S., Gao, X. X., Yao, M., et al. (2019). Chemical composition, antioxidant activity and antitumor activity of tumorous stem mustard leaf and stem extracts. *CyTA-J Food* 17, 272–279. doi: 10.1080/19476337.2019.1577303

Xiao, L., Cao, S., Shang, X. H., Xie, X. Y., Zeng, W. D., Lu, L. Y., et al. (2021). Metabolomic and transcriptomic profiling reveals distinct nutritional properties of cassavas with different flesh colors. *Food Chem.: Mol. Sci.* 2, 100016. doi: 10.1016/j.fochms.2021.100016

Yi, D. B., Zhang, H. N., Lai, B., Liu, L. Q., Pan, X. L., Ma, Z. L., et al. (2021). Integrative analysis of the coloring mechanism of red longan pericarp through metabolome and transcriptome analyses. *J. Agric. Food Chem.* 69, 1806–1815. doi: 10.1021/acs.jafc.0c05023

Zhang, X., Ding, X. L., Ji, Y. X., Wang, S. C., Chen, Y. Y., Luo, J., et al. (2018). Measurement of metabolite variations and analysis of related gene expression in Chinese liquorice (*Glycyrrhiza uralensis*) plants under UV-b irradiation. *Sci. Rep.* 8, 6144. doi: 10.1038/s41598-018-24284-4

Zhang, H. W., Feng, H., Zhang, J. W., Ge, R. C., Zhang, L. Y., Wang, Y. X., et al. (2020a). Emerging crosstalk between two signaling pathways coordinates k^+ and na^+ homeostasis in the halophyte *Hordeum brevisubulatum*. *J. Exp. Bot.* 71, 4345–4358. doi: 10.1093/jxb/eraa191

Zhang, J., Yang, D. S., Li, M. X., and Shi, L. X. (2016). Metabolic profiles reveal changes in wild and cultivated soybean seedling leaves under salt stress. *PloS One* 11, e0159622. doi: 10.1371/journal.pone.0159622

Zhang, H., Zhao, Y., and Zhu, J. K. (2020b). Thriving under stress: how plants balance growth and the stress response. *Dev. Cell* 55, 529–543. doi: 10.1016/j.devcel.2020.10.012

Zhao, M. L., Ren, Y. J., Wei, W., Yang, J. M., Zhong, Q. W., and Li, Z. (2021). Metabolite analysis of jerusalem artichoke (*Helianthus tuberosus* l.) seedlings in response to polyethylene glycol-simulated drought stress. *Int. J. Mol. Sci.* 22, 3294. doi: 10.3390/ijms22073294

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 03 February 2023
DOI 10.3389/fpls.2023.1117868

OPEN ACCESS

EDITED BY
Shangang Jia,
China Agricultural University, China

REVIEWED BY
Maofeng Chai,
Qingdao Agricultural University, China
Chunxiang Fu,
Qingdao Institute of Bioenergy and Bioprocess Technology (CAS), China

*CORRESPONDENCE
Xiaoshan Wang
xswang@yzu.edu.cn

†These authors have contributed equally to this work and share first authorship

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 07 December 2022
ACCEPTED 16 January 2023
PUBLISHED 03 February 2023

CITATION
Wang X, Yin J, Wang J and Li J (2023) Integrative analysis of transcriptome and metabolome revealed the mechanisms by which flavonoids and phytohormones regulated the adaptation of alfalfa roots to NaCl stress.
Front. Plant Sci. 14:1117868.
doi: 10.3389/fpls.2023.1117868

Integrative analysis of transcriptome and metabolome revealed the mechanisms by which flavonoids and phytohormones regulated the adaptation of alfalfa roots to NaCl stress

Xiaoshan Wang*†, Juncheng Yin†, Jing Wang and Junhao Li

Department of Grassland Science, College of Animal Science and Technology, Yangzhou University, Yangzhou, China

Introduction: Salinity critically affects the growth and development of alfalfa (Medicago sativa), making it necessary to understand the molecular mechanism of alfalfa's adaptation to salt stress.

Methods: In this study, alfalfa roots were subjected to salt stress and transcriptomics and metabolomics analyses were performed.

Results: The results showed that flavonoid synthesis, hormone synthesis, and transduction pathways may be involved in the alfalfa salt stress adaptation reaction, and that they are related. Combined analysis of differential genes and differential metabolites found that dihydroquercetin and beta-ring hydroxylase (LUT5), ABA responsive element binding factor 2 (ABF2), protein phosphatase PP2C (PP2C) and abscisic acid (ABA) receptor PYL2 (PYL), luteolinidin was significantly correlated with PP2C and phytochrome-interacting factor 4 (PIF4) and (+)-7-isomethyl jasmonate were significantly correlated with flavonol synthase (FLS) gene. (+)-7-isomethyl jasmonate and homoeriodictyol chalcone were significantly correlated with peroxidase (POD). POD was significantly up-regulated under NaCl stress for 6 and 24 h. Moreover, flavonoids, gibberellin (GA), jasmonic acid (JA) and ABA were suggested to play an important role in alfalfa's response to salt stress. Further, GA, ABA, and JA may be involved in the regulation of flavonoids to improve alfalfa's salt tolerance, and JA may be a key signal to promote the synthesis of flavonoids.

Discussion: This study revealed the possible molecular mechanism of alfalfa adaptation to salt stress, and identified a number of salt-tolerance candidate genes from the synthesis and signal transduction pathways of flavonoids and plant hormones, providing new insights into the regulatory network of alfalfa response to salt stress.

KEYWORDS

flavonoid, *Medicago sativa*, metabolome, NaCl stress, plant hormones, transcriptome

1 Introduction

Soil salinization is one of the important environmental factors that limit the growth and development of plants. Every year, soil salinization causes about 27.3 billion dollars of economic losses worldwide (Morton et al., 2019). When the ion concentration in the soil is very high, the osmotic potential of the soil is reduced, which makes it difficult for plant roots to absorb water, resulting in osmotic stress. In addition, excessive Na^+ and Cl^- accumulation in plant cells causes ion toxicity and disrupts the normal physiological and biochemical processes of cells (Jia et al., 2019). Since plants grow in soil and cannot escape salt stress, specific salt-tolerance mechanisms, such as osmotic homeostasis, ion balance, and changes in hormone levels and antioxidant enzyme activities, have evolved (Li et al., 2019; Gong, 2021).

Flavonoids are common secondary metabolites in plants. Salt stress can significantly change the transcription level of genes related to flavonoid synthesis, increase flavonoid content, and improve the antioxidant enzyme activity of rice and *Cyclocarya paliurus* (Ithal and Reddy, 2004; Lei et al., 2021). The overexpression of the chalcone synthase (CHS) and flavonol synthase (FLS) genes can increase flavonoid content and enhance salt tolerance in plants (Wang et al., 2018; Wang et al., 2021a). Flavonoids have been found to improve salt resistance in plants by affecting the levels of free radicals and the activity of antioxidant enzymes. The expression of the flavanone hydroxylase (F3H) gene has also been up-regulated after salt stress, promoting the accumulation of flavanols to improve antioxidant enzyme activity in tobacco (Mahajan and Yadav, 2014). The dihydroflavonol reductase (DFR) gene in the anthocyanin branching pathway regulates anthocyanin accumulation and was shown to improve salt tolerance in *Brassica napus* (Kim et al., 2017).

Plant hormones are key signaling molecules in plant responses, in particular, defense responses, to abiotic stress. One study has found that after salt stress, the gibberellin 20 oxidase (GA20ox) gene was down-regulated and the activity of GA20ox oxidase decreased synchronously, resulting in a decrease in gibberellin (GA) content and a decrease in plant growth and development (Zhu et al., 2019). Abscisic acid (ABA) can alleviate the degradation of chlorophyll under salt stress, promote the accumulation of osmotic regulatory substances, and regulate stomata to improve water balance, thereby increasing the salt adaptability of tomato and rice (Shahzad et al., 2017). The ABA responsive element binding factor 2 (ABF2) regulates the stress-induced protein kinase 2 (KIN2) gene in the salt stress transduction pathway to enhance salt resistance in *Arabidopsis* (Zhao et al., 2016). Jasmonic acid (JA) can increase the content of sugars, phenols, and soluble proteins in soybeans and alleviate the negative effects of salt stress (Noor et al., 2022). It has been reported that JA interferes with other plant hormones in stress responses, indicating that JA might be a key signal of the regulatory network of plant hormones (Yang et al., 2019). JA methyl ester (MeJA) inhibits the transcription of B-type RR1, RR11, and RR12, and GA transcription factor PIF3 in the GA transduction pathway. MeJA also promotes the up-regulation of ABA synthesis and the transduction genes 9-cis-epoxycarotenoid dioxygenase (NCED) 4 and NCED 5, as well as the ABA receptors PYL2 (PYL2) and PYL6 (PYL6), and serine/threonine-protein kinase SRK2 (SNRK2). Furthermore, the expression levels of peroxidase (POD), superoxide dismutase (SOD), catalase, and vegetative storage protein 2 (VSP2) genes were increased, improving the salt tolerance of *Nitraria tangutorum* Bobr (Gao et al., 2021).

Transcriptome sequencing is considered an effective method for studying molecular regulatory mechanisms and screening candidate genes for salt resistance in plants (Xu et al., 2021). Metabolomics is the bridge between plant genes and phenotypes. With metabolome detection, changes in plant metabolic profiles can be identified and the response mechanisms of metabolic levels can be analyzed. A number of researchers have used transcriptome analysis methods to identify a large number of salt stress-related genes in plants. In addition, transcription factor families such as MYB, bZIP, and WRKY have been identified to be related to salt tolerance (Zhou et al., 2016; Shi and Gu, 2020; Ouertani et al., 2021; Zhang et al., 2021a; Wang et al., 2022). Metabolomic analysis has revealed that amino acids, fatty acids, sugars, organic acids, plant hormones, and flavonoids play important roles in the salt stress resistance of castor bean, *Haloxylon ammodendron*, soybean, and tobacco (Zhang et al., 2011; Li et al., 2017; Jiao et al., 2018; Panda et al., 2021; Wang et al., 2021b).

Alfalfa (*Medicago sativa* L.) is an important genus of legumes that can both provide nutrients to animals and improve soil fertility (Mokhtari et al., 2017). Alfalfa is considered to be a salt-tolerant plant and is widely cultivated in the coastal areas of eastern China. However, owing to the autotetraploid and cross-pollination characteristics of alfalfa, it is difficult to study the mechanism of alfalfa salt resistance. In recent years, with the development of multi-omics integrative analysis technology and its successful application in plant research, the study of plant salt resistance mechanisms has progressed. In this study, transcriptome and metabolome detection were performed on alfalfa roots after NaCl treatment to study the effects of salt stress on alfalfa gene expression and metabolic pathways. The transcriptome and metabolome data was integrated to explore the specific genes and metabolic pathways that alfalfa adapts under salt stress, providing a new understanding of the regulatory mechanism involved in the response.

2 Materials and methods

2.1 Plant materials and salt treatment

Alfalfa (WL525) was selected as the experimental plant. Alfalfa seeds were sterilized by soaking in 5% NaClO solution for 3 min, washed four times with distilled water, and then germinated in sand in a dark incubator at 26°C. After three days of incubation, the seedlings were carefully removed from the sand and transplanted into shrinkage holes on foam boards, three seedlings per hole. The foam boards and transplanted seedlings were then transferred to plastic vessels (35 × 28 ×13 cm^3) with 8.5 L of culture solution. The culture solution contained $Ca(NO_3)_2$ (2.5 mM), KNO_3 (2.5 mM), $MgSO_4$ (1.0 mM), $(NH_4)H_2PO_4$ (0.5 mM), $CuSO_4$ (0.0002 mM), $ZnSO_4$ (0.001 mM), EDTA-Fe-Na (0.1 mM), H_3BO_3 (0.02 mM), $(NH)_4Mo_{27}O_4$ (0.000005 mM), and $MnSO_4$ (0.001 mM), had a pH of 7, and was replaced every seven days (Wang and Han, 2007). A 450 μmol m^{-2} s^{-1}

bio-sodium lamp provided illumination for 15 h (7:30 am–22:30 pm) every day, and the temperature was kept at 26 or 20°C in light or darkness, respectively. After plants were cultivated for 28 days, and they were divided into two groups. One group was cultured in Hoagland complete nutrient solution as the control group, and the other was supplemented with 150 mmol/L NaCl solution in Hoagland complete nutrient solution. At 0, 1, 6 and 24 h, 1-2 g of alfalfa young roots was collected in 1.7 mL centrifuge tubes. After collection, the samples were placed in liquid nitrogen. Then, the samples were divided into two parts. One part was stored in a -80°C refrigerator for subsequent transcriptome sequencing and metabolome detection, and the other part was used to RT-qPCR verification.

2.2 Measurement of physiological and biochemical indicators

Peroxidase (POD) activity was determined by guaiacol colorimetry; superoxide dismutase (SOD) activity was determined by the nitrogen blue tetrazole colorimetric method; catalase (CAT) activity was determined by the ammonium molybdate colorimetric method; malonaldehyde (MDA) was determined by the thiobarbiturate method; proline (PRO) was determined by ninhydrin colorimetry, and soluble protein by the bicinchoninic acid (BCA) method (Cai, 2012). The above indices were determined using the COMIN kit (COMIN, Suzhou, China).

2.3 Transcriptome sequencing

Total RNA was extracted from 12 root samples using a TreliefTM RNAprep Pure Plant Kit (Tsingke, Nanjing, China), and a transcriptome library was constructed. The quality and concentration of RNA were determined using 1% agarose gel electrophoresis, Agilent2100 (BGI, Shenzhen, China), and NanoDrop (BGI, Shenzhen, China). Sequencing was performed on the DNBSEQ platform (BGI, Shenzhen, China), using the latest published genome data from Zhongmu No.1 (https://figshare.com/articles/dataset/Medicago_sativa_genome_and_annotation_files/12623960) as a reference, and using SOAPnuke (BGI, Shenzhen, China) and Trimmomatic (BGI, Shenzhen, China) for statistical filtering. The resulting high-quality reads (clean reads), and datasets can be found in online repositories (https://www.ncbi.nlm.nih.gov/sra/,PRJNA892760, Table S1). The detection of differentially expressed genes (DEGs) was performed according to the Kyoto Encyclopedia of Genes and Genomes (KEGG) and Gene Ontology (GO). The Phyper function in R software was used for enrichment analysis, the P-value was calculated, and then the P-value was corrected by false discovery rate. Generally, a Q-value ≤ 0.05 was regarded as significant enrichment.

2.4 Metabolomic analysis

Metabolites in the alfalfa roots were extracted and analyzed using liquid chromatography (2777CUPLC system; BGI, Shenzhen, China) and mass spectrometry (Xevo G2-XS QTOF; BGI, Shenzhen, China). To obtain reliable data, quality control (QC) samples were added to the detection process. The original mass spectrometry data were imported into the Progenesis QI software (Version 2.2; BGI, Shenzhen, China) for peak extraction. The KNN method was used to fill the missing values and remove low-quality ions (removing the ions missing in more than 50% QC samples or in more than 80% actual samples). Finally, univariate and multivariate analyses were performed using the metabolomic data analysis software Metax (BGI, Shenzhen, China). The multivariate analysis was used to analyze the variable important for the projection of the first two principal components of the PLS-DA model. The univariate analysis was combined with fold-change and Q values to screen metabolites with differential expression (VIP ≥ 1; fold change ≥ 1.2 or ≤ 0.8333; Q-value < 0.05). The metabolites at the intersection of the three were selected as the final differential accumulation metabolites (DAMs), and the thus-obtained inter-group differential metabolites were annotated by KEGG.

2.5 Transcriptome and metabolome correlation network analysis

A Spearman correlation analysis was performed on the transcriptome and metabolome, and the genes and metabolites with absolute correlation coefficient > 0.99 and p-value < 0.01 were screened. Cytoscape was used to construct the network map.

2.6 Real-time quantitative PCR

FastKing gDNA Dispelling RT SuperMix (TIANGEN, Beijing, China) was used to transcribe the RNA to cDNA. Primer Premier 5.0 was used with MSActin as an internal reference. RT-qPCR was performed using the 2 × TSINGKE Master qPCR Mix (SYBR Green I; Tsingke, Nanjing, China) on a MiniOpticon TM Real-Time PCR System (Bio-Rad, Hercules, CA, USA). Each biological replicate consisted of three technical replicates, and relative expression was calculated according to the $2^{-\Delta\Delta CT}$ method.

3 Results

3.1 Transcriptome sequencing for alfalfa roots under short-term NaCl stress

The roots of alfalfa were collected for transcriptome sequencing after treatment with 150 mM NaCl solution for 0, 1, 6, and 24 h, and 77.04 Gb of clean reads were obtained. The sequencing results of the samples treated with NaCl for different periods were highly correlated ($r^2 > 0.91$) and repeatable (Figure 1A). Principal component analysis (PCA) results revealed that the PCA values at 6 h under NaCl stress were significantly different from those at 0 and 1 h. Moreover, PCA results were separated in the direction of PC1 at 24 h (Figure 1B), which was consistent with the correlated heat map results (Figure 1A). The results indicated a significant difference in the transcription levels at 6 and 24 h of NaCl treatment compared to those at 0 and 1 h of treatment. Compared to the NaCl treatment at 0 h, 3385, 4369, and 3948 DEGs were detected at 1, 6, and 24 h of treatment, respectively. At 1 h of NaCl treatment, 2175 DEGs were up-regulated and 1210 DEGs were down-regulated. A total of 2407

10.3389/fpls.2023.1117868

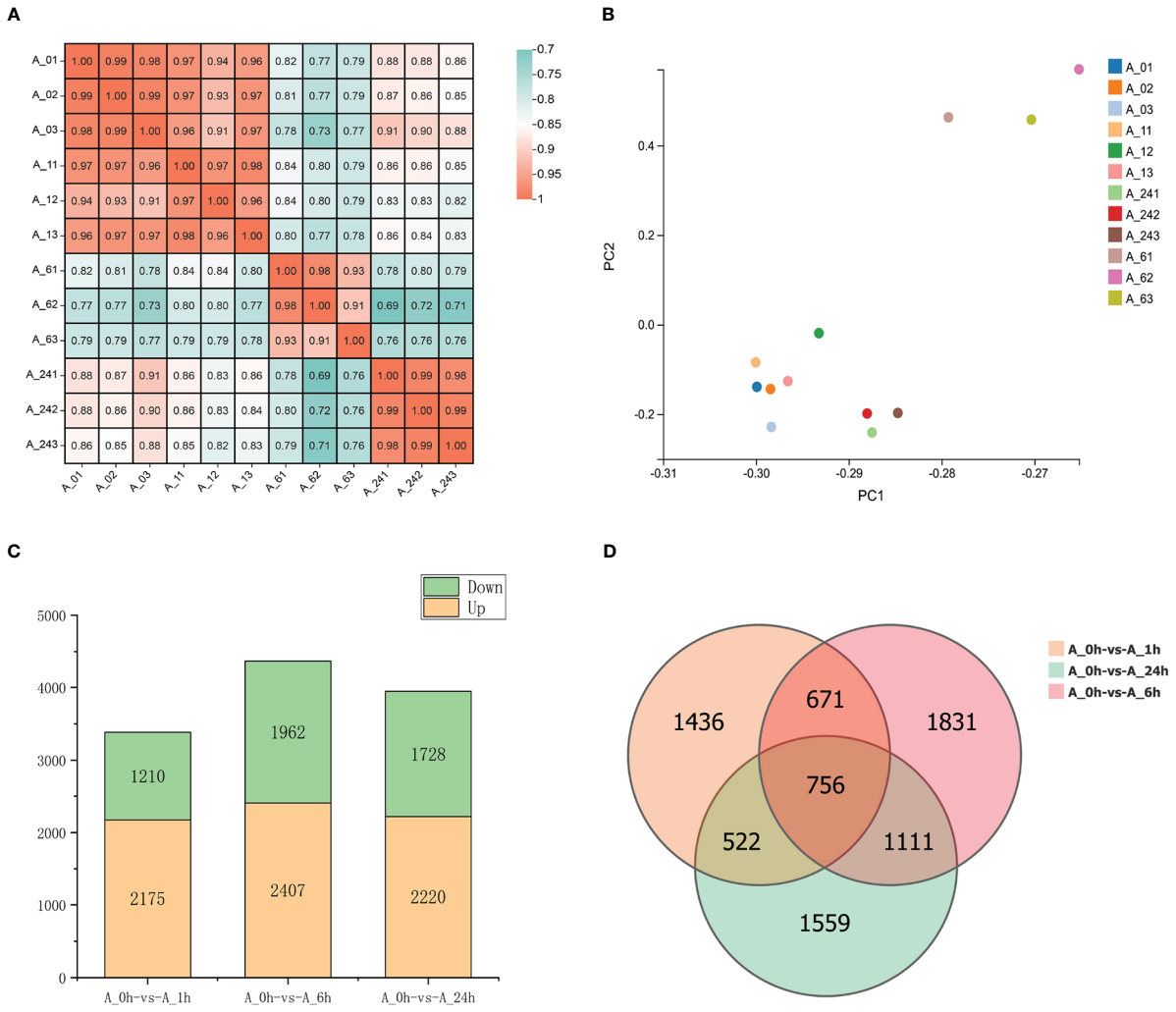


FIGURE 1
(A) Heat map showing the correlation between samples. **(B)** PCA results of transcriptome of each treatment. **(C)** The number of DEGs upregulated and downregulated in different treatments. **(D)** Venn diagram of DEGs. The four time treatments of 0, 1, 6 and 24h were denoted as A_0h, A_1h, A_6h and A_24h respectively.

DEGs were up-regulated and 1962 DEGs were down-regulated at 6 h of treatment. At 24 h, 2220 DEGs were up-regulated and 1728 DEGs were down-regulated. There were 756 common DEGs under NaCl treatment at 1, 6, and 24 h, among which 352 remained up-regulated and 300 remained down-regulated (Figures 1C, D).

GO enrichment analysis revealed that the DEGs were mainly enriched in signal transduction, stimulation response, regulation of cellular processes, protein kinase activity, defense response, naringin CHS activity, cell stimulation response, flavonoid biosynthesis, and other functional units under NaCl stress for 1 h (Figure S1). Under NaCl stress for 6 h, DEGs were mainly enriched in functional units such as oxidoreductase activity, reactive oxygen species metabolism process, oxidative stress response, signal transduction, flavonoid biosynthesis process, JA biosynthesis process, hormone-mediated signaling pathway, and ABA-activated signaling pathway (Figure S1). Under NaCl stress for 24 h, DEGs were mainly enriched in oxidoreductase activity, dioxygenase activity, reactive oxygen species metabolism process, hydrogen peroxide catabolic process, POD activity, antioxidant activity, ABA binding, and hormone-binding functional units (Figure S1). The 756 common DEGs under NaCl stress at 1, 6, and 24 h were mainly enriched in functional units such as oxidoreductase activity, hormone-mediated signaling pathway, ABA binding, protein dephosphorylation regulation, cell response to hormone stimulation, ABA response, hormone binding, and cell response to endogenous stimulus (Figure S1). By comparing the results of DEGs and analyzing the GO enrichment results of common DEGs at 1, 6, and 24 h of NaCl stress, we found that functional units such as the oxidoreductase regulation process, flavonoid synthesis of secondary metabolites, as well as hormone synthesis and transduction, are important pathways for alfalfa adaptation to NaCl stress.

3.2 Analysis of metabolite detection under short-term NaCl stress

After treatment with 150 mM NaCl for 0, 1, 6, and 24 h, the root tips of alfalfa were collected for untargeted metabolite detection using liquid chromatography-mass spectrometry. PCA results showed that the QC samples clustered to one point, indicating that the instrument was stable, and the collected data were of high quality (Figure 2A). The three principal components PC1, PC2, and PC3 accounted for

42.4, 24.8, and 14.1% of the total variation, respectively (Figure 2A). The PCA values of alfalfa root tips under NaCl stress at 0 and 1 h were closely clustered, whereas the PCA values of alfalfa root tips under NaCl stress at 6 and 24 h were separated in the direction of PC1 and PC3 (Figure 2A). The results showed that the metabolite levels at 6 and 24 h of NaCl stress were significantly different from those at 0 and 1 h, and the metabolite levels at 0 h and 1 h of NaCl stress were similar. The PLS-DA results showed that the components at 6 and 24 h were separated from the components at 0 and 1 h of NaCl stress, which is consistent with the PCA results (Figure 2B).

In the present study, 32 DAMs were found under NaCl stress for 1 h, of which 25 were down-regulated and 7 were up-regulated. Twenty-nine DAMs were found under NaCl stress for 6 h, of which 11 were down-regulated and 18 were up-regulated. A total of 153 DAMs were found under NaCl stress for 24 h, of which 50 were down-regulated and 103 were up-regulated. The results showed that differential metabolites increased significantly with increasing NaCl stress duration (Figure S2). A total of 178 common metabolites were detected from the DAMs of alfalfa roots treated by NaCl for 1, 6, and 24 h, which were divided into four categories according to their expression changes. The expression levels of 28 metabolites decreased under NaCl stress for 1 h, 6 h, and 24 h. The expression levels of 32 metabolites remained unchanged at 1 and 6 h under NaCl stress but decreased significantly at 24 h. The expression levels of 15 metabolites increased first and then decreased during 1–24 h of NaCl stress. The expression levels of 103 metabolites remained unchanged at 1 and 6 h under NaCl stress, but significantly increased at 24 h (Figure 2C; Tables S3, S4).

3.3 Integrative analysis of transcriptome and metabolome for alfalfa roots under short-term NaCl stress

To determine the correlation between DEGs and DAMs under NaCl stress, KEGG pathway analysis was used to conduct a correlation analysis of the transcriptome and metabolome detection

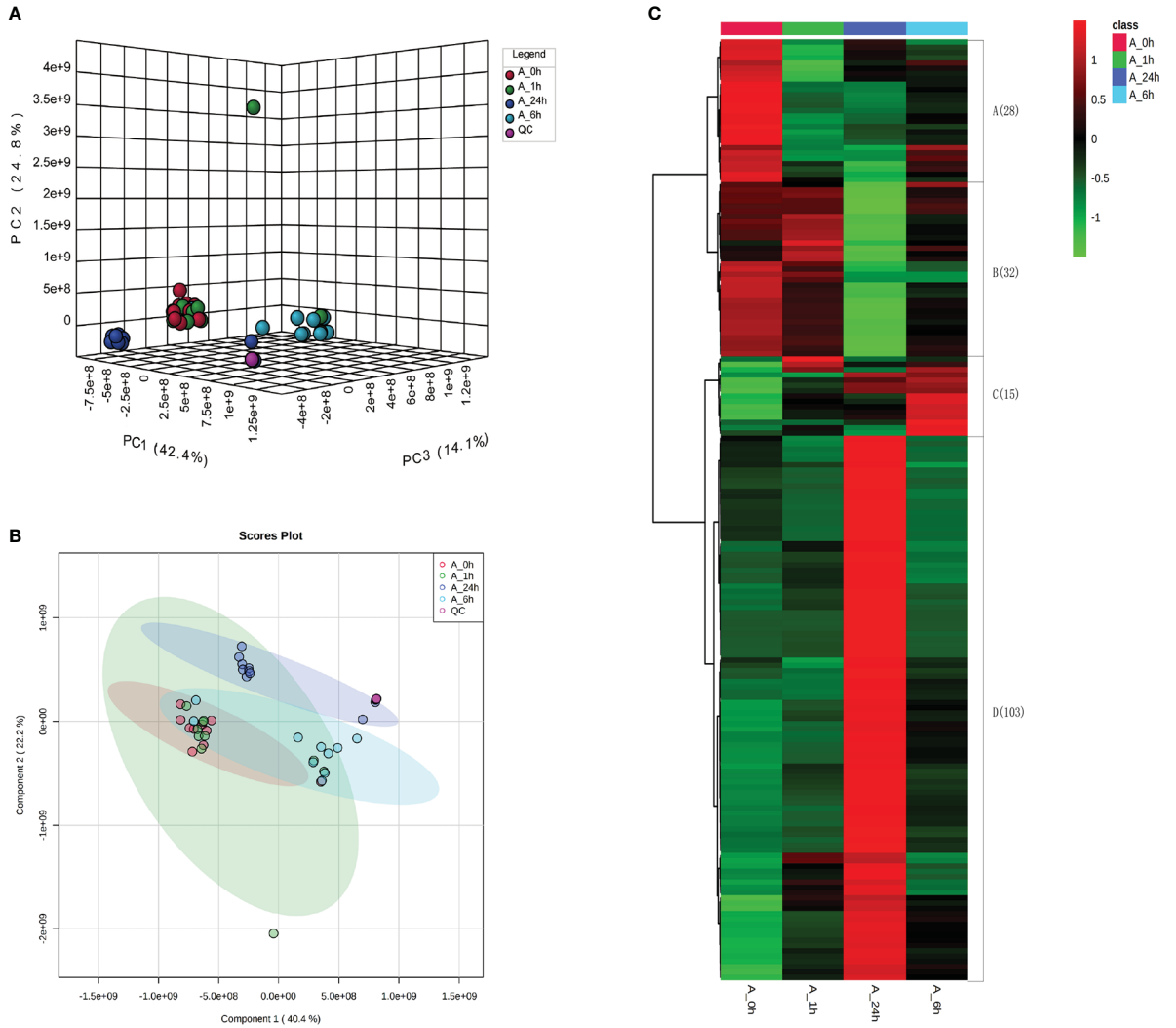


FIGURE 2
(A) Principal component analysis (PCA) results of metabolome of each treatment. **(B)** The results of metabolome partial least squares (PLS-DA) analysis. Partial least squares (PLS-DA) model was constructed with MetaX software for the analysis of the results. The model parameters R2 and Q2 had high values (Table S2). The 200 displacement tests of the model parameters R2 and Q2 showed that the current PLS-DA model was reliable. **(C)** Heat map analysis of DAMs in different treatments.

data to elucidate a common enrichment pathway. KEGG enrichment pathways showed one common enrichment pathway in DEGs and DAMs under NaCl stress at 1 h. There were seven common enrichment pathways at 6 and 24 h of NaCl stress (Figure 3). The data analyses of transcriptome and metabolome also showed that the expression levels of genes and metabolites were similar at 0 and 1 h under NaCl stress, whereas at 6 and 24 h, these were significantly different from those at 0 h and 1 h (Figures 1A, B, 2A, B). Moreover, the contents of MDA, osmotic substances PRO and soluble protein, and the activities of antioxidant enzymes POD, SOD, and CAT were also measured at 0, 1, 6, and 24 h of NaCl stress. The contents of MDA, PRO, and soluble protein and the activities of POD, SOD, and CAT did not change significantly at 1 h of NaCl stress compared with those at 0 h. The content of MDA and the activities of POD and SOD increased significantly at 6 and 24 h of NaCl stress (Figure S3 and Table S5). The results showed that alfalfa plants had no obvious response to NaCl stress before 1 h, but a clear response after 1 h. At 6 and 24 h of NaCl stress, common pathways of DEGs and DAMs were isoflavone biosynthesis, flavonoid biosynthesis, ABC transporter, arachidonic acid metabolism, flavonoid and flavonol biosynthesis, and diterpenoid biosynthesis, as well as degradation of valine, leucine, and isoleucine. In addition, DEGs were also enriched in the MAPK signaling pathway, plant hormone signal transduction, carotenoid biosynthesis, and zeatin biosynthesis. DAMs were also enriched in alpha-linolenic acid metabolism and anthocyanin biosynthesis (Figure 3).

3.4 Effects of short-term NaCl stress on flavonoid biosynthetic pathways

Flavonoids are common secondary metabolites in plants and are closely related to plant stress tolerance. Under NaCl stress, the major metabolites, and regulatory genes involved in the flavonoid biosynthesis pathway of alfalfa were significantly altered (Figure 4; Tables S6, S7). With the duration of NaCl stress, the expression levels of most enzymes involved in flavonoid synthesis first increased and then decreased (Figure 4). Accumulation of most metabolites reached its highest level at 24 h of NaCl stress (Figure 4). This is because in the early stage of NaCl stress, the enzymes genes related to flavonoid synthesis were up-regulated, which led to the rapid initial increase of metabolite synthesis (Figure 4). With continuous NaCl stress, plants gradually adapted, and the expression of related enzymes genes decreased, but the accumulation of metabolites still increased. Chalcone synthase (CHS) is a key rate-limiting enzyme in the flavonoid biosynthesis pathway. In the present study, the expression levels of five genes (MsG0080048348.01, MsG0080048352.01, MsG0180005356.01, MsG0380016127.01, and MsG0580024218.01) encoding CHS were significantly up-regulated at 1 or 6 h of NaCl stress. Two genes (MsG0180003128.01 and MsG0180 003129.01) encoding chalconeisomerase (CHI) always remained up-regulated during NaCl stress. The expression level of the flavanone 3-hydroxylase (F3H, MsG0880045643.01) gene was the highest under 6 h NaCl stress (Figure 4). The accumulation of dihydroquercetin, an intermediate product of flavonoid biosynthesis, first decreased and then increased with the duration of NaCl stress, and reached its maximum value at 24 h (Figure 4). The flavonol synthetase pathway is a branch of the flavonol synthesis pathway. In the present study, four genes (MsG0480022886.01, MsG0580027459.01, MsG0880042881.01, and MsG0880046855.01) encoding flavonol synthetase (FLS) were up-regulated at 6 and 24 h under NaCl stress (Figure 4). In the flavone branch pathway, the accumulation of luteolin first decreased and then increased with the duration of NaCl stress, and accumulation reached its highest value at 24 h (Figure 4). Under NaCl stress, the expression of dihydroflavonol reductase (DFR, MsG0880042687.01) was up-regulated in the anthocyanin synthesis pathway branch, reaching its highest level at

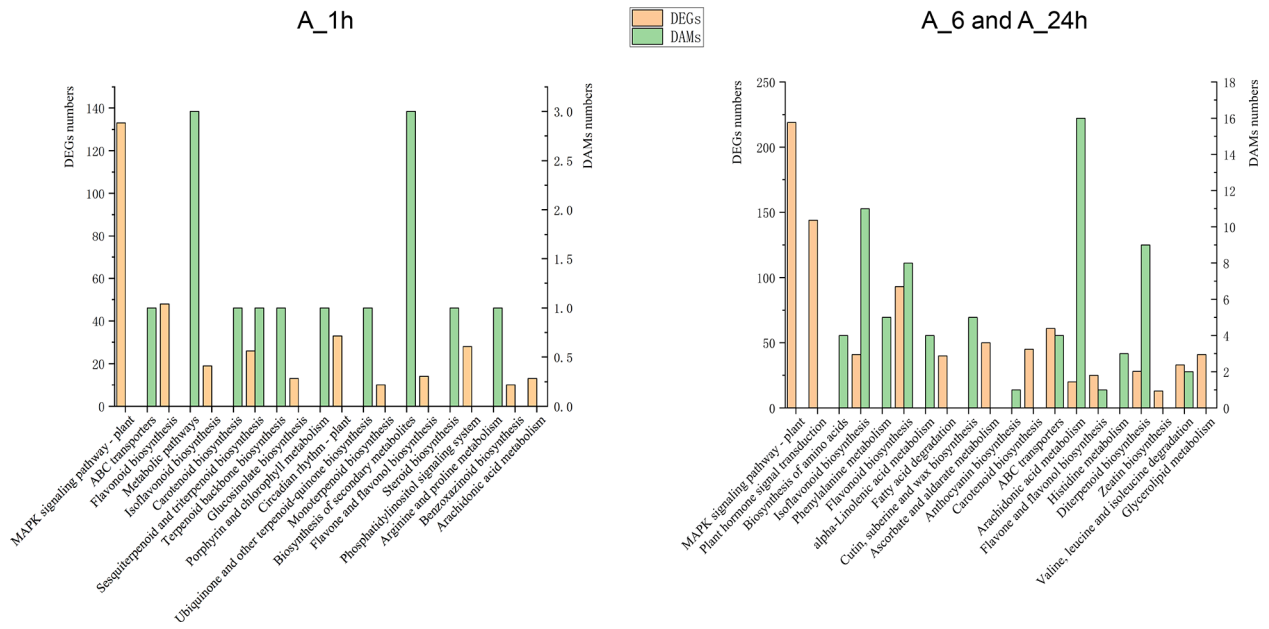


FIGURE 3
KEGG coenrichment analysis of DEGs and DAMs treated with NaCl stress at different times (Q-value < 0.05). According to the changes of transcription and metabolism levels of the samples, alfalfa had no obvious response to salt stress at 1h, so 1h and the following 6 and 24h were analyzed separately. A_6h and A_24h were the two time points of KEGG enrichment pathway, which were combined for metabolome and transcript co-enrichment analysis.

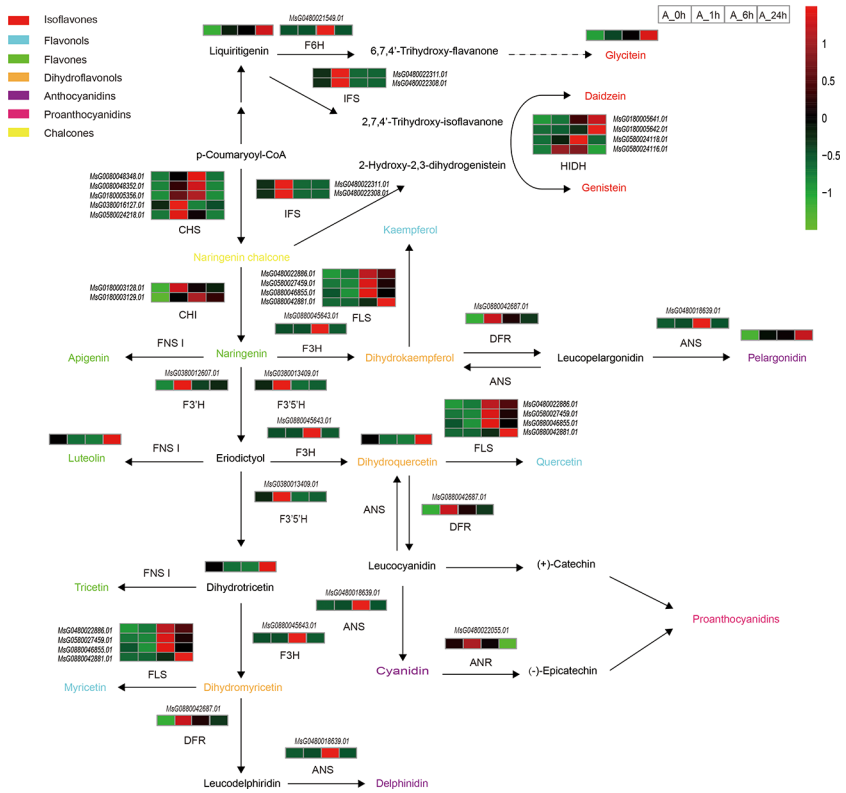


FIGURE 4
Changes of important genes and metabolites in flavonoid biosynthesis under NaCl stress. Path network diagram were mapped based on KEGG and related literature. Italics indicate gene ids that label enzymes or transcription factors, the rest are unlabeled metabolites, and different colors indicate different types of flavonoids.

1 h. Interestingly, the expression of anthocyanin synthase (ANS, MsG0480018639.01) was only up-regulated at 6 h. In the flavonoid synthesis pathway, the accumulation of the metabolite pelargonidin, the intermediate product of flavonoid synthesis, gradually increased with the duration of NaCl stress, and reached its maximum value at 24 h. The anthocyanidin reductase (ANR, MsG0480022055.01) gene of the proanthocyanidin synthesis pathway, another branch pathway of flavonoid synthesis, was up-regulated at 1 h and down-regulated at 24 h of NaCl stress. (Figure 4). The isoflavone synthesis pathway is yet another branch pathway of flavonoid synthesis. The accumulation of liquiritigenin and glycitein in this pathway gradually increased with the duration of NaCl stress (Figure 4).

3.5 Effects of short-term NaCl stress on phytohormone synthesis and transactivation

Phytohormones are the integrated centers of plant response to environmental stress, and the complex signal network formed by different hormones can achieve fine regulation of stress response. In the present study, the expression levels of genes related to the synthesis and transduction pathways of key hormones GA, ABA and JA, and the accumulation of intermediate metabolites during their synthesis were altered in alfalfa under NaCl stress (Tables S8, 9). The gibberellin-20 oxidase (GA20ox, MsG0780040403.01) gene was down-regulated in the GA synthesis pathway for 6 and 24 h. The accumulation of gibberellin 12 (GA 12), gibberellin 1 (GA 1), gibberellin 4 (GA 4), gibberellin 9 (GA 9), gibberellin 14 (GA 14), gibberellin 15 (GA 15), gibberellin 20 (GA 20), and gibberellin 53 (GA 53) showed a trend of first decreasing and then increasing during NaCl stress, with the highest levels reaching at 24 h (Figure 5A). This result may be related to the up-regulation of ent-kaurene oxidase (KO, MsG0680031275.01) and ent-kaurenoic acid monooxygenase (KAO, MsG0180001205.01) genes at 1 and 6 h of NaCl stress (Figure 5A). In the present study, the expression of gibberellin 3beta-dioxygenase (GA3ox, MsG0280011347.01 and MsG0880042366.01) gene first decreased and then increased. After 1 h of NaCl stress, the levels of GA19 and GA36 were significantly decreased, while the levels of GA9, GA1 and GA4 were significantly increased, indicating that active GA1 and GA4 may be synthesized mainly through GA9 after 1 h NaCl stress. In our study, seven of the eight DELLA protein genes (MsG0280011041.01, MsG0280011042.01, MsG0480020976.01, MsG0480021764.01, MsG0580024328.01, MsG0780041365.01, and MsG0880046208.01), four PIF3/4 genes, and one GID1 gene (MsG0780041225.01) were down-regulated at 6 h of NaCl, while the expression levels of most DELLA genes at 24 h were significantly higher than at 0 and 6 h (Figure 5A).

Three genes (MsG0280009782.01, MsG0280009781.01, and MsG0180001827.01) encoding 9-cis-epoxycarotenoid dioxygenase (NCED) in the ABA synthesis pathway of alfalfa were significantly up-regulated under NaCl stress for 1 h and 6 h (Figure 5B), but the ABA content did not change significantly (Figure 5B). Both PYL (MsG0380015520.01 and MsG0480018569.01) and SnRK2

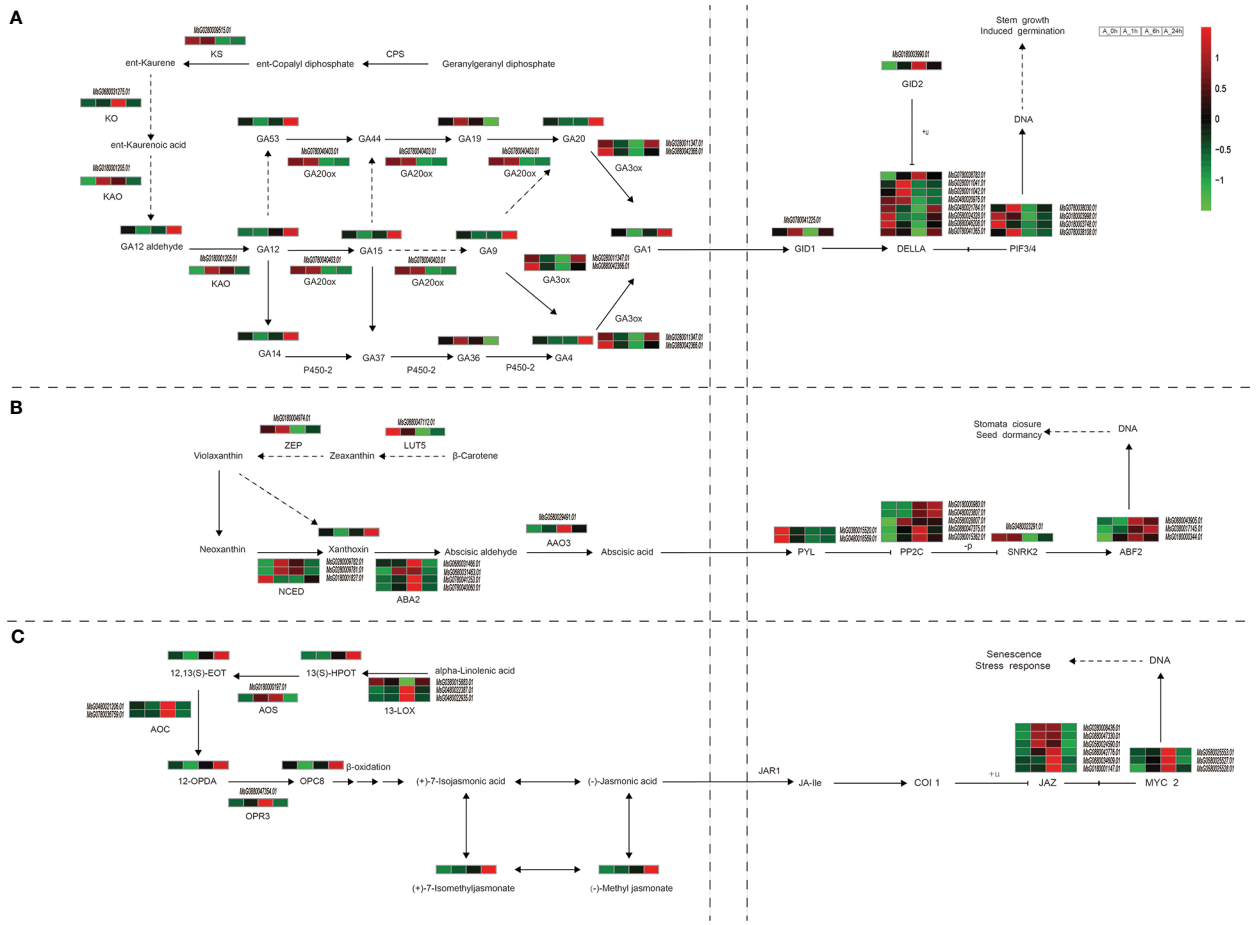


FIGURE 5
Changes of important genes and metabolites in plant hormone synthesis and transduction under NaCl stress. **(A)** GA synthesis and transduction process. **(B)** ABA synthesis and transduction process. **(C)** JA synthesis and transduction process. Pathways were mapped based on KEGG and related literature. Italics indicate gene ids that label enzymes or transcription factors, the rest are unlabeled metabolites.

(MsG0480023291.01) genes were down-regulated under NaCl stress for 6 and 24 h, but all genes encoding PP2C (MsG0180000980.01, MsG0380015362.01, MsG0480023807.01, MsG0580028807.01, and MsG0880047375.01) and ABF2 (MsG0380017145.01, MsG0880043905.01, and MsG0180000344.01) were up-regulated (Figure 5B). These results indicate that salt stress could affect the gene expression of some enzymes and the accumulation of some intermediate metabolites in the ABA synthesis pathway, but has no significant effect on the content of ABA in roots.

In the JA synthesis pathway, two of the three genes encoding 13s-lipoxygenase (13-Lox) (MsG0480022387.01 and MsG0480022935.01), and two genes (MsG0780036759.01 and MsG0480021206.01) encoding allene oxide cyclase (AOC) were up-regulated under NaCl stress for 6 h. Allene oxide synthase (AOS, MsG0180000197.01), and 12-oxophytodienoic acid reductase (OPR3, MsG0880047354.01) in the JA synthesis pathway were up-regulated under NaCl stress for 1 h and 6 h (Figure 5C). This may lead to the accumulation of synthetic intermediates such as 12, 13(S)-EOT, 12-OPDA, and OPC8; these were down-regulated at 1 h compared with those at 0 h, and decreased first and then increased during 24 h NaCl stress, at which time their concentrations were the highest (Figure 5C). The metabolites 13(S)-HPOT, methyl jasmonate, and (+)-7-Isomethyljasmonate showed a gradual increase in accumulation within 24 h of NaCl stress. All of the six genes (MsG0180001147.01, MsG0280008436.01, MsG0580024590.01, MsG0680034609.01, MsG0880042776.01, and MsG0880047330.01) encoding Jasmonate-ZIM domain genes (JAZ) were up-regulated at 6 h NaCl stress, and four of these (MsG0280008436.01, MsG0580024590.01, MsG0880042776.01, and MsG0880047330.01) were up-regulated at 1 h NaCl stress. All the genes (MsG0580025527.01, MsG0580025528.01, and MsG0580025553.01) encoding MYC2 were up-regulated at 1 and 6 h NaCl stress (Figure 5C). The results showed that the expression levels of some enzyme and transduction genes in the JA synthesis pathway were significantly up-regulated by NaCl stress at 6 h, and the accumulation of the intermediates methyl jasmonate and 7-isomethyl jasmonate significantly increased.

3.6 Gene and metabolite related networks

Correlation analysis of DEGs and DAMs on the major enrichment pathways in alfalfa at 6 and 24 h of NaCl stress resulted in the construction of a gene–metabolite association network related to flavonoids and hormones (Table S10). Dihydroquercetin in the flavonoid synthesis pathway; ABF2 (MsG0180000344.01), PP2C (MsG0880047375.01) and PYL (MsG0380015520.01) genes in the

ABA signal transduction pathway; and the β-cyclic hydroxylase (LUT5, MsG0880047112.01) gene in the gibberellin synthesis pathway were significantly correlated (Figure 6A). The flavonoid luteolinidin was significantly correlated with the genes PIF4 (MsG0180003748.01) and PP2C (MsG0180000980.01, MsG0480023807.01) in the gibberellin signal transduction pathway. Homoeriodictyol chalcone in flavonoid synthesis pathway and (+)-7-isomethyl jasmonate in JA synthesis pathway coexisted with POD (MsG0280008067.01) and pathogenesis-related protein 1 (PR1, MsG0480022571.01) (Figure 6B). In addition, (+)-7-isomethyl jasmonate was also associated with FLS (MsG0480022886.01), JA downstream vegetative storage protein 2 (VSP2, MsG0380015438.01) gene, cis zeatin synthesis dimethyl allyl transferase (miaA, MsG0780039542.01) gene, and ethylene responsive transcription factor 1 (ERF1, MsG0780040733.01, Figure 6B). These results suggested that flavonoid synthesis is highly correlated with DEGs and the metabolites of phytohormone syntheses, such as GAs, JA, and ABA. Alfalfa could be resistant to salt stress through the interaction between flavonoids and phytohormones, which may play a key role in the process of salt stress tolerance in plants.

3.7 RT-qPCR validation

A total of 20 DEGs that play key roles in flavonoid synthesis and in hormone synthesis and signal transduction, and some DEGs involved in gene and metabolite related networks, were selected for RT-qPCR to verify the accuracy of RNA sequencing. The results showed that the RNA sequencing data correlated significantly with the qRT-PCR results ($r^2 = 0.78259$, $P < 0.05$; Figures 7, 8; Table S11), indicating that the results were reliable.

4 Discussion

Alfalfa can grow in moderately saline soils; hence, it is cultivated in large areas of coastal China to provide forage for dairy cows. However, alfalfa plants are sensitive to high salt stress, and 50–200 mmol L^{-1} NaCl stress significantly inhibits the growth of alfalfa (Mokhtari et al., 2017). Since alfalfa is cross-pollinated and homologous tetraploid, previous studies on its salt tolerance mechanism have not been carried out in depth. In recent years, with the development of multi-omics and multi-omics data integration analysis technology, transcriptomics, proteomics, and metabolomics analyses have been widely used in the study of alfalfa plant growth and development, and its responses to abiotic stress. This has resulted in new progress in the study of some mechanisms in alfalfa (Song et al., 2016; Chen et al., 2021; Ma et al., 2021; Xu et al., 2021; Yang et al., 2021). In the present study, alfalfa plants were subjected to NaCl stress, root samples were collected for transcriptome sequencing and metabolite detection, and the transcriptome data and metabolome data were integratively analyzed. The transcriptome analysis of alfalfa roots revealed that under NaCl stress DEGs were mainly enriched in isoflavone biosynthesis, flavonoid biosynthesis, ABC transporters, the MAPK signaling pathway, phytohormone signaling, carotenoid biosynthesis, and trans-Zeatin biosynthesis. In addition, metabolomic analysis indicated that under NaCl stress the DAMs in alfalfa were mainly enriched in isoflavone biosynthesis, flavonoid biosynthesis, ABC transporters, alpha-linolenic acid metabolism, and anthocyanin biosynthesis pathways. The integrated analysis of transcriptome and metabolome data showed that isoflavone biosynthesis, flavonoid biosynthesis, and ABC transporter pathways were enriched in alfalfa at 6 and 24 h of NaCl stress. Therefore, based on DEGs and

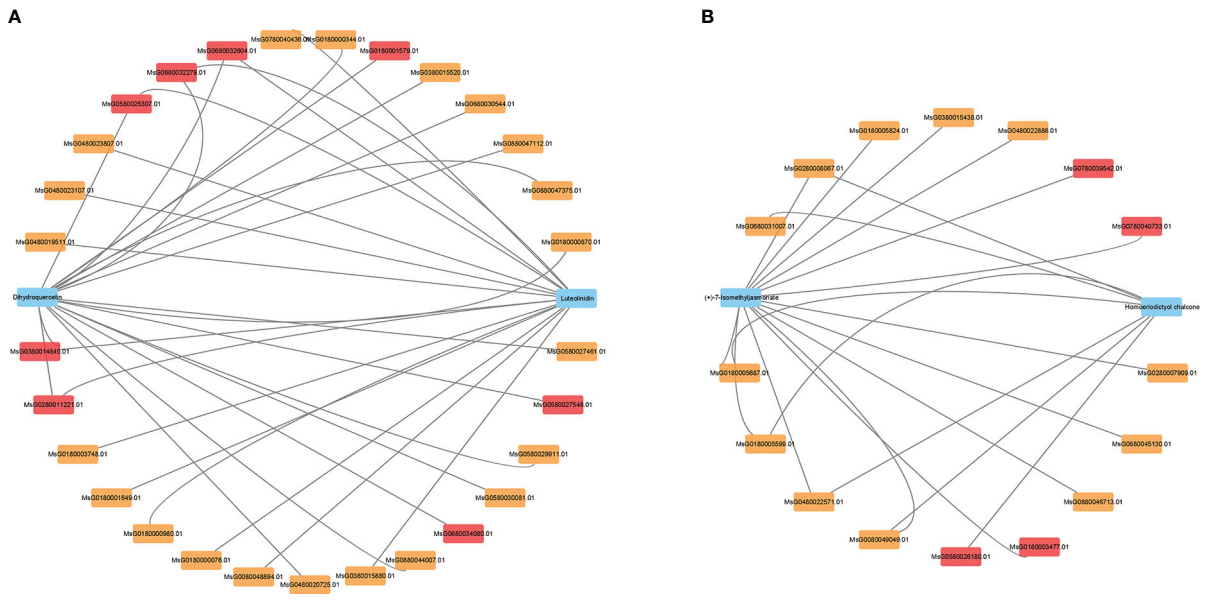


FIGURE 6
DEGs and DAMs association analysis network. **(A, B)** Correlation analysis of DEGs and DAMs at 6 and 24 h NaCl stress. Blue represents metabolites, red ($p < 0.01$) and yellow ($p < 0.05$) are genes.

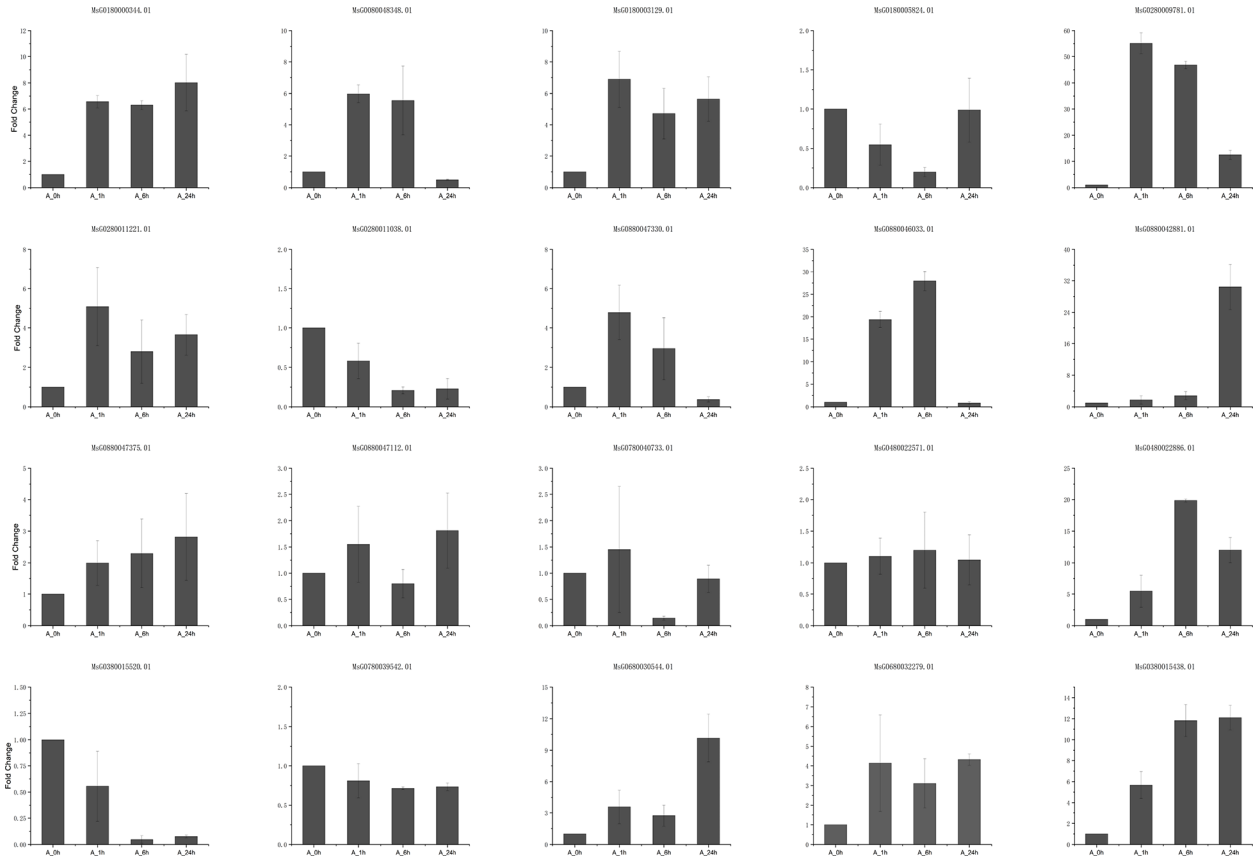


FIGURE 7
The results of RT-qPCR with twenty genes.

DAMs being enriched in common pathways, this study revealed the molecular mechanism of the rapid response of alfalfa to NaCl stress.

4.1 Flavonoid regulatory networks

Flavonoids are common secondary metabolites in plants and can be divided into eight groups according to their molecular structure: flavonols, anthocyanins, proanthocyanidins, isoflavonoids, chalcones, flavones, flavandiols, and aurones (Di Ferdinando et al., 2012). Increased total flavonoid content has been reported in sorghum under moderate and severe salinity stress, with significant changes in related synthetic genes (Ma et al., 2020). The total flavonoid content and the content of seven flavonoids in the leaves of *Cymbopogon* showed an increasing trend with increasing salt concentrations, and short-term salt stress also enhanced antioxidant enzyme activity in the leaves of *Cymbopogon* (Lei et al., 2021). High salt stress significantly increased the expression levels of flavonoid synthesis genes DFR and ANS in rice (Ithal and Reddy, 2004). In this study, the content of some flavonoids such as dihydroquercetin, dihydromyricetin, pelargonidin, luteolin, liquitigenin, and glycitein increased significantly at 24 h of NaCl stress. Moreover, the flavonoid synthesis genes CHS, CHI, FLS, DFR, ANS, F3H, ANR, IFS (2-hydroxyisoflavanone synthase) and F6H (flavonoid 6-hydroxylase) were significantly up-regulated at 1 or 6 h of NaCl stress, which is consistent with the results of Ithal's study. The CHS gene of okra overexpressed in *Arabidopsis thaliana* increased flavonoid content and the expression levels of SOD and POD genes under salt stress, thereby improving the salt-tolerance of *Arabidopsis* (Wang et al., 2018). Overexpression of the F3H gene increased the content of flavan-3-ols (catechin and epicatechin) and directly improved the antioxidant activity of the tobacco plants (Mahajan and Yadav, 2014). DFR gene overexpression can lead to the accumulation of large amounts of anthocyanins and maintain lower levels of reactive oxygen species and higher chlorophyll content after high salt treatment (Kim et al., 2017). Overexpression of the FLS gene in tobacco increased the total flavonoid content by 1.84-fold compared to that in tobacco in which this gene was not overexpressed. Moreover, it was able to take up more potassium ions to maintain sodium–potassium ion homeostasis and have higher salt tolerance (Wang et al., 2021c). These previous reports indicate that the up-regulation of some important genes in the flavonoid synthesis pathway, alone or in combination, can increase the flavonoid content of plants, thereby enhancing their stress resistance. Additional research has shown that flavonoids can improve plant tolerance to abiotic stresses by directly or indirectly regulating the levels of oxygen-containing radicals, peroxides, and reactive oxygen species, such as oxygen ions (Wang et al., 2021a). In our study, we found that NaCl stress significantly changed the expression levels of CHS, CHI, FLS, DFR, ANS, F3H, and IFS in the flavonoid synthesis pathway of alfalfa. Moreover, the concentrations of some flavonoid products such as licorice, dihydroquercetin, and luteolin increased

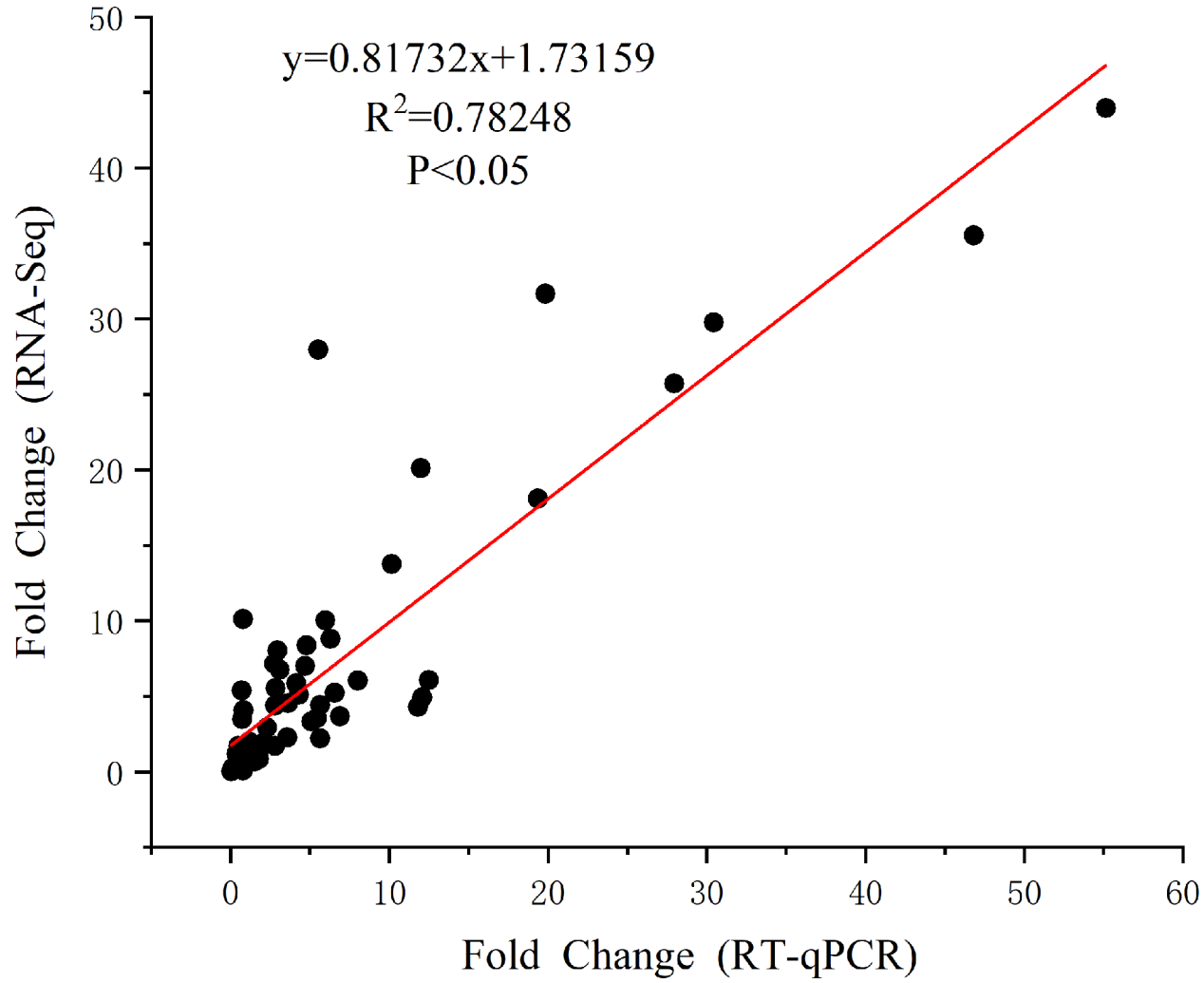


FIGURE 8
Correlation analysis between RNA-Seq results and RT-qPCR results ($p < 0.05$).

significantly under NaCl stress for 24 h. The transcriptome GO was also enriched in some functional units such as redox-regulation process and flavonoid synthesis of secondary metabolites, while the transcriptome and metabolome KEGG were enriched in common pathways such as isoflavone biosynthesis, flavonoid biosynthesis, and ABC transport. The ABCC2 transporter protein can co-transport flavonoid compounds such as anthocyanin (Cy3G) with glutathione (Behrens et al., 2019). Moreover, the ABCG10 transporter protein can transport the isoflavone synthesis precursors, coumarate 4-coumarate, and liquiritigenin (Biala et al., 2017). The results of this study indicated that the regulation of flavonoid synthesis and transport is an important mechanism for alfalfa to adapt to salt stress.

4.2 Plant hormone regulatory networks

Phytohormones are trace endogenous signaling molecules synthesized by plants, which play an important role in stress response. Under salt stress, GA content decreases and the expression of GA20ox, a key gene for GA synthesis, is down-regulated (Zhu et al., 2019). Salt stress inhibits plant growth by reducing GA levels, whereas exogenous application of GA stimulates the synthesis of endogenous GA in plants, thus promoting plant growth (Niharika et al., 2020). The repressor DELLA protein can inhibit the expression of transcription factor PIF3/4 in the gibberellin transduction pathway, and the amycin receptor protein genes 1 and 2 (GID1 and GID2, respectively) can attenuate DELLA protein after binding to GA. In the present study, the GA20ox in alfalfa roots was down-regulated at 6 and 24 h of NaCl stress, while the contents of GA1, GA4, and GA9 were decreased at 1 and 6 h of NaCl stress and significantly increased at 24 h of NaCl stress. The genes of DELLA, GID1, and PIF3/4 in the gibberellin transduction pathway of alfalfa roots were significantly down-regulated at 6 h, but most of them were restored to pre-stress levels, or even up-regulated, at 24 h. These results indicated that alfalfa plants can adapt to a salt stress environment by reducing GA content to inhibit plant growth in the early stage of NaCl stress. With the duration of salt stress, alfalfa plants gradually adapted to its effects, and the GA content was increased again to restore plant growth.

ABA is a hemiterpene phytohormone, and the addition of 10 μM ABA to the root zone of wheat under salt stress significantly increases aboveground biomass and root length, reduces sodium uptake by

more than 80%, and increases antioxidant enzyme activity in roots (Parveen et al., 2021). Endogenous ABA can improve salt tolerance of rice by regulating stomatal closure and improving water balance (Shahzad et al., 2017). In this study, the ABA content in roots did not change significantly after NaCl stress, while ABA synthesis-related genes were up-regulated. A possible explanation of this is that ABA synthesized by roots is transported to the ground and is involved in the regulation of shoot defense against salt stress (Poór et al., 2019). The SnRK2 kinase can induce the expression of transcription factor ABF2 in the ABA transduction pathway of alfalfa. Protein phosphatase PP2C inhibits SnRK2 kinase activity by dephosphorylation, while receptor protein PYL inhibits PP2C activity by ABA. In a previous study, overexpression of ABF2, a transcription factor for ABA synthesis in *Arabidopsis*, was shown to increase the expression of the stress response gene KIN2, thereby improving *Arabidopsis* salt tolerance (Zhao et al., 2016). In the present study, while PYL and SnRK2 genes were down-regulated under 6 and 24 h of NaCl stress, ABF2 and PP2C gene expression increased, suggesting that alfalfa can enhance the expression of downstream resistance genes in roots by increasing the expression of the transcription factor ABF2, thereby enhancing the resistance of the plants to salt stress.

JA is an important stress regulatory hormone in plants. MeJA, and JA isoleucine (JA-ile) are JA analogs (Shimizu et al., 2010; Wang et al., 2021d). Transcriptome analysis of sweet potato roots under salt stress revealed significant up-regulation of genes related to JA synthesis and signal transduction in salt-tolerant lines compared with that in salt-sensitive lines (Zhang et al., 2017). The JA and JA-ile in *Medicago truncatula* increased by 3 and 6 times, respectively, after 1 h salt stress (Domenico et al., 2019). In the present study, the genes related to JA synthesis (LOX, AOS, AOC, and OPR3) were significantly up-regulated in alfalfa roots after 1 or 6 h NaCl treatment, and meJA content increased significantly at 24 h of NaCl treatment. These results indicated that salt stress can lead to an increase in JA content in plants (Prerostova et al., 2017). It has been shown that foliar spraying of JA on soybean under salt stress reduces sodium uptake while increasing betaine content, membrane stability index, and leaf water content (Farhangi-Abriz and Ghassemi-Golezani, 2018). It has been reported that JA acts through and is central to a regulatory network of multiple hormone signals (Yang et al., 2019). Moreover, MeJA can increase GA3 levels in plants (Gao et al., 2021), and significantly increase the expression of ABF2, ABF8, and ABF10 under 24 h of NaCl stress (Lu et al., 2021). In this study, MeJA content started to rise after 1 h of NaCl stress and increased significantly at 24 h of NaCl stress, GA content increased significantly at 24 h of NaCl stress, and ABF2 was significantly up-regulated under NaCl stress for 1 h. These results indicated that GA, ABA, and JA were all involved in the defense response of alfalfa under NaCl stress to some extent, and JA was possibly a key link in this hormonal regulatory network.

4.3 Association between flavonoids and phytohormones

Overexpression of genes in the GA and JA synthesis pathways and the ABA and cytokinesis (CK) transduction pathways can change flavonoid content in plants, suggesting that flavonoid content changes may be achieved through hormonal regulation of the competition between FLS and DFR (Wan et al., 2020). Treatment of soybean with exogenous GA3 under NaCl stress has been similarly shown to increase GAs (GA1 and GA4) and JA, as well as isoflavone (daidzein and genistein) content, and reduce salt stress injury (Hamayun et al., 2010). The ABCC2 transporter protein can transport ABA, which is involved in the transport of flavonoid compounds (Behrens et al., 2019). In this study, the contents of GAs, MeJA, and isoflavones and some flavonoid compounds in alfalfa roots significantly increased at 24 h of NaCl stress, and DEGs and DAMs were also enriched in the ABC transporter pathway after 1 h of NaCl stress. The results also showed that dihydroquercetin and LUT5 of the flavonoid synthesis pathway were significantly correlated with ABF2, PP2C, and PYL genes of the hormone synthesis pathway in alfalfa roots, suggesting that GA and ABA were involved in flavonoid synthesis by regulating the expression of ABF2, PP2C, and PYL to enhance the salt resistance of alfalfa plants. Transcriptomic and metabolomic analyses of alfalfa roots revealed that 250 mM NaCl treatment resulted in alterations in several pathways, including phenylalanine metabolism, flavonoid synthesis, and α-linolenic acid metabolism, with significant up-regulation of the AOS, AOC, LOX, and OPR genes in the JA synthesis pathway. In addition, JA acts as a signaling molecule to promote flavonoid biosynthesis and induce flavonoid accumulation to adapt to salt stress (Zhang et al., 2021b). MeJA increases the isoflavone content and antioxidant activity of *Astragalus trichocarpa* roots and induces up-regulation of the CHI and CHS genes (Gai et al., 2016). MeJA treatment at 25–50 μM significantly induced anthocyanin accumulation in red fenugreek roots (Yasuhiro et al., 2010). In this study, the significant up-regulation of genes related to JA synthesis at 1 or 6 h of salt stress was the same as the expression pattern of genes related to flavonoid synthesis, and the concentration of flavonoids was significantly increased at 24 h of stress. (+)-7-Jasmonate isomethyl ester was significantly correlated with FLS (MsG0480022886.01), which is consistent with the above results.

5 Conclusion

We used transcriptomic and metabolomic approaches to study the changes in gene transcript levels and metabolite content in alfalfa plants under NaCl stress. Through the association analysis of metabolic pathways, it was found that flavonoids and phytohormones interact to jointly participate in the adaptation of plants to salt stress. The results showed that most of the flavonoid synthesis genes, as well as JA synthesis and transduction genes, were significantly up-regulated at 1 or 6 h of NaCl stress. The expression of most genes in the GA and ABA synthesis and transduction pathways also changed significantly during NaCl stress. Moreover, the levels of six flavonoids (dihydroquercetin, dihydromyricetin, anemone, lignan, glycyrrhizin, and glycine), three GA products (GA1, GA4, and GA9), and two jasmonate products ((+)-7-isomethyl jasmonate and methyl jasmonate) increased significantly at 24 h of NaCl stress. A combined analysis of DEGs and DAMs revealed that LUT5 (MsG0880047112.01) and three ABA transduction genes (MsG0180000344.01, MsG0880047375.01, MsG0380015520.01) were significantly correlated with dihydroquercetin. The flavonoid

Luteolinidin was significantly correlated with PIF4 (MsG0180003748.01) and two PP2C (MsG0180000980.01, MsG0480023807.01). (+) -7-isomethyl jasmonate was significantly correlated with the FLS (MsG0480022886.01) gene. In summary, flavonoids, GA, JA and ABA all play important roles in alfalfa's response to salt stress. The flavonoid products dihydroquercetin and luteolin were significantly correlated with several ABA and GA genes, and the JA product (+) -7-isomethyl jasmonate was significantly correlated with FLS genes. In addition, (+) -7-isomethyl jasmonate and homoeriodictyol chalcone were significantly correlated with POD (MsG0280008067.01). POD was also significantly up-regulated under NaCl stress for 6 and 24 h. These results indicate that GA, ABA, and JA may be involved in the regulation of flavonoids to improve the salt tolerance of alfalfa. JA and flavonoid-related genes and metabolites have similar expression patterns under NaCl stress, which may be the key signal to promote the synthesis of flavonoids. In conclusion, these results reveal the possible molecular mechanism of alfalfa's adaptation to salt stress, identify a number of candidate genes, and provide new insights for the study salt tolerance in alfalfa. Future studies will focus on which genes are involved, and how JA links flavonoids and various plant hormones to participate in the salt tolerance of alfalfa.

Data availability statement

The original contributions presented in the study are publicly available. This data can be found here: NCBI, PRJNA892760.

Author contributions

XW conceived and designed the research. JY, JW, and JL performed the experiments. JY analyzed the data, and wrote the paper. XW put forward valuable suggestions on the modifications of the paper. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the National Key P & D Program of China (2018YFE0108100), Jiangsu Modern Agriculture (dairy cow) industry technology system construction project (JATS [2022] 074).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1117868/full#supplementary-material

References

Behrens, C. E., Smith, K. E., Iancu, C. V., Choe, J. Y., and Dean, J. V. (2019). Transport of anthocyanins and other flavonoids by the arabidopsis ATP-binding cassette transporter AtABCC2. *Sci. Rep.* 9 (1), 437. doi: 10.1038/s41598-018-37504-8

Biala, W., Banasiak, J., Jarzyniak, K., Pawela, A., and Jasinski, M. (2017). *Medicago truncatula* ABCG10 is a transporter of 4-coumarate and liquiritigenin in the medicarpin biosynthetic pathway. *J. Exp. Bot.* 68 (12), 3231–3241. doi: 10.1093/jxb/erx059

Di Ferdinando, M., Brunetti, C., Fini, A., and Tattini, M. (2012). Flavonoids as Antioxidants in Plants Under Abiotic Stresses. In: P. Ahmad and M. Prasad (eds) *Abiotic Stress Responses in Plants*. (New York, NY: Springer). 159–179. doi: 10.1007/978-1-4614-0634-1_9

Cai, Y. P. (2012). Method for determination of antioxidant enzyme activity. *Experimental guidance of plant physiology*. (China Agricultural University Press). 108–129.

Chen, L., Xia, F., Wang, M., and Mao, P. (2021). Physiological and proteomic analysis reveals the impact of boron deficiency and surplus on alfalfa (*Medicago sativa* l.) reproductive organs. *Ecotoxicol. Environ. Saf.* 214, 112083. doi: 10.1016/j.ecoenv.2021.112083

Domenico, S. D. , Taurino, M., Gallo, A., Poltronieri, P., Pastor, V., Flors, V., et al. (2019). Oxylipin dynamics in *Medicago truncatula* in response to salt and wounding stresses. *Physiol. plantarum* 165 (2), 198–208. doi: 10.1111/ppl.12810

Farhangi-Abriz, S., and Ghassemi-Golezani, K. (2018). How can salicylic acid and jasmonic acid mitigate salt toxicity in soybean plants? *Ecotoxicol. Environ. Saf.* 147, 1010–1016. doi: 10.1016/j.ecoenv.2017.09.070

Gai, Q. Y., Jiao, J., Luo, M., Wang, W., Gu, C. B., Fu, Y. J., et al. (2016). Tremendous enhancements of isoflavonoid biosynthesis, associated gene expression and antioxidant capacity in astragalus membranaceus hairy root cultures elicited by methyl jasmonate. *Process Biochem.* 51 (5), 642–649. doi: 10.1016/j.procbio.2016.01.012

Gao, Z., Gao, S., Li, P., Zhang, Y., Ma, B., and Wang, Y. (2021). Exogenous methyl jasmonate promotes salt stress-induced growth inhibition and prioritizes defense response of *Nitraria tangutorum* bobr. *Physiol. plantarum* 172 (1), 162–175. doi: 10.1111/ppl.13314

Gong, Z. (2021). Plant abiotic stress: New insights into the factors that activate and modulate plant responses. *J. Integr. Plant Biol.* 63 (3), 429–430. doi: 10.1111/jipb.13079

Hamayun, M., Khan, S. A., Khan, A. L., Shin, J. H., Ahmad, B., Shin, D. H., et al. (2010). Exogenous gibberellic acid reprograms soybean to higher growth and salt stress tolerance. *J. Agric. Food Chem.* 58 (12), 7226–7232. doi: 10.1021/jf101221t

Ithal, N., and Reddy, A. R. (2004). Rice flavonoid pathway genes, OsDfr and OsAns, are induced by dehydration, high salt and ABA, and contain stress responsive promoter elements that interact with the transcription activator, OsC1-MYB. *Plant Sci.* 166 (6), 1505–1513. doi: 10.1016/j.plantsci.2004.02.002

Jiao, Y., Bai, Z., Xu, J., Zhao, M., Khan, Y., Hu, Y., et al. (2018). Metabolomics and its physiological regulation process reveal the salt-tolerant mechanism in glycine soja seedling roots. *Plant Physiol. biochem.: PPB* 126, 187–196. doi: 10.1016/j.plaphy.2018.03.002

Jia, X. M., Zhu, Y. F., Hu, Y., Zhang, R., Cheng, L., Zhu, Z. L., et al. (2019). Integrated physiologic, proteomic, and metabolomic analyses of malus halliana adaptation to saline-alkali stress. *Horticul. Res.* 6, 91. doi: 10.1038/s41438-019-0172-0

Kim, J., Lee, W. J., Vu, T. T., Jeong, C. Y., Hong, S. W., and Lee, H. (2017). High accumulation of anthocyanins *via* the ectopic expression of AtDFR confers significant salt stress tolerance in *Brassica napus* l. *Plant Cell Rep.* 36 (8), 1215–1224. doi: 10.1007/s00299-017-2147-7

Lei, Z. A., Zz, A., Sfa, B., Yang, L. A., and Xsa, B. (2021). Integrative analysis of metabolome and transcriptome reveals molecular regulatory mechanism of flavonoid

biosynthesis in *Cyclocarya paliurus* under salt stress - sciencedirect. *Ind. Crops Products* 170, 113823. doi: 10.1016/j.indcrop.2021.113823

Li, M., Guo, R., Jiao, Y., Jin, X., Zhang, H., and Shi, L. (2017). Comparison of salt tolerance in soja based on metabolomics of seedling roots. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.01101

Li, J., Liu, J., Zhu, T., Zhao, C., Li, L., and Chen, M. (2019). The role of melatonin in salt stress responses. *Int. J. Mol. Sci.* 20 (7), 1735-1735. doi: 10.3390/ijms20071735

Lu, J., Du, J., Tian, L., Li, M., Zhang, X., Zhang, S., et al. (2021). Divergent response strategies of CsABF facing abiotic stress in tea plant: Perspectives from drought-tolerance studies. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.763843

Mahajan, M., and Yadav, S. K. (2014). Overexpression of a tea flavanone 3-hydroxylase gene confers tolerance to salt stress and alternaria solani in transgenic tobacco. *Plant Mol. Biol.* 85 (6), 551–573. doi: 10.1007/s11103-014-0203-z

Ma, S., Lv, L., Meng, C., Zhang, C., and Li, Y. (2020). Integrative analysis of the metabolome and transcriptome of *Sorghum bicolor* reveals dynamic changes in flavonoids accumulation under saline-alkali stress. *J. Agric. Food Chem.* 68 (50), 14781–14789. doi: 10.1021/acs.jafc.0c06249

Ma, Q., Xu, X., Wang, W., Zhao, L., Ma, D., and Xie, Y. (2021). Comparative analysis of alfalfa (*Medicago sativa* l.) seedling transcriptomes reveals genotype-specific drought tolerance mechanisms. *Plant Physiol. biochem.: PPB* 166, 203–214. doi: 10.1016/j.plaphy.2021.05.008

Mokhtari, F., Rafiei, F., Shabani, L., and Shiran, B. (2017). Differential expression pattern of transcription factors across annual *Medicago* genotypes in response to salinity stress. *Biol. Plantarum* 61 (2), 227–234. doi: 10.1007/s10535-016-0666-7

Morton, M., Awlia, M., Al-Tamimi, N., Saade, S., Pailles, Y., Negrão, S., et al. (2019). Salt stress under the scalpel - dissecting the genetics of salt tolerance. *Plant journal: Cell Mol. Biol.* 97 (1), 148–163. doi: 10.1111/tpj.14189

Niharika, S., Singh, A., Khare, S., and Yadav, R. K. (2020). Mitigating strategies of gibberellins in various environmental cues and their crosstalk with other hormonal pathways in plants: a review. *Plant Mol. Biol. Rep.* 53 (4), 39, 34-49. doi: 10.1007/s11105-020-01231-0

Noor, J., Ullah, A., Saleem, M. H., Tariq, A., Ullah, S., Waheed, A., et al. (2022). Effect of jasmonic acid foliar spray on the morpho-physiological mechanism of salt stress tolerance in two soybean varieties (*Glycine max* l.). *Plants (Basel Switzerland)* 11 (5), 651. doi: 10.3390/plants11050651

Ouertani, R. ,. N., Arasappan, D., Abid, G., Ben Chikha, M., Jardak, R., Mahmoudi, H., et al. (2021). Transcriptomic analysis of salt-Stress-Responsive genes in barley roots and leaves. *Int. J. Mol. Sci.* 22 (15), 8155. doi: 10.3390/ijms22158155

Panda, A., Rangani, J., and Parida, A. K. (2021). Unraveling salt responsive metabolites and metabolic pathways using non-targeted metabolomics approach and elucidation of salt tolerance mechanisms in the xero-halophyte *Haloxylon salicornicum*. *Plant Physiol. Biochem.* 158, 284–296. doi: 10.1016/j.plaphy.2020.11.012

Parveen, A., Ahmar, S., Kamran, M., Malik, Z., and Ali, S. (2021). Abscisic acid signaling reduced transpiration flow, regulated na^+ ion homeostasis and antioxidant enzyme activities to induce salinity tolerance in wheat (*triticum aestivum* l.) seedlings. *Environ. Technol. Innovation* 413), 101808. doi: 10.1016/j.eti.2021.101808

Poór, P., Borbély, P., Czékus, Z., Takács, Z., Ördög, A., Popović, B., et al. (2019). Comparison of changes in water status and photosynthetic parameters in wild type and abscisic acid-deficient sitiens mutant of tomato (*Solanum lycopersicum* cv. *Rheinlands ruhm*) exposed to sublethal and lethal salt stress. *J. Plant Physiol.* 232, 130–140. doi: 10.1016/j.jplph.2018.11.015

Prerostova, S., Dobrev, P. I., Gaudinova, A., Hosek, P., Soudek, P., Knirsch, V., et al. (2017). Hormonal dynamics during salt stress responses of salt-sensitive *Arabidopsis thaliana* and salt-tolerant *Thellungiella salsuginea*. *Plant sci.: an Int. J. Exp. Plant Biol.* 264, 188–198. doi: 10.1016/j.plantsci.2017.07.020

Shahzad, R., Khan, A. L., Bilal, S., Waqas, M., Kang, S. M., and Lee, I. J. (2017). Inoculation of abscisic acid-producing endophytic bacteria enhances salinity stress tolerance in *Oryza sativa*. *Environ. Exp. Bot.* 136, 68–77. doi: 10.1016/j.envexpbot.2017.01.010

Shi, P., and Gu, M. (2020). Transcriptome analysis and differential gene expression profiling of two contrasting quinoa genotypes in response to salt stress. *BMC Plant Biol.* 20 (1), 568. doi: 10.1186/s12870-020-02753-1

Shimizu, Y., Maeda, K., and Shimomura, K. K. (2010). Methyl jasmonate induces anthocyanin accumulation in gynura bicolor cultured roots. *In Vitro Cell. Dev. Biol. Plant* 46 (5), 460–465. doi: 10.1007/s11627-010-9294-7.a

Song, L., Jiang, L., Chen, Y., Shu, Y., Bai, Y., and Guo, C. (2016). Deep-sequencing transcriptome analysis of field-grown *Medicago sativa* l. crown buds acclimated to freezing stress. *Funct. Integr. Genomics* 16 (5), 495–511. doi: 10.1007/s10142-016-0500-5

Wang, Y., Liu, J., Yang, F., Zhou, W., Mao, S., Lin, J., et al. (2021b). Untargeted LC-MS-based metabolomics revealed specific metabolic changes in cotyledons and roots of *Ricinus communis* during early seedling establishment under salt stress. *Plant Physiol. biochem.: PPB* 163, 108–118. doi: 10.1016/j.plaphy.2021.03.019

Wang, X. S., and Han, J. G. (2007). Effects of NaCl and silicon on ion distribution in the roots, shoots and leaves of two alfalfa cultivars with different salt tolerance. *Soil Sci. Plant Nutr.* 53, 278–285. doi: 10.1111/j.1747-0765.2007.00135.x

Wang, Y., Mostafa, S., Zeng, W., and Jin, B. (2021d). Function and mechanism of jasmonic acid in plant responses to abiotic and biotic stresses. *Int. J. Mol. Sci.* 22 (16), 8568. doi: 10.3390/ijms22168568

Wang, W., Pang, J., Zhang, F., Sun, L., Yang, L., Fu, T., et al. (2022). Salt–responsive transcriptome analysis of canola roots reveals candidate genes involved in the key metabolic pathway in response to salt stress. *Sci. Rep.* 12 (1), 1666. doi: 10.1038/s41598-022-05700-2

Wang, M., Ren, T., Huang, R., Li, Y., Zhang, C., and Xu, Z. (2021c). Overexpression of an apocynum venetum flavonols synthetase gene confers salinity stress tolerance to transgenic tobacco plants. *Plant Physiol. Biochem.* 162, 667–676. doi: 10.1016/j.plaphy.2021.03.034

Wang, F., Ren, G., Li, F., Qi, S., Xu, Y., Wang, B., et al. (2018). A chalcone synthase gene aechs from abelmoschus esculentus regulates flavonoid accumulation and abiotic stress tolerance in transgenic arabidopsis. *Acta Physiol. Plantarum* 40 (5), 97. doi: 10.1007/s11738-018-2680-1

Wang, M., Zhang, Y., Zhu, C., Yao, X., Zheng, Z., Tian, Z., et al. (2021a). EkFLS overexpression promotes flavonoid accumulation and abiotic stress tolerance in plant. *Physiol. plantarum* 172 (4), 1966–1982. doi: 10.1111/ppl.13407

Wan, L., Lei, Y., Yan, L., Liu, Y., Pandey, M. K., Wan, X., et al. (2020). Transcriptome and metabolome reveal redirection of flavonoids in a white testa peanut mutant. *BMC Plant Biol.* 20 (1), 161. doi: 10.1186/s12870-020-02383-7

Xu, Z., Chen, X., Lu, X., Zhao, B., Yang, Y., and Liu, J. (2021). Integrative analysis of transcriptome and metabolome reveal mechanism of tolerance to salt stress in oat (*Avena sativa* l.). *Plant Physiol. biochem.: PPB* 160, 315–328. doi: 10.1016/j.plaphy.2021.01.027

Yang, J., Duan, G., Li, C., Liu, L., Han, G., Zhang, Y., et al. (2019). The crosstalks between jasmonic acid and other plant hormone signaling highlight the involvement of jasmonic acid as a core component in plant response to biotic and abiotic stresses. *Front. Plant Sci.* 10. doi: 10.3389/fpls.2019.01349

Yang, X., Zhang, Y., Lai, J. L., Luo, X. G., Han, M. W., Zhao, S. P., et al. (2021). Analysis of the biodegradation and phytotoxicity mechanism of TNT, RDX, HMX in alfalfa (*Medicago sativa*). *Chemosphere* 281, 130842. doi: 10.1016/j.chemosphere.2021.130842

Yasuhiro, S., Kazuhiro, M., Mika, K., and Koichiro, S. (2010). Methyl jasmonate induces anthocyanin accumulation in Gynura bicolor cultured roots. In Vitro Cellular and Development Biology. *Plant: Journal of the Tissue Culture Association* 46 (5), 460–465. doi: 10.1007/s11627-010-9294-7

Zhang, X., Liu, P., Qing, C., Yang, C., Shen, Y., and Ma, L. (2021). Comparative transcriptome analyses of maize seedling root responses to salt stress. *Peer J.* 9, e10765. doi: 10.7717/peerj.10765

Zhang, M., Yu, Z., Zeng, D., Si, C., Zhao, C., Wang, H., et al. (2021b). Transcriptome and metabolome reveal salt-stress responses of leaf tissues from dendrobium officinale. *Biomolecules* 11 (5), 736. doi: 10.3390/biom11050736

Zhang, J., Zhang, Y., Du, Y., Chen, S., and Tang, H. (2011). Dynamic metabonomic responses of tobacco (*Nicotiana tabacum*) plants to salt stress. *J. Proteome Res.* 10 (4), 1904–1914. doi: 10.1021/pr101140n

Zhang, H., Zhang, Q., Zhai, H., Li, Y., Wang, X., Liu, Q., et al. (2017). Transcript profile analysis reveals important roles of jasmonic acid signalling pathway in the response of sweet potato to salt stress. *Sci. Rep.* 7, 40819. doi: 10.1038/srep40819

Zhao, B. Y., Hu, Y. F., Li, J. J., Yao, X., and Liu, K. D. (2016). BnaABF2, a bZIP transcription factor from rapeseed (*Brassica napus* l.), enhances drought and salt tolerance in transgenic arabidopsis. *Botanical Stud.* 57 (1), 12. doi: 10.1186/s40529-016-0127-9

Zhou, Y., Yang, P., Cui, F., Zhang, F., Luo, X., and Xie, J. (2016). Transcriptome analysis of salt stress responsiveness in the seedlings of dongxiang wild rice (*Oryza rufipogon* griff.). *PloS One* 11 (1), e0146242. doi: 10.1371/journal.pone.0146242

Zhu, T., Lin, J., Zhang, M., Li, L., and Chen, M. (2019). Phytohormone involved in salt tolerance regulation of elaeagnus angustifolia l. seedlings. *J. For. Res.* 24 (4), 235–242. doi: 10.1080/13416979.2019.1637995

10.3389/fpls.2023.1117868

Glossary

ABA	abscisic acid
ABF2	ABA responsive element binding factor 2
ANS	anthocyanin synthase
AOC	allene oxide cyclase
AOS	allene oxide synthase
CAT	catalase
CHI	chalcone isomerase
CHS	chalcone synthase
DAM	differential accumulation metabolite
DEG	differentially expressed gene
DFR	dihydroflavanol reductase
ERF1	ethylene-responsive transcription factor 1
F3H	flavanone hydroxylase gene
FLS	flavonol synthase
GA	gibberellin
GA20ox	gibberellin 20 oxidase
GA3ox	gibberellin 3beta-dioxygenase
GO	Gene Ontology
IFS	2-hydroxyisoflavanone synthase
JA	jasmonic acid
JA-ile	jasmonic acid isoleucine coupling
KEGG	Kyoto Encyclopedia of Genes and Genomes
KIN2	stress-induced protein kinase 2
LOX	13-lipoxygenase
LUT5	beta-ring hydroxylase
MDA	malonaldehyde
MeJA	jasmonic acid methyl ester
miaA	tRNA dimethylallyl transferase
NCED	9-cis-epoxycarotenoid dioxygenase
OPR3	12-oxophytodienoic acid reductase
PCA	principal component analysis
PIF3/4	phytochrome-interacting factor 3/4
PLS-DA	partial least-squares discriminant analysis
POD	peroxidase
PP2C	protein phosphatase PP2C
PR1	pathogenesis-related protein 1
PRO	proline
PYL	ABA receptor PYL2
QC	quality control
RR	response regulator
SNRK2	serine/threonine-protein kinase SRK2
SOD	superoxide dismutase
VIP	variable important for the projection
VSP2	vegetative storage protein 2

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 07 February 2023
DOI 10.3389/fpls.2023.1105521

OPEN ACCESS

EDITED BY
Shangang Jia,
China Agricultural University, China

REVIEWED BY
Pu Liu,
Anhui Agricultural University, China
Xiaosan Huang,
Nanjing Agricultural University, China
Bin Tan,
Henan Agricultural University, China

*CORRESPONDENCE
Dingli Li
qauldl@163.com

†These authors have contributed equally to this work

SPECIALTY SECTION
This article was submitted to Functional and Applied Plant Genomics, a section of the journal Frontiers in Plant Science

RECEIVED 22 November 2022
ACCEPTED 16 January 2023
PUBLISHED 07 February 2023

CITATION
Cheng Y, Liang C, Qiu Z, Zhou S, Liu J, Yang Y, Wang R, Yin J, Ma C, Cui Z, Song J and Li D (2023) Jasmonic acid negatively regulates branch growth in pear. *Front. Plant Sci.* 14:1105521. doi: 10.3389/fpls.2023.1105521

Jasmonic acid negatively regulates branch growth in pear

Yuanyuan Cheng[1†], Chenglin Liang[2†], Zhiyun Qiu[1], Siqi Zhou[1], Jianlong Liu[1], Yingjie Yang[1], Ran Wang[1], Jie Yin[1], Chunhui Ma[1], Zhenhua Cui[1], Jiankun Song[1] and Dingli Li[1*]

[1]Qingdao Key Laboratory of Genetic Improvement and Breeding in Horticultural Plants, Engineering Laboratory of Genetic Improvement of Horticultural Crops of Shandong Province, College of Horticulture, Qingdao Agricultural University, Qingdao, China, [2]Haidu College, Qingdao Agricultural University, Laiyang, China

The quality of seedlings is an important factor for development of the pear industry. A strong seedling with few branches and suitable internodes is ideal material as a rootstock for grafting and breeding. Several branching mutants of pear rootstocks were identified previously. In the present study, 'QAU-D03' (*Pyrus communis* L.) and it's mutants were used to explore the mechanism that affects branch formation by conducting phenotypic trait assessment, hormone content analysis, and transcriptome analysis. The mutant plant (MP) showed fewer branches, shorter 1-year-old shoots, and longer petiole length, compared to original plants (OP), i.e., wild type. Endogenous hormone analysis revealed that auxin, cytokinin, and jasmonic acid contents in the stem tips of MP were significantly higher than those of the original plants. In particular, the jasmonic acid content of the MP was 1.8 times higher than that of the original plants. Transcriptome analysis revealed that *PcCOI1*, which is a transcriptional regulatory gene downstream of the jasmonic acid signaling pathway, was expressed more highly in the MP than in the original plants, whereas the expression levels of *PcJAZ* and *PcMYC* were reduced in the MP compared with that of the original plants. In response to treatment with exogenous methyl jasmonate, the original plants phenotype was consistent with that of the MP in developing less branches. These results indicate that jasmonic acid negatively regulates branch growth of pear trees and that jasmonic acid downstream regulatory genes play a crucial role in regulating branching.

KEYWORDS

pear, branch, JA, transcriptional regulation, less branching mutant plant

1 Introduction

Grafting is a common method for asexual propagation of plants and for raising fruit tree seedlings. Pear rootstock with a strong stem, few branches, and suitable internodes is ideal material for grafting and breeding. Branch development and growth are regulated by a complex network of factors involving plant growth and development, hormone regulation, transcriptional regulation, and other factors.

Bud mutation is an important means of causing changes in branching through plant somatic variation (Zhao et al., 2021). The meristem cells of buds undergo genetic changes during cell division, leading to the development of lateral branches (Leng et al., 2021). The shoot apical meristem (SAM) is also crucial to plant development and is responsible for the development of leaves, stems, and flowers (Han et al., 2019). The SAM gives rise to leaf primordia during the vegetative growth stage and an inflorescence meristem in the transitional stage from vegetative growth to reproductive growth. An axillary meristem (AM) gives rise to a lateral bud primordium, which develops into a lateral bud and ultimately forms lateral branches (Grbić and Bleecker, 2000; Tanaka et al., 2013;). The development of plant meristems mainly depends on the population of multifunctional stem cells in plants, and is affected by hormones and environmental factors.

Plant growth and development adapt to environmental changes through hormone signals, which provide a biochemical connection between the environment and cellular responses (Mur et al., 2006; Clarke et al., 2009). The auxin and cytokinin regulatory pathways are the main factors that regulate meristem formation and differentiation (Depuydt and Hardtke, 2011). Auxin combines with brassinolide (BR) to induce *SMALL AUXIN UPREGULATED RNA 10 (SAUR10)*, thus participating in the regulation of branching angle (Bemer et al., 2017). Cytokinin is considered to be a second messenger that transmits auxin signals to lateral buds (Leyser, 2003). Cytokinin plays an important role in regulating leaf tip dominance and axillary bud growth. The expression level of *PsIPT*, a critical enzyme involved in cytokinin biosynthesis in pea, increases at the node after decapitation. Bud growth after decapitation is caused by the local accumulation of cytokinin (Tanaka et al., 2006). In poplar, overexpression of *GA2ox* leads to an increase in the number of tillers or branches, which indicates that gibberellin may play a negative role in the control of poplar bud branching (Zawaski and Busov, 2014). *RAX1* promotes early formation of the AM, negatively regulates the gibberellin content of the stem tip, and affects the timing of AM development (Fambrini et al., 2017; Nie et al., 2018). Jasmonic acid is involved in regulating responses to biological and abiotic stresses, as well as growth and development of plants (Wasternack and Hause, 2013; Campos et al., 2014). In the jasmonic acid signaling pathway, *COI1*, as a downstream regulatory element of jasmonic acid, participates in almost all JAs regulatory processes (Ye et al., 2012). The *COI1* gene encodes a F-box protein, which is a component of E3 ubiquitin ligase (Xie et al., 1998). The enzyme is associated with meristem arrest and apical dominance (Zhai et al., 2015). The *coi1* mutant shows strong apical dominance and enhanced meristem longevity (Kim et al., 2013). *Jasmonate ZIM-domain (JAZ)* proteins play an inhibitory role in the JA signaling pathway (Chini et al., 2007; Thines et al., 2007; Yan et al., 2007). Overexpression of *JAZ* can cause increase in the number of lateral branches in tomato (Yu et al., 2018).

The regulation of transcription is important in the control of eukaryotic gene expression. Transcription factors are involved in all aspects of plant growth and development. The *SQUAMOSA PROMOTER BINDING PROTEIN-LIKE 13 (SPL13)* gene encodes a *SBP* transcription factor, which is mainly expressed in meristems and is critical to regulating the branching and vegetative growth of alfalfa plants (Gao et al., 2018). Overexpression of *SPL13* inhibits the growth of axillary buds and reduces the number of lateral branches. *MsMYB112* RNA interference promotes branching, which indicates that *MYB112* inhibits the growth of lateral branches in alfalfa (Stracke et al., 2001; Gao et al., 2018;). In addition, *TCP* transcription factors are involved in lateral meristem growth, cell proliferation, and regulatory hormone effects in many cases (Aguilar-Martínez et al., 2007; Martin-Trillo and Cubas, 2010). *PpTCP18* controls peach branching by positive feedback regulation of SL biosynthesis (Wang et al., 2022). In apple, *MdWUS2* can regulate branching by inhibiting *MdTCP12* expression (Li et al., 2021). The *TB1* transcription factor, also known as *FINECULM 1 (FC1)*, is a member of the *TCP* family that negatively regulates rice tillering and inhibits the subsequent growth of axillary buds (Takeda et al., 2003). The *TB1* gene is expressed at the base of the axillary bud and SAM. Its overexpression leads to a significant reduction in number of tillers, whereas in the *tb1* mutant an increased number of tillers develop (Wai and An, 2017). The transcription factor *ERF BUD ENHANCE* (*EBE*) affects cell proliferation, axillary bud growth, and branching of *Arabidopsis thaliana*. This gene encodes an *AP2/ERF* transcription factor and is highly expressed in proliferating cells (Mehrnia et al., 2013). Genes involved in plant hormone biosynthesis, transduction, and SAM formation are also associated with branching (Kurakawa et al., 2007; Ligerot et al., 2017; Zhang et al., 2018). These studies indicate that branching development is a complex process in plants.

The growth characteristics of pear rootstock have important influence on grafting effect. Sturdy seedlings and few branches make ideal stock material. In this study, the mechanism affecting pear branching formation were explored through phenotypic trait assessment, hormone content analysis and transcriptomic analysis, by using pear 'QAU-D03' (*Pyrus communis* L.) and less-branching mutants as materials. The results provide novel insights to improve understanding of the molecular mechanism of branch development in pear, which is of considerable importance for rootstock breeding.

2 Materials and methods

2.1 Plant materials

The original plant (OP) with more branches was a seedling progeny of 'QAU-D03' (*Pyrus communis* L.). The shoots of OP were exposed to gamma radiation at 50 Gy (^{60}Co source, 10 Gy min^{-1}) and then grafted onto one-year-old rootstocks of *Pyrus betulifolia* Bunge. Among the resulting trees, one less-branching mutant was found and named MP. One-year-old shoots of OP and MP plants were collected from the Jiaozhou Demonstration Park in Qingdao, Shandong Province, China, for tissue culture to obtain tissue-cultured plantlets. The plants were grown on MS medium with 1.0 mg·L^{-1} 6-BA, 0.1 mg·L^{-1} IBA, 30 g·L^{-1} sucrose, and 7 g·L^{-1} agar in a tissue culture room. Tissue-cultured shoots with 20 days of uniform growth were selected for rooting. Once the shoots had produced 3-4 roots of length 2-3 cm, the plantlets were transplanted to a mixture of perlite:vermiculite:peat (1:1:1, v/v/v). Samples were collected at the shoot dormancy in mid-March and bud expansion stages at the end of April, with three biological replicates per sample, for transcriptome sequencing analysis.

2.2 Measurement of growth traits

Annual branches from the periphery of crown of OP and MP plants, and mature leaves were randomly selected to observe leaf

characteristics. Each treatment contains at least three biological replicates. The following quantitative characters were determined, including annual branch length, annual branch thickness, internode length, leaf length, leaf width and petiole length. The OP and MP of phenotypic difference was evaluated based on characteristics of pear plant.

2.3 Histological observation of vegetative organs

Buds and stems of MP and OP plants were sampled for conventional paraffin-embedded sectioning. The samples were fixed in formaldehyde–alcohol–acetic acid for 1 d and then dehydrated in an ethanol series (50%, 60%, 70%, 80%, and 95%) for 60 min at each step. Samples were placed in 100% ethanol and left overnight. After decoloration with dimethylbenzene, the samples were embedded in paraffin. Sections (8 μm) were cut using a rotary microtome (HI1220, Leica, Nussloch, Germany), then dewaxed, rehydrated, cleaned, stained with toluidine blue, and the coverslip mounted with neutral balata. Sections were observed and photographed using an optical microscope (RM2235, Leica, Nussloch, Germany). In addition, buds were observed by scanning electron microscopy. After manual removal of the bud scales, the buds were fixed, washed, and dehydrated as described. The material was dried by carbon dioxide critical point drying and sputter-coated with gold. The material was observed and photographed with a scanning electron microscope (JSM-7500F, JEOL, China).

2.4 Hormone content determination

Fresh plant stem tips material (0.2-1.0 g) was ground in an ice-cooled mortar in 10 mL of 80% (v/v) methanol extraction medium containing 1 mM butylated hydroxytoluene as an antioxidant. The extract was incubated at 4°C for 4 h and then centrifuged at 4000 rpm for 15 min at 4°C. The supernatant was filtered through Chromosep C18 columns (C18 Sep-Park Cartridge, Waters Corp., Milford, MA, USA), and prewashed with 10 mL of 100% (w/v) and 5 mL of 80% (v/v) methanol. The hormone fractions eluted with 10 mL of 100% (v/v) methanol and 10 mL ether from the columns were dried under N2 gas, dissolved in 2 mL phosphate buffer saline containing 0.1% (v/v) Tween 20 and 0.1% (w/v) gelatin (pH 7.5) for analysis by an enzyme linked immunosorbent assay (ELISA) (Bollmark et al., 1988). The ELISA was performed in a 96-well microtitration plate. Each well was coated with 100 μL coating buffer (1.5 g·L^{-1} Na_2CO_3, 2.93g·L^{-1} $NaHCO_3$, and 0.02g·L^{-1} NaN_3, pH 9.6) containing 0.25 μg·mL^{-1} antigens against the hormones. The coated plates were incubated for 4 h at 37°C and then kept at room temperature for 30–40 min. The plate was incubated for 3 h at 28°C for measurement of dihydrozeatin riboside (DHZR), zeatin riboside (ZR), brassinolide (BR), methyl jasmonate (JA-Me), indole-3-acetic acid (IAA), indolepyruvic acid (IPA), gibberellins (GAs), and overnight at 4°C for IAA, and then washed as described above. Color development in each well was detected using an ELISA Reader (EL310, BioTek, Winooski, VT, USA) at an optical density of A490. The contents of DHZR, ZR, BR, JA-Me, IAA, IPA, GAs, and ABA were calculated following the method of Weiler et al. (1981).

2.5 RNA extraction and transcriptome sequencing

Plant total RNA isolation kit (TaKaRa, Beijing, China) was used to extract total RNA from samples, following the manufacturer's instructions, and different samples were subjected to three biological replicates. After qualification using a bioanalyzer (2,100, Agilent, United States), 1 μg of each sample was used for cDNA library construction. Four RNA sequencing libraries were constructed, in which MP1 and OP1 were the dormant shoot tips, MP2 and OP2 were the buds in expansion period. For library construction, 1 μg of RNA per sample was used as the input material. Library quality was assessed with a 2100 Bioanalyzer system (Agilent, Santa Clara, CA, USA). The clean dataset was obtained by removing reads containing the adapter sequence, poly-N, and low-quality reads from the raw data. The Q20, Q30, and GC content of the clean data were calculated. All downstream analyses used the high-quality clean data. The reference genome (*Pyrus communis* Bartlett DH Genome v2.0) and gene model annotation files were downloaded from GDR database (https://www.rosaceae.org/species/pyrus/pyrus_communis/genome_v2.0). An index of the reference genome was generated using HISAT2 v2.0.5 and paired-end clean reads were aligned to the reference genome using HISAT2 v2.0.5. FeatureCounts v1.5.0-p3 was used to count the read numbers mapped to each gene. The fragments per kilobase of exon per million mapped fragments (FPKM) value of each gene was calculated based on the length of the gene and number of reads mapped to the gene. Differential expression analysis of two conditions/groups (two biological replicates per condition) was performed using the DESeq2 R package (v1.20.0). Gene ontology (GO) enrichment analysis of differentially expressed genes (DEGs) was implemented with the clusterProfiler R package for which gene length bias was corrected. Statistical enrichment of DEGs in KEGG pathways was detected with the clusterProfiler R package.

2.6 Real-time quantitative PCR analysis

Total RNA was extracted from pear using the RNAprep Pure Plant Kit (Tiangen Biotech Co., Beijing, China). The cDNA was synthesized using the HiScript II 1st Strand cDNA Synthesis Kit (Vazyme Biotech Co., Nanjing, China). The Lightcycler® 480 II System (Roche, Basel, Switzerland) and the ChamQ SYBR Color qPCR Master Kit (Vazyme Biotech Co.) were used to estimate relative gene expression levels under the different treatments. The reaction system (20 μL total volume) consisted of 2 μL template cDNA, 1 μL each forward and reverse primer, 10 μL Supermix, and 6 μL RNA-free water. The reaction protocol was as follows: 95°C for 5 min, then 45 cycles at 95°C for 15 s, 60°C for 30 s, and 72°C for 30 s. The Actin gene was used as an internal control. The relative expression level for each gene was calculated with the $2^{-\Delta\Delta Ct}$ method. Each sample analysis was repeated three times. The primers used are listed in Supplementary Table S1.

2.7 Hormone treatment

The seedlings of 'OP' and 'MP' clones were selected and subcultured with Murashige and Skoog (MS) medium (Coolaber, Coolaber Science & Technology Co.,Ltd., China). The seedlings were divided into different treatments: (i) 'OP' and 'MP' control: continued use of the MS nutrient solution; (ii) 'OP' and 'MP' treated with JA-Me: MS with 100 μmol L^{-1} Methyl jasmonate (Macklin, Shanghai Macklin Biochemical Technology Co., Ltd., China). After JA-Me treatment, observations were carried out after 10 days.

2.8 Statistical analysis

Statistical analysis was performed using IBM SPSS Statistics 23.0 (IBM Corporation, Armonk, NY, USA). Values are presented as the mean ± SD of at least three independent biological replicates. The significance of differences between means was analyzed with Duncan's multiple range test or Student's t-test. The probability level $p < 0.05$ was considered to be significant.

3 Results

3.1 Comparison of OP and MP phenotypes

The wild type (OP) plants had more branches, whereas the mutant (MP) plants developed fewer branches and the leaves were larger (Figure 1A). The leaf lamina base of MP was broadly wedge-shaped or rounded, and the tip was tapered or blunt, whereas the OP leaf lamina base was mostly wedge-shaped or broadly wedge-shaped, and the tip was acute (Figure 1B). The mutant (MP) developed fewer branches, shorter 1-year-old shoots, and longer petiole, compared to OP, whereas no significant difference in leaf length, leaf width, leaf shape index, and annual branch thickness between OP and MP were observed (Table 1). With further development, the axillary meristem

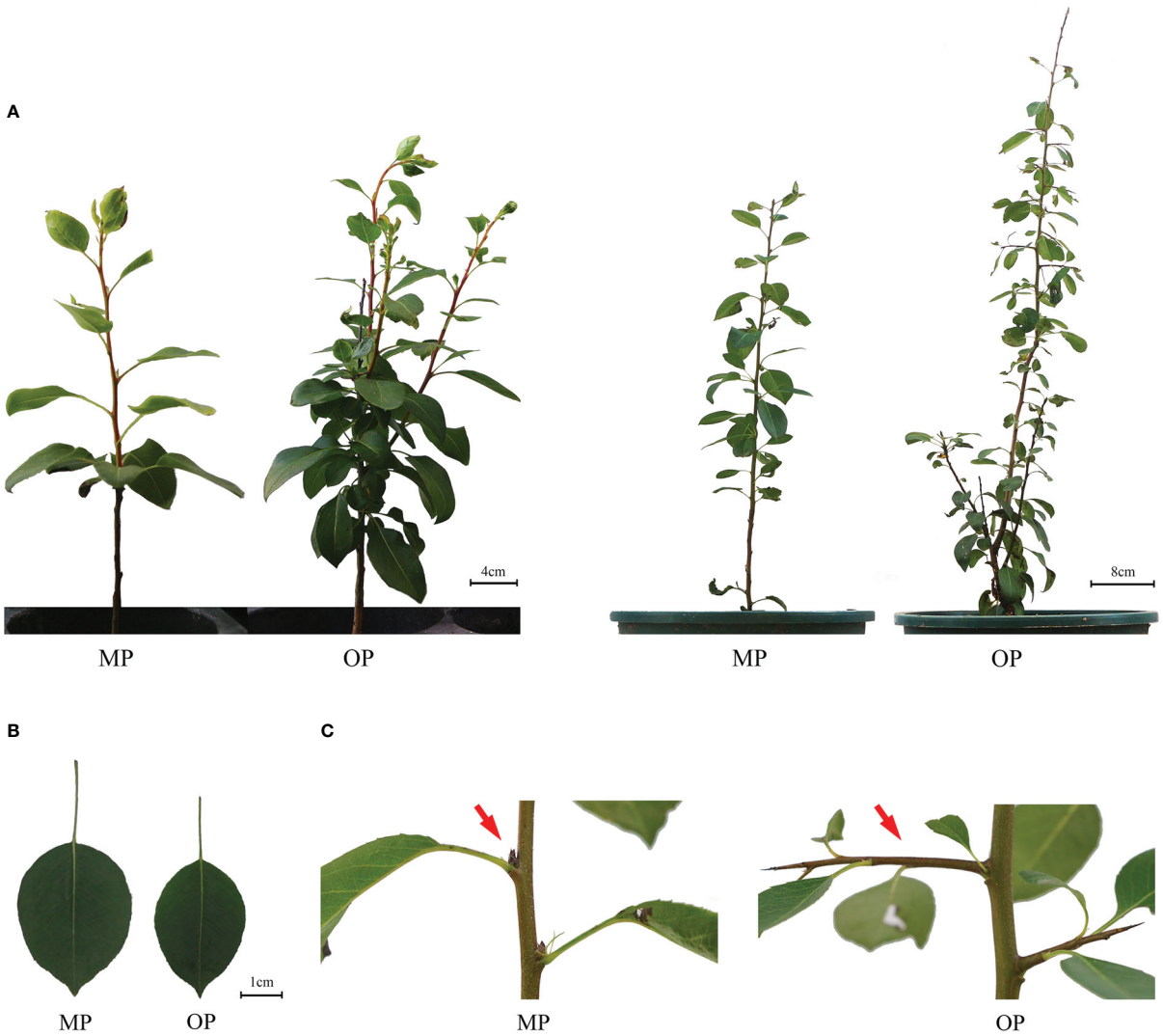


FIGURE 1
Phenotypes of the wild type (OP) and mutant (MP). **(A)** Plant phenotype of the same plant after growth for 9 months (left plant) and 12 months (right plant). **(B)** Leaf phenotype of MP and OP plants. **(C)** Axillary buds of MP and OP. The red arrow indicates transformation of the axillary meristem into a spur shoot (thorn) in OP and a dormant bud containing an undifferentiated axillary meristem in MP.

of OP developed into a spine-tipped spur shoot, whereas in MP axillary bud outgrowth did not occur (Figure 1C). All the above results confirmed that MP has the morphological characteristics of large leaves, short branches, few branches and no spines. Compared with OP, there were fewer branches in the early stage and no thorns in the leaf axils in the late stage, both of which were related to the activity of axillary meristem.

3.2 Anatomy characteristics of the leaf, stem, and axillary bud

The MP vascular bundle was larger than that of OP in the leaf, stem, and axillary buds at different development stages (Figure 2A; Table 2). The pith radius, xylem width and phloem width in the stem of new shoots of OP increased with stem development, and the phloem width changed most obviously (Figure 2B; Table 3). Axillary buds were divided into four developmental stages, namely bud primordium initiation preparation stage (I), bud primordium initiation stage (II), bud primordium formation stage (III), and bud primordium maturity stage (IV). Stage I was observed in the axillary buds of MP and OP; a group of darkly stained cells were present above the bud base, in the leaf axils of the new shoot, which represented the bud primordium (Figure 2C). With the further development of the bud primordium, the bud primordium entered the initiation stage (II). The growth point differentiated scale primordia from the outer to the inner, and the outermost is the scale developed from the scale primordium (LP). At the tip, the bud primordium formed the SAM and gradually differentiated into the scale leaf and bud primordium. When OP buds were in stage II, the MP buds had already entered stage III. In stage III, the bud base began to form the bud primordium, which was round and spherical, and gradually developed into a triangular form with time. In stage IV, the bud base formed a branch primordium and bracts one by one. When MP were in stage IV, the OP buds were still in stage III (Figure 2C). Observation of dormant axillary buds by scanning electron microscopy revealed that branch primordia and bract primordia were present in both MP and OP (Figure 2D). The above results showed that the bud structure of MP and OP were in different developmental stages at the same time, and the bud development of MP was faster than that of OP.

TABLE 1 Analysis of quantitative morphological characters between the mutant (MP) and wild type (OP).

Character	OP	MP
Leaf length (cm)	3.73 ± 0.43 a	4.25 ± 0.41 a
Leaf width (cm)	2.61 ± 0.30 a	3.05 ± 0.27 a
Leaf shape index	1.43 ± 0.17 a	1.41 ± 0.12 a
Petiole length (cm)	1.80 ± 0.40 a	2.97 ± 0.57 b
Internode length (cm)	1.88 ± 0.25 a	2.09 ± 0.25 b
Annual branch length (cm)	3.22 ± 0.92 a	1.36 ± 0.44 b
Annual branch thickness (cm)	0.38 ± 0.06 a	0.34 ± 0.04 a
Number of branches	2.30 ± 0.49 a	0.42 ± 0.53 b

Data are the mean ± standard error (n = 3). Different lowercase letters within a column indicate a statistically significant difference ($p < 0.05$, Student's test).

3.3 Determination of endogenous hormone contents

Endogenous hormone analysis revealed that the content of GA4 in OP was significantly higher than that in MP, whereas the contents of auxin (IAA), methyl jasmonate (JA-Me), indolepro pionic acid (IPA), trans-zeatin-riboside (ZR), brassinolide (BR), and dihydrozeatin-riboside (DHZR) in MP were higher than those in OP. The difference in JA-Me content between OP and MP was most marked (Figure 3). Based on these results and the following transcriptome analysis, we speculate that jasmonic acid may play a role in regulating pear branching.

3.4 Gene ontology and KEGG pathway analysis of DEGs

The MP and OP shoot tips were sampled for transcriptome analysis. A total of 334 DEGs were common to the MP1 vs. OP1 and MP2 vs. OP2 comparison groups. Enrichment analysis of the GO terms indicated that the DEGs were mainly enriched in the apoplast in the biological process category, extracellular regions in the cellular component category, and xyloglucosyl transferase activity in the molecular function category (Figure 4A). The enrichment of KEGG pathways indicated that the DEGs were mainly enriched in the pathways of plant-pathogen interaction, fatty acid elongation, and nitrogen metabolism (Figure 4B). Thirty-eight of the 334 DEGs were associated with plant growth, cell division and differentiation, and SAM activity. Among these 38 DEGs, 23 genes encoded transcription factors, including members of the *AP2/ERF*, *P450/CYP*, *MYB*, *WRKY*, *TCP*, and *NAC* families (Figure 4C).

3.5 DEGs involved in plant hormone signaling pathways

The hormone-related DEGs were screened and the changes in their expression in the two control groups were analyzed. Among these DEGs, most genes were upregulated in OP2, including IAA, JA-Me, Ethylene, ABA, CTK. Genes involved in the gibberellin signaling pathway were highly expressed in OP1 (Figure 5A). With regard to the JA biosynthesis and regulation pathways, 35 detected genes were associated with the octadecane synthesis pathway starting from α-linolenic acid, including genes encoding plant dienoic acid reductase (OPR3), acetyl coenzyme A oxidase (ACX), peroxisome fatty acid oxidized multifunctional protein (MFP2), and ketoethyl coenzyme A thiolase (3-KAT2), which are all involved in JA biosynthesis in the peroxisome. *COI1-related* genes were highly expressed in MP, whereas *JAZ* protein-related genes and *MYC-related* transcription factors were highly expressed in OP (Figure 5B). In OP and MP, the expressions of *COI1-related* genes and *JAZ* protein-related genes showed regular changes. We speculated that the phenomenon of plant branching might be related to *COI1* and *JAZ*.

10.3389/fpls.2023.1105521

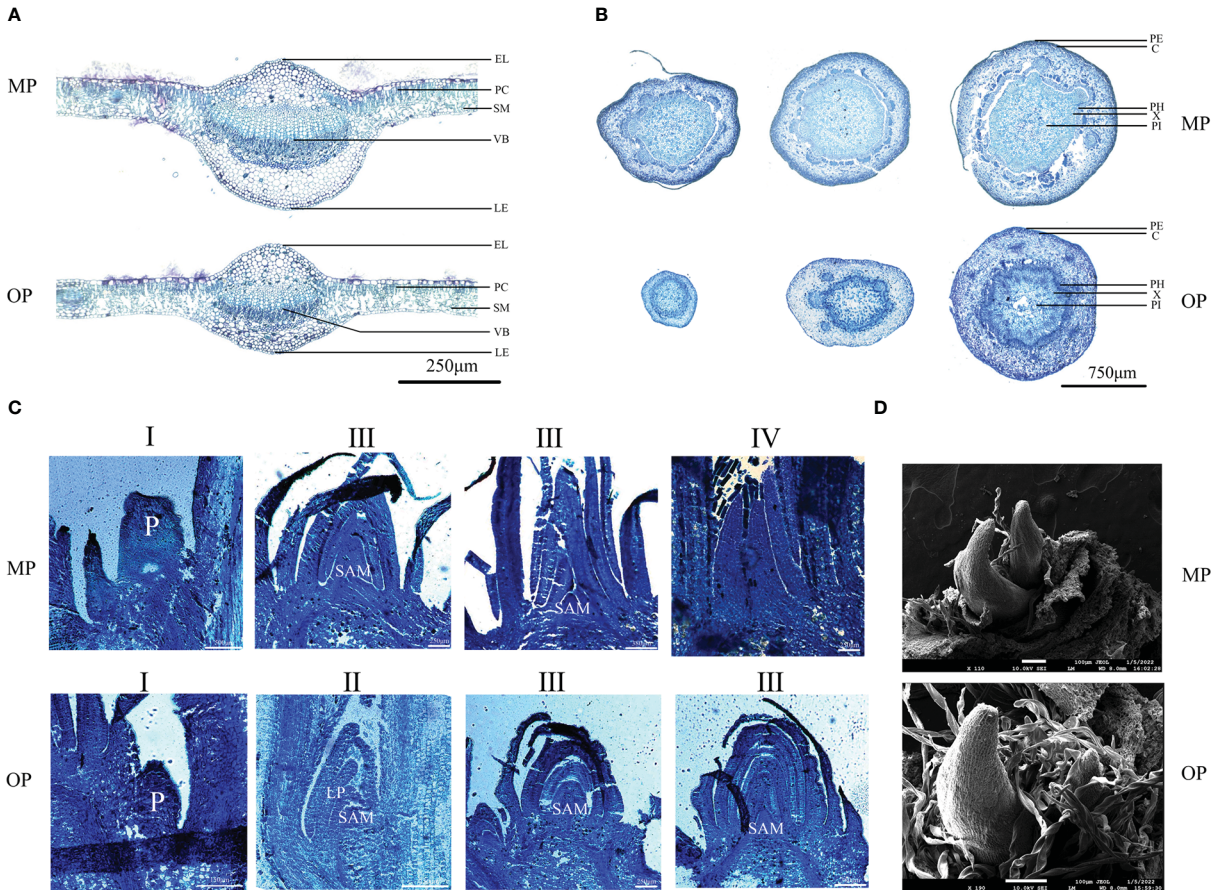


FIGURE 2
Anatomy of the vegetative organs of the wild type (OP) and mutant (MP). **(A)** Transverse sections of the leaf. EL, epithelial layer; PC, palisade mesophyll; SM, spongy mesophyll; VB, vascular bundle; LE, abaxial epidermis. **(B)** Transverse sections of the stem of the new shoots. Sections from leaf to right are from the shoot tip, mid-stem, and the shoot base, respectively. PE, periderm; C, cortex; PH, phloem; X, xylem; PI, pith. **(C)** Longitudinal sections of bud primordia at different stages of development, namely initiation preparation stage (I), initiation stage (II), formation stage (III), and maturation stage (IV). SAM, shoot apical meristem; LP, leaf primordium; P, bud primordium. **(D)** Scanning electron micrographs of dormant axillary buds.

3.6 Jasmonic acid negatively regulates branching formation

Exogenous application of 100 μmol·L^{-1} JA-Me did not notably affect the branching of MP plants, whereas the branching of OP plants was visibly inhibited, in contrast to the control but consistent with the phenotype of MP plants. Expression analysis showed that the expression of *PcOPR3* was significantly up-regulated after JA-Me treatment. In OP plants, *PcCOI1* was not significantly increased in response to JA-Me treatment, whereas *PcCOI1* was significantly upregulated in MP. *PcJAZ1* was significantly upregulated in OP and MP in response to JA-Me treatment, but the extent of upregulation in OP was less than that in MP (Figure 6).

TABLE 2 Leaf anatomical characters of the mutant (MP) and wild type (OP).

Sample	VB thickness (μm)	EL thickness (μm)	LE thickness (μm)
MP	156.60 ± 0.88a	12.19 ± 0.16a	7.39 ± 0.08a
OP	96.34 ± 0.31b	11.42 ± 0.71a	7.06 ± 0.15a

VB, vascular bundle; EL, epithelial layer; LE, abaxial epidermis. Data are the mean ± standard error (n = 3). Different lowercase letters within a column indicate a statistically significant difference ($p < 0.05$, Student's test).

4 Discussion

Pear is an economically important deciduous fruit tree. However, pear has a long growth cycle and complex genetic background, which is usually propagated by grafting. The growth characteristics of rootstocks play an important role in the outcome of grafting. Strong seedlings with few branches are ideal rootstock materials. In the present study, compared with OP, MP developed fewer branches, but the regulatory mechanism for these phenotypic changes remains unclear. Exploration of the regulation of pear branching through comparative analysis is important for the breeding of seedlings suitable for rootstocks.

Morphological assessment is the most common and direct method of identifying bud mutations. In this study, the morphology of OP and MP differed significantly. The MP plants had the phenotypic characteristics of few branches and large leaves, whereas OP had the contrasting traits of many branches and small leaves. Compared with MP, OP developed more branches at an early stage and the leaf axillary buds produced spiny spur shoots. Similar results

10.3389/fpls.2023.1105521

TABLE 3 Anatomical characters of the stem in transverse section for the mutant (MP) and wild type (OP).

Sample	Position in stem	Pith radius (μm)	Xylem width (μm)	Phloem width (μm)
MP	Upper stem	243.14 ± 1.07c	29.57 ± 0.93c	51.25 ± 1.30b
	Middle stem	286.48 ± 0.92b	37.35 ± 0.62b	55.32 ± 0.96b
	Basal stem	327.18 ± 0.91a	62.84 ± 0.77a	89.05 ± 0.73a
OP	Upper stem	128.02 ± 0.97c	27.75 ± 0.51c	2.65 ± 0.63c
	Middle stem	191.51 ± 1.01b	33.04 ± 0.38b	42.25 ± 1.07b
	Basal stem	227.35 ± 0.92a	86.49 ± 0.76a	93.33 ± 1.02a

The upper stem, middle stem and basal stem were selected as 2cm below the tip of the new shoots, the middle of the new shoots and 2cm above the stem base of the new shoots respectively. Data are the mean ± standard error (n = 3). Different lowercase letters within a column indicate a statistically significant difference ($p < 0.05$, Duncan's test).

have been reported in other species, such as birch and larch (Han et al., 2019; Cai et al., 2021). In addition, the vascular bundles and abaxial epidermal cells of MP leaves were larger than those of OP, and the areal proportion of the pith, xylem, and phloem in new shoots of MP from the base to the tip was greater than that of OP, and almost no conducting tissue had differentiated at the shoot tip of OP. These results indicated that MP had a greater capacity for nutrient and water transport than OP, which may be one reason why MP had stronger main branches and fewer branches. The bud structure and development period of MP and OP were different; bud development was more rapid in MP than in OP, thus the bud growth patterns were indicated to differ between OP and MP.

Cell division, expansion, and differentiation affect the basic processes of plant organ growth and development, and ultimately affect plant phenotype (Sugiyama, 2005; Yang et al., 2015). Plant morphogenesis is closely associated with genes involved in cell division, expansion, and differentiation (Lopez-Hernandez et al., 2020). The diversity of plant morphology is largely affected by SAM activity (Sussex and Kerk, 2001). In the current study, 38 DEGs associated with cell division, differentiation, and SAM activity were identified, including 23 transcription factors belonging to the AP2/ERF, P450/CYP, TCP, and NAC families. Transcription factors play a vital role in plant development and regulation of gene expression, forming a complex gene regulatory network (Muiño et al., 2016; Yang et al., 2018). Most of the 23 transcription factors were highly expressed in OP and participate in the regulation of mutant branching traits. The *TCP* transcription factors are involved in the growth of lateral meristems, cell proliferation, and regulation of hormones (Aguilar-Martínez et al., 2007; Martin-Trillo and Cubas, 2010). In citrus, the *TCP* transcription factors *THORN IDENTITY 1*

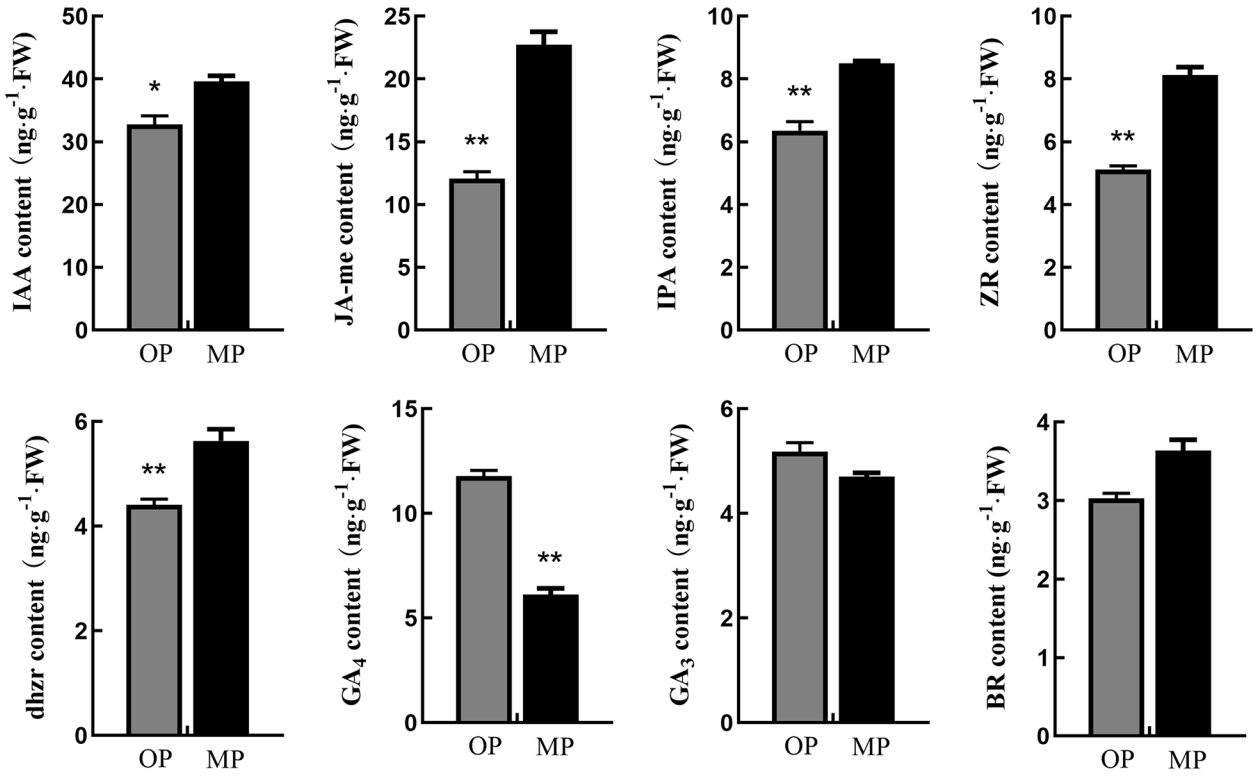


FIGURE 3
Hormone contents in stem tips of the wild type (OP) and mutant (MP). BR, brassinolide; DHZR, dihydrozeatin riboside; GA3 and GA4, gibberellins; IAA, indole-3-acetic acid; IPA, indolepyruvic acid; JA-Me, methyl jasmonate; ZR, zeatin riboside. Error bars indicate the standard error. *$p < 0.05$, **$p < 0.01$ (Duncan's test).

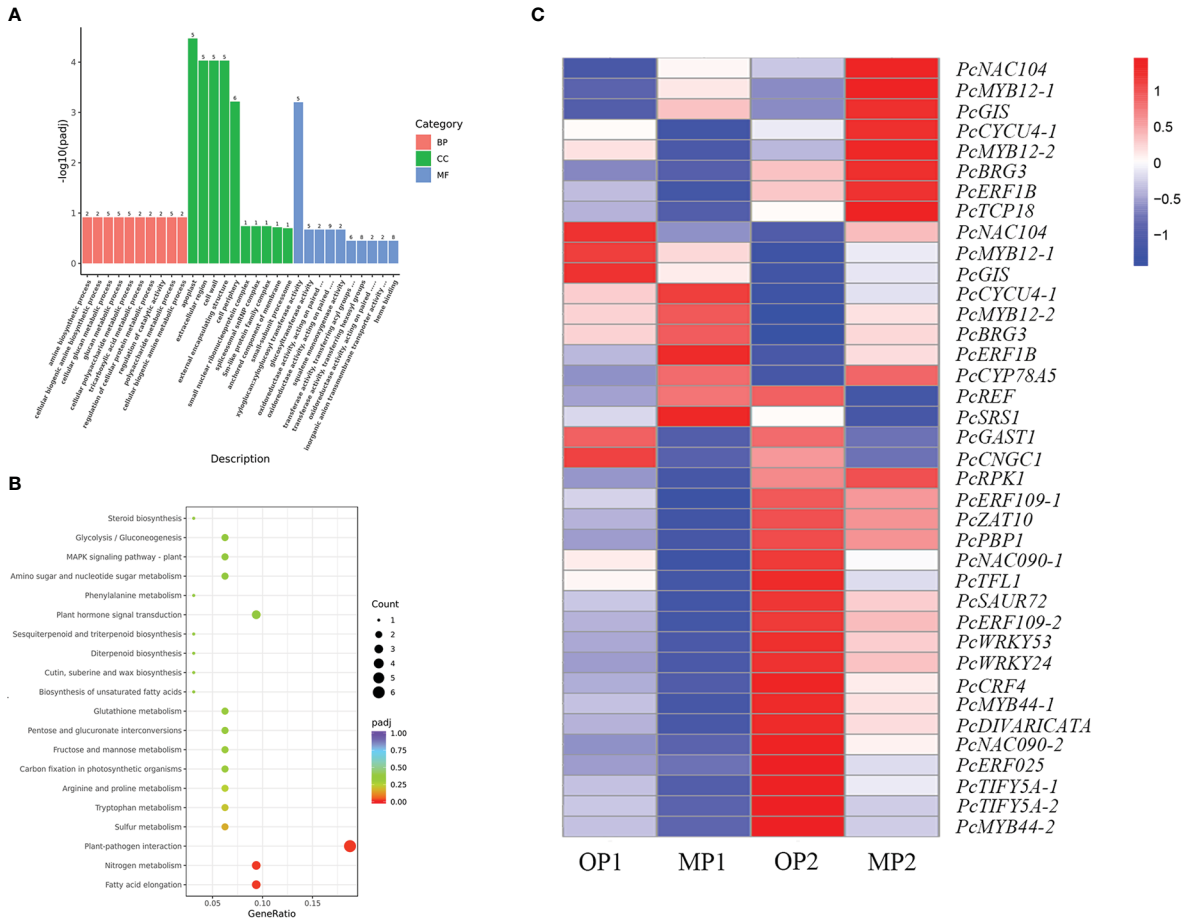


FIGURE 4
Functional annotation of differentially expressed genes (DEGs) in the mutant (MP) and wild type (OP). **(A)** DEGs enriched among the top 30 genes ontology terms. The annotated items are divided into three categories: cellular components(CC), molecular functions(MF), and biological processes(BP). **(B)** DEGs enriched among the top 20 KEGG pathways. **(C)** DEGs involved in cell division, differentiation, and shoot apical meristem activity, in which MP1 and OP1 were the shoot dormancy, MP2 and OP2 were the bud expanding stage. FPKM values are log_2-based.

(TI1) and *TI2* are essential to preclude the proliferation of meristems and simultaneously produce thorns (Zhang et al., 2020; Zhang et al., 2021). In the present study, the expression of *PcTCP18* in OP was higher than that in MP at an early stage of development, whereas expression was higher in MP than in OP at an advanced stage of development. The gene *PcTCP18* has high homology with the citrus *TI1* gene. In peach, *PpTCP18* can reduce secondary branches through brassinolide pathway (Wang et al., 2022). Based on these results, we speculated that *PcTCP18* may be a candidate gene responsible for reducing the branching of MP at an early stage of development.

Plant hormones can regulate plant development processes, such as cell division, bud development, branch branching, and senescence (Heyl et al., 2007). Genes associated with hormone metabolism and signal transduction play an important role in regulating plant and organ size (Guo et al., 2010). In the present study, MP and OP were revealed to differ significantly in hormone contents. The IAA and cytokinin contents in MP were significantly higher than those in OP. Furthermore, the expression level cytokinin-related genes in OP were also higher than that in MP, indicating that cytokinins played an important role in plant branching. This result was consistent with a previous report that cytokinin promotes plant branching, tillering development, and lateral bud growth (Foo et al., 2007). Gibberellins play an important regulatory role in plant growth and development, and can control vertical growth and branching (Arend et al., 2009). In the current study, the expression of DEGs associated with GAs was higher in OP. Similarly, the GA_4 content in OP was higher than that in MP. Gibberellin had been proved to have the effect of inhibiting branching, especially the development of axillary buds was closely related to the content of gibberellin (Martinez-Bello et al., 2015; Tan et al., 2018). This also showed that gibberellin played an important role in the branch development of pear trees. In addition, the JA content in MP was significantly higher than that in OP. Jasmonic acid has diverse roles in regulating developmental processes such as seed germination, root development, and senescence (Jin and Zhu, 2017). Exogenous JA-Me treatment and *JAZ* mutants all can cause the phenotype of branch reduction (Hong et al., 2020). Auxin can also cause changes in plant branching through jasmonic acid (Wang et al., 2013). In our study, 35 genes related to jasmonate biosynthesis and signal transduction were differentially expressed. Among them, in the JA signaling pathway, *PcCOI1-related* genes were highly expressed in both stages of shoot development in MP, whereas *PcJAZs* were expressed at a lower level in MP, indicating that Jasmonate signal transduction is involved in the formation of branching. Application of exogenous JA-Me did not notably change the MP branching

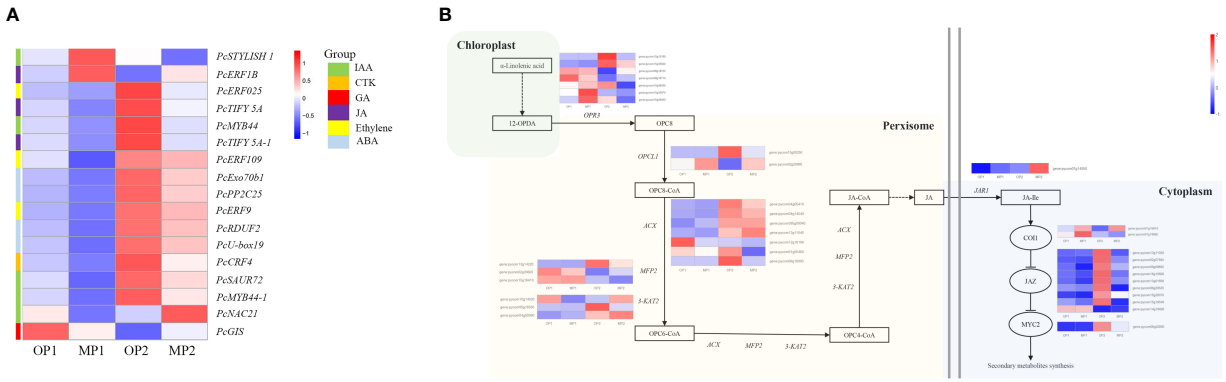


FIGURE 5
Differentially expressed genes involved in plant hormone signaling pathways and in jasmonate biosynthetic and signal transduction pathways. **(A)** Hormone-related DEGs, in which MP1 and OP1 were the shoot dormancy, MP2 and OP2 were the bud expanding stage. **(B)** Pathway constructed based on KEGG pathways and the literature. The enzymes and intermediates are indicated as follows: OPR3, OPDA reductase; ACX, acyl-coenzyme A oxidase; MFP2, peroxisomal fatty acid beta-oxidation multifunctional protein; 3-KAT2, 3-ketoacyl-CoA thiolase 2; JAR1, jasmonate resistant 1; OPC8, 8-(3-oxo-2-(pent-2-enyl)cyclopentyl) octanoic acid; JA, jasmonic acid; JA-Ile, jasmonoyl-L-isoleucine. FPKM values are $\log_2$-based.

phenotype, whereas OP treated with exogenous JA-Me showed a decrease in frequency of lateral branching, resulting in a branching phenotype similar to that of MP. The expression of *PcOPR3* was significantly up-regulated after JA-Me treatment. The downstream JA regulatory gene *PcCOI1* was not highly expressed in OP, but was significantly upregulated in MP plants, in response to JA-Me treatment. *JAZ* negatively regulates the downstream regulatory pathway of JA. In response to JA-Me treatment, the expression level of *PcJAZ* was significantly increased. So far, there were few reports about the effect of jasmonic acid on the branches, but previous studies had confirmed that JAZ could promote growth and inhibit aging and COI1 also participated in plant growth and development

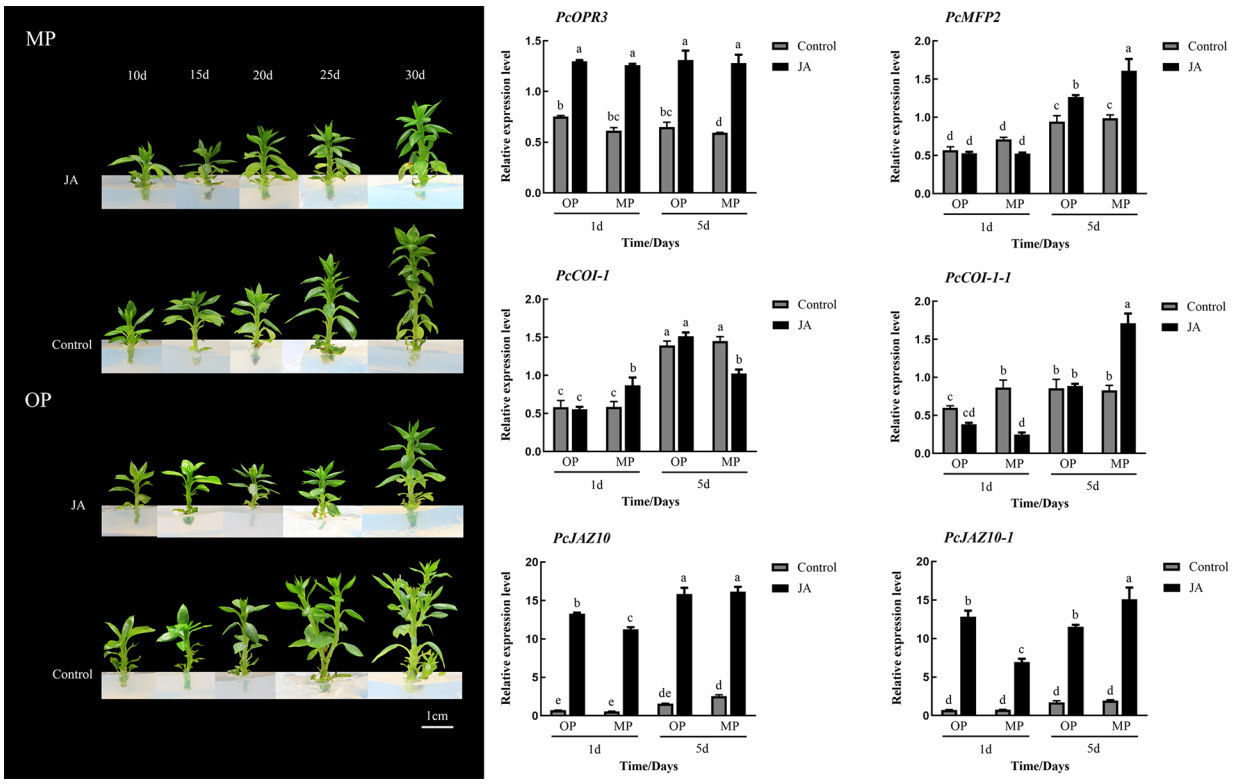


FIGURE 6
Effects of exogenous methyl jasmonate (JA) on plant phenotype and relative expression of six genes associated with the jasmonate signaling pathway at different stages of shoot development. Different lowercase letters above bars within a graph indicate a statistically significant difference ($p < 0.05$, Duncan's multiple range test).

(Huang et al., 2017; Oblessuc et al., 2020;). Therefore, we concluded that JA caused the decrease of plant branch development due to the inhibition of jasmonate on growth, and the jasmonate regulatory gene in OP was not sensitive to jasmonate.

In conclusion, through observation of the phenotypic characters of MP and OP, the mutation in MP results in reduced branching, large leaves, and slower development of bud primordia. The contents of JA, IAA, IPA, and ZR were significantly higher in MP than in OP, especially that of JA. The wild type showed a little-branching phenotype after treatment with exogenous JA-Me. The expression levels of JA synthesis regulatory genes in MP were increased and these genes were responsive to application of exogenous JA-Me.

Data availability statement

The data presented in the study are deposited in the National Center for Biotechnology Information (NCBI) BioProject database repository, accession number PRJNA916958.

Author contributions

YC and CL conducted the experiment, and YC prepared the manuscript. ZQ, SZ, and JL provided important technical help in transcriptional analysis, YY, RW, and CM gave useful guidance and valuable discussion. JY, ZC, and JS provided the plant materials and important technical help in histological Observation. DL conceived the idea and provided financial support. 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 (2019YFD1001404), the earmarked fund for CARS (CARS-28-07), the Agricultural Variety Improvement Project of Shandong Province (2022LZGC011), Qingdao Agricultural University Doctoral Start-Up Fund (2019).

Conflict of interest

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

Publisher's note

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

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1105521/full#supplementary-material

References

Aguilar-Martinez, J. A., Poza-Carrión., C., and Cubas, P. (2007). *Arabidopsis BRANCHED1* acts as an integrator of branching signals within axillary buds. *Plant Cell.* 19, 458–472. doi: 10.1105/tpc.106.048934

Arend, M., Schnitzler, J. P., Ehlting, B., Hänsch, R., Lange, T., Rennenberg, H., et al. (2009). Expression of the *Arabidopsis* mutant *ABI1* gene alters abscisic acid sensitivity, stomatal development, and growth morphology in gray poplars. *Plant Physiol.* 151, 2110–2119. doi: 10.1104/pp.109.144956

Bemer, M., van Mourik, H., Muiño, J. M., Ferrándiz, C., Kaufmann, K., and Angenent, G. C. (2017). *FRUITFULL* controls *SAUR10* expression and regulates *Arabidopsis* growth and architecture. *J. Exp. Bot.* 68, 3391–3403. doi: 10.1093/jxb/erx184

Bollmark, M., Kubát, B., and Eliasson, L. (1988). Variations in endogenous cytokinin content during adventitious root formation in pea cuttings. *J. Plant Physiol.* 132, 262–265. doi: 10.1016/S0176-1617(88)80102-0

Cai, K. W., Zhou, X. Y., Li, X., Kang, Y., Yang, X. M., Cui, Y. H., et al. (2021). Insight into the multiple branches traits of a mutant in larix olgensis by morphological, cytological, and transcriptional analyses. *Front. Plant sci.* 12. doi: 10.3389/fpls.2021.787661

Campos, M. L., Kang, J. H., and Howe, G. A. (2014). Jasmonate-triggered plant immunity. *J. Chem. Ecol.* 40, 657–675. doi: 10.1007/s10886-014-0468-3

Chini, A., Fonseca, S., Fernández, G., Adie, B., Chico, J., Lorenzo, O., et al. (2007). The *JAZ* family of repressors is the missing link in jasmonate signalling. *Nature* 448, 666–671. doi: 10.1038/nature06006

Clarke, S. M., Cristescu, S. M., Miersch, O., Harren, F. J., Wasternack, C., and Mur, L. A. (2009). Jasmonates act with salicylic acid to confer basal thermotolerance in *Arabidopsis thaliana*. *New Phytol.* 182, 175–187. doi: 10.1111/j.1469-8137.2008.02735.x

Depuydt, S., and Hardtke, C. S. (2011). Hormone signalling crosstalk in plant growth regulation. *Curr. Biol.* 21, R365–R373. doi: 10.1016/j.cub.2011.03.013

Fambrini, M., Salvini, M., and Pugliesi, C. (2017). Molecular cloning, phylogenetic analysis, and expression patterns of *LATERAL SUPPRESSOR-LIKE* and *REGULATOR OF AXILLARY MERISTEM FORMATION-LIKE* genes in sunflower (*Helianthus annuus l.*). *Dev. Genes Evol.* 227, 159–170. doi: 10.1007/s00427-016-0571-2

Foo, E., Morris, S. E., Parmenter, K., Young, N., Wang, H., Jones, A., et al. (2007). Feedback regulation of xylem cytokinin content is conserved in pea and *Arabidopsis*. *Plant Physiol.* 143, 1418–1428. doi: 10.1104/pp.106.093708

Gao, R., Gruber, M. Y., Amyot, L., and Hannoufa, A. (2018). *SPL13* regulates shoot branching and flowering time in medicago sativa. *Plant Mol. Biol.* 96, 119–133. doi: 10.1007/s11103-017-0683-8

Grbić, V., and Bleecker, A. B. (2000). Axillary meristem development in *Arabidopsis thaliana*. *Plant J.* 21, 215–223. doi: 10.1046/j.1365-313x.2000.00670.x

Guo, M., Rupe, M. A., Dieter, J. A., Zou, J. J., Spielbauer, D., Duncan, K. E., et al. (2010). Cell number Regulator1 affects plant and organ size in maize: implications for crop yield enhancement and heterosis. *Plant Cell.* 22, 1057–1073. doi: 10.1105/tpc.109.073676

Han, R., Wang, S., Liu, C. Y., Xu, W. D., Bian, X. Y., Liu, G. F., et al. (2019). Transcriptome analysis of a multiple-branches mutant terminal buds in betula platyphylla×B. pendula. *Forests* 10, 374. doi: 10.3390/f10050374

Heyl, A., Wulfetange, K., Pils, B., Nielsen, N., Romanov, G. A., and Schmülling, T. (2007). Evolutionary proteomics identifies amino acids essential for ligand binding of the cytokinin receptor *CHASE* domain. *BMC Evol. Biol.* 7, 62. doi: 10.1186/1471-2148-7-62

Hong, S. Y., Sun, B., Straub, D., Blaakmeer, A., Mineri, L., Koch, J., et al. (2020). Heterologous microProtein expression identifies LITTLE NINJA, a dominant regulator of jasmonic acid signaling. *Proc. Natl. Acad. Sci. U.S.A.* 117, 26197–26205. doi: 10.1073/pnas.2005198117

Huang, H., Liu, B., Liu, L. Y., and Song, S. S. (2017). Jasmonate action in plant growth and development. *J. Exp. Bot.* 6, 1349–1359. doi: 10.1093/jxb/erw495

Jin, H. H., and Zhu, Z. Q. (2017). Temporal and spatial view of jasmonate signaling. *Trends Plant Sci.* 22, 451–454. doi: 10.1016/j.tplants.2017.04.001

Kim, J., Dotson, B., Rey, C., Lindsey, J., Bleecker, A. B., Binder, B. M., et al. (2013). New clothes for the jasmonic acid receptor *COI1*: delayed abscission, meristem arrest and apical dominance. *PloS One* 8 (4), e60505. doi: 10.1371/journal.pone.0060505

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

Leng, F., Ye, Y. L., Zhu, X. H., Zhang, Y., Zhang, Z. Y., Shi, J. Y., et al. (2021). Comparative transcriptomic analysis between 'Summer black' and its bud sport 'Nantaihutezao' during developmental stages. *Planta* 253, 23. doi: 10.1007/s00425-020-03543-7

Leyser, O. (2003). Regulation of shoot branching by auxin. *Trends Plant Sci.* 8, 541–545. doi: 10.1016/j.tplants.2003.09.008

Ligerot, Y., de Saint Germain, A., Waldie, T., Troadec, C., Citerne, S., Kadakia, N., et al. (2017). The pea branching *RMS2* gene encodes the *PsAFB4/5* auxin receptor and is involved in an auxin-strigolactone regulation loop. *PloS Genet.* 13 (12), e1007089. doi: 10.1371/journal.pgen.1007089

Li, G. F., Tan, M., Ma, J. J., Cheng, F., Li, K., Liu, X. J., et al. (2021). Molecular mechanism of *MdWUS2-MdTCP12* interaction in mediating cytokinin signaling to control axillary bud outgrowth. *J. Exp. Bot.* 72, 4822–4838. doi: 10.1093/jxb/erab163

Lopez-Hernandez, F., Tryfona, T., Rizza, A., Yu, X. L., Harris, M. O. B., Webb, A. A. R., et al. (2020). Calcium binding by arabinogalactan polysaccharides is important for normal plant development. *Plant Cell.* 32, 3346–3369. doi: 10.1105/tpc.20.00027

Martinez-Bello, L., Moritz, T., and López-Díaz, I. (2015). Silencing c-19-GA 2-oxidases induces parthenocarpic development and inhibits lateral branching in tomato plants. *J. Exp. Bot.* 19, 5897–5910. doi: 10.1093/jxb/erv300

Martin-Trillo, M., and Cubas, P. (2010). *TCP* genes: a family snapshot ten years later. *Trends Plant Sci.* 15, 31–39. doi: 10.1016/j.tplants.2009.11.003

Mehrnia, M., Balazadeh, S., Zanor, M. I., and Mueller-Roeber, B. (2013). *EBE*, an *AP2/ERF* transcription factor highly expressed in proliferating cells, affects shoot architecture in *Arabidopsis*. *Plant Physiol.* 162, 842–857. doi: 10.1104/pp.113.214049

Muiño, J. M., de Bruijn, S., Pajoro, A., Geuten, K., Vingron, M., Angenent, G. C., et al. (2016). Evolution of DNA-binding sites of a floral master regulatory transcription factor. *Mol. Biol. Evol.* 33, 185–200. doi: 10.1093/molbev/msv210

Mur, L. A., Kenton, P., Atzorn, R., Miersch, O., and Wasternack, C. (2006). The outcomes of concentration-specific interactions between salicylate and jasmonate signaling include synergy, antagonism, and oxidative stress leading to cell death. *Plant Physiol.* 140, 249–262. doi: 10.1104/pp.105.072348

Nie, J., Wen, C., Xi, L., Lv, S. H., Zhao, Q. C., Kou, Y. P., et al. (2018). The *AP2/ERF* transcription factor *CmERF053* of chrysanthemum positively regulates shoot branching, lateral root, and drought tolerance. *Plant Cell Rep.* 37, 1049–1060. doi: 10.1007/s00299-018-2290-9

Oblessuc, P. R., Obulareddy, N., DeMott, L., Matiolli, C. C., Thompson, B. K., and Melotto, M. (2020). JAZ4 is involved in plant defense, growth, and development in *Arabidopsis*. *Plant J.* 101, 371–383. doi: 10.1111/tpj.14548

Stracke, R., Werber, M., and Weisshaar, B. (2001). The *R2R3-MYB* gene family in *Arabidopsis thaliana*. *Curr. Opin. Plant Biol.* 4, 447–456. doi: 10.1016/S1369-5266(00)00199-0

Sugiyama, S. I. (2005). Polyploidy and cellular mechanisms changing leaf size: comparison of diploid and autotetraploid populations in two species of lolium. *Ann. Bot.* 96, 931–938. doi: 10.1093/aob/mci245

Sussex, I. M., and Kerk, N. M. (2001). The evolution of plant architecture. *Curr. Opin. Plant Biol.* 4, 33–37. doi: 10.1016/S1369-5266(00)00132-1

Takeda, T., Suwa, Y., Suzuki, M., Kitano, H., Ueguchi-Tanaka, M., Ashikari, M., et al. (2003). The *OsTB1* gene negatively regulates lateral branching in rice. *Plant J.* 33, 513–520. doi: 10.1046/j.1365-313X.2003.01648.x

Tanaka, W., Pautler, M., Jackson, D., and Hirano, H. Y. (2013). Grass meristems II: Inflorescence architecture, flower development and meristem fate. *Plant Cell Physiol.* 54, 313–324. doi: 10.1093/pcp/pct016

Tanaka, M., Takei, K., Kojima, M., Sakakibara, H., and Mori, H. (2006). Auxin controls local cytokinin biosynthesis in the nodal stem in apical dominance. *Plant J.* 45, 1028–1036. doi: 10.1111/j.1365-313X.2006.02656.x

Tan, M., Li, G. F., Liu, X. J., Cheng, F., Ma, J. J., Zhao, C. P., et al. (2018). Exogenous application of GA(3) inactively regulates axillary bud outgrowth by influencing of branching-inhibitors and bud-regulating hormones in apple (*Malus domestica* borkh.). *Mol. Genet. Genomics* 293, 1547–1563. doi: 10.1007/s00438-018-1481-y

Thines, B., Katsir, L., Melotto, M., Niu, Y., Mandaokar, A., Liu, G. H., et al. (2007). *JAZ* repressor proteins are targets of the *SCFCOI1* complex during jasmonate signaling. *Nature* 448, 661–665. doi: 10.1038/nature05960

Wai, A. H., and An, G. (2017). Axillary meristem initiation and bud growth in rice. *J. Plant Biol.* 60, 440–451. doi: 10.1007/s12374-017-0088-x

Wang, X. B., Wang, Q. P., Yan, L. X., Hao, Y. H., Lian, X. D., Zhang, H. P., et al. (2022). *PpTCP18* is upregulated by lncRNA5 and controls branch number in peach (*Prunus persica*) through positive feedback regulation of strigolactone biosynthesis. *Horticul. Res.* 10, uhac224. doi: 10.1093/hr/uhac224

Wang, J., Yan, D. W., Yuan, T. T., Gao, X., and Lu, Y. T. (2013). A gain-of-function mutation in IAA8 alters *Arabidopsis* floral organ development by change of jasmonic acid level. *Plant Mol. Biol.* 82, 71–83. doi: 10.1007/s11103-013-0039-y

Wasternack, C., and Hause, B. (2013). Jasmonates: Biosynthesis, perception, signal transduction and action in plant stress response, growth and development. an update to the 2007 review in annals of botany. *Ann. Bot.* 111, 1021–1058. doi: 10.1093/aob/mct067

Weiler, E. W., Jordan, P. S., and Conrad, W. (1981). Levels of indole 3-acetic acid in intact and decapitated coleoptiles as determined by a specific and highly sensitive solid-phase enzyme immunoassay. *Planta* 153, 561–571. doi: 10.1007/BF00385542

Xie, D. X., Feys, B. F., James, S., Nieto-Rostro, M., and Turner, J. G. (1998). *COI1*: An *Arabidopsis* gene required for jasmonate-regulated defense and fertility. *Science* 280, 1091–1094. doi: 10.1126/science.280.5366.1091

Yang, Y., Bao, S. Y., Zhou, X. H., Liu, J., and Zhuang, Y. (2018). The key genes and pathways related to male sterility of eggplant revealed by comparative transcriptome analysis. *BMC Plant Biol.* 18, 209. doi: 10.1186/s12870-018-1430-2

Yang, C. X., Gao, Y. N., Gao, S. H., Yu, G., Xiong, C., Chang, J., et al. (2015). Transcriptome profile analysis of cell proliferation molecular processes during multicellular trichome formation induced by tomato *Wov* gene in tobacco. *BMC Genomics* 16, 868. doi: 10.1186/s12864-015-2099-7

Yan, Y. X., Stolz, S., Chételat, A., Reymond, P., Pagni, M., Dubugnon, L., et al. (2007). A downstream mediator in the growth repression limb of the jasmonate pathway. *Plant Cell* 19, 2470–2483. doi: 10.1105/tpc.107.050708

Ye, M., Luo, S. M., Xie, J. F., Li, Y. F., Xu, T., Liu, Y., et al. (2012). Silencing *COI1* in rice increases susceptibility to chewing insects and impairs inducible defense. *PloS One* 7 (4), e36214. doi: 10.1371/journal.pone.0036214

Yu, X., Chen, G. P., Tang, B. Y., Zhang, J. L., Zhou, S. G., and Hu, Z. L. (2018). The jasmonate *ZIM-domain* protein gene *SlJAZ2* regulates plant morphology and accelerates flower initiation in solanum lycopersicum plants. *Plant Sci.* 267, 65–73. doi: 10.1016/j.plantsci.2017.11.008

Zawaski, C., and Busov, V. B. (2014). Roles of gibberellin catabolism and signaling in growth and physiological response to drought and short day photoperiods in *Populus* trees. *PloS One* 9, e86217. doi: 10.1371/journal.pone.0086217

Zhai, Q. Z., Zhang, X., Wu, F. M., Feng, H. L., Deng, L., Xu, L., et al. (2015). Transcriptional mechanism of jasmonate receptor *COI1-mediated* delay of flowering time in *Arabidopsis*. *Plant Cell.* 27, 2814–2828. doi: 10.1105/tpc.15.00619

Zhang, B., Liu, J., Yang, Z. E., Chen, E. Y., Zhang, C. J., Zhang, X. Y., et al. (2018). Genome-wide analysis of *GRAS* transcription factor gene family in gossypium hirsutum l. *BMC Genomics* 19, 348. doi: 10.1186/s12864-018-4722-x

Zhang, F., Rossignol, P., Huang, T., Wang, Y. W., May, A., Dupont, C., et al. (2020). Reprogramming of stem cell activity to convert thorns into branches. *Curr. Biol.* 30, 2951–2961.e5. doi: 10.1016/j.cub.2020.05.068

Zhang, F., Wang, Y. W., and Irish, V. F. (2021). *CENTRORADIALIS* maintains shoot meristem indeterminacy by antagonizing *THORN IDENTITY1* in citrus. *Curr. Biol.* 31, 2237–2242.e4. doi: 10.1016/j.cub.2021.02.051

Zhao, X. H., Wen, B. B., Li, C., Tan, Q. P., Liu, L., Chen, X. D., et al. (2021). Overexpression of the peach transcription factor early bud-break 1 leads to more branches in poplar. *Front. Plant Sci.* 12. doi: 10.3389/fpls.2021.681283

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 08 May 2023
DOI 10.3389/fpls.2023.1186718

OPEN ACCESS

EDITED BY
Qiang Guo,
Beijing Academy of Agricultural and Forestry Sciences, China

REVIEWED BY
Fang Fang,
South China Agricultural University, China
Debatosh Das,
University of Missouri, United States

*CORRESPONDENCE
Fulin Yang
fulin.yang@aliyun.com

RECEIVED 15 March 2023
ACCEPTED 10 April 2023
PUBLISHED 08 May 2023

CITATION
Zhou J, Liu Y, Li Y, Ling W, Fan X, Feng Q, Ming R and Yang F (2023) Combined analyses of transcriptome and metabolome reveal the mechanism of exogenous strigolactone regulating the response of elephant grass to drought stress. *Front. Plant Sci.* 14:1186718. doi: 10.3389/fpls.2023.1186718

Combined analyses of transcriptome and metabolome reveal the mechanism of exogenous strigolactone regulating the response of elephant grass to drought stress

Jing Zhou[1], Yijia Liu[1,2], Yan Li[1,2], Wenqing Ling[2], Xiaoyu Fan[1,2], Qixian Feng[2], Ray Ming[3] and Fulin Yang[2*]

[1]National Engineering Research Center of Juncao Technology, Fujian Agriculture and Forestry University, Fuzhou, China, [2]College of Animal Sciences (College of Bee Science), Fujian Agriculture and Forestry University, Fuzhou, China, [3]Center for Genomics and Biotechnology, Fujian Provincial Key Laboratory of Haixia Applied Plant Systems Biology, Fujian Agriculture and Forestry University, Fuzhou, China

Elephant grass is widely used in feed production and ecological restoration because of its huge biomass and low occurrence of diseases and insect pets. However, drought seriously affects growth and development of this grass. Strigolactone (SL), a small molecular phytohormone, reportedly participates in improving resilience to cope with arid environment. But the mechanism of SL regulating elephant grass to response to drought stress remains unknown and needs further investigation. We conducted RNA-seq experiments and identified 84,296 genes including 765 and 2325 upregulated differential expression genes (DEGs) and 622 and 1826 downregulated DEGs, compared drought rehydration with spraying SL in roots and leaves, respectively. Combined with targeted phytohormones metabolite analysis, five hormones including 6-BA, ABA, MeSA, NAA, and JA had significant changes under re-watering and spraying SL stages. Moreover, a total of 17 co-expression modules were identified, of which eight modules had the most significant correlation with all physiological indicators with weighted gene co-expression network analysis. The venn analysis revealed the common genes between Kyoto Encyclopedia of Genes and Genomes enriched functional DEGs and the top 30 hub genes of higher weights in eight modules, respectively. Finally, 44 DEGs had been identified as key genes which played a major role in SL response to drought stress. After verification of its expression level by qPCR, six key genes in elephant grass including *PpPEPCK*, *PpRuBPC*, *PpPGK*, *PpGAPDH*, *PpFBA*, and *PpSBPase* genes regulated photosynthetic capacity under the SL treatment to respond to drought stress. Meanwhile, *PpACAT*, *PpMFP2*, *PpAGT2*, *PpIVD*, *PpMCCA*, and *PpMCCB* regulated root development and phytohormone crosstalk to respond to water

deficit conditions. Our research led to a more comprehensive understanding about exogenous SL that plays a role in elephant grass response to drought stress and revealed insights into the SL regulating molecular mechanism in plants to adapt to the arid environment.

KEYWORDS

transcriptome, metabolome, exogenous strigolactone, drought resistance, SL crosstalk, elephant grass

Introduction

Drought is one of the important environmental factors affecting plant growth and development. In recent years, with continuous changes in climate worldwide, the frequency and intensity of drought have been increasing (Banks et al., 2019). Severe drought may cause metabolic imbalance of plant cells, resulting in excess light energy and damage of photosynthetic organs. It may also cause the accumulation of reactive oxygen species in plant leaves, accelerate biofilm lipid peroxidation and inhibit plant growth. Plants have evolved various defense and self-protection mechanisms in the long-term evolution process to deal with the influence of drought stress. It could regulate the photosynthetic process of leaves, remove excess reactive oxygen species, and maintain certain turgor pressure in cells by closing stomata in leaves, activating the antioxidant system, and increasing osmotic substances, finally maintain sustainable growth of plants under drought stress (Yamada and Umehara, 2015; Muhammad et al., 2017).

In addition to their own stress response, exogenous growth regulators help improve the drought tolerance of plants. The common growth regulators include abscisic acid (ABA), gibberellin (GA), auxin, cytokinin, brassinosteroids, ethylene, and some small molecular metabolites discovered in recent years, such as nitric oxide (NO), jasmonic acid (JA), salicylic acid (SA), polyamines and strigolactones (SLs) (Koltai et al., 2010; Daviere and Achard, 2016). Plant growth regulators are generally recognized to promote stomatal closure, reduce transpiration rate, and decrease water loss under water-deficit growth environment (Yamada and Umehara, 2015). They could also delay the degradation rate of chlorophyll in leaves and lessen the damage degree of chloroplast caused by drought. In addition, they could enhance the enzymatic activity and increase the proline contents to reduce lipid peroxidation (Visentin et al., 2020). Therefore, growth regulators play an important role in plant response to drought stress.

SLs are a small class of carotenoid-derived compounds and they were first characterized as crystalline germination stimulants in root parasitic plants, including *Striga*, *Phelipanche*, and *Orobanche* species (Cook et al., 1966). Subsequently, SLs were found as a class of phytohormones with root-derived signal that could enhance the symbiosis relationship between plants and arbuscular mycorrhizal fungi, playing a remarkable role in the suppression of shoot branching and tillering (Chen et al., 2014). Meanwhile, studies have shown that SLs could not only affect the development and growth process of plants but also devote to the regulatory systems of plant stress adaptation for establishing multiple abiotic stress tolerances in plants (Bhoi et al., 2021). For example, spraying exogenous SL on arid *Vitis Vinifera* and wheat (*Triticum aestivum*) seedlings could reduce stomatal opening, electrolyte leakage and H_2O_2, reactive oxygen species, and malondialdehyde (MDA) content, to regulate its ability to cope with drought stress (Min et al., 2018).

A single phytohormone regulates several aspects of plant life cycle. Most developmental and growth processes are affected by various hormones simultaneously (de Saint Germain et al., 2013). The crosstalk of SLs with other phytohormone indicates that SLs accumulate in plant tissues to cope with different environmental conditions and actively contribute to multiple hormonal responsive pathways of plant stress adaptation (Bhoi et al., 2021). SLs positively regulate stress- and/or ABA-responsive gene expression in connection with plant development and abiotic stress responses (Ha et al., 2014; Haider et al., 2018). SL application reduced the levels of zeatin riboside and indoleacetic acid (Ghorbel et al., 2021) and increased the level of ABA in leaves and roots under drought stress (Min et al., 2018). Increasing amounts of endogenous SLs in less-branched varieties may have restrained the concentration of endogenous CKs, leading to the inhibition of axillary buds in rice (*Oryza sativa*) (Xu et al., 2015). Similar results have also been demonstrated in *Zantedeschia aethiopica*, where bud growth was activated by SL and CK interaction (Manandhar et al., 2018). In addition to the above crosstalk, SL and GA were involved in seed germination and shoot branching. GA application negatively and independently regulated SL biosynthesis in rice and *Lotus japonicus* (de Saint Germain et al., 2013). Therefore, SLs adjust multiple hormonal reaction pathways, which improve the capacity of plant development and coping with environmental stress.

The mechanism of SL responding to drought stress differ among various plant species. The α/β hydrolases DWARF14 (D14), which acts as receptors of SL, interact with the F-box protein MORE AXILLARY GROWTH 2 (MAX2) to target SUPPRESSOR OF MAX2 1 (SMAX1)-LIKE/D53 family members for degradation *via* the 26S proteasome (Yao et al., 2016; Yang et al., 2019). However, in *Arabidopsis*, the SMAX2 1-LIKE 6, 7, and 8 (SMXL 6, 7, and 8) and SMAX1-LIKE2 (SMXL2) act as negative regulators of water deficit resistance, and destruction of these *SMXL* genes in plants may lead to an original strategy to improve their drought resistance (Li et al., 2020; Feng et al., 2022). In a mutant of

Hordeum vulgare, the *hvd14.d* gene is unusually sensitive to drought stress, may be due to the disruption of ABA signaling and metabolism pathways, leading to the drought-sensitive phenotype in the SL mutant. The correlation between SL signal and drought stress response is dependent on ABA level (Marzec et al., 2020). Under drought stress in tomato (*Solanum lycopersicum*), the level of SL in the roots decreased, and it was considerable related to the down-regulation of *SlCCD7* and *SlCCD8* genes in the roots (Ruiz-Lozano et al., 2015). In rice, *OsTB1* and SL were found to be involved in tillering inhibition under low water-deficit treatment, which is independent of the flowering pathway (Du et al., 2018). Comparative analysis of rice SL mutants (D3, D10, D17 and D27) showed that ABA content was positively correlated with the expression of β-carotene isomerase encoding gene of D27, and the homeostasis of ABA was maintained by regulating this gene to increase the drought tolerance of plants (Haider et al., 2018).

Elephant grass (*Pennisetum purpureum* Schum.) is an ideal gramineae C_4 plant for sand fixation, wind prevention, and water-soil conservation, on the account of its large biomass and developed root system (Zhou et al., 2021). It has been widely planted for full-scale environmental management of regions with vulnerable ecology (Zhou et al., 2021). Meanwhile, arid environment affects the yield and constraints its large-scale plantation. For improvement of its performance and productivity in water deficit conditions, exogenous SL was applied to improve drought resistance. RNA sequencing (RNA-seq) and hormone-targeted metabolism analysis, combined with physiological indices, could comprehensively and quickly acquire the gene expression of a specific tissue or cell in a certain state. The findings will improve the understanding of the exogenous SL on regulating the mechanism of drought stress tolerance in C_4 grasses and offer a theoretical basis for genetic improvement and breeding of elephant grass.

Materials and methods

Experimental materials

Elephant grass sprouts were collected from the plantation base in National Herbage Cultivar Evaluation Station. After growing healthily and producing nine leaves, they were subjected to stop irrigation for 10 days under room temperature (27°C) to obtain drought-stressed plants. The experiment included three groups: normal watering treatment, drought rehydration management, and foliar spraying SL handling. Water and 3 μmol/L GR 24 (Beijing Solarbio Science & Technology Co., Itd, China), which is a synthetic SL, were applied to the leaves to make the surface fully coated with liquid. After 72 h, the three groups of plants were tested.

Physiological index detection

Determination of photosynthetic indices of elephant grass

In this study, the photosynthetic indices of elephant grass were surveyed using a CIRAS-3 portable photosynthesizer (PP-Systems Company Amesbury, MA01913, USA) to test the top third unfolded leaf of the plant. The photosynthetic indices were commanded under the conditions of which light intensity, air relative humidity, and CO_2 concentration were 1200 μmol·$(m^2 \cdot s)^{-1}$, 75%, and 380 μmol·mol^{-1}, respectively. The net photosynthetic rate (Pn), intercellular CO_2 concentration (Ci), transpiration rate (Tr), stomatal conductance (Gs), and water use efficiency (WUE) were measured in each treatment group. Then, the leaves were dark-adapted for 20 min before Chl fluorescence parameter was determined. The performance index on an absorption basis (PIabs) and the maximum photochemical efficiency (Fv/Fm) were determined using a plant efficiency analyzer. Three plants were surveyed for each treatment.

Measurement of physiological indices

Taking the roots and leaves as sample, the physiological indices, including peroxidase (POD), superoxide dismutase (SOD) activity, and proline (PRO), and MDA contents, were determined in terms of a previously reported method. Meanwhile, leaf samples were snap frozen and ground in liquid nitrogen for photosynthetic enzyme activity test. Then, 0.1 g of the sample was placed into a centrifuge tube and added with 1 mL of extraction solution. The mixture was homogenized in an ice bath and centrifuged at 8000 g and 4°C for 10 min. Ten microliters of supernatant were collected for the determination of photosynthetic key enzyme nicotinamide adenine dinucleotide phosphate-malic enzyme (NADP-ME), phosphoenolpyruvate carboxylase (PEPC), and pyruvate phosphate dikinase (PPDK) activities in accordance with the NADP-ME kit, PEPC kit, and PPDK kit purchased from Shanghai Preferred Biotechnology Co., China, respectively. Three plants were taken as repetition for each treatment.

Roots measurements

After the roots were dug out from the soil medium, they were gently washed with running water until the attachment of roots was totally cleaned, and then one of the main roots was randomly selected. Next, 10 cm-long roots were taken from the stem as samples for measurement. The samples were scanned for images, and the root morphology index, including root length (Len), surface area (SA), number of tips (NTips), volume (Vol), and average diameter (AvgD), was measured with WinRHIZO. Every treatment had three replications. After the above physiological indices of plants were detected, the leaves and roots of fresh elephant grass were collected and immediately fixed with liquid nitrogen to prepare for subsequent hormone metabolism and transcriptome analysis.

Targeted metabolite analysis of plant hormone under exogenous SL treatment

Metabolite extraction

About 50 mg of samples of elephant grass, including leaves and roots, was respectively placed in a 1.5 mL Eppendorf tube for phytohormone metabolite testing, with three replicates per treatment. All samples were added with 1 mL of 50% acetonitrile aqueous solution and ground at 4°C for 6 min by freezing tissue

grinding mill (Wonbio-96E, Shanghai, China). The mixture was shocked by ultrasonic for 30 min, stood at 4°C for 30 min, and centrifuged at 13,000 g at 4°Cfor 15 min. All supernatants were taken for purification, and the eluant was placed into 2 mL Eppendorf tube to blow nitrogen to dry. Finally, 50 μL of 30% acetonitrile aqueous solution was added into the tube, mixed together, and centrifuged at 13,000 g at 4°C for 15 min. The solution was the metabolite extraction, and it was transferred to sample vials for LC-MS/MS analysis.

Targeted metabolite analysis by LC-MS/MS

Before the phytohormone concentration of the samples were tested, a total of 22 kinds of standards were melted with methanol and mixed together to obtain the standard stock solution. Phytohormone metabolites were tested using an ultrahigh-performance liquid phase chromatograph mass spectrometer (UHPLC-Qtrap). The mobile phases were 0.01% formic acid in water (phase A) and 0.01% formic acid in acetonitrile (solvent B). The mass spectral conditions were positive and negative modes, curtain gas of 35 psi, medium collision gas, ion spray voltage floating of +5500/−4500, and source temperature of 550°C.

Data preprocessing

By using AB Sciex quantitative software, the default indices were used to automatically authenticate and integrate each ion fragment. With the mass spectral peak area and the concentration of appointed hormone as ordinates and abscissa, respectively, the standard curves of linear regression were drawn. The concentration of sample hormones was calculated by substituting the mass spectral peak area of the samples into the linear equation. Finally, the phytohormone content of the samples (ng/mg) was calculated as being equal to the concentration of samples hormone multiplied by the sample extraction volume and divided by sample weight.

Characteristics of transcriptome in exogenous SL treated drought elephant grass

Total RNA extraction and sequencing library preparation for transcriptome

Total RNA was extracted from the elephant grass seedling and roots by using the TRIzol reagent (TransGen, Beijing, China) following the manufacturer's protocol. RNA integrity was measured using an Agilent Bioanalyzer 2100 (Agilent Technologies, Santa Clara, CA, USA). The concentration and purity of RNA were tested using a NanoDrop 1000 spectrophotometer (Thermo Fisher Scientific, Wilmington, DE, USA). Only high-quality RNA sample was used to construct sequencing library.

The RNA-seq library was constructed with the Illumina TruSeq RNA sample preparation kit. mRNA was linked with poly-T oligo-attached magnetic beads and then fragmented by a fragmentation buffer. Then, double-stranded cDNA was composited using a SuperScript double-stranded cDNA synthesis kit with random hexamer primers. The composite cDNA was subjected to phosphorylation, A base addition, and end repair. Library fragments were singled out with cDNA target fragments of 300 bp and sequenced with the Illumina NovaSeq 6000 sequencer.

Sequence read mapping and assembly

The raw paired-end reads were trimmed and quality controlled by SeqPrep (https://github.com/jstjohn/SeqPrep) and Sickle (https://github.com/najoshi/sickle) with default parameters. The genome of *Cenchrus purpureus* was used as reference gene, and the clean reads were separately aligned to it with orientation mode by using HISAT2 (http://ccb.jhu.edu/software/hisat2/index.shtml) software. The mapped reads of each sample were assembled by StringTie (https://ccb.jhu.edu/software/stringtie) in a reference-based approach.

Gene functional annotation and differential expression analysis

All the assembled transcripts were analyzed by BLAST with six public databases, including Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway database, Gene Ontology (GO) database, Pfam protein family (Pfam) database, non-redundant (NR) protein sequence, eukaryotic orthologous group (KOG) database, and manually annotated and reviewed protein sequence database (Swiss-Prot). The expression of each transcript was calculated in accordance with the transcripts-per-million-reads (TPM) method. After the read count data were standardized, the DESeq2 package analyses were adjusted *via* Benjamini–Hochberg method to determine the false discovery rate (FDR). Differentially expressed genes (DEGs) were determined with q-value ≤ 0.01 and $|\log_2 FC| > 2$.

KEGG pathway analysis of DEGs

The KEGG database was used for KEGG enrichment analysis of DEGs during the three treatments of elephant grass to study the DEGs of SL regulating drought stress. By estimating the variation of gene length, the probability of different gene-enriched KEGG pathways could be calculated more accurately. The statistical concentration of DEGs was examined using KOBAS 2.0 web server, and a corrected p-value < 0.05 was considered to be remarkable abundant in KEGG.

Weighted gene co-expression network analysis in drought resistance of elephant grass

WGCNA was used to probe the relationship between genes and physiological indicators, and between genes and genes. The DEGs with TPM < 1 were removed, and the remaining DEGs were inputted into the WGCNA network. Pearson's correlation matrix and network topology analysis were used to calculate the soft thresholding power and the gene correlation, respectively. Then, the adjacent relationship was converted to a topological overlap matrix. In standard WGCNA networks, the mergeCutHeight, power, and minModuleSize values were set to 0.3, 5, and 20, respectively. The networks were visualized using Cytoscape version 3.7.2.

Quantitative RT-PCR validation

Quantitative RT-PCR analysis was conducted on a CFX Connect qPCR detection system to validate the accuracy of the RNA-seq results. Six randomly selected genes and 12 identified key genes were detected, with *PpACTB* as the reference gene (the primers are shown in Supplementary Table 4 as supplementary data). The $2^{-\Delta\Delta CT}$ method was used to calculate the relative expression levels of genes.

Statistical analysis

Physiological data were statistically analyzed using SPSS 20.0. The significance of differences between every treatment was detected by Duncan's multiple comparative and one-way ANOVA ($P < 0.05$).

Availability of data and materials

All data generated or analyzed during this study are included in this published article. The RNA-Seq data and the datasets presented in this study can be found at the NCBI repository, accession number PRJNA928572, https://submit.ncbi.nlm.nih.gov/subs/sra/SUB12636465/overview.

Results

Physiological evaluation of elephant grass seedlings in response to SL treatments

Exogenous SL treatments affected the photosynthetic indices of elephant grass (Figure 1). Compared with normal watering treatment and drought rehydration management, the plants treated with SL showed a significant difference in Pn, Gs, Tr, and PIabs which were significantly higher than those in plants exposed to drought-rehydration condition and lower than those in plants exposed to normal watering environment. The values of Pn, Gs, Tr and PIabs were 21.77 μmol m^{-2} s^{-1}, 182.67 mol m^{-2} s^{-1}, 3.75 mol m^{-2} s^{-1}, and 3.08, respectively (Figures 1A, C, D, H). The Ci, VPD, and WUE could be restored to CK under SL treatments, but in the drought-rehydration process, the values of Ci and VPD were higher than those in CK and SL treatment. Meanwhile, the values of WUE showed the opposite trend (Figures 1B, E, F). However, among all photosynthetic indices, Fv/Fm was not significantly affected by the three treatments, and the values were between 0.75 and 0.79 (Figure 1G).

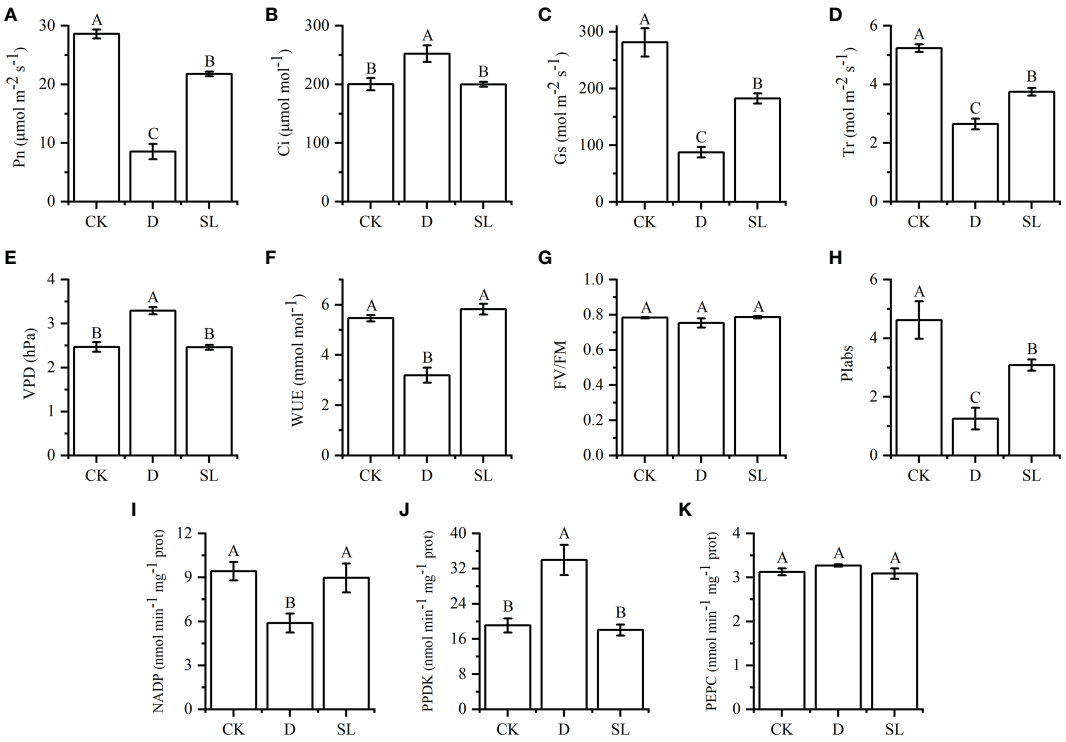


FIGURE 1
Effects of normal, drought-rehydration, and exogenous SL treatments on the growth of elephant grass **(A–K)** Photosynthesis-related indices, including Pn, net photosynthetic rate; Ci, intercellular CO_2 concentration; Gs, stomatal conductance; Tr, transpiration rate; VPD, vapor pressure deficit; WUE, water use efficiency; FV/FM, optimal/maximal quantum yield of PSII; PIabs, photosynthetic performance index; NADP-ME, nicotinamide adenine dinucleotide phosphate-malic enzyme; PPDK, pyruvate phosphate dikinase; and PEPC, phosphoenolpyruvate carboxylase. Different capital letters in the same panel indicate statistical significance ($P < 0.05$).

For the key enzymes in photosynthetic reactions of leaves, NADP-ME and PPDK exhibited significant differences among the three treatments, especially after SL treatment. The values of NADP-ME and PPDK were 8.97 and 18.04 $nmol^{-1}$ min^{-1} mg, respectively, which were similar to those under control condition. On the contrary, the values of NADP-ME and PPDK in drought-rehydration treatment were lower and higher than those in CK and SL treatments, respectively (Figures 1I, G, K).

After drought-rehydration and SL treatments, MDA, PRO, POD, and SOD showed different trends in the leaves and roots. The value of MDA in the leaves and roots under SL treatment were 20.27 and 6.08 $nmol^{-1}$g FW, and both values in leaves and roots were significantly higher and lower than those under control and drought-rehydration conditions, respectively (Figure 2A). PRO and POD had the same decreased trend in the roots, the values of which were significantly less in drought-rehydration and SL treatment than in control conditions. Meanwhile, SOD had the opposite trend under the three different conditions with values significantly higher in drought-rehydration and SL treatments (Figures 2B–D). However, in leaves, the value of PRO and POD were 46.55 μg^{-1} g FW and 79.82 U g^{-1} FW under the drought-rehydration and SL treatment conditions, respectively (Figures 2B, C). SOD had no significant difference in leaves (Figure 2D).

Different treatments created different growth conditions of elephant grass elephant grassroots. Spraying SL in elephant grass yielded more fibrous roots (Figure 3A). The values of Len, SA, NTips and Vol were 580.97 cm, 38.96 cm^3, 1289.33, and 0.21 cm^3 under SL treatment, respectively, which were significantly higher than those under control and drought-rehydration treatment (Figures 3B–E). Compared with normal growth condition, spraying SL had no significant difference and spraying water had remarkable discrepancy in the AvgD of roots, the values of which were 0.22 and 0.20 cm, respectively (Figure 3F).

A total of 22 hormones that are common and important in plants were further analyzed to accurately identify hormone metabolites affected by exogenous SL regulating the response of elephant grass to drought stress (Supplementary File 1). Only five hormones had significant difference in leaves and roots under different treatments. In roots, the concentration of 6-BA significantly decreased in drought-rehydration and SL treatments compared with that in control condition (Figure 4A). Meanwhile, the ABA, MeSA, and NAA concentrations had an opposite trend. In leaves, the values of ABA, MeSA, and NAA concentrations were 0.003, 0.003, and 0.029 ng^{-1} mg in SL treatment, respectively, higher than those in control condition (Figures 4B–D). The JA in leaves had no significant changes under spraying water and SL compared with that under CK (Figure 4E). JA was higher in roots under drought-rehydration treatment than under SL treatment at 0.05 ng^{-1} mg.

Transcriptional analysis revealing key genes affecting SL regulation of elephant grass response to drought stress

A total of 18 RNA-seq libraries from CK, drought-rehydration, and SL treatments were constructed to explore the mechanism underlying SL regulating the response of elephant grass root and leaf to drought stress. All original FASTQ data files were submitted to the NCBI Sequence Read Archive under accession number SUB12636465. An overview of the sequencing is listed in Supplementary Table 1, and more than 98.3% and 95.0% of bases had q-value > 20 and 30 (an error probability of 0.02%),

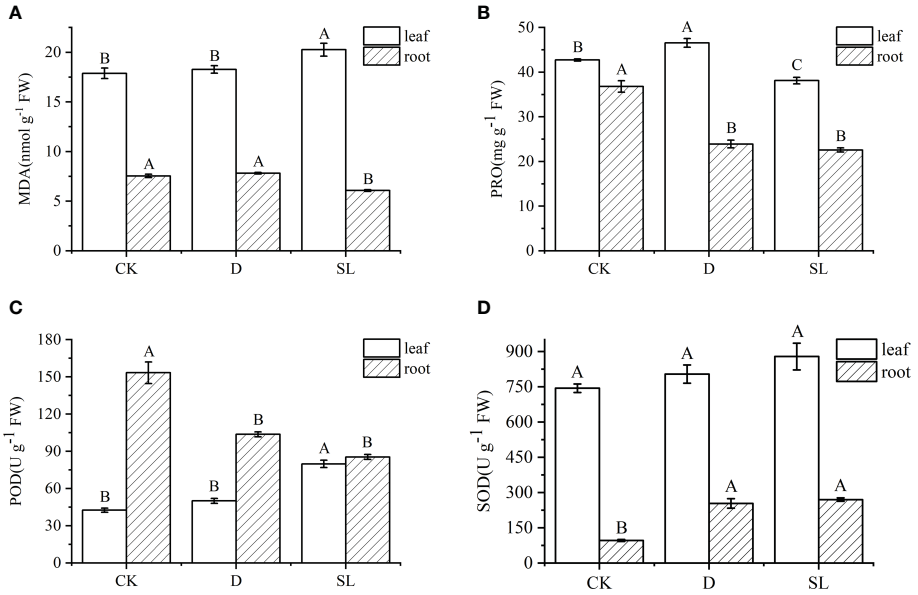


FIGURE 2
Physiological indices of elephant grass regulated under normal, drought-rehydration, and exogenous SL treatment conditions **(A)** Malondialdehyde (MDA), **(B)** proline (PRO), **(C)** peroxidase (POD), and **(D)** superoxide dismutase (SOD). Different capital letters in the same panel indicate statistical significance ($P < 0.05$).

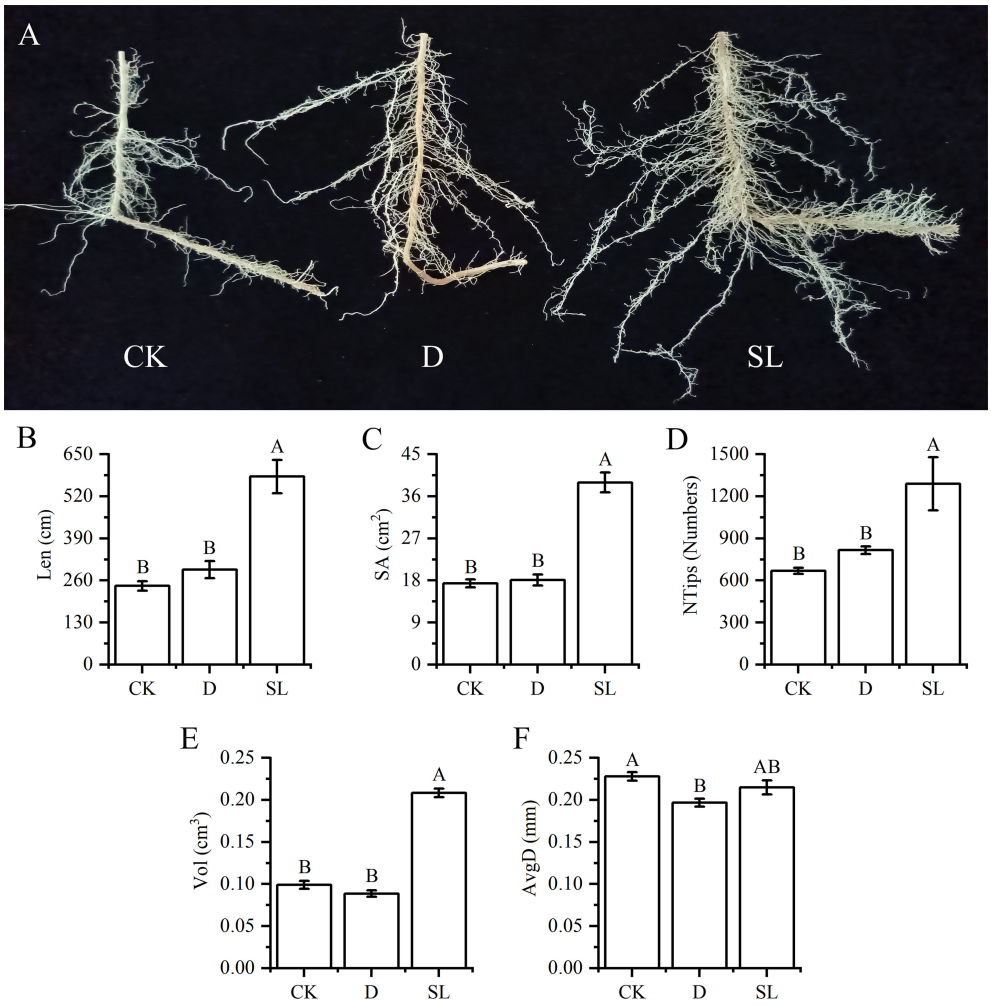


FIGURE 3
Growth of elephant grass roots under drought-rehydration and SL treatment compared with that under CK **(A)** Image of elephant grass roots under different treatment conditions. **(B–F)** Root length (Len), surface area (SA), number of tips (NTips), root volume (Vol) and average diameter (AvgD) among normal, drought-rehydration, and exogenous SL treatments of elephant grass. Different capital letters in the same panel indicate statistical significance ($P < 0.05$).

respectively. The GC-content ranged between 52.84% and 53.57% (Supplementary Table 1). After removing low-quality reads, a total of 797 clean reads were generated from the 18 samples, and 700 million mapped genes were compared with the reference genome (Supplementary Table 1). All 84,296 genes and 137,299 transcripts, including 57,491 known genes, 26,805 new genes, 53,975 known transcripts, and 83,324 new transcripts, were tested (Supplementary Table 2). The number of all known genes in GO, KEGG, KOG, NR, Swiss-Prot, and Pfam databases were 49,452 (86.0%), 25,627 (44.6%), 56,848 (98.9%), 57,400 (99.8%), 49,026 (85.3%), and 50,591 (88.0%), respectively (Supplementary Table 3).

The gene expression levels were calculated and normalized using the TPM method; |log2FC| > 2 and q-value ≤ 0.01 were set for the screening of significant differential expression. In total, 765, 2,756, 2,907, 2,325, 2688 and 1,708 DEGs were upregulated, whereas 622, 1,543, 1,627, 1,826, 2,008 and 2,208 DEGs were downregulated at R1 (RSL vs. RD), R2 (RSL vs. RCK), R3 (RCK vs. RD), L1 (LSL vs. LD), L2 (LSL vs. LCK) and L3 (LCK vs. LD) combinations, respectively (Figure 5A). All the DEGs were annotated into the KEGG database, and 48 pathways were in the network, which had significant enrichment, including top three pathways (phenylpropanoid biosynthesis, nitrogen metabolism and alpha-Linolenic acid metabolism) (Figure 5B). All the upregulated and downregulated DEGs were shown at Venn figures, in which 1,488 and 402 upregulated DEGs were located at L1 and R1, respectively, and 589 and 540 downregulated DEGs were exhibited at the L2 and R2 combinations, respectively (Figures 5C, D). All the 13,656 identified DEGs were classified into nine groups by hierarchical clustering heatmap, and the genes in groups 5 and 8 had significantly high expression under SL treatment in the root and leaf, respectively (Figure 5E). The transcription factors (TFs) of 47 families were included in the DEGs during drought rehydration and spraying SL (Supplementary File 2). The largest three group of TFs was the basic helix-loop-helix family, myeloblastosis and ethylene responsive factor, which have 349, 345 and 292 TFs. Six genes

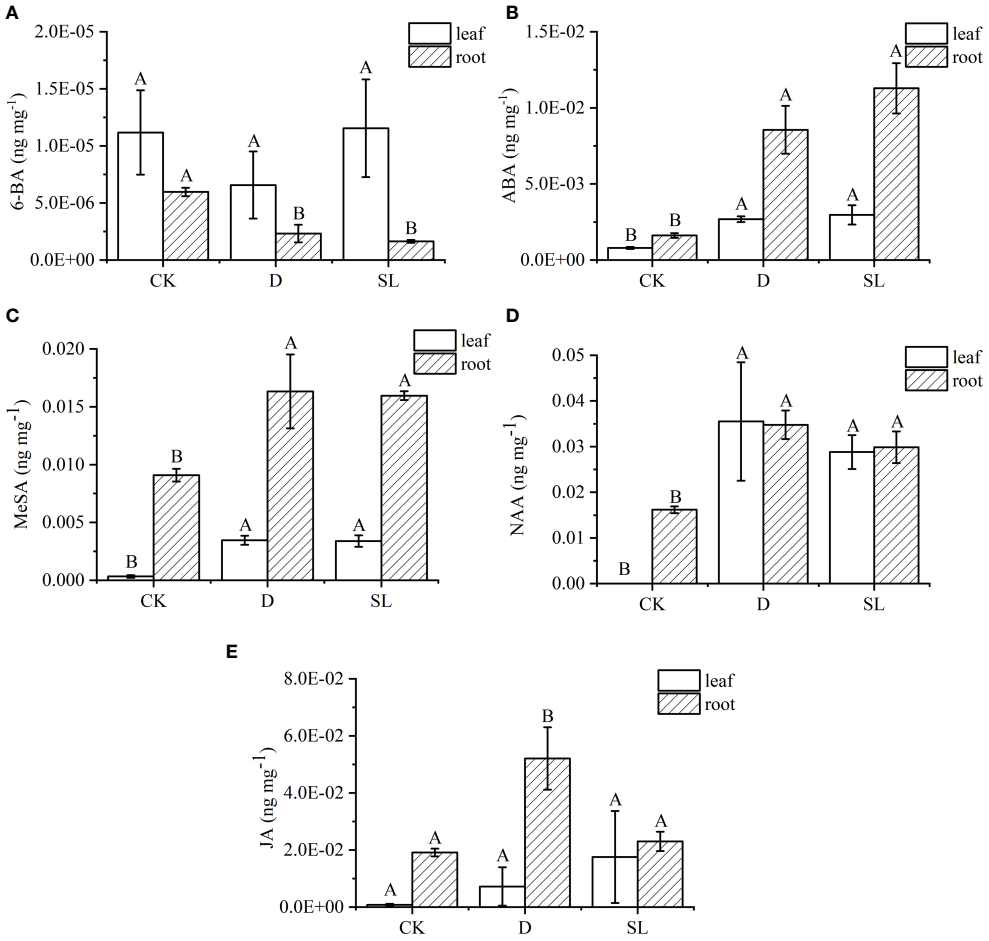


FIGURE 4
Hormone metabolite abundances among three different treatments of elephant grass on the basis of LC-MS analysis **(A–E)** Concentration of 6-benzylaminopurine (6-BA), abscisic acid (ABA), methyl salicylate (MeSA), naphthlcetic acid (NAA), and jasmonic acid (JA) between leaves and roots. Different capital letters in the same panel indicate statistical significance ($P < 0.05$).

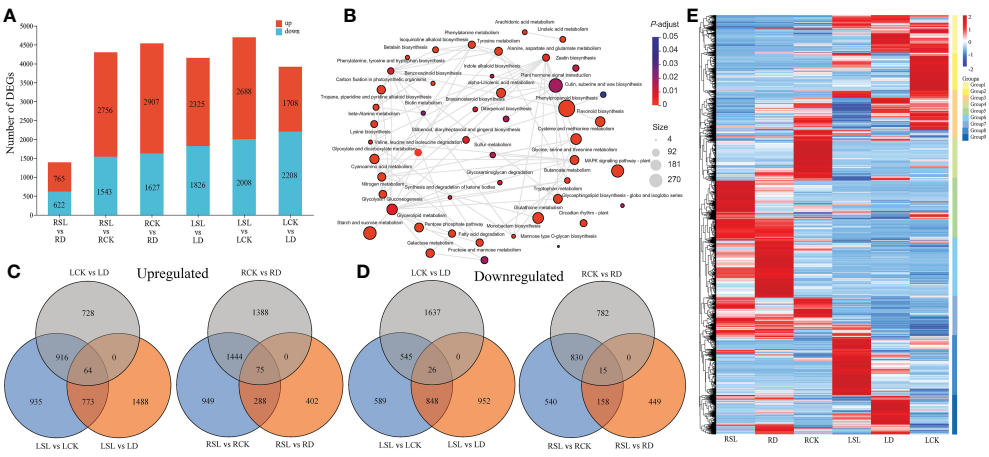


FIGURE 5
Analysis of all identified DEGs **(A)** Upregulated and downregulated DEGs by drought-rehydration and SL treatments in elephant grass. **(B)** Network of significantly enriched KEGG pathways using all detected DEGs. Colors correspond to groups of corrected *p*-value. **(C–D)** Venn diagrams showing the upregulated and downregulated DEGs across six comparisons. **(E)** Hierarchical clustering heatmap of all identified DEGs. Each column corresponds to control, drought-rehydration, and SL treatments in root and leaf, and nine groups were separated.

were randomly selected to validate the accuracy of the RNA-seq results, and the relative expression levels of genes had the same tendency between the qPCR system and the RNA-seq results (Supplementary Figure 1).

Generation of weighted gene co-expression network analysis module and hub gene analysis

The gene co-expression profiles of elephant grass in response to control, drought-rehydration, and drought-SL treatments were analyzed using WGCNA to explore the relationship between genes and physiological indexes, especially genes in modules that were significantly correlated with physiological indicators. A total of 17 co-expression modules were identified and obtained, of which eight modules had the most significant correlation with all physiological indicators (Figure 6). The photosynthetic indices were negatively correlated with brown and turquoise modules and positively correlated with blue, magenta, and yellow module; the correlation coefficients were between -0.75 and -0.90 and between 0.75 and 0.91, respectively. However, the brown and turquoise modules had positive correlation with root conditions and phytohormone changes under different treatments, and the correlation coefficient was between 0.74 and 0.93.

A total of 1,223, 1,167, 2,993, 293, 97, 93, 902, and 110 genes could be found in the blue, brown, turquoise, magenta, tan, salmon, yellow and green-yellow modules, respectively (Supplementary File 3). Further analyses were conducted with the blue module as an example (Figure 7). Totally 1,223 genes were annotated in the KEGG pathway database, and 180 genes were significantly enriched in 16 pathways. The top three most significantly enriched pathways were carbon fixation in photosynthetic organisms, photosynthesis, and glyoxylate and dicarboxylate metabolism (Figure 7A). The top 30 hub genes with a higher weight in the blue module were selected for analysis and drawing the network. Genes *CpA0700513.1*, *CpB0202678.1*, *CpA0504761.1*, *CpA0503348.1*, and *CpA0700959.1* had higher degree than 517 (Figure 7B; Supplementary File 4). The TPM hierarchical clustering heatmap of 30 hub genes and the Venn analyses of 180 genes with KEGG function annotation were shown (Figures 7C, D). Among them, 13 genes were not only hub genes but also genes with significant enrichment function.

In accordance with the blue module analysis methods, 20, 22, two, five, five, 12, and three significantly enriched pathways and 202, 444, 8, 13, 11, 89, and 12 genes with KEGG enrichment function were identified on the other seven modules (Supplementary Files 5, 6). The top 30 hub genes were shown as network and hierarchical clustering heatmap (Supplementary File 7; Supplementary Figures 2, 3). The Venn analysis showed that the common parts of KEGG enrichment function genes and hub genes were five, two, one, eight, eight, zero, and seven genes in the brown, turquoise, magenta, tan, salmon, yellow, and green-yellow modules, respectively (Supplementary Figure 4). In general, in the eight modules that were significantly related to physiological indicators, 44 DEGs that played a major role in SL response to drought stress were identified (Supplementary File 8).

Functional analysis and verification of key genes

For the 44 identified key DEGs, hierarchical clustering heatmaps were analyzed and the genes had more differential expression in leaves than in roots under exogenous SL regulating drought stress (Supplementary Figure 5A). The KEGG significant enrichment analysis network diagram of these genes showed that valine, leucine, and isoleucine degradation; carbon fixation in photosynthetic organisms; and fatty-acid degradation were the top three most significantly enriched pathways (Supplementary Figure 5B). Correlation coefficient analysis was performed to understand the relationship between the 44 key DEGs and various physiological indicators (Supplementary Figure 5C). Genes *CpA0101446.1*, *CpB0105721.1*, *CpB0104060.1*, and *CpB0700538.1* showed positive correlation with photosynthetic indices, physiology indicators, and one kind of phytohormone 6-BA and negative correlation with root growth and other

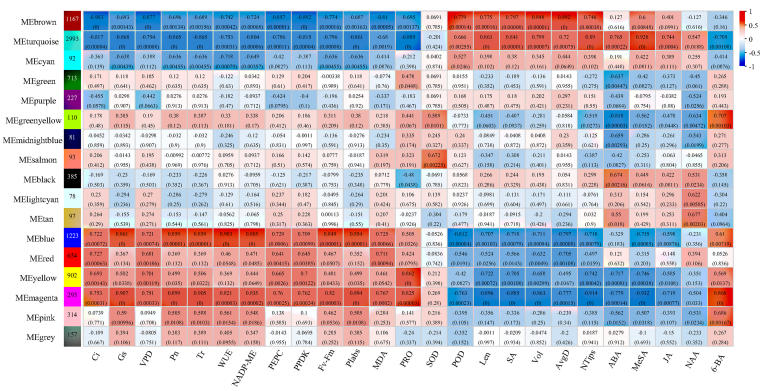


FIGURE 6
Module–trait relationship with physiological, root growth, and hormone indices of elephant grass The number represents the correlation coefficient of modules with physiological indices. The number in the bracket means *p*-value.

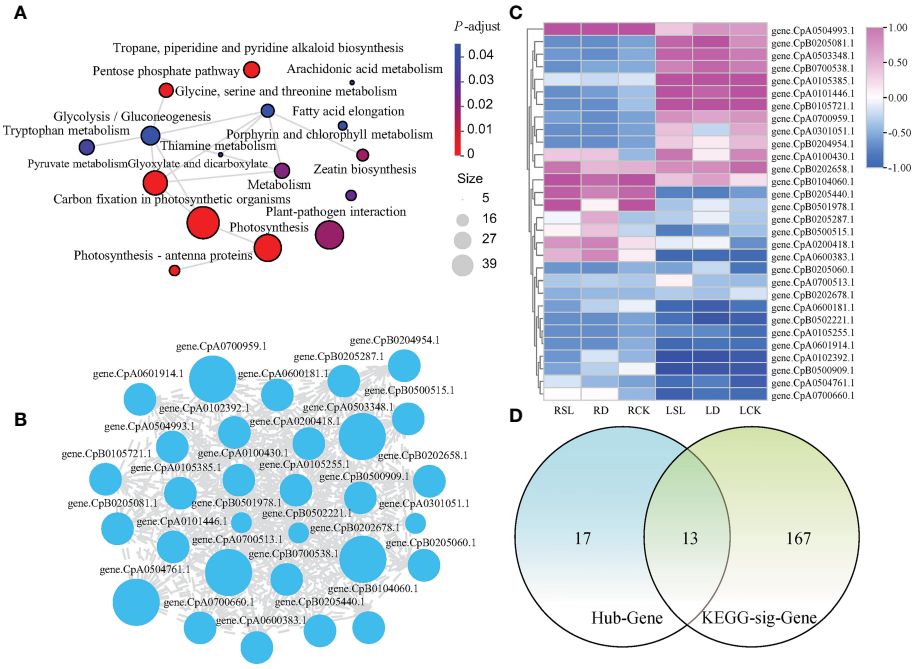


FIGURE 7
Analysis of 1223 genes in blue module **(A)** Network analysis of significant enrichment by KEGG pathway. **(B)** Cytoscape analysis of top 30 hub genes with a higher weight in blue module. **(C)** Hierarchical clustering heatmap of 30 hub genes. **(D)** Venn analysis revealing the common parts between KEGG enriched functional genes and the top 30 hub genes of higher weights in blue module.

phytohormones. In this analysis, the 44 identified DEGs became connected when the Network Analyst program was applied (Supplementary Figure 5D). Three independent groups, including purple, blue, and pink, were shown. In the purple group, many DEGs had high positive correlation coefficient with photosynthesis indices, whereas the DEGs in the pink group had negative correlation coefficient with phytohormone when SL was sprayed on elephant grass (Supplementary Figures 5C, D).

The FDR of 44 key DEGs under four comparisons was analyzed to further understand the response of exogenous SL treatment to drought in elephant grass (Table 1). Compared with drought-rehydration treatment, SL treatment showed many downregulated and upregulated genes located in the leaves. Phosphoenolpyruvate carboxykinase (*PpPEPCK*), ribulose bisphosphate carboxylase small chain A (*PpRuBPC*), phosphoglycerate kinase (*PpPGK*), glyceraldehyde-3-phosphate dehydrogenase A (*PpGADPH*), fructose-bisphosphate aldolase (*PpFBA*), and sedoheptulose-1,7-bisphosphatase (*PpSBPase*) had FDRs between 2.24 and 3.52. However, in roots, methylcrotonoyl-CoA carboxylase subunit alpha (*PpMCCA*), methylcrotonoyl-CoA carboxylase beta chain (*PpMCCB*), alanine-glyoxylate aminotransferase 2 homolog 1 (*PpAGT2*), isovaleryl-CoA dehydrogenase (*PpIVD*), acetyl-CoA acetyltransferase (*PpACAT*), and glyoxysomal fatty acid beta-oxidation multifunctional protein MFP-a (*PpMFP2*) were the most important DEGs, with FDRs of 4.12, 3.87, 2.66, 3.4, 3.82, 2.91, −2.34, −3.33, and −2.03, respectively.

The qPCR results of core genes were observed, and they proved the correctness of gene expression. The results showed that *PpPEPCK*, *PpRuBPC*, *PpPGK*, *PpGADPH*, *PpFBA*, and *PpSBPase* had higher expression in SL treatment than in normal and drought-rehydration treatments in leaves (Supplementary Figure 6). In roots, *PpACAT* and *PpMFP2* had low expression under SL treatment, whereas *PpAGT2*, *PpIVD*, *PpMCCA*, and *PpMCCB* had the opposite expression under spraying SL (Supplementary Figure 6).

Discussion

SL is a recently discovered class of carotenoid-derived phytohormones and has multiple biological functions, from regulating plant architecture and stimulating the germination of parasitic plant seeds to participating in a range of plant development processes, including root growth, leaf senescence, and photomorphogenesis. In addition, SL could be associated with various stress environments, and it supports a positive potential strategy for abiotic stress tolerance in plants (Ha et al., 2014; Li et al., 2019). In this study, a comprehensive approach was taken to reveal the effect of physiological response during drought-rehydration and SL recovery stages in elephant grass seedlings. Combined analyses of transcriptome and metabolome were applied to investigate the response and molecular mechanism of exogenous SL on C_4 plant under drought stress.

10.3389/fpls.2023.1186718

TABLE 1 Analyzed direction of gene expression changes of 44 key DEGs involved in the eight mode during drought-rehydration and spraying SL of elephant grass.

	Direction of gene expression changes					
Gene ID	LCK vs LSL	LSL vs LD	RCK vs RSL	RSL vs RD	Gene description	Abbreviation
gene.Cp0000455.1			4.12		Methylcrotonoyl-CoA carboxylase subunit alpha	*MCCA*
gene.CpA0101446.1		3.55			Phosphoenolpyruvate carboxykinase	*PEPCK*
gene.CpA0102214.1			2.66		Methylcrotonoyl-CoA carboxylase beta chain	*MCCB*
gene.CpA0103469.1	-2.51		-2.26		Probable glycerol-3-phosphate dehydrogenase	*GPDH*
gene.CpA0104386.1					Acyl-coenzyme A oxidase 4	*ACOX4*
gene.CpA0203545.1					5'-nucleotidase SurE	*surE*
gene.CpA0301051.1		3.52			Ribulose bisphosphate carboxylase	*RuBPC*
gene.CpA0403417.1	-2.52		-2.54		3-ketoacyl-CoA thiolase 2	*KAT2*
gene.CpA0504761.1	-2.4	2.16			Photosystem I chlorophyll a/b-binding protein 5	*CAB5*
gene.CpA0504993.1		2.03			Aspartate aminotransferase	*AST*
gene.CpA0700959.1	-2.04	2.84			Sedoheptulose-1,7-bisphosphatase	*SBPase*
gene.CpA0702908.1			2.4		2-oxoacid dehydrogenases acyltransferase	*BLAODA*
gene.CpA0703361.1			-2.34		Acetyl-CoA acetyltransferase	*ACAT*
gene.CpB0103289.1			-3.33		Glyoxysomal fatty acid beta-oxidation multifunctional protein MFP-a	*MFP2*
gene.CpB0103297.1			-2.03		Glyoxysomal fatty acid beta-oxidation multifunctional protein MFP-a	*MFP2*
gene.CpB0104060.1	-2.36	2.24			Phosphoglycerate kinase	*PGK*
gene.CpB0105196.1			2.91		Isovaleryl-CoA dehydrogenase	*IVD*
gene.CpB0105721.1		2.53			Fructose-bisphosphate aldolase	*FBA*
gene.CpB0203332.1	4.12				Chlorophyll a-b binding protein 2	*CAB2*
gene.CpB0205081.1	-2.64				Photosystem II 22 kDa protein 1	*PSBS1*
gene.CpB0500657.1			3.82		Alanine-glyoxylate aminotransferase 2 homolog 1	*AGT2*
gene.CpB0700538.1		2.29			Glyceraldehyde-3-phosphate dehydrogenase A	*GAPDH*
gene.CpA0100504.1			2.1		Xylanase inhibitor protein 1	*XIP1*
gene.CpA0102576.1			2.57		Quinolinate synthase	*QS*
gene.CpA0104511.1					Isovaleryl-CoA dehydrogenase	*IVD*
gene.CpA0105385.1		2.75			Fructose-bisphosphate aldolase	*FBA*
gene.CpA0201611.1			-2.2		Phospholipase D delta	*PLDdelta*
gene.CpA0201626.1			-2.15		Phospholipase D delta	*PLDdelta*
gene.CpA0202256.1	-2.77		-3.77		flavonoid 3'-monooxygenase	*F3'M*
gene.CpA0202265.1	-3.9		-3.11		Trimethyl tridecatetraene synthase	*TMTT*
gene.CpA0203312.1			2.15		Primary amine oxidase 2	*PrAO2*
gene.CpA0301066.1			3.17		Probable O-methyltransferase 2	*OMT2*
gene.CpA0303976.1			3.87		Methylcrotonoyl-CoA carboxylase subunit alpha	*PpMCCA*
gene.CpA0503348.1	-2.43	2.71			Glyceraldehyde-3-phosphate dehydrogenase A	*GAPDH*
gene.CpA0600615.1			2.07		Alcohol dehydrogenase-like 5	*ADH5*

(Continued)

10.3389/fpls.2023.1186718

TABLE 1 Continued

Gene ID	Direction of gene expression changes				Gene description	Abbreviation
	LCK vs LSL	LSL vs LD	RCK vs RSL	RSL vs RD		
gene.CpB0102025.1			3.4		Methylcrotonoyl-CoA carboxylase beta chain	*MCCB*
gene.CpB0103652.1			2.1		Peroxidase 15	*POD15*
gene.CpB0201645.1			2.66		2-oxoacid dehydrogenases acyltransferase	*BLAODA*
gene.CpB0204000.1			-2.58		Mitogen-activated protein kinase kinase kinase 17	*MAPKKK17*
gene.CpB0204954.1		2.86			Sedoheptulose-1,7-bisphosphatase	*SBPase*
gene.CpB0205060.1	-2.53				Photosystem II 22 kDa protein 1	*PSBS1*
gene.CpB0301112.1					Probable uridine nucleosidase 2	*URH2*
gene.CpB0400578.1			3.25	2.29	Indole-3-acetaldehyde oxidase	*IAAId*
gene.CpB0700954.1	-3.8		-2.19		Beta-glucosidase 18	*BGLU18*

red, upregulated DEGs.
green, downregulated DEGs.

Response of photosynthesis to SL regulating drought stress in elephant grass

Foliar application of SL on elephant grass largely ameliorated drought stress symptoms. In particular, photosynthesis had different photosynthetic capacities during the three growth conditions. The recovery ability after re-watering is usually considered important for successful adaptation of plants to arid conditions. Rehydration helps plants offset the damage and recover their physiological functions from drought stress (Zhou et al., 2021). Comparison of different recovering conditions with water and SL demonstrated that spraying SL provided elephant grass with better recovery capacity, with Pn, Gs, Tr, WUE, and PIabs being significantly higher than those under drought-rehydration condition (Figures 1A, C, D, F, H). These five indices are a comprehensive reflection of the photosynthetic and WUE characteristics of plants. The alleviation of drought damage by SL was significantly better than that by rewatering treatment, indicating that SL was more effective in promoting plant recovery under drought stress. SL could improve the photosynthetic capacity of plants, and it showed higher tolerance to drought stress (Li et al., 2023). This result is similar to that in other species, such as *V. vinifera* (Min et al., 2018; Li et al., 2021), maize seedlings (Sattar et al., 2022), and winter wheat (*Triticum aestivum*) (Sedaghat et al., 2021). Therefore, SL has stronger ability to repair drought plants than that of only using water.

Elephant grass as a model C_4 plant species has higher photosynthesis due to a repertoire of C_4 enzymes, including NADP-ME, PPDK, and PEPC, to enhance CO_2 fixation (Li et al., 2022). The activities of the three enzymes above could improve photosynthesis efficiency, enhance stress tolerance, and ultimately increase crop yield under adverse conditions (Chen et al., 2019; He et al., 2020). PPDK and PEPC are the key enzymes that convert CO_2 in the atmosphere into oxaloacetic acid (OAA) after being absorbed through stomata in mesophyll cells, while NADP-ME is a crucial enzyme that concentrates CO_2 in bundle sheath cells (Huang et al., 2016). NADP-ME and PPDK had the opposite variation trend under drought-rehydration and SL treatments, possibly because after spraying SL under drought stress, the Gs of plants was more open than that under drought-rehydration conditions. Therefore, the absorbed CO_2 content was higher, which reduced the activity of PPDK that catalyzes the conversion of CO_2 to OAA. This phenomenon leads to a decrease in intercellular CO_2 concentration and finally promotes an increase in the activity of NADP-ME in the CO_2 concentration mechanism to improve the photosynthetic capacity. This result also showed that different treatment methods could cause diverse enzyme activities at distinct stages, thus leading to different photosynthetic capacity, which is a strategy for plants to adapt to changed environments.

Insights into the root system and SL crosstalk profiling with other phytohormones

The root system is the main organ for nutrient and water absorption in plants, even those faced with severe drought environment. However, water deficit conditions usually lead to self-thinning of the root system, abscise some death roots and decreased root-absorption capacity (Kim et al., 2020; Li et al., 2021). SL is generally considered to inhibit adventitious root formation and regulate the elongation of primary roots in eudicot plants while increasing the number of adventitious roots and promoting the elongation of seminal roots in grass plants in short time (Sun et al., 2022). This opinion also confirmed the results of the present study that SL treatment could significantly increase the Len, SA, NTips, and Vol of elephant grass. Similar results were found in *A. thaliana*, that is, SL positively controls root elongation and regulates the primary root length by regulating the content of *A. thaliana* growth hormone (Ruyter-Spira et al., 2011). The exogenous application of SL increases the length of the original roots, improves root phenotype, and promotes growth and development under stress conditions.

The results exhibited a complex phytohormonal response coordinated from multiple signaling pathways to respond to drought, in which ABA, MeSA, and NAA showed seemingly same increasing trends in roots and leaves under stress, whereas 6-BA showed a decreased tendency in roots. ABA is considered a "stress hormone," whose levels could be constantly adjusted to adapt to environmental conditions (Liu et al., 2022). The ABA–SL interaction has been studied from various perspectives effectively, and inconsistencies in the interaction were found, which could be caused by the crosstalk with other phytohormones that could have antagonism on two components (Kaniganti et al., 2022). The expression levels of SL biosynthetic genes go down in ABA-deficient lines and SL induced mir156 accumulation, which increased guard cell sensitivity to ABA and subsequently generates stomatal closure for decreased water loss (Brun, 2020; Kaniganti et al., 2022). In consequence, the network regulation of SL is still inseparable from ABA.

SL has been proposed to serve as a modulator of auxin transport to regulate root elongation (Ruyter-Spira et al., 2011; Hu et al., 2018), and some research found that plant may increase SL production to promote symbiosis establishment to respond to stress (Ruiz-Lozano et al., 2015). In tall fescue, SL positively promoted crown root elongation under abiotic stress due to the interference of NAA transport and regulation of cell division (Du et al., 2018). Meanwhile, exogenous application of NAA was incapable of restoring root-hair elongation and symmetric root growth in the presence of SL on tomato (Koltai et al., 2010). However, in the present research, the content of NAA in leaves and roots was significantly higher than that in the control in elephant grass, may be due to the indirect correlation between SL and NAA during the interaction of plant growth regulators. Moreover, the increase in NAA content is more obviously related to the external environment.

Overall, the SL crosstalk with other phytohormones is complicated, and it needs further exploration. In this research, the SL contents and SL synthetic pathway were not identified using LC-MS/MS and RNA-Seq in roots and leaves, respectively. First, elephant grass may not produce endogenous SL by themselves. Second, as a signal molecule, SL may be quickly degraded and absorbed in plants through phytohormone metabolism.

Key genes regulating SL response of elephant grass to drought stress

According to the WGCNA and correlation analyses, 18 unigenes encoding 12 genes were identified as the key genes for SL regulating the response of elephant grass to drought stress. Among them, *PpPEPCK*, *PpRuBPC*, *PpPGK*, *PpGAPDH*, *PpFBA*, and *PpSBPase* were demonstrated to be profusely expressed in the leaves, whereas *PpACAT*, *PpMFP2*, *PpAGT2*, *PpIVD*, *PpMCCA*, and *PpMCCB* were abundantly expressed in the roots of elephant grass. A model showing SL affects elephant grass to cope with drought stress from the leaves to the roots was proposed (Figure 8).

SL regulated the response of elephant grass to drought stress from carbon fixation in photosynthetic organisms; valine, leucine and isoleucine degradation; and fatty acid degradation pathways. PEPCK plays a very important role in plant metabolism, and it is involved in regulating gluconeogenesis and catalyzing the transformation of fat or malic acid into sugar. It could also participate in photosynthesis as a key CO_2 concentrating enzyme in C_4 plants (Lee et al., 2022). PEPCK catalyzed the reversible decarboxylation of OAA to form PEP and CO_2, which increased the concentration of CO_2 in the sheath cells of vascular bundles and saturated the concentration of CO_2 in Rubisco substrate, thus improving the efficiency of photosynthesis (Huang et al., 2016).

Meanwhile, GAPDH catalyzes the reduction of 3-phosphoglycerate to triose phosphate, a key step in carbon reduction phase of photosynthesis. Both enzymes have been researched in Kentucky bluegrass, and the results exhibited that

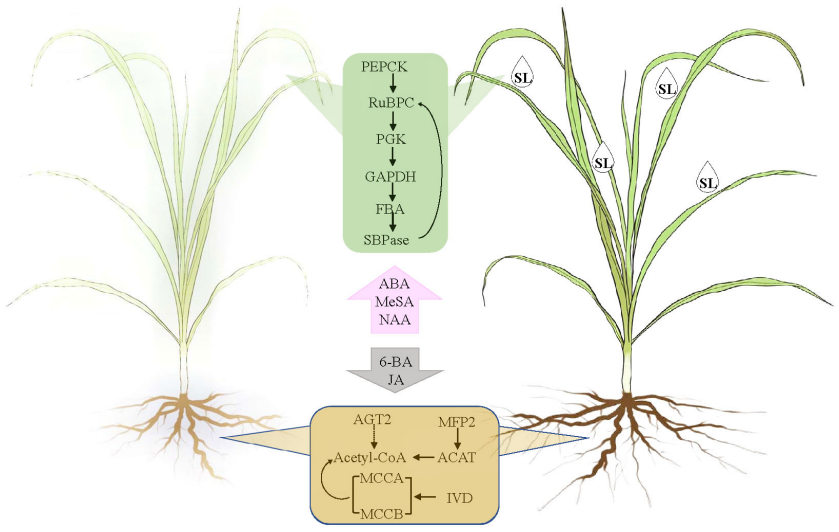


FIGURE 8
Model of the key hormones and gene adjustment of drought root and leaf elephant grass in response to SL treatment.

under drought stress and re-watering recovery conditions, carbon reduction regulated by GAPDH and carboxylation controlled by Rubisco could be the key metabolic processes of genetic variation in Pn responses to drought stress, and active Rubisco and GAPDH could be involved in superior post-drought recovery (Xu et al., 2013). In plants, PGK not only catalyzes the reversible conversion of 1,3-bisphosphoglycerate to 3-phosphoglycerate in glycolysis but also participates in gluconeogenesis and the Calvin-Benson cycle (Rosa-Tellez et al., 2018; Massange-Sánchez et al., 2020). GAPDH and PGK participate in the same compartment, and at the same time, in photosynthetic and glycolytic reactions (Rosa-Tellez et al., 2018). These key genes showed similar performance during the process of plant recovery and the gene expression increased sharply, which also reflected the main regulatory mechanism of SL in the process of regulating the growth of elephant grass.

PpACAT and *PpMFP2* were downregulated in the roots under SL treatment. These genes influence acetyl-CoA, which, as the core enzyme, is the connection site of valine, leucine, and isoleucine degradation and fatty-acid degradation pathways based on the KEGG enrichment analysis. Meanwhile, *PpIVD*, *PpMCCA* and *PpMCCB* were focused on fatty-acid metabolism, and they were upregulated in the roots under SL treatment. Combination of root physiological indices showed that SL treatment considerably improved the growth of roots and promoted the development of fibrous roots. This finding may be the result of the interaction regulation of root development by phytohormone interaction. However, how the identified genes worked on root growth must still be explored.

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, This data can be found here: https://dataview.ncbi.nlm.nih.gov/object/PRJNA928572.

Author contributions

JZ conceived and performed the experiments and wrote the manuscript. YJL, YL, WL, XF, and QF performed parts of the experiments. RM supported the ideas of experiment and modified the manuscript. FY designed the experiments, provided experimental place, and financial assistance. All authors contributed to the article and approved the submitted version.

Funding

This work was supported by the National Natural Science Foundation of China (32101418, 42075116).

Acknowledgments

We thank Shuizhen Huang, National Herbage Cultivar Evaluation Station (Jianyang), Agricultural Ecology Institute, FAAS, for their supporting seedling of elephant grass, and acknowledge National Natural Science Foundation of China for the financial supports.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1186718/full#supplementary-material

SUPPLEMENTARY FILE 1
A total of 22 hormones metabolites.

SUPPLEMENTARY FILE 2
TF analysis results.

SUPPLEMENTARY FILE 3
The genes identified in the eight key modules.

SUPPLEMENTARY FILE 4
The degree of top 30 hub genes with a higher weight in eight modules.

SUPPLEMENTARY FILE 5
KEGG pathways significant enrichment in eight modules.

SUPPLEMENTARY FILE 6
The genes in the KEGG significant enrichment pathway in eight modules.

SUPPLEMENTARY FILE 7
The top 30 hub genes in eight modules.

SUPPLEMENTARY FILE 8
The key 44 genes.

References

Banks, J. M., Percival, G. C., and Rose, G. (2019). Variations in seasonal drought tolerance rankings. *Trees-Struct Funct.* 33 (4), 1063–1072. doi: 10.1007/s00468-019-01842-5

Bhoi, A., Yadu, B., Chandra, J., and Keshavkant, S. (2021). Contribution of strigolactone in plant physiology, hormonal interaction and abiotic stresses. *Planta* 254 (2), 28. doi: 10.1007/s00425-021-03678-1

Brun, G. (2020). At The crossroads of strigolactones and abscisic acid pathways: a role for miR156. *Plant Cell Environ.* 43 (7), 1609–1612. doi: 10.1111/pce.13787

Chen, M., MacGregor, D. R., Dave, A., Florance, H., Moore, K., Paszkiewicz, K., et al. (2014). Maternal temperature history activates flowering locus T in fruits to control progeny dormancy according to time of year. *Proc. Natl. Acad. Sci. U.S.A.* 111 (52), 18787–18792. doi: 10.1073/pnas.1412274111

Chen, Q. Q., Wang, B. P., Ding, H. Y., Zhang, J., and Li, S. C. (2019). Review: the role of NADP-malic enzyme in plants under stress. *Plant Sci.* 281, 206–212. doi: 10.1016/j.plantsci.2019.01.010

Cook, C. E., Whichard, L. P., Turner, B., Wall, M. E., and Egley, G. H. (1966). Germination of witchweed (Striga lutea lour.): isolation and properties of a potent stimulant. *Science* 154 (3753), 1189–1190. doi: 10.1126/science.154.3753.1189

Daviere, J. M., and Achard, P. (2016). A pivotal role of DELLAs in regulating multiple hormone signals. *Mol. Plant* 9 (1), 10–20. doi: 10.1016/j.molp.2015.09.011

de Saint Germain, A., Ligerot, Y., Dun, E. A., Pillot, J. P., Ross, J. J., Beveridge, C. A., et al. (2013). Strigolactones stimulate internode elongation independently of gibberellins. *Plant Physiol.* 163 (2), 1012–1025. doi: 10.1104/pp.113.220541

Du, H., Huang, F., Wu, N., Li, X. H., Hu, H. H., and Xiong, L. H. (2018). Integrative regulation of drought escape through ABA-dependent and-independent pathways in rice. *Mol. Plant* 11 (4), 584–597. doi: 10.1016/j.molp.2018.01.004

Feng, Z., Liang, X., Tian, H., Yasuko, W., Huu, N. K., Duy, T. C., et al. (2022). Suppressor of MAX2 1 (SMAX1) and SMAX1-LIKE2 (SMXL2) negatively regulate drought resistance in arabidopsis thaliana. *Plant Cell Physiol.* 63 (12), 1900–1913. doi: 10.1093/pcp/pcac080

Ghorbel, M., Zribi, I., Missaoui, K., Drira-Fakhfekh, M., Azzouzi, B., and Brini, F. (2021). Differential regulation of the durum wheat pathogenesis-related protein (PR1) by calmodulin TdCaM1.3 protein. *Mol. Biol. Rep.* 48 (1), 347–362. doi: 10.1007/s11033-020-06053-7

Ha, C. V., Leyva-Gonzalez, M. A., Osakabe, Y., Tran, U. T., Nishiyama, R., Watanabe, Y., et al. (2014). Positive regulatory role of strigolactone in plant responses to drought and salt stress. *Proc. Natl. Acad. Sci. U.S.A.* 111 (2), 851–856. doi: 10.1073/pnas.1322135111

Haider, I., Andreo-Jimenez, B., Bruno, M., Bimbo, A., Floková, K., Abuauf, H., et al. (2018). The interaction of strigolactones with abscisic acid during the drought response in rice. *J. Exp. Bot.* 69 (9), 2403–2414. doi: 10.1093/jxb/ery089

He, Y. F., Xie, Y. F., Li, X., and Yang, J. (2020). Drought tolerance of transgenic rice overexpressing maize C4-PEPC gene related to increased anthocyanin synthesis regulated by sucrose and calcium. *Biol. Plant* 64, 136–149. doi: 10.32615/bp.2020.031

Hu, Q. N., Zhang, S. X., and Huang, B. R. (2018). Strigolactones and interaction with auxin regulating root elongation in tall fescue under different temperature regimes. *Plant Sci.* 271, 34–39. doi: 10.1016/j.plantsci.2018.03.008

Huang, C. F., Chang, Y. M., Lin, J. J., Yu, C. P., Lin, H. H., Liu, W. Y., et al. (2016). Insights into the regulation of c-4 leaf development from comparative transcriptomic analysis. *Curr. Opin. Plant Biol.* 30, 1–10. doi: 10.1016/j.pbi.2015.12.011

Kaniganti, S., Bhattacharya, J., Petla, B. P., and Reddy, P. S. (2022). Strigolactone, a neglected plant hormone, with a great potential for crop improvement: crosstalk with other plant hormones. *Environ. Exp. Bot.* 204, 105072. doi: 10.1016/j.envexpbot.2022.105072

Kim, Y., Chung, Y. S., Lee, E., Tripathi, P., Heo, S., and Kim, K. H. (2020). Root response to drought stress in rice (Oryza sativa l.). *Int. J. Mol. Sci.* 21 (4), 1513. doi: 10.3390/ijms21041513

Koltai, H., Dor, E., Hershenhorn, J., Joel, D. M., Weininger, S., Lekalla, S., et al. (2010). Strigolactones' effect on root growth and root-hair elongation may be mediated by auxin-efflux carriers. *J. Plant Growth Regul.* 29 (2), 129–136. doi: 10.1007/s00344-009-9122-7

Lee, M. S., Boyd, R. A., and Ort, D. R. (2022). The photosynthetic response of c-3 and c-4 bioenergy grass species to fluctuating light. *Glob. Chang. Biol.* 14 (1), 37–53. doi: 10.1111/gcbb.12899

Li, C., Dong, L., Durairaj, J., Guan, J. C., Yoshimura, M., Quinodoz, P., et al. (2023). Maize resistance to witchweed through changes in strigolactone biosynthesis. *Science* 379 (6627), 94–99. doi: 10.1126/science.abq4775

Li, W. Q., Herrera-Estrella, L., and Tran, L. S. P. (2019). Do cytokinins and strigolactones crosstalk during drought adaptation? *Trends Plant Sci.* 24 (8), 669–672. doi: 10.1016/j.tplants.2019.06.007

Li, Y., Li, S. T., Feng, Q. X., Zhang, J., Han, X. L., Zhang, L., et al. (2022). Effects of exogenous strigolactone on the physiological and ecological characteristics of pennisetum purpureum schum. seedlings under drought stress. *BMC Plant Biol.* 22 (1), 578. doi: 10.1186/s12870-022-03978-y

Li, C. N., Li, L., Reynolds, M. P., Wang, J. Y., Chang, X. P., Mao, X. G., et al. (2021). Recognizing the hidden half in wheat: root system attributes associated with drought tolerance. *J. Exp. Bot.* 72 (14), 5117–5133. doi: 10.1093/jxb/erab124

Li, W. Q., Nguyen, K. H., Tran, C. D., Watanabe, Y., Tian, C. J., Yin, X. J., et al. (2020). Negative roles of strigolactone-related SMXL6, 7 and 8 proteins in drought resistance in arabidopsis. *Biomolecules* 10 (4), 607. doi: 10.3390/biom10040607

Liu, J., Shu, D. F., Tan, Z. L., Ma, M., Guo, N., Gao, S., et al. (2022). The arabidopsis IDD14 transcription factor interacts with bZIP-type ABFs/AREBs and cooperatively regulates ABA-mediated drought tolerance. *New Phytol.* 236 (3), 929–942. doi: 10.1111/nph.18381

Manandhar, S., Funnell, K. A., Woolley, D. J., and Cooney, J. M. (2018). Interaction between strigolactone and cytokinin on axillary and adventitious bud development in zantedeschia. *Physiol. Mol. Biol. Plants* 6 (1), 1–6. doi: 10.4172/2329-955x.1000172

Marzec, M., Daszkowska-Golec, A., Collin, A., Melzer, M., Eggert, K., and Szarejko, I. (2020). Barley strigolactone signalling mutant hvd14.d reveals the role of strigolactones in abscisic acid-dependent response to drought. *Plant Cell Environ.* 43 (9), 2239–2253. doi: 10.1111/pce.13815

Massange-Sánchez, J. A., Casados-Vázquez, L. E., Juarez-Colunga, S., Sawers, R. J. H., and Tiessen, A. (2020). The phosphoglycerate kinase (PGK) gene family of maize (Zea mays var. B73). *Plants (Basel)* 9 (12), 1639. doi: 10.3390/plants9121639

Min, Z., Li, R. Y., Chen, L., Zhang, Y., Li, Z. Y., Liu, M., et al. (2018). Alleviation of drought stress in grapevine by foliar-applied strigolactones. *Plant Physiol. Biochem.* 135, 99–110. doi: 10.1016/j.plaphy.2018.11.037

Muhammad, N., Muhammad, S. N., Rashid, A., Z.I., M., Y.A., M., Yasir, H., et al. (2017). Foliar calcium spray confers drought stress tolerance in maize *via* modulation of plant growth, water relations, proline content and hydrogen peroxide activity. *Arch. Agron. Soil Sci.* 64 (1), 1–1. doi: 10.1080/03650340.2017.1341108

Rosa-Tellez, S., Djoro Anoman, A., Flores-Tornero, M., Toujani, W., Alseek, S., Fernie, A. R., et al. (2018). Phosphoglycerate kinases are Co-regulated to adjust metabolism and to optimize growth. *Plant Physiol.* 176 (2), 1182–1198. doi: 10.1104/pp.17.01227

Ruiz-Lozano, J. M., Aroca, R., Zamarreno, A. M., Molina, S., Andreo-Jiménez, B., Porcel, R., et al. (2015). Arbuscular mycorrhizal symbiosis induces strigolactone biosynthesis under drought and improves drought tolerance in lettuce and tomato. *Plant Cell Environ.* 39 (2), 441–452. doi: 10.1111/pce.12631

Ruyter-Spira, C., Kohlen, W., Charnikhova, T., Zeijl, A. V., Bezouwen, L. V., Ruijter, N. D., et al. (2011). Physiological effects of the synthetic strigolactone analog GR24 on root system architecture in arabidopsis: another belowground role for strigolactones? *Plant Physiol.* 155 (2), 721–734. doi: 10.1104/pp.110.166645

Sattar, A., Ul-Allah, S., Ijaz, M., Sher, A., Butt, M., Abbas, T., et al. (2022). Exogenous application of strigolactone alleviates drought stress in maize seedlings by regulating the physiological and antioxidants defense mechanisms. *Cereal Res. Commun.* 50 (2), 263–272. doi: 10.1007/s42976-021-00171-z

Sedaghat, M., Emam, Y., Mokhtassi-Bidgoli, A., Hazrati, S., Lovisolo, C., Visentin, I., et al. (2021). The potential of the synthetic strigolactone analogue GR24 for the maintenance of photosynthesis and yield in winter wheat under drought: investigations on the mechanisms of action and delivery modes. *Plants (Basel)* 10 (6), 1223. doi: 10.3390/plants10061223

Sun, H. W., Li, W. Q., Burritt, D. J., Tian, H. T., Zhang, H., Liang, X. H., et al. (2022). Strigolactones interact with other phytohormones to modulate plant root growth and development. *Crop J.* 10 (6), 1517–1527. doi: 10.1016/j.cj.2022.07.014

Visentin, I., Pagliarani, C., Deva, E., Caracci, A., Turečková, V., Novák, O., et al. (2020). A novel strigolactone-miR156 module controls stomatal behaviour during drought recovery. *Plant Cell Environ.* 43 (7), 1613–1624. doi: 10.1111/pce.13758

Xu, L. X., Yu, J. J., Han, L. B., and Huang, B. R. (2013). Photosynthetic enzyme activities and gene expression associated with drought tolerance and post-drought recovery in Kentucky bluegrass. *Environ. Exp. Bot.* 89, 28–35. doi: 10.1016/j.envexpbot.2012.12.001

Xu, J. X., Zha, M. R., Li, Y., Ding, Y. F., Chen, L., Ding, C. Q., et al. (2015). The interaction between nitrogen availability and auxin, cytokinin, and strigolactone in the control of shoot branching in rice (Oryza sativa l.). *Plant Cell Rep.* 34 (9), 1647–1662. doi: 10.1007/s00299-015-1815-8

Yamada, Y., and Umehara, M. (2015). Possible roles of strigolactones during leaf senescence. *Plants (Basel)* 4 (3), 664–677. doi: 10.3390/plants4030664

Yang, T., Lian, Y., and Wang, C. (2019). Comparing and contrasting the multiple roles of butenolide plant growth regulators: strigolactones and karrikins in plant development and adaptation to abiotic stresses. *Int. J. Mol. Sci.* 20 (24), 6270. doi: 10.3390/ijms20246270

Yao, R. F., Ming, Z. H., Yan, L. M., Li, S. H., Wang, F., Ma, S., et al. (2016). DWARF14 is a non-canonical hormone receptor for strigolactone. *Nature* 536 (7617), 469–473. doi: 10.1038/nature19073

Zhou, J., Chen, S. Q., Shi, W. J., David-Schwartz, R., Li, S. T., Yang, F. L., et al. (2021). Transcriptome profiling reveals the effects of drought tolerance in giant juncao. *BMC Plant Biol.* 21 (1), 2. doi: 10.1186/s12870-020-02785-7

10.3389/fpls.2023.1186718

Glossary

SL	Strigolactone
Pn	Net photosynthetic rate
Gs	Stomatal conductance
Tr	Transpiration rate
Ci	Intercellular CO2 concentration
WUE	Water use efciency
Fv/Fm	Maximum photochemical efciency
PIabs	Performance index based on absorbed light energy
PEPC	Phosphoenolpyruvate carboxylase
PPDK	Pyruvate phosphate dikinase
NADP-ME	NADP-malic enzyme
PEPCK	Phosphoenolpyruvate carboxykinase
ABA	Abscisic acid
6-BA	6-Benzylaminopurine
MeSA	methyl salicylate
NAA	naphthlcetic acid
JA	jasmonic acid
Len	Root length
SA	surface area
NTips	number of tips
Vol	root volume
AvgD	average diameter
RuBPC	Ribulose bisphosphate carboxylase
PGK	Phosphoglycerate kinase
FBA	Fructose-bisphosphate aldolase
GAPDH	Glyceraldehyde-3-phosphate dehydrogenase A
SBPase	Sedoheptulose-1
7-bisphosphatase	
MCCB	Methylcrotonoyl-CoA carboxylase beta chain
MCCA	Methylcrotonoyl-CoA carboxylase subunit alpha
MFP2	Glyoxysomal fatty acid beta-oxidation multifunctional protein MFP-a
IVD	Isovaleryl-CoA dehydrogenase
ACAT	Acetyl-CoA acetyltransferase
AGT2	Alanine-glyoxylate aminotransferase

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 15 June 2023
DOI 10.3389/fpls.2023.1215084

OPEN ACCESS

EDITED BY
Qiang Guo,
Beijing Academy of Agricultural and Forestry Sciences, China

REVIEWED BY
Guangyan Feng,
Sichuan Agricultural University, China
Tiejun Zhang,
Beijing Forestry University, China
Xiqiang Liu,
Chinese Academy of Forestry, China

*CORRESPONDENCE
Peisheng Mao
maops@cau.edu.cn

†These authors share first authorship

RECEIVED 01 May 2023
ACCEPTED 31 May 2023
PUBLISHED 15 June 2023

CITATION
Sun M, Sun S, Jia Z, Zhang H, Ou C, Ma W, Wang J, Li M and Mao P (2023) Genome-wide analysis and expression profiling of glyoxalase gene families in oat (*Avena sativa*) indicate their responses to abiotic stress during seed germination. *Front. Plant Sci.* 14:1215084. doi: 10.3389/fpls.2023.1215084

Genome-wide analysis and expression profiling of glyoxalase gene families in oat (*Avena sativa*) indicate their responses to abiotic stress during seed germination

Ming Sun†, Shoujiang Sun†, Zhicheng Jia, Han Zhang, Chengming Ou, Wen Ma, Juan Wang, Manli Li and Peisheng Mao*

College of Grassland Science and Technology, China Agricultural University, Beijing, China

Abiotic stresses have deleterious effects on seed germination and seedling establishment, leading to significant crop yield losses. Adverse environmental conditions can cause the accumulation of methylglyoxal (MG) within plant cells, which can negatively impact plant growth and development. The glyoxalase system, which consists of the glutathione (GSH)-dependent enzymes glyoxalase I (GLX1) and glyoxalase II (GLX2), as well as the GSH-independent glyoxalase III (GLX3 or DJ-1), plays a crucial role in detoxifying MG. However, genome-wide analysis of glyoxalase genes has not been performed for one of the agricultural important species, oat (*Avena sativa*). This study identified a total of 26 *AsGLX1* genes, including 8 genes encoding Ni^{2+}-dependent GLX1s and 2 genes encoding Zn^{2+}-dependent GLX1s. Additionally, 14 *AsGLX2* genes were identified, of which 3 genes encoded proteins with both lactamase B and hydroxyacylglutathione hydrolase C-terminal domains and potential catalytic activity, and 15 *AsGLX3* genes encoding proteins containing double DJ-1 domains. The domain architecture of the three gene families strongly correlates with the clades observed in the phylogenetic trees. The *AsGLX1*, *AsGLX2*, and *AsGLX3* genes were evenly distributed in the A, C, and D subgenomes, and gene duplication of *AsGLX1* and *AsGLX3* genes resulted from tandem duplications. Besides the core cis-elements, hormone responsive elements dominated the promoter regions of the glyoxalase genes, and stress responsive elements were also frequently observed. The subcellular localization of glyoxalases was predicted to be primarily in the cytoplasm, chloroplasts, and mitochondria, with a few presents in the nucleus, which is consistent with their tissue-specific expression. The highest expression levels were observed in leaves and seeds, indicating that these genes may play important roles in maintaining leaf function and ensuring seed vigor. Moreover, based on in silico predication and expression pattern analysis, *AsGLX1-7A*, *AsGLX2-5D*, *AsDJ-1-5D*, *AsGLX1-3D2*, and *AsGLX1-2A* were suggested as promising candidate genes for improving stress resistance or seed vigor in oat. Overall, the identification and analysis of the glyoxalase gene

families in this study can provide new strategies for improving oat stress resistance and seed vigor.

KEYWORDS

Avena sativa, glyoxalase, seed germination, abiotic stress, expression profiling

1 Introduction

Reactive dicarbonyl compounds, including methylglyoxal (MG), glyoxal (GO), and 3-deoxyglucosone (3-DG), are the main toxic metabolites spontaneously produced by various metabolic pathways when plants are exposed to stress, and these compounds can hinder normal growth and development (Chinchansure et al., 2015; Ramu et al., 2020). MG, which is the most prevalent reactive dicarbonyl compound, is a by-product of nonenzymatic reactions through glycolysis and the Calvin cycle, and enzymatic pathways through proteins and fatty acid metabolism (Kaur et al., 2014; Takagi et al., 2014). High concentrations of MG can be cytotoxic to cells as they spontaneously form advanced glycation end-products (AGEs) when interacting with nucleic acids, proteins, and lipids (Kaur et al., 2014; Singla-Pareek et al., 2020). However, despite its potential cytotoxicity at high concentrations, MG also serves as a vital signaling molecule involved in various biological processes (Singla-Pareek et al., 2020). Therefore, the maintenance of MG homeostasis is crucial for the normal growth and development of plants.

Plants maintain MG homeostasis through the glyoxalase system and non-glyoxalase system. The former primarily comprises three enzymes, namely, glyoxalase I (GLX1, GLXI or GLYI), also known as lactoylglutathione lyase (EC 4.4.1.5); glyoxalase II (GLX2, GLXII or GLYII), also known as hydroxyacylglutathione hydrolase (EC 3.1.2.6); and glyoxalase III (GLX3, GLYIII or DJ-1)(Li, 2016). GSH spontaneously reacts with MG to form hemithioacetal (HTA), which is then converted to S-D-lactoylglutathione (SLG) by GLX1. GLX2 hydrolyses SLG to produce D-lactate and GSH, which is then recycled back into the system (Thornalley, 1990). The functions of these two GSH-dependent glyoxalases have been extensively explored in plants and reviewed in terms of MG detoxification, cell aging, signal transduction, cell division and differentiation, starch synthesis, pollination, nutrient response, and stress response (Singla-Pareek et al., 2020). Among these, stress response is considered their primary function (Kaur et al., 2014; Hasanuzzaman et al., 2017). The expression and activity of GLX1 and GLX2 vary significantly under various stress conditions, such as hypoxia, salt, drought, heat, cold, and heavy metal (Kaur et al., 2014; Sankaranarayanan et al., 2017). GLX3, a novel glyoxalase enzyme of the DJ-I protein family, directly catalyzes the conversion of MG to D-lactate (Ghosh et al., 2016). Only a few studies have investigated GLX3 in model plants and cereal crops, demonstrating its participation in stress responses (Jana et al., 2021; Kumar et al., 2021; Gambhir et al., 2023). Recent studies have shown that GLX1 and GLX2 are also involved in regulating seed vigor in *Oryza sativa* and *Arabidopsis* (Schmitz et al., 2017; Liu et al., 2022), but their role in seed development, storage, and germination is still poorly understood.

The identification and functional analysis of the glyoxalase gene family members have gained significant attention due to the diverse functions of glyoxalase enzymes. Although *GLX1* and *GLX2* genes have been genome-wide identified in model plants and important crops, such as *Arabidopsis* (Mustafiz et al., 2011), *O. sativa* (Mustafiz et al., 2011), *Sorghum bicolor* (Bhowal et al., 2020), *Glycine max* (Ghosh and Islam, 2016), and *Brassica rapa* (Yan et al., 2018), the genome-wide identification on GLX3 is limited (Li et al., 2019; Jana et al., 2021; Yan et al., 2023). Notably, both *GLX1* and *GLX2* genes in these species have multiple members with diverse subcellular localizations (Li, 2016). GLX1s are classified into two types: Ni^{2+}-dependent and Zn^{2+}-dependent (Schmitz et al., 2018), while GLX2s belong to the beta-lactamase protein family and have a binuclear metal center consisting of Fe^{3+}, Zn^{2+}, and Mn^{2+} (Schilling et al., 2003). GLX3s belong to the DJ-1/PfpI superfamily and do not require metal ions for their optimal activity (Ghosh et al., 2016). The identification and analysis of *GLX1*, *GLX2*, and *GLX3* genes are crucial for comprehending the regulation of plant growth, development, and stress response.

Oat (*A. sativa*) is an important cereal and feed crop that has recently had its genome data released (Rasane et al., 2015; Kamal et al., 2022; Peng et al., 2022), but has not yet been extensively studied in terms of the *GLX1*, *GLX2*, and *GLX3* genes. Therefore, the identification and expression analysis of these genes at the whole-genome level will help to rapidly advance stress-resistant breeding and seed vigor improvement in oat. In our previous research, we discovered a close association between the GSH-dependent glyoxalases and the AsA-GSH cycle with oat seed vigor, where GSH content serves as a potential marker for seed germination percentage (Sun et al., 2022b). We also conducted a genome-wide identification of the glutathione reductase (*GR*) genes involved in the AsA-GSH cycle and analyzed its expression pattern during seed germination under stress conditions (Sun et al., 2022a). The objective of this study is to identify the *GLX1*, *GLX2*, and *GLX3* genes in oat at the whole-genome level, and to conduct chromosome mapping, phylogenetic analysis, synteny analysis, conserved domain analysis, *cis*-regulatory element and subcellular localization prediction, and tissue-specific analysis. In addition, the role of these members in seed germination under stress conditions will be analyzed using qPCR. This comprehensive analysis of the

glyoxalase gene families in oat will significantly contribute to the genetic enhancement of stress resistance and seed vigor in this essential crop.

2 Materials and methods

2.1 Identification and chromosomal mapping of AsGLX1, AsGLX2 and AsGLX3 genes

The HMM (Hidden Markov Model) files for the conserved domains of GLX1, GLX2, and GLX3 proteins, including conserved glyoxalase domain (PF00903), metallo-beta-lactamase domain (PF00753), and DJ-1/PfpI domain (PF01965), were obtained from the Pfam database (Mustafiz et al., 2011; Ghosh et al., 2016; Bhowal et al., 2020). The Simple HMM Search module in TBtools software was used to perform the alignment of HMM files to the oat genome (PepsiCo_OT3098_v2_genome, https://wheat.pw.usda.gov/GG3/) to initially identify oat *GLX1*, *GLX2*, and *GLX3* genes. The protein sequences were then submitted to the Pfam website for individual confirmation of their conserved domains, and ultimately determine the members of the GLX1, GLX2, and GLX3 families.

The TBtools software was also used to visualize the chromosome distribution of oat *GLX1*, *GLX2*, and *GLX3* genes (Chen et al., 2020). The nomenclature of oat *GLX1*, *GLX2*, and *GLX3* genes followed the international wheat gene nomenclature rules (http://wheat.pw.usda.gov/ggpages/wgc/98/intro.htm), which could reflect the chromosomal locations of these genes in the oat subgenomes.

2.2 Phylogenetic tree, gene duplication, and synteny analysis of AsGLX1, AsGLX2, and AsGLX3 members

The identified AsGLX1, AsGLX2, and AsGLX3 members were aligned with glyoxalase members from *Arabidopsis* and *O. sativa* using ClustalX (v2.1) with default settings (Larkin et al., 2007). The protein sequences of *O. sativa* and *Arabidopsis* GLX1s, GLX2s, and GLX3s used for phylogenetic analysis were sourced from previously published genome-wide identification studies (Mustafiz et al., 2011; Ghosh et al., 2016). The phylogenetic tree was constructed using the neighbor-joining method in MEGA 6.0, with 10,000 bootstrap tests and support values expressed as percentages based on 1000 replications (Tamura et al., 2013). Gene duplication events were analyzed using the Multiple Collinearity Scan toolkit (MCScanX) with default parameters (Wang et al., 2012). To illustrate the interspecies syntenic relationships of *GLX1*, *GLX2*, and *GLX3* genes between oat, *Arabidopsis*, and *O. sativa*, a synteny analysis plot was constructed using Dual Synteny Plotter for MC ScanX in TBtools (Chen et al., 2020).

2.3 Physiochemical properties and subcellular localization of AsGLX1s, AsGLX2s, and AsGLX3s

The online program ExPaSy-ProtParam (https://web.expasy.org/protparam/) was utilized to analyze the physical and chemical properties of AsGLX1, AsGLX2, and AsGLX3 proteins, including the amino acid number (AA), molecular weight (MW), and theoretical isoelectric point (pI), instability index (II), negatively charged residues (NCRs), and positively charged residues (PCRs). Protein subcellular localization was predicted using online tools, including WoLF PSORT (https://wolfpsort.hgc.jp/), CELLov.2.5 (http://CELLO.life.nctu.edu.tw/), and Plant-mPLoc (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/).

2.4 Domain architecture of AsGLX1s, AsGLX2s, and AsGLX3s

The TBtools software was used to analyze the protein domains of AsGLX1s, AsGLX2s, and AsGLX3s through HMM files alignment (Chen et al., 2020).

2.5 Identification of *cis*-regulatory elements in the promoter *AsGLX1*, *AsGLX2*, and *AsGLX3* genes

To comprehensively investigate the potential response of the *AsGLX1*, *AsGLX2*, and *AsGLX3* genes to stress, we extracted a putative promoter region of 2 kb upstream of the genes from the oat genome sequence and identified *cis*-regulatory elements using the PlantCARE online tool (https://bioinformatics.psb.ugent.be/webtools/plantcare/html/). It may be helpful to further explore the transcriptional regulation mechanism of the glyoxalase genes and reveal its roles in plant development and stress response.

2.6 Sampling of different tissues in oat

To perform tissue-specific expression profiling of *AsGLX1*, *AsGLX2*, and *AsGLX3* genes, oat (cv Challenger) samples were collected from seeds, roots, leaves, stems, florets, and lemmas. The seed samples were collected during the imbibition phase (0 h, 12 h, and 24 h) and the development phase (8, 15, and 30 days after flowering, DAF). At the flowering stage, young leaves, lemmas, and florets were collected, while old leaves were obtained from plants at 30 DAF. The roots used in the study were from 10-day-old plants. All seeds used in the study had a germination percentage of 100%.

2.7 Stress treatments and sampling during oat seed germination

Seed germination was conducted in a plant growth chamber using plastic petri dishes (11.5 cm × 11.5 cm) containing three

layers of filter paper and 50 seeds under a 16-h dark and 8-h light cycle according to the guidelines of the International Seed Testing Association (ISTA, 2019). Normal seed imbibed in distilled water at 20°C was used as the control (CK), while salt, drought, and MG stress treatments employed solutions of 150 mM NaCl (Xu et al., 2021), 20% PEG6000 (Xie et al., 2021), and 10 mM MG (Hoque et al., 2012), respectively. Cold treatment was conducted by imbibing seeds in distilled water at 10°C. Seeds aged for 30 days were also imbibed in distilled water at 20°C. The seed aging treatment was based on the method described by Xia et al. (Xia et al., 2020). Seed samples were collected after 0 h, 6 h, 12 h, 24 h, 36 h, and 72 h of imbibition, with each treatment consisting of three biological replicates, and 20 seeds collected as one replicate.

2.8 Gene expression analysis by qRT-PCR and statistical analyses

The total RNA was extracted from oat tissue samples using the Quick RNA isolation Kit (Huayueyang Biotech Co., Ltd., China). The first-strand cDNA was synthesized from 1 μg of RNA using the EasyScript® All-in-One First-Strand cDNA Synthesis SuperMix for qPCR Kit (TransGen Biotech, China). The qRT-PCR was performed on a CFX96 Real-Time System using 2×RealStar Fast SYBR qPCR Mix (Genstar, China), with *AsEIF4A* as the reference gene (Yang et al., 2020). The thermal cycle program was as follows: an initial step at 95°C for 3 min, followed by 40 cycles of 95°C for 15 s and 60°C for 30 s. The relative expression level was determined using the $2^{-\Delta\Delta Ct}$ method.

The comparison of the relative expression levels among different tissues was analyzed using ANOVA and a Duncan's test in SPSS Statistics 22 and was visualized using GraphPad Prism version 8.0. The expression heatmaps for *AsGLX1*, *AsGLX2*, and *AsGLX3* genes during seed germination under stress were created using TBtools. The significant expression changes were calculated using Student's *t* test. The details of the primers used in the qRT-PCR assay are listed in Table S1.

3 Result

3.1 Identification and chromosomal mapping of oat glyoxalase genes

Through whole genome alignment in oat, 26 *AsGLX1* genes, 14 *AsGLX2* genes, and 15 *AsGLX3* genes were identified. Chromosomal mapping showed all 26 *AsGLX1* genes were evenly distributed in the A, C, and D subgenomes, with 10 in the A subgenome and 8 in both the C and D subgenomes. *AsGLX1* genes in chromosome set 1 had 9 genes, but only one gene on chromosome set 5. As for 14 *AsGLX2* genes, they were also evenly distributed across the subgenomes, with 5 in both the A and C subgenomes, and 4 in the D subgenome. *AsGLX2* genes were only present once on each chromosome. Regarding *AsGLX3* genes, 5 genes were located in the A subgenome, 4 in the C subgenome, and 6 in the D subgenome. *AsGLX3* genes were not present in chromosome sets 1 and 2, but were highly represented in chromosome sets 3 and 7, with two genes located on each of 3A, 3C, 3D, 7A, and 7D chromosomes. Additionally, some genes of the GLX1 and GLX3 families in oat are adjacent on chromosomes (Figure 1).

3.2 Phylogenetic analysis of AsGLX1s, AsGLX2s, and AsGLX3s

To elucidate the phylogenetic relationships among members of the oat GLX families, amino acid-based phylogenetic trees were constructed using the sequences of GLX1s, GLX2s, and GLX3s from oat, *Arabidopsis*, and *O. sativa* (Figure 2). Notably, the number of GLX1 members exceeded that of GLX2 and GLX3 in *A. sativa*, *O. sativa*, and *Arabidopsis*. The 26 AsGLX1s were categorized into two clades. Clade I consisted of 11 AsGLX1s, 4 AtGLX1s, and 5 OsGLX1s, while Clade II contained 16 AsGLX1s, 7 AtGLX1s, and 6 OsGLX1s (Figure 2A). The 14 AsGLX2s were classified into four clades, with Clade I and Clade II exclusively containing AsGLX2s, with 2 and 6 members, respectively. Clade III consisted of 1 AtGLX2, 1 OsGLX2s, and 3 AsGLX2s, while Clade IV comprised 3 AsGLX2s, 4 AtGLX2s, and 2 OsGLX2s (Figure 2B). Additionally, the 15 AsGLX3 (AsDJ-1) members were also categorized into four clades, with Clade I containing 3 AtDJ-1s, 2 OsDJ-1s, and 3 AsDJ-1s. Clade II included 1 OsDJ-1 and 3 AsDJ-1s, and Clade III consisted of 1 AtDJ-1s, 1 OsDJ-1s, and 3 AsDJ-1s. Clade IV has the largest number of AtDJ-1 members, including 2 AtDJ-1s, 2 OsDJ-1s, and 6 AsDJ-1s (Figure 2C).

3.3 Gene duplication and synteny analysis of *GLX1*, *GLX2*, and *GLX3* genes

Gene duplication analysis revealed that certain gene pairs of the *GLX1* and *GLX3* gene families, namely *AsGLX1-3A1* and *AsGLX1-3A2*, *AsDJ-1-3A1* and *AsDJ-1-3A2*, and *AsDJ-1-3D1* and *AsDJ-1-3D2*, were the result of tandem duplications, and they were adjacent on the chromosomes. However, no evidence of segmental duplications was found among the genes of the GLX1, GLX2, and GLX3 families (Figure 1).

The synteny analysis between oat and rice, as well as *Arabidopsis*, revealed that there is no collinearity between oat and *Arabidopsis* in the glyoxalase gene families. However, collinearity was observed between oat and rice in the *GLX1*, *GLX2*, and *GLX3* gene families. Specifically, 18 *AsGLX1* genes showed collinearity with 8 *O. sativa* genes, 9 *AsGLX2* genes showed collinearity with 4 *O. sativa* genes, and 8 *AsDJ-1* genes showed collinearity with 3 *O. sativa* genes (Figures 3A–C). Among these collinear gene pairs, *OsGLYI6* has the highest number of homologous genes in oat, with six homologous genes (*AsGLX1-1A3*, *AsGLX1-1C3*, *AsGLX1-1D3*, *AsGLX1-3A1*, *AsGLX1-3C*, and *AsGLX1-3D1*). The second highest is *OsGLYI1*, which has five oat homologous genes (*AsGLX1-1A3*, *AsGLX1-1D3*, *AsGLX1-3A1*, *AsGLX1-3C*, and *AsGLX1-3D1*). *OsGLYII3* has three oat homologous genes including *AsGLX2-5A*, *AsGLX2-5C*, and *AsGLX2-5D*. Similarly, *OsDJ-1A* has three

10.3389/fpls.2023.1215084

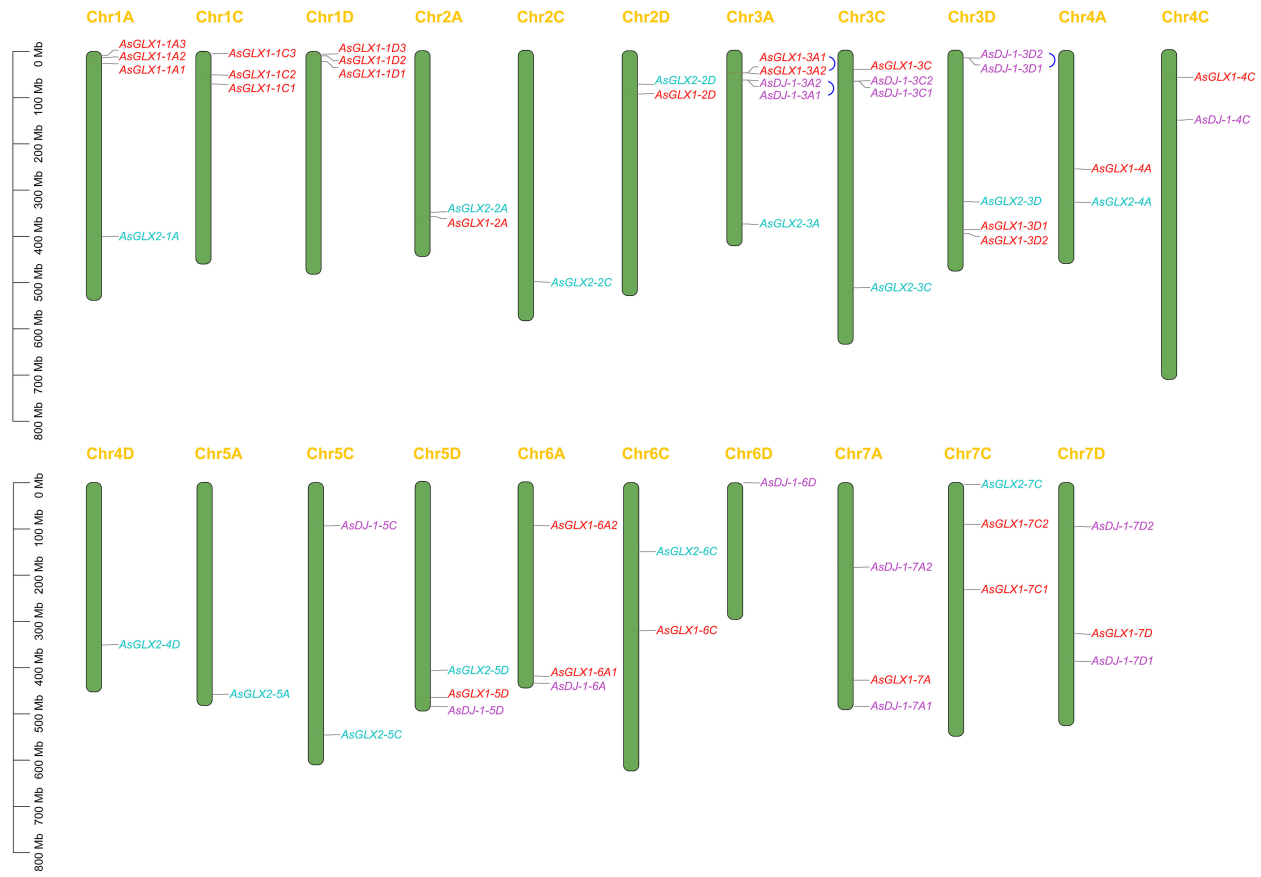


FIGURE 1
The chromosomal mapping of *AsGLX1*, *AsGLX2*, and *AsGLX3* genes across oat chromosomes. The blue connecting lines represent tandem duplications. The chromosome number is indicated at the top of each chromosome and the scale is shown in megabases (Mb).

homologous genes (*AsDJ-1-3A2*, *AsDJ-1-3C2*, and *AsDJ-1-3D2*), and *OsDJ-1E* also has three oat homologous genes (*AsDJ-1-4C*, *AsDJ-1-7A1*, and *AsDJ-1-7D1*).

3.4 Physiochemical properties and subcellular localization of AsGLX1s, AsGLX2s, and AsGLX3s

The online subcellular localization prediction of *AsGLX1*, *AsGLX2*, and *AsGLX3* genes showed that these genes are primarily expressed in the cytoplasm, chloroplasts, and mitochondria (Table S2). Based on the predicted results of two or three online servers, 13 genes of the AsGLX1 family were found to be expressed in the cytoplasm, 17 in chloroplasts, and 4 in mitochondria. While, 5 genes of the AsGLX2 family were found to be expressed in the cytoplasm, 8 in chloroplasts, and 7 in mitochondria. Furthermore, 5 genes of the AsGLX3 family were found to be expressed in the cytoplasm, 13 in chloroplasts, and 3 in mitochondria. In addition, *AsGLX1-2A* was predicted by all three online servers to be expressed in the nucleus. Other members, including 6 *AsGLX1s*, 6 *AsGLX2s*, and 4 *AsGLX3s*, predicted by a single program to potentially be expressed in the nucleus, suggest that these members are important for maintaining nuclear stability.

Through the calculation of amino acid length, molecular weight, theoretical isoelectric points, instability index, and charged residues of AsGLX1, AsGLX2, and AsGLX3 proteins, it was found that among the 26 AsGLX1 members, the molecular weight ranged from 15.219 to 94.136 kDa, the amino acid length ranged from 140 AA to 834 AA, and the pI ranged from 4.81 to 8.98. Among them, 12 proteins had an instability index of less than 40, and 20 proteins were rich in negatively charged residues. For the 14 members of AsGLX2, the molecular weight ranged from 32.495 to 143.432 kDa, the amino acid length ranged from 297 AA to 1300 AA, and the pI ranged from 5.87 to 9.12. Among them, 4 proteins had an instability index of less than 40, and 11 proteins were rich in negatively charged residues. As for the 15 members of AsGLX3, the molecular weight ranged from 41.235 to 76.399 kDa, the amino acid length ranged from 395 AA to 723 AA, and the pI ranged from 4.90 to 9.26. Among them, 5 proteins had an instability index of less than 40, and 11 proteins were rich in negatively charged residues (Table S2).

10.3389/fpls.2023.1215084

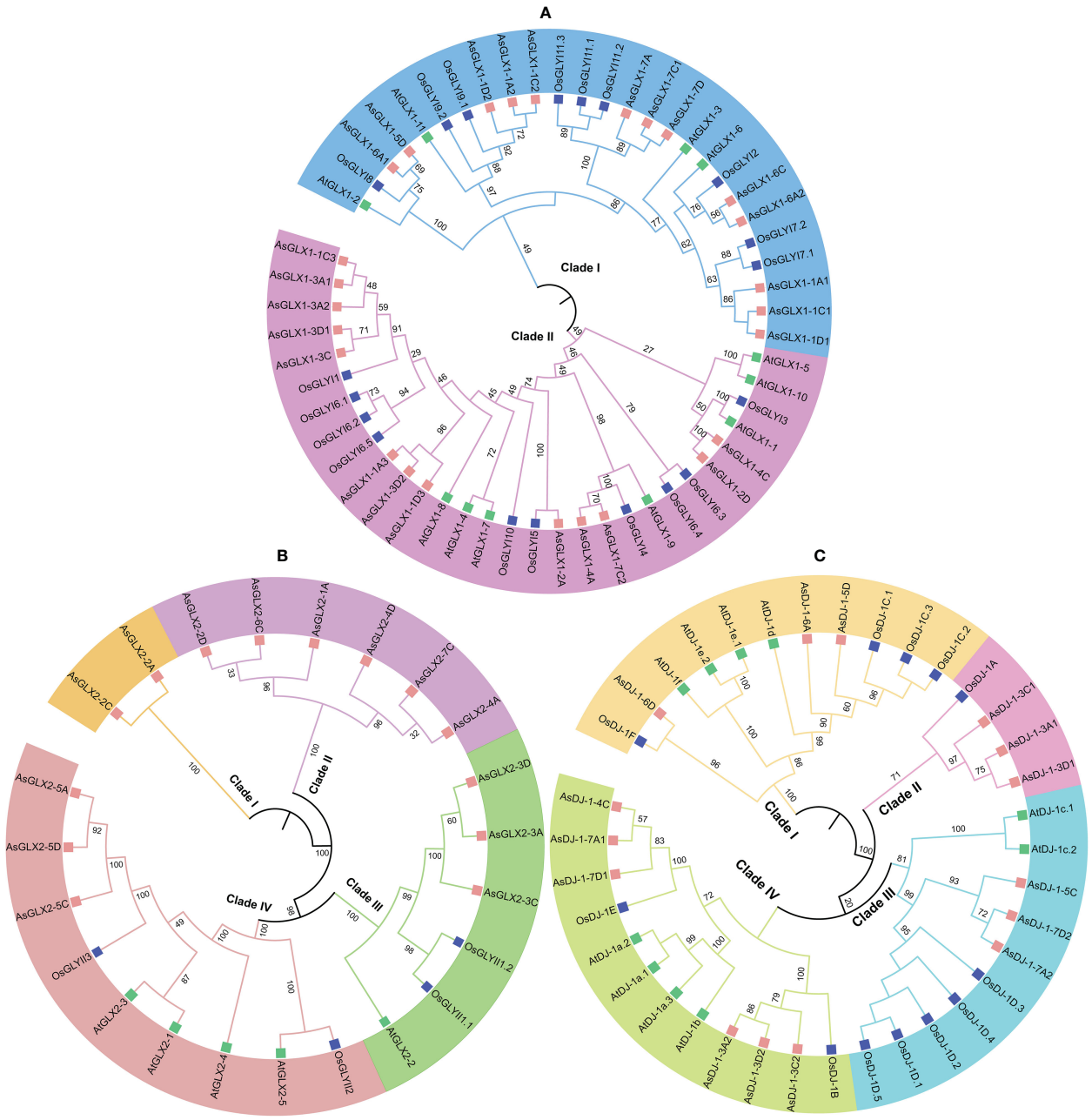


FIGURE 2
Phylogenetic relationships of GLX1s **(A)**, GLX2s **(B)** and GLX3s **(C)**.

3.5 Domain analysis of AsGLX1s, AsGLX2s, and AsGLX3s

Conserved domain analysis of glyoxalase families showed that all AsGLX1s possess conserved glyoxalase domain, AsGLX2s possess conserved metallo-β-lactamase domain, and AsGLX3s possess two conserved DJ-1/PfpI domains (Figure 4). Among AsGLX1 members, AsGLX1-1A1, AsGLX1-1C1, AsGLX1-1D1, AsGLX1-6C, AsGLX1-6A2, AsGLX1-7A, AsGLX1-7C1, and AsGLX1-7D possess two conserved glyoxalase domains, with the first domain consisting of 120 aa and the second domain consisting of 115 aa in AsGLX1-1A1, AsGLX1-1C1, AsGLX1-1D1, AsGLX1-6C, and AsGLX1-6A2, while the first domain is 121 aa and the second domain is 120 aa in AsGLX1-7A, AsGLX1-7C1, and AsGLX1-7D. These eight members, along with AtGLYI-3, AtGLYI-6, OsGLYI2, OsGLYI7, and OsGLYI11, belong to Clade I and are Ni^{2+}-dependent GLX1s. AsGLX1-5D and AsGLX1-6A1 contain only one glyoxalase domain, with a length of 141 aa, and belong to the same branch in Clade I as AtGLYI-2 and OsGLYI8, indicating that they are Zn^{2+}-dependent GLX1s. AsGLX1-1A2, AsGLX1-1D2, and AsGLX1-1C2, which have a domain length of 115 aa, belong to Clade I as well but are not Ni^{2+}-dependent GLX1s. The 13 AsGLX1s with functional domains ranging from 108 to 295 aa in Clade II may be GLX1-like proteins. In addition, AsGLX1-1C3 has one reverse transcriptase domain (RVT_1), and AsGLX1-2A has one

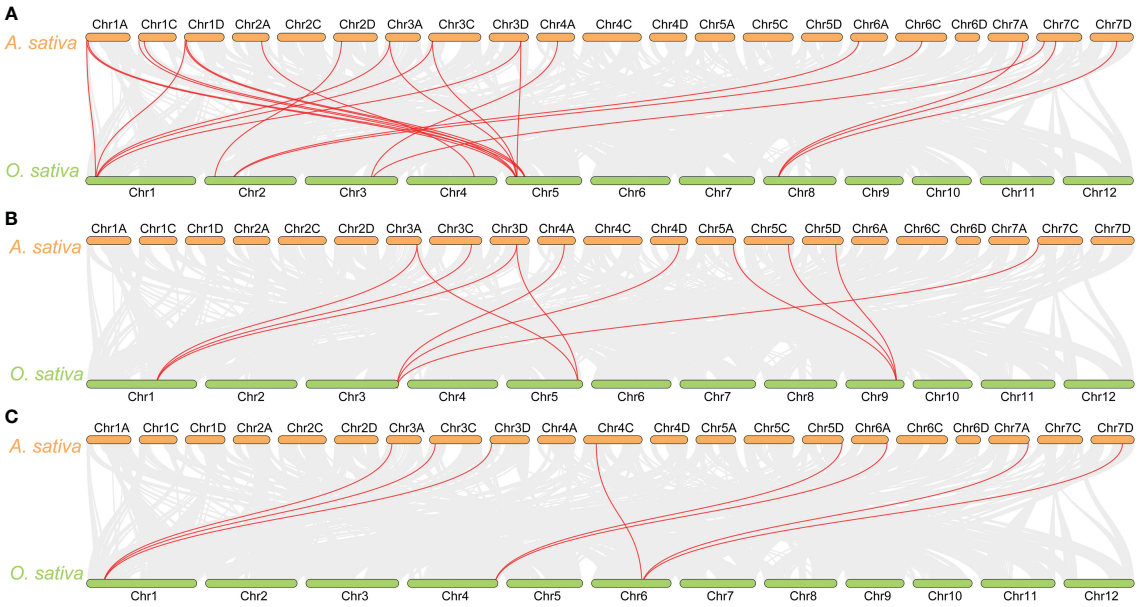


FIGURE 3
Synteny analysis of *GLX1* **(A)**, *GLX2* **(B)**, and *GLX3* **(C)** genes between *A. sativa* and *O. sativa*. The gray lines show colinear blocks in the genomes of *A. sativa* and *O. sativa*, while the red line highlights the colinear *GLX1*, *GLX2*, and *GLX3* gene pairs.

NADH dehydrogenase complex I subunit M domain (NdhM), which is closely related to its predicted chloroplast localization (Figures 2A, 4A).

AsGLX2-5A, AsGLX2-5C, and AsGLX2-5D belong to Clade IV in the GLX2 family, which are characterized by the presence of both a lactamase B domain (148 aa) and a HAGH_C domain (hydroxyacylglutathione hydrolase C-terminal) with 84 aa. These three members share a conserved metal ion-binding site (THHHYDH). Clade III of the GLX2 family comprises members that contain the lactamase B domain with169 aa, while the lactamase B domain of 6 AsGLX2s in Clade II is 154 aa. AsGLX2s of Clade I, including AsGLX2-2C and AsGLX2-2A, have lactamase B domains that are 156 aa and 158 aa in length, respectively. Addition, the active site of AsGLX2-5A, AsGLX2-5C, AsGLX2-5D, AsGLX2-3A, AsGLX2-3C, AsGLX2-3D, AsGLX2-2C, and AsGLX2-2A all feature the GHT residue, which is essential for catalytic activity. In addition to the lactamase B domain and the HAGH_C domain, some AsGLX2s also contain other domains such as the Pre-mRNA 3'-end-processing endonuclease polyadenylation factor C-term domain (CPSF73-100_C), Beta-Casp domain, Zn-dependent metallo-hydrolase RNA specificity domain (RMMBL), GAG-pre-integrase domain, and 4Fe-4S single cluster domain of Ferredoxin I (Fer4_13) (Figures 2B, 4B).

Although all AsDJ-1s have two DJ-1/PfpI domains, there are differences in the amino acid length that connects the two domains among these members. The number of amino acids between the two domains in Clade I members of the AsDJ-1 family is shorter, ranging from 11 to 14 aa, while the number of amino acids between the two domains in members of the other three Clades is longer, ranging from 39 to 41 aa, except for AsDJ-1-3C2, which has 15 aa. Additionally, AsDJ-1-3A2 and AsDJ-1-3D2 each have one Helix-loop-helix DNA binding domain (HLH), AsDJ-1-6A has one F-box associated domain (FBA_3), and AsDJ-1-5D has one F-box domain (F-box) (Figures 2C, 4C).

3.6 *Cis*-regulatory elements in the promoter region of glyoxalase genes

The analysis of the upstream 2 kb promoter regions of the *AsGLX1*, *AsGLX2*, and *AsGLX3* genes showed that the basic eukaryotic promoter elements CAAT-box and TATA-box are widely distributed in the promoter regions of these genes. Other *cis*-regulatory elements mainly related to hormone responsiveness, anaerobic induction, defense and stress responsiveness, light responsiveness, endosperm expression, MYB binding sites, meristem expression, and seed-specific regulation (Figure 5). The hormone responsive elements mainly included ABA, GA, MeJA, SA, and IAA responsiveness. The stress responsiveness was mainly related to low temperature, anaerobic induction, and defense and stress responsiveness. Apart from core elements, hormone responsive elements are the most abundant in the *AsGLX1*, *AsGLX2*, and *AsGLX3* promoters, accounting for 47%, 46%, and 53% of all elements, respectively (Table S3). Light responsive elements are also present at high proportions in the *AsGLX1*, *AsGLX2*, and *AsGLX3* gene promoters, accounting for 32%, 24%, and 23%, respectively. Among hormone responsive elements, ABA responsive elements are the most abundant in the *AsGLX1* gene promoters, accounting for about 50%, while MeJA responsive elements are the most abundant in the *AsGLX2* and *AsGLX3* gene promoters, accounting for 46% and 55%, respectively. Elements associated with adverse conditions are present in descending order of number in *AsGLX1*, *AsGLX2*, and *AsGLX3* gene promoters, namely anaerobic induction, low-temperature responsiveness, and

10.3389/fpls.2023.1215084

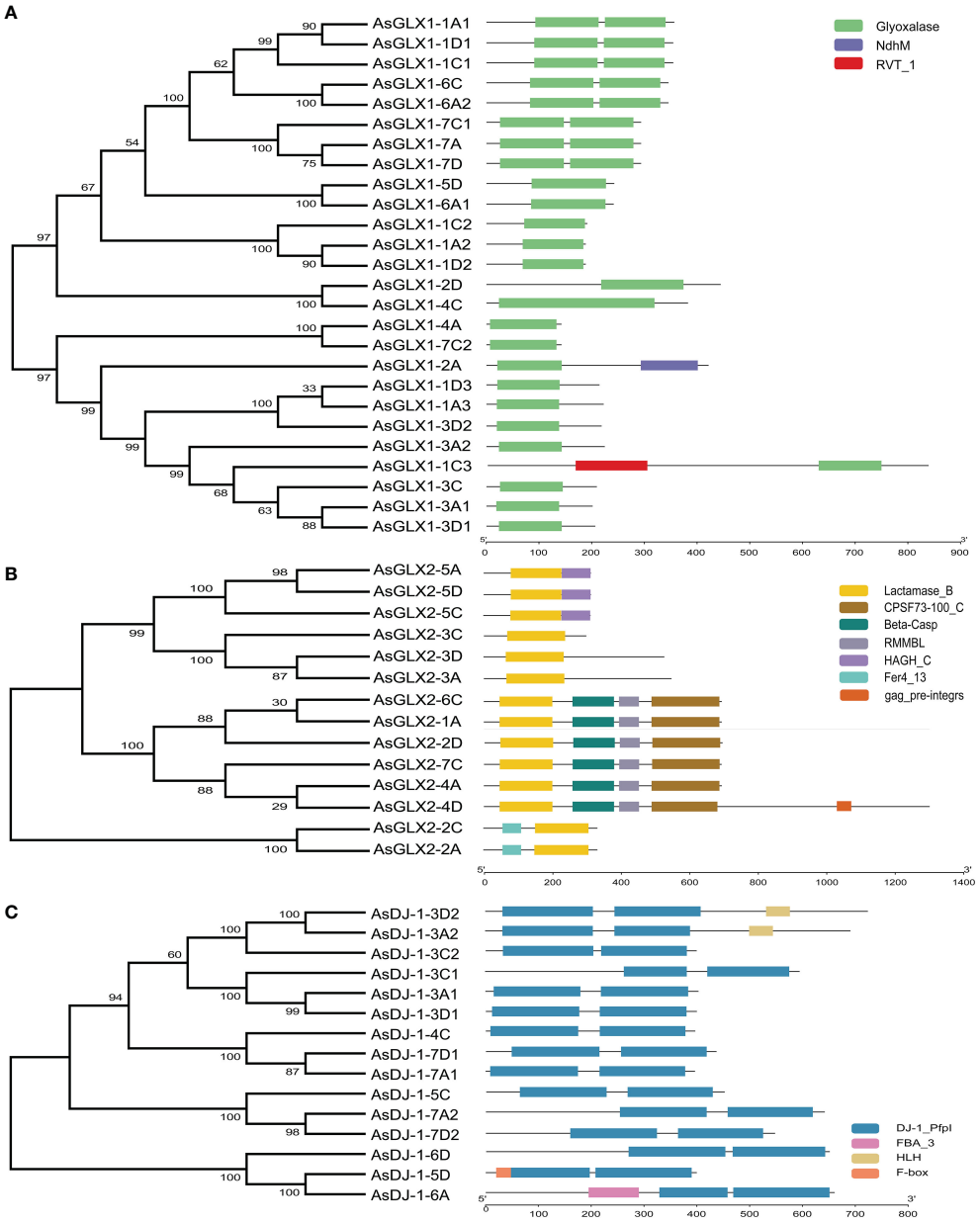


FIGURE 4
Schematic representation of domain architecture of GLX1s **(A)**, GLX2s **(B)** and GLX3s **(C)** from oat.

defense and stress responsiveness (Table S3). Each of the three gene families contains three genes with seed-specific regulatory elements in their promoter regions, namely *AsGLX1-1A3*, *AsGLX1-3A1*, *AsGLX1-4C*, *AsGLX2-2A*, *AsGLX2-3A*, *AsGLX2-4D*, *AsDJ-1-3A1*, *AsDJ-1-3D1*, and *AsDJ-1-3D2*. In addition, MYB binding site elements are also widely distributed in the promoter of genes in the three gene families, particularly in *AsGLX2*, where each gene member's promoter region contains MYB binding site elements (Figure 5).

3.7 Tissue-specific expression analysis of *AsGLX1*, *AsGLX2*, and *AsGLX3* genes

Three members from each of the *AsGLX1*, *AsGLX2*, and *AsGLX3* gene families were randomly selected in different clades with distinct protein domain compositions and *cis*-element compositions to investigate their tissue-specific expression. *AsGLX1-1A1* and *AsGLX2-2D* showed high expression levels in leaves, particularly in old leaves, and were predicted to be localized

10.3389/fpls.2023.1215084

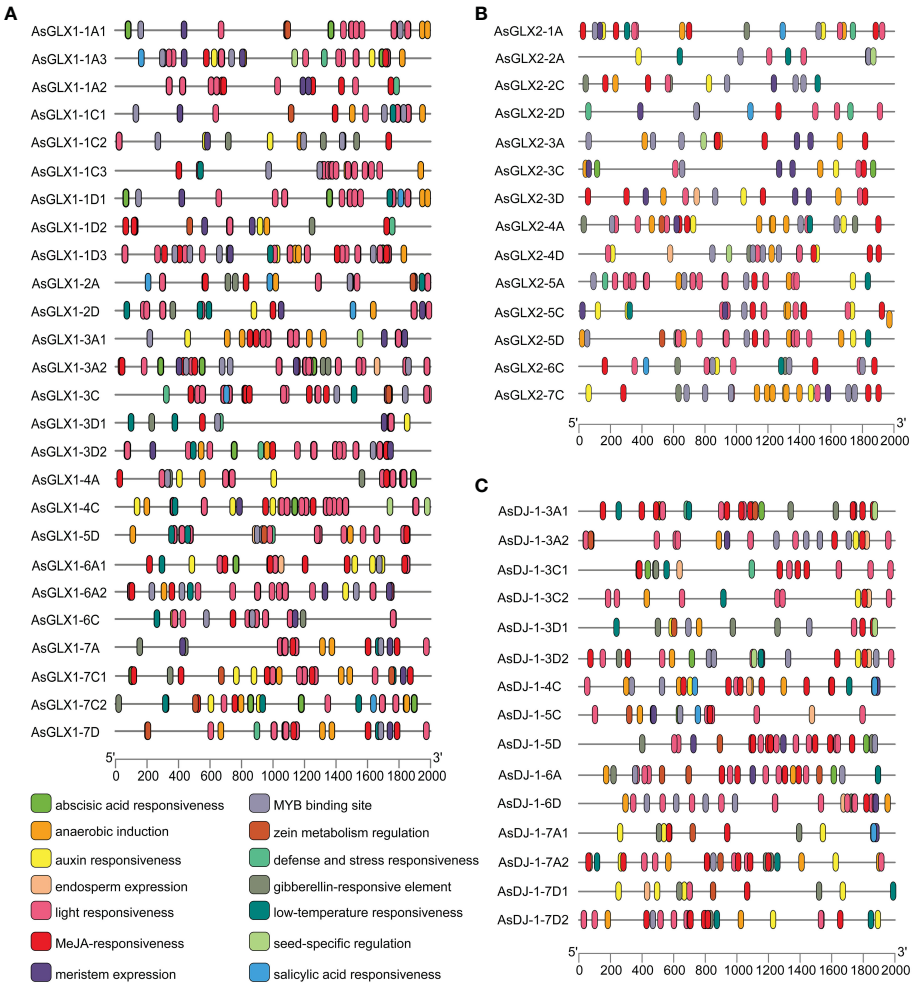


FIGURE 5
Cis-regulatory elements in the promoters of *AsGLX1* **(A)**, *AsGLX2* **(B)**, and *AsGLX3* **(C)** genes.

in chloroplasts, indicating their importance in leaf function and development (Figures 6A, D, Table S2). *AsGLX1-7A* and *AsGLX1-3D2* exhibited the highest expression levels in dry seeds, with increasing expression levels during seed development, indicating their potential roles in seed development and dehydration (Figures 6B, C). *AsGLX2-3C* was highly expressed in germinated seeds and roots, but had low expression levels in leaves, which is consistent with its predicted mitochondrial localization (Figure 6E, Table S2). *AsDJ-1-3D2*, *AsDJ-1-4C*, and *AsDJ-1-5D* showed high expression levels in leaves, glumes, dry seeds, and developing seeds, with expression levels gradually increasing during seed development (Figures 6G–I). Overall, *AsGLX1*, *AsGLX2*, and *AsGLX3* genes exhibited relatively high expression levels in leaves and seeds, suggesting their potential roles in maintaining leaf function and seed vigor.

3.8 Expression analysis of *AsGLX1* genes during seed germination under stresses

The expression patterns of *AsGLX1-1A1* and *AsGLX1-7A* during seed germination showed a similar trend of first decreasing and then increasing, while *AsGLX1-3D2* exhibited a distinct trend of gradual decrease during germination. After aging treatment, the expression levels of *AsGLX1-1A1* and *AsGLX1-7A* in seeds significantly decreased, while the expression level of *AsGLX1-3D2* showed no significant changes (Figure 7, Supplementary Table 4).

Compared to the control, *AsGLX1-1A1* showed significant downregulation during 0-36 h imbibition of aged seed, and it showed relatively small changes during the early imbibition stage (0-12 h) under cold, PEG, salt, and MG treatments. However, after 24 h of treatment, *AsGLX1-1A1* expression was significantly

10.3389/fpls.2023.1215084

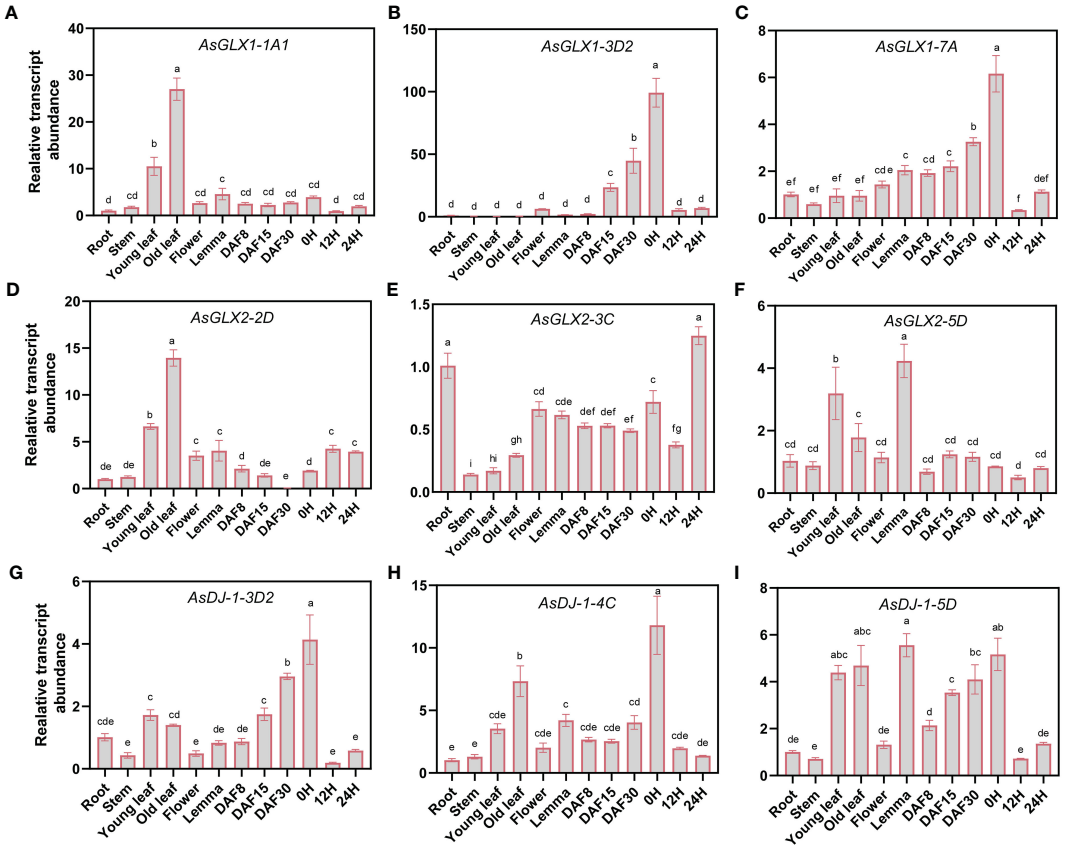


FIGURE 6
Expression profiling of *AsGLX1, AsGLX2*, and *AsGLX3* genes in different tissues of oat by qPCR. The relative expression was calculated using root as reference. Seeds imbibed for 0 h, 12 h, and 24 h in germination phases were marked as 0 H, 12 H, and 24 H, respectively. Developing seeds at 8, 15, and 30 days after flowering were marked as DAF8, DAF15, and DAF30, respectively. The lower-case letters **(A–I)** represent statistical significance among the samples and the vertical bars represent the ± SEM for three replicates. The mean values sharing different letters, obtained from Duncan test, are different significantly at $p < 0.05$ level.

induced, particularly under salt, cold, and MG treatments at 72 h of imbibition (Figure 7A). *AsGLX1-3D2* exhibited relatively small changes under cold treatment, with no significant difference between the control and treatments at 24 h and 72 h. *AsGLX1-3D2* was upregulated at most imbibition time points during PEG and MG treatments, and in aged seeds, indicating its detoxification role during seed germination under PEG and MG stress and after aging treatment (Figure 7B). *AsGLX1-7A* was significantly induced during imbibition for 0-72 h under PEG treatment, with the most significant upregulation occurring at 72 h. However, the expression level of *AsGLX1-7A* in aged dry seeds significantly decreased, being significantly higher than the control during early imbibition (6-12 h), but significantly lower than or not significantly different from the control at 24-72 h of imbibition, suggesting a potential role during the early stages of germination in aged seeds. In addition, *AsGLX1-7A* was significantly induced under cold, salt, and MG treatments at all imbibition time points except for 24 h under cold and salt treatments or 36 h under MG treatment (Figure 7C). Overall, *AsGLX1* genes exhibit specificity in response to different stresses, and *AsGLX1-3D2* and *AsGLX1-7A* may play important detoxification roles during seed germination under stress conditions.

3.9 Expression analysis of *AsGLX2* genes during seed germination under stresses

The response of the three *AsGLX2* genes to stress treatments is relatively smaller compared to the tested *AsGLX1* genes. Following aging treatment, both *AsGLX2-3C* and *AsGLX2-5D* exhibited significant downregulation, while *AsGLX2-2D2* showed no significant difference compared to the CK (Figure 8, Supplementary Table 4).

AsGLX2-2D2 was significantly up-regulated during imbibition under cold treatment, except for 12 h, and showed marked response at 24 h and 36 h. In aged seeds, *AsGLX2-2D2* showed almost no significant response. *AsGLX2-2D2* was up-regulated at 24 h and 36 h of imbibition under salt stress, while it was up-regulated at 72 h under MG treatment (Figure 8A). *AsGLX2-3C* showed a weak response to different treatments in the early stage of imbibition, and was only significantly up-regulated under PEG treatment for 12 h and significantly down-regulated at 6 h of imbibition in aged seeds. During 24-36 h of imbibition, *AsGLX2-3C* was up-regulated under various treatments, except for significant down-regulation at 36 h of imbibition in aged seeds. At 72 h of imbibition, the expression of *AsGLX2-3C* showed no significant difference

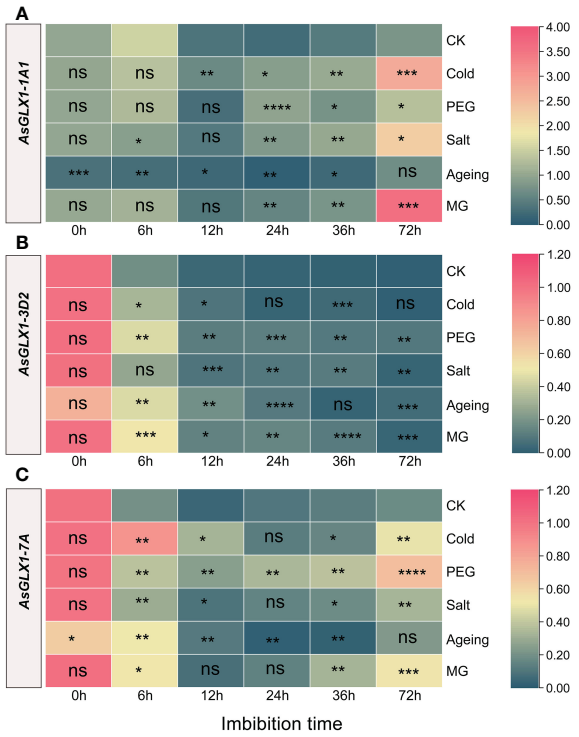


FIGURE 7
Expression profiling of *AsGLX1-1A1* **(A)**, *AsGLX1-3D2* **(B)** and *AsGLX1-7A* **(C)** during seed germination of oat under stress by qPCR. The relative expression was calculated using dry seed (0 h) of CK as reference. The significant change in comparison to CK has been calculated using Student's *t* test. *indicates a significant difference at $p < 0.05$; **indicates a significant difference at $p < 0.01$; ***indicates a significant difference at $p < 0.001$; ****indicates a significant difference at $p < 0.0001$; ns represents not significant.

compared to the control under various treatments (Figure 8B). *AsGLX2-5D* also showed a strong response under cold treatment, and was up-regulated at all-time points except for 36 h of imbibition. Under PEG and salt treatments, *AsGLX2-5D* showed no significant difference in expression level compared to the control at most time points, except for significant down-regulation at 72 h under PEG treatment. *AsGLX2-5D* was significantly up-regulated more than 2-fold at 12 h of imbibition in aged seeds, and its expression level was significantly higher than the control at 6 h of imbibition under MG treatment (Figure 8C). Overall, *AsGLX2* genes showed a more significant response to cold treatment, indicating their important role in maintaining seed vigor or promoting seed germination under cold stress.

3.10 Expression analysis of *AsDJ-1* genes during seed germination under stresses

The *AsDJ-1* genes exhibit a diverse response to various stresses during seed imbibition, with *AsDJ-1-5D* displaying a more pronounced response to stress, whereas the responses of *AsDJ-1-4C* and *AsDJ-1-3D2* are comparatively minor. Following aging treatment, all three tested *AsDJ-1* genes show a substantial downregulation (Figure 9, Supplementary Table 4).

AsDJ-1-3D2 showed the most significant response under MG treatment, with significant upregulation compared to the control at 6-36 h, and the most significant response occurred at 6 h. Under cold treatment, *AsDJ-1-3D2* was significantly upregulated at 12 h and significantly downregulated at 72 h. It was significantly upregulated at 6-24 h under PEG treatment, and showed a significant response at 12 h and 72 h of imbibition under salt treatment. In aged seeds, it was significantly upregulated during the early imbibition stage (6-12 h) (Figure 9A). *AsDJ-1-4C* did not show any significant response under salt treatment. Under cold and PEG treatments, it showed a significant response during the early imbibition stage, and cold treatment significantly induced its expression at 6 h of imbibition. In aged seeds, the expression level of *AsDJ-1-4C* was significantly upregulated during the early imbibition stage (6-12 h), and then showed a significant decrease. Moreover, under MG treatment, its expression level was significantly upregulated at 6 h, 36 h, and 72 h of imbibition (Figure 9B). *AsDJ-1-5D* was significantly upregulated during the imbibition under PEG treatment, especially in the early imbibition stage (6-12 h). There was no significant response under cold

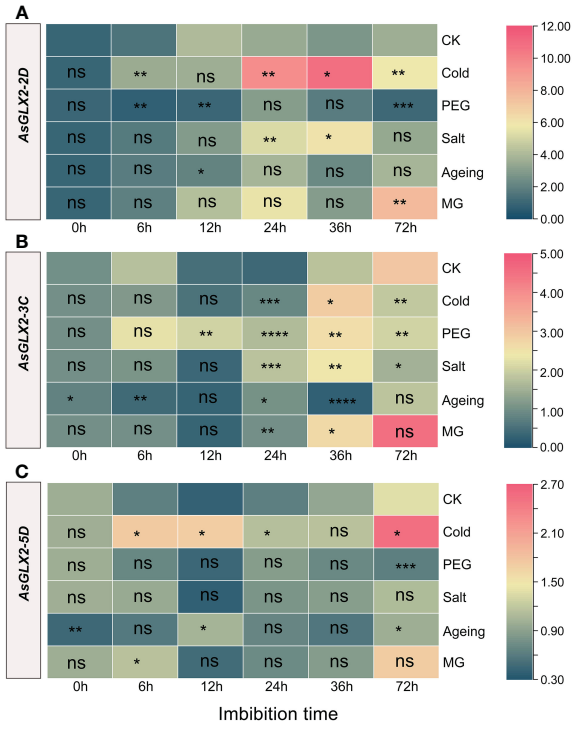


FIGURE 8
Expression profiling of *AsGLX2-2D* **(A)**, *AsGLX2-3C* **(B)**, *AsGLX2-5D* **(C)** during seed germination of oat under stress by qPCR. The relative expression was calculated using dry seed (0 h) of CK as reference. The significant change in comparison to CK has been calculated using Student's *t* test. *indicates a significant difference at $p < 0.05$; **indicates a significant difference at $p < 0.01$; *** indicates a significant difference at $p < 0.001$; ****indicates a significant difference at $p < 0.0001$; ns represents not significant.

treatment at 24 h and under salt treatment at 72 h. At other time points, both cold and salt treatments significantly induced the expression of *AsDJ-1-5D*. During the imbibition of aged seeds, *AsDJ-1-5D* was significantly upregulated during 6-24 h, with the most significant response observed at 6 h, with an increase of more than 5 folds. Under MG treatment, its expression level was only significantly upregulated at 6 h (Figure 9C). Overall, *AsDJ-1-5D* has a relatively strong role in detoxification during the early germination stage of aged seeds and under PEG treatment.

4 Discussion

The glyoxalases participate in various biological processes in plants (Singla-Pareek et al., 2020), such as stress response (Kaur et al., 2014), seed germination (Schmitz et al., 2017), plant senescence (Singla-Pareek et al., 2009), nutrient regulation (Borysiuk et al., 2022), signal transduction (Sankaranarayanan et al., 2017), starch synthesis (You et al., 2019), and pollen development (Sankaranarayanan et al., 2015). Their primary role in plants is to detoxify MG, which is spontaneously produced in plants and significantly accumulates under abiotic stresses such as salinity, drought, heavy metals, and low temperature, thus hindering plant growth and development (Kaur et al., 2014; Hasanuzzaman et al., 2017; Sankaranarayanan et al., 2017). Higher plants often contain multiple members of the *GLX1*, *GLX2*, and *GLX3* gene families, and different members exhibit variations in their subcellular localization, expression patterns, and functional roles (Ghosh and Islam, 2016; Schmitz et al., 2017; Singla-Pareek et al., 2020). Therefore, a comprehensive genome-wide identification of glyoxalase gene families and understanding their chromosome distribution, evolutionary relationships, conserved domains, *cis*-regulatory elements, and gene expression patterns is crucial for exploring the functional diversity of glyoxalase genes and improving plant stress resistance.

The GSH-dependent pathway is the main pathway for clearing MG by GLX1 and GLX2 enzymes, which has led previous studies on plant glyoxalases to focus mainly on these two enzymes. However, the GSH-independent GLX3 has only recently received attention. Genome-wide identification and analysis of GLX1 and GLX2 families have been completed in in various plants. In *Arabidopsis*, a total of 11 *GLX1* genes and 5 *GLX2* genes were identified, while 11 *GLX1* genes and 3 *GLX2* genes were found in *O. sativa* (Mustafiz et al., 2011). In *G. max*, 24 *GLX1* genes and 12 *GLX2* genes were identified (Ghosh and Islam, 2016). In *S. bicolor*, 15 *GLX1* genes and 6 *GLX2* genes were found (Bhowal et al., 2020). However, the extensive whole-genome identification of glyoxalase families has ignored *GLX3* genes, and only a few plants, such as *Medicago truncatula* and *V. vinifera*, have been systematically identified for the GLX1, GLX2, and GLX3 families (Ghosh, 2017; Li et al., 2019). In addition, recent studies have also individually identified GLX3s in some plant species. For example, 217 GLX3s were obtained by using AtDJ-1d (AT3g02720) to compare the Swiss-Prot database, including 8 *Oryza* species, 2 *Triticum* species, *Hordeum vulgare*, *Zea*

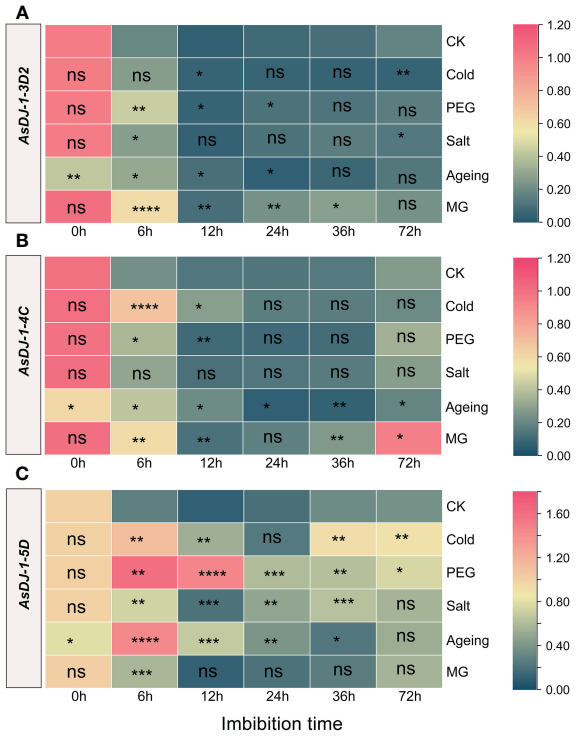


FIGURE 9
Expression profiling of *AsDJ-1-3D2* **(A)**, *AsDJ-1-4C* **(B)**, and *AsDJ-1-5D* **(C)** during seed germination of oat under stress by qPCR. The relative expression was calculated using dry seed (0 h) of CK as reference. The significant change in comparison to CK has been calculated using Student's *t* test. *indicates a significant difference at $p < 0.05$; **indicates a significant difference at $p < 0.01$; ***indicates a significant difference at $p < 0.001$; ****indicates a significant difference at $p < 0.0001$; ns represents not significant.

mays, *S. bicolor*, *Setaria italica*, and *Brachypodium distachyon*. Among them, 12 GLX3s were identified in *O. sativa*, which were encoded by 6 genes and strongly induced by MG (Ghosh et al., 2016). To gain insight into the evolutionary patterns of GLX3, the evolution of GLX3s across prokaryotes and eukaryotes were studied, and 183 GLX3s belonging to 69 species was used for evolutionary analysis in plants (Kumar et al., 2021). In summary, the study of plant GLX3 have become a focus and hotspot, and systematic analysis of GLX1, GLX2, and GLX3 members will play an important role in plant improvement.

A total of 26 *AsGLX1* genes, 14 *AsGLX2* genes, and 15 *AsGLX3* genes were identified in oat. The number of *AsGLX1* and *AsGLX2* genes identified in *Arabidopsis*, *O. sativa*, *G. max*, *S. bicolor*, and *B. rapa* was lower than that in oat, which may be related to genome duplication events during evolution (Ghosh and Islam, 2016; Yan et al., 2018; Bhowal et al., 2020; Peng et al., 2022). In the secondary branches of the evolutionary tree of the GLX1, GLX2, and GLX3, there were two or three oat homologous genes, originating from two or three oat subgenomes (A, C, and D), corresponding to one *Arabidopsis* glyoxalase gene. Oat has diploid and tetraploid ancestors, and genome duplication may be the main reason for the high number of homologous genes in oat (Peng et al., 2022). Some clades contained only GLX members from *O. sativa* and oat, such as OsGLYI11, AsGLX1-7A, AsGLX1-C1, and AsGLX1-D in GLX1, as well as OsDJ-1A, AsDJ-1-3A1, AsDJ-1-3C1, and AsDJ-1-3D1 in GLX3, indicating that the three families evolved asynchronously in the *O. sativa*, oat, and *Arabidopsis* genomes. Furthermore, the gene duplication events analysis of *AsGLX1*, *AsGLX2*, and *AsGLX3* genes showed that only *AsGLX1* and *AsGLX3* gene families had tandem duplications, while none of the three families had segmental duplications. In contrast, in *O. sativa*, *Arabidopsis*, *B. napus* and *G. max*, gene duplications of *GLX1* and *GLX2* genes were caused by the segmental duplication, rather than tandem duplication (Mustafiz et al., 2011; Ghosh and Islam, 2016; Yan et al., 2023). Therefore, there are significant differences in gene duplication events that occur in the glyoxalase families in different plant species.

The AsGLX1 family members, AsGLX1-1A1, AsGLX1-1C1, AsGLX1-1D1, AsGLX1-6C, AsGLX1-6A2, AsGLX1-7A, AsGLX1-7C1, and AsGLX1-7D, contain two conserved glyoxalase domains and are clustered with the Ni^{2+}-dependent GLX1 members in *Arabidopsis*, indicating that they are Ni^{2+}-dependent AsGLX1s. AsGLX1-5D and AsGLX1-6A1 contain a single glyoxalase domain of 141 amino acids in length and are clustered with Zn^{2+}-dependent GLX1s in *Arabidopsis* and *O. sativa*, indicating that they are Zn^{2+}-dependent AsGLX1s (Ghosh and Islam, 2016; Bhowal et al., 2020). The activity of GLX1 depends on divalent metal ions, and early studies suggested that the type of divalent ion required for GLX1 activity varies between prokaryotes and eukaryotes. GLXI in humans and yeast is Zn^{2+}-dependent, while GLX1 in *E. coli* requires Ni^{2+} for optimal activity (He et al., 2000). However, studies on plants have identified two types of divalent

ion-dependent GLX1s. For instance, in *Arabidopsis*, AtGLYI2 is dependent on Zn^{2+}, whereas AtGLYI3 and AtGLYI6 are dependent on Ni^{2+}. Most studies on the function of plant GLX1s have focused on these two types, with less research on other GLX1-like proteins. In *O. sativa*, OsGLYII-1 has ethylmalonic encephalopathy-1 activity, which can be activated by Ca^{2+}, and OsGLYII-2 has a binuclear zinc/iron center at its active site that is crucial for its activity. AsGLX2-5A, AsGLX2-5C, and AsGLX2-5D have both lactamase B and HAGH_C domains, all of which contain conserved THHHYDH metal ion-binding sites and GHT activity sites, indicating that these members encode for putative functionally active AsGLX2 enzymes (Ghosh and Islam, 2016; Bhowal et al., 2020; Singla-Pareek et al., 2020). GLX2 enzymes with these two domains, including the THHHYDH and GHT sites, have been identified in other plants such as sorghum (SbGLYII-3 and SbGLYII-4) and grape (VvGLYII-like1 and VvGLYII-like2) (Ghosh and Islam, 2016; Li et al., 2019).

Subcellular localization prediction analysis revealed that AsGLX1s are mainly expressed in the cytoplasm, chloroplasts, and mitochondria, AsGLX2s are mainly expressed in chloroplasts and mitochondria, and AsGLX3s are mainly expressed in chloroplasts. Chloroplasts and mitochondria are organelles responsible for photosynthesis and respiration, respectively, and are metabolically active. They are not only the main source of ROS, but also the potential organelles for MG production (Hasanuzzaman et al., 2017). Excessive MG has been found to inhibit photosynthesis and disrupt mitochondrial function (Kaur et al., 2016). Under stress conditions, chloroplast and mitochondrial components are often hotspots for glycation (Tripathi et al., 2023). The expression and activity of glyoxalase enzymes in chloroplasts and mitochondria are crucial for protecting the photosynthetic system and mitochondrial function. In addition, there are also some members that function in the nucleus, which may be important for protecting the nucleus and maintaining DNA and RNA stability. For example, in sorghum, SbGLYI-8/8.1 proteins were also found to harbor putative nuclear localization signals and therefore, may catalyze the conversion of nuclear MG to SLG (Bhowal et al., 2020). And OsGLYI-8 is located in the nucleus and can alleviate DNA damage caused by MG in the nucleus (Kaur et al., 2017).

Through tissue-specific expression analysis of some members, it was found that their expression patterns were consistent with subcellular localization predictions. For example, *AsGLX1-7A* was predicted to be expressed in the cytoplasm and was found to be highly expressed in developing seeds and dry seeds. *AsGLX2-3C* was predicted to be primarily expressed in mitochondria and was found to be highly expressed in roots, flowers, dry seeds, and seeds imbibed for 24 h. *AsDJ-1-4C* was predicted to be expressed in chloroplasts and cytoplasm and was found to be highly expressed in dry seeds and leaves. Overall, most of the tested members were expressed in leaves and dry seeds, indicating their important roles in maintaining leaf function and seed vigor. *AsGLX1-3D2* is exclusively expressed in developing and dry seeds, indicating that its function may be seed-specific, such as enhancing seed tolerance to stress during seed maturation and dehydration. In *Arabidopsis* and *O. sativa*, *AtGLYI8*, *OsGLYI3*, and *OsGLYI10* were also found to be highly expressed in developing seeds (Mustafiz et al., 2011). Glyoxalases have been suggested play an important role in seed development, seed germination, and seed vigor regulation. In *Arabidopsis*, a cytosolic GLXI3 isoform works on the elimination of toxic reactive carbonyl species during germination and seedling establishment (Schmitz et al., 2017). And the lack of *AtGLYI2* resulted in severe inhibition of seed germination under MG treatment, and the growth of seedlings was also limited under salt stress (Liu et al., 2022). In rice, OsGLYI7 participates in starch synthesis in the endosperm, and its mutant had significantly reduced starch content and altered expression of starch synthesis genes (You et al., 2019). While, *OsGLYI3* is specifically expressed in rice seeds and contributes to seed longevity and salt stress tolerance (Liu et al., 2022). *AsGLX2-3C* is highly expressed in roots and flowers, showing a significant difference from other members, and may play a detoxifying role during root and flower development. In *B. napus*, GLX1 is required for pollination and is targeted by the self-incompatibility system (Sankaranarayanan et al., 2015). Previous studies have mainly neglected the role of the glyoxalases in seeds. The diverse expression patterns and subcellular localization of *GLX1*, *GLX2*, and *GLX3* genes provide a basis for their functional diversity in plants, and their important roles in seeds should be given more attention in the future.

The expression patterns of glyoxalase gene families, especially *GLX1* and *GLX2* genes, under stresses have been extensively studied. It has proven GLX1 and GLX2 transcript levels and enzyme activities could be induced by various adverse conditions (Kaur et al., 2014; Sankaranarayanan et al., 2017). But their expression patterns during seed germination have been rarely reported. However, from seed sowing to seedling establishment stage, crops often suffer from stress such as drought, low temperature, or salinity, which can lead to a significant reduction in crop yield. Therefore, analyzing the expression patterns of *GLX1*, *GLX2*, and *GLX3* genes during seed germination under stress conditions is crucial for understanding their role in regulating seed vigor and coping with stress. In this study, it was found that the *GLX1* genes *AsGLX1-3D2* and *AsGLX1-7A* may play an important role in detoxification during seed germination under stress conditions. They were significantly induced by cold, drought, salt, MG, and aging treatments in the early stages of germination (6 h and 12 h). *AsGLX2-5D* was upregulated in the early stages of germination by cold, MG, and aging induction, while *AsGLX2-3C* was induced by various stresses in the later stages (24h-72h) of imbibition, indicating that these genes may play a detoxification role during seed germination under stress conditions, but with specificity in terms of stress time and types. GLX1 and GLX2 transgenic plants often exhibit increased stress resistance, while their function-deficient mutants exhibit reduced stress resistance (Kaur et al., 2014; Kaur et al., 2016; Sankaranarayanan et al., 2017). For instance, in *Arabidopsis*, complementation of *AtGLYI2* and the rice homolog *OsGLYI8* significantly enhanced stress resistance in the *atglyI-2* mutant (Kaur et al., 2017). Overexpression of *OsGLYII3* in tobacco significantly improved plant resistance to MG and NaCl (Singla-Pareek et al., 2003). Overexpression of *OsGLYII2* in *E. coli* and tobacco resulted in a significant increase in resistance to both MG and salt stresses, along with improvements in plant

photosynthetic performance and a decrease in oxidative damage (Ghosh et al., 2014). The expression analyzed GLX3 members, *AsDJ-1-3D2*, *AsDJ-1-4C*, and *AsDJ-1-5 AsDJ-1-5D*, show significant responses to stress during the early stages of seed imbibition, with *AsDJ-1-5 AsDJ-1-5D* being significantly upregulated throughout the imbibition process under stress conditions such as cold, drought, salt, and aging. In both *E. arundinaceus* and commercial sugarcane hybrids, *Gly III* respond to drought and salt stress like *Gly II* and *Gly I*, but the expression level of *Gly III* is higher under stress (Manoj et al., 2019). Additionally, the overexpression of *EaGly III* in sorghum confers significant improvements in drought and salt stress resistance, as evidenced by increased levels of proline and soluble sugars, enhanced photosynthetic and antioxidant abilities, and decreased lipid peroxidation in transgenic lines (Mohanan et al., 2020; Mohanan et al., 2021). Although studies on *GLX3* genes in plants are limited, they all indicate the great potential of *GLX3* in combating unfavorable conditions.

In this study, it was found that AsGLX1-7A has two glyoxalase domains and is a Ni^{2+}-dependent GLX1, located in the cytoplasm, and has a high gene expression level in seeds. AsGLX2-5D has THHHYDH metal ion and GHT active sites and may play a protective role in mitochondria, which are important for respiration and stress responses during seed germination. AsDJ-1-5D has two conserved DJ-1 domains and a high gene expression level in seeds. The three genes showed significant responses to various stresses and encode for putative functionally active glyoxalase enzymes. In addition, *AsGLX1-3D2* is seed-specifically expressed, and *AsGLX1-2A* may function in the protection of nucleic acid stability under stress conditions. These members of the glyoxalase gene families can be focused on in future research as potential candidate genes for oat stress resistance breeding and seed vigor improvement.

5 Conclusion

The identification of oat genes involved in stress resistance and seed vigor regulation represents a crucial endeavor in oat molecular breeding, with significant implications for crop yield and germplasm conservation. This study found that oat has more glyoxalase genes than most other plant species due to genome duplication events and tandem duplications during evolution. These genes are generally regulated by hormones and respond to adverse conditions. Their diverse tissue expression patterns and subcellular localizations indicate their functional diversity in plants, especially in leaf development and seed vigor formation. *AsGLX1-3D2* was specifically expressed in seeds, and *AsGLX1-2A* may play an important role in alleviating nucleic acid glycation. In addition, AsGLX1-7A has potential Ni^{2+}-dependent GLX1 activity, and AsDJ-1-5D has double DJ-1 domains, both of which can significantly respond to cold, drought, salt, MG and aging treatments. AsGLX2-5D has potential GLX2 activity and can respond to cold, aging, and MG stress. These highlighted genes are promising candidates for further investigation into oat stress resistance or seed vigor regulation. However, the specific functions of other members of AsGLX1, AsGLX2, and AsGLX3 families require further investigation. Nonetheless, such research may lead to the development of new strategies for breeding oats with improved stress tolerance and seed vigor through the utilization of glyoxalase genes.

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

PM conceived and designed the experiment. MS and SS performed the experiments and analyzed the data. ZJ, CO, JW, HZ, and WM contributed to the experiment. MS and SS wrote the paper, and ML and PM revised the paper. All authors contributed to the article and approved the submitted version.

Funding

This research was supported by R&D of seed coating technology for key pasture grasses in the Qinghai-Tibet Plateau (SJCZFY2022-10) and CARS (CARS-34).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1215084/full#supplementary-material

References

Bhowal, B., Singla-Pareek, S. L., Sopory, S. K., and Kaur, C. (2020). From methylglyoxal to pyruvate: a genome-wide study for the identification of glyoxalases and d-lactate dehydrogenases in *Sorghum bicolor*. *BMC Genomics* 21, 145. doi: 10.1186/s12864-020-6547-7

Borysiuk, K., Ostaszewska-Bugajska, M., Kryzheuskaya, K., Gardeström, P., and Szal, B. (2022). Glyoxalase I activity affects *Arabidopsis* sensitivity to ammonium nutrition. *Plant Cell Rep.* 41, 2393–2413. doi: 10.1007/s00299-022-02931-5

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

Chinchansure, A. A., Korwar, A. M., Kulkarni, M. J., and Joshi, S. P. (2015). Recent development of plant products with anti-glycation activity: a review. *RSC Adv.* 5, 31113–31138. doi: 10.1039/C4RA14211J

Gambhir, P., Singh, V., Raghuvanshi, U., Parida, A. P., Pareek, A., Roychowdhury, A., et al. (2023). A glutathione-independent DJ-1/PfpI domain-containing tomato glyoxalaseIII2, SlGLYIII2, confers enhanced tolerance under salt and osmotic stresses. *Plant Cell Environ.* 46, 518–548. doi: 10.1111/pce.14493

Ghosh, A. (2017). Genome-wide identification of glyoxalase genes in medicago truncatula and their expression profiling in response to various developmental and environmental stimuli. *Front. Plant Sci.* 8. doi: 10.3389/fpls.2017.00836

Ghosh, A., and Islam, T. (2016). Genome-wide analysis and expression profiling of glyoxalase gene families in soybean (*Glycine max*) indicate their development and abiotic stress specific response. *BMC Plant Biol.* 16, 87. doi: 10.1186/s12870-016-0773-9

Ghosh, A., Kushwaha, H. R., Hasan, M. R., Pareek, A., Sopory, S. K., and Singla-Pareek, S. L. (2016). Presence of unique glyoxalase III proteins in plants indicates the existence of shorter route for methylglyoxal detoxification. *Sci. Rep.* 6, 18358. doi: 10.1038/srep18358

Ghosh, A., Pareek, A., Sopory, S. K., and Singla-Pareek, S. L. (2014). A glutathione responsive rice glyoxalase II, OsGLYII-2, functions in salinity adaptation by maintaining better photosynthesis efficiency and anti-oxidant pool. *Plant J.* 80, 93–105. doi: 10.1111/tpj.12621

Hasanuzzaman, M., Nahar, K., Hossain, M., Mahmud, J. A., Rahman, A., Inafuku, M., et al. (2017). Coordinated actions of glyoxalase and antioxidant defense systems in conferring abiotic stress tolerance in plants. *Int. J. Mol. Sci.* 18, 200. doi: 10.3390/ijms18010200

He, M. M., Clugston, S. L., Honek, J. F., and Matthews, B. W. (2000). Determination of the structure of escherichia coli glyoxalase I suggests a structural basis for differential metal activation. *Biochemistry* 39, 8719–8727. doi: 10.1021/bi000856g

Hoque, T. S., Uraji, M., Tuya, A., Nakamura, Y., and Murata, Y. (2012). Methylglyoxal inhibits seed germination and root elongation and up-regulates transcription of stress-responsive genes in ABA-dependent pathway in arabidopsis. *Plant Biol. (Stuttg)* 14, 854–858. doi: 10.1111/j.1438-8677.2012.00607.x

ISTA (2019). *International rules for seed testing*. Bassersdorf: International Seed Testing Association.

Jana, G. A., Krishnamurthy, P., Kumar, P. P., and Yaish, M. W. (2021). Functional characterization and expression profiling of glyoxalase III genes in date palm grown under abiotic stresses. *Physiol. Plant* 172, 780–794. doi: 10.1111/ppl.13239

Kamal, N., Tsardakas Renhuldt, N., Bentzer, J., Gundlach, H., Haberer, G., Juhász, A., et al. (2022). The mosaic oat genome gives insights into a uniquely healthy cereal crop. *Nature* 606, 113–119. doi: 10.1038/s41586-022-04732-y

Kaur, C., Sharma, S., Singla-Pareek, S. L., and Sopory, S. K. (2016). Methylglyoxal detoxification in plants: role of glyoxalase pathway. *Ind. J. Plant Physiol.* 21, 377–390. doi: 10.1007/s40502-016-0260-1

Kaur, C., Singla-Pareek, S. L., and Sopory, S. K. (2014). Glyoxalase and methylglyoxal as biomarkers for plant stress tolerance. *Crit. Rev. Plant Sci.* 33, 429–456. doi: 10.1080/07352689.2014.904147

Kaur, C., Tripathi, A. K., Nutan, K. K., Sharma, S., Ghosh, A., Tripathi, J. K., et al. (2017). A nuclear-localized rice glyoxalase I enzyme, OsGLYI-8, functions in the detoxification of methylglyoxal in the nucleus. *Plant J.* 89, 565–576. doi: 10.1111/tpj.13407

Kumar, B., Kaur, C., Pareek, A., Sopory, S. K., and Singla-Pareek, S. L. (2021). Tracing the evolution of plant glyoxalase III enzymes for structural and functional divergence. *Antioxid. (Basel)* 10, 648. doi: 10.3390/antiox10050648

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., et al. (2007). Clustal W and clustal X version 2.0. *Bioinformatics* 23, 2947–2948. doi: 10.1093/bioinformatics/btm404

Li, Z.-G. (2016). Methylglyoxal and glyoxalase system in plants: old players, new concepts. *Bot. Rev.* 82, 183–203. doi: 10.1007/s12229-016-9167-9

Li, T., Cheng, X., Wang, Y., Yin, X., Li, Z., Liu, R., et al. (2019). Genome-wide analysis of glyoxalase-like gene families in grape (*Vitis vinifera* l.) and their expression profiling in response to downy mildew infection. *BMC Genomics* 20, 362. doi: 10.1186/s12864-019-5733-y

Liu, S., Liu, W., Lai, J., Liu, Q., Zhang, W., Chen, Z., et al. (2022). *OsGLYI3*, a glyoxalase gene expressed in rice seed, contributes to seed longevity and salt stress tolerance. *Plant Physiol. Biochem.* 183, 85–95. doi: 10.1016/j.plaphy.2022.04.028

Manoj, V. M., Anunanthini, P., Swathik, P. C., Dharshini, S., Ashwin Narayan, J., Manickavasagam, M., et al. (2019). Comparative analysis of glyoxalase pathway genes in *Erianthus arundinaceus* and commercial sugarcane hybrid under salinity and drought conditions. *BMC Genomics* 19, 986. doi: 10.1186/s12864-018-5349-7

Mohanan, M. V., Pushpanathan, A., Padmanabhan, S., Sasikumar, T., Jayanarayanan, A. N., Selvarajan, D., et al. (2021). Overexpression of glyoxalase III gene in transgenic sugarcane confers enhanced performance under salinity stress. *J. Plant Res.* 134, 1083–1094. doi: 10.1007/s10265-021-01300-9

Mohanan, M. V., Pushpanathan, A., Sasikumar, S. P. T., Selvarajan, D., Jayanarayanan, A. N., Arun, K. R., et al. (2020). Ectopic expression of DJ-1/PfpI domain containing *Erianthus arundinaceus* glyoxalase III (EaGly III) enhances drought tolerance in sugarcane. *Plant Cell Rep.* 39, 1581–1594. doi: 10.1007/s00299-020-02585-1

Mustafiz, A., Singh, A. K., Pareek, A., Sopory, S. K., and Singla-Pareek, S. L. (2011). Genome-wide analysis of rice and *Arabidopsis* identifies two glyoxalase genes that are highly expressed in abiotic stresses. *Funct. Integr. Genomics* 11, 293–305. doi: 10.1007/s10142-010-0203-2

Peng, Y., Yan, H., Guo, L., Deng, C., Wang, C., Wang, Y., et al. (2022). Reference genome assemblies reveal the origin and evolution of allohexaploid oat. *Nat. Genet.* 54, 1248–1258. doi: 10.1038/s41588-022-01127-7

Ramu, V. S., Preethi, V., Nisarga, K. N., Srivastava, K. R., Sheshshayee, M. S., Mysore, K. S., et al. (2020). Carbonyl cytotoxicity affects plant cellular processes and detoxifying enzymes scavenge these compounds to improve stress tolerance. *J. Agric. Food Chem.* 68, 6237–6247. doi: 10.1021/acs.jafc.0c02005

Rasane, P., Jha, A., Sabikhi, L., Kumar, A., and Unnikrishnan, V. S. (2015). Nutritional advantages of oats and opportunities for its processing as value added foods - a review. *J. Food Sci. Technol.* 52, 662–675. doi: 10.1007/s13197-013-1072-1

Sankaranarayanan, S., Jamshed, M., Kumar, A., Skori, L., Scandola, S., Wang, T., et al. (2017). Glyoxalase goes green: the expanding roles of glyoxalase in plants. *Int. J. Mol. Sci.* 18, 898. doi: 10.3390/ijms18040898

Sankaranarayanan, S., Jamshed, M., and Samuel, M. A. (2015). Degradation of glyoxalase I in *Brassica napus* stigma leads to self-incompatibility response. *Nat. Plants* 1, 1–7. doi: 10.1038/nplants.2015.185

Schilling, O., Wenzel, N., Naylor, M., Vogel, A., Crowder, M., Makaroff, C., et al. (2003). Flexible metal binding of the metallo-β-lactamase domain: glyoxalase II incorporates iron, manganese, and zinc in vivo. *Biochemistry* 42, 11777–11786. doi: 10.1021/bi034672o

Schmitz, J., Dittmar, I. C., Brockmann, J. D., Schmidt, M., Hüdig, M., Rossoni, A. W., et al. (2017). Defense against reactive carbonyl species involves at least three subcellular compartments where individual components of the system respond to cellular sugar status. *Plant Cell* 29, 3234–3254. doi: 10.1105/tpc.17.00258

Schmitz, J., Rossoni, A. W., and Maurino, V. G. (2018). Dissecting the physiological function of plant glyoxalase I and glyoxalase I-like proteins. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.01618

Singla-Pareek, S. L., Kaur, C., Kumar, B., Pareek, A., and Sopory, S. K. (2020). Reassessing plant glyoxalases: large family and expanding functions. *New Phytol.* 227, 714–721. doi: 10.1111/nph.16576

Singla-Pareek, S. L., Reddy, M. K., and Sopory, S. K. (2003). Genetic engineering of the glyoxalase pathway in tobacco leads to enhanced salinity tolerance. *Proc. Natl. Acad. Sci. U S A.* 100, 14672–14677. doi: 10.1073/pnas.2034667100

Singla-Pareek, S. L., Yadav, S. K., Sopory, A. M., and Sopory, S. K. (2009). Role of the glyoxalase pathway in delaying plant senescence under stress conditions. *SEB Exp. Biol. Ser.* 62, 171–185.

Sun, M., Sun, S., Jia, Z., Ma, W., Mao, C., Ou, C., et al. (2022a). Genome-wide analysis and expression profiling of glutathione reductase gene family in oat (*Avena sativa*) indicate their responses to abiotic stress during seed imbibition. *Int. J. Mol. Sci.* 23, 11650. doi: 10.3390/ijms231911650

Sun, M., Sun, S., Mao, C., Zhang, H., Ou, C., Jia, Z., et al. (2022b). Dynamic responses of antioxidant and glyoxalase systems to seed aging based on full-length transcriptome in oat (*Avena sativa* l.). *Antioxidants* 11, 395. doi: 10.3390/antiox11020395

Takagi, D., Inoue, H., Odawara, M., Shimakawa, G., and Miyake, C. (2014). The Calvin cycle inevitably produces sugar-derived reactive carbonyl methylglyoxal during photosynthesis: a potential cause of plant diabetes. *Plant Cell Physiol.* 55, 333–340. doi: 10.1093/pcp/pcu007

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

Thornalley, P. J. (1990). The glyoxalase system: new developments towards functional characterization of a metabolic pathway fundamental to biological life. *Biochem. J.* 269, 1. doi: 10.1042/bj2690001

Tripathi, D., Oldenburg, D. J., and Bendich, A. J. (2023). Oxidative and glycation damage to mitochondrial DNA and plastid DNA during plant development. *Antioxidants* 12, 891. doi: 10.3390/antiox12040891

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

Xia, F., Cheng, H., Chen, L., Zhu, H., Mao, P., and Wang, M. (2020). Influence of exogenous ascorbic acid and glutathione priming on mitochondrial structural and functional systems to alleviate aging damage in oat seeds. *BMC Plant Biol.* 20, 104. doi: 10.1186/s12870-020-2321-x

Xie, H., Li, M., Chen, Y., Zhou, Q., Liu, W., Liang, G., et al. (2021). Important physiological changes due to drought stress on oat. *Front. Ecol. Evol.* 9. doi: 10.3389/fevo.2021.644726

Xu, Z., Chen, X., Lu, X., Zhao, B., Yang, Y., and Liu, J. (2021). Integrative analysis of transcriptome and metabolome reveal mechanism of tolerance to salt stress in oat (*Avena sativa* l.). *Plant Physiol. Biochem.* 160, 315–328. doi: 10.1016/j.plaphy.2021.01.027

Yan, G., Xiao, X., Wang, N., Zhang, F., Gao, G., Xu, K., et al. (2018). Genome-wide analysis and expression profiles of glyoxalase gene families in Chinese cabbage (*Brassica rapa* l). *PloS One* 13, e0191159. doi: 10.1371/journal.pone.0191159

Yan, G., Zhang, M., Guan, W., Zhang, F., Dai, W., Yuan, L., et al. (2023). Genome-wide identification and functional characterization of stress related glyoxalase genes in *Brassica napus* l. *Int. J. Mol. Sci.* 24, 2130. doi: 10.3390/ijms24032130

Yang, Z., Wang, K., Aziz, U., Zhao, C., and Zhang, M. (2020). Evaluation of duplicated reference genes for quantitative real-time PCR analysis in genome unknown hexaploid oat (*Avena sativa* l.). *Plant Methods* 16, 138. doi: 10.1186/s13007-020-00679-1

You, X., Zhang, W., Hu, J., Jing, R., Cai, Y., Feng, Z., et al. (2019). FLOURY ENDOSPERM15 encodes a glyoxalase I involved in compound granule formation and starch synthesis in rice endosperm. *Plant Cell Rep.* 38, 345–359. doi: 10.1007/s00299-019-02370-9

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 28 June 2023
DOI 10.3389/fpls.2023.1209664

OPEN ACCESS

EDITED BY
Jiejun Xi,
Northwest A&F University, China

REVIEWED BY
Jiangqi Wen,
Oklahoma State University, United States
Tianzuo Wang,
Chinese Academy of Sciences (CAS), China
Huairong Pan,
Hunan University, China

*CORRESPONDENCE
Yuhui Chen
cyh@lzu.edu.cn

†These authors have contributed equally to this work and share first authorship

RECEIVED 21 April 2023
ACCEPTED 13 June 2023
PUBLISHED 28 June 2023

CITATION
Shen Y, Ma Y, Li D, Kang M, Pei Y, Zhang R, Tao W, Huang S, Song W, Li Y, Huang W, Wang D and Chen Y (2023) Biological and genomic analysis of a symbiotic nitrogen fixation defective mutant in *Medicago truncatula*. *Front. Plant Sci.* 14:1209664. doi: 10.3389/fpls.2023.1209664

Biological and genomic analysis of a symbiotic nitrogen fixation defective mutant in *Medicago truncatula*

Yitong Shen†, Yelin Ma†, Dengyao Li, Mingming Kang, Yue Pei, Rui Zhang, Weiyu Tao, Shenxi Huang, Wenjie Song, Yuecheng Li, Wanqi Huang, Duanyang Wang and Yuhui Chen*

Ministry of Education Key Laboratory of Cell Activities and Stress Adaptations, School of Life Sciences, Lanzhou University, Lanzhou, Gansu, China

Medicago truncatula has been selected as one of the model legume species for gene functional studies. To elucidate the functions of the very large number of genes present in plant genomes, genetic mutant resources are very useful and necessary tools. Fast Neutron (FN) mutagenesis is effective in inducing deletion mutations in genomes of diverse species. Through this method, we have generated a large mutant resource in *M. truncatula*. This mutant resources have been used to screen for different mutant using a forward genetics methods. We have isolated and identified a large amount of symbiotic nitrogen fixation (SNF) deficiency mutants. Here, we describe the detail procedures that are being used to characterize symbiotic mutants in *M. truncatula*. In recent years, whole genome sequencing has been used to speed up and scale up the deletion identification in the mutant. Using this method, we have successfully isolated a SNF defective mutant FN007 and identified that it has a large segment deletion on chromosome 3. The causal deletion in the mutant was confirmed by tail PCR amplication and sequencing. Our results illustrate the utility of whole genome sequencing analysis in the characterization of FN induced deletion mutants for gene discovery and functional studies in the *M. truncatula*. It is expected to improve our understanding of molecular mechanisms underlying symbiotic nitrogen fixation in legume plants to a great extent.

KEYWORDS

***Medicago truncatula* (barrel medic), symbiotic nitrogen fixation (SNF), nodule development, legume, whole genome sequencing (WGS)**

Introduction

Fast neutron produced from nuclear fission reactors or particle accelerator spallation sources are a type of high-energy ionizing irradiation effective in generating mutations in genomes of diverse organisms (Men et al., 2002). The nature of mutations induced by FN mutagenesis includes deletions of various sizes and other mutation types such as inversion, translocation and duplication (Bruggemann et al., 1998). Deletions cause gene loss or disruption and are especially valuable in functional studies of individual and tandemly-duplicated genes. Several large FN mutant populations have been developed in different plant species including *Arabidopsis thaliana*, *Oryza sativa L.*, *Glycine max*, *Lotus japonicus*, *barley* and *M. truncatula* (Li et al., 2002; Alonso et al., 2003; Mitra et al., 2004; Rostoks et al., 2006; Wang et al., 2006; Hoffmann et al., 2007; Bolon et al., 2011).

M. truncatula is also an ideal model species for studying legume nodule development and symbiotic nitrogen fixation (SNF) (Cook, 1999). The symbiotic interaction between *M. truncatula* and rhizobia forms indeterminate nodules, which can be divided into five main zones: meristem, infection, intermediate transition, nitrogen fixation, and senescence. The apical meristem zone is mainly responsible for cell division and new cell supply for nodule development. The infection zone is the entry point for rhizobia into nodule cells through infection threads, and the transition zone is where rhizobia enter nodule cells and form symbiont. After entering the nitrogen fixation zone, bacteroids play a role in converting atmospheric nitrogen into ammonia, which can be used directly by plants. Finally, the senescence zone is where symbiont and bacteroids degrade and lose their nitrogen-fixing activity. Clearly, the symbiotic nitrogen fixation process is very delicate and complex, and depends on the mutual recognition between plant and rhizobia at the molecular level for signaling and metabolite exchange. The symbiotic relationship will cease if there are problems in any of the above areas. Loss of symbiotic nitrogen fixation in legumes can lead to the formation of different types of mutants, which are characterized as the three main categories: (1) non-nodulating mutants (Nod-), (2) nodulating but not nitrogen-fixing mutants (Nod+/Fix-), and (3) mutants with excessive nodulation (Nod+++) (Oldroyd, 2013).

Years of mutant analyses have led to identification of genes and molecular mechanisms involved in nodule initiation and development, including mutual recognition of rhizobia and root hairs. Mutants from the early recognition process mainly affect nodule formation, but not the nitrogen fixation (Mergaert et al., 2020). However, there are relatively less studies on genes and their mechanisms involved in rhizobia release to infected nodule cells and nodule senescence. Therefore, studies on the late development of legume-rhizobia symbiotic nitrogen fixation have become a major direction in legume research (Roy et al., 2020).

The establishment of an FN mutant library in *M. truncatula* provides opportunities for studying gene functions, and a large number of mutants with symbiotic nitrogen fixation deficiency phenotypes provide a reliable guarantee for identifying gene functions and studying molecular regulatory mechanisms (Chen and Chen, 2018). However, molecular cloning of genes from FN mutants traditionally relies on the map-based cloning approach, which is a time-consuming technique that limits the amount of mutants to be analyzed at the molecular level. Several complementary approaches, such as transcript-based deletion detection and DNA copy number variation detection using a whole genome tiling array comparative genomic hybridization (CGH) platform, have been developed to facilitate the characterization of deletion mutants in diverse organisms including animals and plants (Geschwind et al., 1998; Infante et al., 2003; Gong et al., 2004; Werner et al., 2005; Bejjani and Shaffer, 2006; Lakshmi et al., 2006; Emerson et al., 2008; Perry et al., 2008).

The array-based CGH (aCGH) method is a powerful tool for the analysis of copy number variation, polymorphisms, and structural variations in the genome (Bolon et al., 2011; Haun et al., 2011; Horvath et al., 2015; Kim et al., 2015). However, low accuracy and low sensitivity for small deletions by the CGH approach are apparent. It has a higher probe density in exons than in introns and 5' and 3' UTRs, which limits the detection of sequence variations to only covered regions (Chen et al., 2017).

Recently, the whole genome sequencing (WGS) platform has delivered comprehensive genetic analyses in humans, plants and animals (Sun et al., 2018). Paired-end whole genome sequencing involves sequencing both ends of a DNA fragment, which increases the likelihood of alignment to the reference genome and facilitates detection of genomic rearrangements, repetitive sequences, and gene fusions. Unlike focused approaches such as exome or targeted resequencing, which analyze a limited portion of the genome, whole genome sequencing conveys a comprehensive view of the entire genome. It is ideal for discovery applications, such as identifying genome deletions and duplications. Whole-genome sequencing can detect single nucleotide variants, insertions, deletions, copy number changes, and large structural variants. The application of WGS platforms to FN mutant populations can identify and locate genomic changes at the whole genome level (Du et al., 2021).

Here, we isolated and characterized of the *M. truncatula* FN007 mutant from fast neutron deletion mutant library. The FN007 mutant showed defective in biological nitrogen fixation and nodule development associated with early senescence phenotype. Detailed phenotypic analysis was carried out in the FN007 mutant and would provide new insight of symbiotic nitrogen fixation processes in leguminous plants. Furthermore, the FN007 mutant was analyzed by WGS platform to detect genomic alterations. Our results confirmed that the WGS platform was efficient in mapping genome deletions and would greatly facilitate characterization of a large number of FN mutants for gene functional studies in *M. truncatula*. Using this mutant screening and identification method should improve our understanding of molecular mechanisms underlying symbiotic nitrogen fixation.

Materials and methods

Plant materials and growth condition

M. truncatula seeds were scarified with concentrated sulfuric acid for 5 to 8 min, rinsed with deionized water for five times,

surface sterilized with 5% sodium hypochlorite solution and rinsed with autoclaved deionized water for several times. After vernalization at 4°C for 3 days, seedlings were transplanted to pots and grown in greenhouse with 16 hr/8 hr light/dark cycle and 150 E/m2/sec light intensity. Perlite and sand (volume ratio of 3:1) were used as the planting medium. The perlite and sand were separately rinsed with deionized water, and then mixed in proportion for use. The germinated seeds were transferred to a 72-cell tray (530 mm × 270 mm, with a single well volume of 40 mL), watered with 1 L of 0.125 mM KNO_3 (low-nitrogen nutrient) solution. Rhizobia were inoculated after one week of growth in the greenhouse.

Inoculation with the rhizobial: *Sinorhizobium medicae* ABS7 strain and *Sinorhizobium meliloti* Sm2011-GFP strain

The rhizobia strains were removed from a -80°C refrigerator. Two mL of the rhizobia were inoculated into 50 mL of TY liquid medium containing tetracycline and cultured at 28°C for 2 days. Then, the bacterial cultures were centrifuged at 4,000 rpm for fifteen min in a 50 mL centrifuge tube, re-suspended and diluted to OD600≈0.05 with the low nitrogen nutrient solution. Five mL of the above bacterial suspension were inoculated to each plant.

Microscopic observations

Three to five plants inoculated with *S. medicae* ABS7 strain carrying the *hemA::LacZ* reporter were harvested at 6 dpi, 8 dpi, 10 dpi, 13 dpi and 16 dpi (days post inoculation). Roots were washed with deionized water. The root nodules were fixed as previously described (Xi et al., 2013). For GFP studies, nodules from plants inoculated with the *S. meliloti* Sm2011 strain harboring *hemA::mGFP* reporter gene were washed with deionized water, embedded in 6% agarose, and sectioned with Vibratome™ 1000 Plus (Leica Biosystems, Germany). Nodule sections (50-70 µm) were mounted on glass slides and imaged using MVX10+DP74 microscope (Olympus).

Live/dead staining of nodule cells and potassium iodide staining

Plants inoculated with *S. medicae* ABS7 strain were used for 50-70 µm nodule section preparation. 5 µM SYTO9 and 30 µM PI were added to 50 mM Tris (pH7.0) buffer to make a staining solution. Nodule sections were stained for 20 min at room temperature. Sections were placed on slides covered with coverslips and observed with a confocal microscope [Nikon Confocal laser scanning microscope (A1R+ Ti2-E)].

Five-micrometer semi-thin sections were placed on glass slides and stained with 1% potassium iodide phosphate buffer in dark condition for 10 min. Then, the image was observed with microscope [Axio Imager.Z2 (ZEISS)].

Semithin section of root nodules and toluidine blue staining

The nodules were placed in the fixing solution, vacuumed for 10 min, and fixed overnight at 4°C. The fixed nodules were dehydrated in 70%, 85% and 95% ethanol series for 30 min each, and in 100% ethanol for 2 h. Finally, 0.1% eosin Y3 was added to 100% ethanol to dye for half an hour. After the nodules turned red, samples were replaced in a mixture of 100% ethanol and base liquid Technovit 7100 resin (v/v=1:1) for 2 h. The nodule samples were transferred to the infiltration solution overnight (infiltration solution: 1 g hardener I to 100 mL base liquid Technovit 7100 resin. Allow approximately 5 min for complete dissolution). The nodule samples were embedded after the permeation treatment. The temperature was set to 40°C and the embedding mould was dried beforehand. The infiltrate and hardener II (v/v=15:1) were mixed gently, and 200 µL of the mixture was added to each well of the embedding mould. The nodule materials were carefully picked up with forceps and placed on the tip of the embedding mould, with materials placed vertically to facilitate sectioning. The samples were arranged neatly and then polymerized overnight in a 40°C incubator. The Leica RM2265 rotary microtome was used to cut sections with a thickness of 3-5 µm. The semithin sections were placed on glass slides and stained with 1% toluidine blue for 5 min. Finally, the stained sections were observed and imaged with microscope [Axio Imager.Z2 (ZEISS)].

Whole genome sequencing analysis

The whole genome sequencing (WGS) analysis was performed by Shanghai Oebiotech Co. The compound leaf samples were collected from wild type (WT) A17 and FN007 mutant, respectively. The genomic DNA was extracted using the DNeasy plant mini kit (Qiagen Shanghai, China), following the supplier's instructions. The library was constructed after the DNA quality was passing the bioanalyzer test. The DNA samples were randomly sheared into 350 bp-500 bp fragments, and the TruSeq DNA LT Sample Prep Kit was used for library construction. The DNA fragments were treated with the following steps to complete library construction: end repair, addition of ployA tails, addition of sequencing connectors, purification, and PCR amplification. After library construction and qualification, pair-end sequencing was performed on the sequencer (HiSeq X Ten sequencer).

The raw sequencing data were processed in two stages. Sequencing data quality was mainly examined through the sequencing error rate, data volume, comparison rate, coverage and other statistics to assess whether the sequenced library meet the standard for subsequent analysis. Then, the high-quality sequences were compared to the reference genome to detect the variation information in the samples. SNP (single nucleotide polymorphisms) and InDel (insertion or deletion) detection was performed using the GVCF mode of the Haplotypecaller module of the GATK (Version: 4.1.0.0) software, based on the results of the sample comparison with the reference genome. To reduce the error

rate of SNP and InDel assays, a criterion of QD more than 2.0 was chosen for filtering, and only mutant loci that met this criterion were retained. QD is the ratio of the quality of variation divided by the depth of coverage, which is actually the quality of variation per unit depth. The SNP and InDel results were annotated using SnpEff (Version: 4.4) software. Finally, the genomic variation information detected was visualized using Circos plots.

To eliminate the effect of sequencing errors on the results, the original sequencing data were cleaned using the pre-processing software Fastp. Clean reads were compared to reference genome using BWA, and the comparison results were converted to SAMtools, and de-redundant using Picard (http://broadinstitute.github.io/picard/), and the results were analyzed using the Qualimap software (Li and Durbin, 2009; Li et al., 2009; Chen et al., 2018). *M. truncatula* reference genome can download at the website: https://medicago.toulouse.inra.fr/MtrunA17r5.0-ANR/(MtrunA17r5.0).

TAIL PCR

Thermal asymmetric interlaced PCR (TAIL PCR) assay was carried out essentially as previously described (Liu and Chen, 2007). The specific primer P1 adjacent to the proposed deletion border was combined with five short randomly primers (AD1, AD2, AD3, AD5, AD6) with low Tm values and amplified by PCR with thermally asymmetric temperature cycling, using the FN007 mutant genomic DNA as template.

The resulting PCR product was amplified again to obtain a flanking sequence using the nested primer P2 of P1 (Table S1) and the five random primers mentioned above. The PCR product was sequenced and compared with the genome sequence of *M. truncatula* A17 to confirm the deletion junctions in FN007 mutant.

Results

Phenotypes of FN007 mutant under low nitrogen condition

M. truncatula FN007 mutant was isolated from a Fast Neutron (FN) induced deletion mutant collection in the Jemalong A17 background. Under Nitrogen replete condition, the FN007 mutant is comparable with Wild Type (WT) plants in growth and development processes, including leaf, seed pod, flower and root development (data not shown). But under Nitrogen-limited and symbiotic condition, the FN007 mutant plant develop nitrogen starvation symptoms including stunted growth and yellowish leaves at 16 days post inoculation (dpi) with *S. medicae* ABS7 strain, indicating that these phenotypes were caused by defective symbiotic nitrogen fixation (Figures 1A, B). Nodules developed on the FN007 mutant roots were small and spherical in shape, most with white colors (due to degradation of leghemoglobin), whereas in the WT plant, the pink cylindrical nodules were developed (Figures 1C–F). Moreover, nodule distribution was concentrated in WT plants but scattered in FN007 mutant and the WT plant had fewer nodule numbers than the FN007 mutant.

Time-course observation of nodule phenotype

To further characterize the nodule development processes of the FN007 mutant, time-course experiments were carried out. Under low nitrogen condition (0.125 mM KNO_3), the growth and development of root nodules in the FN007 mutant and WT plant inoculated with *S. medicae* ABS7 strain was continuously monitored and recorded. Morphological and color differences between FN007 mutant and WT nodules became apparent over time (Figure 2). At 6 dpi, WT (Figure 2A) and FN007 mutant (Figure 2F) nodules appeared to be similar, both white and spherical shape. At 9 dpi, WT nodules (Figure 2B) became pale pink color and conical in shape, while FN007 mutant nodules (Figure 2G) were still white color and cone-shaped. At 12dpi, WT nodules (Figure 2C) were pink and cylindrical shape, but FN007 mutant nodules (Figure 2H) gradually increased in size with increasing growth time, and remained white color. At 15 dpi, WT nodules (Figure 2D) continued to grow, almost entirely pink, and were already mature and capable of efficient nitrogen fixation. FN007 mutant nodules (Figure 2I) grew slowly, with one or two white and rod-like, and the remainder were basically white spheres or cones shape. At 19 dpi, WT nodules (Figure 2E) were generally the same as those at 15 dpi, but the number of mature nodules increased, while FN007 mutant nodules (Figure 2J) still appeared as white rod shape. Time-course experiments showed that early nodule development (6 dpi) appears to be normal in the FN007 mutant. However, differences began apparent at 6-9 dpi with leghemoglobin accumulation in WT nodules. Even though the FN007 nodules did not show the pink functional leghemoglobin, some nodules were able to elongate and grow into rod-like shape during later development stage, indicating that the nodule meristem still had normal developmental functions.

Under low nitrogen condition (0.125 mM KNO_3), the FN007 mutant and WT seedlings grown for one week in greenhouse were inoculated with *S. medicae* ABS7 strain containing the *hemA::LacZ* reporter gene. Nodules were harvested at different development stages after inoculation and stained with *LacZ* staining buffer. At 12 dpi, in FN007 mutant nodules (Figure 3B), there were none or few rhizobia in nodule cells. For WT plant at the same time (Figure 3A), rhizobia filled the entire nodule cells. At 15 dpi, the number of rhizobia-containing nodule cells in FN007 mutant (Figure 3D) was decreased, and nodule cell debris was observed, at which point a distinct senescence zone (rather than a nitrogen-fixing zone) had formed, which was significantly different from the WT (Figure 3C). We speculated that between 6-7 dpi, the FN007 mutant failed to form a nitrogen-fixing zone and went directly into senescence, forming a region that accounted for an increasing proportion of nodules over time. To confirm this result, another rhizobia strain with GFP reporter gene, *S. meliloti* Sm2011-GFP, was inoculated. At

10.3389/fpls.2023.1209664

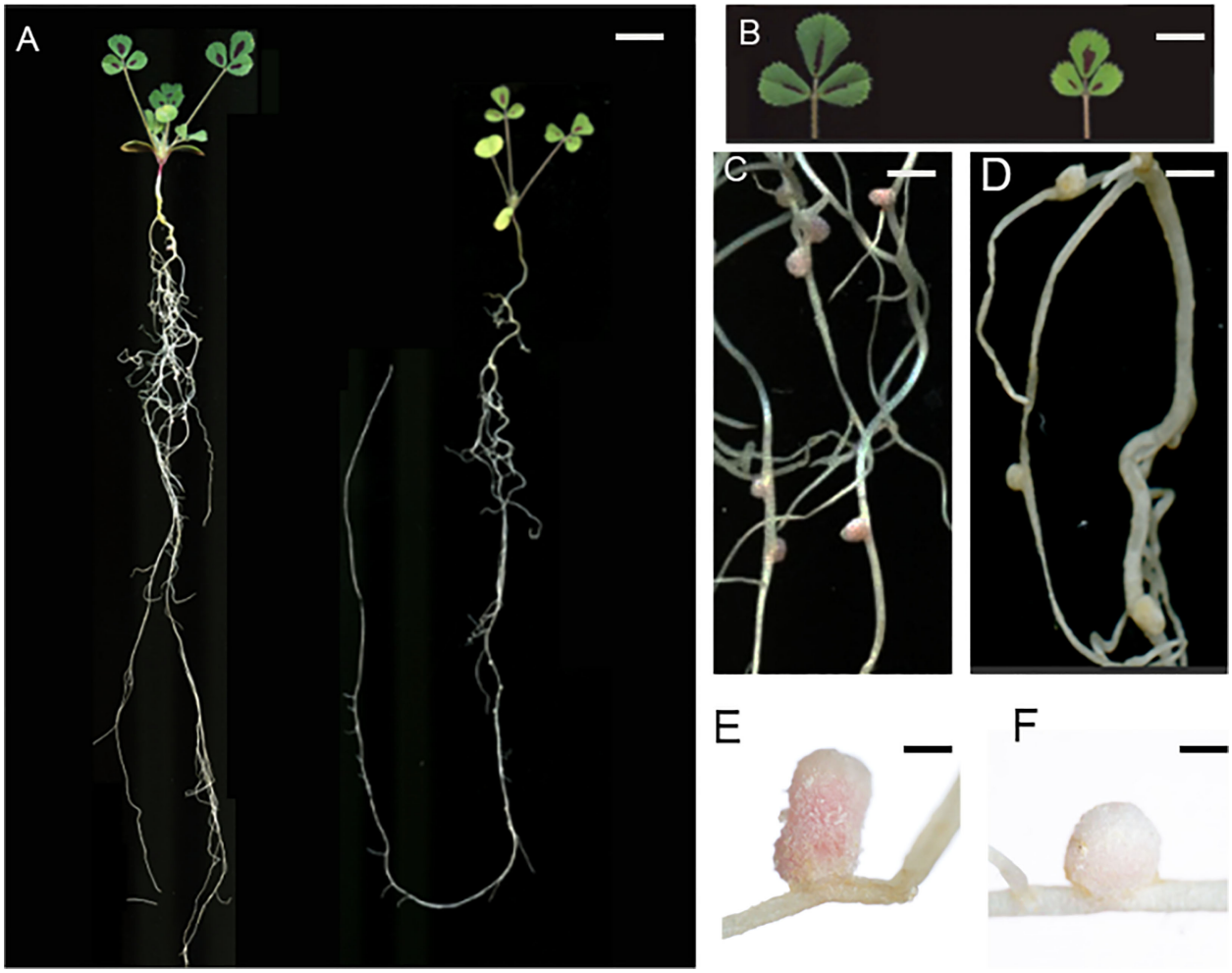


FIGURE 1
M. truncatula FN007 mutant phenotype. **(A)** Growth of FN007 (right) and wild type (WT, left) seedlings 16 days post inoculation (dpi) with *S. medicae* ABS7 strain, showing dwarf phenotype of FN007 mutant. **(B)** leaf phenotype of FN007 mutant (right) and WT (left), yellowish leaves are the sign of nitrogen deficiency phenotype. Nodules developed on **(C)** (WT) and **(D)** (FN007 mutant) were pink and white color, respectively. Close-up view of **(E)**, WT and **(F)**, FN007 mutant nodules. Scale bar: **(A)**, 1 cm; **(B)**, 50 mm; **(C, D)**, 10 mm; **(E, F)**, 1 mm.

11 dpi, the fluorescence signal in wild type nodules was very strong, indicating that rhizobia released and filled with nodule cells (Figure 4A). However, in FN007 mutant nodules, though the fluorescence intensity of zones I and II was no different from the WT, zone III fluorescence was distributed in a punctiform pattern and no mature symbiont was formed (Figure 4B), indicating that rhizobia release into nodule cells during infection was normal, which was consistent with the results of nodule sections inoculated with *S. medicae* ABS7 strain. At 14 dpi, the WT nodules showed characteristic partitioning and a dominant green fluorescent signal in mature symbiotic cells completely filled with bacteroids (Figure 4C), whereas in the FN007 mutant nodules an increasingly severe degradation of the bacterial cycle in the nitrogen-fixing zone was observed (Figure 4D). This observation suggests the FN007 mutant nodules may undergo early senescence.

M. truncatula FN007 mutant is defective in symbiotic nitrogen fixation

In indeterminate nodules, rhizobia released into nodule cells differentiate before nitrogen fixation begins, and the differentiated rhizobia are morphologically longer than free-state rhizobia and have endoreduplication of their chromosomes (Mergaert et al., 2006). From GFP and *LacZ* stained sections experiments, we found that the FN007 mutant was normal in early nodule development stage including infection thread development and rhizobia release. To further determine the development and differentiation of rhizobia released into nodule cells, 5 µm semithin longitudinal nodule sections were stained with Toluidine blue for microscopic observation. In healthy symbionts, rhizobia undergo intracellular replication and differentiation into rod-like

10.3389/fpls.2023.1209664

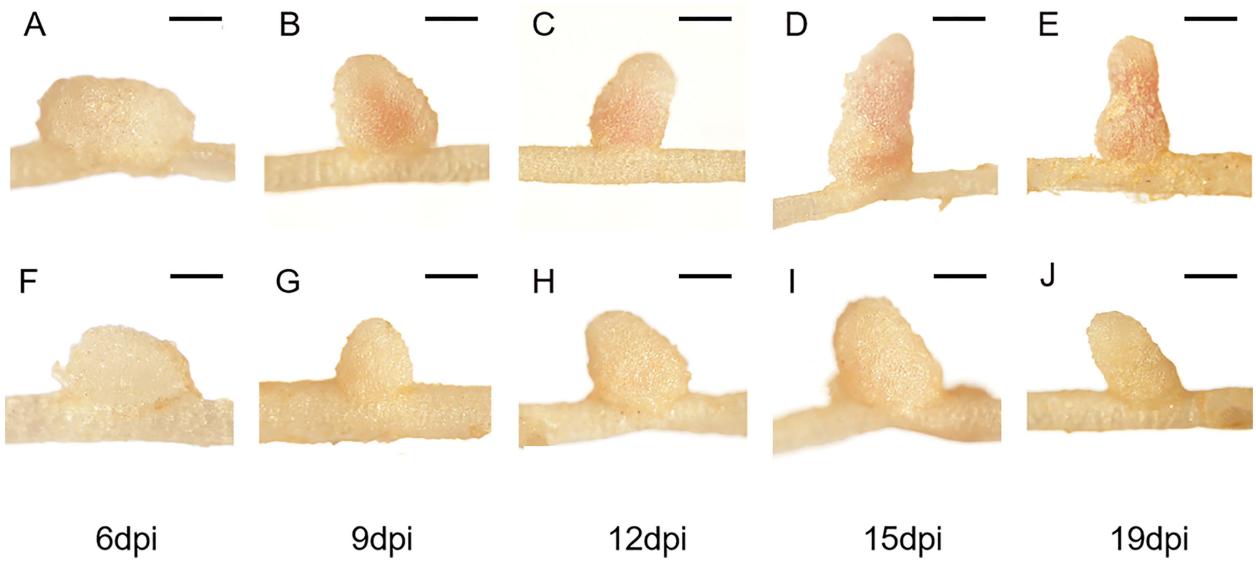


FIGURE 2
Time-course observations of nodule phenotypes. *M.truncatula* WT and FN007 mutant seedlings were inoculated with *S. medicae* ABS7 strain and inoculated root nodules were observed in a time-course assay. Representative images of **(A)**, WT and **(F)**, FN007 mutant roots 6 dpi; **(B)**, WT and **(G)**, FN007 mutant 9 dpi; **(C)**, WT and **(H)**, FN007 mutant 12 dpi; **(D)**, WT and **(I)**, FN007 mutant 15 dpi; **(E)**, WT and **(J)**, FN007 mutant 19 dpi. These observations suggest that FN007 mutant has an abnormal nodule development. Scale bar: 500 μm.

structures evenly distributed around a central vacuole, but rhizobia in infected cells need to differentiate to exert nitrogen-fixing function properly. At 13 dpi, WT nodules (Figure 5A) had formed four distinct zones: zone I (zones of meristem, ZI), zone II (infection, ZII), zone II-III (a transition zone between ZII and ZIII, IZ), and zone III (nitrogen fixation, ZIII). Zone II-III was a transition zone from zone II to III, when rhizobia had already entered into infected cells, and with a compartment area of about one-fifth of the total rhizobia area. Infected cells were filled with intact symbiosomes in WT. In this period, the FN007 mutant nodules appeared to have four zones in terms of cell morphology: I, II, IZ, and IV (senescence zone), but did not appear to have the nitrogen-fixing zone (ZIII) as observed in WT nodules, entering directly into the senescence zone. Symbiosomes in most infected cells were disintegrated at the proximal end of the root in the FN007 mutant (Figure 5D). In the WT, the senescence zone was formed at 35 dpi in the absence of environmental stress, but in the FN007 mutant, the senescence zone was formed at 13 dpi, which clearly showed an early senescence phenotype. Therefore, it was assumed that symbiotic nitrogen fixation defect occurs in the II-III transition zone, and rhizobia cannot complete the transition after its release into the infected cells in the FN007 nutant. In WT nodules (Figures 5B, C), infected cells were morphologically normal, mostly subrounded, with a large number of symbiosomes containing rod-shaped bodies regularly arranged around a large central vacuole. In the FN007 mutant, infected cells (Figures 5E, F) were smaller and irregularly shaped compared to that in WT nodules, with less symbiosomes and most were not rod-shaped, resembling the free-living cells.

To further characterize the symbiotic phenotype of the FN007 mutant, semithin sections of nodules at 13 dpi were stained for starch granules with potassium iodide (KI). The WT nodules had almost no accumulation of starch granules at the distal root end (Figures 6A, B), but in the nitrogen fixation zone (Figure 6C), there was a small number of starch granules deposition around the cell membrane. Conversely, starch granules in the distal end of the FN007 nodules were evenly distributed and surrounded the cell membrane (Figure 6D), where clearly cell contours were observed and the starch granules were similar in size. At the proximal root end of the FN007 mutant nodule (Figures 6E, F), the number of deposited amyloplasts was increased and distributed in a disorganized manner. Overall, the number of starch granules in FN007 mutant nodules were greatly exceeded that of the wild type. These results indicated that the FN007 mutant cannot form a functional fixation zone, and thus many metabolic activities can not be carried out in infected cells, which accumulating more energy material (starch grains).

To confirm whether the FN007 mutant nodules undergo early senescence, we performed live/dead staining experiment, using fluorescence dyes SYTO9, which stains live bacteroids and nucleic acids of plant cells and propidium iodide (PI), which stains dead bacteroids. The FN007 mutant and WT plants were inoculated with *S. medicae* ABS7 strain. A large green fluorescence was observed within the WT nodule (Figures 7A–C), and clearly visible symbiosomes were formed within infected cells and regularly arranged around the central vacuole. At 13 dpi, the FN007 mutant nodules were mostly stained red fluorescence compared with the WT (Figures 7D–F), and high magnification images

10.3389/fpls.2023.1209664

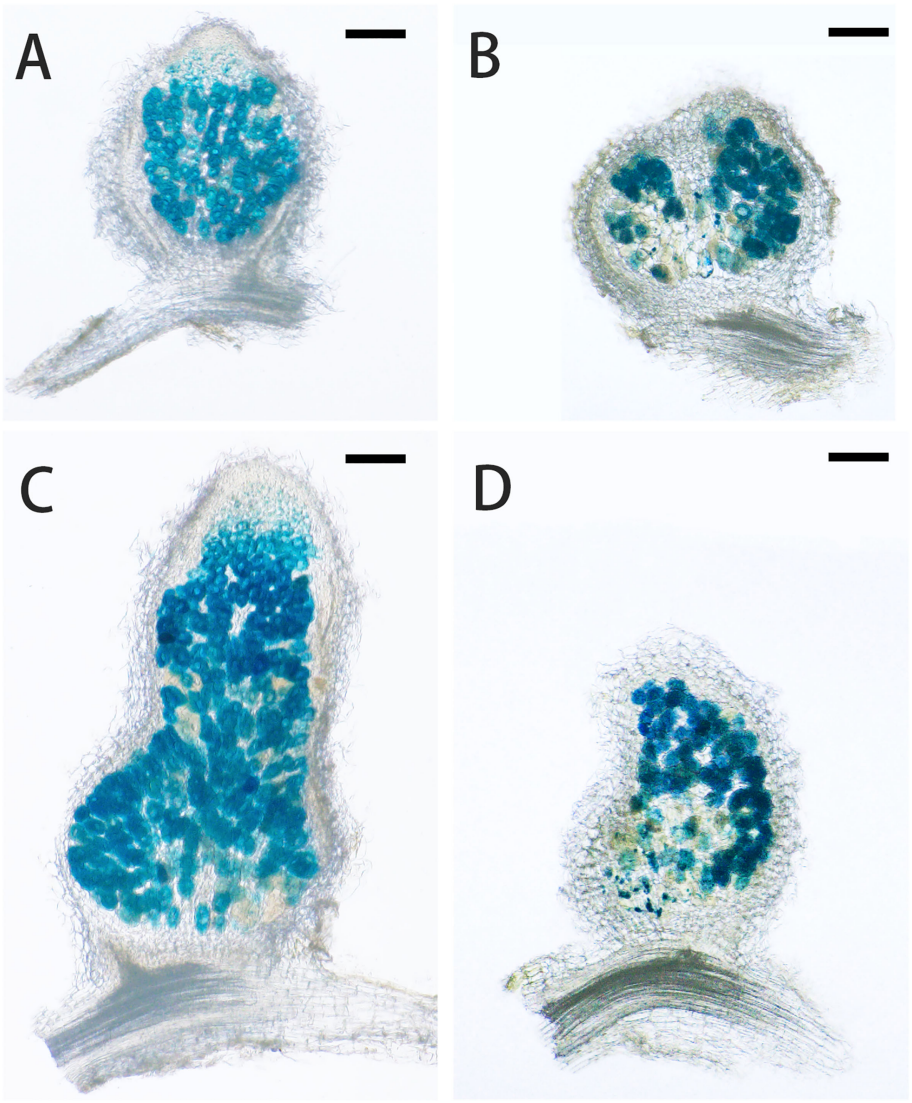


FIGURE 3
LacZ stained sections observation of nodule development. *M.truncatula* WT and FN007 mutant seedlings were inoculated with *S. medicae* ABS7 strain harboring *hemA::LacZ* reporter gene. Longitudinal sections of nodules were observed under a MVX10+DP74 microscope and representative images were shown. **(A)**, WT and **(B)**, FN007 mutant nodules 12 dpi; **(C)**, WT and **(D)**, FN007 mutant nodules 15 dpi. Scale bar: 200 µm.

showed that infected cells contained both red and green scattered bacteroids, indicating that some bacteroids in the FN007 mutant were dead at this time. These observations suggest that nodule cells of the FN007 mutant exhibited signs of an early senescence phenotype (12 dpi).

Whole genome sequencing analysis

To identify mutations in the FN007 mutant, whole genome resequencing was carried out. Raw reads were cleaned after quality control and QC results were obtained (Table 1). Normally, for next generation sequencing values, Q20 should be over 95% (at worst, not below 90%), and Q30 is required to be more than 85% (at worst, not below 80%). The sequencing quality of the FN007 mutant and WT was high, with both Q20 values over 97% and both Q30 values over 93% (Table 1). The distribution of base error rates for pair-end reads in length range of 1 to 150 bp is shown in Figure 8. Regardless of read length, the average base error rate was less than 0.05%, indicating high sequencing quality.

Whole-genome sequencing uses non-strand-specific library building. Theoretically, G and C, and A and T bases should be equal in each sequencing cycle, and the whole sequencing process should be stable. The FN007 mutant base distribution sequencing data is shown in Figure 9. The start position of the sequence is linked to the primer junction during sequencing, so A, C, G and T will fluctuate at the beginning but stabilize later.

For species with a reference genome, their genomic information or complete transcript information can be downloaded from public databases, and then the sequenced reads can be compared to the

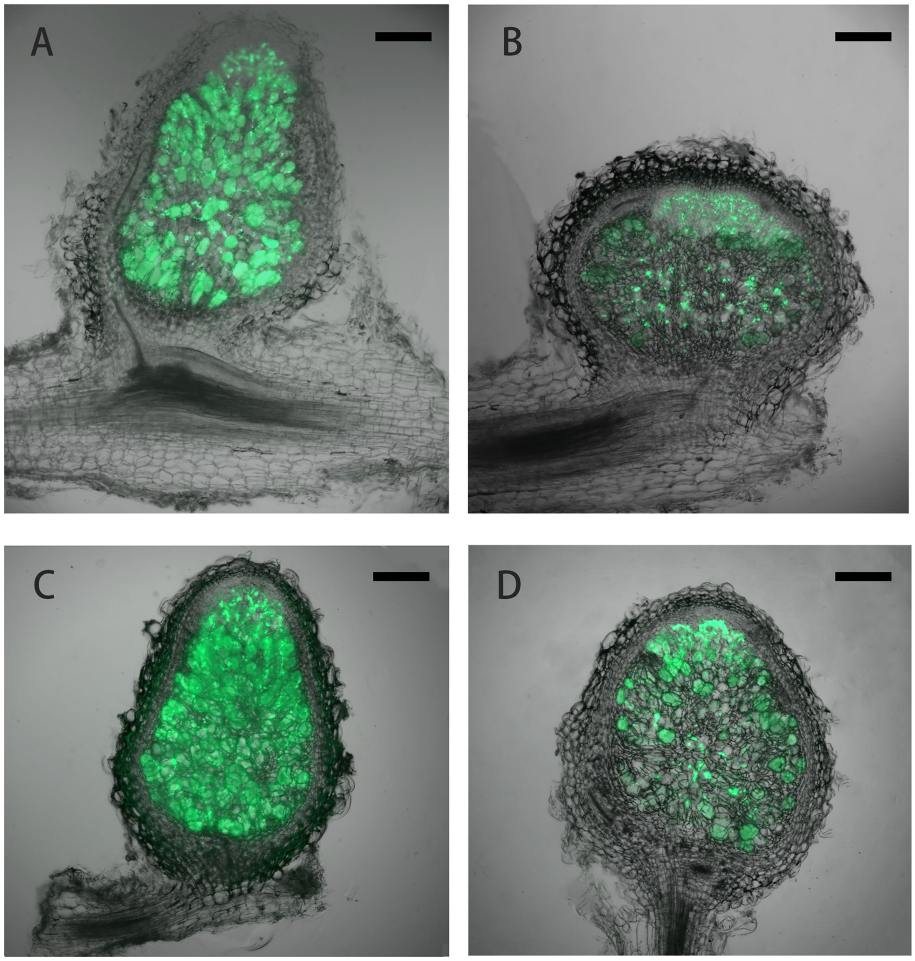


FIGURE 4
Laser confocal microscopic analysis of nodule development. *M.truncatula* WT and FN007 mutant seedlings were inoculated with *S. meliloti* Sm2011 strain harboring *hemA::mGFP* reporter gene. Longitudinal sections of nodules were observed under a MVX10+DP74 microscope and representative images were shown. **(A)**, WT and **(B)**, FN007 mutant nodules 11 dpi; **(C)**, WT and **(D)**, FN007 mutant nodules 14 dpi. Scale bar: 200 µm.

reference genome. By quickly comparing each read to the reference genome, information about read position and match quality on the compared genome or other reference sequences can be obtained, and then the genes or transcripts can be annotated and quantified. After comparing the clean reads of the FN007 mutant and WT with the reference genome, the statistical results are shown in Table 2. It was suggested that reads from the FN007 mutant is comparable with WT. For the FN007 mutant, more than 189 million pair-end reads (2×150 bp) were produced, which represented a 65.2× sequencing depth. For the WT plant, 151 million reads were produced, representing 52.3× sequencing depth (Table 2). These reads were mapped to the latest reference genome of *M. truncatula* (MtrunA17r5.0). The total mean Mapping quality is nearly 50% in both mutant and WT. These result indicated that the WGS data has higher quality of the alignment.

SNP and InDel detection were performed using the GATK4 software Haplotypecaller module, based on the comparison of samples with the reference genome. QD is the ratio of variant quality divided by the depth of coverage, which is actually variant quality per unit depth, and most false positive variants have a QD of less than 2. Compared to transversion, transition is the most common mutation type, with T to C and A to G being the most common substitution types (Figure 10A). Of all the InDels obtained, deletions less than 4 bp accounted for 85.28% and insertions less than 4 bp accounted for 94.53% (Figure 10B). All the mutations detected were integrated into a Circos plot to visualize the genomic data (Figure 10C).

The FN007 mutant resequencing data revealed that there was a copy number variation (CNV) in large segment of chromosome 3. CNVs may be biologically important for species-specific genome composition, species evolution and phylogeny, as well as for the expression and regulation of genes in specific genome regions. We used IGV software to visualize detection CNVs in the FN007 mutant, and found a possible large deletion on chromosome 3 from location 4,518,720 to 4,871,880 (Figure 11A). To determine the specific interval of the mutant deletion, four pairs of primer (Table S1) were first designed on both sides of each proposed boundary provided by the IGV data, and these primers were amplified by PCR in the FN007 mutant and WT (Figure 11B).

10.3389/fpls.2023.1209664

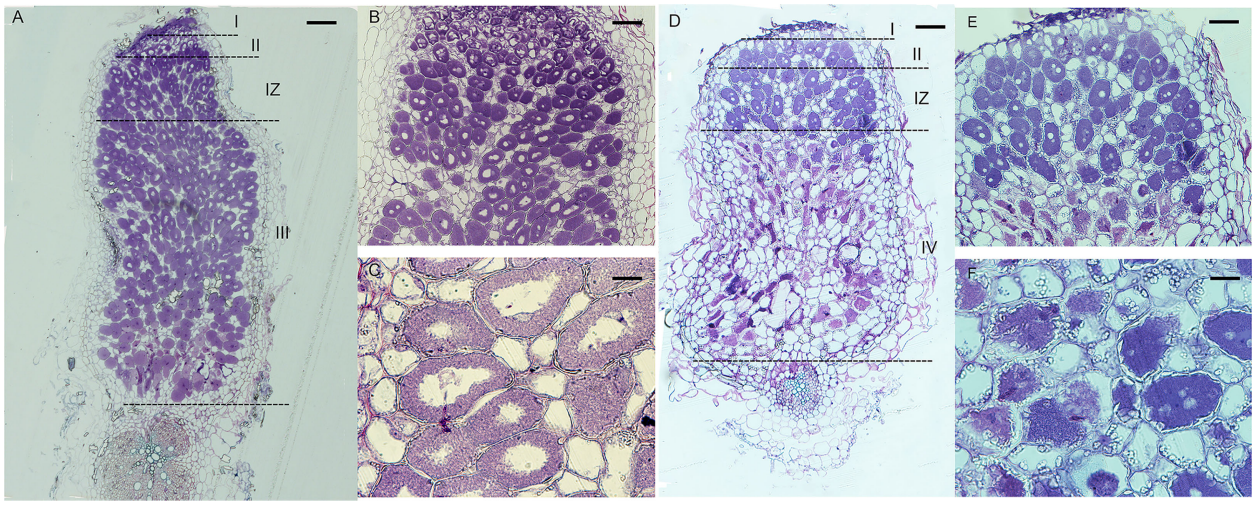


FIGURE 5
Semithin sections of WT and FN007 mutant nodules. *M. truncatula*, **(A)**, WT and D, FN007 mutant nodules 13 days post inoculation (dpi) were sectioned longitudinally and stained with toluidine blue, showing development of meristem **(I)**, infection (II) and transition zones (IZ), and infected cells in both WT and FN007 mutant. Compared with WT nodules, there was no fixation zone (III) in FN007 mutant, but a clearly senescence zone (IV) was emerged. Close-up view of **(B)**, WT and **(E)**, FN007 infected cells. Magnification view of **(C)**, WT and **(F)**, FN007 infected cells, showing degradation of symbiosomes in FN007 infected cells. Scale Bar: **(A, D)**: 100 μm; **(B, E)**: 20 μm; **(C, F)**: 10 μm.

The results showed that primers 3 and 4 could amplify the target bands in the FN007 mutant, but in contrast, primers 1 and 2 could not (Figure 11C). This suggests that the IGV data matches the actual deletion junction in the FN007 mutant. To identify the exactly deletion site, TAIL PCR was performed to identify the deletion regions in the FN007 mutant (Figure 11D). Results showed that the deletion border in the FN007 mutant was located at position 4,519,828-4,872,290 on chromosome 3. Based on the TAIL PCR results, we designed left and right border primers and verified the results, conformed that there was a deletion region in the FN007 mutant compared with WT (Figure 11E).

Genetic linkage analysis

Whole genome sequencing analysis revealed that the FN007 mutant has a large deletion of approximately 348 kb on chromosome 3. To characterize the nature of the FN007

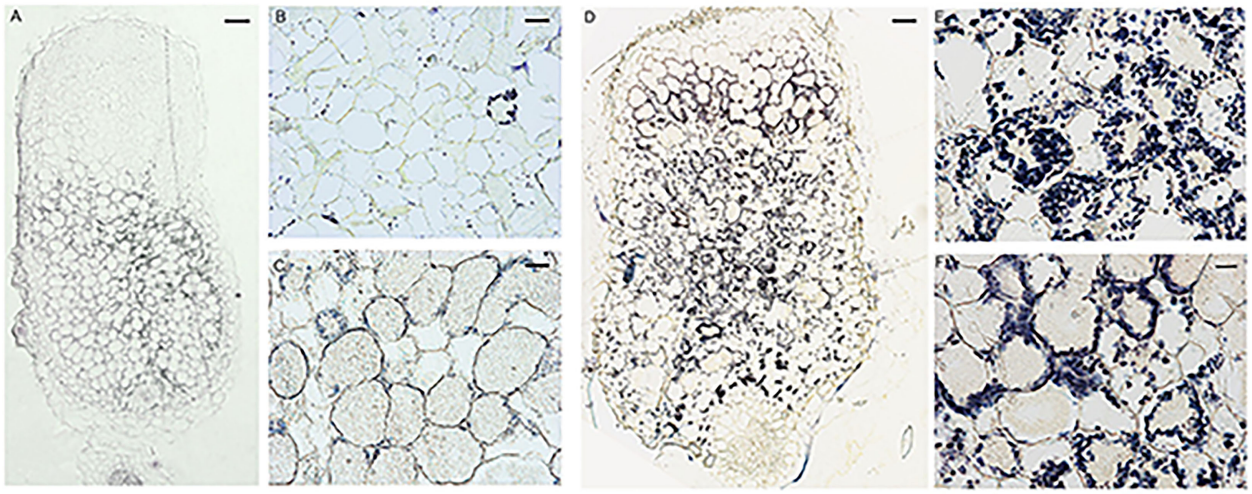


FIGURE 6
Accumulation of starch granules in FN007 mutant nodules. Potassium iodide staining of starch granules in infected cells in WT nodules **(A–C)** and FN007 mutant nodules **(D–F)**. In WT nodules, starch granules was deposited in large amounts only in the II-III interval zone and in small amounts in the IV zone, while in FN007 mutant nodules, the starch granules deposition was greatly expanded. Scale bar: **(A)**: 100 μm; **(D)**:200 μm; **(B, E)**: 20 μm; **(C, F)**: 10 μm.

10.3389/fpls.2023.1209664

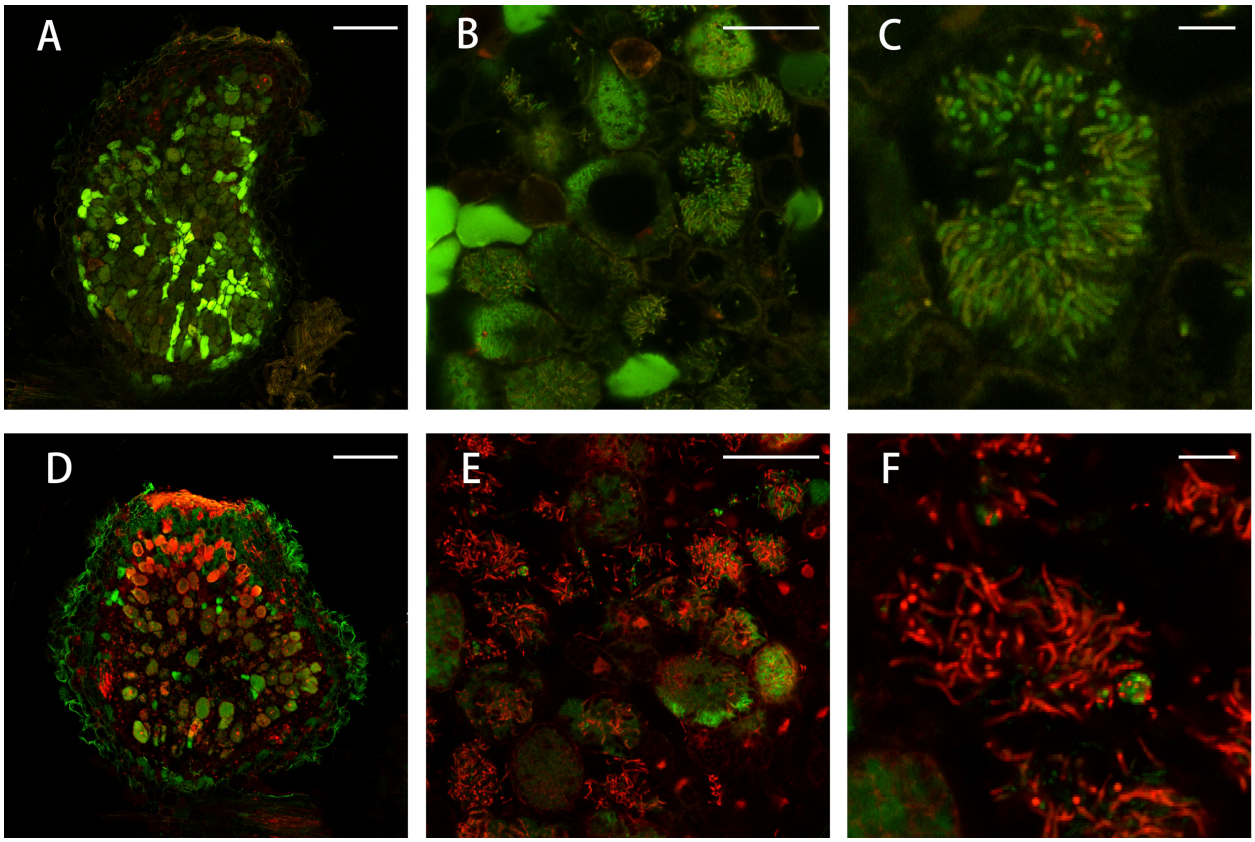


FIGURE 7
Live/dead staining of nodule cells. Live/dead staining of bacteroids in FN007 mutant compared with WT nodules (13 dpi). **(A–C)**, wild type (WT) and **(D–F)**, FN007 mutant. Live bacteroids have a green signal (GFP) and dying bacteroids have a red signal from propidium iodide (PI). PI can only enter bacteroids with compromised membranes. FN007 mutant nodules fail to sustain differentiated bacteroids. Scale bar: **(A, D)**: 200 μm; **(B, E)**: 50 μm; **(C, F)**: 10 μm.

mutation, we generated F2 populations from cross-pollination between the FN007 mutant and its WT parent (Jemalong A17).

We analyzed 163 individual F2 plants and observed a ratio of 3.075:1 segregation of WT and mutant nodulation phenotypes in F2 populations, indicating a single recessive locus for the FN007 mutant phenotype. Next, we performed a linkage analysis, using pools of DNA samples isolated from individual F2 plants. The 123 plants with the WT phenotype were divided into 12 groups and 40 plants with the mutant phenotype were divided into 4 groups (the DNA from each group was pooled together). The deletion regions were identified by PCR using the corresponding primers (Figure 12A). The results showed that all the mutants plant could amplify the deletion border by MDB-F and MDB-R primer (data not shown), but could not get the corresponding band by the MDB-F and DB-R primer, while the WT plants can amplify the correct bands in the deletion region (Figure 12B). These results indicated that the fix$^-$ phenotype of the FN007 mutant was tightly linked to the deleted region on chromosome 3.

Within the deletion region on chromosome 3, there are 53 annotated genes base on *M. truncatula* A17 r5.0 genome (Figure 12C; Table 3). This, together with the results of the genetic linkage analysis, suggests that these genes are candidates

TABLE 1 Statistics of sample sequencing quality control results.

Samples	Raw_ Reads	Clean_ Reads	Clean_Reads_ Percent	Raw_Base	Clean_Base	Clean_Base_ Percent	GC_ Content	>Q20	>Q30
FN007	206334558	204918294	99.31%	30950183700 (30.95G)	30378091695 (30.38G)	98.15%	34.79%	97.53%	93.03%
WT	164747736	163907076	99.49%	24712160400 (24.71G)	24411975186 (24.41G)	98.79%	34.72%	97.80%	93.83%

Raw_Reads represent Number of original sequencing reads; Clean_Reads represent Number of valid reads after quality control; Clean_Reads_Percent represent Clean_Reads as a proportion of Raw_Reads; Raw_Base represent Amount of raw sequencing data in bp; Clean_Base represent The amount of effective data after quality control in bp; Clean_Base_Percent represent the proportion of Clean_Base to Raw_Base; >Q20 represent the proportion of bases with a quality value greater than Q20 to Clean_Base; >Q30 represent the proportion of bases with a quality value greater than Q30 in Clean_Base.

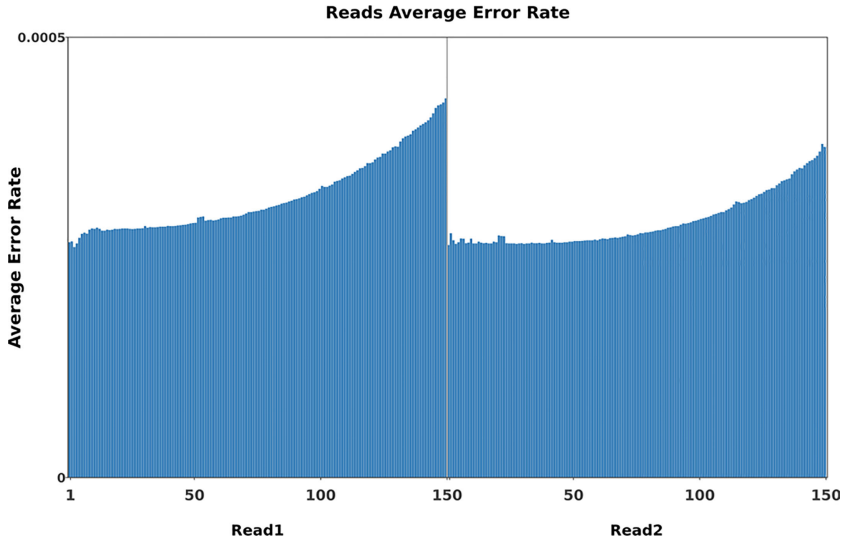


FIGURE 8
Base error rate distribution graph. Horizontal coordinates represent the base positions on reads, and vertical coordinates represent the average error rate of single bases; the left side shows the base quality distribution of R1 Reads in pair-end sequencing, and the right side shows the base quality distribution of R2 Reads.

for the cause of the phenotype of the FN007 mutant. Though we have successfully identified the deletion region in the FN007 mutant, it will take more effort to pinpoint the causal gene. To identify which is the target gene, RNA-seq technology would be used to narrow down the gene list in the deletion region. Finally, genetic complementation assay should be carried out, using *Agrobacterium rhizogenes*-mediated root transformation.

Discussion

Leguminous plants are able to establish symbiotic relationships with the soil bacteria rhizobia and as a result, develop nodules on their roots, in which biological nitrogen fixation take place. The development of nodules on roots of host plants is triggered by the perception of Nod factors by receptor-like kinase complexes of

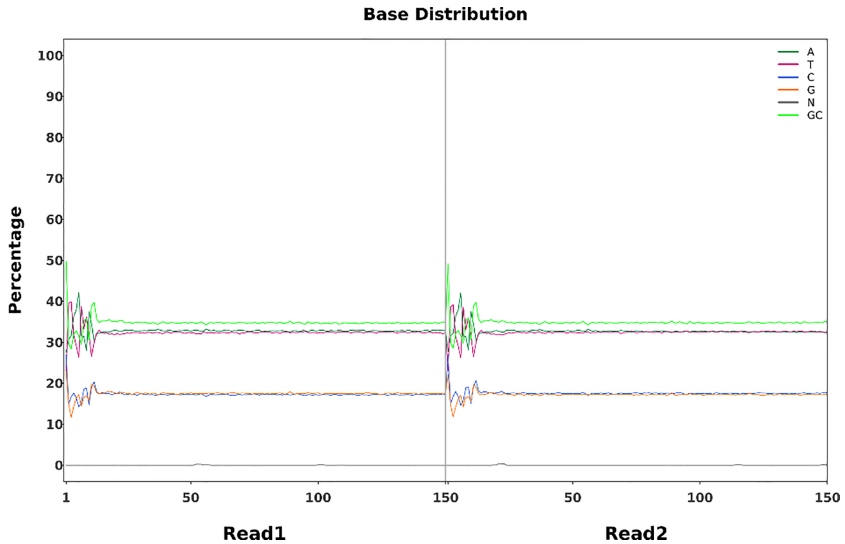


FIGURE 9
Base percentage distribution. Horizontal coordinates represent base position on the reads, and vertical coordinates represent the proportion of single bases, A,T,C,G,N, and GC, respectively.

TABLE 2 Statistical table of comparison information between FN007 and WT.

Samples	WT	FN007
Total_reads	163907076	204918294
Dup_reads	10113696	15706947
Dup_percent	6.17%	7.66%
RmDup_reads	153793380	189211347
Mapped_reads	151287765	188990765
Mapped_Rate	98.37%	99.88%
Mean_coverge	52.3507×	65.2098×
Mean_mapping_quality	49.237	49.1109
1×_coverage	99.75%	99.67%
10×_coverage	99.36%	99.36%

Total_Reads represent the number of valid Reads after quality control; Dup_Reads represent the number of PCR duplicate Reads; Dup_percent represent the ratio of PCR duplicate Reads; RmDup_Reads represent the number of Reads after removing PCR duplicates; Mapped_Reads represent the number of Reads on the reference genome; Mapped_Rate represent the ratio between Mapped_Reads and RmDup_Reads; mean_coverage represent average coverage depth, calculated as the total number of bases paired to the reference genome/reference genome length; mean_mapping_quality represent the higher the value, the better the quality of the alignment; 1×_coverage represent the percentage of loci covered by at least 1 Reads in the reference genome to the whole length of the genome, i.e. the genome coverage; 10×_coverage represent the number of loci covered by at least 10 Reads in the reference genome as a percentage of the total genome length.

compatible host plants (Oldroyd et al., 2011; Zipfel and Oldroyd, 2017). Within cytoplasm of plant cells, rhizobia are enclosed by the plant-derived membrane, symbiosome membrane and differentiate into nitrogen fixing bacteroids, and symbiotic nitrogen fixation occurs. Later on, nodule cells senesce and nitrogen fixation stop (Puppo et al., 2005). In *M. truncatula*, nodules formed on roots are indeterminate. They are cylindrical and contain a series of development zones of different cell types. In mature nodules, there is a senescence zone, where the host cells and bacteroids degenerate. For most leguminous plants, significant nodule senescence is observed during pod filling, when nutrient reallocation and remobilization occur to ensure yield and seed quality. Therefore, increasing active nitrogen fixation in nodules by delaying their senescence might be a powerful strategy to improve nitrogen-use efficiency during seed development (Yu et al., 2023). However, the key factors that orchestrate nodule senescence have not been identified (Wang et al., 2022).

In this study, we isolated and characterized a new nodulation mutant from *M. truncatula*, FN007. The FN007 mutant was isolated as being defective in symbiotic nitrogen fixation from a *M. truncatula* deletion mutant collection generated by fast neutron bombardment. Phenotypic characterization showed that the FN007 mutant was defective in later stage of nodule development. In contrast to WT plants, which develop rod shape nodules, the FN007 mutant plants develop round, spherical nodules. A possible explanation for this could be that the deletion of the target gene affects the expression of the leghemoglobin gene in FN007 mutant (Jiang et al., 2021), resulting in white root nodules at all stages of development, and that the expression of genes involved in root nodule development is restricted (Li et al., 2022).

Microscopic observations of the FN007 mutant nodules showed the presence of the meristem zone, infection zone, transition zone, and senescence zone without a functional fixation zone. The symbiosomes and infected cells were degraded starting about 12 dpi. The number of nodules was much higher and the distribution of nodules was scattered in the FN007 mutant than WT plants after inoculation with rhizobia. These observations suggest that nodules developed on the FN007 mutant roots undergo early senescence.

Mutagenesis is a powerful tool for gene functional studies. Fast neutron, a high energy ionization mutagen, has been used in generating mutants in many plant species. A large portion of mutations derived from FN mutagenesis are due to DNA deletions that range in size from a few base pairs to mega base pairs. Many phenotype-associated genes have been successfully identified and characterized. Previously, molecular cloning of the underlying genes from FN mutants relied on a map-based approach, which is time -consuming and limits the number of mutants to be characterized at the molecular level. Recently, several complimentary approaches including genome tiling array-based comparative genomic hybridization (CGH) for DNA copy number variation detection, has been employed to facilitate the characterization of deletion mutants in diverse organisms including animals and plants. These approaches provided new solutions in characterizing with FN mutants. However, low resolution and low accuracy, especially for small deletions limited these methods to apply in identifying mutations for FN mutants.

With the rapid development of next generation sequencing (NGS) technologies, whole genome sequencing (WGS) has become more affordable, providing a new means of identifying causal deletions in FN mutants (Sun et al., 2018). WGS for FN mutants not only increases the sensitivity in detecting mutations on chromosomes, but also improves the reliability and resolution in identifying deletions (Du et al., 2021). In this study, we used it to analyze symbiotic nitrogen fixation deficiency mutant FN007 in *M. truncatula* and identified the deletion locus that was tightly linked with this phenotype. The causal deletion in the FN007 mutant was confirmed by Tail-PCR amplification, sequencing and genetic linkage analyses.

WGS can be used for large and small deletion identification in many species with reference genome sequences. WGS uses high coverage short reads to examine the large deletions or indel between the reference genome and short reads. In this study, we performed a 30×sequencing depth for mutation detection. Given sufficient sequencing depth, WGS could theoretically detect all potential deletions or indel caused by different mutagenesis. WGS provided a powerful platform to rapidly identify the candidate deletions for FN mutants.

In this study, the FN007 mutant showed a clearly different phenotype from WT under nitrogen-limited and symbiotic condition, so further use of the hairy root transformation technique to rescue the phenotype of the mutant and find the target gene is very important to elucidate the mechanism of symbiotic nitrogen fixation in leguminous plants. This work will provide a reference for most researchers to investigate similar mutant and explore the molecular mechanism of symbiotic nitrogen fixation in leguminous plants.

10.3389/fpls.2023.1209664

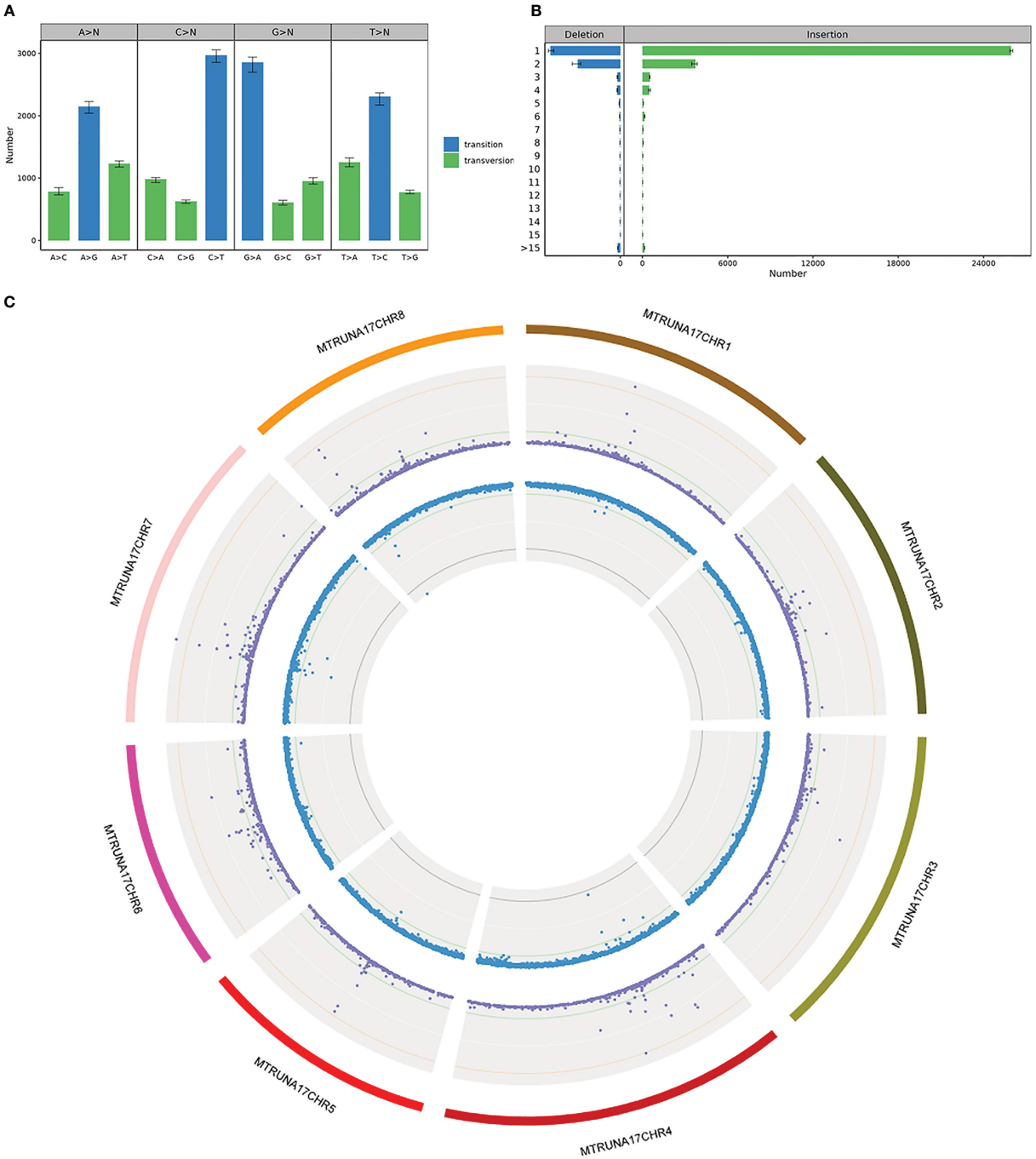


FIGURE 10
SNP & InDel Variant Detection in FN007 mutant **(A)** SNP substitution type statistics in FN007 mutant. Horizontal coordinates represent SNP substitution type, vertical coordinates are the mean, maximum and minimum values for each sample, and colors indicate whether SNP substitution type is a transition or a transversion. **(B)** InDel substitution type length statistics in FN007 mutant. Horizontal coordinates show the mean, maximum and minimum values of the InDel numbers, vertical coordinates show substitution base lengths, and colors indicate whether the substitution type is insertion or deletion. **(C)** Circos visualization of identified deletions in FN007 mutant. SNP and InDel distributions on individual chromosomes, where the first circle represents the FN007 chromosome, the purple dots in the second circle represent SNP density, and blue dots in the third circle represent InDel density.

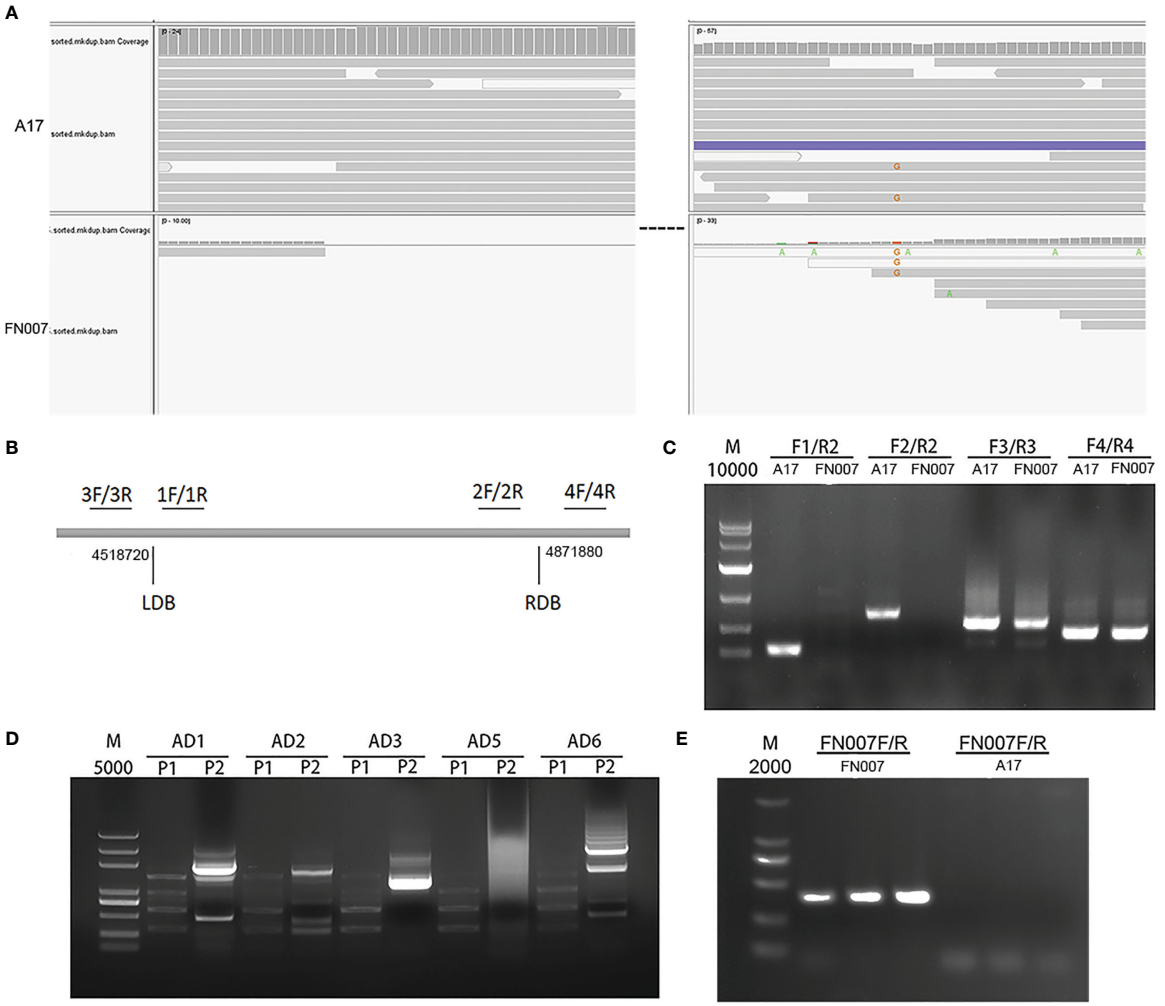


FIGURE 11
Identification of copy number variation (CNV) in FN007 mutant. **(A)** Visualization of the causal deletion using IGV. The reads alignment for WT and FN007 mutant. The sequencing showed a large deletion at the corresponding position in FN007 mutant; **(B)** Coordinate positions between four primers are designed for the identification of deletion border in FN007 mutant; **(C)** PCR detection for deletion junction through the flanking sequencing primers in FN007 mutant; **(D)** Tail-PCR was carried out, and the corresponding band were extracted for sequencing; **(E)** Deletion border confirmed by PCR in FN007 mutant. A pair of primer (FN007F/FN007R) was designed at deletion junction and amplified in FN007 mutant and WT. The results showed that FN007F/R could amplify the target band in FN007 mutant but not in WT.

10.3389/fpls.2023.1209664

10.3389/fpls.2023.1209664

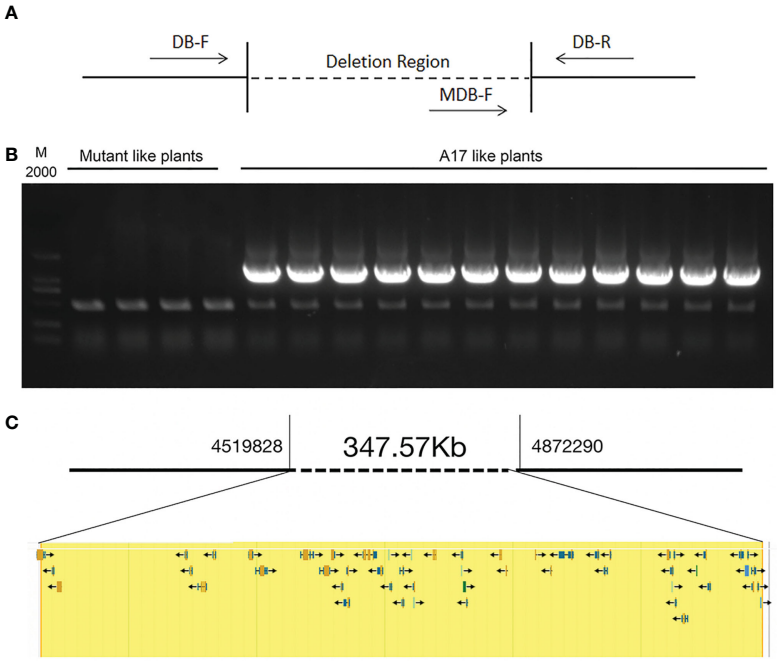


FIGURE 12
Genetic linkage analysis of FN007 mutant A genetic linkage analysis of FN007 F2 segregating population. **(A)** PCR analysis was carried out using primers spanning the deletion border in FN007 F2 population; **(B)** Shown are PCR amplification of deletion border in F2 mutant pools (mutant like plants, 10 plants per pool) and in wild type-like pools (WT like plants, 10 plants per pool). M, DL2000, molecular weight markers. There were no specific bands in F2 mutant pool by MDB-F and DB-R primer sets, while in WT like pools, all of pools could amplify the correct band. These results indicate that the deletion region has linkage with FN007 mutant phenotype; **(C)** A putative 348 kb deletion segment in FN007 mutant.

TABLE 3 Gene list in deletion region of FN007 mutant.

No.	Start	End	Gene name	Annotation
1	4517833	4521967	MtrunA17Chr3g0082791	Putative P-loop containing nucleoside triphosphate hydrolase%2C leucine-rich repeat domain%2C L
2	4525308	4526227	MtrunA17Chr3g0082801	hypothetical protein
3	4527738	4530138	MtrunA17Chr3g0082811	Putative P-loop containing nucleoside triphosphate hydrolase%2C leucine-rich repeat domain%2C L
4	4590083	4591230	MtrunA17Chr3g0082821	hypothetical protein
5	4592379	4593528	MtrunA17Chr3g0082831	hypothetical protein
6	4596357	4600625	MtrunA17Chr3g0082841	Putative tetratricopeptide-like helical domain%2C thioredoxin-like protein
7	4603651	4605160	MtrunA17Chr3g0082851	hypothetical protein
8	4621143	4623476	MtrunA17Chr3g0082861	Putative response regulator and transcription factor RR-A-type family
9	4624537	4630582	MtrunA17Chr3g0082871	Putative P-loop containing nucleoside triphosphate hydrolase%2C leucine-rich repeat domain%2C L
10	4646449	4653195	MtrunA17Chr3g0082881	Putative P-loop containing nucleoside triphosphate hydrolase%2C leucine-rich repeat domain%2C L
11	4655741	4660843	MtrunA17Chr3g0082891	Putative P-loop containing nucleoside triphosphate hydrolase%2C leucine-rich repeat domain%2C L
12	4661521	4663557	MtrunA17Chr3g0082901	Putative transcription factor interactor and regulator CCHC(Zn) family
13	4666281	4667366	MtrunA17Chr3g0082911	hypothetical protein
14	4667573	4670400	MtrunA17Chr3g0082921	Early nodulin-10
15	4669003	4669760	MtrunA17Chr3g0082931	hypothetical protein
16	4676905	4683938	MtrunA17Chr3g0082941	Early nodulin-12A
17	4684026	4686560	MtrunA17Chr3g0082951	Putative Gnk2-like domain-containing protein

(Continued)

10.3389/fpls.2023.1209664

TABLE 3 Continued

No.	Start	End	Gene name	Annotation
18	4689411	4689819	MtrunA17Chr3g1009661	hypothetical protein
19	4689934	4691347	MtrunA17Chr3g0082961	hypothetical protein
20	4694891	4697009	MtrunA17Chr3g0082971	hypothetical protein
21	4695914	4696898	MtrunA17Chr3g0082981	Early nodulin-12A
22	4700523	4700914	MtrunA17Chr3g1009664	hypothetical protein
23	4701367	4702332	MtrunA17Chr3g0082991	hypothetical protein
24	4701570	4702308	MtrunA17Chr3g1009666	hypothetical protein
25	4711121	4712887	MtrunA17Chr3g0083001	Putative Serpin family protein
26	4724803	4725523	MtrunA17Chr3g0083011	Putative Late nodulin
27	4724972	4725480	MtrunA17Chr3g1009670	hypothetical protein
28	4725879	4727343	MtrunA17Chr3g0083021	hypothetical protein
29	4727375	4727763	MtrunA17Chr3g0083031	hypothetical protein
30	4743584	4744548	MtrunA17Chr3g0083041	Putative Serpin family protein
31	4746749	4747099	MtrunA17Chr3g0083051	hypothetical protein
32	4761026	4761329	MtrunA17Chr3g0083061	hypothetical protein
33	4768494	4768931	MtrunA17Chr3g0083071	hypothetical protein
34	4772616	4779909	MtrunA17Chr3g0083081	Putative Late nodulin
35	4790225	4791820	MtrunA17Chr3g0083091	Putative Late nodulin
36	4794477	4796323	MtrunA17Chr3g0083101	hypothetical protein
37	4797383	4797805	MtrunA17Chr3g0083111	Putative Late nodulin
38	4825327	4826727	MtrunA17Chr3g0083121	Putative transcription factor interactor and regulator CCHC(Zn) family
39	4826893	4828195	MtrunA17Chr3g0083131	Putative Late nodulin
40	4827653	4828070	MtrunA17Chr3g1009692	hypothetical protein
41	4828580	4829067	MtrunA17Chr3g0083141	hypothetical protein
42	4829203	4829717	MtrunA17Chr3g0083151	hypothetical protein
43	4832527	4835586	MtrunA17Chr3g0083161	hypothetical protein
44	4839696	4840116	MtrunA17Chr3g0083171	hypothetical protein
45	4843168	4844486	MtrunA17Chr3g0083181	hypothetical protein
46	4845159	4846684	MtrunA17Chr3g0083191	hypothetical protein
47	4860530	4862920	MtrunA17Chr3g0083201	hypothetical protein
48	4863407	4865364	MtrunA17Chr3g0083211	hypothetical protein
49	4864538	4865110	MtrunA17Chr3g0083221	hypothetical protein
50	4866497	4868953	MtrunA17Chr3g0083231	hypothetical protein
51	4866613	4867679	MtrunA17Chr3g0083241	hypothetical protein
52	4869407	4870204	MtrunA17Chr3g0083251	hypothetical protein
53	4870951	4871406	MtrunA17Chr3g0083261	hypothetical protein

Conclusions

M. truncatula has been selected as model species for legume-rhizobial symbiosis and nodule development studies. Here, we describe the detail procedures that are being used to characterize SNF mutants in *M. truncatula*, which provide a reliable solutions to utilize FN mutants for SNF research. Our results demonstrate that the methods used here is feasible in screening and identifying the mutation sites for FN mutants in *M. truncatula*.

Data availability statement

The data presented in the study are deposited in the NCBI Sequence Read Archive (SRA) repository, accession number is PRJNA983295. The release date is: 2023-06-14. The SRA records will be accessible with the following link: https://www.ncbi.nlm.nih.gov/sra/PRJNA983295.

Author contributions

YC, YS and YM conceived the original research plans. YM, YS, DL, YP and MK performed all the phenotype analysis. YM, YS, DL, WT, RZ, SH, WS and YL performed all bioinformatics analysis. WH and DW performed most of the experiments. YC, YS and YM drafted, wrote, and edited the manuscript. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by the foundation of the Ministry of Education Key Laboratory of Cell Activities and Stress Adaptations (lzujbky-2019-kb05; lzujbky-2021-kb05).

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1209664/full#supplementary-material

References

Alonso, J. M., Stepanova, A. N., Solano, R., Wisman, E., Ferrari, S., Ausubel, F. M., et al. (2003). Five components of the ethylene-response pathway identified in a screen for weak ethylene-insensitive mutants in arabidopsis. *Proc. Natl. Acad. Sci. U S A.* 100 (5), 2992–2997. doi: 10.1073/pnas.0438070100

Barker, D., Bianchi, S., Blondon, F., Dattee, Y., Duc, G., Essad, S., et al. (1990). Medicago truncatula, a model plant for studying the molecular genetics of the rhizobium-legume symbiosis. *Plant Mol. Biol. Rep.* 8, 40–49. doi: 10.1007/BF02668879

Bejjani, B. A., and Shaffer, L. G. (2006). Application of array-based comparative genomic hybridization to clinical diagnostics. *J. Mol. Diagn.* 8 (5), 528–533. doi: 10.2353/jmoldx.2006.060029

Bolon, Y. T., Haun, W. J., Xu, W. W., Grant, D., Stacey, M. G., Nelson, R. T., et al. (2011). Phenotypic and genomic analyses of a fast neutron mutant population resource in soybean. *Plant Physiol.* 156 (1), 240–253. doi: 10.1104/pp.110.170811

Bruggemann, E. P., Doan, B., Handwerger, K., and Storz, G. (1998). Characterization of an unstable allele of the arabidopsis HY4 locus. *Genetics.* 149 (3), 1575–1585. doi: 10.1093/genetics/149.3.1575

Chen, Y., and Chen, R. (2018). Physical mutagenesis in medicago truncatula using fast neutron bombardment (FNB) for symbiosis and developmental biology studies. *Methods Mol. Biol.* 1822, 61–69. doi: 10.1007/978-1-4939-8633-0_4

Chen, Y., Wang, X., Lu, S., Wang, H., Li, S., and Chen, R. (2017). An array-based comparative genomic hybridization platform for efficient detection of copy number variations in fast neutron-induced medicago truncatula mutants. *J. Vis. Exp.* 129), 56470. doi: 10.3791/56470

Chen, J., Yu, J., Ge, L., Wang, H., Berbel, A., Liu, Y., et al. (2010). Control of dissected leaf morphology by a Cys(2)His(2) zinc finger transcription factor in the model legume medicago truncatula. *Proc. Natl. Acad. Sci. U S A.* 107 (23), 10754–10759. doi: 10.1073/pnas.1003954107

Chen, S., Zhou, Y., Chen, Y., and Gu, J. (2018). Fastp: an ultra-fast all-in-one FASTQ preprocessor. *Bioinformatics* 34 (17), i884–ii90. doi: 10.1093/bioinformatics/bty560

Cook, D. R. (1999). Medicago truncatula–a model in the making! *Curr. Opin. Plant Biol.* 2 (4), 301–304. doi: 10.1016/s1369-5266(99)80053-3

Du, H., Jiao, Z., Liu, J., Huang, W., and Ge, L. (2021). Rapid identification of mutations caused by fast neutron bombardment in medicago truncatula. *Plant Methods* 17 (1), 62. doi: 10.1186/s13007-021-00765-y

Emerson, J. J., Cardoso-Moreira, M., Borevitz, J. O., and Long, M. (2008). Natural selection shapes genome-wide patterns of copy-number polymorphism in drosophila melanogaster. *Science.* 320 (5883), 1629–1631. doi: 10.1126/science.1158078

Geschwind, D. H., Gregg, J., Boone, K., Karrim, J., Pawlikowska-Haddal, A., Rao, E., et al. (1998). Klinefelter's syndrome as a model of anomalous cerebral laterality: testing gene dosage in the X chromosome pseudoautosomal region using a DNA microarray. *Dev. Genet.* 23 (3), 215–229. doi: 10.1002/(sici)1520-6408(1998)23:3%3C215::aid-dvg7%3E3.0.co;2-x

Gong, J. M., Waner, D. A., Horie, T., Li, S. L., Horie, R., Abid, K. B., et al. (2004). Microarray-based rapid cloning of an ion accumulation deletion mutant in arabidopsis thaliana. *Proc. Natl. Acad. Sci. U S A.* 101 (43), 15404–15409. doi: 10.1073/pnas.0404780101

Haun, W. J., Hyten, D. L., Xu, W. W., Gerhardt, D. J., Albert, T. J., Richmond, T., et al. (2011). The composition and origins of genomic variation among individuals of the soybean reference cultivar williams 82. *Plant Physiol.* 155 (2), 645–655. doi: 10.1104/pp.110.166736

Hoffmann, D., Jiang, Q., Men, A., Kinkema, M., and Gresshoff, P. M. (2007). Nodulation deficiency caused by fast neutron mutagenesis of the model legume lotus japonicus. *J. Plant Physiol.* 164 (4), 460–469. doi: 10.1016/j.jplph.2006.12.005

Horvath, B., Domonkos, A., Kereszt, A., Szucs, A., Abraham, E., Ayaydin, F., et al. (2015). Loss of the nodule-specific cysteine rich peptide, NCR169, abolishes symbiotic nitrogen fixation in the medicago truncatula dnf7 mutant. *Proc. Natl. Acad. Sci. U S A.* 112 (49), 15232–15237. doi: 10.1073/pnas.1500777112

Infante, J. J., Dombek, K. M., Rebordinos, L., Cantoral, J. M., and Young, E. T. (2003). Genome-wide amplifications caused by chromosomal rearrangements play a major role in the adaptive evolution of natural yeast. *Genetics.* 165 (4), 1745–1759. doi: 10.1093/genetics/165.4.1745

Jiang, S., Jardinaud, M.-F., Gao, J., Pecrix, Y., Wen, J., Mysore, K., et al. (2021). NIN-like protein transcription factors regulate leghemoglobin genes in legume nodules. *Science.* 374 (6567), 625–628. doi: 10.1126/science.abg5945

Kim, M., Chen, Y., Xi, J., Waters, C., Chen, R., and Wang, D. (2015). An antimicrobial peptide essential for bacterial survival in the nitrogen-fixing symbiosis. *Proc. Natl. Acad. Sci. U S A.* 112 (49), 15238–15243. doi: 10.1073/pnas.1500123112

Lakshmi, B., Hall, I. M., Egan, C., Alexander, J., Leotta, A., Healy, J., et al. (2006). Mouse genomic representational oligonucleotide microarray analysis: detection of copy number variations in normal and tumor specimens. *Proc. Natl. Acad. Sci. U S A.* 103 (30), 11234–11239. doi: 10.1073/pnas.0602984103

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics.* 25 (14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Li, H., Handsaker, B., Wysoker, A., Fennell, T., Ruan, J., Homer, N., et al. (2009). The sequence Alignment/Map format and SAMtools. *Bioinformatics.* 25 (16), 2078–2079. doi: 10.1093/bioinformatics/btp352

Li, X., Lassner, M., and Zhang, Y. (2002). Deleteagene: a fast neutron deletion mutagenesis-based gene knockout system for plants. *Comp. Funct. Genomics* 3 (2), 158–160. doi: 10.1002/cfg.148

Li, Y., Pei, Y., Shen, Y., Zhang, R., Kang, M., Ma, Y., et al. (2022). Progress in the self-regulation system in legume nodule development-AON (Autoregulation of nodulation). *Int. J. Mol. Sci.* 23 (12). doi: 10.3390/ijms23126676

Liu, Y. G., and Chen, Y. L. (2007). High-efficiency thermal asymmetric interlaced PCR for amplification of unknown flanking sequences. *BioTechniques.* 43 (5), 649–650. doi: 10.2144/000112601

Men, A., Laniya, T., Searle, I., Iturbe-Ormaetxe, I., Gresshoff, I., Jiang, Q., et al. (2002). Fast neutron mutagenesis of soybean (glycine soja l.) produces a supernodulating mutant containing a Large deletion in linkage group h. *Gen. Lett.* 3, (3). doi: 10.1166/gl.2002.017

Mergaert, P., Kereszt, A., and Kondorosi, E. (2020). Gene expression in nitrogen-fixing symbiotic nodule cells in medicago truncatula and other nodulating plants. *Plant Cell.* 32 (1), 42–68. doi: 10.1105/tpc.19.00494

Mergaert, P., Uchiumi, T., Alunni, B., Evanno, G., Cheron, A., Catrice, O., et al. (2006). Eukaryotic control on bacterial cell cycle and differentiation in the rhizobium-legume symbiosis. *Proc. Natl. Acad. Sci. U S A.* 103 (13), 5230–5235. doi: 10.1073/pnas.0600912103

Mitra, R. M., Gleason, C. A., Edwards, A., Hadfield, J., Downie, J. A., Oldroyd, G. E., et al. (2004). A Ca2+/calmodulin-dependent protein kinase required for symbiotic nodule development: gene identification by transcript-based cloning. *Proc. Natl. Acad. Sci. U S A.* 101 (13), 4701–4705. doi: 10.1073/pnas.0400595101

Oldroyd, G. E. (2013). Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nat. Rev. Microbiol.* 11 (4), 252–263. doi: 10.1038/nrmicro2990

Oldroyd, G. E., Murray, J. D., Poole, P. S., and Downie, J. A. (2011). The rules of engagement in the legume-rhizobial symbiosis. *Annu. Rev. Genet.* 45, 119–144. doi: 10.1146/annurev-genet-110410-132549

Pecrix, Y., Staton, S. E., Sallet, E., Lelandais-Briere, C., Moreau, S., Carrere, S., et al. (2018). Whole-genome landscape of medicago truncatula symbiotic genes. *Nat. Plants.* 4 (12), 1017–1025. doi: 10.1038/s41477-018-0286-7

Penmetsa, R. V., and Cook, D. R. (2000). Production and characterization of diverse developmental mutants of medicago truncatula. *Plant Physiol.* 123 (4), 1387–1398. doi: 10.1104/pp.123.4.1387

Perry, G. H., Yang, F., Marques-Bonet, T., Murphy, C., Fitzgerald, T., Lee, A. S., et al. (2008). Copy number variation and evolution in humans and chimpanzees. *Genome Res.* 18 (11), 1698–1710. doi: 10.1101/gr.082016.108

Puppo, A., Groten, K., Bastian, F., Carzaniga, R., Soussi, M., Lucas, M. M., et al. (2005). Legume nodule senescence: roles for redox and hormone signalling in the orchestration of the natural aging process. *New Phytol.* 165, 683–701. doi: 10.1111/j.1469-8137.2004.01285.x

Rogers, C., Wen, J., Chen, R., and Oldroyd, G. (2009). Deletion-based reverse genetics in medicago truncatula. *Plant Physiol.* 151 (3), 1077–1086. doi: 10.1104/pp.109.142919

Rostoks, N., Schmierer, D., Mudie, S., Drader, T., Brueggeman, R., Caldwell, D. G., et al. (2006). Barley necrotic locus nec1 encodes the cyclic nucleotide-gated ion channel 4 homologous to the arabidopsis HLM1. *Mol. Genet. Genomics* 275 (2), 159–168. doi: 10.1007/s00438-005-0073-9

Roy, S., Liu, W., Nandety, R. S., Crook, A., Mysore, K. S., Pislariu, C. I., et al. (2020). Celebrating 20 years of genetic discoveries in legume nodulation and symbiotic nitrogen fixation. *Plant Cell.* 32 (1), 15–41. doi: 10.1105/tpc.19.00279

Stacey, G., Libault, M., Brechenmacher, L., Wan, J., and May, G. D. (2006). Genetics and functional genomics of legume nodulation. *Curr. Opin. Plant Biol.* 9 (2), 110–121. doi: 10.1016/j.pbi.2006.01.005

Sun, L., Ge, Y., Bancroft, A. C., Cheng, X., and Wen, J. (2018). FNBtools: a software to identify homozygous lesions in deletion mutant populations. *Front. Plant Sci.* 9. doi: 10.3389/fpls.2018.00976

Tadege, M., Wen, J., He, J., Tu, H., Kwak, Y., Eschstruth, A., et al. (2008). Large-Scale insertional mutagenesis using the Tnt1 retrotransposon in the model legume medicago truncatula. *Plant J.* 54 (2), 335–347. doi: 10.1111/j.1365-313X.2008.03418.x

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., et al. (2014). An improved genome release (version Mt4.0) for the model legume medicago truncatula. *BMC Genomics* 15, 312. doi: 10.1186/1471-2164-15-312

Wang, H., Chen, J., Wen, J., Tadege, M., Li, G., Liu, Y., et al. (2008). Control of compound leaf development by FLORICAULA/LEAFY ortholog SINGLE LEAFLET1 in medicago truncatula. *Plant Physiol.* 146 (4), 1759–1772. doi: 10.1104/pp.108.117044

Wang, D., Dong, W., Murray, J., and Wang, E. (2022). Innovation and appropriation in mycorrhizal and rhizobial symbioses. *Plant Cell.* 34 (5), 1573–1599. doi: 10.1093/plcell/koac039

Wang, H., Li, G., and Chen, R. (2006). Fast neutron bombardment (FNB) mutagenesis for forward and reverse genetic studies in plants. In *Floriculture, ornamental and plant biotechnology: advances and topical issues*, ed. Teixeira da Silva, J. A. (Isleworth: Global Science Books), 629–639.

Werner, J. D., Borevitz, J. O., Warthmann, N., Trainer, G. T., Ecker, J. R., Chory, J., et al. (2005). Quantitative trait locus mapping and DNA array hybridization identify an FLM deletion as a cause for natural flowering-time variation. *Proc. Natl. Acad. Sci. U S A.* 102 (7), 2460–2465. doi: 10.1073/pnas.0409474102

Xi, J., Chen, Y., Nakashima, J., S-m, W., and Chen, R. (2013). Medicago truncatula esn1 defines a genetic locus involved in nodule senescence and symbiotic nitrogen fixation. *Mol. Plant Microbe Interact.* 26 (8), 893–902. doi: 10.1094/MPMI-02-13-0043-R

Young, N. D., Debelle, F., Oldroyd, G. E., Geurts, R., Cannon, S. B., Udvardi, M. K., et al. (2011). The medicago genome provides insight into the evolution of rhizobial symbioses. *Nature.* 480 (7378), 520–524. doi: 10.1038/nature10625

Yu, H. X., Xiao, A. F., Wu, J. S., Li, H. X., Duan, Y., Chen, Q. S., et al. (2023). GmNAC039 and GmNAC018 activate the expression of cysteine protease genes to promote soybean nodule senescence. *Plant Cell*, koad129. doi: 10.1093/plcell/koad129

Zipfel, C., and Oldroyd, G. E. (2017). Plant signalling in symbiosis and immunity. *Nature.* 543, 328–336. doi: 10.1038/nature22009

frontiers | Frontiers in Plant Science

TYPE Original Research
PUBLISHED 23 August 2023
DOI 10.3389/fpls.2023.1195479

OPEN ACCESS

EDITED BY
Zhenbin Hu,
Agricultural Research Service (USDA),
United States

REVIEWED BY
Sebastian Beier,
Forschungszentrum Jülich GmbH,
Germany
Manosh Kumar Biswas,
University of Leicester, United Kingdom

*CORRESPONDENCE
Yuesen Yue
yysen2008@sina.com
Xifeng Fan
fanxifengcau@163.com

†These authors have contributed equally to this work

RECEIVED 28 March 2023
ACCEPTED 28 July 2023
PUBLISHED 23 August 2023

CITATION
Teng K, Guo Q, Liu L, Guo Y, Xu Y, Hou X, Teng W, Zhang H, Zhao C, Yue Y, Wen H, Wu J and Fan X (2023) Chromosome-level reference genome assembly provides insights into the evolution of *Pennisetum alopecuroides*. *Front. Plant Sci.* 14:1195479. doi: 10.3389/fpls.2023.1195479

Chromosome-level reference genome assembly provides insights into the evolution of *Pennisetum alopecuroides*

Ke Teng†, Qiang Guo†, Lingyun Liu, Yidi Guo, Yue Xu, Xincun Hou, Wenjun Teng, Hui Zhang, Chunqiao Zhao, Yuesen Yue*, Haifeng Wen, Juying Wu and Xifeng Fan*

Institute of Grassland, Flowers, and Ecology, Beijing Academy of Agriculture and Forestry Sciences, Beijing, China

Pennisetum alopecuroides is an important forage grass resource, which plays a vital role in ecological environment improvement. Therefore, the acquisition of *P. alopecuroides* genome resources is conducive to the study of the adaptability of *Pennisetum* species in ecological remediation and forage breeding development. Here we assembled a *P. alopecuroides* cv. 'Liqiu' genome at the chromosome level with a size of approximately 845.71 Mb, contig N50 of 84.83Mb, and genome integrity of 99.13% as assessed by CEGMA. A total of 833.41-Mb sequences were mounted on nine chromosomes by Hi-C technology. In total, 60.66% of the repetitive sequences and 34,312 genes were predicted. The genomic evolution analysis showed that *P. alopecuroides* cv. 'Liqiu' was isolated from Setaria 7.53–13.80 million years ago and from Cenchrus 5.33–8.99 million years ago, respectively. The whole-genome event analysis showed that *P. alopecuroides* cv. 'Liqiu' underwent two whole-genome duplication (WGD) events in the evolution process, and the duplication events occurred at a similar time to that of *Oryza sativa* and *Setaria viridis*. The completion of the genome sequencing of *P. alopecuroides* cv. 'Liqiu' provides data support for mining high-quality genetic resources of *P. alopecuroides* and provides a theoretical basis for the origin and evolutionary characteristics of *Pennisetum*.

KEYWORDS

***P. alopecuroides*, *de novo* assembly, comparative genomics, whole-genome duplication, forage grass**

1 Introduction

Pennisetum, as an excellent forage, has been widely used all over the world (Hou et al., 2022). The main cultivated species in China are perennial elephant grass (*P. purpureum*), pearl millet [*Pennisetum glaucum* (L.) R. Br.], and hybrid *Pennisetum* (*P. glaucum* × *P. purpureum*). Pennisetum plants have strong stress resistance, moderate salt and alkaline

resistance, low water resource consumption, and landscape value (Yue et al., 2020). It could also be used in soil and water conservation, sand fixation, and landscape construction, which means that it has a good ecological restoration and application prospect.

With the rapid improvement of sequencing technology, whole-genome sequencing information have been obtained from more and more species (Marks et al., 2021). At present, whole-genome sequencing has been completed in lots of plants (Zheng et al., 2021). Traditional horticultural plant genomes, such as *Helianthus annuus* (Badouin et al., 2017), *Magnolia biondii* (Dong et al., 2021), *Prunus yedoensis* (Baek et al., 2018), *Phalaenopsis* (Cai et al., 2015), *Dendrobium officinale* (Liang et al., 2015), and *Chrysanthemum nankingense* (Song et al., 2018), have been obtained with high quality. The completion of horticultural plant genome sequencing lays a foundation for the analysis of plant type, flowering, growth and development, and other regulatory mechanisms so as to achieve accurate breeding. In recent years, genome assembly and annotation in Gramineae have been enriched (International Brachypodium, I, 2010; Byrne et al., 2015; Lovell et al., 2018; Huang et al., 2019; Mitros et al., 2020; Wang et al., 2021). The quality of chromosome assembly has also been improved in species closely related to Pennisetum. *Setaria italica* has completed chromosome assembly, which contig N50 increased to 11.2 Mb compared with the prior 1.6 Mb (Bennetzen et al., 2012; Mamidi et al., 2020). The elephant grass (*Cenchrus purpureus*) ($2n = 4x = 28$, AABB) was sequenced with contig N50 of 2.65 Mb higher than the first published genome of *C. purpureus* cv. Purple (contig N50 of 1.83 Mb) (Yan et al., 2020; Zhang et al., 2022a). In addition, according to the whole-genome sequence of *Pennisetum glaucum* (L.) R. Br., syn. *Cenchrus americanus* (L.) Morrone ($2n = 2x = 14$, AA), NAC transcription factors were found to play an important role in drought tolerance (Dudhate et al., 2021; Yan et al., 2023). Recently, a predominantly Pennisetum multiple omics database platform (Milletdb, http://milletdb.novogene.com/) was established (Sun et al., 2023). However, the genome still needs to increase the number of grass plants to further understand the evolutionary history and find more resistance candidate genes to improve breeding.

Pennisetum is a kind of forage (Tenorio and Roque, 2015), food crop (Xu et al., 2021), energy grass (Hayat, 2020), and ecological grass (Gallego et al., 2015). *Pennisetum* is broadly used as an ecological grass around the world (Zhang et al., 2016), such as *Pennisetum alopecuroides* cv. Rubrum (Liu et al., 2022) and *Pennisetum alopecuroides* cv. 'Little Bunny' (Guo et al., 2022). In this study, *P. alopecuroides* cv. 'Liqiu' was selected for genome assembly, which has several advantages, including longer green periods, outstanding cold resistance, and better palatability. It could be used for ecological restoration or landscape construction of barren mountains and slopes, and it could also be used for landscaping.

Here we used PacBio, Illumina, and Hi-C auxiliary assembly strategies to assemble a high-quality genome. The reference genome could provide the basis for functional genomics and improve the molecular basis. In addition, candidate functional genes could be identified to reveal the adaptation mechanism of *Pennisetum* to different adversities.

2 Materials and methods

2.1 Sampling collection and sequencing

P. alopecuroides cv. 'Liqiu' (Figure 1A) was identified and cultivated at the Institute of Grassland, Flowers, and Ecology, Beijing Academy of Agriculture and Forestry Sciences (Beijing, China). Healthy leaves were collected from the plants of varieties growing in National Precision Agriculture Research Demonstration Base, Changping District, Beijing, China (116°28′ E, 39°94′ N). We collected and propagated plants, planted seedling in 20-cm-diameter, 20-cm-deep plastic pots filled with TS1 nutrient medium (Klasmann-Deilmann, Germany). Fresh tissues of *P. alopecuroides* cv. 'Liqiu' (roots, stems, leaves, and spike) were mixed and sampled for PacBio full-length cDNA sequencing.

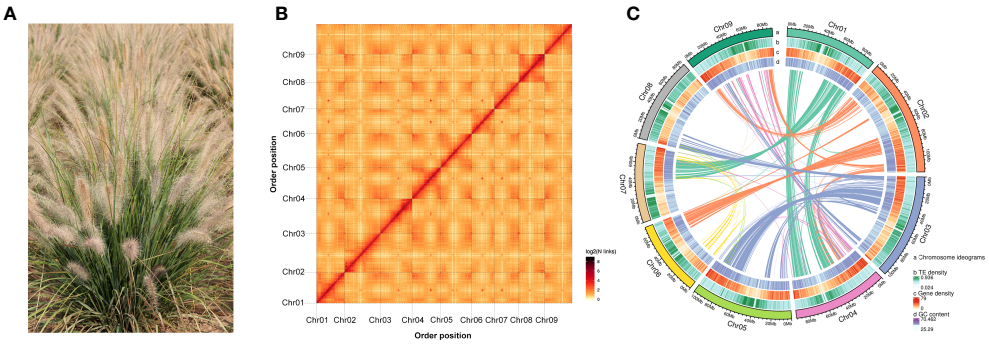


FIGURE 1
Genome assembly and situation of chromosomal collinearity. **(A)** Photograph of *P. alopecuroides* morphology. **(B)** A heat map of chromosome interactions in the *P. alopecuroides* genome shows the interactions between nine chromosomes (Chr1–9). Both horizontal and vertical coordinates represent the order of each bin on the corresponding chromosome group. **(C)** *P. alopecuroides* genome circle map. The outermost layer shows the number of chromosomes, from outside to inside, gene density, repeat sequence density, GC content, and collinearity.

Total RNA was extracted using the Plant RNA Kit (OMEGA, GA, USA, no. R6827-01). The quantity and integrity of RNA samples were assessed in the NanoDrop ND-1000 spectrophotometer (NanoDrop Technologies, DE, USA) and in the 2100 Bioanalyzer (Agilent Technologies, CA, USA), respectively. Qualified RNA samples were then used for constructing PacBio full-length cDNA libraries.

2.2 Library construction and genome sequencing

Genomic DNA was extracted by using DNAsecure Plant Kit (Tiangen, Beijing, China) from fresh young leaves of *P. alopecuroides* cv. 'Liqiu' which were harvested and frozen immediately in liquid nitrogen. For Illumina sequencing, a library with an insert size of 350 bp was constructed and sequenced using an Illumina Novaseq 6000 platform (Illumina, San Diego, CA, United States) to produce a short-read sequencing library (Hu et al., 2022). Short reads generated from the Illumina platform were used for the estimation of the genome size, the level of heterozygosity, and the repeat content of the genome, and long reads from the PacBio platform were used for genome assembly. The short-reads from Illumina platform were quality-filtered by HTQC v1.92.310 (Yang et al., 2013) using the following method. First, the adaptors were removed from the sequencing reads. Second, read pairs were excluded if any one end has an average quality lower than 20. Third, the ends of reads were trimmed if the average quality is lower than 20 in the sliding window size of 5 bp. Finally, read pairs with any end that was shorter than 75 bp were removed. The quality-filtered reads were used for genome size estimation. For PacBio sequencing, 8–20-kb fragment library was used for multi-round sequencing of a single fragment to improve the accuracy. The constructed library was sequenced for third-generation sequencing by PacBio Sequencing II platform (Pacific Biosciences of California, Menlo Park, CA, USA). After reversing the crosslinks, we constructed Hi-C fragment libraries from 300- to 700-bp insert size and performed sequencing through the Illumina platform (Rao et al., 2014).

2.3 Hi-C technology helps anchor contigs

We constructed Hi-C fragment libraries ranging from 300- to 700-bp insert size as illustrated in Rao et al. (2014) and performed sequencing through the Illumina platform. For anchored contigs, clean read pairs were generated from HiC-Pro v2.10.0 (Servant et al., 2015) and were mapped to the polished genome using BWA v.0.7.10-r789 (Li and Durbin, 2009) with the default parameters. Only uniquely aligned pair reads whose mapping quality was more than 20 were retained for further analysis. Invalid read pairs, including dangling-end and self-cycle, re-ligation, and dumped products, were filtered by HiC-Prov2.10.0. Paired reads with mate mapped to a different contig were used to do the Hi-C-associated scaffolding. Self-ligation, non-ligation, and other invalid reads, such as Start NearRsite, PCR amplification, random break, large and small fragments, and extreme fragments, were filtered (Burton et al., 2013). Then, the contigs were clustered into groups with the agglomerative hierarchical clustering method in Lachesis. Lachesis software (Burton et al., 2013) was further applied to order and orient the clustered contigs. The parameters were CLUSTER_MIN_RE_SITES = 86, CLUSTER_MAX_LINK_DENSITY = 2, ORDER_MIN_N_RES_IN_TRUNK = 72, and ORDER_MIN_N_RES_IN_SHREDS = 69. After Hi-C error correction, redundancy removal, and clustering and ordering the contigs, we obtained the final version of the genome.

2.4 Annotation of repetitive sequences

Transposon element (TE) and tandem repeats were annotated by the following workflows. TEs were identified by a combination of homology-based and *de novo* approaches. We first customized a *de novo* repeat library of the genome using RepeatModeler (http://www.repeatmasker.org/RepeatModeler/), which can automatically execute two *de novo* repeat finding programs, including RECON and RepeatScout (Bao and Eddy, 2002; Price et al., 2005). Then, full-length long-terminal repeat retrotransposons (fl-LTR-RTs) were identified using both LTRharvest (v1.5.9) (Ellinghaus et al., 2008) and LTR_finder (v2.8) (Xu and Wang, 2007). The high-quality intact fl-LTR-RTs and non-redundant LTR library were then produced by LTR_retriever (Ou and Jiang, 2018). The non-redundant species-specific TE library was constructed by combining the *de novo* TE sequences library above with the known Dfam (v3.5) database. Final TE sequences in the *P. alopecuroides* genome were identified and classified by homology search against the library using RepeatMasker (v4.12) (Chen, 2004). Tandem repeats were found by Tandem Repeats Finder (TRF 409) (Benson, 1999) and MIcroSAtellite identification tool (MISA v2.1) (Beier et al., 2017).

2.5 Gene prediction

We integrated three approaches, namely, *de novo* prediction, homology search, and transcript-based assembly, to annotate protein-coding genes in the genome. The *de novo* gene models were predicted using two *ab initio* gene prediction software tools, Augustus (v3.1.0) (Stanke et al., 2008) and SNAP (2006-07-28) (Korf, 2004). For the homolog-based approach, GeMoMa (v1.7) (Keilwagen et al., 2016) software was employed by using reference gene models from the other six species, which included *O. sativa*, *C. purpureus*, *S. italica*, *A. thaliana*, *C. americanus*, and *S. viridis*. For the transcript-based prediction, RNA sequencing data were mapped to the reference genome using Hisat (v2.1.0) (Kim et al., 2015) and assembled by Stringtie (v 2.1.4) (Pertea et al., 2015). GeneMarkS-T (v5.1) (Tang et al., 2015) was used to predict genes based on the assembled transcripts. The PASA (v2.4.1) (Haas et al., 2003) software was used to predict genes based on the unigenes [and full-length transcripts from the PacBio (ONT) sequencing] assembled by Trinity (v2.11) (Grabherr et al., 2011). Gene models from these different approaches were combined using the EVM software (v1.1.1) (Haas et al., 2008) and updated by PASA.

2.6 Annotation of protein-coding genes

Gene functions were inferred according to the best match of the alignments to the National Center for Biotechnology Information (NCBI) Non-redundant (NR) (20200921, ftp://ftp.ncbi.nlm.nih.gov/blast/db), EggNOG (5.0, http://eggnog5.embl.de/download/eggnog_5.0/) (Huerta-Cepas et al., 2019), GO, (20200615, http://geneontology.org), Kyoto Encyclopedia of Genes and Genomes (KEGG) (20191220, http://www.genome.jp/kegg) (Kanehisa et al., 2016) with an E-value threshold of 1E-5, Pfam (v33.1, http://pfam.xfam.org) (Finn et al., 2006), and SwissProt (202005, http://ftp.ebi.ac.uk/pub/databases/swissprot) (Boeckmann et al., 2003).

2.7 Annotation of non-coding RNAs

Non-coding RNAs of tRNA were identified using scan-SE (version 1.3.1) (Lowe and Eddy, 1997). Identification of the rRNA genes was conducted by barrnap (v0.9) (Benson, 1999) based on Rfam (v14.5) (Griffiths-Jones et al., 2005). miRNA and snRNA genes were predicted using infernal (v1.1) (Nawrocki and Eddy, 2013) against the Rfam (v14.5) database with default parameters.

2.8 Pseudogene prediction

Pseudogenes usually have similar sequences to functional genes but may have lost their biological function because of some genetic mutations, such as insertion and deletion. The GenBlastA (v1.0.4) (She et al., 2009) program was used to scan the whole genomes after masking predicted functional genes. Putative candidates were then analyzed by searching for premature stop mutations and frame-shift mutations using GeneWise (v2.4.1) (Birney et al., 2004).

2.9 Detection of methylation sites

P. alopecuroides cv. 'Liqiu' was sequenced by PacBio HiFi (CCS), with a total data volume of 41.92 Gb. Based on the quality control of CCS methylation site test data, using SMRTLINK (V11.0) (https://www.pacb.com/support/software-downloads), high-quality CCS reads were obtained. At the same time, kinetic signals were obtained by combining the parameters—all kinetics and kinetic signals in CCS reads were converted into methylation states by the primrose (https://github.com/mattoslmp/primrose) (called 5mC in CpG motifs) program. The 5mC methylation information was stored in the tag tags (MM and ML) of bam files. The CpG/5mC information was analyzed from the above-mentioned data using pb-CpG-tools (v1.0.0) (Ni et al., 2021) software. Each chromosome was divided into multiple Windows with a size of 100 kb, and the average methylation level of each window was calculated.

2.10 Phylogenetic tree reconstruction

Using the obtained single-copy gene family sequences, IQ-TREE (v1.6.11) (Nguyen et al., 2015) software was used to construct the evolutionary tree. Specifically, MAFFT (v7.205) (Katoh et al., 2009) was used to align each single-copy gene family sequence, and then PAL2NAL (v14) (Suyama et al., 2006) was used to convert the aligned protein sequence (protein alignment) into codon alignment (codon alignment). Then, Gblocks (v0.91b) (Talavera and Castresana, 2007) (parameter: -b5 = h) was used to remove regions with poor or large sequence alignment. All the well-ordered gene families of each species were connected end to end, and then each species ends up with a supergene end to end. The optimal model is JTT+F+I+G4 based on the model checking tool ModelFinder (Kalyaanamoorthy et al., 2017). Then, the optimal model was used to construct the evolutionary tree through the maximum likelihood (ML) method, where the bootstrap times were set to 1,000. For the above-mentioned evolutionary tree, we set the outgroup as *A. trichopoda* to obtain a rooted tree and then used the software package MCMCTREE that comes with PAML v4.9i (Yang, 1997) software to calculate the divergence time. The sequences of the single-copy gene families of 16 species were aligned, and a rooted tree was obtained. *A. trichopoda* vs. *S. italica* fossils were obtained at 179–199.1 mya, *S. italica* vs. *S. viridis* fossils at 0.68 to 2.11, and *C. americanus* vs. *C. purpureus* at 3.35–5.01 mya, *S. bicolor* vs. *M. lutarioriparius* at 2.88–8.02 mya, and *O. sativa* vs. *S. italica* at 40.3 to 51.9 mya. The divergence times were calculated using the package MCMCTREE in PAML v4.9i software.

2.11 Genome-wide replication event analysis

Using diamond v0.9.29.130 (Buchfink et al., 2015), we compared the collinear regions in the genomes of *P. alopecuroides* and *O. sativa* and in *S. viridis* to determine similar gene pairs. MCScanX (Wang et al., 2012) software was used to determine whether similar genes were adjacent on chromosomes based on gene location information files. The proportion of base conversion mutations at the fourfold synonymous (degenerative) third-codon transversion (4DTv) locus in each homologous gene was calculated through https://github.com/JinfengChen/Scripts and then corrected 4DTv using the HKY substitution model to obtain the distribution map. The peak was identified as the whole-genome duplication event in this species.

3 Results

3.1 Genome assembly, quality validation, and annotation

A 350-bp library was constructed using the genomic DNA of the sample, and 106.72 Gb of high-quality data was obtained by Illumina sequencing platform. The total sequencing depth was

approximately ×106.55, and the proportion of Q20 was more than 98.82%. The proportion of Q30 reached 93.49%, and the GC content was approximately 44.78%. After the quality assessment of *P. alopecuroides* genome, the average depth of k-mer constructed was 77 (Supplementary Figure S1A). The calculated estimated genome length was approximately 1.00 Gb. According to the distribution of k-mer, the content of repeat sequence was estimated to be approximately 74.46%, and the heterozygosity was estimated to be approximately 0.43%. The data volume of ccs sequencing was 41.92 Gb, and the depth was approximately ×41. The average length of ccs produced by statistics was more than 17,995 bp, and the longest ccs length is 50,190 bp (Supplementary Table S1). The maximum number of ccs is around 18,000 bp in length (Supplementary Table S2). Using Hifiasm (Cheng et al., 2021) splice, the total length of the contig sequence was 845.71 Mb, the contig N50 was 84.83 Mb, and the scaffold N50 was 91.33 Mb (Supplementary Table S4). The CEGMA evaluation completeness was 99.13% (Supplementary Table S3). OrthoDB 10 embryophyta database was selected, and the completeness of BUSCO evaluation was 98.45% (Supplementary Table S3). The short sequence obtained from the second-generation sequencing data was matched with the assembled genome, and 99.22% of the sequences were matched correctly. The response ratio of the third generation of reads was 98.46%, the coverage was 99.99, and the average sequencing depth was 49 (Supplementary Figure S1B).

The Hi-C library was sequenced by the Illumina high-throughput sequencing platform. A total of 125.30 Gb of clean data was generated. A total of 342,693,530 uniquely aligned genomic reads were obtained, of which 142,777,564 were valid Hi-C data, accounting for 41.66% of uniquely aligned genomic data. After Hi-C assembly and manual adjustment of heat map, 833.41 Mb of genome sequence was located on nine chromosomes, accounting for 98.55% (Figure 1B). The longest (115.63 Mb) and shortest (74.04 Mb) chromosomes were chr2 and chr6, respectively (Supplementary Table S5). The TE density, gene density, and GC content of every chromosome are shown together on a circos plot (Figure 1C). The 830.09-Mb sequence that could determine direction and sequence information accounts for 99.6% of the total length of sequences located on chromosomes.

3.2 Repeat annotation

We identified a total of 512,992,680 bp TE sequence, which accounted for 60.66% of the genome. The specific non-LTR-RTs mainly consisted of long interspersed nuclear elements (LINEs) and short interspersed nuclear elements (SINE). Based on the analysis of transposon elements in the repetitive sequence in the *P. alopecuroides* cv. 'Liqiu' genome, we found that LINE, SINE, and DNA transposons accounted for 0.92%, 0.13%, and 8.87%, respectively (Supplementary Table S6). The LTR-RTs accounted for 50.73%, while the most abundant LTRs were *Copia* elements, making up 30.47% of the genome, followed by *Gypsy* elements (8.47%). About 45,931,763 bp of tandem repeats sequence were obtained, accounting for 5.43% of the genome (Supplementary Table S7).

3.3 Gene prediction and functional validation

A total of 34,312 genes were obtained by homology, *de novo*, and transcriptome prediction (Supplementary Figure S2, Supplementary Table S8). The average lengths of the genes and CDS are 1,266.31 and 1,286.37 bp, respectively, and there are 4.95 exons in each gene with a length of 1,400.87 bp per exon (Figure 2A, Supplementary Table S9). Among them, 19,951 gene models were supported by all three lines of evidence. The genetic characteristics were more closely related to the genetic relations of the species. Therefore, compared with the genetic characteristics of the closely related species, we observed the quality of our prediction. The length of most genes is distributed around 1,600 bp, and that of the coding genes is around 1,000 bp (Figures 2B, C; Supplementary Table S12). The distribution of intron length and mRNA length is similar to those of other reference grasses (Figure 2D). The BUSCO database of embryophyta contains 1,614 conserved core genes. We used BUSCO (v5.2.2) software to evaluate the integrity of gene prediction, of which 98.51% of BUSCO genes existed in the genes, indicating the high integrity of gene prediction. We predicted 824 tRNAs, 7,114 rRNAs, and 233 miRNAs, which were non-coding RNAs. In addition, we predicted 276 pseudogenes. A total of 97.5% of genes were annotated based on the prediction of NR, eggNOG, GO, KEGG, KOG, and Pfam databases (Supplementary Figure S5, Supplementary Table S10). The highly abundant tandem repeats dispersed on chromosomes with high TE density like chr1, chr2, chr5 and chr7. The distribution of TE density across chromosomes for chr1–6, chr8, and chr9 shows a tendency for higher density in centromeric and less on pericentromeric. However, chr7 density is high at the beginning of the chromosome and then gradually decreases.

3.4 Detection of methylation sites

A total of 37,729,477 5mC sites were obtained, among which the number of haplotype 1 and 2 methylation sites was 18,363,029 and 14,959,046, respectively (Figure 3A). The distribution of methylation sites across the nine chromosomes was obtained from the average level calculated from the partitioned 100-kb window. The gene density in the window showed high telomere density and low centromere density. The methylation levels and distribution of haplotypes 1 and 2 were similar. The overall methylation level is high. According to the detected methylation sites, chr2, chr6, chr7, and chr8 were found to have concentrated methylation regions in the centromere (Figure 3B).

In the 1,000 5mC sites with the highest methylation level, CG was found to be highly conserved in the analysis of the sequence characteristics of the surrounding 9 bp (Figure 3C). The differentially methylated regions (DMRs) were annotated to different gene regions, and 58.09% were found in distal intergenic regions. It is followed by promoter (≤1 kb) and promoter (1 to 2 kb) (Figure 3D). The DMR annotated to the non-intergenic region of the gene was used for KEGG enrichment. Most of them were

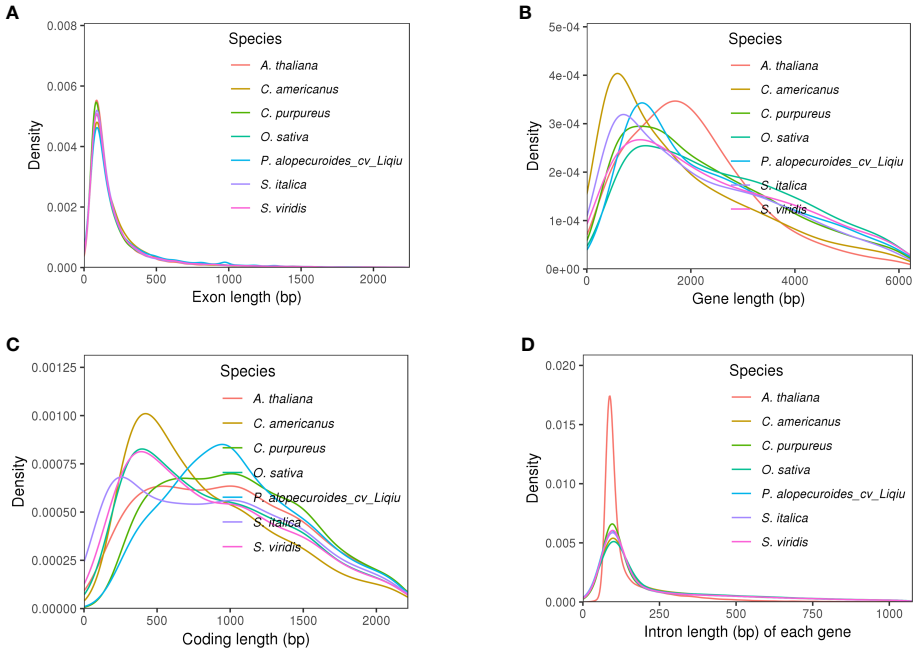


FIGURE 2
Genetic structural characteristics of genomic prediction. **(A)** Extron length of five species, the trend of distribution of which was similar. **(B)** Gene length of five species, and *P. alopecuroides* had the highest density at around 1,750 bp compared with other plants. **(C)** Coding length of five species. **(D)** Intron length of five species, the trend of distribution of which was similar, and *P. alopecuroides* density was relatively small.

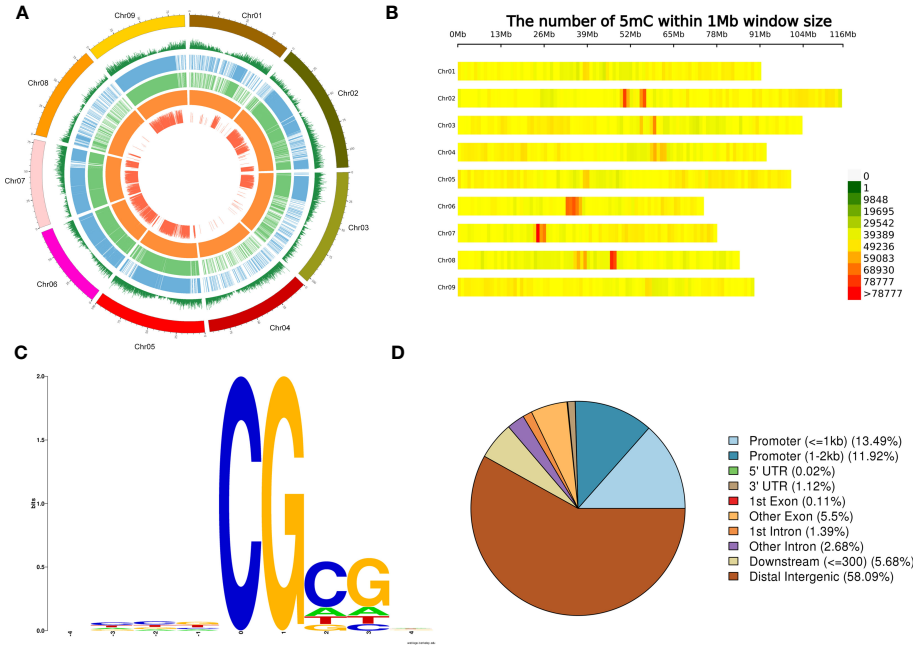


FIGURE 3
Genome-wide methylation levels and distribution of methylation on chromosomes. **(A)** Genome-wide methylation levels. From outside to inside, respectively, of a chromosome: gene density window (the higher level represented a greater number of genes), level 1 of methylation haplotype, level 2 of methylation, haplotype whole-genome methylation level differentially methylated regions, single size or samples. **(B)** Density distribution of methylation sites on each chromosome according to the 1-Mb window. The horizontal coordinate is chromosome length, and each band represents a chromosome. The genome is divided according to the size of 1 Mb. The more methylation sites in each window, the darker the color, and the fewer methylation sites, the lighter the color. The darker the area in the diagram is the one where the methylation sites are concentrated. **(C)** Nine-base-pair sequence distribution around site 5mC. The horizontal coordinate is the relative position of the methylation site, the total height of each position is the sequence conservation of the base at that position, and the height of the base signal represents the relative frequency of the base at that position. **(D)** Functional annotation and enrichment of a differentially methylated region.

enriched to RNA polymerase, followed by tryptophan metabolism (Supplementary Figure S4).

3.5 Gene family analysis and genome evolution

An analysis of gene families in 16 species was performed and annotated, and finally GO and KEGG enrichment analyses were performed for the gene families unique to this species. We obtained 50,510 gene families, and 2,627 gene families were common families (Supplementary Figure S5A). In addition, the gene family information of *P. alopecuroides* cv. 'Liqiu' and *S. viridi*, *S. italica*, *C. purpureus*, and *C. americanus* were analyzed (Figure 4A; Supplementary Figure S5B). *P. alopecuroides* contains a total of 22,007 gene families and 34,312 genes, among which 615 are unique gene families. The unique gene family was obviously enriched in RNA polymerase (Supplementary Figure S6). As for the copy types of genes, the single copy genes accounted for the largest proportion in *P. alopecuroides* (Supplementary Figure S7). From these 16 species, we identified 774 one-to-one single-copy genes that were used to construct a maximum likelihood (ML) tree to show the evolutionary relationships (Figure 4B; Supplementary Table S11). We set the outgroup to *A. trichopoda* and get a rooted tree. According to the phylogenetic tree, *P. alopecuroides* cv. 'Liqiu' separated from Poaceae approximately 7.53–13.8 million years ago, separating from *C. americanus* and *C. purpureus* approximately 5.33–8.99 million years ago. *C. americanus* and *C. purpureu* separated around 3.35–4.98 million years ago.

Through the analysis of expansion and contraction of gene families, it was found that 637 gene families contracted and 223 gene families expanded during the evolution of *P. alopecuroides* (Supplementary Figure S8). The expansion and contraction gene families of this species were analyzed by clusterProfile for GO and KEGG enrichment. The GO enrichment analysis revealed that the contracted gene family was significantly enriched in biological

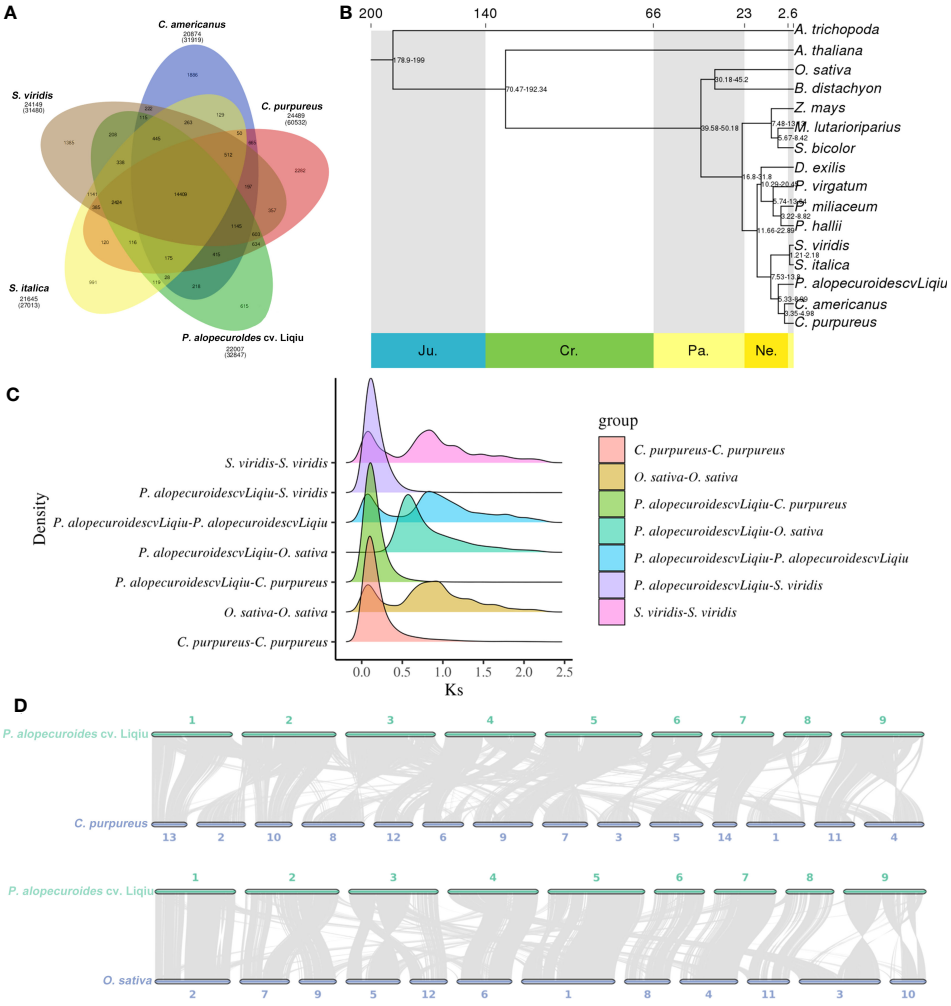


FIGURE 4
Comparative and evolutionary genomic analysis of *P. alopecuroides*. **(A)** Gene family clustering petal map. The number below the species name is the total number of gene families, the corresponding number of genes is in parentheses, and the number of gene families is in the Venn diagram. **(B)** Evolutionary tree with time for differentiation. **(C)** Distribution of average synonymous substitution levels (Ks) between syntenic blocks. Peaks in maps made by interspecies combinations represent species divergence, and peaks produced by assembly-made maps within species represent genome-wide replication events within the species. **(D)** Linear collinear graph between *P. alopecuroides* vs. *C. purpureu* and *P. alopecuroides* vs. *O. sativa*..

processes related to defense response (Supplementary Figure S9A), ADP binding (Supplementary Figure S9B), and extracellular region (Supplementary Figure S9C). The expanded gene family is highly enriched in pathways associated with DNA repair (Supplementary Figure S9D), ribosome (Supplementary Figure S9E), and DNA helicase activity (Supplementary Figure S9F). The KEGG enrichment analysis showed that the contracted gene family was significantly enriched in the biological processes related to flavonoid biosynthesis (Supplementary Figure S10A). The expanded gene family was obviously enriched in "tryptophan metabolism" and "ribosome" (Supplementary Figure S10B).

3.6 Positive selection and collinearity analysis of *P. alopecuroides*

Single-copy gene families in gene family expansion and contraction analysis were extracted, and the CodeML module in PAML was used for positive selection analysis, resulting in 80 positive selection genes in *P. alopecuroides*. The KEGG enrichment analysis of the 80 positively selected genes and the positively selected genes were enriched in "ABC transporters" (Supplementary Figure S11).

The GO enrichment analysis mainly divided 264 positive selections into three categories: biological processes, cell components, and molecular functions (Supplementary Figure S12A). Among them, some genes in the biological process mainly play a role in "DNA repair" (Supplementary Figure S12B). The number of genes in the "ribosome" is the highest in the cellular component (Supplementary Figure S12C). The molecular functions mainly focus on enzyme activity, including "DNA helicase activity" and "nucleotide binding" (Supplementary Figure S12D).

3.7 Whole-genome duplication and collinearity analysis in *P. alopecuroides*

Genome-wide replication events are the major driving force in species divergence. To estimate the timing of WGD events in the *P. alopecuroides* genome, we described synonymous substitution rates at synonymous nucleotide sites (Ks) between *P. alopecuroides* and three other species, including *O. sativa*, *S. viridis*, and *C. purpureus* (Figure 4C). The peaks generated by combinatorial mapping within species represent genome-wide duplication events in this species, which occurred twice during the evolution of *P. alopecuroides*. They occurred when Ks was 0.85 and 0.10, respectively. According to the differentiation time calculation formula: time = Ks/2μ, $\mu = 6.1 \times 10^{-9}$. *P. alopecuroides* was separated from *O. sativa* after the first WGD event and before the most recent. Similarly, *O. sativa*, *S. viridis*, and *P. alopecuroides* have similar peaks, indicating that they both experienced two WGD events and that they occurred at similar times. *C. purpureus* had only one WGD event. The evolutionary process of *S. viridis* was closer to *P. alopecuroides*. These two genome-wide duplication events occurred approximately 69.67 and 8.20 mya. We confirmed that two WGD events had occurred during the evolution of *P. alopecuroides* based on 4DTv method (Supplementary Figure S13A).

The collinearity of chromosomes was obtained by mapping the sequences of *C. americanus*, *C. purpureus*, *O. sativa*, and *S. viridis* to *P. alopecuroides* chromosomes (Figure 4D, Supplementary Figure S13B). *C. purpureus* had the highest collinearity ratio with *P. alopecuroides* (63.12%) and the highest number of collinearity genes (62,987). The number of collinear genes of *S. viridis* was 40,048 (55.42%). After screening, the number of collinearity blocks in non-major corresponding chromosomes was less than 100, and the linear collinearity between *P. alopecuroides* and the four species was obtained. The corresponding collinear gene numbers were as follows: *C. americanus*—170, *C. purpureus*—250, *O. sativa*—925, and *S. viridis*—242. Based on JCVI-filtered information (Tang et al., 2008), the distribution of collinear genes on the chromosomes was provided. Among the collinearity of *S. viridis* and *P. alopecuroides*, chr2, chr5, chr6, chr7, chr8, and chr9 were observed to have a large collinear block, and the chromosomes were highly conserved. chr1, chr3, and chr4 undergo chromosomal rearrangement events. In comparison with rice, all the genetic information of chr7 chromosome comes from chr4 of rice. chr2 is mainly the fusion of chr7 and chr9 of rice, chr3 is mainly the fusion of chr5 and chr12, and chr5 is mainly the fusion of chr1 and chr5.

3.8 Analysis of the LTR insertion time

The analysis of LTR insertion time of five species showed that the insertion time of LTR-RTs in *P. alopecuroides* was different from that in the other four species, but three of them were relatively late. The analysis of insertion time suggested that LTR insertion was a continuous process and that continuous insertion of LTR may be a general phenomenon in angiosperms (Figure 5). The proliferation of LTR-RTs in *P. alopecuroides* (~4.54 mya) and *O. sativa* (~4.08 mya) was earlier than in *S. viridis* (~7.16 mya) but later than in *C. purpureus* (~54.23 mya) and *C. americanus* (~34.79 mya). We observed that LTR insertion was a continuous process that surges at some time, which may have negatively affected the growth of grasslands and other types of vegetation during the Pleistocene when cold weather and limited global atmospheric CO_2 (180 ppm) were present. To adapt to the stressed abiotic environment, most plants activate TEs to cope with environmental stress. These conditions may have led to plant genome recombination during this period. The insertion time of LTR also spikes after the recombination event in plants (Huang et al., 2019). These results suggest that most of the LTR-RTs in these four genomes were recently inserted, and these insertions occurred well after the divergence of these species.

4 Discussion

Pennisetum ecological grasses are increasingly popular because of their elegant ecology form, strong stress resistance, and low maintenance. However, due to the complex genomes of these grasses with high heterozygosity and high amounts of repetitive sequences, there is currently a lack of reference genomes, which hinders the genetic studies of these grasses. Here we assembled the

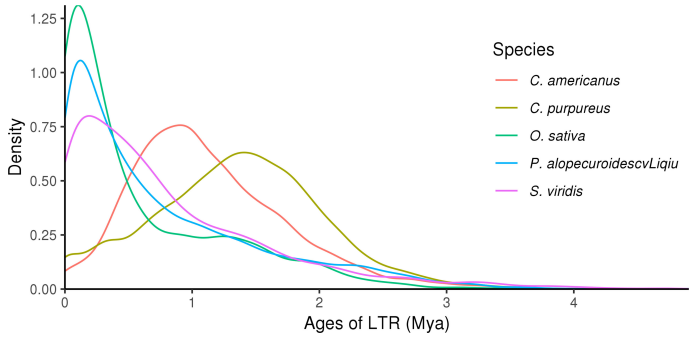


FIGURE 5
Long terminal repeat insertion time analysis.

P. alopecuroides cv. 'Liqiu' genome at the chromosomal level with a size of approximately 845.71 Mb, contig N50 of 84.83 Mb, and genome integrity of 99.13% as assessed by CEGMA. The N50 of this reference genome (contig N50 84.83 Mb) was much higher than that of forage perennial ryegrass genome (contig N50 1,637 kb). Meanwhile, it is better than other sequenced genomes of grasses such as *P. purpureus* (Zhang et al., 2022a), *C. purpureus* (contig N50 of 1.83 Mb) (Yan et al., 2020), *C. americanus* (contig N50 of 1.82 kb) (Varshney et al., 2017), *Dactylis glomerata* L. (contig N50 of 0.93 Mb) (Huang et al., 2019), and other grasses such as *Achnatherum splendens* (contig N50 of 40.30 Mb) (Ren et al., 2022) and *Eremochloa ophiuroides* (contig N50 of 40.30 Mb) (Wang et al., 2021). A total of 833.41 Mb sequences were located on nine chromosomes by Hi-C technology.

Polyploidy or whole-genome duplication events provided the original genetic material for biological evolution, and duplication of plant genomes would be extremely important for plant evolution with rapid recombination, loss of a large number of genes, and increase of structural variation (D'Hont et al., 2012; Shi et al., 2022). The early WGD events that the Gramineae collectively experienced, approximately 60 mya, marked the end of the Cretaceous period (Jiao et al., 2014). The replication genes retained after this stage are associated with heat and cold stress, osmotic pressure maintenance, salt stress, water acquisition, and wound repair (Jiao et al., 2014; Gao et al., 2018). The genome analysis showed that the last WGD event occurred in *P. alopecuroides* after differentiation from other monocotyledonous plants, and the corresponding family and genus had already differentiated, which was estimated to have occurred at around 8.2 mya. Ancient chromosome reconstruction of grasses revealed 12 chromosomes after two WGD events and rearrangement. We also found that *P. alopecuroides* and *Setaria* have similar evolutionary characteristics, providing important sources of genetic, evolutionary, and breeding information.

In addition to polyploidization, TE/LTR activity is an important mechanism for karyotype and genome size evolution in angiosperms (Bennetzen et al., 2005), and it is involved in stress responses and probably important for adaptive evolution in dynamically changing environments (VanBuren et al., 2015). The insertion of LTR-RTs is a common phenomenon in angiosperms, and its number reaches a peak between 0 and 1 mya in the Pleistocene (or ice) age (Huang et al., 2019). In this study, LTR insertion was observed to be a continuous process. Previous studies have shown that TE activity in rice increases the total number of stress-induced genes to help rice adapt to stress (Lisch, 2013). Most terrestrial plants and green algae based on the gypsy type of LTR (Jiao et al., 2020), such as Ty1/*Copia* and Ty3/*Gypsy*, occupied 8.43% and 30.47%, respectively, in the *E. ophiuroides* genome which has 867.43 Mb (Wang et al., 2021) and 7.61% and 22.66% in the hawthorn genome with 823.41 Mb (Zhang et al., 2022b). In *P. alopecuroides*, the proportion of Ty1/*Copia* family was 8.43%, indicating that the Ty1/*Copia* family contributes more to the genomic expansion of *P. alopecuroides* than the Ty3/*Gypsy* family. In addition, we consider that LTR was a major cause of genomic expansion, and differences in LTR activity types and timing of divergence may have led to new functions for *P. alopecuroides* (Jiao et al., 2020). The peak of LTR insertion occurred after the recent WGD event, indicating that the LTR insertion could cooperate with WGD events for chromosome deletion, ectopia, and other rearrangements, thus affecting the *P. alopecuroides* genome size and promoting the adaptation of species to the environment.

5 Conclusion

We constructed a comprehensive and complete high-quality *P. alopecuroides* genome for the first time, which sheds new light on genetic evolution and polyploid breeding of *P. alopecuroides*. We assembled an 833.41-Mb genome located on nine chromosomes, with 84.83-Mb contig N50, higher than the assembly data for other grasses, to provide a reference for the assembly and annotation of grass plants. The phylogenetic analysis revealed that *P. alopecuroides* was a sister group of *Cenchrus* L., and the proliferation of LTR-RTs and WGD event analysis verified that two WGD events occurred in *P. alopecuroides*, which provided a reference genome for *Pennisetum* and laid the foundation for analyzing the molecular mechanism of high yield and stress resistance and mining the functional genes related to the target traits.

Data availability statement

The data presented in the study are deposited in the NCBI repository, accession number PRJNA952524.

Author contributions

YY and XF conceived the study and designed the experiments. JW and WT collected the experimental materials. YG and LL performed genome analyses. KT and QG completed the manuscript with suggestions from HZ, HW, CZ, XH, and YX. All authors contributed to the article and approved the submitted version.

Funding

This study was supported by Scientific Funds of Beijing Academy of Agriculture and Forestry Sciences (KJCX20230119, CZZJ202210) and Scientific Funds of Beijing Innovation Consortium of Agriculture Research System (BAIC09-2023-02). Each of the funding bodies granted the funds based on a research proposal.

Acknowledgments

We acknowledge the Biomarker Corporation (Beijing, China) for the facilities and their expertise of the Illumina platform for library construction and sequencing.

Conflict of interest

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

Publisher's note

Supplementary material

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fpls.2023.1195479/full#supplementary-material

References

Badouin, H., Gouzy, J., Grassa, C. J., Murat, F., Staton, S. E., Cottret, L., et al. (2017). The sunflower genome provides insights into oil metabolism, flowering and Asterid evolution. *Nature* 546 (7656), 148–152. doi: 10.1038/nature22380

Baek, S., Choi, K., Kim, G.-B., Yu, H.-J., Cho, A., Jang, H., et al. (2018). Draft genome sequence of wild Prunus yedoensis reveals massive inter-specific hybridization between sympatric flowering cherries. *Genome Biol.* 19 (1), 127. doi: 10.1186/s13059-018-1497-y

Bao, Z., and Eddy, S. R. (2002). Automated *de novo* identification of repeat sequence families in sequenced genomes. *Genome Res.* 12 (8), 1269–1276. doi: 10.1101/gr.88502

Beier, S., Thiel, T., Münch, T., Scholz, U., and Mascher, M. (2017). MISA-web: a web server for microsatellite prediction. *Bioinformatics* 33 (16), 2583–2585. doi: 10.1093/bioinformatics/btx198

Bennetzen, J. L., Ma, J., and Devos, K. M. (2005). Mechanisms of recent genome size variation in flowering plants. *Ann. Bot.* 95 (1), 127–132. doi: 10.1093/aob/mci008

Bennetzen, J. L., Schmutz, J., Wang, H., Percifield, R., Hawkins, J., Pontaroli, A. C., et al. (2012). Reference genome sequence of the model plant Setaria. *Nat. Biotechnol.* 30 (6), 555–561. doi: 10.1038/nbt.2196

Benson, G. (1999). Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 27 (2), 573–580. doi: 10.1093/nar/27.2.573

Birney, E., Clamp, M., and Durbin, R. (2004). GeneWise and genomewise. *Genome Res.* 14 (5), 988–995. doi: 10.1101/gr.1865504

Boeckmann, B., Bairoch, A., Apweiler, R., Blatter, M. C., Estreicher, A., Gasteiger, E., et al. (2003). The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. *Nucleic Acids Res.* 31 (1), 365–370. doi: 10.1093/nar/gkg095

Buchfink, B., Xie, C., and Huson, D. H. (2015). Fast and sensitive protein alignment using DIAMOND. *Nat. Methods* 12 (1), 59–60. doi: 10.1038/nmeth.3176

Burton, J. N., Adey, A., Patwardhan, R. P., Qiu, R., Kitzman, J. O., and Shendure, J. (2013). Chromosome-scale scaffolding of *de novo* genome assemblies based on chromatin interactions. *Nat. Biotechnol.* 31 (12), 1119–1125. doi: 10.1038/nbt.2727

Byrne, S. L., Nagy, I., Pfeifer, M., Armstead, I., Swain, S., Studer, B., et al. (2015). A synteny-based draft genome sequence of the forage grassLolium perenne. *Plant J.* 84 (4), 816–826. doi: 10.1111/tpj.13037

Cai, J, Liu, X, Vanneste, K, Proost, S, and Liu ZJ. (2015). The genome sequence of the orchid Phalaenopsis equestris. *Nature Genetics* 47, 304-304.

Chen, N. (2004). Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr. Protoc. Bioinf.* doi: 10.1002/0471250953.bi0410s05. Chapter 4, Unit 4.10.

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 (2), 170–175. doi: 10.1038/s41592-020-01056-5

D'Hont, A., Denoeud, F., Aury, J. M., Baurens, F. C., Carreel, F., Garsmeur, O., et al. (2012). The banana (Musa acuminata) genome and the evolution of monocotyledonous plants. *Nature* 488 (7410), 213–217. doi: 10.1038/nature11241

Dong, S., Liu, M., Liu, Y., Chen, F., Yang, T., Chen, L., et al. (2021). The genome of Magnolia biondii Pamp. provides insights into the evolution of Magnoliales and biosynthesis of terpenoids. *Hortic. Res.* 8 (1), 38. doi: 10.1038/s41438-021-00471-9

Dudhate, A., Shinde, H., Yu, P., Tsugama, D., and Takano, T. (2021). Comprehensive analysis of NAC transcription factor family uncovers drought and salinity stress response in pearl millet (Pennisetum glaucum). *BMC Genomics* 22 (1), 70. doi: 10.1186/s12864-021-07382-y

Ellinghaus, D., Kurtz, S., and Willhoeft, U. (2008). LTRharvest, an efficient and flexible software for *de novo* detection of LTR retrotransposons. *BMC Bioinf.* 9, 18. doi: 10.1186/1471-2105-9-18

Finn, R. D., Mistry, J., Schuster-Böckler, B., Griffiths-Jones, S., Hollich, V., Lassmann, T., et al. (2006). Pfam: clans, web tools and services. *Nucleic Acids Res.* 34 (Database issue), D247–D251. doi: 10.1093/nar/gkj149

Gallego, L. J., Escobar, A., Penuela, M., Pena, J. D., and Rios, L. A. (2015). King Grass: A promising material for the production of second-generation butanol. *Fuel* 143, 399–403. doi: 10.1016/j.fuel.2014.11.077

Gao, B., Chen, M., Li, X., Liang, Y., Zhu, F., Liu, T., et al. (2018). Evolution by duplication: paleopolyploidy events in plants reconstructed by deciphering the evolutionary history of VOZ transcription factors. *BMC Plant Biol.* 18 (1), 256. doi: 10.1186/s12870-018-1437-8

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 (7), 644–652. doi: 10.1038/nbt.1883

Griffiths-Jones, S., Moxon, S., Marshall, M., Khanna, A., Eddy, S. R., and Bateman, A. (2005). Rfam: annotating non-coding RNAs in complete genomes. *Nucleic Acids Res.* 33 (Database issue), D121–D124. doi: 10.1093/nar/gki081

Guo, Y. D., Liu, L. Y., Yue, Y. S., Fan, X. F., Teng, W. J., Zhang, H., et al. (2022). Development of SSR markers based on transcriptome sequencing and verification of their conservation across species of ornamental pennisetum rich. (Poaceae). *Agronomy-Basel* 12 (7). doi: 10.3390/agronomy12071683

Haas, B. J., Delcher, A. L., Mount, S. M., Wortman, J. R., Smith, R. K. Jr., Hannick, L. I., et al. (2003). Improving the Arabidopsis genome annotation using maximal

transcript alignment assemblies. *Nucleic Acids Res.* 31 (19), 5654–5666. doi: 10.1093/nar/gkg770

Haas, B. J., Salzberg, S. L., Zhu, W., Pertea, M., Allen, J. E., Orvis, J., et al. (2008). Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* 9 (1), R7. doi: 10.1186/gb-2008-9-1-r7

Hayat, K. J. F. S. (2020). Combating soil salinity with combining saline agriculture and phytOmanagement with salt-accumulating plants. *Crit. Rev. Environ. Sci. Technol.* 50 (7a12). doi: 10.1080/10643389.2019.1646087

Hou, X., Teng, K., Guo, Q., Zhao, C., Gao, K., Yue, Y., et al. (2022). Research advances in forage pennisetum resource. *Chin. Bull. Bot.* 57 (06), 814–825. doi: 10.11983/CBB22195

Hu, G., Cheng, L., Cheng, Y., Mao, W., Qiao, Y., and Lan, Y. (2022). Pan-genome analysis of three main Chinese chestnut varieties. *Front. Plant Sci.* 13. doi: 10.3389/fpls.2022.916550

Huang, L., Feng, G., Yan, H., Zhang, Z., Bushman, B. S., Wang, J., et al. (2019). Genome assembly provides insights into the genome evolution and flowering regulation of orchardgrass. *Plant Biotechnol. J.* 18 (2), 373–388. doi: 10.1111/pbi.13205

Huerta-Cepas, J., Szklarczyk, D., Heller, D., Hernández-Plaza, A., Forslund, S. K., Cook, H., et al. (2019). eggNOG 5.0: a hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses. *Nucleic Acids Res.* 47 (D1), D309–d314. doi: 10.1093/nar/gky1085

International Brachypodium, I (2010). Genome sequencing and analysis of the model grass Brachypodium distachyon. *Nature* 463 (7282), 763–768. doi: 10.1038/nature08747

Jiao, Y., Li, J., Tang, H., and Paterson, A. H. (2014). Integrated syntenic and phylogenomic analyses reveal an ancient genome duplication in monocots. *Plant Cell* 26 (7), 2792–2802. doi: 10.1105/tpc.114.127597

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 (5), 1097–1111.e1012. doi: 10.1016/j.cell.2020.04.019

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

Kanehisa, M., Sato, Y., Kawashima, M., Furumichi, M., and Tanabe, M. (2016). KEGG as a reference resource for gene and protein annotation. *Nucleic Acids Res.* 44 (D1), D457–D462. doi: 10.1093/nar/gkv1070

Katoh, K., Asimenos, G., and Toh, H. (2009). Multiple alignment of DNA sequences with MAFFT. *Methods Mol. Biol.* 537, 39–64. doi: 10.1007/978-1-59745-251-9_3

Keilwagen, J., Wenk, M., Erickson, J. L., Schattat, M. H., Grau, J., and Hartung, F. (2016). Using intron position conservation for homology-based gene prediction. *Nucleic Acids Res.* 44 (9), e89. doi: 10.1093/nar/gkw092

Kim, D., Langmead, B., and Salzberg, S. L. (2015). HISAT: a fast spliced aligner with low memory requirements. *Nat. Methods* 12 (4), 357–360. doi: 10.1038/nmeth.3317

Korf, I. (2004). Gene finding in novel genomes. *BMC Bioinf.* 5, 59. doi: 10.1186/1471-2105-5-59

Li, H., and Durbin, R. (2009). Fast and accurate short read alignment with Burrows-Wheeler transform. *Bioinformatics* 25 (14), 1754–1760. doi: 10.1093/bioinformatics/btp324

Liang, Y., Xiao, W., Hui, L., Yang, T., Lian, J., Yang, R., et al. (2015). The genome of dendrobium officinale illuminates the biology of the important traditional chinese orchid herb. *Mol. Plant* 8 (006), 922–934. doi: 10.1016/j.molp.2014.12.011

Lisch, D. (2013). How important are transposons for plant evolution? *Nat. Rev. Genet.* 14 (1), 49–61. doi: 10.1038/nrg3374

Liu, L. Y., Teng, K., Fan, X. F., Han, C., Zhang, H., Wu, J. Y., et al. (2022). Combination analysis of single-molecule long-read and Illumina sequencing provides insights into the anthocyanin accumulation mechanism in an ornamental grass, Pennisetum setaceum cv. Rubrum. *Plant Mol. Biol.* 109 (1-2), 159–175. doi: 10.1007/s11103-022-01264-x

Lovell, J. T., Jenkins, J., Lowry, D. B., Mamidi, S., Sreedasyam, A., Weng, X., et al. (2018). The genomic landscape of molecular responses to natural drought stress in Panicum hallii. *Nat. Commun.* 9 (1), 5213. doi: 10.1038/s41467-018-07669-x

Lowe, T. M., and Eddy, S. R. (1997). tRNAscan-SE: a program for improved detection of transfer RNA genes in genomic sequence. *Nucleic Acids Res.* 25 (5), 955–964. doi: 10.1093/nar/25.5.955

Mamidi, S., Healey, A., Huang, P., Grimwood, J., Jenkins, J., Barry, K., et al. (2020). A genome resource for green millet Setaria viridis enables discovery of agronomically valuable loci. *Nat. Biotechnol.* 38 (10), 1203–1210. doi: 10.1038/s41587-020-0681-2

Marks, R. A., Hotaling, S., Frandsen, P. B., and Vanburen, R. (2021). Representation and participation across 20 years of plant genome sequencing. *Nat. Plants* 7, 1571–1578. doi: 10.1038/s41477-021-01031-8

Mitros, T., Session, A. M., James, B. T., Wu, G. A., Belaffif, M. B., Clark, L. V., et al. (2020). Genome biology of the paleotetraploid perennial biomass crop Miscanthus. *Nat. Commun.* 11 (1), 5442. doi: 10.1038/s41467-020-18923-6

Nawrocki, E. P., and Eddy, S. R. (2013). Infernal 1.1: 100-fold faster RNA homology searches. *Bioinformatics* 29 (22), 2933–2935. doi: 10.1093/bioinformatics/btt509

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 (1), 268–274. doi: 10.1093/molbev/msu300

Ni, P., Huang, N., Nie, F., Zhang, J., Zhang, Z., Wu, B., et al. (2021). Genome-wide detection of cytosine methylations in plant from Nanopore data using deep learning. *Nat. Commun.* 12 (1), 5976. doi: 10.1038/s41467-021-26278-9

Ou, S., and Jiang, N. (2018). LTR_retriever: A highly accurate and sensitive program for identification of long terminal repeat retrotransposons. *Plant Physiol.* 176 (2), 1410–1422. doi: 10.1104/pp.17.01310

Pertea, M., Pertea, G. M., Antonescu, C. M., Chang, T. C., Mendell, J. T., and Salzberg, S. L. (2015). StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* 33 (3), 290–295. doi: 10.1038/nbt.3122

Price, A. L., Jones, N. C., and Pevzner, P. A. (2005). *De novo* identification of repeat families in large genomes. *Bioinformatics* 21 (Suppl 1), i351–i358. doi: 10.1093/bioinformatics/bti1018

Rao, S. S., Huntley, M. H., Durand, N. C., Stamenova, E. K., Bochkov, I. D., Robinson, J. T., et al. (2014). A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell* 159 (7), 1665–1680. doi: 10.1016/j.cell.2014.11.021

Ren, G., Jiang, Y., Li, A., Yin, M., Li, M., Mu, W., et al. (2022). The genome sequence provides insights into salt tolerance of Achnatherum splendens (Gramineae), a constructive species of alkaline grassland. *Plant Biotechnol. J.* 20 (1), 116–128. doi: 10.1111/pbi.13699

Servant, N., Varoquaux, N., Lajoie, B. R., Viara, E., Chen, C. J., Vert, J. P., et al. (2015). HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* 16, 259. doi: 10.1186/s13059-015-0831-x

She, R., Chu, J. S., Wang, K., Pei, J., and Chen, N. (2009). GenBlastA: enabling BLAST to identify homologous gene sequences. *Genome Res.* 19 (1), 143–149. doi: 10.1101/gr.082081.108

Shi, T., Huneau, C., Zhang, Y., Li, Y., Chen, J., Salse, J., et al. (2022). The slow-evolving Acorus tatarinowii genome sheds light on ancestral monocot evolution. *Nat. Plants* 8 (7), 764–777. doi: 10.1038/s41477-022-01187-x

Song, C., Liu, Y., Song, A., Dong, G., and Chen, S. (2018). The Chrysanthemum nankingense Genome Provides Insights into the Evolution and Diversification of Chrysanthemum Flowers and Medicinal Traits. *Mol. Plant* 11 (12), 1–10. doi: 10.1016/j.molp.2018.10.003

Stanke, M., Diekhans, M., Baertsch, R., and Haussler, D. (2008). Using native and syntenically mapped cDNA alignments to improve *de novo* gene finding. *Bioinformatics* 24 (5), 637–644. doi: 10.1093/bioinformatics/btn013

Sun, M., Yan, H., Zhang, A., Jin, Y., Lin, C., and Lou, L. (2023). Milletdb: a multi-omics database to accelerate the research of functional genomics and molecular breeding of millets. *Plant Biotechnol. J.*, 1–10. doi: 10.1111/pbi.14136

Suyama, M., Torrents, D., and Bork, P. (2006). PAL2NAL: robust conversion of protein sequence alignments into the corresponding codon alignments. *Nucleic Acids Res.* 34 (Web Server issue), W609–W612. doi: 10.1093/nar/gkl315

Talavera, G., and Castresana, J. (2007). Improvement of phylogenies after removing divergent and ambiguously aligned blocks from protein sequence alignments. *Syst. Biol.* 56 (4), 564–577. doi: 10.1080/10635150701472164

Tang, H., Bowers, J. E., Wang, X., Ming, R., Alam, M., and Paterson, A. H. (2008). Synteny and collinearity in plant genomes. *Science* 320 (5875), 486–488. doi: 10.1126/science.1153917

Tang, S., Lomsadze, A., and Borodovsky, M. (2015). Identification of protein coding regions in RNA transcripts. *Nucleic Acids Res.* 43 (12), e78. doi: 10.1093/nar/gkv227

Tenorio, C., and Roque, R. M. (2015). Quality of pellets made from agricultural and forestry crops in Costa Rican tropical climates. *BioResources* 10 (1), 482–498. doi: 10.1016/j.molp.2018.10.003

VanBuren, R., Bryant, D., Edger, P. P., Tang, H., Burgess, D., Challabathula, D., et al. (2015). Single-molecule sequencing of the desiccation-tolerant grass Oropetium thomaeum. *Nature* 527 (7579), 508–511. doi: 10.1038/nature15714

Varshney, R. K., Shi, C., Thudi, M., Mariac, C., Wallace, J., Qi, P., et al. (2017). Pearl millet genome sequence provides a resource to improve agronomic traits in arid environments. *Nat. Biotechnol.* 35 (10), 969–976. doi: 10.1038/nbt.3943

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 (7), e49. doi: 10.1093/nar/gkr1293

Wang, J., Zi, H., Wang, R., Liu, J., Wang, H., Chen, R., et al. (2021). A high-quality chromosome-scale assembly of the centipedegrass [Eremochloa ophiuroides (Munro) Hack.] genome provides insights into chromosomal structural evolution and prostrate growth habit. *Hortic. Res.* 8 (1), 201. doi: 10.1038/s41438-021-00636-6

Xu, J., Liu, C., Song, Y., and Li, M. (2021). Comparative analysis of the chloroplast genome for four pennisetum species: molecular structure and phylogenetic relationships. *Front. Genet.* 12. doi: 10.3389/fgene.2021.687844

Xu, Z., and Wang, H. (2007). LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* 35 (Web Server issue), W265–W268. doi: 10.1093/nar/gkm286

Yan, H., Sun, M., Zhang, Z., Jin, Y., Zhang, A., Lin, C., et al. (2023). Pangenomic analysis identifies structural variation associated with heat tolerance in pearl millet. *Nat. Genet.* 55 (3), 507–518. doi: 10.1038/s41588-023-01302-4

Yan, Q., Wu, F., Xu, P., Sun, Z., Li, J., Gao, L., et al. (2020). The elephant grass (Cenchrus purpureus) genome provides insights into anthocyanidin accumulation and fast growth. *Mol. Ecol. Resour.* 21 (2), 526–542. doi: 10.1111/1755-0998.13271

Yang, Z. (1997). PAML: a program package for phylogenetic analysis by maximum likelihood. *Comput. Appl. Biosci.* 13 (5), 555–556. doi: 10.1093/bioinformatics/13.5.555

Yang, X., Liu, D., Liu, F., Wu, J., Zou, J., Xiao, X., et al. (2013). HTQC: a fast quality control toolkit for Illumina sequencing data. *BMC Bioinf.* 14(1), 33. doi: 10.1186/1471-2105-14-33

Yue, Y., Fan, X., Hu, Y., Han, C., Li, H., Teng, W., et al. (2020). *In vitro* induction and characterization of hexaploid Pennisetum × advena, an ornamental grass. *Plant C* 142 (2), 221–228. doi: 10.1007/s11240-020-01814-5

Zhang, T., Qiao, Q., Du, X., Zhang, X., Hou, Y., Wei, X., et al. (2022b). Cultivated hawthorn (Crataegus pinnatifida var. major) genome sheds light on the evolution of Maleae (apple tribe). *J. Integr. Plant Biol.* 64 (8), 1487–1501. doi: 10.1111/jipb.13318

Zhang, S., Xia, Z., Li, C., Wang, X., Lu, X., Zhang, W., et al. (2022a). Chromosome-scale genome assembly provides insights into speciation of allotetraploid and massive biomass accumulation of elephant grass (Pennisetum purpureum Schum.). *Mol. Ecol. Resour* 22 (6), 2363–2378. doi: 10.1111/1755-0998.13612

Zhang, Y., Yuan, X. H., Teng, W. J., Chen, C., and Wu, J. Y. (2016). Identification and phylogenetic classification of pennisetum (Poaceae) ornamental grasses based on SSR locus polymorphisms. *Plant Mol. Biol. Rep.* 34 (6), 1181–1192. doi: 10.1007/s11105-016-0990-2

Zheng, T., Li, P., Li, L., and Zhang, Q. (2021). Research advances in and prospects of ornamental plant genomics. *Horticult Res.* 8 (1), 65 doi: 10.1038/s41438-021-00499-x

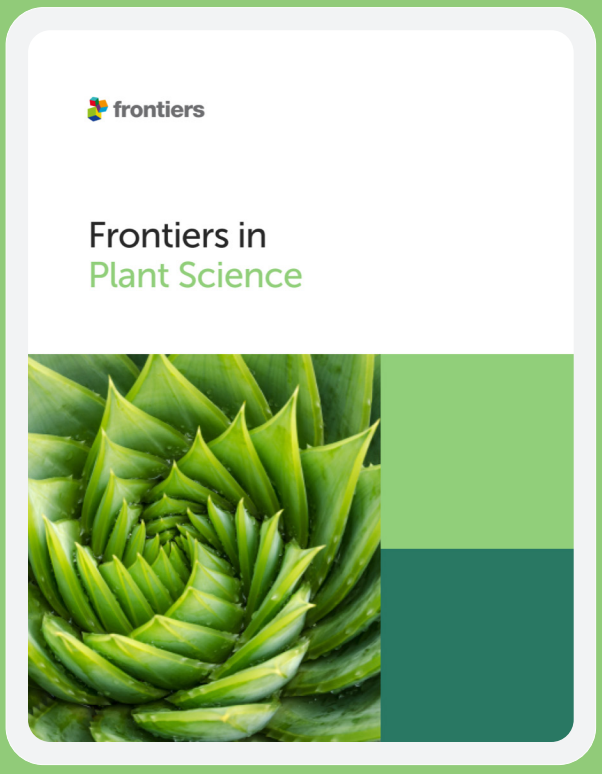
Frontiers in
Plant Science

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