plants

Special Issue Reprint

Crop Functional Genomics and Biological Breeding

Edited by
Yifeng Wang, Jie Huang, Jian Zhang and Jiezheng Ying

mdpi.com/journal/plants

MDPI

Crop Functional Genomics and Biological Breeding

Crop Functional Genomics and Biological Breeding

Guest Editors

Yifeng Wang
Jie Huang
Jian Zhang
Jiezheng Ying

MDPI

Basel • Beijing • Wuhan • Barcelona • Belgrade • Novi Sad • Cluj • Manchester

Guest Editors

Yifeng Wang
State Key Lab of Rice Biology and Breeding
China National Rice Research Institute
Hangzhou
China

Jie Huang
State Key Lab of Rice Biology and Breeding
China National Rice Research Institute
Hangzhou
China

Jian Zhang
State Key Lab of Rice Biology and Breeding
China National Rice Research Institute
Hangzhou
China

Jiezheng Ying
State Key Lab of Rice Biology and Breeding
China National Rice Research Institute
Hangzhou
China

Editorial Office
MDPI AG
Grosspeteranlage 5
4052 Basel, Switzerland

This is a reprint of the Special Issue, published open access by the journal *Plants* (ISSN 2223-7747), freely accessible at: https://www.mdpi.com/journal/plants/special_issues/297G6ZM68I.

For citation purposes, cite each article independently as indicated on the article page online and as indicated below:

Lastname, A.A.; Lastname, B.B. Article Title. *Journal Name* **Year**, *Volume Number*, Page Range.

ISBN 978-3-7258-4113-4 (Hbk)
ISBN 978-3-7258-4114-1 (PDF)
https://doi.org/10.3390/books978-3-7258-4114-1

Cover image courtesy of Jie Huang

Contents

plants

MDPI

Editorial

Crop Functional Genomics and Biological Breeding

Jia Li [1,2], Jie Huang [1], Jiezheng Ying [1], Jian Zhang [1], Dawei Xue [2] and Yifeng Wang [1,*]

[1] State Key Laboratory of Rice Biology and Breeding, China National Rice Research Institute, Hangzhou 311400, China; lij697495@163.com (J.L.); huangjie@caas.cn (J.H.); yingjiezheng@caas.cn (J.Y.); zhangjian@caas.cn (J.Z.)

[2] College of Life and Environmental Sciences, Hangzhou Normal University, Hangzhou 311121, China; dwxue@hznu.edu.cn

* Correspondence: wangyifeng@caas.cn

1. Introduction

Over the past two decades, the rapid development of functional genomics has gradually clarified the regulatory effects of genotypes on phenotypes. Crop breeding researchers have discovered numerous genes that influence critical traits such as yield, resistance, and quality. By targeting these genes, scientists have achieved more customised variety breeding through biological breeding. Crop functional genomics research is a vital pathway for biological breeding. Understanding the functional genomics of crops provides deep insights into the genetic mechanisms controlling key traits like yield, stress resistance, and quality improvement. This article aims to summarise the latest research and discoveries in crop functional genomics from plant journals, including studies on gene responses to stresses through whole-genome and transcriptome analyses and the identification and cloning of new genes related to crop yield, resistance, germination, and quality. This article also covers functional analyses of these genes and their applications in biological breeding.

2. Studying Gene Resistance Through Functional Genomics

Liu et al. conducted a genome-wide analysis of the Nramp gene family in kenaf (*Hibiscus cannabinus*), identifying 15 *HcNramp* genes. Phylogenetic analysis classified these 15 HcNramp proteins into three distinct subfamilies, with genes in the same subfamily sharing similar structures and motifs [1]. The promoters of these genes were rich in hormone-responsive and environmental stress-responsive elements. Collinearity analysis revealed 10 pairs of collinear genes between kenaf and *Arabidopsis* and another 10 pairs between kenaf and tomato. Transcriptomic data showed that all *Nramp* genes in kenaf exhibited tissue-specific expression. Under treatments with varying concentrations of cadmium (Cd) and melatonin, the expression levels of *HcNramp* genes differed across tissues and were significantly affected by Cd stress. *HcNramp3* was preferentially expressed in leaves, while *HcNramp2/12/13/15* were primarily expressed in roots. *HcNramp3/11* may regulate Cd, *HcNramp1/6/7/14* may be involved in Cd transport, and *HcNramp2/6/8/9/12* could enhance kenaf's tolerance to Cd. These findings provide a foundation for understanding the potential role of the *Nramp* gene family in Cd stress and are significant for screening Cd-resistant genes in kenaf.

The *SAUR* (small auxin-up RNA) genes rapidly respond to auxin hormones and play key roles in various biological processes, including plant growth and abiotic stress responses. Zhu et al. conducted a genome-wide analysis in sandalwood (*Santalum album* L.), identifying 43 *SaSAUR* genes. Phylogenetic and gene structure analyses divided these genes into five groups with similar structures, with most lacking introns. Collinearity

Received: 27 March 2025
Revised: 11 April 2025
Accepted: 17 April 2025
Published: 21 April 2025

Citation: Li, J.; Huang, J.; Ying, J.; Zhang, J.; Xue, D.; Wang, Y. Crop Functional Genomics and Biological Breeding. *Plants* **2025**, *14*, 1258. https://doi.org/10.3390/plants14081258

analysis revealed 14 segmental duplications and 9 tandem duplications among *SaSAUR* genes, suggesting that duplication is the primary driver of the evolutionary expansion in the *SAUR* gene family [2]. Transcriptomic data from various tissues of sandalwood showed that all 43 *SAUR* genes were expressed in leaves, with no significant differences in expression levels. *SaSAUR18* was highly expressed in roots, while *SaSAUR08* and *SaSAUR13* showed significantly higher expression in heartwood. Under salt treatments with varying concentrations, *SaSAUR04*, *SaSAUR07*, *SaSAUR08*, *SaSAUR09*, *SaSAUR18*, *SaSAUR27*, and *SaSAUR28* were upregulated considerably, indicating their potential involvement in salt stress response mechanisms. These studies contribute to a better understanding of the *SAUR* gene family, particularly its relationship with salt stress responses, and lay the groundwork for future functional analyses.

Maize lethal necrosis (MLN) poses a significant threat to food security in sub-Saharan Africa (SSA), with only a few commercial inbred lines showing tolerance. Murithi et al. performed transcriptomic analyses on four commercial maize inbred lines and one non-adapted inbred line under MLN infection [3]. They found that the gene expression patterns of different varieties did not fully align with their phenotypic resistance levels. The resistance of KS23-6 was primarily against maize chlorotic mottle virus (MCMV), while the CML lines exhibited resistance to sugarcane mosaic virus (SCMV). Differentially expressed gene (DEG) enrichment results showed that the disease resistance genes were mainly *leucine-rich repeat* (*LRR*) genes and components of the plant innate immune system, such as the RNA interference pathway and the ubiquitin–proteasome system. The expression levels of translation initiation factor *eIF4E* and elongation factor *eIF4G*, which are related to viral replication, were also closely associated with MLN viruses. These findings provide molecular-level theoretical support for MLN disease control and offer candidate gene targets for breeding MLN-resistant maize germplasm in SSA.

Serine hydroxymethyltransferase (SHMT, E.C.2.1.2.1) is a critical enzyme that catalyses serine and glycine conversion and is widely distributed in plants, animals, and microorganisms. However, systematic functional studies on SHMT gene family members in rice remain limited. Pan et al. comprehensively analysed the *SHMT* gene family [4]. Phylogenetic analysis and subcellular localisation results showed that the five *SHMT* genes in rice were distributed across three subgroups: Ia, Ib, and IIb. *OsSHMT3*, classified initially in the cytoplasm-localised Ib subgroup, was localised to chloroplasts in rice protoplasts. Pan et al.'s study also revealed a unique gene loss event in monocots, leading to the absence of the chloroplast-localised IIa subgroup of *SHMT*. Yeast two-hybrid experiments confirmed that all five OsSHMT proteins could form homodimers, with OsSHMT3 capable of forming heterodimers with other members except for OsSHMT1. Molecular docking simulations suggested that OsSHMT3 might coordinate with other members as a mobile protein. Cis-acting element predictions and expression pattern analyses indicated that *OsSHMT* family members might participate in various stress responses and hormone-regulatory pathways. In vitro experiments showed that OsSHMT1/3 exhibited enzymatic activity, while OsSHMT4/5 were inactive. Whether these proteins function as catalytic enzymes or purely regulatory elements remains to be elucidated.

Alfalfa (*Medicago sativa* L.), a crucial perennial leguminous forage crop, is widely recognised as the "King of Forages" due to its exceptional nutritional value, palatability, and adaptability. In Northeast China, where vast areas of saline–alkali land remain underutilised, investigating alfalfa's salt tolerance is vital for developing salt-tolerant cultivars and reclaiming these marginal lands. Through a two-phase screening process of 41 alfalfa varieties, Wang et al. identified Longmu801 as a salt-tolerant cultivar and WL168 as a salt-sensitive one [5]. Transcriptomic analysis revealed that differentially expressed genes (DEGs) in both Longmu801 and WL168 under salt stress conditions showed significant

enrichment in oxidative stress response, nuclear activities, and plasma membrane functions. KEGG pathway analysis further demonstrated close associations between salt tolerance and key metabolic pathways, including phenylpropanoid biosynthesis, endoplasmic reticulum protein processing, and starch and sucrose metabolism. Using weighted gene co-expression network analysis (WGCNA), researchers identified six modules related to salt tolerance and subsequently pinpointed five hub genes directly responsive to salt stress: *Msa085011*, *Msa0605650*, *Msa0397400*, *Msa1258740*, and *Msa0958830*. In conclusion, while salt-tolerant and salt-sensitive cultivars share similar growth and physiological response mechanisms under salt stress, the salt-tolerant Longmu801 exhibits superior tolerance through the enhanced accumulation of osmolytes, elevated antioxidant enzyme activities, and distinct gene expression patterns. This study provides valuable candidate genes for breeding salt-tolerant alfalfa varieties and offers scientific foundations for the sustainable utilisation of saline–alkali soils.

3. Functional Analysis of Plant Trait-Related Genes

The BBX protein family is highly conserved in plants, with 32 BBX genes in *Arabidopsis* and 30 in rice. BBX proteins involve diverse plant regulatory networks, ranging from seedling photomorphogenesis to stress responses. Yang et al. identified a BBX protein, OsBBX2, which is highly expressed in rice roots and exhibits diurnal rhythmic transcription [6]. Overexpressing *OsBBX2* in the *japonica* rice variety Longjing 11 (LJ11), which has weakened or lost functions of *Ghd7* and *PRR37*, delayed heading under long-day (LD) conditions. The flowering phenotype of the *bbx2* mutant was not significantly different from that of the wild type (WT), possibly due to functional redundancy among other BBX family members. In *OsBBX2*-overexpressing plants, *Ehd1*, *Hd3a*, and *RFT* expression levels were significantly reduced. Yeast two-hybrid experiments revealed that OsBBX2 interacts with Hd1 (BBX18), substantially suppressing *Hd3a* expression. OsBBX2 and Hd1 synergistically inhibit *Hd3a* expression to delay rice flowering. Meanwhile, the *bbx2* and *hd1* double mutant exhibited a late-flowering phenotype similar to *hd1*.

Research on rice grain size is crucial for breeding high-yield and high-quality rice varieties. Plant hormones are important factors regulating grain development. Aux/IAA family proteins, as core components of the auxin signalling pathway, are plant-specific transcription factors that play vital roles in regulating plant growth and development. Although *Arabidopsis* Aux/IAA family genes have been extensively studied, the functions of most members in rice remain unclear. Only two genes, *OsIAA3* and *OsIAA10*, have been reported to regulate rice grain size. Jia et al. found that the *osiaa19* mutant exhibited significantly increased grain length and thousand-grain weight, with no significant effects on agronomic traits such as plant height, tiller number, or flag leaf length [7]. However, the chalkiness of *osiaa19* grains increased, which might be related to the transient starch degradation enzyme GWD1 in source tissues. The high expression of *OsIAA19* in source tissues may indirectly cause chalkiness by affecting the carbon supply. This study provides new insights into the auxin signalling regulatory network and has significant application value for rice quality improvement.

The composition and distribution of storage substances in rice endosperm directly affect rice quality. Yang et al. identified a floury endosperm mutant, *wcr* (white-core rice), which exhibited a unique "transparent periphery-floury core" phenotype [8]. Microscopic observations revealed loosely arranged starch granules with large pores in the inner endosperm, while the outer region had densely packed starch granules. Compared to the wild type, brown rice's total starch and amylose content showed no changes, but the levels of four protein components were significantly reduced. In milled rice (inner endosperm), total starch and amylose content decreased significantly, while albumin content nearly tripled.

Through gene mapping and sequence alignment, Yang et al.'s study revealed that the *wcr* mutant phenotype was regulated by the recessive nuclear gene *OsPPDK8*, which encodes pyruvate phosphate dikinase. The mutation in *OsPPDK8* rendered it unable to catalyse pyruvate and inorganic phosphate (Pi), leading to energy deficiency, impaired internal substance transport, and developmental defects. *OsPPDK8* expression peaked at 10 days after flowering and was rapidly inactivated through threonine phosphorylation and protein degradation mechanisms. This period coincides with the critical phase of storage substance accumulation from the endosperm centre to the periphery, potentially explaining the severe floury phenotype in the inner endosperm. However, further experimental validation is required to determine whether this regulatory mechanism is accurate. This study confirms that *wcr* is a key regulator, coordinating protein synthesis and starch metabolism to regulate rice quality formation jointly.

Rice tiller number directly affects final yield and is a key trait for breeding and nitrogen-efficient cultivation. Although extensive research has been conducted, the molecular mechanisms underlying the interaction between nitrogen and plant hormones in regulating tillering remain unclear. Rimi et al. performed genetic analysis and gene mapping through population construction and BSA sequencing, revealing that the dwarf and multi-tiller phenotype of the *p47dt1* mutant was controlled by the single gene *P47DT1 (D10)* [9]. By analysing the tillering response patterns of the wild-type P47-1 and the *p47dt1* mutant under different nitrogen concentrations, Rimi et al. found that the mutation in *P47DT1* caused the loss of nitrogen response capability in rice. Observations throughout the growth cycle confirmed that P47-1 consistently had fewer tillers than *p47dt1*, indicating that *P47DT1* is a key gene regulating nitrogen response in rice tillering. Further experiments with different nitrogen forms revealed that the *P47DT1* mutation altered nitrogen allocation patterns between leaves and stems through the nitrate pathway. Transcription factor binding site predictions in the *D10* gene promoter identified numerous plant hormone-responsive elements and *Arabidopsis TCP20* transcription factor binding sites, suggesting that this site in rice might bind to *TCP19*. However, these results remain speculative and require further experimental exploration to confirm the regulatory pathways and their effects on tillering.

Rice blast is a primary global disease in rice production, causing annual yield losses of 10–50%. Disease-resistant breeding is the core solution to this problem. Jin et al. cloned a new gene, *Pigm-1*, from the double-resistant line 77009, which is allelic to the known broad-spectrum resistance gene *Pigm* [10]. Resistance spectrum analysis revealed that the two genes differed only in their resistance to the JL-37 line: *Pigm* was resistant to JL-37, while *Pigm-1* was relatively susceptible. Using *Pigm-1* donor R20-4 crossed with susceptible glutinous rice S19-118, a new blast-resistant glutinous rice material, Xiannuo 23, was developed through marker-assisted selection, showing significant field resistance. Recent studies found that *PigmR* (the protein encoded by *Pigm*) enriches cell membrane microdomains via the PIBP4-Rab5a transport machinery, activating OsRac1 to trigger reactive oxygen bursts and enhance resistance. To date, *Pigm-1* has been successfully used to develop six blast-resistant restorer lines (using Minghui 63 as the receptor), four new resistant materials (using R20 as the receptor), and five materials resistant to both blast and bacterial blight (combining *Xa23*). The *Pigm* allele has been used in molecular breeding by over 30 domestic institutions, with several varieties entering regional trials. However, the molecular mechanism of JL-37 strain-specific recognition remains unresolved. This study confirms that *Pigm-1* possesses broad-spectrum resistance comparable to *Pigm*, providing a new option for disease-resistant breeding.

4. Conclusions and Perspectives

This paper reviews the latest research progress in crop functional genomics and biological breeding, focusing on applying gene functional genomics in crop resistance, trait regulation, and breeding. Further functional validation of identified genes (e.g., HcNramp, *SaSAUR*, *OsSHMT*, etc.) is needed to elucidate their specific mechanisms in stress resistance and quality regulation while exploring gene interaction networks to uncover the complex mechanisms of multi-gene synergistic regulation. The molecular basis of crop trait formation can be comprehensively analysed by integrating multiple omics methods, including genomics, transcriptomics, proteomics, and metabolomics. Advanced gene-editing technologies (e.g., CRISPR-Cas9) are employed to modify key genes precisely, and high-efficiency molecular marker-assisted selection techniques are developed to breed new crop varieties with enhanced stress tolerance, high yield, and superior quality. In response to global challenges such as climate change and resource scarcity, research prioritises crop response mechanisms to abiotic stresses (e.g., salinity, drought, and heavy metal contamination) to develop more resilient varieties.

Additionally, disease resistance breeding is intensified, particularly in the mining and application of broad-spectrum resistance genes against significant diseases like rice blast. Through functional genomics and biological breeding technologies, crop yield and quality can be improved while reducing reliance on chemical fertilisers and pesticides, promoting sustainable green agriculture. Furthermore, the potential of marginal lands (e.g., saline–alkali soils) is maximised to provide new solutions for food security.

In conclusion, crop functional genomics and biological breeding hold immense potential. Future efforts should strengthen the integration of fundamental research and practical applications to provide robust scientific and technological support for sustainable agricultural development.

Author Contributions: J.H., J.Y., J.Z. and Y.W. participated in the editing of this Research Topic. J.L. wrote the draft, and all the other authors provided suggestive comments on the Editorial. D.X. participated in reviewing and editing the Editorial. All authors have read and agreed to the published version of the manuscript.

Funding: This project was funded by the National Natural Science Foundation of China (32372070).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Liu, Q.; Li, S.C.; Du, G.H.; An, X. Genome-Wide Analysis of the *Nramp* Gene Family in Kenaf (*Hibiscus cannabinus*): Identification, Expression Analysis, and Response to Cadmium Stress. *Plants* **2024**, *13*, 2514. [CrossRef] [PubMed]
2. Zhu, Q.; Zheng, H.Y.; Hu, X.; Liu, Y.; Zheng, X.Y.; Li, L.B.; Tang, M.Q. Genome-Wide Analysis of the SAUR Gene Family and Its Expression Profiles in Response to Salt Stress in *Santalum album*. *Plants* **2024**, *13*, 1286. [CrossRef] [PubMed]
3. Murithi, A.; Panangipalli, G.; Wen, Z.Y.; Olsen, M.S.; Lübberstedt, T.; Dhugga, K.S.; Jung, M.R. Global Transcriptomic Analysis of Inbred Lines Reveal Candidate Genes for Response to Maize Lethal Necrosis. *Plants* **2025**, *14*, 295. [CrossRef] [PubMed]
4. Pan, T.; Jin, H.M.; Zhou, C.H.; Yan, M.Y. Rice Serine Hydroxymethyltransferases: Evolution, Subcellular Localization, Function and Perspectives. *Plants* **2024**, *13*, 1116. [CrossRef] [PubMed]
5. Wang, F.D.; Wu, H.F.; Yang, M.; Xu, W.; Zhao, W.J.; Qiu, R.; Kang, N.; Cui, G.W. Unveiling Salt Tolerance Mechanisms and Hub Genes in Alfalfa (*Medicago sativa* L.) Through Transcriptomic and WGCNA Analysis. *Plants* **2024**, *13*, 3141. [CrossRef] [PubMed]
6. Yang, Y.S.; Wei, J.M.; Tian, X.J.; Liu, C.H.; Li, X.F.; Bu, Q.Y. OsBBX2 Delays Flowering by Repressing *Hd3a* Expression Under Long-Day Conditions in Rice. *Plants* **2025**, *14*, 48. [CrossRef] [PubMed]
7. Jia, S.S.; Ren, X.Y.; Tong, M.N.; Jiang, S.Y.; Zhang, C.Q.; Liu, Q.Q.; Li, Q.F. *OsIAA19*, an *Aux/IAA* Family Gene, Involved in the Regulation of Seed-Specific Traits in Rice. *Plants* **2024**, *13*, 3538. [CrossRef] [PubMed]
8. Yang, Y.H.; Yang, X.Y.; Wu, L.J.; Sun, Z.X.; Zhang, Y.; Shen, Z.Y.; Zhou, J.; Guo, M.; Yan, C.J. Phenotypic Analysis and Gene Cloning of Rice Floury Endosperm Mutant *wcr* (White-Core Rice). *Plants* **2024**, *13*, 2653. [CrossRef] [PubMed]

9. Rimi, T.I.; Zhang, M.R.; Zhang, R.X.; Zhang, Z.; Leng, X.Y.; Han, J.F.; Meng, S.H.; Du, W.; Zhang, Z.C. Putative Allele of *D10* Gene Alters Rice Tiller Response to Nitrogen. *Plants* **2024**, *13*, 3349. [CrossRef] [PubMed]
10. Jin, Y.D.; He, N.Q.; Cheng, Z.P.; Lin, S.J.; Huang, F.H.; Wang, W.X.; Li, Q.Q.; Yang, D.W. Resistance Spectrum Analysis and Breeding Utilization of Rice Blast Resistance Gene *Pigm-1*. *Plants* **2025**, *14*, 535. [CrossRef] [PubMed]

plants

MDPI

Article

Rice Serine Hydroxymethyltransferases: Evolution, Subcellular Localization, Function and Perspectives

Tian Pan *,†, Hongmiao Jin †, Chuanhui Zhou and Mengyuan Yan *

The Key Laboratory for Quality Improvement of Agricultural Products of Zhejiang Province, College of Advanced Agricultural Sciences, Zhejiang A&F University, Hangzhou 311300, China; jhm@stu.zafu.edu.cn (H.J.); 202220020225@stu.zafu.edu.cn (C.Z.)

* Correspondence: tianpan@zafu.edu.cn (T.P.); yanmengyuan@zafu.edu.cn (M.Y.)

† These authors contributed equally to this work.

Abstract: In rice, there is a lack of comprehensive research on the functional aspects of the members of the *serine hydroxymethyltransferase* (*SHMT*) gene family. This study provides a comprehensive investigation of the *SHMT* gene family, covering phylogeny, gene structure, promoter analysis, expression analysis, subcellular localization, and protein interaction. Remarkably, we discovered a specific gene loss event occurred in the chloroplast-localized group IIa SHMTs in monocotyledons. However, OsSHMT3, which originally classified within cytoplasmic-localized group Ib, was found to be situated within chloroplasts in rice protoplasts. All five OsSHMTs are capable of forming homodimers, with OsSHMT3 being the only one able to form dimers with other OsSHMTs, except for OsSHMT1. It is proposed that OsSHMT3 functions as a mobile protein, collaborating with other OsSHMT proteins. Furthermore, the results of *cis*-acting element prediction and expression analysis suggested that members of the *OsSHMT* family could be involved in diverse stress responses and hormone regulation. Our study aims to provide novel insights for the future exploration of SHMTs.

Keywords: serine hydroxymethyltransferase; rice; function

Citation: Pan, T.; Jin, H.; Zhou, C.; Yan, M. Rice Serine Hydroxymethyltransferases: Evolution, Subcellular Localization, Function and Perspectives. *Plants* **2024**, *13*, 1116. https://doi.org/10.3390/plants13081116

Academic Editor: Khalid Meksem

Received: 21 March 2024
Revised: 9 April 2024
Accepted: 15 April 2024
Published: 16 April 2024

1. Introduction

One carbon metabolism is essential to all organisms. In plants, these reactions provide the necessary C1 units for protein synthesis, nucleic acids, pantothenate, and a diverse array of methylated compounds [1]. Despite the initial understanding of one carbon metabolism, many aspects of plant one carbon metabolism, including enzymes, pathways, and regulatory mechanisms, remain to be fully understood. Both serine and glycine serve as potential providers of C1 units in plants. The reciprocal transformation between serine and glycine is extensively studied in plant photorespiration and is integral in the synthesis of crucial compounds like purines, pyrimidines, and one-carbon units [1–4]. Serine hydroxymethyltransferase (SHMT, E.C. 2.1.2.1) is an important enzyme that catalyzes the conversion of serine and glycine and is widely distributed in plants, animals, and microorganisms [1]. SHMT facilitates the reversible transfer of a hydroxymethyl group from serine to $H_4PteGlu_n$ (tetrahydrofolate), resulting in the production of glycine and 5,10-CH_2-$H_4PteGlu_n$ (5,10-methylenetetrahydrofolate) [5]. In plants, SHMT activity has been detected in different organelles (mitochondria, cytosol, plastids, and the nucleus), indicating their diverse roles in metabolic pathways [6,7]. Previous studies have shown that nearly all prokaryotic SHMTs exist in the form of homologous dimers, while eukaryotic SHMTs exist as homotetramers, which are dimers of dimeric structures [8]. It remains unclear whether SHMT proteins localized in different organelles within plants can form homologous or heterologous multimers.

Currently, the most extensively studied SHMTs are those localized within mitochondria. Earlier studies have found that mitochondrial serine hydroxymethyltransferase (mSHMT) has adapted to participate in photorespiration [9]. Within mitochondria, mSHMT

and glycine decarboxylase (GDC) convert photorespiratory glycine into serine [4,10]. Arabidopsis mutants lacking mitochondrial SHMT activity exhibited lethal phenotypes when grown under ambient CO_2 concentrations and are rescued under high CO_2 [11,12]. AtSHMT1 enzyme activity plays a crucial role in controlling cell damage induced by high light and abiotic stress like salt [13]. The phosphorylation of SHMT1 at S31 might play a role in regulating SHMT1 protein stability, consequently regulating salt stress or drought stress in Arabidopsis [14]. The mitochondrially localized OsSHMT1 in rice is also a major protein influencing photorespiration. Studies have found that a mitochondrial *SHMT* gene mutation in rice (*osshmt1*) causes blockage of the photo-respiration pathway. *OsSHMT1* plays a role in reducing chloroplast ROS production to alleviate photoinhibition [15]. Additionally, Arabidopsis transgenic plants overexpressing *OsSHMT3* (actually *OsSHMT1*) exhibit salt stress tolerance, revealing the crucial role of *OsSHMT1* in enhancing plant tolerance to salt stress [16]. Recent research has reported the regulatory role of OsSHMT1 in rice under low-temperature stress, revealing its potential role in scavenging H_2O_2 to enhance cold tolerance in rice [17].

The presence of chloroplasts in plant cells makes one-carbon metabolism more intricate compared to animals and prokaryotes. Enzymes responsible for nucleotide biosynthesis reside within chloroplasts and the principal role of chloroplast SHMT is probably centered on utilizing serine within this space to supply a one-carbon unit [18]. Notably, there is a scarcity of studies focusing on SHMT proteins within chloroplasts. Although chloroplast-localized AtSHMT3 enzymatic activity was analyzed, it lacked corresponding mutant studies to further explore its impact on plant phenotypes and molecular mechanisms. The activity of SHMT in chloroplasts potentially plays a crucial role in light reception and the biosynthesis of purines, pyrimidines, and N-formylmethionine [7].

The one-carbon flux in the plant cytoplasm might predominantly engage in supporting SAM-mediated methylation reactions, as nucleotide synthesis seems not to occur within this compartment [19]. The products of cytosol, mitochondria, and plastid-located SHMTs can interconvert and communicate with each other [19]. Cytoplasmically localized GmSHMT08 in soybeans confers resistance to soybean cyst nematodes (SCN; *Heterodera glycines*) [20]. Recent experimental evidence indicates the involvement of GmSHMT08 in DNA methylation [21]. The GmSHMT08/GmSNAP18 (soluble NSF attachment protein, SNAP)/GmPR08-Bet VI (pathogenesis-related protein, PR) multi-protein complex is fundamental to soybean resistance against SCN [22,23]. Both THF and PLP sites at the GmSHMT08 are involved in resistance to SCN [24]. In conclusion, GmSHMT08 plays a pivotal role in the host defense process.

There is relatively limited research conducted on nuclear-localized SHMT. In Arabidopsis, AtSHMT7 in the cell nucleus regulates SAM biosynthesis and maintains the dynamic balance of sulfur via DNA methylation [6]. *OsSHMT4* negatively regulates the uptake and assimilation of sulfate/selenate, and its mutation enhances rice tolerance to Cd and selenium content in grains [25]. Furthermore, *OsSHMT4* influences SAM level and global DNA differential methylation, which further regulates the endosperm development [26].

In mammals, two SHMT isozymes are encoded by separate genes. *SHMT1* encodes the cytoplasmic/nuclear isozyme (SHMT1), and *SHMT2* encodes the mitochondrial (SHMT2) and the cytoplasmic/nuclear (SHMT2α) isoform through alternative promoter use [27]. The pathway for de novo thymidylate biosynthesis involves the translocation of enzymes like SHMT1 and SHMT2α to the nucleus during DNA replication and repair [27]. During the S and G2/M phases of the cell cycle, cytoplasmic SHMT1 has been demonstrated to relocate to the nucleus and nuclear periphery. This translocation is facilitated by post-translational modification through the attachment of the small ubiquitin-like modifier (SUMO) [28,29]. It remains unknown whether SHMT proteins from different organelles in plants undergo translocation. Although the SHMT family of proteins is relatively conserved, SHMTs with different signal peptides exhibit diversity in structure and function, participating extensively in the cellular life processes. The specific functions of SHMT proteins in different cellular locations within plants require further in-depth research.

In this study, a comprehensive analysis including the identification, evolution analysis, gene structure and promoter analysis, expression patterns in different tissues, responses to various hormones, subcellular localization analysis, and the interaction networks of SHMTs in rice was performed. These results contribute to the clarification of OsSHMT protein functions and provide valuable information for further studies in plants to uncover the functional roles of *SHMT* genes in growth and progression.

2. Results

2.1. Identification and Evolutionary Analysis of SHMT Family Genes in the Rice Genome

In this study, five *SHMT* genes were identified with a certain SHMT domain. The biophysical characteristics of these genes vary (Table S1). The coding sequence length of the five *SHMT* genes was between 1416 and 1800 bp, and the predicted protein lengths of OsSHMTs were between 471 and 599 aa, with molecular weights (MWs) varying from 51.44 to 64.79 kDa. Among the five proteins, two have a pI smaller than 7, which were acidic proteins, while three OsSHMTs were projected to encode basic proteins with pI > 7. All OsSHMTs, excluding OsSHMT2, displayed an instability index (II) greater than 40, indicating a prevalence of unstable proteins within this gene family. It is worth noting that all OsSHMTs had a grand average of hydropathicity (GRAVY) of less than 0, indicating their hydrophilic properties. Among them, OsSHMT2 was the smallest protein of the five, with the lowest instability index and the biggest grand average of hydropathicity. All five SHMTs were distributed on different chromosomes but with similar MWs and hydrophilic characteristics.

To further evaluate the evolutionary relation among OsSHMTs, phylogenetic tree analysis was performed. The tree consists of 42 SHMT proteins, namely 7 in Arabidopsis, 12 in soybean, 8 in tomato, 10 in maize, and 5 in rice. According to the classification of Arabidopsis and soybean SHMTs. Plant SHMTs can be divided into four groups, group Ia, group Ib, group IIa, and group IIb (Figure 1). OsSHMT1 was grouped with five GmSHMT proteins (GmSHMT02m, GmSHMT08m, GmSHMT09m, GmSHMT14m, and GmSHMT18m) and two AtSHMT proteins (AtSHMT1 and AtSHMT2) in group IIb. OsSHMT2 and OsSHMT3 were grouped with AtSHMT4, AtSHMT5, GmSHMT05c, and GmSHMT08c and other five SHMT proteins in tomato, cucumber, and maize, classified into group Ib. OsSHMT4 and OsSHMT5 were placed in group Ia. Additionally, two AtSHMT proteins (AtSHMT6 and AtSHMT7), two CsSHMT proteins (CsSHMT5 and CsSHMT6), two ZmSHMT proteins (ZmSHMT3 and ZmSHMT6), four GmSHMT proteins (GmSHMT04n, GmSHMT06n, GmSHMT08n, and GmSHMT12n), and one SlSHMT1 protein were also included in this group. Notably, no homolog of AtSHMT3 was identified in rice (Figure 1), suggesting that a branch of the rice SHMT gene family is absent.

To verify whether the absence of group IIa SHMT in rice is also observed in other plant species, the extensive exploration of SHMT proteins among several representative species were carried out. SHMTs in *Populus trichocarpa*, *Nymphaea colorata*, *Amborella trichopoda*, *Chlamydomonas reinhardtii*, *Physcomitrium patens*, *Homo sapiens*, and *Saccharomyces cerevisiae* were added. The comprehensive analysis of the SHMT tree revealed a total of 79 SHMT proteins distributed among seven taxa, encompassing Eudicots, Monocots, Amborella, Chlorophyta, Bryophyta, Chordata, and Eumycophyta (Table S3). The evolutionary tree is generally classification consistent with subcellular localization results, and all SHMT proteins can be categorized into two groups and four classes based on their cellular localization (chloroplastic, mitochondrial, cytosolic, and nuleaic) (Figure 2). The chloroplastic- and mitochondrial-localized SHMTs form one clade, while the cytosolic- and nuclear-localized SHMTs constitute another. Surprisingly, the chloroplastic class exclusively comprises proteins from dicotyledons, while the remaining subgroups encompass proteins from both monocotyledons and dicotyledons. Plant SHMTs likely underwent a duplication event, contrasting with their reduced numbers in humans and fungi. Within the same subgroup, the SHMT members of monocotyledons and dicotyledons clustered together, signifying a significant divergence in the evolution of SHMT genes between these two groups. Eudicots

exhibit over three times the number of SHMT proteins compared to monocots, suggesting potential distinct evolutionary strategies adopted by dicot species.

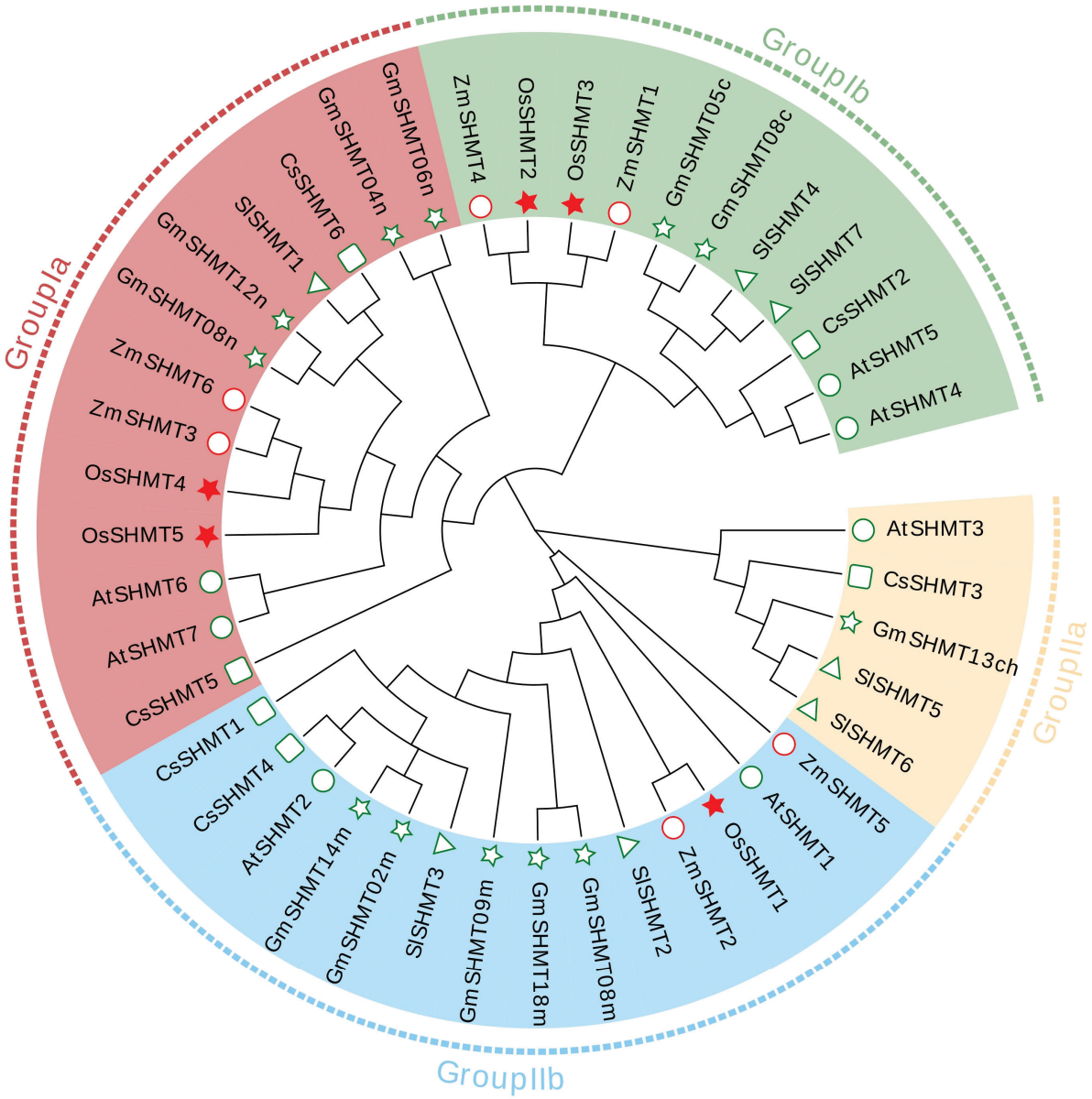

Figure 1. Phylogenetic tree of SHMT proteins from *O. sativa*, *Z. mays*, *A. thaliana*, *G. max*, *S. lycopersicum*, and *C. sativus*. The phylogenetic tree was built using the maximum likelihood method (ML) with 1000 bootstrap replicates by MEGA 7. Different colors represent different subfamilies, as follows: group Ia (red), group Ib (green), group IIa (orange), and group IIb (blue). Symbols with different colors represent the SHMT proteins from different species in this tree: *O. sativa* (red stars), *Z. mays* (red circles), *A. thaliana* (green circles), *G. max* (green stars), *S. lycopersicum* (green triangles), and *C. sativus* (green squares). Gene IDs of the analyzed genes can be found in Table S2.

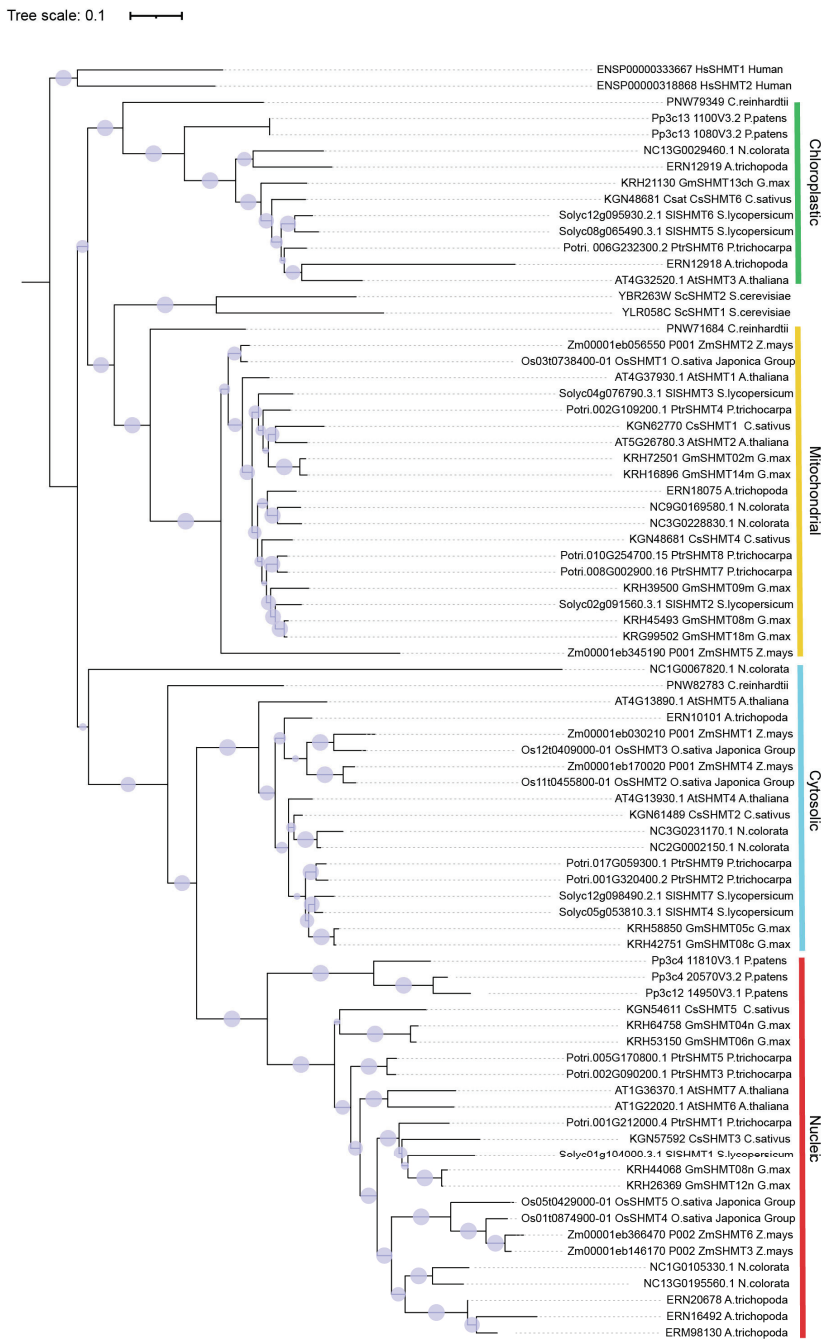

Figure 2. The evolution of SHMT proteins. The tree was constructed using trimmed amino acid sequences obtained from the ensemble (Table S3). The phylogenetic tree was generated using MEGA7 and calculated using the maximum likelihood method. The tree bootstrap values are indicated at the nodes (n = 1000).

2.2. Gene Structure and Promoter Cis-Element Analysis of OsSHMTs

To further understand the *OsSHMTs*, gene structure and conserved motifs were analyzed (Figure 3, Table S4). The functional diversity of OsSHMTs is influenced by the unique compositions of the motifs. The analysis of conservative motifs revealed that five OsSHMTs shared motifs 1–7 and 9, with OsSHMT1 lacking motif 8, while OsSHMT4 and OsSHMT5 lacked motif 10 (Figure 3B). Previous studies have indicated that OsSHMT4 and OsSHMT5 individually lacked enzyme activity in vitro [26], suggesting that the absence of motif 10 could be pivotal for the enzyme activity of the SHMT family. The analysis of conserved domains demonstrated that all five OsSHMT proteins contain the SHMT

domain, underscoring the conservative nature of their structure (Figure 3C). The divergence in exon–intron structure plays an important role in the evolution of duplicated genes. Then, we looked into the structure of the *OsSHMT* genes. The results showed that four *SHMT* genes contain four exons, but *OsSHMT1* has fifteen exons and all five *OsSHMT* genes have 3′ and 5′ UTR (Figure 3D). Although the majority of SHMT family members display some disparities in gene structure, substantial structural alterations have been noted in *OsSHMT1*, implying a differentiation mechanism during its evolution. The combined results of gene structure and phylogenetic trees indicate that OsSHMTs shared the SHMT conserved domain but underwent structural diversifications.

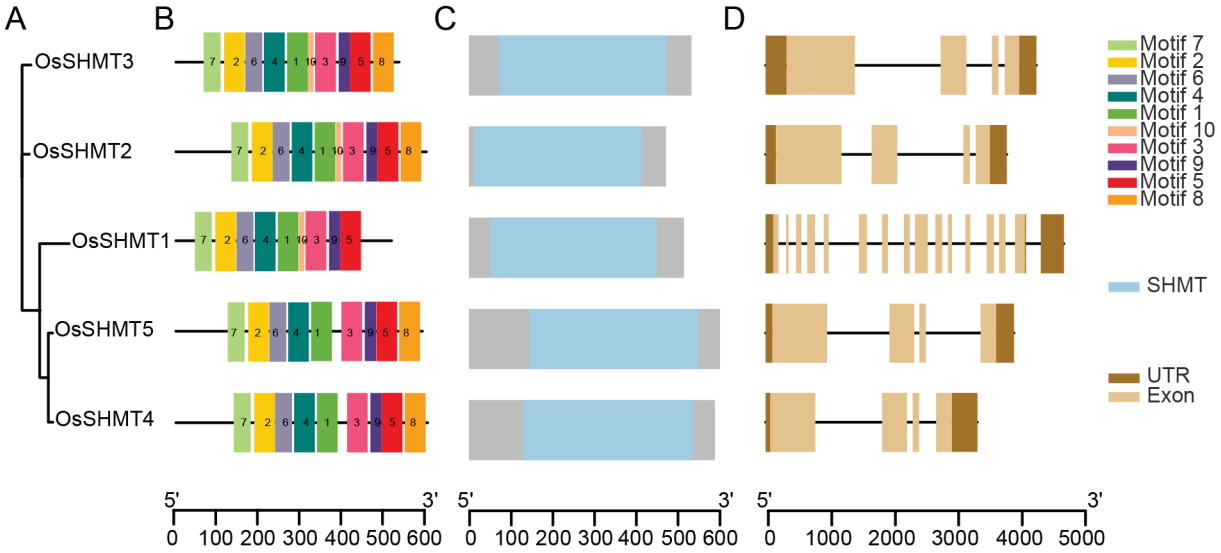

Figure 3. Phylogenetic relationships, conserved motifs, protein structure, and gene structure of *OsSHMTs*. (**A**) The phylogenetic tree was constructed using the maximum likelihood method with 1000 bootstrap replicates by MEGA 7. (**B**) The motif composition of SHMT proteins. The 10 motifs were indicated by colored boxes and numbered 1–10. (**C**) The conserved domain of OsSHMT proteins. Blue boxes for the SHMT domain. (**D**) The intron and exon distribution of *OsSHMT* genes. Light brown boxes indicate exons, dark brown boxes indicate untranslated sequence, and black lines indicate introns.

To elucidate the potential roles of *OsSHMT* family members in different responses and investigate their functions, a 2000 bp sequence upstream of the initial codons was extracted for a deeper understanding of the regulatory mechanisms of *OsSHMTs*. Our research revealed that in addition to light-responsive elements (such as G-Box, GT1-motif, and MRE), the promoter region of *OsSHMTs* contains a multitude of elements associated with growth, development, and responses to abiotic stress (Figure 4A, Table S5). Growth and developmental elements include a meristematic tissue expression element (CAT-box and CCGTCC-box) and an endosperm expression element (RY-element and GCN4 motif). Hormone response elements encompass those that react to ABA (ABRE), methyl jasmonate (CGTCA-motif and TGACG-motif), auxin (TGA-element), gibberellin (GARE-motif and TATC-box), and salicylic acid (TCA). The stress response elements include wound response elements (WUN and WRE3 motif), dehydration response elements (DRE), drought response elements (MBS, Myb/Myc binding sites), stress-response element (STRE), anaerobic induction elements (ARE), and low-temperature response elements (LTR). These *cis*-elements found in the promoters of the *OsSHMT* genes have the ability to regulate the stress responsiveness or tissue-specific expression of the *OsSHMT* gene across various environmental conditions.

Apart from *OsSHMT5*, the promoters of the remaining *OsSHMT* family members harbor numerous ABA response elements. Additionally, the promoters of *OsSHMT2*, *OsSHMT3*, and *OsSHMT4* feature methyl jasmonate response elements. Furthermore, the

promoters of *OsSHMT2* and *OsSHMT4* are rich in pressure response elements (Figure 4A). A large number of AP2/ERF family gene members can bind to the promoters of *OsSHMT2* and *OsSHMT4*, suggesting multiple AP2/ERF transcription factor binding sites within the promoter region. The MYB family was the only member predicted to be present in all five *OsSHMT* promoters (Figure 4B). These discoveries suggest that *OsSHMTs* may be influenced by various hormones involved in rice growth and its response to environmental stresses.

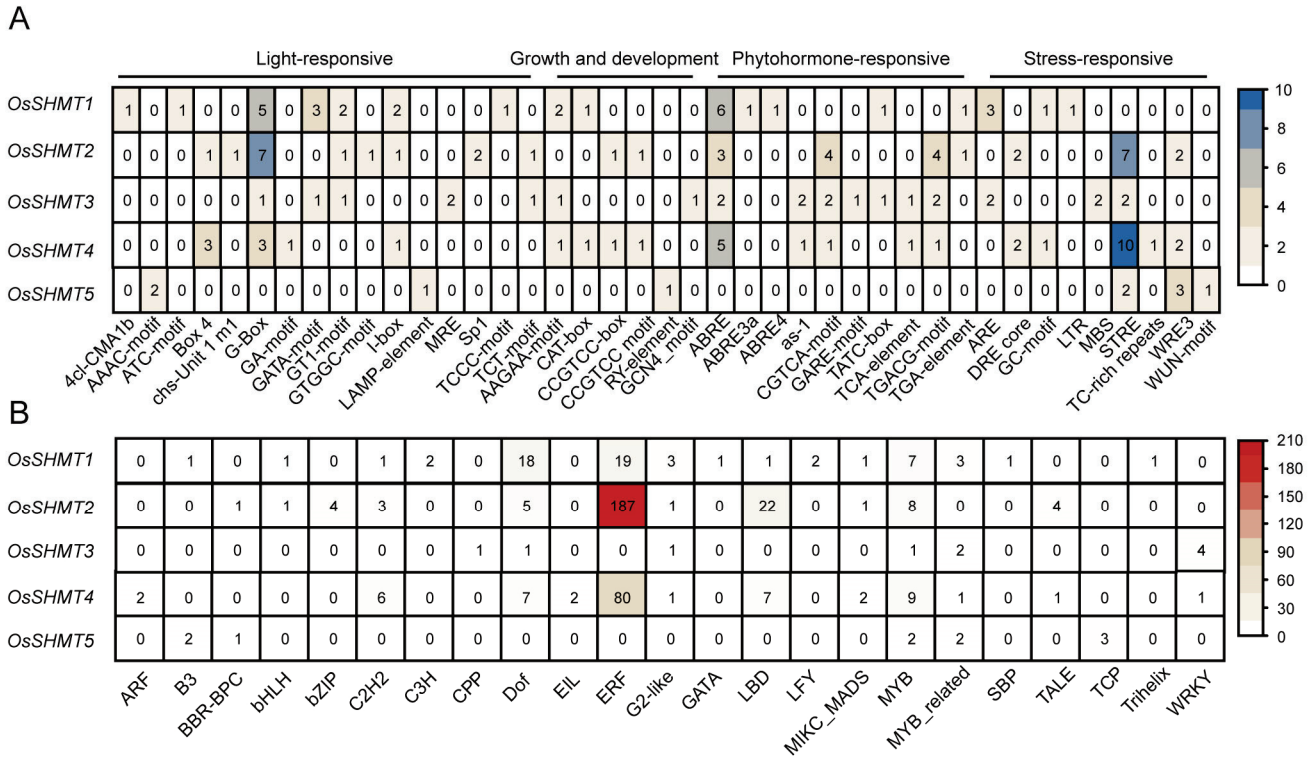

Figure 4. The prediction of *cis*-element and bound transcription factors on *OsSHMT* promoters. (**A**) The *cis*-elements distributed on the promoters of *OsSHMTs*. (**B**) Putative transcription factors binding to the *OsSHMT* promoters.

2.3. Expression Analysis of OsSHMTs in Various Tissues and in Response to Phytohormones

Considering that *cis*-acting elements are closely linked to gene expression, we first examined tissue-specific expression patterns for *OsSHMTs*. There is some variation in the expression patterns of homologous genes across different tissues. *OsSHMT1* was actively expressed in most young and mature tissues listed, especially in seeding roots and mature leaves. *OsSHMT2* only presented high expression in seeding roots, lateral inflorescence, and early seeds. *OsSHMT3* tends to be expressed in developed seeds and shares a similar pattern with OsSHMT4. The peak expression of *OsSHMT5* is shown in both young and mature leaves (Figure 5A). These results indicate that the expression of the *OsSHMTs* varies in different tissues to adapt to their functional differentiation.

As various phytohormone and stress responsive elements were distributed abundantly on *OsSHMT* promotors, we further analyzed the expression of *OsSHMTs* under different phytohormone treatments from 15 min to 6 h. The results showed that ABA treatment repressed the expression of *OsSHMTs*, while Jasmonic acid can induce the expression of *OsSHMT2*, *OsSHMT4*, and *OsSHMT5* and inhibit the expression of *OsSHMT3*, suggested that OsSHMT family proteins play an important role in biological and abiotic stress responses. *OsSHMT5* was suppressed by Cytokinin but induced by Auxin treatment, indicating that *OsSHMT5* may be involved in rice growth and development. Meanwhile, Cytokinin can increase the expression of *OsSHMT1* and *OsSHMT2* (Figure 5B). Treatment with different hormones triggers different changes in the expression of *OsSHMTs*, indicating that *Os-*

SHMTs can respond to multiple hormones, and there is significant functional differentiation among the homologous genes of *OsSHMTs*.

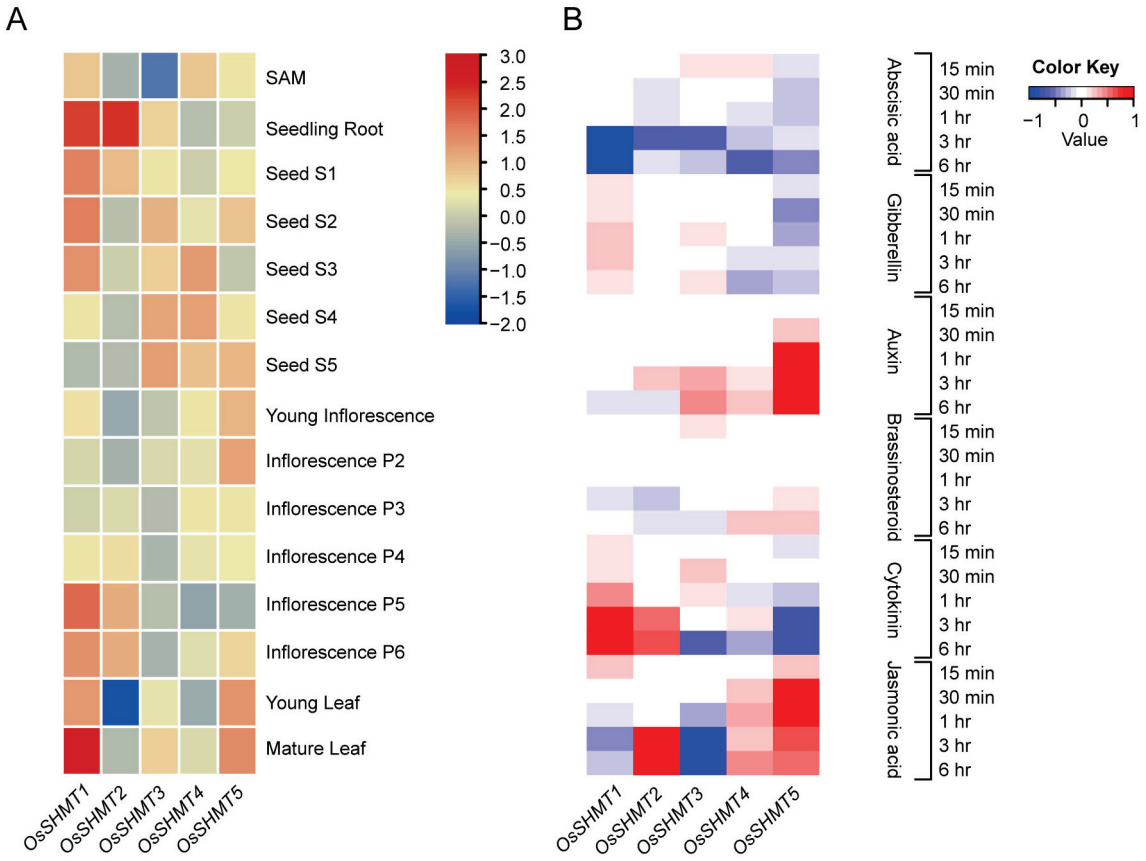

Figure 5. Expression analysis of *OsSHMT* genes during rice development and hormone treatment. (**A**) The expression profile of *OsSHMTs* in tissues. The samples were obtained from shoot apical meristems (SAM), seedling roots, five stages of developing seeds, six stages of developing panicles, young leaves, and mature leaves. The log10(FPKM) value was used to draw this heatmap. (**B**) The expression profile of *OsSHMT* genes under hormone treatment in rice seedling roots with a time course. The heatmap was downloaded from RiceXpro.

2.4. Subcellular Localization of OsSHMTs

Based on the evolutionary classification of OsSHMT family members, OsSHMT1 is localized in mitochondria, OsSHMT2 and OsSHMT3 are located in the cytoplasm, while OsSHMT4 and OsSHMT5 are located in the nucleus, and there are no homologous genes of SHMT localized in chloroplasts in rice. To verify the accuracy of the prediction results, we conducted subcellular localization analysis using rice protoplasts. We constructed the fused vector consisting of the coding sequence of *OsSHMTs* and GFP in the C-terminal driven by the cauliflower mosaic virus 35S (CaMV35S). Vectors to express the OsSHMTs-GFP fusion protein and empty-GFP (as a control) were introduced into rice protoplasts (Figure 6). Transiently expressed in the rice protoplasts, analysis demonstrated that OsSHMT1-GFP fusion protein presented both mitochondrial-like and cytosolic localization (Figure 6A), which is consistent with the previously reported results of the subcellular localization of OsSHMT1 [15]. Furthermore, OsSHMT2-GFP was localized in the cytosol of the transformed rice protoplasts (Figures 6B and S1A), whereas the OsSHMT4-GFP and OsSHMT5-GFP fusion proteins accumulated in the nucleus (Figures 6D,E and S1C,D), which is consistent with evolutionary classification. Remarkably, chloroplast-localized SHMTs are absent in monocotyledonous plants (Figure 2). Interestingly, OsSHMT3, classified within the cytoplasmic localization clade along with OsSHMT2, presented a chloroplastic-like localization

(Figures 6C and S1B). Taken together, the rice SHMT protein family is localized in the mitochondria, plastids, cytoplasm, and nuclei, indicating their diverse roles in metabolic pathways.

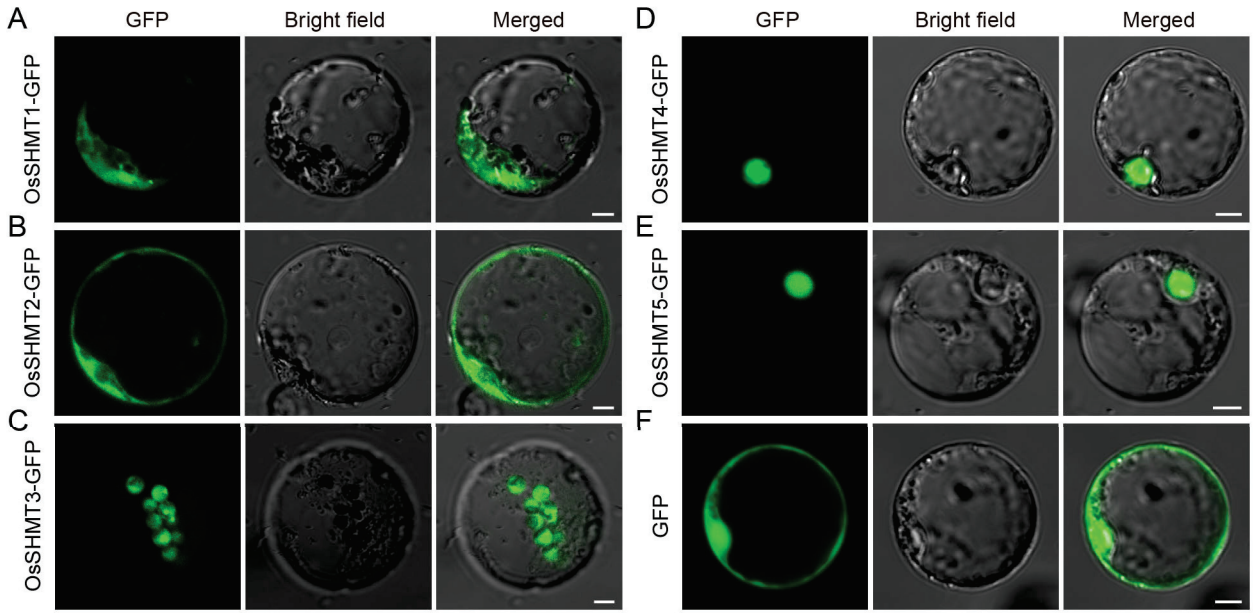

Figure 6. Subcellular localization of five OsSHMTs. (**A**–**F**) Confocal microscopic images showing the subcellular localization of five OsSHMTs in rice protoplast. GFP, fluorescence signals of the fusion proteins or free green fluorescent protein (GFP) proteins. Bars = 5 μm.

2.5. Protein Interaction Networks of OsSHMTs

SHMTs are important enzymes that catalyze the conversion of serine and glycine. To deeply comprehend the potential function of OsSHMT proteins and the interaction between its family members, a protein–protein interaction network was predicted to explore the potential interacting proteins of OsSHMTs. There were five proteins predicted to interact with all the five OsSHMTs (Figure 7). According to the annotation, all five proteins have catalytic activities. Among them, LOC_Os01g51410, LOC_Os04g53230, and LOC_Os06g40940 participated in organic acid and amino acid metabolism. LOC_Os01g51410 and LOC_Os06g40940 were predicted to be glycine decarboxylase and can affect net photosynthetic rates as a key enzyme in the photorespiratory cycle [30,31]. Previous studies have shown that SHMTs in mitochondria cooperate with the glycine decarboxylase complex (GDC) to mediate photorespiratory glycine–serine interconversion in plants [10]. LOC_09g27420 was predicted to encode formate–tetrahydrofolate ligase, which is involved in the one-carbon pool by folate (Table S7). Overall, proteins predicted to interact with OsSHMTs were annotated to be involved in carbon, glycine, and serine metabolism, which aligns with the overall functional role of OsSHMT proteins.

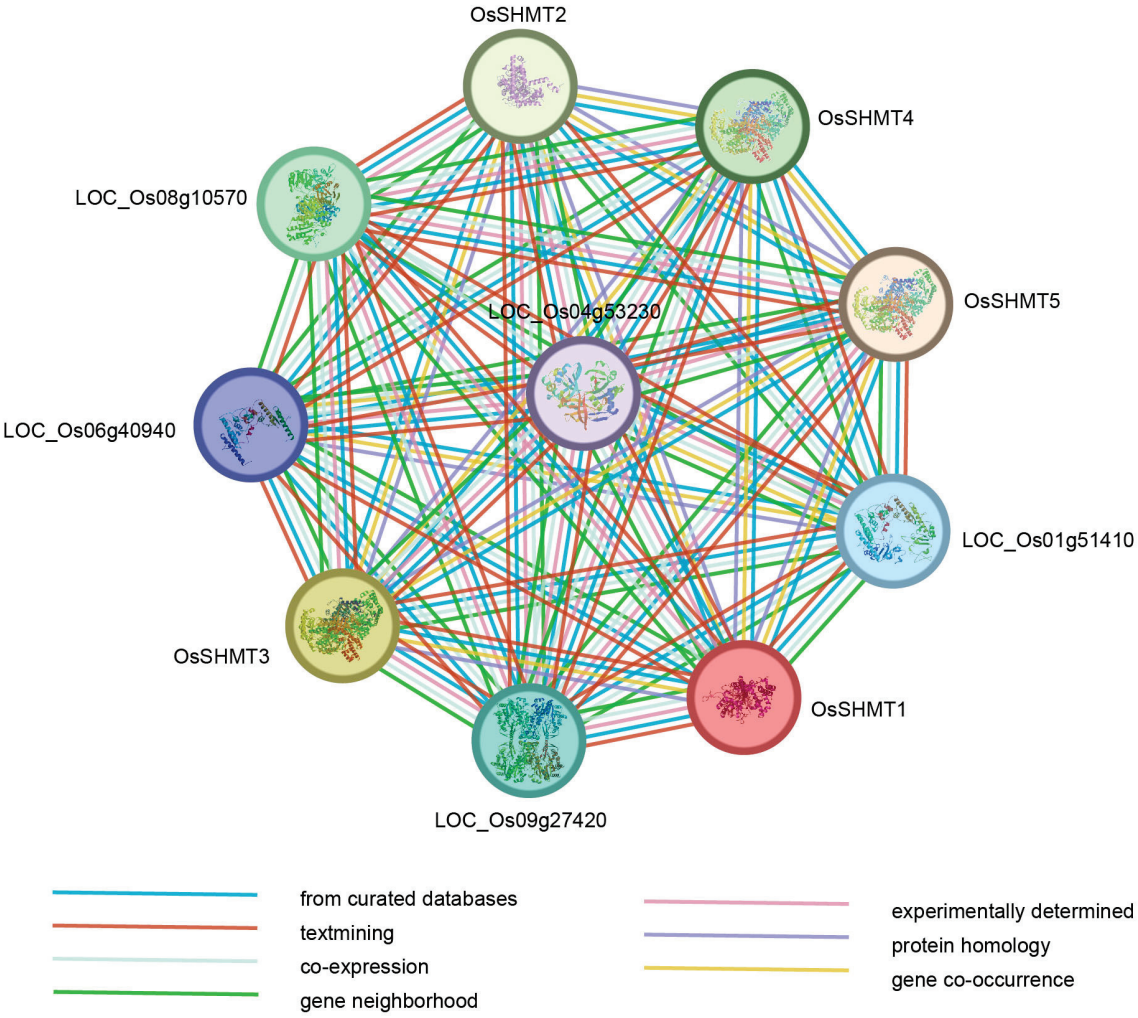

Figure 7. Protein interaction network of OsSHMTs. The nodes represent proteins, and the edges indicate the associations between two proteins. The types of evidence suggesting a functional link are distinguished by edges with different colors.

2.6. OsSHMTs Form Polymers

Studies have reported that the SHMT enzyme occurs as a tetramer in eukaryotes, because the active site is located at the interface of two monomers in an obligate homodimeric structure [32]. It is still uncertain whether SHMT proteins, localized in various organelles within plants, can form homologous or heterologous multimers. Based on the protein–protein interaction network, there is a potential for interaction among the homologous OsSHMT proteins. To further evaluate this possibility, we employed the yeast two-hybrid assay (Y2H) and found that all five OsSHMT proteins can interact with themselves (Figures 8A and S2). Interestingly, OsSHMT3 can interact with other OsSHMT proteins, except OsSHMT1 (Figures 8B and S2). In previous studies, we found that OsSHMT3 and OsSHMT4 can interact in the nucleus [26]. It seems that OsSHMT3 works in a similar way to SHMTs in animals, translocating in different organelles. OsSHMT3 may be the movable protein and cooperate with other OsSHMT proteins. Furthermore, OsSHMT4 and OsSHMT5 can also interact with each other.

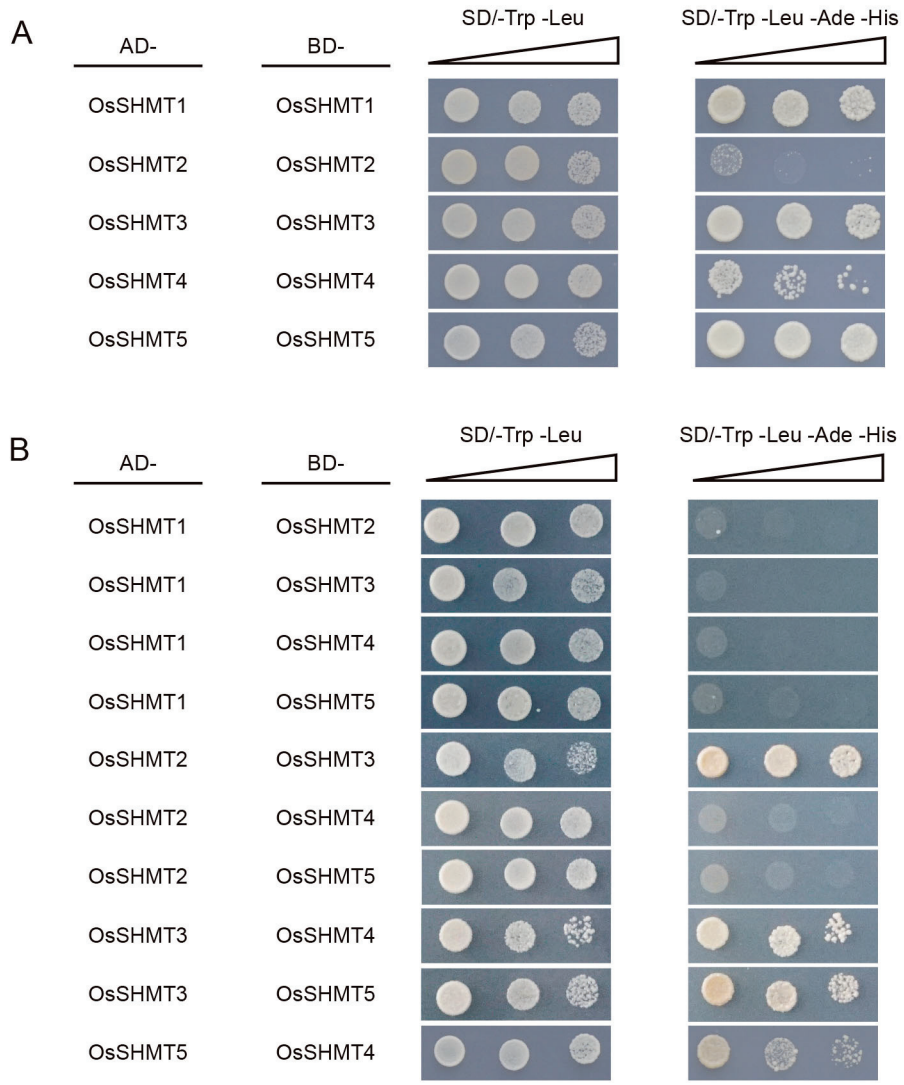

Figure 8. (**A**,**B**) Y2H assay shows that OsSHMTs interact with themselves (**A**) and their homologous proteins (**B**). DDO, SD/-Trp/-Leu; QDO, SD/-Trp/-Leu/-Ade/-His. AD, fused with activation domain; BD, fused with binding domain.

The crystal structure and protein interaction model provide a way to learn how the OsSHMT complex was formed. The crystal structure of OsSHMTs were predicted by Alphafold. HDOCK and PyMOL were applied to imitate and visualize the model [33–35]. The result shows that OsSHMT3 forms tight-binding obligate dimers with OsSHMT2, OsSHMT3, OsSHMT4, and OsSHMT5 (Figure 9 and Table S8). When OsSHMT3 interacts with OsSHMT2, the amino acids Ala75, Asp76, and Glu78 on OsSHMT3 form hydrogen bonds with Gln304 and Arg85 on OsSHMT2 (Figure 9A). Within the OsSHMT3 homodimer, a hydrogen bond is formed between Leu72 and Glu109 (Figure 9B). In the dimers of OsSHMT3 and OsSHMT4, the Thr175, Thr196, Thr210, and Try213 of OsSHMT3 are linked with Arg83 and Try86 of OsSHMT4 (Figure 9C). In the dimers of OsSHMT3 and OsSHMT5, the amino acids Asp76 and Glu78 on the former are linked with both Arg203 and Lys427 on OsSHMT5 (Figure 9D). The aforementioned findings indicate that the Leu72–Glu78 segment of OsSHMT3 plays an important role in the interaction between OsSHMT3 and its homologous proteins.

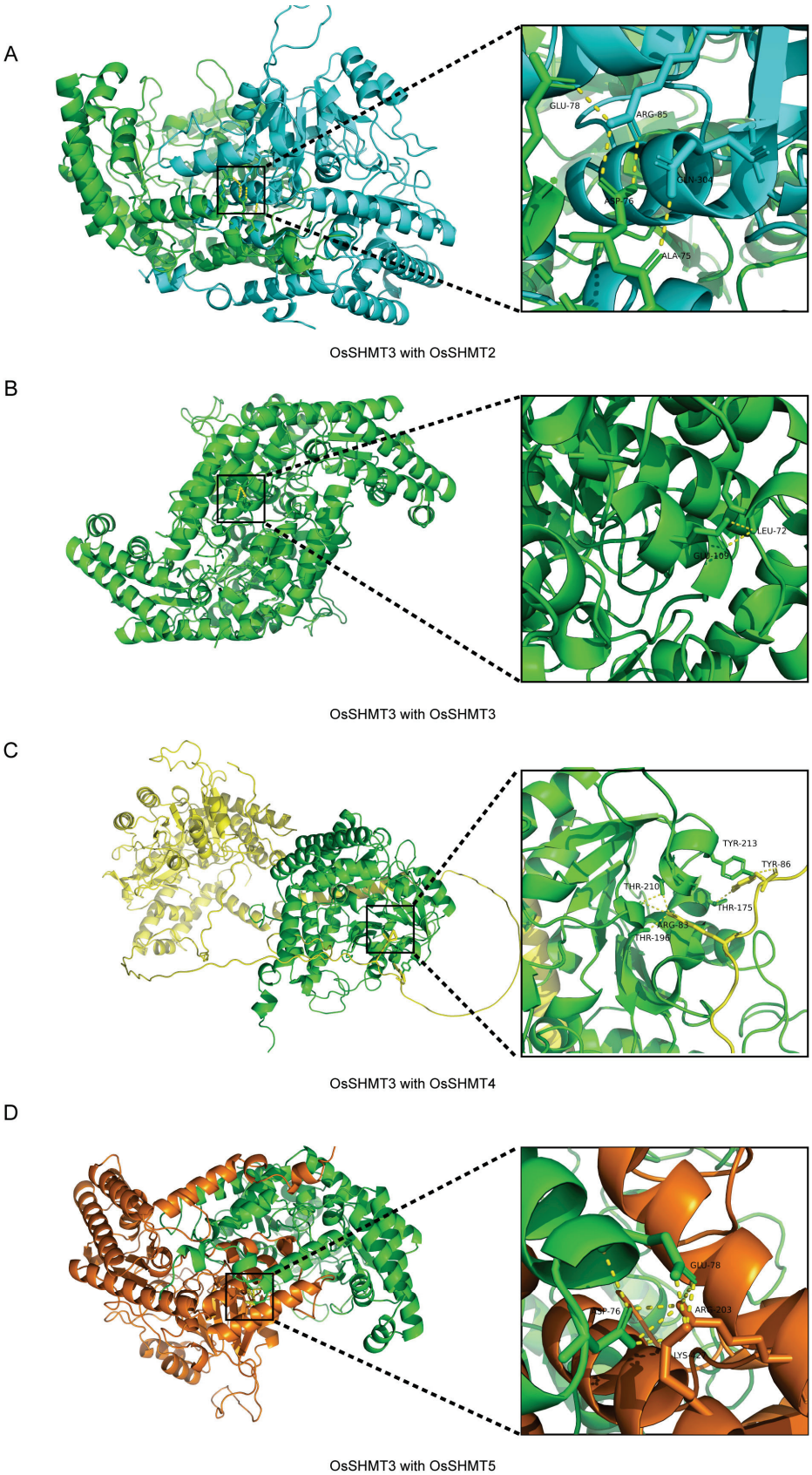

Figure 9. OsSHMT3 dimers and interfaces residues. (**A**–**D**) The interaction of OsSHMT3 with OsSHMT2 (**A**), OsSHMT3 (**B**), OsSHMT4 (**C**), OsSHMT5 (**D**). Differently colored proteins denoted different OsSHMT proteins. OsSHMT2 for blue, OsSHMT3 in green, OsSHMT4 in yellow, and OsSHMT5 in orange. Black squares show the zoomed interface of the dimers, and yellow lines represent the hydron bonds between interacting amino acids.

3. Discussion

3.1. The Functions of OsSHMTs Show Differentiation

Serine hydroxymethyltransferase is a crucial enzyme in the cell–carbon metabolic pathway, as it is responsible for the reversible interconversion of Ser and Gly [1]. Plants' SHMT proteins can be classified into four groups according to their predicted subcellular localization [36]. The five *SHMT* genes in rice are categorized into three groups: group Ia, group Ib, and group IIb, with group IIa (chloroplast group) genes being absent. Remarkably, a specific gene loss event occurred in group IIa within monocotyledons (Figures 1 and 2). The SHMT enzyme in plastids provides a carbon unit for the biosynthesis of 5,10-CH_2-$H_4PteGlu_n$, which can then be oxidized by other catalysts into 5,10-CH^+-$H_4PteGlu_n$ (5,10-methylene tetrahydrofolate) and 10-HCO-$H_4PteGlu_n$ (10-formyl tetrahydrofolate) [5]. The 5,10-CH^+-$H_4PteGlu_n$ serves as a vital light-harvesting cofactor in plastid-localized cryptochrome [37]. Therefore, SHMT activity in plastids has a potentially important role in light reception. The abundance of SHMT proteins in plant species surpasses those found in humans and fungi, indicating a greater complexity in plant SHMT proteins and potentially novel functions (Table S3). In monocotyledonous plants, alternative mechanisms may compensate for the absence of SHMT in group IIa. Interestingly, our subcellular localization analysis revealed that OsSHMT3 was located within chloroplasts in rice protoplasts (Figure 6), suggesting a gene duplication event within group Ib may compensated for this loss in group IIa. The aforementioned findings suggest that SHMT proteins in plants are more intricate and may possess novel functionalities. SHMTs usually function as homozygous or heterozygous complexes. The yeast two-hybrid assay results showed that all five OsSHMTs can form homodimers, and only OsSHMT3 can form dimers with other OsSHMTs, except OsSHMT1 (Figure 8). In silico docking simulations showed that at a certain angle, some residues can form hydron bonds shorter than 3 Å (Figure 9, Table S8). In short, these results indicated that OsSHMT3 may serve as a mobilized protein and cooperate with other proteins. While the functions of SHMT1 and SHMT4 in rice have been relatively well elucidated, the understanding of other SHMTs is still limited. Previous studies have indicated that OsSHMT1 and OsSHMT3 have enzyme activity, whereas OsSHMT4 and OsSHMT5 individually lacked enzyme activity in vitro [16,26]. It is unclear whether these SHMTs primarily act as enzymes or serve purely regulatory roles. However, this hypothesis needs further investigation.

3.2. OsSHMTs Emerge as Promising Candidates for Bolstering Plant Resistance against Both Abiotic and Biotic Stressors

As stationary organisms, plants face a multitude of environmental stresses. Essential to their survival are plant hormones, which govern responses to both biotic and abiotic challenges. Notably, the ABA and JA signaling pathways, along with their interconnections, play pivotal roles in enhancing plant resilience against both biotic and abiotic stresses [38,39]. Environmental stresses induce changes in hormone biosynthesis, initiating the binding of transcription factors to specific *cis*-acting elements within promoter regions of downstream stress-responsive genes. This regulatory process governs gene expression, ultimately fine-tuning plant tolerance in reaction to abiotic stresses. The AP2/ERF and MYB transcription factor families are ubiquitous among plants, playing pivotal roles in regulating responses to abiotic stresses while also influencing plant growth and development [40,41]. In our results, the promoters of *OsSHMT* family members contain numerous ABA and JA response elements (Figure 4A). Moreover, *OsSHMT* promoters are rich in pressure response elements, along with numerous bound AP2/ERF and MYB transcription factors (Figure 4B). The analysis of promoters and the observed responsiveness of *SHMTs* to ABA and JA suggest that OsSHMTs are likely to play a crucial role in plant stress responses, serving as downstream targets for AP2/ERF and MYB.

Prior research has highlighted the significant role of Arabidopsis and rice SHMT1 in regulating cellular damage induced by abiotic stress [13,17]. Moreover, the GmSHMT08 gene in soybeans has been associated with imparting resistance against soybean cyst

nematodes (SCN) [20]. OsSHMT4 negatively regulated the uptake and assimilation of mineral salt and improve rice resistance and grain quality [25]. Its homolog in wheat was found to be responsive to abiotic stress, abscisic acid, methyl jasmonate, hydrogen peroxide, and especially, Fusarium head blight [42]. *OsSHMT2* and *OsSHMT4* exhibit a significant abundance of *cis*-elements and transcription factor binding sites among the five *OsSHMT* genes (Figure 4). This suggests that *OsSHMT2* and *OsSHMT4* potentially hold greater significance in rice development. Hence, exploiting plant SHMTs to enhance crop resilience against both biological and abiotic stress represents a promising strategy. However, further investigation is required to uncover additional naturally occurring variations in *SHMT* genes associated with plant stress responses.

4. Materials and Methods

4.1. Identification and Phylogenetic Analysis of SHMT Family Genes in Rice Genome

The protein sequence of *Oryza sativa*, *Arabidopsis*, *Solanum lycopersicum*, *Zea mays*, and *Glysine Max* were downloaded from the phytozome database (https://phytozome-next.jgi.doe.gov/, accessed on 11 November 2023). The hidden Markov model (HMM) of the SHMT domain was obtained from Pfam (http://pfam-legacy.xfam.org/, accessed on 11 November 2023) with the ID of PF00464. All proteins were searched against the SHMT domain and then checked by NCBI (https://www.ncbi.nlm.nih.gov/, accessed on 11 November 2023). The properties of the OsSHMT proteins were predicted by Expasy (https://www.expasy.org/, accessed on 11 November 2023). Then, all the SHMT proteins were used to build the SHMT phylogenetic tree using MAFFT (version 7.380) [43], FastTree (version 2.1.7) [44], and MEGA 7 (version 7.0.9) [45] with default parameters. Proteins used to build the phylogenetic tree were listed in Table S2. The tree was embellished by the online tool Evolview (https://www.evolgenius.info/evolview/#/, accessed on 25 November 2023) [46]. The aligned sequences of SHMT from different species were obtained from Gramene (https://plants.ensembl.org/Oryza_sativa/Info/Index, accessed on 5 January 2024), the tree was built by the maximal likelihood method using Fasttree and visualized by iTOL (https://itol.embl.de/, accessed on 19 January 2024)

4.2. Gene Structure Analysis of OsSHMTs

The gff file and genome sequence file of rice and corn were downloaded from phytozome (https://phytozome-next.jgi.doe.gov/, accessed on 11 November 2023). Protein and full-length coding sequences of *OsSHMTs* were analyzed by MEME (https://meme-suite.org/meme/, accessed on 25 November 2023) [47] and NCBI (https://www.ncbi.nlm.nih.gov/cdd, accessed on 25 November 2023), respectively. The visualization of the gene structure analysis was accomplished by TBtools-II (version 1.098667) [48].

4.3. Predicted Cis-Element and Transcription Factor Binding Sites on OsSHMT Promoters

The 2 kb upstream elements of the initiation codon were extracted as promoters. Then, Plantcare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 10 December 2023) [49] and PlantTFDB (http://planttfdb.gao-lab.org/prediction.php, accessed on 10 December 2023) [50] were used for *cis*-element and transcription factor binding sites, respectively, with default parameters. Excel and TBtools-II [48] were used to count and visualize the results.

4.4. Expression Analysis

The tissue-specific expression data were obtained from Rice Genome Annotation Project (http://rice.uga.edu/expression.shtml, accessed on 23 January 2024). The original expression data can be found in Table S6. TBtools [48] was used to draw the heatmap with a $\log_{10}$(FPKM) value. The expression data of hormone treatment on 7-day-old seedling roots were downloaded from RiceXpro (https://ricexpro.dna.affrc.go.jp/data-set.html, accessend o 23 January 2024).

4.5. Protein–protein Interaction Prediction

The protein sequences of OsSHMTs were used for protein–protein interaction analysis by STRING (https://cn.string-db.org/, accessed on 25 January 2024) with default parameters. Gene annotation was also downloaded from STRING.

4.6. Subcellular Localization Analysis

The *OsSHMT* coding sequences were cloned into pAN580 vector with a GFP tag and driven by the double 35S (d35S) promoter. The primers used are listed in Table S9. Rice seedlings were germinated and grown in half-strength MS liquid media within a growth chamber (14 h/30 °C light and 10 h/25 °C darkness) for 10 days before being used to generate protoplasts. Protoplast transformation was conducted following previously established procedures [51]. The OsSHMTs-GFP construct was introduced into the protoplasts, which were then incubated in darkness at 28 °C for 16 h prior to examination. Localization of the fusion protein used was determined under a confocal laser scanning microscope (LSM980 laser, Carl Zeiss, Oberkochen, Germany). GFP signals were recorded with the excitation wavelength at 488 nm and emission wavelength ranging from 505 to 530 nm, while mCherry signals were recorded with the excitation wavelength at 587 nm and emission wavelength ranging from 600 to 630 nm.

4.7. Yeast Two-Hybrid

The GAL4 system was selected for yeast two-hybridization. The coding regions of *OsSHMTs* were linked into pGBKT7 vector and pGADT7 vector, respectively. The constructed plasmids were co-transfected into yeast AH109 and cultured on SD/-Leu/-Trp solid medium and then transferred to SD/-Leu/-Trp/-His/-Ade solid medium to verify the interactions (the relevant vectors and reagents were obtained from Clontech). The primers used are listed in Table S9.

4.8. Protein Structure Prediction

Protein sequences of OsSHMT were loaded on alphafold2.0 (https://colab.research.google.com/github/sokrypton/ColabFold/blob/main/AlphaFold2.ipynb, accessed on 5 February 2024) for the three-dimensional structure prediction. Proteins for molecular docking simulations were carried out by HDOCK (http://hdock.phys.hust.edu.cn/, accessed on 5 February 2024) and visualized by PyMOL (version 2.32).

5. Conclusions

In this study, a genome-wide analysis of the *SHMT* gene family in rice was conducted. We examined the physical and chemical characteristics, phylogenetic relationships, gene structure, motif scanning, promoter *cis*-elements, subcellular localization, and the protein interactions of the five SHMT proteins. Phylogenetic analysis revealed that the five OsSHMT proteins could be classified into three groups. Gene structure analysis showed that *OsSHMT1* lacks one motif but contains more introns compared to other *OsSHMTs*. The results from *cis*-acting element prediction and expression analysis suggested that *SHMT* family members may play various roles in multiple stress responses and hormone regulation. Subcellular localization analysis indicated that OsSHMT1–5 are localized in the mitochondria, cytoplasm, chloroplast, and nucleus, respectively. Protein interaction analysis demonstrated that OsSHMT3 can form both homodimers and heterodimers. The accumulation of variations during the evolution of OsSHMT led to neofunctionalization and subfunctionalization events. Our data will serve as a crucial resource for advancing research on the functional aspects of *OsSHMT* genes.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/plants13081116/s1, Figure S1: Subcellular localization of OsSHMTs in rice protoplast; Figure S2: Negative controls for the Y2H assay; Table S1: The characteristics of *SHMT* genes in rice; Table S2: *SHMT* genes used to construct the phylogenetic tree; Table S3: Number

of *SHMT* genes identified in selected species; Table S4: The details of conserved protein motifs; Table S5: The predictive *cis*-acting elements of *OsSHMT* genes; Table S6: Original expression data of *OsSHMTs* in tissues; Table S7: Interacting protein function annotation; Table S8: Interface residues of OsSHMT3 dimers; Table S9: Primers used in this study.

Author Contributions: T.P. and M.Y. conceived the idea and designed the experiments. T.P., H.J. and C.Z. performed the experiments. H.J. and M.Y. prepared the materials. T.P. and H.J. wrote the manuscript, T.P. and M.Y. finalized the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This work was supported by the Program for Research and Development of Zhejiang A&F University (2022LFR109).

Data Availability Statement: The original contributions presented in the study are included in the article/Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Hanson, A.D.; Roje, S. One-Carbon Metabolism in Higher Plants. *Annu. Rev. Plant Physiol. Plant Mol. Biol.* **2001**, *52*, 119–137. [CrossRef] [PubMed]
2. Mentch, S.J.; Locasale, J.W. One-Carbon Metabolism and Epigenetics: Understanding the Specificity. *Ann. N. Y. Acad. Sci.* **2016**, *1363*, 91–98. [CrossRef] [PubMed]
3. Zrenner, R.; Stitt, M.; Sonnewald, U.; Boldt, R. Pyrimidine and Purine Biosynthesis and Degradation in Plants. *Annu. Rev. Plant Biol.* **2006**, *57*, 805–836. [CrossRef] [PubMed]
4. Schulze, S.; Westhoff, P.; Gowik, U. Glycine Decarboxylase in C3, C4 and C3–C4 Intermediate Species. *Curr. Opin. Plant Biol.* **2016**, *31*, 29–35. [CrossRef]
5. Hanson, A.D.; Gage, D.A.; Shachar-Hill, Y. Plant One-Carbon Metabolism and Its Engineering. *Trends Plant Sci.* **2000**, *5*, 206–213. [CrossRef]
6. Huang, X.Y.; Chao, D.Y.; Koprivova, A.; Danku, J.; Wirtz, M.; Müller, S.; Sandoval, F.J.; Bauwe, H.; Roje, S.; Dilkes, B.; et al. Nuclear Localised More Sulphur Accumulation1 Epigenetically Regulates Sulphur Homeostasis in Arabidopsis Thaliana. *PLoS Genet.* **2016**, *12*, e1006298. [CrossRef] [PubMed]
7. Zhang, Y.; Sun, K.; Sandoval, F.J.; Santiago, K.; Roje, S. One-Carbon Metabolism in Plants: Characterization of a Plastid Serine Hydroxymethyltransferase. *Biochem. J.* **2010**, *430*, 97–105. [CrossRef] [PubMed]
8. Wei, Z.; Sun, K.; Sandoval, F.J.; Cross, J.M.; Gordon, C.; Kang, C.; Roje, S. Folate Polyglutamylation Eliminates Dependence of Activity on Enzyme Concentration in Mitochondrial Serine Hydroxymethyltransferases from Arabidopsis Thaliana. *Arch. Biochem. Biophys.* **2013**, *536*, 87–96. [CrossRef] [PubMed]
9. Douce, R.; Neuburger, M. Biochemical Dissection of Photorespiration. *Curr. Opin. Plant Biol.* **1999**, *2*, 214–222. [CrossRef] [PubMed]
10. Rajinikanth, M.; Harding, S.A.; Tsai, C.J. The Glycine Decarboxylase Complex Multienzyme Family in Populus. *J. Exp. Bot.* **2007**, *58*, 1761–1770. [CrossRef]
11. Somerville, C.R.; Ogren, W.L. Photorespiration-Deficient Mutants of Arabidopsis Thaliana Lacking Mitochondrial Serine Transhydroxymethylase Activity. *Plant Physiol.* **1981**, *67*, 666–671. [CrossRef] [PubMed]
12. Voll, L.M.; Jamai, A.; Renné, P.; Voll, H.; McClung, C.R.; Weber, A.P.M. The Photorespiratory Arabidopsis Shm1 Mutant Is Deficient in SHM1. *Plant Physiol.* **2006**, *140*, 59–66. [CrossRef]
13. Moreno, J.I.; Martín, R.; Castresana, C. Arabidopsis SHMT1, a Serine Hydroxymethyltransferase That Functions in the Photorespiratory Pathway Influences Resistance to Biotic and Abiotic Stress. *Plant J.* **2005**, *41*, 451–463. [CrossRef]
14. Liu, Y.; Mauve, C.; Lamothe-Sibold, M.; Guérard, F.; Glab, N.; Hodges, M.; Jossier, M. Photorespiratory Serine Hydroxymethyltransferase 1 Activity Impacts Abiotic Stress Tolerance and Stomatal Closure. *Plant Cell Environ.* **2019**, *42*, 2567–2583. [CrossRef]
15. Wu, J.; Zhang, Z.; Zhang, Q.; Han, X.; Gu, X.; Lu, T. The Molecular Cloning and Clarification of a Photorespiratory Mutant, Oscdm1, Using Enhancer Trapping. *Front. Genet.* **2015**, *6*, 226. [CrossRef]
16. Mishra, P.; Jain, A.; Takabe, T.; Tanaka, Y.; Negi, M.; Singh, N.; Jain, N.; Mishra, V.; Maniraj, R.; Krishnamurthy, S.L.; et al. Heterologous Expression of Serine Hydroxymethyltransferase-3 from Rice Confers Tolerance to Salinity Stress in *E. coli* and Arabidopsis. *Front. Plant Sci.* **2019**, *10*, 217. [CrossRef] [PubMed]
17. Fang, C.; Zhang, P.; Li, L.; Yang, L.; Mu, D.; Yan, X.; Li, Z.; Lin, W. Serine Hydroxymethyltransferase Localised in the Endoplasmic Reticulum Plays a Role in Scavenging H_2O_2 to Enhance Rice Chilling Tolerance. *BMC Plant Biol.* **2020**, *20*, 236. [CrossRef]
18. Besson, V.; Neuburger, M.; Rébeillé, F.; Douce, R. Evidence for Three Serine Hydroxymethyltransferases in Green Leaf Cells. Purification and Characterization of the Mitochondrial and Chloroplastic Isoforms. *Plant Physiol. Biochem.* **1995**, *33*, 665–673.
19. Christensen, K.E.; MacKenzie, R.E. Mitochondrial One-Carbon Metabolism Is Adapted to the Specific Needs of Yeast, Plants and Mammals. *Bioessays* **2006**, *28*, 595–605. [CrossRef]

20. Liu, S.; Kandoth, P.K.; Warren, S.D.; Yeckel, G.; Heinz, R.; Alden, J.; Yang, C.; Jamai, A.; El-Mellouki, T.; Juvale, P.S.; et al. A Soybean Cyst Nematode Resistance Gene Points to a New Mechanism of Plant Resistance to Pathogens. *Nature* **2012**, *492*, 256–260. [CrossRef]
21. Rambani, A.; Pantalone, V.; Yang, S.; Rice, J.H.; Song, Q.; Mazarei, M.; Arelli, P.R.; Meksem, K.; Stewart, C.N.; Hewezi, T. Identification of Introduced and Stably Inherited DNA Methylation Variants in Soybean Associated with Soybean Cyst Nematode Parasitism. *New Phytol.* **2020**, *227*, 168–184. [CrossRef] [PubMed]
22. Lakhssassi, N.; Piya, S.; Knizia, D.; El Baze, A.; Cullen, M.A.; Meksem, J.; Lakhssassi, A.; Hewezi, T.; Meksem, K. Mutations at the Serine Hydroxymethyltransferase Impact Its Interaction with a Soluble NSF Attachment Protein and a Pathogenesis-Related Protein in Soybean. *Vaccines* **2020**, *8*, 349. [CrossRef] [PubMed]
23. Lakhssassi, N.; Piya, S.; Bekal, S.; Liu, S.; Zhou, Z.; Bergounioux, C.; Miao, L.; Meksem, J.; Lakhssassi, A.; Jones, K.; et al. A Pathogenesis-Related Protein GmPR08-Bet VI Promotes a Molecular Interaction between the GmSHMT08 and GmSNAP18 in Resistance to Heterodera Glycines. *Plant Biotechnol. J.* **2020**, *18*, 1810–1829. [CrossRef] [PubMed]
24. Lakhssassi, N.; Knizia, D.; El Baze, A.; Lakhssassi, A.; Meksem, J.; Meksem, K. Proteomic, Transcriptomic, Mutational, and Functional Assays Reveal the Involvement of Both THF and PLP Sites at the GmSHMT08 in Resistance to Soybean Cyst Nematode. *Int. J. Mol. Sci.* **2022**, *23*, 11278. [CrossRef] [PubMed]
25. Chen, J.; Huang, X.Y.; Salt, D.E.; Zhao, F.J. Mutation in *OsCADT1* Enhances Cadmium Tolerance and Enriches Selenium in Rice Grain. *New Phytol.* **2020**, *226*, 838–850. [CrossRef] [PubMed]
26. Yan, M.; Pan, T.; Zhu, Y.; Jiang, X.; Yu, M.; Wang, R.; Zhang, F.; Luo, S.; Bao, X.; Chen, Y.; et al. *FLOURY ENDOSPERM20* Encoding SHMT4 Is Required for Rice Endosperm Development. *Plant Biotechnol. J.* **2022**, *20*, 1438–1440. [CrossRef]
27. Anderson, D.D.; Woeller, C.F.; Chiang, E.P.; Shane, B.; Stover, P.J. Serine Hydroxymethyltransferase Anchors de Novo Thymidylate Synthesis Pathway to Nuclear Lamina for DNA Synthesis. *J. Biol. Chem.* **2012**, *287*, 7051–7062. [CrossRef]
28. Anderson, D.D.; Woeller, C.F.; Stover, P.J. Small Ubiquitin-like Modifier-1 (SUMO-1) Modification of Thymidylate Synthase and Dihydrofolate Reductase. *Clin. Chem. Lab. Med.* **2007**, *45*, 1760–1763. [CrossRef]
29. Woeller, C.F.; Anderson, D.D.; Szebenyi, D.M.; Stover, P.J. Evidence for Small Ubiquitin-like Modifier-Dependent Nuclear Import of the Thymidylate Biosynthesis Pathway. *J. Biol. Chem.* **2007**, *282*, 17623–17631. [CrossRef]
30. Timm, S.; Bauwe, H. The Variety of Photorespiratory Phenotypes—Employing the Current Status for Future Research Directions on Photorespiration. *Plant Biol.* **2013**, *15*, 737–747. [CrossRef]
31. Zhou, Q.; Yu, Q.; Wang, Z.; Pan, Y.; Lv, W.; Zhu, L.; Chen, R.; He, G. Knockdown of GDCH Gene Reveals Reactive Oxygen Species-Induced Leaf Senescence in Rice. *Plant Cell Environ.* **2013**, *36*, 1476–1489. [CrossRef]
32. Schirch, V.; Szebenyi, D.M. Serine Hydroxymethyltransferase Revisited. *Curr. Opin. Chem. Biol.* **2005**, *9*, 482–487. [CrossRef]
33. Yan, Y.; Tao, H.; He, J.; Huang, S.Y. The HDOCK Server for Integrated Protein–Protein Docking. *Nat. Protoc.* **2020**, *15*, 1829–1852. [CrossRef]
34. Mooers, B.H.M. Shortcuts for Faster Image Creation in PyMOL. *Protein Sci.* **2020**, *29*, 268–276. [CrossRef]
35. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly Accurate Protein Structure Prediction with AlphaFold. *Nature* **2021**, *596*, 583–589. [CrossRef] [PubMed]
36. Lakhssassi, N.; Patil, G.; Piya, S.; Zhou, Z.; Baharlouei, A.; Kassem, M.A.; Lightfoot, D.A.; Hewezi, T.; Barakat, A.; Nguyen, H.T.; et al. Genome Reorganization of the *GmSHMT* Gene Family in Soybean Showed a Lack of Functional Redundancy in Resistance to Soybean Cyst Nematode. *Sci. Rep.* **2019**, *9*, 1506. [CrossRef] [PubMed]
37. Huang, Y.; Baxter, R.; Smith, B.S.; Partch, C.L.; Colbert, C.L.; Deisenhofer, J. Crystal Structure of Cryptochrome 3 from *Arabidopsis Thaliana* and Its Implications for Photolyase Activity. *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 17701–17706. [CrossRef]
38. Wang, Y.; Mostafa, S.; Zeng, W.; Jin, B. Function and Mechanism of Jasmonic Acid in Plant Responses to Abiotic and Biotic Stresses. *Int. J. Mol. Sci.* **2021**, *22*, 8568. [CrossRef] [PubMed]
39. Vanitha, P.A.; Vijayaraghavareddy, P.; Uttarkar, A.; Dawane, A.; Sujitha, D.; Ashwin, V.; Babitha, K.C.; Niranjan, V.; Sheshshayee, M.S.; Anuradha, C.V.; et al. Novel Small Molecules Targeting BZIP23 TF Improve Stomatal Conductance and Photosynthesis under Mild Drought Stress by Regulating ABA. *FEBS J.* **2022**, *289*, 6058–6077.
40. Ma, Z.; Hu, L.; Jiang, W. Understanding AP2/ERF Transcription Factor Responses and Tolerance to Various Abiotic Stresses in Plants: A Comprehensive Review. *Int. J. Mol. Sci.* **2024**, *25*, 893. [CrossRef]
41. Wang, X.; Niu, Y.; Zheng, Y. Multiple Functions of Myb Transcription Factors in Abiotic Stress Responses. *Int. J. Mol. Sci.* **2021**, *22*, 6125. [CrossRef]
42. Hu, P.; Song, P.; Xu, J.; Wei, Q.; Tao, Y.; Ren, Y.; Yu, Y.; Li, D.; Hu, H.; Li, C. Genome-Wide Analysis of Serine Hydroxymethyltransferase Genes in Triticeae Species Reveals That *TaSHMT3A-1* Regulates Fusarium Head Blight Resistance in Wheat. *Front. Plant Sci.* **2022**, *13*, 847087. [CrossRef]
43. Katoh, K.; Misawa, K.; Kuma, K.-I.; Miyata, T. MAFFT: A Novel Method for Rapid Multiple Sequence Alignment Based on Fast Fourier Transform. *Nucleic Acids Res.* **2002**, *30*, 3059–3066. [CrossRef] [PubMed]
44. Price, M.N.; Dehal, P.S.; Arkin, A.P. Fasttree: Computing Large Minimum Evolution Trees with Profiles Instead of a Distance Matrix. *Mol. Biol. Evol.* **2009**, *26*, 1641–1650. [CrossRef]
45. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef] [PubMed]

46. Zhang, H.; Gao, S.; Lercher, M.J.; Hu, S.; Chen, W.H. EvolView, an Online Tool for Visualizing, Annotating and Managing Phylogenetic Trees. *Nucleic Acids Res.* **2012**, *40*, W569–W572. [CrossRef]
47. Bailey, T.L.; Boden, M.; Buske, F.A.; Frith, M.; Grant, C.E.; Clementi, L.; Ren, J.; Li, W.W.; Noble, W.S. MEME Suite: Tools for Motif Discovery and Searching. *Nucleic Acids Res.* **2009**, *37*, W202–W208. [CrossRef] [PubMed]
48. Chen, C.; Wu, Y.; Li, J.; Wang, X.; Zeng, Z.; Xu, J.; Liu, Y.; Feng, J.; Chen, H.; He, Y.; et al. TBtools-II: A "One for All, All for One" Bioinformatics Platform for Biological Big-Data Mining. *Mol. Plant* **2023**, *16*, 1733–1742. [CrossRef]
49. Lescot, M.; Déhais, P.; Thijs, G.; Marchal, K.; Moreau, Y.; Van De Peer, Y.; Rouzé, P.; Rombauts, S. PlantCARE, a Database of Plant Cis-Acting Regulatory Elements and a Portal to Tools for in Silico Analysis of Promoter Sequences. *Nucleic Acids Res.* **2002**, *30*, 325–327. [CrossRef]
50. Jin, J.; Tian, F.; Yang, D.C.; Meng, Y.Q.; Kong, L.; Luo, J.; Gao, G. PlantTFDB 4.0: Toward a Central Hub for Transcription Factors and Regulatory Interactions in Plants. *Nucleic Acids Res.* **2017**, *45*, D1040–D1045. [CrossRef]
51. Chen, S.; Tao, L.; Zeng, L.; Vega-Sanchez, M.E.; Umemura, K.; Wang, G.L. A highly efficient transient protoplast system for analyzing defence gene expression and protein-protein interactions in rice. *Mol. Plant Pathol.* **2006**, *7*, 417–427. [CrossRef] [PubMed]

plants

MDPI

Article

Genome-Wide Analysis of the SAUR Gene Family and Its Expression Profiles in Response to Salt Stress in *Santalum album*

Qing Zhu [1,†], Haoyue Zheng [1,†], Xu Hu [1,†], Yi Liu [1], Xinyi Zheng [1], Libei Li [2,*] and Minqiang Tang [1,*]

[1] Key Laboratory of Genetics and Germplasm Innovation of Tropical Special Forest Trees and Ornamental Plants (Ministry of Education), Hainan Key Laboratory for Biology of Tropical Ornamental Plant Germplasm, Collaborative Innovation Center of Ecological Civilization, School of Tropical Agriculture and Forestry, Hainan University, Haikou 570228, China
[2] College of Advanced Agriculture Sciences, Zhejiang A&F University, Hangzhou 311300, China
* Correspondence: libeili@zafu.edu.cn (L.L.); tangminqiang@hainanu.edu.cn (M.T.)
† These authors contributed equally to this work.

Abstract: The SAUR (small auxin-up RNA) family constitutes a category of genes that promptly respond to the hormone auxin and play a pivotal role in diverse biological processes encompassing plant growth and the response to abiotic stress. *Santalum album* L., a semi-parasitic evergreen tree, is renowned for its economically valuable essential oils, positioning it among the most prized tree species. In this study, a meticulous identification and comprehensive analysis of 43 *SAUR* genes was conducted within *S. album*. Based on phylogenetic relationships, the *SaSAUR* genes were systematically categorized into five groups. A collinearity analysis revealed intriguing insights, disclosing 14 segmental duplications and 9 tandem duplications within the *SaSAUR* genes, emphasizing the pivotal role of duplication in the expansion of this gene family. Noteworthy variations in the expression levels of *SaSAUR* genes were observed by delving into the *SaSAUR* transcriptome data from various tissues, including leaves, roots, and heartwood, as well as under salt-stress conditions. Notably, *SaSAUR08* and *SaSAUR13* were significantly upregulated in heartwood compared with roots and leaves, while *SaSAUR18* was markedly more expressed in roots compared with heartwood and leaves. Furthermore, *SaSAUR27* and *SaSAUR28* were found to respond closely to salt stress, hinting at their potential involvement in the salt-stress response mechanism. This research offers a comprehensive investigation of *SAUR* genes in *S. album* and establishes a foundation for future exploration of the *SAUR* gene family, particularly its relation to growth and salt-stress responses.

Keywords: SAUR gene family; *Santalum album*; transcriptome analysis; salt stress; RT-qPCR

Citation: Zhu, Q.; Zheng, H.; Hu, X.; Liu, Y.; Zheng, X.; Li, L.; Tang, M. Genome-Wide Analysis of the SAUR Gene Family and Its Expression Profiles in Response to Salt Stress in *Santalum album*. *Plants* **2024**, *13*, 1286. https://doi.org/10.3390/plants13101286

Academic Editors: Jian Zhang, Jiezheng Ying, Yifeng Wang and Jie Huang

Received: 10 April 2024
Revised: 1 May 2024
Accepted: 5 May 2024
Published: 7 May 2024

1. Introduction

Santalum album L. is an evergreen semi-parasitic tree belonging to the Santalum genus in the Santalaceae family and is mainly distributed in Southeast Asia and the Pacific region [1]. As *S. album* is one of the tree species with remarkable economic worth, all its parts can be utilized as raw materials. The heartwood oil is renowned for its unique aroma [2] and has vital roles as an anti-cancer [3], anti-inflammatory [4], and antioxidative substance [5]. Therefore, it has widespread application in various industries, including in perfumery, pharmaceuticals, cosmetics, etc. Heartwood oil is often referred to as "liquid gold" [6]. *S. album*'s heartwood possesses a remarkable hardness and a beautiful texture, making it one of the important raw materials for wood carving and the production of expensive furniture [7]. Therefore, *S. album* currently stands as one of the most widely commercially developed and utilized tropical tree species. In recent years, the demand for its heartwood and sandalwood essential oils has been increasing in both domestic and foreign markets; however, the growth rate and heartwood formation of *S. album* is relatively slow. Consequently, the prices of *S. album* have risen dramatically, leading to

the overexploitation of this resource [8]. Therefore, proper cultivation and protection management measures of the sandalwood tree is of the utmost need at this time [9].

SAUR (small auxin-up RNA) belongs to one of the three major early auxin-response genes [10]. Early auxin-response genes can be rapidly activated and transcribed upon auxin induction without the need for new protein synthesis. Moreover, the molecular weights of the proteins encoded by *SAUR* genes are generally small, ranging from 9–30 kDa [11]. Auxin, as the earliest discovered plant hormone, can promote rooting, regulate shoot development, and control cell division and differentiation [12]. It plays a crucial role in plant growth and stress resistance [13]. The *SAUR* genes were first discovered in soybean hypocotyls that were stimulated to elongate by auxin [14]. Subsequently, *SAURs* have been found in a variety of plants, including *Arabidopsis* (*Arabidopsis thaliana*) [15], rice (*Oryza sativa*) [16], sorghum (*Sorghum bicolor*) [17], and pineapple (*Ananas comosus*) [18]. Among the early auxin-responsive genes, the *SAUR* gene family has the largest number of genes and a high degree of structural similarity [19]. Most *SAURs* contain only one exon and no introns. They are found in clusters on the chromosome, indicating that high frequencies of tandem and segmental duplications have occurred during evolution [20]. Auxin regulates gene expression by the binding of auxin response factor (ARF) to auxin response elements (AuxREs) in the promoter region of auxin-responsive genes [21]. The number and diversity of AuxREs in different *SAURs* vary significantly. Nine AuxREs have been identified in rice [16] and seven in tomato [22]. The mRNA encoded by the *SAURs* is highly unstable due to the presence of a conserved downstream element (DST) in its 3′ untranslated region (3′ UTR) [23]. This instability reduces the association between auxin and their regulated *SAUR* genes.

The *SAUR* gene family is involved in multiple processes in plants, i.e., plant growth and development. It plays an important role in responding to changes in auxins [24], affecting their synthesis and transport, as well as cell elongation and senescence resistance [25]. In *Arabidopsis thaliana*, overexpression of the *AtSAUR36* [26] and *AtSAUR41* [27] genes results in the significant elongation of hypocotyl epidermal cells, while the *AtSAUR36* gene responds to leaf senescence [28]. The *SAUR* gene family is involved in abiotic stress tolerance [29], and many *SAUR* genes are down-regulated under abiotic stress as an adaptation to adverse environments [30]. In the model species *Arabidopsis thaliana*, seedlings from *AtSAUR41* overexpression lines showed higher sensitivity to salt stress, while the *AtSAUR41* subfamily of genes was found to be associated with salt tolerance [31]. In the perennial plant *Medicago sativa*, several *SAUR* genes have been associated with abiotic-stress tolerance. Specifically, *MsSAUR14*, *MsSAUR94*, *MsSAUR297*, *MsSAUR306*, and *MsSAUR254* respond to both drought and salinity stress [32]. In *Glycine max*, it was discovered that *GmSAUR299* may be related to drought tolerance [33]. In poplar, the PagWOX11/12a proteins directly bind to the promoter of the *SAUR36* gene, while enhanced binding is observed under salt stress, thereby improving the salt tolerance [34]. Expression of the poplar genes *PtSAUR12*, *PtSAUR34*, *PtSAUR54*, *PtSAUR67*, *PtSAUR91*, and *PtSAUR97* was up-regulated at low temperatures [35]. These findings reflect the significant role of the *SAUR* gene family in different plants and trees; however, a comprehensive genome-wide identification of the *SAUR* gene family in *S. album* has yet to be documented.

In this study, bioinformatics methods were employed to identify the members of the *SAUR* gene family in *S. album*, while also delving into their molecular characteristics, chromosomal distribution patterns, gene structures, conserved motifs, and their evolutionary and functional intricacies. Leveraging transcriptome data, we further investigated the expression profiles of *SAUR* genes across various tissues, including heartwood, root, and leaf, as well as their response to salt stress. In conclusion, this study provides valuable resource information for understanding the regulatory roles of *SAUR* gene family members in *S. album*'s growth and stress tolerance.

2. Materials and Methods

2.1. Identification of SAUR Genes and Chromosomal Distribution

A hidden Markov model profile of the SAUR DNA-binding domain (PF00642) was downloaded from the Pfam database (http://pfam.xfam.org/) [36]. The whole genome sequence and annotation files of *S. album* employed in this study were downloaded from the National Genomics Data Center of China under accession number GWHCBHT00000000, which is publicly accessible at https://ngdc.cncb.ac.cn/gwh/Assembly/37800/show (accessed on 9 December 2023). The initial members of the *SAUR* gene family were identified in the genome sequence of *S. album* through an HMM search ($E < 1 \times 10^{-5}$) [37]. Subsequently, we utilized the SMART (http://smart.embl.de/smart/batch.pl/, accessed on 11 December 2023) [38] and CDD online websites (https://www.ncbi.nlm.nih.gov/cdd/, accessed on 11 December 2023) [39] to meticulously screen the initial family members for any potential redundancy. The distribution of SAUR proteins on chromosomes was mapped using the MapChart version 5.4.6 software based on the data from the *S. album* genome annotation files [40]. The *SAUR* genes of *S. album* were renamed based on their chromosomal location, following the nomenclature of the *Arabidopsis thaliana SAUR* gene family members [41].

2.2. Phylogenetic Analysis and Classification of SAUR Genes

Firstly, the protein sequences of the *SAUR* gene family members of *Arabidopsis thaliana* were downloaded from the TAIR database website (http://www.arabidopsis.org/, accessed on 1 December 2023). Subsequently, multiple sequence alignment was performed on SAUR proteins from *A. thaliana* and *S. album* using the MAFFT version 7 software [42]. The resulting alignment was then used to construct a phylogenetic tree using the IQ-TREE version 2 software and the maximum likelihood (ML) method [43]. The phylogenetic tree was bootstrapped 1000 times with default parameters, resulting in the final phylogenetic tree file. The phylogenetic tree was uploaded to the iTOL website (https://itol.embl.de) for visualization and information annotation [44]. The *SAUR* gene family members of *S. album* were classified based on the branch clustering of the phylogenetic tree.

2.3. Gene Duplication and Analysis of Protein Properties

The Multiple Collinearity Scan toolkit X version 1 (MCScanX) [45] with default parameters was used to identify gene duplications, which were then visualized by using the One Step MCScanX tool in TBtools [46]. Non-synonymous substitutions (Ka) and synonymous substitutions (Ks) were determined for each duplicated *SAUR* gene pair using the KaKs_Calculator 2.0 [47]. The length of amino acids, theoretical isoelectric point (pI), and molecular weight of the SAUR proteins were analyzed using the online tool ExPASy (https://web.expasy.org/protparam/, accessed on 13 December 2023) [48]. The subcellular localization of the SAUR proteins was predicted using the Plant-mPLoc predictor (http://www.csbio.sjtu.edu.cn/bioinf/plant-multi/, accessed on 13 December 2023) [49]. Conserved motifs in the *S. album* SAUR proteins were identified using the online software Multiple Em for Motif Elicitation (MEME) suite 5.5.5 (https://meme-suite.org/meme/tools/streme, accessed on 13 December 2023), with a maximum of 10 motifs [50]. Additionally, we extracted exon and intron positions from the *S. album* genome annotation file. We used the TBtools Gene Structural View (Advanced) [46] to visualize the identified conserved motifs and gene structures.

2.4. Transcriptome Data and Expression of SAUR Genes

To analyze the gene functions in growth and development, we downloaded the transcriptome data for various *S. album* tissues (PRJNA243306, PRJNA360335, PRJNA192417) from the NCBI (National Center for Biotechnology Information, http://www.ncbi.nlm.nih.gov/, accessed on 1 December 2023). To investigate the mechanism of the salt-stress response in *S. album* mediated by the *SAUR* genes, approximately 50 cm tall sandalwood seedlings were treated with 50 mmol/L and 100 mmol/L NaCl solutions, while the blank

control group was treated with an aqueous solution. The third or fourth fully stretched leaf samples were taken at 0, 12, 24, and 48 h intervals. All seedlings were placed under controlled growth conditions in an artificial climate chamber set at a constant temperature of 26 °C and 50% relative humidity, with 16 h of light exposure (200 µmol photons m^{-2} s^{-1}, 26 °C) and 8 h in the dark. The experiment was repeated three times, resulting in a total of 36 samples. All the samples were collected, directly frozen in liquid nitrogen, and stored at −80 °C for RNA extraction. Total RNA was extracted from the leaves of each treatment using the TRIzol reagent. The RNA quality and quantity were checked using an Agilent 2100 Bioanalyzer, and then cDNA libraries were constructed. These libraries were sequenced using the Illumina platform at the Novogene Bioinformatics Technology Co., Ltd. (Beijing, China). To ensure data quality, all RNA sequence data were filtered using Fastp with default parameters [51]. Subsequently, the reads were aligned with the cultivated sandalwood reference genome using the HISAT2 version 2.1.0 software with default parameters [52]. To assess the expression levels of individual genes, the fragments per kilobase per million mapped reads (FPKM) of each gene was quantified by FeatureCounts [53]. Subsequently, differentially expressed genes (DEGs) were identified using the DESeq2 package in R, with a stringent false discovery rate (FDR) threshold of $\leq$0.05 and a $|\log2\ FC| \geq 1$ criterion to establish statistically significant differences in gene expression [54].

2.5. Real-Time Quantitative PCR Analysis

Total RNA was extracted using the TRIzol reagent, with subsequent removal of trace amounts of genomic DNA using the RNase-free DNase I method (RNase-free; TakaRa). According to the manufacturer's protocol, the first-strand cDNA was synthesized using a PrimerScript RT MasterMix (Takara, Kusatsu, Japan). The Primer3 web (version 4.1.1) software was used to design RT-qPCR gene-specific primers (Supplementary Table S4), and the primer sequences were synthesized by Shanghai Biosune Platinum Bio-Technology Co. (Shanghai, China) Primers SaSAUR27-1 and SaSAUR28-4 were used in the final data analysis. The data of relative gene expression were analyzed using the $2^{-\Delta\Delta CT}$ method, with the actin gene (GI:156972025) of *S. album* as the internal reference to account for differences. There were three biological and three technical replicates performed for each sample. We used the GraphPad Prism version 8 software for statistical analysis and data visualization.

3. Results

3.1. Identification and Analysis of Physicochemical Properties of the SAUR Genes

The conserved domains in the SAUR proteins of *S. album* were validated using the SMART and CDD databases, resulting in the identification of 43 *SAUR* genes (Supplementary File S1). The physical locations of the *SAUR* genes on the chromosomes were determined by querying the genome annotation files and were named *SaSAUR01* to *SaSAUR43* sequentially (Table 1). The physicochemical characteristics of the 43 identified SAUR proteins were examined using the ExPASy program. The study revealed that the length of these proteins ranged from 75 (*SaSAUR15*) to 434 (*SaSAUR01*) amino acids, and their molecular weight ranged from 8.32 to 48.35 kDa. The average isoelectric point (pI) of all the SAUR proteins was calculated to be 8.37, and ranged from 4.45 (*SaSAUR28*) to 11.16 (*SaSAUR08*) (Table 1). Notably, one-quarter of the SAUR proteins were classified as acidic because their isoelectric point (pI) values were below 7, while the remaining three-quarters were alkaline (Table 1). The subcellular localization analysis of *SAUR* genes using Plant-mPLoc revealed that most *SAUR* genes were predicted to be located in chloroplasts. The remaining *SAUR* genes were predicted to be distributed in the nucleus, while only the *SaSAUR36* gene was predicted to be located in the cytoplasm (Table 1). A thorough examination of the physicochemical characteristics and subcellular localization of the identified SAUR proteins was conducted in detail to gain deeper insight into their functions in *S. album*.

Table 1. Detailed information on the *SAUR* gene family in *S. album.*

Gene Name	Gene Id	Length AA	Molecular Weight (kD)	pI	Chr	Predicted Localization
SaSAUR01	SA01G02339	434	48.35	8.96	1	Chloroplast
SaSAUR02	SA02G00132	151	17.13	10.66	2	Chloroplast
SaSAUR03	SA02G00206	112	12.59	6.33	2	Nucleus
SaSAUR04	SA02G00516	107	11.81	9.34	2	Chloroplast, Nucleus
SaSAUR05	SA02G00677	137	15.15	9.40	2	Chloroplast
SaSAUR06	SA02G01106	102	11.18	8.56	2	Chloroplast
SaSAUR07	SA02G02406	135	15.60	8.45	2	Nucleus
SaSAUR08	SA02G02629	184	20.35	11.16	2	Chloroplast, Mitochondrion
SaSAUR09	SA02G02667	416	46.60	8.76	2	Chloroplast
SaSAUR10	SA02G03132	91	9.90	9.67	2	Chloroplast
SaSAUR11	SA02G03133	109	12.41	4.99	2	Nucleus
SaSAUR12	SA02G03134	127	13.95	9.07	2	Chloroplast
SaSAUR13	SA03G00088	163	18.15	9.48	3	Nucleus
SaSAUR14	SA03G00867	115	13.31	10.02	3	Chloroplast
SaSAUR15	SA04G00058	75	8.32	10.28	4	Chloroplast, Nucleus
SaSAUR16	SA04G00729	97	11.43	9.58	4	Chloroplast, Nucleus
SaSAUR17	SA04G00792	109	12.09	9.75	4	Chloroplast, Nucleus
SaSAUR18	SA04G01353	120	14.15	6.34	4	Nucleus
SaSAUR19	SA04G02219	155	17.14	9.72	4	Chloroplast, Nucleus
SaSAUR20	SA04G02816	131	14.82	8.71	4	Chloroplast
SaSAUR21	SA04G02839	156	17.35	8.58	4	Chloroplast, Cytoplasm, Nucleus
SaSAUR22	SA05G00696	152	17.17	8.82	5	Chloroplast, Nucleus
SaSAUR23	SA06G00626	156	17.87	9.85	6	Chloroplast
SaSAUR24	SA06G00887	149	15.81	6.50	6	Nucleus
SaSAUR25	SA06G00889	158	17.41	5.38	6	Nucleus
SaSAUR26	SA07G00208	146	16.17	8.72	7	Nucleus
SaSAUR27	SA07G00209	146	16.26	8.33	7	Chloroplast, Mitochondrion, Nucleus
SaSAUR28	SA07G00210	88	9.63	4.45	7	Nucleus
SaSAUR29	SA07G00277	115	12.73	8.01	7	Chloroplast, Mitochondrion, Nucleus
SaSAUR30	SA07G00278	109	12.58	6.29	7	Chloroplast, Nucleus
SaSAUR31	SA07G00279	93	10.20	6.82	7	Chloroplast
SaSAUR32	SA07G00280	94	10.76	8.96	7	Chloroplast, Mitochondrion, Nucleus
SaSAUR33	SA07G00281	94	10.73	8.93	7	Chloroplast
SaSAUR34	SA07G00282	94	10.59	6.57	7	Chloroplast
SaSAUR35	SA07G00283	94	10.74	9.37	7	Chloroplast, Nucleus
SaSAUR36	SA07G00284	420	47.02	7.71	7	Cytoplasm
SaSAUR37	SA08G01800	137	14.98	9.10	8	Chloroplast, Cytoplasm, Nucleus
SaSAUR38	SA08G01915	130	14.92	6.33	8	Chloroplast
SaSAUR39	SA08G01928	101	11.61	8.31	8	Nucleus
SaSAUR40	SA09G00119	123	14.21	9.26	9	Chloroplast
SaSAUR41	SA09G00719	119	13.98	5.61	9	Chloroplast
SaSAUR42	SA09G00803	125	14.33	9.18	9	Nucleus
SaSAUR43	SA10G00348	256	28.87	9.66	10	Chloroplast, Nucleus

3.2. Gene Structure and Phylogenetic Analysis of SaSAUR Genes

A phylogenetic tree was constructed for the SAUR proteins from *Santalum album* and *Arabidopsis thaliana* using MAFFT version 7 and IQ-TREE version 2 software. The results of the phylogenetic tree revealed that sandalwood *SAUR* genes could be classified into five groups. The study found that there was a total of 43 SAUR proteins, with one in Group I, four in Group II, five in Group III, ten in Group IV, and twenty-one in Group V. Additionally, there were two *SAUR* genes that could not be classified into any of the groups (Figure 1). Group V had the highest number of family members, while Group I had the lowest. In all groups, the number of *SAUR* genes in *A. thaliana* exceeded that in *S. album*. However, it is worth noting that the number of *SAUR* genes in Group V approximates that in *A. thaliana*, with 28 in *A. thaliana* and 20 in *S. album*. The conserved motifs, conserved domains, and exon–intron structure were analyzed to explore the structural diversity of *SaSAURs* (Figure 2). The online MEME tool identified 10 conserved motifs (motif 1 to motif 10) in SAUR proteins (Figure 2B). Motifs 1, 2, and 3 were widely distributed among the *SAUR* gene family in the order of 2–1–3, indicating strong conservation of the conserved domains of the *SAUR* gene family in *S. album*. Only *SaSAUR15*, *SaSAUR21*, *SaSAUR24*, and *SaSAUR28* do not contain motif 2. *SaSAUR15* and *SaSAUR24* do not have motif 1, while *SaSAUR09*, *SaSAUR10*, and *SaSAUR19* do not have motif 3. However, some motifs are unique to specific *SAUR* genes. For instance, motif 5 and motif 7 were exclusively found in *SaSAUR26* and *SaSAUR27*; motif 9 was identified in *SaSAUR19* and *SaSAUR22*; motif 6 was observed in *SaSAUR10*, *SaSAUR12*, and *SaSAUR37*; and motif 8 was specifically

present in *SaSAUR03*, *SaSAUR07*, and *SaSAUR08*. Members of the same group, such as *SaSAUR32* and *SaSAUR34*, *SaSAUR04* and *SaSAUR06*, and *SaSAUR26* and *SaSAUR27*, exhibited similar motif patterns (Figure 2B). This indicates that they likely share similar functions. All *SaSAUR* genes contain conserved auxin domains, as shown in Figure 2C. In Figure 2D, the exon–intron structure of the *SaSAURs* clearly verify the diversity of the gene structures. The 43 *SaSAUR* genes exhibited different numbers of exons, ranging from 1 to 12. *SaSAUR09* has 12 exon structures, while *SaSAUR01* and *SaSAUR36* have 8 exons. Most of the *SaSAUR* genes did not contain introns. Nine (20.93%) of the *SaSAUR* genes contained one or more introns, while the remaining 34 (79.07%) *SaSAUR* genes were intron-free (Figure 2D). The similar conserved motifs and gene structures of *SAUR* genes within the same group strongly supported the accuracy of the group classification.

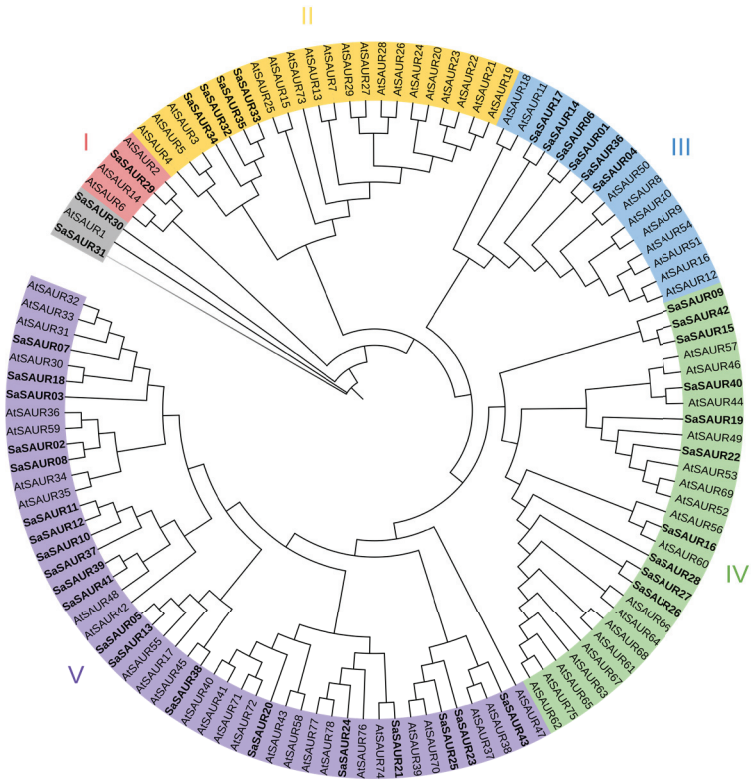

Figure 1. Phylogenetic tree of SAUR proteins sequences in *Santalum album* and *Arabidopsis thaliana* (different-colored shaded areas are used to differentiate distinct groups, and bold text indicates SAUR proteins of *Santalum album*).

3.3. SAUR Gene Duplication and Chromosomal Location

The locations of the *SAUR* genes were extracted from the genome-annotated generic-feature format (GFF) file of *Santalum album*, and the physical locations of these *SAUR* genes on the chromosomes were visualized using the MapChart version 5.4.6 software. Figure 3 illustrates the irregular distribution of the 43 *SAUR* genes across 10 chromosomes. Specifically, one gene was located on chr01 (2.32% of the total), eleven genes on chr02 (25.58%), two genes on chr03 (4.65%), seven genes on chr04 (16.28%), one gene on chr05 (2.32%), three genes on chr06 (6.98%), twelve genes on chr07 (27.91%), three genes on chr08 (6.98%), two genes on chr09 (4.65%), and one gene on chr10 (2.32%). Some of the *SaSAUR* genes can be observed in Figure 3 as gene clusters.

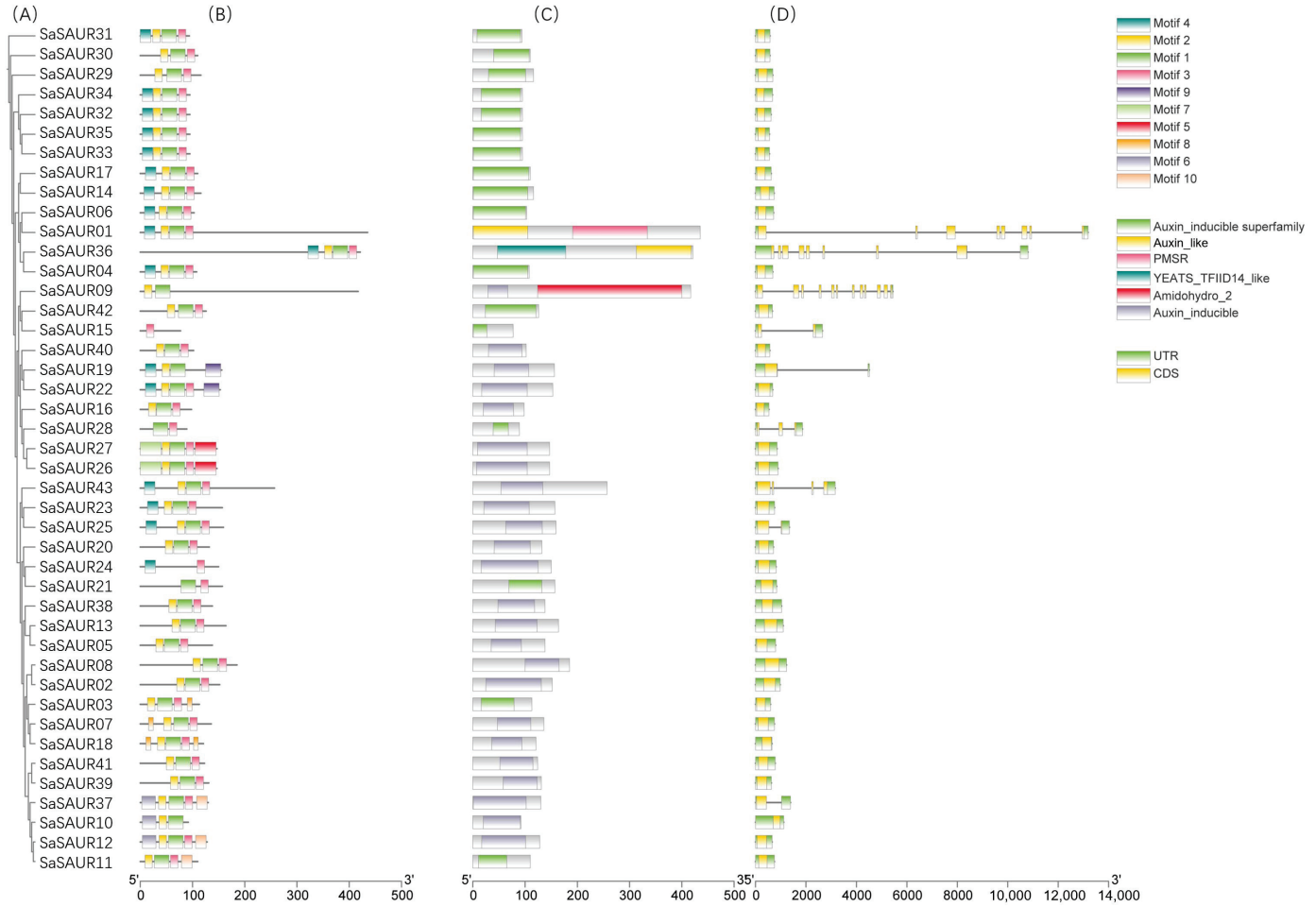

Figure 2. Conserved motifs, conserved domains, and gene structures of the *SaSAUR* genes. (**A**) Phylogenetic tree of the SaSAUR proteins. (**B**) Conserved motifs in the *SaSAUR* genes. (**C**) Conserved domains in the *SaSAUR* genes. (**D**) Exon–intron structure of the *SaSAUR* genes.

Gene duplication events typically include tandem and segmental duplications, and they are important for studying the expansion of gene families. The results revealed the presence of 14 pairs of segmentally duplicated (SD) genes and 9 pairs of tandemly duplicated (TD) genes within the *SaSAUR* genes (Table 2). *SaSAUR* genes involved in tandem duplications typically belong to the same group. The presence of multiple duplication events in *SaSAUR* genes suggests that certain genes have arisen through both tandem and segmental duplication, potentially serving as drivers of gene evolution. The majority of segmental duplications in the *SAUR* gene family of *Santalum album* were located on chr02, chr04, and chr07, as shown in Figure 4. Ka/Ks ratios can predict the selection pressure on genes encoding proteins. Our analysis revealed that the Ka/Ks ratios of all the duplicated gene pairs were below 1, with the exception of three gene pairs where the Ka/Ks value was incalculable (Table 2). A Ka/Ks value below 1 indicates a purifying (negative) selection, while a Ka/Ks value of 1 indicates neutral evolution, and Ka/Ks values above 1 indicate adaptive (positive) selection. Our findings suggest that negative selection may have played a role in the development of these duplicated genes in the *SAUR* gene family in *Santalum album*.

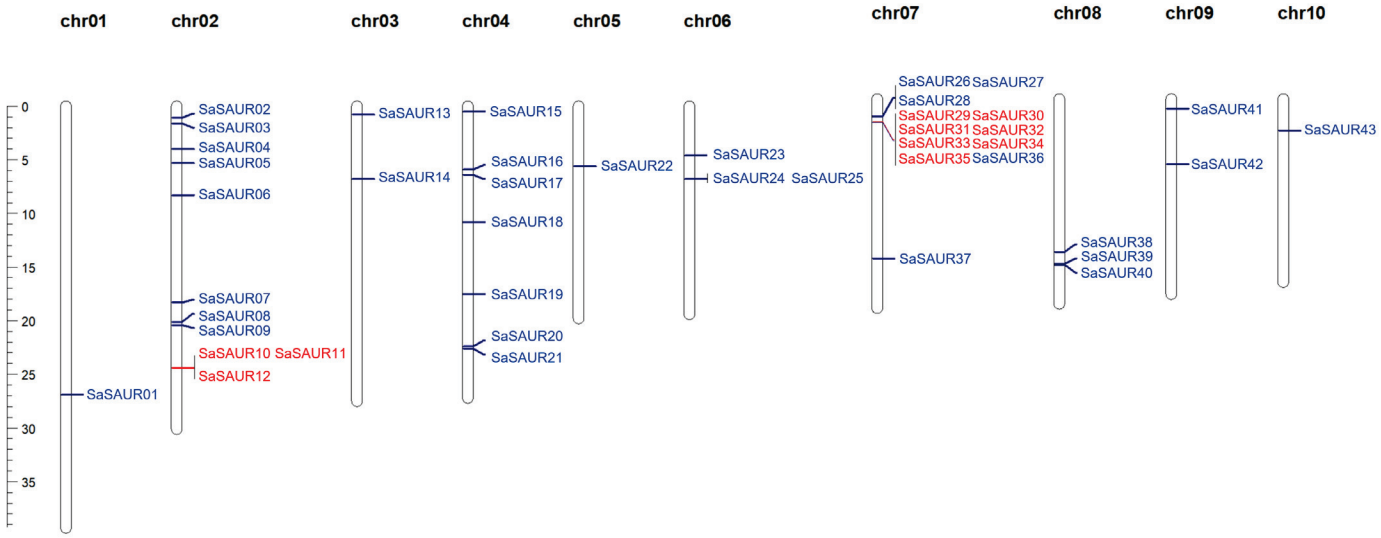

Figure 3. Chromosomal locations of the 43 *SAUR* genes in *S. album*. The ruler on the left indicates the physical position of the reference genome. The pair of tandemly duplicated genes is shown in red font.

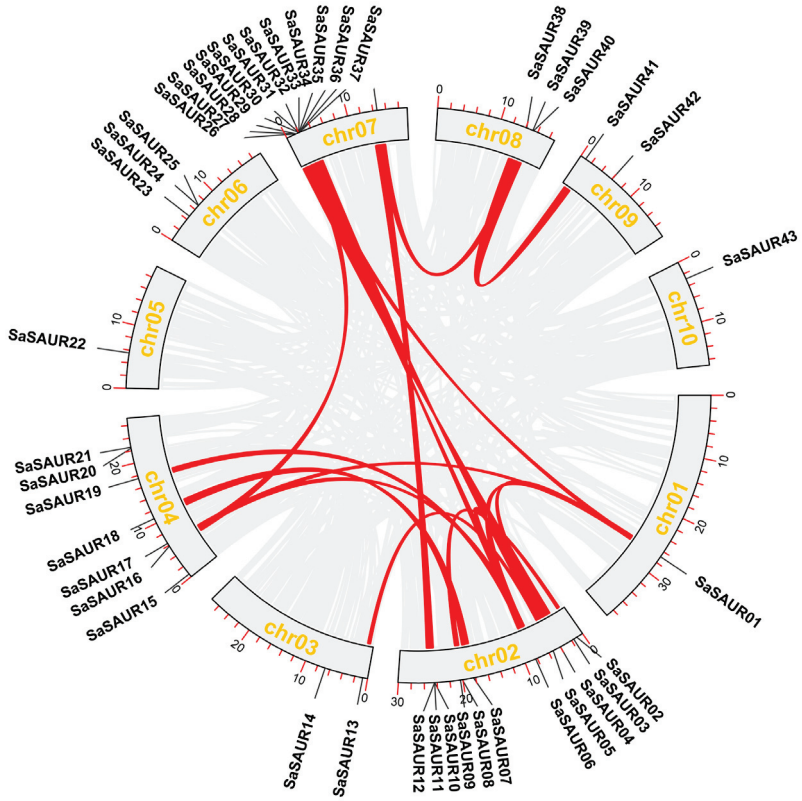

Figure 4. The intrachromosomal segmental duplication map in *S. album*. Gray lines are collinear blocks in *S. album*. Red lines are the syntenic gene pairs of *SaSAUR* genes.

Table 2. Ka, Ks, and Ka/Ks of duplicated gene pairs of the *SaSAURs*.

Duplicated Gene Pairs	Non-Synonymous (Ka)	Synonymous (Ks)	Ka/Ks	Duplicated Type
SaSAUR10&SaSAUR11	0.20	0.45	0.44	tandem
SaSAUR11&SaSAUR12	0.22	0.48	0.46	tandem
SaSAUR26&SaSAUR27	0.10	0.33	0.32	tandem
SaSAUR29&SaSAUR30	0.42	NaN	NaN	tandem
SaSAUR30&SaSAUR31	0.35	NaN	NaN	tandem

Table 2. *Cont.*

Duplicated Gene Pairs	Non-Synonymous (Ka)	Synonymous (Ks)	Ka/Ks	Duplicated Type
SaSAUR31&SaSAUR32	0.30	1.46	0.21	tandem
SaSAUR32&SaSAUR33	0.12	0.94	0.13	tandem
SaSAUR33&SaSAUR34	0.10	0.33	0.30	tandem
SaSAUR34&SaSAUR35	0.11	0.31	0.35	tandem
SaSAUR01&SaSAUR04	0.22	1.88	0.12	segmental
SaSAUR01&SaSAUR06	0.14	0.80	0.17	segmental
SaSAUR01&SaSAUR17	0.30	2.11	0.14	segmental
SaSAUR01&SaSAUR36	0.20	NaN	NaN	segmental
SaSAUR04&SaSAUR06	0.21	1.81	0.11	segmental
SaSAUR02&SaSAUR08	0.25	0.51	0.48	segmental
SaSAUR05&SaSAUR13	0.41	1.71	0.24	segmental
SaSAUR07&SaSAUR18	0.32	1.54	0.21	segmental
SaSAUR06&SaSAUR17	0.26	1.61	0.16	segmental
SaSAUR06&SaSAUR19	0.62	1.59	0.39	segmental
SaSAUR04&SaSAUR36	0.12	1.60	0.07	segmental
SaSAUR06&SaSAUR36	0.18	1.43	0.12	segmental
SaSAUR16&SaSAUR26	0.56	3.45	0.16	segmental
SaSAUR38&SaSAUR40	0.23	1.00	0.23	segmental

3.4. Expression Analysis of SaSAUR Genes

The current research work investigated the expression patterns of all 43 *SaSAUR* genes using transcriptome data from various tissues of *Santalum album* to understand their potential functions (Supplementary Table S2). As shown in Figure 5A, *SaSAUR* genes exhibited distinct tissue-specific expression patterns. Notably, all the *SaSAUR* genes were expressed in leaves, displaying no significant differences in their expression levels in leaves, which was consistent with the results of the subcellular localization indicating that the majority of *SaSAUR* genes are localized in chloroplasts. While approximately half of the *SaSAUR* genes did not show expression in roots, it is worth noting that *SaSAUR18* exhibited high expression levels in roots, hinting a pivotal role in the function of root development. Only a limited number of *SaSAUR* genes (*SaSAUR33*, *SaSAUR07*, *SaSAUR08*, *SaSAUR09*, *SaSAUR11*, *SaSAUR13*, *SaSAUR18*, *SaSAUR19*, and *SaSAUR34*) were expressed in the heartwood, with *SaSAUR08* and *SaSAUR13* showing significantly higher expression levels. *SaSAUR08*, *SaSAUR11*, *SaSAUR13*, and *SaSAUR18* were expressed in all three tissues, indicating their significant roles in plant growth and development.

The *SAUR* gene family plays a crucial role in various abiotic stresses. To comprehend the function of the *SAUR* gene family in salt stress, we exposed seedling leaves of *Santalum album* to different salt concentrations (W: CK; L: 50 mmol/L NaCl; H: 100 mmol/L NaCl) for 0, 12, 24, and 48 h, with three biological replicates. Quality control was conducted on the data obtained from sequencing the libraries on the platform, revealing that all samples adhered to the established standards. The sequencing produced a total of 413.17 Gb of raw data, which was subsequently filtered to produce 402.13 Gb of high-quality data. The specific sequencing quality for each sample is comprehensively presented in Supplementary Table S1. An analysis of the transcriptomic data for *SaSAURs* under salt stress using a bioinformatics approach revealed that most *SaSAUR* genes were expressed both before and after salt stress (Figure 5B and Supplementary Table S3). It is worth noting that *SaSAUR04*, *SaSAUR07*, *SaSAUR08*, *SaSAUR09*, and *SaSAUR18* exhibited significantly higher expression levels. *SaSAUR27* and *SaSAUR28* are salt-stress-associated differential genes ($\log_2$FC > 1), with almost no expression before salt stress and significantly up-regulated expression after salt stress, indicating that these genes may play a role in the response to salt stress. The two predicted differential genes (*SaSAUR27* and *SaSAUR28*) were selected to detect their expression characteristics under salt treatment using real-time quantitative polymerase chain reaction (RT-qPCR). Figure 6 shows that *SaSAUR27* exhibited significant differences from the control at all treatment time periods after treatment with the salt solution. *SaSAUR28* did not exhibited any significant differences at the 12 h treatment time but exhibited significant differences at the 24 h and 48 h treatment times. The RT-qPCR results were consistent with the transcriptome results.

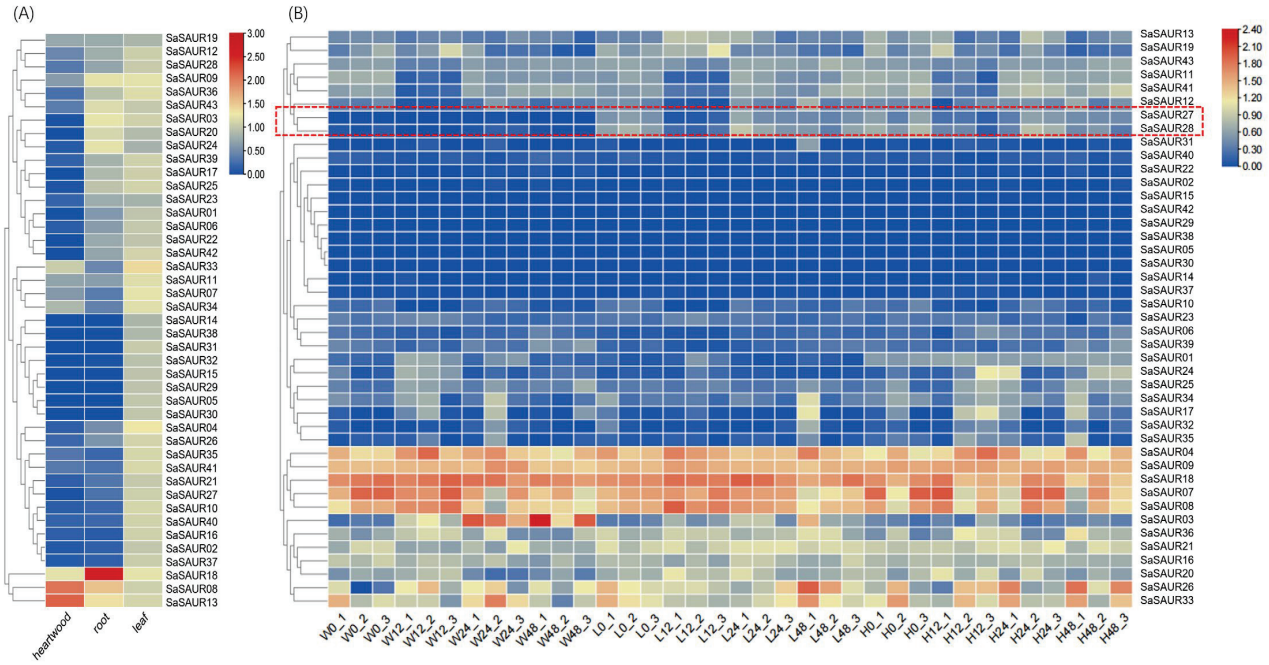

Figure 5. Expression profiles of the *SaSAUR* genes. (**A**) Expression of *SaSAUR* genes in various tissues (heartwood, root, leaf). (**B**) Expression of *SaSAUR* genes under salt stress. Differential genes are labeled in red boxes. The initials of the horizontal coordinates represent the salt solutions (W: CK; L: 50 mmol/L; H: 100 mmol/L), followed by treatment times 0 (0 h), 12 (12 h), 24 (24 h), and 48 (48 h), with the numbers after the underscore indicating the biological replicate. Expression profiles were normalized to $\log_{10}$(FPKM).

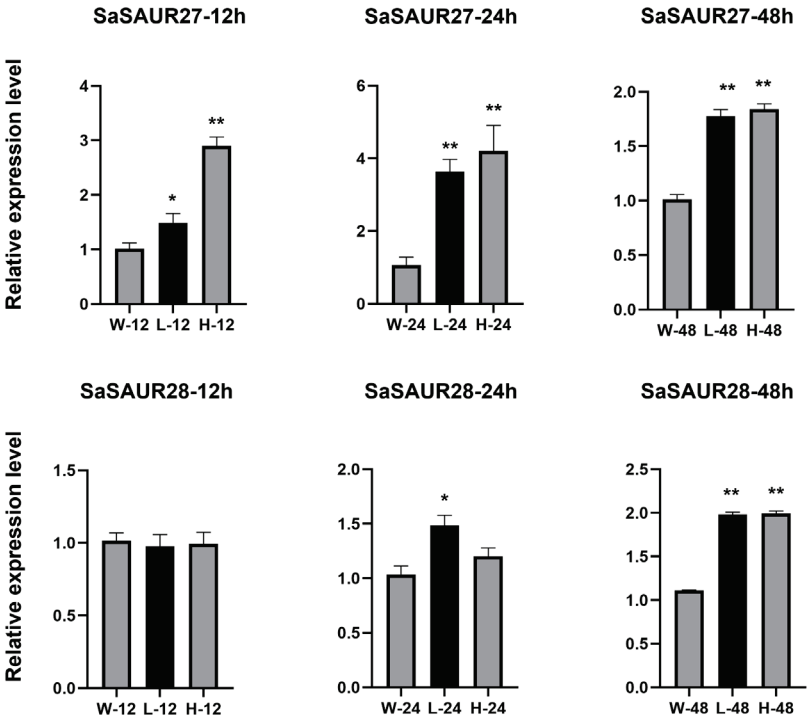

Figure 6. Expression profiles of *SaSAUR27* and *SaSAUR28* in response to salt treatment as determined by RT-qPCR. The *x*-axis indicates the different concentrations of NaCl solution treatments (W: CK; L: 50 mmol/L; H: 100 mmol/L), and the *y*-axis indicates the relative expression levels. The values are the means ± SD of three independent experiments, and the asterisks indicate a significant difference (* $p < 0.05$; ** $p < 0.01$) compared with the corresponding controls.

4. Discussion

The *SAUR* gene family has been identified in a variety of plants and shown to play a critical role in plant growth and stress tolerance [24]. Spartz found that overexpression of *AtSAUR19* in *Arabidopsis thaliana* resulted in increased expression in the leaf area, which positively regulated leaf growth by promoting leaf cell enlargement [55]. Overexpressing both *AtSAUR41* [27] and *AtSAUR63* [26] resulted in enlarged petals and elongated filaments, highlighting their significance in plant growth. In Belladonna, *AbSAUR1* promotes the growth and development of the plant, resulting in increased plant height and root meristem, which significantly enhances yield and alkaloid content [56]. The auxin-induced SAUR protein can regulate PM H^+-ATPase activity by modulating PP2C-D phosphatase, causing an increase in plasma membrane potential, providing expansion pressure for cell growth [24]. Consequently, it promotes cell expansion, plant growth and development. The results of several studies have shown that *TaSAUR75* plays a positive regulatory role in plants' responses to drought and salt stress [57]. In cucumber, *CsaSAUR1*, *CsaSAUR13*, *CsaSAUR15*, *CsaSAUR49*, and *CsaSAUR50* were up-regulated by high-temperature-induced expression, suggesting that they may be involved in high-temperature tolerance response processes [58]. The *OsSAUR51* gene in rice could inhibit the spread of pathogenic bacteria by reducing the accumulation of auxin, which made the plant resistant to bacterial leaf blight (BLB) and played a positive role in biotic-stress tolerance in rice [59]. Guo found that overexpression of the *TaSAUR78* gene could enhance the tolerance of wheat to salt and drought stresses [60]. Analysis of the *S. album* transcriptome data revealed that the majority of *SaSAUR* genes showed no significant changes in their expression level before and after salt stress. Expression patterns were categorized into two distinct cases: one where genes remained unexpressed both before and after salt stress, and another where genes exhibited consistent expression levels before and after salt stress. The *SaSAUR27* and *SaSAUR28* (4.65% of total) genes in *S. album* exhibited significant responses to salt stress. These two genes were classified within Group V alongside *AtSAUR36* and *AtSAUR41*, suggesting their potential functional similarities. These results are confirmed by comparative analysis with other species. Pineapple possesses 52 *SAUR* genes, with 3 genes (5.77%) exhibiting responsiveness to salt stress [18]. In contrast, alfalfa possesses 433 *SAUR* genes, of which 11 genes (2.54%) respond to salt stress [28].

Based on the genomic data of *S. album*, a total of 43 *SaSAUR* genes were identified in this study. The number of *SAUR* gene family members was significantly lower than that of peanut (162) [61], poplar (105) [35], and maize (79) [22], but higher than that of oil palm (40) [62]. The limited number of members might be associated with the extended growth cycle characteristic of *S. album*. *SaSAUR* genes were classified into five groups depending on their phylogenetic relationships, and *SAUR* genes of melons [63] have also been classified into five groups. In rice [16], tomato [22], and watermelon [64], five, three, and two clusters of *SAUR* gene family members have been identified, respectively. Chromosomal localization revealed that the *SaSAUR* genes were unevenly distributed among chr01 to chr10, with *SaSAUR26* to *SaSAUR36* clustered in the short arm of chr07. In addition, there was a high degree of sequence similarity among the members of these gene clusters. Four genes (*SaSAUR32*, *SaSAUR33*, *SaSAUR34*, *SaSAUR35*) on chr07 were classified in group II, and these genes were clustered together on the same branch of the phylogenetic tree, exhibiting a high level of similarity in motif arrangement and gene structure. In terms of gene structure, the *SAUR* genes contain almost no introns, but AUX and GH3, which also belong to the three large early auxin-responsive families, have a high number of introns. The number and location of introns can often influence the level of gene expression. The incidence of alternative splicing in intron-less genes tends to be low, so the *SAUR* genes may be relatively functionally conserved. The identified members of *SAUR* gene family with intronic structures accounted for 20.9% of the total, which is higher than maize (7.5%) and poplar (10.5%), and similar to that in millet (18.1%). Segmental and tandem duplications are the primary drivers of gene family expansion and evolution [65]. We identified nine tandem duplications and 14 segmental duplication events in 43 *SaSAUR*

genes (Table 2). This suggests that segmental duplication plays a critical role in amplifying and evolving the *SAUR* gene family of *S. album*. Motif 1, Motif 2, and Motif 3 were widely distributed among the *SaSAUR* genes and were arranged in the order of 2–1–3, which is consistent with the study of the cabbage *SAUR* gene family by Zhao [66]. The exon–intron structures and motif arrangements were highly conserved, and the genes of close relatives exhibited similar genetic structures. The similarity and variation in gene structure, motifs, and conserved domains of *SaSAUR* genes are certainly related to the evolutionary and gene duplication of *S. album*. Previous studies documented the subcellular location of *SAUR* genes within the nucleus [67] and cytoplasm [68]. Contrary to these findings, our study indicated that a predominant fraction, exceeding 70%, of the SaSAUR proteins was predicted to be localized within the chloroplasts, a pattern similarly observed in maize [69]. Plant chloroplasts are important intracellular organelles that produce ATP (Adenosine triphosphate) through photosynthetic phosphorylation, suggesting that auxins may play a role in the energy transfer process that takes place at the chloroplast membrane [70].

To explore the principal functions of *SaSAUR* genes and the role of salt stress on them, we embarked on an expression analysis across distinct tissues—roots, leaves, heartwood—and at incremental salt exposure durations (0, 12, 24, and 48 h), leveraging transcriptome datasets. A meticulous tissue-specific expression scrutiny unveiled that disparate *SAUR* genes assume unique roles in the ambit of plant growth and morphogenesis. Notably, the expression profile of *SAUR* genes in the leaves of *S. album* was significantly higher than in other parts of the plant, a pattern similar to that observed in watermelon [64]. This suggests a conserved role of *SAUR* genes in foliar development across species. Specifically, *SaSAUR27* and *SaSAUR28*, both localized on chr07 and categorized under group IV, emerged as salt-stress-responsive genes ($\log_2$FC > 1), suggesting their potential involvement in the salt-stress response mechanism. Subsequent RT-qPCR analysis results were consistent with those of the transcriptome. *SaSAUR27* differed from the blank control group at every time interval, and *SaSAUR28* also showed differential performance at all time intervals, except at 12 h. This pattern intimates a robust involvement of these genes in salt-stress adaptation. This study has revealed the expression response of *SaSAUR* genes in different tissues and under various salt treatments. With the advent of comprehensive plant genomic sequencing, it is anticipated that an expanded repertoire of *SAUR* genes will be identified across diverse taxa. These findings will contribute to a more in-depth study of the functional characterization and the evolutionary genesis and diversification of the *SAUR* gene family.

5. Conclusions

The *SAUR* gene family in *Santalum album* was comprehensively analyzed for the first time in this study. A total of 43 *SaSAUR* genes were identified, which were unevenly distributed across 10 chromosomes. The phylogenetic tree and gene structure analyses revealed that *SaSAUR* genes can be classified into five groups, each exhibiting similar gene structures. Additionally, most *SaSAURs* were found to lack introns. The duplication events of *SaSAUR* genes suggests that segmental duplication served as the primary means of evolutionary expansion for *SAUR* gene family. Analysis of *SaSAUR* gene expression levels in various tissues and under salt-stress conditions uncovered that all *SaSAUR* genes are expressed in leaves. *SaSAUR08* and *SaSAUR13* exhibit significantly higher expression levels in heartwood compared with roots and leaves. Conversely, *SaSAUR18* shows significantly higher expression in roots than in heartwood and leaves. Additionally, *SaSAUR27* and *SaSAUR28* have emerged as potential differential genes in response to salt stress. Overall, our study will contribute to a better understanding of the complexity of the *S. album SAUR* gene family and will serve as a valuable guide for future functional analysis studies.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/plants13101286/s1.

Author Contributions: M.T. and L.L. designed this study and revised the manuscript. Q.Z. carried out the study and wrote the manuscript. H.Z. and X.H. revised the manuscript and provided constructive suggestions. Y.L. and X.Z. revised the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This project was supported by the National Natural Science Foundation of China (32201624), the Hainan Provincial Natural Science Foundation of China (321RC469), the Hainan Province Science and Technology Special Fund (ZDYF2023XDNY078), the Collaborative Innovation Center Project of Nanfan and High-Efficiency Tropical Agriculture at Hainan University (XTCX2022NYB08), the Collaborative Innovation Center Project of Ecological Civilization in Hainan University (XTCX2022STC10), and the Priming Scientific Research Foundation at Hainan University (KYQD(ZR)-21039).

Data Availability Statement: The original contributions presented in this study are included in the article and Supplementary Material. The whole genome sequence and annotation files of *S. album* can be downloaded from the National Genomics Data Center of China at https://ngdc.cncb.ac.cn/gwh/Assembly/37800/show (accessed on 3 April 2024). The RNA-seq data under salt stress of *S. album* can be downloaded from the National Genomics Data Center of China at https://bigd.big.ac.cn/gsa/browse/CRA015677 (accessed on 3 April 2024).

Acknowledgments: We are sincerely thankful to the contributors of the RNA sequence data uploaded to the NCBI (accession numbers: PRJNA243306, PRJNA360335, PRJNA192417). We would like to thank the editors and reviewers for their efforts and constructive comments.

Conflicts of Interest: The authors declare that they have no conflicts of interest.

References

1. Solanki, N.S.; Chauhan, C.S.; Vyas, B.; Marothia, D. *Santalum album* Linn: A review. *Int. J. PharmTech Res.* **2015**, *7*, 629–640.
2. Baldovini, N.; Delasalle, C.; Joulain, D. Phytochemistry of the Heartwood from Fragrant Santalum Species: A Review. *Flavour. Fragr. J.* **2011**, *26*, 7–26. [CrossRef]
3. Santha, S.; Dwivedi, C. Anticancer Effects of Sandalwood (*Santalum Album*). *Anticancer. Res.* **2015**, *35*, 3137–3145. [PubMed]
4. Saneja, A.; Kaushik, P.; Kaushik, D.; Kumar, S.; Kumar, D. Antioxidant, Analgesic and Anti-Inflammatory Activities of *Santalum album* Linn. *Planta Med.* **2009**, *75*, 102. [CrossRef]
5. Mohankumar, A.; Kalaiselvi, D.; Levenson, C.; Shanmugam, G.; Thiruppathi, G.; Nivitha, S.; Sundararaj, P. Antioxidant and stress modulatory efficacy of essential oil extracted from plantation-grown *Santalum album* L. *Ind. Crops Prod.* **2019**, *140*, 111623. [CrossRef]
6. Ouyang, Y.; Zhang, X.; Chen, Y.; Teixeira da Silva, J.A.; Ma, G. Growth, Photosynthesis and Haustorial Development of Semiparasitic *Santalum album* L. Penetrating into Roots of Three Hosts: A Comparative Study. *Trees* **2016**, *30*, 317–328. [CrossRef]
7. Tah, J. White Sandal *(Santalum album* L.), A Precious Mediclnal and Tiber yielding Plant: A Short Review. *World J. Pharm. Pharm. Sci.* **2017**, *18*, 616–634. [CrossRef]
8. Nageswara-Rao, M.; Soneji, J.R. While Exploitation Threatens Indian Sandalwood, a Plea for Conservation and Regeneration Resounds. *Flavour. Fragr. J.* **2008**, *33*, 38–43.
9. Dhanya, B.; Viswanath, S.; Purushothaman, S. Sandal (*Santalum album* L.) Conservation in Southern India: A Review of Policies and Their Impacts. *J. Trop. Agric.* **2010**, *48*, 1–10.
10. Hagen, G.; Guilfoyle, T. Auxin-Responsive Gene Expression: Genes, Promoters and Regulatory Factors. *Plant Mol. Biol.* **2002**, *49*, 373–385. [CrossRef]
11. Gil, P.; Liu, Y.; Orbović, V.; Verkamp, E.; Poff, K.L.; Green, P.J. Characterization of the Auxin-Inducible SAUR-AC1 Gene for Use as a Molecular Genetic Tool in *Arabidopsis*. *Plant Physiol.* **1994**, *104*, 777–784. [CrossRef]
12. Woodward, A.W.; Bartel, B. Auxin: Regulation, Action, and Interaction. *Ann. Bot.* **2005**, *95*, 707–735. [CrossRef]
13. Vanneste, S.; Friml, J. Auxin: A Trigger for Change in Plant Development. *Cell* **2009**, *136*, 1005–1016. [CrossRef] [PubMed]
14. McClure, B.A.; Guilfoyle, T. Characterization of a Class of Small Auxin-Inducible Soybean Polyadenylated RNAs. *Plant Mol. Biol.* **1987**, *9*, 611–623. [CrossRef] [PubMed]
15. Goda, H.; Sawa, S.; Asami, T.; Fujioka, S.; Shimada, Y.; Yoshida, S. Comprehensive Comparison of Auxin-Regulated and Brassinosteroid-Regulated Genes in Arabidopsis. *Plant Physiol.* **2004**, *134*, 1555–1573. [CrossRef]
16. Jain, M.; Tyagi, A.K.; Khurana, J.P. Genome-Wide Analysis, Evolutionary Expansion, and Expression of Early Auxin-Responsive SAUR Gene Family in Rice (*Oryza sativa*). *Genomics* **2006**, *88*, 360–371. [CrossRef]
17. Wang, S.; Bai, Y.; Shen, C.; Wu, Y.; Zhang, S.; Jiang, D.; Guilfoyle, T.J.; Chen, M.; Qi, Y. Auxin-Related Gene Families in Abiotic Stress Response in Sorghum Bicolor. *Funct. Integr. Genom.* **2010**, *10*, 533–546. [CrossRef]
18. Zhang, Y.; Ye, T.; She, Z.; Huang, S.; Wang, L.; Aslam, M.; Qin, R.; Wang, X.; Qin, Y.; Niu, X. Small Auxin Up RNA (SAUR) Gene Family Identification and Functional Genes Exploration during the Floral Organ and Fruit Developmental Stages in Pineapple (*Ananas comosus* L.) and Its Response to Salinity and Drought Stresses. *Int. J. Biol. Macromol.* **2023**, *237*, 124061. [CrossRef] [PubMed]

19. Stortenbeker, N.; Bemer, M. The SAUR Gene Family: The Plant's Toolbox for Adaptation of Growth and Development. *J. Exp. Bot.* **2019**, *70*, 17–27. [CrossRef]
20. Zhang, H.; Yu, Z.; Yao, X.; Chen, J.; Chen, X.; Zhou, H.; Lou, Y.; Ming, F.; Jin, Y. Genome-Wide Identification and Characterization of Small Auxin-up RNA (SAUR) Gene Family in Plants: Evolution and Expression Profiles during Normal Growth and Stress Response. *BMC Plant Biol.* **2021**, *21*, 4. [CrossRef]
21. Ghanashyam, C.; Jain, M. Role of Auxin-Responsive Genes in Biotic Stress Responses. *Plant Signal. Behav.* **2009**, *4*, 846–848. [CrossRef] [PubMed]
22. Wu, J.; Liu, S.; He, Y.; Guan, X.; Zhu, X.; Cheng, L.; Wang, J.; Lu, G. Genome-Wide Analysis of SAUR Gene Family in *Solanaceae* Species. *Gene* **2012**, *509*, 38–50. [CrossRef] [PubMed]
23. Newman, T.C.; Ohme-Takagi, M.; Taylor, C.B.; Green, P.J. DST Sequences, Highly Conserved among Plant SAUR Genes, Target Reporter Transcripts for Rapid Decay in Tobacco. *Plant Cell* **1993**, *5*, 701–714. [PubMed]
24. Ren, H.; Gray, W.M. SAUR Proteins as Effectors of Hormonal and Environmental Signals in Plant Growth. *Mol. Plant* **2015**, *8*, 1153–1164. [CrossRef]
25. Mueller-Roeber, B.; Balazadeh, S. Auxin and Its Role in Plant Senescence. *J. Plant Growth Regul.* **2014**, *33*, 21–33. [CrossRef]
26. Chae, K.; Isaacs, C.G.; Reeves, P.H.; Maloney, G.S.; Muday, G.K.; Nagpal, P.; Reed, J.W. Arabidopsis Small Auxin up RNA63 Promotes Hypocotyl and Stamen Filament Elongation. *Plant J. For. Cell Mol. Biol.* **2012**, *71*, 684–697. [CrossRef] [PubMed]
27. Kong, Y.; Zhu, Y.; Gao, C.; She, W.; Lin, W.; Chen, Y.; Han, N.; Bian, H.; Zhu, M.; Wang, J. Tissue-Specific Expression of Small Auxin up RNA41 Differentially Regulates Cell Expansion and Root Meristem Patterning in Arabidopsis. *Plant Cell Physiol.* **2013**, *54*, 609–621. [CrossRef]
28. Hou, K.; Wu, W.; Gan, S.-S. *SAUR36*, a Small Auxin up RNA Gene, Is Involved in the Promotion of Leaf Senescence in Arabidopsis. *Plant Physiol.* **2013**, *161*, 1002–1009. [CrossRef]
29. Du, M.; Spalding, E.P.; Gray, W.M. Rapid Auxin-Mediated Cell Expansion. *Annu. Rev. Plant Biol.* **2020**, *71*, 379–402. [CrossRef]
30. Kodaira, K.-S.; Qin, F.; Tran, L.-S.P.; Maruyama, K.; Kidokoro, S.; Fujita, Y.; Shinozaki, K.; Yamaguchi-Shinozaki, K. Arabidopsis Cys2/His2 Zinc-Finger Proteins AZF1 and AZF2 Negatively Regulate Abscisic Acid-Repressive and Auxin-Inducible Genes under Abiotic Stress Conditions1. *Plant Physiol.* **2011**, *157*, 742–756. [CrossRef]
31. Qiu, T.; Qi, M.; Ding, X.; Zheng, Y.; Zhou, T.; Chen, Y.; Han, N.; Zhu, M.; Bian, H.; Wang, J. The SAUR41 Subfamily of Small Auxin up RNA Genes Is Abscisic Acid Inducible to Modulate Cell Expansion and Salt Tolerance in *Arabidopsis Thaliana* Seedlings. *Ann. Bot.* **2020**, *125*, 805–819. [CrossRef]
32. Liu, H.; Li, X.; He, F.; Wang, X.; Li, M.; Long, R.; Kang, J.; Yang, Q.; Chen, L. Identification of the alfalfa SAUR gene family and its expression pattern under abiotic stress. *Acta Prataculturae Sin.* **2024**, *4*, 135–153. [CrossRef]
33. Li, G.; Wang, Q.; Lu, L.; Wang, S.; Chen, X.; Khan, M.H.U.; Zhang, Y.; Yang, S. Identification of the Soybean Small Auxin Upregulated RNA (SAUR) Gene Family and Specific Haplotype for Drought Tolerance. *Biologia* **2022**, *77*, 1197–1217. [CrossRef]
34. Liu, R.; Wen, S.S.; Sun, T.-T.; Wang, R.; Zuo, W.-T.; Yang, T.; Wang, C.; Hu, J.J.; Lu, M.Z.; Wang, L.Q. PagWOX11/12a Positively Regulates the *PagSAUR36* Gene That Enhances Adventitious Root Development in Poplar. *J. Exp. Bot.* **2022**, *73*, 7298–7311. [CrossRef] [PubMed]
35. Hu, W.; Yan, H.; Luo, S.; Pan, F.; Wang, Y.; Xiang, Y. Genome-Wide Analysis of Poplar SAUR Gene Family and Expression Profiles under Cold, Polyethylene Glycol and Indole-3-Acetic Acid Treatments. *Plant Physiol. Biochem.* **2018**, *128*, 50–65. [CrossRef]
36. Finn, R.D.; Bateman, A.; Clements, J.; Coggill, P.; Eberhardt, R.Y.; Eddy, S.R.; Heger, A.; Hetherington, K.; Holm, L.; Mistry, J.; et al. Pfam: The Protein Families Database. *Nucleic Acids Res.* **2014**, *42*, 222–230. [CrossRef]
37. Finn, R.D.; Clements, J.; Eddy, S.R. HMMER Web Server: Interactive Sequence Similarity Searching. *Nucleic Acids Res.* **2011**, *39*, 29–37. [CrossRef] [PubMed]
38. Schultz, J.; Milpetz, F.; Bork, P.; Ponting, C.P. SMART, a Simple Modular Architecture Research Tool: Identification of Signaling Domains. *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 5857–5864. [CrossRef]
39. Marchler-Bauer, A.; Anderson, J.B.; Chitsaz, F.; Derbyshire, M.K.; DeWeese-Scott, C.; Fong, J.H.; Geer, L.Y.; Geer, R.C.; Gonzales, N.R.; Gwadz, M.; et al. CDD: Specific Functional Annotation with the Conserved Domain Database. *Nucleic Acids Res.* **2009**, *37*, 205–210. [CrossRef]
40. Voorrips, R.E. MapChart: Software for the Graphical Presentation of Linkage Maps and QTLs. *J. Hered.* **2002**, *93*, 77–78. [CrossRef]
41. Markakis, M.N.; Boron, A.K.; Loock, B.V.; Saini, K.; Cirera, S.; Verbelen, J.-P.; Vissenberg, K. Characterization of a Small Auxin-Up RNA (SAUR)-Like Gene Involved in *Arabidopsis Thaliana* Development. *PLoS ONE* **2013**, *8*, e82596. [CrossRef] [PubMed]
42. Katoh, K.; Standley, D.M. MAFFT Multiple Sequence Alignment Software Version 7: Improvements in Performance and Usability. *Mol. Biol. Evol.* **2013**, *30*, 772–780. [CrossRef] [PubMed]
43. Nguyen, L.T.; Schmidt, H.A.; von Haeseler, A.; Minh, B.Q. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. *Mol. Biol. Evol.* **2015**, *32*, 268–274. [CrossRef] [PubMed]
44. Letunic, I.; Bork, P. Interactive Tree of Life (iTOL) v5: An Online Tool for Phylogenetic Tree Display and Annotation. *Nucleic Acids Res.* **2021**, *49*, 293–296. [CrossRef] [PubMed]
45. Wang, Y.; Tang, H.; Debarry, J.D.; Tan, X.; Li, J.; Wang, X.; Lee, T.; Jin, H.; Marler, B.; Guo, H.; et al. MCScanX: A Toolkit for Detection and Evolutionary Analysis of Gene Synteny and Collinearity. *Nucleic Acids Res.* **2012**, *40*, e49. [CrossRef] [PubMed]
46. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef]

47. Wang, D.; Zhang, Y.; Zhang, Z.; Zhu, J.; Yu, J. KaKs_Calculator 2.0: A Toolkit Incorporating Gamma-Series Methods and Sliding Window Strategies. *Genom. Proteom. Bioinform.* **2010**, *8*, 77–80. [CrossRef] [PubMed]
48. Gasteiger, E.; Gattiker, A.; Hoogland, C.; Ivanyi, I.; Appel, R.D.; Bairoch, A. ExPASy: The Proteomics Server for in-Depth Protein Knowledge and Analysis. *Nucleic Acids Res.* **2003**, *31*, 3784–3788. [CrossRef] [PubMed]
49. Chou, K.C.; Shen, H.B. Plant-mPLoc: A Top-down Strategy to Augment the Power for Predicting Plant Protein Subcellular Localization. *PLoS ONE* **2010**, *5*, e11335. [CrossRef]
50. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. *Nucleic Acids Res.* **2015**, *43*, 39–49. [CrossRef]
51. Chen, S.; Zhou, Y.; Chen, Y.; Gu, J. Fastp: An Ultra-Fast All-in-One FASTQ Preprocessor. *Bioinformatics* **2018**, *34*, 884–890. [CrossRef]
52. Kim, D.; Paggi, J.M.; Park, C.; Bennett, C.; Salzberg, S.L. Graph-Based Genome Alignment and Genotyping with HISAT2 and HISAT-Genotype. *Nat. Biotechnol.* **2019**, *37*, 907–915. [CrossRef]
53. Liao, Y.; Smyth, G.K.; Shi, W. featureCounts: An Efficient General Purpose Program for Assigning Sequence Reads to Genomic Features. *Bioinformatics* **2014**, *30*, 923–930. [CrossRef]
54. Love, M.I.; Huber, W.; Anders, S. Moderated Estimation of Fold Change and Dispersion for RNA-Seq Data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef]
55. Spartz, A.K.; Lee, S.H.; Wenger, J.P.; Gonzalez, N.; Itoh, H.; Inzé, D.; Peer, W.A.; Murphy, A.S.; Overvoorde, P.J.; Gray, W.M. The SAUR19 Subfamily of Small Auxin up RNA Genes Promote Cell Expansion. *Plant J. For. Cell Mol. Biol.* **2012**, *70*, 978–990. [CrossRef] [PubMed]
56. Bai, F.; Li, S.; Yang, C.; Zhao, T.; Zhang, T.; Lan, X.; Liao, Z. Overexpression of the *AbSAUR1* gene enhanced biomass production and alkaloid yield in Atropa belladonna. *Ind. Crops Prod.* **2019**, *140*, 111705. [CrossRef]
57. Guo, Y.; Jiang, Q.; Hu, Z.; Sun, X.; Fan, S.; Zhang, H. Function of the Auxin-Responsive Gene TaSAUR75 under Salt and Drought Stress. *Crop J.* **2018**, *6*, 181–190. [CrossRef]
58. Wang, H.; Shang, Q. Genome-wide Identification and Expression Analysis of the SAUR Gene Family in *Cucumis sativus*. *Acta Hortic. Sin.* **2019**, *46*, 1093–1111. [CrossRef]
59. Aoki, H.; Onishi, A.; Miyashita, M.; Miyagawa, H.; Yatou, O.; Saito, K. Involvement of the Rice *OsSAUR51* Gene in the Auxin-Related Field Resistance Mechanism against Bacterial Blight Disease. *Jpn. Agric. Res. Q. JARQ* **2016**, *50*, 219–227. [CrossRef]
60. Guo, Y.; Xu, C.; Sun, X.; Hu, Z.; Fan, S.; Jiang, Q.; Zhang, H. *TaSAUR78* Enhances Multiple Abiotic Stress Tolerance by Regulating the Interacting Gene *TaVDAC1*. *J. Integr. Agric.* **2019**, *18*, 2682–2690. [CrossRef]
61. Liu, Y.; Xiao, L.; Chi, J.; Li, R.; Han, Y.; Cui, F.; Peng, Z.; Wan, S.; Li, G. Genome-Wide Identification and Expression of SAUR Gene Family in Peanut (*Arachis hypogaea* L.) and Functional Identification of *AhSAUR3* in Drought Tolerance. *BMC Plant Biol.* **2022**, *22*, 178. [CrossRef] [PubMed]
62. Zhou, L.X.; Yang, M.D.; Cao, H.X.; Yang, Y.D. Genome-Wide Identification and Bioinformatics Analysis of the Oil Palm *SAUR* Gene Family. *J. South. Agric.* **2022**, *23*, 13271. [CrossRef]
63. Jiao, L.; Si, X.; Guo, Z.; Li, P.; Liu, D. Identification and bioinformatics analysis of the SAUR gene family in melon. *China Veg.* **2023**, *5*, 42–55. [CrossRef]
64. Zhang, N.; Huang, X.; Bao, Y.; Wang, B.; Zeng, H.; Cheng, W.; Tang, M.; Li, Y.; Ren, J.; Sun, Y. Genome-Wide Identification of SAUR Genes in Watermelon (*Citrullus lanatus*). *Physiol. Mol. Biol. Plants* **2017**, *23*, 619–628. [CrossRef]
65. Cannon, S.B.; Mitra, A.; Baumgarten, A.; Young, N.D.; May, G. The Roles of Segmental and Tandem Gene Duplication in the Evolution of Large Gene Families in Arabidopsis Thaliana. *BMC Plant Biol.* **2004**, *4*, 10. [CrossRef] [PubMed]
66. Zhao, J.; Zhang, T. Bioinformatics analysis of SAUR gene family in *Brassica rapa*. *Chin. Agric. Sci. Bull.* **2012**, *22*, 130–137.
67. Knauss, S.; Rohrmeier, T.; Lehle, L. The Auxin-Induced Maize Gene *ZmSAUR2* Encodes a Short-Lived Nuclear Protein Expressed in Elongating Tissues. *J. Biol. Chem.* **2003**, *278*, 23936–23943. [CrossRef]
68. Kant, S.; Bi, Y.-M.; Zhu, T.; Rothstein, S.J. SAUR39, a Small Auxin-up RNA Gene, Acts as a Negative Regulator of Auxin Synthesis and Transport in Rice. *Plant Physiol.* **2009**, *151*, 691–701. [CrossRef] [PubMed]
69. Chen, Y.; Hao, X.; Cao, J. Small Auxin Upregulated RNA (SAUR) Gene Family in Maize: Identification, Evolution, and Its Phylogenetic Comparison with Arabidopsis, Rice, and Sorghum. *J. Integr. Plant Biol.* **2014**, *56*, 133–150. [CrossRef]
70. Chen, D.; Deng, Y.; Zhao, J. Distribution and Change Patterns of Free IAA, ABP 1 and PM H^+-ATPase during Ovary and Ovule Development of *Nicotiana tabacum* L. *J. Plant Physiol.* **2012**, *169*, 127–136. [CrossRef]

plants

MDPI

Article

Genome-Wide Analysis of the *Nramp* Gene Family in Kenaf (*Hibiscus cannabinus*): Identification, Expression Analysis, and Response to Cadmium Stress

Qin Liu [1,2,†], Shaocui Li [2,†], Guanghui Du [1,*] and Xia An [2,*]

[1] School of Agriculture, Yunnan University, Kunming 650500, China; lq18388746921@163.com
[2] Zhejiang Xiaoshan Institute of Cotton&Bast Fiber Crops, Zhejiang Institute of Landscape Plants and Flowers, Zhejiang Academy of Agricultural Sciences, Hangzhou 311251, China; lishaocui1992@163.com
* Correspondence: dgh2012@ynu.edu.cn (G.D.); anxia@zaas.ac.cn (X.A.)
† These authors contributed equally to this work.

Abstract: Kenaf (*Hibiscus cannabinu*) is a grass bast fiber crop that has the ability to tolerate and accumulate heavy metals, and it has been considered as a potential heavy metal accumulator and remediation plant. Nramp is a natural resistance-related macrophage, which plays an important role in the transport of divalent metal ions, plant growth and development, and abiotic stress. In this study, the *Nramp* gene family of kenaf was analyzed at the whole genome level. A total of 15 *HcNramp* genes were identified. They are distributed unevenly on chromosomes. Phylogenetic analysis classified 15 HcNramp proteins into 3 different subfamilies. All proteins share specific motif 4 and motif 6, and the genes belonging to the same subfamily are similar in structure and motif. The promoters are rich in hormone response, meristem expression, and environmental stress response elements. Under different treatments, the expression levels of *HcNramp* genes vary in different tissues, and most of them are expressed in roots first. These findings can provide a basis for understanding the potential role of the *Nramp* gene family in kenaf in response to cadmium (Cd) stress, and are of great significance for screening related Cd tolerance genes in kenaf.

Keywords: *Hibiscus cannabinus*; *Nramp* gene family; cadmium; phylogenetic analysis

Citation: Liu, Q.; Li, S.; Du, G.; An, X. Genome-Wide Analysis of the *Nramp* Gene Family in Kenaf (*Hibiscus cannabinus*): Identification, Expression Analysis, and Response to Cadmium Stress. *Plants* **2024**, *13*, 2514. https://doi.org/10.3390/plants13172514

Academic Editors: Jian Zhang, Jiezheng Ying, Yifeng Wang, Richard R.C. Wang and Jie Huang

Received: 25 July 2024
Revised: 22 August 2024
Accepted: 5 September 2024
Published: 7 September 2024

1. Introduction

Heavy metal pollution is one of the most serious environmental problems, with cadmium (Cd) pollution being a primary concern among heavy metal pollutants. Cd stress significantly impacts plant growth and development, altering various physiological and metabolic processes [1–4]. Unlike other elements, Cd does not have a dedicated transporter in plants. It enters the roots and is transported to the shoots by competing with other nutrients for transport pathways [5]. Cd stress can trigger the accumulation of reactive oxygen species in plants, leading to decreased activities of superoxide dismutase (SOD) and peroxidase (POD) enzymes, thereby reducing antioxidant and detoxification capacities. In turn, it disrupts the antioxidant system and can cause membrane lipid peroxidation [1,4]. Hyperaccumulator plants have efficient Cd absorption, transportation, and detoxification systems, and are the main materials for remediation of heavy-metal-contaminated soil, which can be used to study the adaptability of plants in extreme environments [6].

Kenaf (*Hibiscus cannabinus* L.), also known as hibiscus, is an annual herbaceous bast fiber crop in the Malvaceae [7]. It boasts a short growth cycle coupled with substantial biomass yield [8,9]. Kenaf and other bast fiber crops possess the characteristics of wide cultivation adaptability, making them versatile and multifunctional. Kenaf is recognized as a new type of papermaking raw material that can replace wood pulp. Additionally, it serves as an important raw material in the traditional bast fiber textile industry, and has diverse applications, such as automobile lining, paper mulching, fluff pulp, sewage

purification material, soil conditioner, plastic filler, activated carbon, and environmental protection adsorption material, and is known as a "potential superior crop in the 21st century" and "futuristic crop" [6,8–11]. Kenaf exhibits strong heavy metal tolerance and accumulation ability [1], and is not involved in the circulation of the food chain. It is considered to be a potential plant for repairing heavy metal pollution [1,12]. Cultivating crops in soils contaminated with heavy metals facilitates the remediation of degraded agricultural lands and enables the productive utilization of otherwise uncultivated areas [6,12]. In the treatment of heavy metal Cd stress, adding different forms and concentrations of nitrogen fertilizer can promote the growth of kenaf [13]. These crops can simultaneously provide fiber biomass raw materials for both traditional and innovative industries, offering new avenues for agricultural and industrial chain innovation [1].

Nramp (nature resistance-associated macrophage protein) is a natural resistance-associated macrophage protein [14]. It is an important proton/metal transporter in plants [15,16] and participates in the transport of heavy metals, such as zinc (Zn), iron (Fe), manganese (Mn), Cd, and arsenic (As) [5,17–23]. At the same time, it plays an important role in the transmembrane transport of divalent metals [5,10,15]. The Nramp sequence has a similar secondary structure among species [18], with 10–12 highly conserved transmembrane domains (TMD) [16,18,22,24], and has the structural characteristics of ion channels and ion transporters in evolution [18]. The gene family plays an important regulatory role in maintaining cell homeostasis, photosynthesis, protein activity, and plant response to environmental stress [18,24,25]. It keeps the balance of metal ions in plants through the transport of metal ions. It plays an important role in the detoxification of heavy metals [18,26,27] and has ion-selective changes in different genes [24]. Nramp proteins have been discovered in a variety of plants, including *Arabidopsis thaliana* [28], tomato (*Solanum lycopersicum*) [21], seabuckthorn (*Hippophae rhamnoides*) [25], soybean (*Glycine max*) [20], rice (*Oryza sativa*) [29], *Populus trichocarpa* [14], and potato (*Solanum tuberosum*) [21]. Nramp protein is the key to plant uptake and transport of Cd [6].

There is a lot of evidence that *Nramp* transcription factors in plants play a key role in various biological processes. *AtNramp1* is a transporter encoding Fe, Mn, and Cd [21,23,30], which is a key transporter of Mn absorption under low Mn conditions [18]. In response to Cd treatment, *AtNramp3* undergoes overexpression, thereby assuming a pivotal role in facilitating Fe transport and modulating the sensitivity of plants to Cd [28]. *AtNramp6* functions as a transporter for both Cd and Fe [23], exhibiting a dual role in nutrient and metal homeostasis. When Fe is deficient, *AtNramp6* has been shown to influence the growth of lateral roots in mustard plants [30], suggesting its importance in adapting to Fe limiting conditions. Despite its high sensitivity to Cd, *AtNramp6* does not significantly alter the overall Cd content within plants. Rather, it appears to enhance the tolerance of plants to Cd stress, allowing them to better withstand exposure to this toxic metal [18,25,30]. Rice *OsNramp1* is a transporter that plays a pivotal role in the absorption and translocation of Cd. Its involvement promotes the accumulation of both As and Cd within the plant. Specifically, when *OsNramp1* is overexpressed, it enhances the accumulation of Cd in leaves, contributing to elevated Cd levels in the photosynthetic tissues. Furthermore, when overexpressed in roots, *OsNramp1* increases the translocation of Cd to stems, indicating its role in facilitating the long-distance movement of this toxic metal within the rice plant [18]. *OsNramp6* is a Fe and Mn transporter located in the plasma membrane, which is helpful for disease resistance [5,19]. Tomato *SlNramp2* and *SlNramp3* are genes related to Cd absorption by roots, and they are highly expressed under Cd stress [31]. *SlNramp3* is a heavy metal transporter, which is involved in the transport of heavy metals from roots to stems [32]. *SaNramp1* from *Sedum alfredii* can transport Cd, Mn, and Zn, and its overexpression in tobacco significantly increases the contents of Cd, Zn, and Mn above ground [33]. Overexpression of *TtNramp1* can improve Cd transport, while overexpression of *TtNramp6* can increase Cd content in plants and promote Cd accumulation [34]. Tian et al. identified and analyzed the *Nramp* gene family of potato and its response to five kinds of heavy metal stresses, such as Cu^{2+}, Cd^{2+}, and Zn^{2+} [21]. Chen et al. identified the whole genome of

the *Nramp* gene family of *Spirodela polyrhiza* and analyzed the expression of Cd stress [19]. Hussain et al. identified and analyzed the whole genome of the *Nramp* gene family in *Kandelia obovate* and its response to copper (Cu) stress [35]. Genome-wide identification and evolutionary analysis of the *Nramp* gene family in the AC genome of *Brassica* species was performed by Zhao et al. [27]. However, there is no systematic analysis of *Nramp* transcription factors in kenaf to clarify their role in responding to heavy metal stress.

In this study, the *Nramp* of kenaf was comprehensively analyzed. Fifteen *Nramp* genes were identified by bioinformatics, and the phylogenetic analysis, conserved motifs, domains, gene structure, subcellular localization, chromosome position, cis-regulatory elements, and expression profile were deeply analyzed at the whole genome level. In addition, qRT-PCR was used to detect the effects of four treatments on the expression pattern of *HcNramps*. Our findings serve as a solid foundation for future endeavors aimed at unraveling the multifaceted functions of the *Nramp* gene family in kenaf's resilience against heavy metal stress. They offer invaluable insights into the evolutionary mechanisms governing *HcNramps*, while simultaneously presenting novel strategies to tackle the challenge of heavy metal accumulation in plants. In essence, this study provides insights into a deeper understanding of the interactions between plants and metals and the development of innovative solutions to reduce environmental pollution.

2. Results

2.1. Identification and Characterization of Nramp Genes in Kenaf

Following a search of the kenaf genome database using HMMER 3.0, the identified candidate sequences were examined using NCBI Batch CDD, Pfam, and SMART to confirm the presence of the Nramp domain. A total of 15 *HcNramp* gene family members were identified from the kenaf genome (Table S1), which were systematically named based on their chromosomal positions, designated as *HcNramp1* through *HcNramp15* (Table S2). The analysis of the physicochemical properties revealed a diverse range of characteristics among the identified proteins. Specifically, the length varied from 221 to 1379 amino acids, with the relative molecular weight range from 24,223.94 to 150,763.33 Da. The theoretical isoelectric point (pI) ranged from 5.08 to 10.18, with one member having a pI value of 7. Seven were alkaline proteins and another eight were acidic proteins. The instability coefficient ranged from 30.06 to 44.88. Among them, HcNramp2/8/9/10/12 protein was unstable. From the Aliphatic index, HcNramp7 protein was the most stable, with a reading of 124.21. HcNramp12 protein was the most unstable, with a reading of 89.04. The average total hydrophilicity (GRAVY) ranged from −0.092 to 0.656, and HcNramp2/9/10/12 were hydrophilic proteins, and the remaining proteins were hydrophobic proteins. Subcellular localization showed that all HcNramp proteins were located on the plasma membrane, HcNramp3/5 proteins were also located in the vacuole, and HcNramp11 proteins were located on the endoplasmic reticulum.

2.2. Secondary Structure Analysis of HcNramp Family Proteins

The secondary structures of HcNramp family proteins were mainly composed of α-helix, extended chain, folding, and random coil (Table S3). Among them, the proportion of α-helix was the highest, ranging from 39.85% to 59.08%, and the proportion of extended chain was the least, ranging from 5.15% to 13.57%. Only HcNramp2 protein had folding, no random coil, and the rest of the proteins had random coil, indicating that the protein structures were stable. The proportion of secondary structures of HcNramp8/9/10 was extended strand < alpha-helix < random coil, which indicates that the three protein structures were variable and may be the sites of proteins with special functions.

2.3. Classification, Phylogenetic Analysis, Gene Structure, and Conserved Motifs of the HcNramp Gene Family

In order to understand the evolutionary relationship of the *Nramp* gene family in kenaf and other species, a phylogenetic tree of 29 Nramp protein sequences was constructed,

including 15 HcNramp kenaf proteins, 7 *A. thaliana* proteins, and 7 tomato proteins (Table S4). These three species were divided into three categories (Figure 1), which were the I, II, and SLC5 subfamilies. HcNramp1/6/7/14 belong to subfamily I, with proteins ranging in length from 501 to 564 amino acids, theoretical pI ranging from 5.08 to 5.41, instability coefficient ranging from 32.56 to 37.38, and Aliphatic index ranging from 116.17 to 124.21, indicating that the proteins in this subfamily were stable, all of which were hydrophobic proteins, and this subfamily was secondary in protein. HcNramp3/4/5/11/13/15 belong to subfamily II, with amino acids ranging from 221 to 649, pI ranging from 7.12 to 10.18, instability coefficient ranging from 30.06 to 33.82, and Aliphatic index ranging from 122.53 to 106.97. The proteins were stable and all hydrophobic. HcNramp2/8/9/10/12 belong to the SCL5 subfamily, with the proteins ranging in length from 1300 to 1379 amino acids, the theoretical pI ranging from 5.43 to 7, the instability coefficient ranging from 41.11 to 44.88, and the Aliphatic index ranging from 89.04 to 94.10. All proteins were unstable except HcNramp in the SCL5 subfamily. In the secondary structure of proteins, the proportion of folding was the highest in HcNramp2, and the proportion of random curl was the highest in HcNramp8/9/10. The proportion of α-helix and random curl in HcNramp12 is the same. Protein in the same subfamily is species-specific, indicating that these proteins may have similar structures. The cross-species phylogenetic analysis of Nramps in kenaf, *A. thaliana*, and tomato showed that all the homologous Nramps could be well clustered together (Figure 1). HcNramp1/6/7/14 in kenaf, AtNramp2/5/4/7 in *A. thaliana*, and SlNramp1/4 in tomato belong to subfamily I. HcNramp3/4/5/11/13/15 in kenaf, AtNramp1/3 in *A. thaliana*, and SlNramp2/8 in tomato belong to subfamily II. HcNramp2/8/9/10/12 in kenaf, AtNramp6 in *A. thaliana*, and SlNramp5/6/7 in tomato belong to subfamily SCL5.

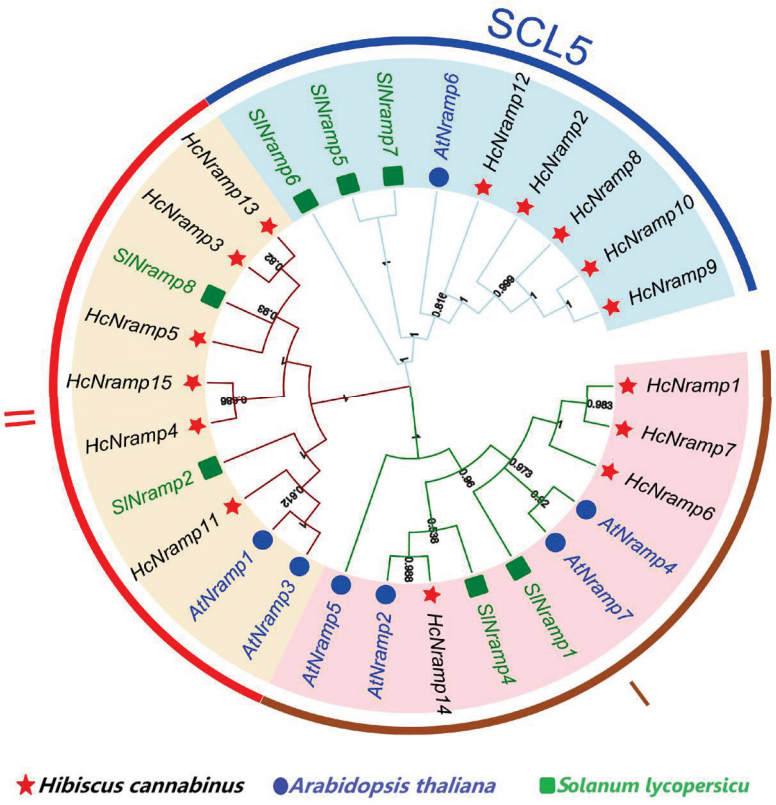

Figure 1. Phylogenetic analysis of 29 Nramp proteins in kenaf, *A. thaliana*, and tomato. Fifteen proteins of kenaf are labeled in black font, seven proteins of *A. thaliana* are in blue font, and seven proteins of tomato are in green font. The maximum likelihood tree was created using MEGAX (bootstrap value = 1000) and the bootstrap value of each branch is displayed.

The conserved motif prediction of the kenaf Nramp protein sequence was performed using MEME Suite 5.5.5 online software. The results showed that the motifs between different subfamilies were different, and the motifs between members of the same subfamily were relatively conserved (Figure 2 and Table S5). In the Nramp gene family of kenaf, all 15 *HcNramps* contained motif 4 and motif 6, indicating that the motif of the kenaf Nramp gene is relatively conservative. Most of the HcNramps contained motifs 1/2/4/5/6/8/9, where only *HcNramp4* and *HcNramp5* contained only motifs 4 and 5, respectively. The difference in the number of conserved motifs of the gene indicates that the gene has functional differences with other genes and may have different biochemical characteristics and biological functions [3]. Gene structure analysis showed that each gene contained introns, and the number of introns varied greatly, with a maximum of 17 introns and a minimum of 3 introns. The *HcNramp* genes clustered in the same branch of the phylogenetic tree were similar in exon–intron tissues. The *HcNramp3/4/5/13/15* genes in the same branch did not contain untranslated regions (UTRs) at the 5′ and 3′ ends. Among them, *HcNramp12* did not contain untranslated regions (UTRs) at the 3′ end, and the remaining nine genes had complete gene structures.

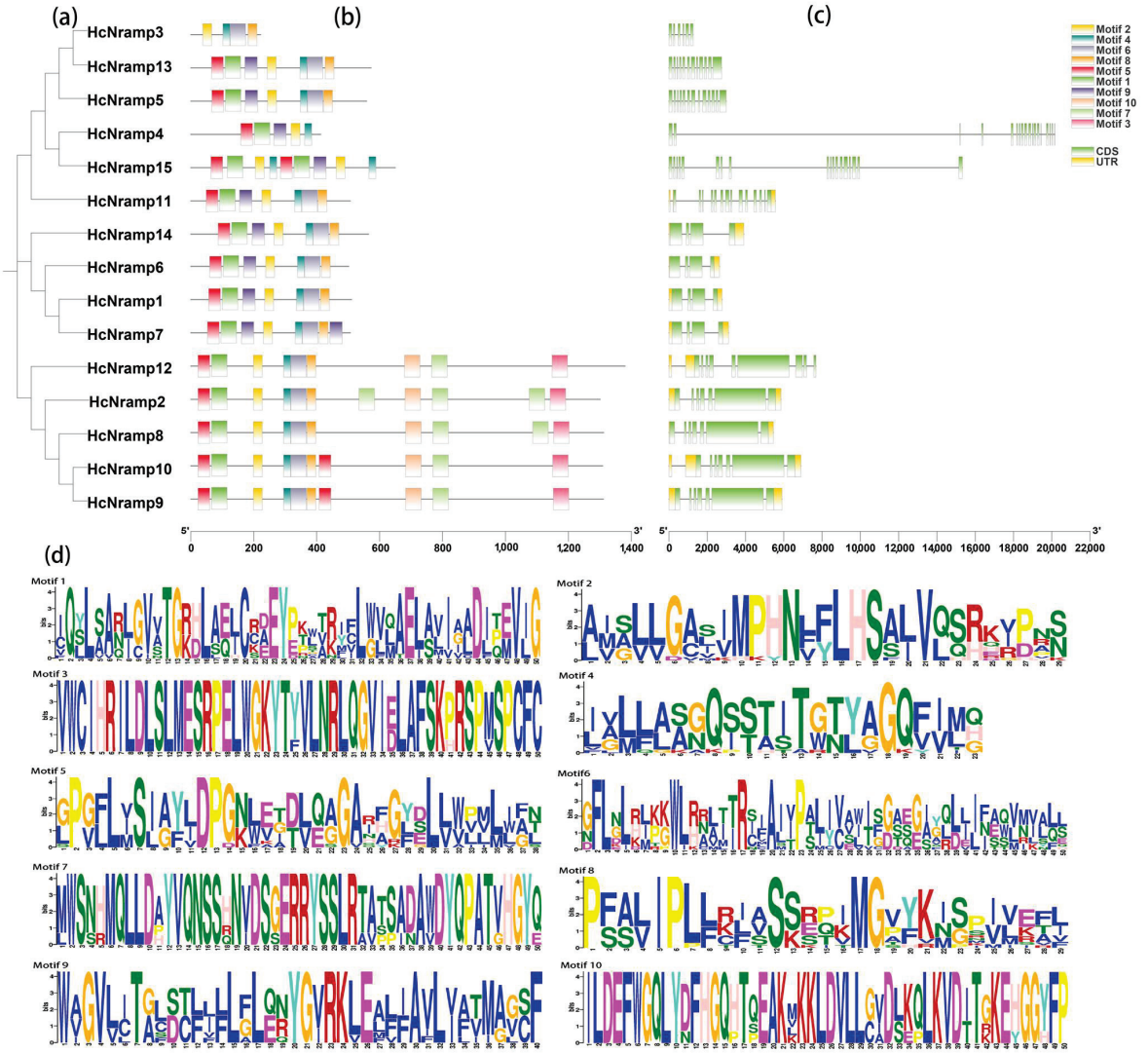

Figure 2. Conservation motif and gene structure analysis of *HcNramp* genes according to the phylogenetic relationship. (**a**) The phylogenetic relationship of HcNramps. (**b**) Conserved motifs of HcNramps. Different colors represent different motifs. (**c**) Exon and intron structures of the *HcNramp* genes in kenaf. The grey lines indicate introns, the green boxes represent exons, and the orange boxes indicate untranslated regions. (**d**) The amino acid composition of each motif.

2.4. Analysis of Cis-Acting Elements of HcNramp Promoter

The cis-regulatory elements in the promoter region (including 2,000 base pairs upstream) were comprehensively analyzed and screened, and the online software PlantCare was used for analysis (Figure 3). The results showed that 156 cis-acting elements were divided into 3 parts, including abiotic stress response elements (67 elements accounted for 42.95%), hormone response elements (85 elements accounted for 54.49%), and growth and development response elements (4 elements accounted for 2.56%). *HcNramp* gene was involved in a complex regulatory network and was regulated by various environmental, developmental, and physiological factors.

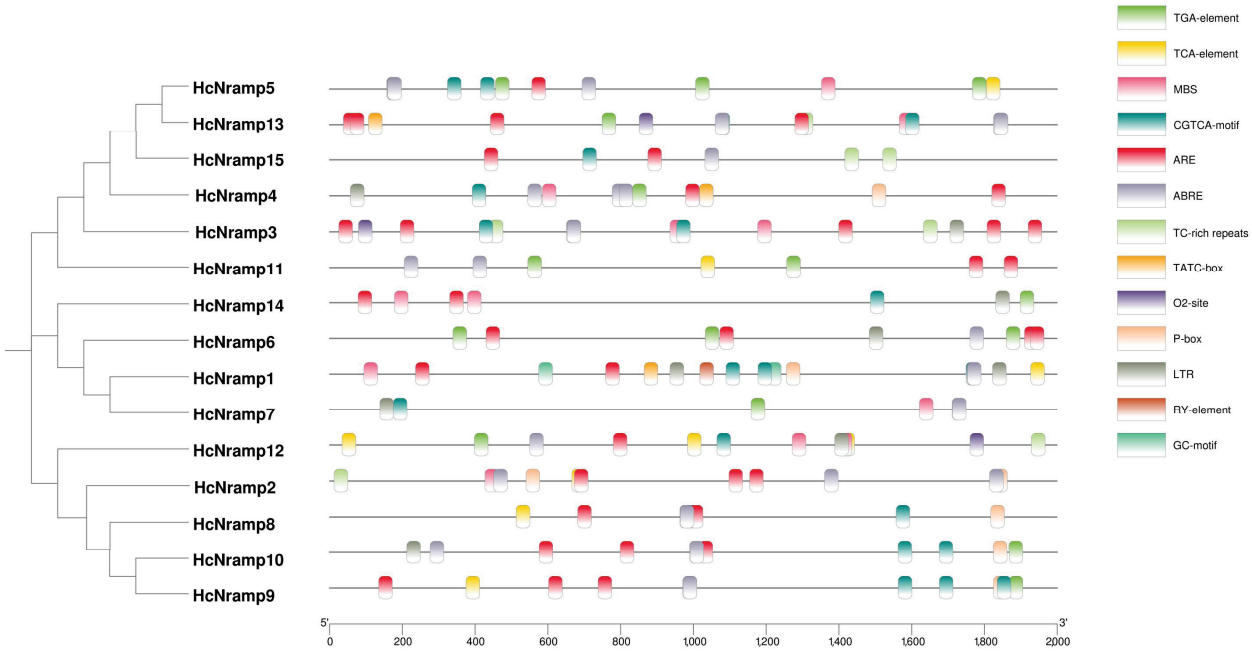

Figure 3. Identification of the cis-acting elements in the promoter of *HcNramp genes*.

The abiotic stress elements are composed of hypoxia/anaerobic induction, defense and stress elements, and cis-elements necessary for drought and low temperature. These abiotic stresses were composed of 2 GC-motifs (2.99%), 37 AREs (55.22%), 7 TC-rich repeats (10.45%), 12 MBSs (17.91%), and 9 LTRs (13.43%). *HcNramp3* contains ten abiotic stress response elements, including two TC-rich repeats, two MBSs, one LTR, and five AREs; secondly, *HcNramp1* contains seven abiotic stress response elements, including one MBS, two LTRs, two GC-motifs, and two AREs. This indicates that *HcNramp3* may play an important role in anaerobic stress.

The hormone response elements include auxin, gibberellin, salicylic acid, abscisic acid, and methyl jasmonate response elements. They were composed of 15 TGA elements (17.65%), 3 TATC-box (3.53%), 7 P-box (8.24%), 9 TCA elements (10.59%), 31 ABREs (36.47%), and 20 CGTCA motifs (23.53%). *HcNramp5* has 10 hormone response elements, including 3 TGA elements, 1 TCA element, 2 CGTCA motifs, and 4 ABREs, indicating that the gene may be involved in the regulation of abscisic acid. Secondly, there were two genes containing eight hormone response elements, namely, *HcNramp1* (including one TCA element, one TATC-box, one p-box, three CGTCA motifs, and two ABREs) and *HcNramp9* (one TGA element, one TCA element, one P-box, three CGTCA motifs, and two ABREs), indicating that *HcNramp1* and *HcNramp9* genes are involved in the regulation of methyl jasmonate.

Growth and development response elements, including seed-specific cis-acting elements and zein metabolic regulation, were composed of one RY element (25%) and three O2 sites (75%). Among them, *HcNramp1* contains a RY element, which indicates that the gene is involved in the specific regulation of seed expression. *HcNramp3/2/13* all contained one O2 site, indicating that these three genes may be involved in the metabolism of gliadin.

2.5. Chromosomal Localization and Collinearity Analysis of HcNramp Genes

The 15 *HcNramp* genes were located on 9 chromosomes, and the distribution of each gene on the chromosome was uneven. The *HcNramp* gene was most distributed on Chr3 and Chr6 chromosomes, with three genes located, followed by two genes located on both Chr2 and Chr5 chromosomes, and only one gene located on Chr7, Chr8, Chr12, Chr13, and Chr14 (Figure 4).

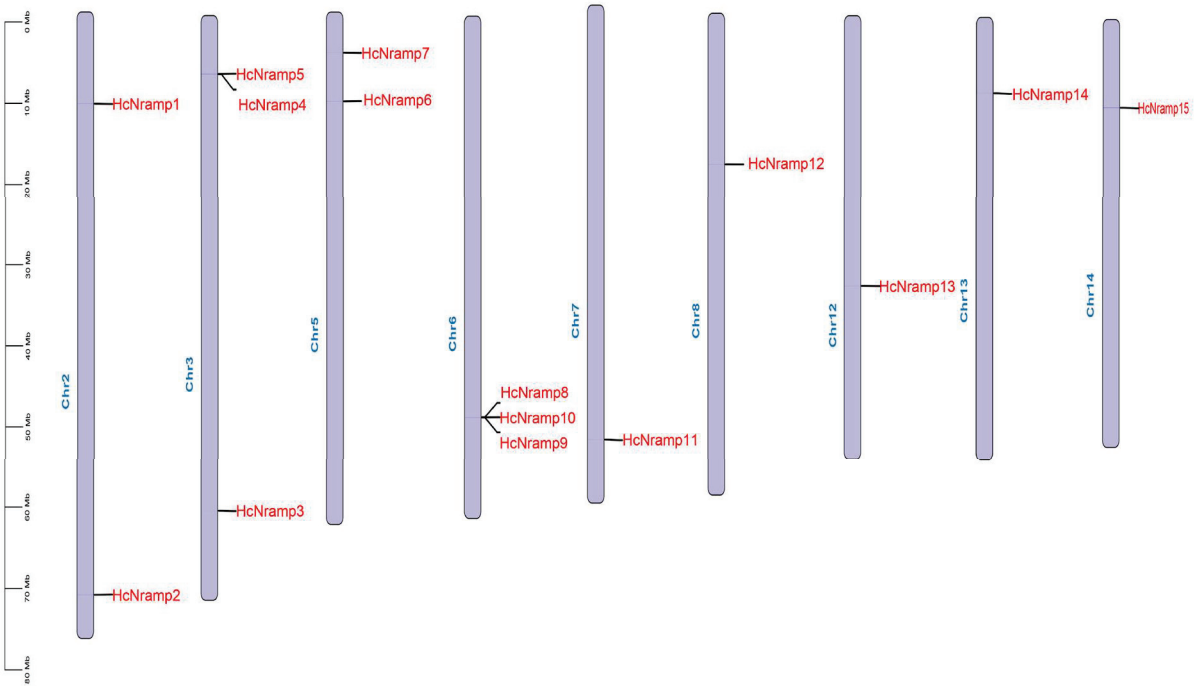

Figure 4. Chromosome location of *HcNramp* genes. Chromosome numbers are shown to the left of the chromosomes. *Nramp* genes are labeled to the right of the chromosomes. Scale bar on the left indicates the chromosome lengths (Mb).

In order to clarify the evolutionary relationship, the replication of the *HcNramp* gene was studied. It was found that there were six collinear gene pairs in the kenaf species (Figure 5), including H*cNramp1*:*HcNramp6*, *HcNramp1*:*HcNramp7*, and *HcNramp6*:*HcNramp7* in subfamily I, *HcNramp3*:*HcNramp15* in subfamily II, and *HcNramp2*:*HcNramp8* in subfamily SCL5. In order to further study the homology of *Nramp* gene members among different species, collinear maps of *Nramp* gene families of kenaf and *A. thaliana*, and kenaf and tomato, were drawn. There were ten collinear gene pairs in kenaf and *A. thaliana*, and ten collinear gene pairs in kenaf and tomato (Figure 6 and Table S6).

2.6. Tissue-Specific Expression Analysis of the HcNramp Genes

The expression pattern of a gene is crucial for elucidating its function [27]. In order to further explore the role of *HcNramp* genes, a heat map showing the expression pattern of the *HcNramp* gene was constructed by using the transcriptome data (Accession: PRJNA1134624). Tissue-specific expression analysis was carried out by calculating the TPM value of the gene. All *HcNramp* genes showed tissue-specific expression (Figure 7 and Table S7). Fifteen *HcNramp* genes were divided into three different groups: the first group was *HcNramp3/4/6/7*, the second group was *HcNramp2/8/9/14*, and the third group was *HcNramp1/5/10/11/12/13/15*. The gene expression levels were different in different treatments, with significant differences between normal treatment and Cd treatment. The expression trends of genes in CK-H and CK-M, and Cd-H and Cd-M treatments were similar, but the specific expression levels were different. *HcNramp1/6* were expressed relatively high in Cd treatment. *HcNramp2/14* was highly expressed in the roots and stems of CK-H and

CK-M treatments. *HcNramp7* was preferentially expressed in leaves, *HcNramp2/6/14* in stems, and *HcNramp1/5/8/9/11/12/13/15* in roots. *HcNramp3/4/7* was low expressed in roots and *HcNramp10* in stems. *HcNramp2/14/8/9/11* was low or not expressed in leaves, while *HcNramp5/13/15* was low expressed in stems and leaves. *HcNramp13* was highly expressed in Cd-M-R, and *HcNramp15* was highly expressed in CK-H-R. *HcNramp4* was low expressed in CK-H-R, *HcNramp6* was low expressed in CK-M-R, *HcNramp9* was low expressed in Cd-M-L, and *HcNramp12* was low expressed in CK-H-L.

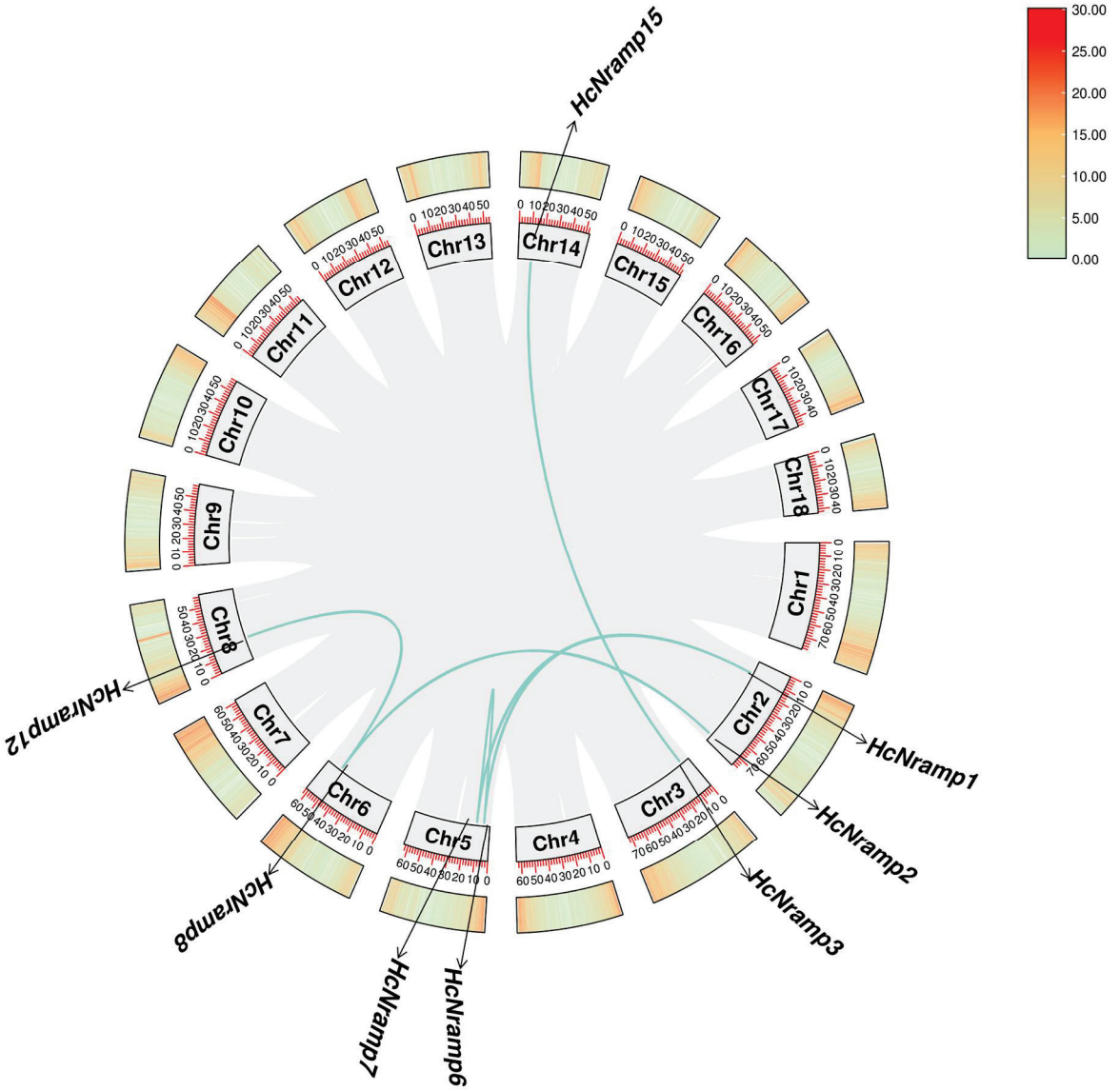

Figure 5. Intraspecific collinearity analysis of the kenaf *Nramp* gene family. The green lines in the figure are common genes, and the gray lines represent collinear blocks of the plant genome.

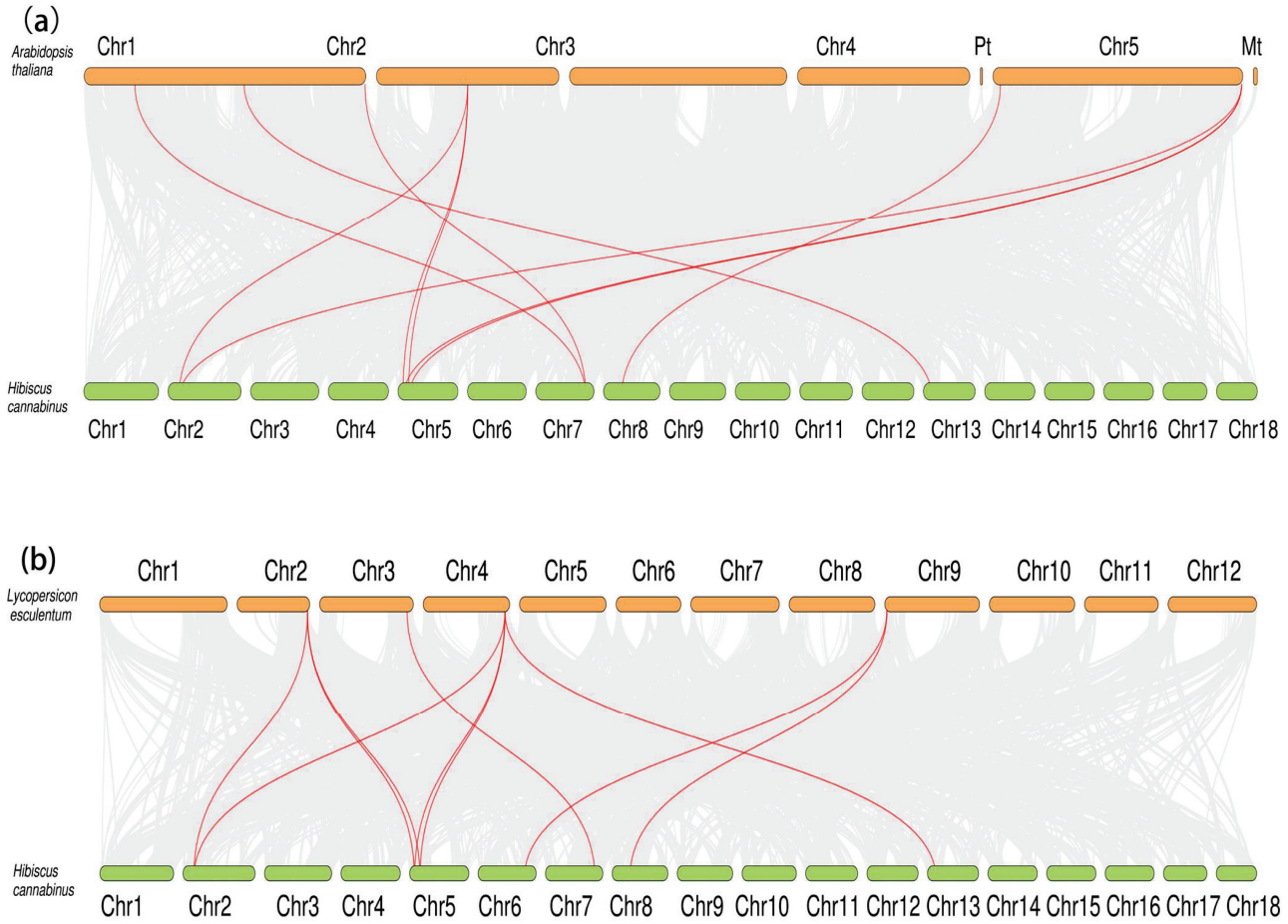

Figure 6. Interspecific collinearity analysis of the kenaf *Nramp* gene family. (**a**) Interspecific collinearity analysis of kenaf and *A. thaliana*. (**b**) Interspecific collinearity analysis of kenaf and tomato. The red lines in the figure are the common genes between kenaf, *A. thaliana,* and tomato, and the gray lines represent the collinear blocks of the plant genome.

2.7. HcNramp Gene Expressions under Exposure to Different Treatments of Heavy Metals

To clarify the response of the *HcNramp* genes to Cd stress, we used qRT-PCR to analyze the gene expression level in roots, stems, and leaves under four different treatment conditions (Figures 8–10). There were notable differences in the expression patterns of *HcNramp* genes across roots, stems, and leaves, all significantly influenced by Cd stress. According to the results of the relative gene expression levels in kenaf roots (Figure 8), Cd treatment alone significantly upregulated the expression of *HcNramp1/4/7/13/14* compared to the control. Furthermore, combined treatment of Cd and melatonin significantly enhanced the expression of HcNramp1/4/6/7/13, but the expression levels of other genes did not change significantly. In the kenaf stems (Figure 9), melatonin treatment alone significantly increased the expression of HcNramp3/4/5/6/7/8/9/10/11/12/13/14 compared to the control. On the other hand, Cd treatment significantly upregulated the expression of HcNramp3/5/11. In the kenaf leaves (Figure 10), melatonin treatment significantly increased the expression of HcNramp7 and 11 compared to the control. Cd treatment alone significantly upregulated the expression of HcNramp11/15, while combined Cd and melatonin treatment significantly enhanced the expression of *HcNramp11*. These findings provide valuable insights into the differential regulation of *HcNramp* genes in response to Cd stress and melatonin treatment, highlighting their potential roles in Cd tolerance and detoxification mechanisms in kenaf.

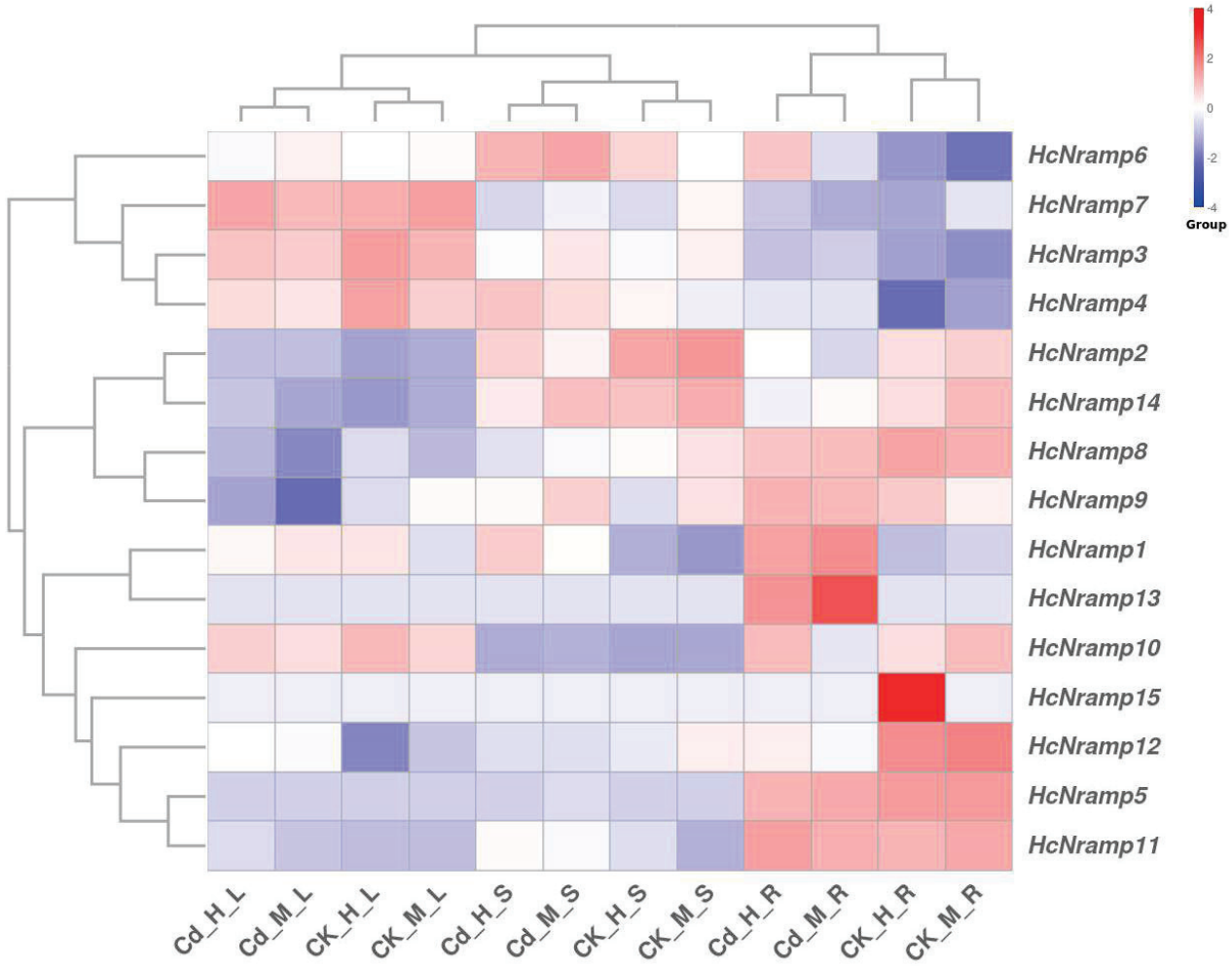

Figure 7. Tissue-specific expression profiles of *HcNramp* genes in kenaf. CK-H-R represents the root samples treated by CK + water spraying, CK-M-R represents the root samples treated by CK + melatonin spraying, Cd-H-R represents the root samples treated by Cd + water spraying, and Cd-M-R represents the root samples treated by Cd + melatonin spraying. L, S, and R represent the leaves, stems, and roots of kenaf samples. Red and blue boxes indicate high and low expression levels of genes, respectively.

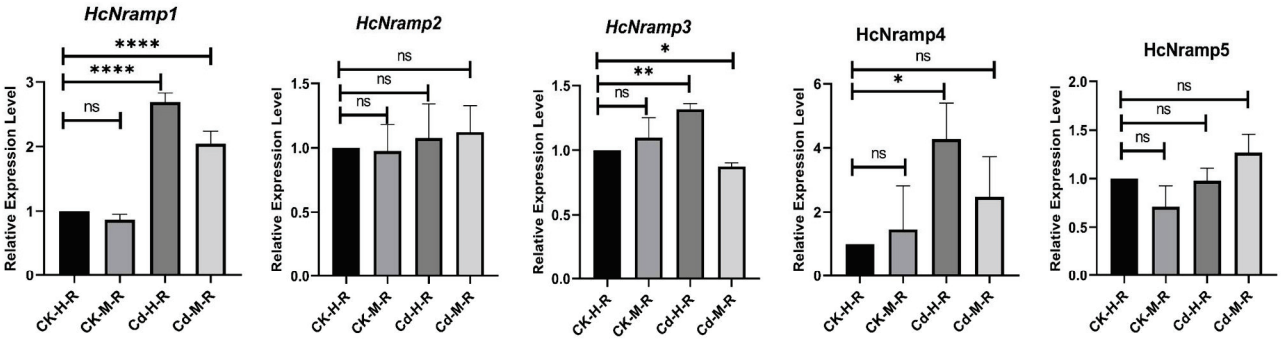

Figure 8. *Cont.*

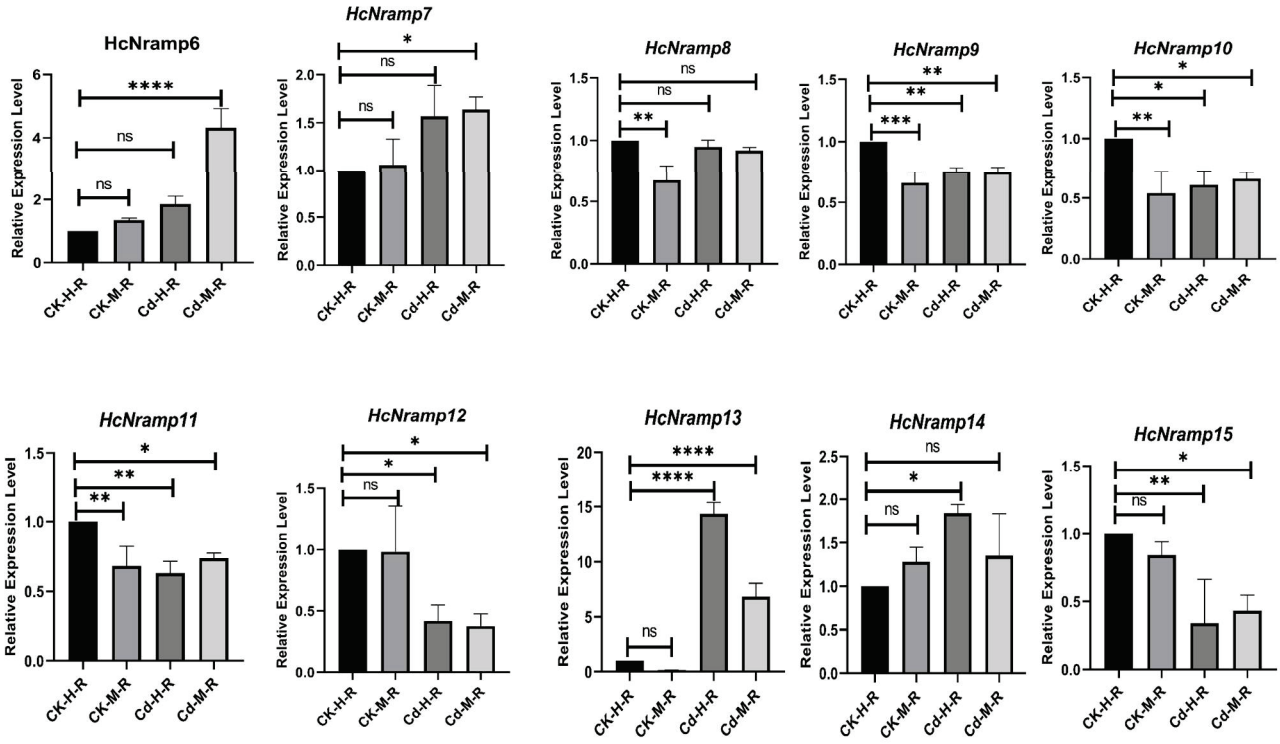

Figure 8. The expression levels of *HcNramp* genes in kenaf roots under different treatments. Data represent means ± SD of three biological replicates. CK-H-R represents control samples. The expression levels of each gene were normalized against its own expression level in the root tissue in control (CK-H-R) conditions. Error bars indicate mean ± SD and asterisks indicate statistical differences between the treatment samples and the corresponding control samples, the roots (n = 3; * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$, **** $p < 0.0001$ and ns not significant; Student's *t*-test).

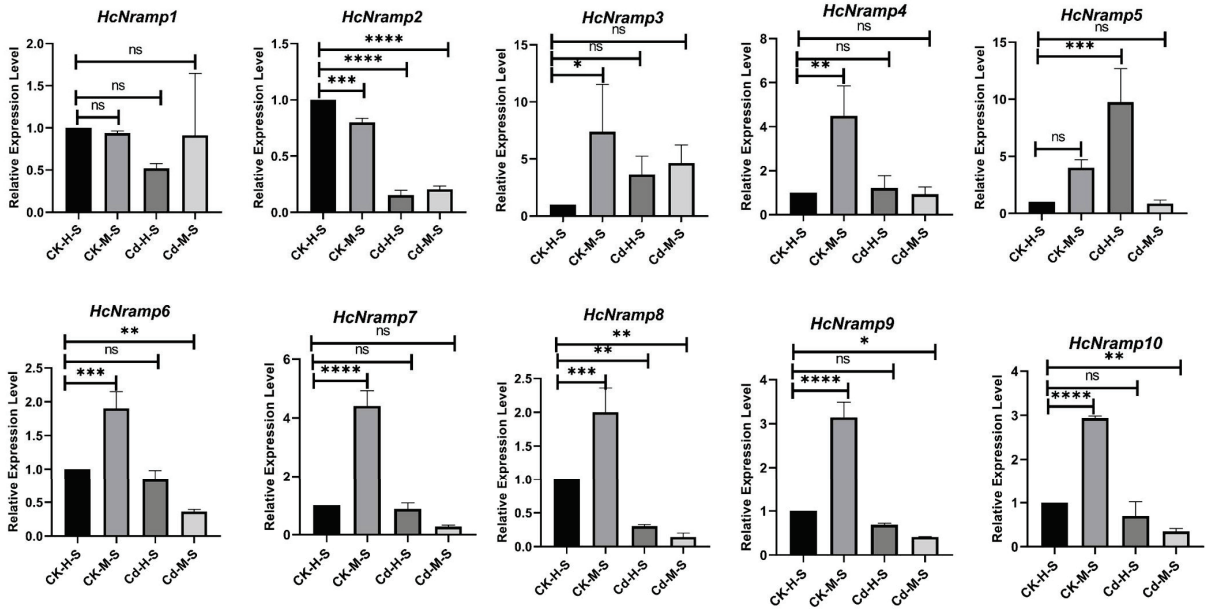

Figure 9. *Cont.*

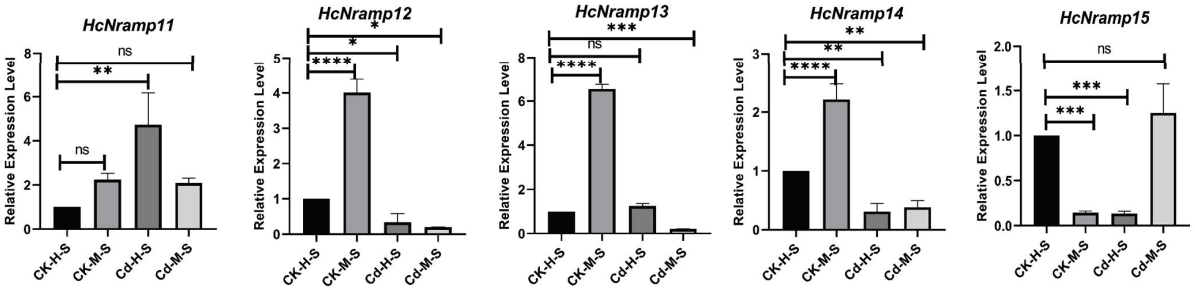

Figure 9. The expression levels of *HcNramp* genes in kenaf stems under different treatments. Data represent means ± SD of three biological replicates. CK-H-S represents control samples. The expression levels of each gene were normalized against its own expression level in the stem tissue in control (CK-H-S) conditions. Error bars indicate mean ± SD and asterisks indicate statistical differences between the treatment samples and the corresponding control samples, the stems (n = 3; * $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$, **** $p < 0.0001$ and ns not significant; Student's *t*-test).

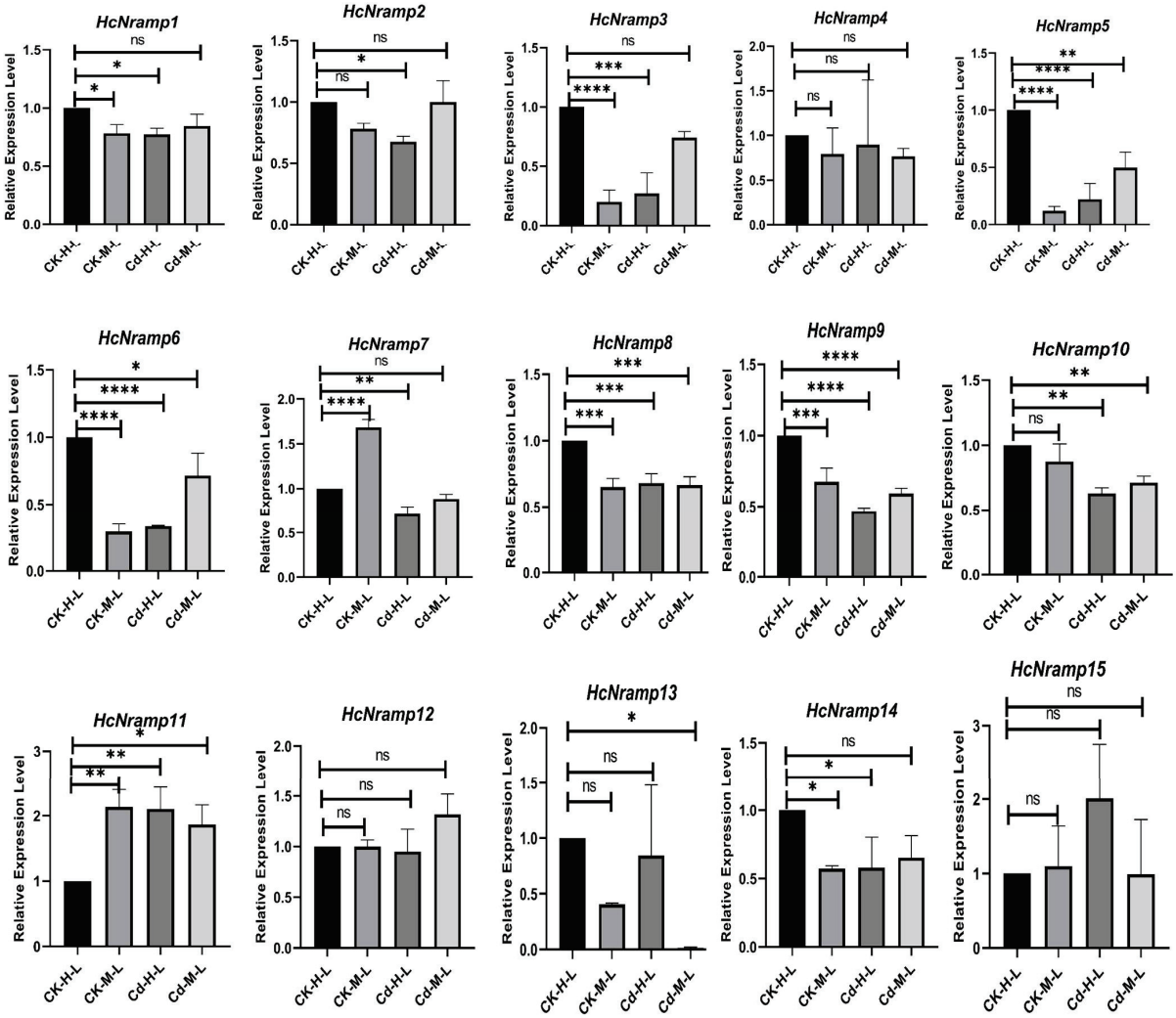

Figure 10. The expression levels of *HcNramp* genes in kenaf leaves under different treatments. Data represent means ± SD of three biological replicates. CK-H-L represents control samples. The expression levels of each gene were normalized against its own expression level in the leaf tissue in control (CK-H-L) conditions. Error bars indicate mean ± SD and asterisks indicate statistical differences between the treatment samples and the corresponding control samples, the leaves (n = 3; * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$, **** $p < 0.0001$ and ns not significant; Student's *t*-test).

3. Discussion

Kenaf is an annual herbaceous fiber crop, which has the characteristics of a short growth cycle, large biomass, and not participating in the circulation of the food chain [6]. It can not only be used to remediate polluted soil, but also has certain economic benefits. Kenaf is considered as a potential plant to remediate soil contaminated by heavy metals and has attracted much attention in the field of heavy metals [8,10]. Therefore, it is of great significance to analyze the *Nramp* gene family of kenaf using bioinformatics.

The *Nramp* gene plays an important role in the absorption and transport of metal ions. A large number of *Nramp* genes have been identified functionally [34] and have been analyzed in many species [24]. In our study, fifteen members of the *Nramp* gene family were identified from the whole genome of kenaf. Compared with *A. thaliana* (7) and tomato (9) [21], the number of *Nramp* in kenaf was the largest. The subcellular localization of the *Nramp* gene family in kenaf is mostly located in the plasma membrane, and a few are located in the vacuole membrane and endoplasmic reticulum. Protein located in the plasma membrane mainly absorbs metal ions from plants in vitro, which may be related to its function of transporting heavy metal ions across membranes [23]. Protein located in the vacuole membrane is responsible for transporting metal ions and their steady state [25]. Protein located in endoplasmic reticulum is mainly involved in the process of cell synthesis and modification of protein. Therefore, it can be inferred that kenaf is a heavy-metal-rich plant with a strong absorption ability and weak transport ability.

Phylogenetic analysis clustered the Nramp proteins into three distinct subfamilies, which was consistent with the assessment that subdivided *A. thaliana* Nramp proteins into three subgroups (Figure 1) [21]. According to previous studies, different structures and subfamilies have different functions [7]. Further *HcNramp* gene structure analysis (Figure 2) showed that the members of the most closely related subfamilies usually shared similar exon–intron structures, indicating they have similar evolutionary relationships. All members of the *HcNramp* gene family contained introns, and conserved motifs 4 and 6, and only the protein in the SCL5 subfamily had motif 3. All members contained motif 4 (GQSSTITGTYAGQFIMQ), a transport motif of the *Nramp* family, which indicated that the *Nramp* family is highly conserved [25]. Furthermore, the *HcNramp* genes in the different subfamilies displayed different exon–intron structures, while subtle differences may play a critical role in the gene evolution progress. Therefore, it is speculated that different gene structures are involved in divergent functions.

Twenty-nine protein sequences of kenaf, *A. thaliana*, and tomato were used to construct the phylogenetic tree. According to the phylogenetic tree, they can be divided into three subfamilies. According to the grouping of *A. thaliana* and tomato, these three subfamilies were named I, II, and SCL5, respectively [21]. The protein grouping of *A. thaliana* and tomato in each branch of the phylogenetic tree was consistent with Tian's findings [21]. Different species have different evolutionary speeds, so there are differences in the classification of evolutionary trees. The distribution of *HcNramp* genes in the three subfamilies of kenaf was uneven. Subfamily I contained four *HcNramp* genes, subfamily II contained six *HcNramp* genes, and subfamily SCL5 contained five *HcNramp* genes. These differences indicate that there is asymmetric evolution among the three subfamilies of kenaf [5]. In this study, *SlNramp3* and *SlNramp9* proteins were not mentioned because *SlNramp3* and *SlNramp9* in tomato belong to the Mnth family, while *HcNramp* in kenaf has only three subfamilies, and there are no members of the Mnth family in *A. thaliana* and kenaf [21]. The *Nramps* of *A. thaliana* and tomato were collinear with 15 *HcNramp* genes of kenaf, and there were 10 collinear gene pairs between the 2 species and kenaf, which indicated that there was a close evolutionary relationship between them. Among them, *HcNramp1/6/7/11/12/14* genes of kenaf had collinear gene pairs with *AtNramp1/2/3/4/6/7* of *A. thaliana*, and *HcNramp1/2/6/7/8/11/12/14* had collinear gene pairs with *SlNramp1/2/3/7* of tomato. Previous studies have mentioned that there are multiple collinear gene pairs between species, which indicates that there are multiple gene copies in the process of species evolution [36]. There were 6 pairs of collinear genes in kenaf, 10 pairs in kenaf and *A. thaliana*, and 10 pairs

in kenaf and tomato. Previous studies have mentioned that there are multiple collinear gene pairs between species, which indicated that there are multiple gene copies in species evolution [21].

The *Nramp* family is a family of transporters, which can transport divalent metal cations. All *HcNramp* have transport motifs [25], so it is speculated that these proteins may transport Cd related to transport motifs. The cis-acting promoter elements were identified to address *HcNramp* gene expression regulation. The analysis of cis-acting elements of promoters showed that the promoter sequences of genes in the same subfamily were similar. The promoter region of *HcNramps* contained a variety of stress response elements and hormone response elements, so it is speculated that *HcNramps* genes may be regulated by complex regulatory networks [21]. Studies have shown that some phytohormones, such as jasmonic acid (JA) and SA, participate in the plant response to different metal stress [37,38]. We also found that in the *HcNramp* gene promoters, there were many cis elements associated with heavy metal stress, which may be involved in the transport and absorption of metals. For example, SA significantly reduces Cd in rice grains by regulating the expression levels of *OsNramp2* associated with Cd translocation and accumulation [39]. Among the various *HcNramp* genes, *HcNramp2/3/12/13/15* are unique in possessing TC-rich repetitive sequences, which may confer specific regulatory or structural properties.

The expression of 15 *HcNramp* genes was divided into 3 different groups. The expression trends of the genes in the same group were similar, with significant differences between the different groups [24]. In different treatments, the expression level of the genes was different, and there was a significant difference between CK and Cd treatments. However, the expression trends of CK-H and CK-M, and Cd-H and Cd-M, were similar, but the specific expression levels were different. It is speculated that the reason for the above situation may be that the *Nramp* gene family is mainly involved in the transport of divalent cations, and *HcNramps* may be involved in the transport of Cd^{2+}, so there is a significant difference in expression between CK and Cd treatments, and the expressions in CK-H and CK-M, and Cd-H and Cd-M, are similar. There are significant differences between CK and Cd treatments in *BnNRAMPs* (*Brassica napus*) [27], between CK and Pb (lead) stress in *HrLNRAMPs* [25], and between CK and Cu treatments in *KoNRAMPs* (*Kandelia obovata*) [35], which are consistent with the significant differences in the expression of *HcNramps* between CK and Cd treatments in this paper. These results were consistent with previous studies, indicating that *Nramps* were responsible for Cd uptake and transport.

HcNramp3 exhibited preferential expression in leaves, whereas the remaining four genes, *HcNramp2/12/13/15*, were predominantly expressed in roots. This differential expression pattern suggests that these genes may play an important role in roots, as they may be induced by stress. This function is similar to the mechanisms observed in peanuts [24]. *AtNramp1* is a Cd transporter [21,30], and *AtNramp3* plays a role in Cd sensitivity [28]. *HcNramp3/11* are homologous genes, so it can be speculated that these genes may be candidates for Cd regulation of kenaf. *AtNramp5* is involved in Cd transport [19], and it has homology with four *HcNramp5* genes in subfamily I, and all of them are located in stable hydrophobic proteins on the plasma membrane, so it is speculated that *HcNramp1/6/7/14* genes may be involved in Cd transport. *AtNramp6* is a transport protein of Cd [23], and it can also improve the tolerance of plants to heavy metal Cd [18,25,30]. It has homology with five genes of the SCL5 subfamily, indicating that *HcNramp2/6/8/9/12* may also improve the tolerance of kenaf to Cd. The tolerance of heavy metal hyperaccumulator plants and their ability to accumulate metals mainly depend on their adaptation to the dynamic balance of metals. In some hyperaccumulator plants, the *Nramp* metal transporter is highly expressed [40]. High-accumulation plants can enrich toxic heavy metals through the *Nramp* gene, which can improve the soil where heavy metals accumulate and is beneficial to the development of phytoremediation biotechnology of transgenic plants [18]. Therefore, the *Nramp* gene family can be excavated more comprehensively and systematically, and its biological function can be deeply analyzed.

4. Materials and Methods

4.1. Identification and Sequence Analysis of Nramp Genes in Kenaf

The genomic data of kenaf were obtained from the nucleic acid sequence archiving system (CNCB; https://download.cncb.ac.cn/gwh/Plants/Hibiscus_cannabinus_Fuhong952_GWHACDB00000000.1/, accessed on 6 May 2024) of the China National Center for Bioinformation [41], including genome-wide files, protein sequence files, and GFF annotation files. The genomic files of *A. thaliana* and tomato were download from EnsembPlants (https://ftp.Ensemblgenomes.ebi.ac.uk/pub/plants/release-59/, accessed on 16 May 2024). The protein sequence files of kenaf, *A. thaliana*, and tomato were obtained by TBtools-II (Toolbox for Biologists) v2.119 software [42,43]. The Hidden Markov Model (HMM) profile of the Nramp domain (PF01566) was obtained from the Pfam database (https://pfam.xfam.org/, accessed on 10 May 2024) [20], and we used TBtools software [42,43] to localize Blast to obtain the protein sequences of members of the *Nramp* gene family. We obtain the candidate proteins of kenaf, *A. thaliana*, and tomato by Hmm search with an E value of <0.001, and named them according to the chromosome structures. The Nramp protein sequence of kenaf was submitted to the SMART (http://smart.embl.de/, accessed on 17 July 2024) database as a search query sequence to verify the existence of the Nramp domain [44].

The candidate protein sequence files of kenaf were submitted to NCBI Batch CDD (https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi, accessed on 31 May 2024) to obtain the conservation domain. We used the protein molecular weight analysis website ExPASy (http://web.expasy.org/protparam/, accessed on 1 June 2024) to analyze the physical and chemical properties of kenaf Nramp gene family proteins [24]. We used the WoLF PSORT website (https://wolfpsort.hgc.jp/, accessed on 31 May 2024) to determine the subcellular location of the family members [44]. Through the online website SOPMA (https:npsa-prabi.ibcp.fr/cgi-bin/npsa_automat.pl?Page=/NPSA/npsa_sopma.html, accessed on 10 May 2024), we predicted the secondary structures of the family proteins.

4.2. Sequence Alignment, Phylogenetics, Gene Structure, and Conserved Motif Analysis of Nramp Genes

To study the phylogenetic relationships of kenaf, *A. thaliana*, and tomato, the protein full-length sequences of plants were compared using the MUSCLE program provided by MEGA11.0 software [45], using the neighbor-joining (NJ) method with 1000 bootstrap replicates to build a multi-species phylogenetic tree [24,46,47]. The phylogenetic tree used Evolview v3 online (https://www.evolgenius.info/evolview/#/, accessed on 27 June 2024) software for visual analysis and beautification [24,36].

The gene structure was analyzed by Tbtools software. Conserved domains and motifs of the HcNramp proteins were analyzed with the Pfam 37.0 tool (http://pfam.xfam.org/,accessed on 17 July 2024) and the MEME Suite 5.5.5 (https://meme-suite.org/me/me/tools/meme, accessed on 14 May 2024), respectively. The relevant parameters were set as a maximum base ordinal number of 20, and the others used the default parameters. TBtools software [42,43] was used to visualize the phylogenetic trees and conserved motifs.

4.3. Cis Element Analysis of HcNramp Genes

The ATG sequence of the genomic start codon was extracted from the *HcNramp* gene family as the promoter sequence, and then we used the online software of PlantCare (http://bioinformatics.psb.ugent.be/webtools/plantcare/html/, accessed on 30 May 2024) to analyze the promoter sequence [24,47]. The target cis-acting elements were screened in Excel, and then drawn with TBtools software [42,43].

4.4. Chromosome Localization and Collinearity Analyses of the Kenaf Nramp Gene Family

We extracted the position information of the *HcNramp* sequence from the kenaf GFF file and used TBtools software [42,43] to map the position distribution on the kenaf chromosome.

To explore the collinearity of *Nramp* genes in kenaf, *A. thaliana*, and tomato, whole-genome information was analyzed using TBtools software [42,43,48].

4.5. Plant Materials and Treatments

The kenaf "H368" variety was used as the test material, provided by the Institute of Bast Fiber Crops, Chinese Academy of Agricultural Sciences. Seeds were surface-sterilized with 75% alcohol for 2 min, washed with deionized water three times, germinated for one week, then transplanted into Hoagland nutrient solution for another week. Following this, they were cultured in 0/30 μmol/L of $CdCl_2$ for two weeks. Subsequently, the plants were sprayed with 100 μmol/L of melatonin solution once every two days for a total of three times over a period of one week. The culture conditions were 25 °C/20 °C for 16 h/8 h (light/dark), relative humidity of 65%, and illumination of 12,000 lx. Finally, the roots, stems, and leaves (the third leaf from the top) of kenaf seedlings were sampled with liquid nitrogen and put in a refrigerator at −80 °C for further experiments. There were four treatments: control treatment (CK + water, CK-H), melatonin treatment (CK + melatonin, CK-M), Cd treatment (30 μmol/L $CdCl_2$ + water, Cd-H), and Cd + melatonin treatment (30 μmol/L $CdCl_2$ + melatonin, Cd-M). Nutrients were renewed every three days. There were three replicates for each treatment.

4.6. RNA Extraction and qRT-PCR Analysis

The total RNA was extracted from the root, stem, and leaf samples of kenaf by using the EASYspin Plus plant RNA rapid extraction kit. The NanoReadyF-1100 was used to determine the quality and concentration of RNA samples. Then, the first-strand cDNA was synthesized by using 5 × Prime Script RT Master Mix (TaKaRa). Taking the *Histone3* of kenaf as the reference gene [49], the primer sequence of *HcNramp* was designed by using TBtools and listed in Table S8. The qRT-PCR reaction was detected in the Step OnePlus Real-Time PCR System. The total reaction volume of PCR was 10 μL, including 1 μL of diluted cDNA, 5 μL of TB Green Premix Ex Taq (Tli RNaseH Plus) (2×) reagent, 0.2 μL of ROX Reference Dye, 0.2 μL of forward primer, 0.2 μL of reverse primer, and 3.4 μL of RNA-free H_2O. The amplification conditions were as follows: in the first stage, 95 °C acted for 30 s; in the second stage, 95 °C for 5 s and 60 °C for 30 min, with 40 cycles; in the third stage, 95 °C for 15 s and 60 °C for 1 min. The expression of *HcNramp* genes was calculated using the $2^{-\triangle\triangle ct}$ method. Three replicates were set for each reaction.

4.7. Statistical Analyses

GraphPad Prism was used for statistical analyses. Error bars indicate mean ± SD and asterisks indicate statistical differences (n = 3; * $p < 0.05$, ** $p < 0.01$, and *** $p < 0.001$; ns, not significant; Student's *t*-test) [50].

5. Conclusions

A comprehensive analysis of the *Nramp* gene family in kenaf was carried out in this study. Fifteen *HcNramp* genes were identified and further classified into three main phylogenetic subfamilies. A phylogenetic comparison of *Nramp* genes among kenaf, *A. thaliana*, and tomato was constructed, which provided some clues for understanding the evolution of *HcNramp* genes. All proteins of the HcNramp family were located in the plasma membrane, and they had highly similar exon–intron structures, motifs, and promoter elements in the same subfamilies. The distribution of chromosomes was not uniform, and the expression levels of different tissues in different treatments were also different. Most *HcNramp* genes were preferentially expressed in roots. These findings provide necessary information for understanding the potential role of the *HcNramp* genes in Cd stress and are also of great significance for screening Cd-tolerant genes in kenaf.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/plants13172514/s1. Table S1: 15 protein sequences of HcNramp identified from kenaf. Table S2: Physical and chemical properties of *HcNramp* gene family members. Table S3: Secondary structure of HcNramp family protein. Table S4: List of 29 Nramp protein sequences. Table S5: List of the putative motifs of HcNramp proteins. Table S6: Collinearity information. Table S7: Transcriptome data of 15 *HcNramp* genes. Table S8: Primer sequences for ten target genes.

Author Contributions: Conceptualization, Q.L. and S.L.; methodology, Q.L. and S.L.; software, Q.L. and S.L.; validation, X.A. and G.D.; formal analysis, X.A. and G.D.; investigation, Q.L. and S.L.; resources, Q.L. and S.L.; data curation, Q.L. and S.L.; writing—original draft preparation, Q.L.; writing—review and editing, Q.L. and S.L.; visualization, X.A. and G.D.; supervision, X.A. and G.D.; project administration, X.A. and G.D.; funding acquisition, X.A. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the China Agriculture Research System of MOF and MARA, China Agriculture Research System for Bast and Leaf Fiber Crops (CARS-16-S05).

Data Availability Statement: All datasets generated for this study are included in the article/Supplementary Materials.

Conflicts 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.

References

1. Chen, P.; Li, Z.; Luo, D.; Jia, R.; Lu, H.; Tang, M.; Hu, Y.; Yue, J.; Huang, Z. Comparative transcriptomic analysis reveals key genes and pathways in two different cadmium tolerance kenaf (*Hibiscus cannabinus* L.) cultivars. *Chemosphere* **2021**, *263*, 128211. [CrossRef] [PubMed]
2. Gong, Z.; Duan, Y.; Liu, D.; Zong, Y.; Zhang, D.; Shi, X.; Hao, X.; Li, P. Physiological and transcriptome analysis of response of soybean (*Glycine max*) to cadmium stress under elevated CO_2 concentration. *J. Hazard. Mater.* **2023**, *448*, 130950. [CrossRef]
3. Meng, J.G.; Zhang, X.D.; Tan, S.K.; Zhao, K.X.; Yang, Z.M. Genome-wide identification of Cd-responsive *NRAMP* transporter genes and analyzing expression of *NRAMP 1* mediated by miR167 in *Brassica napus*. *BioMetals* **2017**, *30*, 917–931. [CrossRef]
4. Zhang, X.; Yang, M.; Yang, H.; Pian, R.; Wang, J.; Wu, A.-M. The Uptake, Transfer, and Detoxification of Cadmium in Plants and Its Exogenous Effects. *Cells* **2024**, *13*, 907. [CrossRef]
5. Tan, Z.; Li, J.; Guan, J.; Wang, C.; Zhang, Z.; Shi, G. Genome-Wide Identification and Expression Analysis Reveals Roles of the *NRAMP* Gene Family in Iron/Cadmium Interactions in Peanut. *Int. J. Mol. Sci.* **2023**, *24*, 1713. [CrossRef]
6. Chen, P.; Chen, T.; Li, Z.; Jia, R.; Luo, D.; Tang, M.; Lu, H.; Hu, Y.; Yue, J.; Huang, Z. Transcriptome analysis revealed key genes and pathways related to cadmium-stress tolerance in Kenaf (*Hibiscus cannabinus* L.). *Ind. Crops Prod.* **2020**, *158*, 112970. [CrossRef]
7. Li, H.; Yang, X.; Niyitanga, S.; He, Q.; Chen, S.; Xu, J.; Qi, J.; Tao, A.; Fang, P.; Zhang, L. Transcriptomes of Different Tissues for Expression Characteristics Analysis of *MYB* gene Family in Kenaf (*Hibiscus cannabinus* L.). *Trop. Plant Biol.* **2022**, *15*, 261–275. [CrossRef]
8. An, X.; Jin, G.; Zhang, J.; Ma, G.; Jin, L.; Luo, X.; Chen, C.; Shi, X.; Zhou, J.; Wei, W.; et al. Research Progress on Tissue Culture and Genetic Transformation of Kenaf (*Hibiscus cannabinus*). *Open Life Sci.* **2017**, *12*, 465–472. [CrossRef]
9. Chen, M.; She, Z.; Aslam, M.; Liu, T.; Wang, Z.; Qi, J.; Niu, X. Genomic insights of the *WRKY* genes in kenaf (*Hibiscus cannabinus* L.) reveal that *HcWRKY44* improves the plant's tolerance to the salinity stress. *Front. Plant Sci.* **2022**, *13*, 984233. [CrossRef]
10. An, X.; Chen, J.; Liu, T.; Li, W.; Luo, X.; Zou, L. Transcriptomic and Metabolic Profiling of Kenaf Stems under Salinity Stress. *Plants* **2022**, *11*, 1448. [CrossRef]
11. Niu, X.; Chen, M.; She, Z.; Aslam, M.; Qi, J.; Qin, Y. Ectopic Expression of Kenaf (*Hibiscus cannabinus* L.) HcWRKY50 Improves Plants' Tolerance to Drought Stress and Regulates ABA Signaling in *Arabidopsis*. *Agronomy* **2022**, *12*, 1176. [CrossRef]
12. Deng, Y.; Li, D.; Huang, Y.; Huang, S. Physiological response to cadmium stress in kenaf (*Hibiscus cannabinus* L.) seedlings. *Ind. Crops Prod.* **2017**, *107*, 453–457. [CrossRef]
13. Li, W.; Chen, C.; Deng, Y.; Luo, X.; Liu, T.; An, X.; Zou, L.; Luan, M.; Li, D. Influence of Nitrogen Supply on Growth, Antioxidant Capacity and Cadmium Absorption of Kenaf (*Hibiscus cannabinus* L.) Seedlings. *Plants* **2023**, *12*, 4067. [CrossRef] [PubMed]
14. Liu, S.; Long, T.; Chen, Z.; Liu, J.; Cui, W.; Leng, H.; Xing, Y.; Rodriguez, L.G.; Gao, Y.; Yao, Y. Genome-wide identification of *NRAMP* family genes in *Populus trichocarpa* and their roles in transport of heavy metals. *Tree Genet. Genomes* **2023**, *19*, 51. [CrossRef]
15. Segond, D.; Dellagi, A.; Lanquar, V.; Rigault, M.; Patrit, O.; Thomine, S.; Expert, D. *NRAMP* genes function in *Arabidopsis thaliana* resistance to Erwinia chrysanthemi infection. *Plant J.* **2009**, *58*, 195–207. [CrossRef]
16. Wei, W.; Chai, T.; Zhang, Y.; Han, L.; Xu, J.; Guan, Z. The *Thlaspi caerulescens NRAMP* Homologue *TcNRAMP3* is Capable of Divalent Cation Transport. *Mol. Biotechnol.* **2008**, *41*, 15–21. [CrossRef]

17. Cailliatte, R.; Schikora, A.; Briat, J.-F.; Mari, S.; Curie, C. High-Affinity Manganese Uptake by the Metal Transporter *NRAMP1* Is Essential for *Arabidopsis* Growth in Low Manganese Conditions. *Plant Cell* **2010**, *22*, 904–917. [CrossRef] [PubMed]
18. Chen, K.; Jiang, X.; Zhu, Z.; Wang, H.; Feng, S. Advances in the study of plant *Nramp* family involved in metal ion absorption and distribution. *Plant Physiol. J.* **2020**, *56*, 345–355. [CrossRef]
19. Chen, Y.; Zhao, X.; Li, G.; Kumar, S.; Sun, Z.; Li, Y.; Guo, W.; Yang, J.; Hou, H. Genome-Wide Identification of the *Nramp* Gene Family in *Spirodela polyrhiza* and Expression Analysis under Cadmium Stress. *Int. J. Mol. Sci.* **2021**, *22*, 6414. [CrossRef]
20. Qin, L.; Han, P.; Chen, L.; Walk, T.C.; Li, Y.; Hu, X.; Xie, L.; Liao, H.; Liao, X. Genome-Wide Identification and Expression Analysis of *NRAMP* Family Genes in Soybean (*Glycine max* L.). *Front. Plant Sci.* **2017**, *8*, 1436. [CrossRef]
21. Tian, W.; He, G.; Qin, L.; Li, D.; Meng, L.; Huang, Y.; He, T. Genome-wide analysis of the *NRAMP* gene family in potato (*Solanum tuberosum*): Identification, expression analysis and response to five heavy metals stress. *Ecotoxicol. Environ. Saf.* **2021**, *208*, 111661. [CrossRef] [PubMed]
22. Tiwari, M.; Sharma, D.; Dwivedi, S.; Singh, M.; Tripathi, R.D.; Trivedi, P.K. Expression in *Arabidopsis* and cellular localization reveal involvement of rice *NRAMP*, *OsNRAMP1*, in arsenic transport and tolerance. *Plant Cell Environ.* **2013**, *37*, 140–152. [CrossRef] [PubMed]
23. Zhao, Z.; Yuan, M.; Wu, J.; Men, Y.; Jiang, M.; Gu, L.; Wang, J. Genome-wide Identification and Bioinformatics Analysis of Poplar *NRAMP* Gene Family. *J. Anhui Agric. Sci.* **2023**, *51*, 90–94+141. [CrossRef]
24. Yan, L.; Jin, H.; Raza, A.; Huang, Y.; Gu, D.p.; Zou, X. Natural resistance-associated macrophage proteins (NRAMPs) are involved in cadmium enrichment in peanut (*Arachis hypogaea* L.) under cadmium stress. *Plant Growth Regul.* **2023**, *102*, 619–632. [CrossRef]
25. Li, J.; Li, H.; Liu, L.-E.; Zhang, T.; Zhou, W. Identification and Analysis of the *NRAMP* Family in Seabuckthorn Under Lead Stress. *Biotechnol. Bull.* **2024**, *40*, 191–202. [CrossRef]
26. Kumar, A.; Singh, G.; Prasad, B.; Sharma, P.; Kumar, R. Genome wide analysis and identification of NRAMP gene family in wheat (Triticum aestivum L.). *Pharma Innov. J.* **2022**, *SP-11*, 499–504.
27. Zhao, Y.; Xie, Q.; Yang, Q.; Cui, J.; Tan, W.; Zhang, D.; Xiang, J.; Deng, L.; Guo, Y.; Li, M.; et al. Genome-wide identification and evolutionary analysis of the *NRAMP* gene family in the AC genomes of Brassica species. *BMC Plant Biol.* **2024**, *24*, 311. [CrossRef]
28. Thomine, S.B.; Wang, R.W.; John, M.; Crawford, N.M.; Schroeder, J.I. Cadmium and iron transport by members of a plant metal transporter family in *Arabidopsis* with homology to *Nramp* genes. *Proc. Natl. Acad. Sci. USA* **2000**, *97*, 4991–4996. [CrossRef]
29. Mani, A.; Sankaranarayanan, K. In Silico Analysis of Natural Resistance-Associated Macrophage Protein (NRAMP) Family of Transporters in Rice. *Protein J.* **2018**, *37*, 237–247. [CrossRef]
30. Cailliatte, R.; Lapeyre, B.; Briat, J.-F.; Mari, S.; Curie, C. The *NRAMP6* metal transporter contributes to cadmium toxicity. *Biochem. J.* **2009**, *422*, 217–228. [CrossRef]
31. Zhao, S.; Zhang, Y.; Zhang, Q.; Gangjun, W.; Ye, X. Differential responses of two tomato cultivars to cadmium stress. *J. Plant Nutr. Fertil.* **2015**, *21*, 1261–1268. [CrossRef]
32. Wei, T.; Sun, Y.; Yashir, N.; Li, X.; Guo, J.; Liu, X.; Jia, H.; Ren, X.; Hua, L. Inoculation with Rhizobacteria Enhanced Tolerance of Tomato (*Solanum lycopersicum* L.) Plants in Response to Cadmium Stress. *J. Plant Growth Regul.* **2021**, *41*, 445–460. [CrossRef]
33. Zhang, J.; Zhang, M.; Song, H.; Zhao, J.; Shabala, S.; Tian, S.; Yang, X. A novel plasma membrane-based *NRAMP* transporter contributes to Cd and Zn hyperaccumulation in Sedum alfredii Hance. *Environ. Exp. Bot.* **2020**, *176*, 104121. [CrossRef]
34. Wang, C.; Chen, X.; Yao, Q.; Long, D.; Fan, X.; Kang, H.; Zeng, J.; Sha, L.; Zhang, H.; Zhou, Y.; et al. Overexpression of *TtNRAMP6* enhances the accumulation of Cd in *Arabidopsis*. *Gene* **2019**, *696*, 225–232. [CrossRef] [PubMed]
35. Hussain, Q.; Ye, T.; Shang, C.; Li, S.; Khan, A.; Nkoh, J.N.; Mustafa, A.E.-Z.M.A.; Elshikh, M.S. *NRAMP* gene family in *Kandelia obovata*: Genome-wide identification, expression analysis, and response to five different copper stress conditions. *Front. Plant Sci.* **2024**, *14*, 1318383. [CrossRef] [PubMed]
36. He, Z.; Zhang, H.; Gao, S.; Lercher, M.J.; Chen, W.-H.; Hu, S. Evolview v2: An online visualization and management tool for customized and annotated phylogenetic trees. *Nucleic Acids Res.* **2016**, *44*, W236–W241. [CrossRef]
37. Guo, B.; Liu, C.; Li, H.; Yi, K.; Ding, N.; Li, N.; Lin, Y.; Fu, Q. Endogenous salicylic acid is required for promoting cadmium tolerance of *Arabidopsis* by modulating glutathione metabolisms. *J. Hazard. Mater.* **2016**, *316*, 77–86. [CrossRef]
38. Lei, G.J.; Sun, L.; Sun, Y.; Zhu, X.F.; Li, G.X.; Zheng, S.J. Jasmonic acid alleviates cadmium toxicity in *Arabidopsis* via suppression of cadmium uptake and translocation. *J. Integr. Plant Biol.* **2019**, *62*, 218–227. [CrossRef]
39. Wang, F.; Tan, H.; Huang, L.; Cai, C.; Ding, Y.; Bao, H.; Chen, Z.; Zhu, C. Application of exogenous salicylic acid reduces Cd toxicity and Cd accumulation in rice. *Ecotoxicol. Environ. Saf.* **2021**, *207*, 111198. [CrossRef]
40. Oomen, R.J.F.J.; Wu, J.; Lelièvre, F.; Blanchet, S.; Richaud, P.; Barbier-Brygoo, H.; Aarts, M.G.M.; Thomine, S. Functional characterization of *NRAMP3* and *NRAMP4* from the metal hyperaccumulator *Thlaspi caerulescens*. *New Phytol.* **2008**, *181*, 637–650. [CrossRef]
41. Wu, Q.; Chen, C.; Yue, J.; Mubeen, S.; Cao, S.; Li, X.; Wang, M.; Zhang, H.; Wu, X.; Wang, C.; et al. Genome-wide identification of *CUC* gene family and functional analysis of *HcCUC1* in kenaf. *Plant Cell Tissue Organ Cult.* **2023**, *155*, 91–102. [CrossRef]
42. Chen, C.; Chen, H.; Zhang, Y.; Thomas, H.R.; Frank, M.H.; He, Y.; Xia, R. TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Mol. Plant* **2020**, *13*, 1194–1202. [CrossRef] [PubMed]
43. Chen, C.; Wu, Y.; Li, J.; Wang, X.; Zeng, Z.; Xu, J.; Liu, Y.; Feng, J.; Chen, H.; He, Y.; et al. TBtools-II: A "one for all, all for one" bioinformatics platform for biological big-data mining. *Mol. Plant* **2023**, *16*, 1733–1742. [CrossRef] [PubMed]

44. Kasapoğlu, A.G.; Muslu, S.; Aygören, A.S.; Öner, B.M.; Güneş, E.; İlhan, E.; Yiğider, E.; Aydin, M. Genome-wide characterization of the *GPAT* gene family in bean (*Phaseolus vulgaris* L.) and expression analysis under abiotic stress and melatonin. *Genet. Resour. Crop Evol.* **2024**, *25*, 6101. [CrossRef]
45. Kumar, S.; Stecher, G.; Li, M.; Knyaz, C.; Tamura, K.; Battistuzzi, F.U. MEGA X: Molecular Evolutionary Genetics Analysis across Computing Platforms. *Mol. Biol. Evol.* **2018**, *35*, 1547–1549. [CrossRef]
46. Liu, J.; Wang, X.; Chen, Y.; Liu, Y.; Wu, Y.; Ren, S.; Li, L. Identification, evolution and expression analysis of *WRKY* gene family in *Eucommia ulmoides*. *Genomics* **2021**, *113*, 3294–3309. [CrossRef]
47. Yao, Y.; He, Z.; Li, X.; Xu, J.; Han, X.; Liang, H.; Zhuo, R.; Qiu, W. Genome-wide identification of *bHLH* gene family and its response to cadmium stress in *Populus* × *canescens*. *PeerJ* **2024**, *12*, e17410. [CrossRef]
48. Yang, W.; Zhao, L.; Tang, B.; Fu, W.; Wang, Q.; Mo, C.; Zhang, Y.; Ao, N.; Li, Y.; Li, F.; et al. Genome-wide identification and expression analysis of the *SBP-box* gene family in *Brassica juncea* L. *J. Northeast. Agric. Univ.* **2023**, *54*, 23. [CrossRef]
49. Jia, R. Comparative Morphological and Physiological Responses and Differential Gene Discovery in Kenaf (*Hibiscus cannabinus* L.) under Cadmium Stress. Master's Thesis, Guangxi University, Nanning, China, 2019.
50. Singh, R.; Jha, A.B.; Misra, A.N.; Sharma, P. Differential responses of growth, photosynthesis, oxidative stress, metals accumulation and *NRAMP* genes in contrasting Ricinus communis genotypes under arsenic stress. *Environ. Sci. Pollut. Res.* **2019**, *26*, 31166–31177. [CrossRef]

plants

MDPI

Article

Phenotypic Analysis and Gene Cloning of Rice Floury Endosperm Mutant *wcr* (White-Core Rice)

Yihao Yang [1,2], Xiaoyi Yang [1], Lingjun Wu [1], Zixing Sun [1], Yi Zhang [1], Ziyan Shen [1], Juan Zhou [1,2], Min Guo [1,2] and Changjie Yan [1,2,*]

[1] Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding/Zhongshan Biological Breeding Laboratory/Key Laboratory of Plant Functional Genomics of the Ministry of Education, Agricultural College, Yangzhou University, Yangzhou 225009, China
[2] Jiangsu Co-Innovation Center for Modern Production Technology of Grain Crops/Jiangsu Key Laboratory of Crop Genetics and Physiology, Yangzhou University, Yangzhou 225009, China
* Correspondence: cjyan@yzu.edu.cn

Abstract: The composition and distribution of storage substances in rice endosperm directly affect grain quality. A floury endosperm mutant, *wcr* (white-core rice), was identified, exhibiting a loose arrangement of starch granules with a floury opaque appearance in the inner layer of mature grains, resulting in reduced grain weight. The total starch and amylose content remained unchanged, but the levels of the four component proteins in the mutant brown rice significantly decreased. Additionally, the milled rice (inner endosperm) showed a significant decrease in total starch and amylose content, accompanied by a nearly threefold increase in albumin content. The swelling capacity of mutant starch was reduced, and its chain length distribution was altered. The target gene was mapped on chromosome 5 within a 65 kb region. A frameshift mutation occurred due to an insertion of an extra C base in the second exon of the *cyOsPPDKB* gene, which encodes pyruvate phosphate dikinase. Expression analysis revealed that *wcr* not only affected genes involved in starch metabolism but also downregulated expression levels of genes associated with storage protein synthesis. Overall, *wcr* plays a crucial role as a regulator factor influencing protein synthesis and starch metabolism in rice grains.

Keywords: grain quality; storage substances; *OsPPDKB*

Citation: Yang, Y.; Yang, X.; Wu, L.; Sun, Z.; Zhang, Y.; Shen, Z.; Zhou, J.; Guo, M.; Yan, C. Phenotypic Analysis and Gene Cloning of Rice Floury Endosperm Mutant *wcr* (White-Core Rice). *Plants* **2024**, *13*, 2653. https://doi.org/10.3390/plants13182653

Academic Editor: Naoki Hirotsu

Received: 18 August 2024
Revised: 18 September 2024
Accepted: 21 September 2024
Published: 22 September 2024

1. Introduction

Rice (*Oryza sativa* L.) is a crucial staple crop worldwide, and the advancements in dwarf breeding during the 1960s, as well as the theoretical and technical breakthroughs in hybrid rice breeding during the 1970s, have significantly improved rice production in China, effectively addressing food scarcity issues [1]. With an improvement in living standards, there has been an increasing demand among consumers for higher rice grain quality. Starch and protein are the primary storage substances within the endosperm of rice grains. Starch constitutes a majority portion that determines both grain quality and yield, followed by protein content which influences rice palatability and nutritional value [2,3]. Therefore, it is of immense significance to comprehensively analyze regulatory mechanisms governing starch and protein synthesis as well as accumulation while exploring associated genes to enhance rice grain quality.

The rice floury endosperm mutants are valuable genetic resources for studying the complex networks involved in endosperm development and quality regulation due to their abnormal synthesis and accumulation of starch or storage proteins, resulting in floury and opaque endosperms. Currently, a multitude of genes associated with floury endosperm mutants have been successfully cloned, encompassing diverse aspects of cellular metabolic processes including starch synthesis and sucrose metabolism, amyloplastic development, energy supply, glycolytic metabolism, protein processing and transport, lipid transport,

epigenetics, transcriptional regulation, and protein interaction (Table 1). The cloning and molecular mechanism analysis of the aforementioned genes provide a theoretical foundation for understanding starch and protein synthesis, as well as the biochemical metabolic pathways in rice. This enhances our knowledge of endosperm development regulation and offers the potential for improving rice quality through genetic engineering.

Table 1. Reported floury endosperm genes.

Classification	Name	Gene ID	Annotation
Starch synthesis and sucrose metabolism	*FLO8*	*Os09g0553200*	UTP-glucose-1-phosphate uridylyltransferase [4]
	OsAGPL2	*Os01g0633100*	Glucose-1-phosphate adenylyltransferase large subunit [5]
	OsAGPS	*Os09g0298200*	Glucose-1-phosphate adenylyltransferase large subunit [5]
	OsBT1	*Os02g0202400*	ADP-Glucose Transporter [6]
	OsBEIIb	*Os02g0528200*	1,4-alpha-glucan-branching enzyme [7]
	FLO5	*Os08g0191433*	Starch synthase III [8]
	Pho1	*Os03g0758100*	Alpha-glucan phosphorylase isozyme [9]
	GIF1	*Os04g0413500*	Cell wall invertase [10]
	Wx	*Os06g0133000*	Starch synthase [11]
	FLO24	*Os03g0426900*	Heat shock protein 101 [12]
Amyloplast development	*SSG4*	*Os01g0179400*	Protein containing a DUF490 domain [13]
	SSG6	*Os06g0130400*	Aminotransferase [14]
	SSG7	*Os11g0524300*	Plant-specific DUF1001 domain-containing protein [15]
	FSE1	*Os08g0110700*	Phospholipase-like protein [16]
Energy supply	*FLO13*	*Os02g0816800*	Mitochondrial complex I subunit [17]
	OGR1	*Os12g0270200*	Pentatricopeptide repeat–DYW protein [18]
	FLO10	*Os03g0168400*	Pentatricopeptide repeat protein [19]
	OsNPPR1	*Os08g0290000*	Pentatricopeptide repeat protein [20]
	FLO18	*Os07g0688100*	Pentatricopeptide repeat protein [21]
	FLO22	*Os07g0179000*	P-type pentatricopeptide repeat (PPR) protein [22]
Glycolytic metabolism	*PFPβ*	*Os06g0247500*	Pyrophosphate-fructose 6-phosphate 1-phosphotransferase [23]
	FLO4	*Os05g0405000*	Pyruvate, phosphate dikinase [24]
	OsPK2	*Os07g0181000*	Plastidic pyruvate kinase [25]
	FLO12	*Os10g0390500*	Aminotransferase [26]
	FLO15	*Os05g0230900*	Glyoxalase family protein [27]
	FLO16	*Os10g0478200*	Lactate/malate dehydrogenase [28]
	FLO23	*Os03g0294200*	Fructose-6-phosphate-2-kinase/fructose-2, 6-bisphosphatase [29]
	FLO19	*Os04g0119400*	Pyruvate dehydrogenase complex E1 component subunit α1 [30]
	FLO19	*Os03g0685300*	Class I glutamine amidotransferase [31]
Protein processing and transport	*PDIL1-1*	*Os11g0199200*	Protein disulfide isomerase-like enzyme [32]
	GPA1	*Os12g0631100*	Small GTPase [33]
	GPA2	*Os03g0262900*	Guanine nucleotide exchange factor [34]
	GPA3	*Os03g0835800*	Regulator of post-Golgi vesicular traffic [35]
	GPA4	*Os03g0209400*	Golgi Transport 1 [36]
	GPA5	*Os06g0643000*	Rab5a Effector [37]
	GPA6	*Os09g0286400*	Vacuolar Na^+/H^+ antiporter [38]
	GPA7	*Os08g0427300*	Homolog of Arabidopsis CCZ1a and CCZ1b [39]
	GPA8	*Os01g0659200*	Subunit E isoform 1 of vacuolar H^+-ATPase [40]
Lipid transport	*ESG1*	*Os04g0553000*	Bacterial MlaE lipid transfer protein [41]
Epigenetics	*OsROS1*	*Os01g0218032*	DNA demethylase [42]
	FLO20-1	*Os01g0874900*	Serine hydroxymethyltransferase [43]
Transcriptional regulation and protein interaction	*RISBZ1*	*Os07g0182000*	bZIP transcription factor [44]
	RSR1	*Os05g0121600*	Transcription factor of the AP2/EREBP family [45]
	NF-YB1	*Os02g0725900*	Nuclear transcription factor Y subunit B [46]
	NF-YC12	*Os05g0304800*	CCAAT-box-binding transcription factor [46,47]
	bHLH144	*Os04g0429400*	Helix-loop-helix DNA-binding domain containing protein [46]
	FLO2	*Os04g0645100*	Tetratricopeptide repeat domain-containing protein [48]
	FLO6	*Os03g0686900*	CBM48 domain-containing protein [49]
	FLO7	*Os10g0463800*	DUF1388 domain protein [50]
	FLO11	*Os12g0244100*	Plastid heat shock protein 70 [51]
	FLO9	*Os11g0586300*	Homologous to Arabidopsis LIKE EARLY STARVATION1 [52]

The synthesis and transportation of storage substances in cereal grain endosperm require sufficient energy supply [53]. However, the low oxygen environment within the endosperm hampers oxidative phosphorylation, thereby limiting ATP production [54,55]. The glycolytic pathway serves as a crucial means for biological organisms to generate energy under anaerobic conditions, with pyruvate kinase (PK) and pyruvate phosphate dikinase (PPDK) catalyzing this reaction [56,57]. In plants, PPDK can be classified into chloroplast chPPDK and cytosolic cyPPDK. ChPPDK is primarily found in chloroplasts and expressed predominantly in photosynthetic tissues such as leaves of C4 plants. Conversely, cyPPDK is mainly localized within the cytoplasm and expressed in non-photosynthetic tissues like grains and roots [58]. In rice, there are only two genes encoding PPDK: one encodes

OsPPDKA while another produces two transcripts of C4-type chloroplastic chOsPPDKB and cytosolic cyOsPPDKB [24]. CyOsPPDKB is responsible for the reversible conversion of pyruvate and inorganic phosphate (Pi) into phosphoenolpyruvate (PEP) and inorganic pyrophosphate (PPi), which can serve as an alternative phosphate donor for ATP in plant cells [24,59–61]. This process provides essential carbon skeletons for amino acid and lipid biosynthesis, directly influencing ADP-glucose allocation towards starch and lipids [24,62–68].

In this study, we identified a white-core endosperm mutant (*wcr*) derived from the rice variety Sasanishiki through tissue culture. We present a comprehensive investigation of the starch morphology, physicochemical properties, and protein composition content of the grains. Additionally, we precisely mapped the mutant gene and explored its involvement in regulating the network that governs rice grain starch and storage protein content. This research contributes valuable genetic resources and provides a theoretical foundation for unraveling mechanisms underlying rice endosperm development and grain quality regulation.

2. Results

2.1. Analysis of Phenotypic and Crop Traits of the wcr Mutant

The *wcr* mutant was derived from the *Japonica* rice variety Sasanishiki through plant tissue culture. Overall, there were minimal changes in the plant architecture of the *wcr* mutant (Figure 1A). Upon hull removal, the mature grain of the *wcr* mutant exhibits a starchy and translucent state, with opaqueness primarily localized in the inner endosperm (Figure 1B,C). The starch granules in the endosperm of wild-type exhibited a tightly and evenly arranged pattern in both inner and outer layers under electron microscope scanning (Figure 1D). In contrast, the starch granules in *wcr*'s endosperm showed a loosely arranged structure with larger inter-granular gaps and an increased presence of single-grain type starch granules within the inner layer (Figure 1D). The mutant showed significantly increased plant height, grain length, and width compared to the wild type (Figure 1A,E,G,H), while exhibiting a significant decrease in tiller number, grain thickness, and thousand-grain weight (Figure 1F,I,J).

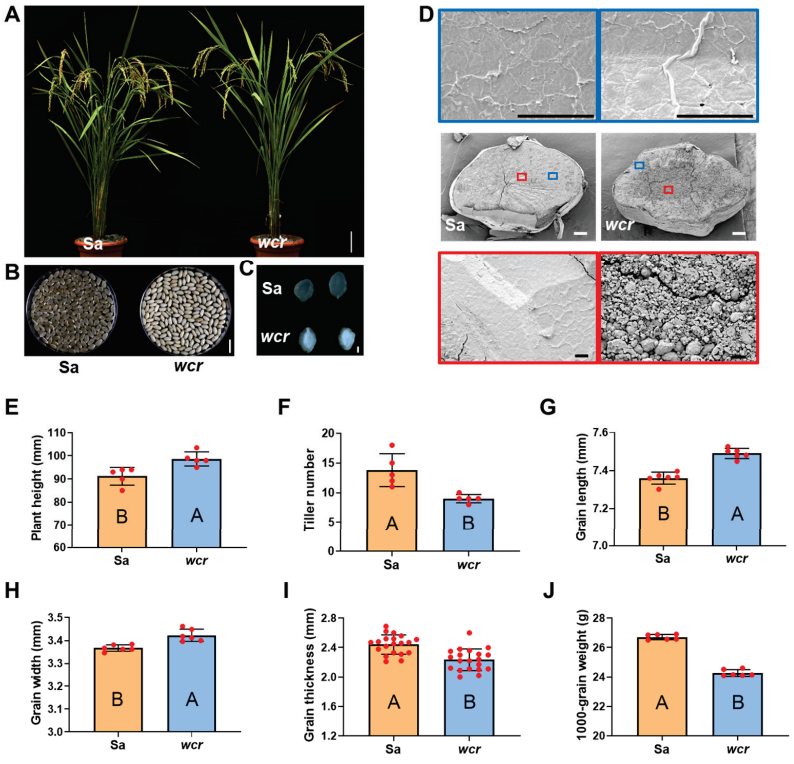

Figure 1. Phenotypic and crop traits analysis of the *wcr* mutant and wild type. (**A**) The whole plants of the Sa-Sasanishiki wild-type and *wcr*, scale bar = 10 cm. (**B**) The brown rice of the Sa and *wcr*,

scale bar = 10 mm. (**C**) The cross-section of the Sa and *wcr* grains, scale bar = 1 mm. (**D**) Electron microscopic scanning of the endosperm of Sa and *wcr*, black scale bar = 10 μm, white scale bar = 200 μm. Different colored boxes represent local enlargements of starch structures. (**E**–**J**) The crop traits of Sa and *wcr*. Different upper case letters denote significant statistical differences between Sa (orange) and *wcr* (blue) plants, with the *p*-value < 0.01.

2.2. Physicochemical Characteristics of Mature Grains of wcr Mutant

The physicochemical properties of brown rice and milled rice from the *wcr* mutant were further investigated in comparison to those of the wild type, due to its abnormal endosperm development and significant differences observed in starch granule structure between the inner and outer layers (Figure 1B–D). No significant differences were observed in the total starch content and amylose content of brown rice flour between *wcr* and wild type (Figure 2A,B). However, both traits showed a significant decrease when analyzing milled rice flour (Figure 2A,B). Based on these results, the total starch content and amylose content in the peripheral endosperm of *wcr* significantly increased compared to the wild type (Figure 2H). The analysis of storage proteins showed a significant decrease in all four component proteins (albumin, globulin, prolamin, and glutelin) in mutant brown rice flour (Figure 2C–F). However, apart from a nearly threefold increase in the content of albumin compared to the wild type (Figure 2C), there is no difference in the protein content of the other three components when analyzed in milled rice flour (Figure 2D–F). These changes were also reflected in the SDS-PAGE analysis of storage proteins (Figure 2G). It can be inferred that unlike the changing trends of total starch content and amylose content in the *wcr* peripheral endosperm, protein contents for all four components significantly decrease compared to the wild-type with albumin showing the largest decrease (Figure 2H).

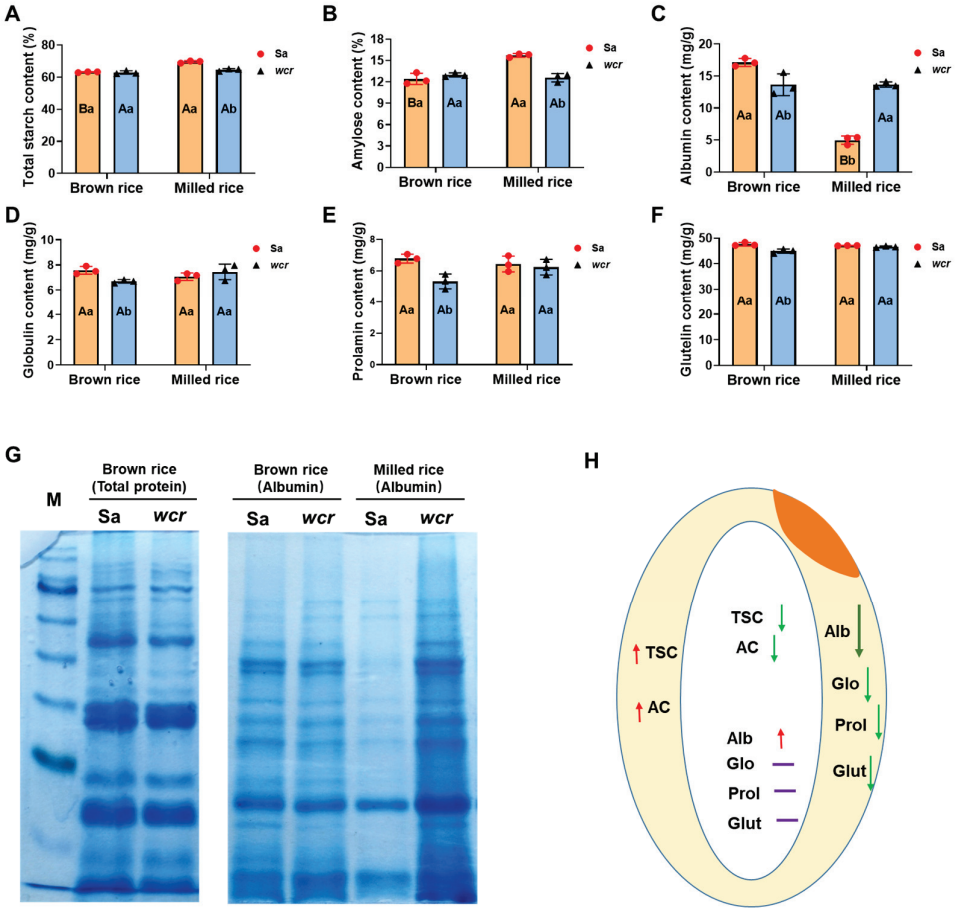

Figure 2. Physicochemical properties of mature grains of Sa and *wcr* mutant. (**A**–**F**) The contents of total starch, amylose, albumin, globulin, prolamin, and glutelin in brown rice flour and milled rice

flour of Sa and *wcr* mutant. Different small case letters denote statistical differences between Sa (orange) and *wcr* (blue) plants in the same rice type (brown or milled); different upper case letters denote statistical differences between brown or milled rice types in the same wild type or *wcr* plants. (**G**) The SDS-PAGE analysis of storage proteins of Sa and *wcr* mutant rice flour. (**H**) Changes in the contents of storage substances in the inner and outer endosperm of *wcr* compared to wild type (the yellow part indicates the outer endosperm; the white part indicates the inner endosperm; Red arrows represent up-regulated levels, light green and dark green arrows represent down-regulated levels, and purple horizontal lines indicate unchanged levels; TSC, total starch content; AC, amylose content; Alb, albumin; Glo, globulin; Prol, prolamin; Glut, glutelin).

2.3. Gelatinization Characteristics and Amylopectin Structure Analysis of *wcr*

Although the total starch content and amylose content in the brown rice flour of the mutant did not show significant differences compared to those of the wild type (Figure 2A,B), results from starch gelatinization tests using varying concentrations of urea solution revealed that the *wcr* mutant exhibited reduced solubility in urea. As shown in Figure 3A,B, the swelling volume of the mutant's brown rice flour was smaller than that of the wild-type under different urea concentrations, with a noticeable difference observed at a concentration of 1 mol/L (Figure 3B). At a urea concentration of 8 mol/L, no expansion was observed for the mutant while it continued for the wild type (Figure 3A,B). Furthermore, the analysis of the amylopectin chain length distribution revealed a reduction in the content of chain lengths ranging from 6 to 35 in the degree of polymerization (DP) in the brown rice flour of the *wcr* mutant, while an increase was observed in the chain lengths greater than 36 DP (Figure 3C).

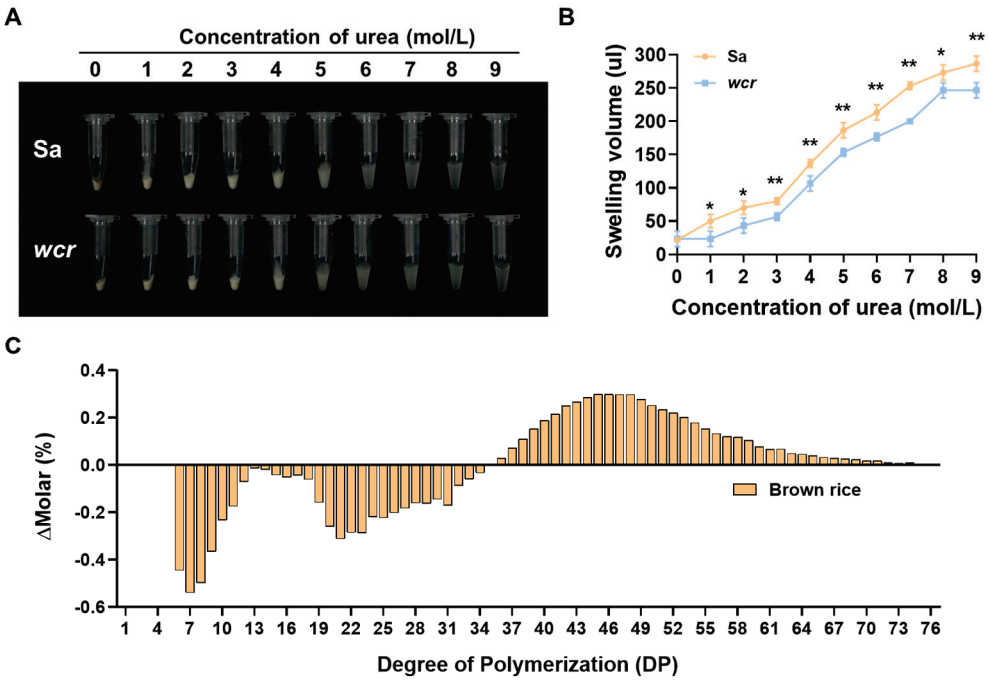

Figure 3. Gelatinization characteristics and amylopectin structure analysis of *wcr* mutant. (**A**,**B**) Comparison of the swelling volume of brown rice flour between Sa and *wcr* mutant under different urea concentrations. The *p*-values < 0.05 * and <0.01 ** calculated using an independent-samples *t*-test. (**C**) Determination of amylopectin chain length distribution in Sa and *wcr* mutant brown rice flour.

2.4. Genetic Analysis of the *wcr* Mutant

The *wcr* mutant was crossed with the background Sasanishiki, resulting in F_1 plants with normal grain development. In the subsequent F_2 population, segregation occurred between plants with normal endosperm and those with floury endosperm. A random

survey of 232 F_2 plants showed that 177 had normal phenotypes, while 55 had mutant phenotypes. The observed segregation pattern of 3:1 ($x^2 = 0.143 < 3.84$) in the F_2 population suggests that the *wcr* mutant is governed by a pair of recessive nuclear genes.

2.5. Fine Mapping of the wcr Gene

We successfully mapped the *wcr* gene to chromosome 5 at approximately 5.41 Mb (19,268,954 bp–24,757,870 bp) physical interval using MutMap analysis (Figure 4A). To further fine-map the *wcr* gene, we crossed the *wcr* mutant with the single-segment substitution line SL418 [69], resulting in an F_2 population of 2074 plants. From this population, we selected 108 plants with floury endosperm and narrowed down the location of the target gene within a 76 kb region delimited by markers C5.8 and C5.9 (Table 2, Figure 4B), which encompasses eight *ORFs*. Among these *ORFs*, *LOC_Os05g33570* encoding a functional pyruvate phosphate dikinase *OsPPDKB* was found to be allelic to the *flo4* gene [24] (Figure 4C). Sequencing results revealed a frameshift mutation caused by an insertion of an additional base C into the second exon of *cyOsPPDKB* (Figure 4D). Therefore, *wcr* represents a novel allelic variation of *flo4*.

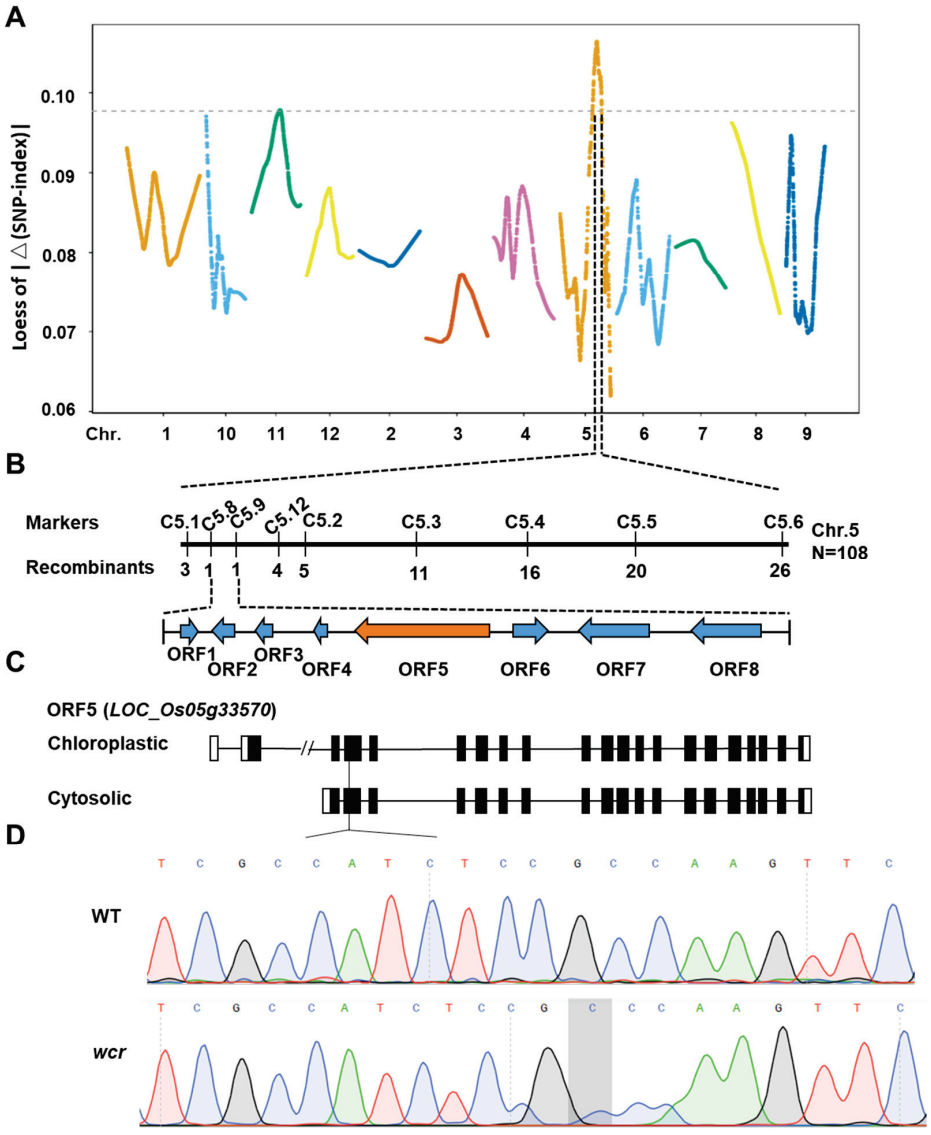

Figure 4. Fine mapping of the *wcr* gene. (**A**) The MutMap analysis of the *wcr* gene. (**B**) Fine mapping of the *wcr* gene using linkage analysis. (**C**) The gene structure of the *LOC_Os05g33570*. (**D**) The Sanger chromatogram of the WT and *wcr*.

Table 2. The primers used for fine-mapping the *wcr.*

Name	Forward Primer (5′-3′)	Reverse Primer (5′-3′)
C5.5	CTATGCAGTGCAGTGTGCAC	AGCCGAAGGAGGTGTGAATC
C5.4	GCTCAAGCAAGGTCCATTCC	CAGCTACTAGGCCCCATTTG
C5.3	CCTGGCGTCAAACACATCTG	CTGAGGGTGTTCTTTTGGGC
C5.2	ATGGGAGAAGTGTCCAGCAG	GTGTGGACTGTGGATTGTGG
C5.1	AGAACGGAGGGAGTAGGATC	TCGCGGCTCTGAATTACCAG
C5.8	GTCCACCCGTTTCTTGCATG	CCACCCGTTTCTTGCATACC
C5.9	CCGGATTGTAGCTGTAGCTC	GGGTCACAGCATCAAAGCAG
C5.12	GTGCTGGAAACTCCATGTCG	ATGGCTCTATCGGTGTCAGC

2.6. Gene Expression Analysis of Storage Substance-Related Genes

Considering the alterations in storage substances (starch and protein) composition and distribution in the *wcr* mutant (Figure 2), we conducted a comprehensive investigation into the expression levels of 35 genes involved in grain protein synthesis and 18 genes associated with starch metabolism in both wild type and mutant rice. As shown in Figure 5, the expression levels of most rice protein synthesis genes in the mutant were significantly reduced (Figure 5A), while the expression levels of starch metabolism-related genes exhibited varying degrees of downregulation, but the overall magnitude was small (Figure 5B).

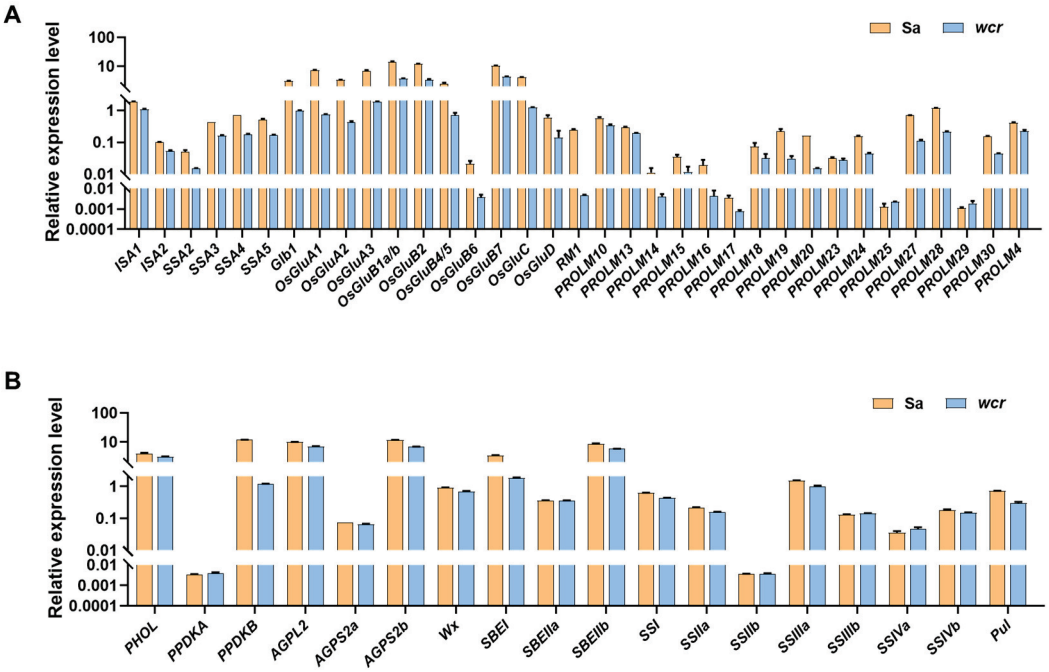

Figure 5. Gene expression analysis of storage substance-related genes of Sa and the *wcr* mutant. (**A**) Gene expression analysis of grain protein biosynthesis genes. (**B**) Gene expression analysis of grain starch metabolism genes.

3. Discussion

Through gene mapping and sequencing comparison, it has been determined that the *wcr* mutant is governed by a pair of recessive nuclear genes encoding pyruvate phosphate dikinase *OsPPDKB*. A comprehensive literature search has identified a total of nine distinct allelic mutant genes for *OsPPDKB*, including knockout mutations (*flo4-1* [24], *flo4-2* [24], *flo4-3* [24], and *flo4-303* [63]) and missense mutations (*flo4-4* [64], *flo4-5* [65], *flo4-6* [66], *M14* [67], and *floTR1* [68]) (Figure 6). The *wcr* mutation in this study was caused by an additional C base insertion in the second exon of *cyOsPPDKB*, resulting in a frameshift mutation. Therefore, the *wcr* can be considered as a novel allelic gene variant of *cyOsPPDKB*. Although there have been some changes in the crop traits of *wcr,* such as increased plant

height and reduced tillering, which differ from previous studies, it should be noted that the *wcr* mutant originated from plant tissue culture, where stable heritable variations were likely generated during the process.

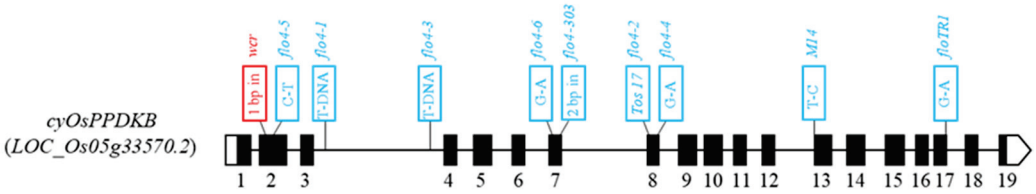

Figure 6. The structure of *cyOsPPDKB*. The *wcr* mutant is highlighted in red boxes. Previously reported allelic mutants of *cyOsPPDKB* are indicated in blue boxes.

The 10 identified mutants exhibit a similar floury endosperm phenotype but show significant variation in the physicochemical properties of rice flour, including amylose content, protein content, and amylopectin structure. In terms of amylose content, most allelic mutant grains show a significant reduction ranging from 13% to 46%, with *flo4-4* showing a 5% significant increase and *flo4-5* having no impact on amylose levels [24,63–68]. Interestingly, while the total starch and amylose content in the brown rice of the *wcr* mutant was similar to those of the wild type, there was a significant decrease observed in milled rice (Figure 2A,B). Regarding total protein content, *flo4-1*, *flo4-2*, and *floTR1* mutant grains showed an increase ranging from 5% to 16%, while *flo4-4* and *flo4-5* alleles resulted in a decrease of more than 15% [24,65]. No effect on protein was observed for *M14* rice grains [67]. Notably, the levels of four component proteins were significantly reduced in the *wcr* mutant brown rice; however, apart from a notable increase in albumin content observed for milled rice (Figure 2C–F), minimal changes were seen for the other three component proteins. The analysis of amylopectin structure revealed that both *floTR1* and *M14* grains contained fewer short chains (DP < 16) but more middle-length chains (DP > 18) [67,68], which diverged from our study's findings. In contrast, the *wcr* mutant showed fewer middle-length chains (DP < 36) but more long chains (DP > 36) (Figure 3C). In summary, different allelic mutations of *cyOsPPDKB* resulted in significant changes in the content and distribution of rice storage substances across various genetic backgrounds; however, these changes were not entirely consistent. We propose that the diverse physicochemical properties observed among different allelic mutants of *cyOsPPDKB* may be attributed to several potential factors. Firstly, distinct allelic variants of *cyOsPPDKB* may generate varying protein functionalities; however, further confirmation through molecular genetic experiments is required for validation. Secondly, dissimilar genetic backgrounds among the mutants are likely to contribute to the observed differences in physicochemical properties, which could be influenced by combined effects with other genes within the genome. Therefore, future investigations can explore utilizing CRISPR/Cas9 technology to generate saturated mutations of *cyOsPPDKB* within a consistent genetic background and fully elucidate its molecular functions.

The distinguishing feature of the *OsPPDK* mutant, compared to other mutants with rice floury endosperm, is its translucent periphery and internal flouriness. In this study, the arrangement of amyloplasts in the inner endosperm and peripheral regions of *wcr* grains was observed using scanning electron microscopy. It was found that internal amyloplasts had a loose structure with larger pores, while external ones were densely packed (Figure 1D). The reasons for this phenomenon may be as follows: (1) Previous research has shown that OsPPDKB plays a compensatory role in ATP deficiency under anaerobic conditions [24,57–61]. In normal rice grains, the level of hypoxia is more severe in the inner endosperm than in the peripheral region. When there is a mutation in *OsPPDKB* and it loses its ability to convert pyruvate and inorganic phosphate (PI) into phosphoenolpyruvate (PEP) and inorganic pyrophosphate (PPI), which can serve as an alternative phosphate donor for ATP production [24,57–61], substance transport within the inner endosperm is disrupted, leading to poor endosperm development. However, there is an improvement in aerobic conditions within the grain

peripheral region, likely providing sufficient energy for material transport through other oxidative phosphorylation pathways, resulting in relatively normal development. (2) The expression of *cyOsPPDKB* in rice grains peaks at 10 days after flowering, followed by a rapid decrease in protein level and activity through threonyl-phosphorylation and protein degradation mechanisms around approximately 20 days after anthesis [61]. This suggests that *cyOsPPDKB* likely plays a crucial role during the initial 20-day period of grain filling, which coincides with the accumulation of storage substances from the central region to the periphery of the endosperm [70]. Therefore, *cyOsPPDKB*'s temporal and spatial expression pattern is likely associated with severe flouriness in the inner endosperm and light flouriness in the outer part. However, these reasons are speculative hypotheses and require rigorous scientific experiments for validation.

With the diversification in dietary habits and rapid economic development, rice-processing products like rice cakes, noodles, crackers, and wine have become increasingly popular. Compared to regular rice cultivars, floury mutants with loosely packed starch granules are particularly useful for dry-milled flour production or sake brewing due to their easily breakable soft endosperm, finer particle size and less damaged starch, and strong water absorption properties [71]. For example, South Korea has developed new floury rice varieties such as suweon542 [72], Hangaru [73], Shingil [74], and Garumi2 [71] in recent years, which successfully reduced the rice milling cost. In this study, the starch granules of *wcr* endosperm were also loosely arranged, providing new genetic material for breeding low milling cost rice material. Additionally, previous research has shown that the spatial distribution and characteristics of protein content and composition in *Japonica* rice grains play a crucial role in determining the taste quality of sake [75]. In this study, the total protein content and component protein content in the *wcr* mutant brown rice were significantly reduced, accompanied by alterations in the distribution of different components. These findings highlight its potential as a valuable germplasm resource for further research aimed at enhancing sake quality.

4. Materials and Methods

4.1. Experimental Materials and Field Design

The experimental materials utilized in this study included a rice endosperm mutant *wcr* obtained through plant tissue culture, the parental variety Sasanishiki as the background control, an F_2 segregating population derived from crossing *wcr* with Sasanishiki for MutMap analysis, and an F_2 population obtained by crossing *wcr* with single-segment substitution line SL418 [69] for precise gene mapping. Each genotype was individually planted with 10 plants per row. The spacing between plants was maintained at 20 cm × 20 cm. Standardized water and fertilizer management practices were followed.

4.2. Main Crop Characteristics of Rice Plants

The main panicles of 5 mutants and 5 wild-type rice plants were selected at the mature stage for the investigation of height, tiller, grain length, grain width, and 1000-grain weight. The grain length and width of 5 mutants and 5 wild-type rice grains were analyzed using a seed-measuring instrument (Model SC-G, Wanshen, Hangzhou, China).

4.3. Scanning Electron Microscopy Analysis

The mature grains were transversely cut by a knife, coated with gold, and examined under a scanning electron microscope (SEM, S-4800, Hitachi, Tokyo, Japan). For the observation of compound starch granules, transverse sections of WT and *wcr* mature grains were fixed overnight in 2% (*v*/*v*) glutaraldehyde (CAS: P1126, Solarbio, Beijing, China). The samples were dehydrated in an alcohol series, embedded in LR White resin (Heraeus Kulzer, Wehrheim, Germany), and sectioned using an ultrathin microtome (EM UC7, Leica, Wetzlar, Germany). Semi-thin sections were dried at 40 °C, stained with I2-KI solution for visualization purposes, and observed using a light microscope (BX53 Olympus, Tokyo, Japan).

4.4. Protein Extraction and SDS-PAGE Analysis

The four kinds of storage protein were extracted from the floury grains of 3 mutants and 3 wild-type rice grains as described previously [76], and the protein content of each component was quantified by means of the Bradford assay [77]. SDS-PAGE and protein gel blot analysis were performed as described previously [78].

4.5. Analysis of the Amylose and Total Starch Contents

The total starch content of 3 mutants and 3 wild-type rice grains was measured using a Megazyme Inc. kit (Bray, Ireland) following the procedure provided by the manufacturer. The apparent amylose content of 3 mutants and 3 wild-type rice grains was measured using the iodine colorimetric method [79].

4.6. Determination of Chain Length of Amylopectin

Starch (10 mg) was dissolved in 5 mL water in a water bath (100 °C) for 60 min; 10 microliters (μL) of sodium azide solution (2% *w*/*v*) (CAS: 26628-22-8, Anpel, Shanghai, China), 50 μL acetate buffer (0.6 M, pH 4.4) (CAS: 127-09-3, Anpel, China), and 10 μL isoamylase (1400 U) (08124, Sigma, Darmstadt, Germany) were added to the starch dispersion, and the mixture was incubated in a water bath at 37 °C for 24 h. The hydroxyl groups of the debranched glucans were reduced by treatment with 0.5% (*w*/*v*) of sodium borohydride (16940-66-2, Sigma, Germany) under alkaline conditions for 20 h. The preparation of about 600 μL was dried in a vacuum at room temperature and allowed to dissolve in 30 μL of 1 M NaOH (16940-66-2, Anpel, China) for 60 min. Then, the solution was diluted with 570 μL of distilled water.

The sample extracts were analyzed by high-performance anion-exchange chromatography (HPAEC) on a CarboPac PA-200 anion-exchange column (4.0 × 250 mm; Dionex, Sunnyvale, CA, USA) using a pulsed amperometric detector (PAD; Dionex ICS 5000 system): flow rate, 0.4 mL/min; injection volume, 5 μL; solvent system, 0.2 M NaOH: (0.2 M NaOH, 0.2 M NaAc); gradient program, 90:10 *v*/*v* at 0 min, 90:10 *v*/*v* at 10 min, 40:60 *v*/*v* at 30 min, 40:60 *v*/*v* at 50 min, 90:10 *v*/*v* at 50.1 min, 90:10 *v*/*v* at 60 min.

Data were acquired on the ICS5000 (Thermo Scientific, Waltham, MA, USA) and processed using chromeleon 7.2 CDS (Thermo Scientific). Quantified data were output into Excel format.

4.7. Genome Mapping of the wcr Locus

To clone the *wcr* gene, the *wcr* mutant was used as the parent for a cross with the background Sasanishiki to produce an F_2 population. In the F_2 population, we selected 20 plants that showed *wcr* phenotypes and pooled their DNA in an equal ratio for whole-genome resequencing using NextSeq 500 (Illumina, San Diego, CA, USA, http://www.illumina.com/). The MutMap was performed according to a previous study [80]. To further fine-mapping the *wcr* gene, we crossed the *wcr* mutant with the single-segment substitution line SL418 to create an F_2 population of 2074 plants. By utilizing nine polymorphic markers, we successfully fine-mapped the target gene. The primers are listed in Table 2.

4.8. RNA Extraction, cDNA Preparation, and qRT-PCR

Total RNA was extracted from endosperm 15 days after flowering using an RNA extraction kit (Tiangen, Beijing, China). First-strand cDNA was synthesized using a reverse transcription kit (Vazyme, Nanjing, China). Quantitative reverse transcription (qRT-PCR) was performed with a CFX96 Real-Time PCR System (Bio-Rad, Hercules, CA, USA) using an SYBR qPCR Master Mix (Vazyme, Nanjing, China). The PCR procedure was carried out using the following program: 95 °C for 10 min, then 40 cycles of 95 °C for 15 s, and 60 °C for 1 min. All assays were performed with at least three biological replicates; each biological study repeats the setup three times for technical replication. The rice *actin* gene served as the internal control to normalize gene expression. Quantitative gene expression

was analyzed from three biological replicates by the $2^{-\Delta\Delta Ct}$ method. The primers were used as described previously [76].

4.9. Data Analysis

The experiment was conducted with three biological replicates for each sample. Microsoft Excel 2016 was used for data collection, while the statistical analysis software SPSS 15.0 was used for variance analysis. All experimental data are presented as means ± SD.

Author Contributions: Y.Y. and C.Y. designed this research and analyzed the data; Z.S. (Zixing Sun), L.W., Y.Z., X.Y., Z.S. (Ziyan Shen), M.G. and J.Z. performed this research; Y.Y. and C.Y. wrote this manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Biological Breeding-National Science and Technology Major Project (2023ZD04068), the National Natural Science Foundation of China (32301827), the China Postdoctoral Science Foundation (2023M733005, 2024T170754), the programs of the Jiangsu Province Government (BK20210798, BE2022365-2, BE2022335, JBGS [2021]001, and BE2021334-1), the Project of Zhongshan Biological Breeding Laboratory (ZSBBL-KY2023-01), the Open Project Program (20240106) of State Key Laboratory of Rice Biology and Breeding, the Open Funds of the Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding (PL202306), and the Priority Academic Program Development of Jiangsu Higher Education Institutions (PAP).

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Zheng, X.M.; Wei, F.; Cheng, C.; Qian, Q. A historical review of hybrid rice breeding. *J. Integr. Plant Biol.* **2024**, *66*, 532–545. [CrossRef] [PubMed]
2. Ren, D.Y.; Ding, C.Q.; Qian, Q. Molecular bases of rice grain size and quality for optimized productivity. *Sci. Bull.* **2023**, *68*, 314–350. [CrossRef] [PubMed]
3. Zhao, D.S.; Zhang, C.Q.; Li, Q.F.; Liu, Q.Q. Genetic control of grain appearance quality in rice. *Biotechnol. Adv.* **2022**, *60*, 108014. [CrossRef] [PubMed]
4. Long, W.H.; Dong, B.N.; Wang, Y.H.; Pan, P.Y.; Wang, Y.L.; Liu, L.L.; Chen, X.L.; Liu, X.; Liu, S.J.; Tian, Y.L.; et al. *FLOURY ENDOSPERM 8*, encoding the UDP-glucose pyro phosphorylase 1, affects the synthesis and structure of starch in rice endosperm. *J. Plant Biol.* **2017**, *60*, 513–522. [CrossRef]
5. Tsuneo, K.; Ryutaro, M.; Shinjiro, O.; Yu, W.; Naohiro, A.; Akira, H. Evaluation of alleles at *OsAGPS2*, *OsAGPL2*, and *OsSUT1* related to grain filling in rice in a common genetic background. *Crop Sci.* **2020**, *61*, 1154–1167.
6. Bilal, C.; Shota, S.; Aytug, T.; Hiroaki, M.; Ryosuke, S.; Salvinder, S.; Naoko, C.; Yuko, H.; Naoko, F.; Seon-Kap, H.; et al. Analysis of the Rice ADP-Glucose Transporter (OsBT1) Indicates the Presence of Regulatory Processes in the Amyloplast Stroma That Control ADP-Glucose Flux into Starch. *Plant Physiol.* **2016**, *170*, 1271–1283.
7. Yang, R.F.; Sun, C.L.; Bai, J.J.; Luo, Z.X.; Shi, B.; Zhang, J.M.; Yan, W.G.; Piao, Z.Z. A Putative Gene *sbe3-rs* for Resistant Starch Mutated from *SBE3* for Starch Branching Enzyme in Rice (*Oryza sativa L.*). *PLoS ONE* **2012**, *7*, e43026. [CrossRef]
8. Ryoo, N.; Yu, C.; Park, C.; Baik, M.; Park, I.M.; Cho, M.; Bhoo, S.H.; An, G.; Hahn, T.; Jeon, J. Knockout of a starch synthase gene *OsSSIIIa/Flo5* causes white-core floury endosperm in rice (*Oryza sativa L.*). *Plant Cell Rep.* **2007**, *26*, 1083–1095. [CrossRef]
9. Satoh, H.; Shibahara, K.; Tokunaga, T.; Nishi, A.; Tasaki, M.; Hwang, S.; Okita, T.; Kaneko, N.; Fujita, N.; Yoshida, M.; et al. Mutation of the plastidial α-glucan phosphorylase gene in rice affects the synthesis and structure of starch in the endosperm. *Plant Cell* **2008**, *20*, 1833–1849. [CrossRef]
10. Sun, L.; Yang, D.L.; Kong, Y.; Chen, Y.; Li, X.Z.; Zeng, L.J.; Li, Q.; Wang, E.T.; He, Z.H. Sugar homeostasis mediated by cell wall invertase *GRAIN INCOMPLETE FILLING 1* (*GIF1*) plays a role in pre-existing and induced defence in rice. *Mol. Plant Pathol.* **2014**, *15*, 161–173. [CrossRef]
11. Huang, L.C.; Sreenivasulu, N.; Liu, Q.Q. Waxy editing: Old meets new. *Trends Plant Sci.* **2020**, *25*, 963–966. [CrossRef] [PubMed]
12. Wu, H.M.; Ren, Y.L.; Dong, H.; Xie, C.; Zhao, L.; Wang, X.; Zhang, F.L.; Zhang, B.L.; Jiang, X.K.; Huang, Y.S.; et al. FLOURY ENDOSPERM24, a heat shock protein 101 (HSP101), is required for starch biosynthesis and endosperm development in rice. *New Phytol.* **2024**, *242*, 2635–2651. [CrossRef] [PubMed]
13. Matsushima, R.; Maekawa, M.; Kusano, M.; Kondo, H.; Fujita, N.; Kawagoe, Y.; Sakamoto, W. Amyloplast-Localized SUBSTANDARD STARCH GRAIN4 Protein Influences the Size of Starch Grains in Rice Endosperm. *Plant Physiol.* **2013**, *164*, 623–636. [CrossRef]
14. Matsushima, R.; Maekawa, M.; Kusano, M.; Tomita, K.; Kondo, H.; Nishimura, H.; Crofts, N.; Fujita, N.; Sakamoto, W. Amyloplast Membrane Protein SUBSTANDARD STARCH GRAIN6 Controls Starch Grain Size in Rice Endosperm. *Plant Physiol.* **2016**, *170*, 1445–1459. [CrossRef]

15. Yang, H.G.; Ren, Y.L.; Zhang, B.L.; Jin, J.; Du, F.L.; Shang, Z.Z.; Fu, Y.S.; Zhu, Y.; Wang, X.; Zhu, C.Y.; et al. SUBSTANDRAD STARCH GRAIN7 regulates starch grain size and endosperm development in rice. *Plant Biotechnol. J.* **2024**, 1–15. [CrossRef]
16. Long, W.H.; Wang, Y.L.; Zhu, S.S.; Jing, W.; Wang, Y.H.; Ren, Y.L.; Tian, Y.L.; Liu, S.J.; Liu, X.; Chen, L.M.; et al. *FLOURY SHRUNKEN ENDOSPERM1* connects phospholipid metabolism and amyloplast development in rice. *Plant Physiol.* **2018**, *177*, 698–712. [CrossRef] [PubMed]
17. Hu, T.T.; Tian, Y.L.; Zhu, J.P.; Wang, Y.L.; Jing, R.N.; Lei, J.; Sun, Y.L.; Yu, Y.F.; Li, J.F.; Chen, X.L.; et al. *OsNDUFA9* encoding a mitochondrial complex I subunit is essential for embryo development and starch synthesis in rice. *Plant Cell Rep.* **2018**, *37*, 1667–1679. [CrossRef]
18. Kim, S.R.; Yang, J.I.; Moon, S.; Ryu, C.H.; An, K.; Kim, K.M.; Yim, J.; An, G. Rice *OGR1* encodes a pentatricopeptide repeat-DYW protein and is essential for RNA editing in mitochondria. *Plant J.* **2009**, *59*, 738–749. [CrossRef]
19. Wu, M.M.; Ren, Y.L.; Cai, M.H.; Wang, Y.L.; Zhu, S.S.; Zhu, J.P.; Hao, Y.Y.; Teng, X.; Zhu, X.P.; Jing, R.N.; et al. Rice *FLOURY ENDOSPERM10* encodes a pentatricopeptide repeat protein that is essential for the trans-splicing of mitochondrial nad1 intron 1 and endosperm development. *New Phytol.* **2019**, *223*, 736–750. [CrossRef]
20. Hao, Y.Y.; Wang, Y.L.; Wu, M.M.; Zhu, X.P.; Teng, X.; Sun, Y.L.; Zhu, J.P.; Zhang, Y.Y.; Jing, R.N.; Lei, J.; et al. The nuclear-localized PPR protein OsNPPR1 is important for mitochondrial function and endosperm development in rice. *J. Exp. Bot.* **2019**, *70*, 4705–4720. [CrossRef]
21. Yu, M.Z.; Wu, M.M.; Ren, Y.L.; Wang, Y.H.; Li, J.F.; Lei, C.L.; Sun, Y.L.; Bao, X.H.; Wu, H.M.; Yang, H.; et al. Rice *FLOURY ENDOSPERM 18* encodes a pentatricopeptide repeat protein required for 5′ processing of mitochondrial nad5 mRNA and endosperm development. *J. Integr. Plant Biol.* **2020**, *63*, 834–847. [CrossRef] [PubMed]
22. Yang, H.; Wang, Y.L.; Tian, Y.L.; Teng, X.; Lv, Z.H.; Lei, J.; Duan, E.C.; Dong, H.; Yang, X.; Zhang, Y.Y.; et al. Rice *FLOURY ENDOSPERM22*, encoding a pentatricopeptide repeat protein, is involved in both mitochondrial RNA splicing and editing and is crucial for endosperm development. *J. Integr. Plant Biol.* **2022**, *65*, 755–771. [CrossRef] [PubMed]
23. Chen, C.; He, B.S.; Liu, X.X.; Ma, X.D.; Liu, Y.J.; Yao, H.Y.; Zhang, P.; Yin, J.L.; Wei, X.; Koh, H.J.; et al. Pyrophosphate-fructose 6-phosphate 1-phosphotransferase (PFP1) regulates starch biosynthesis and seed development via heterotetramer formation in rice (*Oryza sativa* L.). *Plant Biotechnol. J.* **2020**, *18*, 83–95. [CrossRef]
24. Kang, H.G.; Park, S.; Matsuoka, M.; An, G. White-core endosperm *floury endosperm-4* in rice is generated by knockout mutations in the C-type pyruvate orthophosphate dikinase gene (*OsPPDKb*). *Plant J.* **2005**, *42*, 901–911. [CrossRef]
25. Cai, Y.C.; Li, S.F.; Jiao, G.A.; Sheng, Z.H.; Wu, Y.W.; Shao, G.N.; Xie, L.H.; Peng, C.; Xu, J.F.; Tang, S.Q.; et al. *OsPK2* encodes a plastidic pyruvate kinase involved in rice endosperm starch synthesis, compound granule formation and grain filling. *Plant Biotechnol. J.* **2018**, *16*, 1878–1891. [CrossRef]
26. Zhong, M.S.; Liu, X.; Liu, F.; Ren, Y.L.; Wang, Y.L.; Zhu, J.P.; Teng, X.; Duan, E.C.; Wang, F.; Zhang, H.; et al. *FLOURY ENDOSPERM12* Encoding Alanine Aminotransferase 1 Regulates Carbon and Nitrogen Metabolism in Rice. *J. Plant Biol.* **2019**, *62*, 61–73. [CrossRef]
27. You, X.M.; Zhang, W.W.; Hu, J.L.; Jing, R.N.; Cai, Y.; Feng, Z.M.; Kong, F.; Zhang, J.; Yan, H.G.; Chen, W.W.; et al. *FLOURY ENDOSPERM15* encodes a glyoxalase I involved in compound granule formation and starch synthesis in rice endosperm. *Plant Cell Rep.* **2019**, *38*, 345–359. [CrossRef]
28. Teng, X.; Zhong, M.S.; Zhu, X.P.; Wang, C.M.; Ren, Y.L.; Wang, Y.L.; Zhang, H.; Jiang, L.; Wang, D.; Hao, Y.Y.; et al. *FLOURY ENDOSPERM16* encoding a NAD-dependent cytosolic malate dehydrogenase plays an important role in starch synthesis and seed development in rice. *Plant Biotechnol. J.* **2019**, *17*, 1914–1927. [CrossRef] [PubMed]
29. Chen, X.L.; Ji, Y.; Zhao, W.Y.; Niu, H.Y.; Yang, X.; Jiang, X.K.L.; Zhang, Y.P.; Lei, J.; Yang, H.; Chen, R.B.; et al. Fructose-6-phosphate-2-kinase/fructose-2,6-bisphosphatase regulates energy metabolism and synthesis of storage products in developing rice endosperm. *Plant Sci.* **2023**, *326*, 111503. [CrossRef]
30. Lei, J.; Teng, X.; Wang, Y.F.; Jiang, X.K.; Zhao, H.H.; Zheng, X.M.; Ren, Y.L.; Dong, H.; Wang, Y.L.; Duan, E.C.; et al. Plastidic pyruvate dehydrogenase complex E1 componentsubunit Alpha1 is involved in galactolipid biosynthesisrequired for amyloplast development in rice. *Plant Biotechnol. J.* **2022**, *20*, 437–453. [CrossRef]
31. Lou, G.M.; Chen, P.L.; Zhou, H.; Li, P.B.; Xiong, J.W.; Wan, S.S.; Zheng, Y.Y.; Alam, M.; Liu, R.J.; Zhou, Y.; et al. *FLOURY ENDOSPERM19* encoding a class I glutamine amidotransferase affects grain quality in rice. *Mol. Breed.* **2021**, *41*, 36. [CrossRef] [PubMed]
32. Han, X.H.; Wang, Y.H.; Liu, X.; Jiang, L.; Ren, Y.L.; Liu, F.; Peng, C.; Li, J.J.; Jin, X.M.; Wu, F.P.; et al. The failure to express a protein disulphide isomerase-like protein results in a floury endosperm and an endoplasmic reticulum stress response in rice. *J. Exp. Bot.* **2012**, *63*, 121–130. [CrossRef] [PubMed]
33. Ren, Y.L.; Wang, Y.H.; Liu, F.; Zhou, K.N.; Ding, Y.; Zhou, F.; Wang, Y.; Liu, K.; Gan, L.; Ma, W.W.; et al. *GLUTELIN PRECURSOR ACCUMULATION3* encodes a regulator of post-Golgi vesicular traffic essential for vacuolar protein sorting in rice endosperm. *Plant Cell* **2014**, *26*, 410–425. [CrossRef] [PubMed]
34. Liu, F.; Ren, Y.L.; Wang, Y.H.; Peng, C.; Zhou, K.N.; Lv, J.; Guo, X.P.; Zhang, X.; Zhong, M.S.; Zhao, S.L.; et al. *OsVPS9A* Functions Cooperatively with *OsRAB5A* to Regulate Post-Golgi Dense Vesicle-Mediated Storage Protein Trafficking to the Protein Storage Vacuole in Rice Endosperm Cells. *Mol. Plant* **2013**, *6*, 1918–1932. [CrossRef] [PubMed]
35. Wang, Y.H.; Ren, Y.L.; Liu, X.; Jiang, L.; Chen, L.M.; Han, X.H.; Jin, M.N.; Liu, S.J.; Liu, F.; Lv, J.; et al. *OsRab5a* regulates endomembrane organization and storage protein trafficking in rice endosperm cells. *Plant J.* **2010**, *64*, 812–824. [CrossRef]

36. Wang, Y.H.; Liu, F.; Ren, Y.L.; Wang, Y.L.; Liu, X.; Long, W.H.; Wang, D.; Zhu, J.P.; Zhu, X.P.; Jing, R.N.; et al. *GOLGI TRANSPORT 1B* Regulates Protein Export from the Endoplasmic Reticulum in Rice Endosperm Cells. *Plant Cell* **2016**, *28*, 2850–2865. [CrossRef]
37. Ren, Y.L.; Wang, Y.H.; Pan, T.; Wang, Y.L.; Wang, Y.F.; Gan, L.; Wei, Z.Y.; Wang, F.; Wu, M.M.; Jing, R.N.; et al. *GPA5* Encodes a Rab5a Effector Required for Post-Golgi Trafficking of Rice Storage Proteins. *Plant Cell* **2020**, *32*, 758–777. [CrossRef]
38. Zhu, J.P.; Ren, Y.L.; Wang, Y.L.; Liu, F.; Teng, X.; Zhang, Y.Y.; Duan, E.C.; Wu, M.M.; Zhong, M.S.; Hao, Y.Y.; et al. OsNHX5-mediated pH homeostasis is required for post-Golgi trafficking of seed storage proteins in rice endosperm cells. *BMC Plant Biol.* **2019**, *19*, 295. [CrossRef]
39. Pan, T.; Wang, Y.H.; Jing, R.N.; Wang, Y.F.; Wei, Z.Y.; Zhang, B.L.; Lei, C.L.; Qi, Y.Z.; Wang, F.; Bao, X.H.; et al. Post-Golgi trafficking of rice storage proteins requires the small GTPase Rab7 activation complex MON1–CCZ1. *Plant Physiol.* **2021**, *187*, 2174–2191. [CrossRef]
40. Zhu, J.P.; Ren, Y.L.; Zhang, Y.Y.; Yang, J.; Duan, E.C.; Wang, Y.L.; Liu, F.; Wu, M.M.; Pan, T.; Wang, Y.F.; et al. Subunit E isoform 1 of vacuolar H+-ATPase OsVHA enables post-Golgi trafficking of rice seed storage proteins. *Plant Physiol.* **2021**, *187*, 2192–2208. [CrossRef]
41. Wang, R.Q.; Ren, Y.L.; Yan, H.G.; Teng, X.; Zhu, X.P.; Wang, Y.P.; Zhang, X.; Guo, X.P.; Lin, Q.B.; Cheng, Z.J.; et al. ENLARGED STARCH GRAIN1 affects amyloplast development and starch biosynthesis in rice endosperm. *Plant Sci.* **2021**, *305*, 110831. [CrossRef] [PubMed]
42. Liu, J.X.; Wu, X.B.; Yao, X.F.; Yu, R.; Larkin, P.; Liu, C.M. DNA demethylation by *ROS1a* in rice vegetative cells promotes methylation in sperm. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 11327–11332. [CrossRef] [PubMed]
43. Yan, M.; Pan, T.; Zhu, Y.; Jiang, X.; Yu, M.; Wang, R.; Zhang, F.; Luo, S.; Bao, X.; Chen, Y.; et al. *FLOURY ENDOSPERM20* encoding SHMT4 is required for rice endosperm development. *Plant Biotechnol. J.* **2022**, *20*, 1438–1440. [CrossRef] [PubMed]
44. Wang, J.C.; Xu, H.; Zhu, Y.; Liu, Q.Q.; Cai, X.L. *OsbZIP58*, a basic leucine zipper transcription factor, regulates starch biosynthesis in rice endosperm. *J. Exp. Bot.* **2013**, *64*, 3453–3466. [CrossRef]
45. Fu, F.F.; Xue, H.W. Coexpression analysis identifies Rice Starch Regulator1, a rice AP2/EREBP family transcription factor, as a novel rice starch biosynthesis regulator. *Plant Physiol.* **2010**, *154*, 927–938. [CrossRef]
46. Bello, B.K.; Hou, Y.X.; Zhao, J.; Jiao, G.A.; Wu, Y.W.; Li, Z.Y.; Wang, Y.F.; Tong, X.H.; Wang, W.; Yuan, W.Y.; et al. NF-YB1-YC12-bHLH144 complex directly activates *Wx* to regulate grain quality in rice (*Oryza sativa* L.). *Plant Biotechnol. J.* **2019**, *17*, 1222–1235. [CrossRef]
47. Xiong, Y.F.; Ren, Y.; Li, W.; Wu, F.S.; Yang, W.J.; Huang, X.L.; Yao, J.L. *NF-YC12* is a key multi-functional regulator of accumulation of seed storage substances in rice. *J. Exp. Bot.* **2019**, *70*, 3765–3780. [CrossRef]
48. She, K.C.; Kusano, H.; Koizumi, K.; Yamakawa, H.; Hakata, M.; Imamura, T.; Fukuda, M.; Naito, N.; Tsurumaki, Y.; Yaeshima, M.; et al. A novel factor *FLOURY ENDOSPERM2* is involved in regulation of rice grain size and starch quality. *Plant Cell* **2010**, *22*, 3280–3294. [CrossRef]
49. Zhang, L.; Li, N.; Zhang, J.; Zhao, L.; Wei, C. The CBM48 domain-containing protein *FLO6* regulates starch synthesis by interacting with SSIVb and GBSS in rice. *Plant Mol. Biol.* **2022**, *108*, 343–361. [CrossRef]
50. Zhang, L.; Ren, Y.L.; Lu, B.Y.; Wan, J.M. *FLOURY ENDOSPERM7* encodes a regulator of starch synthesis and amyloplast development essential for peripheral endosperm development in rice. *J. Exp. Bot.* **2016**, *67*, 633–647. [CrossRef]
51. Zhu, X.P.; Teng, X.; Wang, Y.L.; Hao, Y.Y.; Jing, R.N.; Wang, Y.F.; Liu, Y.; Zhu, J.P.; Wu, M.M.; Zhong, M.S.; et al. *FLOURY ENDOSPERM11* encoding a plastid heat shock protein 70 is essential for amyloplast development in rice. *Plant Sci.* **2018**, *277*, 89–99. [CrossRef] [PubMed]
52. Yan, H.G.; Zhang, W.W.; Wang, Y.H.; Jin, J.; Xu, H.C.; Fu, Y.S.; Shan, Z.Z.; Wang, X.; Teng, X.; Li, X.; et al. LIKE EARLY STARVATION1 cooperates with FLOURY ENDOSPERM6 to modulate starch biosynthesis and endosperm development. *Plant Cell* **2024**, *36*, 1892–1912. [CrossRef] [PubMed]
53. Huang, L.C.; Tan, H.Y.; Zhang, C.Q.; Li, Q.F.; Liu, Q.Q. Starch biosynthesis in cereal endosperms: An updated review over the last decade. *Plant Commun.* **2021**, *2*, 100237. [CrossRef]
54. Borisjuk, L.; Rolletschek, H. The oxygen status of the developing seed. *New Phytol.* **2009**, *182*, 17–30. [CrossRef] [PubMed]
55. Rolletschek, H.; Koch, K.; Wobus, U.; Borisjuk, L. Positional cues for the starch/lipid balance in maize kernels and resource partitioning to the embryo. *Plant J.* **2005**, *42*, 69–83. [CrossRef]
56. Hu, L.; Tu, B.; Yang, W.; Yuan, H.; Li, J.L.; Guo, L.N.; Zheng, L.; Chen, W.L.; Zhu, X.B.; Wang, Y.P.; et al. Mitochondria-Associated Pyruvate Kinase Complexes Regulate Grain Filling in Rice. *Plant Physiol.* **2020**, *183*, 1073–1087. [CrossRef]
57. Lee, S.K.; Jeon, J.S. Review: Crucial Role of Inorganic Pyrophosphate in Integrating Carbon Metabolism from Sucrose Breakdown to Starch Synthesis in Rice Endosperm. *Plant Sci.* **2020**, *298*, 110572. [CrossRef]
58. Imaizumi, N.; Ku, M.S.; Ishihara, K.; Samejima, M.; Kaneko, S.; Matsuoka, M. Characterization of the gene for pyruvate, orthophosphate dikinase from rice, a C3 plant, and a comparison of structure and expression between C3 and C4 genes for this protein. *Plant Mol. Biol.* **1997**, *34*, 701–716. [CrossRef] [PubMed]
59. Plaxton, W.C.; Podestá, F.E. The functional organization and control of plant respiration. *Crit. Rev. Plant Sci.* **2006**, *25*, 159–198. [CrossRef]
60. Lappe, R.R.; Baier, J.W.; Boehlein, S.K.; Huffman, R.; Lin, Q.H.; Wattebled, F.; Settles, A.M.; Hannah, L.C.; Borisjuk, L.; Rolletschek, H.; et al. Functions of maize genes encoding pyruvate phosphate dikinase in developing endosperm. *Proc. Natl. Acad. Sci. USA* **2018**, *115*, 24–33. [CrossRef]

61. Chastain, C.J.; Heck, J.W.; Colquhoun, T.A.; Voge, D.G.; Gu, X.-Y. Posttranslational regulation of pyruvate, orthophosphate dikinase in developing rice (*Oryza sativa*) seeds. *Planta* **2006**, *224*, 924–934. [CrossRef] [PubMed]
62. Zhou, H.J.; Wang, L.J.; Liu, G.F.; Meng, X.B.; Jing, Y.H.; Shu, X.L.; Kong, X.L.; Sun, J.; Yu, H.; Smith, S.M.; et al. Critical roles of soluble starch synthase SSIIIa and granule-bound starch synthase Waxy in synthesizing resistant starch in rice. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 12844–12849. [CrossRef] [PubMed]
63. Matsuba, S.; Maruyama-Funatsuki, W.; Umemoto, T.; Kato, H.; Kuroki, M.; Yokogami, N.; Ikegaya, T.; Shimizu, H.; Iriki, N. The Induced Mutant Allele *flo4-303* Confers Floury Characteristics on the *Japonica* Rice Cultivar 'Hoshinoko'. *Breed. Sci.* **2022**, *72*, 383–388. [CrossRef] [PubMed]
64. Wang, H.; Mo, Y.; Im, D.E.; Jang, S.G.; Ham, T.H.; Lee, J.; Jeung, J.U.; Kwon, S.W. A New SNP in *CyOsPPDK* Gene is Associated with Floury Endosperm in Suweon 542. *Mol. Genet. Genom.* **2018**, *293*, 1151–1158. [CrossRef] [PubMed]
65. Wang, H.; Ham, T.H.; Im, D.E.; Lar, S.M.; Jang, S.G.; Lee, J.; Mo, Y.; Jeung, J.U.; Kim, S.T.; Kwon, S.W. A New SNP in Rice Gene Encoding Pyruvate Phosphate Dikinase (PPDK) Associated with Floury Endosperm. *Genes* **2020**, *11*, 465. [CrossRef]
66. Ha, S.K.; Lee, H.S.; Lee, S.Y.; Lee, C.M.; Mo, Y.J.; Jeung, J.U. Characterization of *flo4-6*, a novel *cyOsPPDKB* allele conferring floury endosperm characteristics suitable for dry-milled rice flour production. *Agronomy* **2023**, *13*, 1306. [CrossRef]
67. Zhang, L.; Zhao, L.L.; Lin, L.S.; Zhao, L.X.; Liu, Q.Q.; Wei, C.X. A Novel Mutation of *OsPPDKB*, Encoding Pyruvate Orthophosphate Dikinase, Affects Metabolism and Structure of Starch in the Rice Endosperm. *Int. J. Mol. Sci.* **2018**, *19*, 2268. [CrossRef]
68. Muroyama, R.; Ito, H.; Takahashi, S.; Kang, D.J.; Hamada, S. Biochemical Analysis of a Novel Allele of the *OsPPDKB* Gene Associated with Floury Endosperm. *J. Cereal Sci.* **2022**, *107*, 103529. [CrossRef]
69. Ando, T.; Yamamoto, T.; Shimizu, T.; Ma, X.F.; Shomura, A.; Takeuchi, Y.; Lin, S.Y.; Yano, M. Genetic dissection and pyramiding of quantitative traits for panicle architecture by using chromosomal segment substitution lines in rice. *Theor. Appl. Genet.* **2008**, *116*, 881–890. [CrossRef]
70. Liu, J.X.; Wu, M.W.; Liu, C.M. Cereal Endosperms: Development and Storage Product Accumulation. *Annu. Rev. Plant Biol.* **2022**, *73*, 255–291. [CrossRef]
71. Mo, Y.; Jeung, J.-U. The Use of Floury Endosperm Mutants to Develop Rice Cultivars Suitable for Dry Milling. *Plant Biotechnol. Rep.* **2020**, *14*, 185–191. [CrossRef]
72. Mo, Y.; Jeung, J.-U.; Shin, Y.-S.; Park, C.S.; Kang, K.-H.; Kim, B.-K. Agronomic and Genetic Analysis of Suweon 542, a Rice Floury Mutant Line Suitable for Dry Milling. *Rice* **2013**, *6*, 37. [CrossRef] [PubMed]
73. Won, Y.-J.; Ahn, E.-K.; Jeong, E.-G.; Chang, J.-K.; Lee, J.-H.; Jung, K.-H.; Hyun, U.-J.; Cho, Y.-C.; Oh, S.-K.; Yoon, M.-R.; et al. An Opaque Endosperm Rice Cultivar, 'Hangaru', Suitable for Exclusive Dry-Milling Rice Flour Production. *Korean J. Breed. Sci.* **2019**, *51*, 134–139. [CrossRef]
74. Cho, Y.; Baek, M.; Park, H.; Cho, J.; Ahn, E.; Suh, J.; Jeung, J. 'Shingil (Milyang317)', Tongil-Type Variety Specialized for Rice Flour. *Korean J. Breed. Sci.* **2020**, *72*, 58–72. [CrossRef]
75. Takahashi, K.; Kohno, H.; Okuda, M. Spatial Distribution and Characteristics of Protein Content and Composition in Japonica Rice Grains: Implications for Sake Quality. *Rice* **2024**, *17*, 26. [CrossRef]
76. Yang, Y.H.; Zhang, Y.; Sun, Z.X.; Shen, Z.Y.; Li, Y.G.; Guo, Y.F.; Feng, Y.T.; Sun, S.Y.; Guo, M.; Hu, Z.; et al. Knocking out *OsAAP11* to improve rice grain quality using CRISPR/Cas9 system. *Int. J. Mol. Sci.* **2023**, *24*, 14360. [CrossRef]
77. Bradford, M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **1976**, *72*, 248–254. [CrossRef] [PubMed]
78. Saito, Y.; Shigemitsu, T.; Yamasaki, R.; Sasou, A.; Goto, F.; Kishida, K.; Kuroda, M.; Tanaka, K.; Morita, S.; Satoh, S.; et al. Formation mechanism of the internal structure of type I protein bodies in rice endosperm: Relationship between the localization of prolamin species and the expression of individual genes. *Plant J.* **2012**, *70*, 1043–1055. [CrossRef] [PubMed]
79. Sun, J.; Wang, Y.; Zhang, X.Q.; Rasmussen, S.K.; Jiang, X.T.; Song, W.J.; Wu, D.X.; Shu, X.L. Dependence of physiochemical, functional and textural properties of high–resistant starch rice on endogenous nonstarch polysaccharides. *Inte. J. Food Sci. Technol.* **2018**, *53*, 1079–1086. [CrossRef]
80. Abe, A.; Kosugi, S.; Yoshida, K.; Natsume, S.; Takagi, H.; Kanzaki, H.; Matsumura, H.; Yoshida, K.; Mitsuoka, C.; Tamiru, M.; et al. Genome sequencing reveals agronomically important loci in rice using MutMap. *Nat. Biotechnol.* **2012**, *30*, 174–178. [CrossRef]

plants

MDPI

Article

Unveiling Salt Tolerance Mechanisms and Hub Genes in Alfalfa (*Medicago sativa* L.) Through Transcriptomic and WGCNA Analysis

Fengdan Wang [1], Hanfu Wu [1], Mei Yang [1], Wen Xu [1], Wenjie Zhao [1], Rui Qiu [1], Ning Kang [2] and Guowen Cui [1,*]

[1] Department of Grassland Science, College of Animal Science and Technology, Northeast Agricultural University, Harbin 150030, China; wangfengdan122@163.com (F.W.); w995450205@sina.com (H.W.); yangmei199611@163.com (M.Y.); xw20220921@163.com (W.X.); 13310467886@163.com (W.Z.); qr020327@163.com (R.Q.)

[2] Department of Animal Science, College of Animal Science and Technology, Inner Mongolia Agricultural University, Hohhot 010018, China; 13789734560@163.com

* Correspondence: cgw603@163.com

Abstract: Alfalfa (*Medicago sativa* L.), an important forage crop with high nutritional value and good palatability, plays a vital role in the development of animal husbandry in China. In Northeast China, there are vast areas of saline–alkali land that remain undeveloped. Given that alfalfa is a highly adaptable forage crop, exploring its salt tolerance at the molecular transcriptional level and identifying salt-tolerant genes has great significance for breeding salt-resistant alfalfa varieties. This also provides valuable genetic resources for better utilization of saline–alkali land. In this study, we conducted two rounds of screening on 41 alfalfa varieties and identified WL168 as a salt-sensitive variety and Longmu801 as a salt-tolerant variety. After 7 days of 300 mM salt stress, both varieties showed a decreasing trend in plant height, fresh weight, and dry weight over time, but Longmu801 demonstrated better water retention ability compared to WL168. Chlorophyll content also declined, but chlorophyll a and total chlorophyll levels in Longmu801 were higher than in WL168. Hydrogen peroxide and malondialdehyde levels increased overall, but Longmu801 had significantly lower levels than WL168 under prolonged stress. Both varieties showed increasing trends in soluble sugars, proline, and antioxidant enzymes (SOD, POD, CAT), with Longmu801 significantly outperforming WL168. This suggests that the two varieties share similar growth and physiological response mechanisms, with their differences primarily arising from variations in indicator levels. In the above, comparisons between varieties were conducted based on the relative values of the indicators in relation to their controls. Transcriptomic analysis revealed that under salt stress, Longmu801 had 16,485 differentially expressed genes (DEGs) relative to its control, while WL168 had 18,726 DEGs compared to its control. Among these, 2164 DEGs shared the same expression trend, with GO functions enriched in response to oxidative stress, nucleus, plasma membrane, and others. The KEGG pathways were enriched in phenylpropanoid biosynthesis, protein processing in the endoplasmic reticulum, starch and sucrose metabolism, and others. This suggests that alfalfa's transcriptional response mechanism to salt stress involves these pathways. Additionally, the variety-specific DEGs were also enriched in the same KEGG pathways and GO functions, indicating that the differences between the two varieties stem from their unique stress-responsive DEGs, while their overall mechanisms for coping with stress remain similar. To further identify salt stress-related genes, this study conducted WGCNA analysis using 32,683 genes and physiological indicators. Six modules closely related to physiological traits were identified, and the top five genes ranked by degree in each module were selected as hub genes. Further analysis of these hub genes identified five genes directly related to salt stress: *Msa085011*, *Msa0605650*, *Msa0397400*, *Msa1258740*, and *Msa0958830*. Mantel test analysis revealed that these genes showed strong correlations with physiological indicators. This study will provide important insights for breeding salt-tolerant alfalfa varieties.

Keywords: alfalfa; salt stress; transcriptomic analysis; WGCNA; hub gene

Citation: Wang, F.; Wu, H.; Yang, M.; Xu, W.; Zhao, W.; Qiu, R.; Kang, N.; Cui, G. Unveiling Salt Tolerance Mechanisms and Hub Genes in Alfalfa (*Medicago sativa* L.) Through Transcriptomic and WGCNA Analysis. *Plants* **2024**, *13*, 3141. https://doi.org/10.3390/plants13223141

Academic Editor: Roberto Barbato

Received: 16 September 2024
Revised: 24 October 2024
Accepted: 6 November 2024
Published: 8 November 2024

1. Introduction

Alfalfa (*Medicago sativa* L.) is an important perennial legume forage crop that is widely cultivated globally, with a planting area of approximately 45 million hectares. The history of its use as forage for feeding livestock dates back over 3300 years [1,2]. Alfalfa is known as the "king of forages" due to its high crude protein content, good palatability, and high biomass yield [3,4]. It also has a well-developed root system and strong adaptability [5], demonstrating characteristics such as cold tolerance [6], drought tolerance [7], and salt tolerance [8]. Additionally, alfalfa roots host nitrogen-fixing rhizobia, which play a role in nitrogen fixation [9], contributing to its ecological value [10]. Due to these superior traits, alfalfa has become the most widely planted forage crop in the world [11].

In China, alfalfa has also been utilized for a long time. As early as 2000 years ago, during the Western Han Dynasty, Zhang Qian introduced alfalfa to China via the Silk Road [12]. With the development of China's animal husbandry industry and the increasing domestic demand for livestock products, the demand for high-quality alfalfa as forage has steadily increased [13]. However, domestic alfalfa production is far from meeting market demand. The average self-sufficiency rate of high-quality alfalfa in China is only 64% [14]. Between 2012 and 2020, China's alfalfa hay imports increased from 440,000 tons to 1.36 million tons [15], highlighting a significant gap in alfalfa production. The development of China's alfalfa industry faces several challenges. For example, the germplasm resources of alfalfa are relatively poor, and modern forage breeding in China began relatively late in the second half of the 20th century, compared to developed countries [16]. Good arable land is typically used for traditional crops, and to alleviate land-use conflicts and prevent competition between grain and forage crops, alfalfa and other forages are often planted on marginal lands, such as those prone to drought or salinity [17,18]. This poses significant challenges for the production of high-quality alfalfa, making it urgent to accelerate alfalfa breeding and obtain salt-tolerant germplasm resources. Molecular breeding techniques provide an effective solution to this challenge [19,20].

Currently, many researchers have improved alfalfa's stress resistance by identifying stress-related genes and using transgenic molecular breeding methods. Tang et al. heterologously expressed a zinc-finger protein (Cys2/His2-type zinc-finger protein) from Glycine soja in alfalfa and found that the transgenic lines exhibited better salt and drought tolerance [21]. Zhang et al. introduced the *ScABI3* gene from *Syntrichia caninervis* into alfalfa and found that the transgenic alfalfa exhibited better salt tolerance by enhancing antioxidant enzyme activities and improving photosynthetic parameters [22]. Zhang et al. also identified the key gene *P5CDH1* from *C. songorica* and successfully developed transgenic *CsP5CDH1* alfalfa lines. Through drought stress tests, they found that the transgenic lines accumulated large amounts of proline under drought conditions, effectively reducing leaf water loss and enhancing alfalfa's drought tolerance [23]. Luo et al. developed *MsPPCK1* overexpression plants and demonstrated that the *MsPPCK1* gene could enhance alfalfa's alkaline tolerance and improve crop yield [24].

Northeast China is one of the key regions for the development of the animal husbandry industry and an important production base for alfalfa [25]. However, alfalfa breeding in the region faces practical challenges, such as the region's cold climate and saline–alkali soils. A report indicates that when selecting and breeding alfalfa varieties for cultivation in Northeast China, in addition to cold tolerance, it is highly desirable for the varieties to possess other stress-resistance traits, such as drought, salt, and alkali tolerance [26]. Based on this perspective, this study adopted a two-step variety screening process. First, 41 alfalfa varieties were pre-screened for cold tolerance. Then, 15 pre-screened varieties were evaluated for salt tolerance and clustered. Finally, one salt-tolerant alfalfa variety and one salt-sensitive alfalfa variety were selected for further analysis under salt stress. Growth and physiological parameters were measured, and transcriptomic analysis was conducted to reveal the salt-tolerance regulatory mechanisms in alfalfa with contrasting salt tolerance. Furthermore, weighted gene co-expression network analysis (WGCNA) was performed to identify salt-responsive hub genes. This study not only provides a theoretical foundation

for the study of alfalfa salt-tolerance mechanisms but also offers valuable insights for molecular breeding of salt-tolerant alfalfa in Northeast China.

2. Results

2.1. Results of Variety Screening

2.1.1. Growth and Cold Tolerance Analysis of Different Varieties under Cold Stress

In this study, we first performed cold tolerance screening on 41 alfalfa varieties, followed by salt tolerance screening on those with good cold tolerance. As shown in Tables S1–S5, we measured plant height, growth rate, fresh weight, increase in biomass, and dry weight of the 41 alfalfa varieties under cold stress (4 °C) and normal conditions (CK) after 10 days of treatment. To assess cold tolerance, we calculated the relative values of these indicators (Table S6). For a more comprehensive comparison of the cold tolerance of the 41 varieties, we analyzed the relative values of all indicators using grey correlation analysis. As a result, 15 alfalfa varieties were identified (Figure 1a). Correlation analysis between the unweighted correlation and the weighted correlation (Figure 1b) revealed an R^2 value of 0.99, indicating that both unweighted and weighted correlations can effectively explain the ranking, allowing for further experiments. The grey correlation rankings for all 41 varieties are listed in Table S7.

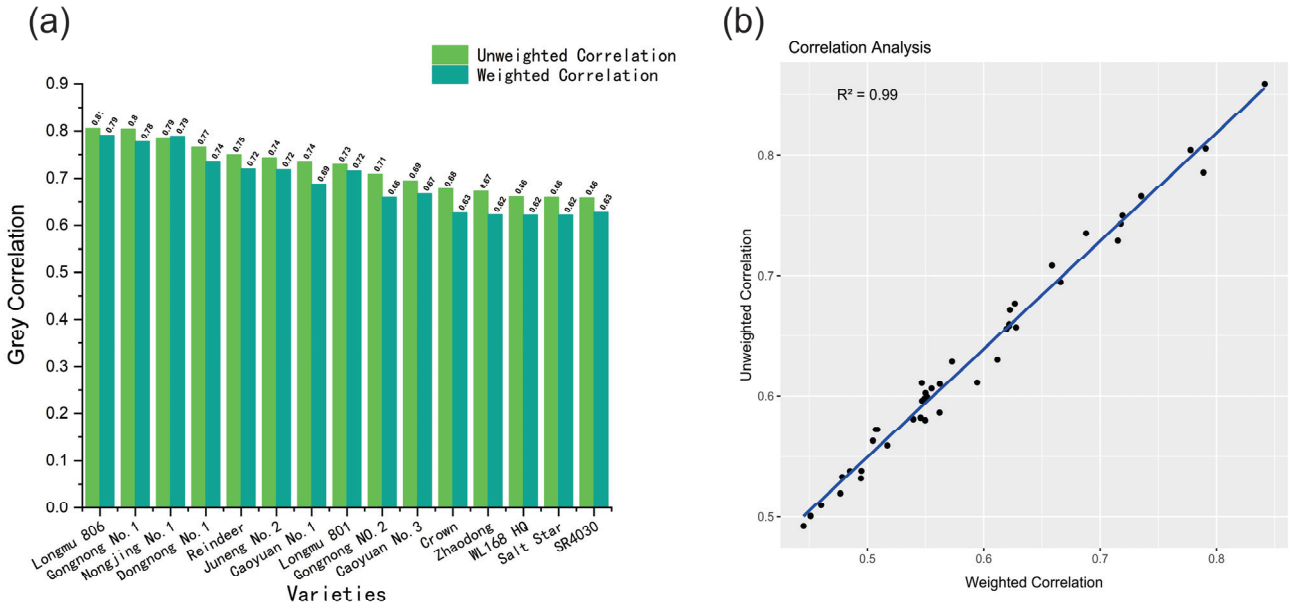

Figure 1. Cold tolerance analysis of alfalfa. (**a**) Grey correlation analysis of cold tolerance in different alfalfa varieties. (**b**) Correlation analysis between weighted correlation and unweighted correlation. Note: For ease of presentation, only the top 15 alfalfa varieties are shown in Figure 1.

2.1.2. Analysis of the Half-Lethal Concentration Under Salt Concentrations

In this study, the mortality rates of 15 alfalfa varieties under different NaCl concentrations (Figure S1) and their respective half-lethal concentrations (LC_{50}) (Figure 2) were measured. No mortality was observed at 0 mM and 100 mM NaCl concentrations for any of the varieties, while mortality occurred at both 200 mM and 300 mM NaCl, with a significantly higher rate at 300 mM compared to 200 mM NaCl (Figure S1). Notably, some alfalfa varieties exhibited very low mortality rates at 200 mM NaCl, ranging from 2% to 6%, including Longmu806 (4%), Reindeer (2%), Juneng No.2 (6%), Longmu801 (2%), and SR4030 (2%). The LC_{50} for each variety was calculated based on their mortality rates at 300 mM after 20 days, and the LC_{50} values ranged from 202.3 to 324.3 (Figure 2). The variety with the highest LC_{50} was SR4030, while the lowest was Gongnong No.2.

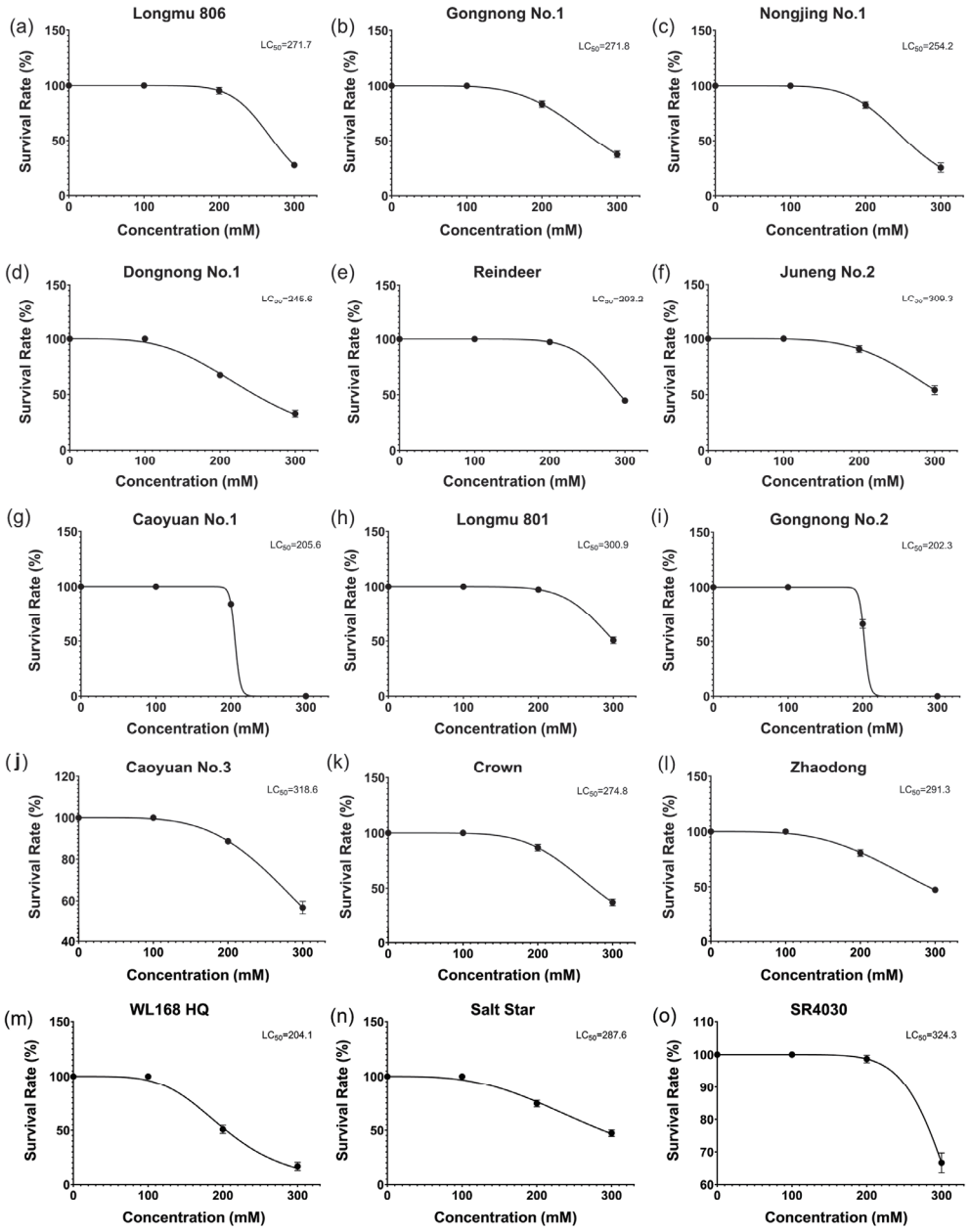

Figure 2. LC_{50} of different alfalfa varieties. (**a**–**o**) represent different alfalfa varieties.

2.1.3. Cluster Analysis of the Half-Lethal Concentration of Alfalfa

Figure 3a presents the LC_{50} values of the tested alfalfa varieties, which are ranked in descending order as follows: SR4030 > Caoyuan No.3 > Juneng No.2 > Longmu801 > Reindeer > Zhaodong > Salt Star > Crown > Gongnong No.1 > Longmu806 > Nongjing No.1 > Dongnong No.1 > Caoyuan No.1 > WL168 HQ (hereinafter referred to as WL168) > Gongnong No.2. K-means clustering analysis was performed based on LC_{50} values (Figure 3b), categorizing the 15 varieties into three groups: salt-tolerant, moderate, and salt-sensitive. Based on this classification, Longmu801 (salt-tolerant) and WL168 (salt-sensitive) were selected for further physiological and transcriptomic analysis.

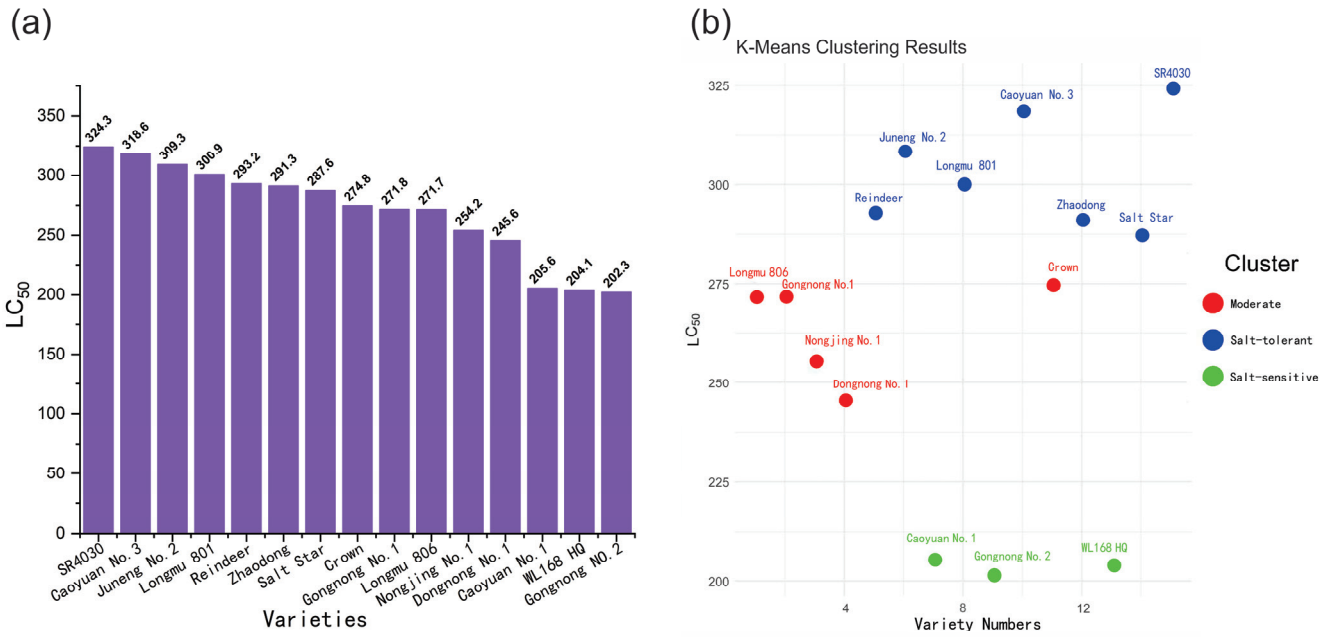

Figure 3. (**a**) LC_{50} ranking of different alfalfa varieties. (**b**) LC_{50} K-means clustering analysis.

2.2. Effects of Salt Stress on the Growth and Physiology of Alfalfa

2.2.1. Effects of Salt Stress on the Growth Indicators

With increasing stress duration, there was no significant difference in plant height between the treatment and control groups of Longmu801 and WL168 on day 3 (Figure 4a). However, by day 7, the plant height in both varieties under stress was significantly lower than that of the control groups, with reductions of 7.38% and 9.50%, respectively, compared to their controls. In Figure 4b, the fresh weight of both varieties decreased significantly compared to their respective controls as stress duration increased. On day 3, the fresh weight of the treatment group for Longmu801 and WL168 decreased by 12.83% and 16.36%, respectively, compared to the control. By day 7, the reductions were 25.30% and 32.24%, respectively. Figure 4c shows that the dry weight of both varieties significantly decreased compared to their controls as the stress duration increased. On day 3, the dry weight of Longmu801 and WL168 decreased by 13.14% and 11.65%, respectively, compared to the control, and by day 7, the reductions were 7.12% and 32.49%, respectively.

To better compare the response differences between Longmu801 and WL168 under salt stress, the relative values of each indicator were calculated as the ratio of the treatment group value to the control group value for the same indicator (i.e., "relative indicator = treatment group (indicator)/control group (indicator)"), to minimize the differences caused by the variety baseline. Figure 4d illustrates that relative plant height showed a decreasing trend as stress duration increased, but there was no significant difference between the two varieties at different time points. In Figure 4e, relative fresh weight also decreased with increasing stress duration. While no significant differences were observed on day 3, by day 7, Longmu801 had a significantly higher relative fresh weight than WL168, being 110.28% of WL168's relative fresh weight. Relative dry weight decreased with increasing stress duration but there were no significant differences between the two varieties at any time point (Figure 4f).

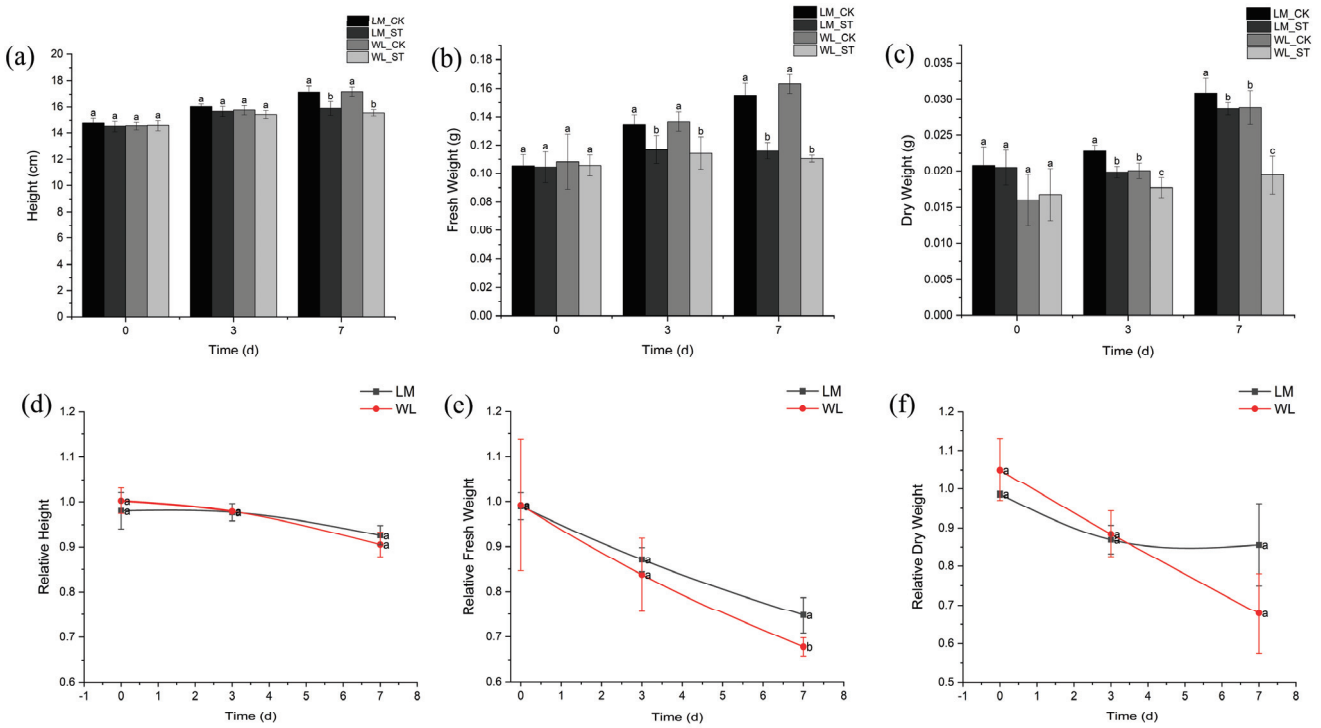

Figure 4. Selected growth indicators of alfalfa: (**a**) height; (**b**) fresh weight; (**c**) dry weight; (**d**) relative height; (**e**) relative fresh weight; and (**f**) relative dry weight. Note: For ease of data presentation and visualization, Longmu801 is abbreviated as LM and WL168 as WL in the Figures. Their respective treatment and control groups are labeled as LM_ST, LM_CK, WL_ST, and WL_CK. In (**a**–**c**), the lowercase letters indicate significant differences among the four different groups at the same time point. In (**d**–**f**), the lowercase letters indicate significant differences between the two varieties at the same time point.

2.2.2. Effects of Salt Stress on the Photosynthetic Pigments

The chlorophyll a content of Longmu801 and WL168 decreased as the duration of stress increased, with the treatment group showing significantly lower levels than the control group at all time points (Figure 5a). The most significant differences were observed on day 7, with reductions of 26.64% and 33.63%, respectively, compared to the control. Figure 5b illustrates a similar decreasing trend in chlorophyll b content, where the treatment groups were significantly lower than the controls at all time points. On day 7, the most significant reductions of 20.62% and 25.19% were observed. The total chlorophyll content also exhibited a decreasing trend over time, with significant reductions at each time point (Figure 5c). On day 7, the treatment groups were reduced by 24.80% and 30.91%, respectively, compared to the control. The relative chlorophyll a content of Longmu801 and WL168 decreased with the duration of stress (Figure 5d). From day 3 onward, a significant difference was observed between the two varieties, with Longmu801 showing 7.85%, 9.92%, and 10.58% higher relative chlorophyll a content than WL168 on days 3, 5, and 7, respectively. The relative chlorophyll b content of both varieties also decreased with the duration of stress (Figure 5e), but no significant differences between the varieties were observed at any time point. The relative total chlorophyll content of both varieties decreased over time, with significant differences appearing on days 5 and 7 (Figure 5f). On these days, Longmu801 had 8.13% and 8.95% higher relative total chlorophyll content than WL168, respectively.

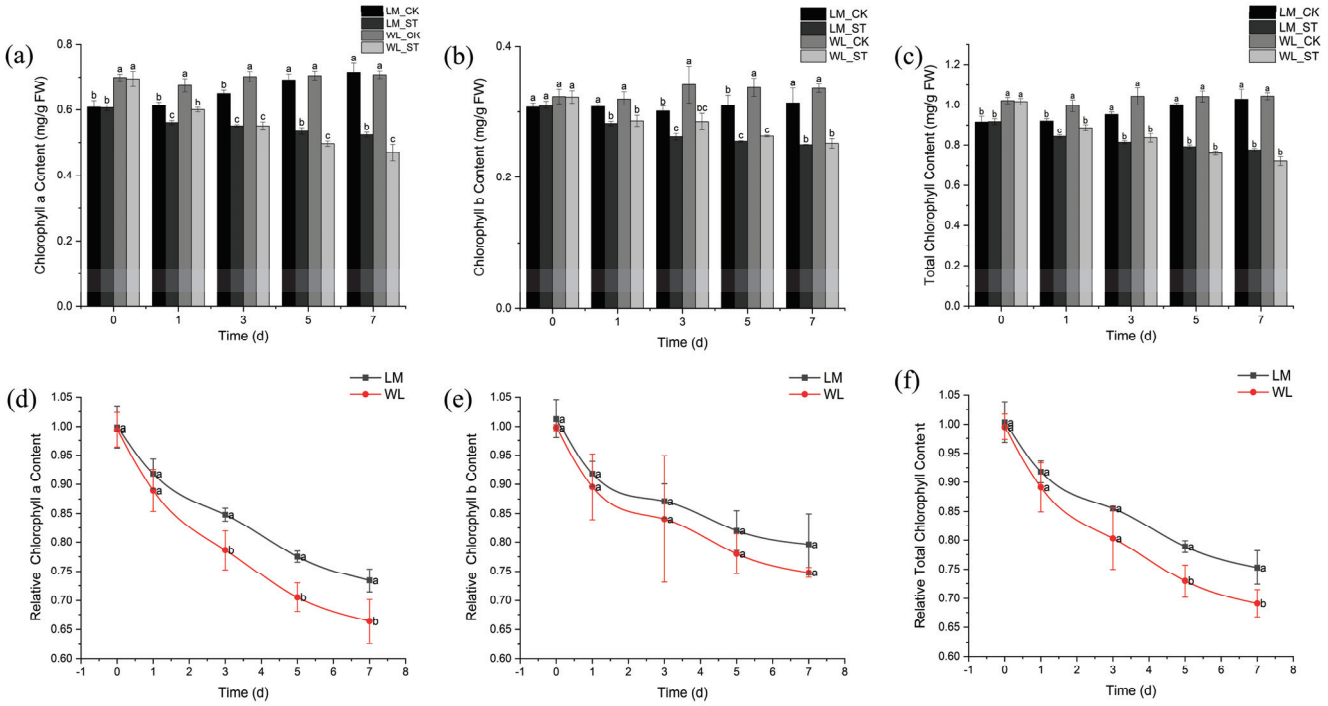

Figure 5. Selected photosynthetic pigment content of alfalfa: (**a**) chlorophyll a content; (**b**) chlorophyll b content; (**c**) total chlorophyll content; (**d**) relative chlorophyll a content; (**e**) relative chlorophyll b content; and (**f**) relative total chlorophyll content. In (**a**–**c**), the lowercase letters indicate significant differences among the four different groups at the same time point. In (**d**–**f**), the lowercase letters indicate significant differences between the two varieties at the same time point.

2.2.3. Effects of Salt Stress on Oxidative Stress Markers

The malondialdehyde (MDA) content of Longmu801 and WL168 increased as the duration of stress increased, with the treatment groups significantly higher than their respective controls at all time points (Figure 6a). The most significant difference was observed on day 7, with the MDA content of Longmu801 and WL168 increasing by 40.4% and 47.29%, respectively, compared to the controls. On days 1 and 7, WL168 had significantly higher MDA content than Longmu801. As shown in Figure 6b, the hydrogen peroxide (H_2O_2) content of both Longmu801 and WL168 also increased over time, with the treatment groups significantly higher than the controls at all time points. The most pronounced difference was on day 7, with the H_2O_2 content in the salt-treated Longmu801 and WL168 groups increasing by 26.92% and 39.03%, respectively, compared to their controls. Additionally, on days 1, 3, 5, and 7, WL168 had significantly higher H_2O_2 content than Longmu801. The relative MDA content of both varieties increased over time, with WL168 showing significantly higher values than Longmu801 on days 1 and 7, by 2.68% and 4.87%, respectively (Figure 6c). The relative H_2O_2 content of both varieties also increased over time, with WL168 showing a significantly higher relative H_2O_2 content than Longmu801 on day 7, with an increase of 9.55% (Figure 6d).

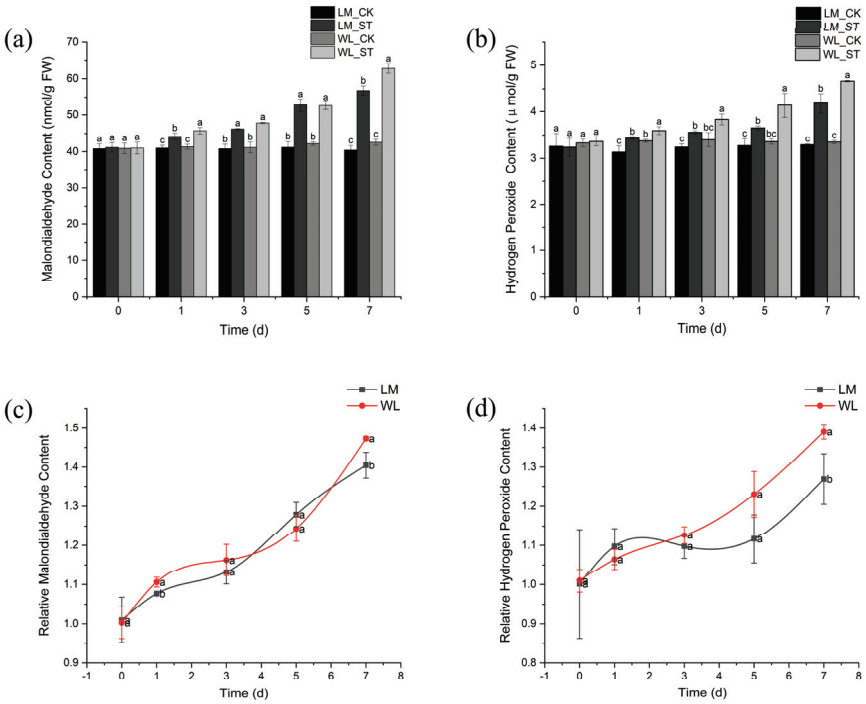

Figure 6. Selected oxidative stress marker content of alfalfa: (**a**) MDA content; (**b**) H_2O_2 content; (**c**) relative MDA content; and (**d**) relative hydrogen peroxide content. In (**a**,**b**), the lowercase letters indicate significant differences among the four different groups at the same time point. In (**c**,**d**), the lowercase letters indicate significant differences between the two varieties at the same time point.

2.2.4. Effects of Salt Stress on Osmolytes

The proline content of Longmu801 and WL168 increased with the duration of salt stress (Figure 7a). The proline content in the treatment groups was significantly higher than that in the control groups at all time points, with the most significant difference observed on day 7, where proline levels in the treatment groups of Longmu801 and WL168 were 19.35 and 14.67 times higher than their respective controls. On days 5 and 7, the proline content in Longmu801 was significantly higher than in WL168. The soluble sugar content also increased over time (Figure 7b). On day 1, the soluble sugar content in the WL168 salt treatment group was significantly higher than in the control group, while no significant change was observed in Longmu801. On days 3, 5, and 7, the soluble sugar content in the treatment groups of both varieties was significantly higher than in their respective controls, with the most significant difference on day 7, where the salt treatment groups of Longmu801 and WL168 showed increases of 40% and 25.87%, respectively, compared to their controls. The relative proline content of both varieties increased with the duration of stress (Figure 7c). On day 1, the relative proline content of Longmu801 was significantly higher than that of WL168, reaching a maximum on day 7, with an increase of 29.96% compared to WL168. The relative soluble sugar content also increased over time (Figure 7d). On days 5 and 7, the relative soluble sugar content of Longmu801 was significantly higher than that of WL168, with the most significant difference observed on day 5, showing an increase of 21.67%.

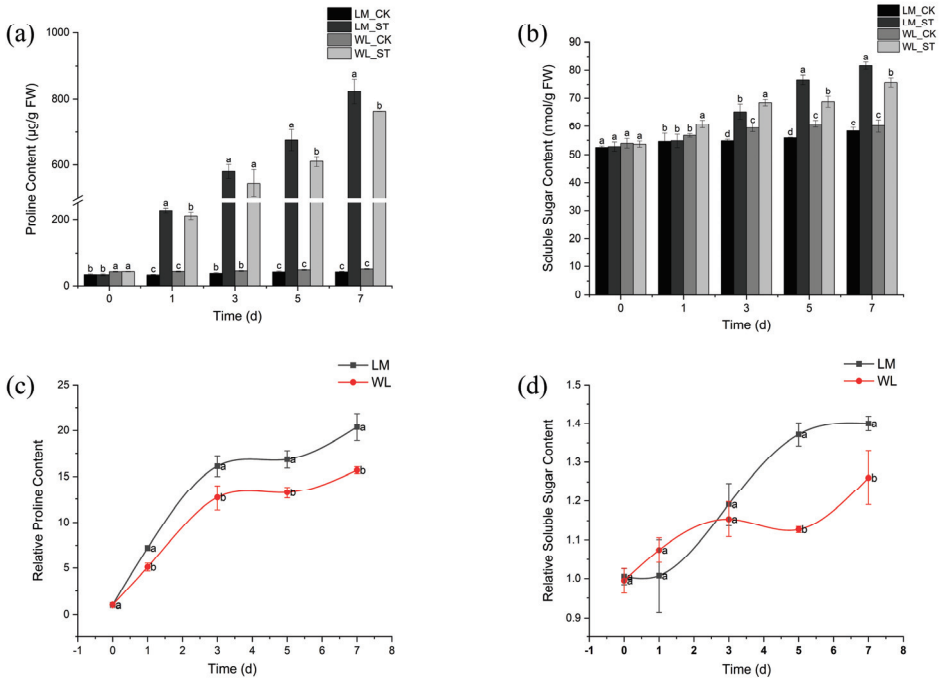

Figure 7. Selected osmolyte content of alfalfa: (**a**) proline content; (**b**) soluble sugar content; (**c**) relative proline content; and (**d**) relative soluble sugar content. In (**a**,**b**), the lowercase letters indicate significant differences among the four different groups at the same time point. In (**c**,**d**), the lowercase letters indicate significant differences between the two varieties at the same time.

2.2.5. Effects of Salt Stress on Antioxidant Enzymes

The superoxide dismutase (SOD) activity of Longmu801 and WL168 remained at a relatively high level as stress duration increased (Figure 8a). On day 1, the SOD activity of the treatment groups was higher than their respective controls, with Longmu801 increasing by 37.19% compared to the control. WL168 reached its peak SOD activity on day 3, increasing by 27.49% compared to the control, but showed a decreasing trend on days 5 and 7. Longmu801, on the other hand, maintained a consistent SOD activity throughout the time points, remaining significantly higher than the control group. The peroxidase (POD) activity of Longmu801 and WL168 initially increased and then decreased over time (Figure 8b). On day 1, the POD activity of the treatment groups was highest, with Longmu801 and WL168 increasing by 84.63% and 40.65%, respectively, compared to their controls. Both varieties showed a decreasing trend on days 3, 5, and 7, but their POD activity remained significantly higher than their respective controls. Figure 8c shows that catalase (CAT) activity remained at a high level as stress duration increased, with both Longmu801 and WL168 reaching their highest CAT activity on day 1, increasing by 61.68% and 26.8%, respectively, compared to the control. A decreasing trend followed, but CAT activity remained significantly higher than the control in both varieties. The relative SOD activity of both Longmu801 and WL168 was significantly higher than on day 0 (Figure 8d). On days 1, 5, and 7, Longmu801 had significantly higher relative SOD activity than WL168, while no significant difference was observed on day 3. Figure 8e shows a similar trend in relative POD activity, with Longmu801 showing a 31.36% higher relative POD activity than WL168 on day 1. No significant difference was observed on day 3, but Longmu801 had significantly higher POD activity on days 5 and 7. The relative CAT activity of both varieties was higher than on day 0, with Longmu801 showing significantly higher CAT activity than WL168 at all time points (Figure 8f).

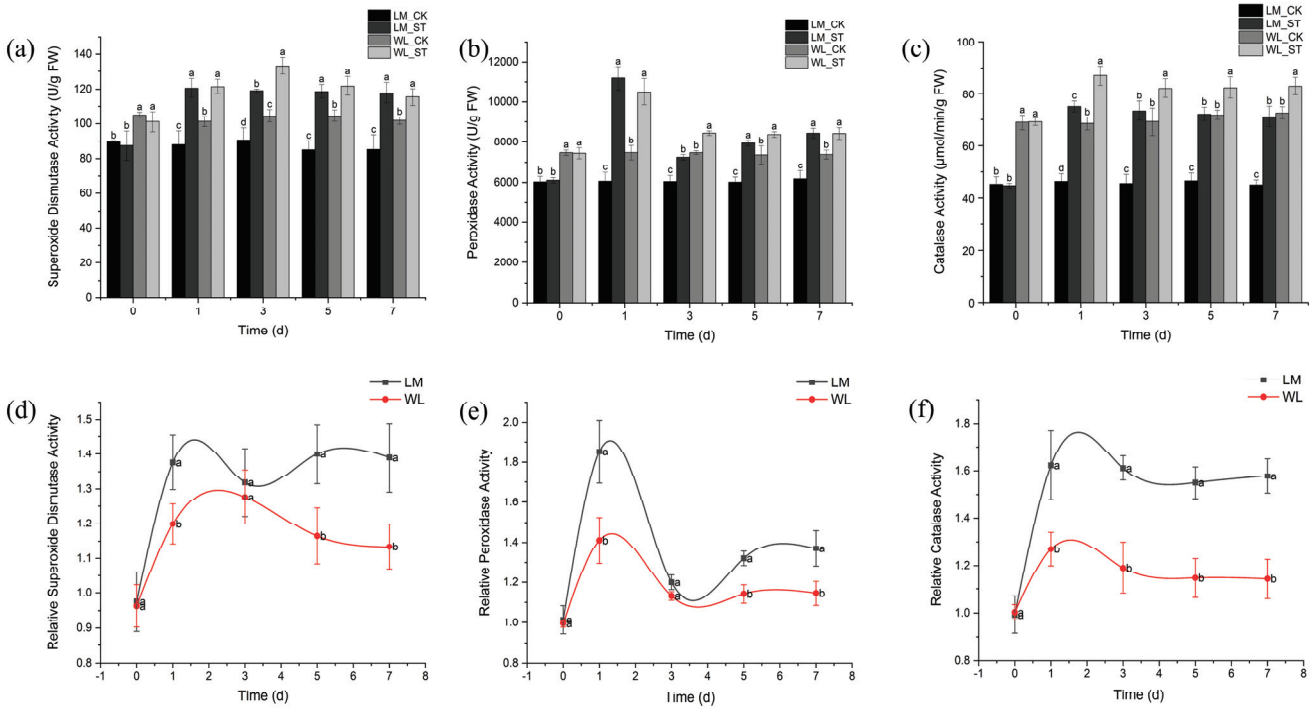

Figure 8. Selected antioxidant enzymes activity of alfalfa: (**a**) SOD activity; (**b**) POD activity; (**c**) CAT activity; (**d**) relative SOD activity; (**e**) relative POD activity; and (**f**) relative CAT activity. In (**a**–**c**), the lowercase letters indicate significant differences among the four different groups at the same time point. In (**d**–**f**), the lowercase letters indicate significant differences between the two varieties at the same time point.

2.2.6. Comprehensive Grey Correlation Analysis Based on Physiological Indicators

To determine the key time point for transcriptomic sequencing, this study conducted a grey correlation analysis of physiological indicators. First, the unweighted correlation ranking was obtained for each time point. The critic method was then used to calculate weighted values, and based on this, the weighted correlation was determined. The ranking for each time point is shown in Figure 9a. Next, Pearson correlation analysis was performed between the equal-weight and weighted grey correlation degrees, with the results showing a low correlation (R^2 = 0.23) in Figure 9b, indicating that the weight values had a significant impact on the grey correlation. Therefore, data from both methods were considered in decision-making. Figure 9c,d show the differences between the unweighted and weighted correlation at different time points. The largest difference was observed on day 7 (0.163514 and 0.303217), while the smallest difference was on day 0. After comprehensive consideration, day 7 was chosen as the sampling time point for transcriptomic sequencing.

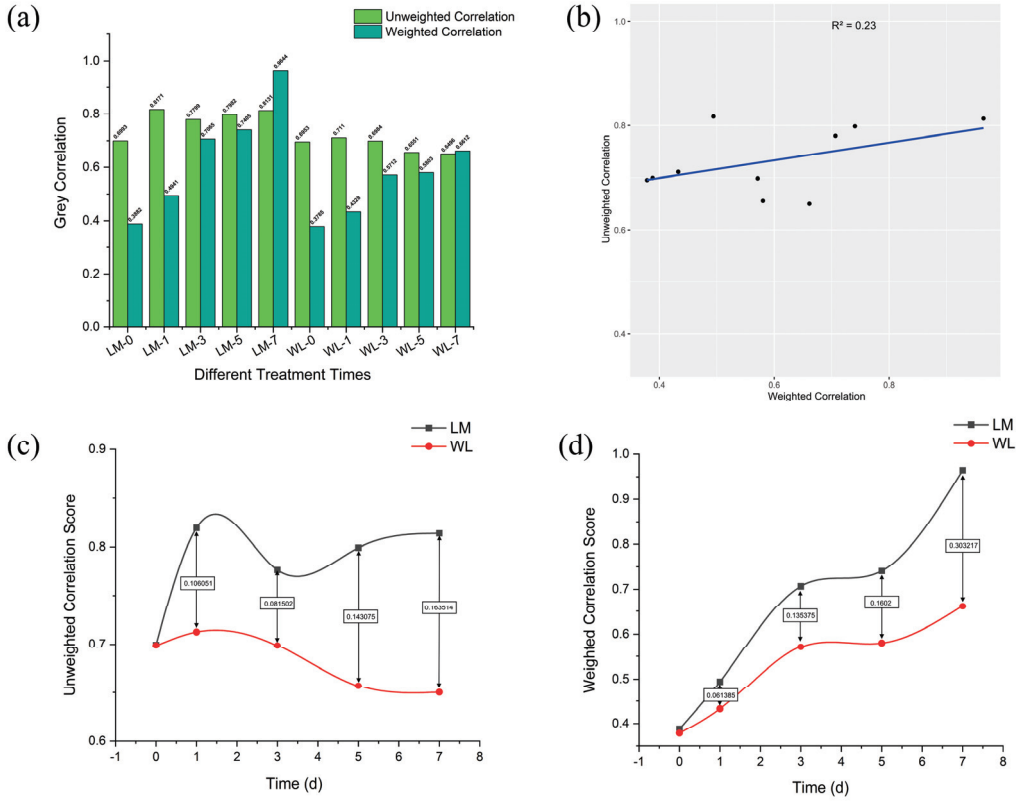

Figure 9. (**a**) Grey correlation of different time points. (**b**) Correlation between unweighted and weighted correlation degrees. (**c**) Unweighted correlation at different time points. (**d**) Weighted correlation at different time points. Note: In (**a**), LM-0,1,3,5,7 and WL-0,1,3,5,7 represent the different time points for the two varieties.

2.3. Transcriptomic Analysis of Different Alfalfa Varieties Under Salt Stress

2.3.1. Overview of Transcriptomic Sequencing Statistics

First, cDNA libraries were constructed for the experimental and control groups of different varieties under salt stress, and transcriptomic sequencing was performed using the Illumina Hiseq 4000 platform, with RNA-seq used to obtain transcriptomic data (details of the specific methods can be found in Section 4.4). The total number of raw reads ranged from 35,512,472 to 43,229,200, and the sequencing data size ranged from 5.33 G to 6.48 G (Table S8). After filtering, the number of valid reads ranged from 33,821,714 to 41,592,130, with a sequencing data size between 5.07 G and 6.24 G. The valid ratio was between 94.24% and 96.91%, Q20 and Q30 values were 99.20% to 99.44% and 96.59% to 97.50%, respectively, and the GC content of the samples ranged from 41.5% to 42.50%, with an average GC content of 42%.

The valid data were aligned to the reference genome using Hisat (details of the specific methods can be found in Section 4.4) after preprocessing. First, all available alfalfa reference genomes were used for alignment. The average alignment rate with the "Zhongmu No. 4" genome was slightly higher than that of the "Xinjiang Daye" genome at 87.17% and significantly higher than the "Zhongmu No. 1" genome at 71.32% (Table S9) [27]. Therefore, the "Zhongmu No. 4" alfalfa reference genome was selected for further analysis. As shown in Table S10, by aligning the valid reads to the reference genome, the genomic alignment for each sample was obtained, allowing the identification of valid reads' positions on the reference genome as well as specific sequence information. The aligned reads (mapped reads) ranged from 30,258,802 to 34,413,892, with an alignment rate between 84.21% and 89.94% and an average alignment rate of 87.38%. An alignment rate greater than 80% indicates that the selected reference genome is appropriate and the alignment was effective. Of these, the proportion of uniquely mapped reads (unique mapped) ranged from 36.98%

to 39.54%, the proportion of paired-end mapped reads (PE mapped) ranged from 72.21% to 78.28%, non-spliced reads ranged from 22.46% to 25.04%, and spliced reads ranged from 13.81% to 15%.

For ease of data processing, Longmu801 is abbreviated as LM, and WL168 is abbreviated as WL in the transcriptomic analysis. Their respective treatment and control groups are LM_ST, LM_CK, WL_ST, and WL_CK. The gene expression levels of the samples in the different treatment groups are represented by darker colors, with correlation coefficients close to 1, indicating a high correlation among the samples within each group (Figure 10a). From Figure 10b, it can be observed that in the Principal Component Analysis (PCA), there is a larger degree of dispersion between different varieties, such as LM_CK and WL_CK, compared to LM_ST and WL_ST, which is due to the different experimental treatment conditions. The relatively smaller dispersion between LM_CK and WL_CK, as well as between LM_ST and WL_ST.

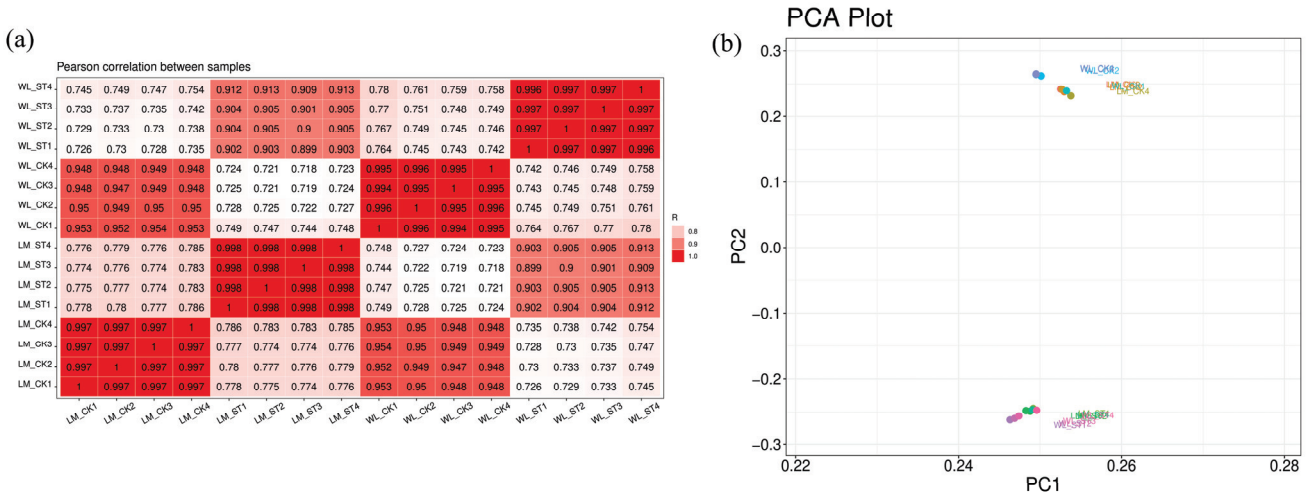

Figure 10. (**a**) Correlation analysis between samples and (**b**) principal component analysis (PCA) between samples.

2.3.2. Transcriptomic Analysis of Longmu801 and WL168 Under Salt Stress

There were 16,485 differentially expressed genes (DEGs) in the LM_ST vs. LM_CK group, of which 5929 DEGs were upregulated and 10,556 DEGs were downregulated (Figure 11a). In the group, a total of 14,350 DEGs were annotated with GO functions, which were enriched in 3568 GO functions. These functions are classified into three categories: 1917 in Biological Process, 438 in Cellular Component, and 1213 in Molecular Function. A total of 799 GO functions were significantly enriched ($p < 0.05$). The GO functions based on *p*-values, including kinesin complex (GO:0005871), microtubule-based movement (GO:0007018), plant-type secondary cell wall biogenesis (GO:0009834), extracellular region (GO:0005576), microtubule motor activity (GO:0003777), and cell wall (GO:0005618), among others (Figure 11b). A total of 5396 DEGs were annotated into 135 KEGG pathways. These 135 pathways are classified into six primary metabolic categories: 4 in Environmental Information Processing, 4 in Cellular Processes, 21 in Genetic Information Processing, 2 in Human Diseases, 102 in Metabolism, and 2 in Organismal Systems. A total of 37 KEGG pathways were significantly enriched ($p < 0.05$). The pathways based on *p*-values are displayed, including the KOK signaling pathway—plant (ko04016), ether lipid metabolism (ko00565), glutathione metabolism (ko00480), sphingolipid metabolism (ko00600), plant–pathogen interaction (ko04626), and phenylpropanoid biosynthesis (ko00940), among others (Figure 11c).

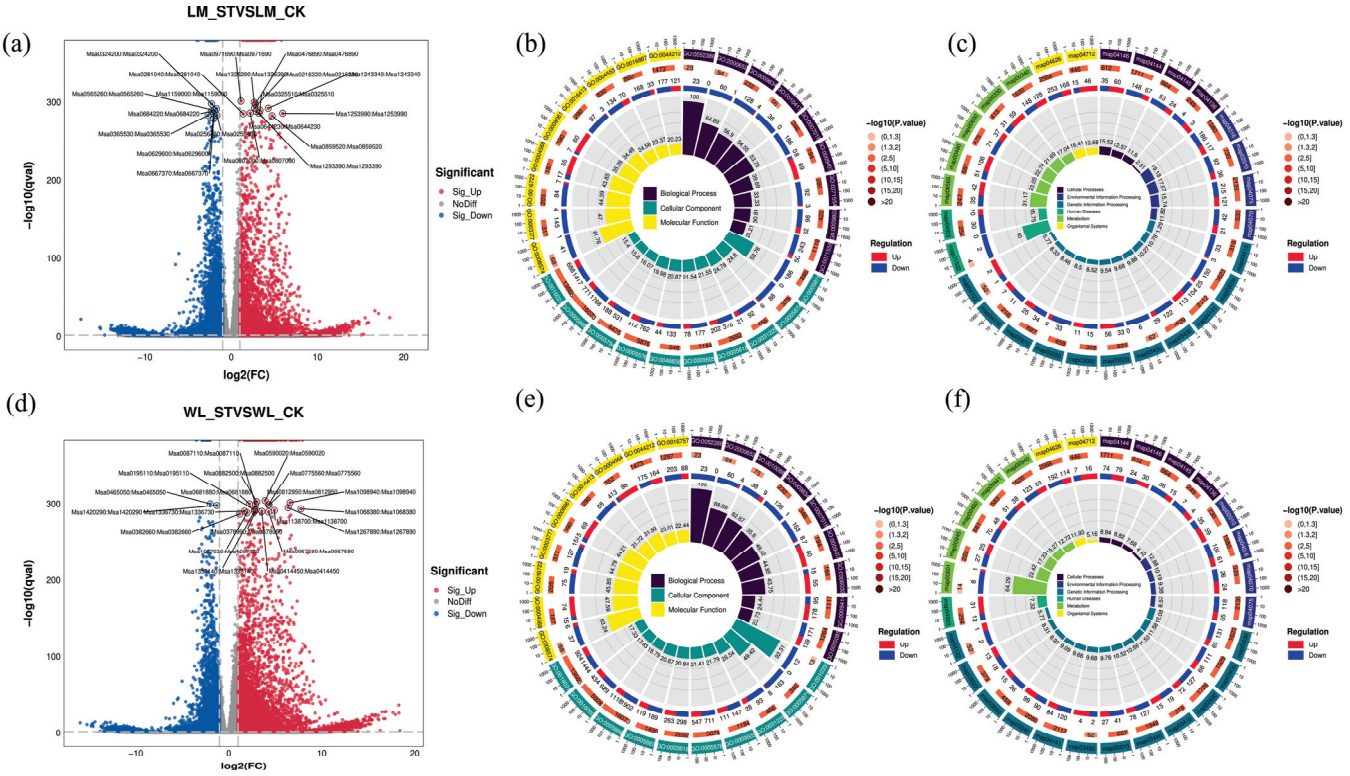

Figure 11. Transcriptomic Analysis of Longmu801 and WL168 under Salt Stress. (**a**) Differentially Expressed Genes in Longmu 801. (**b**) GO Enrichment Function of Longmu801. (**c**) KEGG Pathways of Longmu801. (**d**) Differentially Expressed Genes in WL168. (**e**) GO Enrichment Function of WL168. (**f**) KEGG Pathways of WL168.

Figure 11d illustrates that the WL_ST vs. WL_CK group contained a total of 18,726 DEGs, with 7798 showing upregulation and 10,928 exhibiting downregulation. In the group, a total of 16,307 DEGs were annotated with GO functions, enriching 3817 GO entries. These entries are classified into three categories: 2052 in Biological Process, 473 in Cellular Component, and 1292 in Molecular Function. A total of 782 GO functions were significantly enriched ($p < 0.05$). The GO functions based on p-values, including plasma membrane (GO:0005886), kinesin complex (GO:0005871), microtubule-based movement (GO:0007018), plant-type secondary cell wall biogenesis (GO:0009834), extracellular region (GO:0005576), and microtubule motor activity (GO:0003777), among others (Figure 11e). A total of 6043 DEGs were annotated into 133 KEGG pathways. These 133 pathways are classified into six primary metabolic categories: 4 in Environmental Information Processing, 4 in Cellular Processes, 21 in Genetic Information Processing, 2 in Human Diseases, 100 in Metabolism, and 2 in Organismal Systems. A total of 39 KEGG pathways were significantly enriched ($p < 0.05$). KEGG pathways based on p-values are displayed, including ether lipid metabolism (ko00565), glutathione metabolism (ko00480), ABC transporters (ko02010), arginine and proline metabolism (ko00330), glycerolipid metabolism (ko00561), and isoflavonoid biosynthesis (ko00943), among others (Figure 11e).

2.3.3. Analysis of DEGs Between Longmu801 and WL168

To better explore the differences in response to salt stress between the varieties, Venn diagrams were generated for the different comparison groups, as shown in Figure 12a. Among the groups LM_ST vs. LM_CK, LM_ST vs. WL_ST, and WL_ST vs. WL_CK, a total of 2164 DEGs were common across all comparisons. These 2164 DEGs represent genes that respond to salt stress in both varieties. There are 1850 DEGs that are shared between LM_ST vs. LM_CK and LM_ST vs. WL_ST but are absent in WL_ST vs. WL_CK, indicating that these 1850 DEGs are specifically responsive to salt stress in the LM variety but not in

the WL variety. The expression of these genes in LM is significantly different from that in WL, suggesting that these 1850 DEGs are unique to the LM variety's response to salt stress. Additionally, there are 2679 DEGs that are shared between WL_ST vs. WL_CK and LM_ST vs. WL_ST but are absent in LM_ST vs. LM_CK. These 2679 DEGs are specifically responsive to salt stress in the WL variety but not in the LM variety, and their expression levels in WL are significantly different from those in LM, indicating that these 2679 DEGs are unique to the WL variety's response to salt stress.

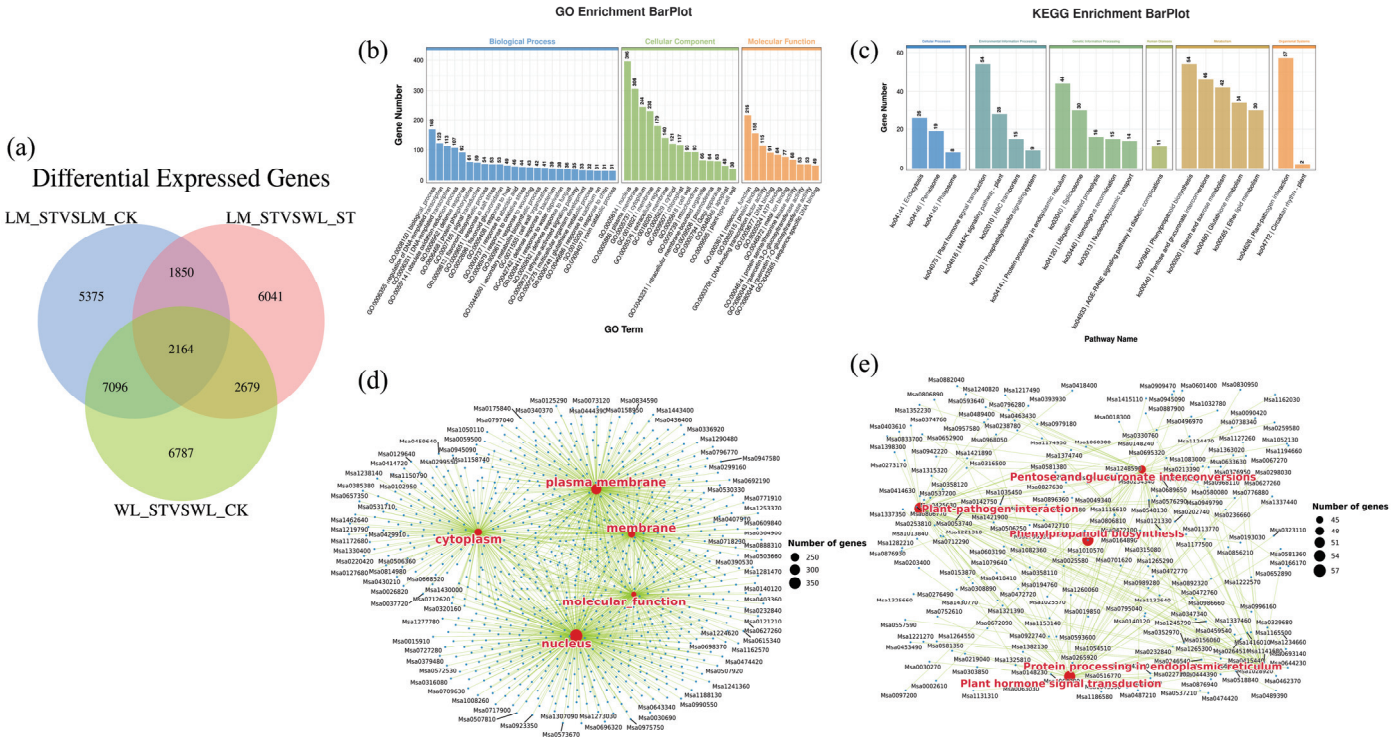

Figure 12. Analysis of DEGs between Longmu801 and WL168. (**a**) DEGs Venn diagrams. (**b**) GO functions of 2164 DEGs. (**c**) KEGG pathways of 2164 DEGs. (**d**) The relationship between GO functions and DEGs. (**e**) The relationship between KEGG pathways and DEGs.

As shown in Figure 12b, the 2164 shared DEGs were enriched in 1291 GO functions, which can be categorized into three groups: 689 in Biological Process, 150 in Cellular Component, and 452 in Molecular Function. Among these, 27 GO functions were significantly enriched ($p < 0.05$). The GO functions under the Biological Process category include: biological process (GO:0008150), regulation of DNA-templated transcription (GO:0006355), DNA-templated transcription (GO:0006351), obsolete oxidation-reduction process (GO:0055114), defense response (GO:0006952), and others. Under the Cellular Component category, the GO functions include: nucleus (GO:0005634), plasma membrane (GO:0005886), cytoplasm (GO:0005737), membrane (GO:0016021), and extracellular region (GO:0005576), among others. Under the Molecular Function category, the GO functions include: molecular function (GO:0003674), protein binding (GO:0005515), DNA-binding transcription factor activity (GO:0003700), DNA binding (GO:0003677), and ATP binding (GO:0005524), among others. For the relationship between the GO functions and DEGs, the top five GO functions with a p-value < 0.05 were selected for display (Figure 12d).

As shown in Figure 11c, these DEGs were enriched in 111 KEGG pathways, of which 85 were significantly enriched ($p < 0.05$). These include plant–pathogen interaction (ko04626); plant hormone signal transduction (ko04075); phenylpropanoid biosynthesis (ko00940); pentose and glucuronate interconversions (ko00040), folding, sorting, and degradation in the endoplasmic reticulum (ko04141); starch and sucrose metabolism (ko00500); glutathione metabolism (ko00480); spliceosome (ko03040); ether lipid metabolism (ko00565);

and glycerolipid metabolism (ko00561), among others. The relationship between the KEGG pathways and DEGs, the top five GO functions with a p-value < 0.05 were selected for display (Figure 12e). The GO function enrichment and KEGG pathway enrichment for the 1850 DEGs and 2679 DEGs are shown in the Supplementary Figures (Figures S2 and S3).

2.3.4. Results of qRT-PCR Validation

In the transcriptomic analysis, we identified a class of genes referred to as "abscisic acid and environmental stress-inducible proteins", which exhibited significant changes following salt treatment. These genes, *Msa0819320*, *Msa0738880*, *Msa0783290*, *Msa0819370*, and *Msa0783280*, were selected as candidate genes for further validation using qPCR technology (Figure S4a). A correlation analysis between the transcriptomic (RNA-seq) results and qPCR was conducted, with an R^2 value of 0.99, indicating that the transcriptomic results are reliable (Figure S4b).

2.4. Weighted Gene Co-Expression Network Analysis

2.4.1. Sample Processing and WGCNA Module Correlation Analysis

To discover additional core genes involved in the response to salt stress, we conducted a weighted gene co-expression network analysis (WGCNA). First, genes from the transcriptomic analysis with an expression level below 2 were filtered out, resulting in 32,684 genes being included in the WGCNA. A sample clustering dendrogram was then constructed from the expression levels of these genes across 16 samples to assess for any potential outliers. All samples from the different groups clustered well, indicating no outliers, and thus, no samples needed to be removed (Figure 13a). In this WGCNA analysis, the plateau threshold was set to 0.85, and the soft threshold was set to 1. In Figure 13b, the x-axis represents the soft threshold, and the y-axis represents the evaluation parameter of the scale-free network, which indicates the squared correlation coefficient of the corresponding network. In Figure 13c, the x-axis represents the soft threshold, and the y-axis represents the average connectivity (adjacency function) of all genes under different β values.

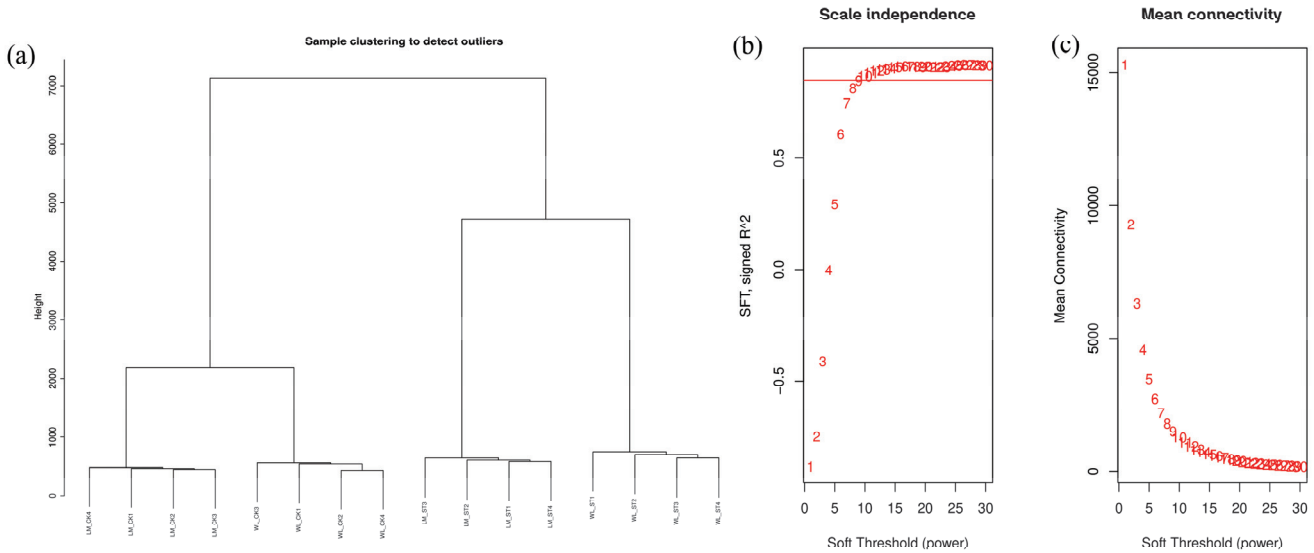

Figure 13. (**a**) Clustering analysis of different samples. (**b**,**c**) Soft threshold and gene connectivity.

To increase biological relevance, the topological overlap measure (TOM) was used to calculate the degree of association between genes. In addition to analyzing the relationship between two genes, the connections between these two genes and other genes were also considered. The minimum module size for module detection was set to 30. In this WGCNA analysis, genes were clustered into eight modules: green, yellow, red, blue, turquoise, black, brown, and grey (Figure 14a). The grey module contains all unclustered genes and thus

holds little analytical significance, so it was not analyzed in this study. Figure 14b shows the correlation analysis between the genes within each module. Generally, genes within the same module are represented by deeper colors, while the color between modules is lighter. The consistent color representation of genes within each module indicates good clustering, offering greater research value and potential insights.

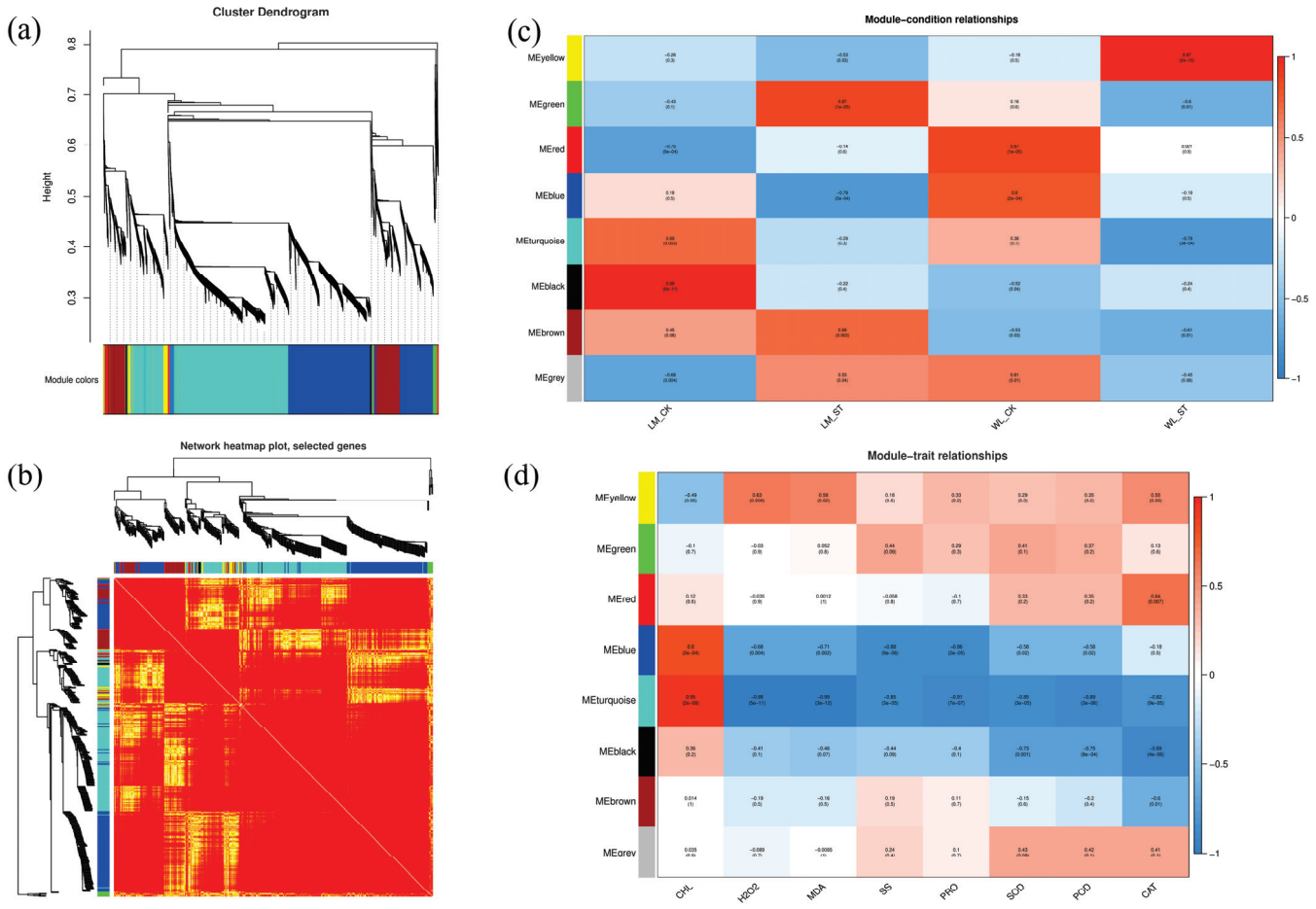

Figure 14. Module correlation analysis. (**a**) Module hierarchical clustering dendrogram. (**b**) Module correlation. (**c**) Module–sample correlation. (**d**) Module–physiological correlation.

A correlation analysis was conducted between different modules, experimental groups, and physiological indicators. Figure 14c illustrates the correlation between different modules and experimental groups. The color variations across modules and groups reflect the strength of these correlations, enabling a visual assessment of gene associations within each module. The values in parentheses denote *p*-values, with smaller values indicating higher statistical significance. Figure 14d displays the correlation between different modules and physiological indicators. These correlations divide the modules into two distinct groups: one group, including the yellow, green, and red modules, shows positive correlations with indicators such as soluble sugars (SS), proline (PRO), superoxide dismutase (SOD), peroxidase (POD), and catalase (CAT). The other group, comprising the blue, turquoise, and black modules, is negatively correlated with hydrogen peroxide (H2O2) and malondialdehyde (MDA). Further analysis will focus on these six modules.

2.4.2. Study of Modules Gene Functions and Hub Gene Screening

Figure S5 provides an overview of the GO functions and KEGG pathways enriched in genes from six distinct color modules. The yellow module is significantly enriched for 331 GO functions, including cytoplasm (GO:0005737), chloroplast (GO:0009507), cytosol

(GO:0005829), response to cold (GO:0009409), and protein transport (GO:0015031). The yellow module is also significantly enriched for 51 KEGG pathways, including starch and sucrose metabolism (ko00500), protein export (ko03060), phenylpropanoid biosynthesis (ko00940), and protein processing in the endoplasmic reticulum (ko04141). The hub genes identified in this module are: *Msa0835190*, *Msa0644060*, *Msa0559800*, *Msa1079730*, and *Msa0192380* (Figure 15a).

The green module is significantly enriched for 291 GO functions, including cytosol (GO:0005829), metal ion binding (GO:0046872), chloroplast thylakoid membrane (GO:0009535), plastoglobuli (GO:0010287), P-body (GO:0000932), and photosynthesis (GO:0009768). The green module is also significantly enriched for 49 KEGG pathways, including propanoate metabolism (ko00640), photosynthesis—antenna proteins (ko00196), glycolysis/gluconeogenesis (ko00010), and valine, leucine, and isoleucine degradation (ko00280). Figure 15b presents the network diagram of hub genes in the green module, with the selected hub genes being: *Msa1168040*, *Msa0367140*, *Msa0458690*, *Msa0066410*, and *Msa125874*.

The red module is significantly enriched for 215 GO functions, including cytosol (GO:0005829), cytosolic large ribosomal subunit (GO:0022625), nucleus (GO:0005634), mitochondrion (GO:0005739), and plastid (GO:0009536). The red module is also significantly enriched for 38 KEGG pathways, including glycolysis/gluconeogenesis (ko00010), phenylpropanoid biosynthesis (ko00940), fatty acid elongation (ko00062), oxidative phosphorylation (ko00190), and tryptophan metabolism (ko00380). Figure 15c presents the network diagram of hub genes in the red module, with the selected hub genes being: *Msa0489620*, *Msa1242210*, *Msa0776160*, *Msa1078300*, and *Msa0397400*.

The GO enrichment analysis for genes in the blue module revealed 1171 significantly enriched GO functions, including catalase activity (GO:0004096), xanthophyll binding (GO:0051738), glyceraldehyde-3-phosphate dehydrogenase (NADP+) (phosphorylating) activity (GO:0047100), peroxiredoxin activity (GO:0051920), and antioxidant activity (GO:0016209). The KEGG enrichment analysis of the blue module identified 115 significantly enriched pathways, including plant hormone signal transduction (ko04075), ribosome (ko03010), protein processing in the endoplasmic reticulum (ko04141), nucleocytoplasmic transport (ko03013), and endocytosis (ko04144). As shown in Figure 15d, genes in the blue module were clustered into two groups, with the second cluster having the highest degree values. The top five genes based on degree ranking were selected as hub genes: *Msa0933220*, *Msa0850110*, *Msa0605650*, *Msa0257490*, and *Msa0218780*.

The GO enrichment analysis for genes in the turquoise module revealed 1306 significantly enriched functions, including nucleus (GO:0005634), cytoplasm (GO:0005737), plasma membrane (GO:0005886), protein binding (GO:0005515), membrane (GO:0016021), and biological process (GO:0008150). The KEGG enrichment analysis of the turquoise module identified 113 significantly enriched pathways, including plant hormone signal transduction (ko04075), ribosome (ko03010), amino sugar and nucleotide sugar metabolism (ko00520), phagosome (ko04145), starch and sucrose metabolism (ko00500), and MAPK signaling pathway—plant (ko04016). Figure 15e presents the network diagram of hub genes in the turquoise module, with the top five genes based on degree ranking selected as hub genes: *Msa0631990*, *Msa1335970*, *Msa1031990*, *Msa0564100*, and *Msa0242790*.

The GO enrichment analysis for genes in the black module revealed 192 significantly enriched functions, including inorganic triphosphate phosphatase activity (GO:0050355), structural constituent of ribosome (GO:0003735), translation (GO:0006412), and cellular response to singlet oxygen (GO:0071452). The KEGG enrichment analysis of the black module identified 35 significantly enriched pathways, including ribosome (ko03010), spliceosome (ko03040), purine metabolism (ko00230), pyrimidine metabolism (ko00240), photosynthesis (ko00195), and flavonoid biosynthesis (ko00941). Figure 15f presents the network diagram of hub genes in the black module, with the top five genes based on degree ranking selected as hub genes: *Msa1242220*, *Msa0958830*, *Msa0168360*, *Msa1314080*, and *Msa0761640*.

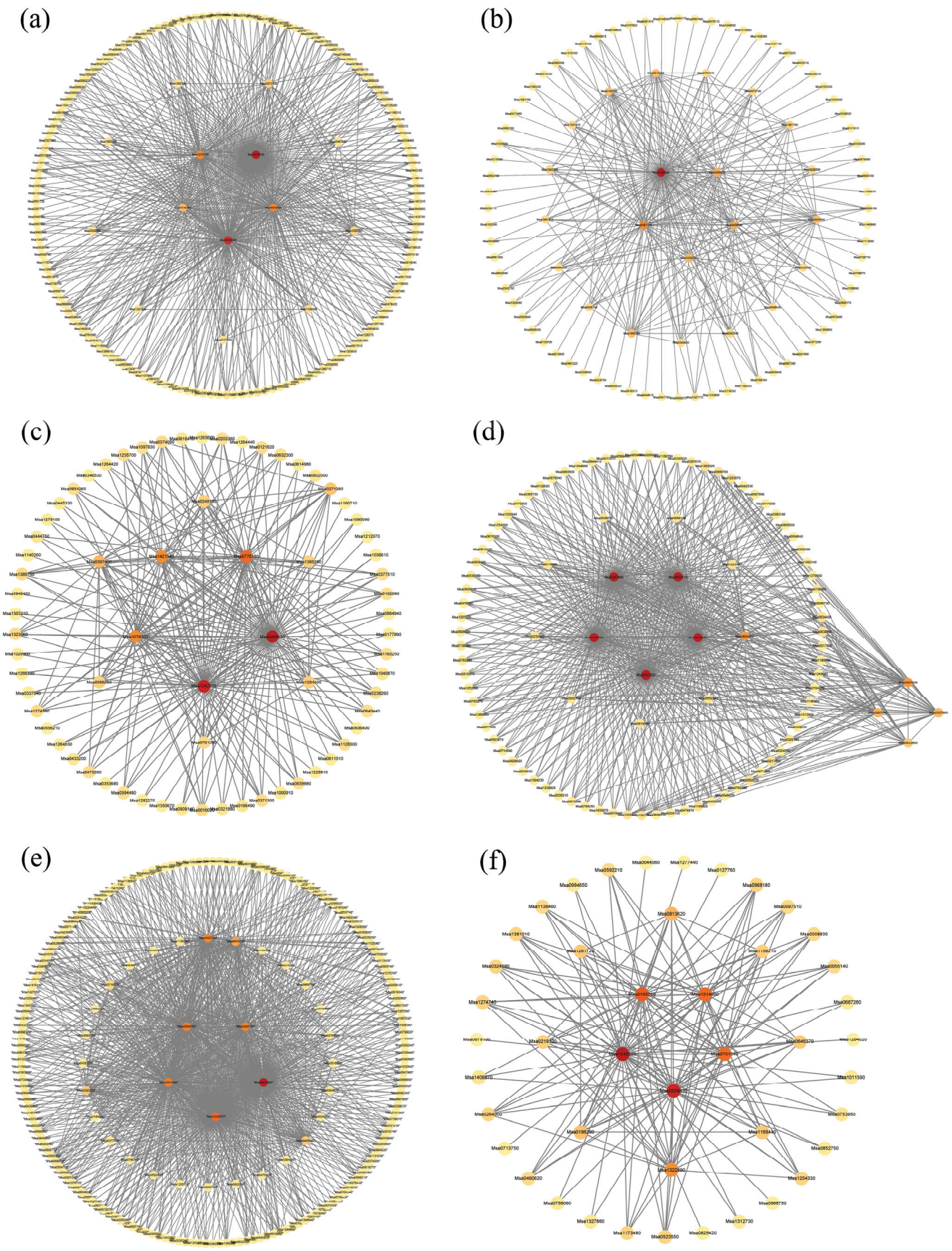

Figure 15. Hub genes of different modules. (**a**) Hub genes in yellow modules. (**b**) Hub genes in green modules. (**c**) Hub genes in red modules. (**d**) Hub genes in blue modules. (**e**) Hub genes in turquoise

modules. (**f**) Hub genes in black modules. Note: The hub genes were clustered using the Mcode plugin in Cytoscape (Details of the specific methods can be found in Section 4.6.), and the top five genes with the highest degree values within each cluster were selected. The darker the color, the higher the degree value.

2.4.3. Correlation Analysis of HUB Genes

The correlation analysis of the 35 hub genes identified from the six different WGCNA color modules and those annotated as "abscisic acid and environmental stress-inducible proteins" in the transcriptomic analysis revealed that the genes *Msa0218780* and *Msa0605650* have the highest degree values, with 22 hub genes co-expressed with these two genes. In contrast, *Msa1242210*, *Msa1078300*, and *Msa1314080* have the lowest degree values, with only 10 hub genes co-expressed (Figure 16a). The blue module exhibited the highest co-expression connectivity, while the yellow and black modules had lower connectivity. Figure 16b,c show the GO and KEGG enrichment analysis of the selected hub genes. The GO term "response to salt stress" (GO:0009651) was enriched, including five hub genes: *Msa085011*, *Msa0605650*, *Msa0397400*, *Msa1258740*, and *Msa0958830*.

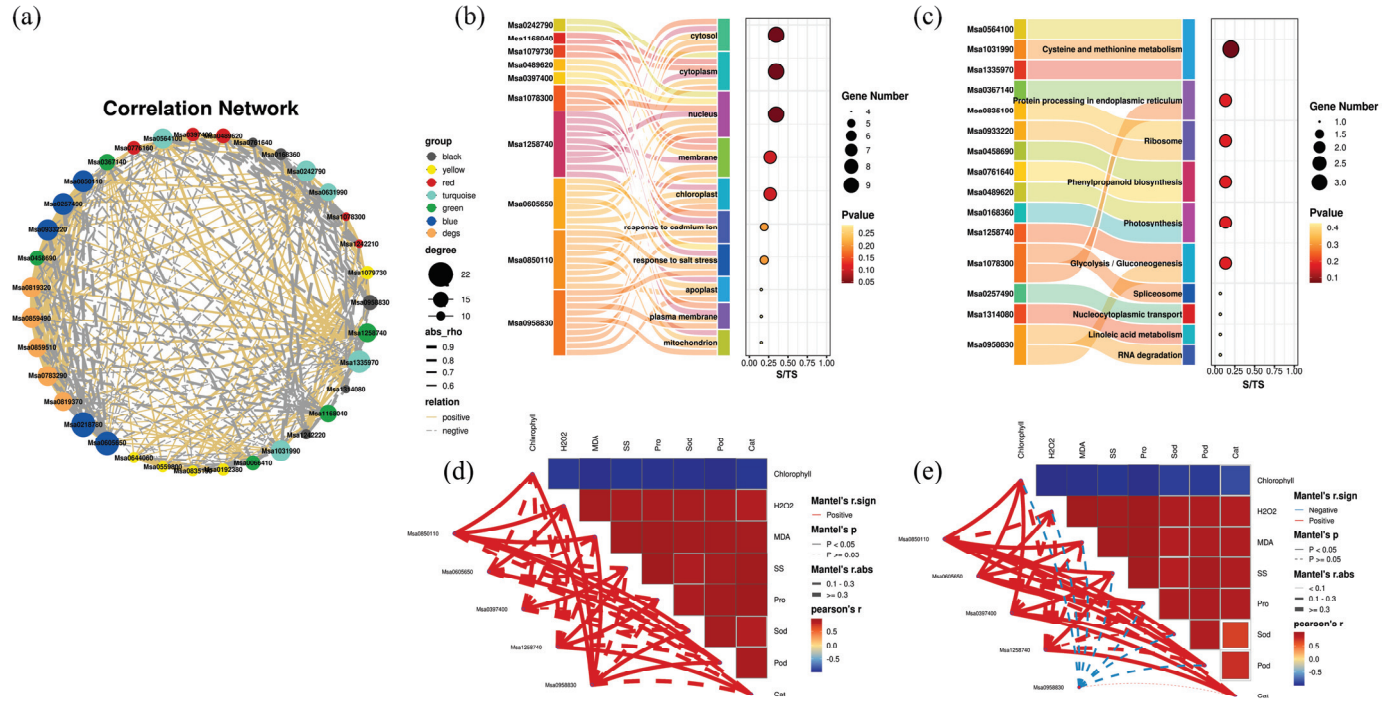

Figure 16. Analysis of HUB Genes. (**a**) Co-expression network relationship of hub genes. (**b**) GO enrichment Sankey diagram of hub genes. (**c**) KEGG enrichment Sankey diagram of hub genes. (**d**) Relationship between hub genes responding to salt stress and physiological indicators in Longmu801. (**e**) Relationship between hub genes responding to salt stress and physiological indicators in WL168.

A Mantel test was conducted to analyze the correlation between these five hub genes and physiological indicators (Figure 16d,e). In Longmu801, H_2O_2, MDA, SS, Pro, Sod, Pod, and Cat were negatively correlated with chlorophyll content. All five hub genes showed a positive correlation with the physiological indicators. In WL168, the same physiological indicators were negatively correlated with chlorophyll content. Specifically, *Msa0958830* was negatively correlated with H_2O_2, MDA, SS, Pro, Sod, and Pod but positively correlated with Cat, although this correlation was weak. The remaining hub genes showed positive correlations with the physiological indicators.

3. Discussion

3.1. Interpretation of Variety Screening Outcomes

In this study, we adopted the research approach proposed by Li et al. [26], which suggests that, based on a certain level of resistance, further screening for other resistance genes can help identify alfalfa resistance genes better suited to the local ecological environment. Therefore, we implemented a two-stage screening process: the first stage involved cold tolerance screening, and the second focused on salt tolerance.

In the first stage, a series of growth indicators, such as growth rates and fresh weight, were measured. Although these are standard growth indicators, they effectively reflect the growth and development of plants under 4 °C stress [28,29]. To rank the cold tolerance of different varieties, we employed a grey correlation analysis, a method widely recognized for its effectiveness in integrating multiple indicators for comprehensive ranking. This approach has already been extensively applied in plants, such as tea (*Camellia sinensis* L.) [30], soybean (*Glycine max* L.) [31], and potato (*Solanum tuberosum* L.) [32]. Two ranking modes were generated: weighted grey correlation degree [33] and unweighted grey correlation degree, and 15 varieties were selected for further salt tolerance screening. We applied a subjective–objective approach to screen the 15 alfalfa varieties. Subjectively, we selected varieties that were already cultivated in Northeast China, such as Longmu801, Longmu806, Gongnong No.1, Gongnong No.2, and Zhao Dong [34,35]. Objectively, we used grey correlation degree scores for the screening process. By combining subjective and objective selection, we concluded that these 15 varieties possess cold tolerance.

In the second stage, we evaluated the salt tolerance of different 15 alfalfa varieties using the half-lethal rate [36] under salt stress to ensure the rationality and purposefulness of the experimental data. Assessing plant resistance using the half-lethal concentration is a well-established method. For example, Yang et al. evaluated cold tolerance in various Liliaceae species by measuring cell mortality [37]. We treated 15 alfalfa varieties with different salt concentrations (0, 100, 200, and 300 mM) and observed that, as the duration of stress increased, a large number of plants died under 300 mM salt stress. After prolonged treatment, some deaths also occurred under 200 mM, but the mortality rate was significantly lower. No deaths were observed under 100 mM and 0 mM conditions. The absence of mortality at 0 mM was expected and indicates that external environmental factors did not affect the experimental results. Similarly, the lack of mortality under 100 mM suggests that alfalfa, being a highly adaptable species with inherent salt tolerance [38], gradually adjusted to the 100 mM environment over time, leading to no deaths. Mortality occurred at both 200 mM and 300 mM, indicating that, over time, the stress exerted by these concentrations exceeded the plant's tolerance limits, resulting in death [39,40]. The higher and earlier mortality observed at 300 mM compared to 200 mM further supports this conclusion. We used the 20-day mortality data to perform a log-logistic fit of the half-lethal concentration. The reason for selecting the 20-day data for the fit was that by this time, varieties Gongnong No.2 and Caoyuan No.1 had exhibited complete mortality, making this time point suitable for data fitting.

After ranking the LC_{50}, we applied K-means clustering to categorize the varieties into salt-tolerant, moderate, and sensitive groups, thus providing better differentiation between salt-tolerant and sensitive alfalfa varieties. K-means clustering is a valuable tool for researchers to classify varieties, enhancing the focus and purpose of studies. For example, Atsa'am used this method in studies on West African cereals, and Mondal applied it in mango (*Mangifera indica* L.) research [41,42]. The salt-tolerant varieties largely included those already confirmed to be cultivable in saline–alkali regions of Northeast China, such as Zhaodong and Longmu801 [33,43]. To further investigate the salt tolerance mechanisms of alfalfa and identify salt tolerance genes, we selected Longmu801, a salt-tolerant variety, and WL168, a sensitive variety, as experimental varieties for further physiological measurements. As shown in Table S11, Longmu801 is a Chinese variety, while WL168 originates from the United States. Due to geographical differences and distinct breeding methods, the two

varieties are expected to have significant genotypic differences [44]. Further research on these two varieties will provide more distinct and meaningful results.

In the variety screening section, we applied rigorous methods to identify cold-tolerant and salt-tolerant varieties. This not only establishes a solid theoretical foundation for our study but also provides a valuable reference for other researchers conducting alfalfa studies in the region. They can build upon our selected varieties for their research or introduce additional selection criteria to identify varieties that better align with their specific goals.

3.2. Interpretation of Salt Stress Effects on Alfalfa Growth and Physiology

To better understand the physiological responses and transcriptional regulatory mechanisms of alfalfa under salt stress, we employed 300 mM salt stress in this study. This concentration poses significant harm to alfalfa and induces clear physiological responses. To eliminate the inherent differences between the two varieties, we compared the relative values of the stress and control groups, which more accurately reflect the response of each variety to salt stress and highlight differences in salt tolerance.

First, both varieties showed slow or stagnant growth in terms of plant height. However, the relative plant height between the two varieties did not differ significantly. It is well known that plants reduce their growth rate and photosynthetic activity when subjected to stress as an adaptive response [45]. This suggests that both varieties experienced significant stress, leading to inhibited growth. For fresh weight and dry weight, both varieties exhibited an increasing trend over time, but the treated groups showed lower growth compared to the control groups. Notably, there was no significant difference in fresh weight between the two varieties on day 5. However, by day 7, a significant difference emerged. In contrast, dry weight did not show significant differences at any time point. This indicates that, under prolonged stress, Longmu801 has a better capacity for water regulation compared to WL168, allowing it to maintain its fresh weight and respond more effectively to salt stress.

Physiological indicators also showed significant differences under prolonged salt stress. In terms of photosynthetic pigments, chlorophyll a and total chlorophyll are important indicators of plant stress resistance [46,47] and are commonly studied in stress experiments with crops like maize [48], tomato [49], and rice [50]. In this study, both chlorophyll a, chlorophyll b, and total chlorophyll levels showed a decreasing trend under prolonged stress. Significant differences in total chlorophyll and chlorophyll a content were observed between the two varieties. Previous research suggests that chlorophyll a is more suitable for detecting abiotic stress [51], further highlighting the differences in salt tolerance between the two varieties. Hydrogen peroxide (H_2O_2) and malondialdehyde (MDA) are key indicators of oxidative stress in plants [52], and their levels are commonly used to assess the degree of stress [53–55]. In this experiment, the H_2O_2 and MDA levels of both varieties increased at 0, 1, 3, and 5 days. However, by day 7, WL168 had significantly higher levels compared to Longmu801, suggesting that Longmu801 has stronger salt tolerance under prolonged stress. Osmolytes are also key indicators of a plant's stress resistance [56]. Proline, a typical osmotic regulator, plays a crucial role in osmotic adjustment under drought and salt stress [57]. In this study, proline levels showed significant changes. After one day of stress, both varieties began accumulating large amounts of proline, but Longmu801 exhibited significantly greater proline accumulation, indicating its stronger capacity to maintain cellular osmotic pressure under salt stress. Soluble sugars, which are critical for maintaining plant life activities, have been shown to increase significantly under abiotic stress due to a series of physiological activities [58]. In this study, the soluble sugar content of both varieties increased over time, but Longmu801 accumulated significantly more than WL168, and this difference became more pronounced as time progressed. This suggests that Longmu801 has stronger osmotic regulation capacity under salt stress. The higher accumulation of soluble sugars may also provide additional energy, helping Longmu801 better cope with the stress.

Additionally, antioxidant enzymes are important indicators of plant resistance to salt stress and reflect the plant's ability to resist oxidative damage. Although WL168 had higher antioxidant enzyme activity before stress, the increase in antioxidant enzyme activity in

Longmu801 under salt stress was significantly greater, suggesting that Longmu801 may have a stronger antioxidant capacity, allowing it to better cope with oxidative damage induced by salt stress. The antioxidant enzyme activity in both varieties exhibited an initial increase followed by a subsequent decline, which is a normal physiological response in plants. During the early stages of salt stress, alfalfa produces antioxidant enzymes to mitigate the stress, resulting in elevated enzyme activity. However, as the stress persists, the plant sustains damage, impairing its ability to maintain enzyme secretion, leading to a decline in enzyme activity. Nonetheless, the activity remains higher than that of the control (CK). If the stress continues until the plant dies, the enzyme activity would drop significantly below that of the CK. The differences in enzyme activity between the two varieties (Figure 8d−f) directly reflect their varying antioxidant capacities. This indicates that Longmu801 is better at sustaining antioxidant enzyme activity compared to WL168.

In this section, the growth and physiological indicators we measured are commonly used to evaluate the physiological state of plants under stress. For example, Feng et al. measured these indicators in a comparative study of salt tolerance in four buckwheat varieties and used a membership function to integrate the indicators for ranking [59]. Similarly, Rasel et al. utilized these physiological indicators to study different rice (*Oryza sativa* L.) genotypes under salt stress and applied clustering methods to screen salt-tolerant genotypes [60]. Quan et al. also employed these indicators in their research on salt tolerance among different alfalfa varieties and successfully obtained the desired results [61].

In terms of the salt stress experiment, what sets our research apart from others is that we used LC_{50} and clustering analysis to identify salt-tolerant and salt-sensitive varieties. We then selected specific varieties from these groups for further study, which significantly saved both experimental time and cost. It is also worth mentioning that in some variety studies, many researchers overlook the inherent differences between varieties and directly compare them, which we believe is not rigorous. In contrast, researchers such as Rasel [60] employed the susceptibility index (SI) for each variety, calculated as "(control value − salt treatment value)/control value × 100". This approach is similar to the use of relative values in our study, as it helps eliminate the inherent differences between varieties when making comparisons, ensuring a more rigorous analysis. In summary, our findings indicate that the alfalfa varieties with differing salt tolerance exhibited similar trends in their growth and physiological response mechanisms. The differences between them were primarily reflected in the varying levels of the growth and physiological indicators.

Additionally, in this section, we used the relative values of physiological indicators from the two varieties to conduct grey correlation analysis to evaluate the key time points with the most significant differences. Grey correlation analysis has been widely used in various fields to optimize experiments [62–64]. In this study, day 7 was selected as the key time point for transcriptome sequencing, allowing for the identification of more critical differential genes.

3.3. Analysis of the Transcriptional Regulatory Mechanism of Alfalfa Under Salt Stress

Transcriptomic analysis revealed that, although the same reference genome, "Zhongmu No.4" was used, there were considerable differences in the DEGs between LM_ST vs. LM_CK and WL_ST vs. WL_CK under stress. This is due to the different gene expression levels among the varieties after salt stress. These differences can be attributed to the long-term independent breeding of the two varieties [44]. The results showed that WL had more DEGs than LM; however, WL exhibited poorer salt tolerance. This suggests that the additional DEGs may accelerate plant death, possibly because the increased genetic activity leads to faster energy metabolism and hastens the plant's decline [65]. Alternatively, these genes may negatively regulate the plant's stress tolerance, further contributing to its accelerated death [66]—a hypothesis that requires further validation.

To better compare the differences between LM and WL, we did not directly use LM_ST vs. WL_ST. Instead, we explored this aspect through LM_ST vs. LM_CK, WL_ST vs. WL_CK, and LM_ST vs. WL_ST comparisons. This approach effectively eliminates

errors and biases caused by the inherent differences between the varieties. In this part, 2146 DEGs with similar expression trends were identified between the two alfalfa varieties. Among these, 1850 DEGs exhibited significant differential expression in Longmu801 compared to WL168, while 2679 DEGs showed significant differences in WL168 compared to Longmu801, likely reflecting genetic differences due to their respective evolutionary or breeding histories [44]. The enrichment of the 2146 DEGs in GO functions suggests that, under salt stress, alfalfa experiences extensive cellular activity in the nucleus, cytoplasm, plasma membrane, and membranes. ATP binding plays a crucial role in providing the energy required for these cellular activities [67], enabling the activation of stress-responsive transcription factors that, in turn, drive the expression of more stress-resistance genes [68]. Additionally, protein binding facilitates the formation of complexes to defend against salt stress, with these cellular activities being part of the plant's response to oxidative stress.

The KEGG pathway enrichment analysis of these 2146 DEGs highlighted pathways such as phenylpropanoid biosynthesis (ko00940), protein processing in the endoplasmic reticulum (ko04141), and starch and sucrose metabolism (ko00500). The phenylpropanoid biosynthesis pathway is not only involved in the synthesis of lignin and flavonoids but also plays a critical role in gene regulation and DNA binding [69], providing a positive contribution to alfalfa's defense against stress. The protein processing and starch and sucrose metabolism pathways further suggest that the plant requires more resistance-related or osmotic regulation proteins to combat stress. The increase in soluble sugar content observed in the physiological analysis is likely due to these metabolic activities. In addition to serving as a carbon source, sucrose also functions as a soluble sugar that regulates cellular osmotic pressure and helps maintain normal cellular activities.

Although 1850 DEGs were significantly different in Longmu801 and 2679 DEGs were significantly different in WL168, the GO functions and KEGG pathways of these DEGs were largely similar to the 2146 shared DEGs (as shown in Figures S2 and S3). This suggests that, despite genetic differences resulting from their evolutionary and breeding histories, both varieties rely on similar metabolic pathways and biological functions to resist salt stress. This contrasts with Lei's study [70], where Lei posits that salt-tolerant varieties possess unique regulatory mechanisms, which contradicts our conclusions. In Lei's research, the regulatory mechanisms refer more to the differences in transcription levels of salt-tolerant genes within salt-tolerant varieties compared to non-tolerant varieties, suggesting that the mechanisms are distinct. In our study, we prefer to use KEGG and GO analyses to explain the transcriptional regulatory mechanisms in alfalfa rather than focusing solely on a few specific genes. Our approach helps provide a more comprehensive explanation of the transcriptional regulatory mechanisms in alfalfa under salt stress. A conclusion consistent with Lei's is that both studies agree that the differences between salt-tolerant and sensitive varieties arise from the expression levels of DEGs.

The conclusion we can draw from the transcriptomic analysis is that, although salt-tolerant and sensitive varieties exhibit similar transcriptional regulatory mechanisms in response to salt stress, the differences in DEGs between the varieties determine their variations in salt tolerance.

3.4. WGCNA-Based Hub Gene Mining and Functional Analysis

In this study, to better identify the key genes involved in salt stress resistance, we conducted WGCNA analysis using physiological indicators and transcriptomic data [71]. The WGCNA results identified seven modules (excluding the non-informative grey module), and we selected six modules that showed strong correlations with stress-resistant physiological traits for further analysis. Using WGCNA to identify hub genes is a common and effective approach. Zhu applied WGCNA in the study of salt-responsive hub genes in rice (*Oryza sativa* L.), identifying key transcription factors and genes within the relevant modules [72]. Similarly, Meng used WGCNA in their research on drought tolerance in Tartary buckwheat (*Fagopyrum tataricum* Gaertn.), discovering that HD-ZIP and MYB transcription factors may serve as key downstream regulators of drought resistance in buckwheat [73].

We conducted further research on the 30 hub genes identified and the 5 "abscisic acid and environmental stress-inducible protein" genes obtained from the transcriptomic analysis. To better understand the relationships among these 35 genes in alfalfa, we analyzed their correlations to gain insights into potential interactions, providing a more comprehensive understanding of their roles [74,75].

Further enrichment analysis of these genes revealed five hub genes directly related to salt stress. The results of the Mantel test showed a strong correlation between these five genes and physiological indicators. The Mantel test is a practical method for exploring the correlation between two different variables and has been widely applied in studies investigating the relationship between population genetics and environmental variables [76,77]. In this study, the Mantel test was used to analyze the correlation between the FPKM values of these five genes and physiological indicators, aiming to better explore the potential relationships between these genes and physiological traits.

Through WGCNA and further correlation analysis, this study not only identified hub genes but also explored the potential relationships among the hub genes and between the hub genes and physiological indicators. This deepens our understanding of these genes and provides a solid theoretical foundation for future research on them.

4. Materials and Methods

4.1. Preparation of Experimental Materials

Alfalfa (*Medicago sativa* L.) seeds were provided by the Forage Laboratory of Northeast Agricultural University. The variety and origin of the seeds are listed in Table S11. All seeds were sterilized with 15% sodium hypochlorite for 30 min and rinsed three times with sterile water. The seeds were then transferred to Petri dishes and incubated at a constant temperature of 25 °C in a plant incubator (GPL-250, Labotery, Tianjin, China) in darkness for one week, with the humidity maintained at 85%. After germination, the seedlings were transferred to 50-cell trays filled with pure vermiculite and irrigated with 1/5 Hoagland nutrient solution every two days. The seedlings were grown to the five-leaf stage before proceeding with subsequent experiments.

4.2. Variety Screening Process

4.2.1. Cold Tolerance Screening Experiment

The alfalfa plants grown to the five-leaf stage were subjected to low-temperature stress at 4 °C. The experiment was conducted in the Grass Science Laboratory at Northeast Agricultural University. The specific procedure is as follows: First, the five-leaf stage alfalfa plants were placed in a plant incubator (GPL-250, Labotery, Tianjin, China) at room temperature (25 °C). The temperature was then gradually reduced by 2 °C per hour until it reached 4 °C, marking the beginning of the treatment period. Measurements of the plants were taken on day ten. Additionally, the light cycle in the incubator was set to 16 h of light and 8 h of darkness, with a light intensity of 7500 LX and humidity at 75%. For each variety, a treatment group was established at 4 °C, alongside a control group at room temperature (25 °C). Each group contained 50 plants, and three independent plants were randomly selected for measurement. The following indicators were measured: plant height, growth rate, fresh weight, dry weight, and biomass accumulation. Plant height was measured directly using a ruler. The growth rate was calculated by measuring the height on day 10, subtracting the initial height, and dividing by the height on day 1. Fresh weight was obtained by directly weighing the above-ground parts of the alfalfa. Dry weight was measured after drying the above-ground parts in a constant temperature drying oven (GZX-9023MBE, Boxun, Shanghai, China) at 65 °C until constant weight was achieved. The biomass accumulation was calculated as the fresh weight on day 10 minus the fresh weight on day 1. Cold tolerance was evaluated using the method by Xiang [78], where the relative values of various indicators were calculated to represent cold tolerance. Grey correlation analysis [79] was then used to evaluate cold tolerance, and the top 15 varieties were selected for the next stage of screening. The formula is as follows:

$$Cold\ tolerance(indicator) = \frac{Cold\ treatment\ group\ (indicator)}{Control\ group\ (indicator)}$$

4.2.2. Salt Tolerance Screening Experiment

The selected varieties were subjected to salt tolerance screening using pot experiments in salt trays. To keep the vermiculite relatively dry while ensuring that the five-leaf stage alfalfa is not subjected to drought stress, we maintained the moisture needs of the alfalfa primarily by spraying water daily for about five days at the beginning of the experiment rather than using irrigation. We placed the trays containing five-leaf stage alfalfa, with carefully controlled vermiculite moisture into large trays filled with salt solution, allowing the vermiculite to fully absorb the salt solution. Starting on the first day after treatment, we recorded the number of dead plants under different NaCl concentrations (0, 100 mM, 200 mM, 300 mM). Each variety was established with three independent replicates, each consisting of fifty individual plants. A log-logistic half-lethal concentration fitting was then performed [80], followed by K-means clustering analysis to identify salt-tolerant and salt-sensitive varieties [81]. To ensure the accuracy of the salt treatments, the salt concentration, salt content, and electrical conductivity of the vermiculite suspension in the trays treated with different salt concentrations were measured according to the methods outlined in "Soil Agrochemical Analysis" [82]. The results are shown in Supplementary Figure S6. A correlation analysis between the salt concentration of the vermiculite extract and the salt treatment solution showed $R^2 = 0.99$, indicating that the salt concentration in the vermiculite was consistent with that of the treatment solution.

4.3. Measurement of Growth and Physiological Indicators Under Salt Stress

A salt stress experiment was conducted on salt-tolerant and salt-sensitive varieties using 300 mM NaCl. As mentioned in Section 4.2.2 regarding the salt stress experiment, we employed the method of pot experiments in salt trays for this experiment. Each variety had a salt treatment group and a control group (the control group used distilled water instead of salt solution). Each group contained 50 alfalfa plants, and three independent plants were randomly selected from these 50 for measurement. Growth indicators were measured at 0, 3, and 7 days, and physiological indicators were measured at 0, 1, 3, 5, and 7 days. Grey correlation analysis was performed at different time points to identify the optimal time for transcriptomic analysis.

Growth indicators included measuring plant height using a ruler and assessing fresh and dry weights through a weighing method. Fresh weight was measured by taking the above-ground parts of the alfalfa and directly weighing them. Dry weight was obtained by placing the weighed fresh samples in a constant temperature drying oven (GZX-9023MBE, Boxun, Shanghai, China) at 65 °C until a constant weight was achieved, after which the weight was measured.

The physiological indicators were measured using kits purchased from Suzhou Keming Biological Co., Ltd., Suzhou, China (http://www.cominbio.com/index.html (accessed on 10 September 2023)). The above-ground parts of the alfalfa were used for sample grinding, followed by homogenization for measurement. Specifically, proline content was measured using the sulfosalicylic acid method, malondialdehyde (MDA) and soluble sugars were measured using the thiobarbituric acid method, chlorophyll was measured using the anhydrous ethanol extraction method, and hydrogen peroxide was extracted with acetone. Superoxide dismutase (SOD), peroxidase (POD), and catalase (CAT) activities were measured using spectrophotometry (N5000Plus, Yoke Instrument, Shanghai, China) [83]. Grey correlation analysis [79] was then used to select the key time point for transcriptomic sequencing.

4.4. Transcriptomic Analysis

Sampling was performed at the key time points determined from the salt stress experiment, followed by transcriptomic analysis. The specific steps are as follows: After 7 days of salt stress and control treatments, samples were collected from Longmu801 and WL168, primarily taking the above-ground parts of the alfalfa, which were then ground and homogenized. For each variety, one group of control treatment and one group of salt treatment were collected, totaling four groups. Four biological replicates were collected for each treatment and sent to LC-Bio Technologies Co., Ltd. (Hangzhou, China) for testing under dry ice conditions, with approximately 0.5 g of sample per replicate.

Total RNA was isolated from tissue samples using TRIzol reagent (Invitrogen, Waltham, MA, USA) according to the manufacturer's protocol. RNA quantity and purity were assessed using a NanoDrop ND-1000 spectrophotometer (ThermoFisher, Waltham, MA, USA), and RNA integrity was evaluated using the Agilent 2100 Bioanalyzer system (Agilent, Santa Clara, CA, USA). Samples with an RNA integrity number (RIN) greater than 7.0 were used for downstream analysis, primarily assessed using the RNA 6000 Nano LabChip Kit (Agilent, Santa Clara, CA, USA, 5067-1511). Polyadenylated mRNA was enriched using PolyA and Oligo(dT) magnetic beads. The mRNA was fragmented using divalent cations, and first-strand cDNA was synthesized using reverse transcriptase. Double-stranded cDNA was generated using DNA polymerase I and RNase H. The ends of the double-stranded cDNA were blunted, an A-base was added, and adapters were ligated. The resulting fragments (~300 bp) were purified and amplified by PCR. Libraries were sequenced on the Illumina HiSeq 4000 platform to generate 150 bp paired-end reads.

The raw data were further filtered by Cutadapt (https://cutadapt.readthedocs.io/en/stable (accessed on 11 December 2023), version 1.9). After deleting low-quality and repeated sequences, high-quality clean reads were produced. The clean data were then compared to the alfalfa genome (https://modms.lzu.edu.cn/alfalfa/browse/genome/ZM-4, https://modms.lzu.edu.cn/alfalfa/browse/genome/ZMNO, https://modms.lzu.edu.cn/alfalfa/browse/genome/XJDY, accessed on 11 December 2023) [27,84] using HISAT2 (hierarchical indexing for spliced alignment of transcripts) (https://daehwankimlab.github.io/hisat2 (accessed on 11 December 2023), version 2.0.4) to obtain the BAM file. The initial assembly of transcripts for each sample was accomplished using StringTie (http://ccb.jhu.edu/software/stringtie/ (accessed on 11 December 2023), version 2.1.6) with default parameters, and the mapped reads were merged to reconstruct a comprehensive transcriptome. Subsequently, gffcompare software (http://ccb.jhu.edu/software/stringtie/gffcompare.shtml (accessed on 11 December 2023), version 0.9.8) was employed to compare the assembled transcripts with the reference annotation, generating the final assembly annotation. After producing the final transcriptome, StringTie and ballgown (http://www.bioconductor.org/packages/release/bioc/html/ballgown.html (accessed on 11 December 2023), version 2.12.0) were used to estimate expression levels of all transcripts and calculate FPKM (fragments per kilobase of transcript per million mapped reads) values for mRNAs, with ballgown facilitating the FPKM quantification file input [85–87]. The R package edgeR was used to assess the statistically significant differences between samples. Differentially expressed genes (DEGs) were identified as those with a multiple-fold change of more than 2 (upregulate) or less than 0.5 (downregulate) and a p-value < 0.05. The biological functions and signaling pathways of DEGs were classified using GO (Gene Ontology) and KEGG (Kyoto Encyclopedia of Genes and Genomes) enrichment analyses.

4.5. qRT-PCR Validation Procedure

The RNA used for qPCR was extracted from samples collected after 7 days of salt treatment, ensuring consistency with the samples used for transcriptomic analysis. The RNA extraction was carried out using the Ultrapure RNA Kit (Cowin Biotech, Taizhou, China) from Cowin Biotech. cDNA was synthesized using the HiScript II Q RT SuperMix for qPCR (+gDNA wiper) kit (Vazyme Biotech, Nanjing, China) from Vazyme Biotech. This was followed by qPCR analysis.

For qPCR, GAPDH was used as the reference gene [88], and the primers for validating the DEGs from the transcriptome data can be found in Supplementary Table S12. The qPCR instrument used for quantification was the q225 model (q225, Kubo Technology, Beijing, China) from Kubo Technology. The qPCR assays followed the three-step protocol, using the 2× SYBR Green qPCR Master Mix II (Universal) (Seven Biotech, Beijing, China). The detailed reaction protocol is provided in Table S13.

4.6. Gene Co-Expression Network Analysis (WGCNA)

In this study, we performed WGCNA using the WGCNA cloud tool (https://www.omicstudio.cn/analysis, accessed on 27 December 2023) [89] provided by the LC-Bio cloud platform. The specific procedures are as follows:

First, we compiled the FPKM values of all detected genes from the transcriptomic data of LM_CK, LM_ST, WL_CK, and WL_ST samples, with each group comprising four biological replicates, resulting in a total of 16 datasets. This sample size meets the requirement for conducting WGCNA (a minimum of 15 samples). We subsequently filtered out genes with an expression level below 2, yielding a final count of 32,684 genes.

Next, we conducted a clustering analysis of the different samples to eliminate any outlier data points, followed by further WGCNA analysis. We then determined the soft threshold, setting the platform threshold to 0.85 and the soft threshold to 1. To enhance the biological significance of our findings, we employed the topological overlap measure (TOM) to evaluate the degree of association between genes, taking into account not only the interactions between pairs of genes but also their connections with other genes in the network. We also calculated the correlation between modules and samples, generating a correlation heatmap to identify modules most closely associated with the samples. Additionally, GO and KEGG enrichment analyses were performed to further elucidate the functional roles of genes within the modules. Finally, hub genes within different modules were identified using Cytoscape (v3.9.1) in conjunction with the MCODE plugin. Finally, key genes were identified based on degree values within each module using Cytoscape (v3.9.1) and the MCODE plugin [90].

4.7. Hub Gene Correlation Analysis and Functional Study

The expression levels of hub genes under different treatments were analyzed using the LC-Bio cloud platform (https://www.omicstudio.cn/tool, accessed on 10 March 2024). Specifically, the FPKM values of hub genes in LM_CK, LM_ST, WL_CK, and WL_ST were calculated and further analyzed using the platform's correlation analysis tool (https://www.omicstudio.cn/tool/62, accessed on 10 March 2024). The results were visualized using the platform's visualization tool (https://www.omicstudio.cn/tool/64, accessed on 10 March 2024) [89].

Additionally, the GO functions and KEGG pathways of hub genes were visualized using a Sankey diagram on the same platform (https://www.omicstudio.cn/tool/155, accessed on 10 March 2024). To evaluate the correlation between salt stress-related genes' FPKM and physiological indicators across different varieties, a Mantel test was performed (https://www.omicstudio.cn/tool/109, accessed on 10 March 2024) [91].

5. Conclusions

In this study, Longmu801 was identified as a salt-tolerant alfalfa variety, while WL168 was identified as a salt-sensitive variety. The two varieties exhibited the same growth and physiological response mechanisms, with their differences primarily reflected in the varying levels of the indicators. When subjected to salt stress, Longmu801 maintained higher levels of osmolytes (proline and soluble sugars), antioxidant enzymes (SOD, POD, and CAT), chlorophyll, and fresh weight compared to WL168, along with lower levels of oxidative stress markers (malondialdehyde and hydrogen peroxide). Transcriptomic analysis revealed significant differences in DEGs between the two varieties; however, the enriched GO functions (such as oxidative stress, nucleus, protein binding, etc.) and

KEGG pathways (including phenylpropanoid biosynthesis, protein processing in the endoplasmic reticulum, and starch and sucrose metabolism, etc.) were remarkably similar. This suggests that both varieties share analogous transcriptional regulatory mechanisms in response to salt stress. Furthermore, through weighted gene co-expression network analysis (WGCNA), five hub genes directly associated with salt response were identified: *Msa085011*, *Msa0605650*, *Msa0397400*, *Msa1258740*, and *Msa0958830*.

In summary, the two varieties exhibit similar growth and physiological response mechanisms, as well as analogous transcriptional regulatory mechanisms under salt stress. Their differences in salt tolerance are primarily attributed to variations in the levels of specific indicators and the expression levels of DEGs. This will provide a theoretical basis and research foundation for breeding salt-tolerant alfalfa.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/plants13223141/s1. Figure S1. Mortality rates of different alfalfa varieties; Figure S2. Analysis of 1850 DEGs; Figure S3. Analysis of 2679 DEGs; Figure S4. qRT-PCR validation of transcriptomic data; Figure S5. Functional study of genes in different WGCNA modules; Figure S6. Verification of salt concentrations in the salt stress experiment; Table S1. Effects of cold stress on the plant height of different alfalfa varieties.; Table S2. Effects of cold stress on the growth rate of different alfalfa Varieties; Table S3. Effects of cold stress on the fresh weight of different alfalfa varieties; Table S4. Effects of cold stress on the increase in biomass of different alfalfa varieties; Table S5. Effects of cold stress on the dry weight of different alfalfa varieties; Table S6. Effects of cold stress on relative growth indicators of different alfalfa varieties; Table S7. Unweighted and weighted grey correlation of different alfalfa varieties; Table S8. Statistical table for assessment of sample sequencing data; Table S9. Comparative analysis of different alfalfa genomes; Table S10. Comparison table between sample and reference genome; Table S11. Origin and location of alfalfa varieties; Table S12. Primers of five selected DEGs for qRT-PCR analysis; Table S13. qRT-PCR reaction protocol.

Author Contributions: G.C. and F.W. designed the research; F.W., H.W., M.Y., W.X., W.Z., and R.Q. conducted the experiments and analyzed the data; resources, G.C.; writing—original draft preparation, F.W.; data curation, H.W.; validation, M.Y.; visualization, W.X.; writing—review and editing, W.Z.; supervision, R.Q.; software, N.K. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by "The Science and Technology Innovation 2030-Major Project (2022ZD040120401)".

Data Availability Statement: The RNA-seq datasets generated during the current study have been submitted to the NCBI Sequence Read Archive under the accession number PRJNA1110792 (https://dataview.ncbi.nlm.nih.gov/object/PRJNA1110792. Release date: 31 October 2024).

Acknowledgments: We would like to thank the Forage Laboratory for providing essential equipment and instruments, and the College of Animal Science and Technology for supplying necessary experimental reagents and materials.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Iannucci, A.; Di Fonzo, N.; Martiniello, P. Alfalfa (*Medicago sativa* L.) seed yield and quality under different forage management systems and irrigation treatments in a Mediterranean environment. *Field Crops Res.* **2002**, *78*, 65–74. [CrossRef]
2. Mielmann, A. The utilisation of lucerne (*Medicago sativa*): A review. *Br. Food J.* **2013**, *115*, 590–600. [CrossRef]
3. Wang, Y.; You, Z.; Ren, P. Research progress on alfalfa (*Medicago sativa* L.) adaptability under drought stress. *Agric. Sci. Technol.* **2017**, *18*, 219–222.
4. Karayilanli, E.; Ayhan, V. Investigation of feed value of alfalfa (*Medicago sativa* L.) harvested at different maturity stages. *Legume Res. Int. J.* **2016**, *39*, 237–247. [CrossRef]
5. Fan, J.W.; Du, Y.L.; Turner, N.C.; Wang, B.R.; Fang, Y.; Xi, Y.; Guo, X.R.; Li, F.M. Changes in root morphology and physiology to limited phosphorus and moisture in a locally-selected cultivar and an introduced cultivar of *Medicago sativa* L. growing in alkaline soil. *Plant Soil* **2015**, *392*, 215–226. [CrossRef]
6. Liu, Z.Y.; Li, X.L.; Li, F.; Wang, Z.L.; Sun, Q.Z. Mechanisms underlying the effects of fall dormancy on the cold acclimation and winter hard-iness of *Medicago sativa*. *Chin. J. Plant Ecol.* **2015**, *39*, 635.

7. Annicchiarico, P.; Pecetti, L.; Tava, A. Physiological and morphological traits associated with adaptation of lucerne (*Medicago sativa*) to severely drought-stressed and to irrigated environments. *Ann. Appl. Biol.* **2013**, *162*, 27–40. [CrossRef]
8. Wang, Y.; Wang, J.; Guo, D.; Zhang, H.; Che, Y.; Li, Y.; Tian, B.; Wang, Z.; Sun, G.; Zhang, H. Physiological and comparative transcriptome analysis of leaf response and physiological adaption to saline alkali stress across pH values in alfalfa (*Medicago sativa*). *Plant Physiol. Biochem.* **2021**, *167*, 140–152. [CrossRef]
9. de Aguilar, C.A.G.; Azcón, R.; Barea, J.M. Endomycorrhizal fungi and Rhizobium as biological fertilisers for *Medicago sativa* in normal cultivation. *Nature* **1979**, *279*, 325–327. [CrossRef]
10. Song, Y.; Lv, J.; Ma, Z.; Dong, W. The mechanism of alfalfa (*Medicago sativa* L.) response to abiotic stress. *Plant Growth Regul.* **2019**, *89*, 239–249. [CrossRef]
11. Guo, Z.; Zhang, T.; Chen, Z.; Niu, J.; Cui, X.; Mao, Y.; Hassan, M.U.; Kareem, H.A.; Xu, N.; Sui, X.; et al. Occurrence, distribution, and genetic diversity of alfalfa (*Medicago sativa* L.) viruses in four major alfalfa-producing provinces of China. *Front. Microbiol.* **2022**, *12*, 771361. [CrossRef] [PubMed]
12. Sun, Q.; Liu, Q.; Tao, Y.; Xu, L. A preliminary study on alfalfa cultivation and utilization during the Han, Wei, Jin, and Southern and Northern Dynasties. *Acta Prataculturae Sin.* **2017**, *26*, 185.
13. Feng, Y.; Shi, Y.; Zhao, M.; Shen, H.; Xu, L.; Luo, Y.; Liu, Y.; Xing, A.; Kang, J.; Jing, H.; et al. Yield and quality properties of alfalfa (*Medicago sativa* L.) and their influencing factors in China. *Eur. J. Agron.* **2022**, *141*, 126637. [CrossRef]
14. Jin, J.; Wang, T.; Cheng, Y.; Wang, L.; Zhang, J.; Jing, H.; Chong, K. Current situation and prospect of forage breeding in China. *Bull. Chin. Acad. Sci.* **2021**, *36*, 660–665.
15. Wang, T.; Zhong, J. Creating modern technological system for grass product processing to guarantee macroscopic food security. *Bull. Chin. Acad. Sci.* **2021**, *36*, 675–684.
16. Shi, S.; Nan, L.; Smith, K.F. The current status, problems, and prospects of alfalfa (*Medicago sativa* L.) breeding in China. *Agronomy* **2017**, *7*, 1. [CrossRef]
17. Shi, J.; Lu, Q.; Zhang, G.; Wang, S.; Hou, M.; Liu, L.; Sa, D. Research progress on the effects of salt stress on the growth, development and nutritional quality of alfalfa. *Grassl. Grassl. Ind.* **2024**, *36*, 1–7.
18. Song, M. Making Good Use of Reserve Arable Land Resources. *Economic Daily*, 10 May 2024; p. 11.
19. Moose, S.P.; Mumm, R.H. Molecular plant breeding as the foundation for 21st century crop improvement. *Plant Physiol.* **2008**, *147*, 969–977. [CrossRef] [PubMed]
20. Tester, M.; Langridge, P. Breeding technologies to increase crop production in a changing world. *Science* **2010**, *327*, 818–822. [CrossRef]
21. Tang, L.; Cai, H.; Ji, W.; Luo, X.; Wang, Z.; Wu, J.; Wang, X.; Cui, L.; Wang, Y.; Zhu, Y.; et al. Overexpression of *GsZFP1* enhances salt and drought tolerance in transgenic alfalfa (*Medicago sativa* L.). *Plant Physiol. Biochem.* **2013**, *71*, 22–30. [CrossRef]
22. Zhang, Y.; Zhang, Y.; Wang, C.; Xiao, J.; Huang, M.; Zhuo, L.; Zhang, D. Enhancement of salt tolerance of alfalfa: Physiological and molecular responses of transgenic alfalfa plants expressing *Syntrichia caninervis*-derived *ScABI3*. *Plant Physiol. Biochem.* **2024**, *207*, 108335. [CrossRef] [PubMed]
23. Zhang, Z.; Xu, P.; Duan, Z.; Lu, L.; Nan, Z.; Zhang, J. Overexpression of P5CDH from *Cleistogenes songorica* improves alfalfa growth performance under field drought conditions. *Plant Physiol. Biochem.* **2024**, *209*, 108551. [CrossRef] [PubMed]
24. Luo, Y.; Wang, X.; Zhang, D.; Zhan, L.; Li, D.; Li, C.; Cong, C.; Cai, H. Overexpression of phosphoenolpyruvate carboxylase kinase gene *MsPPCK1* from *Medicago sativa* L. increased alkali tolerance of alfalfa by enhancing photosynthetic efficiency and promoting nodule development. *Plant Physiol. Biochem.* **2024**, *213*, 108764. [CrossRef]
25. Han, W. Current status and prospects of alfalfa industry in the high latitude cold regions of Heilongjiang Province. *Grass Sci.* **2020**, *5*, 5–8.
26. Li, Z.; Zhou, B.; Ba, L. Several issues to be noted in alfalfa cultivation in Northeast my country. In Proceedings of the 7th (2017) China Alfalfa Development Conference, Jiuquan, China, 15–18 August 2017; pp. 387–392.
27. Fang, L.; Liu, T.; Li, M.; Dong, X.; Han, Y.; Xu, C.; Li, S.; Zhang, J.; He, X.; Zhou, Q.; et al. MODMS: A multi-omics database for facilitating biological studies on alfalfa (*Medicago sativa* L.). *Hortic. Res.* **2024**, *11*, 245. [CrossRef]
28. Parkash, V.; Singh, S. A review on potential plant-based water stress indicators for vegetable crops. *Sustainability* **2020**, *12*, 3945. [CrossRef]
29. Jouyban, Z. The effects of salt stress on plant growth. *Tech. J. Eng. Appl. Sci.* **2012**, *2*, 7–10.
30. Wang, Z.; Yang, P.; Peng, H.; Li, C.; Yue, C.; Li, W.; Jiang, X. Comprehensive evaluation of 47 tea [*Camellia sinensis* (L.) O. Kuntze] germplasm based on entropy weight method and grey correlation degree. *Genet. Resour. Crop Evol.* **2021**, *68*, 3257–3270. [CrossRef]
31. Shi, X.; Ding, S.; Cong, H.; Zhang, J.; Qu, K.; Wang, X.; Han, D.; Yan, Y. Grey correlation analysis of yield-related traits of summer-sown soybean. *Xinjiang Agric. Sci.* **2023**, *60*, 1641–1652.
32. Wang, N.; Xing, Y.; Wang, X. Exploring options for improving potato productivity through reducing crop yield gap in Loess Plateau of China based on grey correlation analysis. *Sustainability* **2019**, *11*, 5621. [CrossRef]
33. Guo, Y.; Lan, J.; Li, C.; Tang, X.; Yu, D.; Chen, X. Analysis and evaluation of the nutritional value of wild forage resources in Medog. *Feed. Res.* **2024**, *1*, 1–13.
34. Liu, J.; Wang, X.; Kong, X.; Zhang, Q.; Qu, H. Preliminary report on the production performance evaluation of 25 alfalfa varieties in Suihua area. *Feed. Res.* **2022**, *45*, 102–107. [CrossRef]

35. Wen, L.; Gao, X.; Fan, Y.; Wang, B.; Li, W.; Chen, J.; Wang, Y.; Bi, S. Evaluation and screening of 28 alfalfa varieties in Xing'an League. *Feed. Res.* **2024**, *19*, 133–139. [CrossRef]
36. Stephan, C.E. Methods for calculating an LC_{50}//Aquatic toxicology and hazard evaluation. In *Aquatic Toxicology and Hazard Evaluation*; ASTM International: West Conshohocken, PA, USA, 1977.
37. Yang, W.H.; Yong, S.H.; Park, D.J.; Ahn, S.J.; Kim, D.H.; Park, K.B.; Jin, E.J.; Choi, M.S. Efficient Cold Tolerance Evaluation of Four Species of Liliaceae Plants through Cell Death Measurement and Lethal Temperature Prediction. *Horticulturae* **2023**, *9*, 751. [CrossRef]
38. Bhattarai, S.; Biswas, D.; Fu, Y.B.; Biligetu, B. Morphological, physiological, and genetic responses to salt stress in alfalfa: A review. *Agronomy* **2020**, *10*, 577. [CrossRef]
39. May, L.H.; Milthorpe, F.L. Drought resistance of crop plants. *Field Crop Abstr.* **1962**, *15*, 3.
40. Maximov, N.A. Internal factors of frost and drought resistance in plants. *Protoplasma* **1929**, *7*, 259–291. [CrossRef]
41. Atsa'am, D.D.; Oyelere, S.S.; Balogun, O.S.; Wario, R.; Blamah, N.V. K-means cluster analysis of the West African species of cereals based on nutritional value composition. *Afr. J. Food Agric. Nutr. Dev.* **2021**, *21*, 17195–17212. [CrossRef]
42. Mondal, B.; Paul Choudhury, J. A comparative study on k means and pam algorithm using physical characters of different varieties of mango in india. *Int. J. Comput. Appl.* **2013**, *78*, 21–24. [CrossRef]
43. Cai, D. *Introduction of High-Quality Cold-Resistant and High-Yield Alfalfa Germplasm Resources and Integrated Demonstration of Cultivation Technology*; Heilongjiang Academy of Agricultural Sciences: Harbin, China, 2018.
44. Rauf, S.; da Silva, J.T.; Khan, A.A.; Naveed, A. Consequences of plant breeding on genetic diversity. *Int. J. Plant Breed* **2010**, *4*, 1–21.
45. Läuchli, A.; Grattan, S.R. Plant growth and development under salinity stress. In *Advances in Molecular Breeding Toward Drought and Salt Tolerant Crops*; Springer: Cham, Switzerland, 2007; pp. 1–32.
46. Lichtenthaler, H.K.; Rinderle, U. The role of chlorophyll fluorescence in the detection of stress conditions in plants. *CRC Crit. Rev. Anal. Chem.* **1988**, *19* (Suppl. 1), S29–S85. [CrossRef]
47. Kancheva, R.H.; Borisova, D.S.; Iliev, I.T. Chlorophyll fluorescence as a plant stress indicator. *Recent Dev. Remote Sens. Space* **2008**, *5*, 301–306.
48. Qi, M.; Liu, X.; Li, Y.; Song, H.; Yin, Z.; Zhang, F.; He, Q.; Xu, Z.; Zhou, G. Photosynthetic resistance and resilience under drought, flooding and rewatering in maize plants. *Photosynth. Res.* **2021**, *148*, 1–15. [CrossRef] [PubMed]
49. Sperdouli, I.; Mellidou, I.; Moustakas, M. Harnessing chlorophyll fluorescence for phenotyping analysis of wild and cultivated tomato for high photochemical efficiency under water deficit for climate change resilience. *Climate* **2021**, *9*, 154. [CrossRef]
50. Khaleghi, E.; Arzani, K.; Moallemi, N.; Barzegar, M. Evaluation of chlorophyll content and chlorophyll fluorescence parameters and relationships between chlorophyll a, b and chlorophyll content index under water stress in *Olea europaea* cv. Dezful. *Int. J. Agric. Biosyst. Eng.* **2012**, *6*, 636–639.
51. Kalaji, H.M.; Jajoo, A.; Oukarroum, A.; Brestic, M.; Zivcak, M.; Samborska, I.A.; Cetner, M.D.; Łukasik, I.; Goltsev, V.; Ladle, R.J. Chlorophyll a fluorescence as a tool to monitor physiological status of plants under abiotic stress conditions. *Acta Physiol. Plant.* **2016**, *38*, 1–11. [CrossRef]
52. Sairam, R.K.; Srivastava, G.C. Water stress tolerance of wheat (*Triticum aestivum* L.): Variations in hydrogen peroxide accumulation and antioxidant activity in tolerant and susceptible genotypes. *J. Agron. Crop Sci.* **2001**, *186*, 63–70. [CrossRef]
53. Pongprayoon, W.; Roytrakul, S.; Pichayangkura, R.; Chadchawan, S. The role of hydrogen peroxide in chitosan-induced resistance to osmotic stress in rice (*Oryza sativa* L.). *Plant Growth Regul.* **2013**, *70*, 159–173. [CrossRef]
54. Chowdhury, S.R.; Choudhuri, M.A. Hydrogen peroxide metabolism as an index of water stress tolerance in jute. *Physiol. Plant.* **1985**, *65*, 476–480. [CrossRef]
55. Jiang, J.; Zhuang, J.; Fan, Y.; Shen, B. Mapping of QTLs for leaf malondialdehyde content associated with stress tolerance in rice. *Rice Sci.* **2009**, *16*, 72–74. [CrossRef]
56. Wang, J.; Li, D. Accumulation of osmotic regulating substances and reactive oxygen metabolism in plants under adverse conditions. *Acta Bot. Sin.* **2001**, *18*, 459.
57. Delauney, A.J.; Verma, D.P.S. Proline biosynthesis and osmoregulation in plants. *Plant J.* **1993**, *4*, 215–223. [CrossRef]
58. Rosa, M.; Prado, C.; Podazza, G.; Interdonato, R.; González, J.A.; Hilal, M.; Prado, F.E. Soluble sugars: Metabolism, sensing and abiotic stress: A complex network in the life of plants. *Plant Signal. Behav.* **2009**, *4*, 388–393. [CrossRef] [PubMed]
59. Feng, Z.; Yang, H.; Wang, B.; Wu, J. Physiological response to salt stress and evaluation of salt tolerance of four buckwheat varieties. *J. Gansu Agric. Univ.* **2019**, *59*, 97–103. [CrossRef]
60. Rasel, M.; Tahjib-Ul-Arif, M.; Hossain, M.A.; Sayed, M.A.; Hassan, L. Discerning of rice landraces (*Oryza sativa* L.) for morpho-physiological, antioxidant enzyme activity, and molecular markers' responses to induced salt stress at the seedling stage. *J. Plant Growth Regul.* **2020**, *39*, 41–59. [CrossRef]
61. Quan, W.; Liu, X.; Wang, H.; Chan, Z. Physiological and transcriptional responses of contrasting alfalfa (*Medicago sativa* L.) varieties to salt stress. *Plant Cell Tissue Organ Cult.* **2016**, *126*, 105–115. [CrossRef]
62. Cao, X. Hyperspectral Grey Correlation Local Regression Estimation Model for Soil Organic Matter Content. Master's Thesis, Shandong Agricultural University, Taian, China, 2021. [CrossRef]
63. Hu, X. Analysis of Temperature Field in Laying Hen House and Optimization of Temperature Sensor Layout. Master's Thesis, Hebei Agricultural University, Baoding, China, 2022. [CrossRef]

64. Wang, J.; Pan, L.; Xing, D.; Ding, H. Multi-objective optimization of cutting parameters for SiCp/Al composites based on energy efficiency. *J. Cent. South Univ. (Sci. Technol.)* **2020**, *51*, 1565–1574.
65. Huang, L.; Yuan, Z.; Yu, J.; Zhou, T. Fundamental principles of energy consumption for gene expression. Chaos: An Interdisciplinary. *J. Nonlinear Sci.* **2015**, *25*, 123101.
66. Dietrich, R.A.; Richberg, M.H.; Schmidt, R.; Dean, C.; Dangl, J.L. A novel zinc finger protein is encoded by the Arabidopsis LSD1 gene and functions as a negative regulator of plant cell death. *Cell* **1997**, *88*, 685–694. [CrossRef]
67. Jorgensen, P.L.; Pedersen, P.A. Structure–function relationships of Na^+, K^+, ATP, or Mg^{2+} binding and energy transduction in Na, K-ATPase. *Biochim. Biophys. Acta (BBA)-Bioenerg.* **2001**, *1505*, 57–74. [CrossRef]
68. Papavassiliou, A.G. Transcription factors. *N. Engl. J. Med.* **1995**, *332*, 45–47. [CrossRef] [PubMed]
69. Vogt, T. Phenylpropanoid biosynthesis. *Mol. Plant* **2010**, *3*, 2–20. [CrossRef] [PubMed]
70. Lei, Y.; Xu, Y.; Hettenhausen, C.; Lu, C.; Shen, G.; Zhang, C.; Li, J.; Song, J.; Lin, H.; Wu, J. Comparative analysis of alfalfa (*Medicago sativa* L.) leaf transcriptomes reveals genotype-specific salt tolerance mechanisms. *BMC Plant Biol.* **2018**, *18*, 35. [CrossRef] [PubMed]
71. Lu, C.; Pu, Y.; Liu, Y.; Li, Y.; Qu, J.; Huang, H.; Dai, S. Comparative transcriptomics and weighted gene co-expression correlation network analysis (WGCNA) reveal potential regulation mechanism of carotenoid accumulation in Chrysanthemum× morifolium. *Plant Physiol. Biochem.* **2019**, *142*, 415–428. [CrossRef]
72. Zhu, M.; Xie, H.; Wei, X.; Dossa, K.; Yu, Y.; Hui, S.; Tang, G.; Zeng, X.; Yu, Y.; Hu, P.; et al. WGCNA analysis of salt-responsive core transcriptome identifies novel hub genes in rice. *Genes* **2019**, *10*, 719. [CrossRef]
73. Meng, H.L.; Sun, P.Y.; Wang, J.R.; Sun, X.Q.; Zheng, C.Z.; Fan, T.; Chen, Q.F.; Li, H.Y. Comparative physiological, transcriptomic, and WGCNA analyses reveal the key genes and regulatory pathways associated with drought tolerance in Tartary buckwheat. *Front. Plant Sci.* **2022**, *13*, 985088. [CrossRef]
74. Ali, Q.; Ahsan, M. Estimation of genetic variability and correlation analysis for quantitative traits in chickpea (*Cicer arietinum* L.). *Int. J. Agro Vet. Med. Sci.* **2012**, *6*, 241–249.
75. Malik, H.N.; Malik, S.I.; Hussain, M.O.Z.A.M.I.L.; Chughtai, S.U.R.; Javed, H.I. Genetic correlation among various quantitative characters in maize (*Zea mays* L.) hybrids. *J. Agric. Social. Sci.* **2005**, *3*, 262–265.
76. Diniz-Filho, J.A.F.; Soares, T.N.; Lima, J.S.; Dobrovolski, R.; Landeiro, V.L.; Telles, M.P.D.C.; Rangel, T.F.; Bini, L.M. Mantel test in population genetics. *Genet. Mol. Biol.* **2013**, *36*, 475–485. [CrossRef]
77. Prysiazhniuk, L.; Starychenko, Y.; Hryniv, S.; Shytikova, Y.; Leshchuk, N.; Melnyk, S.; Kliachenko, O. The comparison analysis of software for Mantel test between DNA markers and morphological traits of plant varieties. *AGROFOR Int. J.* **2021**, *6*, 132–143. [CrossRef]
78. Xiang, S.; Wang, H.; Tian, L.; Chen, Y.; Sun, W.; Zhou, Q. Effects of temperature on germination characteristics of different varieties of alfalfa seeds. *Grassl. Turf* **2022**, *42*, 74–82. [CrossRef]
79. Liang, W.; He, Z.; Li, X.; Li, S.; Shi, H.; Xu, Z.; Yue, Y.; Lu, Z. Grey correlation analysis of main agronomic traits and yield of new alfalfa varieties. *Feed. Res.* **2022**, *45*, 86–89.
80. Ritz, C. Toward a unified approach to dose–response modeling in ecotoxicology. *Environ. Toxicol. Chem.* **2010**, *29*, 220–229. [CrossRef] [PubMed]
81. Kodinariya, T.M.; Makwana, P.R. Review on determining number of Cluster in K-Means Clustering. *Int. J.* **2013**, *1*, 90–95.
82. Bao, S. *Soil Agrochemical Analysis*, 3rd ed.; China Agriculture Press: Beijing, China, 2000.
83. Pan, R. *Plant Physiology*; Higher Education Press: Beijing, China, 2005.
84. Long, R.; Zhang, F.; Zhang, Z.; Li, M.; Chen, L.; Wang, X.; Liu, W.; Zhang, T.; Yu, L.X.; He, F.; et al. Genome assembly of alfalfa cultivar zhongmu-4 and identification of SNPs associated with agronomic traits. *Genom. Proteom. Bioinform.* **2022**, *20*, 14–28. [CrossRef]
85. Pertea, M.; Kim, D.; Pertea, G.M.; Leek, J.T.; Salzberg, S.L. Transcript-level expression analysis of RNA-seq experiments with HISAT, StringTie and Ballgown. *Nat. Protoc.* **2016**, *11*, 1650–1667. [CrossRef]
86. Kovaka, S.; Zimin, A.V.; Pertea, G.M.; Razaghi, R.; Salzberg, S.L.; Pertea, M. Transcriptome assembly from long-read RNA-seq alignments with StringTie2. *Genome Biol.* **2019**, *20*, 278. [CrossRef]
87. Pertea, M.; Pertea, G.M.; Antonescu, C.M.; Chang, T.C.; Mendell, J.T.; Salzberg, S.L. StringTie enables improved reconstruction of a transcriptome from RNA-seq reads. *Nat. Biotechnol.* **2015**, *33*, 290–295. [CrossRef]
88. Wang, X.; Fu, Y.; Ban, L.; Wang, Z.; Feng, G.; Li, J.; Gao, H. Selection of reliable reference genes for quantitative real-time RT-PCR in alfalfa. *Genes Genet. Syst.* **2015**, *90*, 175–180. [CrossRef]
89. Lyu, F.; Han, F.; Ge, C.; Mao, W.; Chen, L.; Hu, H.; Chen, G.; Lang, Q.; Fang, C. OmicStudio: A composable bioinformatics cloud platform with real-time feedback that can generate high-quality graphs for publication. *Imeta* **2023**, *2*, e85. [CrossRef]
90. Niissalo, A. *Cytoscape and Its Plugins*; Department of Computer Science University of Helsinki Finland: Helsinki, Finland, 2007.
91. Somers, K.M.; Jackson, D.A. Putting the Mantel test back together again. *Ecology* **2022**, *103*, e3780. [CrossRef] [PubMed]

plants

MDPI

Article

Putative Allele of *D10* Gene Alters Rice Tiller Response to Nitrogen

Tamanna Islam Rimi [1,2,†], Meirong Zhang [1,2,†], Ruixin Zhang [1,†], Zhe Zhang [1], Xueyu Leng [1], Jiafang Han [1], Sihan Meng [1], Wen Du [1] and Zhongchen Zhang [1,2,*]

[1] College of Agriculture, Northeast Agricultural University, 600 Changjiang Road, Harbin 150030, China; tamannaislam2209@yahoo.com (T.I.R.); 18846914861@163.com (M.Z.); 19917625056@163.com (R.Z.); jiudragon1215@163.com (Z.Z.); 17561746702@163.com (X.L.); hanjiafang0@163.com (J.H.); msh528172023@163.com (S.M.); 18245666208@163.com (W.D.)

[2] National Key Laboratory of Smart Farm Technologies and Systems, Harbin 150030, China

* Correspondence: zzcneau@neau.edu.cn

† These authors contributed equally to this work.

Abstract: The number of tillers in rice significantly affects final yield, making it a key trait for breeding and nitrogen-efficient cultivation. By investigating agronomic characteristics, we analyzed phenotypic differences between the wild-type P47-1 and the mutant *p47dt1*, performing genetic analysis and gene mapping through population construction and BSA sequencing. The *p47dt1* mutant, exhibiting dwarfism and multiple tillering, is controlled by a single gene, *P47DT1*, which is tightly linked to *D10*. A single base mutation (T to G) on chromosome 1 alters methionine to arginine, supporting *D10* as the candidate gene for *p47dt1*. To investigate nitrogen response in tillering, KY131 (nitrogen-inefficient) and KY131$^{OsTCP19\text{-H}}$ (nitrogen-efficient) materials differing in *TCP19* expression levels were analyzed. Promoter analysis of *D10* identified TCP19 as a nitrogen-responsive transcription factor, suggesting *D10*'s potential role in a TCP19-mediated nitrogen response pathway. Further analysis of P47-1, p47dt1, KY131, and KY131$^{OsTCP19\text{-H}}$ under different nitrogen concentrations revealed *p47dt1*'s distinct tiller response to nitrogen, altered nitrogen content in stems and leaves, and changes in *TCP19* expression. Additionally, *D10* and *TCP19* expression levels were lower in KY131$^{OsTCP19\text{-H}}$ than KY131 under identical conditions. In summary, *P47DT1*/*D10* appears to modulate nitrogen response and distribution in rice, affecting tiller response, possibly under TCP19's regulatory influence.

Keywords: rice (*Oryza sativa* L.); dwarf and multiple tillering; *P47DT1*; nitrogen response

Citation: Rimi, T.I.; Zhang, M.; Zhang, R.; Zhang, Z.; Leng, X.; Han, J.; Meng, S.; Du, W.; Zhang, Z. Putative Allele of *D10* Gene Alters Rice Tiller Response to Nitrogen. *Plants* **2024**, *13*, 3349. https://doi.org/10.3390/plants13233349

Academic Editor: Alberto Gianinetti

Received: 6 July 2024
Revised: 23 November 2024
Accepted: 25 November 2024
Published: 29 November 2024

1. Introduction

Rice (*Oryza sativa* L.) is one of the world's most important staple crops, providing food for over half of the global population [1]. In China, rice is a primary grain crop, and its production is closely tied to the nation's food security. The number of tillers, which are side shoots that grow from the base of the rice plant, is a key determinant of rice yield. Each tiller can produce a panicle, which directly contributes to grain production. Optimizing the tiller number is a critical objective in rice breeding programs to enhance yield potential [2,3]. As one of the essential nutrient elements in plant growth and development, nitrogen is very important for the growth of rice, especially the growth of rice tillers [4]. Therefore, studying the genes related to nitrogen response in rice is crucial for improving rice yield.

The conditions affecting rice tillering are complex, involving both endogenous hormones and external environmental factors [5]. Plant hormones, such as gibberellins (GAs), cytokinins (CKs), brassinosteroids (BRs), auxins, and strigolactones (SLs), regulate plant structure in rice [6]. Current research indicates that cytokinin, auxin, and strigolactone are the main factors affecting rice tillering [7–11]. Cytokinins promote tiller initiation by stimulating the outgrowth of axillary buds, while auxins, which are transported downward from the shoot apex, suppress tillering by maintaining apical dominance, preventing axillary

bud growth. Strigolactones, discovered as a new class of plant hormones in 2008, play an inhibitory role in tiller development by suppressing the outgrowth of axillary buds [12,13]. Mutations in SL biosynthesis or signaling can lead to plant dwarfing and an increased tiller number [6].

Several key genes involved in the strigolactone synthesis pathway have been identified in rice. For example, the *D17*, *D10*, and *D27* genes encode enzymes that are critical for SL biosynthesis. *D17* encodes OsCCD7, which inhibits tiller growth by limiting lateral branching, thereby negatively regulating tillering [14,15]. Similarly, *D10*, encoding OsCCD8, functions in the carotenoid cleavage pathway [16], and its expression is induced by auxin, which can reduce auxin transport capacity and promote cytokinin synthesis [17,18]. *D27*, a newly identified gene in the strigolactone synthesis pathway, enhances auxin transport polarity when mutated, leading to an increased tiller number and a dwarfing multi-tillering phenotype [19,20].

In addition to these hormonal controls, brassinosteroids (BRs), a class of polyhydroxy steroids, also regulate rice tillering. BR signaling is mediated by the DLT transcription factor (involved in brassinosteroid (BR) signaling, essential for regulating cell growth and tillering), which promotes cell division and elongation, impacting tiller formation [21,22]. The interaction between nitrogen availability and BR signaling is evident through the regulation of OsTCP19, a transcription factor that combines with the DLT promoter to control brassinosteroid synthesis and tiller development [23]. Understanding how nitrogen interacts with BRs is crucial for unraveling the hormonal regulation of rice tillering.

The main environmental factors affecting rice tillering include temperature, light, water, and nutrients [24]. However, in actual field production, temperature, light, water, and other factors are usually stable and are not factors that can cause significant changes in the tillering of rice. In general, producers can directly and effectively regulate nutrient levels only, especially nitrogen levels. Nitrogen affects plant growth and development mainly through the nutrient level and signal transduction pathway, and there is a cooperative regulation relationship between these two aspects [25]. It has been shown to decrease strigolactone content in rice nodes, indicating that nitrogen may promote tillering by inhibiting the SL pathway. This nitrogen–strigolactone interaction is further supported by the finding that nitrogen application increases cytokinin levels, promoting tiller formation [26]. Moreover, in strigolactone-deficient mutants, such as *d3* and *d10*, nitrogen content in the leaves is reduced compared to wild-type plants. Treatment with GR24 (a synthetic strigolactone analog used to study SL signaling and its effects on plant development) restores normal nitrogen content in these mutants, highlighting the importance of strigolactones in coordinating nitrogen distribution and tillering [27,28].

Despite these advances, the molecular mechanisms through which nitrogen and plant hormones interact to regulate tillering remain largely unknown. This study aims to explore the relationship between nitrogen, plant hormones, and rice tillering by focusing on the dwarf and multi-tillering mutant *p47dt1*, thereby contributing to a better understanding of how nitrogen-responsive genes influence rice tillering.

2. Results

2.1. p47dt1 Mutant Showed Dwarfing Multiple Tillers

This study found a dwarf and multiple tillering mutant *p47dt1,* isolated from high-generation line P47, in the field environment. The tillering number of *p47dt1* was higher than that of P47-1 under outdoor potted planting conditions. Plant height of *p47dt1* was significantly lower than that of P47-1 (Figure 1a,b). The tillers of P47-1 were tall and stout, while those of *p47dt1* were many, short and thin. In addition, there were significant differences between P47-1 and *p47dt1* in grain phenotypes after maturity; P47-1 was larger than *p47dt1* in length, width, and thickness (Figure 1c).

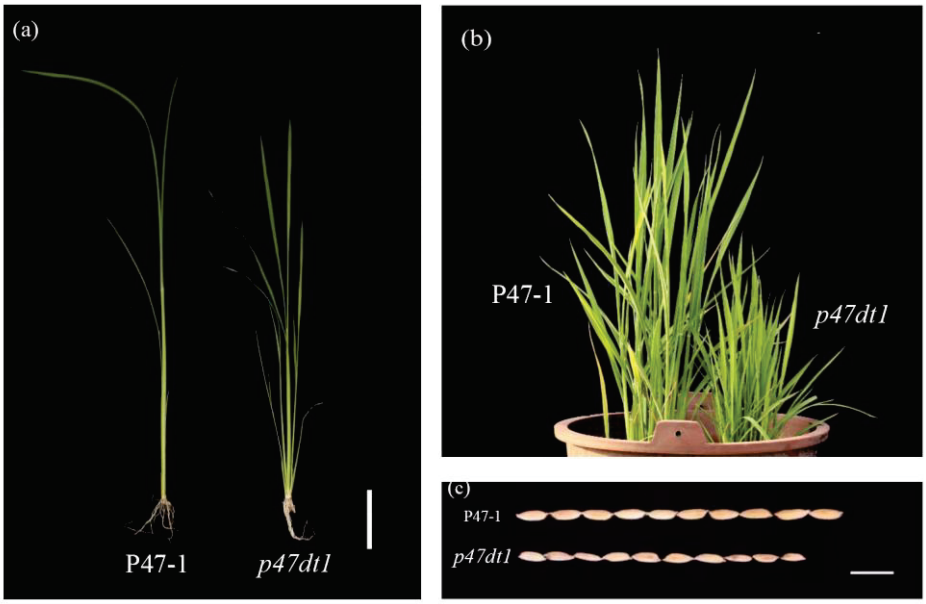

Figure 1. Phenotypes of P47-1 and *p47dt1*. (**a**) Seedling phenotypes of P47-1 and *p47dt1*, with a scale of 2 cm; (**b**) tiller stage phenotypes of P47-1and *p47dt1*, with a scale of 10 cm; (**c**) grain phenotypes of P47-1 and *p47dt1*, with a scale of 1 cm.

Due to the obvious differences in phenotype, agronomic trait parameters of the mutant and wild type were measured, and it was found that the main agronomic traits of the two showed significant differences in many aspects (Table 1). The average plant height of P47-1was 94.80 cm, and that of *p47dt1* was 55.33 cm, 41.6% lower than that of P47-1. The average tiller number of P47-1 was 16.1, and that of *p47dt1* was 84.5, 5.26 times that of P47-1. The average effective panicle number of P47-1 was 16.3, and that of *p47dt1* was 34.3, 2.1 times that of P47-1. Although the number of effective panicles of *p47dt1* was more than that of P47-1, the number of solid grains, 1000-grain weight and seed setting rate of *p47dt1* were significantly lower than those of P47-1, which indicated that, although the number of tillers of *p47dt1* was significantly increased, most of them were ineffective tillers, and the seed setting rate and 1000-grain weight were very low. This mutation would lead to a decrease in yield.

Table 1. Comparison of agronomic traits between P47-1 and *p47dt1*.

Agronomic Traits	P47-1	*p47dt1*
Plant height (cm)	94.80 ± 2.73	55.33 ± 3.79 **
Number of tillers	16.1 ± 1.9	84.5 ± 14.0 ***
Panicle length (cm)	16.33 ± 0.57	10.73 ± 0.21 **
Effective panicles	16.3 ± 0.5	34.3 ± 4.9 **
Total grains per plant	1884.33 ± 116.42	1219.66 ± 132.34 **
Number of solid grains	1505.33 ± 72.96	369.33 ± 65.49 ***
1000-grain-weight (g)	25.03 ± 0.12	13.53 ± 0.24 **
Seed setting rate (%)	80.00 ± 4.54	30.66 ± 9.56 ***

Data are mean ± standard deviation (n = 10); student's t-test, ** means significant at 0.01 level; *** means extremely significant at 0.001 level.

2.2. Dwarfing Multiple Tiller Traits of p47dt1 Were Controlled by a Single Recessive Gene

To identify the genetic characteristics of the mutant, the F_2 population, which was obtained by backcrossing wild-type P47-1 and mutant *p47dt1* and then self-crossing, was used as the genetic population. One hundred and fifty seeds from the isolated family were sown, and phenotypic observation was carried out. It was found that wild-type phenotypes and dwarf and multiple tillering phenotypes were isolated from F_2 families, including

113 individuals with the P47-1 phenotype and 37 individuals with the *p47dt1* phenotype. The SPSS Chi-Square test showed that the ratio of wild type to mutant supertype followed the 3:1 separation law (Table 2), indicating that a single recessive gene controlled the dwarfing and multi-tillering traits of the mutant *p47dt1*.

Table 2. Genetic analysis of tiller trait in the F_2 population.

Population	Population Size	The Phenotype of P47-1	The Phenotype of *p47dt1*	Value of Expectation	χ^2	$\chi^2_{(0.05)}$
F_2	150	113	37	3:1	0.009	3.84

2.3. P47DT1 as a Putative Allele of D10

SNP data were filtered, and the SNP index was used for correlation analysis. The 99th percentile of the fitted ΔSNP-index value <0.3 was used as the association threshold to extract genes in the association region (Figure 2). This threshold was chosen to focus on SNPs likely to be associated with the genetic trait being studied, improving the accuracy of candidate region mapping. A total of 62 SNPs were identified after screening SNP sites.

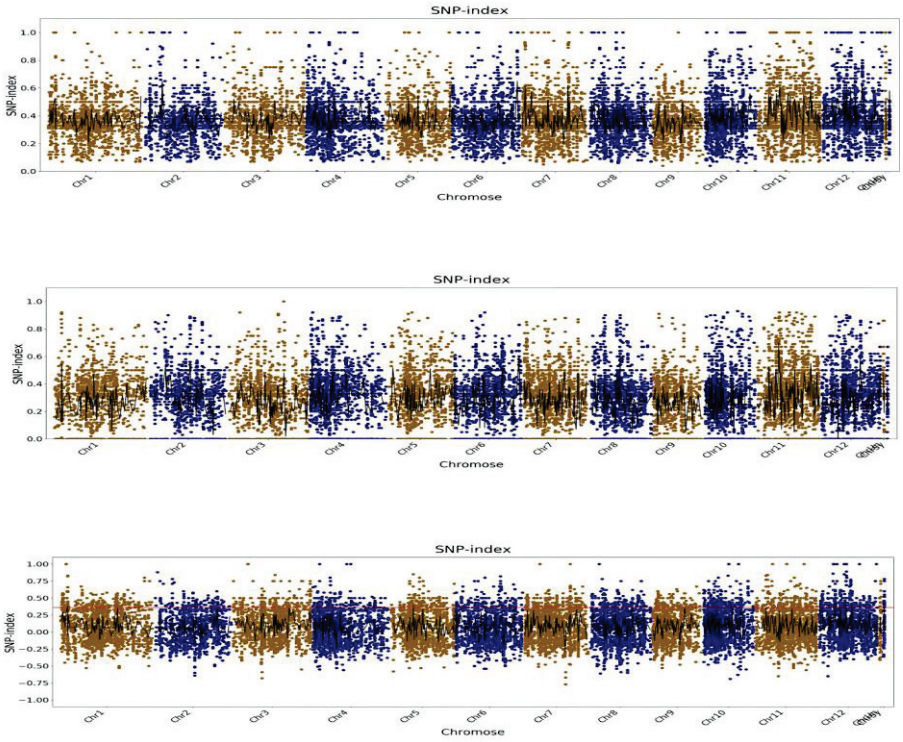

Figure 2. Distribution map of Chromosome SNP-index. The abscissa is the chromosome name, the colored dots represent the calculated SNP-index (or ΔSNP-index) value, and the black line is the fitted SNP-index (or ΔSNP-index) value. The upper picture is the distribution of SNP-index values in the F_2M pool; the middle picture is the distribution of SNP-index values in the F_2W pool; the lower picture is the distribution of ΔSNP-index values, in which the red lines represent 99 percent, respectively. Threshold line for digits.

Upon comparing marker sites, it was found that, among the possible SNP candidate markers, there was a point mutation (T to G) at position 31,226,439 bp on chromosome 1 within the cloned *D10* gene region, located in an exon (Figure 3). This mutation caused the encoded amino acid to change from methionine (M) to arginine (R). To validate this mutation, PCR amplification and gene sequencing were performed on P47-1, *p47dt1*, and

p47dt1's parental mixed DNA samples. The samples were then sent to Shanghai Sangon Biotech Co., Ltd. (Shanghai, China) for gene sequencing.

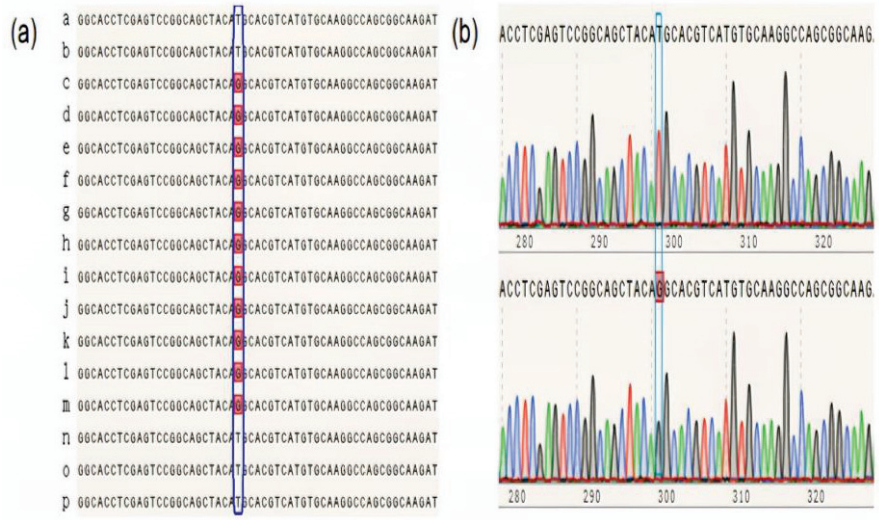

Figure 3. Sequencing analysis of mutation sites in wild-type and mutant and their progeny populations. In figure (**a**), a-b is the mutation site sequence of P47-1; c-d is the mutation site sequence of *p47dt1*; e–m is the mutation site sequence of the *p47dt1* phenotype in the F_2 population; n–p is the mutation site of the P47-1 phenotype in the F_2 population Point sequence. (**b**) Shows the mutation site sequence and peak plot of P47-1; the lower figure shows the mutation site sequence and peak plot of *p47dt1*. The blue squares indicate the positions of mutation sites in the sequences, while the red squares highlight the corresponding peaks in the plot that represent these mutations.

Sequencing results aligned with the BSA findings, confirming a T-to-G mutation at position 31,226,439 on chromosome 1 in *p47dt1*. To investigate the potential linkage between this mutation and the candidate gene *D10*, we selected wild-type and mutant phenotypes from the F_2 segregating population of P47-1/*p47dt1* for pooled sequencing analysis (Figure 3). The results indicated that this mutation site also displayed a T-to-G change in the *p47dt1* offspring population, with a strong association to the dwarf and multi-tillering phenotype (Figures 3 and S1, Table S1). Therefore, these results suggest that *P47DT1* is tightly linked to the candidate gene *D10*, making it a putative allele of *D10*.

2.4. P47DT1 Changes the Response Pathway of Nitrogen Concentration in Rice

Hormone regulation and nutrient level are the main factors regulating the tillering number of rice plants [4,29,30], so we speculated that *P47DT1* may also change the normal nitrogen response pathway of rice. By treating P47-1 and *p47dt1* with three nitrogen concentrations, it was found that P47-1 increased by an average of 6 tillers from low nitrogen to normal nitrogen with a strong nitrogen response, and increased by an average of 0.33 tillers from normal nitrogen to high nitrogen with a weak nitrogen response. However, there was no difference in the tillering number of *p47dt1* under low nitrogen, normal nitrogen and high nitrogen treatments (Figure 4), and the response to nitrogen was always weak, indicating that the mutation of *P47DT1* changed the normal nitrogen response of rice. In addition, according to the difference in tillering number between P47-1 and *p47dt1* under different nitrogen concentrations, the tillering number of P47-1 was less than that of *p47dt1* with the advance of the growth period, suggesting that *P47DT1* was the key gene controlling the tiller response to nitrogen in rice.

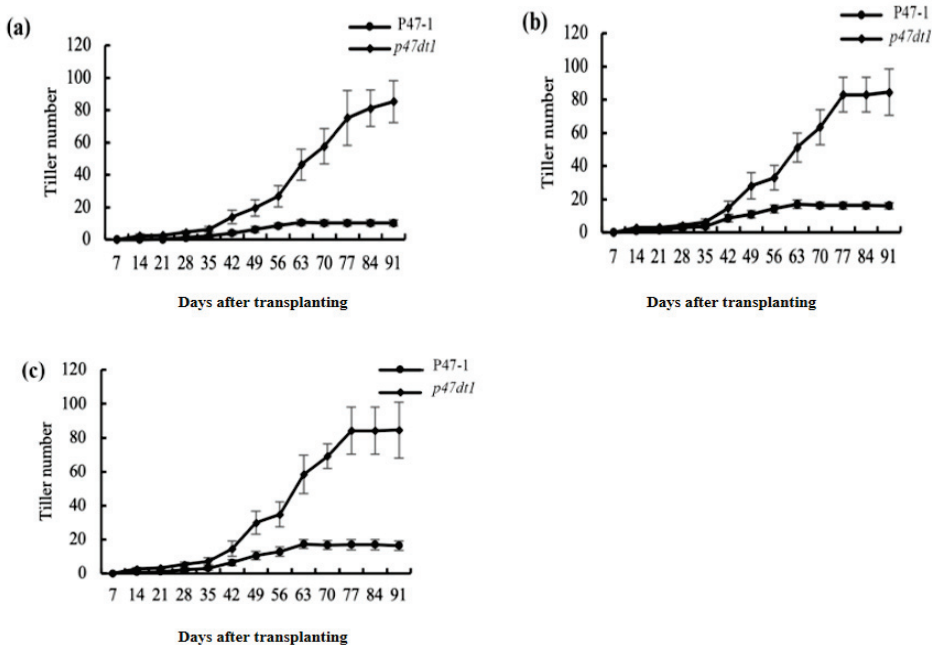

Figure 4. Tiller number and its difference between P47-1 and *p47dt1* under different nitrogen concentrations. (**a**–**c**) Tiller number of P47-1 and *p47dt1* in different periods under low nitrogen (5 g/m^2), normal nitrogen (10 g/m^2) and high nitrogen (15 g/m^2) treatments. The tiller number was measured at different time points (7 to 91 days) after transplanting. Error bars represent the standard error of the mean (SEM) from three biological replicates.

Tillering number determines effective panicle number (the panicles that successfully produce grains, excluding sterile or empty panicles) to some extent, and effective panicle number is the basis of yield formation. Under different nitrogen concentrations, the effective panicle number of P47-1 showed a significant difference under low nitrogen and normal nitrogen treatments but no significant difference under normal nitrogen and high nitrogen treatments, which was consistent with the trend of the tillering number of P47-1 under different nitrogen concentrations. The effective panicle number of *p47dt1* showed significant differences and different trends under different nitrogen concentrations, which increased significantly from low nitrogen to normal nitrogen and decreased significantly from normal nitrogen to high nitrogen (Figure 5). In conclusion, the mutation of *P47DT1* made a significant difference compared to P47-1 in effective panicle number, which also indicated that the mutation affected the normal nitrogen absorption and transformation function of rice.

In addition, the number of solid grains of P47-1 under normal nitrogen treatment was significantly higher than that under low nitrogen treatment, but there was no significant difference between normal nitrogen treatment and high nitrogen treatment, which was consistent with the change law of effective panicle number under different nitrogen treatments. The number of solid grains of *p47dt1* under normal nitrogen treatment was significantly higher than that under low nitrogen and high nitrogen treatment. Different from P47-1, the number of solid grains of *p47dt1* from normal nitrogen to high nitrogen decreased instead of increasing, and the number of solid grains under high nitrogen treatment was less than that under low nitrogen treatment (Figure 6). As the nitrogen concentration increases, the change in grain weight per plant of wild-type and mutant is similar to the change in the number of solid grains (Figure 7). This result indicated that the mutation of *P47DT1* changed the response mode of rice to exogenous nitrogen concentration.

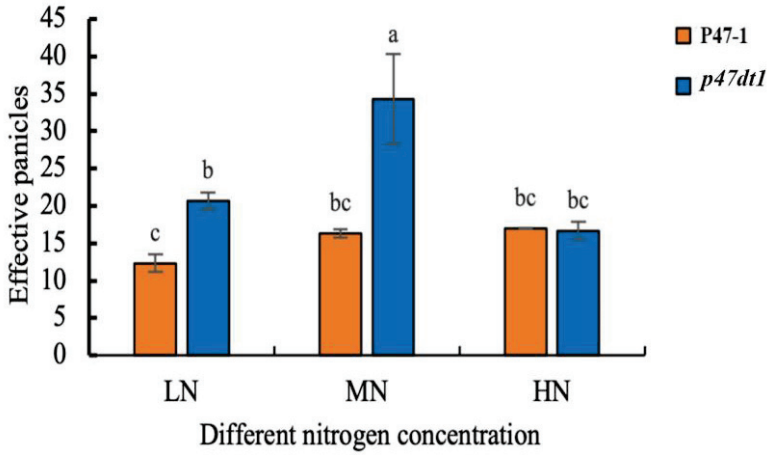

Figure 5. Effective panicles (the panicles that successfully produce grains, excluding sterile or empty panicles) of P47-1 and *p47dt1* under different nitrogen concentrations. Different lowercase letters on the same bar chart indicate significant differences at the 0.05 level. LN represents low nitrogen treatment (5 g/m^2), MN represents normal nitrogen treatment (10 g/m^2), and HN represents high nitrogen treatment (15 g/m^2), as below.

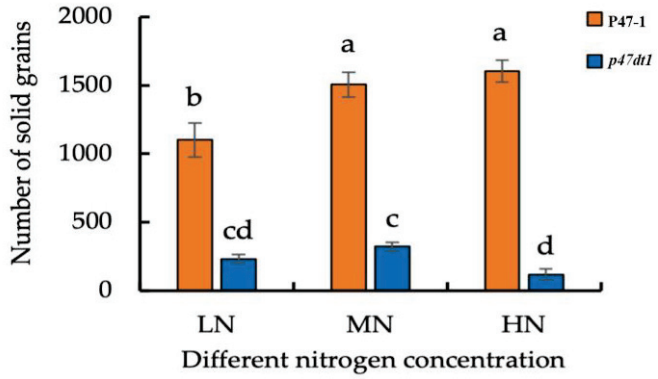

Figure 6. Number of solid grains of P47-1 and *p47dt1* under different nitrogen concentrations. Different lowercase letters on the same bar chart indicate significant differences at the 0.05 level. LN represents low nitrogen treatment (5 g/m^2), MN represents normal nitrogen treatment (10 g/m^2), and HN represents high nitrogen treatment (15 g/m^2).

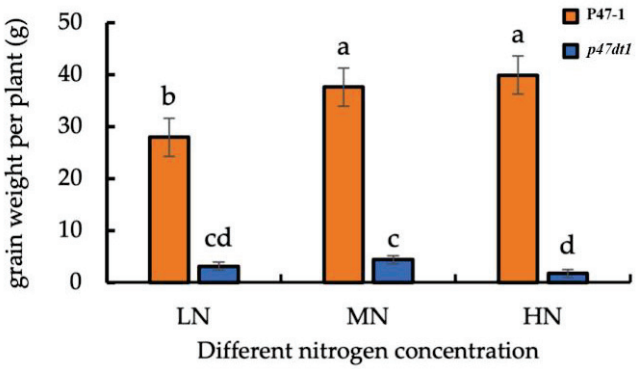

Figure 7. Grain weight per plant for P47-1 and *p47dt1* under different nitrogen concentrations. Different lowercase letters on the same bar chart indicate significant differences at the 0.05 level. LN represents low nitrogen treatment (5 g/m^2), MN represents normal nitrogen treatment (10 g/m^2), and HN represents high nitrogen treatment (15 g/m^2).

2.5. P47DT1 Alters Nitrogen Allocation Patterns in Rice

Rice primarily absorbs nitrogen in the forms of nitrate and ammonium, with each form following distinct uptake and assimilation pathways [31,32]. We hypothesized that the *p47dt1* mutation may affect specific pathways, leading to altered nitrogen responses. To investigate these potential pathway changes, P47-1 and *p47dt1* were treated with potassium nitrate, ammonium nitrate, and ammonium chloride. Results showed that, for P47-1, plant height increased significantly with high potassium nitrate concentration, while tiller numbers remained consistent across treatments. In contrast, treatments with ammonium nitrate and ammonium chloride resulted in significantly greater plant height and tiller numbers under high concentrations. In *p47dt1*, tiller numbers showed minimal differences between high and low concentrations across all nitrogen sources but were notably lower with potassium nitrate compared to ammonium nitrate or ammonium chloride (Figure 8). This suggests that both P47-1 and *p47dt1* display insensitivity to nitrate nitrogen, with *p47dt1* exhibiting an even greater insensitivity. Thus, we inferred that *P47DT1* may influence nitrate nitrogen absorption pathways, making the mutant highly insensitive to nitrate nitrogen.

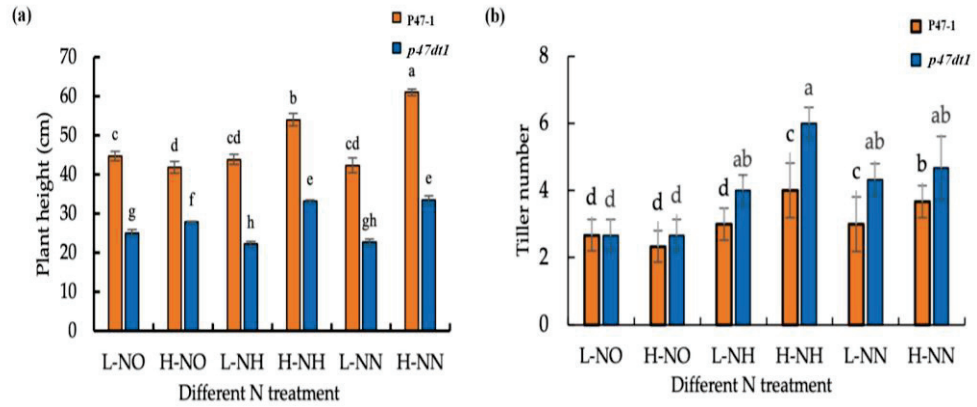

Figure 8. Phenotypic comparison of P47-1and *p47dt1* under hydroponic treatment with different nitrogen conditions. (**a**) Plant height of P47-1 and *p47dt1* under different nitrogen treatments; (**b**) tiller number of P47-1 and *p47dt1* under different nitrogen conditions. The scales in Figures (**a**,**b**) are 15 cm. Different lowercase letters on the same bar chart indicate significant differences at the 0.05 level. L-NO represents 0.2 mM potassium nitrate treatment, H-NO represents 1.5 mM potassium nitrate treatment, L-NH represents 0.2 mM ammonium chloride treatment, H-NH represents 1.5 mM ammonium chloride treatment, L-NN represents 0.2 mM ammonium nitrate treatment, and H-NN represents 1.5 mM ammonium nitrate treatment.

To further assess if *P47DT1* mutation affects nitrogen allocation within the plant, both P47-1 and *p47dt1* were treated under six nitrogen conditions: 0.2 mM potassium nitrate, 1.5 mM potassium nitrate, 0.2 mM ammonium chloride, 1.5 mM ammonium chloride, 0.2 mM ammonium nitrate and 1.5 mM ammonium nitrate. Nitrogen content in P47-1 leaves was significantly higher than in stems across all treatments, while *p47dt1* showed higher nitrogen content in leaves under ammonium-based treatments but had elevated stem nitrogen under low potassium nitrate conditions (Figure 9). These findings indicate that *P47DT1* mutation affects nitrogen allocation between leaves and stems through nitrate nitrogen pathways.

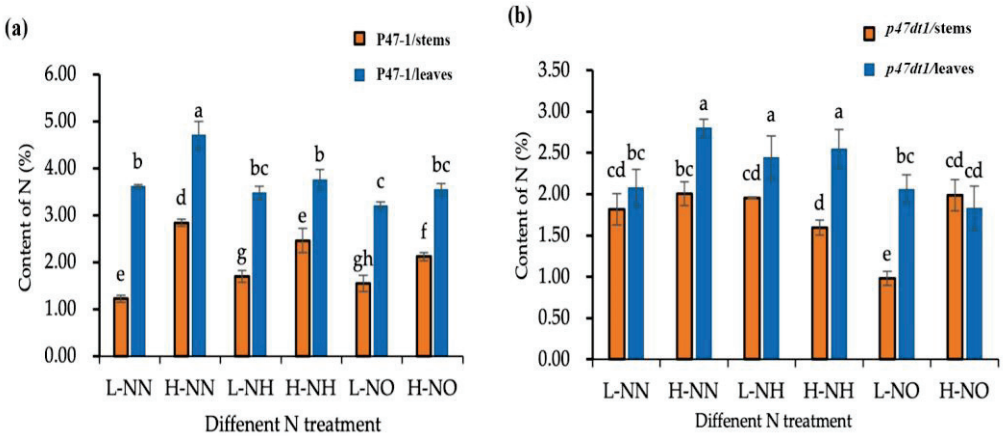

Figure 9. Nitrogen content of P47-1 and *p47dt1* under different nitrogen conditions. (**a**) Comparison of nitrogen content in stems and leaves of wild type P47-1 under different nitrogen conditions; (**b**) comparison of nitrogen contents in stems and leaves of mutant *p47dt1* under different nitrogen conditions. Different lowercase letters on the same bar chart indicate significant differences at the 0.05 level. L-NO represents 0.2 mM potassium nitrate treatment, H-NO represents 1.5 mM potassium nitrate treatment, L-NH represents 0.2 mM ammonium chloride treatment, H-NH represents 1.5 mM ammonium chloride treatment, L-NN represents 0.2 mM ammonium nitrate treatment, and H-NN represents 1.5 mM ammonium nitrate treatment.

2.6. Potential Role of TCP19 in Regulating D10 Expression

To analyze how the *p47dt1* mutation might influence the nitrogen absorption pathway, the transcription factor prediction of the promoter of the *D10* gene was performed by the bioinformatics method. As *D10* is completely linked to *P47DT1*, the 2.4 kb promoter sequence upstream of the start codon of the *D10* gene was predicted, revealing that the promoter of this gene contains a large number of elements related to plant hormone response. *D10* was involved in the biosynthesis of strigolactone derivative SL and should also be involved in biological processes regulated by plant hormones. According to previous experiments, *D10* was also found to participate in nitrogen response [23]. Therefore, transcription factors related to nitrogen and plant hormones were selected as screening conditions for prediction results analysis (Figure 10). A TCP family transcription factor TCP20 (*Arabidopsis thaliana*) was found at Matrix ID 0424. It was speculated that this position might be combined with other TCP family transcription factors in rice, among which the transcription expression of *Os06g0226700 (TCP19)* was negatively regulated by external nitrogen level. As a transcription factor, TCP19 can promote the expression of gene *DLT* and thus realize the regulation of rice tiller development. Both *D10* and *DLT* regulate tillering number in rice through the hormone pathway, and both were related to nitrogen. Therefore, it was speculated that *D10* may be involved in another related pathway in the OsTCP19-mediated nitrogen regulation network.

The expression of *D10* in the wild-type plant P47-1 significantly increased with higher nitrogen concentrations across different nitrogen forms (Figure 11a–c). This is consistent with the feedback mechanism of strigolactone biosynthesis, where reduced strigolactone synthesis caused by nitrogen promotes *D10* expression [33]. However, in the mutant *p47dt1*, this trend was not observed, suggesting that the mutation disrupted the strigolactone synthesis pathway, leading to altered feedback and elevated *D10* expression (Figure 11c).

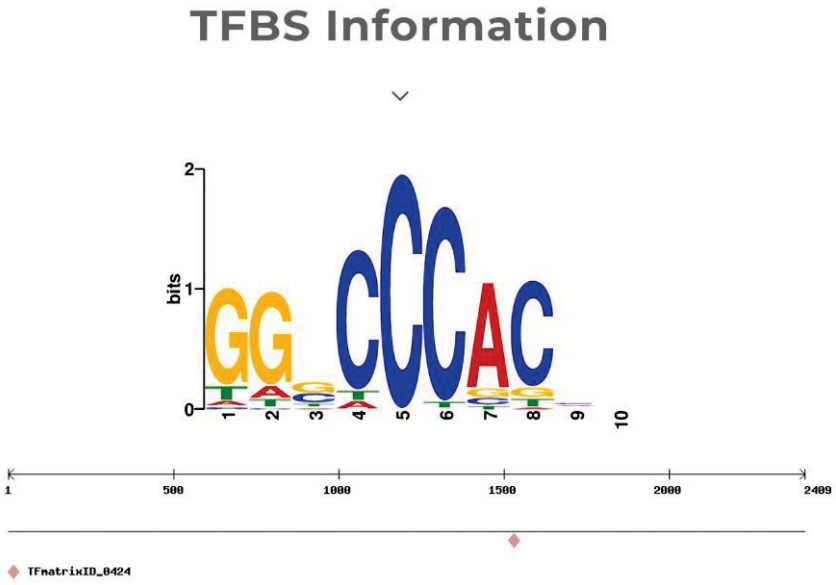

Figure 10. The target sequence of TCP19 and the putative binding position on the *D10* promoter. Here TFBS represents Transcription Factor Binding Site. The TFBS Information visualizes the motif recognized by a particular transcription factor, identified by a matrix (TFmatrixID_0424), with the sequence "GGCCCAC". This shows the binding motif that the transcription factor is likely to interact with in the genome. Blue highlights indicate the regions of the promoter where TCP19 is predicted to bind, while red highlights indicate regions where the transcription factor may have a stronger binding affinity. The figure also illustrates the position of the binding motif in the genome.

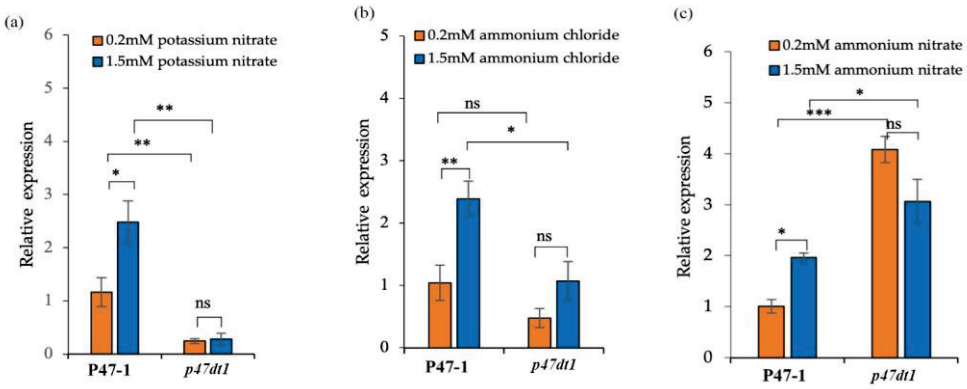

Figure 11. Relative expressions of *D10* in P47-1 and *p47dt1* under different nitrogen conditions. (**a**) Relative expressions of *D10* of P47-1 and *p47dt1* treated with potassium nitrate of 0.2 mM and potassium nitrate of 1.5 mM; (**b**) relative expressions of *D10* of P47-1 and *p47dt1* treated with 0.2 mM ammonium chloride and 1.5 mM ammonium chloride; (**c**) relative expressions of *D10* of P47-1 and *p47dt1* treated with 0.2 mM ammonium nitrate and 1.5 mM ammonium nitrate. Data are mean ± standard deviation (n = 10); student's t-test, * means significant at 0.05 level; ** means significant at 0.01 level; *** means extremely significant at 0.001 level; ns means no significant difference.

To investigate the relationship between TCP19 and *D10*, we used nitrogen-efficient KY131$^{OsTCP19\text{-H}}$ and nitrogen-inefficient KY131. These materials were selected for their contrasting nitrogen efficiency and differences in *TCP19* expression levels. Notably, *TCP19* expression was significantly lower in KY131$^{OsTCP19\text{-H}}$ compared to KY131 under the same nitrogen conditions (Figure 12c) [23]. Additionally, *D10* expression in KY131 and KY131$^{OsTCP19\text{-H}}$ showed a conserved response to increasing nitrogen concentration, aligning with patterns observed in P47-1. This trend suggests a possible involvement of TCP19 in the regulation of *D10* under nitrogen conditions. However, these findings demonstrate an association and are not definitive evidence of a direct regulatory mechanism.

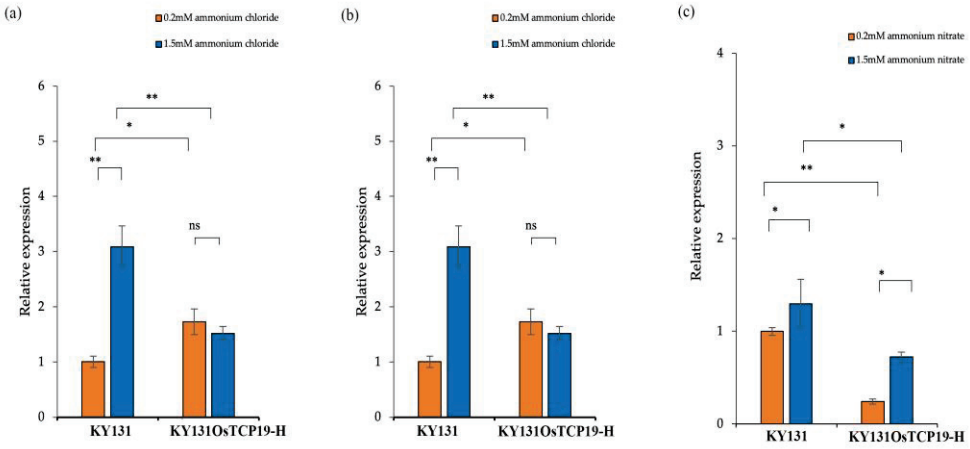

Figure 12. Relative expressions of *D10* in KY131 and KY131$^{OsTCP19\text{-}H}$ under different nitrogen conditions. (**a**) Relative expressions of *D10* of KY131 and KY131$^{OsTCP19\text{-}H}$ treated with potassium nitrate of 0.2 mM and 1.5 mM; (**b**) relative expressions of *D10* of KY131 and KY131$^{OsTCP19\text{-}H}$ treated with 0.2 mM ammonium chloride and 1.5 mM ammonium chloride; (**c**) relative expressions of *D10* of KY131 and KY131$^{OsTCP19\text{-}H}$ treated with 0.2 mM ammonium nitrate and 1.5 mM ammonium nitrate. Data are mean ± standard deviation (*n* = 10); student's *t*-test, * means significant at 0.05 level; ** means significant at 0.01 level; ns means no significant difference.

Previous research has shown that nitrogen negatively regulates TCP19 primarily via the nitrate pathway [23]. As TCP19 can inhibit the expression of the tiller-promoting gene *DLT* through the brassinosteroid pathway, it is speculated that TCP19 might also influence nitrogen responses involving *D10* and strigolactones. However, the exact mechanisms linking TCP19 to *D10* remain speculative and require further validation (Figure 13).

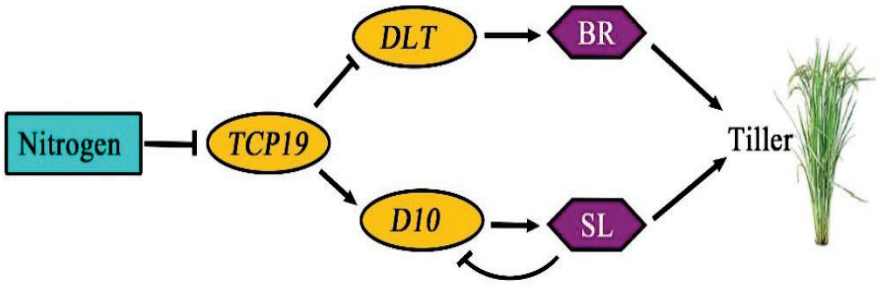

Figure 13. A proposed model for *D10* in TCP19-mediated tiller development response to nitrogen. The diagram shows the nitrogen response pathway influencing rice tillering. TCP19 represents a transcription factor that regulates tiller formation in response to nitrogen. DLT (Dwarf and Low-Tillering transcription factor involved in brassinosteroid (BR) signaling, essential for regulating cell growth and tillering), SL (strigolactones), (BR) brassinosteroid.

In conclusion, while the results suggest that TCP19 may influence *D10* expression and nitrogen response in rice, these findings remain speculative. Further experimental studies are needed to confirm the regulatory pathways and their implications for tiller development.

3. Discussion

3.1. P47DT1 Mutation Affects Nitrogen Absorption Pathways

Many agronomic characteristics, such as plant height, tiller number, effective panicle number, and grain number, were significantly different between P47-1 and *p47dt1* after potted soil culture with low nitrogen, normal nitrogen, and high nitrogen. P47-1 showed a significant tiller response to nitrogen from low nitrogen to normal nitrogen but no significant change in tiller response to nitrogen from normal nitrogen to high nitrogen. This

verifies the conclusion of previous studies that rice's nitrogen absorption has a threshold and, when the nitrogen absorbed by rice meets its own growth and development needs, it will not continue to absorb and utilize soil nitrogen [34]. Therefore, in the actual process of rice production, we should not blindly think that more nitrogen fertilizer can increase the yield. Overuse of nitrogen can result in groundwater contamination, higher production costs, decreased crop yields, and overall environmental pollution [35].

However, *p47dt1* showed almost the same tiller response to nitrogen under low nitrogen, normal nitrogen, and high nitrogen conditions, indicating that the mutation also affected the normal nitrogen response of rice. Although the *p47dt1* mutant exhibited a higher tiller number under high nitrogen concentration, the number of effective panicles did not increase correspondingly. This suggests that the mutation impacts the plant's ability to transform the increased tillering into productive panicles, likely due to reduced nitrogen absorption efficiency and transformation into grains.

When *p47dt1* was treated with nitrate nitrogen, its phenotype changes at high and low concentrations were different from those under other nitrogen conditions. It was preliminarily concluded that *P47DT1/D10* might be related to the nitrate nitrogen absorption pathway, and it was extremely insensitive to nitrate nitrogen after mutation. The *p47dt1* mutation appears to affect nitrogen use efficiency rather than total nitrogen uptake, as evidenced by the lower nitrogen content in the mutant tissues (Figure 9). Although the mutant *p47dt1* produces more tillers, these tillers are less efficient at utilizing nitrogen to form effective panicles, resulting in fewer grains and lower overall yield.

Additionally, although the *p47dt1* mutant forms more tillers, the number of solid grains and the grain weight per plant are significantly lower than in the wild type (Figures 7 and 8). This may be attributed to inefficient nutrient partitioning and nitrogen use in the mutant, where increased tiller production does not result in higher yield. The reduced ability to absorb and effectively utilize nitrogen for grain development may explain the lower solid grain numbers and grain weight in the mutant.

Thus, the *p47dt1* mutation affects nitrogen use efficiency, leading to lower grain yield despite increased tillering. This reduced efficiency is reflected in lower nitrogen content in the plant tissues (Figure 9) and fewer solid grains, as shown in (Figures 6 and 7). This demonstrates that the mutation altered the normal nitrogen uptake and distribution rules of rice, impacting the overall agronomic performance.

3.2. D10's Possible Role in the TCP19-Mediated Nitrogen Response Pathway

There is a conserved feedback inhibition mechanism between the strigolactone synthesis gene *D10* and strigolactone (SL). Under nitrogen limitation, enhanced SL biosynthesis suppresses *D10* expression due to negative feedback regulation. Conversely, when nitrogen is abundant, SL synthesis decreases, relieving feedback inhibition and increasing *D10* expression. In the mutant *p47dt1*, the disruption of the SL synthesis pathway appears to alter this feedback mechanism. As a result, *D10* expression remains elevated, likely due to enhanced feedback effects, with expression levels significantly higher than in the wild-type plant P47-1.

To further explore potential regulatory pathways involving *D10*, transcription factors related to nitrogen and plant hormone signaling were predicted based on the *D10* promoter sequence. Among these, TCP19 was identified as a candidate transcription factor potentially influencing nitrogen responses in rice. Previous studies have suggested that TCP19 might regulate the expression of tiller-promoting genes through pathways such as brassinosteroids, and its expression is responsive to nitrogen. However, the specific interaction between TCP19 and *D10* remains speculative.

The relevance of KY131 and KY131$^{OsTCP19\text{-H}}$ to the present study lies in their contrasting nitrogen efficiency, which provides a comparative framework for investigating the role of TCP19 and *D10* in nitrogen response and tillering regulation. KY131 is nitrogen-inefficient, while KY131$^{OsTCP19\text{-H}}$ is nitrogen-efficient, presumably due to alterations in *TCP19* expression levels. This distinction allows researchers to explore the relationship

between nitrogen-responsive transcription factors and their downstream effects on genes like *D10*, which are involved in the nitrogen absorption and tillering pathways. Using these materials, it was observed that *TCP19* expression was significantly lower in KY131$^{OsTCP19\text{-}H}$ under the same nitrogen conditions. This trend was consistent across various nitrogen treatments, aligning with observations in P47-1. These findings suggest that TCP19 may influence *D10* expression under nitrogen-regulated conditions. However, it should be noted that the observed association between *TCP19* and *D10* expression patterns is not definitive evidence of a direct regulatory relationship. Instead, these findings highlight a possible involvement of TCP19 in the nitrogen response pathway in which *D10* might participate.

Although *TCP19* and *D10* exhibit significant expression differences between nitrogen-efficient and nitrogen-inefficient materials, the exact regulatory mechanism remains unclear. Further studies are needed to determine whether TCP19 directly regulates *D10* or if the observed patterns result from broader transcriptional networks responding to nitrogen. Therefore, while our results suggest that *P47DT1/D10* modulates nitrogen response and tiller development, this regulation may occur indirectly or through additional pathways influenced by TCP19.

In conclusion, *D10* is likely to participate in a TCP19-mediated nitrogen response pathway, potentially affecting tiller responses under varying nitrogen conditions. However, this conclusion is speculative, and additional studies are required to validate these interactions and their implications for rice tillering regulation.

4. Materials and Methods

4.1. Experimental Materials

The experiment was conducted in the pot farm of Northeast Agricultural University (126.72° E, 45.75° N) in 2021. The mutant *p47dt1* was used as the test material and the wild-typeP47-1 was used as the control material for phenotypic analysis, genetic analysis, gene mapping and nitrogen response analysis. KY131$^{OsTCP19\text{-}H}$ was used as test material and KY131 was used as control material for nitrogen response analysis The lines KY131 and KY131$^{OsTCP19\text{-}H}$ were provided by Professor Chu Chengcai from South China Agricultural University and are used for nitrogen response analysis, and the natural mutant *p47dt1* of the high-generation rice material P47-1 was discovered in a field environment.

4.2. Population Construction and Separation Statistics

F_2 population used was derived from a cross between the P47-1 wild-type rice and the *p47dt1* mutant, which exhibits a dwarf and multi-tillering phenotype. After obtaining F_1 seeds from the cross, these seeds were self-pollinated to generate the F_2 population.

All the F_2 generation seeds harvested underwent broken dormancy for 16 h with 0.6% dilute HNO_3, then washed three times with clean water to wash away the dilute nitric acid and replaced with distilled water to promote germination in the dark at 37 °C. After the seeds' dewing, they were sown in the rice nutrient soil for seedling cultivation (Culture temperature: 30 °C in the daytime, 22 °C at night; photoperiod: 12 h during the day, 12 h at night; ambient humidity: relative humidity (RH) is about 80%; light intensity: 30,000 Lux). At the tillering stage, the number of individuals with P47-1 phenotype and *p47dt1* phenotype isolated from F_2 generation was counted based on phenotypic differences.

4.3. Gene Mapping

The CTAB method was used to extract DNA from 30 independent P47-1 phenotypic individuals, 30 independent *p47dt1* phenotypic individuals, and the parental line (P47-1). For each genotype, a total of 30 DNA samples were collected. After extraction, the DNA samples were pooled by genotype, resulting in three bulk DNA samples: (1) a bulk sample consisting of DNA from 30 P47-1 phenotypic individuals; (2) a bulk sample consisting of DNA from 30 *p47dt1* phenotypic individuals; (3) a bulk sample consisting of DNA from the P47-1 parental line.

These three pooled DNA samples were then sent to Frasergen Bioinformatics Co., Ltd. (Wuhan, China) for Bulked Segregant Analysis (BSA) mixed sequencing analysis. The test process includes sample testing, library preparation, library quality testing, and computer sequencing. In order to ensure the quality of information analysis, clean reads were obtained by fine filtering of raw reads, and subsequent analysis was conducted based on clean reads. Specific filtration standards are (1) SNP loci with multiple genotypes, unidentified SNP loci; (2) SNP loci with support of less than 4; (3) SNP loci of mutant progeny with mixed pool genotype consistent with wild type parent genotype or homozygous loci with genotypic consistency among offspring mixed pools; (4) wild type parents heterozygous loci; (5) the SNP-index of the mixed pool genotypes of the mutant and wild offspring was less than 0.3.

The SNP-index method was used to select the candidate regions related to the target traits and their mutation locations and mutation types. SNP-index difference for two offspring pools with extreme traits (ΔSNP index) were used to obtain the ΔSNP-index distribution. Candidate regions were screened within a 0.01 confidence interval. SNP markers and InDel markers were screened. After screening, primers were designed according to the location of candidate sites, and PCR amplification was performed. The amplified products were sent to Shanghai Sangong Biological Co., Ltd. (Shanghai, China) for sequencing.

4.4. Pot Test Scheme

Three gradient fertilizer treatments were set, respectively: high nitrogen (50% higher than the normal amount, HN): 15 g/m^2, normal nitrogen (control, MN): 10 g/m^2, low nitrogen (50% lower than the normal amount, LN): 5 g/m^2; 10 replicates were set for each treatment, and all treatments of phosphate fertilizer, potassium fertilizer and organic fertilizer were consistent. Nitrogen fertilizer is urea (N $\geq$ 46%), phosphate fertilizer is granular heavy superphosphate calcium ($P_2O_5 \geq 46\%$), and potassium fertilizer is granular potassium chloride ($K_2O \geq 60\%$). According to the surface area of the basin, to convert the nitrogen, phosphorus, and potassium fertilizer application amount of each basin, the ratio of nitrogen, phosphorus, and potassium was 1:0.5:0.8. The fertilization method uses all phosphate fertilizer as base fertilizer; 50% of potassium fertilizer is used as base fertilizer and 50% as ear fertilizer; 50% of Nitrogen fertilizer is used as base fertilizer, 30% as tillering fertilizer, and 20% as ear fertilizer. The application was divided into three stages: before transplanting (base fertilizer application in mid-May), at the end of rejuvenation (tiller fertilizer application in late May), and at the booting stage (ear fertilizer application in early July).

4.5. Hydroponics Test Scheme

Because the content of $(NH_4)_6Mo_7O_{24}{\cdot}4H_2O$ in the formula of the International Rice Research Institute [31] is very low, it is only necessary to replace and adjust the amount and type of NH_4NO_3 in the mother liquor for nitrogen-treated medium with different forms and concentrations. For specific adjustments, please refer to (Table 3). Seeds of P47-1, *p47dt1*, KY131, and KY131$^{OsTCP19\text{-}H}$ were sowed on the mesh plates, and the mesh plates were placed in Syngenta transparent boxes without any treated tap water. Replace the water every 3 days until the nutrition of the seeds is exhausted (pinch the seeds to show a withered state), then use different nitrogen treatments of culture medium for culture, each treatment consisting of 3 repeats every 3 days to replace the culture medium. On the 21st day of culture in the growth chamber, the plant height and tiller number of P47-1 and *p47dt1* were measured, and materials consistent with the mean value were selected for photo recording. On the 33rd day, roots of four kinds of materials were taken under different treatment conditions for RNA extraction, reverse transcription, fluorescence quantification, and other tests.

Table 3. Adjustment table of nitrogen nutrient mother solution.

Stock-2 Drugs	0.2 (mol/L)	1.5 (mol/L)
NH_4Cl (53.49 g/mol)	10.69 (g/L)	80.23 (g/L)
KNO_3 (101.1 g/mol)	20.22 (g/L)	151.65 (g/L)
NH_4NO_3 (80 g/mol)	16 (g/L)	120 (g/L)

4.6. Nitrogen Content Determination

Leaves and stems of P47-1 and *p47dt1* treated with different nitrogen conditions were dried, ground into powder, and boiled. The total nitrogen content in leaves and stems was determined by the Kjeldahl nitrogen determination method [36].

4.7. Transcription Factor Prediction

The *D10* promoter gene sequence was obtained from the NCBI website (https://www.ncbi.nlm.nih.gov), and then this sequence was entered into the promoter analysis function dialogue box of PLANTPAN online website (https://plantpan.itps.ncku.edu.tw), and rice was selected as the analyzed species to perform transcription factor prediction analysis of promoter binding.

4.8. Determination of Gene Expression

RNA from the roots of P47-1, *p47dt1*, KY131, and KY131$^{OsTCP19\text{-H}}$, treated with different nitrogen conditions, was extracted using the Trizol method. The sequence of gene *D10* was obtained at NCBI, and the accuracy of the sequence was tested by BLAST. The online website Primer 3 (https://primer3.ut.ee) was used to screen the CDS region sequence of *D10* and, after designing the primer length and product size (Table 4), appropriate *D10* and *TCP19* primers were selected and synthesized by Beijing Ruiboxingke Biological Co., Ltd. (Beijing, China) [20]. RNA reverse transcription was performed using Novozan HiScript III 1st Strand cDNA Synthesis Kit (+gDNA wiper), and then quantitative qRT–PCR was performed on cDNA of rice root samples. *D10* and *TCP19* gene expression data were analyzed and processed by the $2^{-\Delta\Delta Ct}$ method. the QUBI gene was used as the reference gene for normalizing the expression levels of the target genes.

Table 4. qRT–PCR primer information.

Gene Name	Accession No.	Forward-Primer (5′-3′)	Reverse-Primer (5′-3′)
D10	LOC_Os01g54270	GGAAGAGTGTACGGCAGGAG	GTAGTCGCCGAGGTTCCATA
TCP19	LOC_Os06g12230	GACAGTGTACCGTGGCGT	CGCCGGGAAGTTCATGAAAT
QUBI	LOC_Os03g13170	GCTCCGTGGCGGTATCAT	CGGCAGTTGACAGCCCTAG

4.9. Data Analysis

Analysis of variance (ANOVA), followed by post hoc least significant difference (LSD), was carried out using SPSS software (IBM Corp., Armonk, NY, USA). LSD at 5% ($p \leq 0.05$) probability levels was employed for pairwise means comparison analysis. A Two-Way ANOVA was performed to analyze the effects of two independent variables: nitrogen concentration (levels: LN, MN, HN for Figures 5–7; and levels: L-NO, H-NO, L-NH, H-NH, L-NN, H-NN for Figures 9 and 10) and variety (wild-type P47-1 and mutant *p47dt1*) on the dependent variables. This approach allowed precise identification of treatment effects and interactions under controlled conditions [37].

4.10. Graphical Software

The experimental data were mainly processed using Excel 2007, SPSS 22.0, Graph prism 8, Photoshop CS5, and other software for data significance analysis and chart making.

5. Conclusions

The dwarf and multiple tillering traits of the *p47dt1* mutant are controlled by a single recessive gene, and gene mapping suggests that *P47DT1* is a putative allele of *D10*. The mutation in *P47DT1* appears to alter the nitrogen response pathway and the general pattern of nitrogen allocation in rice, potentially involving the nitrate nitrogen pathway. This mutation may also interact with pathways regulating tillering in rice, including those mediated by TCP19. While our findings suggest a possible positive regulation of *P47DT1*/*D10* by TCP19, further experimental validation is required to clarify the underlying mechanisms.

This study also demonstrated that more nitrogen does not always lead to better yield, highlighting the importance of efficient nitrogen management. The mutation increases tillering but does not enhance effective grain production, suggesting that future breeding efforts should focus on achieving a balance between tiller number and grain formation for improved productivity. Furthermore, insights into *P47DT1* and TCP19 can guide breeding programs to develop rice varieties with better nitrogen use efficiency, contributing to higher yields with fewer resources. This research underscores the importance of nitrogen-efficient crop varieties for sustainable agriculture, reducing fertilizer use while maintaining high productivity.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/plants13233349/s1, Table S1: Primer information for gene identification and Figure S1: Mutation positions of the *d10* mutant versus the *p47dt1* mutant.

Author Contributions: Conceptualization, T.I.R. and S.M.; methodology, M.Z. and R.Z.; software, T.I.R. and W.D.; validation, M.Z. and R.Z.; formal analysis R.Z., Z.Z. (Zhe Zhang) and Z.Z. (Zhongchen Zhang); investigation, Z.Z. (Zhongchen Zhang); writing—original draft preparation, T.I.R., M.Z., Z.Z. (Zhe Zhang) and X.L.; writing—review and editing, T.I.R. and J.H.; supervision, Z.Z. (Zhongchen Zhang); funding acquisition, Z.Z. (Zhongchen Zhang). All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Key Laboratory of Smart Farm Technologies and Systems (JD2023GJ01-12), the National Key Research and Development Program (2016YFD0300604-4) and the Natural Science Foundation of Heilongjiang Province (LH2020C006).

Data Availability Statement: The datasets presented in this study can be found in online repositories. The names of the repository/repositories and accession number(s) can be found below: NGDC-GSA database under the BioProject no. PRJCA031448 and accession no. CRA019863 for the BSA-seq data.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Liu, T.; Zhang, X.; Zhang, H.; Cheng, Z.; Liu, J.; Zhou, C.; Luo, S.; Luo, W.; Li, S.; Xing, X. Dwarf and High Tillering1 represses rice tillering through mediating the splicing of *D14* pre-mRNA. *Plant Cell* **2022**, *34*, 3301–3318. [CrossRef] [PubMed]
2. Li, H.; He, W.; Lian, L.; Wei, Y.; Cai, Q. Research progress of rice plant type. *Nat. Sci. Ed.* **2020**, *38*, 61–66.
3. Yan, Y.; Ding, C.; Zhang, G.; Hu, J.; Zhu, L.; Zeng, D.; Qian, Q.; Ren, D. Genetic and environmental control of rice tillering. *Crop J.* **2023**, *11*, 1287–1302. [CrossRef]
4. Wei, G.; Xu, R.; Sun, H.; Wang, S. Quantitative relationship between tillering and leaf size of main stem and its response to nitrogen. *J. Jiangsu Agric.* **2014**, *30*, 950–958.
5. Riou-Khamlichi, C.; Huntley, R.; Jacqmard, A.; Murray, J.A. Cytokinin activation of Arabidopsis cell division through a D-type cyclin. *Science* **1999**, *283*, 1541–1544. [CrossRef]
6. Chen, T.; Xiao, W.; Huang, C.; Zhou, D.; Liu, Y.; Guo, T.; Chen, Z.; Wang, H. Fine Mapping of the Affecting Tillering and Plant Height Gene *CHA-1* in Rice. *Plants* **2023**, *12*, 1507. [CrossRef] [PubMed]
7. Hao, X.; Xiang, Y.; Zeng, M.; Yang, Y.; Xie, P.; Li, M.; Li, D.; Tian, L. Cytological characterization and gene mapping of rice dwarf rod mutants. *Life Sci Res.* **2021**, *25*, 39–47. Available online: http://smkx.hunnu.edu.cn/CN/abstract/abstract2341.shtml (accessed on 28 June 2024).
8. Huang, W.; Weng, F.; Cha, M.; Ding, Y.; Wang, S. Relationship between multiple tiller phenotype production and cytokinin in rice mutant D12W191. *J. Nanjing Agric. Univ.* **2016**, *39*, 711–721.
9. Sang, D.; Chen, D.; Liu, G.; Liang, Y.; Huang, L.; Meng, X.; Chu, J.; Sun, X.; Dong, G.; Yuan, Y. Strigolactones regulate rice tiller angle by attenuating shoot gravitropism through inhibiting auxin biosynthesis. *Proc. Natl. Acad. Sci. USA* **2014**, *111*, 11199–11204. [CrossRef]

10. Sasaki, A.; Ashikari, M.; Ueguchi-Tanaka, m.; Itoh, H.; Nishimura, A.; Swapan, D.; Ishiyama, K.; Saito, T.; Kobayashi, M.; Khush, G.S.; et al. A mutant gibberellin-synthesis gene in rice. *Nature* **2002**, *416*, 701–702. [CrossRef] [PubMed]
11. Umehara, M.; Hanada, A.; Yoshida, S.; Akiyama, K.; Arite, T.; Takeda-Kamiya, N.; Magome, H.; Kamiya, Y.; Shirasu, K.; Yoneyama, K. Inhibition of shoot branching by new terpenoid plant hormones. *Nature* **2008**, *455*, 195–200. [CrossRef] [PubMed]
12. Gomez-Roldan, V.; Fermas, S.; Brewer, P.B.; Puech-Pagès, V.; Dun, E.A.; Pillot, J.P.; Letisse, F.; Matusova, R.; Danoun, S.; Portais, J.C. Strigolactone inhibition of shoot branching. *Nature* **2008**, *455*, 189–194. [CrossRef] [PubMed]
13. Li, S.; Gao, J.; Li, J.; Wang, Y. Advances in regulating rice tillers by strigolactones. *Chin. Bull. Bot.* **2015**, *50*, 539.
14. Wang, Y.; Shang, L.; Yu, H.; Zeng, L.; Hu, J.; Ni, S.; Rao, Y.; Li, S.; Chu, J.; Meng, X. A strigolactone biosynthesis gene contributed to the green revolution in rice. *Mol. Plant* **2020**, *13*, 923–932. [CrossRef] [PubMed]
15. Zou, J.; Zhang, S.; Zhang, W.; Li, G.; Chen, Z.; Zhai, W.; Zhao, X.; Pan, X.; Xie, Q.; Zhu, L. The rice HIGH-TILLERING *DWARF1* encoding an ortholog of Arabidopsis *MAX3* is required for negative regulation of the outgrowth of axillary buds. *Plant J.* **2006**, *48*, 687–698. [CrossRef] [PubMed]
16. Arite, T.; Iwata, H.; Ohshima, K.; Maekawa, M.; Nakajima, M.; Kojima, M.; Sakakibara, H.; Kyozuka, J. *DWARF10*, an *RMS1/MAX4/DAD1* ortholog, controls lateral bud outgrowth in rice. *Plant J.* **2007**, *51*, 1019–1029. [CrossRef]
17. Ito, S.; Kitahata, N.; Umehara, M.; Hanada, A.; Kato, A.; Ueno, K.; Mashiguchi, K.; Kyozuka, J.; Yoneyama, K.; Yamaguchi, S.; et al. A new lead chemical for strigolactone biosynthesis inhibitors. *Plant Cell Physiol.* **2010**, *51*, 1143–1150. [CrossRef] [PubMed]
18. Zhang, S.; Li, G.; Fang, J.; Chen, W.; Jiang, H.; Zou, J.; Liu, X.; Zhao, X.; Li, X.; Chu, C. The interactions among *DWARF10*, auxin and cytokinin underlie lateral bud outgrowth in rice. *J. Integr. Plant Biol.* **2010**, *52*, 626–638. [CrossRef]
19. Lin, H.; Wang, R.; Qian, Q.; Yan, M.; Meng, X.; Fu, Z.; Yan, C.; Jiang, B.; Su, Z.; Li, J.; et al. *DWARF27*, an iron-containing protein required for the biosynthesis of strigolactones, regulates rice tiller bud outgrowth. *Plant Cell* **2009**, *21*, 1512–1525. [CrossRef]
20. Liu, W.; Kohlen, W.; Lillo, A.; Op den Camp, R.; Ivanov, S.; Hartog, M.; Limpens, E.; Jamil, M.; Smaczniak, C.; Kaufmann, K. Strigolactone biosynthesis in Medicago truncatula and rice requires the symbiotic GRAS-type transcription factors *NSP1* and *NSP2*. *Plant Cell* **2011**, *23*, 3853–3865. [CrossRef]
21. Tong, H.; Liu, L.; Jin, Y.; Du, L.; Yin, Y.; Qian, Q.; Zhu, L.; Chu, C. DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate brassinosteroid responses in rice. *Plant Cell* **2012**, *24*, 2562–2577. [CrossRef] [PubMed]
22. Xu, Y.; Huang, Z.; Wang, Z.; Liu, L. Progress of brassinosterols in growth, development and stress resistance in rice. *Chin. Agric. Bull.* **2012**, *28*, 1–5. Available online: https://www.casb.org.cn/CN/Y2012/V28/I9/1 (accessed on 28 June 2024).
23. Liu, Y.; Wang, H.; Jiang, Z.; Wang, W.; Xu, R.; Wang, Q.; Zhang, Z.; Li, A.; Liang, Y.; Ou, S. Genomic basis of geographical adaptation to soil nitrogen in rice. *Nature* **2021**, *590*, 600–605. [CrossRef] [PubMed]
24. Shou, P.; Le, L.; Xiaohong, L.; Hongxuan, W.; Guohua, X.U. Effects of nitrogen on rice tiller bud development and its mechanism. *J. Nanjing Agric. Univ.* **2016**, *39*, 973–978.
25. Forde, B.G. Local and long-range signaling pathways regulating plant responses to nitrate. *Annu. Rev. Plant Biol.* **2002**, *53*, 203–224. [CrossRef] [PubMed]
26. Liu, Y.; Gu, D.; Ding, Y.; Wang, Q.; Li, G.; Wang, S. The relationship between nitrogen, auxin and cytokinin in the growth regulation of rice (*Oryza sativa* L.) tiller buds. *Aust. J. Crop Sci.* **2011**, *5*, 1019–1026.
27. Pan, S. Regulation of the Growth and Development of Rice Tiller Buds by Nitrogen and Monogolactone. Ph.D. Dissertation, Nanjing Agricultural University, Nanjing, China, 1 June 2016.
28. Xu, J.; Zha, M.; Li, Y.; Ding, Y.; Chen, L.; Ding, C.; Wang, S. The interaction between nitrogen availability and auxin, cytokinin, and strigolactone in the control of shoot branching in rice *(Oryza sativa* L.). *Plant Cell Rep.* **2015**, *34*, 1647–1662. [CrossRef]
29. Hong, Z.; Ueguchi-Tanaka, M.; Shimizu-Sato, S.; Inukai, Y.; Fujioka, S.; Shimada, Y.; Takatsuto, S.; Agetsuma, M.; Yoshida, S.; Watanabe, Y. Loss-of-function of a rice brassinosteroid biosynthetic enzyme, C-6 oxidase, prevents the organized arrangement and polar elongation of cells in the leaves and stem. *Plant J.* **2002**, *32*, 495–508. [CrossRef]
30. Leyser, O. The fall and rise of apical dominance. *Curr. Opin. Genet. Dev.* **2005**, *15*, 468–471. [CrossRef]
31. Li, B. Analysis of the Physiological Characteristics of Different N and Expression Regulation of Related Genes in Rice. Ph.D. Dissertation, Nanjing Agricultural University, Nanjing, China, 1 June 2007.
32. Sun, L.; Zhu, L.; Zheng, G.; Zhu, F.; Guo, X.; Lugo, O.; Zhang, Z.; Kim, J.H. Study on the characteristics of nitrate and ammonium nitrogen accumulation and the regulation of nitrogen fertilizer in rice grain. *Rice China* **2016**, *22*, 25–29.
33. Liu, X.; Hu, Q.; Yan, J.; Sun, K.; Liang, Y.; Jia, M.; Meng, X.; Fang, S.; Wang, Y.; Jing, Y.; et al. ζ-Carotene isomerase suppresses tillering in rice through the coordinated biosynthesis of strigolactone and abscisic acid. *Mol. Plant* **2020**, *13*, 1784–1801. [CrossRef] [PubMed]
34. Dong, H.; Liu, Q. Analysis on the rationality of nitrogen and phosphorus chemical fertilizer in farmland of Anhui Province. *Soil Fertil. Sci. China* **2018**, 1673–6257.
35. Jahan, A.; Islam, A.; Sarkar, M.I.U.; Iqbal, M.; Ahmed, M.N.; Islam, M.R. Nitrogen response of two high yielding rice varieties as influenced by nitrogen levels and growing seasons. *Geol. Ecol. Landsc.* **2022**, *6*, 24–31. [CrossRef]

36. Li, W.; Zhou, Y.; Hu, J.; Wang, L. Determination of an unknown nitrogen compound by KN method. *China Health Stand. Manag.* **2016**, *7*, 110–112.
37. Sedri, M.H.; Niedbała, G.; Roohi, E.; Niazian, M.; Szulc, P.; Rahmani, H.A.; Feiziasl, V. Comparative analysis of plant growth-promoting rhizobacteria (PGPR) and chemical fertilizers on quantitative and qualitative characteristics of rainfed wheat. *Agronomy* **2022**, *12*, 1524. [CrossRef]

plants

MDPI

Article

OsIAA19, an *Aux/IAA* Family Gene, Involved in the Regulation of Seed-Specific Traits in Rice

Sha-Sha Jia [1,†], Xin-Yu Ren [1,†], Man-Ni Tong [1], Si-Yao Jiang [1], Chang-Quan Zhang [1,2], Qiao-Quan Liu [1,2] and Qian-Feng Li [1,2,*]

[1] Jiangsu Key Laboratory of Crop Genomics and Molecular Breeding/Zhongshan Biological Breeding Laboratory/Key Laboratory of Plant Functional Genomics of the Ministry of Education, Agricultural College of Yangzhou University, Yangzhou 225009, China; jiashasha0217@126.com (S.-S.J.); renxinyucherry@126.com (X.-Y.R.); tongmanni0814@163.com (M.-N.T.); jiangsiyao202411@163.com (S.-Y.J.); cqzhang@yzu.edu.cn (C.-Q.Z.); qqliu@yzu.edu.cn (Q.-Q.L.)

[2] Co-Innovation Center for Modern Production Technology of Grain Crops of Jiangsu Province/Jiangsu Key Laboratory of Crop Genetics and Physiology, Yangzhou University, Yangzhou 225009, China

* Correspondence: qfli@yzu.edu.cn

† These authors contributed equally to this work.

Abstract: The Aux/IAA family proteins, key components of the auxin signaling pathway, are plant-specific transcription factors with important roles in regulating a wide range of plant growth and developmental events. The *Aux/IAA* family genes have been extensively studied in Arabidopsis. However, most of the *Aux/IAA* family genes in rice have not been functionally studied. Only two *IAA* genes have been reported to be involved in the regulation of rice grain size. Grain size is a key factor affecting both rice yield and quality. Therefore, we selected an unreported *IAA* member, *OsIAA19*, based on bioinformatics analysis to investigate its potential role in grain size control. Our study showed that *OsIAA19* was constitutively expressed in all tissues tested and that the encoding protein was nuclear localized. The *osiaa19* mutants were then generated using CRISPR/Cas9 gene editing. Agronomic trait analyses showed that the *OsIAA19* mutation significantly increased rice grain length and weight, but had no significant effect on plant height, number of tillers, flag leaf length and width. In addition, the chalkiness of the *osiaa19* mutant seeds also increased, but their eating and cooking quality (ECQ) was not altered. Finally, seed germination analysis showed that knocking out *OsIAA19* slightly suppressed rice seed germination. These results suggest that *OsIAA19* may specifically regulate rice seed-related traits, such as grain shape, rice chalkiness and seed germination. This study not only enriched the functional study of the *Aux/IAA* genes and the auxin signaling pathway in rice, but also provided valuable genetic resources for breeding elite rice varieties.

Keywords: *Aux/IAA* gene; *OsIAA19*; auxin; grain shape; chalkiness; seed germination; rice

Citation: Jia, S.-S.; Ren, X.-Y.; Tong, M.-N.; Jiang, S.-Y.; Zhang, C.-Q.; Liu, Q.-Q.; Li, Q.-F. *OsIAA19*, an *Aux/IAA* Family Gene, Involved in the Regulation of Seed-Specific Traits in Rice. *Plants* **2024**, *13*, 3538. https://doi.org/10.3390/plants13243538

Academic Editor: Shigeto Morita

Received: 12 November 2024
Revised: 6 December 2024
Accepted: 16 December 2024
Published: 18 December 2024

1. Introduction

Phytohormones, trace amounts of organic substances produced by plants, play a key regulatory role in plant growth and development events as well as in adaptation to environmental changes [1]. The phytohormone family mainly includes auxin, gibberellin (GA), cytokinin, ethylene, abscisic acid (ABA), brassinosteroid (BR), jasmonic acid (JA), strigolactone and salicylic acid [2]. Auxin, as the first identified plant hormone, is required for normal plant growth and development processes, such as vascular tissue formation, apical dominance, root formation, tropism, cell division, differentiation and flower and fruit development [3–5]. Therefore, auxin is quite important and deserves a lot of attention from scientists.

The elucidation of the biosynthesis and signal transduction mechanism of auxin is of great importance for a better understanding of plant growth and development as well as their response to environmental changes. In general, the synthesis of auxin involves

two main pathways, namely the tryptophan-dependent and the tryptophan-independent pathways [6]. Compared to the complex auxin biosynthetic pathways, great progress has been made in elucidating the mechanisms of auxin signaling and transport [2,7]. *TIR1/AFB*-mediated auxin transcriptional regulation, also known as classical auxin signaling, is the best-studied and characterized auxin signaling pathway [8]. This pathway comprises three core transduction components: the auxin receptors, TIR1/AFBs; the transcriptional repressors, Aux/IAA; and the transcription factors, ARFs [9–11]. Briefly, in the presence of low levels of auxin, the repressor protein Aux/IAA binds to the active transcription factor ARF to form a heterodimer, which in turn inhibits the transcriptional activating activity of ARF [12]. In contrast, in a high auxin environment, auxin binds to the TIR1/AFB receptor and triggers the degradation of the Aux/IAA proteins via the ubiquitin–proteasome pathway, thereby releasing the ARF transcription factors, which can activate the transcription of auxin-responsive genes [13].

The Aux/IAA protein family is a group of plant-specific transcriptional repressors that are mainly involved in the regulation of root and leaf formation, phototropism, gravitropism, and apical dominance of buds [14]. There are a total of 29 members of the *Aux/IAA* gene family in Arabidopsis thaliana, distributed on five chromosomes [15]. Due to the similarity of IAA proteins, the knock-out of some genes in the *Aux/IAA* family has little effect on the plant phenotype, indicating the existence of functional redundancy among some *Aux/IAA* members [16]. Therefore, the functions of many *Aux/IAA* genes in Arabidopsis have been characterized using the gain-of-function mutation strategy [17–22].

Rice is not only a model monocotyledonous plant, but also one of the most important food crops in the world. More than half of the world's population feeds on rice. The production of elite rice varieties with superior quality and high yield is the main goal of rice breeding. Given the critical role of auxin in plants, extensive functional studies of rice *Aux/IAA* genes are essential. According to phylogenetic sequence analysis, there are 31 members of the *Aux/IAA* gene family in rice [23]. However, most of the *Aux/IAA* genes are uncharacterized and their biological functions are still unclear.

To date, the *Aux/IAA* genes reported in rice include *OsIAA1*, *OsIAA3*, *OsIAA4*, *OsIAA6*, *OsIAA10*, *OsIAA11*, *OsIAA12*, *OsIAA13*, *OsIAA18*, *OsIAA20*, *OsIAA23* and *OsIAA31.* Of these identified *Aux/IAA* genes, almost half are involved in the control of rice root formation. In particular, overexpression of *OsIAA1* resulted in reduced inhibition of root elongation in response to auxin treatment [24]. Similarly, overexpression of *OsIAA31* resulted in an auxin-insensitive phenotype, including a reduced number of coronal roots, abnormal gravity of roots and aboveground parts [25]. Regarding the regulation of lateral root development, different *Aux/IAA* genes have different or even opposite effects. For example, a gain-of-function mutation in *OsIAA11* inhibits lateral root initiation in rice [26]. On the other hand, the mutation of either *OsIAA13* or *OsIAA23* leads to a reduction in the number of lateral roots [27,28]. In addition, some *Aux/IAA* genes are also involved in the regulation of plant architecture [29]. For example, overexpression of *OsIAA1* reduced plant height and loosened plant architecture [24]. Similarly, overexpression of *OsIAA4* resulted in reduced plant height, increased tiller and leaf angles, and reduced gravity response, although expression level is low throughout the life cycle of rice [30]. *OsIAA6* controls the number of tillers by inhibiting the growth of tiller buds, while *OsIAA12* plays a positive role in the control of leaf angle in rice [31,32]. In addition, *Aux/IAA* genes can also mediate rice responses to abiotic stresses [33]. *OsIAA6* can induce an auxin-mediated drought response by regulating the expression of auxin synthesis genes, and its overexpression enhances drought resistance [31]. *OsIAA18* positively regulates drought and salt stress responses in plants by modulating stress-induced ABA signaling [34]. *OsIAA20* is also a positive regulator of drought tolerance and salt tolerance in rice [35]. Finally, only two *Aux/IAA* genes have been reported to function in the regulation of rice grain size. In more detail, the knock-down of *OsIAA3* resulted in longer rice grains [36]; conversely, overexpression of *OsIAA10* increased grain length and grain weight [37], suggesting that *OsIAA3* is a

negative regulator of rice grain length, whereas *OsIAA10* is a positive regulator of rice grain length.

Grain size, a key agronomic trait closely related to both rice yield and quality, has attracted considerable attention from rice geneticists and breeders. To date, more than 80 grain size-related genes have been cloned in rice, and a number of regulatory mechanisms have been identified, including the G-protein pathways, the ubiquitin–proteasome pathway, the mitogen-activated protein kinase (MAPK) signaling pathway, and plant hormones [38]. Within the plant hormone family, BR is known to be an important positive regulator of grain size in rice, and a number of BR-biosynthetic or signaling components are involved in controlling grain size [39]. With respect to auxin, the research progress on rice grain size regulation is still lagging behind. For example, although the number of functionally dissected *Aux/IAA* genes in rice is increasing, the functions of most members have not yet been identified so far. Furthermore, as mentioned above, only *OsIAA3* and *OsIAA10* are involved in grain size control [36,37].

To further identify novel grain size regulators in the *Aux/IAA* gene family, we first performed a phylogenetic analysis of the *Aux/IAA* genes. Then, using expression pattern assay, we successfully identified a potential target, *OsIAA19*, which is an unreported gene with relatively high expression in rice panicles and seeds. We speculated that *OsIAA19* may play an important role in controlling grain size and other seed-related traits. Therefore, a series of analyses were then conducted in this study to investigate the expression and biological functions of *OsIAA19*, including bioinformatics analysis, expression pattern assay and examination of rice yield or grain quality related key agronomic traits of *osiaa19* mutants. Our study not only revealed the biological function of *OsIAA19* and improved the auxin signaling pathway in rice, but also provided a valuable gene resource for rice grain size improvement.

2. Results

2.1. Bioinformatic Analysis of the Aux/IAA Gene Family in Rice

First, we constructed the phylogenetic tree of the *OsIAA* gene family in rice using MEGA7 software [40], which divided the whole family into four subfamilies according to the degree of evolution (Figure 1A). Subfamilies I to IV contain eight, seven, six and ten genes, respectively. The conserved motifs of the Aux/IAA family were analyzed using the MEME website [41]. The result showed that 25 out of 31 members shared four conserved motifs (Figure 1B). Subfamily IV had the largest number of members, but only three of the 10 members have been reported so far. Interestingly, the subfamily includes *OsIAA3*, the negative regulator of rice grain. In order to isolate target genes with a potential role in grain size regulation, the ePlants Rice website [42] was used to predict the expression patterns of the uncharacterized genes in this branch. The result showed that the expression level of *OsIAA19* in panicle and seed was significantly higher than that of the other genes (Supplementary Figure S1A). To further confirm their expression patterns, RT-qPCR analysis was performed. The data showed that *OsIAA17* had the highest expression in young rice panicles, while the expression of *OsIAA19* was close to that of *OsIAA17*. The expression of the other five genes was quite low (Supplementary Figure S1B). A similar result was obtained in the expression analysis of developing seeds. A major difference is that *OsIAA19* had the highest expression, which was about two times higher than that of *OsIAA17*, which had the second highest expression level (Supplementary Figure S1C), suggesting that *OsIAA19* may be involved not only in the regulation of rice grain size, but also in the control of grain filling or other seed-related traits. Therefore, *OsIAA19* was selected for the following expression and functional analysis.

The full-length cDNA of the *OsIAA19* gene is 1261 bp, contains 5 exons and encodes a 281-amino-acid Aux/IAA protein. We then analyzed the promoter of *OsIAA19* by using PlantCARE [43]. The result showed that the *OsIAA19* promoter contained a number of motifs related to plant hormones, including two P-box elements, two ABRE motifs, a TGACG motif and a TGA motif (Figure 1C). In particular, the P-box element

is involved in the GA response and the ABRE element is involved in the ABA response. The TGACG motif and the TGA element are involved in the ME-JA response and the auxin response, respectively. The information suggested that the *OsIAA19* gene may be involved in several plant hormone pathways. At present, studies have confirmed that genes related to GA and auxin pathways, such as *GW6* and *OsARF6*, are involved in the regulation of rice grain size [38]. However, no such genes have been reported in the ABA and MEJA pathways.

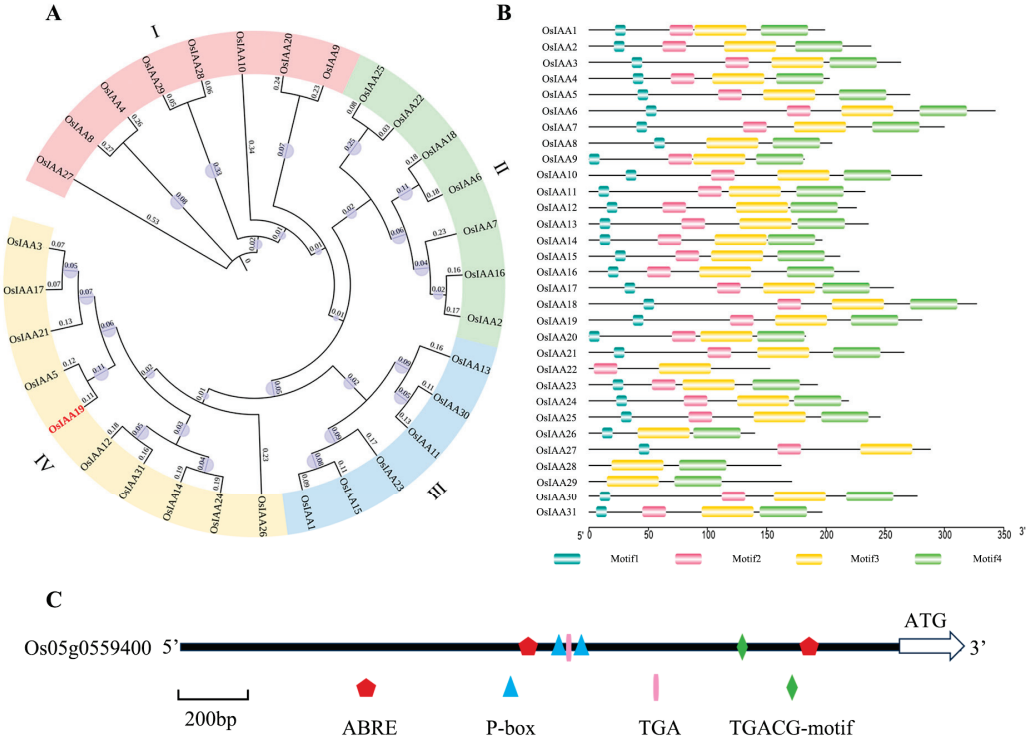

Figure 1. Bioinformatic analysis of the Aux/IAA protein family and the *OsIAA19* promoter in rice. (**A**) Phylogenetic tree of the Aux/IAA protein family in rice. The tree was constructed using the adjacency method in the MEGA7 software. The bootstrap is set to 1000. (**B**) Conserved motif analysis of the OsIAA protein in rice. (**C**) Analysis of hormone and stress response elements in the 2 kb upstream region of the *OsIAA19* transcription start site (ATG). P-box is a gibberellin response element. ABRE is an ABA response element. TGACG is related to MeJA and TGA is an auxin response element.

2.2. The OsIAA19 Gene Has a Constitutive Expression Pattern and Its Protein Is Localized in the Nucleus

To clarify the actual spatial and temporal expression pattern of the *OsIAA19* gene, total RNA was extracted from various tissues as well as developing panicles and seeds at different stages of wild-type rice ZH11. After the corresponding cDNA was generated by reverse transcription, the expression pattern of *OsIAA19* was analyzed by RT-qPCR. The results showed that *OsIAA19* was expressed in all rice samples tested, including roots, stems, leaves and leaf sheaths at the heading stage, as well as developing spikelets and seeds, indicating that the *OsIAA19* gene has a constitutive expression pattern (Figure 2A). More specifically, the expression abundance of *OsIAA19* was maintained throughout the developmental stages of spikelet tissues. However, the expression of the *OsIAA19* gene gradually increased with the development of rice seeds (Figure 2A), suggesting that OsIAA19 plays multiple roles in the development of rice spikelets and seeds.

To further investigate the subcellular localization of the OsIAA19 protein, the *OsIAA19* gene sequence was fused with the *eGFP* coding gene and transformed into the tobacco epidermal cells. The GFP fluorescence signal was then observed using laser confocal microscopy. The result showed that the OsIAA19 protein was only localized in the nucleus, whereas the eGFP protein alone was present in both the cytoplasm and the nucleus (Supplementary Figure S2). To further confirm the nuclear localization of OsIAA19, a nuclear-specific marker DLT protein was selected for co-localization assay [44]. DLT was fused to the RFP protein and co-expressed with the OsIAA19-eGFP protein. The results showed that OsIAA19-eGFP and DLT-RFP co-localized well in the nucleus (Figure 2B), demonstrating that OsIAA19 is a nuclear-localized protein.

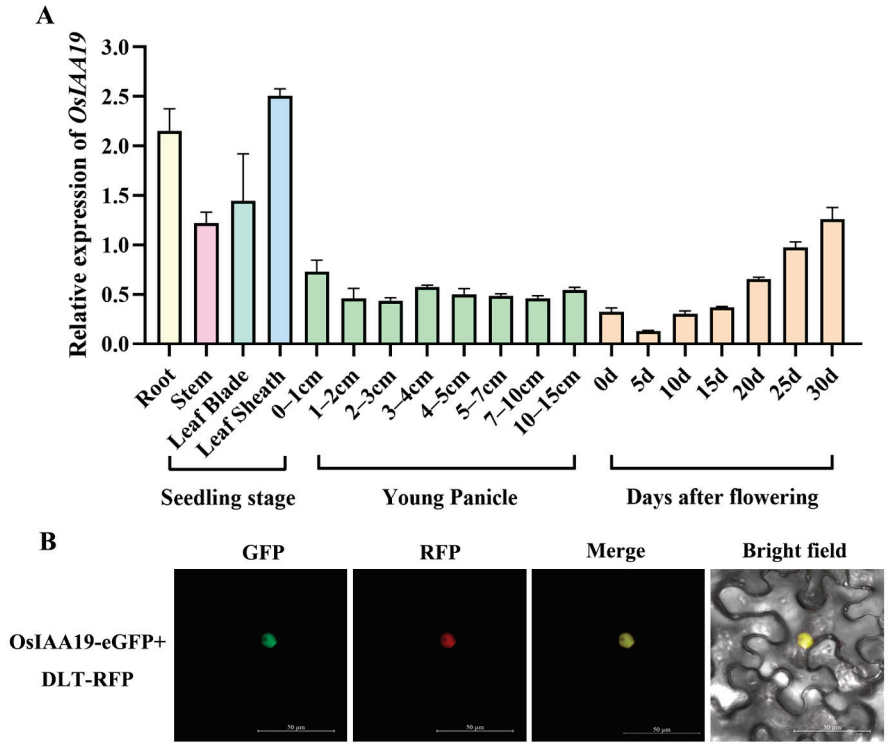

Figure 2. Expression analysis of *OsIAA19*. (**A**) The expression of *OsIAA19* was analyzed in different tissues, including developing spikelets and seeds. *Actin01* was used as an internal control for normalization. Data are means ± SD (n = 3 biological replicates). (**B**) Co-localization analysis of OsIAA19-eGFP protein and DLT-RFP protein in tobacco epidermal cells. The scale bar is 50 μm.

2.3. Knockout of OsIAA19 Had Little Effect on Rice Morphology

To further investigate the biological function of *OsIAA19*, CRISPR/Cas9 gene editing technology was used to generate *osiaa19* rice mutants. The editing target site of the *OsIAA19* gene in the generated transgenic rice was then confirmed by sequencing and finally, two homozygous *osiaa19* mutants were obtained, named *osiaa19-1* and *osiaa19-2*, respectively (Supplementary Figure S3). Specifically, two bases were deleted in the target region of *OsIAA19* in mutant *osiaa19-1*, and eight bases were deleted in mutant *osiaa19-2*, both resulting in frameshift mutations (Figure 3A). We then examined the main agronomic traits of the *osiaa19* mutants, including plant height, number of tillers, leaf length and width. The analysis showed that no significant difference was observed between the *osiaa19* mutants and the wild-type control ZH11 in these traits (Figure 3B–F), indicating that the *OsIAA19* mutation had no effect on normal plant growth.

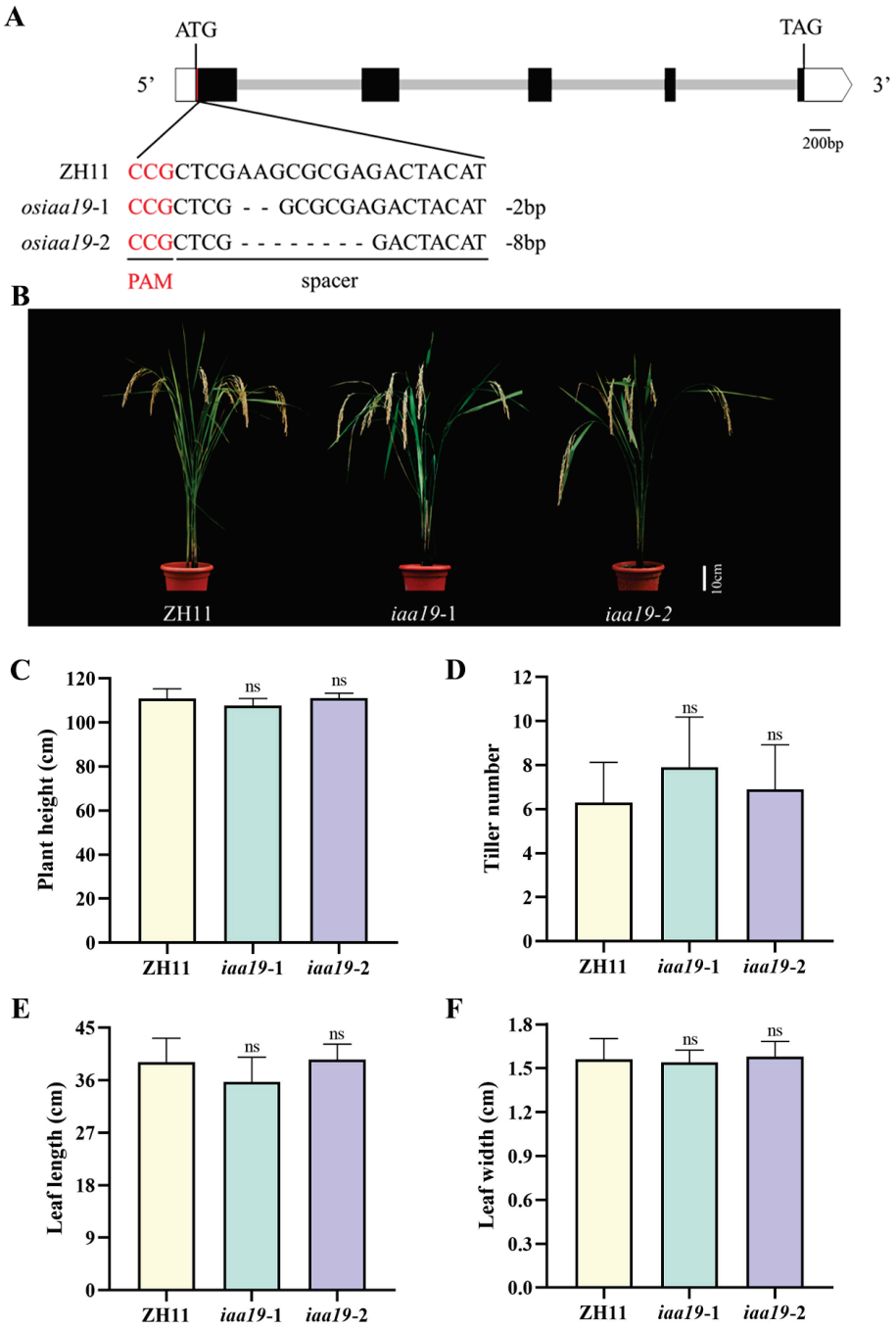

Figure 3. Analysis of key agronomic traits of *osiaa19* mutants and the wild-type control. (**A**) Schematic diagram of the vector used for CRISPR/Cas9-mediated gene editing of *OsIAA19*. The target site is underlined and the protospacer adjacent motif (PAM) is highlighted in red. Deletions are indicated by hyphens. (**B**) Morphology of *osiaa19* mutants and wild type ZH11. The scale bar is 10 cm. Quantitative data of plant height (**C**), number of tillers (**D**), leaf length (**E**) and leaf width (**F**) of the *osiaa19* mutants and wild-type control. Data are means ± SD (n = 5 rice plants). ns, not significant (Student's *t*-test).

2.4. OsIAA19 Mutation Increased Rice Grain Length and Weight

Since the regulation of grain size is important for both rice yield and grain quality, we then analyzed the grain shape of both *osiaa19* mutants and wild type (Figure 4A). The results showed that the grain length of the *osiaa19* mutants was significantly longer than that of the ZH11 control (Figure 4A,B). There was no difference in grain width between *osiaa19* mutants and wild type (Figure 4C). As a consequence, the 1000-grain weight of the *osiaa19* mutants was significantly increased (Figure 4D). Similarly, the glume length

of the *osiaa19* mutants was also increased at the heading stage, while the width remained unchanged (Supplementary Figure S4).

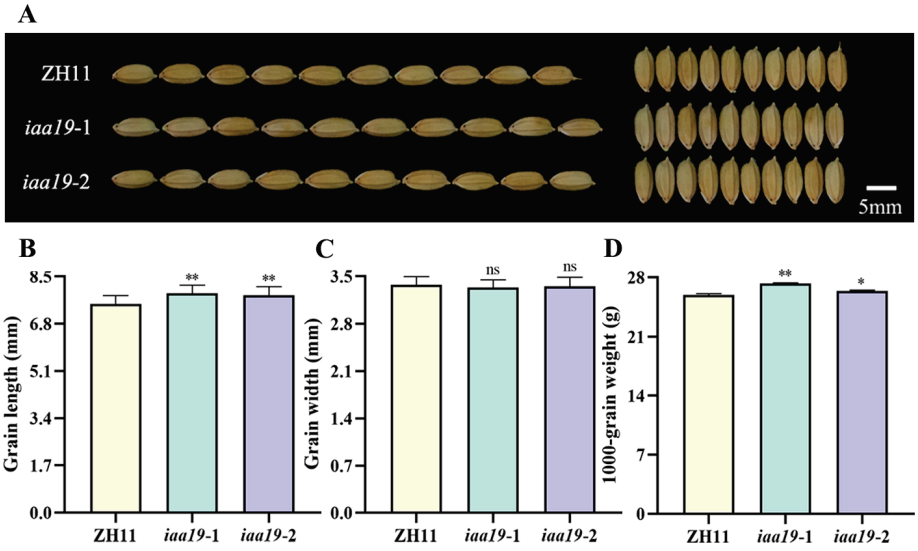

Figure 4. Grain size analysis of *osiaa19* mutants and ZH11 control. Grain morphology (**A**), grain length (**B**), grain width (**C**) and 1000-grain weight (**D**) of *osiaa19* mutants and ZH11 control were analyzed. Data are means ± SD (n = 100 seeds). * $p < 0.05$; ** $p < 0.01$; ns, not significant (Student's *t*-test).

To preliminarily explore the potential mechanism of *OsIAA19* in regulating rice grain length, we examined the expression of two representative grain size genes, *GS9* [45] and *GW7* [46]. The results showed that the expression of *GS9*, a negative regulator of grain length, was significantly reduced in *iaa19* mutants (Figure 5A). In contrast, the expression of *GW7*, a positive regulator of grain length, was significantly increased in *iaa19* mutants (Figure 5B). These data suggest that *OsIAA19* negatively regulates rice grain length, at least in part, by modulating the expression of *GS9* and *GW7*.

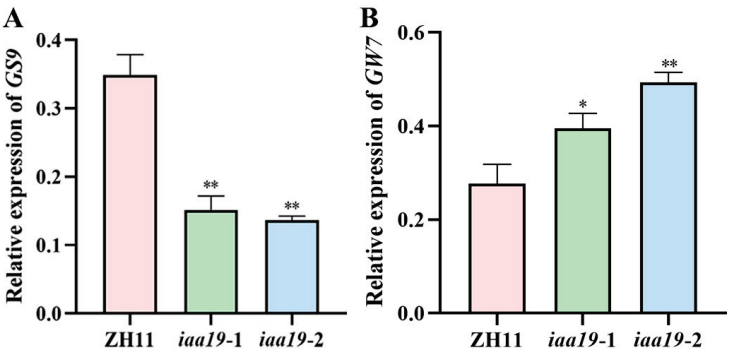

Figure 5. Expression analysis of the grain size genes *GS9* (**A**) and *GW7* (**B**) in rice spikelets of *osiaa19* mutants and ZH11 control. *Actin01* was used as an internal control for gene expression. Data are means ± SD (n = 3 biological replicates). * $p < 0.05$; ** $p < 0.01$ (Student's *t*-test).

2.5. OsIAA19 Mutation Increased Rice Chalkiness but Without Significant Effect on Eating and Cooking Quality (ECQ)

Chalkiness refers to opaque areas in the endosperm of rice that reduce the processing and appearance quality of rice [47]. The degree of chalkiness can be measured by two indices: chalky grain rate and chalkiness degree [48]. The result showed that both the chalky grain rate and chalkiness degree of the *osiaa19* mutants were significantly increased (Figure 6A–C). In addition to rice appearance, two important physicochemical properties of rice, amylose content (AC) and gel consistency (GC), were also investigated. AC is the most

important determinant of rice ECQ, and GC is another important index for evaluating rice ECQ. The data showed that both the AC and GC of *osiaa19* mutants were not significantly changed compared to the control ZH11 (Figure 6D,E).

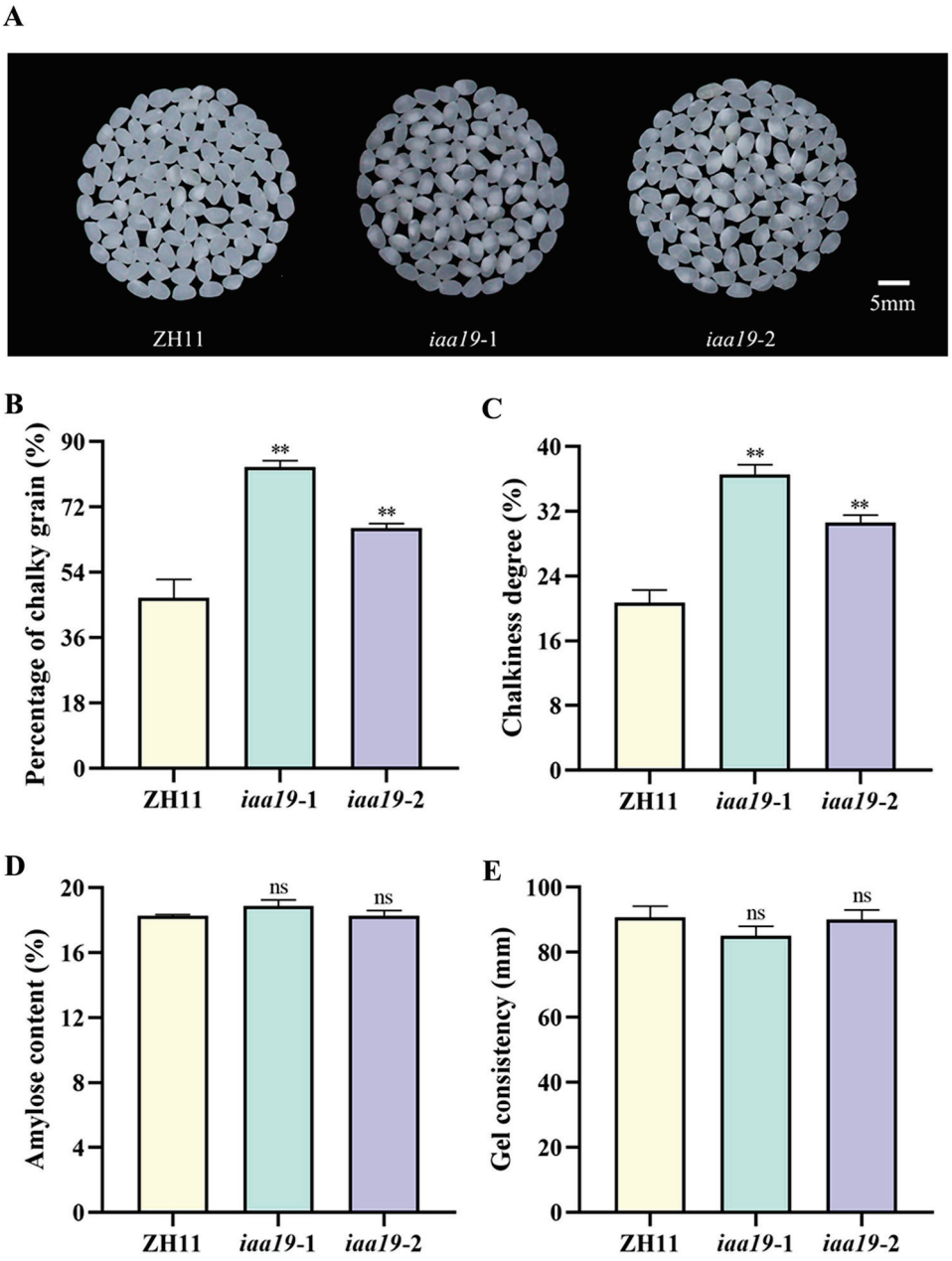

Figure 6. Analysis of the appearance quality and physicochemical properties of the *osiaa19* mutants and the wild type ZH11. (**A**) The appearance of the milled rice from *osiaa19* mutants and ZH11. Chalky grain rate (**B**), chalkiness degree (**C**), AC (**D**), and GC (**E**) of the *osiaa19* mutants and ZH11 control. Data are means ± SD (n = 3 biological replicates). In each biological replicate, 150–200 grains were used for seed chalkiness analysis. ** $p < 0.01$; ns, not significant (Student's *t*-test).

2.6. OsIAA19 Mutation Slightly Suppresses Seed Germination

The ability of rice seeds to germinate is critical for direct seeding and ensuring high rice yields. Since OsIAA19 is a component of the auxin signaling pathway and a number of other hormone-related cis-elements exist in its promoter, we further investigated whether the *OsIAA19* mutation would affect the germination property of rice seeds. The results showed that the germination rate of the *osiaa19* mutants was slower than ZH11 at the early

stage of germination, but the final germination rate was not affected (Figure 7). In addition, the bud length of the *osiaa19* mutants was shorter (Supplementary Figure S5).

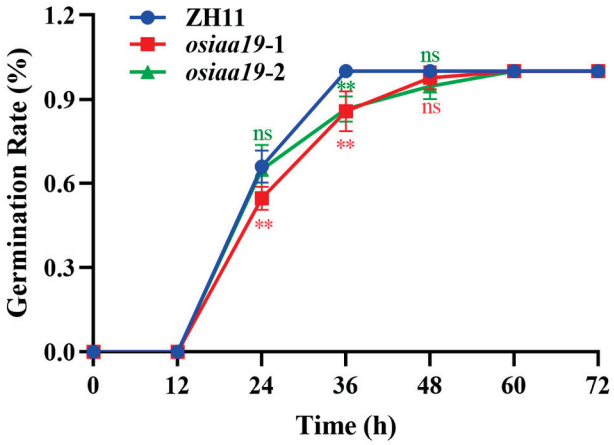

Figure 7. Seed germination analysis of the *osiaa19* mutants and the wild-type control ZH11. Data are means ± SD (n = 3 biological replicates). In each biological replicate, 30 seeds were used for analysis. ** $p < 0.01$; ns, not significant (Student's *t*-test).

3. Discussion

The study of rice grain size is of great importance for the cultivation of rice varieties with high yield and superior quality. Grain size includes the length, width and thickness of grains, as well as the ratio of length to width. Grain size not only affects rice yield but also the appearance quality and processing quality of rice [49–51]. More than 80 rice grain shape genes have been cloned, and a number of different regulatory mechanisms have been successfully revealed [52,53]. Among them, plant hormones are important regulators of rice grain size. Recently, a number of publications have reported that the plant hormones BR, auxin, GA and cytokinin play important roles in regulating grain size [38].

Auxin plays an important role in a wide range of growth and developmental events in rice. Several auxin-related genes have been reported to regulate grain size in rice. *BG1*, a primary response gene of auxin, functions as a positive regulator of auxin response and transport and consequently regulates grain size by modulating cell division and elongation [54]. *TGW6* encodes indole-3-acetic acid (IAA)-glucose hydrolase for the production of free IAA. Knocking out *TGW6* increases grain length and weight [55]. Further studies showed that *TGW6* is expressed for only a short time during early inflorescence development, suggesting that *TGW6* may also play an important role in regulating pollen development [56]. *qTGW3* encodes a GSK3/SHAGGY-like kinase OsSK41/OsGSK5 that interacts with and phosphorylates the transcriptional repressor OsARF4 in the auxin pathway. The OsSK41–OsARF4 interaction module may repress rice grain size and weight by negatively regulating auxin signaling [57]. In addition, a recent study revealed a novel mechanism of grain size regulation centered on the OsTIR1–OsIAA10–OsARF4 module [37]. Finally, as a transcription factor, OsARF6 binds directly to the promoter of the auxin influx transporter gene *OsAUX3* to negatively regulate rice grain length and weight. Mechanistically, the OsARF6–OsAUX3 module affects the longitudinal elongation of glume cells by altering both the level and distribution of auxin in glume cells. In addition, miR167a can directly direct OsARF6 mRNA silencing to increase rice grain length and weight [58].

There are a total of 31 *Aux/IAA* gene members in rice. However, only two of them, *OsIAA3* and *OsIAA10*, have been reported to be involved in grain size regulation. Interestingly, the two genes play opposite roles in controlling grain size. Specifically, *OsIAA3* is a negative regulator, whereas *OsIAA10* is a positive regulator of grain size [36,37]. In this study, we demonstrated that, like *OsIAA3*, *OsIAA19* also functions as a negative regulator of rice grain size because the knock-out of *OsIAA19* promoted rice grain length and weight (Figure 4). In addition, our data showed that *OsIAA19* mutation had no significant effect on

rice plant height, tiller, leaf length and width (Figure 3). Therefore, *OsIAA19* is considered to be a novel gene with a specific regulatory role in seed-related traits, including grain size. However, the underlying molecular mechanism still requires further investigation.

In addition to grain size, some other seed-related traits of the *osiaa19* mutants were also investigated, including rice quality and germination characteristics. In general, rice quality includes processing quality, appearance quality, eating and cooking quality, and nutritional quality. Appearance quality is mainly influenced by grain shape, chalkiness and transparency [38]. In addition, grain shape is closely related to chalkiness. In general, the slender grain, i.e., with longer grain length and narrower grain width, usually has a better appearance quality with less chalkiness. For example, *GW2* encodes an E3 ubiquitin ligase that negatively affects grain width and grain weight in rice [59]. NIL-*gw2.1*, the near-isogenic line of *GW2*, not only increased grain length and width, but also reduced chalkiness [60]. OsSPL16/GW8, a transcription factor containing the SBP domain, can bind directly to the promoter of *GW7* and inhibit its expression, thereby regulating rice grain width [61,62]. In addition, phytohormones are also involved in the regulation of chalkiness. *DG1* is mainly responsible for the long-distance transport of ABA from rice leaves to grains, and its mutation resulted in abnormal grain filling and subsequent silty endosperm [63]. OsNF-YB1, an endosperm-specific transcription factor, binds to the promoter of *OsYUC11* and positively regulates auxin synthesis. Therefore, IAA biosynthesis was reduced in *osyuc11* and *osnf-yb1* mutants, corresponding to smaller seeds and increased chalkiness [64]. Surprisingly, although grain length increased in *osiaa19*, grain chalkiness also increased (Figure 6). However, it is still unclear whether the increased chalkiness in the *osiaa19* mutants is caused by a change in grain size or a grain filling problem, b66ecause *OsIAA19* was also highly expressed in rice source tissues such as rice leaf. In fact, GWD1, an enzyme involved in controlling transient starch degradation in source tissues, also plays an important regulatory role in seed traits such as grain size, rice chalkiness and seed germination [65], suggesting the importance of source–sink interactions in determining important seed traits.

4. Materials and Methods

4.1. Bioinformatic Analysis

According to the gene IDs of 31 members of the *Aux/IAA* family in rice (Supplementary Table S1), their corresponding protein sequences were used to construct the phylogenetic tree by using the MEGA7 software with the neighbor-joining method [66], and the bootstrap was set to 1000. The constructed phylogenetic tree was then further refined by ITOL (https://itol.embl.de/, accessed on 30 October 2024). The conserved motifs of the Aux/IAA family were analyzed using the MEME websites (https://meme-suite.org/meme/, accessed on 30 October 2024).

4.2. Plant Materials and Growth Conditions

The rice materials used in this study were *osiaa19* mutants and their corresponding wild-type control japonica rice variety Zhonghua 11 (ZH11). All rice materials were grown in the field of the Agricultural College of Yangzhou University. During the planting process, each experimental material was planted in two rows according to the standard wide and narrow rows, with ten seedlings in each row, and the conventional crop management methods were followed for uniform management.

4.3. Vector Construction

We designed specific targets for gene editing in the first exon of the *OsIAA19* gene. The target DNA sequence was then cloned into the sgRNA-Cas9 expression vector. After confirmation by sequencing, the correct target plasmid was transformed into rice callus by the Agrobacterium-mediated transformation method. The CRISPR/Cas9-related vector system and the detailed procedure can be found in a previous publication [67]. The primers used for sequencing are shown in Supplementary Table S2.

4.4. Agronomic Trait Analysis

The agronomic traits of the *osiaa19* mutants and the wild-type control ZH11 were investigated in the experimental field at Yangzhou University. The agronomic traits investigated included plant height, number of tillers, flag leaf length and width. As to the seed-related traits, grain length, grain width and 1000-grain weight were measured using the rice appearance quality detector (SC-E, Wanshen, Hangzhou, China).

4.5. RT-qPCR Assay

Samples of different rice tissues and developing spikelets or seeds were collected for *OsIAA19* gene expression analysis. Total RNA was carefully extracted from the above samples using an RNA extraction kit (RC401, Vazyme, Nanjing, China). Then, 1 μg of total RNA was used for reverse transcription. Finally, RT-qPCR analysis was performed using the 2 × ChamQ SYBR qPCR Master Mix Kit (Q711, Vazyme, Nanjing, China), and the *Actin01* gene was used as a reference gene for normalization. Three biological replicates were used in each experiment. All primer sequences used are listed in Supplementary Table S2.

4.6. Subcellular Localization Analysis

To study the subcellular localization of the OsIAA19 protein, the full-length OsIAA19 CDS sequence was ligated into the p2300-35S-eGFP vector to construct the 35S::OsIAA19-GFP plasmid, then transfected into the Agrobacterium strain GV3101 for transient expression in tobacco leaf epidermal cells. DLT-RFP, as a nuclear localization marker, was co-expressed with OsIAA19-eGFP in tobacco leaf epidermal cells. Finally, the GFP and RFP signals in the transferred tobacco leaves were observed by using the laser confocal microscope (LSM 710, Carl Zeiss AG, Jena, Germany).

4.7. Rice Flour Preparation and Rice Appearance Analysis

First, a rice hulling machine (SY88-TH, SsangYong, Hirazawa, Republic of Korea) was used to dehull the ripe seeds to obtain brown rice. Then, the brown seeds were manually selected to remove the moldy and incompletely matured seeds. Next, the brown rice was polished using a grain polisher (Pearlest, KETT, Tokyo, Japan) to obtain polished rice [68]. After removing the broken rice, the whole milled rice was scanned by the rice appearance quality detector (SC-E, Wanshen, Hangzhou, China) to evaluate the chalky rice.

4.8. Rice Quality Analysis

The rice flour prepared above was used to analyze the amylose content (AC) and gel consistency (GC) of rice. The detailed procedure could follow the previous publication [69].

4.9. Seed Germination Test

The mature rice seeds were manually hulled, sterilized and incubated in the dark in an artificial climate incubator at a temperature of 26 °C for seed germination analysis. A detailed method of seed germination testing can be found in previously published literature [70].

4.10. Statistical Analysis

The data were analyzed using GraphPad Prism 8.0.2. Data in this study were presented as mean ± SD. The level of significance was assessed using Student's *t*-test (* $p < 0.05$, ** $p < 0.01$).

5. Conclusions

In this study, we performed a bioinformatic analysis of the Aux/IAA family proteins. Based on the phylogenetic tree data and the expression pattern of *OsIAA* genes, *OsIAA19* was selected as a candidate gene for further analysis of its role in controlling rice grain size. First, expression analysis showed that the *OsIAA19* gene was constitutively expressed in all tissues tested, including the developing spikelets. Next, the subcellular localization

assay showed that the OsIAA19 protein was specifically localized in the nucleus. *osiaa19* mutants were then generated using CRISPR/Cas9 gene editing. Agronomic trait analysis showed that, except for increased grain length, the other rice traits tested were unchanged. In addition, grain quality analysis showed that the *OsIAA19* mutation had no effect on rice ECQ, but increased grain chalkiness. Finally, the germination of *osiaa19* mutants was slightly inhibited. In conclusion, this research not only dissected the expression pattern and biological function of a new *Aux/IAA* gene *OsIAA19* in rice, but also provides useful information on the regulatory network of the auxin pathway and is an important gene resource for future rice quality improvement.

Supplementary Materials: The following supporting information can be downloaded at: https://www.mdpi.com/article/10.3390/plants13243538/s1, Supplementary Figure S1: Prediction of the expression pattern of *OsIAA* genes; Supplementary Figure S2: Subcellular localization of OsIAA19 protein in tobacco epidermal cells; Supplementary Figure S3: Sequencing data of *osiaa19* mutants; Supplementary Figure S4: Morphology of spikelet hulls of *osiaa19* mutants and ZH11 control at heading stage; Supplementary Figure S5: Photographs of germinating seeds 96 h after imbibition (HAI); Supplementary Table S1: Gene IDs of *OsIAA* family members in rice; Supplementary Table S2: Sequence of the primers used in the study.

Author Contributions: Conceptualization, Q.-F.L.; Methodology, S.-S.J., X.-Y.R. and C.-Q.Z.; Investigation, S.-S.J., X.-Y.R., M.-N.T. and S.-Y.J.; Writing—original draft, S.-S.J.; Writing—review and editing, Q.-F.L.; Supervision, Q.-F.L. and Q.-Q.L.; Project administration, Q.-F.L.; Funding acquisition, Q.-F.L. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the National Natural Science Foundation of China (32270575, 32071984), the Project of Zhongshan Biological Breeding Laboratory (BM2022008-02), the Programs from Government of Jiangsu Province (JBGS [2021]035-1, BK20200045, PAPD) and the Post-graduate Research & Practice Innovation Program of Jiangsu Province (KYCX21_3238).

Data Availability Statement: All data supporting the findings of this study are available within the paper and Supplementary Materials.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wawrzyńska, A.; Sirko, A. Sulfate Availability and Hormonal Signaling in the Coordination of Plant Growth and Development. *Int. J. Mol. Sci.* **2024**, *25*, 3978. [CrossRef] [PubMed]
2. Anfang, M.; Shani, E. Transport mechanisms of plant hormones. *Curr. Opin. Plant Biol.* **2021**, *63*, 102055. [CrossRef] [PubMed]
3. Zhao, Y. Auxin biosynthesis and its role in plant development. *Annu. Rev. Plant Biol.* **2010**, *61*, 49–64. [CrossRef] [PubMed]
4. Kawai, T.; Akahoshi, R.; Shelley, I.J.; Kojima, T.; Sato, M.; Tsuji, H.; Inukai, Y. Auxin Distribution in Lateral Root Primordium Development Affects the Size and Lateral Root Diameter of Ric\e. *Front. Plant Sci.* **2022**, *13*, 834378. [CrossRef] [PubMed]
5. Rutten, J.; van den Berg, T.; Tusscher, K.T. Modeling Auxin Signaling in Roots: Auxin Computations. *Cold Spring Harb. Perspect. Biol.* **2022**, *14*, a040089. [CrossRef]
6. Gomes, G.L.B.; Scortecci, K.C. Auxin and its role in plant development: Structure, signalling, regulation and response mechanisms. *Plant Biol.* **2021**, *23*, 894–904. [CrossRef] [PubMed]
7. Mockaitis, K.; Estelle, M. Auxin receptors and plant development: A new signaling paradigm. *Annu. Rev. Cell Dev. Biol.* **2008**, *24*, 55–80. [CrossRef]
8. Yu, Z.; Zhang, F.; Friml, J.; Ding, Z. Auxin signaling: Research advances over the past 30 years. *J. Integr. Plant Biol.* **2022**, *64*, 371–392. [CrossRef] [PubMed]
9. Leyser, O. Auxin Signaling. *Plant Physiol.* **2018**, *176*, 465–479. [CrossRef]
10. Matthes, M.S.; Best, N.B.; Robil, J.M.; Malcomber, S.; Gallavotti, A.; McSteen, P. Auxin EvoDevo: Conservation and Diversification of Genes Regulating Auxin Biosynthesis, Transport, and Signaling. *Mol. Plant* **2019**, *12*, 298–320. [CrossRef] [PubMed]
11. Niemeyer, M.; Moreno Castillo, E.; Ihling, C.H.; Iacobucci, C.; Wilde, V.; Hellmuth, A.; Hoehenwarter, W.; Samodelov, S.L.; Zurbriggen, M.D.; Kastritis, P.L.; et al. Flexibility of intrinsically disordered degrons in AUX/IAA proteins reinforces auxin co-receptor assemblies. *Nat. Commun.* **2020**, *11*, 2277. [CrossRef] [PubMed]
12. Szemenyei, H.; Hannon, M.; Long, J.A. TOPLESS mediates auxin-dependent transcriptional repression during Arabidopsis embryogenesis. *Science* **2008**, *319*, 1384–1386. [CrossRef] [PubMed]
13. Chapman, E.J.; Estelle, M. Mechanism of auxin-regulated gene expression in plants. *Annu. Rev. Genet.* **2009**, *43*, 265–285. [CrossRef] [PubMed]

14. Luo, J.; Zhou, J.J.; Zhang, J.Z. Aux/IAA Gene Family in Plants: Molecular Structure, Regulation, and Function. *Int. J. Mol. Sci.* **2018**, *19*, 259. [CrossRef] [PubMed]
15. Shani, E.; Salehin, M.; Zhang, Y.; Sanchez, S.E.; Doherty, C.; Wang, R.; Mangado, C.C.; Song, L.; Tal, I.; Pisanty, O.; et al. Plant Stress Tolerance Requires Auxin-Sensitive Aux/IAA Transcriptional Repressors. *Curr. Biol.* **2017**, *27*, 437–444. [CrossRef] [PubMed]
16. Overvoorde, P.J.; Okushima, Y.; Alonso, J.M.; Chan, A.; Chang, C.; Ecker, J.R.; Hughes, B.; Liu, A.; Onodera, C.; Quach, H.; et al. Functional genomic analysis of the AUXIN/INDOLE-3-ACETIC ACID gene family members in *Arabidopsis thaliana*. *Plant Cell* **2005**, *17*, 3282–3300. [CrossRef] [PubMed]
17. Tian, Q.; Reed, J.W. Control of auxin-regulated root development by the *Arabidopsis thaliana* SHY2/IAA3 gene. *Development* **1999**, *126*, 711–721. [CrossRef] [PubMed]
18. Rouse, D.; Mackay, P.; Stirnberg, P.; Estelle, M.; Leyser, O. Changes in auxin response from mutations in an AUX/IAA gene. *Science* **1998**, *279*, 1371–1373. [CrossRef]
19. Nagpal, P.; Walker, L.M.; Young, J.C.; Sonawala, A.; Timpte, C.; Estelle, M.; Reed, J.W. AXR2 encodes a member of the Aux/IAA protein family. *Plant Physiol.* **2000**, *123*, 563–574. [CrossRef] [PubMed]
20. Rogg, L.E.; Lasswell, J.; Bartel, B. A gain-of-function mutation in IAA28 suppresses lateral root development. *Plant Cell* **2001**, *13*, 465–480. [CrossRef]
21. López-Bucio, J.; Ortiz-Castro, R.; Ruíz-Herrera, L.F.; Juárez, C.V.; Hernández-Madrigal, F.; Carreón-Abud, Y.; Martínez-Trujillo, M. Chromate induces adventitious root formation via auxin signalling and SOLITARY-ROOT/IAA14 gene function in *Arabidopsis thaliana*. *BioMetals* **2015**, *28*, 353–365. [CrossRef] [PubMed]
22. Uehara, T.; Okushima, Y.; Mimura, T.; Tasaka, M.; Fukaki, H. Domain II mutations in CRANE/IAA18 suppress lateral root formation and affect shoot development in *Arabidopsis thaliana*. *Plant Cell Physiol.* **2008**, *49*, 1025–1038. [CrossRef] [PubMed]
23. Jain, M.; Kaur, N.; Garg, R.; Thakur, J.K.; Tyagi, A.K.; Khurana, J.P. Structure and expression analysis of early auxin-responsive Aux/IAA gene family in rice (Oryza sativa). *Funct. Integr. Genom.* **2006**, *6*, 47–59. [CrossRef] [PubMed]
24. Song, Y.; You, J.; Xiong, L. Characterization of OsIAA1 gene, a member of rice Aux/IAA family involved in auxin and brassinosteroid hormone responses and plant morphogenesis. *Plant Mol. Biol.* **2009**, *70*, 297–309. [CrossRef]
25. Nakamura, A.; Umemura, I.; Gomi, K.; Hasegawa, Y.; Kitano, H.; Sazuka, T.; Matsuoka, M. Production and characterization of auxin-insensitive rice by overexpression of a mutagenized rice IAA protein. *Plant J.* **2006**, *46*, 297–306. [CrossRef] [PubMed]
26. Zhu, Z.X.; Liu, Y.; Liu, S.J.; Mao, C.Z.; Wu, Y.R.; Wu, P. A gain-of-function mutation in OsIAA11 affects lateral root development in rice. *Mol. Plant* **2012**, *5*, 154–161. [CrossRef] [PubMed]
27. Kitomi, Y.; Inahashi, H.; Takehisa, H.; Sato, Y.; Inukai, Y. OsIAA13-mediated auxin signaling is involved in lateral root initiation in rice. *Plant Sci.* **2012**, *190*, 116–122. [CrossRef]
28. Jun, N.; Gaohang, W.; Zhenxing, Z.; Huanhuan, Z.; Yunrong, W.; Ping, W. OsIAA23-mediated auxin signaling defines postembryonic maintenance of QC in rice. *Plant J.* **2011**, *68*, 433–442. [CrossRef]
29. Chen, R.; Deng, Y.; Ding, Y.; Guo, J.; Qiu, J.; Wang, B.; Wang, C.; Xie, Y.; Zhang, Z.; Chen, J.; et al. Rice functional genomics: Decades' efforts and roads ahead. *Sci. China Life Sci.* **2022**, *65*, 33–92.
30. Song, Y.; Xu, Z.F. Ectopic overexpression of an AUXIN/INDOLE-3-ACETIC ACID (Aux/IAA) gene OsIAA4 in rice induces morphological changes and reduces responsiveness to Auxin. *Int. J. Mol. Sci.* **2013**, *14*, 13645–13656. [CrossRef]
31. Jung, H.; Lee, D.K.; Choi, Y.D.; Kim, J.K. OsIAA6, a member of the rice Aux/IAA gene family, is involved in drought tolerance and tiller outgrowth. *Plant Sci.* **2015**, *236*, 304–312. [CrossRef] [PubMed]
32. Chen, S.H.; Zhou, L.J.; Xu, P.; Xue, H.W. SPOC domain-containing protein Leaf inclination3 interacts with LIP1 to regulate rice leaf inclination through auxin signaling. *PLoS Genet.* **2018**, *14*, e1007829. [CrossRef] [PubMed]
33. Hannah, M.A.; Heyer, A.G.; Hincha, D.K. A global survey of gene regulation during cold acclimation in *Arabidopsis thaliana*. *PLoS Genet.* **2005**, *1*, e26. [CrossRef]
34. Wang, F.; Niu, H.; Xin, D.; Long, Y.; Wang, G.; Liu, Z.; Li, G.; Zhang, F.; Qi, M.; Ye, Y.; et al. OsIAA18, an Aux/IAA Transcription Factor Gene, Is Involved in Salt and Drought Tolerance in Rice. *Front. Plant Sci.* **2021**, *12*, 738660. [CrossRef]
35. Zhang, A.; Yang, X.; Lu, J.; Song, F.; Sun, J.; Wang, C.; Lian, J.; Zhao, L.; Zhao, B. OsIAA20, an Aux/IAA protein, mediates abiotic stress tolerance in rice through an ABA pathway. *Plant Sci.* **2021**, *308*, 110903. [CrossRef] [PubMed]
36. Zhang, Z.; Li, J.; Tang, Z.; Sun, X.; Zhang, H.; Yu, J.; Yao, G.; Li, G.; Guo, H.; Li, J.; et al. Gnp4/LAX2, a RAWUL protein, interferes with the OsIAA3-OsARF25 interaction to regulate grain length via the auxin signaling pathway in rice. *J. Exp. Bot.* **2018**, *69*, 4723–4737. [CrossRef]
37. Ma, M.; Shen, S.Y.; Bai, C.; Wang, W.Q.; Feng, X.H.; Ying, J.Z.; Song, X.J. Control of grain size in rice by TGW3 phosphorylation of OsIAA10 through potentiation of OsIAA10-OsARF4-mediated auxin signaling. *Cell Rep.* **2023**, *42*, 112187. [CrossRef]
38. Li, P.; Chen, Y.H.; Lu, J.; Zhang, C.Q.; Liu, Q.Q.; Li, Q.F. Genes and Their Molecular Functions Determining Seed Structure, Components, and Quality of Rice. *Rice* **2022**, *15*, 18. [CrossRef]
39. Zhang, X.; Meng, W.; Liu, D.; Pan, D.; Yang, Y.; Chen, Z.; Ma, X.; Yin, W.; Niu, M.; Dong, N.; et al. Brassinosteroid regulation in rice seed biology. *Seed Biol.* **2022**, *1*, 2.
40. Kumar, S.; Stecher, G.; Tamura, K. MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Mol. Biol. Evol.* **2016**, *33*, 1870–1874. [CrossRef]
41. Bailey, T.L.; Johnson, J.; Grant, C.E.; Noble, W.S. The MEME Suite. *Nucleic Acids Res.* **2015**, *43*, W39–W49. [CrossRef]

42. Waese, J.; Fan, J.; Pasha, A.; Yu, H.; Fucile, G.; Shi, R.; Cumming, M.; Kelley, L.A.; Sternberg, M.J.; Krishnakumar, V.; et al. ePlant: Visualizing and Exploring Multiple Levels of Data for Hypothesis Generation in Plant Biology. *Plant Cell* **2017**, *29*, 1806–1821. [CrossRef] [PubMed]
43. Rombauts, S.; Déhais, P.; Van Montagu, M.; Rouzé, P. PlantCARE, a plant cis-acting regulatory element database. *Nucleic Acids Res.* **1999**, *27*, 295–296. [CrossRef] [PubMed]
44. Tong, H.; Liu, L.; Jin, Y.; Du, L.; Yin, Y.; Qian, Q.; Zhu, L.; Chu, C. DWARF AND LOW-TILLERING acts as a direct downstream target of a GSK3/SHAGGY-like kinase to mediate brassinosteroid responses in rice. *Plant Cell* **2012**, *24*, 2562–2577. [CrossRef] [PubMed]
45. Zhao, D.S.; Li, Q.F.; Zhang, C.Q.; Zhang, C.; Yang, Q.Q.; Pan, L.X.; Ren, X.Y.; Lu, J.; Gu, M.H.; Liu, Q.Q. GS9 acts as a transcriptional activator to regulate rice grain shape and appearance quality. *Nat. Commun.* **2018**, *9*, 1240. [CrossRef]
46. Zhou, Y.; Miao, J.; Gu, H.; Peng, X.; Leburu, M.; Yuan, F.; Gu, H.; Gao, Y.; Tao, Y.; Zhu, J.; et al. Natural Variations in SLG7 Regulate Grain Shape in Rice. *Genetics* **2015**, *201*, 1591–1599. [CrossRef] [PubMed]
47. Wu, B.; Yun, P.; Zhou, H.; Xia, D.; Gu, Y.; Li, P.; Yao, J.; Zhou, Z.; Chen, J.; Liu, R.; et al. Natural variation in WHITE-CORE RATE 1 regulates redox homeostasis in rice endosperm to affect grain quality. *Plant Cell* **2022**, *34*, 1912–1932. [CrossRef] [PubMed]
48. Yang, W.; Xu, P.; Zhang, J.; Zhang, S.; Li, Z.; Yang, K.; Chang, X.; Li, Y. OsbZIP60-mediated unfolded protein response regulates grain chalkiness in rice. *J. Genet. Genom.* **2022**, *49*, 414–426. [CrossRef] [PubMed]
49. Sakamoto, T.; Matsuoka, M. Identifying and exploiting grain yield genes in rice. *Curr. Opin. Plant Biol.* **2008**, *11*, 209–214. [CrossRef]
50. Harberd, N.P. Shaping Taste: The Molecular Discovery of Rice Genes Improving Grain Size, Shape and Quality. *J. Genet. Genom.* **2015**, *42*, 597–599. [CrossRef]
51. Zheng, J.; Zhang, Y.; Wang, C. Molecular functions of genes related to grain shape in rice. *Breed. Sci.* **2015**, *65*, 120–126. [CrossRef] [PubMed]
52. Zuo, J.; Li, J. Molecular genetic dissection of quantitative trait loci regulating rice grain size. *Annu. Rev. Genet.* **2014**, *48*, 99–118. [CrossRef] [PubMed]
53. Li, N.; Li, Y. Signaling pathways of seed size control in plants. *Curr. Opin. Plant Biol.* **2016**, *33*, 23–32. [CrossRef]
54. Liu, L.; Tong, H.; Xiao, Y.; Che, R.; Xu, F.; Hu, B.; Liang, C.; Chu, J.; Li, J.; Chu, C. Activation of Big Grain1 significantly improves grain size by regulating auxin transport in rice. *Proc. Natl. Acad. Sci. USA* **2015**, *112*, 11102–11107. [CrossRef]
55. Ishimaru, K.; Hirotsu, N.; Madoka, Y.; Murakami, N.; Hara, N.; Onodera, H.; Kashiwagi, T.; Ujiie, K.; Shimizu, B.; Onishi, A.; et al. Loss of function of the IAA-glucose hydrolase gene TGW6 enhances rice grain weight and increases yield. *Nat. Genet.* **2013**, *45*, 707–711. [CrossRef] [PubMed]
56. Kabir, M.R.; Nonhebel, H.M. Reinvestigation of THOUSAND-GRAIN WEIGHT 6 grain weight genes in wheat and rice indicates a role in pollen development rather than regulation of auxin content in grains. *Theor. Appl. Genet.* **2021**, *134*, 2051–2062. [CrossRef] [PubMed]
57. Hu, Z.; Lu, S.J.; Wang, M.J.; He, H.; Sun, L.; Wang, H.; Liu, X.H.; Jiang, L.; Sun, J.L.; Xin, X.; et al. A Novel QTL qTGW3 Encodes the GSK3/SHAGGY-Like Kinase OsGSK5/OsSK41 that Interacts with OsARF4 to Negatively Regulate Grain Size and Weight in Rice. *Mol. Plant* **2018**, *11*, 736–749. [CrossRef]
58. Qiao, J.; Jiang, H.; Lin, Y.; Shang, L.; Wang, M.; Li, D.; Fu, X.; Geisler, M.; Qi, Y.; Gao, Z.; et al. A novel miR167a-OsARF6-OsAUX3 module regulates grain length and weight in rice. *Mol. Plant* **2021**, *14*, 1683–1698. [CrossRef] [PubMed]
59. Song, X.J.; Huang, W.; Shi, M.; Zhu, M.Z.; Lin, H.X. A QTL for rice grain width and weight encodes a previously unknown RING-type E3 ubiquitin ligase. *Nat. Genet.* **2007**, *39*, 623–630. [CrossRef]
60. Huang, J.; Chen, Z.; Lin, J.; Guan, B.; Chen, J.; Zhang, Z.; Chen, F.; Jiang, L.; Zheng, J.; Wang, T.; et al. gw2.1, a new allele of GW2, improves grain weight and grain yield in rice. *Plant Sci.* **2022**, *325*, 111495. [CrossRef]
61. Wang, S.; Wu, K.; Yuan, Q.; Liu, X.; Liu, Z.; Lin, X.; Zeng, R.; Zhu, H.; Dong, G.; Qian, Q.; et al. Control of grain size, shape and quality by OsSPL16 in rice. *Nat. Genet.* **2012**, *44*, 950–954. [CrossRef] [PubMed]
62. Wang, S.; Li, S.; Liu, Q.; Wu, K.; Zhang, J.; Wang, S.; Wang, Y.; Chen, X.; Zhang, Y.; Gao, C.; et al. The OsSPL16-GW7 regulatory module determines grain shape and simultaneously improves rice yield and grain quality. *Nat. Genet.* **2015**, *47*, 949–954. [CrossRef] [PubMed]
63. Wang, S.; Li, S.; Liu, Q.; Wu, K.; Zhang, J.; Wang, S.; Wang, Y.; Chen, X.; Zhang, Y.; Gao, C.; et al. Leaf-derived ABA regulates rice seed development via a transporter-mediated and temperature-sensitive mechanism. *Sci. Adv.* **2021**, *7*, eabc8873.
64. Xu, X.; E, Z.; Zhang, D.; Yun, Q.; Zhou, Y.; Niu, B.; Chen, C. OsYUC11-mediated auxin biosynthesis is essential for endosperm development of rice. *Plant Physiol.* **2021**, *185*, 934–950. [CrossRef] [PubMed]
65. Wang, Z.; Wei, K.; Xiong, M.; Wang, J.D.; Zhang, C.Q.; Fan, X.L.; Huang, L.C.; Zhao, D.S.; Liu, Q.Q.; Li, Q.F. Glucan, Water-Dikinase 1 (GWD1), an ideal biotechnological target for potential improving yield and quality in rice. *Plant Biotechnol. J.* **2021**, *19*, 2606–2618. [CrossRef]
66. Saitou, N.; Nei, M. The neighbor-joining method—A new method for reconstructing phylogenetic trees. *Mol. Biol. Evol.* **1987**, *4*, 406–425. [PubMed]
67. Wang, C.; Liu, Q.; Shen, Y.; Hua, Y.; Wang, J.; Lin, J.; Wu, M.; Sun, T.; Cheng, Z.; Mercier, R.; et al. Clonal seeds from hybrid rice by simultaneous genome engineering of meiosis and fertilization genes. *Nat. Biotechnol.* **2019**, *37*, 283–286. [CrossRef] [PubMed]

68. Yang, X.; Lu, J.; Shi, W.J.; Chen, Y.H.; Yu, J.W.; Chen, S.H.; Zhao, D.S.; Huang, L.C.; Fan, X.L.; Zhang, C.Q.; et al. RGA1 regulates grain size, rice quality and seed germination in the small and round grain mutant srg5. *BMC Plant Biol.* **2024**, *24*, 16. [CrossRef]
69. Wang, Z.; Zhou, Y.; Ren, X.Y.; Wei, K.; Fan, X.L.; Huang, L.C.; Zhao, D.S.; Zhang, L.; Zhang, C.Q.; Liu, Q.Q.; et al. Co-Overexpression of Two Key Source Genes, OsBMY4 and OsISA3, Improves Multiple Key Traits of Rice Seeds. *J. Agric. Food Chem.* **2023**, *71*, 615–625. [CrossRef] [PubMed]
70. Xiong, M.; Yu, J.; Wang, J.; Gao, Q.; Huang, L.; Chen, C.; Zhang, C.; Fan, X.; Zhao, D.; Liu, Q.Q.; et al. Brassinosteroids regulate rice seed germination through the BZR1-RAmy3D transcriptional module. *Plant Physiol.* **2022**, *189*, 402–418. [CrossRef] [PubMed]

plants

MDPI

Communication

OsBBX2 Delays Flowering by Repressing *Hd3a* Expression Under Long-Day Conditions in Rice

Yusi Yang [1], Jiaming Wei [2,3], Xiaojie Tian [2], Changhua Liu [1,*], Xiufeng Li [2,*] and Qingyun Bu [2,*]

[1] College of Advanced Agriculture and Ecological Environment, Heilongjiang University, Harbin 150080, China
[2] State Key Laboratory of Black Soils Conservation and Utilization, Key Laboratory of Soybean Molecular Design Breeding, NortheastInstitute of Geography and Agroecology, Chinese Academy of Sciences, Harbin 150081, China
[3] University of Chinese Academy of Sciences, Beijing 100049, China
* Correspondence: liuchanghua70@163.com (C.L.); lixiufeng@iga.ac.cn (X.L.); buqingyun@iga.ac.cn (Q.B.)

Abstract: Members of the B-Box (BBX) family of proteins play crucial roles in the growth and development of rice. Here, we identified a rice BBX protein, Oryza sativa BBX2 (OsBBX2), which exhibits the highest expression in the root. The transcription of *OsBBX2* follows a diurnal rhythm under photoperiodic conditions, peaking at dawn. Functional analysis revealed that *OsBBX2* possesses transcriptional repression activity. The *BBX2* was overexpressed in the rice japonica cultivar Longjing 11 (LJ11), in which *Ghd7* and *PRR37* were non-functional or exhibited weak functionality. The overexpression of *OsBBX2* (*OsBBX2 OE*) resulted in a delayed heading date under a long-day (LD) condition, whereas the *bbx2* mutant exhibited flowering patterns similar to the wild type (WT). Additionally, transcripts of *Ehd1*, *Hd3a*, and *RFT1* were downregulated in the *OsBBX2 OE* lines under the LD condition. *OsBBX2* interacted with *Hd1* (*BBX18*), and the *bbx2 hd1* double mutant displayed a late flowering phenotype comparable to that of *hd1*. Furthermore, *OsBBX2* enhanced the transcriptional repression of *Hd3a* through its interaction with *Hd1*, as demonstrated in the protoplast-based assay. Taken together, these findings suggest that the *OsBBX2* delays flowering by interacting with *Hd1* and co-repressing *Hd3a* transcription.

Keywords: rice; heading date; BBX2; Hd1; Hd3a

Academic Editor: Konstantin V. Kiselev

Received: 23 November 2024
Revised: 17 December 2024
Accepted: 24 December 2024
Published: 27 December 2024

Citation: Yang, Y.; Wei, J.; Tian, X.; Liu, C.; Li, X.; Bu, Q. OsBBX2 Delays Flowering by Repressing *Hd3a* Expression Under Long-Day Conditions in Rice. *Plants* **2025**, *14*, 48. https://doi.org/10.3390/plants14010048

The BBX family of proteins is highly conserved across green plants. They are characterized by one or two B-Box motifs in the N-terminal domain and a specific CCT (CONSTANS, CO-like, and TO) domain at the C-terminus [1]. They are involved in many plant regulatory networks, ranging from seedling photomorphogenesis to stress response [2]. In *Arabidopsis thaliana*, a total of 32 genes encode BBX proteins, while in rice, the BBX family comprises 30 proteins, designated as OsBBX1 to OsBBX30, based on their chromosomal locations. Among these, seven OsBBXs contain two B-box domains and a conserved CCT domain, ten possess one B-box and one CCT domain, three have a single B-box domain, and ten contain two B-box domains [1].

Most *BBX* family members have been reported to play a vital role in floral transition. *CONSTANS (CO)/BBX1* was the first *BBX* family member identified in *Arabidopsis*. *CO* induces flowering by activating the expression of *FLOWERING LOCUS T* (*FT*) and *SUPPRESSOR OF CONSTANS1* (*SOC1*) [3]. Subsequent studies have identified several BBX proteins that regulate flowering. For instance, *BBX7* (*COL9*) delays flowering under long-day (LD) conditions by reducing *CO* expression [4], while *bbx4* (*col3*) plants flower earlier during both short and long days, and *COL3* is a positive regulator of photomorphogenesis [5]. The late flowering phenotypes of overexpressed *BBX32* (*EIP6*) transgenic plants

suggest that its interaction with *EMBRYONIC FLOWER1* (*EMF1*) can regulate flowering time in *Arabidopsis* [6]. *BBX19* functions as a negative regulator of flowering time, and it also functions as a regulator of circadian rhythm by complexing PRR proteins to enhance their repressive effect on *CCA1* transcription [7]. *BBX24* and *BBX6* (*COL5*) function as positive regulators of flowering, with *BBX24* overexpressing lines exhibiting accelerated flowering by striving for *FLOWERING LOCUS C* (*FLC*) to impact downstream flowering genes [8]. *BBX6* overexpression induces early flowering by promoting FT expression [9].

Rice *HEADING DATE 1* (*Hd1*/*OsBBX18*), which is an orthologue of *CO* in *Arabidopsis*, promotes heading under short-day (SD) conditions and inhibiting under long-day (LD) conditions [10]. In addition, a group of B-box-containing proteins were proven to repress flowering in rice. *BBX27* (*OsCO3*), which has two incomplete B-boxes, serves as a flowering repressor upstream of *Hd3a* under SD [11]. The overexpression of *OsBBX26* (*OsCOL15*) suppresses flowering by promoting *Ghd7* and repressing *RID1* under both SD and LD conditions [12]. The overexpression of *OsBBX17* (*OsCOL16*) delays heading under both SD and LD conditions by upregulating *Ghd7* transcripts [13]. *OsCOL13* functions as a negative regulator of flowering downstream of *OsphyB* and upstream of *Ehd1*. On the other hand, OsCOL13 is functionally redundant with BBX5 (OsCOL4), which is a constitutive flowering repressor upstream of *Ehd1* and downstream of *OsphyB* [14]. Both the overexpression of *BBX7* (*OsCOL9*) and *BBX10* (*OsCOL10*) delay flowering under SD and LD conditions by repressing *Ehd1*, *Hd3a*, and *RFT1* expression. OsCOL10 acts as a suppressor of rice flowering time by bridging *Ghd7* and *Ehd1* [15]. *OsBBX14* suppresses rice flowering by regulating either *Hd1* or *Ehd1* under LD or SD conditions [16].

Here, we show that *BBX2* acts as a negative regulator of flowering time under long-day (LD) conditions. *BBX2* physically interacts with *Hd1* to repress the transcription of *Hd3a*.

The *BBX2* (LOC_Os02g07930) gene was overexpressed in the rice *japonica* cultivar Longjing 11 (LJ11), resulting in two independent *BBX2* overexpressing (*OE)* transgenic lines, which exhibited significantly higher levels of *BBX2* expression (Supplemental Figure S1A). The *BBX2 OE* plants flowered outstandingly later than LJ11 under the LD (14 h light/10 h dark) condition, with delays of 9.35 and 10.65 days for the two approaches (Figure 1A,B). An RT-qPCR interpretation exhibited that the mRNA levels of *Ehd1*, *Hd3a*, and *RFT1* were dramatically lessened in the *BBX2 OE* plants compared to LJ11 (Supplemental Figure S1B–D). The decreased expression of the florigen genes and their catalyst *Ehd1* is consistent with the later flowering phenotypes observed in *BBX2 OE* plants (Figure 1A,B).

Most BBX proteins serve as flowering repressors [2]. To inspect the attributes of *BBX2*, we first assessed its transcriptional activity in a rice protoplast system; *BBX2* reduced *LUC* transcription activity compared with GAL4BD, revealing it to be a transcriptional repressor (Figure 1C,D). As BBX2 functions as a negative regulator of flowering, we investigated the diurnal expression modes of *BBX2* under LD and SD conditions. The *BBX2* manifestation began to accumulate slowly at dusk before peaking at dawn and sharply decreasing until dusk to its lowest level under both SD and LD conditions (Supplemental Figure S2A,B). In addition, we examined *BBX2*'s pattern of temporal and spatial expression in diverse tissues, including roots, stems, leaves, panicles, and seeds, by RT-qPCR. *BBX2* was shown mainly in the roots, with a comparative expression in the leaves, stem, and spike, and no expression was observed in the panicles or seeds (Supplemental Figure S2C).

To analyze the underlying mechanism for BBX2 in regulating flowering, we investigated the potential interacting proteins of BBX2 by the STRING database (https://cn.string-db.org/). We first used a yeast two-hybrid system to identify whether BBX2 could physically interact with Hd1, PRR1, and DTH2. The results demonstrated that BBX2 interacts with Hd1, but not with PRR1 or DTH2 (Figure 1E). To confirm this interaction, we applied split luciferase complementation assays by creating merger constructs of *BBX2* and *Hd1* to C- and N-

terminal fragments of luciferase. *Agroinfiltration*-based transient assays in *N. benthamiana* showed the restoration of luciferase activity in the leaves co-infiltrated with BBX2 and Hd1 (Figure 1F).

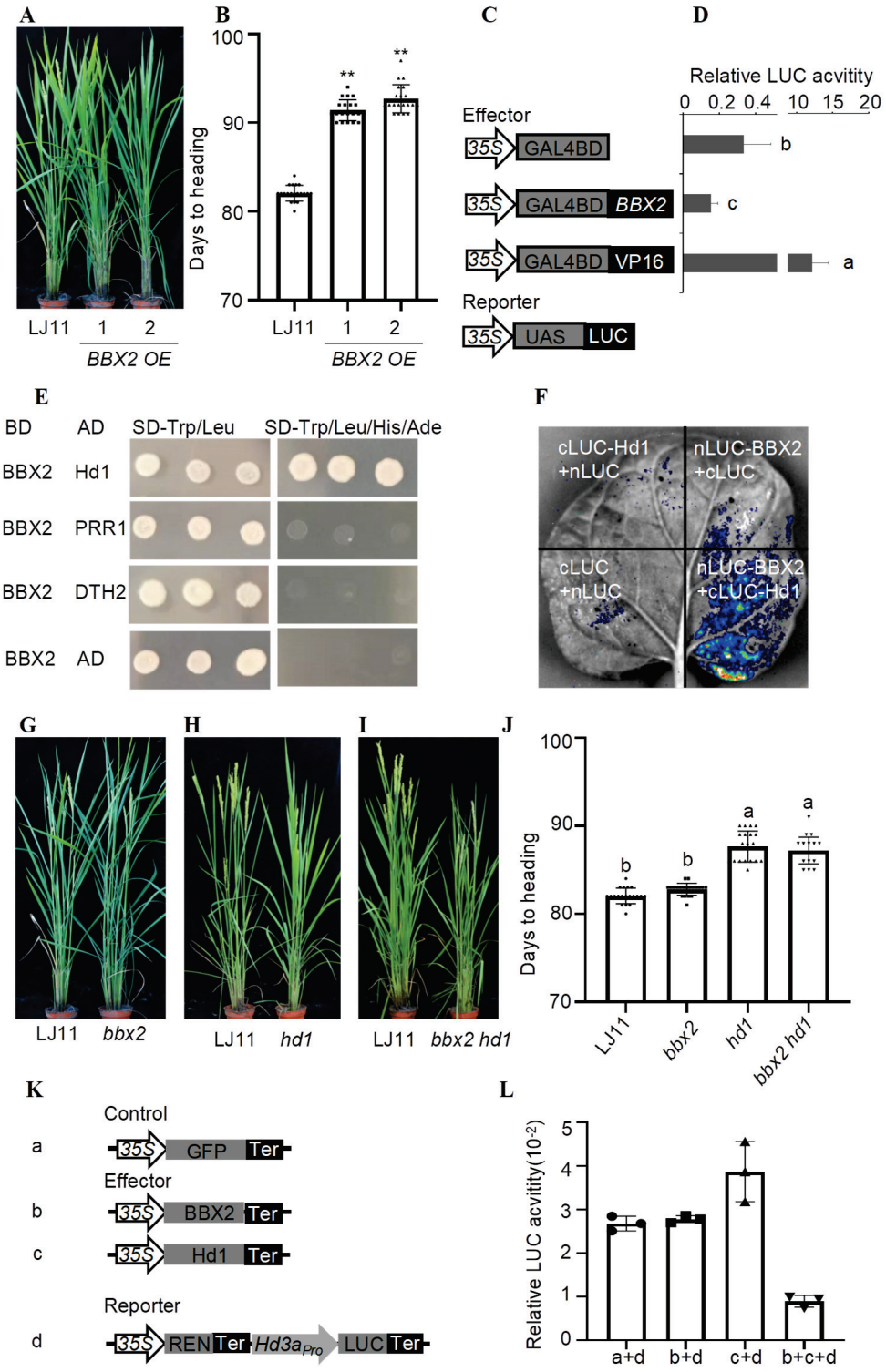

Figure 1. *OsBBX2* delays flowering by repressing *Hd3a* expression. (**A**) Representative image of LJ11 and *BBX OE* plants grown under LD conditions at the heading stage. (**B**) Flowering time of LJ11 and *BBX OE* under LD conditions. Data are means ± standard error (SE; $n = 20$). p values were calculated by Student's t test compared to LJ11; **: $p < 0.01$. (**C**) Schematic diagrams of the reporter plasmids used in the rice protoplast transient assay. REN, Renilla luciferase; LUC, firefly luciferase. (**D**) The

LUC activity in rice protoplasts with indicated reporter plasmids. Data are means ± SE (n = 3). Statistically significant differences are indicated by different lowercase letters ($p < 0.05$, one-way ANOVA with Tukey's significant difference test). (**E**) The yeast two-hybrid assay showed that BBX2 interacts with Hd1. Yeast grew at 30 °C for 3 days. Empty vectors were used as the negative controls. AD, activation domain. BD, binding domain. (**F**) The LCI assay of the *BBX2* interaction with *Hd1* in *N. benthamiana* leaves. The co-transformation of cLUC-*Hd1* and nLUC-*BBX2* led to the re-constitution of the LUC signal, whereas no signal was detected when cLUC-*Hd1* and nLUC, cLUC and nLUC-*BBX2*, and cLUC and nLUC were co-expressed. In each experiment, at least five independent *N. benthamiana* leaves were infiltrated and analyzed. (**G**–**I**) A representative image of LJ11, *bbx2* (**G**), *hd1* (**H**), and *bbx2 hd1* (**I**) mutants grown under the LD condition at the heading stage. (**J**) The flowering time of LJ11, *bbx2*, *hd1*, and *bbx2 hd1* mutants under LD conditions. Data are means ± standard error (SE; n = 20). Statistically significant differences are indicated by different lowercase letters ($p < 0.05$, one-way ANOVA with Tukey's significant difference test). (**K**) Schematic diagrams of the reporter plasmids used in the rice protoplast transient assay. $35S_{Pro}$*:GFP* was used as the control and $35S_{Pro}$*:BBX2*, $35S_{Pro}$*:Hd1*, and $Hd3a_{Pro}$*:LUC* were used as the effectors and reporters. (**L**) Relative LUC activity expressed with reporters and effectors. The expression level of Renilla (REN) was used as an internal control. The LUC/REN ratio represents the relative activity of the *Hd3a* promoter. Data are shown as means ± SE (n = 3). Statistically significant differences are indicated by different lowercase letters ($p < 0.05$, one-way ANOVA with Tukey's significant difference test).

To explore the genetic relationship between *BBX2* and *Hd1*, we generated *bbx2*, *hd1*, and *bbx2 hd1* mutants using CRISPR/Cas9-mediated genome editing in the LJ11 background [17]. Target sites were designed for the first exons of *BBX2* and *Hd1* (Supplemental Figure S3A,B). Despite the different mutation types in the single and double mutants, *BBX2* inserted a G base in the single mutant of *bbx2*, and a G base was missing in the *bbx2 hd1* double mutant. *Hd1* inserted a T base in the single mutant, and the A base was inserted in the *bbx2 hd1* double mutant, for which successful gene knockout was achieved (Supplemental Figure S3C,D). Phenotypic analysis showed that *bbx2* flowers at a similar time to LJ11 under the LD condition, while *hd1* and *bbx2 hd1* mutants flower approximately 5.1 to 5.6 days later than LJ11 under the LD condition (Figure 1G–J). These results suggest that Hd1 functions genetically downstream of BBX2. As a flowering repressor, Hd1 controls the expression of florigen genes under the LD condition. We hypothesized that BBX2 enhances the inhibitory effect on flowering by physically interacting with Hd1. Considering that BBX2 can interact with Hd1, which represses *Hd3a* expression, we performed a transient expression assay in rice protoplasts using $35S_{Pro}$*:GFP* as a control and $35S_{Pro}$*:BBX2*, $35S_{Pro}$*:Hd1*, and $Hd3a_{Pro}$*:LUC* as effectors and reporters, respectively (Figure 1K). The results indicate that $35S_{Pro}$*:Hd1* efficiently suppressed *Hd3a* expression compared to the control $35S_{Pro}$*:GFP*, and this suppression was further enhanced by the addition of $35S_{Pro}$*:BBX2* (Figure 1L).

Ehd1 and *Hd1* are key factors in regulating flowering in rice, particularly under long-day conditions [18]. It has been shown that Hd1 and Ghd7 interact to form a complex, which specifically binds to the cis-regulatory region of *Ehd1*, thereby suppressing its expression [19]. Additionally, studies have shown interactions between Hd1, Ghd7, and PRR37. Notably, Hd1 promotes flowering in both long- and short-day conditions in the *ghd7 prr37* double mutant [20]. This suggests that Hd1 functions in conjunction with Ghd7 and PRR37 under long-day conditions [21], while the *OsELF4s-OsELF3-1-OsLUX* complex reduces the expression of *Ghd7* and *PRR37* under short-day conditions. Therefore, Hd1 plays an independent role in the regulation of flowering [22]. Interestingly, it has been reported that *Ghd7* and *PRR37* are non-functional or exhibit weak functionality in the LJ11 background [23]. Therefore, knocking out *Hd1* in the LJ11 background results in a late flowering phenotype.

As *bbx2* flowered at a similar time to LJ11, this was probably due to functional redundancy with other *BBX* family members. To explore this, we investigated the evolutionary

relationships within the *BBX* gene family. We identified *BBX21* as a homolog of *BBX2*; however, its role in regulating rice flowering remains unknown. Future research will involve creating a *bbx2 bbx21* double mutant to assess potential functional redundancy.

In summary, our study reveals that *BBX2* functions as a negative regulator of flowering, and the expressions of *Ehd1*, *Hd3a*, and *RFT1* are significantly reduced in *BBX2 OE* plants. *BBX2* exhibits transcriptional repression activity and can interact with Hd1. These results suggest that BBX2 interacts with Hd1, co-suppressing the expression of *Hd3a* and consequently delaying flowering in rice.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/plants14010048/s1. Supplemental Figure S1: Characterization of *BBX2 OE* plants in the LJ11 background. Supplemental Figure S2: Expression profiles of *BBX2*. Supplemental Figure S3: Characterization of *bbx2*, *hd1*, and *bbx2 hd1* mutants in LJ11 background. Supplemental Table S1: Primers used in this study. Refs. [24,25] was cited in the Supplementary Materials.

Author Contributions: Q.B. and X.L. designed and supervised the research. Y.Y. and J.W. performed experiments, including material planting and the heading time survey. X.T. and C.L. analyzed the data. X.L. wrote the paper. All authors have read and agreed to the published version of the manuscript.

Funding: This study was supported by the Youth Innovation Promotion Association CAS (Grant No. 2022231), the Natural Science Foundation of Heilongjiang (YQ2022C039), and the Young Scientist Group Project of Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences (2023QNXZ02).

Data Availability Statement: All data generated or analyzed during this study are included in this published article and its Supplementary Materials.

Acknowledgments: We thank Yaoguang Liu for providing the vector *pYLCRISPR/Cas9Pubi-H*.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Huang, J.; Zhao, X.; Weng, X.; Wang, L.; Xie, W. The rice B-box zinc finger gene family: Genomic identification, characterization, expression profiling and diurnal analysis. *PLoS ONE* **2012**, *7*, e48242. [CrossRef]
2. Cao, J.; Yuan, J.; Zhang, Y.; Chen, C.; Zhang, B.; Shi, X.; Niu, R.; Lin, F. Multi-layered roles of BBX proteins in plant growth and development. *Stress Biol.* **2023**, *3*, 1. [CrossRef] [PubMed]
3. Samach, A.; Onouchi, H.; Gold, S.E.; Ditta, G.S.; Schwarz-Sommer, Z.; Yanofsky, M.F.; Coupland, G. Distinct roles of CONSTANS target genes in reproductive development of. *Science* **2000**, *288*, 1613–1616. [CrossRef] [PubMed]
4. Cheng, X.F.; Wang, Z.Y. Overexpression of COL9, a CONSTANS-LIKE gene, delays flowering by reducing expression of CO and FT in *Arabidopsis thaliana*. *Plant. J.* **2005**, *43*, 758–768. [CrossRef]
5. Datta, S.; Hettiarachchi, G.H.; Deng, X.W.; Holm, M. *Arabidopsis* CONSTANS-LIKE3 is a positive regulator of red light signaling and root growth. *Plant Cell* **2006**, *18*, 70–84. [CrossRef] [PubMed]
6. Park, H.Y.; Lee, S.Y.; Seok, H.Y.; Kim, S.H.; Sung, Z.R.; Moon, Y.H. EMF1 interacts with EIP1, EIP6 or EIP9 involved in the regulation of flowering time in *Arabidopsis*. *Plant Cell Physiol.* **2011**, *52*, 1376–1388. [CrossRef] [PubMed]
7. Yuan, L.; Yu, Y.; Liu, M.; Song, Y.; Li, H.; Sun, J.; Wang, Q.; Xie, Q.; Wang, L.; Xu, X. BBX19 fine-tunes the circadian rhythm by interacting with PSEUDO-RESPONSE REGULATOR proteins to facilitate their repressive effect on morning-phased clock genes. *Plant Cell* **2021**, *33*, 2602–2617. [CrossRef] [PubMed]
8. Wang, C.Q.; Guthrie, C.; Sarmast, M.K.; Dehesh, K. BBX19 interacts with CONSTANS to repress FLOWERING LOCUS T transcription, defining a flowering time checkpoint in *Arabidopsis*. *Plant Cell* **2014**, *26*, 3589–3602. [CrossRef] [PubMed]
9. Li, F.; Sun, J.J.; Wang, D.H.; Bai, S.N.; Clarke, A.K.; Holm, M. The B-box family gene STO (BBX24) in *Arabidopsis thaliana* regulates flowering time in different pathways. *PLoS ONE* **2014**, *9*, e87544. [CrossRef] [PubMed]
10. Hassidim, M.; Harir, Y.; Yakir, E.; Kron, I.; Green, R.M. Over-expression of CONSTANS-LIKE 5 can induce flowering in short-day grown *Arabidopsis*. *Planta* **2009**, *230*, 481–491. [CrossRef] [PubMed]
11. Yano, M.; Katayose, Y.; Ashikari, M.; Yamanouchi, U.; Monna, L.; Fuse, T.; Baba, T.; Yamamoto, K.; Umehara, Y.; Nagamura, Y.; et al. Hd1, a Major Photoperiod Sensitivity Quantitative Trait Locus in Rice, Is Closely Related to the *Arabidopsis* Flowering Time Gene CONSTANS. *Plant Cell* **2000**, *12*, 2473–2483. [CrossRef] [PubMed]

12. Kim, S.-K.; Yun, C.-H.; Lee, J.H.; Jang, Y.H.; Park, H.-Y.; Kim, J.-K. OsCO3, a CONSTANS-LIKE gene, controls flowering by negatively regulating the expression of FT-like genes under SD conditions in rice. *Planta* **2008**, *228*, 355. [CrossRef] [PubMed]
13. Wu, W.; Zhang, Y.; Zhang, M.; Zhan, X.; Shen, X.; Yu, P.; Chen, D.; Liu, Q.; Sinumporn, S.; Hussain, K.; et al. The rice CONSTANS-like protein OsCOL15 suppresses flowering by promoting Ghd7 and repressing RID1. *Biochem. Biophys. Res. Commun.* **2018**, *495*, 1349–1355. [CrossRef] [PubMed]
14. Wu, W.; Zheng, X.M.; Chen, D.; Zhang, Y.; Ma, W.; Zhang, H.; Sun, L.; Yang, Z.; Zhao, C.; Zhan, X.; et al. OsCOL16, encoding a CONSTANS-like protein, represses flowering by up-regulating Ghd7 expression in rice. *Plant Sci.* **2017**, *260*, 60–69. [CrossRef] [PubMed]
15. Sheng, P.; Wu, F.; Tan, J.; Zhang, H.; Ma, W.; Chen, L.; Wang, J.; Wang, J.; Zhu, S.; Guo, X.; et al. A CONSTANS-like transcriptional activator, OsCOL13, functions as a negative regulator of flowering downstream of OsphyB and upstream of Ehd1 in rice. *Plant Mol. Biol.* **2016**, *92*, 209–222. [CrossRef]
16. Lee, Y.S.; Jeong, D.H.; Lee, D.Y.; Yi, J.; Ryu, C.H.; Kim, S.L.; Jeong, H.J.; Choi, S.C.; Jin, P.; Yang, J.; et al. OsCOL4 is a constitutive flowering repressor upstream of Ehd1 and downstream of OsphyB. *Plant. J.* **2010**, *63*, 18–30. [CrossRef]
17. Tan, J.; Jin, M.; Wang, J.; Wu, F.; Sheng, P.; Cheng, Z.; Wang, J.; Zheng, X.; Chen, L.; Wang, M.; et al. OsCOL10, a CONSTANS-Like Gene, Functions as a Flowering Time Repressor Downstream of Ghd7 in Rice. *Plant Cell Physiol.* **2016**, *57*, 798–812. [CrossRef]
18. Liu, H.; Gu, F.; Dong, S.; Liu, W.; Wang, H.; Chen, Z.; Wang, J. CONSTANS-like 9 (COL9) delays the flowering time in Oryza sativa by repressing the Ehd1 pathway. *Biochem. Biophys. Res. Commun.* **2016**, *479*, 173–178. [CrossRef] [PubMed]
19. Bai, B.; Zhao, J.; Li, Y.; Zhang, F.; Zhou, J.; Chen, F.; Xie, X. OsBBX14 delays heading date by repressing florigen gene expression under long and short-day conditions in rice. *Plant Sci.* **2016**, *247*, 25–34. [CrossRef] [PubMed]
20. Li, X.F.; Liu, H.Z.; Wang, M.Q.; Liu, H.L.; Tian, X.J.; Zhou, W.J.; Lu, T.X.; Wang, Z.Y.; Chu, C.C.; Fang, J.; et al. Combinations of Hd2 and Hd4 genes determine rice adaptability to Heilongjiang Province, northern limit of China. *J. Integr. Plant Biol.* **2015**, *57*, 698–707. [CrossRef]
21. Nemoto, Y.; Nonoue, Y.; Yano, M.; Izawa, T. Hd1, a CONSTANS ortholog in rice, functions as an Ehd1 repressor through interaction with monocot-specific CCT-domain protein Ghd7. *Plant. J.* **2016**, *86*, 221–233. [CrossRef] [PubMed]
22. Xu, P.; Zhang, Y.X.; Wen, X.X.; Yang, Q.Q.; Liu, L.; Hao, S.L.; Li, J.X.; Wu, Z.Z.; Shah, L.Q.; Sohail, A.; et al. The clock component OsLUX regulates rice heading through recruiting OsELF3-1 and OsELF4s to repress Hd1 and Ghd7. *J. Adv. Res.* **2023**, *48*, 17–31. [CrossRef] [PubMed]
23. Andrade, L.; Lu, Y.L.; Cordeiro, A.; Costa, J.M.F.; Wigge, P.A.; Saibo, N.J.M.; Jaeger, K.E. The evening complex integrates photoperiod signals to control flowering in rice. *Proc. Natl. Acad. Sci. USA* **2009**, *119*, e2122582119. [CrossRef] [PubMed]
24. Ma, X.L.; Zhang, Q.Y.; Zhu, Q.L.; Liu, W.; Chen, Y.; Qiu, R.; Wang, B.; Yang, Z.F.; Li, H.Y.; Lin, Y.R.; et al. A Robust CRISPR/Cas9 System for Convenient, High-Efficiency Multiplex Genome Editing in Monocot and Dicot Plants. *Mol. Plant* **2015**, *8*, 1274–1284. [CrossRef]
25. Tian, X.; Wang, Z.; Li, X.; Lv, T.; Liu, H.; Wang, L.; Niu, H.; Bu, Q. Characterization and Functional Analysis of Pyrabactin Resistance-Like Abscisic Acid Receptor Family in Rice. *Rice* **2015**, *8*, 28. [CrossRef]

plants

MDPI

Article

Global Transcriptomic Analysis of Inbred Lines Reveal Candidate Genes for Response to Maize Lethal Necrosis

Ann Murithi [1,2,3,*], Gayathri Panangipalli [1], Zhengyu Wen [2,4], Michael S. Olsen [5], Thomas Lübberstedt [6], Kanwarpal S. Dhugga [2] and Mark Jung [1]

[1] Corteva Agriscience, 7000 NW 62nd Ave, Johnston, IA 50131, USA; gayathri.panangipalli@corteva.com (G.P.); mark.t.jung@corteva.com (M.J.)
[2] International Maize and Wheat Improvement Center (CIMMYT), Carretera México-Veracruz, Km. 45, Texcoco 56237, Mexico; allen.wen@keygene.com (Z.W.); dhuggaks@gmail.com (K.S.D.)
[3] Genetics and Genomics Graduate Program, Iowa State University, 2014 Molecular Building, 2437 Pammel Dr., Ames, IA 5001, USA
[4] Keygene, Inc., 9600 Gudelsky Dr., Rockville, MD 20850, USA
[5] Bayer, Crop Science Division, 800 N. Lindbergh Blvd., St. Louis, MO 63167, USA; michael.olsen@bayer.com
[6] Department of Agronomy, Iowa State University, 716 Farm House Lane, Ames, IA 50011, USA; thomasl@iastate.edu
* Correspondence: amurithi@iastate.edu

Abstract: Maize lethal necrosis (MLN) is a significant threat to food security in Sub-Saharan Africa (SSA), with limited commercial inbred lines displaying tolerance. This study analyzed the transcriptomes of four commercially used maize inbred lines and a non-adapted inbred line, all with varying response levels to MLN. RNA-Seq revealed differentially expressed genes in response to infection by maize chlorotic mottle virus (MCMV) and sugarcane mosaic virus (SCMV), the causative agents of MLN. Key findings included the identification of components of the plant innate immune system, such as differentially regulated R genes (mainly LRRs), and activation/deactivation of virus resistance pathways, including RNA interference (RNAi) via *Argonaute* (AGO), *Dicer-like proteins*, and the ubiquitin–proteasome system (UPS) via *RING/U-box* and *ubiquitin ligases*. Genes associated with redox signaling, *WRKY* transcription factors, and cell modification were also differentially expressed. Additionally, the expression of translation initiation and elongation factors, *eIF4E* and *eIF4G*, correlated with the presence of MLN viruses. These findings provide valuable insights into the molecular mechanisms of MLN resistance and highlight potential gene candidates for engineering or selecting MLN-resistant maize germplasm for SSA.

Keywords: maize; viruses; MCMV; SCMV; MLN; gene expression; RNA-Seq; disease resistance

Academic Editors: Jian Zhang, Jiezheng Ying, Yifeng Wang and Jie Huang

Received: 30 November 2024
Revised: 10 January 2025
Accepted: 15 January 2025
Published: 20 January 2025

Citation: Murithi, A.; Panangipalli, G.; Wen, Z.; Olsen, M.S.; Lübberstedt, T.; Dhugga, K.S.; Jung, M. Global Transcriptomic Analysis of Inbred Lines Reveal Candidate Genes for Response to Maize Lethal Necrosis. *Plants* **2025**, *14*, 295. https://doi.org/10.3390/plants14020295

1. Introduction

Maize is grown on over 35 million ha and is a staple crop for more than 70 million people in Sub-Saharan Africa (SSA) [1–3]. One of the most serious viral diseases is maize lethal necrosis (MLN) [1]. MLN emerged in Kenya in 2011, wreaking havoc on maize production, with estimated losses ranging between 25 and 100% [1,4]. A lack of MLN-resistant hybrids threatens food security [1]. MLN is caused by the synergistic interaction between maize chlorotic mottle virus (MCMV) from the family *Tombusviridae*, genus *Machlomovirus*, and one of several viruses from the family *Potyviridae*, most commonly sugarcane mosaic virus (SCMV) [4–6].

Coordinated efforts between the International Maize and Wheat Improvement Centre (CIMMYT), Kenya Agricultural Livestock Research Organization (KALRO), and other

organizations to develop MLN-resistant maize varieties proved to be the most economically feasible and environmentally sustainable approach for MLN control [1,4,7]. Lack of resistance in SSA germplasm meant identification of the sources of MLN resistance first and then introgression into the locally adapted hybrids. Association mapping studies and Joint Linkage Association mapping (JLAM) using Drought-Tolerant Maize for Africa (DTMAS), Improved Maize for African Soils (IMAS) panels, and a DH population revealed Quantitative Trait Loci (QTL) associated with MLN resistance on chromosomes 1, 3, 5, 6, 7, and 9 [8–10]. Recently, two parallel genome mapping studies using populations derived from crosses with two MLN-resistant sister lines, KS23-5 and KS23-6, revealed a recessively inherited major-effect QTL on chromosome 6, qMLN.06_157 [11]. The donor lines were developed at Kasetsart University in Thailand [12,13]. Other researchers have worked to identify genes scmv1 and scmv2 associated with SCMV resistance through genetic mapping [14–16]. Recently, an experiment investigating the role of the translation initiation factor 4E (eIF4E) in resistance to SCMV produced edited eIF4E variants that effectively confer resistance to the virus [17]. High-throughput experimental studies such as transcriptomics can provide insights into the complexities of the biological processes activated in plants in response to various treatments [18]. Transcriptomic studies have previously been reported in maize to evaluate the expression of genes in response to various stresses [19–23]. Expression studies on the response of maize plants infected with SCMV identified an atypical h-type thioredoxin, a candidate gene associated with the Scmv1 QTL on chromosome 6 [24–27]. A proteomic study in maize identified several differentially abundant proteins (DAPs) in response to MCMV, including two proteins, disulfide isomerase-like protein and ZmPDIL-1 and peroxiredoxin family protein ZmPrx5. Gene-silencing experiments further validated these findings, demonstrating that the absence of these proteins suppressed MCMV, suggesting they act as positive regulators of the virus [28].

Although qMLN.06_157 and the edited eIF4E offer novel sources of resistance to MLN, more sources are crucial to ensure durable disease management to mitigate the risk of pathogen adaptation and resistance breakdown. We, therefore, hypothesize that the response of the maize transcriptome to MLN infection could help reveal the biological processes involved in MLN defense response and possibly identify new gene targets for breeding or gene manipulation. Understanding the mechanisms and, thus, the plant response to the viral infection will facilitate the identification of novel breeding approaches toward MLN resistance. RNA sequencing (RNA-Seq) is a method used to quantify changes in expression levels in the transcriptome in response to external (biotic or abiotic) or internal (growth and development) stimuli [29,30]. The knowledge of the complete transcriptome in a cell is necessary to interpret the functions of genes, reveal the molecular constituents of tissue and cells, understand development, and, in this case, disease response [29]. Consequently, the objectives of this study were to perform an RNA-Seq analysis to (i) analyze transcriptome changes across five tropical maize inbred lines with varying levels of resistance to MLN, (ii) identify potential candidate genes for MLN resistance, and (iii) associate candidate genes to either SCMV or MCMV resistance.

2. Results and Discussion

2.1. Overview of the RNA-Seq Results

To generate a whole-genome view of changes in the transcriptome of maize plants in response to MLN infection, the expression of mock/null maize samples was compared to MLN-infected plants via high-throughput sequencing. The lines tested were four CIMMYT lines (CMLs) and one line non-adapted to sub-Saharan Africa. Samples infected with MLN were sampled 72 h after inoculation. After low-quality reads and adapter sequence were trimmed, RNA-Seq yielded 33,561,693 to 92,029,813 and 56,165,105 to 114,853,577

clean reads for the mock group and MLN-infected samples, respectively. With a mapping rate >78% in both groups, mapped reads ranged from 27,513,876 to 78,372,589 in the mock group and 46,167,716 to 95,282,527 in the MLN-infected group (Table S1).

After a row-wise normalization of gene expression data in DESeq2, hierarchical clustering was performed to demonstrate technical reproducibility across all samples. This approach grouped samples with similar expression patterns into clusters (Figure S1).

2.2. Gene Expression Levels in Response to MLN in Maize

In this study, transcriptome analysis of leaves inoculated with MLN revealed differentially expressed gene profiles in different genotypes (Supplementary Tables S1 and S2). From field experiments, the average disease scores of the genotypes based on a rating scale of 1 (tolerant) to 9 (susceptible) ranged from 2.3 in KS23-6 to 6.6 in CML536 (Figure 1). KS23-6 is a non-SSA-adapted donor line from Thailand identified during diversity screening for MLN resistance. All the differentially expressed genes were selected based on a log2foldchange ≥ 2 or ≤ -2 with a *p*-value adjusted to ≤ 0.05. Cluster analysis of the differentially expressed genes (DEGs) showed that DEGs can be divided into two groups based on the experimental conditions, indicating a significant change in gene expression after MLN (Figure 2). In the null status, genes with high expression showed a marked decrease in expression after inoculation, while genes that were low in the null showed an increased expression in the treated condition.

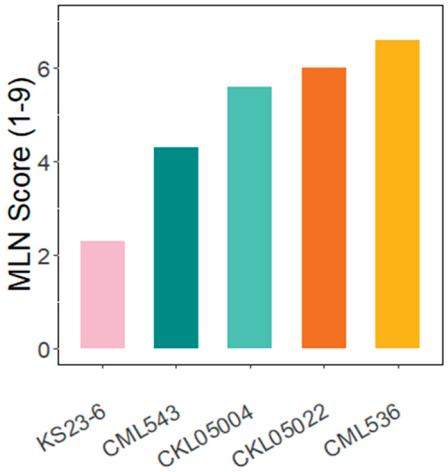

Figure 1. Phenotypic effect of MLN. Average field MLN phenotypic scores of five lines collected in Naivasha, Kenya.

Cluster analysis of the differentially expressed genes (DEGs) showed the presence of two major groups, divided by gene expression levels in the control group and gene levels in the treated group (Figure 2).

Principal component analysis (PCA) of the 39,000 gene models based on their normalized counts is shown in Figure 3. PC1 showed that CML543 and CKL05004 were closely related but divergent from the other CMLs. PC1 showed that CML543 and CKL05004 were closely related but divergent from the other CMLs. CKL05002 appeared more different across all the lines. KS23-6 clustered closer to CML536; however, there was a significant difference in the other lines. KS23-6 is a tropical line from Thailand. This proximity may indicate shared ancestry or a common breeding history influencing their genetic composition.

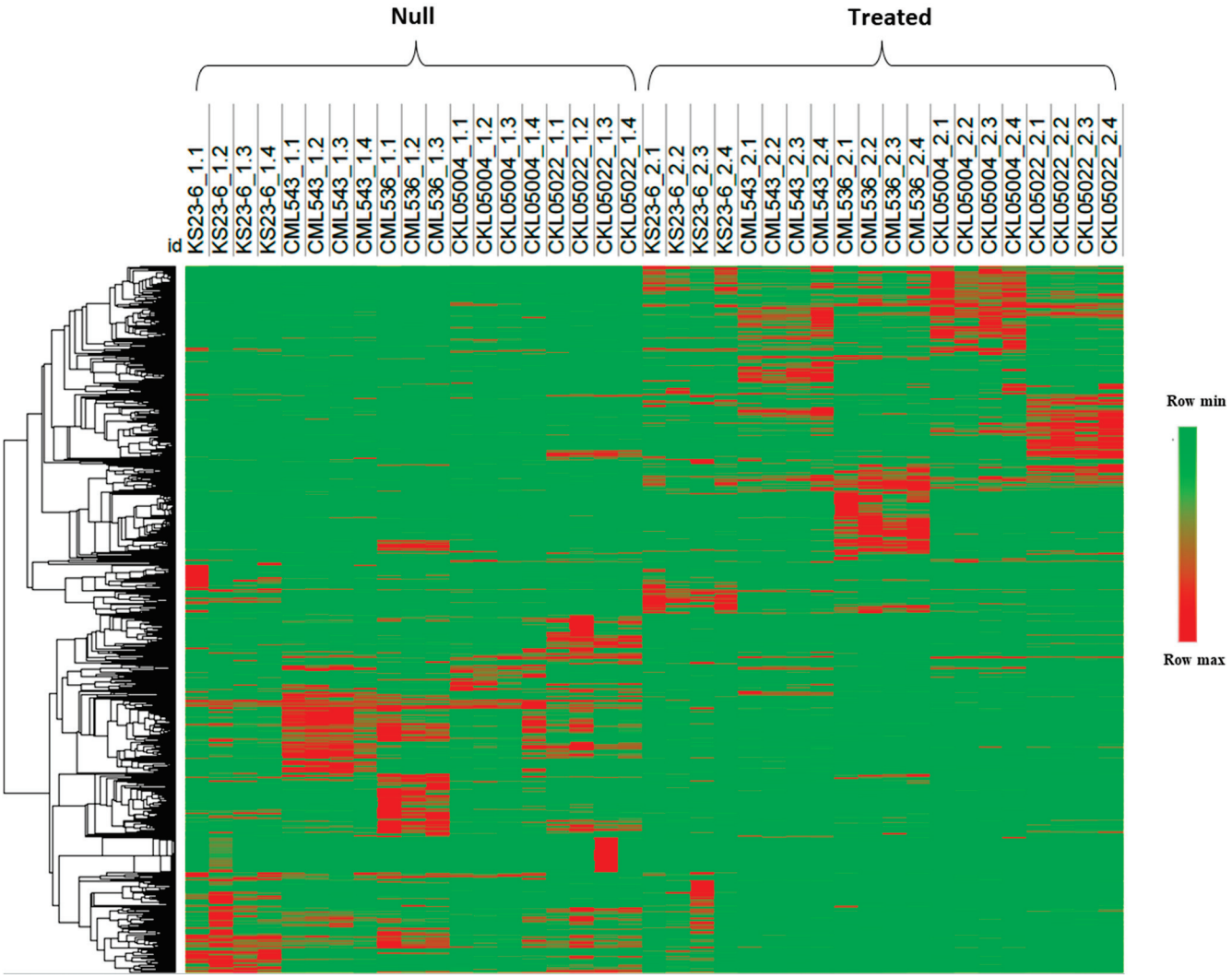

Figure 2. Heatmap clustering of 2503 differentially expressed genes based on their normalized counts. Each row of the heatmap represents the gene count in each genotype before and after inoculation. Red indicates high counts of a gene in a biological replicate, while green indicates lower counts of a gene. "Null" indicates the expression before inoculation, and "treated" indicates expression after MLN inoculation.

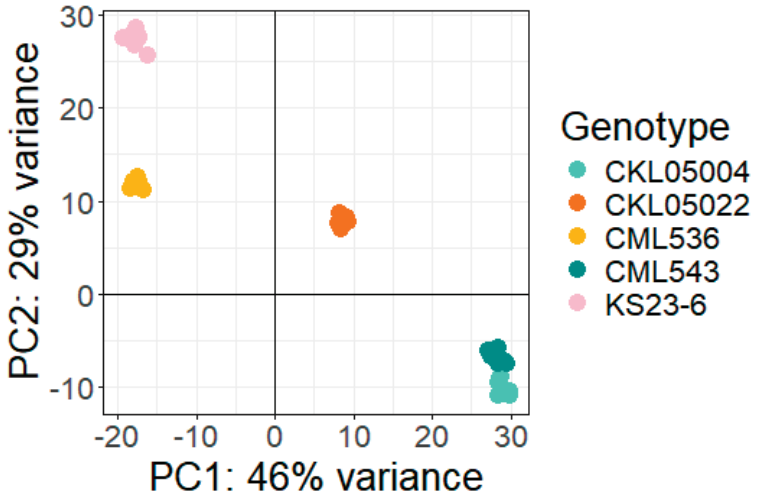

Figure 3. Principal component analysis: Plot generated from normalized gene expression counts of each genotype under MLN.

Following the variability seen in the PCA plot, differential expression was organized into three comparisons: (1) differential expression between the control and inoculated samples in each line (control vs. MLN treatment), (2) the differential expression after contrasting DEGs across CMLs, and (3) differential expression after contrasting resistant versus susceptible samples. The control vs. treatment experiment indicated the DEGs identified in each line (Table S2). Differential expression of the control versus treatment of ~39,000 maize transcripts that mapped on the five lines in this study revealed 341 DEGs in KS23-6, 721 in CML536, 476 in CML543, 537 in CKL05022, and 428 in CKL05004 (Figure 4A, Table S2). The CML comparisons were designed based on the relatedness using the PCA output, creating four comparisons: (a) CML536 vs. CML543, (b) CML543 vs. CKL05002, (c) CML536 vs. CKL05004, and (d) CKL05004 vs. CKL05022. The relationship between CML543 vs. CKL05004 was not considered because of their relatedness in the PCA, while CML536 vs. CKL05022 was eliminated because there was already more variability detected between CML43 vs. CKL050022 and CML536 vs. CKL05004. In genotype vs. genotype comparisons, there were 3768 DEGs in CML536 vs. CML543 (1820 upregulated and 1948 downregulated), 1762 in CML543 vs. CKL05002 (885 upregulated and 877 downregulated), 3679 in CML536 vs. CKL05004 (1828 upregulated and 1851 downregulated), and 1867 in CKL05004 vs. CKL05022 (926 upregulated and 941 downregulated) (Figure 4B, Table S2). The smallest number of DEGs were detected in the control vs. treatment comparisons, where CML536 had more unique genes than all the other lines (Figure 4A, Table S2). This outcome, where DEGs increase when contrasting two genotypes, corresponds with the work performed by [31]. In the comparison of resistant vs. susceptible samples, there were 3907 DEGs in CML536 vs. KS23-6 (1613 upregulated and 2294 downregulated) and 3851 DEGs in CML543 vs. KS23-6 (1647 upregulated and 2204 downregulated) (Figure 4C, Table S3). CML536 and CML543 were chosen as the susceptible genotypes because, although susceptible, their phenotypic scores differed significantly compared to CKL05004 and CKL05022 (Figure 1). Based on the contrast analysis, upregulated genes in one line were downregulated or had no differential in the contrasted line. For instance, since the contrast was coded (1,0) in CML536 vs. CML543, when a gene was upregulated in CML536, it was either downregulated or undetected in CML543 and vice versa.

2.3. Gene Ontology (GO) Enrichment Analysis of Differentially Expressed Genes (DEGs)

Functional enrichment analysis was performed to study the DEGs further. We considered three categories of GO enrichment analysis (biological process (BP), molecular function (MF), and cellular component (CC)). Across all the comparisons and lines, GO terms related to stress/defense response were enriched; for instance, in the control vs. treatment group, GO terms significantly enriched in CML543 included the hydrogen peroxide catabolic pathway (GO:0042744) and cellular response to hydrogen peroxide (GO:0070301) (Figure 5; Table S4). CKL05004 showed enrichment in RNA modification (GO:0009451), while CKL05022, among the many GO terms enriched, was cell surface receptor-linked signaling pathways (GO:0007166) (Figure 5; Table S4). CML536 showed enrichment in DNA duplex unwinding (GO:0032508), the hydrogen peroxide catabolic process (GO:0042744), and protein amino acid autophosphorylation (GO:0046777), while in molecular function, DNA helicase activity (GO:0003678) was enriched (Figure 5; Table S3). On the other hand, KS23-6 showed enrichment of secondary metabolic processes, including flavonoids and phenylpropanoid biosynthetic processes, which were also seen in CML543. Across all the lines, oxidoreductase activity (GO:0016709) was enriched in molecular function.

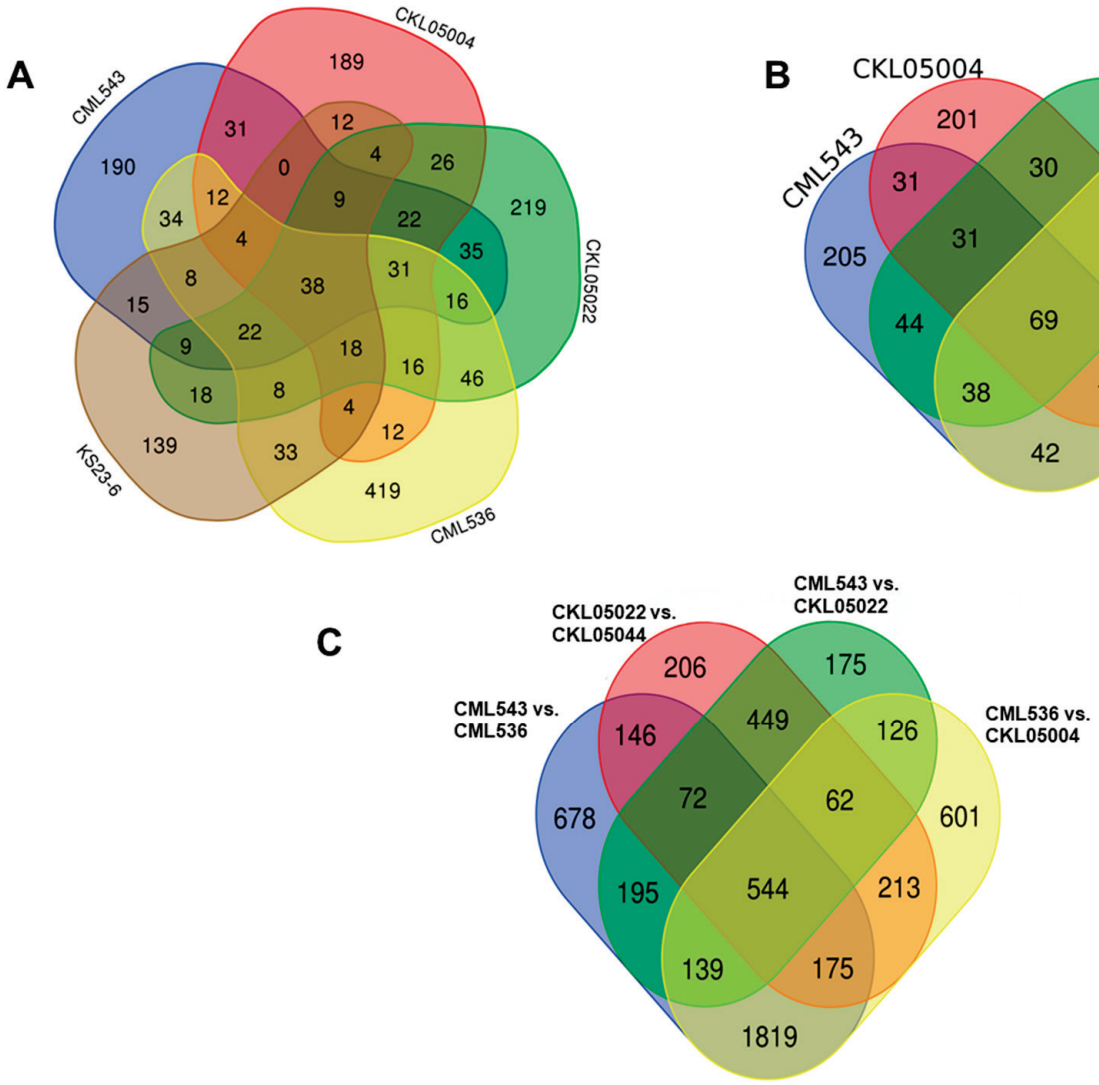

Figure 4. Differentially expressed genes (DEGs). (**A**) Venn diagram of the DEGs in five lines under MLN stress, (**B**) Venn diagram of the DEGs in CIMMYT lines, and (**C**) Venn diagram of the DEGs of the genotype vs. genotype comparison.

There was significant enrichment of DNA recombination (GO:0006310) GO terms as well as GO terms associated with the cell cycle (GO:0007049) across all the genotype vs. genotype comparisons. In the molecular function GO term category, DNA helicase activity (GO:0003678) was enriched in all the CML comparisons (Figure 5; Table S5). In all the comparisons without CML536, the cell surface receptor-linked signaling pathway (GO:0007166) was enriched (Figures 5 and 6; Table S5). Transcription from the RNA polymerase II promoter (GO:0006366) and regulation of transcription from the RNA polymerase II promoter (GO:0006357) were enriched in all the comparisons except for CKL05022 vs. CKL05004. More importantly, the term gene-silencing by RNA (GO:0031047) and the defense response to oomycetes (GO:0002229) were enriched in CML543 vs. CKL05022, while the term cellular response to stress was enriched in CML536 vs. CKL05004, CML536 vs. KS23-6, and CML543 vs. KS23-6 (Figures 5 and 6; Table S5).

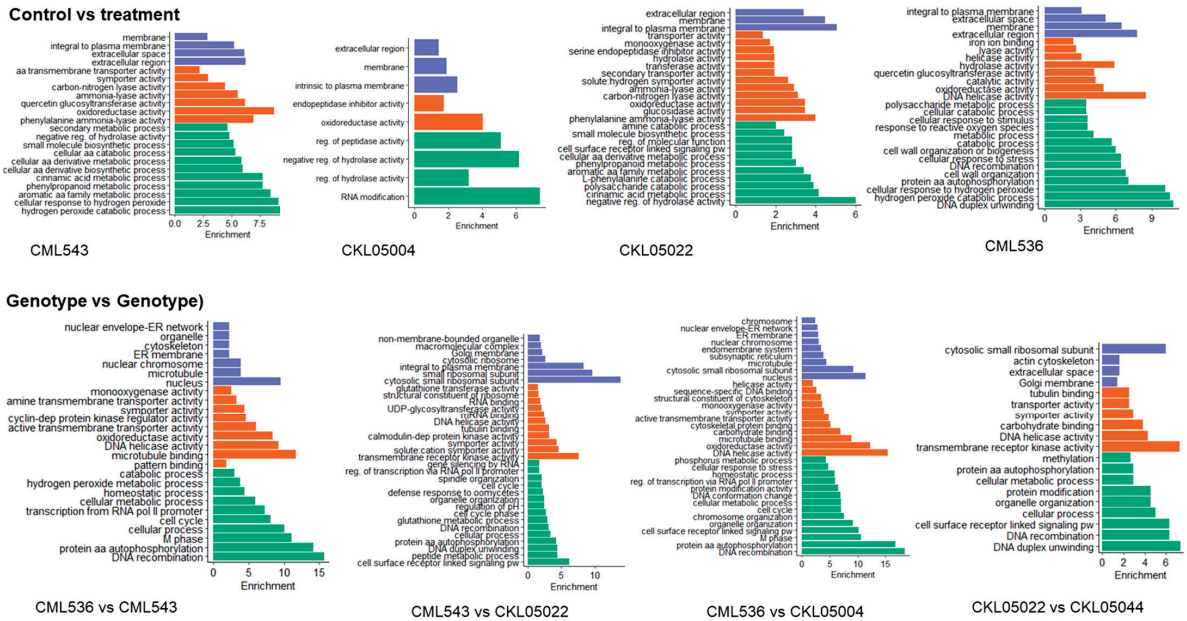

Figure 5. Gene Ontology (GO) enrichment of differentially expressed genes (DEGs) after MLN. A Fisher's exact test and Bonferroni correction were used to identify the significantly (p-value < 0.05) enriched GO terms from the total DEGs across the CIMMYT lines relative to all genes in the maize genome. The first group represents enriched GO terms in each genotype, while the second group represents GO enrichment after contrasting expressions between the lines. The Y-axis represents the DEGs' biological functions, biological process (green), molecular function (orange), and cellular component (purple). The X-axis represents the positive values of the estimated p-values, calculated as $-\log10(p\text{-value})$ via GO term analysis.

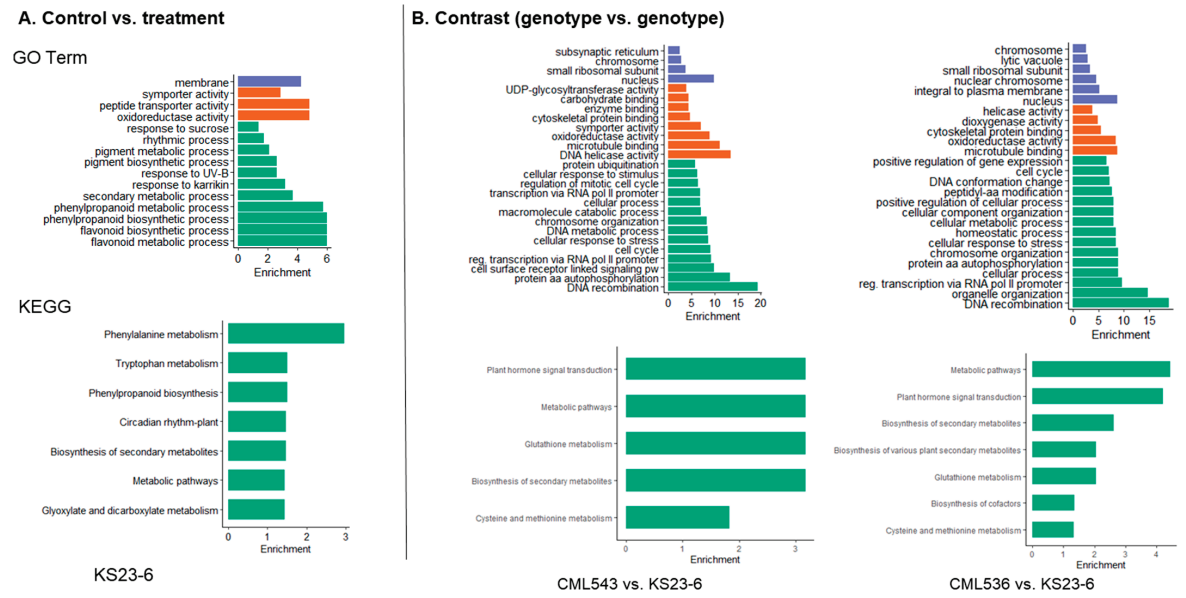

Figure 6. Gene Ontology (GO) and KEGG enrichment of DEGs in KS23-6. Column (**A**) shows the GO terms and KEGG of DEGs in KS23-6 after comparing the control vs. treatment groups, and column (**B**) shows GO terms and KEGG pathways after comparing KS23-6 DEGs to CML543 and CML536. In both, a Fisher's exact test and Bonferroni correction were used to identify the significantly (p-value < 0.05) enriched GO terms and KEGG pathways. In the GO term bar charts, the Y-axis represents the DEGs' biological process (green), molecular function (orange), and cellular component (purple), while the X-axis represents the positive values of the estimated p-values, calculated as $-\log10(p\text{-value})$ via Go term analysis.

To characterize the complex behavior of the maize transcriptome, all the DEGs were subjected to a KEGG pathway enrichment analysis. Similar to GO enrichment, in KEGG, pathways associated with stress/defense response were enriched. In the control vs. treatment comparison, the biosynthesis of secondary metabolites was central, where biosynthesis and metabolism of various secondary metabolites were indicated across all the CMLs and KS23-6 (Figures 6 and 7; Table S6). Some secondary metabolite pathways observed were flavonoid biosynthesis, phenylpropanoid biosynthesis, carotenoid biosynthesis, and diterpenoid biosynthesis (Figure 7; Table S6). Only one KEGG pathway, Ribosome biogenesis in eukaryotes, was identified in CKL05004 (Table S6). In the genotype vs. genotype comparisons, including the biosynthesis of secondary metabolites, plant hormone signal transduction was also central. In CML536 vs. CML543, CML536 vs. KS23-6, and CML543 vs. KS23-6 comparisons, the pathway involving plant–pathogen interactions was enriched (Figures 6 and 7; Table S6). The glutathione metabolism pathway was enriched only in CML543 vs. CKL05022, CML536 vs. KS23-6, and CML543 vs. KS23-6 comparisons.

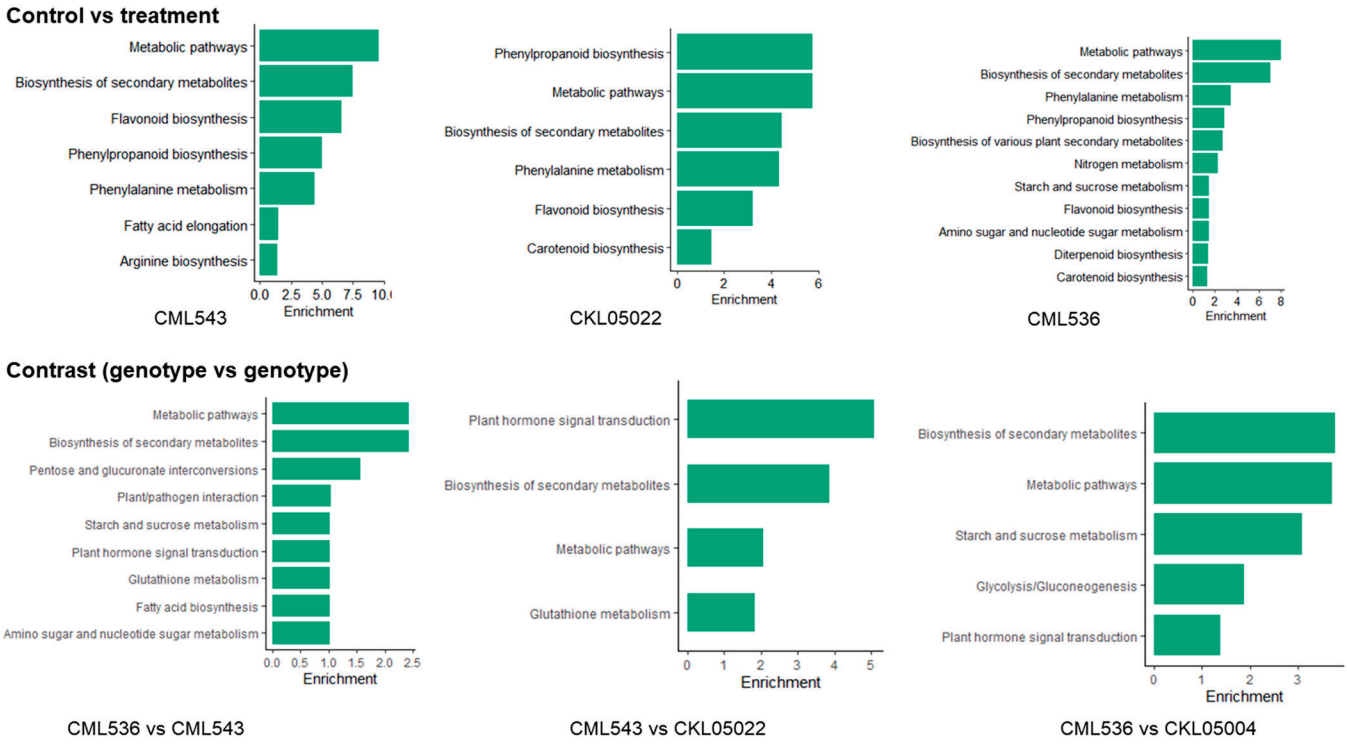

Figure 7. Kyoto Encyclopedia of Genes and Genomes (KEGG) pathway enrichment analyses. A Fisher's exact test and Bonferroni correction were used to identify the significantly (p-value < 0.05) enriched KEGG terms from the total DEGs. The first group represents enriched GO terms in each genotype, while the second group represents GO enrichment after contrasting expressions between the lines. The *Y*-axis represents the KEGG terms. The *X*-axis represents the positive values of the estimated p-values, calculated as $-\log10(p\text{-value})$.

2.4. Differentially Expressed Genes (DEGs) Involved in Stress-Related Pathways

To understand the key DEGs regulating MLN response, a functional annotation of all the DEGs was performed. Plants combat viral infection by restricting viral replication and movement through defense mechanisms such as immune receptor signaling (innate immunity), gene silencing, protein degradation (autophagy), and the regulation of metabolic components [32,33]. On the other hand, plant susceptibility genes (S-genes) promote establishment by providing the viruses with translation machinery, aiding in forming replication compartments, and facilitating replication and cell-to-cell movement. The comparisons of control vs. treatment and genotype vs. genotype presented DEGs that may play an

important role in either positively or negatively influencing the establishment of MCMV or SCMV. Although KS23-6 is the most resistant line, its resistance phenotype is primarily associated with MCMV, while the CMLs, though not significantly resistant, have been linked to resistance against SCMV through various QTL mapping studies, and in some lines like CML543, MCMV-resistance QTLs were identified [34]. This is not surprising, as CMLs are adapted to sub-Saharan Africa, where SCMV is endemic, whereas KS23-6 originates from Thailand, a region where MCMV was reported prior to its first documented occurrence in Kenya and Africa in 2012. Consequently, it is important to recognize that the resistant vs. susceptible comparisons may highlight genes associated with resistance to either MCMV or SCMV. Together, the individual genotype response (control vs. treatment) (Table S2) and the contrast of genotype vs. genotype (Table S3) revealed gene models associated with disease and stress-/defense-related pathways such as the plant's innate immune system, DNA damage repair, cell wall modifications, redox signaling, phytohormone expression, the ubiquitin–proteasome system (UPS), RNAi silencing, and metabolic disruption. Differential expression of translation initiation factors was also considered as a potential indicator of susceptibility genes.

2.5. Innate Plant Immune Response

The pathogen-/microbe-/damage-associated molecular pattern (PAMP/MAMP/DAMP)-triggered immunity (PTI) is the first line of defense against a pathogenic attack, while effector-triggered immunity (ETI), the second line, is triggered by interactions of plant R proteins and pathogen effector proteins [32,35]. The R proteins prevent pathogen movement either by directly or indirectly recognizing the virulence effectors emitted by the pathogen, thus activating immune responses and a cascade of reactions, resulting in the induction of genes associated with defense responses [32,35]. From GO terms, the cell surface receptor-linked signaling pathway, protein amino acid autophosphorylation, cellular response to stress, and cellular response to stimulus were enriched. Many of these receptor kinases, like *leucine-rich repeat* (LRR) protein kinase family protein, were identified in CML536 (Table S2); however, in the genotype vs. genotype comparisons, the number of differentially expressed LRRs was >20 (Table S3). The pattern of expression of the LRR- and NB-ARC-domain-containing disease resistance protein was 50/50 across the CML comparisons, where the number of upregulated and downregulated LRRs was almost equal; for instance, of the 52 LRRs identified in the CML536 vs. CML543 comparison, 25 were downregulated, while the other 27 were upregulated. Similarly, in CML536, of the 16 LRRs identified, 8 were upregulated, and the other 8 were downregulated. On the other hand, of the 43 LRRs identified in the CML536 vs. Ks23-6 comparison, 26 were downregulated, while of the 47 identified in the CML543 vs. KS23-6 comparison, 26 were also downregulated (Table S3). This trend shows that more LRRs were upregulated in KS23-6 compared to the CMLs. Resistance to viruses via LRRs is rare; however, one well-characterized NB-LRR is the tobacco N protein that directly interacts with the helicase domain of the *tobacco mosaic virus* (TMV) replicase to establish resistance [36]. Furthermore, extreme resistance associated with NLR genes has also been characterized in *potato virus X* (PVX) and resistance to certain *soybean mosaic virus* (SMV) strains [37]. Gene models for disease resistance family protein with domain, LRR, CC-NBS-LRR, or dirigent-like protein were also differentially expressed in the genotype vs. genotype comparisons (Table S3). This expression pattern of LRRs and other cell receptor kinases seen here implies a possible resistant vs. susceptible mechanism, where the host produces genes to divert the viruses. In contrast, the viruses may have a stronger mechanism to suppress/bypass the host.

Differential expression of transcription factors was notable across the comparisons, indicating diverse transcriptional regulations in response to MLN. Slavokhotova [38] showed differential expressions of TFFs, including *WRKY*, *NAC*, and *MYB*, in a transcriptomic

analysis of cucumber plants infected with cucumber mottle mosaic virus. Similarly, under SCMV, Akbar [39] demonstrated an upregulation of *WRKY*, *AP2*, *NAC*, *bZIP*, and *bHLH* in the resistance genotype, indicating a possible mechanism for establishing resistance. The distribution and regulation of *WRKY* TF show an upregulation in the susceptible genotypes and a downregulation in the more tolerant genotypes. For instance, all the WRKY identified in three CMLs were upregulated except in CML543 (Table S2). In the CML536 vs. CML543 comparison, *WRKY* was mainly (4 out of 13 were downregulated) upregulated (Table S3). Interestingly, *WRKY* in KS23-6 was upregulated, while the *WRKY* gene models detected in the CML543 vs. KS23-6 comparison were mainly (5 out of 17 were upregulated) downregulated. This comparison shows that nine *WRKY* genes were downregulated in the CML536 vs. CML543 comparison, while 12 were downregulated in the CML543 vs. KS23-6 comparison. This means there was more upregulation of *WRKY* gene models in KS23-6 compared to CML543, while the CML536 vs. CML543 comparison showed more upregulation in CML543. In virus resistance, *WRKY1-WRKY3* are attributed to N-gene-mediated resistance to *tobacco mosaic virus* (TMV), while *WRKY8* functions by restricting the long-distance movement of crucifer-infecting TMV (TM-cg) [40,41]. The expression pattern identified here could indicate a function of *WRKY* TFs in facilitating tolerance or susceptibility to the MLN virus.

Redox Signaling

The production of reactive oxygen species (ROS) is one of the first defense responses employed by plants, which can contribute to plants' defense against viral infections or facilitate viral infection and spread [42,43]. Moderate ROS levels serve as adaptive and defensive signals, whereas excessive ROS can lead to oxidative damage, impair cellular functions, and exacerbate symptom development during viral infections [42]. In all the comparisons, GO terms associated with the hydrogen peroxide catabolic process and oxidoreductase activity were enriched (Figure 5). In CML536, 10 peroxidase genes were identified, where only one was upregulated, while four of the nine identified in CML543 were upregulated (Table S2). Across the lines, more peroxidase gene models were upregulated in CML543 compared to the other CMLs, while CKL05022 and CK05004 showed more upregulation patterns than CML536. In the CML536 vs. CML543 comparison, more peroxidase gene models were shown to be differentially expressed in favor of CML543; however, in the comparison involving KS23-6, all the peroxidases are downregulated, meaning that they were majorly upregulated in KS23-6. This ROS regulation confirms the findings on redox regulation in the study of proteomic changes in maize infected with MCMV [28]. Although their exact functions are not fully understood, Dang demonstrated that ROS response genes, such as *protein disulfide isomerases* (PDIs) and *peroxiredoxin proteins* (Prx), positively regulate MCMV accumulation. Interestingly, the upregulation of peroxidase in KS23-6 and CML543 contradicts these findings; we hypothesize that the presence of SCMV may also influence the regulation of peroxidases. Since the accumulation of ROS is a main symptom of virus infection leading to cell death and necrosis, the expression pattern of peroxidases seen here could explain why the symptoms of MLN are less severe in CML543 than in other CMLs. Furthermore, these results are consistent with Dang's [28] finding where *2-oxoglutarate (2OG)-* and *Fe (II)-dependent oxygenase superfamily protein*, a component that plays a vital role in the generation of ROS, are upregulated in CML536 (susceptible) and downregulated in KS23-6 and CML543 (tolerant). Here, the expression pattern of these two gene families likely points to their possible involvement in preventing the accumulation of MLN viruses and, more specifically, MCMV.

Conversely, gene models of the *thioredoxin* family protein appear to be mainly upregulated in CML536 and downregulated in KS23-6 (Table S2). The gene, which encodes an

atypical *h-type thioredoxin*, was found to be the causal gene for SCMV-resistant QTL, scmv1, on chromosome 6 [24]. The *thioredoxin* family protein is unique to the comparisons with CML536, where five out of seven identified in the CML536 vs. CML543 comparison were upregulated, while eight out of the nine identified in the CML536 vs. CKL05004 comparison were upregulated (Table S3). More importantly, in the CML536 vs. KS23-6 comparison, 8 out of 12 *thioredoxins* were upregulated, while in the CML543 vs. KS23-6 comparison, 6 out of 11 were upregulated. Following the function of *thioredoxin* in SCMV resistance, this regulation indicates it may be related to resistance to SCMV. However, it is worth noting that although the induction of *thioredoxin* is higher in CML536, it is still the most susceptible genotype, implying other possible mechanisms of SCMV that subvert the host.

Another important class of ROS-scavenging protein identified was the *Glutathione S transferase* (GST) family protein identified in CKL05004, where it is upregulated and downregulated in KS23-6 (Table S2). Increased GST enzymatic activity has been linked to SCMV resistance in sorghum, where higher GST levels contribute to stronger, early-stage resistance, while reduced GST activity is associated with susceptibility [44]. Interestingly, GO terms associated with the glutathione metabolic process were enriched in the CML543 vs. CKL05004 comparison (Figure 5). In contrast, the KEGG pathway of Glutathione metabolism was enriched in the CML536 vs. CML543 and CML543 vs. CKL05022 and both the KS23-6 comparisons (Figures 6 and 7). In the CML536 vs. CML543 comparison, GST was mainly downregulated (Table S3). On the other hand, there seems to be a balanced regulation of GST in the CML536 vs. KS23-6 comparison; however, in the CML543 vs. KS23-6 comparison, GST was upregulated (Table S3). More so, in the CML543 vs. CKL05004 comparison, GST was mainly upregulated. The expression of GST showed an upregulation in CML543, followed by CKL05022 and CKL05004, with more downregulation in CML536 and KS23-6. The regulation of these ROS-scavenging proteins (thioredoxin and GST) further confirms that SCMV resistance is majorly associated with the CMLs, while MCMV resistance is stronger in KS23-6.

2.6. RNAi and Ubiquitin–Proteasome System (UPS)

To achieve successful infection, viruses hijack the UPS to keep the viral protein stable, while plants utilize the UPS system as a defense mechanism to eliminate viral protein [45]. The *RING/U-box superfamily protein*, along with an F-box protein ubiquitin, is a component of the proteasome ubiquitination complex, which has been shown to interact with begomovirus [46]. In this analysis, the *RING/U-box superfamily protein* gene models were mainly upregulated across the lines, including KS23-6; however, in the CML536 vs. KS23-6 and CML543 vs. KS23-6 comparisons, the RING/U-box gene models were more upregulated in KS23-6 (Tables S2 and S3). On the other hand, 8 out of 26 were downregulated in the CML536 vs. CML543 comparison, indicating a higher upregulation in CML543. This trend was also seen in the CML536 vs. CKL05004 comparison, where more upregulation was seen in CKL05004. A similar expression pattern was also seen in gene models associated with ubiquitination, where they were upregulated across the lines, but more induction was seen in KS23-6 and CML543 in the genotype vs. genotype comparisons (Table S3). Therefore, the trend shows higher induction of RING/U-box and ubiquitin gene models in KS23-6 and CML543 compared to the other CMLs, where CML536 showed a downregulated trend in the genotype vs. genotype comparisons. The role of the UPS system in resistance to MLN viruses has not been established. Still, the differential expression seen here indicates a possible interaction of the viruses with these proteins that either help with their establishment (CML536) or, to a low degree, provide some tolerance (KS23-6/CML543) to the viruses.

Alteration in the expression of gene models associated with the RNA-silencing pathway (RNAi) is another potential antiviral defense mechanism observed here that results in

the translational repression of viral RNA [47–49]. A gene model involved in gene silencing identified here is the *Argonaute*, which was detected and downregulated in CML543, CML536, and KS23-6 (Table S3). However, the pattern of induction is higher in CKL05004 and CKL05022 than in CML543 and CML536. Alternatively, in the comparisons with KS23-6, *Argonaute* gene models were mainly upregulated in KS23-6 (Table S3). Overall, in relation to the detection, *Argonaute* was mainly downregulated in the CMLs. Consistent with the regulation of *Argonaute*, the *dicer-like protein* was identified and downregulated in the genotype vs. genotype comparisons with CML543. Like the dicer-like protein, gene models associated with *double-stranded RNA-binding protein* (DsRBD protein-related) were detected in the CML543 comparison, where they were downregulated (Table S3). The DsRBD protein-related was also detected in the CML536 vs. CKL05004 comparison, where it was upregulated, indicating a downregulation in CKL05004. To promote their establishment, well adapted plant viruses have evolved mechanisms to produce silencing-suppressor proteins, which can inhibit the host antiviral response based on silencing [50,51]. The potyviral *helper component proteinase* (HCPro) is a well-studied suppressor of antiviral RNA silencing. It performs various functions to inhibit the production of virus-derived siRNAs (vsiRNAs), including binding to ds-RNA, binding to HEN1, inhibiting the activity of HEN1 methyltransferase, interacting with RAV2 to block primary siRNA production, and downregulating RDR6 [49]. From the expression pattern indicated here, especially in CML543, it is possible that the HC-Pro of SCMV plays a major role in downregulating the host silencing machinery to facilitate its colonization, indicating the activation of the host RNA-silencing cascade is important in establishing SCMV.

2.7. Cell Cycle Regulation and DNA Damage and Repair

Viruses can alter the host cell cycle to establish conditions that support their replication by disrupting cell cycle checkpoints [52]. Across the genotype vs. genotype comparison, at least one GO term associated with cell cycle, M phase, chromosome organization, DNA recombination, and DNA conformation change was enriched (Figures 5 and S7 and Tables S4 and S5). Including CML536 and the genotype vs. genotype comparisons, the regulation of PIF1 helicase was central (Tables S2 and S3). The *PIF1 helicase* gene models were identified across all the comparisons, where of the 18 identified in the CML536 vs. CML543 comparison, 5 were upregulated, while in the CML536 vs. CKL05004 comparison, of the 22 identified, 19 were upregulated (Table S3). In the CML543 vs. CKL05022 comparison, of the seven helicases, five were upregulated, while in the CKL05022 vs. CKL05004 comparison, all the nine *PIF1 helicases* identified were downregulated (Table S3). Interestingly, in the comparisons with KS23-6, the expression of PIF1 helicase was 50/50, where half of the identified gene models were upregulated, and the other half were downregulated (Table S2). The high differential expression of gene models linked to PIF1 helicases was not clear in regard to SCMV and MCMV, but helicases have been shown to be required by viruses for the transcription of viral mRNAs, translation, disruption of RNA–protein complexes, and packaging of nucleic acids into virions [53]. The *DNA repair (Rad51) family protein, RAD3-like DNA-binding helicase* protein and *RPA70-kDa subunit B* were other DNA repair genes that were identified across the comparisons (Table S3). This pattern of regulation here shows an increased induction of genes associated with damage repair in the more tolerant genotype compared to the susceptible genotype (CML536).

Plant *cyclin* or *cyclin-dependent kinase* CDK members are known for their roles in cell cycle progression, transcriptional regulation, DNA repair, and defense responses [54]. Cui [55] showed that in Arabidopsis thaliana, CDK proteins play a direct role in phosphorylating the C-terminal domain (CTD) of RNA polymerase II and regulating the transcription cycle and are essential for *cauliflower mosaic virus* (CaMV) infection, as the virus utilizes these proteins

to activate the transcription of its genes. Across the lines, all gene models for cyclin family protein were downregulated (Table S2). In the CML536 vs. CML543 comparison, cyclin was downregulated, and six cyclin genes identified in the CML543 vs. CKL05022 comparison were upregulated. This indicates the upregulation of CDK in CML543. All the seven cyclin gene models identified in the CML536 vs. CKL05004 comparison were downregulated, indicating an upregulation in CKL05004 (Table S3). In the KS23-6 comparisons, CDK appears to be upregulated in KS23-6 compared to CML543 and CML536. The expression pattern observed here shows an upregulation of CDK gene models in KS23-6 and CML543 compared to CML536, aligning with the increased expression of DNA damage repair genes. This suggests a heightened repair response in the resistant genotypes compared to the susceptible ones. Further research is required to determine the connection between SCMV or MCMV and the suppression of damage repair mechanisms observed in the susceptible genotypes.

2.8. Eukaryotic Translation Initiation Factor Expression Under MLN

The translation initiation factors (eIFs) constitute another possible viral resistance system identified in this analysis. The host eIFs assist in the translation of the viral genome, facilitating viral multiplication [56–58]. Factors such as *eIF3, eIF4A, eIF4E*, and *eIF5A* are associated with vital processes in vegetative and reproductive growth [59]. In contrast, *eIF1A, eIF2, eIF4*, and *eIF5A* are associated with abiotic and biotic stresses, and eIF4E and eIF4G have been linked to direct interaction with many viruses [60,61]. The eIFs are considered susceptibility (S) genes as they are crucial in viral replication. Both natural and induced recessive resistance associated with mutations in eIF4s (*eIF4E, eIF4G*, and their isoforms) have been exploited in many plants to confer viral resistance [60] (Table 1).

Table 1. Differential expression of eukaryotic initiation factor 4s (eIF4s) and their isoforms after MLN infection in each genotype. The color show a gradient of downregulation (negative) and upregulation (positive), where red is downregulated and Green is upregulated.

Gene_Name	Gene_ID	KS23	CML543	CKL05004	CKL05022	CML536	Protein Name	Chr
eIF4B_10	Zm00001d025692	−0.22	−0.21	−0.09	−0.01	−0.07	Eukaryotic translation initiation factor 4B1	Chr10
eIF4B2_1	Zm00001d003288	−0.01	−0.22	0.08	0.14	−0.15	Eukaryotic translation initiation factor 4B1	Chr2
eIF4E1_3	Zm00001d041682	−0.20	−0.16	0.06	0.12	−0.06	Eukaryotic translation initiation factor 4E	Chr3
eIF4E2_3	Zm00001d041973	0.32	−0.12	−0.06	0.34	−0.01	Eukaryotic translation initiation factor 4E	Chr3
eIF(iso)4E2_1	Zm00001d032775	−0.23	−0.24	0.10	−0.05	−0.30	Eukaryotic initiation factor 4E protein	Chr1
eIF4G_2	Zm00001d006573	−0.21	−0.02	0.17	0.12	0.24	Eukaryotic translation initiation factor 4G	Chr2
eIF4G_5	Zm00001d017310	0.03	0.00	−0.02	−0.14	−0.12	Eukaryotic translation initiation factor 4G	Chr5
eIF4G_7	Zm00001d021741	−0.38	−0.17	0.09	0.10	0.15	Eukaryotic translation initiation factor 4G	Chr7
eIF4G_10	Zm00001d025979	−0.36	−0.07	0.07	0.04	0.18	Eukaryotic translation initiation factor 4G	Chr10
eIF(Iso)4G_2	Zm00001d003147	−0.14	0.00	0.07	0.07	0.08	MIF4G domain-containing protein	Chr2
eIF(Iso)4G_10	Zm00001d025777	−0.06	0.00	0.06	0.16	0.12	MIF4G domain-containing protein	Chr10

Table S8 shows the expression pattern of different translation initiation factors. We found 32 eIFs, including 1 *eIF1, 6 eIF2s, 8 eIF3s, 11 eIF4s* (4B, 4E, 4G, and their isoforms), 3 *eIF5s, 2 eIF6s, and 1 eIF7*, respectively. Although the expression values are not significant based on the base mean change in gene count between the treatments (log2foldchange > +2 or <−2), there is a detectable differential expression of translation initiation factors correlated to MLN stress (Table S8). For instance, the expression pattern of eIF4E, eIF4E1_3 (Zm00001d041682) and *eIF4E2_3* (Zm00001d041973), and *eIF(iso)4E2_1* (Zm00001d032775) varied across the genotypes (Table 1, Figure 8). In KS23-6, the expression pattern shows an upregulation of *eIF4E1_3* and *eIF(iso)4E2_1* but a downregulation in eIF4E2_3, while CKL05004 portrays an opposite expression pattern to that of KS23-6 (Table 1, Figure 8). In contrast, all the eIF4Es are downregulated in CML543. The locations of *eIF4E1_3* and *eIF4E2_3* on chromosome 3 could be linked to previously identified QTL associated with MLN, mapped on CML543 [8,61]. A recent CRISPR-Cas9 knockout study

targeting eIF4E1 and mutations in exon 4 of *eIF4E2* conferred resistance to SCMV, suggesting that the observed downregulation of *eIF4E* in CML543 likely disrupts SCMV growth and replication [17]. Of note, the upregulation of some of eIF4E in KS23-6 further confirms that KS23-6 does not confer resistance to SCMV. Although the differential expression of *eIF(iso)4G_2* (Zm00001d003147) is not indicated in CML543, the expression pattern of the gene shows an expression difference in SSA non-adapted lines (KS23-6), where it is downregulated, and the CMLs, where it is upregulated. The expression of 4Gs in the CML543 is only notable in *eIF4G_2* (Zm00001d006573), *eIF4G_7* (Zm00001d021741), and *eIF4G_10* (Zm00001d025979), where they are downregulated. *eIF4G_2* and *eIF4G_7* show definitive differences in CML543 and KS23-6 with the other CMLs, portraying a clear difference in phenotypic response. This regulation pattern indicates a response of eIF4 to MLN. Given this difference in expression, it would be interesting to examine these genes further as candidate genes for MLN resistance. Understanding the significance of these expression patterns can provide more insight into their relevance to response to viral infection and can be harnessed for managing MLN.

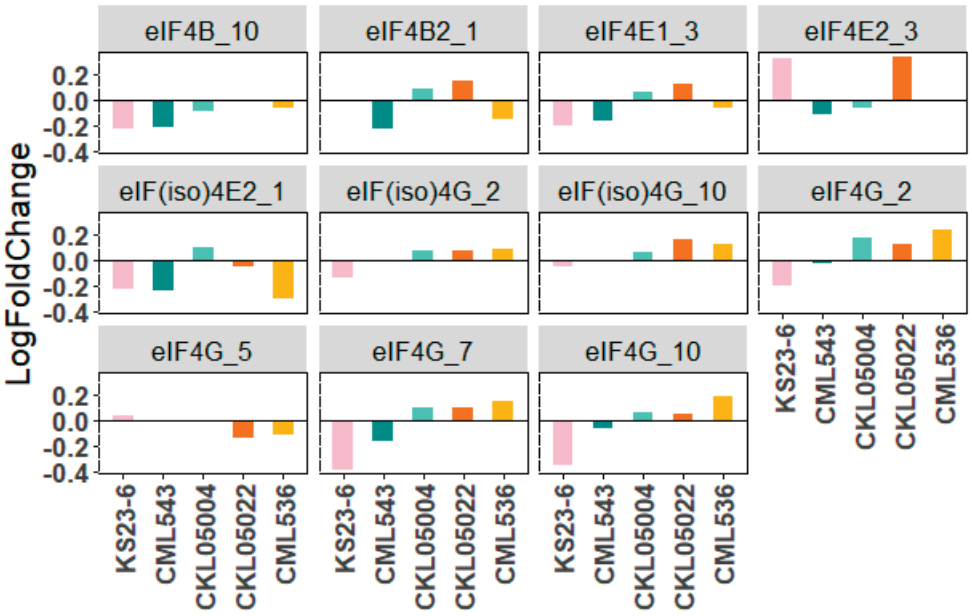

Figure 8. Differential expression of eukaryotic translation initiation factors (eIF4) across the genotypes based on the log2foldchange. The *x*-axis represents the genotypes. The *Y*-axis represents the expression value of each gene within a genotype based on the log2foldchange. (The numbers after the underscores indicate the chromosome locations).

In conclusion, this study demonstrated differences in the responses of the maize inbred lines to MLN, with distinct gene expression patterns observed among the most resistant, moderately resistant, and most susceptible lines. This is incongruent with the established difference in their phenotypic response to MLN, especially between KS23-6 and the CMLs. Our results highlight that KS23-6 resistance is related to MCMV, while the CMLs exhibit resistance to SCMV. This transcriptome analysis also showed the importance of expanding the expression study to more than one line with varying phenotypic responses. The contrast analysis showed that there is a clear transcriptome difference within the CMLs and across them when compared to KS23-6. Gene Ontology and KEGG analysis revealed specific pathways that play a role in stress responses and helped to identify candidate genes that may play an important role in the management of MLN. We identified gene models likely associated with resistance to either MCMV or SCMV, providing insights into the genetic basis of MLN tolerance. Notably, we identified key candidate genes for further

investigation, including those involved in innate immune responses (e.g., cell receptor recognition), redox signaling pathways (e.g., thioredoxin, peroxidase, and GST), RNA interference (e.g., *Argonaute, Dicer-like*, and *RNA-binding proteins)*, *WRKY* transcription factors, *ubiquitin E3-ligases*, and *translation initiation factors 4E* and *4G*. Relating these genes to known MLN-resistant QTL can fast-track the process of validating the differentially expressed genes and selecting target genes. Experimental studies on the genes, either through gene knockout of downregulated genes or overexpression, can be employed to understand gene function in response to MLN. Therefore, these genes present promising targets for advancing MLN tolerance through mutational genetics, gene editing, and other biotechnological approaches. While this experiment offers valuable insights into the maize transcriptome at 72 h post-inoculation (hpi), incorporating additional time points would provide a broader understanding of transcriptome changes as viral load increases. Expanding the time points could identify genes activated early versus those expressed later, offering a clearer understanding of genes that positively or negatively influence the MLN viruses.

3. Materials and Methods

3.1. Plant Materials and Stress Treatment

Four MLN-susceptible CIMMYT lines (CMLs), CML536, CML543, CKL05022, and CKL05004, and one Thailand genotype, KS23-6, were used in this experiment. KS23-6 is known for resistance to MCMV with QTL mapped on chromosome 6, while the CMLs are known for their potential tolerance to SCMV. All the susceptible lines are a set of elite lines used in breeding programs of CIMMYT for hybrid production for SSA. The design followed a randomized block design, where two conditions and four biological replicates were used. All materials were planted in greenhouses, and leaf tissue was sampled twice for RNA extraction. The first tissue sampling was performed immediately before MLN inoculation. Inoculation was then performed on the 3rd and 4th leaves. The second sampling was performed 72 h after inoculation; leaf punches were taken from the 5th and 6th leaves.

MLN inoculum was prepared by combining MCMV and SCMV inoculums. The maize chlorotic mottle virus (MCMV) strain used in this study was isolated by Corteva Agriscience in Hawaii. The sugarcane mosaic virus (SCMV) strain used originated in Ohio and was sourced from USDA at Wooster, OH. MCMV is maintained by serial inoculation in growth chambers, while SCMV is stored in infected tissues at −80 °C at Corteva Agriscience, Johnston, Iowa. The detailed protocol for MLN artificial inoculation has been described in previous studies [9,11,61]. Inoculum was prepared by grinding SCMV and MCMV-infected leaf tissue in 1x phosphate-buffered saline with Tween (PBST) buffer to create a leaf sap. In a Petri dish, MLN inoculum was established by combining/mixing the MCMV and SCMV leaf saps in a ratio of 1:4, respectively. Carborundum was added to facilitate leaf wounding. Plant infection was initiated by gently rubbing the leaf with the sap/carborundum solution.

3.2. RNA Sequencing and Data Analysis

To prepare samples for RNA extraction, 8 ¼" leaf punches were taken, frozen immediately on dry ice, and stored at −80 °C before extraction. According to the manufacturer's protocol, total RNA was isolated from ground-frozen tissue with RNeasy (Qiagen®, Germantown, MD, USA). Total RNA was then analyzed for quality and quantity using an Agilent 5300 Bioanalyzer (Agilent Technologies®, Santa Clara, CA, USA).

A total of 48 RNA samples (6 genotypes × 4 bio reps × 2 time points) were submitted for library preparation and sequencing at Corteva Agriscience. According to the vendor's protocol, sequencing libraries were prepared using the Illumina TruSeq stranded mRNA library kit (Illumina Inc., San Diego, CA, USA). The detailed method for RNA-Seq library

preparation is described in a study method article [62]. Briefly, mRNAs were isolated via attachment to oligo(dT) beads, fragmented, and reverse-transcribed into cDNA by random hexamer primers with Superscript II reverse transcriptase (Life Technologies, Carlsbad, CA, USA). The resulting cDNAs were end-repaired, 3′-A-tailed, and ligated with Illumina-indexed TruSeq adapters. Ligated cDNA fragments were PCR-amplified with Illumina TruSeq primers, purified with AmpureXP Beads (Beckman Coulter Genomics, Brea, CA, USA), and checked for quality and quantity with the Agilent TapeStation 4200 system with D1000 ScreenTape. Libraries were combined into one sequencing pool and normalized to 2 nM. The pool was denatured according to Illumina sequencing protocols, hybridized, and clustered on an appropriate Illumina flow cell targeting 20M reads per sample. Samples were sequenced using single-read 50bp reads. After sequencing, data were trimmed for quality with a minimum threshold of Q13, and the resulting sequences were split by index identifier.

3.3. RNA-Seq Transcript Quantification and Identification of Differentially Expressed Genes (DEGs)

Analysis of the raw reads followed an established RNA-Seq analysis pipeline using the Linux command line. The pipeline begins with preprocessing, where FastQC is used to evaluate the quality of the raw reads (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/). The FastQC filtering identified adapter sequences and assessed the percentage of unknown base sequences and low-quality bases. Trimming of the adapters and sequence repeats was performed using FastX. The downstream analysis was based on quality-controlled clean reads. The publicly available *Zea mays* genome assembly B73 RefGen_v4 (https://www.ncbi.nlm.nih.gov/) was used as the reference genome. Quantification of read abundance was performed using the Salmon tool by aligning the reads to the indexed B73 reference transcriptome [63]. Count data of the reads per gene were generated as transcripts per million (TPM). The count data were subjected to differential expression analysis using DESeq2, a Bioconductor package running on R, version 4.2.1 to identify differentially expressed genes [64]. Differential expression analysis was performed using the general linear model (GLM) implemented in DESeq2. Before fitting and running the model, DESeq2 estimates size factors to normalize the gene counts to account for technical variation, i.e., library size and composition. The dispersion parameter for each gene is also estimated to account for variability between biological replicates using the maximum linear estimation (MLE). The model uses a logarithm link between relative gene abundance and the design matrix:

$$Log_2(q_{ij}) = \sum_r x_{jr}\beta_{ir}$$

DESeq2 integrates the design matrix elements x_{jr}, of whether sample j is treated, and returns a coefficient β_{ir}, indicating the overall expression strength of gene i as a log2foldchange between the treatment and control. q_{ij} is the expected true concentration of gene fragments in a sample [64,65].

The Wald test was used for significance testing, and p-values that passed independent filtering were adjusted for multiple testing using the Benjamini and Hochberg False Discovery Rate (FDR) procedure [64]. Differential expression results were filtered to select differentially expressed genes (DEGs), and only genes with a 2-fold change in expression level and $\leq$5% FDR were selected. Clustering tools, heatmaps, and principal component analysis (PCA) embedded in the DESeq2 package were used to generate PCA plots and heatmaps based on transformed and normalized gene expression counts. An online Venn generator (https://bioinformatics.psb.ugent.be/webtools/Venn/) was used to identify common and unique DEGs across genotypes. K-mean clusters were derived via normalized gene counts of the DEGs using the K-means clustering tool provided by Morpheus (https://software.broadinstitute.org/morpheus/).

3.4. DEG Functional Annotation and Enrichment Analysis

To perform the functional annotation of the genes, Gene Ontology, KEGG, and manual annotation were performed using public databases. Using AgriGO [66], the DEGs were subjected to Gene Ontology (GO) analysis for biological process (BP), molecular function (MF), and cellular composition (CC). Maize AGPv4 provided by the software was used as the annotation reference. Fisher's exact test with multi-test adjustment was used, and GO terms were selected based on $p_{adj} < 0.05$. Further, a Kyoto Encyclopedia of Genes and Genomes (KEGG) analysis was performed in ShinyGO (https://doi.org/10.1093/bioinformatics/btz931) to identify enriched pathways. To identify gene annotations associated with the IDs of the DEGs, a maize gene ID list (based on B73 v4 gene models) with functional annotations sourced from Ensemble Plants biomart (https://plants.ensembl.org) was used to match the DEGs' gene IDs to protein functions.

Supplementary Materials: The following supporting information can be downloaded at https://www.mdpi.com/article/10.3390/plants14020295/s1, Table S1: Details of data generated by sequencing across all the biological replicates; Table S2: Differentially expressed genes (DEGs) of each line in control vs. treatment analysis; Table S3: Differentially expressed genes (DEGs) of each genotype vs. genotype comparison; Table S4: GO terms associated with differentially expressed genes in each line from control vs. treatment analysis; Table S5: GO terms associated with differentially expressed genes in genotype vs. genotype comparisons; Table S6: KEGG pathways associated with differentially expressed genes in each line from control vs. treatment analysis; Table S7: KEGG pathways associated with differentially expressed genes of genotype vs. genotype comparisons; Table S8: Differential expression of eukaryotic translation initiation factors (eIFs) across genotypes based on log2foldchange; Figure S1: Heatmap clustering of all biological replicates based on their normalized gene expression counts across all six lines. Red indicates a higher similarity index between biological reps, while blue indicates a lower similarity index, where 1 represents 4 of the null replicates and 2 represents 4 of the treated replicates.

Author Contributions: A.M.: investigation, formal analysis, visualization, writing—original draft, and writing—review and editing. Z.W.: methodology and formal analysis. G.P.: data curation, formal analysis, software, and writing—review and editing. M.S.O.: project administration, resources, supervision, methodology, and writing—review and editing. T.L.: supervision, resources, and writing—review and editing. K.S.D.: conceptualization, project administration, funding acquisition, supervision, resources, visualization, and writing—review and editing. M.J.: project administration, supervision, resources, methodology, investigation, data curation, software, and writing—review and editing. All authors have read and agreed to the published version of the manuscript.

Funding: Supported by a grant from the Bill & Melinda Gates Foundation (OPP1193833 a n d INV-007833) to the International Center for Maize and Wheat Improvement (CIMMYT).

Data Availability Statement: The authors confirm that the raw data supporting the findings in this study will be accessible once the permissions for public sharing have been obtained. Meanwhile, the raw outputs from the read quantification and DESeq2 analysis are available at the following GitHub link: https://github.com/AnnM-511/Transcriptome_profiling_6maize_inbred_lines.

Acknowledgments: The authors thank Erik Ohlson (USDA—Ohio State University) for providing the SCMV inoculum used in this study.

Conflicts of Interest: Ann Murithi, Gayathri Panangipalli and Mark Jung are employed by Corteva Agriscience, Zhengyu Wei is employed by Keygene, Inc. and Michael S. Olsen is empoloyed by Bayer. The authors declare that the research was conducted without any commercial or financial relationships that could be construed as a potential conflict of interest.

References

1. Boddupalli, P.; Suresh, L.M.; Mwatuni, F.; Beyene, Y.; Makumbi, D.; Gowda, M.; Olsen, M.S.; Hodson, D.; Worku, M.; Mezzalama, M. Maize lethal necrosis (MLN): Efforts toward containing the spread and impact of a devastating transboundary disease in sub-Saharan Africa. *Virus Res.* **2020**, *282*, 197943. [CrossRef]
2. Mwatuni, F.M.N.; Nyende, A.B.; Njuguna, J.; Xiong, Z.; Machuka, E.; Francesca, S. Occurrence, genetic diversity, and recombination of maize lethal necrosis disease-causing viruses in Kenya. *Virus Res.* **2020**, *286*, 198081. [CrossRef] [PubMed]
3. Redinbaugh, M.G.; Stewart, L.R. Maize lethal necrosis: An emerging, synergistic viral disease. *Annu. Rev. Virol.* **2018**, *5*, 301–322. [CrossRef] [PubMed]
4. Mahuku, G.; Lockhart, B.E.; Wanjala, B.; Jones, M.W.; Kimunye, J.N.; Stewart, L.R.; Cassone, B.J.; Sevgan, S.; Nyasani, J.O.; Kusia, E.; et al. Maize Lethal Necrosis (MLN), an Emerging Threat to Maize-Based Food Security in Sub-Saharan Africa. *Phytopathology* **2015**, *105*, 956–965. [CrossRef] [PubMed]
5. Adams, I.P.; Miano, D.W.; Kinyua, Z.M.; Wangai, A.; Kimani, E.; Phiri, N.; Reeder, R.; Harju, V.; Glover, R.; Hany, U.; et al. Use of next-generation sequencing for the identification and characterization of Maize chlorotic mottle virus and Sugarcane mosaic virus causing maize lethal necrosis in Kenya. *Plant Pathol.* **2013**, *62*, 741–749. [CrossRef]
6. Wangai, A.W.; Redinbaugh, M.G.; Kinyua, Z.M.; Miano, D.W.; Leley, P.K.; Kasina, M.; Mahuku, G.; Scheets, K.; Jeffers, D. First Report of Maize chlorotic mottle virus and Maize Lethal Necrosis in Kenya. *APS Publ.* **2012**, *96*, 1582. [CrossRef]
7. Semagn, K.; Beyene, Y.; Babu, R.; Nair, S.; Gowda, M.; Das, B.; Tarekegne, A.; Mugo, S.; Mahuku, G.; Worku, M.; et al. Quantitative Trait Loci Mapping and Molecular Breeding for Developing Stress Resilient Maize for Sub-Saharan Africa. *Crop Sci.* **2015**, *55*, 1449–1459. [CrossRef]
8. Gowda, M.; Beyene, Y.; Makumbi, D.; Semagn, K.; Olsen, M.S.; Bright, J.M.; Das, B.; Mugo, S.; Suresh, L.M.; Prasanna, B.M. Discovery and validation of genomic regions associated with resistance to maize lethal necrosis in four biparental populations. *Mol. Breed.* **2018**, *38*, 66. [CrossRef]
9. Gowda, M.; Das, B.; Makumbi, D.; Babu, R.; Semagn, K.; Mahuku, G.; Olsen, M.S.; Bright, J.M.; Beyene, Y.; Prasanna, B.M. Genome-wide association and genomic prediction of resistance to maize lethal necrosis disease in tropical maize germplasm. *Theor. Appl. Genet.* **2015**, *128*, 1957–1968. [CrossRef] [PubMed]
10. Sitonik, C.A.; Suresh, L.M.; Beyene, Y.; Olsen, M.S.; Makumbi, D.; Oliver, K.; Das, B.; Bright, J.M.; Mugo, S.; Crossa, J.; et al. Genetic architecture of maize chlorotic mottle virus and maize lethal necrosis through GWAS, linkage analysis and genomic prediction in tropical maize germplasm. *Theor. Appl. Genet.* **2019**, *132*, 2381–2399. [CrossRef] [PubMed]
11. Murithi, A.; Olsen, M.S.; Kwemoi, D.B.; Veronica, O.; Ertiro, B.T.; Suresh, L.M.; Beyene, Y.; Das, B.; Prasanna, B.M.; Gowda, M. Discovery and validation of a recessively inherited major-effect QTL conferring resistance to Maize Lethal Necrosis (MLN) disease. *Front. Genet.* **2021**, *12*, 2269. [CrossRef] [PubMed]
12. Brewbaker, J.L. Registration of Nine Maize Populations Resistant to Tropical Diseases. *J. Plant Regist.* **2009**, *3*, 10–13. [CrossRef]
13. Jampatong, S.; Thung-Ngean, M.; Balla, C.; Boonrumpun, P.; Mekarun, A.; Jompuk, C.; Kaveeta, R. Evaluation of improved maize populations and their diallel crosses for yield. *Agric. Nat. Resour.* **2010**, *44*, 523–528.
14. Leng, P.; Ji, Q.; Asp, T.; Frei, U.K.; Ingvardsen, C.; Xing, Y.; Studer, B.; Redinbaugh, M.G.; Jones, M.W.; Gajjar, P.; et al. Auxin binding protein 1 reinforces resistance to Sugarcane mosaic virus in maize. *Mol. Plant* **2017**, *10*, 1357–1360. [CrossRef]
15. Leng, P.; Ji, Q.; Tao, Y.; Ibrahim, R.; Pan, G.; Xu, M.; Lübberstedt, T. Characterization of Sugarcane Mosaic Virus Scmv1 and Scmv2 Resistance Regions by Regional Association Analysis in Maize. *PLoS ONE* **2015**, *10*, e0140617. [CrossRef] [PubMed]
16. Tao, Y.; Jiang, L.; Liu, Q.; Zhang, Y.; Zhang, R.; Ingvardsen, C.R.; Frei, U.K.; Wang, L.; Lai, J.; Lübberstedt, T.; et al. Combined linkage and association mapping reveals candidates for scmv1, a major locus involved in resistance to sugarcane mosaic virus (SCMV) in maize. *BMC Plant Biol.* **2013**, *13*, 162. [CrossRef] [PubMed]
17. Wen, Z.; Lu, F.; Jung, M.; Humbert, S.; Marshall, L.; Hastings, C.; Wu, E.; Jones, T.; Pacheco, M.; Martinez, I.; et al. Edited eukaryotic translation initiation factors confer resistance against maize lethal necrosis. *Plant Biotechnol. J.* **2024**, *22*, 3523–3535. [CrossRef] [PubMed]
18. Chen, Z.; Zhao, J.; Song, J.; Han, S.; Du, Y.; Qiao, Y.; Liu, Z.; Qiao, J.; Li, W.; Li, J. Influence of graphene on the multiple metabolic pathways of *Zea mays* roots based on transcriptome analysis. *PLoS ONE* **2021**, *16*, e0244856. [CrossRef] [PubMed]
19. Stelpflug, S.C.; Sekhon, R.S.; Vaillancourt, B.; Hirsch, C.N.; Buell, C.R.; Leon, N.; Kaeppler, S.M. An Expanded Maize Gene Expression Atlas based on RNA Sequencing and its Use to Explore Root Development. *Plant Genome* **2016**, *9*, plantgenome2015. [CrossRef] [PubMed]
20. Ben Ali, S.-E.; Draxler, A.; Poelzl, D.; Agapito-Tenfen, S.; Hochegger, R.; Haslberger, A.G.; Brandes, C. Analysis of transcriptomic differences between NK603 maize and near-isogenic varieties using RNA sequencing and RT-qPCR. *Environ. Sci. Eur.* **2020**, *32*, 132. [CrossRef]
21. Lambarey, H.; Moola, N.; Veenstra, A.; Murray, S.; Suhail, R.M. Transcriptomic Analysis of a Susceptible African Maize Line to Fusarium verticillioides Infection. *Plants* **2020**, *9*, 1112. [CrossRef] [PubMed]

22. Zhang, X.; Wang, Y.; Yan, Y.; Peng, H.; Long, Y.; Zhang, Y.; Jiang, Z.; Liu, P.; Zou, C.; Peng, H.; et al. Transcriptome sequencing analysis of maize embryonic callus during early redifferentiation. *BMC Genom.* **2019**, *20*, 159. [CrossRef] [PubMed]
23. Barua, D.; Mishra, A.; Kirti, P.B.; Barah, P. Identifying Signal-Crosstalk Mechanism in Maize Plants during Combined Salinity and Boron Stress Using Integrative Systems Biology Approaches. *BioMed Res. Int.* **2022**, *2022*, 1027288. [CrossRef] [PubMed]
24. Liu, Q.; Liu, H.; Gong, Y.; Tao, Y.; Jiang, L.; Zuo, W.; Yang, Q.; Ye, J.; Lai, J.; Wu, J.; et al. An Atypical Thioredoxin Imparts Early Resistance to Sugarcane Mosaic Virus in Maize. *Mol. Plant* **2017**, *10*, 483–497. [CrossRef] [PubMed]
25. Shi, C.; Ingvardsen, C.; Thümmler, F.; Melchinger, A.E.; Wenzel, G.; Lübberstedt, T. Identification by suppression subtractive hybridization of genes that are differentially expressed between near-isogenic maize lines in association with sugarcane mosaic virus resistance. *Mol. Genet. Genom.* **2005**, *273*, 450–461. [CrossRef] [PubMed]
26. Shi, C.; Thümmler, F.; Melchinger, A.E.; Wenzel, G.; Lübberstedt, T. Association between SCMV Resistance and Macroarray-based Expression Patterns in Maize Inbreds. *Mol. Breed.* **2005**, *16*, 173–184. [CrossRef]
27. Shi, C.; Thümmler, F.; Albrecht, E.M.; Wenzel, G.; Lübberstedt, T. Comparison of transcript profiles between near-isogenic maize lines in association with SCMV resistance based on unigene-microarrays. *Plant Sci.* **2006**, *170*, 159–169. [CrossRef]
28. Dang, M.; Parkash, J.; Mehra, R.; Sharma, N.; Singh, B.; Raigond, P.; Joshi, A.; Chopra, S.; Singh, B.P. Proteomic Changes during MCMV Infection Revealed by iTRAQ Quantitative Proteomic Analysis in Maize. *Int. J. Mol. Sci.* **2019**, *21*, 35. [CrossRef] [PubMed]
29. Wang, Z.; Gerstein, M.; Snyder, M. RNA-Seq: A revolutionary tool for transcriptomics. *Nat. Rev. Genet.* **2009**, *10*, 57–63. [CrossRef] [PubMed]
30. Zhao, S.; Zhang, B.; Zhang, Y.; Gordon, W.; Du, S.; Paradis, T.; Vincent, M.; von Schack, D. Bioinformatics for RNA-Seq Data Analysis. In *Bioinformatics*; Abdurakhmonov, I.Y., Ed.; IntechOpen: Rijeka, Croatia, 2016.
31. Kaundal, R.; Parkash, J.; Mehra, R.; Sharma, N.; Singh, B.; Raigond, P.; Joshi, A.; Chopra, S.; Singh, B.P. Transcriptional profiling of two contrasting genotypes uncovers molecular mechanisms underlying salt tolerance in alfalfa. *Sci. Rep.* **2021**, *11*, 5210. [CrossRef] [PubMed]
32. Calil, I.P.; Fontes, E.P. Plant immunity against viruses: Antiviral immune receptors in focus. *Ann. Bot.* **2017**, *119*, 711–723. [CrossRef] [PubMed]
33. Garcia-Ruiz, H. Susceptibility Genes to Plant Viruses. *Viruses* **2018**, *10*, 484. [CrossRef] [PubMed]
34. Biswal, A.K.; Alakonya, A.E.; Mottaleb K., A.; Hearne, S.J.; Sonder, K.; Molnar, T.L.; Jones, A.M.; Pixley, K.; Boddupalli, P. Maize Lethal Necrosis disease: Review of molecular and genetic resistance mechanisms, socio-economic impacts, and mitigation strategies in sub-Saharan Africa. *BMC Plant Biol.* **2022**, *22*, 542. [CrossRef]
35. Yuan, X.; Wang, H.; Cai, J.; Li, D.; Song, F. NAC transcription factors in plant immunity. *Phytopathol. Res.* **2019**, *1*, 3. [CrossRef]
36. Mestre, P.; Baulcombe, D.C. Elicitor-mediated oligomerization of the tobacco N disease resistance protein. *Plant Cell* **2006**, *18*, 491–501. [CrossRef]
37. Ross, B.T.; Zidack, N.H.; Flenniken, M.L. Extreme Resistance to Viruses in Potato and Soybean. *Front. Plant Sci.* **2021**, *12*, 658981. [CrossRef]
38. Slavokhotova, A.; Korostyleva, T.; Shelenkov, A.; Pukhalskiy, V.; Korottseva, I.; Slezina, M.; Istomina, E.; Odintsova, T. Transcriptomic analysis of genes involved in plant defense response to the cucumber green mottle mosaic virus infection. *Life* **2021**, *11*, 1064. [CrossRef]
39. Akbar, S.; Wei, Y.; Khan, M.T.; Qin, L.; Powell, C.A.; Chen, B.; Zhang, M. Gene expression profiling of reactive oxygen species (ROS) and antioxidant defense system following Sugarcane mosaic virus (SCMV) infection. *BMC Plant Biol.* **2020**, *20*, 532. [CrossRef]
40. Chen, L.; Chen, L.; Zhang, L.; Li, D.; Wang, F.; Yu, D. WRKY8 transcription factor functions in the TMV-cg defense response by mediating both abscisic acid and ethylene signaling in Arabidopsis. *Proc. Natl. Acad. Sci. USA* **2013**, *110*, E1963–E1971. [CrossRef]
41. Liu, Y.; Schiff, M.; Dinesh-Kumar, S.P. Involvement of MEK1 MAPKK, NTF6 MAPK, WRKY/MYB transcription factors, COI1 and CTR1 in N-mediated resistance to tobacco mosaic virus. *Plant J.* **2004**, *38*, 800–809. [CrossRef]
42. Hernández, J.A.; Gulliner, G.; Clemente-Moreno, M.J.; Künstler, A.; Juhász, C.; Díaz-Vivancos, P.; Király, L. Oxidative stress and antioxidative responses in plant–virus interactions. *Physiol. Mol. Plant Pathol.* **2016**, *94*, 134–148. [CrossRef]
43. Xu, Y.; Zhang, S.; Zhang, M.; Jiao, G.; Jiang, T. The role of reactive oxygen species in plant-virus interactions. *Plant Cell Rep.* **2024**, *43*, 197. [CrossRef] [PubMed]
44. Gullner, G.; Gullner, G.; Komives, T.; Király, L.; Schröder, P. Glutathione S-Transferase Enzymes in Plant-Pathogen Interactions. *Front. Plant Sci.* **2018**, *9*, 1836. [CrossRef] [PubMed]
45. Alcaide-Loridan, C.; Jupin, I. Ubiquitin and Plant Viruses, Let's Play Together! *Plant Physiol.* **2012**, *160*, 72–82. [CrossRef]
46. Correa, R.L.; Bruckner, F.P.; De Souza C., R.; Alfenas-Zerbini, P. The role of F-box proteins during viral infection. *Int. J. Mol. Sci.* **2013**, *14*, 4030–4049. [CrossRef]
47. Bateman, A. The SGS3 protein involved in PTGS finds a family. *BMC Bioinform.* **2002**, *3*, 21. [CrossRef]
48. Cheng, X.; Wang, A. The potyvirus silencing suppressor protein VPg mediates degradation of SGS3 via ubiquitination and autophagy pathways. *J. Virol.* **2017**, *91*, e01478-16. [CrossRef] [PubMed]

49. Zhang, C.; Wu, Z.; Li, Y.; Wu, J. Biogenesis, function, and applications of virus-derived small RNAs in plants. *Front. Microbiol.* **2015**, *6*, 1237. [CrossRef] [PubMed]
50. Csorba, T.; Kontra, L.; Burgyán, J. Viral silencing suppressors: Tools forged to fine-tune host-pathogen coexistence. *Virology* **2015**, *479–480*, 85–103. [CrossRef]
51. Wieczorek, P.; Obrępalska-Stęplowska, A. Suppress to Survive-Implication of Plant Viruses in PTGS. *Plant Mol. Biol.* **2015**, *33*, 335–346. [CrossRef]
52. Fan, Y.; Sanyal, S.; Bruzzone, R. Breaking Bad: How Viruses Subvert the Cell Cycle. *Front. Cell. Infect. Microbiol.* **2018**, *8*, 396. [CrossRef] [PubMed]
53. Frick, D.N.; Lam, A.M.I. Understanding Helicases as a Means of Virus Control. *Curr. Pharm. Des.* **2013**, *12*, 1315–1338. [CrossRef] [PubMed]
54. Babu, G.S.; Gohil, D.S.; Choudhury, S.R. Genome-wide identification, evolutionary and expression analysis of the cyclin-dependent kinase gene family in peanut. *BMC Plant Biol.* **2023**, *23*, 43.
55. Cui, X.; Fan, B.; Scholz, J.; Chen, Z. Roles of Arabidopsis Cyclin-Dependent Kinase C Complexes in Cauliflower Mosaic Virus Infection, Plant Growth, and Development. *Plant Cell* **2007**, *19*, 1388–1402. [CrossRef] [PubMed]
56. Stern-Ginossar, N.; Thompson, S.R.; Mathews, M.B.; Mohr, I. Translational Control in Virus-Infected Cells. *Cold Spring Harb. Perspect. Biol.* **2019**, *11*, a033001. [CrossRef] [PubMed]
57. Sanfaçon, H. Plant Translation Factors and Virus Resistance. *Viruses* **2015**, *7*, 3392–3419. [CrossRef] [PubMed]
58. Schmitt-Keichinger, C. Manipulating cellular factors to combat viruses: A case study from the plant eukaryotic translation initiation factors eIF4. *Front. Microbiol.* **2019**, *10*, 17. [CrossRef]
59. Dutt, S.; Parkash, J.; Mehra, R.; Sharma, N.; Singh, B.; Raigond, P.; Joshi, A.; Chopra, S.; Singh, B.P. Translation initiation in plants: Roles and implications beyond protein synthesis. *Biol. Plant.* **2015**, *59*, 401–412. [CrossRef]
60. Merchante, C.; Stepanova, A.N.; Alonso, J.M. Translation regulation in plants: An interesting past, an exciting present and a promising future. *Plant J.* **2017**, *90*, 628–653. [CrossRef] [PubMed]
61. Awata, L.A.O.; Beyene, Y.; Gowda, M.; L. M., S.; Jumbo, M.B.; Tongoona, P.; Danquah, E.; Ifie, B.E.; Marchelo-Dragga, P.W.; Olsen, M.; et al. Genetic Analysis of QTL for Resistance to Maize Lethal Necrosis in Multiple Mapping Populations. *Genes* **2019**, *11*, 32. [CrossRef] [PubMed]
62. Kumar, R.; Ichihashi, Y.; Kimura, S.; Chitwood, D.H.; Headland, L.R.; Peng, J.; Maloof, J.N.; Sinha, N.R. A High-Throughput Method for Illumina RNA-Seq Library Preparation. *Front. Plant Sci.* **2012**, *3*, 28988. [CrossRef] [PubMed]
63. Patro, R.; Duggal, G.; Love, M.I.; Irizarry, R.A.; Kingsford, C. Salmon provides fast and bias-aware quantification of transcript expression. *Nat. Methods* **2017**, *14*, 417–419. [CrossRef] [PubMed]
64. Love, M.I.; Huber, W.; Anders, S. Moderated estimation of fold change and dispersion for RNA-seq data with DESeq2. *Genome Biol.* **2014**, *15*, 550. [CrossRef] [PubMed]
65. Li, D.; Dye, T.D.; Goniewicz, M.L.; Rahman, I.; Xie, Z. An evaluation of RNA-seq differential analysis methods. *PLoS ONE* **2022**, *17*, e0264246. [CrossRef] [PubMed]
66. Tian, T.; Liu, Y.; Yan, H.; You, Q.; Yi, X.; Du, Z.; Xu, W.; Su, Z. agriGO v2. 0: A GO analysis toolkit for the agricultural community, 2017 update. *Nucleic Acids Res.* **2017**, *45*, W122–W129. [CrossRef] [PubMed]

plants

MDPI

Article

Resistance Spectrum Analysis and Breeding Utilization of Rice Blast Resistance Gene *Pigm-1*

Yidan Jin [1,2], Niqing He [1], Zhaoping Cheng [1], Shaojun Lin [1], Fenghuang Huang [1], Wenxiao Wang [1,2], Qingshun Q. Li [1,3] and Dewei Yang [1,*]

[1] Institute of Rice, Fujian Academy of Agricultural Sciences, Fuzhou 350018, China; jyd0121@163.com (Y.J.); heniqing430@163.com (N.H.); chengzhaoping@126.com (Z.C.); shaojunlin1993@163.com (S.L.); hhf13305022536@163.com (F.H.); wenxiaowang98@163.com (W.W.); qqli@westernu.edu (Q.Q.L.)
[2] College of Agriculture, Fujian Agriculture and Forestry University, Fuzhou 350002, China
[3] Biomedical Sciences, College of Dental Medicine, Western University of Health Sciences, Pomona, CA 91766, USA
* Correspondence: dewei-y@163.com; Tel.: +86-591-88030184; Fax: +86-591-88030184

Abstract: Rice blast is one of the most important diseases of rice, causing significant economic losses to agricultural production. A new gene, *Pigm-1*, which is allelic to *Pigm*, was cloned from Shuangkang 77009 using map based cloning. However, it is unclear whether there is a difference in the resistance spectrum between *Pigm* and *Pigm-1*. In this study, using 195 rice blast isolates collected from different areas of the Fujian Province, the *Pigm-1* and *Pigm* single gene lines were inoculated to test their resistance. There was only one blast fungus JL-37 that showed a differential response in *Pigm* and *Pigm-1* single gene lines, while the remaining 194 showed no difference. To further explore the application range of *Pigm-1*, the resistant rice R20-4 containing *Pigm-1* was used as the donor, and a sensitive sticky rice S19-118 was used as the receptor. The hybrid F_1 was first backcrossed with S19-118 using a molecular marker-assisted selection breeding method, and a strain containing the *Pigm-1* gene was selected to continue to backcross with S19-118 until BC_3F_1. A new blast resistance rice material, Xiannuo 23, containing *Pigm-1* was developed and confirmed by laboratory and field tests. This material can be broadly used for the future breeding of rice blast resistant cultivars to reduce the loss of rice production.

Keywords: rice; blast disease; *Pigm-1*; resistance spectrum; breeding utilization

Academic Editor: Won Kyong Cho

Received: 10 December 2024
Revised: 30 January 2025
Accepted: 7 February 2025
Published: 10 February 2025

Citation: Jin, Y.; He, N.; Cheng, Z.; Lin, S.; Huang, F.; Wang, W.; Li, Q.Q.; Yang, D. Resistance Spectrum Analysis and Breeding Utilization of Rice Blast Resistance Gene *Pigm-1*. *Plants* **2025**, *14*, 535. https://doi.org/10.3390/plants14040535

1. Introduction

Rice (*Oryza sativa*) is one of the most important food crops in the world, with about half of the world's population using it as their staple food [1]. Rice blast is a worldwide disease caused by *Magnaporthe oryzae* [2], which can cause at least 10% yield loss of rice every year. In severe cases, the yield reduction in rice can be as high as 40% to 50%, or even total loss [3,4]. At present, the main method of rice blast control is chemical application, but excessive use of chemical agents has caused serious damage to the environment [5,6]. Under normal circumstances, new varieties that are resistant to rice blast will gradually lose resistance after 3 to 5 years of field production, mainly due to the narrow spectrum of resistance genes carried by cultivars. The physiological variants of rice blast have strong variability, leading to rapid disease development under suitable environmental conditions [7]. Rice blast resistance breeding is still the most effective and equitable means to limit the losses in rice production [8]. Therefore, it is of great significance to explore more blast resistance genes and analyze their resistance spectra, so as to create new resistant cultivars for ensuring food security [9].

So far, over 100 rice blast resistance loci have been genetically characterized, and more than 50 blast resistance genes have been identified [5,8]. These genes are distributed on different chromosomes of rice. At present, several major genes such as *Pi2*, *Pi9* (*Pi9-type-3/4/5/6/9/10/11*), *Piz-t* (*Pizh*) and *Pigm* (*Pi50*) have been cloned at the *Piz* locus on chromosome 6 of rice [10–12]. Using molecular marker-assisted selection, the resistant rice varieties can be produced by introducing the broad-spectrum blast resistance genes into susceptible genotypes [13]. Different alleles of the blast resistance gene at one locus have different species-specific spectra [8]. For example, seven alleles with different resistance profiles were identified at *Pik* resistance sites [14,15]. Therefore, it is particularly important to explore and isolate more resistance genes and their alleles. Previously, our research team cloned a broad-spectrum blast resistance gene *Pigm-1* from the disease-resistant resource material Shuangkang 77009 by map based cloning. This gene is a new allele of *Pigm*, which is different from *PigmR* by 3 bases [16]. In recent years, through molecular breeding, researchers have created a series of new restorer line materials containing *Pigm* and its allele *Pigm-1* [16–18].

Pigm-1 and *Pigm* are alleles and differ from each other by only a few bases. It is not clear whether their resistance spectra are different. On the other hand, although we have used *Pigm-1* to create a series of new disease-resistant materials, there are few reports on the creation of new special rice materials containing *Pigm-1*. In this study, different blast fungi were isolated from Shanghang County, Jianyang County and Jiangle County, Fujian Province, and inoculated with *Pigm-1* and *Pigm* single gene lines in order to find the resistance spectrum of these resistance genes. At the same time, to further explore the application range of *Pigm-1*, this study created a new disease-resistant glutinous rice material containing *Pigm-1* through molecular breeding methods, and provided excellent disease-resistant new materials for specialty rice breeding.

2. Results

2.1. Creation of Single Gene Lines Containing Pigm-1 and Pigm

To obtain *Pigm-1* and *Pigm* single gene lines, the parent Shuangkang 77009 containing *Pigm-1* and the parent Digu B containing *Pigm* were used as the donors, and 9311 as the acceptor. The F_1 plants were generated from Shuangkang 77009 and Digu B as female and 9311 as male. The F_1 plants were backcrossed with 9311 to produce the BC_1F_1 generation. The BC_1F_1 plants were backcrossed to 9311 to produce BC_2F_1, and the BC_2F_1 plants were backcrossed to 9311 to produce BC_3F_1, and then further backcrossed all the way to BC_4F_1. At the same time, *Pigm-1* and the molecular marker corresponding to *Pigm* were detected in each backcross. From BC_4F_1 to BC_4F_5, molecular markers were tracked for detection in each generation and, finally, *Pigm-1* and *Pigm* single gene lines were obtained, named 9311 (*Pigm-1*) and 9311 (*Pigm*), respectively (Figure 1).

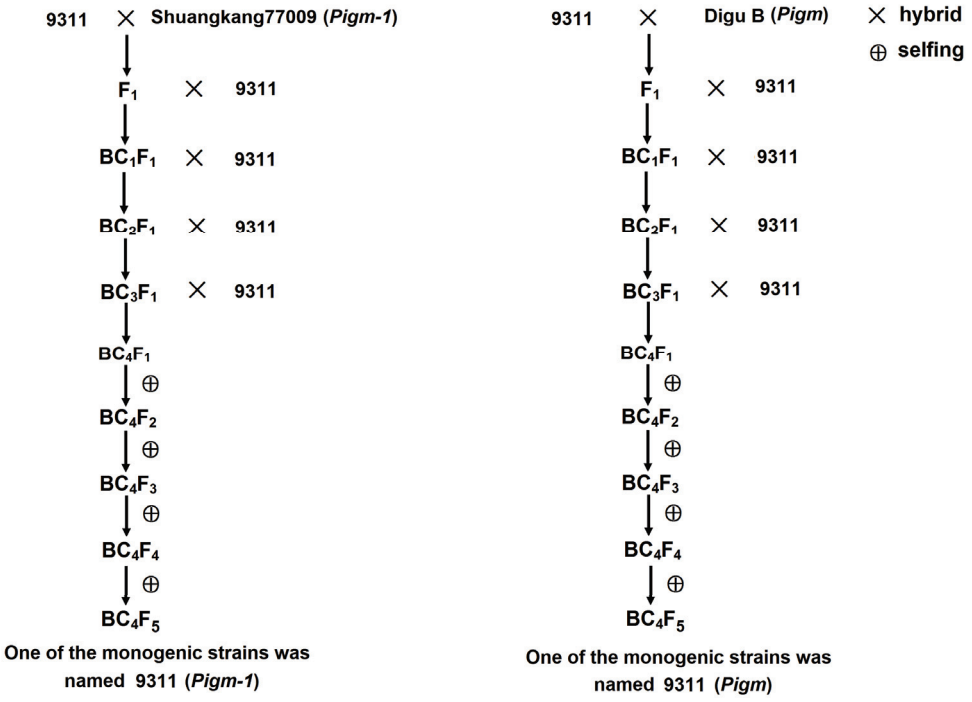

Figure 1. Construction roadmap of *Pigm-1* and *Pigm* monogenic lines. BC, backcross; F, familial.

2.2. Analysis of the Main Agronomic Traits of Pigm-1 and Pigm Single Gene Lines

To further analyze the changes in other agronomic traits of the improved line after the integration of blast resistance, we investigated the agronomic traits of 9311 (*Pigm-1*), 9311 (*Pigm*) and 9311 at maturity. The results showed that there was no difference between 9311 and 9311 (*Pigm*) in main agronomic traits, including plant height, panicle length, spikelets per panicle, seed setting rate, effective panicle number, grain length, grain width and 1000-grain weight. However, there was no difference in agronomic traits between 9311 and monogenic line 9311 (*Pigm-1*) except plant height, panicle length and spikelets per panicle (Table 1). These results indicate that the genetic background of single-gene line 9311 (*Pigm-1*) and 9311 (*Pigm*) basically restore the receptor parent 9311.

Table 1. Comparison of the main agronomical traits between 9311 (*Pigm-1*), 9311 (*Pigm*) and 9311.

Traits	9311 (*Pigm-1*)	9311 (*Pigm*)	9311
Plant height (cm)	115.4 ± 2.24 *	107.3 ± 2.14	106.8 ± 2.11
Panicle length (cm)	23.1 ± 0.34 *	20.7 ± 0.31	19.4 ± 0.30
Number of effective panicle	7.68 ± 1.06	7.91 ± 1.13	7.84 ± 1.19
Spikelets per panicle	197.6 ± 6.46 **	163.7 ± 5.96	160.8 ± 5.16
Seed setting rate (%)	91.06 ± 2.23	90.15 ± 1.92	88.85 ± 1.62
1000-grain weight (g)	32.13± 0.91	30.98± 0.81	31.16± 0.81
Grain length (mm)	9.75 ± 0.18	9.68 ± 0.21	9.71 ± 0.21
Grain width (mm)	3.26 ± 0.10	3.23 ± 0.10	3.24 ± 0.11

* and ** indicate the significance levels of the differences between 9311 (*Pigm-1*), 9311 (*Pigm*) and 9311 were revealed by the *t*-test at $p < 0.05$ and $p < 0.01$, respectively. The data were derived from the trial that was performed at the Fuzhou experimental station in October 2023.

2.3. Analysis of Single Gene Line Pigm-1 and Pigm Blast Resistance Spectrum

These 195 isolates were used to inoculate 9311 (*Pigm-1*), 9311 (*Pigm*) and 9311. The results showed that the three materials were resistant to 5 blast isolates (JL-12, SH-17, SH-26, SH-68 and JY-7). However, 9311 (*Pigm-1*) and 9311 (*Pigm*) were resistant, but 9311 was susceptible, to other 189 isolates. Only one blast strain JL-37 showed differential virulence

where 9311 (*Pigm*) shows resistance, but 9311 (*Pigm-1*) and 9311 were susceptible (Table 2 and Figure 2).

Table 2. The results of indoor inoculation of single gene line 9311 (*Pigm-1*), 9311 (*Pigm*) and 9311 by *M. oryzae* isolates collected from different areas of the Fujian Province.

Name	9311 (*Pigm-1*)	9311 (*Pigm*)	9311	Name	9311 (*Pigm-1*)	9311 (*Pigm*)	9311
JL-1	R	R	S	SH-41	R	R	S
JL-2	R	R	S	SH-42	R	R	S
JL-3	R	R	S	SH-43	R	R	S
JL-4	R	R	S	SH-44	R	R	S
JL-5	R	R	S	SH-45	R	R	S
JL-6	R	R	S	SH-46	R	R	S
JL-7	R	R	S	SH-47	R	R	S
JL-8	R	R	S	SH-48	R	R	S
JL-9	R	R	S	SH-49	R	R	S
JL-10	R	R	S	SH-50	R	R	S
JL-11	R	R	R	SH-51	R	R	S
JL-12	R	R	R	SH-52	R	R	S
JL-13	R	R	S	SH-53	R	R	S
JL-14	R	R	S	SH-54	R	R	S
JL-15	R	R	S	SH-55	R	R	S
JL-16	R	R	S	SH-56	R	R	S
JL-17	R	R	S	SH-57	R	R	S
JL-18	R	R	S	SH-58	R	R	S
JL-19	R	R	S	SH-59	R	R	S
JL-20	R	R	S	SH-60	R	R	S
JL-21	R	R	S	SH-61	R	R	S
JL-22	R	R	S	SH-62	R	R	S
JL-23	R	R	S	SH-63	R	R	S
JL-24	R	R	S	SH-64	R	R	S
JL-25	R	R	S	SH-65	R	R	S
JL-26	R	R	S	SH-66	R	R	S
JL-27	R	R	S	SH-67	R	R	S
JL-28	R	R	S	SH-68	R	R	R
JL-29	R	R	S	SH-69	R	R	S
JL-30	R	R	S	SH-70	R	R	S
JL-31	R	R	S	SH-71	R	R	S
JL-32	R	R	S	SH-72	R	R	S
JL-33	R	R	S	SH-73	R	R	S
JL-34	R	R	S	SH-74	R	R	S
JL-35	R	R	S	SH-75	R	R	S
JL-36	R	R	S	SH-76	R	R	S
JL-37	S	R	S	SH-77	R	R	S
JL-38	R	R	S	SH-78	R	R	S
JL-39	R	R	S	SH-79	R	R	S
JL-40	R	R	S	SH-80	R	R	S
JL-41	R	R	S	SH-81	R	R	S
JL-42	R	R	S	SH-82	R	R	S
JL-43	R	R	S	SH-83	R	R	S
JL-44	R	R	S	SH-84	R	R	S
JL-45	R	R	S	SH-85	R	R	S
JL-46	R	R	S	SH-86	R	R	S
JL-47	R	R	S	SH-87	R	R	S
JL-48	R	R	S	SH-88	R	R	S
JL-49	R	R	S	SH-89	R	R	S
JL-50	R	R	S	JY-1	R	R	S

Table 2. *Cont.*

Name	9311 (*Pigm-1*)	9311 (*Pigm*)	9311	Name	9311 (*Pigm-1*)	9311 (*Pigm*)	9311
JL-51	R	R	S	JY-2	R	R	S
JL-52	R	R	S	JY-3	R	R	S
JL-53	R	R	S	JY-4	R	R	S
JL-54	R	R	S	JY-5	R	R	S
JL-55	R	R	S	JY-6	R	R	S
JL-56	R	R	S	JY-7	R	R	R
JL-57	R	R	S	JY-8	R	R	S
JL-58	R	R	S	JY-9	R	R	S
JL-59	R	R	S	JY-10	R	R	S
JL-60	R	R	S	JY-11	R	R	S
SH-1	R	R	S	JY-12	R	R	S
SH-2	R	R	S	JY-13	R	R	S
SH-3	R	R	S	JY-14	R	R	S
SH-4	R	R	S	JY-15	R	R	S
SH-5	R	R	S	JY-16	R	R	S
SH-6	R	R	S	JY-17	R	R	S
SH-7	R	R	S	JY-18	R	R	S
SH-8	R	R	S	JY-19	R	R	S
SH-9	R	R	S	JY-20	R	R	S
SH-10	R	R	S	JY-21	R	R	S
SH-11	R	R	S	JY-22	R	R	S
SH-12	R	R	S	JY-23	R	R	S
SH-13	R	R	S	JY-24	R	R	S
SH-14	R	R	S	JY-25	R	R	S
SH-15	R	R	S	JY-26	R	R	S
SH-16	R	R	S	JY-27	R	R	S
SH-17	R	R	R	JY-28	R	R	S
SH-18	R	R	S	JY-29	R	R	S
SH-19	R	R	S	JY-30	R	R	S
SH-20	R	R	S	JY-31	R	R	S
SH-21	R	R	S	JY-32	R	R	S
SH-22	R	R	S	JY-33	R	R	S
SH-23	R	R	S	JY-34	R	R	S
SH-24	R	R	S	JY-35	R	R	S
SH-25	R	R	S	JY-36	R	R	S
SH-26	R	R	R	JY-37	R	R	S
SH-27	R	R	S	JY-38	R	R	S
SH-28	R	R	S	JY-39	R	R	S
SH-29	R	R	S	JY-40	R	R	S
SH-30	R	R	S	JY-41	R	R	S
SH-31	R	R	S	JY-42	R	R	S
SH-32	R	R	S	JY-43	R	R	S
SH-33	R	R	S	JY-44	R	R	S
SH-34	R	R	S	JY-45	R	R	S
SH-35	R	R	S	JY-46	R	R	S
SH-36	R	R	S				
SH-37	R	R	S				
SH-38	R	R	S				
SH-39	R	R	S				
SH-40	R	R	S				

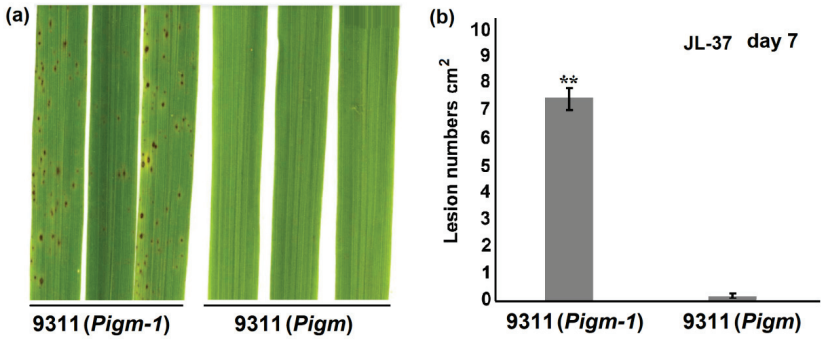

Figure 2. The disease symptoms of the rice cultivars 9311 (*Pigm-1*) and 9311 (*Pigm*) infected by blast fungi. (**a**) Blast resistance of 9311 (*Pigm-1*) and 9311 (*Pigm*) plants using spraying inoculation in a greenhouse. Representative leaves obtained 7 days after inoculation with blast strain JL-37. (**b**) Lesion numbers cm^2 on the rice leaves (M ± SD, $n > 10$ leaves) in (**a**), Asterisk (**) indicates statistical significance ($p < 0.01$) determined by Student's t test.

2.4. An Improved Breeding Line Containing Pigm-1 Gene Was Obtained

In this study, S19-118 was selected and used as the maternal parent, and R20-4, a selected restorer line containing *Pigm-1*, was used as the male parent for hybridization. Then the hybrid F_1 was backcrossed with S19-118, and a single strain containing the *Pigm-1* gene was selected and further backcrossed with S19-118 by combining molecular markers until BC_3F_1 was followed by selfing. Each generation was tracked by molecular markers and planting was continued for 3 generations until BC_3F_9. Other agronomic traits of the line were stable.

In the BC_3F_9, a line T22-23 with excellent comprehensive agronomic traits was selected and named Xiannuo 23. Finally, target molecular markers were used to genotype Xiannuo 23. The results showed that Xiannuo 23 contained the target gene *Pigm-1* (Figure 3a).

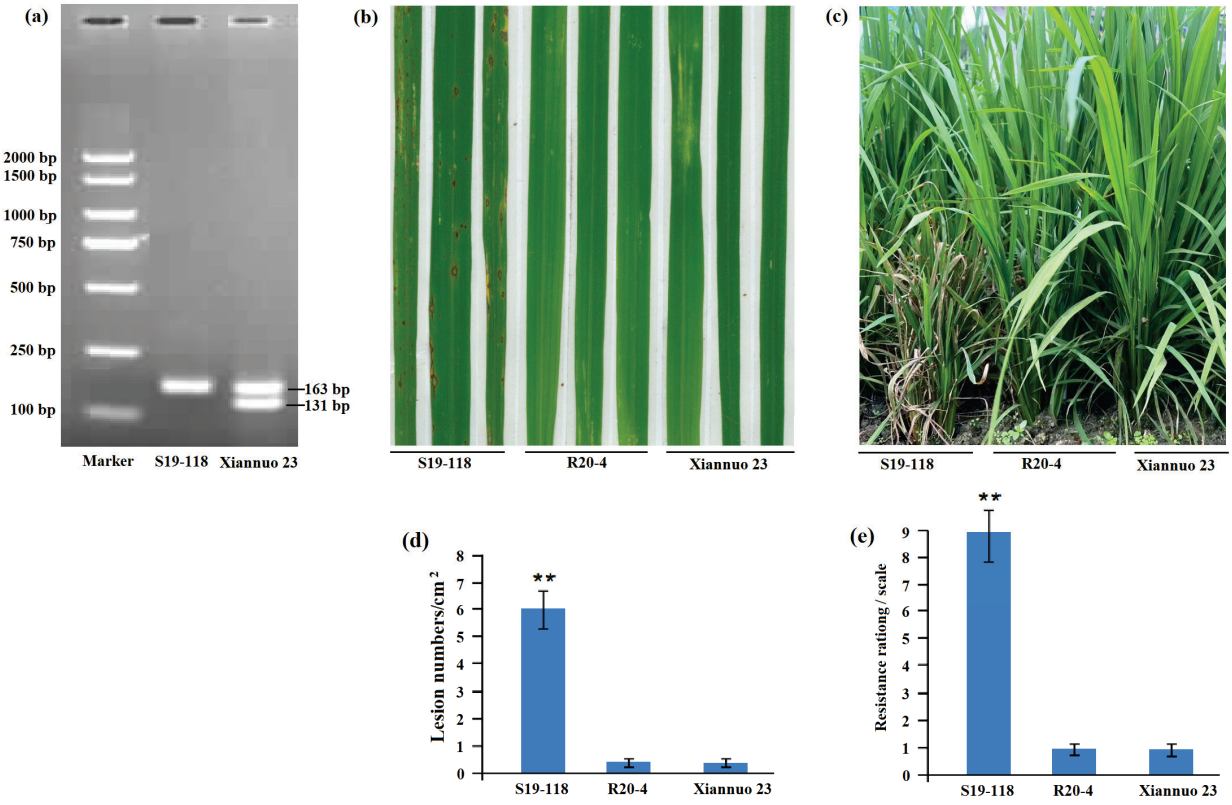

Figure 3. Detection of molecular markers in improved lines and phenotype of blast resistance. (**a**) The results of *Pigm-1* molecular marker detection showed that Xiannuo 23 contained *Pigm-1* that

had two expected bands of 163 bp and 131 bp, while S19-118 without *Pigm-1* had only one band of 163 bp. (**b**) The greenhouse test results of the improved lines showed that S19-118 was susceptible to rice blast, while R20-4 and Xiannuo 23 were resistant to rice blast. (**c**) The field test results of the improved line showed that S19-118 was susceptible to, but R20-4 and Xiannuo 23 were resistant to, the rice blast variant Guy 11. (**d**) Lesion numbers per cm^2 on the rice leaves (M ± SD, $n > 10$ leaves) in (**b**). (**e**) The incidence scores of the strains are based on the results in (**c**), the disease score of S19-118 is 9, while both R20-4 and Xiannuo 23 have a score of 1. Asterisks (**) indicate statistical significance ($p < 0.01$) determined by Student's *t* test.

2.5. Analysis of Rice Blast Resistance of Xiannuo 23 Improved Line

In order to phenotypically identify the rice blast resistance of the created Xiannuo 23, we used the rice blast fungus Guy11 for resistance identification at the seedling stage. The results showed that the recipient parent S19-118 was susceptible to the disease, and the donor parent R20-4 and the improved line Xiannuo 23 were resistant to the disease (Figure 3b). To further evaluate the resistance of the improved line Xiannuo 23 under natural conditions, recipient parent S19-118, donor parent R20-4 and improved line Xiannuo 23 were all planted in the National Blast Resistance Identification Center in Shanghang County, Fujian Province. The test results were consistent with the laboratory test results, and the improved line Xiannuo 23 showed resistance to the disease (Figure 3c).

2.6. Analysis of Main Agronomic Characters of Xiannuo 23

To further analyze the changes of other agronomic traits of Xiannuo 23 while improving resistance, we investigated and analyzed the agronomic traits of the donor parent R20-4, the recipient parent S19-118 and the improved line Xiannuo 23 at maturity. The results showed that, compared with the recipient parent S19-118, the improved line Xiannuo 23 had no change in plant height, panicle length, effective panicle, number of spiklets per panicle and grain length, but there were significant differences in seed setting rate, grain width and 1000-grain weight. Grain width and 1000-grain weight were significantly increased, while seed setting rate decreased. However, there was an increase in average yield per plant (Figure 4).

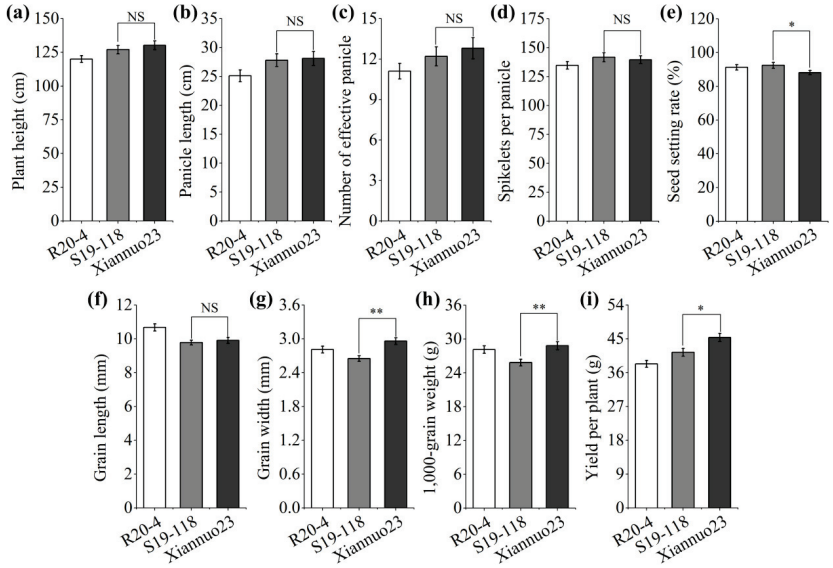

Figure 4. Comparison of main agronomic traits between parents and improved line Xiannuo 23. (**a**) Plant height. (**b**) Panicle length. (**c**) Number of effective panicle. (**d**) Spikelets per panicle. (**e**) Seed setting rate. (**f**) Grain length (**g**) Grain width. (**h**) 1,000-grain weight. (**i**) Yield per plant. NS: no statistically significant difference; * and ** show significant at $p < 0.05$ and 0.01 levels, respectively.

3. Discussion

3.1. *Pigm-1 Antispectral Analysis*

Studies have shown that different alleles of rice blast may have different resistance spectra and resistance levels [8]. Several alleles have been identified at the *Piz* locus, such as *Pi2*, *Pi9*, *Piz-t* (*Pizh*), *Pigm* (*Pi50*) and *Pigm-1* [10–12,16]. Xie et al. [19] used 31 representative *M. oryzae* isolates from different regions to inoculate disease-resistant materials containing *Pi9*, *Pi2* and *Pigm* and found that Gumei4 (*Pigm*) was resistant to all the isolates tested, C101A51 (*Pi2*) was susceptible to eight isolates and 75-1-127 (*Pi9*) was susceptible to five isolates. Peng et al. [20] used 444 representative *M. oryzae* isolates from the Hunan Province to inoculate monogenic rice lines carrying the resistance genes *Pita*, *Pizt*, *Pikm*, *Pib* or *Pi9*. The result shows that both *Pi9* and *Pikm* conferred resistance to >75% of the tested isolates, while *Pizt*, *Pita* and *Pib* were effective against 55.63%, 15.31% and 3.15% of the isolates, respectively. Wu et al. [21] showed in a field inoculation that the resistance profiles of *Pi40* and *Pigm* were higher than *Pi2*, while *Pi2* was higher than *Pi9*. These results indicated that the resistance spectra of *Pigm* and *Pi40* were higher than those of *Pi9* and *Pi2*, and the resistance spectra of *Pi9* and *Pi2* were different in different regions or at different growth stages of rice.

The rice blast resistance genes *Pigm* and *Pigm-1* are both more broad-spectrum resistant in production [12,16], but it is not clear whether the resistance levels between them are different. In this study, 195 different blast fungi from the Fujian Province were isolated. The results showed that only one blast fungus JL-37 had differences in resistance to *Pigm* and *Pigm-1*, while the remaining 194 had no differences. In our previous studies, it was speculated that *Pigm* and *Pigm-1* had certain structural differences, and it was also speculated that this difference might be related to the evolution of rice blast fungus and the recognition of different pathogens by the R protein [16]. However, the identification of resistance proteins *Pigm* and *Pigm-1* with blast fungus JL-37 and the formation of a resistance mechanism still need to be further studied.

3.2. *Analysis of Pigm-1 Breeding and Utilization Prospects*

Since cloning *Pigm-1*, we have created a series of new restorer materials using this gene. For example, using Minghui 63 as the receptor and Shuangkang 77009, which carries the resistance gene *Pigm-1*, as the donor, six new materials for resistance to rice blast were created [16]. Using restorer line R20 as the receptor and Shuangkang 77009, which carries the resistance gene *Pigm-1*, as the donor, four new materials for resistance to rice blast were generated [22]. Five new materials for resistance to rice blast and white leaf blight were prepared using Shuangkang 77009 with rice blast resistance gene *Pigm-1* and IRBBL23, with white leaf blight resistance gene *Xa23* as the donor [23]. Among these materials, the setting rate is above 90%. The allele of *Pigm-1*, *Pigm*, has been applied in rice disease resistance molecular breeding by more than 30 seed subsidiaries and breeding units in China, and new varieties have participated in regional trials and variety approval [12,16,24,25], which indicates that *Pigm-1* and *Pigm* have good application prospects in rice blast resistance breeding.

The latest study revealed that the PIBP4-Rab5a transport machine is involved in regulating the accumulation of the NLR protein PigmR in the cell membrane microregion, and the PigmR protein can activate the OsRac1 protein in the microregion, promote the production of reactive oxygen species and enhance the resistance of rice to blast disease [26]. The study identified a new NLR immune signaling pathway, providing a new perspective and target for PigmR disease resistance molecular breeding.

3.1. Further Study on the Relationship Between Pigm-1 and Waxy Genes

In this study, R20-4 containing *Pigm-1* was more than 90%. However, when Xiannuo 23 containing *Pigm-1* was cultivated with R20-4 containing *Pigm-1* as the donor, the seed setting rate was below 90% (Figure 4e). At the same time, in the process of creating glutinous rice materials, it was found that the seed setting rate of glutinous rice lines containing *Pigm-1* was generally lower than 90%, while the setting rate of glutinous rice lines without *Pigm-1* was relatively higher (>90%). Moreover, studies have shown that some rice blast *R* genes are linked to adverse agronomic traits often at the expense of yield [27,28]. Therefore, it is still uncertain whether the reason for the decreased seed setting rate in this study is the adverse linkage of *Pigm-1* and Waxy genes, and further research is needed to resolve this issue.

4. Materials and Methods

4.1. Isolation of Rice Blast Fungus

Rice blast fungus was isolated by referring to the method reported by Zhao et al. [29]. Fresh leaves with severe disease were collected from Shanghang County, Jianyang County and Jiangle County in the Fujian Province, and then brought to the experiment for single packet separation.

We collected blast fungi in areas with high blast incidence in the Fujian Province and isolated them by single packet isolation. Finally, we obtained 60 monospore blast fungi from Jiangle County (JL-1 to JL-60), 89 from Shanghang County (SH-1 to SH-89) and 46 from Jianyang County (JY-1 to JY-46, Table 2).

4.2. Rice Blast Fungus Incubation

For the internal resistance identification of rice blast, the inoculation method was improved with reference to the previous methods [30]. For spray inoculation, two-week-old rice seedling leaves were sprayed with a spore suspension (1×10^5 spores/mL). The inoculated rice was placed in a moist chamber at 28 °C for 24 h in darkness and then transferred to another moist chamber with a photoperiod of 12 h under light. Leaves of 4-week-old rice seedlings were lightly wounded with a mouse ear punch, and 5 μL of spore suspension (5×10^5 spores/mL) was added to the wound. The lesion size was measured after 5 days of incubation. Among the inoculated materials, there were 24 strains per material and 3 replicates were used.

The recipient parent S19-118, the donor parent R20-4 and the improved line Xiannuo 23 were incubated in the National Blast Resistance Identification Center in Shanghang County, Fujian Province on 18 May 2023. On 5 June 2023, the investigation and statistics of blast resistance of each strain at seedling stage were conducted. Plants were rated on a scale of 0–9 from the rice standard evaluation system (International Rice Research Institute 1996). Grades 0–3 were resistant, and grades 4–9 were susceptible.

4.3. Detection and Analysis of Pigm-1 and Pigm Functional Markers

The resistance function marker of *Pigm-1* refers to the marker developed by Yang et al. [16]. *Pigm-1*F: TTATTTCGTTTGCTATG; *Pigm-1*R: GGACTATGTGATCGGTTA. Specific primers were designed based on the *Pigm-1* sequence, and two bands of 163 bp and 131 bp were amplified with *Pigm-1* (Figure 5).

The resistance function marker of *Pigm* refers to the marker developed by Chen et al. [31]. T-*Pigm*-O-F: TAAGAATTGAGGTGGTTAGTTGAACGGAGA; T-*Pigm*-O-R: TTGCATGGCTCCACTACCCACTATAAG. For the extraction of genomic DNA from rice leaves, and the detection and analysis of molecular marker genes, refer to Yang et al. [16].

Figure 5. Primers designed for detecting the *Pigm-1* gene. The bases in red represent the location of the designed primer. (**a**) The *Pigm-1* specific primers amplified 163 bp bands. (**b**) The *Pigm-1* specific primers simultaneously amplified 131 bp bands.

4.4. Investigation of Agronomic Traits

In July 2023, experimental materials, such as R20-4, S19-118, Xiannuo 23, 9311 (*Pigm-1*), 9311 (*Pigm*) and 9311, were planted in the experimental farm of the Rice Research Institute of Fujian Academy of Agricultural Sciences. The materials were all planted in the same plot, 3 replicates were set in each plot, 5 rows in each plot and 8 plants in each row. Field cultivation techniques are managed according to conventional cultivation techniques.

The main agronomic traits of rice, such as plant height, panicle length and seed setting rate, were investigated and analyzed with reference to Beena et al. [32].

Author Contributions: All the authors contributed to the conception and design of the study. D.Y. planned the project, performed the experiments and data collection and wrote the manuscript with input from all authors. The collection and analysis of the data were performed by Y.J., N.H., Z.C., S.L., F.H., W.W. and D.Y. The first draft of the manuscript was written by D.Y., N.H., Y.J. and Q.Q.L. All authors discussed the results and contributed to the final manuscript. All authors have read and agreed to the published version of the manuscript.

Funding: The work was supported by the Fujian public welfare competitive project (No. 2024R1055), the Special Fund for Agro-scientific Research in the Public Interest of Fujian Province (2023R1021006; 2024R1022001), the Free Exploration Project of Fujian Academy of Agricultural Sciences (No. ZYTS2023001), the National Natural Science Foundation of China (No. 32402387), the National Natural Science Foundation project extension research project (No. GJYS202302), 5511 Collaborative Engineering Project (No. XTCXGC2021001), the Fujian Provincial Natural Science Foundation of China (No. 2022J01450 and 2024J01171), and the Hundred Talent Plans of Fujian Province.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data are contained within the article.

Conflicts of Interest: The authors declare that they have no competing interests.

References

1. Elert, E. Rice by the numbers: A good grain. *Nature* **2014**, *514*, 50–51. [CrossRef] [PubMed]
2. Kamarudin, S.A.A.; Ahmad, F.; Hasan, N.; Hisham, S.N.; Yusof, S.N.; Abu, H.A.; Hussein, S.; Harun, A.R.; Wan, C.W.D.; Md, S.M.; et al. Whole genome resequencing data and grain quality traits of the rice cultivar Mahsuri and its blast disease resistant mutant line, Mahsuri Mutant. *Data Brief* **2023**, *52*, 109974. [CrossRef] [PubMed]
3. Zhou, H.P.; Zhang, F.; Chen, K.; Shen, C.C.; Zhu, S.B.; Qiu, X.J.; Xu, J.L. Identification of rice blast resistance in *xian* and *geng* germplasms by genomewide association study. *Acta Agron. Sin.* **2023**, *49*, 1170–1183.
4. He, F.; Zhang, H.; Liu, J.L.; Wang, Z.L.; Wang, G.L. Recent advances in understanding the innate immune mechanisms and developing new disease resistance breeding strategies against the rice blast fungus *Magnaporthe oryzae* in rice. *Yi Chuan* **2014**, *36*, 756–765. [PubMed]
5. Kou, Y.J.; Shi, H.B.; Qiu, J.H.; Tao, Z.; Wang, W.M. Effectors and environment modulating rice blast disease: From understanding to effective control. *Trends Microbiol.* **2024**, *32*, 1007–1020. [CrossRef] [PubMed]
6. Annegowda, D.C.; Prasannakumar, M.K.; Mahesh, H.B.; Siddabasappa, C.B.; Devanna, P.; Banakar, S.N.; Manojkumar, H.B.; Prasad, S.R. Rice blast disease in India: Present status and future challenges. *Integr. Adv. Rice Res.* **2021**, *98847*, 157–197.

7. An, Z.S.; Liu, G.L.; Mei, H.W.; Li, L.T.; Wang, J.H.; Luo, L.J. Improving blast resistance of parental lines of drought resistant hybrid rice by marker-assisted selection. *Mol. Plant Breed.* **2010**, *8*, 1172–1176.
8. Greenwood, J.R.; Lacorte-Apostol, V.; Kroj, T.; Padilla, J.; Telebanco-Yanoria, M.J.; Glaus, A.N.; Roulin, A.; Padilla, A.; Zhou, B.; Keller, B.; et al. Genome-wide association analysis uncovers rice blast resistance alleles of *Ptr* and *Pia*. *Commun. Biol.* **2024**, *7*, 607. [CrossRef] [PubMed]
9. Mutiga, S.K.; Orwa, P.; Nganga, E.M.; Kyallo, M.M.; Rotich, F.; Gichuhi, E.; Kimani, J.M.; Mwongera, D.T.; Were, V.M.; Yanoria, M.J.; et al. Characterization of blast resistance in a diverse rice panel from sub-saharan africa. *Phytopathology* **2023**, *113*, 1278–1288. [CrossRef]
10. Xiao, N.; Wu, Y.Y.; Li, A.H. Strategy for use of rice blast resistance genes in rice molecular breeding. *Rice Sci.* **2020**, *4*, 263–277.
11. Wu, Y.Y.; Xiao, N.; Yu, L.; Pan, C.H.; Li, Y.H.; Zhang, X.X.; Liu, G.Q.; Dai, Z.Y.; Pan, X.B.; Li, A.H. Combination patterns of major *R* genes determine the level of resistance to the *M. oryzae* in rice (*Oryza sativa* L.). *PLoS ONE* **2015**, *10*, e126130. [CrossRef] [PubMed]
12. Deng, Y.W.; Zhai, K.R.; Xie, Z.; Yang, D.Y.; Zhu, X.D.; Liu, J.Z.; Wang, X.; Qin, P.; Yang, Y.Z.; Zhang, G.M.; et al. Epigenetic regulation of antagonistic receptors confers rice blast resistance with yield balance. *Science* **2017**, *355*, 962–965. [CrossRef]
13. Zhou, Y.; Wan, T.; Yuan, B.; Lei, F.; Chen, M.J.; Wang, Q.; Huang, P.; Kou, S.Y.; Qiu, W.X.; Liu, L. Improving rice blast resistance by mining broad-spectrum resistance genes at the *Pik* locus. *Rice Sci.* **2022**, *29*, 133–142.
14. Yoshida, K.; Saitoh, H.; Fujisawa, S.; Kanzaki, H.; Matsumura, H.; Yoshida, K.; Tosa, Y.; Chuma, I.; Takano, Y.; Win, J.; et al. Association genetics reveals three novel avirulence genes from the rice blast fungal pathogen *Magnaporthe oryzae*. *Plant Cell* **2009**, *21*, 1573–1591. [CrossRef] [PubMed]
15. Kanzaki, H.; Yoshida, K.; Saitoh, H.; Fujisaki, K.; Hirabuchi, A.; Alaux, L.; Fournier, E.E.; Tharreau, D.; Terauchi, R. Arms race co-evolution of *Magnaporthe oryzae AVR-Pik* and rice *Pik* genes driven by their physical interactions. *Plant J.* **2012**, *72*, 894–907. [CrossRef]
16. Yang, D.W.; Li, S.P.; Lu, L.; Fang, J.B.; Wang, W.; Cui, H.T.; Tang, D.Z. Identification and application of the *Pigm-1* gene in rice disease-resistance breeding. *Plant Biol.* **2020**, *22*, 1022–1029. [CrossRef]
17. Feng, Z.M.; Li, M.Y.; Xu, Z.W.; Gao, P.; Wu, Y.Y.; Wu, K.T.; Zhao, J.H.; Wang, X.Q.; Wang, J.N.; Li, M.C.; et al. Development of rice variety with durable and broad-spectrum resistance to blast disease through marker-assisted introduction of *Pigm* gene. *Front. Plant Sci.* **2022**, *13*, 937767. [CrossRef] [PubMed]
18. Wu, Y.; Xiao, N.; Li, Y.; Gao, Q.; Ning, Y.; Yu, L.; Cai, Y.; Pan, C.; Zhang, X.; Huang, N.; et al. Identification and fine mapping of *qPBR10-1*, a novel locus controlling panicle blast resistance in *Pigm*-containing P/TGMS line. *Mol. Breed.* **2021**, *41*, 75. [CrossRef] [PubMed]
19. Xie, Z.; Yan, B.X.; Shou, J.Y.; Tang, J.; Wang, X.; Zhai, K.R.; Liu, J.Y.; Li, Q.; Luo, M.Z.; Deng, Y.W.; et al. A nucleotide-binding site-leucine-rich repeat receptor pair confers broad-spectrum disease resistance through physical association in rice. *Philos. Trans. R. Soc. London Biol. Sci.* **2019**, *4*, 374. [CrossRef]
20. Peng, Z.R.; Li, L.; Wu, S.H.; Chen, X.L.; Shi, Y.F.; He, Q.; Shu, F.; Zhang, W.H.; Sun, P.Y.; Deng, H.F.; et al. Frequencies and variations of *Magnaporthe oryzae* avirulence genes in Hunan province, China. *Plant Dis.* **2021**, *105*, 3829–3834. [CrossRef]
21. Wu, Y.Y.; Chen, Y.; Pan, C.H.; Xiao, N.; Yu, L.; Li, Y.H.; Zhang, X.X.; Pan, X.B.; Chen, X.J.; Liang, C.Z.; et al. Development and evaluation of near-isogenic lines with different blast resistance alleles at the *Piz* locus in japonica rice from the lower region of the Yangtze River, China. *Plant Dis.* **2017**, *101*, 1283–1291. [CrossRef]
22. He, N.Q.; Yang, D.W.; Zheng, X.H.; Huang, F.H.; Cheng, C.P.; Ye, N. Improving blast resistance of R20 by molecular marker-assisted selection of *Pigm-1* gene. *J. Nucl. Agric. Sci.* **2022**, *36*, 245–250.
23. Yang, D.W.; He, N.Q.; Huang, F.H. Pyramiding *Pigm-1* and *Xa23* genes in rice (*Oryza sativa* L.) by marker-assisted selection. *J. Northwest A F Univ. (Nat. Sci. Ed.)* **2023**, *51*, 37–45.
24. Peng, P.; Jiang, H.Y.; Luo, L.H.; Ye, C.R.; Xiao, Y.H. Pyramiding of multiple genes to improve rice blast resistance of photo-thermo sensitive male sterile line, without yield penalty in hybrid rice production. *Plants* **2023**, *12*, 1389. [CrossRef] [PubMed]
25. Wang, H.P.; Wang, G.; Qin, R.; Gong, C.Q.; Zhou, D.; Li, D.K.; Luo, B.J.; Jin, J.H.; Deng, Q.M.; Wang, S.Q.; et al. Improvement of quality and disease resistance for a heavy-panicle hybrid restorer line, R600, in rice (*Oryza sativa* L.) by gene pyramiding breeding. *Curr. Issues Mol. Biol.* **2024**, *46*, 10762–10778. [CrossRef]
26. Liang, D.; Yang, D.Y.; Li, T.; Zhu, Z.; Yan, B.X.; He, Y.; Li, X.Y.; Zhai, K.R.; Liu, J.Y.; Kawano, Y.; et al. A PRA-Rab trafficking machinery modulates NLR immune receptor plasma membrane microdomain anchoring and blast resistance in rice. *Sci. Bull.* **2024**, *9*, S2095-9273(24)00911-3. [CrossRef]
27. Nelson, R.; Wiesner-Hanks, T.; Wisser, R.; Balint-Kurti, P. Navigating complexity to breed disease-resistant crops. *Nat. Rev. Genet.* **2018**, *19*, 21–33. [CrossRef] [PubMed]
28. Wang, J.; Zhou, L.; Shi, H.; Chern, M.; Yu, H.; Yi, H.; He, M.; Yin, J.J.; Zhu, X.B.; Li, Y.; et al. A single transcription factor promotes both yield and immunity in rice. *Science* **2018**, *361*, 1026–1028. [CrossRef]
29. Zhao, D.D.; Chung, H.; Jang, Y.H.; Farooq, M.; Choi, S.Y.; Du, X.X.; Kim, K.M. Analysis of rice blast fungus genetic diversity and identification of a novel blast resistance *OsDRq12* gene. *Phytopathology* **2024**, *114*, 1917–1925. [CrossRef] [PubMed]

30. Yang, D.W.; Li, S.P.; Lu, L.; Zheng, Z.C.; Tang, D.Z.; Cui, H.T. Transcriptome analysis of rice response to blast fungus identified core genes involved in immunity. *Plant Cell Environ.* **2021**, *44*, 3103–3121. [CrossRef] [PubMed]
31. Chen, T.; Sun, X.C.; Zhang, S.L.; Liang, W.H.; Zhou, L.H.; Zhao, Q.Y.; Yao, S.; Zhao, L.; Zhao, C.F.; Zhu, Z.; et al. Development and application of Pigm-specific molecular markers for broad-spectrum resistance gene in rice blast. *Chin. J. Rice Sci.* **2020**, *34*, 28–36.
32. Beena, R.; Kirubakaran, S.; Nithya, N.; Manickavelu, A.; Sah, R.P.; Abida, P.S.; Sreekumar, J.; Jaslam, P.M.; Rejeth, R.; Jayalekshmy, V.G.; et al. Association mapping of drought tolerance and agronomic traits in rice (*Oryza sativa* L.) landraces. *BMC Plant Biol.* **2021**, *21*, 484. [CrossRef] [PubMed]

MDPI AG
Grosspeteranlage 5
4052 Basel
Switzerland
Tel.: +41 61 683 77 34

Plants Editorial Office
E-mail: plants@mdpi.com
www.mdpi.com/journal/plants

Academic Open
Access Publishing

mdpi.com

ISBN 978-3-7258-4114-1