

Identification and Haplotype Diversity Analysis of DVL Family Genes in Rice

Xinting Liu

Anhui Agricultural University

You Zhou

Anhui Agricultural University

Manqiong Zhu

Anhui Agricultural University

Wenhao Lv

Anhui Agricultural University

Chunni Wang

Anhui Agricultural University

Meng Zhang

Anhui Agricultural University

Chenghang Tang

Anhui Agricultural University

Shuo Yang

Anhui Agricultural University

Yingyao Shi

Shiyy123@163.com

Anhui Agricultural University

Research Article

Keywords:

Posted Date: July 2nd, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-9770845/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Additional Declarations: No competing interests reported.

Version of Record: A version of this preprint was published at BMC Plant Biology on September 26th, 2026. See the published version at <https://doi.org/10.1186/s12870-026-10051-5>.

Abstract

Background: The DEVIL (DVL) family encodes small signaling peptides that are functionally important for coordinating plant development and stress responses. However, the evolutionary dynamics, genetic diversity, and agronomic relevance of DVL genes in rice remain poorly understood.

Results: In this study, we identified five DVL genes in the rice genome through comprehensive bioinformatic analyses. Promoter characterization revealed multiple stress-responsive cis-regulatory elements. Expression profiling demonstrated tissue-specific patterns, with transcript levels of selected OsDVL genes showing significant changes under salt and cold stress. Haplotype analysis across 3,010 rice accessions uncovered substantial allelic diversity at OsDVL loci, which appears to have contributed to population differentiation and genetic improvement during rice breeding. Importantly, the predominant haplotypes at most OsDVL loci were significantly associated with key yield-related traits.

Conclusions: Our findings suggest that OsDVL genes may play a role in abiotic stress tolerance and provide new insights into their adaptive evolution. These results also highlight the potential utility of DVL-associated haplotypes as targets for marker-assisted breeding in rice.

1. Introduction

Rice (*Oryza sativa* L.) is one of the most important food crops in the world, as well as a cereal model plant with high intraspecific genetic diversity (Li & Rutger, 2000, Yu *et al.*, 2003). Numerous studies have shown that during growth, rice is affected by different abiotic stresses such as heat, cold, salt, and drought and biotic stresses such as hormones and pathogens. Heat stress is a crucial abiotic factor with particularly significant impacts on rice production. A thorough understanding of the physiological mechanisms underlying heat tolerance is essential for mitigating the adverse effects of high-temperature stress on rice (Beena *et al.*, 2025). Cold stress also poses a significant threat to rice production, with seedlings during germination being particularly susceptible (Jo *et al.*, 2025). Salinity represents another critical abiotic constraint that severely impairs both vegetative growth and grain yield in rice (Zhao *et al.*, 2025). Therefore, these environmental stressors, compounded by ongoing climate change and soil deterioration, present mounting challenges to global rice production (Wang *et al.*, 2025).

Plant peptides play pivotal roles in various aspects of plant growth and development (Guo *et al.*, 2015). Small peptides, in particular, are regulatory molecules that are important for coordinating the cellular responses necessary for plant differentiation, growth, and development (Wen *et al.*, 2004). DVL (DEVIL) is one family of small peptides. It has been found that *DVL* genes are related to the fuzz development in cotton (Jiao *et al.*, 2024). In *Arabidopsis thaliana*, the *DVL* genes are associated with cell proliferation (Ikeuchi *et al.*, 2011), and their overexpression can lead to severe developmental changes (Valdivia *et al.*, 2012). Another study confirmed that overexpression of *DVL* genes in *A. thaliana* led to sharp decreases in root length and pleiotropic phenotypes in shoots (Yang *et al.*, 2024). Moreover, the *DVL* genes in foxtail millet regulate root growth by enhancing the expression of Aux/IAAs and reducing the protein accumulation of PINs to inhibit auxin signaling (Wang *et al.*, 2024).

The release of the genomic sequences for 3,010 rice accessions (3KRG), originating from 89 countries across the globe, has generated a wealth of genetic resources for rice research (Wang *et al.*, 2018). The 3KRG resource enables the integration of haplotype data for functional exploration of rice genes (Zeng *et al.*, 2023, Cheng *et al.*, 2023). In recent years, growing attention has been paid to the expression patterns, biological functions, and stress-related roles of DVL genes, driven by advances in molecular biology. Nevertheless, knowledge regarding the DVL gene family in rice—including its phylogenetic relationships, functional specificity, genetic diversity, allelic variation, and associations with agronomic traits—remains limited. Achieving further gains in rice yield will depend on the adoption of innovative breeding approaches in future improvement programs. Such efforts, however, necessitate a thorough elucidation of the functional roles and diversity of rice genes and their regulatory networks, along with their relationships to key agronomic traits.

In the present work, a systematic genome-wide identification of the DVL gene family was performed in the rice cultivar Nipponbare, leading to the discovery of five OsDVL members. We further characterized these genes with respect to their structural features, functional annotations, expression profiles, and phylogenetic relationships, thereby establishing a foundation for subsequent investigations into their biological roles, regulatory networks, and potential applications. Moreover, we examined the gcHap diversity of OsDVL loci and evaluated their correlations with key agronomic traits, which may offer useful theoretical insights for the development of future breeding strategies.

2. Materials and Methods

2.1 Identification of Rice DVL Family Genes

For genome-wide identification of the DVL gene family in rice, the genomic sequence of the japonica variety Nipponbare was downloaded from the Rice Resource Center database (<https://ricerc.sicau.edu.cn>) on April 7, 2025, and is hereafter designated as the Os reference genome in this work. The PF08137 profile of the hidden Markov model (HMM) was retrieved from the Pfam database (accessed April 7, 2025; <http://pfam.xfam.org/>). We then applied the Simple Hmm Search tool within TBsoftwares to derive candidate DVL gene and protein sequences (Chen *et al.*, 2020). The NCBI online platform (<https://www.ncbi.nlm.nih.gov/Structure/bwrpsb/bwrpsb.cgi>) was subsequently used to examine the conserved domains of these genes. By integrating functional annotations, we performed an exhaustive evaluation to pinpoint genuine family members, which were later renamed based on their chromosomal locations. Subcellular localization was predicted using the Cell-PLoc2.0 web server (Euk-mPLoc 2.0) (Chou & Shen, 2008). Furthermore, the physicochemical parameters of the OsDVL genes, such as theoretical isoelectric point (pI) and molecular mass (in Da), were calculated using the Protein Parameter Calc module implemented in TBtools (Chen *et al.*, 2020).

2.2 Phylogenetic analysis of *DVL* genes

To examine the evolutionary relationships of DVL genes among diverse plant species, we retrieved DVL gene sequences from several representative plants, namely *Arabidopsis thaliana*, indica rice, maize (*Zea*

mays), wheat (*Triticum aestivum*), foxtail millet (*Setaria italica*), and barley (*Hordeum vulgare*), using the Ensembl Plants database (<https://plants.ensembl.org/index.html>). A phylogenetic tree was generated via MEGA11 software based on the Neighbor-joining algorithm (Kumar et al., 2016). The resulting tree was subsequently visualized and enhanced using the iTOL online platform (accessed April 7, 2025; <https://itol.embl.de/>) (Letunic & Bork, 2016). The detailed experimental workflow followed that described in an earlier publication (Rui et al., 2022).

2.3 Analysis of *Cis*-acting Elements, Conserved Motifs, Conserved Domains, and Promoter Regions of DVL Proteins

Gene structure annotation files retrieved from the Rice Resource Center website were visualized using the Gene Structure View (Advanced) module of TBtools (Chen et al., 2020). For motif discovery and visualization, we combined the MEME suite (<https://meme-suite.org/meme/tools/meme>; accessed April 7, 2025) with TBtools (Bailey et al., 2015). Conserved domains were identified through the NCBI CDD search tool, and the resulting data were displayed with TBtools. To analyze *cis*-regulatory elements, the 2,000-bp sequences immediately upstream of the CDS—corresponding to the putative promoter regions of DVL genes—were extracted using the Gtf/Gff3 sequence extraction utility in TBtools. These promoter sequences were subsequently submitted to the PlantCARE database (<http://bioinformatics.psb.ugent.be/webtools/plantcare/html/access>) for promoter analysis (Lescot et al., 2002). The output data were filtered according to the information provided in the result table, and the curated files were retained for further examination. Final visualization was performed with the Simple BioSequence Viewer tool of TBtools (Chen et al., 2020).

2.4 Collinearity Analysis of *OsDVL* Genes

To begin with, the gff and DNA files for the selected species were downloaded from the EnsemblPlants database. TBtools was then used to construct pairwise gene datasets for six representative plant species: *Arabidopsis thaliana*, *Oryza sativa indica*, *Zea mays*, *Triticum aestivum*, *Setaria italica*, and *Hordeum vulgare* (hereafter abbreviated as At, Xi, Zm, Ta, Si, Hv). The syntenic relationships among duplicated gene pairs across these species were assessed using the one-step MCScanX and AdvanceCircos modules integrated in TBtools (Wang et al., 2012). Ultimately, the resulting synteny patterns were displayed graphically with the aid of chromosome length files and genome comparison datasets.

2.5 RNA-seq-based Gene Expression Analysis

Expression data (FPKM values, representing transcripts per kilobase per million mapped reads) for *OsDVL* genes were obtained from the PPRD RNA-seq database (<http://ipf.sustech.edu.cn/pub/plantrna/>; accessed April 7, 2025) (Yu et al., 2022). For species or conditions with multiple sequencing libraries, the arithmetic mean of FPKM values was calculated and applied to evaluate expression across various

tissues, distinct developmental stages, and diverse abiotic stress treatments. Heatmaps were then generated via TBtools software to graphically display the expression profiles (Chen et al., 2020).

2.6 Material treatment

Seeds of the Nipponbare cultivar were surface-sterilized with 3% sodium hypochlorite for 30 min, incubated at 28°C for 3 days to allow germination, and subsequently transferred to hydroponic containers supplemented with Hoagland solution. The plants were cultivated in a programmable growth chamber under controlled conditions: a 12-h photoperiod at 28°C (day) and 26°C (night), 70% relative humidity, and a photosynthetic photon flux density of 18,000 lux. At the appropriate growth stage, the seedlings were exposed to either chilling (4°C) or salt (200 mM NaCl) stress. Leaf tissue was sampled at 0, 1, 3, 6, 12, and 24 h post-treatment, immediately frozen in liquid nitrogen, and stored at – 80°C until RNA isolation.

2.7 Quantitative Real-Time PCR Analysis

Each harvested sample was ground to a fine powder in liquid nitrogen with a pre-chilled mortar and pestle. Total RNA was extracted with the TaKaRa MiniBEST Plant RNA Extraction Kit, and subsequently reverse-transcribed into complementary DNA (cDNA) using a reverse transcription kit. Gene-specific primers were designed for the target genes *OsDVL2* and *OsDVL4* (see Table S7). The relative transcript levels were normalized against the internal control gene *OsActin1* (LOC-Os03g61970) (Cui et al., 2020). The qRT-PCR protocol followed that described in a previous study (Zhang et al., 2024). The $2^{-\Delta\Delta Ct}$ method was used to analyze the expression of each gene (Livak & Schmittgen, 2001). Statistical analysis of the data was performed using WPS2024 software, while one-way ANOVA and graphical representation were conducted using GraphPad Prism8 software.

2.8 gcHaps and gcHap Diversity of *OsDVLs* in Modern and Local Rice Varieties

Shannon's evenness index (EH) was employed to evaluate gcHap diversity at the *OsDLH* loci across various rice population groups. For each gene, Nei's genetic diversity parameter (*INei*) was derived from the gcHap dataset, allowing for the assessment of pairwise genetic divergence and the overall extent of population differentiation (Castelli et al., 2019). To investigate the influence of recent breeding activities on gcHap variation in *OsDLH* genes, we assembled a comprehensive dataset comprising 3,010 rice accessions from the 3KRG panel. Of these, 732 were designated as indica landraces (LANs-Xian), 328 as japonica landraces (LANs-Geng), 358 as modern indica varieties (MVs-Xian), and 139 as modern japonica varieties (MVs-Geng). On April 7, 2025, gcHap information for *OsDLH* was downloaded from the RFGB database (available at <https://www.rmbreeding.cn/Index>). R scripts were then used to compute the frequency shifts of predominant gcHaps for each *OsDLH* gene by comparing MVs with LANs. Subsequently, the gcHap distributions observed in contemporary Xian and Geng cultivars were compared with those of their respective landrace counterparts. All graphical outputs were generated using GraphPad Prism 8 software.

2.9 Phenotypic Determination of Major gcHaps of *OsDVL* Genes

To begin with, a total of 15 agronomic traits were phenotypically characterized across the 3,010 rice accessions. The measured parameters included days to heading (DTH, in days), plant height (PH, cm), flag leaf length (FLL, cm) and width (FLW, cm), panicle number (PN, count) and length (PL, cm), culm number (CN, count) and length (CL, cm), grain length (GL, mm), grain width (GW, mm), grain length-to-width ratio (GLWR), thousand-grain weight (TGW, g), leaf rolling index (LRI, %), seedling height (SH, cm), and lemma length (LL, mm) (Wang et al., 2020). These phenotypic data were downloaded from the RFGB database. R scripts were then applied to identify the predominant gcHaps for each *OsDLH* gene and to conduct association tests linking these major haplotypes with the above-mentioned traits. One-way ANOVA was applied to assess statistical significance, and Tukey's multiple comparison test was used for pairwise comparisons among the predominant gcHaps. The final figures were compiled and arranged with Adobe Illustrator 2020 software.

2.10 Construction of gcHap Network for *OsDVL* Genes

The gcHaps for *OsDVL* genes were initially constructed with the pegas package in R (Paradis, 2010). Subsequently, a statistical parsimony algorithm was applied to generate gcHap networks for each *OsDVL* locus. This approach connects haplotypes that are most closely related, based on the minimum number of mutational steps (Templeton *et al.*, 1992). The resulting network diagrams were finalized and formatted using Adobe Illustrator 2023 software.

3. Results

3.1. Genome-wide identification and characterization of *OsDVL* genes

Five *OsDVL* genes were identified in the Nipponbare genome. To further reveal the characteristics of these genes, we further characterized the physicochemical parameters of these proteins, including length (in amino acids), molecular mass (Da), theoretical pI, instability index, aliphatic index, grand average of hydropathicity (GRAVY), and predicted subcellular localization (Table S1). The *OsDVL* proteins varied in length from 297 aa (*OsDVL2*) to 677 aa (*OsDVL1*), with molecular masses spanning 23,527.58 Da (*OsDVL1*) to 56,993.56 Da (*OsDVL5*) and an average of roughly 46,452.66 Da. Their pI values ranged between 5.10 (*OsDVL1*) and 5.24 (*OsDVL2*), with a mean of 5.14. All five proteins exhibited instability indices below 40, indicating overall structural stability. The aliphatic index reached a minimum of 15.82, suggesting a moderate degree of thermostability. Positive GRAVY scores were observed across all *OsDVL* members, implying their hydrophobic nature. Subcellular localization predictions indicated that these proteins are distributed among distinct organelles, which may reflect functional diversification within rice.

3.2. Evolutionary relationships of *DVL* genes

The chromosomal localization map showed that the five *OsDVL* genes were located on chromosomes 1, 3, and 5 of Nipponbare (Fig. 1a). To gain deeper insights into the evolutionary and structural features of the *OsDVL* genes, we constructed a phylogenetic tree incorporating DVL homologs from multiple plant species, including *Oryza sativa* L. (Nipponbare), *Triticum aestivum*, *Arabidopsis thaliana*, *Hordeum vulgare*, *Sorghum bicolor*, *Zea mays*, *Oryza sativa indica*, and *Setaria italica* (Table S2). Based on their evolutionary affinities, these sequences were clustered into four distinct subgroups, designated I–IV (Fig. 1b). This phylogenetic framework provides a useful basis for understanding the evolutionary trajectories of DVL genes across species, as well as their potential structural and functional divergence.

Moreover, Synteny analysis uncovered 15 pairs of DVL orthologs shared between Nipponbare and wheat (*Ta*), 13 pairs with indica rice (*Xi*), 11 pairs with *A. thaliana* (*At*), 7 pairs with barley (*Hv*), 6 pairs with foxtail millet (*Si*), and 5 pairs with maize (*Zm*) (Fig. 1c). These data suggested that *OsDVL* genes exhibit extensive collinearity with both monocotyledons and dicotyledons, demonstrating the conservation of DVL proteins during species evolution. Comparative genomic analyses across diverse species provide valuable insights into the evolutionary trajectories of the DVL gene family and their potentially conserved functions in plant biology (Griffiths *et al.*, 2006).

Motif composition analysis (Fig. 1d) indicated that the *OsDVL* proteins have relatively similar types and quantities of motifs. All *OsDVL* proteins contained motif 1 and motif 2 and also the DVL domain. In addition, the *OsDVL* proteins also showed relatively consistent gene lengths and exon numbers, which is showing general agreement with the conserved motif and domain features identified in the analysis

3.3. Analysis of *cis*-acting elements of *OsDVL* genes

Promoter sequence analysis of the five *OsDVL* genes revealed a total of 174 *cis*-regulatory elements, which were classified into 38 distinct types. Among these, 19 elements (10.9%) were putatively involved in stress responsiveness. Hormone-related motifs were predominantly represented by those responding to methyl jasmonate (MeJA) and abscisic acid (ABA), accounting for 60.75% and 24.05% of the hormone-responsive elements, respectively. Light-responsive elements constituted approximately 39.08% (68/174) of the total, whereas growth-related motifs were less frequent, representing only 4.5% (8/174). As illustrated in the heatmap of *cis*-element distribution across the *OsDVL* promoters (Fig. 2), motifs such as G-box and ABRE were detected in all members of the gene family. This widespread presence suggests that *OsDVL* genes may contribute to plant stress adaptation through these regulatory elements.

3.4. Expression profiles of *OsDVL* genes in different tissues and under abiotic stresses

The *OsDVL* genes exhibited pronounced tissue-specific expression patterns (Fig. 3a). Specifically, *OsDVL2* and *OsDVL4* were preferentially transcribed in buds and leaves, whereas *OsDVL3* showed

elevated transcript abundance in endosperm and seeds. The expression patterns of *OsDVL* genes also varied greatly under different abiotic stresses. Under drought stress, the expression of *OsDVL4* increased significantly in shoots after 1, 3, and 6 h, while that of *OsDVL3* exhibited nearly no change in shoots (Figure S1a). Upon high-salinity treatment, transcript levels of *OsDVL2* and *OsDVL4* were upregulated, whereas *OsDVL3* expression became undetectable in stems and leaves (Fig. 3b). Following ABA or MeJA application, *OsDVL4* showed increased expression in buds (Figure S1c,d), in contrast to *OsDVL1* and *OsDVL5*, which remained nearly silent. The *OsDVL* genes also displayed distinct responses to temperature stress. Heat exposure led to reduced expression of *OsDVL1*, *OsDVL3*, and *OsDVL5* in stems and leaves (Figure S1b), whereas under chilling conditions, *OsDVL5* exhibited very low or undetectable transcript levels in roots and buds (Fig. 3c).

3.5. Quantitative real-time PCR analysis of *OsDVL* genes

Guided by the cis-element analysis and RNA-seq transcriptomic data, we selected *OsDVL2* and *OsDVL4* for further characterization of their transcriptional responses to diverse environmental stimuli (Fig. 4). Both genes exhibited reduced transcript levels under cold stress. In response to salt stress, however, *OsDVL2* expression declined markedly, whereas *OsDVL4* showed an opposing pattern with increased transcript abundance. These results indicated that the two genes have specific expression patterns in response to different environmental signals, implying that they may play different roles in rice growth and development and its stress response mechanism.

3.6. Diversity of *OsDVL* alleles in different rice populations

To explore the breeding potential of *OsDVL* alleles, we analyzed the gcHap dataset derived from the 3KRG population. Using the gcHap dataset, we computed Shannon's evenness index (EH), gcHap frequency, and the number of major gcHaps (gcHapN, defined as those with frequencies $\geq 1\%$ in 3KRG) for the five *OsDVL* genes across four major rice subpopulations (Table S3). Averaged across the five genes, the gcHap count, gcHapN, and EH value were 27.2, 5.8, and 0.110, respectively, pointing to varying levels of genetic diversity among these loci. Notably, *OsDVL4* exhibited the highest EH (0.191), whereas *OsDVL3* showed the lowest (0.0005), consistent with their respective gcHap numbers—81 for *OsDVL4* versus only 20 for *OsDVL3* (Table S3).

Further analysis of genetic diversity across rice subpopulations uncovered marked variations in the average EH values of the five *OsDVL* genes among different groups. The mean EH values for the Xian, Geng, Aus, Bas, and mixed populations were 0.086, 0.073, 0.113, 0.111, and 0.219, respectively, with the mixed group showing the highest diversity. Substantial differences were also observed in the average gcHap counts and gcHapN across these populations: 17 and 3.6 (Xian), 10.6 and 1.8 (Geng), 6.4 and 1.6 (Aus), 6.4 and 1.0 (Bas), and 10.4 and 2.2 (mixed), respectively (Table S3).

To evaluate genetic divergence of *OsDVL* genes across major rice groups, we performed pairwise comparisons of the five *OsDVL* loci using their gcHap data and calculated the genetic diversity index (I_{Nei}). The results showed that *OsDVL1*, *OsDVL2*, and *OsDVL5* exhibited significant genetic differentiation (I_{Nei} < 0.35) in the Aus-Bas, XI-Adm, and GJ-Bas comparisons (Fig. 5, Table S4). In particular, *OsDVL1*

also showed strong genetic differentiation in the Aus-XI comparison. These results indicated that allelic variations at the *OsDVL* loci have significant contribution to the differentiation of rice populations and their adaptation to different environments.

3.7. Impacts of modern breeding on the gcHap diversity of *OsDVL* genes

Table 1
Comparison of Genetic Diversity of 5 *OsDVL* Genes in Xian between Local and Modern Varieties

Gene name	<i>Xian (indica)</i>							
	LANs		MVs		MVs - LANs		statistical significance	Artificial selection effect
	E_H	gcHapN	E_H	gcHapN	E_H	gcHapN		
OsDVL1	0.049	7	0.091	129	0.042	122	**	up
OsDVL2	0.118	7	0.127	5	0.009	-2	*	up
OsDVL3	0.103	12	0.004	11	-0.099	-1	ns	down
OsDVL4	0.135	28	0.169	25	0.034	-3	*	up
OsDVL5	0.018	8	0.043	7	0.025	-1	ns	down
Mean	0.085	12.4	0.087	35.4	0.002	23		

EH and gcHapN (major gcHapN) represent Shannon's evenness and the number of identified gcHaps (in varieties accounting for ≥ 1%), respectively.

Table 2
Comparison of Genetic Diversity of 5 OsDVL Genes in Geng between Local and Modern Varieties

Gene name	Geng(japonica)							
	LANs		MVs		MVs - LANs		statistical significance	Artificial selection effect
	EH	gcHapN	EH	gcHapN	EH	gcHapN		
OsDVL1	0.045	4	0.087	4	0.042	0	ns	down
OsDVL2	0.029	3	0.064	5	0.035	2	ns	down
OsDVL3	0.165	8	0.173	9	0.008	1	*	up
OsDVL4	0.140	12	0.224	9	0.084	-3	ns	down
OsDVL5	0.051	8	0.065	5	0.014	-3	ns	down
Mean	0.086	7.0	0.123	6.4	0.037	-0.6		

EH and gcHapN (major gcHapN) refer to Shannon's evenness and the number of identified gcHaps (in varieties with a proportion of $\geq 1\%$), respectively.

To assess the effect of modern breeding on gcHap diversity of OsDVL genes over recent decades, we compared the genetic diversity parameters between modern varieties (MVs) and landraces (LANs). The comparison involved 732 Xian landraces (LANs-Xian) versus 358 Xian modern varieties (MVs-Xian), and 328 Geng landraces (LANs-Geng) versus 139 Geng modern varieties (MVs-Geng) (Table 1). Within the Xian subgroup, the average EH across the five OsDVL genes was higher in MVs-Xian (0.087) than in LANs-Xian (0.002), implying that modern breeding has maintained or even enhanced genetic diversity at most OsDVL loci relative to landraces. However, the average value of gcHapN in MVs-Xian was 23 less than that in LANs-Xian. This increase in genetic diversity while decrease in gcHapN at multiple *OsDVL* loci reflect significant changes in the frequency of gcHaps at these loci during modern rice breeding. In fact, there were significant changes in the gcHapN at all *OsDVL* loci except for *OsDVL5* (Table 1).

In the Geng population, the average EH of the five *OsDVL* genes was 0.123 in MVs-Geng while 0.086 in LANs-Geng. Notably, a significant increase in genetic diversity for MVs-Geng was observed exclusively at the *OsDVL3* locus. The average gcHapN in MVs-Geng was 6.4, representing a reduction of 0.6 compared with that in LANs-Geng (Table 2). Furthermore, Hap1 was present across all *OsDVL* loci within the Xian subgroup, whereas Hap2 was found in all *OsDVL* loci of the Geng subgroup, suggesting that these two haplotypes may serve important functions in their respective populations (Table S5).

3.8. Favorable and unfavorable gcHaps of *OsDVL* Genes

The predominant gcHap at each *OsDVL* locus—defined as the one with the highest frequency—is generally regarded as having been favored by natural selection over the course of evolution. In contrast, the least frequent major gcHap within a population is more likely to represent a deleterious allele (Zeng

et al., 2023). Among the five *OsDVL* genes, one had only one haplotype, making it impossible to analyze its haplotype diversity. We therefore examined the phenotypic effects of favorable versus unfavorable gcHaps at the remaining four *OsDVL* loci across 15 agronomic traits (Figures S2–S5). Of the 60 trait-locus comparisons conducted, half (30) revealed significant differences between the two gcHap classes (Table S6). Among the four *OsDVL* genes, three showed TGW-associated phenotypic divergence between favorable and unfavorable gcHaps. *OsDVL2* exhibited the most pronounced effects, with significant gcHap-associated differences detected for all 15 traits examined.

3.9. Relationships between the major gcHaps of *OsDVL* genes and important agronomic traits

To further evaluate the functional relevance of the four *OsDVL* genes, we constructed gcHap networks for their dominant alleles across five rice populations and examined the associations between these alleles and four agronomic traits—panicle number per plant (PN), panicle length (PL), plant height (PH), and thousand-grain weight (TGW)—using the 3KRG panel (Fig. 6; Figure S6). Strong correlations ($P < 10^{-7}$) were detected in 12 of the 16 (4 genes $\times$ 4 traits) trait-locus combinations, with major alleles of multiple *OsDVL* genes showing significant associations with at least one trait. Specifically, *OsDVL1* harbored two major gcHaps, with Hap2 being relatively frequent in LANs-Xian. *OsDVL3* contained five major gcHaps, among which Hap1 was relatively abundant in LANs-Xian. *OsDVL4* also presented five major gcHaps, with Hap1 showing relatively high frequencies in both LANs-Xian and MVs-Xian. *OsDVL5* had three major gcHaps, where Hap1 was relatively frequent in both LANs-Geng and MVs-Geng.

4. Discussion

Based on the genomic sequence of the Nipponbare cultivar, we identified five *OsDVL* genes in this study. Chromosomal localization showed that these genes reside on chromosomes 1, 3, and 5. Phylogenetic analyses across diverse plant species further suggested both evolutionary conservation and possible functional divergence among DVL homologs.

Transcriptome analysis provides reliable information about stress-induced genes and the molecular mechanisms of their actions (Cao et al., 2022, Gu et al., 2023). For instance, GhMDVL17 transcript levels in cotton rose at 6 h post-inoculation but declined by 12 h, suggesting a possible role in the early defense response against *Verticillium dahliae* (Jiao et al., 2024). Tissue-specific expression profiling of *OsDVL* genes revealed their preferential accumulation in floral organs, leaves, roots, and stems, implying functional relevance to rice growth and development. In addition, both RNA-seq (Fig. 3) and qPCR (Fig. 4) analyses revealed that the *OsDVL* genes responded differentially to various environmental cues and stress treatments, including ABA, MeJA, drought, heat, cold, and salinity. *OsDVL4* exhibited relatively high transcript abundance under multiple stress conditions, whereas *OsDVL2* displayed a distinct expression profile specifically under drought and salt stresses. Notably, *OsDVL2* and *OsDVL4* were both upregulated in response to hormone application and showed rapid induction upon salt exposure. Although the

expression profiles of these genes are tightly linked to their biological roles, the underlying molecular mechanisms remain to be fully elucidated(Wong *et al.*, 2014).

Analysis of the 3KRG gcHap dataset revealed substantial genetic diversity of the OsDVL gene family across different rice populations, with the admixed group displaying the highest level(Wang *et al.*, 2020). Genetic differentiation analyses further indicated that OsDVL2 and OsDVL4 harbor significant inter-population variation, which may contribute to rice adaptability and population structure. Comparison between modern varieties and landraces showed that Xian modern cultivars maintain high genetic diversity but possess fewer gcHaps than their landrace counterparts, a pattern potentially attributable to recombination events. In the Geng subgroup, modern varieties exhibited reduced diversity specifically at the OsDVL3 locus, whereas the predominant gcHap (Hap1) remained consistently frequent across all loci, suggesting its stability across different populations. Collectively, these findings highlight the profound impact of breeding on rice genetic diversity and underscore the importance of integrating allelic diversity and environmental adaptability into future breeding strategies(Zhang *et al.*, 2022).

Haplotype breeding assists in improvement of traits through introgression of superior haplotypes(Sahoo *et al.*, 2025). For instance, Singh *et al* reveals that the performance of lines possessing superior haplotypes is significantly higher ($p \leq 0.05$) as measured by single plant yield (SPY), for the OsGSK1-H4, OsDSR2-H3, OsDIL1-H22, OsDREB1C-H3, ASR3-H88, DSM3-H4 and ZFP182-H4 genes as compared to lines without the superior haplotypes(Singh *et al.*, 2024). Moreover, Singh *et al* reveals that introgressing superior haplotypes from the OsIRO2-H3 gene into an elite rice cultivar led to a yield advantage of 25.0–27.3% over recurrent parents(Singh *et al.*, 2025). By evaluating the phenotypic performance of favorable versus unfavorable gcHaps at five OsDVL loci across 15 agronomic traits, we found that OsDVL3 exhibited the largest phenotypic divergence between its two haplotype classes, implying a substantial effect of this gene on multiple traits. With the exception of OsDVL1 and OsDVL4, the remaining OsDVL genes each possessed specific gcHaps that were significantly associated with traits including TGW (thousand-grain weight) and DTH (days to heading), underscoring the importance of incorporating allelic diversity into rice breeding programs. An extended association study further revealed that the gcHaps representing the major alleles of four OsDVL genes were significantly correlated with several agronomic traits—including PN, PL, PH, and TGW—across five rice populations, with highly significant correlations observed in approximately 75% of the tested combinations. In addition, the frequencies of certain favorable gcHaps at OsDVL loci varied among rice populations, suggesting that these variations may represent adaptive genetic responses to distinct ecological conditions(Zeng *et al.*, 2023).

Our results highlight that the major gcHaps at the *OsDVL* loci have significant impacts on some key agronomic traits, and the frequencies of these alleles vary significantly among different rice populations and subspecies. These findings indicate the importance of integrating allelic diversity and environmental adaptability when developing rice varieties, and to select and utilize favorable alleles to enhance the adaptability and yield of rice.

Declarations

Ethics Approval and Consent to Participate

Not applicable.

Consent for Publication

Not applicable.

Data Availability

Generated Statement: The original contributions presented in the study are included in the article/supplementary material, further inquiries can be directed to the corresponding author/s.

Competing Interests

The authors declare no competing interests.

Funding

This research was funded by the Science and Technology of Innovative research program of Anhui Province (202423m10050002), the Major scientific and technological innovation research tasks of Bio-breeding Laboratory of Anhui Province(2025SWYZ0110), the Key Research and Development Program of Anhui Province (2023n06020013), This research was funded by the National Key Research and Development Program of China (2023YFD1200900), the National Natural Science Foundation of China [grant numbers U21A20214].

Author Contributions

Y.S. and X.L. designed the experiments, X.L., Y.Z., M.Z., W.L., C.W., M.Z., C.T. and S.Y. performed the experiments, collated and statistically analyzed the data, X.L., Y.Z., M.Z., W.L., C.W. and M.Z. constructed the graphics, X.L., Y.Z. and M.Z. wrote the paper, Y.S. reviewed the paper. All authors read and approved the final manuscript.

Acknowledgments

Not applicable.

References

1. Bailey, T. L., Johnson, J., Grant, C. E. and Noble, W. S. (2015) The MEME Suite. *Nucleic Acids Research*, **43**, W39-W49.
2. Beena, R., Sowmiya, S., Visakh, R. L., Shelvy, S., Sasmita, B. and Sah, R. P. (2025) Unravelling genetic mechanisms for heat tolerance in rice landraces of Kerala, India. *Euphytica*, **221**.

3. Cao, Y. Y., Shan, T. Y., Fang, H., Sun, K. T., Shi, W., Tang, B., et al. (2022) Genome-wide analysis reveals the spatiotemporal expression patterns of SOS3 genes in the maize B73 genome in response to salt stress. *Bmc Genomics*, **23**.
4. Castelli, M., Cattaneo, G., Manzoni, L. and Vanneschi, L. (2019) A distance between populations for *n*-points crossover in genetic algorithms. *Swarm and Evolutionary Computation*, **44**, 636-645.
5. Chen, C. J., Chen, H., Zhang, Y., Thomas, H. R., Frank, M. H., He, Y. H., et al. (2020) TBtools: An Integrative Toolkit Developed for Interactive Analyses of Big Biological Data. *Molecular Plant*, **13**, 1194-1202.
6. Cheng, M. X., Yuan, H. R., Wang, R. H., Wang, W., Zhang, L. C., Fan, F. F., et al. (2023) Identification and characterization of BES1 genes involved in grain size development of *Oryza sativa* L. *International Journal of Biological Macromolecules*, **253**.
7. Chou, K. C. and Shen, H. B. (2008) Cell-PLoc: a package of Web servers for predicting subcellular localization of proteins in various organisms. *Nature Protocols*, **3**, 153-162.
8. Cui, Y., Lu, S., Li, Z., Cheng, J., Hu, P., Zhu, T., et al. (2020) CYCLIC NUCLEOTIDE-GATED ION CHANNELS 14 and 16 Promote Tolerance to Heat and Chilling in Rice. *Plant physiology*, **183**, 1794-1808.
9. Griffiths, J., Murase, K., Rieu, I., Zentella, R., Zhang, Z. L., Powers, S. J., et al. (2006) Genetic characterization and functional analysis of the GID1 gibberellin receptors in *Arabidopsis*. *Plant Cell*, **18**, 3399-3414.
10. Gu, S. C., Abid, M., Bai, D. F., Chen, C., Sun, L. M., Qi, X. J., et al. (2023) Transcriptome-Wide Identification and Functional Characterization of CIPK Gene Family Members in *Actinidia valvata* under Salt Stress. *International Journal of Molecular Sciences*, **24**.
11. Guo, P., Yoshimura, A., Ishikawa, N., Yamaguchi, T., Guo, Y. and Tsukaya, H. (2015) Comparative analysis of the RTFL peptide family on the control of plant organogenesis. *Journal of Plant Research*, **128**, 497-510.
12. Ikeuchi, M., Yamaguchi, T., Kazama, T., Ito, T., Horiguchi, G. and Tsukaya, H. (2011) ROTUNDIFOLIA4 Regulates Cell Proliferation Along the Body Axis in Arabidopsis Shoot. *Plant and Cell Physiology*, **52**, 59-69.
13. Jiao, Y., Zhao, F. X., Geng, S. W., Li, S. M., Su, Z. L., Chen, Q. J., et al. (2024) Genome-Wide and Expression Pattern Analysis of the DVL Gene Family Reveals GhM_A05G1032 Is Involved in Fuzz Development in *G. hirsutum*. *International Journal of Molecular Sciences*, **25**.
14. Jo, S., Jang, S. G., Lee, S. B., Lee, J. Y., Cho, J. H., Kang, J. W., et al. (2025) Analysis of quantitative trait loci and candidate gene exploration associated with cold tolerance in rice (*Oryza sativa* L.) during the seedling stage. *Frontiers in Plant Science*, **15**.
15. Kumar, S., Stecher, G. and Tamura, K. (2016) MEGA7: Molecular Evolutionary Genetics Analysis Version 7.0 for Bigger Datasets. *Molecular Biology and Evolution*, **33**, 1870-1874.
16. Lescot, M., Déhais, P., Thijs, G., Marchal, K., Moreau, Y., Van de Peer, Y., et al. (2002) PlantCARE, a database of plant *cis*-acting regulatory elements and a portal to tools for *in silico*

- analysis of promoter sequences. *Nucleic Acids Research*, **30**, 325-327.
17. Letunic, I. and Bork, P. (2016) Interactive tree of life (iTOL) v3: an online tool for the display and annotation of phylogenetic and other trees. *Nucleic Acids Research*, **44**, W242-W245.
 18. Li, Z. and Rutger, J. N. (2000) Geographic distribution and multilocus organization of isozyme variation of rice (*Oryza sativa* L.). *Theoretical and Applied Genetics*, **101**, 379-387.
 19. Livak, K. J. and Schmittgen, T. D. (2001) Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta C_T}$ method. *Methods*, **25**, 402-408.
 20. Paradis, E. (2010) pegas: an R package for population genetics with an integrated-modular approach. *Bioinformatics*, **26**, 419-420.
 21. Rui, C., Peng, F. J., Fan, Y. P., Zhang, Y. X., Zhang, Z. G., Xu, N., et al. (2022) Genome-wide expression analysis of *carboxylesterase* (*CXE*) gene family implies *GBCXE49* functional responding to alkaline stress in cotton. *Bmc Plant Biology*, **22**.
 22. Sahoo, R. K., Swain, N., Selvaraj, S., Nayak, G., Sarkar, S., Singh, N. R., et al. (2025) Haplotypes Differences in *Growth Regulating Factor 4* (*GRF4*) for Yield and Biomass Traits in Rice (*Oryza sativa* L.). *Tropical Plant Biology*, **18**.
 23. Singh, P., Pahi, B., Sundaram, K. T., Gnanapragasam, N., Habde, S. V., Kumar, N. V., et al. (2025) Superior haplotypes of genes associated with higher grain yield under reproductive stage drought stress in rice. *Journal of Experimental Botany*.
 24. Singh, P., Sundaram, K. T., Vinukonda, V. P., Venkateshwarlu, C., Paul, P. J., Pahi, B., et al. (2024) Superior haplotypes of key drought-responsive genes reveal opportunities for the development of climate-resilient rice varieties. *Communications Biology*, **7**.
 25. Templeton, A. R., Crandall, K. A. and Sing, C. F. (1992) A cladistic analysis of phenotypic associations with haplotypes inferred from restriction endonuclease mapping and DNA sequence data. III. Cladogram estimation. *Genetics*, **132**, 619-633.
 26. Valdivia, E. R., Chevalier, D., Sampedro, J., Taylor, I., Niederhuth, C. E. and Walker, J. C. (2012) *DVL* genes play a role in the coordination of socket cell recruitment and differentiation. *Journal of Experimental Botany*, **63**, 1405-1412.
 27. Wang, C. C., Yu, H., Huang, J., Wang, W. S., Faruquee, M., Zhang, F., et al. (2020) Towards a deeper haplotype mining of complex traits in rice with RFGB v2.0. *Plant Biotechnology Journal*, **18**, 14-16.
 28. Wang, C. Y., Wang, T. T., Liu, M., Zhang, S. Z. and Wu, C. A. (2024) Small peptide SiDVL/RTFLs from foxtail millet inhibit root growth through repressing auxin signaling in transgenic *Arabidopsis*. *Plant Cell Reports*, **43**.
 29. Wang, S. Y., Liu, Y. J., Asseng, S., Harrison, M. T., Tang, L., Liu, B., et al. (2025) Rice yield stability and its determinants across different rice-cropping systems in China. *Agricultural and Forest Meteorology*, **364**.
 30. Wang, W. S., Mauleon, R., Hu, Z. Q., Chebotarov, D., Tai, S. S., Wu, Z. C., et al. (2018) Genomic variation in 3,010 diverse accessions of Asian cultivated rice. *Nature*, **557**, 43+.

31. Wang, Y. P., Tang, H. B., DeBarry, J. D., Tan, X., Li, J. P., Wang, X. Y., et al. (2012) *MCSscanX*: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. *Nucleic Acids Research*, **40**.
32. Wen, J. Q., Lease, K. A. and Walker, J. C. (2004) DVL, a novel class of small polypeptides:: overexpression alters *Arabidopsis* development. *Plant Journal*, **37**, 668-677.
33. Wong, D. C. J., Sweetman, C. and Ford, C. M. (2014) Annotation of gene function in citrus using gene expression information and co-expression networks. *Bmc Plant Biology*, **14**.
34. Yang, L. Y., Hu, X. C., Ren, M. F., Ma, F., Fu, J. and Cui, H. C. (2024) Stem-cell-expressed DEVIL-like small peptides maintain root growth under abiotic stress via abscisic acid signaling. *Plant Physiology*, **194**, 2372-2386.
35. Yu, S. B., Xu, W. J., Vijayakumar, C. H. M., Ali, J., Fu, B. Y., Xu, J. L., et al. (2003) Molecular diversity and multilocus organization of the parental lines used in the International Rice Molecular Breeding Program. *Theoretical and Applied Genetics*, **108**, 131-140.
36. Yu, Y. M., Zhang, H., Long, Y. P., Shu, Y. and Zhai, J. X. (2022) Plant Public RNA-seq Database: a comprehensive online database for expression analysis of ~45 000 plant public RNA-Seq libraries. *Plant Biotechnology Journal*, **20**, 806-808.
37. Zeng, W., Li, H., Zhang, F. L., Wang, X. C., Rehman, S., Huang, S. J., et al. (2023) Functional characterization and allelic mining of OsGLR genes for potential uses in rice improvement. *Frontiers in Plant Science*, **14**.
38. Zhang, J. G., Meng, D. Z., Li, J. F., Bao, Y. L., Yu, P., Dou, G. H., et al. (2024) Analysis of the Rice Raffinose Synthase (OsRS) Gene Family and Haplotype Diversity. *International Journal of Molecular Sciences*, **25**.
39. Zhang, Y., He, Q. Q., Zhou, X. X., Zheng, S. M., Wang, Y. W., Li, P. J., et al. (2022) Genetic diversity and population structure of 93 rice cultivars (lines) (*Oryza sativa* Xian group) in Qinba in China by 3 types of genetic markers. *Bmc Genomics*, **23**.
40. Zhao, W. J., Wen, J. L., Zhao, J., Liu, L. L., Wang, M., Huang, M. H., et al. (2025) E3 Ubiquitin Ligase OsRFI2 Regulates Salinity Tolerance by Targeting Ascorbate Peroxidase OsAPX8 for its Degradation in Rice. *Rice*, **18**.

Figures

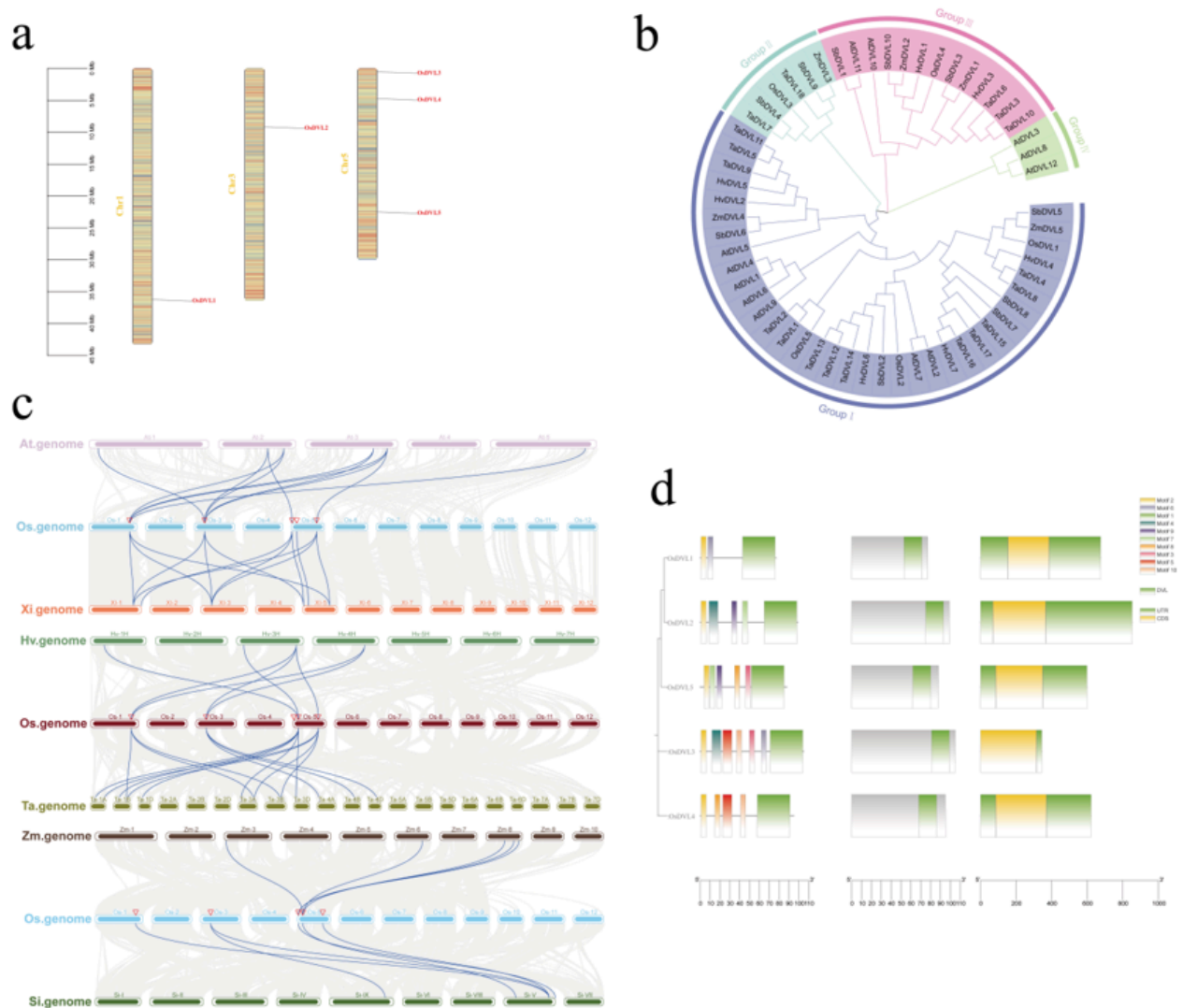


Figure 1.

Figure 1

(a) Chromosomal localization of OsDLH genes.(b) Phylogenetic tree of OsDLH genes from Nipponbare (Os), Triticum aestivum (Ta), Arabidopsis thaliana (At), Hordeum vulgare (Hv), Sorghum bicolor (Sb) and Zea mays (Zm).(c) Collinearity relationships between OsDVL genes and genes of species At, Xi, Hv, Ta, Zm, Setaria italica (Si).(d) Prediction of motifs, domains and gene structures of OsDLH genes.

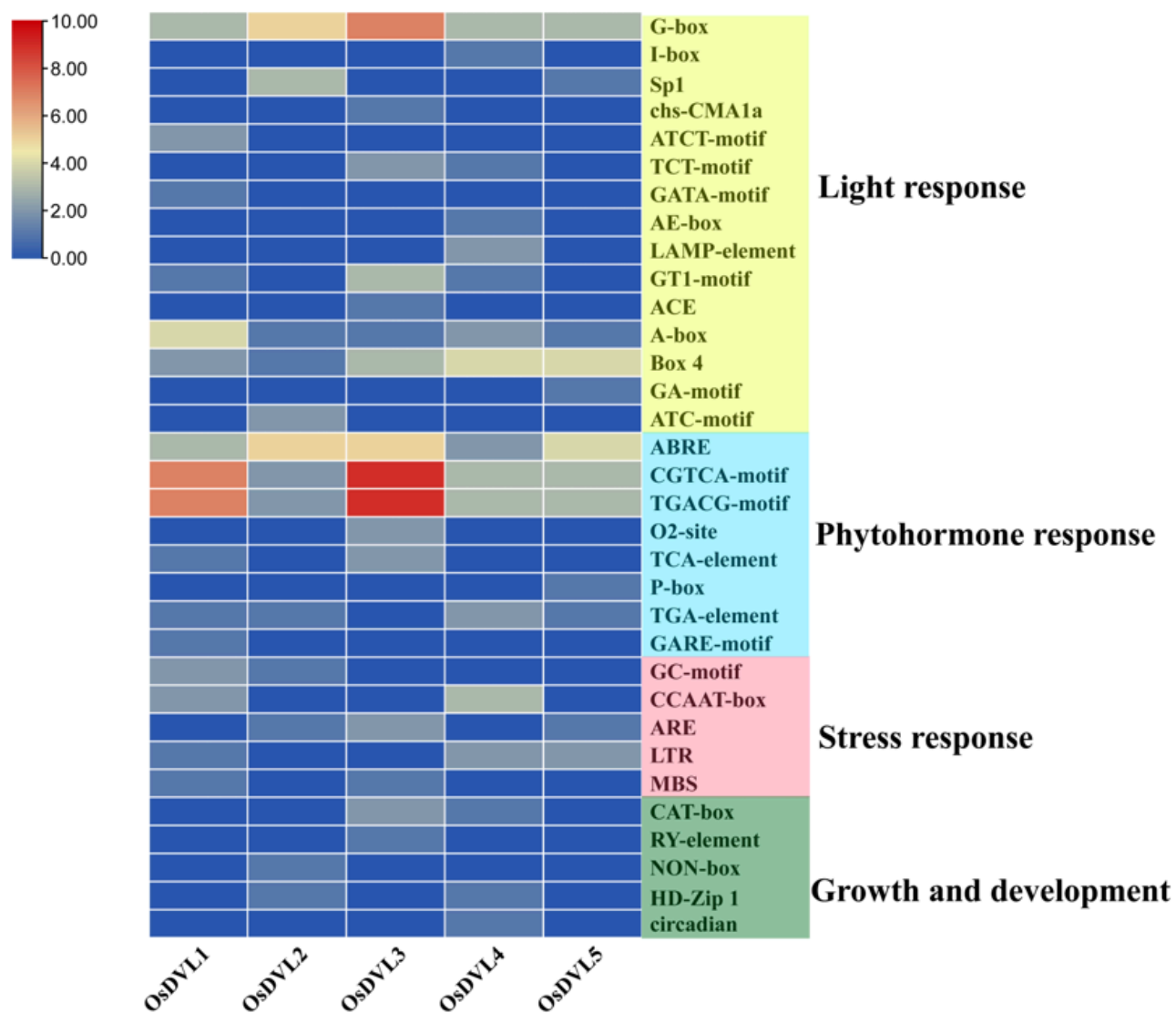


Figure 2

Heat - map analysis of cis - acting elements in the promoter regions of OsDVL genes. In the heat - map, the values represent the numbers of different cis - acting elements. The darker the color, the higher the number.

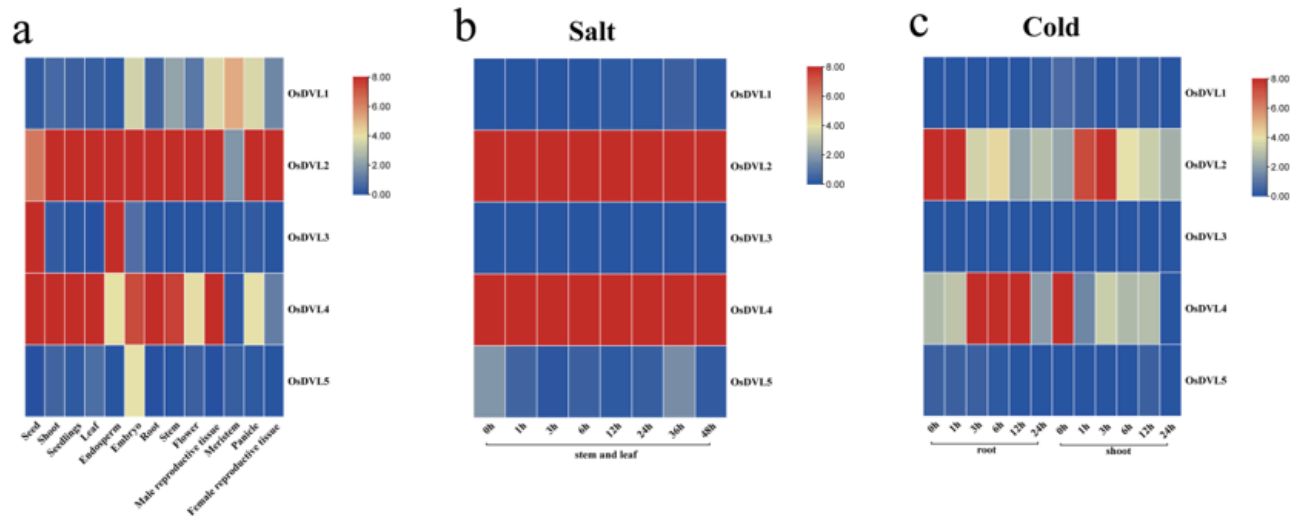


Figure 3

Expression analysis of OsDVL genes. Color coding indicates changes in gene expression. Red represents high expression and blue represents low expression. (a) Expression of OsDVL genes in leaves, roots, seedlings, stems, flowers, embryos, buds, meristems, male reproductive tissues, female reproductive tissues, panicles, and seeds. (b) Expression levels of OsDVL genes in rice stems and leaves after high-salt stress. (c) Expression levels of OsDVL genes in roots and buds after cold stress.

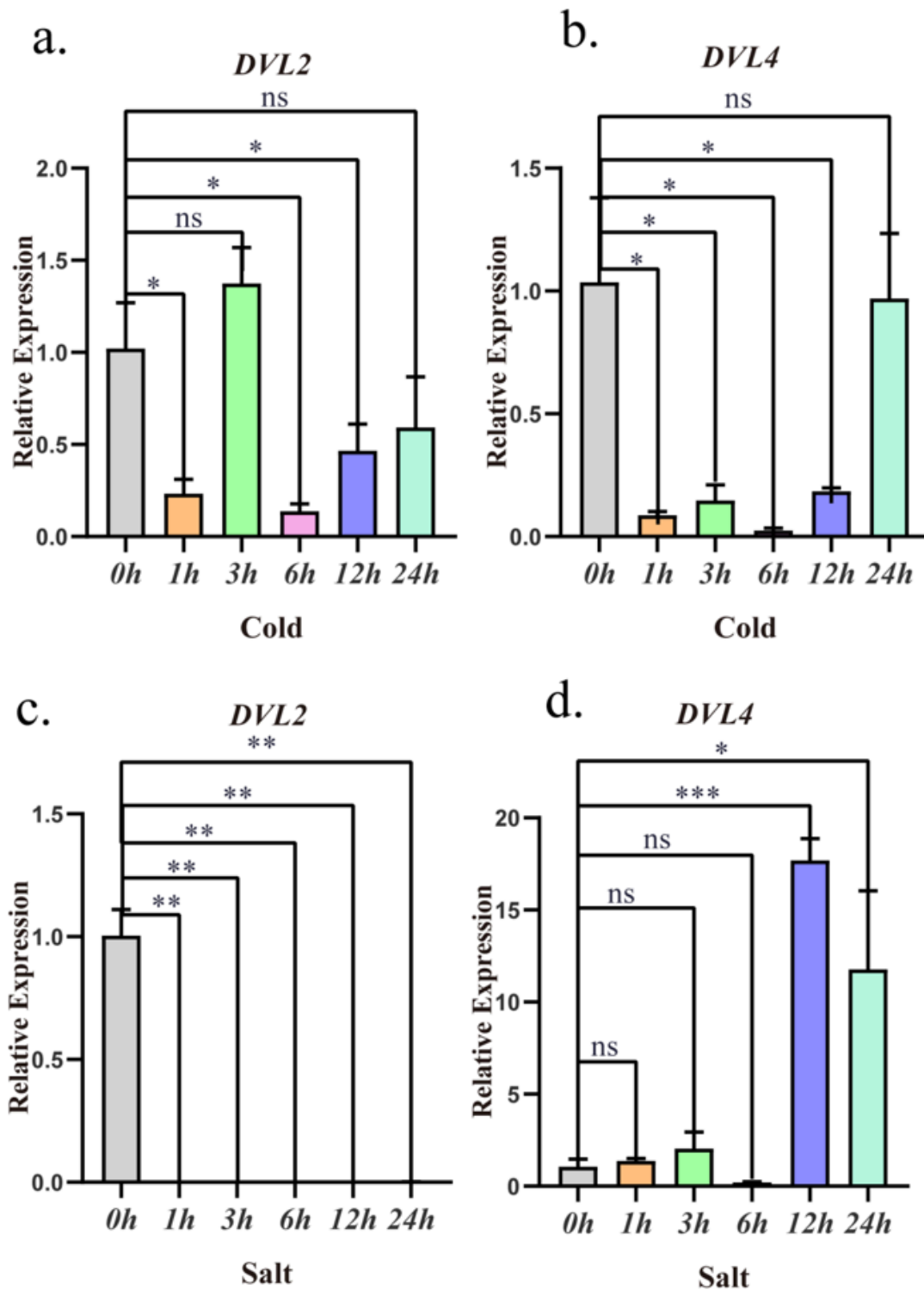


Figure 4

(a) Analysis of the expression level of DVL2 under low - temperature stress;(b) Analysis of the expression level of DVL4 under low - temperature stress;(c) Analysis of the expression level of DVL2 under salt stress;(d) Analysis of the expression level of DVL4 under salt stress

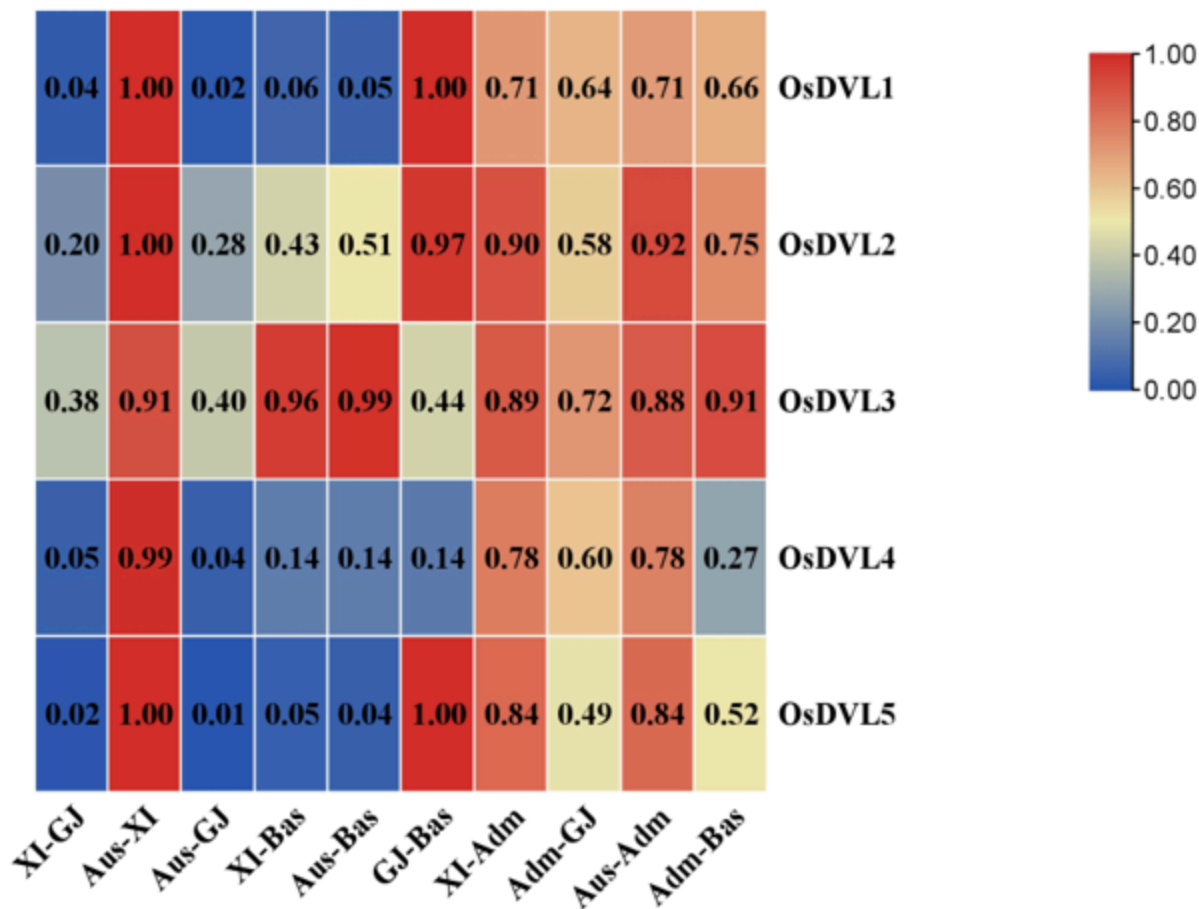


Figure 5

Nei's genetic identity (INei) of rice OsDVL genes among all pairwise populations calculated from gcHap data.

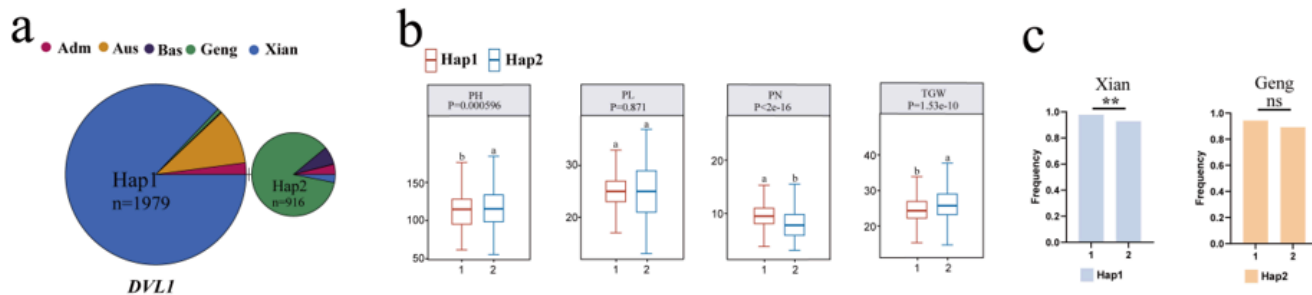


Figure 6

Haplotype network of the OsDVL1 gene and four related agronomic traits in 3KRG. Letters indicate the differences among haplotypes evaluated by two-way ANOVA, where different letters on the box plots indicate statistically significant differences at $p < 0.05$ according to the Duncan's multiple range test. The bar chart on the right shows the frequency differences of the main gcHaps between the landraces (LAN) and modern varieties (MV) of Xian (indica) and Geng (japonica). The chi-square test was used to

determine the significant differences in the proportions of gcHaps among different populations at $p < 0.0001$.

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryFigure.docx](#)
- [SupplementaryTable.xlsx](#)