BMC Plant Biology

RESEARCH

Open Access

# Comprehensive characterization and diversity analysis of the *HIS1* gene family in rice subpopulations for herbicide resistance

May Htet Aung[1†], Sang-Ho Chu[1†], Bhagwat Nawade[1] and Yong-Jin Park[1*]

## Abstract

**Background** Understanding the genetic diversity and functional roles of key resistance genes is crucial for developing sustainable weed management strategies in rice cultivation. *HIS1 (HPPD INHIBITOR SENSITIVE 1)* confers broad-spectrum resistance to β-triketone herbicides in rice. However, despite its importance, the family of *HIS1-like* genes (*HSLs*) in rice remains largely uncharacterized.

**Results** Here, we identified 25 *HIS1* gene family members across four rice subpopulations, including 13 in Nipponbare, 4 in Minghui 63, 6 in Zhenshan 97, and 4 in Nagina-22. Phylogenetic analysis grouped these members into seven distinct subfamilies (*HIS1*, *HSL1* ~ *HSL6*). While *HIS1*, *HSL2*, and *HSL3* were present across all subpopulations, others exhibited subpopulation-specific presence/absence, underscoring the influence of evolutionary pressures on the *HIS1* gene family. Haplotype analysis of family genes within a collection of 475 rice accessions revealed natural genetic variation for only three genes: *HIS1*, *HSL2*, and *HSL3*, with *japonica* accessions exhibiting high conservation across all genes, while *indica* accessions displayed diversity, forming 25 haplotypes for *HSL3* and four for *HSL2*. However, these haplotypes did not strongly correlate with Benzobicyclon (BBC) resistance, suggesting that natural variations of these genes are not primary determinants of herbicide response. Expression profiling under BBC treatment revealed ecotype-specific regulation, with notable upregulation of *HSL5* in BBC-resistant accessions.

**Conclusions** Our analysis identified tandem duplication as a major driver for the expansion and diversification of *HIS1* family members on chromosome 6 in *japonica* rice. The upregulation of *HSL5* in herbicide response points to its potential role in mediating BBC resistance in *indica* ecotypes. Employing an ecotype-specific reference genome could further enhance insights into herbicide resistance mechanisms. This comprehensive analysis enhances understanding of the evolutionary patterns and functions of *HIS1* family genes in rice, offering valuable knowledge for future herbicide resistance breeding programs.

**Clinical trial number** Not applicable.

**Keywords** Rice, Herbicides, HPPD-inhibitors, Gene family, Diversity, Haplotypes

[†]May Htet Aung and Sang-Ho Chu contributed equally to this work.

*Correspondence:
Yong-Jin Park
yjpark@kongju.ac.kr
[1]Department of Plant Resources, Kongju National University, Yesan 32439, Republic of Korea

## Background

Rice (*Oryza sativa L.*) is an important staple crop, sustaining billions of people globally and serving as a vital food source. The genetic diversity within *O. sativa* has been shaped over millennia through domestication, natural adaptation, and breeding efforts to meet diverse ecological and agricultural needs [1]. Recently, a reanalysis of population structure across the 3,000 Rice Genomes (3 K-RG) dataset, capturing genetic variations on a pan-genome scale, divided this dataset into 15 distinct subpopulations, including *GJ*-temp: Nipponbare, *indica* – adm: Minghui 63 *(*MH63*)*, *indica* – 1 A: Zhenshan 97 (ZS97), and *circum-Aus1*: Nagina-22 (N22) [2–4]. The ZS97 and MH63 are prominent hybrid rice cultivars extensively grown in Southeast Asia and China, valued for their higher yield, disease resistance, adaptability, and good eating quality. These varieties are not only essential in rice breeding programs but also serve as model systems for genomic research [5]. While, the *aus* Nagina-22, originating from India, is recognized for its resilience to environmental stressors such as drought and heat [6]. The diversity within these subpopulations underscores the adaptive potential of *O. sativa*, enabling the identification of valuable genes for crop breeding and serving as a key focus in efforts to enhance crop resilience and productivity [7].

Weeds pose multiple challenges for the agriculture industry, impacting crop quality and yield, imposing an annual economic burden of over US$100 billion, in addition to the loss of almost 200 million metric tons of grains globally [8]. Particularly, weedy rice (*Oryza spp.*) poses significant challenges as a weed, often distinguished by its red-pericarp grains [9]. Currently, around 350 species from various genera, including seven species within the *Oryza* genus, are classified as weedy rice. Infestations of weedy rice have led to significant yield losses, affecting over 50% of rice fields globally [9].

β-Triketone herbicides (bTHs), which inhibit 4-hydroxyphenylpyruvate dioxygenase (HPPD), are effective for pre- and post-emergence weed control [10]. These herbicides primarily target the HPPD enzyme, which is essential in synthesizing plastoquinone and tocopherol in plants. By inhibiting HPPD, bTHs induce bleaching symptoms in growing plant tissue, which progresses to chlorosis, ultimately resulting in death [10]. Commonly used bTHs, such as tefuryltrione (TFT) and Benzobicyclon (BBC), are particularly effective against resistant weedy rice populations that evade control by herbicides like quizalofop and imidazolinone [11]. HPPD-inhibiting herbicides offer several advantages, including broad-spectrum efficacy against dicot weeds, inherent crop selectivity, low toxicity to mammals, low application rates, and flexibility in application timing, making them valuable tools in integrated weed management [12].

*HPPD Inhibitor Sensitive 1* (*HIS1*) gene has been identified as a key determinant conferring resistance to β-triketone herbicides in rice [13]. This gene encodes an oxidase with a pivotal role in detoxifying bTHs. The detoxification process involves hydroxylation of benzobicyclon hydrolysate, converting it to less toxic forms, thereby mitigating herbicidal effects [13]. Research has shown that the *HIS1* gene in *japonica* rice confers resistance to HPPD-inhibiting herbicides. Conversely, *indica* rice accessions with a 28 bp deletion in *HIS1* are susceptible to these inhibitors [13]. Analysis of the *HIS1* gene across 631 *indica* varieties revealed a mutation, changing valine to glycine in the predicted substrate binding site of the *HIS1* gene [14]. This amino acid change potentially alters the substrate affinity, affecting the herbicide detoxification capability of the HIS1 enzyme. Additionally, responses have been recorded in Korean cultivated rice accessions, where *japonica* accessions exhibit resistance to TFT, and *aus* accessions show TFT susceptibility, regardless of the presence of functional SNPs [15]. The bTH sensitive *Arabidopsis thaliana* model plant transformed with the *OsHIS1* gene conferred resistance to multiple triketone herbicides [13]. Moreover, overexpression of Os*HIS1* in soybean has shown tolerance to triketone herbicides, such as temborione, sulcotrione, and mesotrione, highlighting its potential application in enhancing herbicide tolerance in crops [16, 17].

Phylogenetic analysis has identified homologs to *HIS1*, revealing both conservation and diversification within the *HIS1* gene family, which includes multiple HIS1-like (HSL) proteins [13]. While these HSL proteins share structural similarities with *HIS1*, their catalytic efficiency in detoxifying β-THs is limited [18]. However, a comprehensive understanding of the *HIS1* gene family for evolutionary dynamics and diversity across different rice subpopulations and ecotypes remains limited, highlighting the importance of further research to reveal the ecological and agricultural implications of these genes.

In light of these insights, our research focuses on identifying the *HIS1* gene family across four rice subpopulations and conducting comprehensive in silico assessments, including genomic distribution, motif composition, conserved domains, and collinearity. Additionally, we investigated the functional haplotypes of *HIS1* gene family members across the Korean Rice collection (KRICE). To further explore their potential role in herbicide resistance, we assessed their expression profiles following HPPD-inhibitor treatment. This comprehensive study aims to elucidate the genetic diversity and evolutionary dynamics of the *HIS1* gene family across rice subpopulations, with a particular emphasis on their roles in influencing herbicide resistance.

## Methods

### Mining of *HIS1* family members in rice

The protein sequence of the *HIS1* gene was retrieved from the Rice Annotation Project Database (RAP-DB, https://rapdb.dna.affrc.go.jp/index.html) of the Nipponbare reference genome, and functional domains were identified using InterProScan (http://www.ebi.ac.uk/interpro/interproscan.html). Hidden Markov Models (HMMs) corresponding to the identified domains, 2OG-FeII_Oxy (PF03171) [19] and DIOX_N (PF14226) [20], were acquired from Pfam (http://pfam.xfam.org/). Additionally, the complete amino acid sequence assembly and gene annotations for rice were obtained from EnsemblPlants (http://plants.ensembl.org/index.html).

Candidate HIS1 proteins containing the specified domains were identified through Simple HMM search using TBtools v2.069 [21]. Extracted domain sequences were used to construct a rice-specific HMM. Proteins with an E-value < 0.01 were considered potential candidates for 2OG-FeII_Oxy and DIOX_N domains. Their presence was further confirmed using Pfam and InterPro databases, and proteins passing this validation were classified as *HIS1* family members.

Orthologous *HIS1* genes in other domesticated rice ecotypes—*indica* (MH63 and ZS97), and *aus* (N22::IRGC 19379-1)—were identified using Blastp searches against their reference genomes available at the Rice Gene Index (https://riceome.hzau.edu.cn/) [2]. Additionally, the NCBI Conserved Domain Database (https://www.ncbi.nlm.nih.gov/Structure/cdd/cdd.shtml) was used to examine the conserved domains of the discovered HIS1 proteins, and the domain structure diagram was drawn using TBtools v2.069 [21].

### Phylogenetic analysis

OrthoFinder (v2.5.4) [22] was utilized to identify orthologs in order to examine the evolutionary links of the *HIS1* gene family among different rice species as well as monocot and dicot crops. DIAMOND (v0.9.25) was employed for rapid BLAST searches using peptide sequences obtained from EnsemblPlants and the Rice Gene Index [2].

The full-length HIS1 protein sequences were aligned in ClustalW with the default parameters (https://www.genome.jp/tools-bin/clustalw). MEGA (v11.0) was employed to build the phylogenetic tree, applying 1,000 bootstrap repetitions, pairwise deletion, p-distance, and neighbor-joining (NJ) methods [23]. The iTOL (https://itol.embl.de/) program was used to show the generated unrooted phylogenetic tree [24].

### Chromosomal localization

To map the chromosomal locations of *HIS1* gene family members, three files were prepared: (1) chromosome length data, (2) gene position data (including gene ID, chromosome ID, start and end positions), and (3) a gene color map. These files were loaded into TBtools v2.069 [21] for visualization. Genome and GFF3 annotation files for four rice subpopulations were sourced from the Rice Gene Index database [2].

### Analysis of conserved motifs and intron/exon structures

The Rice Gene Index database was used to retrieve the entire gene sequences, protein sequences, and coding DNA sequences (CDS) of the *HIS1* gene family members from the four rice subpopulations [2]. The GSDS web server (https://gsds.gao-lab.org/) was used to analyze exon and intron configurations [25]. To investigate conserved motifs within the protein sequences, we utilized the MEME suite (https://meme-suite.org/meme/tools/meme, Version 5.5.5), applying classical mode, with a motif width range from 6 to 50 amino acids.

### Predicting physicochemical properties and subcellular locations

The physicochemical characteristics of *HIS1* family proteins, such as the grand average of hydropathicity (GRAVY), aliphatic index, theoretical isoelectric point (pI), molecular weight (Mw), and instability index, were predicted employing the ProtParam tool (https://web.expasy.org/protparam/) [26]. Subcellular localization predictions were performed using DeepLoc 2.0 (https://services.healthtech.dtu.dk/services/DeepLoc-2.0/) [27].

### Synteny, gene duplication, and Ka/Ks analysis

The "one-step MCScanX" program in TBtools v2.069 [21] was used to analyze the synteny and gene duplication patterns of the *HIS1* gene family across four rice subpopulations. Duplication types of each gene were classified using 'Duplicate_gene_classifier' function [28]. The syntenic relationships between genes were visualized with the "Multiple Synteny View" tool in TBtools, and microcollinearity was analyzed and visualized using the Rice Gene Index database [2]. TBtools was used to calculate synonymous (Ks) and nonsynonymous (Ka) substitution rates along with Ka/Ks ratios. The following formula was used to determine the divergence period (T) between duplicated genes: $T = Ks / (2 \times 9.1 \times 10^{-9}) \times 10^{-6}$ million years ago (Mya) [29].

### Haplotype analysis

Using available whole-genome resequencing data from 475 KRICE accessions, haplotype analysis of the *HIS1* gene family was carried out (Table S1). The Nipponbare reference genome (IRGSP 1.0) was used to align these resequencing data [30]. Variations within each gene were extracted using BCFtools (version 1.8), and the resulting

variant call format (VCF) file was used for subsequent analyses [31].

The VCF file containing gene-specific variations was filtered using VCFtools (version 0.1.15) with 0.3 MDR (maximum missing data ratio) and 0.02 MAF (minor allele frequency) [32]. To find genetic variations among the accessions, the filtered VCF file was imported into TASSEL 5.0. For haplotype analysis, sequences in FASTA format were aligned with MEGA 11 [23]. The number of haplotypes was ascertained by importing the multiple sequence alignment file in nexus format into DnaSP (version 6.12.03) [33].

### Phenotyping of KRICE for BBC response

To evaluate the herbicide response in cultivated rice accessions, we conducted a preliminary screening on selected rice accessions with documented responses to HPPD-inhibiting herbicides. This allowed us to optimize BBC concentration and determine the most effective seedling stage for application. Five-day-old seedlings were grown in plastic trays that were submerged in about 5 cm of water. BBC, solubilized in water, was applied to the flooded soil at concentrations of 0, 20, 40, and 80 g active ingredient per hectare (a.i./ha) during the third and fourth leaf stages (Figure S1).

Following this optimization, BBC was applied during the third leaf stage at 80 g a.i./ha to screen 421 cultivated rice accessions. Ten days following treatment, the herbicide reaction was evaluated using a standardized rating scale: 1 (no damage, consistent with the control treatment), 3 (slight weak seedling growth with whitening in few old leaves), 5 (obvious growth inhibition and approximately 50% whitening of seedling area), 7 (stunted growth, bleaching in approximately 70%, and few necrosis), and 9 (complete seedling bleaching and necrosis) [15]. All treatments were conducted at the Plant Genetic Resources greenhouse facility (36.3959°N, 126.5143°E).

### RNA extraction and qRT-PCR analysis

To investigate the response of *HIS1* family genes to BBC treatment, we performed expression analysis on five accessions (2 *japonica* and 3 *indica*), selected based on their herbicide response. At the third leaf stage, seedlings were treated with 80 g a.i./ha of BBC, and six hours after treatment [34], leaf samples were taken from both untreated and treated plants.

For RNA extraction, a commercial kit was utilized, and cDNA synthesis was carried out using iScript™ cDNA Synthesis Kit (Bio-Rad, USA) and one microgram of total RNA. BIOFACT™ 2X real-time PCR master mix (SYBR-Green I; BioFACT, Korea) was employed for quantitative real-time PCR (qRT-PCR) on a CFX96 Touch™ Real-Time PCR Detection System (Bio-Rad, USA). Each gene was analyzed in triplicate technical replicates thermal cycling set to: pre-denaturation (95 °C, 10 min), denaturation (95 °C, 10 s, 4o cycles), annealing (60 °C, 10 s), and extension (72 °C, 15 s). *OsActin* was selected as the reference gene, with primers customized via NCBI Primer-BLAST (Table S2). Using three biological replicates, the $2^{-\Delta\Delta CT}$ technique was used to determine the relative expression levels of the target genes.

## Results

### Distribution of *HIS1* gene family

Our search for conserved domains using HMM and BLAST, a total of 13 *HIS1* proteins were found in the *japonica* genome (Nipponbare). Additionally, orthologous genes were identified across different rice subpopulations: 4 in MH63, 6 in ZS97, and 4 in N22 (Fig. 1a). 140 orthologous genes were identified across monocots and dicots in OrthoFinder, including seven in *O. glaberrima*, 35 in wild rice species (*O. punctate*, *O. meridionalis*, *O. glumaepatula*, *O. brachyantha*, *O. barthii*, *O. nivara*, and *O. rufipogon*), 41 in monocots (*Zea mays*, *Brachypodium distachyon*, *Hordeum vulgare*, *Sorghum bicolor*) and 32 in dicots (*Solanum lycopersicum*, *Helianthus annuus*, *Gossypium raimondii*, *Glycine max*, *Brassica napus*, and *Arabidopsis thaliana*) (Table S3).

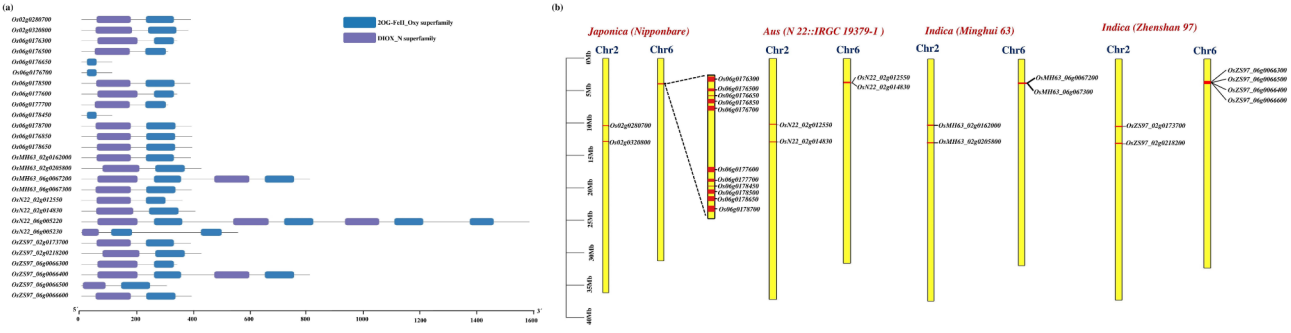



**Fig. 1** Identification and distribution of *HIS1* gene family members across four rice subpopulations. (**a**) Conservation of functional domains within *HIS1* gene family members, highlighting shared and unique domain architectures. (**b**) Chromosomal distribution of HIS1 family genes in each subpopulation

We mapped the chromosomal locations of the *HIS1* family genes across the reference genomes of the four rice subpopulations. Notably, two genes were consistently located on chromosome 2 in all four subpopulations. In contrast, chromosome 6 exhibited variability in the number of *HIS1* genes: *japonica* had 11 *HIS1* genes, the *indica* group had two genes in MH63 and four genes in ZS97, while N22 had two genes. This uneven distribution suggests significant differences in the genomic organization of *HIS1* genes, highlighting potential evolutionary divergences and functional roles specific to each rice subpopulation (Fig. 1b).

### Phylogenetic relationships

An unrooted phylogenetic tree was constructed using 140 proteins retrieved from four rice subpopulations, wild rice species, and other monocot and dicot species (Fig. 2). The *HIS1* gene family members were clustered

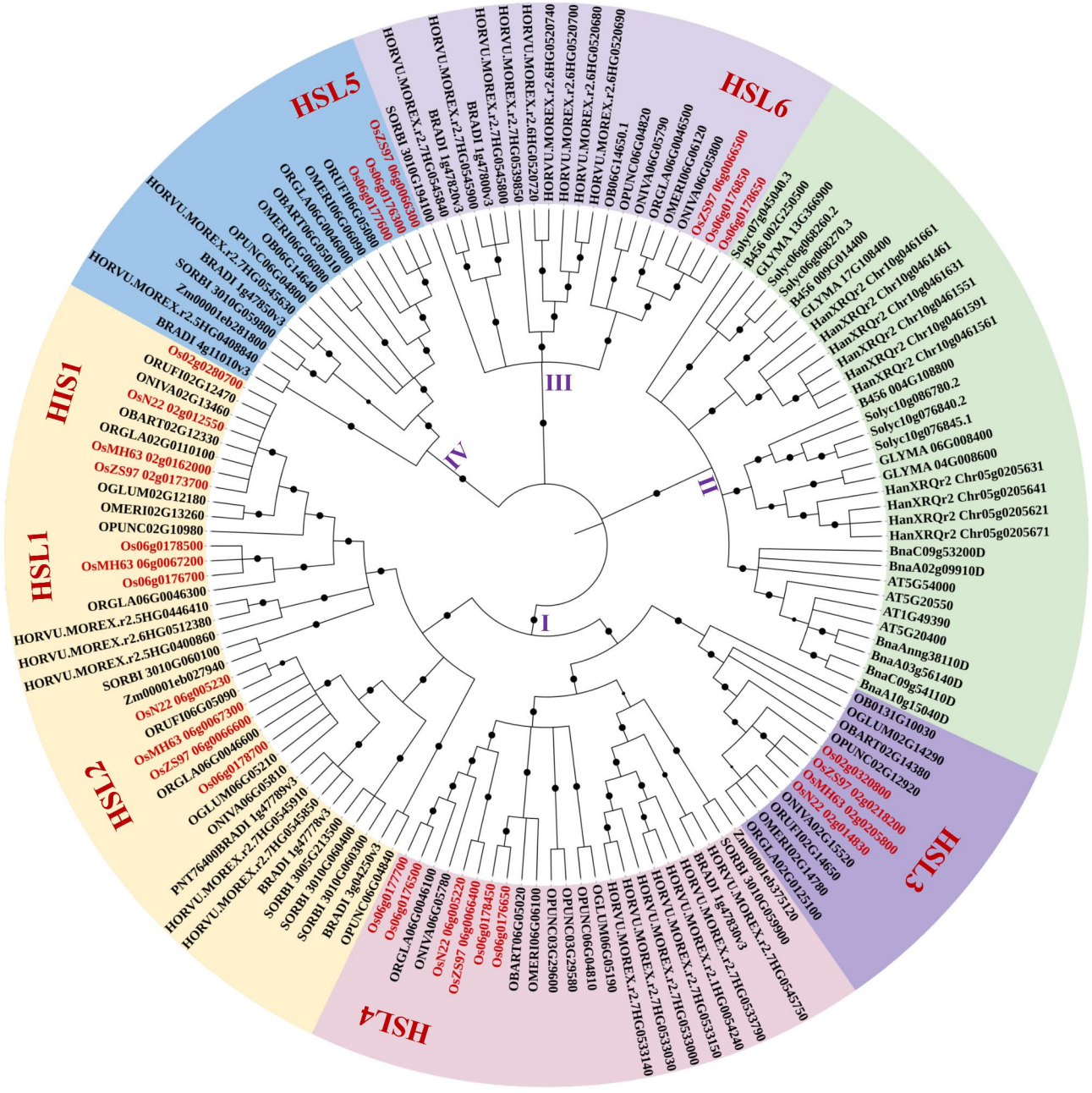



**Fig. 2** Phylogenetic analysis of *HIS1* gene family member proteins. Phylogenetic tree illustrating relationships among 140 *HIS1* gene family proteins from various monocot and dicot species, categorizing them into distinct clades. Refer to Table S3 for detailed species information. HIS1 gene family members identified across the four rice subpopulations are highlighted in red font

into four major clades, designated as Clades I to IV, each containing distinct subclades. Based on their phylogenetic grouping and RAP-DB annotations, the *HIS1* family members were grouped into seven subfamilies, designated as *HIS1* and *HSL1* to *HSL6*.

Clade I formed four distinct subclades, including HIS1 and HSL1-HSL4 proteins. Within Clade I, HIS1 and HSL1 were grouped together in one subclade, while HSL2 formed a separate subclade. Similarly, HSL3 and HSL4 each formed their own subclades. The distinct clustering of HIS1 and HSL proteins within clade I suggests functional diversification while maintaining some evolutionary conservation among the HIS1 family members in rice. Clade II was exclusively composed of dicot-specific proteins; this clade highlights an evolutionary separation between monocots and dicots in the *HIS1* gene family. Clade III and Clade IV were characterized by the presence of HSL6 and HSL5 proteins, respectively.

### Gene architecture and conserved motifs

To gain insight into the structural organization of the *HIS1* gene family, we analyzed gene structures across four rice subpopulations (Fig. 3a). The intron-exon structures within a gene family often reveal both conservation and diversity. Most *HIS1* family members contained between 3 and 6 introns, with a few exceptions. Notably, four genes from *japonica* (*Os06g0176650*, *Os06g0178450*, *Os06g0176850*, and *Os06g0178650*) lacked introns, while the *HSL4* gene from N22 (*OsN22_06g005220*) exhibited a complex structure containing 16 introns (Fig. 3a). Genes with an intronless architecture may have evolutionary advantages under selective pressures, as the absence of introns can lead to faster transcription and reduced processing time [35]. In contrast, the presence of numerous introns is frequently associated with alternative splicing, which enables the production of multiple protein isoforms, contributing to functional diversity and adaptation [36].

In the motif composition analysis, nearly all HIS1 family proteins exhibited a wide range of motifs, with sites ranging from 2 to 28 (Fig. 3b). InterProScan analysis indicated that Motifs 3, 6, 7, and 10 are associated with the Plant_2OG-oxidoreductases (IPR050295) family domain. At least one of these motifs was present in an HIS1 family protein (Table S4). Motifs 1 and 2 were the most conserved and present across nearly all *HIS1* family members, except for three genes (*Os06g0176700*, *Os06g0176650*, and *Os06g0178450*). *HIS1* (*Os02g0280700*) and HSL1 proteins (*Os06g0178500* and *Os06g0176700*) displayed a similar motif composition, with all 10 motifs present consistently, highlighting their close evolutionary relationship. HSL5 lacked Motifs 3 and 5, while HSL6 lacked Motif 6 in *Os06g0176850* and *Os06g0178650* and Motifs 4, 8, and 9 in OsZS97_Os06g0066500. Overall, the MEME motif analysis reveals both conserved and divergent patterns among the *HIS1* gene family members, highlighting the functional diversity and evolutionary adaptation within the rice subspecies.

### Physicochemical properties of *HIS1* family proteins

To comprehensively examine the physicochemical properties of HIS1 family member proteins across four rice subpopulations, we analyzed various parameters, as summarized in Table S5. The lengths of HIS1 proteins varied from 63 to 1441 residues, with corresponding predicted molecular weights spanning from 11.67 to 162.55 kDa. The predicted aliphatic index, an indicator of protein thermostability and half-life [37], was ranged from 71.11 to 101.33. Proteins with higher aliphatic index values, such as Os06g0176700 (101.33), are likely to be more thermostable, while lower values may suggest reduced stability under high-temperature conditions [38]. Most

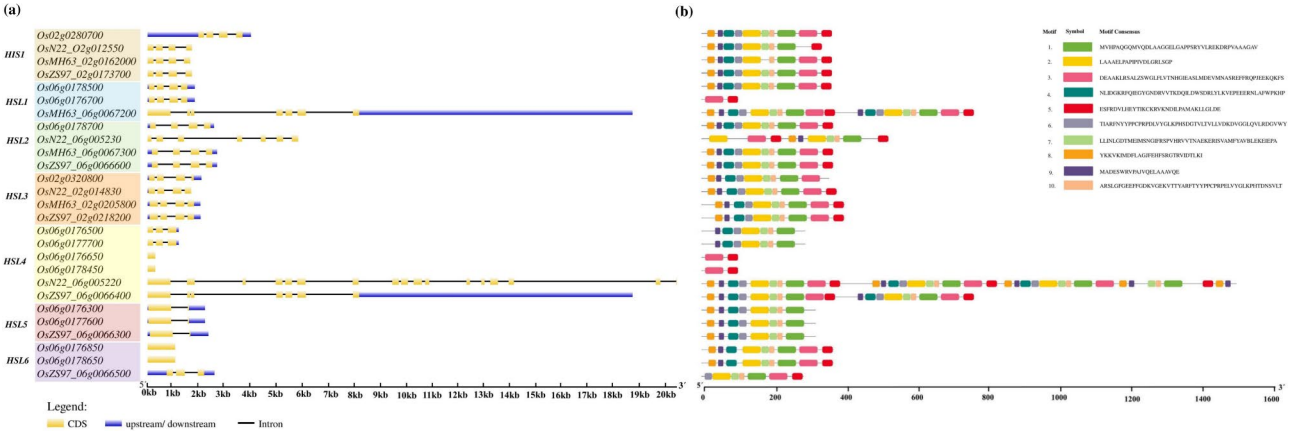



**Fig. 3** Gene structure and motif composition of *HIS1* gene family members. (**a**) Exon-intron structure of *HIS1* family genes. (**b**) Distribution of conserved motifs within *HIS1* family genes, as identified by MEME analysis, with colored boxes representing specific motifs

HIS1 proteins exhibited instability index values indicating variable stability across members. Notably, 15 proteins showed stability (instability index < 40), while others were predicted to be unstable (instability index > 40), suggesting differences in protein turnover rates within cellular environments. HIS1 family members exhibit acidic isoelectric points (pI < 7) and negative GRAVY scores, indicating hydrophilicity. These properties suggest they may play key roles in cellular processes, especially in stress response pathways, through interactions with charged molecules in aqueous environments [37].

Predicted subcellular localization indicated that HIS1 family proteins are predominantly found in the cytoplasm and nucleus, with some also predicted to localize in the plastid. Proteins localized in both the cytoplasm and nucleus, such as *Os06g0178500* and *Os06g0176700*, may have dual roles in cellular regulation and metabolic functions, while plastid-localized proteins could be involved in photosynthetic or biosynthetic pathways.

### Evaluation of selective forces

We computed the Ka/Ks ratios for duplicated gene pairs in order to investigate the evolutionary pressures on the *HIS1* gene family. This metric is critical for interpreting selection pressures on genes [39]. Ka/Ks values above 1 signify adaptive evolution through positive selection, and values below 1 imply negative selection and functional stability, while a ratio of 1 suggests neutral evolution [39].

In our study, the Ka/Ks values varied significantly, ranging from 0.000 to 3.76 (Table 1). For example, the pair *Os06g0176300/OsN22_06g005220* exhibited a Ka/Ks ratio of 3.76, suggesting strong positive selection and possible functional divergence. Conversely, gene pairs such as *Os02g0320800/OsZS97_02g0218200* showed purifying selection and functional conservation with a value of 0.39 (Table 1). These varying Ka/Ks ratios suggest a mix of adaptive and conserved functions among the *HIS1* family members, likely contributing to the diversity and resilience of this gene family across different rice subpopulations.

The estimated timing of duplication events, encompassing both segmental duplication (SD) and whole-genome duplication (WGD), ranged from approximately 0.12 to 4.36 Mya. For instance, the duplication event of *Os02g0320800/OsMH63_02g0205800* was estimated to have occurred around 4.36 Mya, while the duplication of *Os06g0176300/OsMH63_06g0067200* occurred approximately 1.07 Mya (Table 1). These findings provide insights into the evolutionary forces shaping the *HIS1* gene family, indicating that selective pressures have influenced the preservation and diversification of certain gene pairs over time.

### Syntenic relationships

To further investigate the evolutionary relationships and genomic organization of *HIS1* family members across the four rice subpopulations, we conducted a synteny analysis at the chromosome level. This analysis aimed to assess gene duplication patterns and potential evolutionary conservation among these genes. On chromosome 2, two *HIS1* family genes displayed a conserved syntenic relationship across all four rice subpopulations, indicating evolutionary stability in this region. However, the gene arrangement on chromosome 6 exhibited more variability in synteny across subpopulations (Figure S2). To explore this in greater detail, a microcollinearity analysis was performed, revealing distinct gene organization patterns within the *HIS1* family (Fig. 4). Our analysis identified multiple duplication mechanisms contributing to the expansion and diversification of *HIS1* family members on chromosome 6. We observed tandem and proximal duplication events in the *HSL1*, *HSL4*, and *HSL5* genes in *japonica*. Identifying various duplication types within *HSL1*, *HSL4*, and *HSL5* supports the hypothesis that these genes have undergone functional specialization, potentially driven by selective pressures in different ecological or physiological contexts within rice.

These duplication events showed significant sequence similarity within the genes of specific HSL groups (Figure S3). The protein sequences of all genes within the *HSL5* group were identical, while among three *HSL6* genes, *OsZS97_Os06g0066500* differed from other *HSL6* genes only by protein length. Among the four *HSL4* genes in *japonica*, *Os06g176500* and *Os06g177700*

**Table 1** The Ka/Ks and type of duplication gene pairs among four cultivated subspecies

| Subpopulation | Gene pairs | Ka | Ks | Ka_Ks | Time (Mya) | Type of Duplication |
|---|---|---|---|---|---|---|
| Nipponbare/Minghui63 | *Os02g0280700/OsMH63_02g0162000* | 0.0143 | 0.0200 | 0.7170 | 1.0997 | WGD/segmental |
| | *Os02g0320800/OsMH63_02g0205800* | 0.0308 | 0.0793 | 0.3889 | 4.3559 | WGD/segmental |
| | *Os06g0176300/OsMH63_06g0067200* | 0.0320 | 0.0195 | 1.6425 | 1.0696 | WGD/segmental |
| Nipponbare/Zhenshan97 | *Os02g0320800/OsZS97_02g0218200* | 0.0308 | 0.0793 | 0.3889 | 4.3559 | WGD/segmental |
| | *Os06g0176300/OsZS97_06g0066300* | 0.0000 | 0.0043 | 0.0000 | 0.2360 | WGD/segmental |
| Nipponbare/N22 | *Os02g0320800/OsN22_02g014830* | 0.0308 | 0.0793 | 0.3889 | 4.3559 | WGD/segmental |
| | *Os06g0176300/OsN22_06g005220* | 0.0082 | 0.0022 | 3.7619 | 0.1203 | WGD/segmental |
| | *Os06g0178700/OsN22_06g005230* | 0.0000 | 0.0099 | 0.0000 | 0.5418 | WGD/segmental |

Ka = nonsynonymous substitution rate, Ks = synonymous substitution rate, Mya = Millions of years ago

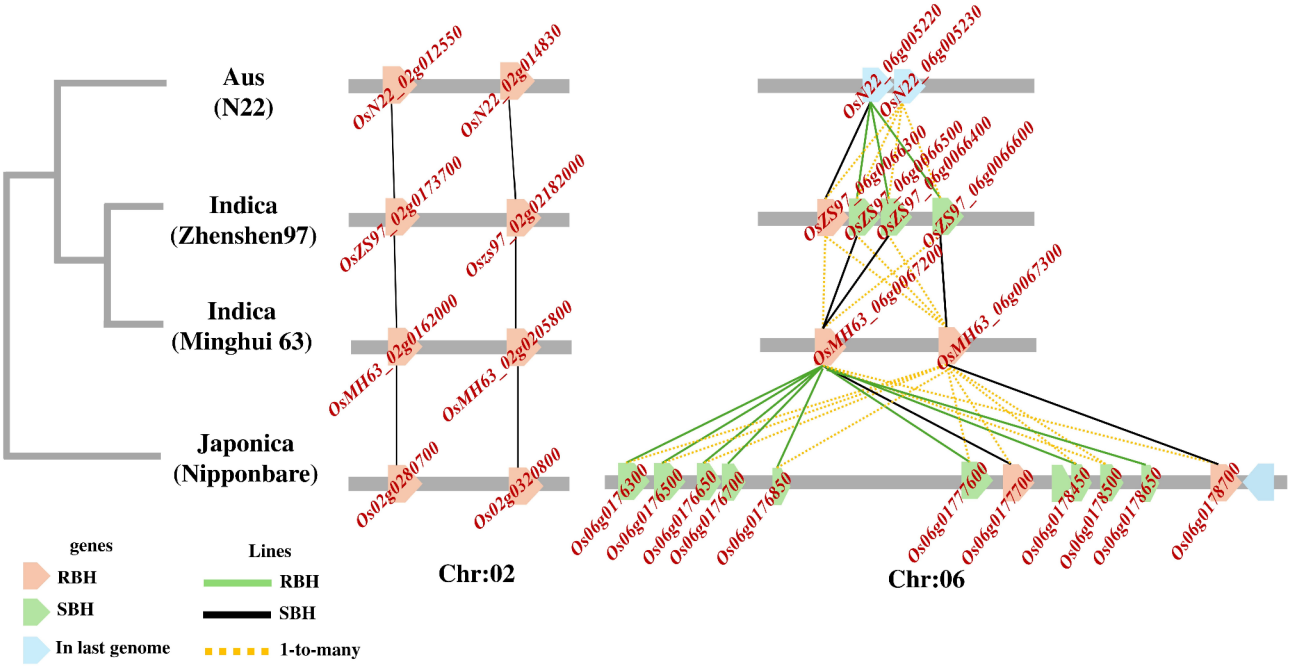



**Fig. 4** Microcollinearity analysis of *HIS1* gene family members among four subspecies. RBH refers to the reciprocal best hits and SBH refers to the single-side best hits

displayed identical sequences, as did *Os06g176650* and *Os06g178450.* Similarly, *HSL1* genes *Os06g0178500 and Os06g176700* showed duplicated sequences (Figure S3).

### Haplotype variation in KRICE for *HIS1* gene family

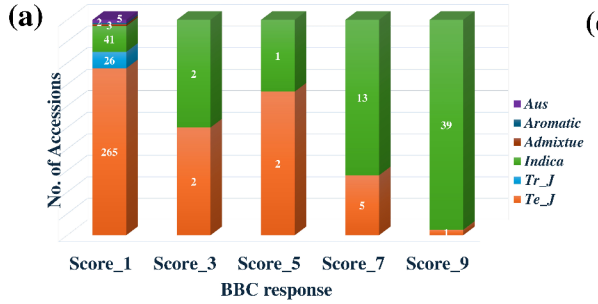
Further, we studied the genetic diversity of the *HIS1* gene family across the 475 KRICE accessions. For the *HIS1* gene, our previous study identified 42 haplotypes, four of these were common in wild and cultivated accessions, 34 were specific to wild rice, and four were unique to cultivated rice [15]. To extend this analysis to the remaining genes of the *HIS1* family, we investigated haplotype diversity in KRICE, using Nipponbare as a reference. Out of 12 genes analyzed, 10 showed no haplotype variation among the 475 accessions, with all sequences identical to the reference genome. In contrast, two genes—*HSL2* (*Os06g0178700*) and *HSL3* (*Os02g0320800*)—exhibited significant genetic variations.

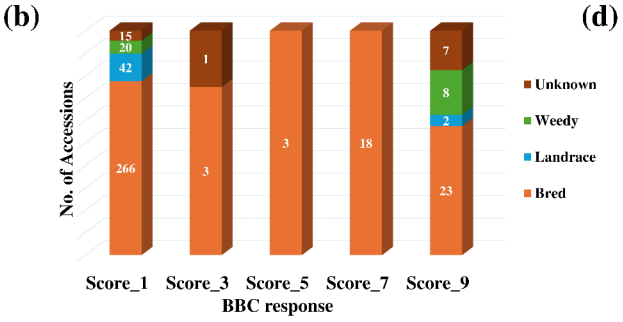
In *HSL2*, 40 distinct haplotypes were identified, of which three haplotypes were shared between both cultivated and wild rice accessions (Figure S4). 35 haplotypes were unique to wild rice accessions, whereas Hap_3 and Hap_4 were distinct to cultivated accessions. Hap_1 (the reference haplotype) included the highest number of accessions (405), encompassing all temperate *japonica*, all tropical *japonica*, 84 *indica*, 11 wild rice, two aromatic, and three *aus* accessions. Hap_2 was found in nine accessions, including one wild accession, five *indica* accessions, and three *aus* accessions. Hap_3 had a single accession, while Hap_4 included eight *indica* and four *aus* accessions. Hap_5 comprised four wild, one *indica*, and one *aus* accession. non-synonymous SNP G/A at 3,901,465 position causing arginine to glutamine amino acid change. In wild-specific haplotypes, 42 wild accessions were distributed across 35 distinct haplotypes, reflecting the diversity and evolutionary divergence within the wild rice populations.

In *HSL3*, a total of 45 distinct haplotypes were detected (Figure S5). Hap_1 (reference haplotype) contained all tropical and temperate *japonica* accessions, along with two aromatic, 12 wild, and two *indica* accessions. Hap_10 was the second largest haplotype, consisting of 76 accessions, including six wild, 66 *indica*, three *aus*, and one admixture accession. Three haplotypes were common between wild and cultivated accessions. Notably, 25 haplotypes were specific to cultivated accessions, predominantly represented by *indica* rice. These haplotypes are characterized by multiple non-synonymous SNPs and indels (Figure S5).

### BBC response of KRICE and phenotyping performances of haplotypes

We analyzed the response of the cultivated rice accessions from KRICE to the BBC application (Table S1). Most *japonica* accessions were resistant to BBC; 267 accessions received a score of 1 (no damage), whereas two accessions received a score of 3 (minor damage) (Fig. 5a). Among *indica* accessions, 41 exhibited high resistance to BBC, along with three *admixtures*, two *aromatic*, and five *aus* accessions. Moderate sensitivity,




(c)

| *HSL2* | Total screened | Score (no. accessions) | | | | | Mean score |
|---|---|---|---|---|---|---|---|
| | | 1 | 3 | 5 | 7 | 9 | |
| ***Indica*** | | | | | | | |
| Hap_1 | 82 | 36 | 2 | 1 | 13 | 30 | 4.98 |
| Hap_2 | 5 | 1 | | 1 | | 3 | 7 |
| Hap_4 | 8 | 3 | | | | 5 | 6 |
| Hap_5 | 1 | 1 | | | | | 1 |
| ***Japonica*** | | | | | | | |
| Hap_1 | 302 | 291 | 2 | 2 | 5 | 2 | 2.54 |
| ***Aus*** | | | | | | | |
| Hap_2 | 1 | 1 | | | | | 1 |
| Hap_4 | 4 | 4 | | | | | 1 |
| Hap_5 | 1 | 1 | | | | | 1 |

(d)

| *HSL3* | Total screened | Score (no. accessions) | | | | | Mean score |
|---|---|---|---|---|---|---|---|
| | | 1 | 3 | 5 | 7 | 9 | |
| ***Indica*** | | | | | | | |
| Hap_1 | 2 | 1 | | | | 1 | 5 |
| Hap_3 | 3 | | | | | 3 | 9 |
| Hap_5 | 2 | 1 | | | | 1 | 5 |
| Hap_10 | 63 | 25 | 1 | 1 | 8 | 28 | 5.4 |
| Hap_21 | 2 | 2 | | | | | 1 |
| Hap_31 | 3 | 3 | | | | | 1 |
| ***Japonica*** | | | | | | | |
| Hap_1 | 302 | 291 | 2 | 2 | 5 | 2 | 2.54 |
| ***Aus*** | | | | | | | |
| Hap_3 | 3 | 1 | | | | | 1 |
| Hap_10 | 1 | 1 | | | | | 1 |
| Hap_16 | 1 | 1 | | | | | 1 |

**Fig. 5** Phenotypic analysis of BBC response among cultivated KRICE accessions. (**a**) BBC response distribution among KRICE accessions, categorized by ecotypes, with Te_J representing temperate *japonica*, and Tr_J for tropical *japonica*. (**b**) Classification of BBC response based on varietal types among KRICE accessions. (**c**-**d**) BBC response variation across different ecotypes and haplotypes of (**c**) *HSL2* and (**d**) *HSL3*

characterized by growth inhibition and bleaching symptoms, was observed in only two temperate *japonica* and one tropical *japonica* accession. The most susceptible accessions included six temperate *japonica* and 52 *indica* accessions, displaying the highest sensitivity to BBC (score 9) (Fig. 5a). In terms of varietal types, cultivated KRICE accessions are categorized into landrace (49), bred (320), weedy (28), and unknown status (24) (Table S1). Among these, BBC susceptibility was noted in eight weedy, 23 bred, and two landrace accessions (Fig. 5b).

We further analyzed the phenotypic response to BBC within specific haplotypes of the *HSL2* and *HSL3* genes across different rice ecotypes. For both *HSL2* and *HSL3*, *japonica* accessions were grouped in Hap_1, the reference haplotype, with a mean score of 2.54, reflecting resistance. For *HSL2*, among 82 *indica* accessions in Hap_1, 38 exhibited resistance with a score of 1 or 3, while 43 accessions were susceptible with scores of 7 or 9, showing a clear contrast within the same haplotype. Other *indica* haplotypes, Hap_2 (5 accessions) and Hap_4 (8 accessions), showed mixed responses without clear separation of resistance or susceptibility. However, Hap_5 (1 *indica* accession) consistently demonstrated resistance to BBC (score 1) (Fig. 5c).

In the *HSL3* gene, 25 of the 66 indica accessions in Hap_10 were highly resistant (score 1), while 28 were highly susceptible (score 9). All three *indica* accessions in Hap_3 showed the highest sensitivity to BBC, scoring 9, whereas two *indica* accessions in Hap_21 demonstrated strong resistance with a score of 1. In *aus*, different haplotypes of both *HSL2* and *HSL3* genes consistently exhibited resistance to BBC, with a score of 1 across all accessions (Fig. 5d).

### Expression profiles under BBC treatment

We examined the gene expression profiles of five representative rice accessions selected based on their BBC responses to better understand how the *HIS1* gene family responds to BBC treatment. By comparing the levels of gene expression in control and BBC-treated samples, we were able to assess the expression profiles of the members of the *HIS1* gene family (*HIS1*, *HSL1* to *HSL5*) (Fig. 6).

In the two *japonica* accessions, Ondami and Milseong, most *HIS1* family genes displayed elevated expression levels under BBC exposure, except *HSL5*. In particular, Milseong, a highly susceptible, had notably lower *HSL5* expression than Ondami, which has high resistance to BBC (BBC response score = 1). In Cheongdo-Donggok-4, a highly BBC-resistant *indica* accession carrying the 28 bp deletion in *HIS1*, *HSL5* was distinctly upregulated, while the expression of all other *HIS1* family genes was downregulated. While highly susceptible *indica* accessions, Dasan1, and Chiem Chanh exhibited downregulation of all *HIS1* family genes following BBC treatment. Overall, *japonica* accessions generally show higher fold expression in multiple genes, while *indica* accessions exhibit more variable expression profiles. These findings underscore the diverse responses of *HIS1* gene family members to BBC stress in rice, with notable differences

**(a)**

| Accession | Ecotype | BBC response | Allele/haplotype HIS1 | HSL2 | HISL3 |
|---|---|---|---|---|---|
| Ondami | Japonica | 1 | Ref | Ref | Ref |
| Cheongdo-Donggok-4 | Indica | 1 | 28bp_del | Hap_4 | Hap_21 |
| Dasan1 | Indica | 9 | 28bp_del | Ref | Hap_10 |
| Chiem Chanh | Indica | 9 | 28bp_del | Hap_2 | Hap_3 |
| Milseong | Japonica | 7 | Ref | Ref | Ref |

**(b)** Fold expression change

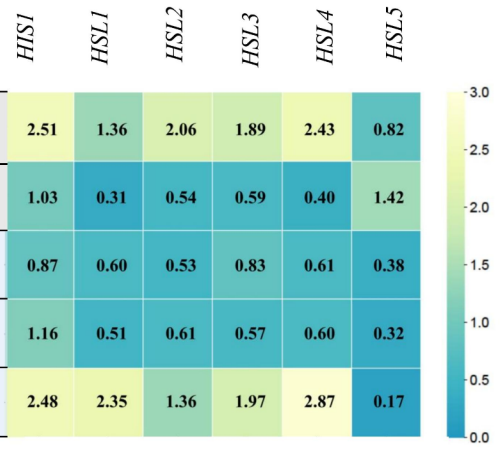



**Fig. 6** Differential expression patterns of *HIS1* gene family members following BBC treatment. (**a**) Overview of selected accessions with BBC response ratings and their haplotypes. (**b**) Heatmap showing the expression changes of *HIS1* gene family members post-BBC application, with expression levels normalized to the actin gene. Data was collected at 6 h post-treatment. Ref: reference haplotype with no sequence variation, del: deletion

between *japonica* and *indica* ecotypes. The unique expression patterns of *HSL5* highlight its potential role in BBC resistance, particularly in *indica* accessions with the 28 bp deletion in *HIS1*. Further exploration of the specific functions of *HIS1* gene family members and their regulatory mechanisms will provide deeper insights into herbicide resistance and susceptibility pathways in rice.

## Discussion

The identification and characterization of herbicide-resistant genes represent a critical area in modern crop breeding, providing a foundation for improved weed management and boosting crop productivity [40, 41]. Our investigation into the *HIS1* gene family across four rice subpopulations highlighted significant genetic diversity within *O. sativa*. Among the family members, *HIS1*, *HSL2*, and *HSL3* were found across four *Oryza* subpopulations, suggesting their fundamental roles in different rice ecotypes (Fig. 2). In contrast, lineage-specific gene loss or divergence was evident, with MH63 lacking *HSL4*, *HSL5*, and *HSL6* genes, while *HSL1* was absent in both N22 and ZS97. These variations align with findings from previous studies, demonstrating that rice subspecies exhibit unique gene presence/absence patterns that influence their stress response and adaptive traits [42–44]. Interestingly, the absence of *HSL4* genes in wild rice (*O. rufipogon*) but its presence in *O. nivara* and *O. glaberrima* suggests that *HSL4* may have arisen or become fixed in the lineage leading to cultivated *japonica* rice after divergence from its wild ancestor [45, 46].

Collinearity analysis revealed that the expansion of the *HIS1* family was driven by both segmental/whole-genome duplications (SD/WGD) and tandem duplications, consistent with patterns seen in other gene families [47]. While SD/WGD duplications occur more frequently in slowly evolving gene families, tandem duplications are common in rapidly evolving clusters, driving functional diversification under selective pressures such as herbicide exposure [48]. This dual mechanism of duplication underscores the evolutionary flexibility of the *HIS1* family, enabling its adaptation to varying ecological challenges. Tandem duplications play a pivotal role in gene family expansion and functional diversification, as they often lead to increased gene dosage or redundancy [49]. In the case of the *HIS1* family, tandem duplication events on chromosome 6 may have enhanced herbicide tolerance in *japonica* rice by amplifying detoxification-related genes, thereby enabling plants to better withstand herbicide-induced selective pressures. Comparative insights from other crops strengthen the significance of tandem duplications in conferring herbicide resistance. For instance, tandemly duplicated *glutathione S-transferase* genes (*TtGSTU1* and *TtGSTU2*) in *Triticum tauschii* have been shown to enhance herbicide detoxification by increasing the production of detoxifying enzymes [50]. Similarly, tandem duplication of the *5-enolpyruvylshikimate-3-phosphate synthase* (*EPSPS*) gene has been identified as a key mechanism for glyphosate resistance in multiple weed species, where elevated gene copy numbers confer robust herbicide tolerance [51]. In *Alopecurus aequalis*, *acetolactate synthase* genes (*ALS3* and *ALS4*) underwent tandem duplication, contributing to resistance against acetolactate synthase-inhibiting herbicides [52]. Additionally, the observed range of Ka/Ks ratios (0.39–3.76) (Table 1) suggests varying degrees of evolutionary constraint and adaptive divergence within the duplicated gene pairs [49]. These evolutionary dynamics likely played a role in shaping the *HIS1* gene family,

allowing it to contribute to plant adaptation and diversification [50].

Population genetics factors, like selection pressures, recombination, and mutation rates, contribute to haplotypic diversity within genomic regions [53–55]. In our analysis, 10 out of 12 genes showed no haplotype variations across 421 cultivated accessions and 54 wild accessions, while five distinct haplotypes were recorded for *HSL2* in cultivated rice (Figure S4). Limited haplotypic diversity in KRICE suggests strong purifying selection or low evolutionary pressure for functional variation in most *HIS1* family genes. Notably, *japonica* accessions were predominantly clustered within a single reference haplotype, whereas *indica* accessions exhibited greater diversity, with 4 and 25 haplotypes identified for *HSL2* and *HSL3*, respectively. These differences underscore the differential selection pressures or a wider adaptive range within *indica* rice (Figure S5).

Consistent with our previous findings on TFT response [15], most *japonica* accessions demonstrated resistance to BBC, while *indica* accessions displayed phenotypic diversity. The most frequent resistant haplotype (Hap_1) was identified in *japonica* rice for both *HSL2* and *HSL3*, with 291 accessions showing resistance (score 1 and 3) (Fig. 5). In *indica*, the most frequent haplotype for *HSL2* (Hap_1) was observed in 82 accessions, of which 38 accessions were resistant (score 1 and 3), while 43 accessions were susceptible (score 7 and 9), indicating a heterogeneous response. Similarly, for *HSL3*, the most common *indica* haplotype (Hap_10) was detected in 66 accessions, with 26 resistant and 36 susceptible accessions. Notably, rare haplotypes such as Hap_5 in *HSL2* and Hap_21 and Hap_31 in *HSL3* were associated with BBC resistance, with all accessions carrying these haplotypes scoring 1 (Table S1). These rare but beneficial haplotypes could serve as valuable genetic resources for marker-assisted selection (MAS) in herbicide-resistant rice breeding. Overall, phenotypic evaluation of *HSL2* and *HSL3* major haplotypes did not reveal distinct patterns of resistance or susceptibility, suggesting that genetic variations within these loci in *indica* and *aus* ecotypes may not be the primary determinants of BBC response (Fig. 5). These findings underscore the possibility that other factors, such as transcriptional regulation, epigenetic modifications, or interactions with other metabolic pathways, could play a critical role in mediating herbicide resistance. Future studies integrating transcriptomic, proteomic, and metabolomic approaches could uncover additional molecular networks governing herbicide detoxification and tolerance. Moreover, the differential expression patterns of *HIS1* family genes in *japonica* and *indica* accessions point to ecotype-specific regulatory mechanisms that may influence herbicide resistance (Fig. 6). In *japonica*, higher expression levels on most of *HIS1* family genes suggest the presence of a coordinated regulatory network enhancing BBC resistance. Conversely, in *indica* accessions, expression patterns were more variable. Notably, the selective upregulation of *HSL5* in a BBC-resistant *indica* accession, coupled with its downregulation in highly susceptible *japonica* accession, highlights its unique role in mediating herbicide resistance. Furthermore, all *HIS1* family genes were downregulated in susceptible *indica* accessions, pointing to potential transcriptional repression mechanisms under herbicide stress. The specific upregulation of *HSL5* in resistant *indica* accession carrying a 28 bp deletion in *HIS1* underscores its potential as a key determinant of resistance. This finding positions *HSL5* as a potential candidate for targeted breeding programs or CRISPR/Cas9-based interventions aimed at optimizing its expression to enhance herbicide resistance in rice. Fine-tuning the regulatory pathways of *HSL5* could provide an effective strategy for developing rice varieties with improved adaptability and resilience to herbicide applications, particularly in *indica* ecotypes where phenotypic responses to herbicides are highly variable.

Interestingly, differences in peptide lengths of HSL1, HSL4, and HSL6 were evident among subpopulations when analyzed using their respective reference genomes (Fig. 3, S3). However, these variations were not as apparent when analyzed through the Nipponbare reference in the KRICE dataset, where no variations were detected, even among 54 wild rice accessions. This suggests that using a *japonica*-centric reference genome may obscure ecotype-specific variations. Employing ecotype-specific or pan-genomic reference genomes could better capture the genetic diversity and structural differences in the *HIS1* family, facilitating the development of more precise breeding strategies tailored to specific subpopulations.

Target site mutagenesis provides a promising strategy for generating plants resistant to HPPD inhibitors. A recent study demonstrated that catalytic selectivity and efficient herbicide binding are essential for effectively detoxifying β-triketone herbicides [18]. Using an artificial evolution approach, researchers developed the *HSL1*-M10 mutant (F140H/L204F/F298L/I335F) lines by modifying four key residues involved in herbicide binding. This enhanced the catalytic efficiency of *HSL1*, and the overexpression of M10 in *japonica* rice resulted in significant improvement to β-triketone herbicide resistance, including a 32-fold increase in resistance to methyl-benquitrione [18]. In our study, *HSL1* is conserved across KRICE accessions, marking it as a potential target for cisgenesis and artificial evolution approaches, along with *HSL5*, to develop herbicide-resistant rice varieties.

## Conclusions

Here, we present an in-depth characterization of the *HIS1* gene family across four rice subpopulations, identifying 25 members located on chromosomes 2 and 6, which were further classified into seven subfamilies (*HIS1*, *HSL1*~*HSL6*). Furthermore, we characterized these genes based on their structural composition, conserved domains, physicochemical traits, collinearity, and selective pressures shedding light on their evolutionary patterns and functional diversification. Our analysis of haplotype variations among *HIS1* family members across 475 KRICE accessions revealed that natural genetic variations were present in only three genes: *HIS1*, *HSL2*, and *HSL3*. Conducting ecotype-specific analyses or using a reference genome tailored to the genetic makeup of each ecotype could improve our ability to identify distinct patterns of variation that are potentially linked to herbicide response. These findings lay a foundation for future research focused on improving herbicide resistance through selective breeding and gene-editing approaches, which could contribute to more effective and sustainable weed management in rice cultivation.

#### Abbreviations

| | |
|---|---|
| a.i. | Active ingredient |
| BBC | Benzobicyclon |
| DNA | Deoxyribonucleic acid |
| GRAVY | Grand average of hydropathicity |
| Hap | Haplotype |
| HIS1 | HPPD inhibitor sensitive 1 |
| HMM | Hidden markov model |
| HPPD | 4-Hydroxyphenylpyruvate dioxygenase |
| HSL | HIS1-like |
| Ka | Nonsynonymous substitution rates |
| kDa | Kilodalton |
| KRICE | Korean rice collection |
| Ks | Synonymous substitution rates |
| MAS | Marker-assisted selection |
| MEME | Multiple EM for Motif Elicitation |
| MH63 | Minghui 63 |
| Mya | Million years ago |
| N22 | Nagina-22 |
| NGS | Next-Generation Sequencing |
| pI | Isoelectric point |
| qRT-PCR | Quantitative reverse transcription polymerase chain reaction |
| RAP | DB-Rice Annotation Project Database |
| RBH | Reciprocal Best Hits |
| RNA | Ribonucleic Acid |
| SBH | Single-side Best Hits |
| SD/WGD | Segmental Duplication/Whole Genome Duplication |
| SNP | Single Nucleotide Polymorphism |
| TFT | Tefuryltrione |
| VCF | Variant call format |
| ZS91 | Zhenshan 97 |
| β-TH | β-Triketone Herbicide |

## Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12870-025-06379-z.

Supplementary Material 1

Supplementary Material 2

#### Acknowledgements

Not applicable.

#### Author contributions

MHA: investigation, methodology, formal analysis, data curation, software, visualization, writing - original draft. S-HC: data curation, formal analysis, resources, visualization, manuscript review and editing. BN: conceptualization, formal analysis, validation, visualization, manuscript review and editing. Y-JP: conceptualization, fund acquisition, project administration, supervision, validation, resources, manuscript review and editing. All authors read and approved the final manuscript.

#### Funding

This work was supported by National Research Foundation of Korea (NRF) grants funded by the Korean government (MSIT) (NRF-2023R1A2C1004432 and NRF-2022R1A4A1030348), Korea Institute of Planning and Evaluation for Technology in Food, Agriculture and Forestry (IPET) through Digital Breeding Transformation Technology Development Program, funded by Ministry of Agriculture, Food and Rural Affairs (MAFRA)(322060031HD020).

#### Data availability

All the data supporting the findings of this study are available within the paper and within its supplementary materials.

## Declarations

#### Ethics approval and consent to participate

Not applicable.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare no competing interests.

Received: 15 November 2024 / Accepted: 10 March 2025
Published online: 24 March 2025

#### References

1. Wing RA, Purugganan MD, Zhang Q. The rice genome revolution: from an ancient grain to green super rice. Nat Rev Genet. 2018;19:505–17.
2. Yu Z, Chen Y, Zhou Y, Zhang Y, Li M, Ouyang Y, et al. Rice gene index: A comprehensive pan-genome database for comparative and functional genomics of Asian rice. Mol Plant. 2023;16:798–801.
3. Zhou Y, Chebotarov D, Kudrna D, Llaca V, Lee S, Rajasekar S, et al. A platinum standard pan-genome resource that represents the population structure of Asian rice. Sci Data. 2020;7:113.
4. Zhou Y, Yu Z, Chebotarov D, Chougule K, Lu Z, Rivera LF, et al. Pan-genome inversion index reveals evolutionary insights into the subpopulation structure of Asian rice. Nat Commun. 2023;14:1567.
5. Zhang J, Chen L-L, Xing F, Kudrna DA, Yao W, Copetti D, et al. Extensive sequence divergence between the reference genomes of two elite indica rice varieties Zhenshan 97 and Minghui 63. Proc Natl Acad Sci U S A. 2016;113:E5163–71.
6. Yadav N, Sevanthi ACMV, Pandey R, Chinnusamy V, Singh AK, Singh NK. Physiological response and agronomic performance of drought tolerance mutants of Aus rice cultivar nagina 22 (*Oryza sativa* L). Field Crops Res. 2023;290:108760.
7. Bai Z, Chen J, Liao Y, Wang M, Liu R, Ge S, et al. The impact and origin of copy number variations in the Oryza species. BMC Genomics. 2016;17:261.
8. Chauhan BS. Strategies to manage weedy rice in Asia. Crop Prot. 2013;48:51–6.
9. Nadir S, Xiong H-B, Zhu Q, Zhang X-L, Xu H-Y, Li J, et al. Weedy rice in sustainable rice production. A review. Agron Sustain Dev. 2017;37:46.
10. Komatsubara K, Sekino K, Yamada Y, Koyanagi H, Nakahara S. Discovery and development of a new herbicide, benzobicyclon. J Pestic Sci. 2009;34:113–4.
11. Patterson JA, Norsworthy JK, Butts TR, Gbur EE. Benzobicyclon for weedy rice control in quizalofop- and imidazolinone-resistant rice systems. Weed Technol. 2022;36:497–505.

12. Jhala AJ, Kumar V, Yadav R, Jha P, Jugulam M, Ii MMW, et al. 4-Hydroxyphenylpyruvate dioxygenase (HPPD)-inhibiting herbicides: past, present, and future. Weed Technol. 2023;37:1–14.
13. Maeda H, Murata K, Sakuma N, Takei S, Yamazaki A, Karim MR, et al. A rice gene that confers broad-spectrum resistance to β-triketone herbicides. Science. 2019;365:393–6.
14. Lv Q, Zhang X, Yuan D, Zhiyuan H, Peng R, Peng J et al. Exploring natural allelic variations of the β-triketone herbicide resistance gene HIS1 for application in indica rice and particularly in two-line hybrid rice. 2020.
15. Aung KM, Chu S-H, Nawade B, Lee C-Y, Myung EJ, Park Y-J. Analyzing the response of rice to Tefuryltrione herbicide: haplotype variation and evolutionary dynamics of the *HIS1* gene. Environ Res. 2024;252:118839.
16. Dai S, Georgelis N, Bedair M, Hong Y, Qi Q, Larue CT, et al. Ectopic expression of a rice triketone dioxygenase gene confers mesotrione tolerance in soybean. Pest Manag Sci. 2022;78:2816–27.
17. Komatsu A, Ohtake M, Shimatani Z, Nishida K. Production of Herbicide-Sensitive strain to prevent volunteer rice infestation using a CRISPR-Cas9 cytidine deaminase fusion. Front Plant Sci. 2020;11.
18. Dong J, Yu X-H, Dong J, Wang G-H, Wang X-L, Wang D-W et al. An artificially evolved gene for herbicide-resistant rice breeding. Proceedings of the National Academy of Sciences. 2024;121:e2407285121.
19. Aravind L, Koonin EV. The DNA-repair protein AlkB, EGL-9, and leprecan define new families of 2-oxoglutarate- and iron-dependent dioxygenases. Genome Biol. 2001;2:RESEARCH0007.
20. Chauhan H, Aiana, Singh K. Genome-wide identification of 2-oxoglutarate and Fe (II)-dependent dioxygenase family genes and their expression profiling under drought and salt stress in potato. PeerJ. 2023;11:e16449.
21. Chen C, Chen H, Zhang Y, Thomas HR, Frank MH, He Y, et al. TBtools: an integrative toolkit developed for interactive analyses of big biological data. Mol Plant. 2020;13:1194–202.
22. Emms DM, Kelly S. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biol. 2015;16:157.
23. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38:3022–7.
24. Letunic I, Bork P. 20 Years of the SMART protein domain annotation resource. Nucleic Acids Res. 2018;46:D493–6.
25. Hu B, Jin J, Guo A-Y, Zhang H, Luo J, Gao G. GSDS 2.0: an upgraded gene feature visualization server. Bioinformatics. 2015;31:1296–7.
26. Gasteiger E, Gattiker A, Hoogland C, Ivanyi I, Appel RD, Bairoch A, ExPASy. The proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 2003;31:3784–8.
27. Thumuluri V, Almagro Armenteros JJ, Johansen AR, Nielsen H, Winther O. DeepLoc 2.0: multi-label subcellular localization prediction using protein Language models. Nucleic Acids Res. 2022;50:W228–34.
28. Wang Y, Tang H, Debarry JD, Tan X, Li J, Wang X, et al. MCScanX: a toolkit for detection and evolutionary analysis of gene synteny and collinearity. Nucleic Acids Res. 2012;40:e49.
29. Zhou M, Jiang Y, Liu X, Kong W, Zhang C, Yang J, et al. Genome-Wide identification and evolution analysis of the CYP76 subfamily in rice (Oryza sativa). IJMS. 2023;24:8522.
30. Phitaktansakul R, Kim K-W, Aung KM, Maung TZ, Min M-H, Somsri A, et al. Multi-omics analysis reveals the genetic basis of rice fragrance mediated by *betaine aldehyde dehydrogenase 2*. J Adv Res. 2022;42:303–14.
31. Somsri A, Chu S-H, Nawade B, Lee C-Y, Park Y-J. Harnessing γ-TMT genetic variations and haplotypes for vitamin E diversity in the Korean rice collection. Antioxidants. 2024;13:234.
32. Aung KM, Oo WH, Maung TZ, Min M-H, Somsri A, Nam J et al. Genomic landscape of the OsTPP7 gene in its haplotype diversity and association with anaerobic germination tolerance in rice. Front Plant Sci. 2023;14.
33. Rozas J, Ferrer-Mata A, Sánchez-DelBarrio JC, Guirao-Rico S, Librado P, Ramos-Onsins SE, et al. DnaSP 6: DNA sequence polymorphism analysis of large data sets. Mol Biol Evol. 2017;34:3299–302.
34. Brabham C, Norsworthy J, Sha X, Varanasi VK, González-Torralva F. Benzobicyclon efficacy is affected by plant growth stage, HPPD inhibitor sensitive 1 (HIS1) expression and zygosity in weedy rice (Oryza sativa). Weed Sci. 2022;70:328–34.
35. Jeffares DC, Penkett CJ, Bähler J. Rapidly regulated genes are intron poor. Trends Genet. 2008;24:375–8.
36. Keren H, Lev-Maor G, Ast G. Alternative splicing and evolution: diversification, exon definition and function. Nat Rev Genet. 2010;11:345–55.
37. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157:105–32.
38. Nawade B, Bosamia TC, Lee JH, Jang JH, Lee OR. Genome-wide characterization of the soybean DOMAIN OF UNKNOWN FUNCTION 679 membrane protein gene family highlights their potential involvement in growth and stress response. Front Plant Sci. 2023;14:1216082.
39. Wang W, Zhao P, Zhou X, Xiong H, Sun M. Genome-wide identification and characterization of Cystatin family genes in rice (Oryza sativa L). Plant Cell Rep. 2015;34:1579–92.
40. Luo Q, Liu Y-G. Breeding herbicide-resistant rice (*Oryza sativa*) using CRISPR/Cas gene editing and other technologies. Plant Commun. 2024;:101172.
41. Dong H, Huang Y, Wang K. The development of herbicide resistance crop plants using CRISPR/Cas9-Mediated gene editing. Genes. 2021;12:912.
42. Jiang L, Wu L, Wang Y, Xu Q, Xu Z, Chen W. Research progress on the divergence and genetic basis of agronomic traits in *Xian* and *Geng* rice. Crop J. 2022;10:924–31.
43. Yu P, Wang C-H, Xu Q, Feng Y, Yuan X-P, Yu H-Y, et al. Genome-wide copy number variations in Oryza sativa L. BMC Genomics. 2013;14:649.
44. Shanmugam S, Samonte SOPB, Sanchez DL, Yabes JC. Genetic assessment of diverse rice germplasm for yield-related and stress tolerance traits. Genet Resour Crop Evol. 2024. https://doi.org/10.1007/s10722-024-02004-4.
45. Fornasiero A, Wing RA, Ronald P. Rice domestication. Curr Biol. 2022;32:R20–4.
46. Molina J, Sikora M, Garud N, Flowers JM, Rubinstein S, Reynolds A, et al. Molecular evidence for a single evolutionary origin of domesticated rice. Proc Natl Acad Sci. 2011;108:8351–6.
47. Huang J, Zhao X, Yu H, Ouyang Y, Wang L, Zhang Q. The Ankyrin repeat gene family in rice: genome-wide identification, classification and expression profiling. Plant Mol Biol. 2009;71:207–26.
48. Hanada K, Zou C, Lehti-Shiu MD, Shinozaki K, Shiu S-H. Importance of Lineage-Specific expansion of plant tandem duplicates in the adaptive response to environmental stimuli. Plant Physiol. 2008;148:993–1003.
49. Gaines TA, Duke SO, Morran S, Rigon CAG, Tranel PJ, Küpper A, et al. Mechanisms of evolved herbicide resistance. J Biol Chem. 2020;295:10307–30.
50. Xu F, Lagudah ES, Moose SP, Riechers DE. Tandemly duplicated Safener-Induced glutathione S-Transferase genes from triticum Tauschii contribute to Genome- and Organ-Specific expression in hexaploid wheat. Plant Physiol. 2002;130:362–73.
51. Patterson EL, Pettinga DJ, Ravet K, Neve P, Gaines TA. Glyphosate resistance and EPSPS gene duplication: convergent evolution in multiple plant species. J Hered. 2018;109:117–25.
52. Iwakami S, Shimono Y, Manabe Y, Endo M, Shibaike H, Uchino A et al. Copy number variation in acetolactate synthase genes of Thifensulfuron-Methyl resistant alopecurus Aequalis (Shortawn Foxtail) accessions in Japan. Front Plant Sci. 2017;8.
53. Nawade B, Chu S-H, Kim S, Lee C-Y, Bao J, Park Y-J. Nested association mapping population in *Japonica* rice: development, characterization, and application in genome-wide association studies. Curr Plant Biology. 2024;40:100401.
54. Gupta P. Population genetics. In: Kar D, Sarkar S, editors. Genetics fundamentals notes. Singapore: Springer Nature; 2022. pp. 1077–103.
55. Bhat JA, Karikari B, Adeboye KA, Ganie SA, Barmukh R, Hu D, et al. Identification of superior haplotypes in a diverse natural population for breeding desirable plant height in soybean. Theor Appl Genet. 2022;135:2407–22.


## Publisher's note