PLANT SCIENCE TODAY
Vol 13(1): 1-9
https://doi.org/10.14719/pst.11864

eISSN 2348-1900

RESEARCH ARTICLE

Comparative *in-silico* analysis of the bHLH protein family in rice *Oryza sativa* subsp. *japonica* cv. Nipponbare

Roshani S[1], Parshant K S[2*], Ruchi Singh[1], Sharad Singh[3] & Vedant Gautam[4]

[1]Department of Genetics and Plant Breeding, Chandra Shekhar Azad University of Agriculture and Technology, Kanpur 208 002, Uttar Pradesh, India
[2]Department of Genetics and Plant Breeding, Banaras Hindu University Varanasi 221 005, Uttar Pradesh, India
[3]Department of Genetics and Plant Breeding, NDUAT University, Ayodhya 224229, Uttar Pradesh, India
[4]Department of Plant Pathology, Banaras Hindu University, Varanasi 221 005, Uttar Pradesh, India

*Correspondence email - prashantgpb@bhu.ac.in

Received: 19 September 2025; Accepted: 24 December 2025; Available online: Version 1.0: 10 February 2026; Version 2.0: 28 February 2026

Cite this article: Roshani S, Parshant KS, Ruchi S, Sharad S, Vedant G. Comparative *in-silico* analysis of the bHLH protein family in rice *Oryza sativa subsp. japonica cv.* Nipponbare. Plant Science Today. 2026; 13(1): 1-9. https://doi.org/10.14719/pst.11864

Abstract

The basic Helix-Loop-Helix (bHLH) proteins are key regulators of gene expression, development and responses to environmental stimuli in plants. In this study, we performed a comprehensive computational analysis of six bHLH proteins (OsbHLH056, OsbHLH057, OsbHLH058, OsbHLH059, OsbHLH062 and OsbHLH063) in *Oryza sativa subsp. japonica* cv. Nipponbare. Their physicochemical properties, secondary and tertiary structures, domains, phylogenetic relationships, subcellular localisation and protein-protein interactions were investigated. Results showed that OsbHLH062 had the highest molecular weight (29,686.27 kilodalton (kDa) and amino acid number (265), while OsbHLH058 had the lowest (7,896.08 kDa; 77 amino acids). Isoelectric point analysis indicated five proteins were acidic, while OsbHLH058 was basic. All proteins were predicted to be unstable, reflecting the flexibility essential for regulation. Aliphatic index values (61.13–85.26) suggested moderate thermo-stability. The secondary structure was dominated by α-helices, which enhance structural stability and the extinction coefficients suggested enrichment of cysteine, tryptophan and tyrosine residues. Phylogenetic analysis showed OsbHLH057 as the earliest ancestor among the six proteins. Subcellular localisation predictions identified nuclear targeting for all proteins, with OsbHLH059 showing the highest nuclear localisation probability. Protein-protein interaction analysis highlighted potential partners, implying roles in diverse cellular pathways. This study provides valuable insights into the molecular characteristics, structure and interactions of rice bHLH proteins. These findings form a foundation for experimental validation and functional characterisation. Understanding these proteins may enable the development of innovative strategies to enhance abiotic stress tolerance and crop productivity in rice.

Keywords: bHLH protein; *in-silico*; phylogenetic analysis; physiochemical properties

Introduction

More than half of the global population depends on rice (*Oryza sativa*), a staple crop increasingly threatened by erratic weather patterns and emerging diseases. Understanding the molecular basis of rice development and stress responses is essential for breeding new varieties with enhanced adaptability and yield. As global efforts intensify to enhance rice productivity, quality and stress resilience, a comprehensive understanding of the molecular mechanisms governing rice growth and development becomes increasingly essential (1). In recent times, the field of genomics and proteomics has seen the emergence of bioinformatics tools, including freely accessible online servers and databases. These tools have proven to be valuable resources for studying various organisms, spanning microorganisms, humans, animals and plants. They offer predictive capabilities and serve as foundational components for genomic and proteomic analyses. Notably, these tools contribute significantly to the exploration of hypothetical and conserved proteins by providing abundant data and insights into their characteristics and potential functions.

Comparative studies of rice bHLH (basic helix–loop–helix) proteins can significantly enhance our understanding of the intricate regulatory mechanisms that control rice development and stress responses. The bHLH transcription factors constitute second largest transcription factor family in plants and are involved in a wide and diverse array of biological processes (2). Research has demonstrated that bHLHs play a role in morphogenesis, iron homeostasis, root vascular cell proliferation, shoot branching and grain yield among other aspects of plant growth and development (3-6). Additionally, bHLHs have become key participants in plants' responses to biotic and abiotic stress (7, 8). Until now, 16 bHLH transcription factors of this type have been shown to be involved in the control of cellular iron homeostasis; it can be expected that this number is not yet complete (9). Despite this, studies elucidating the functions of plant bHLH proteins have been limited. Although, the substantial advances have been made in understanding the roles of particular bHLH proteins in different plant species, a thorough comparative investigation spanning a wider variety of rice bHLH proteins is still lacking. A significant knowledge gap remains in understanding the

broader landscape of bHLH-mediated transcriptional regulation, as most previous studies have focused on individual bHLH proteins or specific functional aspects in rice.

In the present study, we investigated six conserved rice bHLH proteins, namely OsbHLH059 (OsPRI3), OsbHLH058 (OsPRI2), OsbHLH056 (OsIRO2), OsbHLH063 (OsIRO3), OsbHLH057 (OsPRI4) and OsbHLH062 (OsbHLH1), selected for their established roles in iron homeostasis and stress regulation. Although OsPRI2, OsPRI3, OsIRO2 and OsIRO3 are known regulators of iron uptake and signalling, their interaction networks and regulatory behaviour remain largely unexplored. This study addresses this gap using integrative bioinformatics analyses to predict protein interactions, instability regions and potential cross-link with jasmonate signalling, providing new insights into bHLH-mediated regulation during iron stress in rice. Our approach involves determining the three-dimensional structure of these proteins to establish a connection between their structure and function. We also aim to predict how these proteins might interact with other proteins and to analyse their physical and chemical properties. By comparing their structures with those of known proteins, we can gain insights into the potential cellular roles of these bHLH proteins. Overall, our research aims to uncover the functions and characteristics of these unfamiliar proteins, contributing to our knowledge of cellular processes and their implications across various fields.

OsbHLH058/OsPRI2 and OsbHLH059/OsPRI3 have been discovered as positive regulators of iron deficiency responses (10). The OsHRZ2s' interactions with OsbHLH058 and OsbHLH059 control iron homeostasis reactions (11). Numerous iron-responsive genes, such as OsYSL2, OsNAS1 and OsTOM1, are activated by OsbHLH058. Whereas, OsbHLH059 has a beneficial impact on genes involved in a variety of processes, including internal iron translocation, iron (II) absorption, iron (III)-DMA uptake and translocation. It is found that OsbHLH056 to be an intriguing post-transcriptional component that contributes to improved salinity tolerance (12). Similar to this, *OsIRO3/OsbHLH063* is essential for maintaining plant survival in rice under Fe-deficient conditions and plays an important role in signal transmission from shoots to roots which is critical for plants to prevent Fe overload (13). Beyond iron metabolism, OsbHLH057 has been linked to the control of disease resistance and drought tolerance in rice. Deletion of OsbHLH057 renders rice more susceptible to infection and drought, whereas OsbHLH057 overexpression improves disease resistance and drought tolerance (14). In addition, OsbHLH062/OsbHLH1 is suggested to act as a transcription factor in the cold-signal transduction pathway and its interaction with OsIRO3 has been hypothesised to hold particular significance in iron-deficient plants. A key focus of the present study is abiotic stress, particularly iron deficiency. Given the central role of bHLH transcription factors in regulating gene expression networks, we hypothesise that these proteins play important roles in mediating plant responses to iron deficiency. Their complex regulatory functions likely contribute to stress adaptation by modulating downstream signalling and metabolic pathways. Using a bioinformatics-based approach, this study aims to identify candidate bHLH proteins involved in iron deficiency responses, thereby improving our understanding of the molecular mechanisms underlying plant adaptation to challenging environmental conditions.

Materials and Methods

Collection of data / Retrieval of amino acid sequences

The amino acid sequences of six bHLH proteins from *Oryza sativa* subsp. japonica cv. Nipponbare, namely OsbHLH059/OsPRI3, OsbHLH058/OsPRI2, OsbHLH056/OsIRO2, OsbHLH063/OsIRO3, OsbHLH057/OsPRI4 and OsbHLH062/OsbHLH1 (Table 1), were obtained from the Rice Genome Annotation Project Database. The Rice Genome Annotation Project Database is recognized globally for its excellence in computational research networks (15).

Primary structure prediction, physico-chemical properties and functional characterisation

ExPasy (Expert Protein Analysis System) ProtParam web application was used to predict the fundamental structural characteristics of protein sequences and calculate various physical and chemical properties (16). This online tool provides a range of analyses for proteins. These include determining the molecular weight, theoretical isoelectric point (pI), amino acid composition (including the ratio of positively and negatively charged amino acids), aliphatic index, extinction coefficient and the grand average of hydropathicity (GRAVY) score (17-19). Comparing these physicochemical properties

Table 1. The six bHLH proteins used in this study

Protein name/Gene symbol	Gene name (locus id)	Position	Length	Role	Reference
OsbHLH059/OsPRI3	Os02g0116600	Chr02 (-strand)	236AA	Positive regulator of iron deficiency response3 (Maintenance of Fe homeostasis).	(11)
OsbHLH058/OsPRI2	Os05g0455400	Chr05 (+strand)	77AA	Positive regulator of iron deficiency response 2 (Maintenance of Fe homeostasis).	(11)
OsbHLH056/OsIRO2	Os01g0952800	Chr01 (-strand)	247AA	Iron related transcription factor 2. Tolerance to Fe deficiency, Regulation of Fe uptake from soil, Fe transport during germination, Fe translocation to grain during seed maturation	(13, 43)
OsbHLH063/OsIRO3	Os03g0379300	Chr03(+strand)	252AA	Iron-related transcription factor 3 Iron-regulated bHLH transcription factor, Response to iron deficiency, Regulation of iron homeostasis	(11)
OsbHLH057/OsPRI4	Os07g0543000	Chr07 (-strand)	256AA	Basic helix-loop-helix (bHLH) transcription factor 057, Positive regulation of disease resistance, Drought tolerance	(14)
OsbHLH062/OsbHLH1	Os07g0628500	Chr07 (-strand)	265AA	bHLH transcription factor, Cold stress response. Basic helix-loop-helix dimerisation region bHLH domain containing protein.	(44)

Amino acid (AA)

provides valuable insights into the proteins' functional roles and molecular characteristics. This analytical approach is crucial for understanding proteins and their roles in various biological processes.

Secondary structure prediction

Using the Self-optimised prediction method with alignment (SOPMA), the secondary structural characteristics of proteins were predicted. This tool evaluates numerous aspects of the proteins' structure, including the percentage of alpha helix, beta turn, extended strands and random coils (20).

3D structure prediction and Validation

The Iterative Threading Assembly Refinement (I-TASSER) tool is widely used for predicting three-dimensional protein structures from amino acid sequences. I-TASSER is easily accessible and produces five thorough models for each protein (21). All five of the generated models through the I-TASSER methodology underwent rigorous validation using diverse computational techniques, including ERRAT and PROCHECK. The predicted models were verified through the ERRAT tool (22). ERRAT score shows the overall quality factor for non -bonded atomic interactions and the higher the score means better the quality of models (accepted range for a high-quality model is >50). The PROCHECK programme was used to analyse the protein structure with particular attention to its energy and stereochemical geometry. It resolves geometry residue by residue and undergoes pattern analysis to ascertain its dependability and internal consistency. The evaluation of the psi/phi Ramachandran plot obtained with PROCHECK allowed for the confirmation and examination of the backbones' existence.

Protein–protein interaction prediction

The Search Tool for the Retrieval of Interacting Genes/Proteins, STRING version 12.1, was utilised to predict protein-protein interactions. Using a scoring system to integrate information, this database combines known and predicted associations to produce thorough protein networks covering about 67'592'464 proteins from 14'094 organisms. These interactions derive from a variety of sources, including genetic context, high-throughput research, conserved expression patterns and known knowledge. They include both physical and functional relationships. The quantitative harmonisation of this plethora of data across different organisms allows for the transfer of knowledge between them when appropriate (23, 24). Proteins are assigned scores indicating the reliability of interactions to support confident interpretation: scores above 0.40 indicate medium confidence, scores above 0.70 indicate high confidence and scores above 0.90 indicate the highest level of confidence. As a result, proteins with scores higher than 0.444 were included in the subsequent studies.

Determination of soluble or transmembrane proteins

The SOSUI server (using SOSUI engine version 1.10) was used to determine a proteins' type, namely, whether it is a soluble entity or a transmembrane resident (25).

Prediction of protein sub-cellular localisation

Each hypothetical and conserved proteins' subcellular location was predicted by Deeplock 2.0. By submitting sequences in FASTA format, it is the most accurate web-based tool for the prediction of sub-cellular localisation in all protein sequences (26).

Identification of protein family and domain

Using InterPro version 95.0, protein families, domains and clans were identified. The InterPro platform provides a complex functional analysis of proteins by classifying them into different families and predicting their domains and important locations (27).

Phylogenetic analysis

The software Molecular Evolutionary Genetic study (MEGA version 11.0.13) was used to examine the evolutionary relationship between different bHLH proteins selected in this study (28). A neighbour-joining (NJ) phylogenetic tree was constructed. The reliability of the branching was tested using bootstrap resampling (with 1000 pseudo -replicates).

Results and Discussion

Physicochemical properties of bHLH proteins

The physicochemical properties of each bHLH protein computed by ProtParam, including the amino acid composition (Table S1), molecular weight, theoretical pI, extinction coefficient, instability index, aliphatic index, grand average of hydropathicity (GRAVY) and total number of negatively and positively charged residues, are mentioned in Table S2.

Molecular weight

Proteins with a high molecular weight contain a higher proportion of amino acids in them than proteins with a lower molecular weight (29). Among the *O. sativa* bHLH proteins analysed, OsbHLH058 constituted 77 amino acids with the lowest molecular weight (7896.08kDa), while OsbHLH062 had 265 amino acids with the highest molecular weight (29686.27kDa) (Table S2).

Isoelectric point

In order to comprehend the charge stability of proteins, it is helpful to know the theoretical pI, which is the pH at which a specific molecule or surface carries no net electrical charge. If a proteins' pI value is lower than 7, it is considered acidic, while a pI value greater than 7 denotes the proteins' basic nature. According to estimated pI values from the ProtParam programme, OsbHLH056 was observed to be a basic protein (pI>7), but OsbHLH057, OsbHLH059, OsbHLH062 and OsbHLH063 were determined to have acidic nature (pI<7) (Table S2). The protein purification procedure heavily depends on this isoelectric point.

Instability Index

The stability index, an innate property of proteins, helps predict protein in vivo stability. If a proteins' instability index value is less than 40, it is said to be stable; if it is greater than 40, it is said to be unstable. In the present study, we observed that the instability score of all bHLH proteins is greater than 40 (Table S2), which suggests that all the proteins are unstable.

Aliphatic index (AI)

The aliphatic index of a protein is defined as the relative volume occupied by aliphatic side chains [alanine (Ala), valine (Val), isoleucine (Ile) and leucine (Leu)]. It could be seen as a contributing factor to the rise in globular proteins' thermo-stability. The 6 bHLH protein sequences used in this study had aliphatic indices ranging from 61.13 to 85.26 (Table S2). A high aliphatic index indicates that a protein is thermostable over a wide temperature range. (17)

GRAVY (Grand average of hydropathy)

The GRAVY value for a peptide or protein is determined as the total

hydropathy value of all the amino acids divided by the total number of residues in the sequence. Proteins with a negative GRAVY value are non-polar, while those with a positive value are polar. The GRAVY indices of the bHLH proteins in this investigation ranged from -0.824 to 0.194 (Table S2). The low GRAVY range suggests that the protein might be globular (hydrophilic), as opposed to membranous (hydrophobic). This information might be useful for localising these proteins.

Extinction coefficient

For the six bHLH proteins, their respective extinction coefficients at 280 nm range from 13200 to 22460 (as shown in Table S2) and this measurement is linked to the levels of cysteine (Cys), tryptophan (Trp) and tyrosine (Tyr) within the proteins. Essentially, it gauges how effectively a protein captures light at a given wavelength. A higher extinction coefficient suggests an elevated concentration of Cys, Trp and Tyr within the presumed or conserved proteins (16). These calculated extinction coefficients are valuable for studies examining interactions between proteins and interactions between proteins and ligands in a solution.

Secondary structure prediction

The secondary structure of six bHLH proteins was predicted using SOPMA, wherein we found that the percentage of random coils was higher when compared with alpha helices and extended turns in all the selected bHLH proteins except OsbHLH056(Table 2). Random coils are inherently flexible and lack a stable structure. These flexible regions might play crucial roles in protein function, such as in binding to other molecules, protein folding, translocation and stability. The analysis of the amino acid content of the different proteins reveals that five essential amino acids predominate. Alanine holds the top spot among them with a prevalence of 27.3 %, followed by Proline (18.2 %), Serine (13.4 %), Glycine (8.9 %) and Glutamic Acid (8.7 %) (Table S1).

Tertiary(3D) structure prediction and model validation

Top threading Template used by I-TASSER

For six bHLH proteins, I-TASSER determined the identity and coverage area for the top ten best template PDB structures (Table S3 -S8f). I-TASSER only employs the threading alignment templates with the greatest importance, as determined by the Z-score, which is the difference between the raw and average scores expressed in units of standard deviation. Therefore, a categorised Z-score greater than 1 denotes a certain alignment. If there is often at least one template per threading programme with a categorised Z-score > 1, the query protein is categorised as an "Easy" target; otherwise, it is a "Hard" target (30). Hence, on the basis of this, we only consider those threading templates whose score is greater than 1.

Top 5 model

For every bHLH protein, the I-TASSER server produces several models, each assigned a unique C-score. The C-score is a metric used to assess alignment accuracy and assembly convergence during protein structure prediction, with a range between -5 and 2. Higher C -scores indicate more accurate predictions. Among the various models created for each protein, the one with the highest C-score is chosen as the top model, reflecting superior alignment quality and convergence in the prediction process. In the case of the OsbHLH059 protein, the model with the highest C-score (-1.89) out of all generated models is selected as the best model (Fig. 1). The estimated accuracy of this model is assessed through TM-score and RMSD measurements, yielding values of 0.49 ± 0.15 and 10.0 ± 4.6Å, respectively. Similarly, for the other proteins, OsbHLH058, OsbHLH056, OsbHLH063, OsbHLH057 and OsbHLH062, the top models possess C-scores of -3.23 (with TM score and RMSD values of 0.35 ± 0.12 and 10.5 ± 4.6Å), -3.55 (TM score and RMSD values of 0.32 ± 0.11 and 14.3 ± 3.8Å), -3.56 (TM score and RMSD values of 0.32 ± 0.11 and 14.4 ± 3.7Å), -3.36 (TM score and RMSD values of 0.34 ± 0.11 and 13.9 ± 3.9Å) and 2.70 (TM score and RMSD values of 0.40 ± 0.14 and 12.3 ± 4.4Å) (Figure 1). By employing TM-score and RMSD data, representing structural similarity and deviation from the actual

Table 2. The predicted secondary structural characteristics of bHLH proteins (in %) by SOPMA

Protein	Alpha helix (%)	Extended strand (%)	Beta turn (%)	Random coil (%)
OsbHLH056	44.94	9.31	4.05	41.7
OsbHLH057	37.89	3.12	3.52	55.47
OsbHLH058	19.48	7.79	3.9	68.83
OsbHLH059	37.29	8.47	3.81	50.42
OsbHLH062	33.58	6.42	0.75	59.25
OsbHLH063	35.71	3.17	0.79	60.32

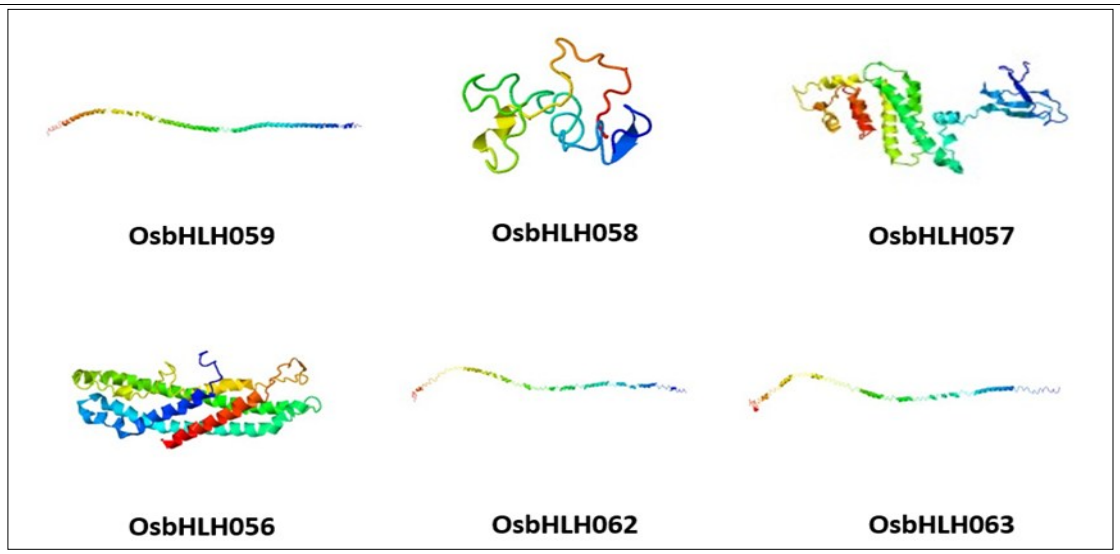

Fig. 1. Models with the highest C score of each bHLH protein.

structure, respectively, the predicted accuracy of each model is quantified. This process of model selection based on C-scores, followed by accuracy estimation, offers valuable insights into the quality of protein structure predictions obtained through the I-TASSER methodology.

Homology modelling and Validation

For homology modelling, the I-TASSER approach was employed. Five 3D models were created for each of the six bHLH proteins. For validation and to evaluate the quality of these models, ERRAT and PROCHECK (Table 3) were used. ERRAT (Overall Quality Factor), which rates the models' overall quality by statistically analysing non-bonded interactions among various atom types in their corresponding atomic interactions, was one of the approaches used for validation. This methodology was used for model validation (31, 32). The generated quality score is used to evaluate the developed models' dependability. The typical range for a high-quality model is a Quality score greater than 50. In this work, the best models of the six bHLH proteins had ERRAT scores of 95.17 %, 65.21 %, 84.10 %, 73.36 %, 82.91 % and 91.28 %, respectively (Table 3).

PROCHECK was also used to assess the stereochemical quality of protein structures. This evaluation involved using the Ramachandran plot, which has separate zones showing most favoured, allowed and prohibited residues, to analyse residue-by-residue geometry as well as overall structure geometry. PROCHECK computations suggest that superior models should have more than 90 % of amino acids in the most desired locations. The highest PROCHECK scores for the bHLH proteins under study (representing the proportion of residues in the preferred region for the best model) were, in order, OsbHLH059, OsbHLH056, OsbHLH057, OsbHLH063, OsbHLH062 and OsbHLH058 (Table 3). Ramachandran plot for each protein is shown in Fig. 2.

Protein–protein interaction prediction

Predicting protein-protein interactions is essential to understanding the complex web of cellular and biological processes. A proteins' interaction with two or more other proteins is frequently a crucial aspect of that proteins' biological function. We used STRING to predict connections between our bHLH proteins and their counterparts (Fig. 3) as we were interested in functional protein

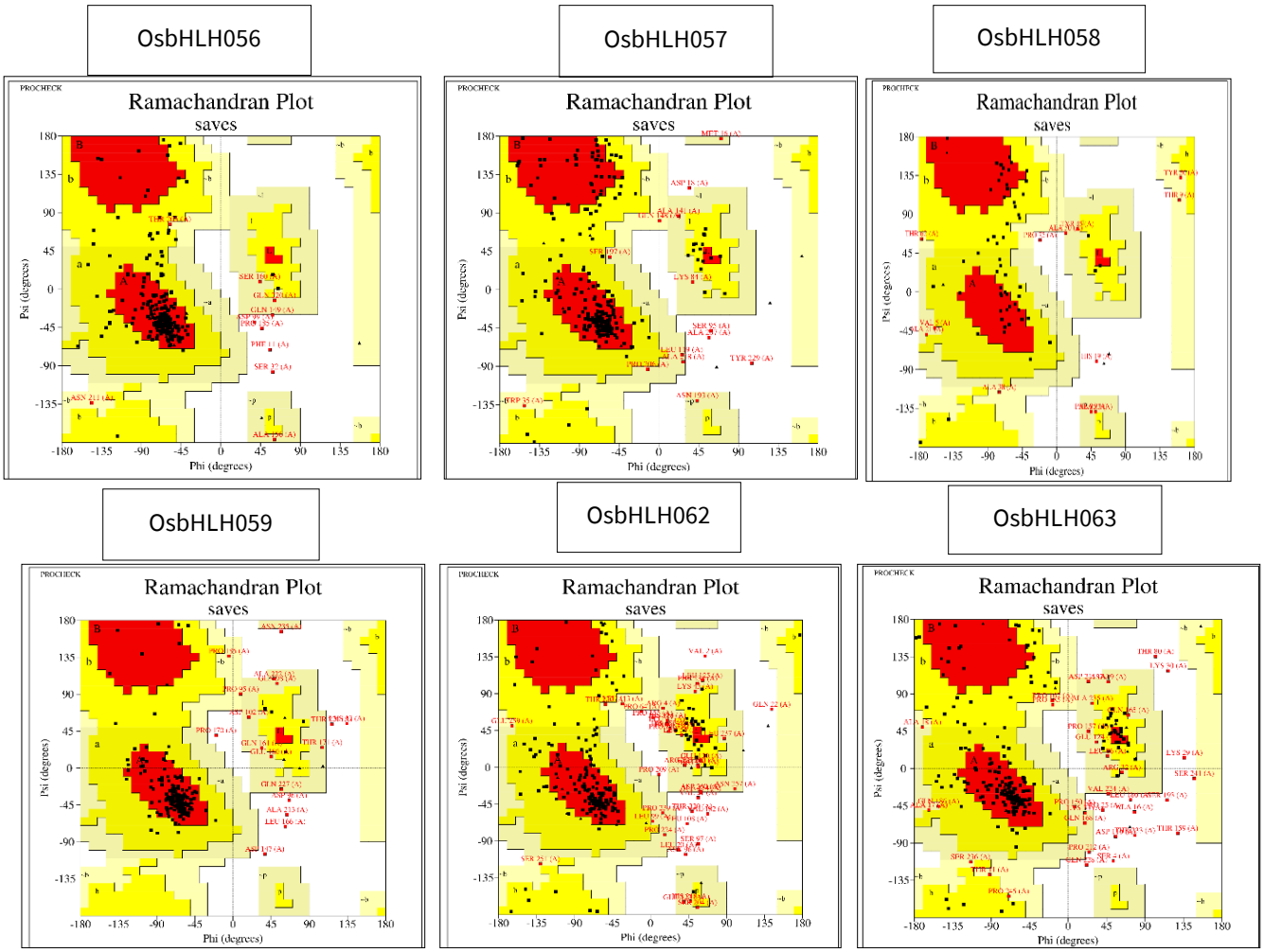

Fig. 2. Ramachandran plot analysis showing the structural validation of modelled bHLH proteins.

Table 3. Validation Scores of different 3D models of six bHLH proteins generated via I-TASSER.

Proteins	Model C Score	ERRAT Score	PROCHECK		
			Ramachandran plot result		
			Most favourable region (%)	Allowed region (%)	Disallowed region (%)
OsbHLH059	-1.89	95.17	83.9	8.9	3.6
OsbHLH058	-3.23	65.21	35.1	45.6	1.8
OsbHLH056	-3.55	84.10	74.4	21.5	1.8
OsbHLH063	-3.56	73.36	56.3	30.6	6.3
OsbHLH057	-3.36	82.91	65.7	28.1	1.9
OsbHLH062	-2.70	91.82	53.8	31.6	4.4

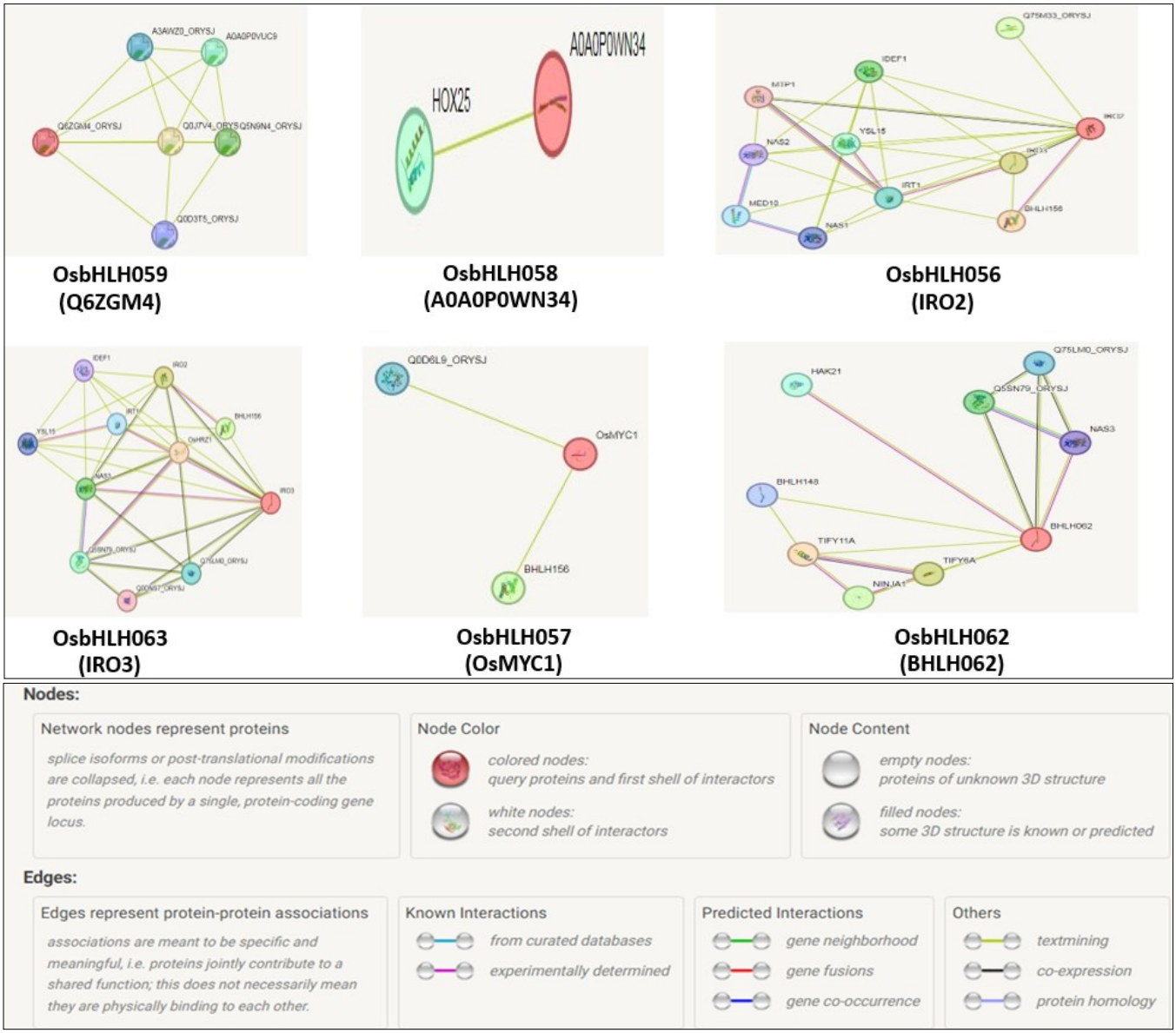

Fig. 3. Protein–protein interaction network established by STRING (Version 12.0) for six bHLH proteins and their partners.

association networks. Our analysis specifically identified interaction partners with scores greater than 0.4. While certain proteins function alone, the majority of proteins interact with other proteins to carry out their intended biological roles. Therefore, the predicted interaction networks provide insights into the biological pathways and regulatory mechanisms in which these bHLH transcription factors may participate. This interaction suggests that OsbHLH059 may be regulated through palmitoylation, a lipid-based post-translational modification known to influence protein stability, subcellular localisation and transcriptional activity. Such regulation may fine-tune OsbHLH059 function during plant development or stress responses. In eukaryotes, protein S-acylation (palmitoylation) plays a critical role in plant growth and responses to both abiotic and biotic stresses (33). Additionally, the homeodomain–leucine zipper (HD-ZIP) gene family is involved in multiple physiological processes in rice. The predicted interaction of OsbHLH058 with an HD-ZIP family member suggests a possible link between bHLH- and HD-ZIP-mediated transcriptional regulation, potentially coordinating developmental or stress-responsive gene expression pathways. However, comprehensive analysis of HD-ZIP genes in rice remains limited (34).With respect to OsbHLH056, four different proteins interact with it: B3 domain-containing protein IDEF1, a transcription regulator crucial for iron deficiency response and tolerance; iron-phyto siderophore transporter YSL15, important in Fe(3+) uptake from the root; Transcription factor BHLH156, linked to positive regulation of genes crucial to strategy II iron acquisition and an iron-related transcription factor. The association of OsbHLH056 with multiple iron homeostasis-related proteins strongly indicates its involvement in strategy II iron acquisition and transcriptional regulation of iron deficiency-inducible genes, highlighting its potential biological role in maintaining iron balance under nutrient-limited conditions (35, 36).

Similarly, OsbHLH063 interacts with five proteins, including iron-related transcription factor 2, which activates Fe deficiency-responsive genes by binding to the CACGTGG motif and transcription factor bHLH156 (36, 37). These interactions suggest that OsbHLH063 may function as a transcriptional hub within iron-responsive regulatory networks, possibly contributing to adaptive responses under iron-deficient conditions. OsbHLH062 interacts with Protein TIFY11a and NINJA homolog 1, along with transcription factor bHLH148. The involvement of TIFY/JAZ and NINJA proteins, which are central repressors in the jasmonic acid (JA) signalling pathway, indicates that OsbHLH062 may participate in JA-mediated stress and defence signalling. This interaction implies a biological role in coordinating hormonal signalling with transcriptional regulation during drought or stress responses. Jasmonates regulate

Table 4. 6bHLH proteins with their functionally important partner proteins

Protein	Predicted functional partner proteins	Score
OsbHLH059	S-acyltransferase; belongs to the DHHC palmitoyltransferase family.	0.456
OsbHLH058	Homeobox-leucine zipper protein HOX25	0.405
OsbHLH056	Transcription factor BHLH156	0.774
	Protein iron-related transcription factor 3	0.623
	B3 domain-containing protein IDEF1	0.560
	Iron-phytosiderophore transporter YSL15	0.509
OsbHLH063	Hemerythrin motif	0.629
	Protein iron-related transcription factor 2	0.623
OsbHLH057	Transcription factor bHLH 156	0.446
OsbHLH062	Protein TIFY 11a	0.970
	Protein NINJA homolog 1	0.796
	Transcription factor bHLH148	0.444

plant development, defence mechanisms and abiotic stress tolerance (38, 39). Further, Table 4 shows the predicted partner proteins and corresponding confidence scores for each of the bHLH proteins examined in the present study. Overall, the interaction networks suggest that these OsbHLH proteins are not isolated regulators but are integrated into key biological pathways, particularly iron homeostasis and jasmonate signalling, underscoring their functional relevance in rice growth and stress adaptation.

Prediction of soluble and transmembrane proteins

SOSUI analysis predicted that all six bHLH proteins are soluble, indicating their hydrophilic nature. The soluble nature of these proteins is consistent with their predicted nuclear localisation and transcriptional regulatory functions, as soluble proteins are more suitable for dynamic interactions with DNA and other nuclear proteins. Membrane proteins are generally more challenging to characterise experimentally due to their intrinsic hydrophobicity (40).

Prediction of protein sub-cellular localisation

The localisation of the proteins within the cell was predicted using the Deeplock2.0 programme (Table S9). OsbHLH059 had the highest likelihood score of all six bHLH proteins, followed by OsbHLH057, OsbHLH056 and OsbHLH063 in that order. All six bHLH proteins were shown to be localised in the nucleus. This nuclear localisation supports their predicted function as transcription factors and aligns with their interaction partners, many of which are also transcriptional regulators involved in iron homeostasis and JA signalling.

Protein family and domains

Proteins often have one or more domains, each of which has unique structural and/or functional characteristics. We searched the Pfam (InterPro) database, which made it easier to classify the proteins in our investigation into different families based on the presence or lack of particular domains in their sequences. Table S10 provides a thorough breakdown of the Pfam domains and families found in the six OsbHLH proteins. Notably, the Myc-type basic helix-loop-helix (bHLH) domain was clearly identified by the Pfam search found in all proteins except for Osbhlh058. The presence of this conserved domain confirms their DNA-binding and dimerisation capacity, which is essential for transcriptional regulation. *OSB1* gene encodes a myc-type basic helix-loop-helix (bHLH) transcription factor, which controls the expression of several structural genes involved in anthocyanin biosynthesis in rice (41). In *Oryza sativa*, the OsMYC2-RNAi lines showed enhanced resistance against bacterial pathogen *Xanthomonas oryzae* (42). This suggests that bHLH proteins may also contribute to defence-related gene regulation.

The Transcription Factor ILR3-like Family was further elucidated by the Pfam (InterPro) search. This family of homo- and heterodimeric transcription factors is intimately related to the synthesis of bHLH38, bHLH39, bHLH100 and bHLH101, directing the positive control of Fe homeostasis (13). Notably, three of the six proteins, OsbHLH056, OsbHLH063 and OsbHLH062, which make up the total, do not belong to this specific family. Table S10 provides a thorough explanation of each Pfam (InterPro) family and domain. A group of conserved domain models with overlapping annotations on the same protein sequences is referred to as a superfamily. These models may cross paths with one another, indicating their shared evolutionary ancestry. Fortunately, the super-families that corresponded to the discovered Pfam families showed a substantial degree of closeness, as seen in Tables S10-S11.

Phylogenetic analysis

Using MEGA 11 software, a phylogenetic tree was constructed to investigate the evolutionary relationship between the chosen bHLH proteins. The findings show links between several bHLH within the context of evolution. OsbHLH057 developed earlier than the other five bHLH proteins, according to the phylogenetic tree. From OsbHLH057, the remaining five bHLH proteins developed.

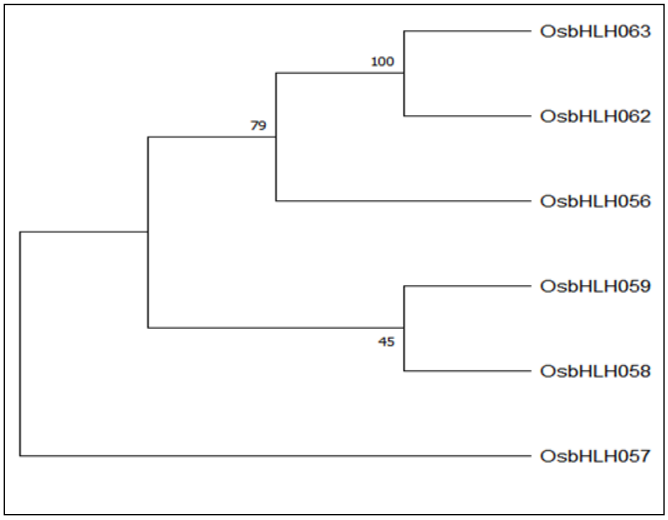

Fig. 4. Phylogenetic tree prepared with 6 bHLH proteins by the maximum likelihood method using MEGA 11 software.

OsbHLH056 is the common ancestor of OsbHLH063 and OsbHLH062. Additionally, OsbHLH059 and OsbHLH058 have a common ancestor that once split from OsbHLH057. Due to the similarity of their divergence branches, OsbHLH063 and OsbHLH062 may have evolved simultaneously (Fig. 4).

Conclusion

This study presents a comprehensive computational analysis of six rice bHLH proteins (OsbHLH056, OsbHLH057, OsbHLH058, OsbHLH059, OsbHLH062 and OsbHLH063), revealing distinct molecular, structural and evolutionary characteristics. The proteins differed in size, charge and stability, with α-helix dominance indicating conserved structural organisation and predicted nuclear localisation supporting their roles as transcription factors. Phylogenetic analysis suggested evolutionary divergence from a common ancestor, while protein–protein interaction predictions indicated involvement in multiple regulatory pathways. Collectively, these findings provide valuable insights into the functional complexity of rice bHLH proteins and establish a foundation for experimental validation. Future studies should employ yeast two-hybrid or bimolecular fluorescence complementation assays to confirm predicted protein interactions and site-directed mutagenesis to evaluate the functional significance of unstable regions, thereby advancing the understanding of bHLH-mediated stress regulation and its potential application in crop improvement.

Acknowledgements

The authors would like to express his sincere gratitude to the Department of Genetics and Plant Breeding at the Institute of Agricultural Sciences, Banaras Hindu University (Varanasi, UP, India), for their invaluable support and guidance. Additionally, we are deeply thankful to the UGC for their generous provision of the NET fellowship, which has been instrumental in facilitating my research and academic pursuits. Your assistance has been instrumental in my academic success.

Authors' contributions

RS, PKS, RS, SS and VG designed the study and performed the data analysis. RS carried out the *in-silico* experiments and prepared the initial manuscript draft. PKS, RS, SS and VG supervised the research, contributed to the interpretation of results, and critically revised the manuscript . All authors read and approved the final manuscript.

Compliance with ethical standards

Conflict of interest: Authors do not have any conflict of interests to declare.

Ethical issues: None

References

1. Mthiyane P, Aycan M, Mitsui T. Strategic advancements in rice cultivation: combating heat stress through genetic innovation and sustainable practices: A review. Stresses. 2024;4(3):452–80. https://doi.org/10.3390/biology13090659
2. Sun PW, Gao ZH, Lv FF, Yu CC, Jin Y, Xu YH, et al. Genome-wide analysis of basic helix–loop–helix (bHLH) transcription factors in *Aquilaria sinensis*. Sci Rep. 2022;12(1):7194. https://doi.org/10.1038/s41598-022-10785-w
3. Khan I, Asaf S, Jan R, Bilal S, Khan AL, Kim KM, et al. Genome-wide annotation and expression analysis of WRKY and bHLH transcription factor families reveal their involvement under cadmium stress in tomato (*Solanum lycopersicum* L.). Front Plant Sci. 2023; 14:1100895. https://doi.org/10.3389/fpls.2023.1100895
4. Berger N, Marin AJ, Stassen MJ, Lourenço T, Li M, Watanabe S, et al. Molecular regulation of iron homeostasis in plants. In: Lüttge U, Cánovas FM, Risueño Almeida MC, Leuschner C, Pretzsch H, editors. Progress in Botany Vol. 85. Cham: Springer; 2023. p. 75–103. https://doi.org/10.1007/124_2023_76
5. Xufeng X, Yuanfeng H, Ming Z, Shucheng S, Haonan Z, Weifeng Z, et al. Transcriptome profiling reveals the genes involved in tuberous root expansion in *Pueraria montana* var. *thomsonii*. BMC Plant Biol. 2023;23(1):1–19. https://doi.org/10.1186/s12870-023-04303-x
6. Zhang L, Xiang Z, Li J, Wang S, Chen Y, Liu Y, et al. bHLH57 confers chilling tolerance and grain yield improvement in rice. Plant Cell Environ. 2023;46(4):1402–18. https://doi.org/10.1111/pce.14513
7. Zou T, Wang X, Sun T, Rong H, Wu L, Deng J, et al. MYB transcription factor OsC1PLSr involves the regulation of purple leaf sheath in rice. Int J Mol Sci. 2023;24(7):6655. https://doi.org/10.3390/ijms24076655
8. Zuo ZF, Lee HY, Kang HG. Basic helix–loop–helix transcription factors: Regulators for plant growth, development and abiotic stress responses. Int J Mol Sci. 2023;24(2):1419. https://doi.org/10.3390/ijms24021419
9. Schmidt W, Thomine S, Buckhout TJ. Iron nutrition and interactions in plants. Front Plant Sci. 2020; 10:1670. https://doi.org/10.3389/fpls.2019.01670
10. Shekhawat PK, Ram H, Soni P. The regulatory circuit of iron homeostasis in rice: a tale of transcription factors. In: Srivastava V, Mishra S, Mehrotra S, Upadhyay SK, editors. Plant transcription factors. London: Academic Press; 2023. p. 251–68. https://doi.org/10.1016/B978-0-323-90613-5.00015-7
11. Zhang H, Li Y, Pu M, Xu P, Liang G, Yu D, et al. *Oryza sativa* positive regulator of iron deficiency response 2 (OsPRI2) and OsPRI3 are involved in the maintenance of Fe homeostasis. Plant Cell Environ. 2020;43(1):261–74. https://doi.org/10.1016/B978-0-323-90613-5.00015-7
12. Aycan M, Nahar L, Baslam M, Mitsui T. B-type response regulator hst1 controls salinity tolerance in rice by regulating transcription factors and antioxidant mechanisms. Plant Physiol Biochem. 2023; 196:542–55. https://doi.org/10.1016/j.plaphy.2023.02.008
13. Wang F, Itai RN, Nozoye T, Kobayashi T, Nishizawa NK, Nakanishi H. The bHLH protein OsIRO3 is critical for plant survival and iron homeostasis in rice (*Oryza sativa L.*) under Fe-deficient conditions. Soil Sci Plant Nutr. 2020;66(4):579–92. https://doi.org/10.1080/00380768.2020.1783966
14. Liu J, Shen Y, Cao H, He K, Chu Z, Li N. OsbHLH057 targets the AATCA cis-element to regulate disease resistance and drought tolerance in rice. Plant Cell Rep. 2022;41(5):1285–99. https://doi.org/10.1007/s00299-022-02859-w
15. Sakai H, Lee SS, Tanaka T, Numa H, Kim J, Kawahara Y, et al. Rice Annotation Project Database (RAP-DB): an integrative and interactive database for rice genomics. Plant Cell Physiol. 2013;54 (2): e6. https://doi.org/10.1093/pcp/pcs183
16. Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. In: Walker JM, editor. The proteomics protocols handbook. Totowa (NJ): Humana Press; 2005.
17. Ikai A. Thermostability and aliphatic index of globular proteins. J Biochem. 1980;88(6):1895–98. https://doi.org/10.1093/oxfordjournals.jbchem.a133168
18. Gill SC, Von Hippel PH. Calculation of protein extinction coefficients from amino acid sequence data. Anal Biochem. 1989;182(2):319–26. https://doi.org/10.1016/0003-2697(89)90602-7

19. Kyte J, Doolittle RF. A simple method for displaying the hydropathic character of a protein. J Mol Biol. 1982;157(1):105–32. https://doi.org/10.1016/0022-2836(82)90515-0

20. Geourjon C, Deleage G. SOPMA: significant improvements in protein secondary structure prediction by consensus prediction from multiple alignments. Bioinformatics. 1995;11(6):681–84. https://doi.org/10.1093/bioinformatics/11.6.681

21. Zhang Y. I-TASSER server for protein 3D structure prediction. BMC Bioinformatics. 2008; 9:40. https://doi.org/10.1186/1471-2105-9-40

22. Colovos C, Yeates TO. Verification of protein structures: patterns of nonbonded atomic interactions. Protein Sci. 1993;2(9):1511–9. https://doi.org/10.1002/pro.5560020916

23. Szklarczyk D, Franceschini A, Kuhn M, Simonovic M, Roth A, Minguez P, et al. The string database in 2011: functional interaction networks of proteins, globally integrated and scored. Nucleic Acids Res. 2011;39(Suppl_1): D561–8. https://doi.org/10.1093/nar/gkq973

24. Franceschini A, Szklarczyk D, Frankild S, Kuhn M, Simonovic M, Roth A, et al. STRING v9.1: protein-protein interaction networks, with increased coverage and integration. Nucleic Acids Res. 2012;41(D1): D808–15. https://doi.org/10.1093/nar/gks1094

25. Hirokawa T, Boon-Chieng S, Mitaku S. SOSUI: classification and secondary structure prediction system for membrane proteins. Bioinformatics. 1998;14(4):378–79. https://doi.org/10.1093/bioinformatics/14.4.378

26. Gardy JL, Laird MR, Chen F, Rey S, Walsh CJ, Ester M, et al. PSORTb v. 2.0: expanded prediction of bacterial protein subcellular localization and insights gained from comparative proteome analysis. Bioinformatics. 2005;21(5):617–23. https://doi.org/10.1093/bioinformatics/bti057

27. Paysan-Lafosse T, Blum M, Chuguransky S, Grego T, Pinto BL, Salazar GA, et al. InterPro in 2022. Nucleic Acids Res. 2023;51(D1): D418–27. https://doi.org/10.1093/nar/gkac993

28. Tamura K, Stecher G, Kumar S. MEGA11: molecular evolutionary genetics analysis version 11. Mol Biol Evol. 2021;38(7):3022–7. https://doi.org/10.1093/molbev/msab120

29. Guan Y, Zhu Q, Huang D, Zhao S, Lo LJ, Peng J. An equation to estimate the difference between theoretically predicted and SDS PAGE-displayed molecular weights for an acidic peptide. Sci Rep. 2015; 5:13370. https://doi.org/10.1038/srep13370

30. Yang J, Zhang Y. I-TASSER server: new development for protein structure and function predictions. Nucleic Acids Res. 2015;43: W174–81. https://doi.org/10.1093/nar/gkv342

31. Bowie JU, Luthy R, Eisenberg D. A method to identify protein sequences that fold into a known three-dimensional structure. Science. 1991; 253:164–70. DOI: 10.1126/science.1853201

32. Sohail R, Kalsoom S, Masood R, Tahir Y, Aftab AA, Muhammad I, et al. *In silico* analysis of four structural proteins of aphthovirus serotypes revealed significant B and T cell epitopes. Microb Pathog. 2019; 128:254–62. https://doi.org/10.1016/j.micpath.2019.01.007

33. Li J, Zhang M, Zhou L. Protein S-acyltransferases and acyl protein thioesterases, regulation executors of protein S-acylation in plants. Front Plant Sci. 2022; 13:956231. https://doi.org/10.3389/fpls.2022.956231

34. Zhang Y, Ge J, Huang B, Tong L, Shu Y, Qi C, et al. Genome-wide identification and expression analysis of the homeodomain–leucine zipper family in rice (*Oryza sativa L.*) under abiotic stress and in seed development. Preprint. 2022 Nov 30; Version. https://doi.org/10.21203/rs.3.rs-2324496/v1

35. Kobayashi T, Ogo Y, Nakanishi Itai R, Nakanishi H, Takahashi M, Mori S, et al. The novel transcription factor IDEF1 regulates iron-deficiency response and tolerance. Proc Int Plant Nutr Colloq XVI. 2009.

36. Liang G, Zhang H, Li Y, Pu M, Yang Y, Li C, et al. Fer-like fe deficiency-induced transcription factor (OsFIT) interacts with OsIRO2 to regulate iron homeostasis. bioRxiv. 2020;2020-03. https://doi.org/10.1101/2020.03.06.981126

37. Jang YH, Park JR, Kim EG, Kim KM. OsbHLHq11, the basic helix-loop-helix transcription factor, involved in regulation of chlorophyll content in rice. Biology. 2022;11(7):1000. https://doi.org/10.3390/biology11071000

38. Ebel C, BenFeki A, Hanin M, Solano R, Chini A. Characterization of wheat (*Triticum aestivum*) TIFY family and role of Triticum durum Td TIFY11a in salt stress tolerance. PLoS One. 2018;13(7): e0200566. https://doi.org/10.1371/journal.pone.0200566

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

40. Hedin LE, Illergard K, Elofsson A. An introduction to membrane proteins. J Proteome Res. 2011;10(8):3324–31. https://doi.org/10.1021/pr200145a

41. Sakulsingharoj C, Inta P, Sukkasem R, Pongjaroenkit S, Chowpongpang S, Sangtong V. Cloning and characterization of OSB1 gene controlling anthocyanin biosynthesis from Thai black rice. Genomics Genet. 2016;9(1):7–18. https://doi.org/10.14456/gag.2016.2

42. Giri MK, Gautam JK, Babu Rajendra Prasad V, Chattopadhyay S, Nandi AK. Rice MYC2 (OsMYC2) modulates light-dependent seedling phenotype, disease defence but not ABA signalling. J Biosci. 2017; 42:501–08. https://doi.org/10.1007/s12038-017-9703-8

43. Ogo Y, Itai RN, Kobayashi T, Aung MS, Nakanishi H, Nishizawa NK. OsIRO2 is responsible for iron utilization in rice and improves growth and yield in calcareous soil. Plant Mol Biol. 2011; 75:593–605. https://doi.org/10.1007/s11103-011-9752-6

44. Wang YJ, Zhang ZG, He XJ, Zhou HL, Wen YX, Dai JX, et al. A rice transcription factor OsbHLH1 is involved in cold stress response. Theor Appl Genet. 2003; 107:1402–9. https://doi.org/10.1007/s00122-003-1378-x

Additional information

Peer review: Publisher thanks Sectional Editor and the other anonymous reviewers for their contribution to the peer review of this work.

Reprints & permissions information is available at https://horizonepublishing.com/journals/index.php/PST/open_access_policy

Publisher's Note: Horizon e-Publishing Group remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Indexing: Plant Science Today, published by Horizon e-Publishing Group, is covered by Scopus, Web of Science, BIOSIS Previews, Clarivate Analytics, NAAS, UGC Care, etc
See https://horizonepublishing.com/journals/index.php/PST/indexing_abstracting

Copyright: © The Author(s). This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited (https://creativecommons.org/licenses/by/4.0/)

Publisher information: Plant Science Today is published by HORIZON e-Publishing Group with support from Empirion Publishers Private Limited, Thiruvananthapuram, India.