



Screening of heat stress-tolerant weedy rice and SNP identification of heat-tolerance-related genes

Mainul Hasan Sarker¹ · Md Hamed Hussain¹ · Ting Xiang Neik² · Md Zobaer Hasan^{1,3} · Wei Yee Wee¹ · Hock Siew Tan¹ · Swee-suak Ko^{4,5} · Beng-Kah Song¹

Received: 7 February 2024 / Revised: 9 July 2024 / Accepted: 16 July 2024 / Published online: 31 July 2024
© The Author(s) 2024

Abstract

Rice, a staple crop that feeds more than one-third of the world's population, encounters a wide range of biotic and abiotic stresses due to climate change. Rising temperature is one of the significant abiotic stresses affecting rice productivity worldwide. The development of heat-tolerant rice cultivars is critical in this regard. Weedy rice could potentially serve as a natural resource for genes conferring agronomically important traits beneficial to cultivated rice. However, heat tolerance in both cultivated and weedy rice is still understudied. This study screened a set of 180 weedy rice accessions for heat stress tolerance and further characterised them using genome-wide single-nucleotide polymorphisms (SNPs) analysis. Five heat-tolerant (HT) accessions (MU244, MU235, MU249, MU260 and MU237), along with five heat-susceptible (HS) accessions (MU100, MU114, MU264, MU251 and MU005), were subjected to relative electrical conductivity (REC) test and reactive oxidative species assay (ROS). These tests verified that the five HT accessions performed better under heat stress than their HS counterparts. In addition, whole-genome sequences of three HT (MU235, MU237 and MU066) and four HS (MU100, MU114, MU022 and MU005) accessions were selected for the genome-wide SNPs comparison, revealing substantial amino acid variation in the heat-tolerance-related genes between the HT and HS rice groups. The proposed genes and genome-wide SNP markers may help rice breeders better understand how different rice cultivars respond to heat stress.

Keywords Genome-wide SNPs · Heat shock protein · Heat stress · Weedy rice

Abbreviations

BH	Blackhull-awnless
BHA	Blackhull-awned
BR	Browhull-awnless
BRA	Browhull-awned
DAB	3, 3'-Diaminobenzidine
HS	Heat-susceptible
HSF	Heat shock factor

HSP	Heat shock protein
HT	Heat-tolerant
HTRGs	Heat-tolerance-related genes
HTI	Heat tolerance index
mSHA	Intermediate strawhull-awned
NBT	Nitroblue tetrazolium
NCBI GEO	National Center for Biotechnology Information Gene Expression Omnibus
nsSNP	Non-synonymous SNP
<i>O. rufipogon</i>	<i>Oryza rufipogon</i>
<i>O. sativa</i>	<i>Oryza sativa</i>
REC	Relative electrical conductivity
ROS	Reactive oxygen species
SH	Strawhull-awnless
sHSPs	Small heat shock proteins
SHA	Strawhull-awned
sSNPs	Synonymous SNP
VCF	Variant call file
VEP	Variant Effect Predictor

✉ Beng-Kah Song
song.beng.kah@monash.edu

- ¹ School of Science, Monash University Malaysia, Bandar Sunway, Selangor, Malaysia
- ² Department of Biological Sciences, National University of Singapore, Singapore, Republic of Singapore
- ³ General Educational Development, Daffodil International University, Daffodil Smart City, Ashulia, Dhaka, Bangladesh
- ⁴ Agricultural Biotechnology Research Center, Academia Sinica, Taipei, Taiwan
- ⁵ Academia Sinica Biotechnology Center in Southern Taiwan, Tainan 711, Taiwan

Introduction

Rice (*Oryza sativa* L.) is a crucial staple food globally, consumed by more than one-third of the world's population, meaning over 2.5 billion people rely on rice for their primary caloric intake. Furthermore, rice contributes to the nutritional needs of over half of the world's population, providing significant nutritional benefits such as carbohydrates for energy and essential minerals (Sen et al. 2020). Therefore, the role of rice in global food security and nutrition is immense and multifaceted. (Fukagawa and Ziska 2019). The Green Revolution in the 1960s saw significant progress in high-yielding rice varieties that dramatically increased world rice production. However, climate change and rising temperatures worldwide threaten rice production and grain quality (Tang et al. 2023), particularly in Asian regions where more than 90% of the global rice is planted and consumed (Hasanuzzaman et al. 2013). Screening and incorporating heat tolerance genes into elite commercial rice cultivars and breeding rice cultivars adaptable to warmer conditions are critical for maintaining global food security.

Recent studies have demonstrated that agricultural weed relatives of crop species contain genes related to stress tolerance that could be potentially valuable for improving the performance of crops (Ghanizadeh and Harrington 2021; Hopper et al. 2019). Weedy rice (*Oryza* spp.) is a close weedy relative of domesticated rice and a globally noxious weed prevalent in rice planting areas that competes with the rice crop and reduces its yield worldwide (Roma-Burgos et al. 2021). Weedy rice typically has taller stature, easy shattering grains, and strong seed dormancy, with some biotypes showing intermediate characteristics between wild rice (*Oryza rufipogon*) and cultivated rice (Sudianto et al. 2016). Studies have indicated that weedy rice from Southeast Asia and South Asia are genetically very diverse in comparison to weedy rice from other parts of the world, with frequent detection of wild genome introgression into weedy rice populations in these regions (Huang et al. 2017; Pusadee et al. 2013; Qiu et al. 2020a; Song et al. 2014; Sudo et al. 2021; Vigueira et al. 2019).

These wild characteristics and their potential for rapid evolution have proven favourable for weed adaptation in agricultural settings. Among the traits, a wild-derived shattering allele has been found in Malaysian weedy rice (Song et al. 2014), while other wild alleles controlling various traits have also been reported in the weed, such as *An-1* that controls awn development (Neik et al. 2019); *bh4* that maintains hull colour (Song et al. 2014); *Rc* that controls pericarp colour (Cui et al. 2016); *OsLGI*, *Progl*, *Phr1*, and *Waxy* that control panicle structure, plant architecture, phenol reaction, amylose content, respectively (Li

et al. 2024). Crop-to-weed evolutionary processes have also been reported in Asian regions, with tolerance to imidazolinone herbicides in Malaysian weedy rice populations found to have introgressed from the Clearfield® rice (Yean et al. 2021). Such continuous introgression from cultivars to the weedy rice was also reported in Thailand (Pusadee et al. 2013), Taiwan (Huang et al. 2021), and South China (Qiu et al. 2020b), highlighting the adaptive nature of weedy rice in the planting areas.

Although the germplasm of weedy rice may exhibit inferior phenotypic performance compared to modern rice cultivars, recent studies have shown that these weedy relatives of rice possess valuable genes conferring agronomically important traits beneficial to cultivated rice. These include cold tolerance (Borjas et al. 2016), resistance to blast disease (Liu et al. 2015), dark spot pigmentation of leaves (Bres-Patry et al. 2001), seed dormancy (Feng et al. 2016), and rapid seedling growth (Lee et al. 1999). Among these, weedy rice from China and the United States have been found to be more resistant to abiotic stresses such as cold, salt, drought, and water resiliency, than cultivated rice lines (Chen and Suh 2015). More recently, Stallworth et al. (2021) identified 3–4 leaf stage weedy rice genotypes that survived heat treatment at 38 °C for 21 days. Given that the optimum temperature of modern rice cultivar seedlings ranges between 28 °C and 30 °C, and rice seedlings are unable to survive above 35 °C (Yoshida et al. 1981), the incorporation of heat-tolerance-related genes from these weedy rice types into the breeding of elite cultivars could be of great benefit.

Many genetic studies on heat tolerance and its effect on crop yield (Janni et al. 2020) have highlighted the importance of genes controlling the biomolecules, such as heat shock transcription factors (HSFs), heat shock proteins (HSPs), molecular chaperones, metabolic catalysts, phytohormones, and other protective molecules. Plants rely heavily on a complex regulatory network to manage their responses to heat stress, and HSFs play a vital role in this network. They are involved in multiple transcriptional regulatory pathways and have significant roles in heat stress signalling and responses to other stresses (Zhang et al. 2011). The transcription of heat shock protein genes is rapidly and significantly upregulated in response to heat stress events, with HSF binding to HSP promoters increasing at a similar rate (Zhang et al. 2021).

Analysing the genes responsible for the production of these HSFs, HSPs and various biomolecules in plants would enhance the characterisation of heat tolerance traits in crop species such as rice, ultimately leading to the breeding climate-smart rice. Nevertheless, it remains unclear how the molecular mechanisms underlying heat-stress adaptation would affect weedy rice.

The optimal temperature range for rice seedling growth is typically between 25 and 28 °C. Heat stress (42–45 °C)

during the seedling stage can result in increased water loss, wilted and yellowing leaves, decreased root growth, and eventually seedling death (Xu et al. 2021). A similar heat stress test was conducted on Chinese rice lines at 45 °C for 24 h (He et al. 2023). However, neither study investigated plant response toward prolonged and continuous heat stress. In light of these, the present study subjected weedy rice accessions from Malaysia to a heat treatment of 44 °C over an extended period, with the aim of identifying heat-tolerant weedy rice and characterising the underlying molecular mechanisms of heat tolerance. Specifically, a panel of 180 accessions of previously collected weedy rice seedlings (Song et al. 2014) were subjected to heat stress screening, followed by a combinatorial examination strategy that employed relative electrical conductivity (REC), 3, 3'-diaminobenzidine (DAB), and nitro-blue tetrazolium (NBT) tests to evaluate and verify heat tolerance in weedy rice. Additionally, the study used a genome-wide single nucleotide polymorphisms (SNPs) approach to explore the genetic variation pattern of heat-tolerant resistance genes. Through this combination of analyses, we aim to answer two main research questions: (i) Do any Malaysian weedy rice types exhibit heat tolerance trait? (ii) Are there SNP variants associated with the heat tolerance response? Our findings unveil a panel of heat-tolerant weedy rice and genetic variants that can be utilized for future rice breeding purposes.

Results

Physiological performance of the weedy rice in response to high temperature stress

A total of 180 weedy rice accessions from Malaysia (Supplementary Table S1) were screened for heat tolerance. All seedling-stage rice plants were treated with a constant high temperature of 44 °C for 72 h based on previous study (Xu et al. 2021). Five accessions, namely MU244, MU235, MU260, MU249 and MU237 (survival rates of 90%, 83%, 60%, 58% and 58%, respectively), survived after 2 weeks of the recovery at ambient temperature (24 to 35 °C) (Supplementary Table S2) and were termed heat tolerant (HT) group. Five other accessions (i.e., MU005, MU100, MU114, MU251 and MU264) with 0% survival rate were named the heat-susceptible (HS) group. The REC values of the HT weedy rice ranged from 3.2 to 29.4, with an average of 11.36. In contrast, the range of REC value of the HS group was between 46.4 and 51.65 with an average value of 48.93 (Supplementary Table S3). Results of the independent *t*-test (based on survival rate and REC) showed a significant difference between the HS and the HT groups ($P < 0.05$). As expected, the average REC scores of the HT accessions were lower than those of HS, indicating greater tolerance towards

heat. Overall, HT weedy rice exhibited less ion leakage under high temperature, while heat treatment caused higher ion leakage in HS accessions (Fig. 1). The continuous high temperature at 44 °C for three days used in the present study is considered harsher than the heat treatment of 45 °C/28 °C, 6 h light/18 h dark night in a heat-tolerant rice study reported by Sukkeoa et al. (2017).

For further investigation on ROS production in weedy rice in response to temperature stress, DAB and NBT staining were used to visualize the presence of H₂O₂ and superoxide, respectively. A 24 h heat stress at 44 °C resulted in a higher pile of ROS in HS samples compared to HT samples (Fig. 1), indicating that HS samples were more vulnerable to oxidative stress under heat treatment. In contrast, HT samples showed a lower accumulation of ROS molecules, suggesting better defence against heat and oxidative stress (Gechev et al. 2006).

Statistical analysis

The Shapiro–Wilk test (Shapiro and Wilk 1965) showed that the heat screening data (survival rate) and REC data are normally distributed. Levene's test (Levene 1960) and independent sample *t*-test concluded that there was a statistically significant difference between HT and HS groups in terms of the survival rate and REC scores (Supplementary Table S4A, B).

Analysis of the variants in the heat-tolerance-related genes (HTRGs)

The genome sequences of the HT and HS weedy rice groups were individually mapped to the Nipponbare (*temperate japonica*) reference genome (MSU 6.1 release), which is known to be a heat-tolerant genotype (Woldegiorgis et al. 2022). The percentage of mapped reads was 98.41% and 98.36% in HT and HS groups, respectively. 4,450,281 and 4,841,370 SNPs were identified in the HT and HS groups, respectively (Supplementary Table S5A, B). The transition to transversion SNPs ratio was 2.4691 in HT and 2.4814 in HS with more transition SNPs (5,025,258 in HS and 4,848,540 in HT) than transversion SNPs (2,025,151 in HS and 1,963,692 in HT), which is in line with studies in elite and hybrid rice (Hu et al. 2014; Hwang et al. 2014; Subbaiyan et al. 2012). For transition mutation, C/T was more common than A/G, similar to the findings in monocots (Batley et al. 2003; Rathinasabapathi et al. 2015; Subbaiyan et al. 2012). T/A transversions were found to be more common than A/C, G/T, and C/G transversions, which is consistent with past studies on rice (Jain et al. 2014; Subbaiyan et al. 2012).

Out of the 4,825,062 SNPs identified in the HT group, 4,101,724 SNPs passed quality filtration (QUAL < 30,

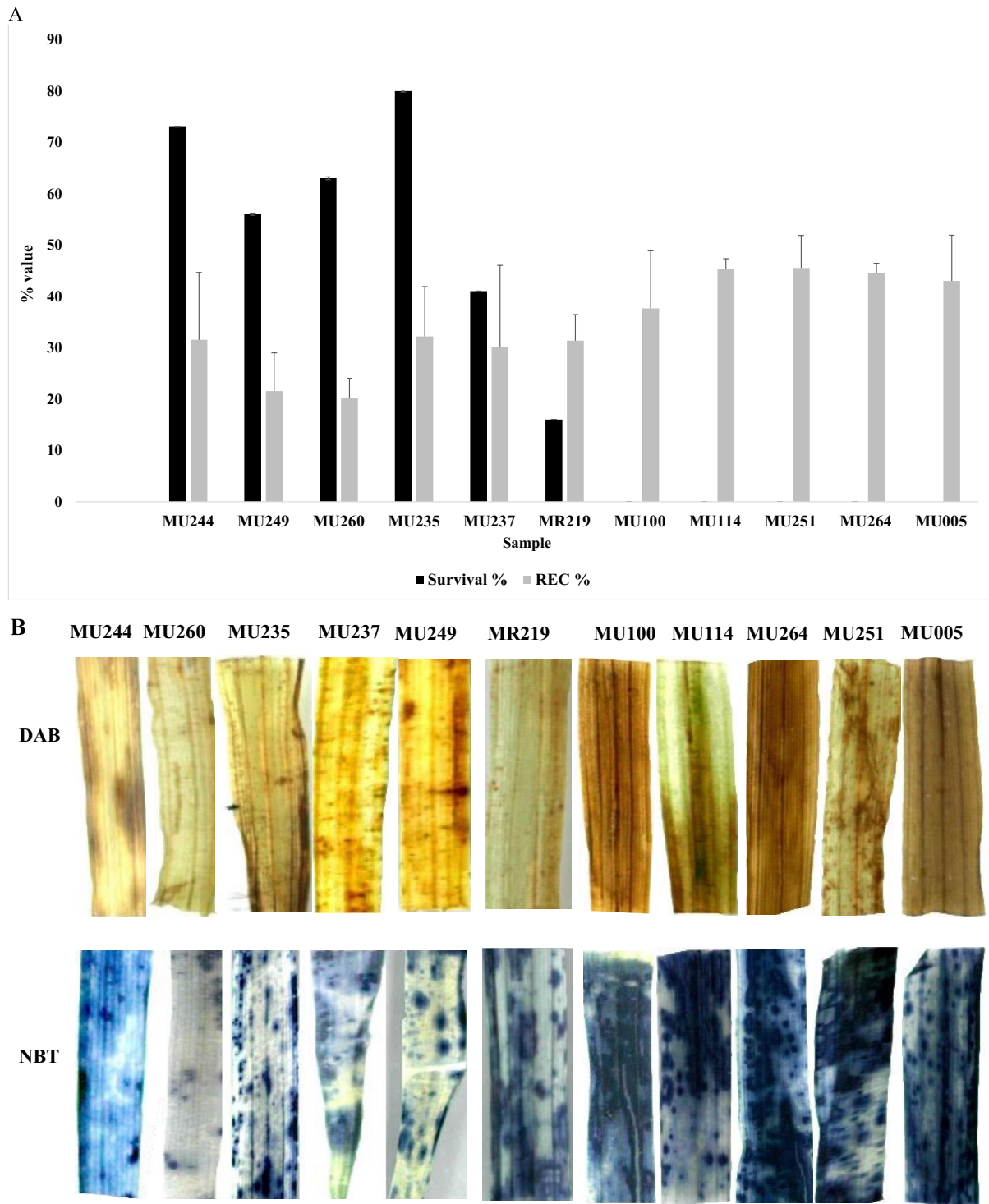


Fig. 1 Validation of heat stress tolerance trait in heat-tolerant (HT, comprising MU235, MU237, MU244, MU249 and MU260) and heat-susceptible (HS, comprising MU005, MU100, MU114, MU251 and MU264) weedy rice. **A** Heat stress screening and REC histogram of five HT and HS accessions. Student's *t*-test was used to compare

HT and HS weedy rice groups. The normality of data was assessed using an independent sample *t*-test and the Shapiro–Wilk (S–W) test. **B** Staining of ROS. The brown colour shows the presence of H_2O_2 after DAB staining. The blue colour represents the NBT staining results of the accumulation of superoxide

QD < 2, MQ < 30, MQ0/DP > 0.1). Of these, 746,992 (18.21%) and 3,354,732 (81.79%) were located in genic and non-genic regions, respectively. Further investigation was conducted on synonymous SNP (sSNP) and non-synonymous SNP (nsSNP) detected in the genic areas. Within the genic regions, a total of 315,907 sSNPs and 405,573 nsSNPs were identified, representing 42.29% and 54.31% of the genic SNPs, respectively. In the HS rice group, 4,137,909 out of a total of 5,011,042 SNPs met the quality filtration requirement. Of these, 758,741 (18.34%) and 3,379,168 (82.29%) were respectively found in genic and non-genic regions. Among the genic areas, 412,055 (54.29%) and

320,675 (42.26%) were identified as nsSNPs and sSNPs, respectively (Table 1).

A total of 53,506 genes were shared among the HT and HS accessions, with only 443 (0.8%) genes being unique to HT and 519 (1%) unique to HS (Fig. 2). Following filtration to select for common heat stress response pathways, five HSP families were detected: HSP101, HSP90, HSP70, HSP20 and HSP32. A comparison of the five HSPs families in the HT group with the Nipponbare reference genome resulted in 657 SNPs, while in the HS group, 591 SNPs were detected. Of these, 610 and 582 variants were identified respectively in the non-coding areas of both the HT and HS groups. The remaining SNPs (47 and nine, respectively in HT and HS) were found in the coding regions. Within the 47 SNPs in the HT group, five nsSNPs were identified in the coding regions of stress tolerance genes (*OsHSP16.9B*, *OsHSP16.9C*, *OsHSP17.9*, *OsHSP71*, *OsHSP70*, *ClpB-c*). In the HS group, of the nine SNPs observed, two nsSNPs were detected in two stress tolerance genes (*OsHSP17.1* and *OsHSP70*) (Supplementary Table S7A, B).

Using gene expression profiling data obtained from the Rice Stress-Resistant SNP Database (Tareke Woldegiorgis et al. 2019), a total of 57 heat tolerance genes were identified and analysed in this study, (Supplementary Table S8), resulting in 3985 and 3725 SNPs in HT and HS rice, respectively. Most of the SNPs reported were silent effect, but a limited number were found to perform function. Notably, we found 334 non-synonymous and 227 synonymous SNPs in the HT

Table 1 SNPs substitution information of heat-tolerant (HT) and heat-susceptible (HS) weedy rice groups

Description	HT	HS	HT%	HS%
Total SNPs	4,101,724	4,137,909		
Silent SNPs	3,354,732	3,379,168	81.79	81.66
Genic SNPs	746,992	758,741	18.21	18.34
Non-synonymous	405,573	412,055	42.29	42.26
Synonymous	315,907	320,675	54.29	54.31
Stop gain, stop loss & start*	25,512	26,011	3.42	3.43

*Stop gain: Mutations that introduce a premature stop codon in the coding sequence, often resulting in a truncated and typically non-functional protein. *Stop loss: Rare mutations that affect the final amino acid of a protein

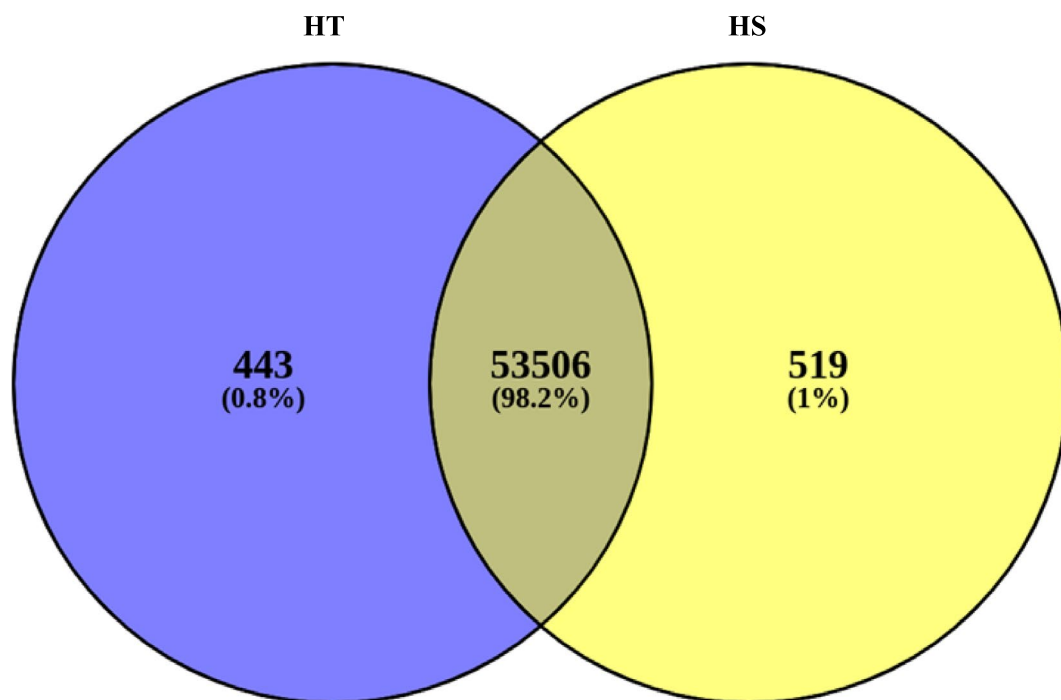


Fig. 2 Stress responsive genes in heat-tolerant (HT) and heat-susceptible (HS) weedy rice groups. Blue colour represents HT, and yellow colour represents HS

sample, and 312 non-synonymous and 227 synonymous SNPs in the HS group (Supplementary Table S9). We found a higher number of nsSNPs between HSP families found in both samples. nsSNPs ranged from 1 to a maximum of 28 per gene in HS and 1 to a maximum of 36 per gene in the HT sample (Supplementary Table S9).

The HT group showed higher diversity in the HSPs and HSFs, while a limited number of HSFs were found in the HS sample. HSFs are crucial in regulating HSPs and switching different stress mechanisms (Kawakami et al. 2020; Meng et al. 2022). Our analysis also revealed several noteworthy patterns in the nsSNPs.

Characterising the distribution of the impact of non-synonymous variants

We further assessed the impact of non-synonymous SNPs (nsSNPs) using PolyPhen-2 (Hicks et al. 2011) and SIFT (Ng and Henikoff 2003) on five selected potential heat-tolerant genes (HSP70, HSP20, *HSFA2C*, *HSFB1* and *HSFB2C*). Higher scores in PolyPhen-2 or lower scores in SIFT indicate that the variants are more likely to be deleterious. From there, we found a total of 55 nsSNPs within the HT group and 37 in the HS counterpart. From the PolyPhen-2 results, we identified 33 nsSNPs as “probably damaging”, 5 nsSNPs as “possibly damaging”, and 13 as “benign”. Meanwhile, using SIFT, 23 were identified as “tolerable”. 32 nsSNPs were predicted by both to be probably damaging or alter the amino acid function (Supplementary Tables S10, S11, S12). Multiple sequence alignment was performed on the coding and amino acid sequences of HSP70, HSP20, *HSFA2C*, *HSFB1*, and *HSFB2C* from the HT and HS groups, and they were compared with the Nipponbare reference genome. The results revealed substantial amino acid variation in the five HTRGs between the HT and HS groups, when compared to the reference (Fig. 3). These SNPs may contribute to heat tolerance in rice and are crucial for understanding how plants adapt to high-temperature stress. The presence of unique amino acid substitutions could indicate the involvement of these genes in the heat stress response. The three-dimensional structures of HSP70, HSP20, *HSFA2C*, *HSFB1*, and *HSFB2C* proteins were modelled using PyMOL (The PyMOL Molecular Graphics System, Version 2.5 Schrödinger, LLC.) (Fig. 4). The modelled HSP70 protein featured two primary domains: the HSP binding domain and the ATP binding domain. Structural analysis revealed substantial changes in residues concentrated within the core region of the HSP binding domain, with the remaining residue changes occurring within the ATP binding domain. The projected three-dimensional structure of HSP20 showcased both the HSP binding domain and the transmembrane (TM) helix. Conversely, the predicted structures of *HSFB1*,

HSFB2C, and *HSFA2C* encompassed a single HSP binding domain.

Discussion

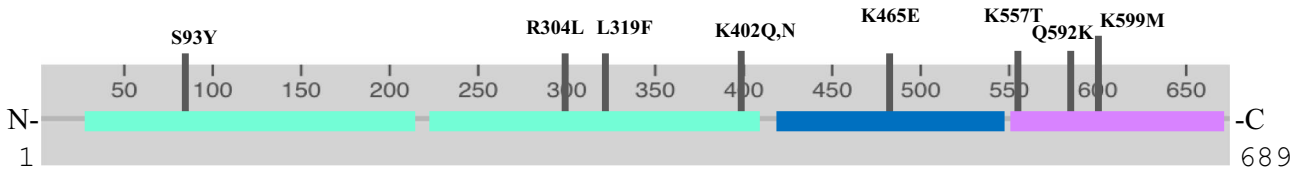
The present study employed a comprehensive heat screening test and a genome-wide SNP analysis to identify heat tolerant Malaysian weedy rice accessions, with further evaluation of the synonymous and non-synonymous substitutions, as well as structural prediction of heat-tolerant related proteins in HT and HS weedy rice groups. Of the 180 weedy rice accessions subjected to heat stress at 44 °C for three consecutive days, only nine accessions were able to survive the rigorous treatment. Among them, four were strawhull-awned, two were blackhull-awned, and three were brown-stripe-hulled weedy rice. Considering no prior studies have reported weedy rice that could withstand such high heat treatment, findings from this study demonstrate that Malaysian weedy rice presents a promising and novel source of heat-resistant alleles, enabling us to breed superior elite rice cultivars that are better adapted to elevated temperatures.

Stallworth and colleagues (2021) discovered that the strawhull trait is linked to heat-stress tolerance in weedy rice varieties in the United States. In the current study, blackhulled accession MU235 and the morphologically intermediate weedy rice MU244, with seed traits between strawhull trait and brown-striped hull, exhibited the highest survival rate of 77.3% and 78.3%, respectively (Fig. 1A). Given that blackhulled and brown-stripe-hulled weedy rice accessions are genetically more similar to *O. rufipogon* (Song et al. 2014), it is suggestive that introgression of wild-type genes plays a significant role in conferring heat stress tolerance in these weedy accessions. Nonetheless, this is the first study reporting heat stress traits in brown-stripe-hulled weed accessions. Indeed, the abundance of brown-stripe-hulled weedy rice and the rich genetic and morphological diversity found in the Malaysian weedy rice population provide a good source of germplasm for breeding elite cultivars with improved heat tolerance.

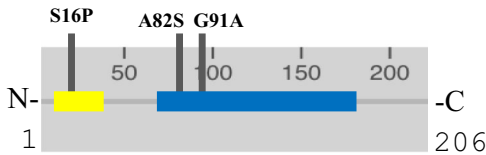
The measurement of REC is a widely accepted indicator for stresses tolerance in plants (Li et al. 2020; Su et al. 2017; Li et al. 2008). The significant reduction in REC values observed in the HT accessions compared to the HS groups of weedy rice accessions (Fig. 1B) not only serve as a strong indicator for heat tolerance characteristics, but may also implicate other abiotic stress tolerance traits, particularly drought tolerance. This is because the responses to drought and heat stress in plants often overlap and interact, especially under the effects of drier and hotter climate change conditions (Priya et al. 2019).

In this study, we compared the genomes of seven previously published Malaysian weedy rice accessions (Qiu et al.

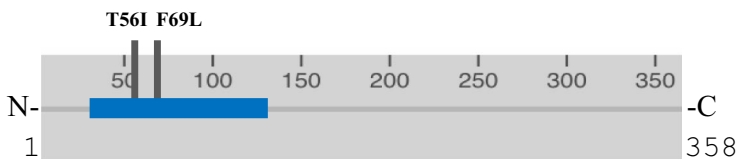
LOC_Os05g23740.1 (HSP70)



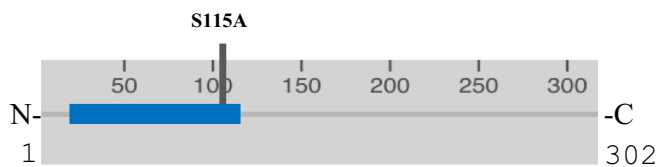
LOC_Os11g13980.1 (HSP20.2)



LOC_Os10g28340.1 (HSFA2C)



LOC_Os09g28354.1 (HSFB1)



LOC_Os09g35790.1 (HSFB2C)

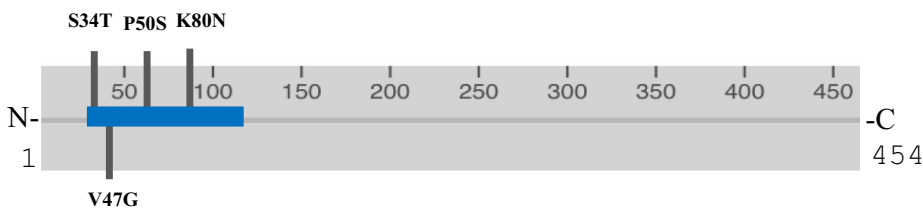


Fig. 3 Structure of five representative HTRGs and unique amino acid changes (from Nipponbare reference to HT rice type). Positions of the substituted amino acids within the domains are indicated by black vertical lines. Each horizontal bar represents domains of the proteins.

The blue colour represent HSP and HSF domain, cyan colour represent ATPase domain, purple colour represent C-terminal domain and the yellow colour represent TM helix

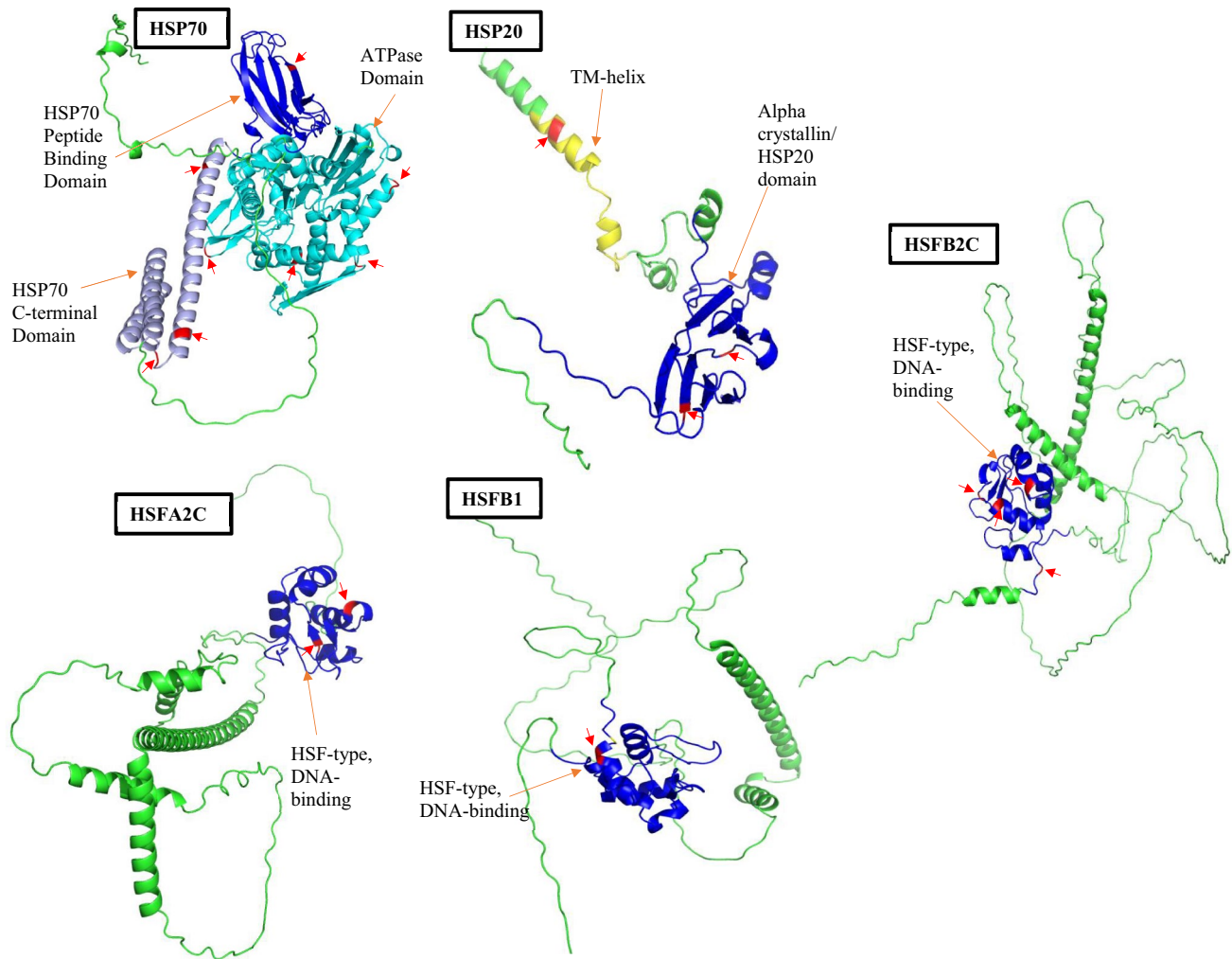


Fig. 4 Predicted three-dimensional structures of HSP70, HSF1, HSFA2C, HSP20 and HSFB2C. The domains are denoted by colour: blue represents HSP and HSF domains; purple represents the C-terminal domain; cyan represents the ATPase domain; and yellow represents the TM-helix. Mutations within a domain are represented with small red arrows

mineral domain; cyan represents the ATPase domain; and yellow represents the TM-helix. Mutations within a domain are represented with small red arrows

2020a, b) to identify differences in SNPs between HT (consisting of MU235, MU237, and MU066 accessions) and HS (consisting of MU022, MU100, MU144, and MU005 accessions) rice groups. The average error rate for seven genome sequences mapped to the Nipponbare reference genome was approximately 8%, consistent with previous research (Jain et al. 2014; Rathinasabapathi et al. 2015). The genomes of weedy rice contained various SNPs deserts, which may reflect the lack of SNPs in certain locations (Hu et al. 2014; Rathinasabapathi et al. 2015; Wang et al. 2017).

The present study also indicated that the HS group is unlikely to have a lower number of mutations on the HSF genes. HSPs and HSFs are crucial for plant heat response and they exhibit significant interconnectivity, highlighting the critical role of the HSF-HSP network in conferring heat tolerance (Sarkar et al. 2014). In a long-term heat treatment of hybrid rice II YOU838, HSF1 and HSFB2C

were evidently induced (Wang et al. 2020). The *OsHsfA* and *OsHsfB* subfamilies of genes are known to play crucial role in cellular defence against abiotic stress and heat stress in plants (Zhang et al. 2012). Overexpression of *HsfA2d* has been shown to enhance tolerance to salt and drought stress in the presence of ongoing heat stress, making it a potential candidate for improving wheat yield under abiotic stress (Chauhan et al. 2013). In plants, ANN2 is recognized to contribute to cellular defence against heat stress (Qiao et al. 2015). In contrast, potentially beneficial mutations were observed in the HT sample groups, possibly enabling them to endure heat stress. The HS group exhibit only the most commonly occurring HSP mutations, which have been shown to be beneficial in all groups when exposed to heat stress. While the HT group showed some stress tolerance abilities, as evidenced by the presence of certain HSP families, these were insufficient to manage chronic stress.

The Variant Effect Predictor (VEP) analysis revealed a relatively higher number of SNPs (47 in HT) in the five HSP families—HSP20, HSP32, HSP70, HSP90, and HSP101, compared to the lower number of SNPs (nine in HS) observed (Supplementary Table S7A, B). The HS group reported only two sHSPs as non-synonymous, such as the HSPs from the class I small heat shock proteins (e.g. OsHSP16.9B, OsHSP17.9), which are members of one of the most abundant HSP families found in various cell organelles and are highly synthesised under high temperatures. In particular, OsHSP17.9 functions as a chaperone to prevent the heat-susceptible citrate synthase in rice seedlings from aggregating at high temperatures (45–50 °C) (Sarkar et al. 2020). Other candidate genes, including *OsHSP71* and *OsHSP70*, were not only involved in heat stress but also in other abiotic stress tolerances, including high salt stress response (Wang et al. 2020; Zou et al. 2009). HSP100 was absent in the HS group but present in HT in our study samples. The 5'-UTR region of *ClpB-c* (HSP100) is required for enhancing protein translation under stressful circumstances (Mishra et al. 2016; Singh et al. 2010). Indeed, HSP100 is an essential component of plant growth and development, aiding in the ability of plants to withstand high temperature stress.

Additionally, this study characterised the putative functional features of nsSNPs in the coding regions of five HTRG representatives. Based on the three dimensional structure of the five selected HTRGs, it was observed that the identified nsSNPs corresponded to significant amino acid changes exist in the predicted domain structures within the HT weedy group (Figs. 3 and 4). Most of the nsSNPs showed higher tolerance or altered protein function. A higher number of nsSNPs were reported in the HT group, with 55 nsSNPs in HT compared to 37 in HS rice group, indicating changes in functional effects that may enhance the stress survival adaptability of the HT rice group (Supplementary Tables S10, S11). Interestingly, when comparing the distribution of variant effects across different HSPs and HSFs, it was observed that exonic regions were more likely to contain SNPs with large effect variants, while regulatory regions tended to be enriched in HSF sequences. In a prior investigation, it was proposed that non-canonical heat shock elements (HSEs), marked by nsSNPs at specific positions, play a crucial role in the heat response and adaptation process by interacting with HSPs (Zhao et al. 2020). The regulatory variants, specifically HSFs and HSPs, have a notable and vital function when influenced by positive mutations. These regulatory variants contribute significantly to strengthening a plant's ability to withstand and endure challenging conditions, thereby enhancing its capacity for resilience and survival during environmental stresses. (Liu et al. 2020). This suggest that regulatory variants are more flexible in enhancing plant environmental adaptability.

The expression of HSP genes is predominantly governed by HSFs, constituting an evolutionarily conserved mechanism for mitigating heat stress (Tian et al. 2021). Within the five HTRGs studied, a substantial proportion (80%) of variation was discerned within the HT subgroup compared to the HS group. This discrepancy could potentially be attributed to a more severe bottleneck in the HS subpopulation, leading to the fixation of numerous deleterious and neutral variants. Remarkably, research consistently underscores the impact of higher genomic heterozygosity for deleterious mutations on crop fitness, thereby limiting plant breeding efficacy (Zhu et al. 2022).

Of these, this study highlights the role of mutations in enhancing HSP20 functionality, particularly within the HSP binding domain at positions Ala82Ser and Gly91Ala, thereby fortifying the chaperone domain. Additionally, the TM helix at position Ser16Pro might facilitate expanded interactions and potential novel functions. These mutations confer increased stress response capabilities to HSP20, which serves as a defence mechanism against heat-induced protein damage. In the context of HSP70, three significant functional domains are present: the HSP70 peptide binding domain, the N-terminal ATPase domain, and the HSP70 C-terminal domain (Alderson et al. 2016). Noteworthy amino acid alterations occurred within the ATPase domain, affecting Ser93Tyr, Arg304Leu, Leu319Phe and Lys402. Similarly, the HSP70 C-terminal domain exhibited changes at positions Lys557Thr, Gln592Lys and Lys599Met. These modifications within the ATPase and HSP70 C-terminal domains potentially influenced the role of HSP70 in thermotolerance mechanisms in our study.

Positive allelic changes within HSP70 likely optimize interactions, thereby enhancing chaperone activity, protein folding and the prevention of aggregation under heat stress. The presence of distinct biological processes within genes of various HSP types underscores the pivotal role of sequence variations within HSPs and HSEs in driving the evolution of heat responses and adaptation (Zhao et al. 2020). Notably, HSF1 displayed changes at positions Ser115Ala and Ile117Asn, HSF2C exhibited alterations at positions Ser34Thr, Val47Gly, Pro50Ser, and Lys80Asn, while HSF2A featured mutations at Thr56Ile and Phe69Leu. These beneficial alleles potentially enhanced the DNA-binding affinity of HSFs, including HSF1, HSF2C, and HSF2A, thereby amplifying heat shock gene activation and cellular responses to elevated temperatures (Wang et al. 2018). Collectively, these mutations could fortify the defence against heat-induced cellular damage, boosting heat tolerance ability. While the limited sample size may introduce a potential constraint to the overall significance of the findings in the present study, the results nevertheless offer valuable foundational insights, particularly molecular basis of heat tolerance mechanisms, to the rice community. Future

research should focus on identifying gene expression patterns in response to temperature stress to gain a better understanding of the underlying mechanisms of heat tolerance.

Conclusion

This study highlights the potential of heat-tolerant weedy rice accessions, including MU244, MU235, MU249, MU260 and MU237, to enhance the heat tolerance of cultivated rice. These accessions consistently outperformed heat-susceptible counterparts in heat stress tests. Genome-wide SNP analysis revealed genetic variation patterns in heat tolerance genes and HSP loci in these HT and HS accessions of weedy rice, shedding light on the rice molecular basis of heat stress response. This research provides valuable insights for developing resilient rice varieties, offering a promising strategy to tackle climate-induced challenges and improve global food security.

Materials and methods

Plant materials

A diverse collection of 180 weedy rice (Supplementary Table 1) accessions were employed, which are a subset of the samples investigated by Song et al. (2014). To cover as much genetic variation of representatives of all weedy rice populations in Malaysia, the sample collection comprises eight morphological groups (strawhull-awnless, SH; strawhull-awned, SHA; intermediate strawhull-awned, mSHA; browhull-awnless, BR; browhull-awned, BRA; blackhull-awnless, BH; and blackhull-awned, BHA). This study used a commonly planted local elite cultivar, MR219, as a control. The inclusion of *Oryza sativa* MR219 in this study is significant due to its prominence in Malaysia's agricultural landscape, where it is grown in 53% of the country's major granary areas (Ahmad et al. 2023). However, its susceptibility to abiotic stress during the reproductive stage presents a challenge that could impact yield and, consequently, food security. Rice seeds were soaked in water and incubated at 32 °C for 3 to 5 days to promote germination (Patil and Khan 2011). Germinated seedlings were transferred to pots and grown in the open environment of a plant house located on the rooftop of Monash University Malaysia. The farmland soil was used to maintain proper natural conditions.

Screening for heat-tolerant rice

Preliminary heat stress tests with varying temperatures (33 °C, 13 h light/11 h dark; 26 °C, 13 h light; 45 °C, 11 h dark; and 44 °C, 13 h light/11 h dark) were performed on

a group of five randomly selected weedy rice accessions. Based on these preliminary tests, we determined that 44 °C was the most suitable temperature for our experiment. At three-leaf stage, 20 healthy seedlings per accession were transferred to the growth chamber and treated with constant high temperature at 44 °C, 13 h light/11 h dark with light intensity of 300 lx and 80% relative humidity conditions for three days (Xu et al. 2021). The same number of seedlings were grown under plant house conditions as a control. The plant house conditions have average temperature ranged from 25 °C to 39 °C, 13 h light/11 h dark with natural sunlight (300 lx measured at noon), and $84 \pm x\%$ relative humidity (Stuerz and Asch 2019).

Heat tolerance index (HTI)

Heat-treated rice plants were relocated after 3 days of heat treatment, placed under optimal plant house conditions, and placed together with control plants. This study aimed to test the survival and recovery of plants under extreme heat a three-day heat treatment at 44 °C. After 2 weeks, the survival and recovery of the heat-treated plants were compared to those grown under normal conditions. The findings revealed that the plants that could tolerate the heat showed significant recovery and healing, which was confirmed using statistical tests (Supplementary Table 2). Using measures of seedling survival rate after three days at normal temperature (control) and 44 °C (treated; after two weeks of recovery phase) demonstrating strong correlations and relative heat tolerance, a heat-screening scale based on the stress response was devised. Parameters were tested under control to allow calculation of heat stress and stress index. The heat tolerance indices (HTIs) were calculated by dividing the sum of individual scores for each parameter by the total score and multiplying by 100% (Sarsu 2018) (Supplementary Tables 1, 2).

$$\text{HTI} = (\text{Value under stress} / \text{Value at the control}) \times 100$$

*value = full recovery score.

The sample that showed an HTI value $\geq 50\%$ was defined as heat tolerant, while the sample below 50% is heat susceptible based on the three biological replicates with continuous heat exposure treatment at 44 °C.

Relative electrical conductivity (REC) measurement

To measure the cell membrane stability of the leaves under heat stress, the relative electrical conductivity (REC) of the leaves was measured as described previously (Ilík et al. 2018). Fresh leaves were harvested and washed with deionized water. From each leaf sample, three rectangular segments of 2 cm \times 2 cm were cut out. The plant materials in the treatment group were subjected to a temperature of 44 °C for 24 h, followed by harvesting and placement into a 20 ml tube containing deionized water. The sample was preheated at

50 °C for one hour and subsequently incubated at room temperature for 10 min. The initial conductivity was measured from the leaf tissue (R1) and water (R2), and the maximal conductivity of both was determined by heating the leaf segment and water at 98 °C for 30 min. The final conductivity of the leaf tissue (R3) and water (R4) was then measured at room temperature. The REC was calculated based on the equation: $REC (\%) = [(R1 - R2)/(R3 - R4)] \times 100$.

Reactive oxygen species (ROS) assays

Nitroblue tetrazolium (NBT) and 3, 3'-diaminobenzidine (DAB) staining assays were used to detect superoxide radical anions (O_2^-) and hydrogen peroxide (H_2O_2) content, respectively as previously described (Gechev et al. 2006). Briefly, 3 cm long segments of weedy rice leaves from both heat-treated (44 °C) and non-treated samples were rinsed with deionized water and immersed in NBT and DAB staining solutions overnight for O_2^- and H_2O_2 detection, respectively. Following staining, the samples were rinsed with absolute ethanol and boiling water for 10 min, then transferred to paper towels saturated with 60% glycerol, and imaged using stereomicroscope (Jambunathan 2010).

Sequence analysis and variants identification

The paired-end Illumina genome sequencing raw read used in this study were obtained from Qiu et al. (2020a, b), comprising three genomes (MU066, MU237, and MU235) with heat tolerance and four genomes (MU022, MU100, MU114, and MU005) susceptible to heat. These seven genotypes were selected based on the availability of their whole genome sequences in Qiu et al. (2020a, b) among the list of accessions studied in this research. Pre-processing of the raw sequencing data was carried out using FastQC (Andrews 2010) and Trimmomatic tools (Bolger et al. 2014). Clean reads were then aligned to a reference rice genome (*O. sativa* ssp. *japonica* cv. Nipponbare; MSUv6.0) (Woldegiorgis et al. 2022) using BWA assembler (Li and Durbin 2010). Nipponbare was selected as the reference genome due to its extensive study, high sequencing accuracy, comprehensive annotation, and widespread use as a standard in rice research, facilitating comparisons across various studies. SNP variants were called using Samtools (Li et al. 2009), and variant detection was performed using Picard v2.2.7., GATK v4.1.9 and bcftools v0.1.13 (McKenna et al., 2010). To minimize false discovery of variants, the variant call file (VCF) was filtered using the following parameters: mapping quality ≥ 55 , base quality ≥ 30 , and variant quality ≥ 90 . The resulting SNP and InDel variants were annotated using the SnpEff database for *Oryza sativa* and SnpEff V3.6 tools (Cingolani et al. 2012), and validated using the Variant Effect Predictor (VEP) (McLaren et al. 2016).

Point mutation detection and structural prediction of proteins

Point mutation analysis was carried out with the existing SNPs database. The dataset of stress-resistant SNPs in rice was obtained from the Rice stress-resistance SNPs database (Tareke Woldegiorgis et al. 2019). The filtered SNPs from weedy rice genomes were further compared with 57 known heat stress-resistant genes in the SNPs database. The ns-SNPs and the amino acids were extracted from the annotation results of point mutations described in the Sect. 5.7. Five potential candidate genes, specifically HSP70, HSP20, HSFA2C, HSB1, and HSFB2C, were chosen based on their evolutionary mutation ratio (≥ 1), indicating potential benefits or neutral effects, aiming to elucidate their impact on HT group's stress resilience, along with their highly conserved and crucial role in the cellular mechanism to combat heat stress, where heat shock factors (HSFs) regulate the expression of heat shock proteins (HSPs) in response to elevated temperatures. For PolyPhen-2 prediction, we first extracted five amino acid sequences harbouring non-synonymous variations and identified homologous proteins in the UniRef100 database (Suzek et al. 2015). We then extracted their amino acid sequences and generated multiple sequence alignments to predict the impacts of the variations. The precomputed homologous sequences in SIFT 4G (<http://sift-dna.org/sift4g>) were used for SIFT prediction (Vaser et al. 2016). We noted that a few of the mutations were annotated by only one of the software tools. An absolute value of PolyPhen-2 score of less than 1.5 is considered benign, 1.5–2 is considered possibly damaging, and a score greater than 2 is considered probably damaging. A SIFT score between 0 and 0.05 is considered deleterious, while a score between 0.05 and 1 is considered tolerable. The selected HTRGs was inspected using AlphaFoldDB structures database model predictor (Varadi et al. 2022). The predicted structures were visualized in PyMOL (The PyMOL Molecular Graphics System, Version 2.5 Schrödinger, LLC.).

Statistical analysis

The normality of the data was assessed using the Shapiro–Wilk (S–W) test. An independent sample *t*-test was utilized to test whether there is any statistically significant differences between the HT and HS groups in terms of REC and HTI scores. Furthermore, Levene's test for equality of variance was performed to determine whether the *t*-test should assume equal or unequal variances. Thus, two independent sample *t*-tests were conducted. Statistical significance was defined as a *p*-value of less than 0.05. IBM SPSS Statistics (version 28.0) was used to conduct the statistical analysis.

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1007/s11816-024-00920-6>.

Acknowledgements We thank Mr. Dilipkumar Masilamany from Malaysian Agricultural Research and Development Institute (MARDI) for providing valuable insights into the study. The authors acknowledge funding support from the Malaysian Ministry of Education (FRGS/1/2020/STG01/MUSM/02/4) and Monash University Malaysia (MUM)-ASEAN Sustainable Development Research Grant (ASEAN-2019-01-SCI).

Author contributions M.H.S. designed and carried out the experiments and bioinformatic data analysis, drafted and reviewed the manuscript; M.H.H. and T.X.N. reviewed and edited the manuscript; W.W. Y., H.S.T., and S.S.K. reviewed and edited the manuscript and supported in project design; M.Z.H. helped in statistical analyses and reviewed the manuscript; B.K.S. conceived and coordinated the study, designed, and supervised the project, contributed to the interpretation of the data, reviewed, edited and written the manuscript. All authors have read and agreed to the published version of the manuscript.

Funding Open Access funding enabled and organized by CAUL and its Member Institutions. The authors acknowledge funding support from the Malaysian Ministry of Education (FRGS/1/2020/STG01/MUSM/02/4) and Monash University Malaysia (MUM)-ASEAN Sustainable Development Research Grant (ASEAN-2019-01-SCI).

Data availability Data supporting the findings of this study that are not available in the supplementary materials will be available from the corresponding author upon request.

Declarations

Conflict of interest The authors declare that they have no competing interests.

Ethics approval Not applicable.

Consent to participate Not applicable.

Consent for publication Not applicable.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

- Ahmad F, Hasan N, Hisham SN, Yusof SN, Hassan AA, Hussein S, Harun AR, Shamsudin NAA (2023) Data on genome resequencing of the rice cultivar MR219 and its abiotic stress tolerant mutant lines, NMR152 and NMR151. Data Brief. <https://doi.org/10.1016/j.dib.2022.108815>
- Alderson TR, Kim JH, Markley JL (2016) Dynamical structures of Hsp70 and Hsp70-Hsp40 complexes. Structure 24:1014–1030. <https://doi.org/10.1016/j.str.2016.05.011>
- Andrews S (2010) FastQC: a quality control tool for high throughput sequence data. Available online at: <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Batley J, Barker G, O'Sullivan H, Edwards KJ, Edwards D (2003) Mining for single nucleotide polymorphisms and insertions/deletions in maize expressed sequence tag data. Plant Physiol 132:84–91
- Bolger AM, Lohse M, Usadel B (2014) Trimmomatic: a flexible trimmer for Illumina sequence data. Bioinformatics 30(15):2114–2120
- Borjas AH, De Leon TB, Subudhi PK (2016) Genetic analysis of germinating ability and seedling vigor under cold stress in US weedy rice. Euphytica 208:251–264
- Bres-Patry C, Lorieux M, Clement G, Bangratz M, Ghesquière A (2001) Heredity and genetic mapping of domestication-related traits in a *temperate japonica* weedy rice. Theor Appl Genet 102:118–126
- Chauhan H, Khurana N, Agarwal P, Khurana JP, Khurana P (2013) A Seed preferential heat shock transcription factor from wheat provides abiotic stress tolerance and yield enhancement in transgenic Arabidopsis under heat stress environment. PLoS ONE. <https://doi.org/10.1371/journal.pone.0079577>
- Chen L, Suh H (2015) Weedy rice—origin and dissemination. Yunnan Publishing Group Corporation, Yunnan Science and Technology Press, China. pp 234
- Cingolani P, Platts A, Wang LL, Coon M, Nguyen T, Wang L, Land SJ, Lu X, Ruden DM (2012) A program for annotating and predicting the effects of single nucleotide polymorphisms, SnpEff: SNPs in the genome of drosophila melanogaster strain w 1118; iso-2; iso-3. Fly (austin) 6:80–92. <https://doi.org/10.4161/fly.19695>
- Cui Y, Song BK, Li L-F, Li Y-L, Huang Z, Caicedo AL, Jia Y, Olsen KM (2016) Little white lies: pericarp color provides insights into the origins and evolution of southeast Asian weedy rice. G3 Genes Genomes Genet. <https://doi.org/10.1534/g3.116.035881>
- Feng J, Ye H, Gu XY (2016) Developmental mechanisms of seed dormancy revealed by map-based cloning of a major quantitative trait locus from rice (*Oryza sativa*). Plant & Animal Genome Conference XXIV. Jan. 8–13, San Diego (Poster P0742)
- Fukagawa NK, Ziska LH (2019) Rice: Importance for global nutrition. J Nutr Sci Vitaminol (tokyo) 65:S2–S3. <https://doi.org/10.3177/jnsv.65.S2>
- Gechev TS, Van Breusegem F, Stone JM, Denev I, Laloi C (2006) Reactive oxygen species as signals that modulate plant stress responses and programmed cell death. BioEssays 28:1091–1101. <https://doi.org/10.1002/bies.20493>
- Ghanizadeh H, Harrington KC (2021) Herbicide resistant weeds in New Zealand: state of knowledge. N Z J Agric Res 64:471–482. <https://doi.org/10.1080/00288233.2019.1705863>
- Hasanuzzaman M, Nahar K, Alam M, Roychowdhury R, Fujita M (2013) Physiological, biochemical, and molecular mechanisms of heat stress tolerance in plants. Int J Mol Sci 14:9643–9684
- He Y, Guan H, Li B, Zhang S, Xu Y, Yao Y, Yang X, Zha Z, Guo Y, Jiao C, Cai H (2023) Transcriptome analysis reveals the dynamic and rapid transcriptional reprogramming involved in heat stress and identification of heat response genes in rice. Int J Mol Sci 24:14802. <https://doi.org/10.3390/ijms241914802>
- Hicks S, Wheeler DA, Plon SE, Kimmel M (2011) Prediction of missense mutation functionality depends on both the algorithm and sequence alignment employed. Hum Mutat 32:661–668. <https://doi.org/10.1002/humu.21490>
- Hopper JV, McCue KF, Pratt PD, Duchesne P, Grosholz ED, Hufbauer RA (2019) Into the weeds: matching importation history to genetic consequences and pathways in two widely used biological

- control agents. *Evol Appl* 12:773–790. <https://doi.org/10.1111/eva.12755>
- Hu Y, Mao B, Peng Y, Sun Y, Pan Y, Xia Y, Sheng X, Li Y, Tang L, Yuan L (2014) Deep re-sequencing of a widely used maintainer line of hybrid rice for discovery of DNA polymorphisms and evaluation of genetic diversity. *Mol Genet Genomics* 289:303–315
- Huang Z, Young ND, Reagon M, Hyma KE, Olsen KM, Jia Y, Caicedo AL (2017) All roads lead to weediness: patterns of genomic divergence reveal extensive recurrent weedy rice origins from South Asian *Oryza*. *Mol Ecol* 26:3151–3167. <https://doi.org/10.1111/mec.14120>
- Huang Y-F, Wu D-H, Wang C-L, Du P-R, Cheng C-Y, Cheng C-C (2021) Survey of rice production practices and perception of weedy red rice (*Oryza sativa* f. *spontanea*) in Taiwan. *Weed Sci* 69:526–535. <https://doi.org/10.1017/wsc.2020.73>
- Hwang S-G, Hwang JG, Kim DS, Jang CS (2014) Genome-wide DNA polymorphism and transcriptome analysis of an early-maturing rice mutant. *Genetica* 142:73–85
- Ilík P, Špundová M, Šicner M, Melkovičová H, Kučerová Z, Krchňák P et al (2018) Estimating heat tolerance of plants by ion leakage: a new method based on gradual heating. *New Phytol* 218(3):1278–1287
- Jain M, Moharana KC, Shankar R, Kumari R, Garg R (2014) Genomewide discovery of DNA polymorphisms in rice cultivars with contrasting drought and salinity stress response and their functional relevance. *Plant Biotechnol J* 12:253–264
- Jambunathan N (2010) Determination and detection of reactive oxygen species (ROS), lipid peroxidation, and electrolyte leakage in plants. In: Sunkar R (ed) *Plant stress tolerance, methods in molecular biology*. Humana Press, Totowa, pp 291–297
- Janni M, Gulli M, Maestri E, Marmioli M, Valliyodan B, Nguyen HT, Marmioli N (2020) Molecular and genetic bases of heat stress responses in crop plants and breeding for increased resilience and productivity. *J Exp Bot* 71:3780–3802. <https://doi.org/10.1093/jxb/eraa034>
- Kawakami T, Goto H, Abe Y, Chuba M, Watanabe M, Hoshino T (2020) High frequency of transversion mutations in the rice (*Oryza sativa* L.) mutant population produced by diepoxybutane mutagenesis. *Genet Resour Crop Evol* 67:1355–1365. <https://doi.org/10.1007/s10722-020-00950-3>
- Lee S-S, Kim J-H, Hong S-B (1999) Effects of priming and growth regulator treatment of seed on emergence and seedling growth of rice. *Korean J Crop Sci* 44:134–137
- Levene H (1960) Contributions to probability and statistics. *Essays Honor Harold Hotell* 278:292
- Li H, Durbin R (2010) Fast and accurate long-read alignment with Burrows–Wheeler transform. *Bioinformatics* 26:589–595
- Li G-W, Zhang M-H, Cai W-M, Sun W-N, Su W-A (2008) Characterization of OsPIP2;7, a water channel protein in rice. *Plant Cell Physiol* 49:1851–1858. <https://doi.org/10.1093/pcp/pcn166>
- Li H, Handsaker B, Wysoker A, Fennell T, Ruan J, Homer N, Marth G, Abecasis G, Durbin R (2009) The sequence alignment/map format and SAMtools. *Bioinformatics* 25(16):2078–2079
- Li J, Zhang M, Sun J, Mao X, Wang J, Liu H, Zheng H, Li X, Zhao H, Zou D (2020) Heavy metal stress-associated proteins in rice and arabidopsis: genome-wide identification, phylogenetics, duplication, and expression profiles analysis. *Front Genet* 11:477. <https://doi.org/10.3389/fgene.2020.00477>
- Li L-F, Pusadee T, Wedger MJ, Li Y-L, Li M-R, Lau Y-L, Yap S-J, Jamjod S, Rerkasem B, Hao Y, Song B-K, Olsen KM (2024) Porous borders at the wild-crop interface promote weed adaptation in Southeast Asia. *Nat Commun* 15:1182. <https://doi.org/10.1038/s41467-024-45447-0>
- Liu Y, Qi X, Gealy DR, Olsen KM, Caicedo AL, Jia Y (2015) QTL analysis for resistance to blast disease in US weedy rice. *Mol Plant Microbe Interact* 28:834–844
- Liu S, Li C, Wang H, Wang S, Yang S, Liu X, Yan J, Li B, Beatty M, Zastrow-Hayes G, Song S, Qin F (2020) Mapping regulatory variants controlling gene expression in drought response and tolerance in maize. *Genome Biol* 21:163. <https://doi.org/10.1186/s13059-020-02069-1>
- McLaren W, Gil L, Hunt SE, Riat HS, Ritchie GRS, Thormann A, Flicek P, Cunningham F (2016) The ensembl variant effect predictor. *Genome Biol* 17:122. <https://doi.org/10.1186/s13059-016-0974-4>
- Meng F, Zhao Q, Zhao X, Yang C, Liu R, Pang J, Zhao W, Wang Q, Liu M, Zhang Z, Kong Z, Liu J (2022) A rice protein modulates endoplasmic reticulum homeostasis and coordinates with a transcription factor to initiate blast disease resistance. *Cell Rep*. <https://doi.org/10.1016/j.celrep.2022.110941>
- Mishra RC, Singh A, Tiwari LD, Grover A (2016) Characterization of 5'UTR of rice ClpB-C/Hsp100 gene: evidence of its involvement in post-transcriptional regulation. *Cell Stress Chaperon* 21:271–283. <https://doi.org/10.1007/s12192-015-0657-1>
- Neik TX, Chai J-Y, Tan S-Y, Sudo MPS, Cui Y, Jayaraj J, Teo S-S, Olsen KM, Song B-K (2019) When west meets origins and spread of weedy rice between continental and island Southeast Asia. *G3 Genes Genomes Genet*. <https://doi.org/10.1534/g3.119.400021>
- Ng PC, Henikoff S (2003) SIFT: predicting amino acid changes that affect protein function. *Nucleic Acids Res* 31:3812–3814. <https://doi.org/10.1093/nar/gkg509>
- Patil SB, Khan MdK (2011) Germinated brown rice as a value added rice product: a review. *J Food Sci Technol* 48:661–667. <https://doi.org/10.1007/s13197-011-0232-4>
- Priya M, Dhanker OP, Siddique KHM, HanumanthaRao B, Nair RM, Pandey S, Singh S, Varshney RK, Prasad PVV, Nayyar H (2019) Drought and heat stress-related proteins: an update about their functional relevance in imparting stress tolerance in agricultural crops. *TAG Theor. Appl Genet Theor Angew Genet* 132:1607–1638. <https://doi.org/10.1007/s00122-019-03331-2>
- Pusadee T, Schaal BA, Rerkasem B, Jamjod S (2013) Population structure of the primary gene pool of *Oryza sativa* in Thailand. *Genet Resour Crop Evol* 60:335–353. <https://doi.org/10.1007/s10722-012-9839-1>
- Qiao B, Zhang Q, Liu D, Wang H, Yin J, Wang R, He M, Cui M, Shang Z, Wang D, Zhu Z (2015) A calcium-binding protein, rice annexin OsANN1, enhances heat stress tolerance by modulating the production of H₂O₂. *J Exp Bot* 66:5853–5866. <https://doi.org/10.1093/jxb/erv294>
- Qiu J, Jia L, Wu D, Weng X, Chen L, Sun J, Chen M, Mao L, Jiang B, Ye C (2020a) Diverse genetic mechanisms underlie worldwide convergent rice feralization. *Genome Biol* 21:1–11
- Qiu J, Jia L, Wu D, Weng X, Chen L, Sun J, Chen M, Mao L, Jiang B, Ye C, Turra GM, Guo L, Ye G, Zhu Q-H, Imaizumi T, Song B-K, Scarabel L, Merotto A, Olsen KM, Fan L (2020b) Diverse genetic mechanisms underlie worldwide convergent rice feralization. *Genome Biol* 21:70. <https://doi.org/10.1186/s13059-020-01980-x>
- Rathinasabapathi P, Purushothaman N, Ramprasad V, Parani M (2015) Whole genome sequencing and analysis of swarna, a widely cultivated indica rice variety with low glycemic index. *Sci Rep* 5:1–10
- Roma-Burgos N, San Sudo MP, Olsen KM, Werle I, Song B-K (2021) Weedy rice (*Oryza spp.*): what's in a name? *Weed Sci* 69:505–513. <https://doi.org/10.1017/wsc.2021.22>
- Sarkar NK, Kim Y-K, Grover A (2014) Coexpression network analysis associated with call of rice seedlings for encountering heat stress. *Plant Mol Biol* 84:125–143. <https://doi.org/10.1007/s11103-013-0123-3>
- Sarkar NK, Kotak S, Agarwal M, Kim Y-K, Grover A (2020) Silencing of class I small heat shock proteins affects seed-related attributes and thermotolerance in rice seedlings. *Planta* 251:26. <https://doi.org/10.1007/s00425-019-03318-9>

- Sarsu F (2018) Correction to: Pre-field screening protocols for heat-tolerant mutants in rice. Pre-field screening protocols for heat-tolerant mutants in rice. Springer, Cham, pp E1–E2
- Sen S, Chakraborty R, Kalita P (2020) Rice—not just a staple food: a comprehensive review on its phytochemicals and therapeutic potential. *Trends Food Sci Technol* 97:265–285. <https://doi.org/10.1016/j.tifs.2020.01.022>
- Shapiro SS, Wilk MB (1965) An analysis of variance test for normality (complete samples). *Biometrika* 52:591–611. <https://doi.org/10.1093/biomet/52.3-4.591>
- Singh A, Singh U, Mittal D, Grover A (2010) Genome-wide analysis of rice ClpB/HSP100. *BMC Genom* 11:95. <https://doi.org/10.1186/1471-2164-11-95>
- Song B-K, Chuah T-S, Tam SM, Olsen KM (2014) Malaysian weedy rice shows its true stripes: wild *Oryza* and elite rice cultivars shape agricultural weed evolution in Southeast Asia. *Mol Ecol* 23:5003–5017. <https://doi.org/10.1111/mec.12922>
- Stallworth S, Shrestha S, Schumaker B, Roma-Burgos N, Tseng T-M (2021) Screening diverse weedy rice (*Oryza sativa* ssp.) mini germplasm for tolerance to heat and complete submergence stress during seedling stage. *Front Agron* 3:642335. <https://doi.org/10.3389/fagro.2021.642335>
- Stuerz A (2019) Responses of rice growth to day and night temperature and relative air humidity—dry matter, leaf area, and partitioning. *Plants* 8:521. <https://doi.org/10.3390/plants8110521>
- Su A-Y, Niu S-Q, Liu Y-Z, He A-L, Zhao Q, Paré P, Li M-F, Han Q-Q, Ali Khan S, Zhang J-L (2017) Synergistic effects of bacillus amyloliquefaciens (GB03) and water retaining agent on drought tolerance of perennial ryegrass. *Int J Mol Sci* 18:2651. <https://doi.org/10.3390/ijms18122651>
- Subbaiyan GK, Waters DL, Katiyar SK, Sadananda AR, Vaddadi S, Henry RJ (2012) Genome-wide DNA polymorphisms in elite indica rice inbreds discovered by whole-genome sequencing. *Plant Biotechnol J* 10:623–634
- Sudianto E, Neik T, Tam S, Chuah T-S, Idris A, Olsen K, Song B-K (2016) Morphology of Malaysian weedy rice (*oryza sativa*): diversity, origin and implications for weed management. *Weed Sci*. <https://doi.org/10.1614/WS-D-15-00168.1>
- Sudo MPS, Yesudasan R, Neik TX, Masilamany D, Jayaraj J, Teo S-S, Rahman S, Song B-K (2021) The details are in the genome-wide SNPs: fine scale evolution of the Malaysian weedy rice. *Plant Sci* 310:110985. <https://doi.org/10.1016/j.plantsci.2021.110985>
- Sukkeoa S, Rerkasemb B, Jamjoda S (2017) Heat tolerance in Thai rice varieties. *ScienceAsia* 43:61–69
- Suzek BE, Wang Y, Huang H, McGarvey PB, Wu CH (2015) UniRef clusters: a comprehensive and scalable alternative for improving sequence similarity searches. *Bioinformatics*. <https://doi.org/10.1093/bioinformatics/btu739>
- Tang L, Wu A, Li S, Tuerdimaimaiti M, Zhang G (2023) Impacts of climate change on rice grain: a literature review on what is happening, and how should we proceed? *Foods* 12:536. <https://doi.org/10.3390/foods12030536>
- Tareke Woldegiorgis S, Wang S, He Y, Xu Z, Chen L, Tao H, Zhang Y, Zou Y, Harrison A, Zhang L, Ai Y, Liu W, He H (2019) Rice stress-resistant SNP database. *Rice* 12:97. <https://doi.org/10.1186/s12284-019-0356-0>
- Tian F, Hu X-L, Yao T, Yang X, Chen J-G, Lu M-Z, Zhang J (2021) Recent advances in the roles of HSFs and HSPs in heat stress response in woody plants. *Front Plant Sci* 12:704905. <https://doi.org/10.3389/fpls.2021.704905>
- Varadi M, Anyango S, Deshpande M, Nair S, Natassia C, Yordanova G, Yuan D, Stroe O, Wood G, Laydon A (2022) AlphaFold protein structure database: massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res* 50:D439–D444
- Vaser R, Adusumalli S, Leng SN, Sikic M, Ng PC (2016) SIFT missense predictions for genomes. *Nat Protoc* 11:1–9. <https://doi.org/10.1038/nprot.2015.123>
- Vigueira CC, Qi X, Song B, Li L, Caicedo AL, Jia Y, Olsen KM (2019) Call of the wild rice: *Oryza rufipogon* shapes weedy rice evolution in Southeast Asia. *Evol Appl* 12:93–104. <https://doi.org/10.1111/eva.12581>
- Wang P, Song H, Li C, Li P, Li A, Guan H, Hou L, Wang X (2017) Genome-wide dissection of the heat shock transcription factor family genes in arachis. *Front Plant Sci* 8:106. <https://doi.org/10.3389/fpls.2017.00106>
- Wang X, Shi X, Chen S, Ma C, Xu S (2018) Evolutionary origin, gradual accumulation and functional divergence of heat shock factor gene family with plant evolution. *Front Plant Sci* 9:71. <https://doi.org/10.3389/fpls.2018.00071>
- Wang Y, Yu Y, Huang M, Gao P, Chen H, Liu M, Chen Q, Yang Z, Sun Q (2020) Transcriptomic and proteomic profiles of II YOU 838 (*Oryza sativa*) provide insights into heat stress tolerance in hybrid rice. *PeerJ*. <https://doi.org/10.7717/peerj.8306>
- Woldegiorgis ST, Wu T, Gao L, Huang Y, Zheng Y, Qiu F, Xu S, Tao H, Harrison A, Liu W, He H (2022) Identification of heat-tolerant genes in non-reference sequences in rice by integrating pan-genome, transcriptomics, and QTLs. *Genes* 13:1353. <https://doi.org/10.3390/genes13081353>
- Xu Y, Chu C, Yao S (2021) The impact of high-temperature stress on rice: challenges and solutions. *Crop J* 9:963–976. <https://doi.org/10.1016/j.cj.2021.02.011>
- Yean R-A, Dilipkumar M, Rahman S, Song B-K (2021) A two-in-one strategy: target and nontarget site mechanisms both play important role in IMI-resistant weedy rice. *Int J Mol Sci* 22:982. <https://doi.org/10.3390/ijms22030982>
- Yoshida S, Satake T, Mackill D (1981) High-temperature stress in rice [study conducted at IRRI, Philippines]. *IRRI Res. Pap Ser Philipp*.
- Zhang Y, Chou S-D, Murshid A, Prince TL, Schreiner S, Stevenson MA, Calderwood SK (2011) The role of heat shock factors in stress-induced transcription. In: Calderwood SK, Prince TL (eds) *Molecular chaperones, methods in molecular biology*. Humana Press, Totowa, pp 21–32
- Zhang X, Li J, Liu A, Zou J, Zhou X, Xiang J, Rerksiri W, Peng Y, Xiong X, Chen X (2012) Expression profile in rice panicle: insights into heat response mechanism at reproductive stage. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0049652>
- Zhang Y, Min H, Shi C, Xia G, Lai Z (2021) Transcriptome analysis of the role of autophagy in plant response to heat stress. *PLoS ONE*. <https://doi.org/10.1371/journal.pone.0247783>
- Zhao P, Javed S, Shi X, Wu B, Zhang D, Xu S, Wang X (2020) Varying architecture of heat shock elements contributes to distinct magnitudes of target gene expression and diverged biological pathways in heat stress response of bread wheat. *Front Genet*. <https://doi.org/10.3389/fgene.2020.00030>
- Zhu M, Cheng Y, Wu S, Huang X, Qiu J (2022) Deleterious mutations are characterized by higher genomic heterozygosity than other genic variants in plant genomes. *Genomics*. <https://doi.org/10.1016/j.ygeno.2022.110290>
- Zou J, Liu A, Chen X, Zhou X, Gao G, Wang W, Zhang X (2009) Expression analysis of nine rice heat shock protein genes under abiotic stresses and ABA treatment. *J Plant Physiol* 166:851–861. <https://doi.org/10.1016/j.jplph.2008.11.007>

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.