communications biology

Article

https://doi.org/10.1038/s42003-024-06663-y

Horizontal gene transfer and symbiotic microorganisms regulate the adaptive evolution of intertidal algae, Porphyra sense lato

Wenlei Wang[1,2,3,7], Qijin Ge[4,7], Jian Wen[1,2,7], Han Zhang[1,2], Yanling Guo[1,2], Zongtang Li[1,2], Yan Xu[1,2], Dehua Ji[1,2], Changsheng Chen[1,2], Lidong Guo[4], Mengyang Xu[4], Chengcheng Shi[4,5], Guangyi Fan[4,5,6] ✉ & Chaotian Xie[1,2,3] ✉

Intertidal algae may adapt to environmental challenges by acquiring genes from other organisms and relying on symbiotic microorganisms. Here, we obtained a symbiont-free and chromosome-level genome of *Pyropia haitanensis* (47.2 Mb), a type of intertidal algae, by using multiple symbiont screening methods. We identified 286 horizontal gene transfer (HGT) genes, 251 of which harbored transposable elements (TEs), reflecting the importance of TEs for facilitating the transfer of genes into *P. haitanensis*. Notably, the bulked segregant analysis revealed that two HGT genes, sirohydrochlorin ferrochelatase and peptide-methionine (R)-S-oxide reductase, play a significant role in the adaptation of *P. haitanensis* to heat stress. Besides, we found Pseudomonas, Actinobacteria, and Bacteroidetes are the major taxa among the symbiotic bacteria of *P. haitanensis* (nearly 50% of the HGT gene donors). Among of them, a heat-tolerant actinobacterial strain (*Saccharothrix* sp.) was isolated and revealed to be associated with the heat tolerance of *P. haitanensis* through its regulatory effects on the genes involved in proline synthesis (*proC*), redox homeostasis (*ggt*), and protein folding (*HSP20*). These findings contribute to our understanding of the adaptive evolution of intertidal algae, expanding our knowledge of the HGT genes and symbiotic microorganisms to enhance their resilience and survival in challenging intertidal environments.

Red algae have thrived on Earth for more than 1.2 billion years, evolving into a unique lineage of photosynthetic eukaryotes[1] that were the first to adapt to intertidal environments[2]. They are characterized by their low evolutionary status, rich diversity, structurally simple yet diverse forms, strong resistance to stress, and economic value[3,4]. Moreover, red algae play a dual role as both the first eukaryotic hosts involved in primary endosymbiosis and as the plastid donors involved in secondary endosymbiosis, thereby endowing them with a combination of ancestral land plant-like stress resistance mechanisms and their own distinct adaptations[5,6]. Therefore, intertidal red algae are crucial subjects for studies on adaptive evolution. In recent years, global climate changes and associated factors, such as high temperatures, high light intensity, hyposaline, and disease outbreaks, have significantly restricted the sustainable development of seaweed cultivation[7,8]. A previous study found that temperature significantly affects seaweed cultivation, leading to a gradual northward shift in the cultivation zones of various red algal species[9]. Thus, elucidating the adaptive evolutionary mechanisms of intertidal red algae can provide new insights into plant evolution and genetic resources for breeding stress-resistant varieties.

To adapt to intertidal environments, red algae may have undergone genome reduction[2,10,11]. However, intertidal red algae remain highly resistant to stress, implying they may have acquired genes related to stress adaptations at critical points during evolution. Horizontal gene transfer (HGT) might promote functional convergence and the adaptability of red algae[11–13]. For example, the HGT genes associated with heat stress responses in

[1]Fisheries College, Jimei University, Xiamen, 361021, China. [2]State Key Laboratory of Mariculture Breeding, Fisheries College of Jimei university, Ningde, China. [3]Fujian Engineering Research Center of Aquatic Breeding and Healthy Aquaculture, Xiamen, 361021, China. [4]BGI Research, Qingdao, 266555, China. [5]Qingdao Key Laboratory of Marine Genomics, BGI Research, Qingdao, 266555, China. [6]BGI Research, Shenzhen, 518083, China. [7]These authors contributed equally: Wenlei Wang, Qijin Ge, Jian Wen. ✉e-mail: fanguangyi@genomics.cn; ctxie@jmu.edu.cn

Cyanidiophyceae (*Galdieria* and *Cyanidioschyzon*) include homologs of genes encoding heat shock protein 20, thermostable α-xylosidase, thermostable β-xylosidase, thioredoxin oxidoreductase, and a putative glutathione-specific γ-glutamylcyclotransferase 2, which can decrease the damage caused by high-temperature-induced free radicals[14]. Similarly, *Pyropia haitanensis*, an intertidal seaweed species, acquired unique lipoxygenase genes that mediate complex chemical defenses from bacteria, while also obtaining carbonic anhydrase (CA) genes that enhance survival during the sporophyte stage as well as carbon utilization[13]. Wang et al. [12] identified 51 genes from prokaryotes that were acquired by *Pyropia yezoensis* through HGT, including genes encoding superoxide dismutase (SOD), CAs, N-acyl-D-glucosamine 2-epimerase, glycine N-methyltransferase, and peroxidase. Thus, in addition to genome simplification, adaptive HGT may have also been a major factor contributing to the evolution of the genomes of red algae.

The evolutionary interaction between plants and microorganisms enabled plants to colonize land through the evolution of ancient gene modules and lineage-specific specializations more than 450 million years ago[15]. Wang et al.[16] integrated genome-wide association studies (GWAS), microbiome-wide association studies (MWAS), and microbiome genome-wide association studies (mGWAS) to identify 257 rhizoplane microbial biomarkers associated with six key agronomic traits (e.g., top second leaf width, main stem width, panicle diameter of the main stem) of *Setaria italica*, and screened four beneficial microorganisms that promote seedling height and root length through selective separation culture medium. Therefore, precise microbiome management facilitated the development of microbial inoculants that can increase crop yield and quality[16,17]. Similarly, the seaweed surface harbors highly diverse microbial communities, creating a unique microenvironment known as the "phycosphere"[18]. Phycosphere microorganisms differ significantly from those in the surrounding seawater in terms of composition and function, and the community structure can also undergo significant changes under different developmental or growth conditions[19–22]. Kessler et al. [23] found that *Ulva* is able to cultivate its microbial community by releasing chemical attractants and carbon sources. Among them, the dimethylsulfonopropionate (DMSP) released by *Ulva* can be perceived and utilized by bacteria with the DMSP demethylase gene (dmdA), contributing to *Ulva's* growth and morphogenesis. Deutsch et al. [24] demonstrated that endophytic bacteria (*Bacillus subtilis*) isolated from *Ulva* sp. can be reintroduced into their original host and influence the biodiversity of associated microbial communities. This interaction may aid in protecting *Ulva* sp. from pathogens and other opportunistic microorganisms. In summary, highly diverse and taxonomically specific microbial communities may be closely related to the adaptation of seaweeds, and identifying a pathway through microbiota manipulation is also important for improving seaweed aquaculture[25]. However, it is unclear which core phycosphere microorganisms are involved in key HGT events and how they regulate seaweed stress responses.

Here, we propose that specific microbial taxa within the phycosphere community may play pivotal roles as potential donors or facilitators of HGT events, thus impacting the abiotic stress tolerance of *P. haitanensis*. To test this hypothesis, this study constructed a high-quality *P. haitanensis* genome and identified HGT events and the donors of HGT genes. Furthermore, a bulked segregant analysis (BSA) was performed to identify linked quantitative trait loci and candidate HGT genes associated with the high-temperature tolerance of *P. haitanensis*. Additionally, we analyzed the changes in the microbial community structure in different *P. haitanensis* strains that varied regarding temperature tolerance. We also isolated the beneficial bacteria via selective medium culture, enhancing the high-temperature tolerance of *P. haitanensis*. The study results provide new insights into the stress resistance of intertidal seaweeds as well as genes potentially useful for breeding novel stress-resistant germplasm.

Results

Symbiont-free and chromosome-level genome assembly

To assemble a symbiont-free *P. haitanensis* genome, we sequenced 51.6 Gb (109-fold) Oxford Nanopore long reads and 23.2 Gb (49-fold) PacBio HiFi-CCS reads of DNA isolated from nuclei (Supplementary Table 1). The draft genome assembly consisted of 176.5 Mb with 174 contigs, which is notably larger than the previously estimated *P. haitanensis* genome size (49.67 Mb) because of the considerable contamination by sequences from symbiont bacteria. To distinguish the symbiont sequences from the host sequences, we classified seven distinct clusters according to the Hi-C contact matrix (Supplementary Fig. 1), including five *P. haitanensis* chromosomes (47.2 Mb). The number of chromosomes is consistent with the results of the karyotype analysis (Supplementary Tables 2, 3). The *P. haitanensis* genome is notably smaller than the genomes of *Porphyra umbilicalis* (87.9 Mb) and *P. yezoensis* (107.6 Mb). We also aligned the candidate symbiont sequences (129.3 Mb) from the remaining clusters to the sequences in the NR database, which found 98.11% of these sequences matched sequences from various bacteria, such as *Hyphomonas* sp. Mor2, *Fuerstia marisgermanicae*, and *Silicimonas algicola*. This observation indicates that the high quality of the symbiont-free *P. haitanensis* genome was obtained (Supplementary Data 1). Interestingly, the analysis of the Hi–C contact matrix detected an uneven segmented pattern in the interaction map of chromosome 1, which consisted of three segments (A/B/C) (Supplementary Fig. 1). To validate the assembly accuracy, we thoroughly examined the links between two weak interaction points among the three segments of chromosome 1 and found no breakpoints. The normalized Hi–C matrices at different resolutions and the syntenic relationship with the published *P. yezoensis* genome provided additional evidences that the chromosome-level genome was accurately assembled.

The proportion of repetitive sequences in *P. haitanensis* genome (24.21%) was less than the corresponding proportions for *P. umbilicalis* (49.78%) and *P. yezoensis* (46.25%) (Supplementary Table 4). This observation suggests a positive correlation between the repetitive sequence content and genome size, and also highlights the considerable diversity in the genome sizes and repetitive sequence proportions in the family Bangiaceae. However, despite the general decrease in the length of LTR-RT sequences, the number and total length of endogenous retrovirus-K (ERVK) sequences within the LTR transposable elements (TEs) are notably higher in *P. haitanensis* (109 sequences, 21.21 kb) compared to *P. yezoensis* (4 sequences, 0.31 kb) and *P. umbilicalis* (10 sequences, 1.01 kb) (Supplementary Table 5). The phylogenetic analysis of ERVK sequences further indicated these sequences expanded specifically in the *P. haitanensis* genome (Supplementary Fig. 2). We also compared the distribution of TEs between *P. haitanensis* and *P. yezoensis*. In *P. yezoensis*, TEs were evenly distributed across all chromosomes, whereas in *P. haitanensis*, the TEs (mainly Gypsy, LTR, and Copia) tended to cluster in segment B of chromosome 1 (Supplementary Table 6-8 and Supplementary Data 2). Interestingly, the lengths of TEs on chromosome 1 were negatively correlated with gene expression, suggesting the TEs inserted in this chromosome can influence gene expression (Supplementary Fig. 3).

Horizontal gene transfer, which is a natural process during which one organism transfers its genetic material to another organism, contributed significantly to the evolution of symbiont species like algae. We identified 2,325 HGT genes across 10 species, including *Cyanidioschyzon merolae*, *Cyanidiococcus yangmingshanensis*, *Galdieria sulphuraria*, *Porphyridium purpureum*, *P. umbilicalis*, *P. yezoensis*, *Chondrus crispus*, *Gracilariopsis chorda*, and *Emiliania huxleyi* (Supplementary Data 3). A total of 54 HGT-related gene families were associated with the ancestral HGTs in the basal species (Fig. 1a, Supplementary Fig. 4). A total of 286 HGT genes were detected in *P. haitanensis*, accounting for approximately 2.25% of the entire genome. We also examined the relationships between TE insertions and HGT genes (Fig. 1b). Among these HGT genes, 251 showed TE presence within their upstream and downstream 5 kb regions (Supplementary Data 4). This suggests that TE might have mediated the occurrence of

horizontal gene transfer events. Moreover, according to sequence similarities and phylogenetic relationships, 97.9% of the HGT genes originated from bacteria. Specifically, the main sources of HGT genes were Pseudomonadota, Actinomycetota, Bacteroidota, and Planctomycetota species (54.9%), each of which contributing 75, 39, 23, and 20 HGT genes, respectively (Fig. 1c, Supplementary Fig. 5).

HGT genes for adapting the high temperature

To identify the major HGT genes responsible for the adaptation of *P. haitanensis* to heat stress, we performed a BSA sequencing (BSA-seq) analysis, which involved genotyping individuals with extreme phenotypes. The heat-resistant mixed pool (HR-Pool) and heat-sensitive mixed pool (HS-Pool) were constructed for the BSA-seq analysis of algal cell morphology and various phenotypic traits, including *Fv/Fm*, length, and fresh weight under high-temperature conditions. Specifically, compared with the control treatment, the 7-day high-temperature treatment resulted in no significant cell morphological changes in the HR-Pool, whereas almost all of the cells in the HS-Pool were dead (Fig. 2a, b). Additionally, the *Fv/Fm* value, daily increase in length, and daily increase in fresh weight were significantly higher in the HR-Pool than in the HS-Pool (Fig. 2c–e). According to the 95% confidence interval of Δ(SNP-index), we identified two candidate genomic regions on chromosomes 2 and 3, which consisted of 1.46 Mb (1–1460 kb) and 0.22 Mb (6110–6330 kb),

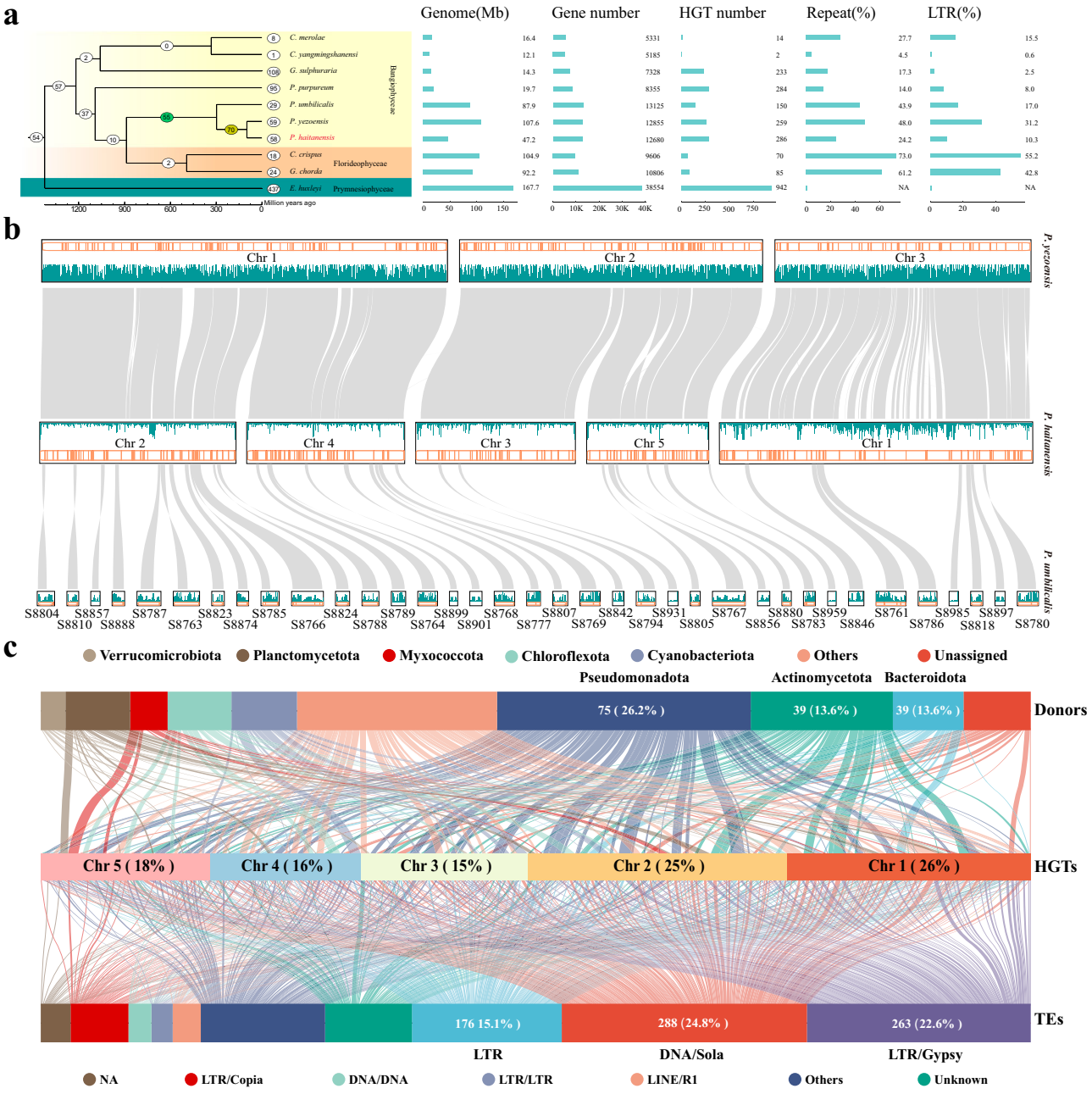

Fig. 1 | Comparative analysis of horizontal gene transfer (HGT). a Phylogenetic tree presenting the relationships among ten algal species. The tree also provides information on their respective genome sizes, gene numbers, HGT gene numbers, repeat proportions, and LTR proportions. The numbers displayed to the left of each species in the tree represent the count of HGT genes in unique gene families. **b** Genomic collinearity demonstrated by the comparison of *P. yezoensis*, *P. haitanensis*, and *P. umbilicalis*. The figure displays the distribution of HGT (orange) and TE density in 20-kb windows (green, values range from 0 to 1) on chromosomes. **c** Sankey diagram illustrating the flow of HGT genes from donor species to TE classes in *P. haitanensis*.

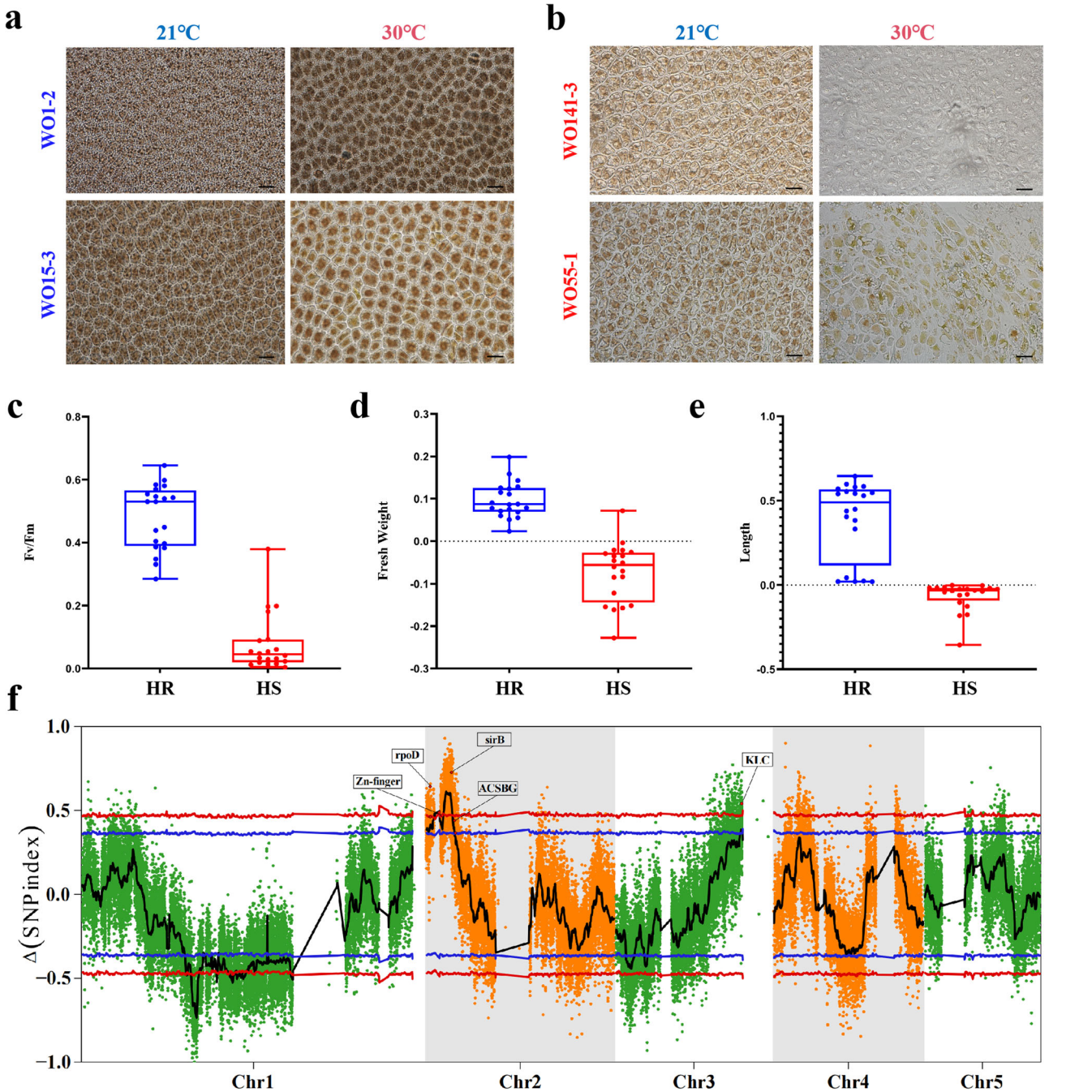

Fig. 2 | Bulked segregant analysis of candidate HGT genes related to the high-temperature tolerance of *P. haitanensis*. **a**, **b** Morphological characteristics of the heat-resistant strain (**a**) and the heat-sensitive strain (**b**) observed before (21 °C) and after the high-temperature treatment (7 days at 30 °C). **c**–**e** Physiological and biochemical indices of the HR-Pool (heat-resistant mixed pool) and HS-Pool (heat-sensitive mixed pool) following the high-temperature treatment (7 days at 30 °C). Measurements include **c** *Fv/Fm*, **d** daily increase in fresh weight, and **e** daily increase in length. **f** Bulked segregant analysis [Δ(SNP-index)] on the candidate intervals related to *P. haitanensis* thallus thermostability. The *x*-axis represents the physical positions of the *P. haitanensis* chromosomes. Red and blue lines indicate statistical significance at $P < 0.01$ and $P < 0.05$, respectively. Abbreviations in the box: sirohydrochlorin ferrochelatase (sirB), RNA polymerase primary sigma factor (rpoD), long-chain-fatty-acid-CoA ligase ACSBG (ACSBG), kinesin light chain (KLC).

respectively; these regions included 468 candidate genes (Fig. 2f). On the basis of the 95% confidence interval of the Euclidean distance (ED) and G-statistic (G), we identified 10 and 8 candidate genomic regions, respectively, which included 871 candidate genes (Supplementary Fig. 6). Among these candidate genes, ten HGT genes, such as sirohydrochlorin ferrochelatase (*SIRB*) and peptide-methionine (R)-S-oxide reductase (*MSRB*), were HGT genes related to the adaptation of *P. haitanensis* to heat stress (Fig. 2f, Supplementary Fig. 6).

Symbiotic bacteria and genes contribute to the high-temperature adaptation

To investigate the symbiotic relationships between intertidal algae and microorganisms, we found that Pseudomonadota (55.1%), Planctomycetota (26.3%), Bacteroidota (10.9%) and Actinomycetota (3.2%) were the main symbiotic bacteria of *P. haitanensis* genome based on the assembled symbiotic sequences (Supplementary Fig. 5b, Supplementary Data 1). To further examine how the symbiotic microbial community affected the tolerance of *P. haitanensis* to heat stress, we conducted a 16 S rDNA

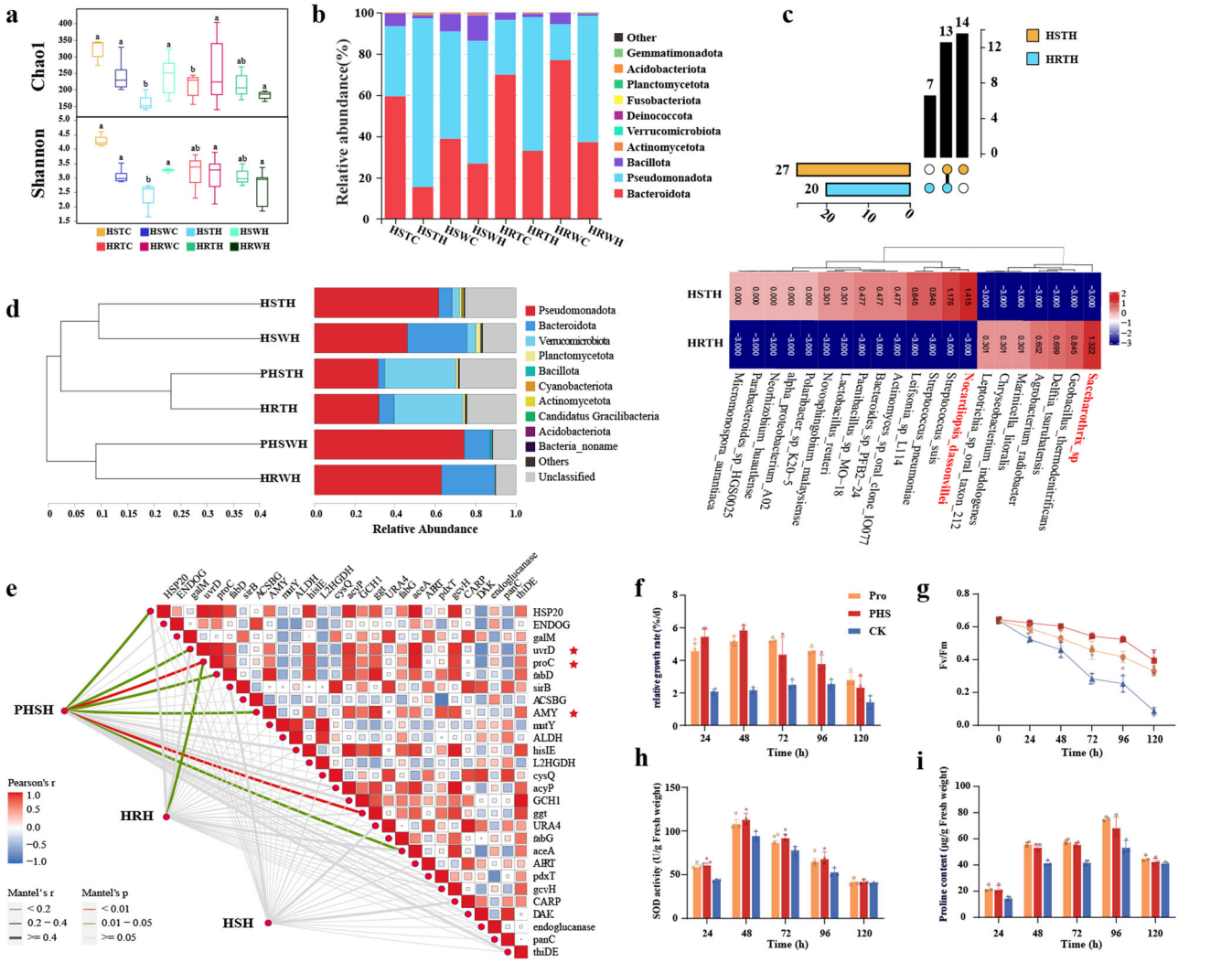

Fig. 3 | Screening, isolation, and regulation of key HGT donor bacteria in *P. haitanensis* and verification of their ability to enhance the thermal tolerance of thalli. a Changes in the Chao1 and Shannon indices in water and thallus samples from the HS strain and HR strain under high-temperature conditions. T and W represent thallus and water samples, respectively, whereas C and H represent normal- and high-temperature treatments, respectively. **b** Relative abundance of the community composition (phylum) in the algal intermicrobial environment. **c** Venn diagram and heatmap illustrating the differences in the number and abundance of shared and unique bacteria between the thallus samples of the HS strain and HR strain. **d** Analysis of the composition of the macrogenomic community (phylum) and sample clusters. *P* represents the addition of *Saccharothrix* sp. **e** Analysis of the correlation between different treatments of intermicrobial microorganisms and HGT homologous genes. The thickness of the line segment represents the size of Mantel's *r*. The color of the line segment represents the size of Mantel's *p*, with red representing $p < 0.01$, green representing $0.01 < p < 0.05$, and gray representing $p \geq 0.05$; The color of the square represents the correlation between genes, with red indicating a positive correlation and blue indicating a negative correlation, and the size of the square reflects the magnitude of the correlation. HRH: HR strain; HSH: HS strain; PHSH: HS strain supplemented with *Saccharothrix* sp. "★" denotes actinomycetota as the candidate donors for these HGT genes. Gene abbreviations are provided in Supplementary Data 6f–i) Effects of the addition of proline (Pro) and *Saccharothrix* sp. (PHS) on the physiological and biochemical indices of the HS strain (CK) thallus under high-temperature conditions. Measurements include **f** relative growth rate, **g** maximum quantum yield (*Fv*/*Fm*), **h** superoxide dismutase (SOD) activity, and **i** free proline content.

sequencing analysis to determine the differences in the microbial communities of the heat-resistant (HR) strain and the heat-sensitive (HS) strain under normal (21 °C) and high-temperature (30 °C) conditions. The Chao1 and Shannon indices indicated that the microbial diversity of the HS strain decreased significantly following the high-temperature treatment. Conversely, the microbial diversity of the HR strain was unchanged by heat stress, indicative of a relatively stable symbiotic microbial community (Fig. 3a). The principal coordinate analysis (PCoA) of the algal samples revealed a significant difference between the high-temperature treatment group and the normal-temperature treatment group, with some overlapping regions, suggestive of the presence of both fixed and specific microbes within each group (Supplementary Fig. 7). The dominant bacterial taxa differed between the normal- and high-temperature treatment groups. Specifically, under normal conditions, Bacteroidota was the dominant taxon, with average relative abundances of 49.06% and 73.35% in the HS and HR strains, respectively. In contrast, under high-temperature conditions, the dominant taxon was Pseudomonadota, with average relative abundances of 70.50% and 49.06% in the HS and HR strains, respectively (Fig. 3b).

To detect the common and specific operational taxonomic unit (OTU, a classification unit based on biological sequence analysis, used to describe different biological species or subspecies) in different samples, we constructed a Venn diagram using the OTU abundance information. Seven bacteria were specific to the surface of the HR strain (*Saccharothrix* sp., *Geobacillus thermodenitrificans*, *Delftia tsuruhatensis*, *Agrobacterium radiobacter*, *Leptotrichia* sp. oral taxon 212, *Marinicella litoralis*, and

Chryseobacterium indologenes) (Fig. 3c). Among the HR strain-specific bacteria, *Saccharothrix* sp. was the most abundant. To clarify the role of *Saccharothrix* sp. during the adaptation to high temperatures, we isolated a *Saccharothrix* sp. strain from the HR strain using a carefully selected culture medium (Supplementary Fig. 8). The HS strain phenotype and cell morphology were affected by the presence (experimental group) and absence (control group) of *Saccharothrix* sp. in the culture medium. Additionally, the high-temperature treatment (30 °C) of the HS strain in the control group resulted in extensive ulceration and cell death. However, in the experimental group, there was no obvious ulceration and the cells remained relatively normal (Supplementary Fig. 9). Moreover, the HS strain in the experimental group had a significantly higher relative growth rate, maximum quantum yield (*Fv/Fm*), SOD activity, and proline content than the HS strain in the control group (Supplementary Fig. 10). Accordingly, *Saccharothrix* sp. can enhance the heat resistance of the HS strain.

To explore the mechanism through which *Saccharothrix* sp. improves the heat resistance of *P. haitanensis*, we analyzed the metagenomic data of the HS strain supplemented with *Saccharothrix* sp. under high-temperature conditions. Interestingly, the community composition and cluster analysis showed that PHSTH (experimental group thallus sample) and HRTH (HR strain thallus sample) clustered together, as did PHSWH (experimental group water sample) and HRWH (HR strain water sample) (Fig. 3d), indicating that the community composition of the HS strain supplemented with *Saccharothrix* sp. was similar to that of the HR strain. Considered together, these results suggest *Saccharothrix* sp. may be an important factor contributing to the differences in heat tolerance between the HS and HR strains.

To identify the genes responsive to high temperatures, we analyzed the association between the composition of algal symbiotic bacteria and gene transcription levels in the HR strain (HRH), HS strain (control group, HSH), and HS strain supplemented with *Saccharothrix* sp. (experimental group, PHSH). The experimental group showed a significant positive correlation with the genes *proC*, *ggt*, *HSP20*, *fabd*, *uvrD*, *AMY*, and *aceA* ($P < 0.05$), while the control group exhibited a negative correlation with these genes (Fig. 3e). Interestingly, all of the seven genes were identified as HGT gene homologs in the *P. haitanensis* genome. Further expression analysis of the genes depicted in Fig. 3e revealed that, compared to the HSH group (HSTH and HSWH), the PHSH group (PHSTH and PHSWH) showed upregulated expression of 23 genes, including *proC, ggt, HSP20, fabd, uvrD, AMY*, and *aceA* (Supplementary Fig. 11A). Among these genes, *proC* is involved in the proline synthesis pathway. It was observed that in this pathway, 1-pyrroline-2-carboxylate reductase [NAD(P)H] (*lhpI*), proline iminopeptidase (*pip*), and proline racemase (*prdF*) also showed upregulated expression in the PHSH group(Supplementary Fig. 11B). Surprisingly, treating the HS strain with proline significantly increased the relative growth rate under high-temperature conditions (Fig. 3e). Moreover, *Fv/Fm*, SOD activity, and the free proline content were higher in the experimental group than in the control group (Fig. 3f–i). Overall, our findings demonstrate that the introduction of *Saccharothrix* sp. isolated from the HR strain substantially increased the heat resistance of the HS strain.

HGT genes involved in other environmental adaptation

To further investigate the relationship between bacteria and algae, we collected in situ samples of *P. haitanensis* from four sea areas in Fujian, China: Xiapu, Nanri Island, Huian, and Zhangpu for 16 S amplicon analysis. Based on the abundance-occupancy analysis, we determined that 148 OTUs belonging to Proteobacteria (core bacterial proportion: 58.8%), Bacteroidota (33.8%), Actinobacteriota (6.8%), and Desulfobacterota (0.7%) were identified as core bacterial communities (Supplementary Fig. 12A, Supplementary Data 5). Further analysis revealed significant differences in the Shannon index and Bray-Curtis index between core and occasional bacterial communities (Supplementary Fig. 12B, C), indicating distinct community composition structures between the two groups. Additionally, analysis of habitat niche breadths showed that core bacterial communities have a broader ecological niche compared to occasional bacterial communities (Supplementary Fig. 12D), implying that core bacterial communities can adapt to a wider range of ecological environments. Random forest and Mantel analyses demonstrated significant correlations between algal-associated core bacteria and salinity, pH, dissolved organic carbon (DOC), dissolved oxygen (DO), particulate organic carbon (POC), temperature, and phosphate (PO_4-P) (Fig. 4a,b). In summary, 16sRNA amplicon analysis indicates that the core bacteria in these samples rely on these environmental factors to maintain community structure stability. The investigation also involved analyzing the effects of different stressors, including heat stress, dehydration stress, N and P availability, light intensity stress, acidification, salt stress, and conchocelis maturation. As a result, changes in the expression levels of 283 HGT genes were detected (fold-change > 1.5, $q < 0.05$) (Supplementary Fig.13, Supplementary Data 4). Among these stressors, salt stress, heat stress, and conchocelis maturation treatment induced differential expression in 256, 236, and 206 HGT genes, respectively. This indicates that these three stressors rank as the top three in terms of their impact (Fig. 4c). Notably, 60 genes, including those encoding HSP20 (HSP20 family protein), fabD ([acyl-carrier-protein] S-malonyltransferase), AMY (alpha-amylase), groES (chaperonin GroES), RDH12 (retinol dehydrogenase 12), and OPR (12-oxophytodienoic acid reductase), were affected by all seven conditions, suggesting these HGT genes may have diverse functions or mediate a universal mechanism underlying the responses of *Pyropia* species to environmental stress (Fig. 4c, d). In addition, the expression levels of five, three, two, and one HGT genes were exclusively affected by conchocelis maturation, heat stress, salt stress, and acidification, respectively. Specifically, genes in the acetyltransferase (GNAT) family as well as genes encoding ADCK (aarF domain-containing kinase), a SNARE-interacting protein-like protein, ST2A (hydroxyjasmonate sulfotransferase), and a VOC domain profile were responsive only to conchocelis maturation. The high-temperature treatment increased the expression of genes encoding acyP (acylphosphatase), rnhB (ribonuclease HII), and a winged helix DNA-binding domain-containing protein, whereas salt stress induced the expression of genes encoding an AAA-ATPase and an unknown protein (evm.model.ctg000033_arrow_pilon.60). Acidification increased the expression of a gene encoding an S4 domain-containing protein.

Discussion

In land plants, HGT genes frequently accumulate, replicate, or undergo functional divergence within descendant populations, thereby contributing to diversification and adaptive evolution. Researchers have identified 593 gene families transferred to both charophytes and land plants, with two major HGT events associated with the early evolution of streptophytes (first events) and the origin of land plants (second events), explaining how land plants acquired a large number of genes in their early evolutionary stages[26]. In this study, 55 HGT gene families were identified in *Pyropia/Porphyra* spp., which is the most HGT family among the 9 red algae genomes that have been reported. Specifically, the HGT genes in *Pyropia/Porphyra* spp. are mainly related to the metabolism (energy metabolism, lipid metabolism, and amino acid metabolism, etc.), and also participate in the regulation of cellular processes and genetic information processing, such as replication and repair, fold, sorting and degradation, as well as transport and catabolism (Supplementary Fig. 14). Moreover, these pathways have been confirmed to participate in the response mechanisms of *Pyropia/Porphyra* spp. to environmental stressors such as high temperature[27], hypersaline stress[28], hyposaline stress[29], dehydration[30] and high light intensity[31]. This indicated that HGTs enabled *Pyropia/Porphyra* spp. to acquire crucial genes with diverse functions that facilitated adaptations to intertidal zone environments, especially under genome reduction conditions. Twenty-four HGT gene families are common to the three *Pyropia/Porphyra* spp. examined in this study. For example, the SOD[32], methyltransferase[33], and E3 ubiquitin–protein ligase[34,35] gene families are closely related to the abiotic stress resistance of *Pyropia/Porphyra* spp. Similarly, the HGT genes

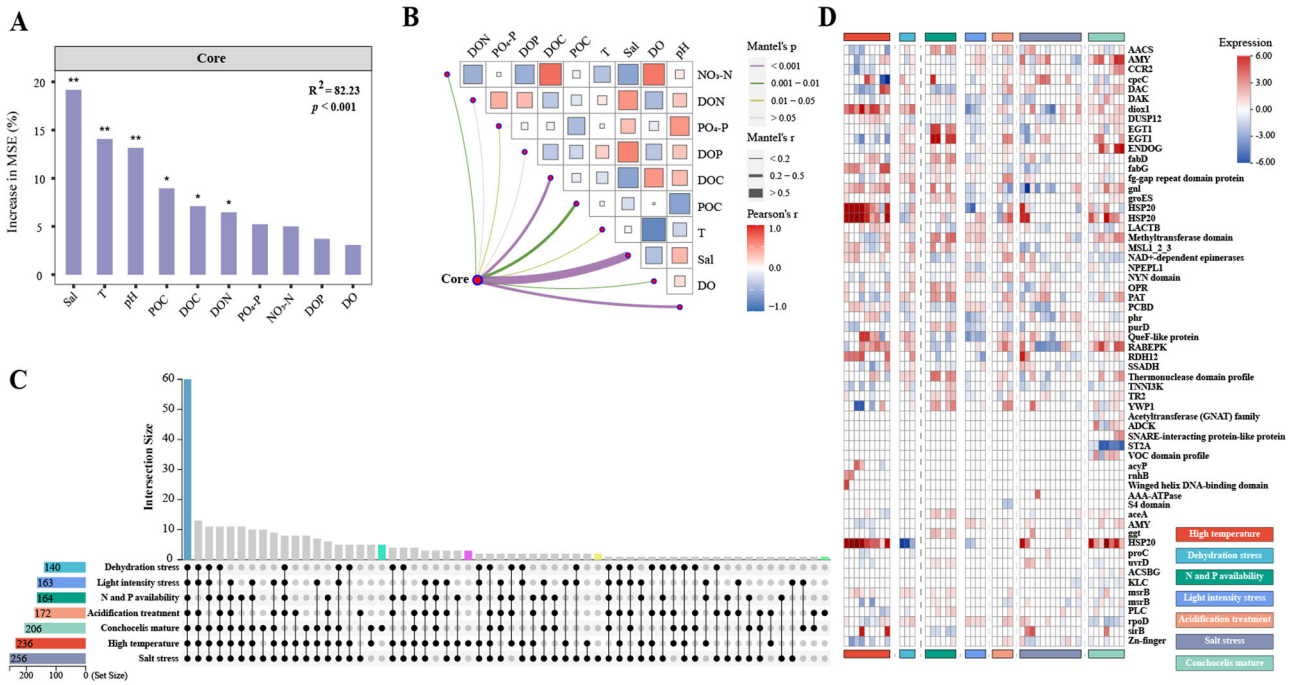

Fig. 4 | Screening of core microorganisms in the phycosphere of *P. haitanensis* collected from different cultivation areas and analysis of the influence of environmental factors on the core microorganisms and the expression of HGT genes. **a** The correlation between core bacterial communities and environmental variables was assessed using a random forest analysis. Significance levels are denoted as $***P < 0.001$, $**P < 0.01$, and $*P < 0.05$. R^2 and P values represent the goodness of fit and significance of the global model, respectively. **b** Correlations between core bacterial communities and environmental variables were evaluated by Mantel tests. The width and color of the edge corresponding to the R value and significance, respectively. The color gradient presents Pearson correlation coefficients for environmental variables. **c** Visualization of the intersection of differentially expressed HGT genes under heat stress, dehydration stress, N and P availability, light intensity stress, acidification, salt stress, and conchocelis maturation treatments. **d** Heatmap analysis of HGT genes with expression levels that were affected by all seven conditions or specifically by conchocelis maturation, heat stress, salt stress, or acidification.

encoding methionine S-methyltransferase-like proteins are associated with plant salt tolerance[26,36]. The HGT genes unique to *P. haitanensis* (i.e., not in *P. yezoensis* and *P. umbilicalis*), including HSP-encoding genes, autophagy-related genes, and DNA repair-related genes, may have contributed to the adaptation of *P. haitanensis* to high-temperature conditions at relatively low latitudes[27,37,38]. Horizontal gene transfer events occur in streptophytes, and they are also a major reason for the functional convergence in intertidal red algae with a unique evolutionary status. Additionally, TEs considered vital components of genomes, often promote HGTs and genome rearrangements while also facilitating the acquisition of genes that confer selective advantages to the host[39,40]. In the present study, we determined that of the 286 HGT genes identified in *P. haitanensis*, 265 (e.g., *HSP20*, *proC*, *sirB*, *MSRB*, and *rpoD*) are associated with TE insertions (Fig. 1c). These findings reflect the close association between either TE insertions and HGT events in *P. haitanensis*.

For many multicellular eukaryotes, the number of genes in the symbiotic microbial community exceeds the number of genes in their own genomes, making the associated microbes important sources of genetic diversity and genes mediating adaptive evolution[41,42]. Bacteria are typically the primary donors of genes that are horizontally transferred to plants[26,43]. In accordance with this earlier finding, in the current study, 98% of the HGT genes in *P. haitanensis* were derived from bacteria (Fig. 1c, Supplementary Fig. 5). A combined analysis of bacterial donors, genome-filtered microbial datasets, and the core microbiota in the phycosphere of *P. haitanensis* collected from natural coastal regions revealed that the symbiotic bacteria in *P. haitanensis* are primarily from three taxa (*Pseudomonas*, Actinobacteria, and Bacteroidetes), accounting for nearly 50% of the potential HGT gene donors (Supplementary Fig. 5, Supplementary Fig. 11). Lu et al. [22] recently reported that 14 core genera from eight families belonging to the phyla Proteobacteria, Bacteroidota, Verrucomicrobiota, and Actinobacteriota account for an average of 43.5% (*Gelidium* sp.), 53.9% (*Grateloupia* sp.), 58.3% (*Ulva* sp.), and 48.8% (*Saccharina* sp.) of all phycosphere bacteria, but only 5.7% and 1.5% of the bacteria in seawater and sediment samples, respectively. Accordingly, there appears to be a close relationship between these core phycosphere bacteria and seaweeds. Therefore, a comprehensive analysis of the key HGT gene donors within the core bacteria in the *P. haitanensis* phycosphere is critical for further clarifying the interactions between algae and microbes as well as their adaptation to intertidal zones.

Further analyses involving amplicon sequencing, isolation, and co-culturing showed that the actinobacterium *Saccharothrix* sp. significantly enhances the high-temperature tolerance of *P. haitanensis* (Supplementary Fig. 9, Supplementary Fig. 10). Similarly, Hmani et al. [44] found that inoculating *Rathayibacter festucae* IH2 (Actinobacteria) and *Roseovarius aestuarii* G8 (Proteobacteria) at 18 °C promotes *Ulva* growth but inhibits it at 30 °C. Interestingly, this inhibitory effect is alleviated by additional inoculation of *Roseovarius* sp. MS2 (Proteobacteria). This indicates a close and complex relationship between seaweeds and their associated microorganisms. The metagenomic analysis of the community composition and functional genes revealed that the *Saccharothrix* sp. treatment group had the strongest correlation with *proC* and *ggt* ($P < 0.01$; Fig. 3e). The *proC* gene, which was acquired from actinobacteria via HGT, is related to proline synthesis. Under heat stress conditions, treatment with *Saccharothrix sp.* not only promoted the upregulation of genes related to the proline synthesis pathway but also increased the proline content of *P. haitanensis* (Supplementary Fig. 10D, Supplementary Fig. 11B). Proline serves as an osmoprotectant, antioxidant, and signaling molecule that regulates the stress responses of higher plants and *Pyropia* spp[45–47]. Similarly, studies on *Ulva*[48], *Chlamydomonas reinhardtii*[49] and *Ectocarpus siliculosus*[50] have found that abiotic stress also significantly induces the accumulation of proline in organisms. Earlier research indicated GGT (E.C.2.3.2.2) metabolizes extracellular reduced glutathione, thereby helping to salvage extracellular glutathione, while also potentially contributing to the control of the

extracellular redox state[51]. Glutathione levels must increase rapidly during the initial exposure to salt stress for *P. yezoensis* thalli to mitigate the effects of ROS bursts[52]. Additionally, *HSP20*, *uvrD*, *AMY*, and *fabD* were also correlated with the *Saccharothrix* sp. treatment ($P < 0.05$; Fig. 3e). This finding highlights the importance of certain genes acquired through HGT for regulating the synthesis of critical metabolites, DNA repair, and protein folding during the response of intertidal algae to abiotic stress. Furthermore, the addition of *Saccharothrix* sp. stabilizes the complex microbial community structure of *P. haitanensis* under heat-stress conditions. The complexity and stability of microbial networks tend to increase in response to global warming[53]. Hence, appropriately regulating the core microbiota associated with the stability and complexity of symbiotic communities is an important approach. This helps ensure that *P. haitanensis* can adapt to changes in intertidal environmental conditions. In conclusion, this study suggests how HGT genes from bacteria, especially actinobacteria (e.g., *Saccharothrix* sp.), can mediate the adaptation of *P. haitanensis* to heat stress by enhancing proline synthesis and stabilizing the phycosphere microbial community.

The present study confirmed the crucial role of key microbiota in the resistance of seaweed to environmental stresses. However, it is based solely on the screening results of specific microbes associated with high-temperature resistant strains. A bacterial strain was isolated using selective culture media, which is currently a common utilization method for rhizosphere or gut microbiota, namely by isolating superior strains to develop microbial fertilizers, microbial pesticides, probiotics, prebiotics, and other bioproducts[54–56]. However, how to fully utilize the entire microbiome to enhance the host's stress resistance at a holistic level and how to transition from applying exogenous beneficial microbes to activating the inherent overall function of the microbiome in situ are core issues that urgently need to be addressed[54]. For the environmentally dominated variable rhizosphere microbiome, developing high-throughput microbial culture-omics technologies to isolate and identify a richer set of beneficial microbes and enhance their application potential is essential. For the host-genetics-determined microbiome, employing mGWAS and gene editing techniques to explore genetic variations and functional genes that dominate the symbiotic microbiome and elucidating the molecular mechanisms and regulatory networks by which hosts recruit beneficial microbes are crucial for harnessing the full function of the entire microbiome in situ[16,57].

Methods

Materials and culture conditions

The experimental strains, including pure paternal line DH115-2, pure maternal line W28, and pure hybrid offspring WO57-1-1, WO57-1-2, WO57-1-3, and WO57-1-4, were selected from a doubled haploid (DH) population developed by the Laboratory of Germplasm Improvements and Applications of *Pyropia haitanensis* at Jimei University. The thalli were cultured at 21 ± 0.5 °C with a 12-h light (50–60 μmol photons m^{-2} s^{-1})/12-h dark photoperiod. The culture medium (Provasoli's enrichment solution) was replaced every 2 days. Genomic DNA was extracted from the thalli of DH115-2, W28, and the four hybrid offspring lines for the subsequent sequencing.

De novo genome assembly via long-read sequencing

The 51.6 Gb Oxford Nanopore long reads (N50 = 25.2 kb) and 23.2 Gb PacBio HiFi reads (N50 = 18.3 kb) were used along with NextDenovo (https://github.com/Nextomics/NextDenovo/) for the de novo genome assembly. Errors in the draft genome assembly were corrected using HiFi reads and GCpp Arrow[58] and then NGS short reads and Pilon[59]. The 174 contigs were further clustered to scaffolds using 3D-DNA and 231.5 Gb Hi–C reads[60]. The read pairs were pre-validated using HiCPro[61] and pre-analyzed using Juicer[62]. The scaffolding order and orientation were manually checked and re-adjusted via Juicebox visualization[62].

The contamination of the target seaweed genome was first prevented by isolating nuclei because the prokaryotic cells of symbiotic microbes do not contain nuclei. The subsequent bioinformatic-based decontamination pipeline comprised four parallel parts. First, Symbiont-Screener was used to separate host contigs and contaminants without references. The clustering was completed using an unsupervised machine learning method to analyze the disparity in statistical and biological features among different species, including GC contents and 3-mer frequencies[63]. Notably, the haplotype-specific information obtained from the cultured samples increased the clustering efficiency, but not substantially because of the relatively small host genome. Second, on the basis of the BLAST search of the NCBI reference database, the taxon-annotated GC blobplots generated by Blobology were used to visualize and identify the seaweed sequences[64]. Third, significant chromatin contacts were selected from pair-wise contact frequencies available in the Hi–C dataset. The spatial segregation in the heatmap was used to construct the chromosomes and isolate the host and symbiont. Alternatively, contigs that sufficiently overlapped previously published *Neoporphyra* genome sequences were marked[13]. The sequence similarity was evaluated using QUAST[65]. The final contig group was determined according to the results of all four methods. The sequences that were not aligned to the published genome during the BLAST search were considered bacterial sequences. Five pseudo-chromosomes were assembled.

To validate the accuracy of the chromosome-level assembly, we realigned the Hi–C reads from the parental materials to the chromosome-level genome assembly using the HiCpro software and generated Hi-C matrices at various resolutions (5, 10, 15, 20, 25, 30, 35, 40, 45, 50, 100, and 500 kb). Using the hicCorrectMatrix diagnostic_plot function from HiCExplorer, we calculated the bin count threshold at different resolutions and normalized the matrices according to the Iterative Correction and Eigenvector decomposition (ICE) method. Finally, we used the hicPlotMatrix function to visualize the normalized matrices at different resolutions. Furthermore, to compute the A/B compartments of the chromosomal regions, we transformed the normalized Hi–C interaction matrix into an observed/expected matrix and used hicPCA to calculate the eigenvectors of the covariance matrix.

Repetitive sequence identification and functional gene annotation

Repeat elements were identified via homolog and de novo strategies. RepeatMasker (v4.0.6) and RepeatProteinMask (http://www.repeatmasker.org/) as well as the Repbase (v21.01) database were used to search for repetitive sequences. The de novo identification of TEs was completed using LTR-Finder (v1.0.6)[66] and RepeatModeler (v1.0.8) (http://www.repeatmasker.org/RepeatModeler/) based on our assembled genome to generate a de novo repeat library. The databases with repeat elements generated by both methods were combined to produce a non-redundant library, which served as the input library for the classification into the corresponding superfamilies using RepeatMasker (http://www.repeatmasker.org/RepeatMasker/). Tandem repeat elements were detected using TRF (v4.07)[67].

To determine the LTR insertion times during evolution, the intact LTR pair sequences were extracted after processing the LTR-Finder results using LTR-Retriever and then aligned using MUSCLE (v3.8.31)[68]. The genetic distance was calculated using the Kimura nucleotide substitution model and the default parameters of DISTMAT (EMBOSS-6.5.7). Finally, the insertion times were calculated using the following formula as previously described: $T = K/2r$, where r represents the number of substitutions per synonymous mutation site per year (7×10^{-9}).

Gene models were predicted using homology-based, ab initio-based, and RNA-seq-based methods. For the ab initio predictions, the default settings of AUGUSTUS (v3.4.0)[69] and GENEMARK (v4.68)[70] were applied. Moreover, the Maker (v2.31.8)[71] annotation pipeline was used with the repeat-masked genome assembly. For the homology-based annotation, five algal species (*P. umbilicalis*, *Cyanidiococcus yangmingshanensis*, *Cyanidioschyzon merolae*, *Galdieria sulphuraria*, and *Porphyridium purpureum*) were selected to infer the protein coding genes using GEMOMA (v1.7.1)[72]. The RNA-seq data were mapped to the genome assembly using HISAT2 (v2.1.0) (Kim et al. 2019)[73]. TransDecoder (v5.5.0) (https://github.com/TransDecoder/TransDecoder) was used to identify potential coding regions

in transcript sequences. Finally, EVidenceModeler[74] was used to integrate the gene models predicted using the three methods.

Identification of HGT genes and their donors

To identify potential HGT genes, we conducted a targeted analysis of the protein sequences encoded by the genomes of 10 algae using the Alien Index. Each sequence was aligned to sequences in the Nr database using Diamond, after which the top 1000 significant hits, sorted by bit-score, were used to calculate the AI score for each query gene as previously described[43,75]. We then retrieved up to 60 corresponding BLAST hit sequences from the database for the genes with AI scores >0 (no more than 3 and 12 sequences were retrieved for each genus and class, respectively) for the phylogenetic analysis. The comparison task was performed using Prank (v170427) and the matrix was modified using TrimAl (v1.4) (-automated1). We selected alignment matrices with a length exceeding 50 amino acids, more than five aligned genes, and a scaffold longer than 100 kb. The phylogenetic tree was constructed using IQ-TREE (v2.0.3). The shape of the HGT gene phylogenetic tree was determined as follows: if a query gene was nested within a branch of a prokaryotic organism, it was considered to have been derived via prokaryotic transfer. Nested nodes were defined as nodes consisting of one or more individual system branches containing both query genes and prokaryotic genes, but lacking genes from eukaryotes other than Chlorophyta, Streptophyta, Rhodophyta, Haptophyta, Oomycota, Bacillariophyta, Cercozoa, Endomyxa, Foraminifera, Ciliophora, Perkinsozoa, and Apicomplexa. The analysis of species-specific or shared HGT gene families is based on the OrthoMCL classification of homologous genes across all species, followed by an upset analysis of the gene families containing HGT genes in each species to determine the number of gene families with species-specific HGT genes and their proportion among all gene families (Supplementary Fig. 4).

In HGT gene phylogenetic trees, the species closest to the recipient species is generally believed to be the donor species for the HGT gene. In this study, we constructed a phylogenetic tree for each HGT gene in *P. haitanensis* using the top five hits for each genus (no more than three sequences were retrieved) in the original database as well as the top five hits for each homologous gene in the remaining nine species (no more than three sequences were retrieved). In the phylogenetic tree, the prokaryotic species that formed a monophyletic group with *P. haitanensis* and had the closest clustering relationship was considered to be the donor of the gene. If many prokaryotic species formed a cluster that was grouped with *P. haitanensis*, the prokaryotic species with the gene that was most similar to the *P. haitanensis* HGT gene was considered to be the donor of the gene.

Differential gene expression analysis

We used HISAT2 (v2.1.0) to align the RNA-seq raw reads for the different treatments to the mRNA sequences derived from the *P. haitanensis* genome. The aligned reads from bam files were counted using featureCounts (v2.0.6). To analyze differential gene expression, we used DESeq2 with the parameters set as follows: $\log_2$(fold-change) ≥ 0.585 and adjusted $P < 0.05$.

16S rDNA and metagenomic analysis

Two *P. haitanensis* strains (WO72-4 and WO49-1) that differed regarding their tolerance to heat stress were obtained from a *P. haitanensis* germplasm resource bank in Fujian province. The initial experiments confirmed WO72-4 and WO49-1 were HR and HS strains, respectively. For the high-temperature treatment, samples were maintained at 30 °C, with all other cultivation conditions the same as those used for cultivating the algae included in the genome sequencing experiments. After the high-temperature treatment, the thalli and culture medium were collected. The culture medium was filtered using a membrane with 0.22 μm pores to obtain the filtrate. The collected samples were immediately frozen using liquid nitrogen and then stored at −80 °C. The DNeasy PowerWater Kit (Qiagen, Hilden, Germany) was used to extract DNA from the water samples. The hypervariable regions (V5–7) of the bacterial 16 S rRNA gene were amplified by PCR using the barcoded bacteria-specific primers 799F (AACMGGATTAGATACCCKG) and 1193R (ACGTCATCCCCACCTTCC). Raw reads were filtered for quality using Cutadapt (v1.9.1; http://cutadapt.readthedocs.io/en/stable/) with specific filtering conditions to obtain high-quality clean reads. The reads were compared with a reference database (Silva database; https://www.arb-silva.de/) to detect and remove chimeric sequences. The retained clean reads were used for OTU clustering and species annotation. The OTU abundance data were normalized using a standard number of sequences (i.e., the sample with the fewest sequences). The normalized data were used for the subsequent analyses of alpha diversity and beta diversity and for predicting functions.

Probiotics screening and preparation

The collected culture medium was vigorously mixed using a vortex mixer, after which 100 μL was added to 900 μL sterilized seawater. The resulting solution was thoroughly mixed. Simultaneously, the thalli were repeatedly rinsed with sterilized seawater to remove loose impurities from the surface. Two thalli that underwent the high-temperature treatment were placed in a sterilized centrifuge tube containing 5 mL of sterilized seawater. Approximately 4 mL of sterilized seawater was added before the sample was vortexed for 10 min. The thalli were removed to obtain the thallus suspension, which was added to the solid culture medium (Gauze's Medium No. 1). The inoculated medium was inverted and incubated at 30 °C for 2–3 days. Bacterial colonies with differing morphological characteristics were selected and isolated via streaking on fresh medium 2–3 times. The isolates were transferred to a slanted medium in tubes prior to storage at 4 °C. Additionally, they were added to glycerol (200 μL) in tubes for long-term storage at −20 °C. Gauze's Medium No. 1 consisted of the following (per liter): 20 g soluble starch, 0.5 g NaCl, 1 g KNO_3, 0.5 g $K_2HPO_4{\cdot}3H_2O$, 0.5 g $MgSO_4{\cdot}7H_2O$, 0.01 g $FeSO_4{\cdot}7H_2O$, and 15–20 g agar, pH 7.4–7.6. Bacterial cultures were incubated at 30 °C for 18–24 h, and then 1–3 mL was collected for the extraction of genomic DNA using the Ezup Column Bacteria Genomic DNA Purification Kit (Sangon Biotech Co., Ltd., Shanghai, China). Bacterial species were identified by sequencing and then comparing the obtained genomic DNA sequences with publicly available bacterial genomes.

Co-culture treatment of symbiotic bacteria

Three healthy WO49-1 (HS strain) thalli that were 3.8–4.2 cm long were inoculated with 1 mL bacterial inoculum in a 500 mL culture medium. The cultivation conditions were the same as those used for the high-temperature treatment. The effects of the isolated bacteria on the thermal tolerance of the thalli were evaluated by analyzing specific parameters, including the relative thallus growth rate, maximum photochemical quantum yield (*Fv*/*Fm*), and SOD activity (Wang et al. 2019)[76].

Metagenome analysis and data processing

The HiPure Bacterial DNA Kit (Guangzhou, China) was used to extract genomic DNA from samples. The DNA quality was assessed using the Qubit fluorometer (Thermo Fisher Scientific, Waltham, MA) and the NanoDrop spectrophotometer (Thermo Fisher Scientific). Genomic DNA that satisfied the Illumina sequencing quality requirements was initially fragmented by sonication to obtain 350 bp fragments. The ends of the fragmented DNA were repaired and an A-tail and a sequencing adapter were added using the NEBNext® Ultra™ DNA Library Prep Kit (NEB, USA) for the subsequent Illumina sequencing analysis. The DNA fragments (300–400 bp) were enriched by PCR amplification and purified using the AMPure XP system (Beckman Coulter, Brea, CA, USA). The quality of the sequencing library was assessed using the 2100 Bioanalyzer (Agilent, Santa Clara, CA), whereas the sequencing library was quantified by real-time PCR. FASTP (v0.18.0) was used to filter the raw data from the Illumina platform as follows: reads containing ≥10% unknown nucleotides (N) were removed; bases with a Phred quality score ≤20 in ≥50% of the reads were removed; and reads containing adapters were discarded. The clean reads were assembled using MEGAHIT (v1.1.2) with k-mer values of 21, 41, 61, 81, 101, 121, and 141. Genes were predicted using contigs >500 bp and MetaGeneMark (v3.38). All gene sequences that were at least 300 bp long were selected. Sequences with ≥95% similarity and >90% coverage were clustered together

using CD-HIT (v4.6). The longest sequence in each cluster was selected as the representative sequence and designated as a unigene. Unigene abundance was calculated using Pathoscope (v2.0.7). Moreover, the unigenes were annotated by aligning them to sequences in the Nr (RefSeq non-redundant protein database) and KEGG (Kyoto Encyclopedia of Genes and Genomes) databases using DIAMOND (v0.9.24).

Bulked segregant analysis

A DH population ($N = 480$) was constructed using DH115-2 (male parent) and W28 (female parent). The thalli (8 ± 1 cm long) of 200 strains were cultured at 20 °C and 30 °C for 7 days. The daily growth rate in terms of length fresh weight, and *Fv*/*Fm* were determined. In addition, samples were examined using a microscope (40× magnification). For the BSA-seq analysis, 20 lines were selected from HR-Pool and HS-Pool according to the phenotypic data. Genomic DNA was extracted from the free-living conchocelis of the parents and the F_2 populations. The genomic DNA (at least 3 μg) extracted from the bulks was used to construct paired-end sequencing libraries (insert size of 500 bp) using the Paired-End DNA Sample Prep kit (Illumina Inc., San Diego, CA, USA). The resequencing data for the parental materials and mixed pools (HR-Pool and HS-Pool) were aligned to the *P. haitanensis* reference genome using the Burrows–Wheeler Aligner[77]. The SAMtools program was used to detect SNPs and insertions/deletions (InDels), which were filtered using the GATK VariantFiltration program and appropriate standards. The SNPs and InDels associated with distorted segregation or those containing sequencing errors were discarded. To determine the physical positions of each SNP, the ANNOVAR software[78] was used to align and annotate SNPs and InDels.

A homozygous SNP from W28 was used to modify the genotype of the publicly available reference genome, resulting in a custom reference sequence for calculating the SNP-index[79]. The SNP-index was calculated only for the homozygous SNP loci between the parents for the BSA-based localization. The SNP-index represents the frequency of the parental (DH115-2) allele in the population. The SNPs and InDels were filtered using a standard pipeline. Briefly, markers with a read depth <5× in each parent and markers with missing genotypes were removed. Additionally, markers with a read depth <10× or >500× in each bulk were excluded to eliminate low-confidence markers because of low coverage or markers that may be in repetitive regions (i.e., inflated read depth). Markers with a SNP-index of 0.7 in both pools were also excluded. The Δ(SNP-index) was calculated for both HR-Pool and HS-Pool. The average Δ(SNP-index) and SNP-index of each pool were calculated using a 100 kb sliding window with a step size of 10 kb. The window in which the average Δ(SNP-index) was greater than the 95th percentile of the genome-wide average Δ(SNP-index) was designated as a significant window. The genes in this interval were identified as candidate genes. Similarly, the average *G* value and ED^5 value were calculated using the same window parameters, but windows with fewer than five SNPs/InDels were discarded. For the *G* value and ED^5 value, windows with an average greater than the 95th percentile of the genome-wide average value were treated as significant windows. Overlapping or adjacent significant windows were then merged into large significant genomic regions. The genes in these intervals were considered to be candidate genes.

Environmental DNA sequencing and analysis

Pyropia haitanensis thalli and water samples were collected from 10 sites in four coastal aquaculture areas in Fujian province, namely Xiapu, Nanri Island, Huian, and Zhangpu (east longitude: 120.096°–117.823°; north latitude: 26.839°–23.970°). The phycosphere bacteria were washed with 500 mL PBS solution and filtered through a nitrocellulose filter membrane. The filter membranes were stored at −80 °C prior to the DNA extraction. The water temperature (*T*), pH, salinity (Salt), and dissolved oxygen (DO) levels were measured at all sites. Each water sample was examined in terms of the nitrate-nitrogen (NO_3-N), PO_4-P, dissolved organic matter (DOM) [including DOC, dissolved organic nitrogen (DON), and dissolved organic phosphorus (DOP)], and POC contents. These analyses were conducted as previously described[80].

The PowerWater DNA isolation kit (Qiagen) was used to extract DNA from each sample. The quality and quantity of DNA were determined using a NanoDrop 2000 spectrophotometer (Bio-Rad Laboratories Inc., Hercules, CA, USA). Four metabarcoding primers were used to amplify 80 eDNA samples. The V3–V4 region of the bacterial 16 S ribosomal RNA gene was amplified by PCR using the primer pair 341F (ACTCCTACGGGAGGCAGCAG) and 806R (GGACTACHVGGGTWTCTAAT)[81]. The PCR amplifications were performed in triplicate. Each 20 μL reaction mixture consisted of 4 μL 5× FastPfu Buffer, 2 μL 2.5 mM dNTPs, 0.8 μL each primer (5 μM), 0.4 μL FastPfu Polymerase, and 10 ng template DNA. The thermal cycling conditions were as follows: 95 °C for 2 min; 25 cycles of 95 °C for 30 s, 55 °C for 30 s, and 72 °C for 30 s; 72 °C for 5 min (final extension). Amplicons were analyzed in 2% agarose gels, purified using the AxyPrep DNA Gel Extraction Kit (Axygen Biosciences, Union City, CA, USA), and quantified using QuantiFluor™-ST (Promega, Madison, WI, USA).

All analyses were conducted using R (v4.1.1), with the results visualized using the "ggplot2" R package[82]. The R package "EasyStat" (v0.1.0; https://github.com/taowenmicro/EasyStat) was used for statistical analyses. The R package "vegan" was used to calculate the Bray–Curtis distance of the biological communities and the ED of the environmental variables. According to the method described by Shade et al. [83], core and occasional bacteria were identified, with the Bray–Curtis distance and Shannon index of the bacterial community calculated using the R package "vegan". Mantel tests were conducted to assess the effects of environmental variables on the core bacterial community structure. The R package "Random Forest" was used to determine the correlation between the environmental variables and core bacterial communities, after which Levins' niche breadth index was calculated using the R package "spaa"[84].

Reporting summary

Further information on research design is available in the Nature Portfolio Reporting Summary linked to this article.

Data availability

The source data associated with the figures are provided in Supplementary Data 1-6. The genome data has been deposited in the CNGB Sequence Archive (CNSA, https://db.cngb.org/cnsa/) under accession number CNP0003571. The BSA sequencing results of high-temperature tolerance traits of *Pyropia haitanensis* have been uploaded to NCBI, with the accession number PRJNA1068197. Amplified raw sequences of phycosphere bacteria reported in this paper have been deposited in the National Center for Biotechnology Information's Short Reads Archive (SRA) database under the BioProject numbers PRJNA1068807. The Metagenomics sequencing results of high-temperature tolerance traits of *Pyropia haitanensis* phycosphere bacteria have been uploaded to NCBI, with the accession number PRJNA1111680.

Received: 10 February 2024; Accepted: 31 July 2024;
Published online: 11 August 2024

References

1. Butterfield, N. J. Bangiomorpha pubescens n. gen., n. sp. implications for the evolution of sex, multicellularity, and the Mesoproterozoic/Neoproterozoic radiation of eukaryotes. *Paleobiology* **26**, 386–404 (2000).
2. Bhattacharya, D. et al. When less is more: red algae as models for studying gene loss and genome evolution in eukaryotes. *Crit. Rev. Plant Sci.* **37**, 81–99 (2018).
3. Blouin, N. A., Brodie, J. A., Grossman, A. C., Xu, P. & Brawley, S. H. Porphyra: a marine crop shaped by stress. *Trends Plant Sci.* **16**, 29–37 (2011).
4. He, P., Zhang, Z., Zhang, X. & Ma, J. *Seaweed Cultivation (in Chinese)*. (Science Press, BeiJing, 2018).
5. Bhattacharya, D., Yoon, H. S. & Hackett, J. D. Photosynthetic eukaryotes unite: endosymbiosis connects the dots. *Bioessays* **26**, 50–60 (2004).
6. Moreira, D., Le Guyader, H. & Philippe, H. J. N. The origin of red algae and the evolution of chloroplasts. *Nature* **405**, 69–72 (2000).

7. Chung, I. K., Sondak, C. F. & Beardall, J. The future of seaweed aquaculture in a rapidly changing world. *Eur. J. Phycol.* **52**, 495–505 (2017).
8. Duarte, C. M., Bruhn, A. & Krause-Jensen, D. A seaweed aquaculture imperative to meet global sustainability targets. *Nat. Sustain.* **5**, 185–193 (2022).
9. Zhang, Y.-Y. et al. Temperature sensitivity of marine macroalgae for aquaculture in China. *Aquaculture* **567**, 739262 (2023).
10. Lee, J. et al. Analysis of the draft genome of the red seaweed Gracilariopsis chorda provides insights into genome size evolution in Rhodophyta. *Mol. Biol. Evolut.* **35**, 1869–1886 (2018).
11. Cho, C. H. et al. Genome-wide signatures of adaptation to extreme environments in red algae. *Nat. Commun.* **14**, 10 (2023).
12. Wang, D. et al. Pyropia yezoensis genome reveals diverse mechanisms of carbon acquisition in the intertidal environment. *Nat. Commun.* **11**, 4028 (2020).
13. Chen, H. et al. Insights into the ancient adaptation to intertidal environments by red algae based on a genomic and multiomics investigation of Neoporphyra haitanensis. *Mol. Biol. Evolut.* **39**, msab315 (2022).
14. Rossoni, A. W. et al. The genomes of polyextremophilic cyanidiales contain 1% horizontally transferred genes with diverse adaptive functions. *Elife* **8**, e45017 (2019).
15. Delaux, P.-M. & Schornack, S. Plant evolution driven by interactions with symbiotic and pathogenic microbes. *Science* **371**, eaba6605 (2021).
16. Wang, Y. et al. GWAS, MWAS and mGWAS provide insights into precision agriculture based on genotype-dependent microbial effects in foxtail millet. *Nat. Commun.* **13**, 5913 (2022).
17. Bai, B. et al. The root microbiome: community assembly and its contributions to plant fitness. *J. Integr. Plant Biol.* **64**, 230–243 (2022).
18. Bell, W. & Mitchell, R. Chemotactic and growth responses of marine bacteria to algal extracellular products. *Biol. Bull.* **143**, 265–277 (1972).
19. Saha, M. & Weinberger, F. Microbial "gardening" by a seaweed holobiont: surface metabolites attract protective and deter pathogenic epibacterial settlement. *J. Ecol.* **107**, 2255–2265 (2019).
20. Wang, W. et al. The cultivation of *Pyropia haitanensis* has important impacts on the seawater microbial community. *J. Appl. Phycol.* **32**, 2561–2573 (2020).
21. Ghaderiardakani, F., Quartino, M. L. & Wichard, T. Microbiome-dependent adaptation of seaweeds under environmental stresses: a perspective. *Front. Mar. Sci.* **7**, 575228 (2020).
22. Lu, D.-C., Wang, F.-Q., Amann, R. I., Teeling, H. & Du, Z.-J. Epiphytic common core bacteria in the microbiomes of co-located green (Ulva), brown (Saccharina) and red (Grateloupia, Gelidium) macroalgae. *Microbiome* **11**, 126 (2023).
23. Kessler, R. W., Weiss, A., Kuegler, S., Hermes, C. & Wichard, T. Macroalgal–bacterial interactions: role of dimethylsulfoniopropionate in microbial gardening by Ulva (Chlorophyta). *Mol. Ecol.* **27**, 1808–1819 (2018).
24. Deutsch, Y., Ofek-Lalzar, M., Borenstein, M., Berman-Frank, I. & Ezra, D. Re-introduction of a bioactive bacterial endophyte back to its seaweed (Ulva sp.) host, influences the host's microbiome. *Front. Mar. Sci.* **10**, 1099478 (2023).
25. Li, J., Weinberger, F., de Nys, R., Thomas, T. & Egan, S. A pathway to improve seaweed aquaculture through microbiota manipulation. *Trends Biotechnol.* **41**, 545–556 (2023).
26. Ma, J. et al. Major episodes of horizontal gene transfer drove the evolution of land plants. *Mol. Plant* **15**, 857–871 (2022).
27. Wang, W. et al. Comparative transcriptome analysis between heat-tolerant and sensitive Pyropia haitanensis strains in response to high temperature stress. *Algal Res.* **29**, 104–112 (2018).
28. Wen, J. et al. Comparative analysis of proteins involved in energy metabolism and protein processing in *Pyropia haitanensis* at different salinity levels. *Front. Mar. Sci.* **7**, 415 (2020).
29. Wen, J. et al. The mechanism of maintaining intracellular homeostasis in the red alga *Pyropia haitanensis* under hyposaline stress. *Front. Mar. Sci.* **9**, 928617 (2022).
30. Shi, J. et al. Insight into transketolase of *Pyropia haitanensis* under desiccation stress based on integrative analysis of omics and transformation. *BMC Plant Biol.* **19**, 1–16 (2019).
31. Ji, D. et al. Investigating the mechanisms underlying the low irradiance-tolerance of the economically important seaweed species *Pyropia haitanensis*. *Life* **13**, 481 (2023).
32. Ji, D. et al. Superoxide dismutase genes in *Pyropia haitanensis*: molecular cloning, characterization and mRNA expression. *Acta Oceanol. Sin.* **35**, 101–111 (2016).
33. Yu, C. et al. Detection of changes in DNA methylation patterns in *Pyropia haitanensis* under high-temperature stress using a methylation-sensitive amplified polymorphism assay. *J. Appl. Phycol.* **30**, 2091–2100 (2018).
34. Wang, W. et al. Transcriptomic study to understand thermal adaptation in a high temperature-tolerant strain of *Pyropia haitanensis*. *PLoS ONE* **13**, e0195842 (2018).
35. Wang, W. et al. A RING type ubiquitin ligase PhCUL4 is involved in thermotolerance of *Pyropia haitanensis*. *Algal Res.* **59**, 102448 (2021).
36. Ogawa, S. & Mitsuya, S. S-methylmethionine is involved in the salinity tolerance of Arabidopsis thaliana plants at germination and early growth stages. *Physiol. Plant.* **144**, 13–19 (2012).
37. Wang, W. et al. Early signaling events in the heat stress response of *Pyropia haitanensis* revealed by phosphoproteomic and lipidomic analyses. *Algal Res.* **67**, 102837 (2022).
38. Chang, J. et al. Molecular mechanism underlying *Pyropia haitanensis* PhHsp22-mediated increase in the high-temperature tolerance of *Chlamydomonas reinhardtii*. *J. Appl. Phycol.* **33**, 1137–1148 (2021).
39. Gao, C. et al. Horizontal gene transfer in plants. *Funct. Integr. Genomics* **14**, 23–29 (2014).
40. Horne, T., Orr, V. T. & Hall, J. P. How do interactions between mobile genetic elements affect horizontal gene transfer? *Curr. Opin. Microbiol.* **73**, 102282 (2023).
41. Rosenberg, E., Sharon, G., Atad, I. & Zilber-Rosenberg, I. The evolution of animals and plants via symbiosis with microorganisms. *Environ. Microbiol. Rep.* **2**, 500–506 (2010).
42. Schönknecht, G., Weber, A. P. & Lercher, M. J. Horizontal gene acquisitions by eukaryotes as drivers of adaptive evolution. *Bioessays* **36**, 9–20 (2014).
43. Fan, X. et al. Phytoplankton pangenome reveals extensive prokaryotic horizontal gene transfer of diverse functions. *Sci. Adv.* **6**, eaba0111 (2020).
44. Hmani, I., Ghaderiardakani, F., Ktari, L., El Bour, M. & Wichard, T. High-temperature stress induces bacteria-specific adverse and reversible effects on Ulva (Chlorophyta) growth and its chemosphere in a reductionist model system. *Bot. Mar.* **67**, 131–138 (2024).
45. Hayat, S. et al. Role of proline under changing environments: a review. *Plant Signal. Behav.* **7**, 1456–1466 (2012).
46. Wang, W. et al. Investigating the mechanisms underlying the hyposaline tolerance of intertidal seaweed, *Pyropia haitanensis*. *Algal Res.* **47**, 101886 (2020).
47. Alvarez, M. E., Savouré, A. & Szabados, L. Proline metabolism as regulatory hub. *Trends Plant Sci.* **27**, 39–55 (2022).
48. Ghaderiardakani, F., Langhans, L., Kurbel, V. B., Fenizia, S. & Wichard, T. Metabolite profiling reveals insights into the species-dependent cold stress response of the green seaweed holobiont Ulva (Chlorophyta). *Environ. Exp. Bot.* **200**, 104913 (2022).
49. Fal, S., Aasfar, A., Rabie, R., Smouni, A. & Arroussi, H. E. Salt induced oxidative stress alters physiological, biochemical and metabolomic responses of green microalga *Chlamydomonas reinhardtii*. *Heliyon* **8**, e08811 (2022).
50. Dittami, S. M. et al. Integrative analysis of metabolite and transcript abundance during the short-term response to saline and oxidative

stress in the brown alga Ectocarpus siliculosus. *Plant, cell Environ.* **34**, 629–642 (2011).
51. Ferretti, M. et al. Gamma-glutamyl transferase in the cell wall participates in extracellular glutathione salvage from the root apoplast. *N. Phytol.* **181**, 115–126 (2009).
52. Yu, B., Yang, J., Niu, J. & Wang, G. Antioxidant responses to hyperosmolarity stress in the intertidal *Pyropia yezoensis* (Bangiales, Rhodophyta). *Algal Res.* **48**, 101930 (2020).
53. Yuan, M. M. et al. Climate warming enhances microbial network complexity and stability. *Nat. Clim. Change* **11**, 343–348 (2021).
54. Xun, W. et al. Dissection of rhizosphere microbiome and exploiting strategies for sustainable agriculture. *N. Phytol.***242**, 2401–2410 (2024).
55. Li, X. et al. Rational management of the plant microbiome for the Second Green Revolution. *Plant Commun.* (2024).
56. Yang, H. et al. ABO genotype alters the gut microbiota by regulating GalNAc levels in pigs. *Nature* **606**, 358–367 (2022).
57. Edwards, J. A. et al. Genetic determinants of switchgrass-root-associated microbiota in field sites spanning its natural range. *Curr. Biol.* **33**, 1926–1938.e1926 (2023).
58. Kronenberg, Z. N. et al. High-resolution comparative analysis of great ape genomes. *Science* **360**, eaar6343 (2018).
59. Walker, B. J. et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PloS ONE* **9**, e112963 (2014).
60. Dudchenko, O. et al. De novo assembly of the Aedes aegypti genome using Hi-C yields chromosome-length scaffolds. *Science* **356**, 92–95 (2017).
61. Servant, N. et al. HiC-Pro: an optimized and flexible pipeline for Hi-C data processing. *Genome Biol.* **16**, 1–11 (2015).
62. Durand, N. C. et al. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* **3**, 99–101 (2016).
63. Xu, M. et al. Symbiont-screener: a reference-free tool to separate host sequences from symbionts for error-prone long reads. *Front. Mar. Sci.***10**, 1087447 (2023).
64. Kumar, S., Jones, M., Koutsovoulos, G., Clarke, M. & Blaxter, M. Blobology: exploring raw genome data for contaminants, symbionts and parasites using taxon-annotated GC-coverage plots. *Front. Genet.* **4**, 237 (2013).
65. Gurevich, A., Saveliev, V., Vyahhi, N. & Tesler, G. QUAST: quality assessment tool for genome assemblies. *Bioinformatics* **29**, 1072–1075 (2013).
66. Xu, Z. & Wang, H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* **35**, W265–W268 (2007).
67. Benson, G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* **27**, 573–580 (1999).
68. Edgar, R. C. Quality measures for protein alignment benchmarks. *Nucleic Acids Res.* **38**, 2145–2153 (2010).
69. Stanke, M., Diekhans, M., Baertsch, R. & Haussler, D. Using native and syntenically mapped cDNA alignments to improve de novo gene finding. *Bioinformatics* **24**, 637–644 (2008).
70. Brůna, T., Lomsadze, A. & Borodovsky, M. GeneMark-EP+: eukaryotic gene prediction with self-training in the space of genes and proteins. *NAR Genomics Bioinforma.* **2**, lqaa026 (2020).
71. Cantarel, B. L. et al. MAKER: an easy-to-use annotation pipeline designed for emerging model organism genomes. *Genome Res.* **18**, 188–196 (2008).
72. Keilwagen, J., Hartung, F., Paulini, M., Twardziok, S. O. & Grau, J. Combining RNA-seq data and homology-based gene prediction for plants, animals and fungi. *BMC Bioinforma.* **19**, 1–12 (2018).
73. Kim, D., Paggi, J. M., Park, C., Bennett, C. & Salzberg, S. L. Graph-based genome alignment and genotyping with HISAT2 and HISAT-genotype. *Nat. Biotechnol.* **37**, 907–915 (2019).
74. Haas, B. J. et al. Automated eukaryotic gene structure annotation using EVidenceModeler and the Program to Assemble Spliced Alignments. *Genome Biol.* **9**, 1–22 (2008).
75. Gladyshev, E. A., Meselson, M. & Arkhipova, I. R. Massive horizontal gene transfer in bdelloid rotifers. *Science* **320**, 1210–1213 (2008).
76. Wang, W. et al. Regulatory mechanisms underlying the maintenance of homeostasis in *Pyropia haitanensis* under hypersaline stress conditions. *Sci. Total Environ.* **662**, 168–179 (2019).
77. Li, H. & Durbin, R. Fast and accurate short read alignment with Burrows–Wheeler transform. *Bioinformatics* **25**, 1754–1760 (2009).
78. Wang, K., Li, M. & Hakonarson, H. ANNOVAR: functional annotation of genetic variants from high-throughput sequencing data. *Nucleic Acids Res.* **38**, e164–e164 (2010).
79. Takagi, H. et al. QTL-seq: rapid mapping of quantitative trait loci in rice by whole genome resequencing of DNA from two bulked populations. *Plant J.* **74**, 174–183 (2013).
80. Chen, S. et al. Release of dissolved and particulate organic matter by marine macroalgae and its biogeochemical implications. *Algal Res.* **52**, 102096 (2020).
81. Berg, J. et al. Selection for Cu-tolerant bacterial communities with altered composition, but unaltered richness, via long-term Cu exposure. *Appl. Environ. Microbiol.* **78**, 7438–7446 (2012).
82. Wickham, H. ggplot2. *Wiley Interdiscip. Rev.* **3**, 180–185 (2011).
83. Shade, A. & Stopnisek, N. Abundance-occupancy distributions to prioritize plant core microbiome membership. *Curr. Opin. Microbiol.* **49**, 50–58 (2019).
84. Zhang, J., Qiong, D. & Jihong, H. Package 'spaa'. *R package version* **1** (2013).

Acknowledgements

This work was supported by the National Natural Science Foundation of China (grant numbers: U21A20265, 42176117, and 32100514), the Young Elite Scientists Sponsorship Program by the China Association of Science and Technology (grant number: 2021QNRC001), and the China Agriculture Research System of MOF and MARA (grant number: CARS-50).

Author contributions

Wenlei Wang, Qijin Ge, Jian Wen, Guangyi Fan, and Chaotian Xie conceived the study and designed the experiments. Wenlei Wang, Qijin Ge, and Jian Wen conducted the experiments, analyzed the data, and wrote the first draft of the manuscript. Han Zhang, Yanling Guo, Zongtang Li, Changsheng Chen, Dehua Ji and Yan Xu helped collect data and prepare consumables. Lidong Guo, Mengyang Xu, and Chengcheng Shi assisted with the analysis of the genomic data.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at https://doi.org/10.1038/s42003-024-06663-y.

Correspondence and requests for materials should be addressed to Guangyi Fan or Chaotian Xie.

Peer review information *Communications Biology* thanks Dongmei Wang, Federica Montesanto and the other, anonymous, reviewer for their contribution to the peer review of this work. Primary Handling Editor: Tobias Goris. A peer review file is available.