Han et al. *Zool. Res.* 2022, 43(6): 927−930
https://doi.org/10.24272/j.issn.2095-8137.2022.155

ZR | Zoological Research

Letter to the editor Open Access

Charybdis japonica genome provides insights into desiccation adaptation and sex-determining region

DEAR EDITOR,

Shore swimming crabs, such as *Charybdis japonica*, predominantly inhabit the intertidal zones and exhibit high tolerance to desiccation. Here, we present the first chromosome-level *C. japonica* genome, containing 51 pseudochromosomes, with a revised genome length of 1 431.02 Mb and scaffold N50 size of 29.67 Mb. In addition, 824.02 Mb of repetitive elements, 30 900 protein-coding genes, 474 microRNAs (miRNAs), 15 570 transfer RNAs (tRNAs), 309 ribosomal RNAs (rRNAs), and 157 small nuclear RNAs (snRNAs) were identified in the genome. Whole-genome resequencing data were used to identify sex-related single-nucleotide polymorphisms (SNPs) and insertion-deletion (indel) mutations. The 0–10 120 000 bp region of chromosome 37 was identified as the sex-determining region of *C. japonica*. Comparative genomic analysis identified 832 *C. japonica*-specific gene families. Phylogenetic analysis indicated that *C. japonica* was closely related to *Portunus trituberculatus*, which also belongs to the family Portunidae. Compared with other species, metabolic rate, oxygen supply, oxidative stress, and various transporter-related genes were expanded or underwent positive selection in *C. japonica*, which may contribute to its ability to overcome multiple stresses in dry environments. Decoding this genome provides valuable information for revealing the mechanisms underlying desiccation adaptation and sex determination in *C. japonica* and enriches current genetic information to explore the evolutionary history and environmental adaptation strategies of other Portunidae crabs.

In this study, we characterized a high-quality chromosome-anchored reference genome of *Charybdis japonica* (H. Milne Edwards, 1861)(Figure 1A) using Illumina short reads, PacBio long reads, and Hi-C reads (see Supplementary Materials and Methods). We correlated the reported genomic data with the biology, evolutionary history, and desiccation tolerance mechanisms of *C. japonica*. In addition, we predicted the genomic regions associated with sex determination based on whole-genome resequencing data (see Supplementary Materials and Methods). These data provide insights into the genetic mechanisms associated with sex differentiation and desiccation tolerance in *C. japonica* and other crustaceans.

The chromosome-level *C. japonica* genome was 1 431.02 Mb long, with a scaffold N50 of 29.67 Mb (Figure 1B). The genome size was confirmed by survey sequencing (1 398 Mb) after Illumina short read correction and PacBio long read correction. In total, 51 pseudochromosomes were assembled in the genome based on Hi-C reads. Some chromosomes were extremely small and rich in repetitive elements, and most chromosomes in the mitosis metaphase were spot-shaped and difficult to distinguish; therefore, identifying all *C. japonica* chromosomes through Hi-C interactions was difficult. Many sequencing reads (Illumina short reads and PacBio long reads) were successfully mapped to the assembled genome sequence (Supplementary Table S1), indicating high read consistency. In addition, 86.40% complete BUSCOs (Benchmarking Universal Single-Copy Orthologs) were found in the *C. japonica* genome, suggesting that the assembled genome was relatively complete (Supplementary Table S2). A total of 824.02 Mb of repetitive elements were identified, accounting for 57.65% of the genome (Supplementary Table S3). The high proportion of repetitive elements may be a key reason for the large-sized genome of *C. japonica*. Previous analysis of whole-genome repetitive sequences and genome size in 44 plants and 68 vertebrates confirmed that the proportion of repeats is positively correlated with genome size (Gao et al., 2018). We also predicted 824.02 Mb of repetitive elements, 30 900 protein-coding genes (Supplementary Table S4), 474 miRNAs, 15 570 tRNAs, 309 rRNAs, and 157 snRNAs (Supplementary Table S5). In conclusion, the assembled genome provides a useful resource for exploring the biological processes of *C. japonica* at the genomic level.

Received: 17 June 2022; Accepted: 19 September 2022; Online: 20 September 2022

Foundation items: This work was supported by the Zhejiang Provincial Natural Science Foundation of China (LR21D060003) and National Natural Science Foundation of China (32070513)

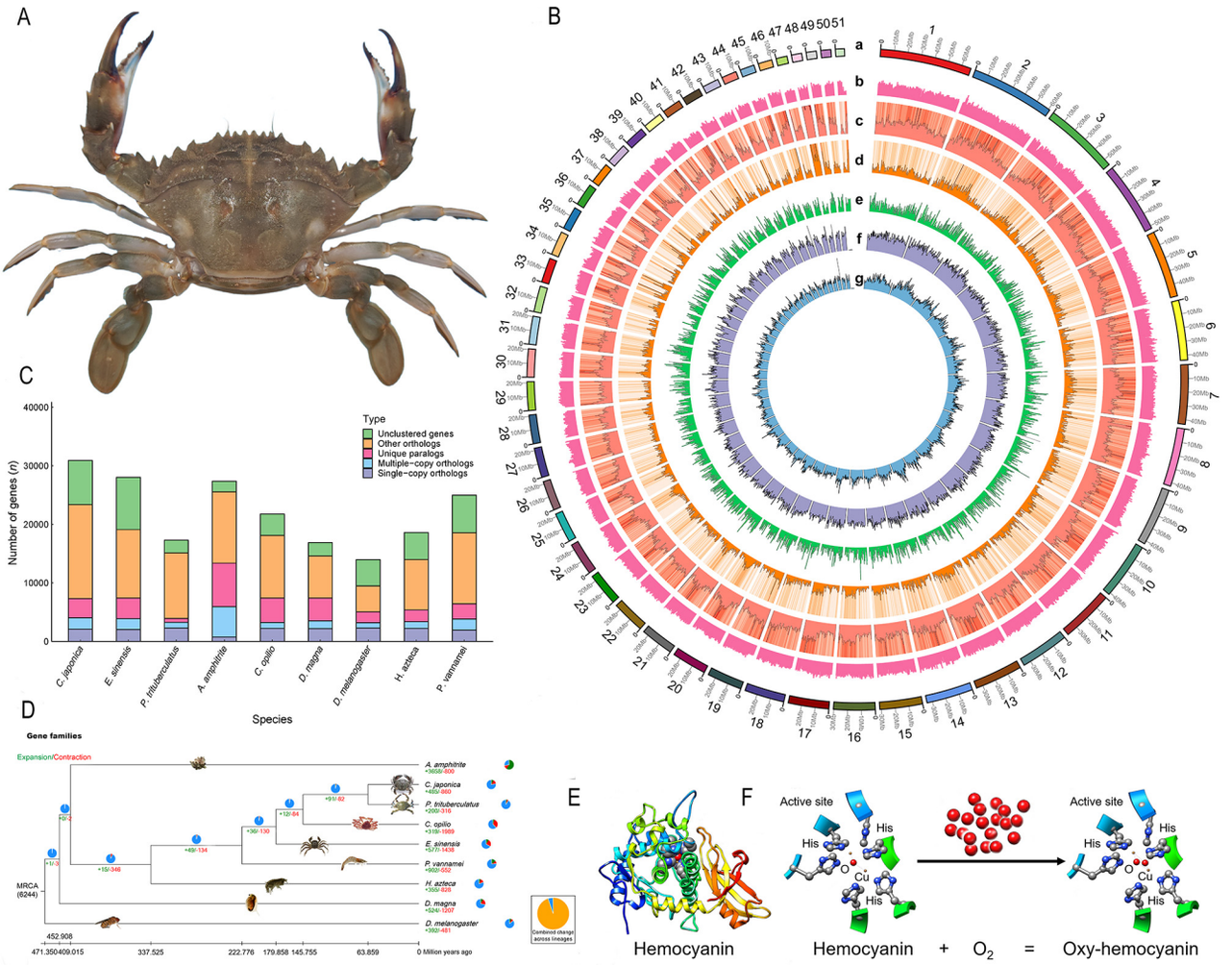

Figure 1 Genome analyses of *Charybdis japonica*

A: *Charybdis japonica*. B: Genome-wide Hi-C heatmap of *C. japonica*. From outer to inner circle: chromosome (a), gene distribution (b), GC content of genome (c), Illumina short read sequencing depth (d), PacBio long read depth (e), DNA transposable element (f), and long terminal repeat (g). C: Number of homologous genes among genomes of related species. D: Phylogenetic analysis based on 340 single-copy orthologous genes. E: Hemocyanin. F: Schematic of hemocyanin binding to oxygen.

Based on whole-genome resequencing of 20 *C. japonica* individuals (both male and female), we identified nine SNP variants (number: 1 811 614; Supplementary Table S6) and 12 indel variants (number: 1 671 995; Supplementary Table S7). We also predicted the sex-determining region (approximately 10 Mb) on chromosome 37 (Supplementary Figure S1A, B). Notably, several genes in this region may be involved in sex differentiation, including the *F-box* and WD repeat domain (*WDR*) genes. Previous research has indicated that the F-box protein encoded by the recombinant *FBXW* gene (containing *F-box* and *WDR* genes) is specific for substrate recognition in ubiquitin-mediated proteolysis (Roberts et al., 2020). Furthermore, the protein can bind with S-phase kinase-associated protein 1 (SKP1) and cullin to form an evolutionarily conserved Skp-Cullin-F-box (SCF) ubiquitin ligase complex (Supplementary Figure S2), which can selectively bind to degraded proteins and participate in the ubiquitin-proteasome pathway (UPP) (Roberts et al., 2020). In plants, the UPP system causes functional loss of male germ cells by regulating the conversion of functional proteins (Kipreos & Pagano, 2000). The mechanism of sex-determination in crustaceans is still in the early stage of evolution, and the strength of genetic factors is lower than that of environmental factors. Therefore, the recombinant *FBXW*-regulated UPP may potentially limit testis development, reduce sperm production, and even promote vitellus formation in male *C. japonica*. However, this mechanism is speculative and requires further verification.

Our assembled genome can be a useful resource for better understanding the evolutionary history of *C. japonica*. Differentiation of conserved single-copy orthologous genes can lead to species divergence. Such genes can therefore be used to explore evolutionary history. In the present study, 340 single-copy orthologous genes were obtained (Figure 1C) and used to determine the phylogenetics of *C. japonica* (Figure 1D). The constructed phylogeny was congruent with

the prevailing morphological and molecular view that crabs from the family Portunidae (e.g., *P. trituberculatus* and *C. japonica*) cluster together on a single branch. We also identified 832 unique gene families in the genome (Supplementary Table S8), which likely contribute to *C. japonica*-specific adaptation.

Comparative genomics can provide insights into the unique adaptive plasticity of marine species (Lou et al., 2022). Previous studies have shown that crustaceans may be under constant stress from both dehydration and hypoxia in arid environments. Specifically, in crustaceans, desiccation can disrupt internal osmotic pressure and metabolic capacity and reduce the oxygen-binding ability of hemoglobin, resulting in hypoxic stress (Allen et al., 2012). Prolonged disruption of the respiratory mechanism may further impair immune function, eventually leading to death. Furthermore, antioxidant enzyme activity directly determines the desiccation tolerance of crustaceans. Here, based on annotations, the desiccation-adaptive mechanisms of *C. japonica* were supported by the marked expansion of gene families or positive selection of genes related to metabolic rate, oxygen supply, oxidative stress, and transporters.

Carbohydrates provide a carbon skeleton and energy for growth, metabolic activity, and stress resistance of organisms. Glycometabolism connects the metabolism of proteins, lipids, nucleic acids, and secondary substances. Based on comparative genomic analysis, abundant glycometabolism-related genes were expanded or positively selected in *C. japonica*, likely contributing to desiccation tolerance. The reasons why carbohydrates may improve desiccation tolerance in *C. japonica* include the following: (i) carbohydrates can accumulate and regulate osmotic pressure and enhance the water retention of cells; (ii) carbohydrates can protect biological substances, such as biofilms and biological macromolecules, to maintain cytoskeletal integrity under drought environments; (iii) carbohydrates can produce energy for resistance to desiccation; and (iv) carbohydrates can be used as a signaling molecule in certain desiccation-resistant regulatory mechanisms.

Although hemocyanin (Figure 1E) has a low affinity for oxygen, it remains an essential oxidizing medium for crustacean tissues in the absence of oxygen (Figure 1F). Desiccation can inhibit oxygen supply and induce respiratory acidosis in crustaceans. Respiratory acidosis further reduces the ability of oxygen and hemoglobin to bind, thereby exacerbating oxygen deficiency and tissue hypoxia (Allen et al., 2012). However, desiccation-tolerant crustaceans can regulate ion (including Ca^{2+} and Mg^{2+}) concentrations to improve the oxygen-binding ability of hemocyanin and promote oxygen transfer between tissues. The expansion and positive selection of *IP3* and *PRPS1* indicated that *C. japonica* may maintain Ca^{2+} and Mg^{2+} homeostasis to ensure the oxygen-binding capacity of hemocyanin to resist hypoxia in drought environments. Hypoxia-inducible factors (HIFs) also play important roles in hypoxia response by binding with hypoxia response elements near the downstream gene promoter for transcriptional activation (Martens et al., 2007). The expansion and positive selection of *STK* in *C. japonica* may enhance HIF phosphorylation and increase the molecular weight of HIFs. The *GSR* gene, which prevents the oxidative decomposition of hemocyanin, was also positively selected and expanded in *C. japonica*. Glutathione reductase encoded by *GSR* can effectively maintain reduced glutathione in cells, which plays an important role in preventing the oxidative decomposition of hemocyanin (Kanzok et al., 2001). In conclusion, the maintenance of hemocyanin levels and oxygen-carrying capacity should help *C. japonica* cope with hypoxic conditions in drought environments.

We also detected the expansion or positive selection of several *HSP* genes (including *HSP70* and *HSP90*), which may encode prototypical chaperonin to prevent the accumulation of unfolded proteins and/or assist in the refolding or degradation of denatured proteins (Beck et al., 2009). The zinc finger (*ZNF*) gene family was also enriched in the *C. japonica* genome. Zinc finger proteins are the most abundant proteins in eukaryotes and exhibit a wide range of structures and functions (Laity et al., 2001). Interestingly, positive selection/expansion of RING-type zinc finger proteins occurred in *C. japonica*. As a zinc finger protein widely found in plants and animals, the RING-type zinc finger protein confers strong drought tolerance in *Arabidopsis thaliana* (Ryu et al., 2010), because it can regulate the UPP system to degrade misfolded proteins and thus improve organismal defense against environmental stress (Takai et al., 2002).

DATA AVAILABILITY

Raw sequencing data (including PacBio, Hi-C, Survey, and RNA-seq reads) for *C. japonica* genome were deposited at the National Center for Biotechnology Information (NCBI) sequence read archive (SRA) (SRR16072374, SRR16072376, SRR16072375, and SRR16078881) under BioProjectID PRJNA766329. Raw whole-genome resequencing data of 20 *C. japonica* were also deposited in the SRA (SRR1689 4624–SRR16894643) under BioProjectID PRJNA779101. Genomic data were collected from the Genome Sequence Archive (GSA) database (https://ngdc.cncb.ac.cn/gwh/Assembly/25212/show) (Accession No. GWHBISL00000000). Genomic data were also archived in the Science Data Bank database (DOI: 10.57760/sciencedb.j00139.00026).

SUPPLEMENTARY DATA

Supplementary data to this article can be found online.

COMPETING INTERESTS

The authors declare that they have no competing interests.

AUTHORS' CONTRIBUTIONS

Z.Q.H. and F.R.L. conceived and managed the project. Z.Q.H. and T.X.G. collected the sequencing samples. Q.L. extracted the DNA/RNA and performed genome sequencing. F.R.L. analyzed the data. Z.Q.H. and F.R.L. wrote the manuscript. All authors read and approved the final version of the manuscript.

Zhi-Qiang Han[1], Tian-Xiang Gao[1], Fang-Rui Lou[2,*], Qi Liu[3]

[1] *Fishery College*, *Zhejiang Ocean University*, *Zhoushan*, *Zhejiang* 316022, *China*

[2] *School of Ocean*, *Yantai University*, *Yantai*, *Shandong* 264001, *China*
[3] *Wuhan Onemore-tech Co.*, *Ltd.*, *Wuhan*, *Hubei* 430076, *China*
*Corresponding author, E-mail: lfr199202@163.com

REFERENCES

Allen BJ, Rodgers B, Tuan YH, Levinton JS. 2012. Size-dependent temperature and desiccation constraints on performance capacity: implications for sexual selection in a fiddler crab. *Journal of Experimental Marine Biology and Ecology*, **438**: 93−99.

Beck M, Malmström JA, Lange V, Schmidt A, Deutsch EW, Aebersold R. 2009. Visual proteomics of the human pathogen *Leptospira interrogans*. *Nature Methods*, **6**(11): 817–823.

Gao SH, Yu HY, Wu SY, Wang S, Geng JN, Luo YF, et al. 2018. Advances of sequencing and assembling technologies for complex genomes. *Hereditas*, **40**(11): 944−963. (in Chinese)

Kanzok SM, Fechner A, Bauer H, Ulschmid JK, Müller HM, Botella-Munoz J, et al. 2001. Substitution of the thioredoxin system for glutathione reductase in *Drosophila melanogaster*. *Science*, **291**(5504): 643–646.

Laity JH, Lee BM, Wright PE. 2001. Zinc finger proteins: new insights into structural and functional diversity. *Current Opinion in Structural Biology*, **11**(1): 39−46.

Lou FR, Wang YJ, Han ZQ, Shui BN. 2022. Comparative transcriptome reveals the molecular regulation mechanism of *Charybdis japonica* to high- and low-temperature stresses. *Frontiers in Marine Science*, **9**: 849485.

Kipreos ET, Pagano M. 2000. The F-box protein family. *Genome Biology*, **1**: 30021−30027.

Martens LK, Kirschner KM, Warnecke C, Scholz H. 2007. Hypoxia-inducible factor-1 (HIF-1) is a transcriptional activator of the *TrkB* neurotrophin receptor gene. *Journal of Biological Chemistry*, **282**(19): 14379−14388.

Roberts JZ, Holohan C, Sessler T, Fox J, Crawford N, Riley JS, et al. 2020. The SCF^{Skp2} ubiquitin ligase complex modulates TRAIL-R2-induced apoptosis by regulating FLIP(L). *Cell Death & Differentiation*, **27**(9): 2726−2741.

Ryu MY, Cho SK, Kim WT. 2010. The Arabidopsis C3H2C3-type RING E3 ubiquitin ligase AtAIRP1 is a positive regulator of an abscisic acid-dependent response to drought stress. *Plant Physiology*, **154**(4): 1983−1997.

Takai R, Matsuda N, Nakano A, Hasegawa K, Akimoto C, Shibuya N, et al. 2002. EL5, a rice N-acetylchitooligosaccharide elicitor-responsive RING-H2 finger protein, is a ubiquitin ligase which functions *in vitro* in co-operation with an elicitor-responsive ubiquitin-conjugating enzyme, *Os*UBC5b. *The Plant Journal*, **30**(4): 447−455.