BMC Genomics

RESEARCH Open Access

Whole-genome survey information revealed the genome characteristics, mitochondrial genome diversity and evolutionary patterns of Stomatopoda species

Xiaowen Duan[1†], Yuan Li[2†], Yue Ding[1], Xinxian Huang[2], Jiaguang Xiao[2], Yifei Wang[1] and Fangrui Lou[1*]

Abstract

Background Stomatopoda possess extremely unique biological characteristics and evolutionary process; however, their phylogenetic relationships remain controversial. To avoid gene tree conflicts caused by a small number of genetic markers in previous phylogenetic studies of Stomatopoda, this study provides the first whole-genome survey sequencing data for three Stomatopoda species (*Odontodactylus japonicus*, *Odontodactylus scyllarus*, and *Lysiosquillina maculata*).

Results Genome size assessment indicates that the genomes of the three species are extraordinarily large (*O. japonicus*: 8,942 Mb; *O. scyllarus*: 11,068 Mb; *L. maculata*: 12,743 Mb), and characterized by high heterozygosity (1.25%–1.51%) and extensive repeats (83.26%–86.22%). We further assembled and annotated the mitochondrial genomes of three Stomatopoda species, and results showed that the total length of the *O. japonicus* mitochondrial genome was 16,237 bp and contained 37 genes, while *O. scyllarus* and *L. maculata* have an *ATP8* deletion and their lengths were 16,110 bp and 16,204 bp, respectively. In addition, mitochondrial coding sequences exhibited strong A/T bias and species-specific codon preferences, and structural modeling of tRNAs confirmed conserved cloverleaf conformations with lineage-specific deviations. Furthermore, we have found that the phylogenetic relationships within Stomatopoda remain inconsistent when inferred from different genetic markers. The further phylogenetic tree that incorporates multiple genetic markers (including mitochondrial genome, coding DNA sequences and amino acids) suggests that *L. maculata* occupies a more basal position. Additionally, the divergence time between Stomatopoda and other three orders (Decapoda, Amphipoda, and Isopoda) was approximately 385.00 Mya.

Conclusions Overall, our whole-genome survey sequencing and integrative phylogenomic analyses provide new insights into resolving the controversies on the phylogenetic relationships within Stomatopoda. These results can improve our understanding of the evolutionary process of Stomatopoda genome and provide valuable resources for future phylogenetic studies of Stomatopoda.

[†]Xiaowen Duan and Yuan Li contributed equally to this work.

*Correspondence:
Fangrui Lou
lfr199202@163.com

Full list of author information is available at the end of the article

Keywords Stomatopoda, Whole-genome survey sequencing, Mitochondrial genome, Phylogenomics

Background

The Stomatopoda (Crustacea: Malacostraca: Hoplocarida) are a highly diverse group found in habitats ranging across the shallow sandy seabed and coral reef areas to the deep-sea continental slopes [1]. Since Brooks (1886) first recorded eight species of the Stomatopoda in the "Challenger" global expedition report, new Stomatopoda species have been continuously discovered [2]. At present, more than 450 Stomatopoda species have been recognized, classified into 7 superfamilies (Erythrosquilloidea, Bathysquilloidea, Parasquilloidea, Eurysquilloidea, Gonodactyloidea, Lysiosquilloidea, and Squilloidea) and 17 families [3]. As the sole animal group of the subclass Hoplocarida, Stomatopoda exhibit a distinctive body segmentation pattern, specialized raptorial appendages, and other biological characteristics that differ markedly from those of other Malacostracan groups [4]. Therefore, the appearance of the Stomatopoda represents an important milestone in the phyletic evolutionary processes of the crustacean and consequently has attracted the attention of evolutionary scientists.

Although the taxonomy of Stomatopoda has been repeatedly revised, controversies remain over the delimitation of their taxonomic status and the identification of the most basal superfamilies [1]. Previous studies based on the telson have revealed that the Bathysquilloidea possesses transitional telson characteristics that lie between ancient and the extant Stomatopoda may represent the most primitive Stomatopoda [5]. In contrast, Ahyong (1997) suggested that the Squilloidea is more primitive than the Bathysquilloidea, and hypothesized that the Stomatopoda with hammer-shaped raptorial appendage originated from the decapods with spear-shaped raptorial appendage [6]. However, the cladistic classification method provided evidence against Ahyongs' (1997) hypothesis, with results indicating the Gonodactyloidea is located at the bottom of the phylogenetic tree of the Stomatopoda [7, 8]. Gonodactyloidea has also been hypothesized to have ancestral forms with compound eyes [9]. The involvement of molecular markers (such as mitochondrial genes and nuclear genes, etc.) has shifted the systematic taxonomy of Stomatopoda from a macroscopic perspective to a microscopic perspective, but has also generated disputes regarding species classification and phylogenetic relationships. Specifically, the Parasquilloidea is more primitive than the Gonodactyloidea, and the Stomatopoda contains numerous cryptic species [9–12]. However, it is worth noting that sample collection constraints and the use of different genetic markers lead to conflicts in the inferred phylogenetic trees. In summary, there is still no consensus on the phylogenetic position and relationships within the Stomatopoda, indicating that more relevant evidence is urgently needed.

Currently, the emergence and widespread application of high-throughput sequencing technology has provided opportunities for the acquisition of the complete genetic information [13, 14], which is conducive to screening core conserved genes or other high-quality genetic markers, ultimately resolving phylogenetic disputes with high resolution. Although many whole-genome resources for crustaceans have been constructed at present [15–17], the presence of a large number of tandemly repeated sequences makes the crustacean genomes extremely complex and ultimately poses a significant challenge to the genome assembly [15]. In fact, the Stomatopoda genome appears to be more complex, containing more repetitive sequences than many crustaceans, resulting in higher sequencing costs; thus only the *Oratosquilla oratoria* whole-genome has previously been successfully assembled [18]. In the context of extremely scarce reference genome information, conducting an examination of the whole-genome survey can predict the genome characteristic and obtain relatively abundant genetic information [19], providing an opportunity to resolve the controversial genetic relationships among Stomatopoda.

The manuscript performed the first genome survey sequencing of three Stomatopoda species (including *O. scyllarus*, *O. japonicus*, and *L. maculate*) collected from the coastal waters of Fujian, China. These species were selected because they represent two phylogenetically distinct superfamilies (Gonodactyloidea and Lysiosquilloidea), exhibit pronounced differences in morphology and ecological niches and encompass contrasting functional types of raptorial appendages—hammer-shaped for smashing versus spear-shaped for piercing. We analyzed the genomic characteristics of the three Stomatopoda species and re-constructed a more comprehensive and higher-resolution phylogenetic relationship. The results of this study can provide new evidence to address the controversial phylogenetic relationships of the Stomatopoda, and enrich the future research on the Stomatopoda.

Materials and methods

Sampling, library construction, and sequencing

Odontodactylus scyllarus, Odontodactylus japonicus, and Lysiosquillina maculate were collected from the coastal waters of Fujian, China (24°48′N, 119°55′E) on November 1, 2024. There were five *O. scyllarus*, five *O. japonicus*, and five *L. maculata*, all of which were transported to the Fisheries Resources and Ecological Environment Laboratory of Yantai University. The Animal Ethics Committee of Yantai University approved the subsequent animal

experiment procedures. After conducting a preliminary morphological identification of all the animals, we then amplified the *COXI* gene (forward primer: 5′-GGTCAACAAATCATAAAGATATTGG-3′; reverse primer: 5′-TAAACTTCAGGGTGACCAAAAAATCA-3′) to ensure the accuracy of experimental animals' identification. Furthermore, one male individual was selected from each Stomatopoda species for cryogenic anesthesia. Subsequently, the fresh muscle was dissected, snap-frozen in liquid nitrogen, and then transferred to −80 °C for subsequent DNA extraction.

High-quality genomic DNA was extracted from muscle tissue using the Phenol/Chloroform extraction method [20]. The DNA concentration was determined using NanoDrop ND-1000 (Thermo Fisher Scientific Inc, USA) spectrophotometer, and the DNA integrity was confirmed by 1% agarose gel electrophoresis. The qualified DNA was ultrasonically fragmented into 300–350 bp fragments using a Covaris S220 ultrasonicator (Covaris, Woburn, MA, USA). These fragments underwent end repair, poly-A addition, sequencing linker connection, purification, and PCR amplification to construct the DNA libraries, using the NEBNext Ultra II DNA Library Prep Kit (New England Biolabs, Ipswich, MA, USA). Finally, all the cDNA libraries were subjected to genome survey paired-end (2 × 150 bp) sequencing on the Illumina NovaSeq 6000 platform (Illumina, San Diego, CA, USA).

Data quality controlling

All raw sequencing data generated by the Illumina NovaSeq 6000 platform were stored in the Short Read Archive (SRA) database of NCBI under BioProject PRJNA1205873. Low-quality data can interfere with downstream data analyses, therefore the fastp software (version 0.23.2, [21]) was applied to perform low-quality data filtering and deduplication on the raw data to generate clean data for subsequent analyses. Our pre-processing methodology included: (1) removal of short reads and low-quality reads containing numerous of ambiguous bases; (2) correction of the low-quality bases located at the 5' and 3' ends of each read using the sliding window algorithm [22]; (3) removal of the adapter sequences; (4) correction of the bases in the overlapping regions of the paired-end reads based on the base quality scores; (5) trimmed the sequenator-specific poly (G) or poly (X) located at the 3' end; and (6) eliminating the repetitive sequences caused by PCR amplification. Meanwhile, the percentages of bases with quality values higher than 20 (Q20) and 30 (Q30), as well as the percentage of GC content, were calculated using FastQC software (https://www.bioinformatics.babraham.ac.uk/projects/fastqc/) to assess the quality of the sequencing data.

K-mer analysis

We set the value of K-mer to 17, and used the GCE (version 1.0.0; [23]) software to evaluate the number and depth of K-mer. Furthermore, we employed the GCE (version1.0.0; [23]) software to predict the size, heterozygosity rate and repeat sequence ratio of the three Stomatopoda genomes based on K-mer characteristic values.

Mitochondrial genome assembly

We randomly extracted 100 Gb of sequencing data from the aforementioned raw data of the three Stomatopoda species, then, we used the *COX1* sequence as the seed sequence and employed the NOVOplasty (version 4.2; [24]) software (the K-mer setting is 33) to assembled the mitochondrial genome. Simultaneously, the MitoZ (version 3.3; [25]) software was employed for annotating the generated mitochondrial genome, followed by manual curation to refine gene boundaries and correction of potential annotation errors. Finally, we visualize these mitochondrial genome results as circular diagrams using Circos (version 0.69, [26]).

Calculation of relative synonymous codon usage (RSCU) and 3D prediction of tRNA secondary structures

The RSCU of mitochondrial protein-coding genes was calculated using the European Molecular Biology Open Software Suite (EMBOSS; [27]), and the results were visualized as stacked bar charts using the OriginPro 2022 [28] software. Three-dimensional models of the mitochondrial tRNA secondary structures were predicted with trRosettaRNA [29] software, and the generated PDB files were subsequently rendered and examined using PyMOLWIN (version2.5.2, [30]) software.

Phylogenetic relationship construction

To construct the phylogenetic relationship of three Stomatopoda species, the published whole genome assembly sequences of *Armadillidium vulgare* (accession number: SAUD01000000; [31]), *A. nasatum* (accession number: SEYY00000000; [32]), *Cherax destructor* (accession number: GCA_009830355.1; [33]), *Eriocheir sinensis* (accession number: GCA_024679095.1; [34]), *Exopalaemon carinicauda* (accession number: GCA_036898095.2; [35]), *Fenneropenaeus indicus* (accession number: GCA_018983055.3; [36]), *Gammarus roeselii* (accession number: GCA_016164225.1; [37]), *Homarus americanus* (accession number: GCA_018991925.1; [38]), *Hyalella Azteca* (accession number: GCA_000764305.4; [39]), *O. oratoria* (accession number: GCA_046742065.1; [18]), *Panulirus ornatus* (accession number: GCA_036320965.1; [40]), *Paralithodes camtschaticus* (accession number: GCA_018397895.1; [41]), *Procambarus clarkii* (accession number: GCA_040958095.1; [42]), *Trachelipus rathkii* (accession number:

GCA_015478945.1; [43]), and *Trinorchestia longiramus* (accession number: GCA_006783055.1; [44]) with relatively close genetic relationships were first downloaded from European Nucleotide Archive (ENA) and National Center for Biotechnology Information (NCBI) databases. Here, "relatively close" relationships were defined operationally at the class level—Malacostraca, which also includes Stomatopoda—by selecting representative taxa from Decapoda, Isopoda, and Amphipoda that are commonly used as near outgroups in crustacean phylogenies. Genome assemblies were then filtered for quality, retaining only those with BUSCO completeness ≥ 85% (Metazoa/Arthropoda datasets) and excluding assemblies showing evidence of contamination or severe fragmentation. Subsequently, we assembled the mitochondrial genome sequences of above-mentioned closely-related species by following the method described in section (Mitochondrial genome assembly) and then extracted the coding sequence (CDS) and protein sequence of the single-copy gene encoding the protein (PEP).

Subsequently, MAFFT (version 7; [45]) software was employed to conduct the multi-sequence alignment of all mitochondrial genomes, and the MUSCLE (version 5; [46]) software was used to align the CDSs (*ATP6*, *COX1*, *COX2*, *COX3*, *CYTB*, *ND3*, *ND4*, *ND4L*, *ND5*, and *ND6*) and PEPs. Low-quality sites were trimmed using trimAl [47]. The cleaned, gene-wise CDS alignments were concatenated into a supermatrix, and an initial partitioning scheme based on gene × codon position (30 subsets). ModelFinder as implemented in IQ-TREE (version 2.2.2.7, [48]) (option -m MFP + MERGE, BIC criteria) was used to select the best-fit substitution models and merge similar subsets, yielding the final optimal partitioning scheme (Supplementary Table 1). For all PCGs, first and second codon positions (pos1/pos2) were best described by GTR + F + I + G4, whereas third positions (pos3) favored GTR + F + G4. Accordingly, Maximum Likelihood analyses were conducted in the RAxML (version 8.2.12; [49]) under GTR + Γ with partition-specific rates (and base-frequency parameters), using 1,000 bootstrap replicates to assess node support. Relative to the unpartitioned analysis, the partitioned model substantially improved fit (more favorable ΔBIC and ΔlnL) and produced equal or higher bootstrap support across the major clades. Meanwhile, all single-copy genes were merged into one super gene based on the multiple sequence alignment results of each gene and then constructed a phylogenetic tree. Then, the Astral-III [50] software was used to integrate the multiple phylogenetic trees mentioned above to construct a high-resolution composite phylogenetic tree. Finally, we obtained the estimated molecular clock data of *F. indicus* and *E. carinicauda* (adjusted-time is 365 Mya), *E. carinicauda* and *E. sinensis* (adjusted-time is 332 Mya), *E. carinicauda* and *T. rathkii* (adjusted-time is 370 Mya), *G. roeselii* and *H. azteca* (adjusted-time is 163 Mya), *P. camtschaticus* and *G. roeselii* (adjusted-time is 370 Mya), *A. nasatum* and *G. roeselii* (adjusted-time is 319 Mya), *P. ornatus* and *E. sinensis* (adjusted-time is 330 Mya), *P. camtschaticus* and *E. sinensis* (adjusted-time is 330 Mya), *C. destructor* and *H. americanus* (adjusted-time is 278 Mya), *O. scyllarus* and *E. carinicauda* (adjusted-time is 385 Mya) from the TimeTree database [51] and then estimated the divergence time of Stomatopoda species using the r8s software [52] (Table 1).

Table 1 Gene abbreviations table

Abbreviation	Gene Name
ATP6	ATP Synthase Subunit 6
COX1	Cytochrome c Oxidase I
COX2	Cytochrome c Oxidase II
COX3	Cytochrome c Oxidase III
CYTB	Cytochrome b
ND3	NADH Dehydrogenase 3
ND4	NADH Dehydrogenase 4
ND4L	NADH Dehydrogenase 4L
ND5	NADH Dehydrogenase 5
ND6	NADH Dehydrogenase 6

Results

The *COXI* gene-based experimental animals' identification

To ensure the reliability of the animals used for subsequent genomic survey sequencing, we first amplified and sequenced *COX1* gene of these experimental animals. Then, we conducted a blast search of the obtained sequences in the NCBI database. Results showed that the query cover rate of the *COX1* sequences to the corresponding Stomatopoda sequences in NCBI ranging from 97% ~ 99%, with E values all being 0 and the percentage identity ranging from 98.61% ~ 100%. Such a high degree of sequence similarity ensures the credibility of the animal used for subsequent genome survey sequencing.

Genome sequencing data of three Stomatopoda species

We conducted genome survey sequencing on *O. japonicus*, *O. scyllarus*, and *L. maculata*, and ultimately generated 213.08 Gb, 442.63 Gb, and 381.92 Gb of raw data, respectively, corresponding to 1,420,536,952, 2,950,877,140, and 2,546,149,612 reads. The ranges of Q20, Q30 and GC contents for the three sets of raw data are 94.21% ~ 97.29%, 89.04% ~ 91.88%, and 40.83% ~ 41.49%. After filtering out the low-quality raw data, a total of 201.63 Gb, 374.54 Gb, 314.15 Gb and of clean data were produced in *O. japonicus*, *O. scyllarus*, and *L. maculata*, respectively, corresponding to 1,350,175,634, 2,554,506,510, and 2,154,406,768 reads. The ranges of Q20, Q30 and GC contents for the three

Table 2 Genome survey sequencing information of *Odontodactylus japonicus*, *Odontodactylus scyllarus*, and *Lysiosquillina maculata*

Species		Read Number	BaseCount (Gb)	Q20 (%)	Q30 (%)	GC Content (%)
O. japonicus	Raw data	1,420,536,952	213.08	97.29	91.62	40.83
	Clean data	1,350,175,634	201.63	97.29	91.62	40.83
O. scyllarus	Raw data	2,950,877,140	442.63	95.98	91.88	41.49
	Clean data	2,554,506,510	374.54	96.70	92.84	41.26
L. maculata	Raw data	2,546,149,612	381.92	94.21	89.04	41.10
	Clean data	2,154,406,768	314.15	96.11	91.62	40.93

Table 3 K-mer analysis results of *O. japonicus*, *O. scyllarus*, and *L. maculata*. The table shows the total K-mer number, K-mer depth, estimated genome size, revised genome size, heterozygosity ratio, and repeat content for each species

Sample	K-mer number	K-mer depth (×)	Genome size (Mb)	Revised genome size (Mb)	Heterozygous ratio (%)	Repeat (%)
O. japonicus	179,932,150,262	20	8,997	8,942	1.25	83.26
O. scyllarus	333,531,047,067	30	11,117	11,068	1.42	84.79
L. maculata	205,264,334,555	16	12,829	12,743	1.51	86.22

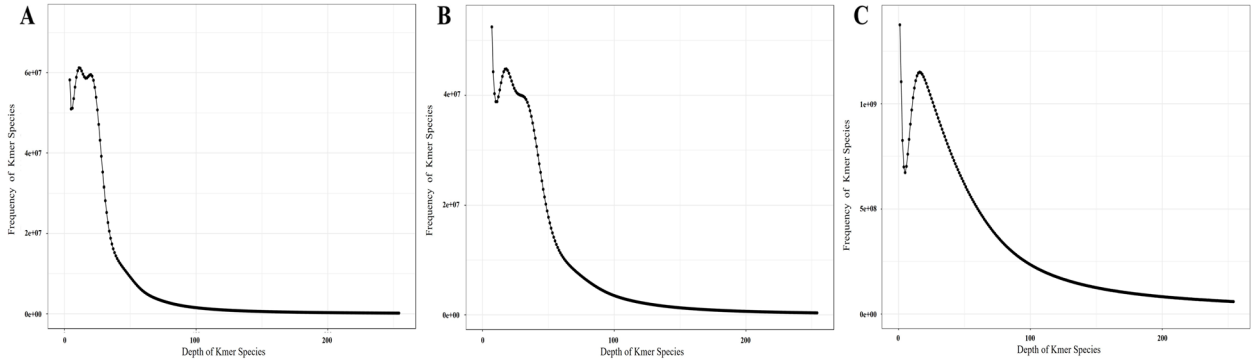

Fig. 1 The correlation between the depth and frequency distribution of K-mer of *Odontodactylus japonicus* (**A**), *Odontodactylus scyllarus* (**B**), and *Lysiosquillina maculata* (**C**). *O. japonicus* and *O. scyllarus* exhibit clear bimodal patterns with heterozygous and homozygous peaks, while *L. maculata* shows a smoother unimodal curve with a weak shoulder

sets of clean data are 96.11% ~ 97.29%, 91.62% ~ 92.84%, and 40.83% ~ 41.26%, respectively (Table 2).

Genomic feature inference based on K-mer distribution

We set the K-mer to 17, and found that 179,932,150,262, 333,531,047,067, and 205,264,334,555 K-mers were distributed in the clean data of *O. japonicus*, *O. scyllarus*, and *L. maculata*, respectively (Table 3). Meanwhile, the K-mer depth of *O. japonicus*, *O. scyllarus*, and *L. maculata* was 20×, 30×, and 16×, respectively. The *O. japonicus*, *O. scyllarus*, and *L. maculata* genome size assessed by K-mer analysis was 8,997 Mb, 11,117 Mb, and 12,829 Mb. After the negative effect caused by incorrect K-mer was eliminated, the revised *O. japonicus*, *O. scyllarus*, and *L. maculata* genome size was 8,942 Mb, 11,068 Mb, and 12,743 Mb. Meanwhile, the heterozygous ratio and repeat of clean data in three Stomatopoda species ranging from 1.25% ~ 1.51% and 83.26% ~ 86.22%. Additionally, the correlation (Fig. 1) between the depth and frequency distribution of K-mer further indicates that all three species have large, highly repetitive genomes. Meanwhile, the K-mer spectra of *O. japonicus* and *O. scyllarus* show clear bimodal distributions with pronounced heterozygous peaks, whereas *L. maculata* displays a smoother unimodal pattern with a weaker shoulder. However, model-based estimates reveal that *L. maculata* has the highest heterozygosity (1.51%), implying that its less distinct peak shape is more likely due to lower sequencing depth and extremely high repeat content than to reduced heterozygosity.

Mitochondrial genome characteristics of three Stomatopoda species

We selected and assembled the mitochondrial genomes of three Stomatopoda species from the survey data. The full lengths of the mitochondrial genome for *O. japonicus*, *O. scyllarus*, and *L. maculata* were 16,237 bp, 16,110 bp, and 16,204 bp, respectively (Fig. 2). The mitochondrial genome of *O. japonicus* contains 37 genes (including 13 protein-coding genes, 22 tRNA genes, and 2 rRNA genes), while those of *O. scyllarus* and *L. maculata* encompasses 36 genes (including 12 protein-coding genes, 22 tRNA genes, and 2 rRNA genes). Specifically, the mitochondrial genomes of *O. scyllarus* and *L. maculata* lack the *ATP8* gene compared to that of *O. japonicus*. Among the three mitochondrial genomes, 14 genes

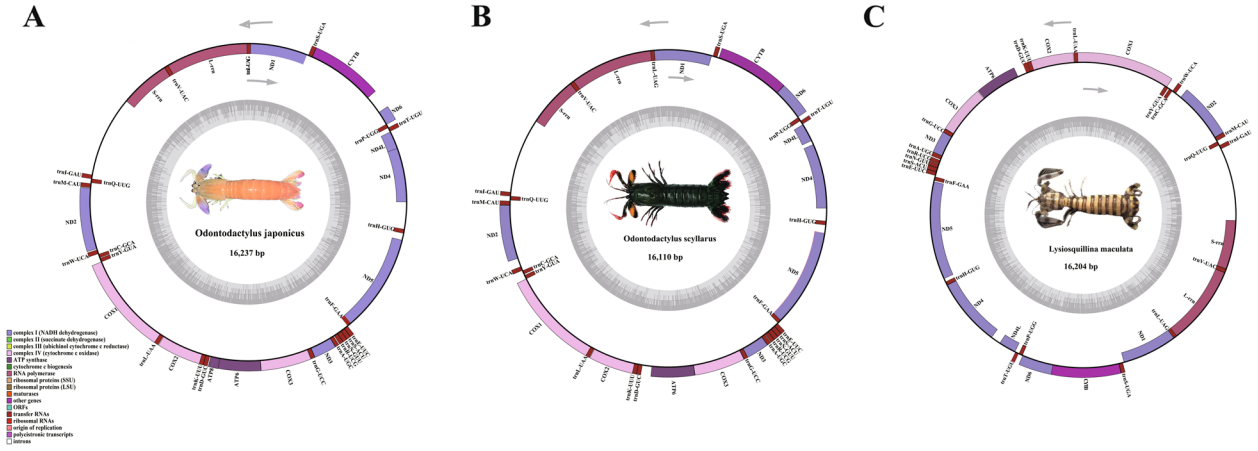

Fig. 2 The circos plots of the *O. japonicus* (**A**), *O. scyllarus* (**B**), and *L. maculata* (**C**) mitochondrial genome

Table 4 Nucleotide composition of the mitochondrial genomes, PCGs, tRNAs and rRNAs of *O. japonicus*, *O. scyllarus*, and *L. maculata*

Sample		A%	T/U%	G%	C%	A + T%	G + C%	AT/U-skew	GC-skew
O. japonicus	mitochondrial genome	31.90	34.90	14.30	18.90	66.80	33.20	−0.04	−0.14
	protein-coding gene	27.22	38.18	16.82	17.80	65.40	34.62	−0.17	−0.03
	tRNA gene	33.90	34.44	17.92	13.72	68.34	31.64	−0.01	0.14
	rRNA gene	37.20	31.95	18.45	12.50	69.15	30.95	0.08	0.19
O. scyllarus	mitochondrial genome	31.80	34.70	14.80	18.70	66.50	33.50	−0.04	−0.12
	Protein-coding gene	27.23	37.80	17.26	17.73	65.03	34.98	−0.16	−0.01
	tRNA gene	34.62	33.44	17.70	14.23	68.06	31.93	0.02	0.11
	rRNA gene	38.05	31.15	17.10	13.75	69.20	30.85	0.10	0.11
L. maculata	mitochondrial genome	32.60	20.20	32.90	14.20	65.50	34.40	0.00	−0.17
	Protein-coding gene	36.71	18.65	26.97	17.69	63.68	36.34	−0.15	−0.03
	tRNA gene	33.34	13.77	35.15	17.70	68.49	31.47	0.03	0.13
	rRNA gene	32.85	11.65	36.75	18.80	69.60	30.45	0.06	0.24

(including *s-rRNA*, *l-rRNA*, $tRNA^{V(UAC)}$, $tRNA^{L(UAG)}$, $tRNA^{P(UGG)}$, $tRNA^{H(GUG)}$, $tRNA^{F(GAA)}$, $tRNA^{Y(GUA)}$, $tRNA^{C(GCA)}$, $tRNA^{Q(UUG)}$, *ND1*, *ND4L*, *ND4*, and *ND5*) are located in the L-strand, while the remaining genes are located in the H-strand.

The AT content of the mitochondrial genomes, protein-coding genes, tRNA genes and rRNA genes of three Stomatopoda species ranged from 65.50 ~ 66.80%, 63.68 ~ 65.40%, 68.06 ~ 68.49%, and 69.15 ~ 69.60%, respectively (Table 4). Meanwhile, the mitochondrial genomes and protein-coding genes of three Stomatopoda species principally exhibit negative AT- and GC- skew, while the tRNA genes and rRNA genes principally show positive AT/U- and GC- skew. Additionally, it is worth noting that the absolute values of GC-skew for mitochondrial genomes, tRNA genes, and rRNA genes are relatively large, while the absolute value of AT-skew for protein-coding genes is larger.

We further analyzed the frequency of codon usage and found that all three mitochondrial genomes contained 64 types of codons (Supplementary Table 2; Fig. 3). Among them, 61 types of codons encoded 24 types of amino acids, and the remaining 3 types of codons (UAA, UAG, and UGA) were termination codons. The total usage count of codons in the *O. japonicus* mitochondrial genome is 3,538, and the most frequently used codon is UUA (count = 206, RSCU = 1.73), while the least frequently used codon is CGA (count = 8, RSCU = 0.57). The total usage count of codons in the *O. scyllarus* mitochondrial genome is 3,497, and the most frequently used codon is UUU (count = 191, RSCU = 1.37), while the least frequently used codon is GCG (count = 6, RSCU = 0.26). The total usage count of codons in the *L. maculata* mitochondrial genome is 3,757, and the most frequently used codon is UUA (count = 216, RSCU = 2.80), while the least frequently used codon is UAG (count = 8, RSCU = 0.10). Additionally, the number of codon types with relatively high usage frequencies (RSCU > 1) in the mitochondrial genomes of *O. japonicus*, *O. scyllarus*, and *L. maculata* were 30, 28, and 29, respectively.

The secondary structure modeling and three-dimensional (3D) structural prediction (Fig. 4) of 22 tRNAs from three Stomatopoda species indicated that all tRNAs could fold into the typical cloverleaf secondary structure and possess the typical L-shaped tertiary conformation. Although the overall structure of all tRNAs is conserved,

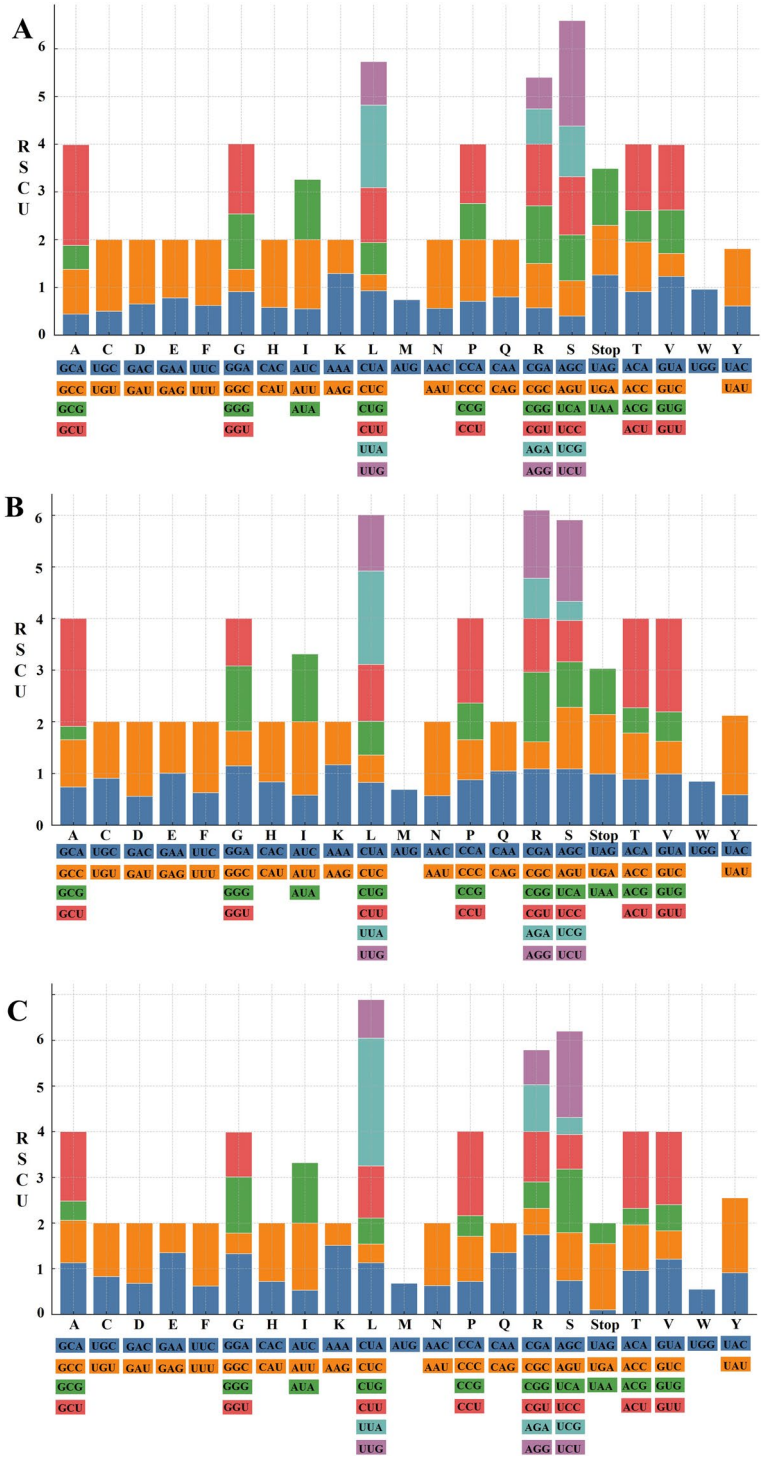

Fig. 3 Relative Synonymous Codon Usage (RSCU) of protein-coding genes *O. japonicus* (**A**), *O. scyllarus* (**B**), and *L. maculata* (**C**) mitochondrial genomes

significant differences in local folding patterns and spatial configurations can still be observed among three Stomatopoda species. Specifically, tRNAs of *O. japonica* exhibited a more compact conformation, with closer proximity between the TψC and DHU arms; tRNAs of *L. maculata* displayed greater structural variability and extended loop regions, suggesting increased spatial tension, whereas *O. scyllarus* tRNAs demonstrated higher symmetry and conformational stability. Further comparative analyses indicated that $tRNA^{L(UAA)}$ and $tRNA^{V(UAC)}$ were structurally conserved across all three species, reflecting their essential and evolutionarily constrained

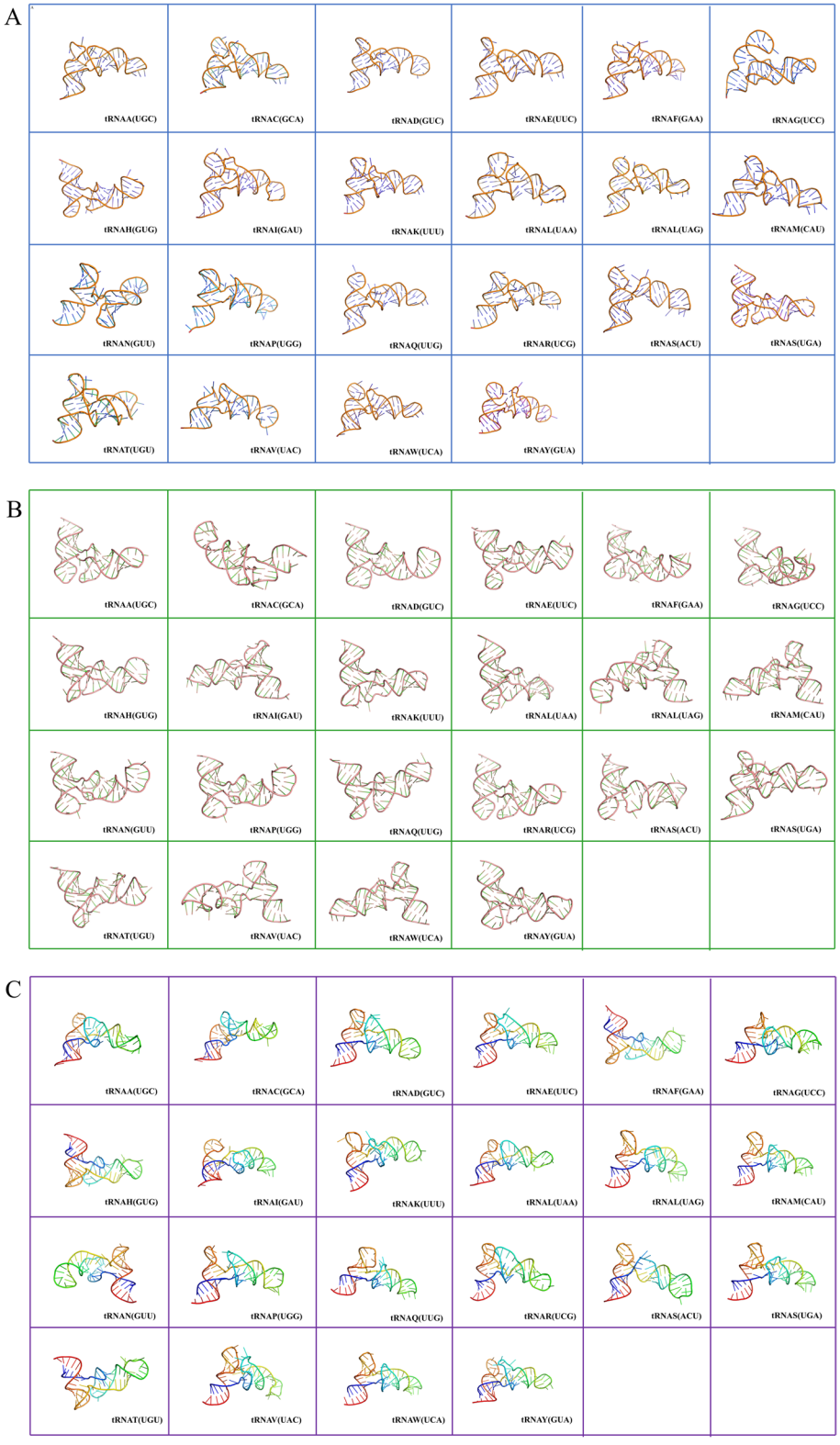

Fig. 4 Predicted 3D models of mitochondrial tRNA secondary structures in *O. japonicus* (**A**), *O. scyllarus* (**B**), and *L. maculate* (**C**)

functional roles. In contrast, substantial deviations were detected in $tRNA^{S(ACU)}$ and $tRNA^{C(GCA)}$ (particularly in *L. maculata*), and the DHU arm of $tRNA^{S(ACU)}$ was notably reduced or absent, implying potential structural degeneration or functional replacement. Additionally, the folding pattern of $tRNA^{P(UGG)}$ shows significant differences among three species, which potentially reflect the species-specific adaptations to ecological or physiological requirements.

Multi-perspective phylogenetic analyses

The mitochondrial genome, CDSs and PEPs were first used separately to construct the ML tree for the 18 research species. Interestingly, gene tree conflict occurs in phylogenetic relationships based on different markers. Specifically, the phylogenetic relationships based on CDSs and PEPs indicate that species belonging to the same order (Stomatopoda, Decapoda, Amphipoda, and Isopoda) are first clustered together, however, the phylogenetic relationships based on the mitochondrial genome result in a confusion between the Stomatopoda and Decapoda. In fact, this confusion also occurs when we use the CDS markers of *COX2* (Supplementary Fig. 1-C) and *ND6* (Supplementary Fig. 1-J), as well as the PEP markers of *ND5* (Supplementary Fig. 2-E), *ND4* (Supplementary Fig. 2-G), and *ND6* (Supplementary Fig. 2-J). It is worth noting that when constructing phylogenetic trees using different genetic markers, each Stomatopoda species (*O. scyllarus*, *O. japonicus*, *L.maculata*, and *O. oratoria*) used in the present study may potentially be located at the root of the Stomatopoda clustering relationship (Fig. 5; Supplementary Fig. 1; Supplementary Fig. 2).

To improve the resolution of the phylogenetic relationship, we integrated the above-mentioned mitochondrial genome-, CDS- and PEP- based phylogenetic trees to reconstruct a coalescent-based phylogenetic tree (Fig. 6). According to all phylogenetic analysis, species from the same order eventually cluster into one branch. In the integrated phylogenetic tree, the *L. maculata* is positioned at the root of the evolutionary lineage of the four Stomatopoda species, which is consistent with the phylogenetic relationship inferred from PEP markers.

Furthermore, the differentiation times of all research animals were further estimated (Fig. 7). Results showed that the divergence time between Stomatopod and other three orders (Decapoda, Amphipoda, and Isopoda) was approximately 385.00 Mya. Meanwhile, the divergence time between *L. maculata* and the other three Stomatopod species (*O. oratoria*, *O. japonicus*, and *O. scyllarus*) was approximately 287.02 Mya. The divergence time among *O. oratoria*, *O. japonicus*, and *O. scyllarus* was approximately 259.14 Mya. Additionally, the divergence time between *O. japonicus* and *O. scyllarus* was approximately 169.08 Mya.

Discussion

Genome structural characteristics in three Stomatopod species

This present study is the first to systematically investigate the whole-genomic structural characteristics of three Stomatopod species (including *O. japonicus*, *O. scyllarus*, and *L. maculata*) by combining whole-genome survey sequencing with the K-mer algorithm. All three Stomatopod species were found to have significantly larger whole-genome sizes (revised *O. japonicus*, *O. scyllarus*, and *L. maculata* genome size was 8,942 Mb, 11,068 Mb, and 12,743 Mb, respectively), far exceeding those of crustaceans (i.e. *P. clarkii* [2.75 Gb; [42]], *C. quadricarinatus* [5.26 Gb; [53]], *Scylla paramamosain* [1.21 Gb; [54]], *H. americanus* [4.31 Gb; [55]], *P. ornatus* [2.65 Gb; [40]], *O. oratoria* [2.97 Gb; [18]]) other than *Euphausia superba*[48.01 Gb; [56]] that have been reported. Further K-mer analyses revealed that the genomes of all three stomatopod species contained high proportions of repetitive sequences (83.26%–86.22%), which are substantially greater than those reported for other crustaceans, such as *P. ornatus* (~69% [40]) and *O. oratoria* (~75% [18]). This indicates that these three stomatopod species may have undergone unique genome expansion events or accumulated more repetitive elements in comparison to many crustaceans.

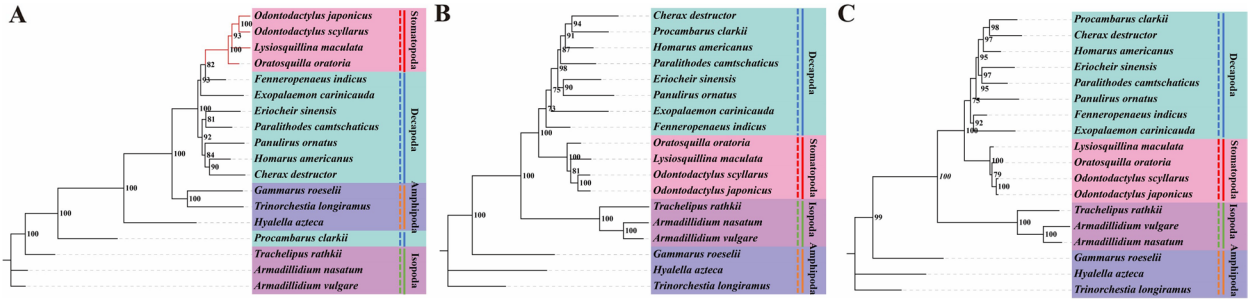

Fig. 5 Maximum likelihood (ML) trees based on the mitochondrial genome (**A**), coding sequences (CDS) (**B**), and protein-coding proteins (PEP) (**C**)

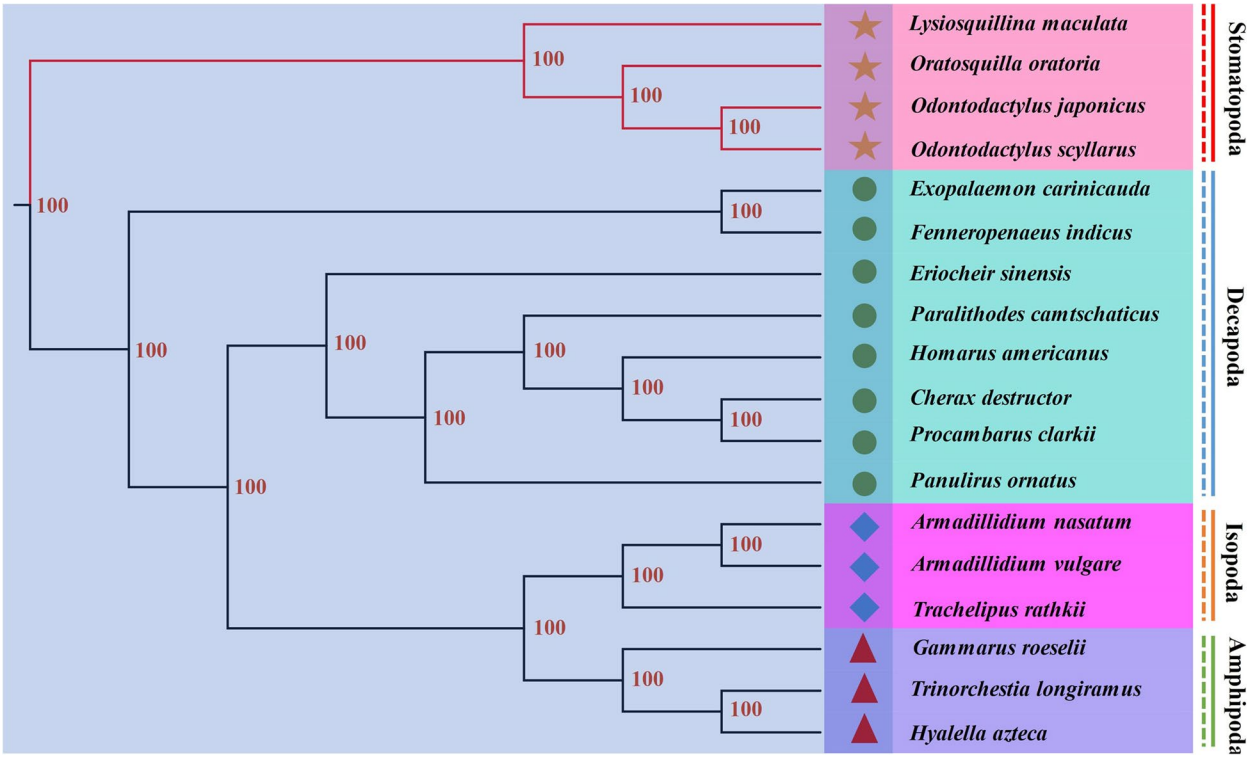

Fig. 6 Maximum likelihood (ML) phylogenetic tree of representative Malacostraca species, constructed by integrating trees based on the mitochondrial genome, CDS, and PEP datasets. The tree comprises species from four orders: Stomatopoda, Decapoda, Isopoda, and Amphipoda. Bootstrap support values derived from 1,000 replicates are shown at each node, and colored shapes denote taxonomic groups as illustrated in the legend on the right

It is worth noting that the differences in genome size among the three Stomatopod species may reflect their long-term adaptation to different habitats [1, 7]. Particularly, the *O. scyllarus* and the *L. maculata* usually live in coral reef or reef-associated environment, while the *O. japonicus* has a more diverse living environment, capable of living in sandy or shelly substrates ([3, 57], https://www.sealifebase.ca/). We hypothesize that different ecologic al response strategies might cause their genomes to follow different evolutionary paths [58, 59]. In other words, the relatively compact genome of *O. japonicus* may be associated with stronger selection pressures due to its adaptation to multiple habitat types. A reduced genome size could minimize energy expenditure and replication costs, thereby enhancing ecological tolerance and competitive ability for survival [60]. In contrast, *O. scyllarus* and *L. maculata*, which inhabit relatively uniform habitats, may have experienced weaker selection pressures. Such conditions can facilitate the accumulation of non-functional repetitive elements and thereby promote genome expansion. Comparable evolutionary patterns have also been documented in other taxa [61, 62]. Notably, the astonishing genome size of *O. scyllarus* may be related to its highly specialized behaviors and physiological functions during the evolutionary process. It is well known that *O. scyllarus* possesses remarkable visual and neural complexity capabilities, which may require expansion in gene families related to sensory perception, neuromuscular control, and regulatory architecture, ultimately leading to repetitive accumulation and genomic expansion [63, 64].

From a methodological perspective, this study demonstrates that whole-genome survey sequencing has an advantage over traditional sequencing [40] in analyzing the genomic structures of Stomatopod species with large and highly repetitive genomes. The whole-genome survey sequencing data obtained in this study can provide valuable genetic information for subsequent studies on phylogenomic relationships, functional genomics, and adaptive evolution of stomatopods.

Investigation of the three Stomatopod species mitochondrial genome

We assembled the mitochondrial genomes of three Stomatopod species using whole-genome survey data and identified species-specific features, including genome structure, RSCU and tRNA conformations. The mitochondrial genomes of the three Stomatopod species exhibit typical and evolutionarily conserved metazoan characteristics, with lengths ranging from 16,110 to 16,237 bp, a strong A/T nucleotide bias (65.50%—66.80%), and a high degree of gene sequence similarity.

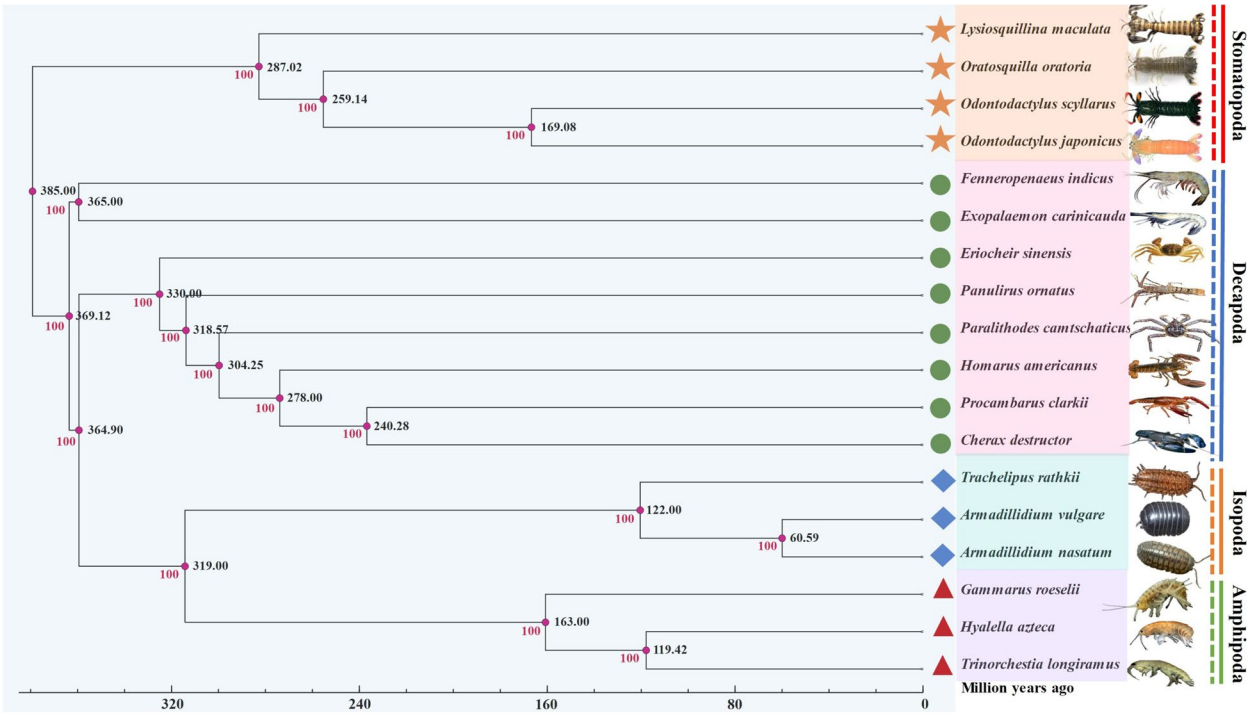

Fig. 7 Divergence time estimates and an integrated phylogenetic tree of representative Malacostraca species were generated through coalescent analyses of complete mitochondrial genome, CDS, and PEP datasets. The tree encompasses species from four orders: Stomatopoda, Decapoda, Isopoda, and Amphipoda. Black numbers above the branches indicate estimated divergence times (in million years ago), while red numbers at the nodes denote bootstrap support values based on 1,000 replicates. Colored shapes and illustrations correspond to the respective taxonomic groups as shown in the legend on the right

Meanwhile, the analysis of AT- and GC- skews revealed strand-specific mutational asymmetric pattern during the mitochondrial genome replication, consistent with other crustaceans, thereby supporting a conserved replication mechanism [65, 66]. Despite these similarities, notable differences remain. Specifically, *O. japonicus* retains the complete set of 37 standard mitochondrial genes, whereas *O. scyllarus* and *L. maculata* lack the *ATP8* gene. Previous studies have shown that *ATP8* is one of the most labile and structurally flexible mitochondrial genes. Its function is redundant with *ATP6* or can be compensated by nuclear-encoded subunits, and thus it is frequently lost in metazoans [67, 68]. Nevertheless, *ATP8* is highly divergent and variable in length, so its apparent absence may reflect annotation challenges rather than true gene loss [69]. In conclusion, we cannot definitively conclude that *ATP8* is absent from the mitochondrial genomes of *O. scyllarus* and *L. maculata.*

Further RSCU analysis indicates that the three mitochondrial genomes generally tend towards A- or T-terminal codons and this bias is mainly driven by the AT-rich base composition [70, 71]. Meanwhile, the most frequently used codons exhibited species-specificity, specifically, *O. japonicus* prefers UUA (RSCU = 1.73), *O. scyllarus* prefers UUU (RSCU = 1.37), and *L. maculata* displays the highest usage of UUA (RSCU = 2.80), which may be related to species-specific translation efficiency, metabolic rate or ecological niche differentiation [67]. 3D model of the secondary structure predictions indicated that all 22 mitochondrial tRNAs in the three stomatopod species adopt the canonical cloverleaf and L-shaped conformations. Nevertheless, we observed substantial variation in folding compactness and loop length: tRNAs of *O. japonicus* display more compact folds; those of *L. maculata* show longer loops and greater structural tension; and those of *O. scyllarus* exhibit higher symmetry and apparent structural stability. The structures of $tRNA^{L(UAA)}$and $tRNA^{V(UAC)}$ are highly conserved in all three Stomatopod species, which further emphasizes their crucial translational roles and evolutionary constraints. However, the $tRNA^{S(ACU)}$ of *L. maculata* exhibits a markedly reduced or absent DHU arm, a pattern that accords with the well-documented "structural minimization–functional compensation" model of mitochondrial tRNAs [69, 70]. Additionally, the interspecific differences in $tRNA^{P(UGG)}$ structure may also reflect species-specific physiological or ecological adaptations [63, 64, 72].

From an ecological perspective, the mitochondrial genome features may be closely related to the habitat preference, rather than stating that they seem to be closely related. The mitochondrial genome of *O. japonicus* is compact and complete, which may facilitate

efficient transcription and translation even under relatively intense environmental selection pressures [61]. By contrast, *L. maculata* and *O. scyllarus* inhabit relatively simple and stable environments and may therefore experience weaker selection pressures; correspondingly, their tRNA structures appear to tolerate greater variation and redundancy. It is worth noting that the relatively large mitochondrial genome and the more complex codon usage pattern of *O. scyllarus* may be related to its special predatory behavior, rapid metabolism, and highly developed sensory system. These physiological characteristics usually require mitochondrial gene expansion and more complex regulatory regions [38]. Overall, these findings highlight the evolutionary conservation of Stomatopod mitochondrial genome while revealing species-specific ecological adaptations.

Multi-marker coalescent framework updates the phylogenetic conflicts and divergence history in Stomatopoda species

Our multi-marker phylogenetic analyses (mitochondrial genome, CDS, PEP, and an integrated dataset) confirm substantial gene-tree conflict among stomatopod phylogenies, a phenomenon widely recognized in phylogenomics [67, 71]. Although the CDS- and PEP- based ML trees in this study clearly recovered the monophyly of the orders (Stomatopoda, Decapoda, Amphipoda, and Isopoda), the mitochondrial genome–based analysis yielded inconsistent placements, leading to an intermingling of Stomatopoda and Decapoda (Fig. 5). Similar discordance is observed for topologies inferred from *COX2*, *ND4*, *ND5*, and *ND6* markers (Supplementary Figs. 1 and 2). Existing studies suggest that these conflicts may arise from incomplete lineage sorting, introgression, or heterogeneity in substitution rates [65, 66, 69]; differences in modes of inheritance, recombination rates, and selection pressures between mitochondrial and nuclear genomes may further amplify the discrepancies [33, 34]. To address this, we integrated the mitochondrial-genome, CDS, and PEP datasets within a coalescent-based framework and obtained a more reliable species tree. This integrative, multi-marker approach is particularly effective for stomatopods shaped by rapid radiations or deep divergences, thereby mitigating the inconsistencies introduced by single-marker inference [45–48]. However, the phylogenetic position conflicts in this study (mainly the divergence between *E. carinicauda* and *F. indicus*) may reflect recent gene flow divergence or incomplete sampling [34, 40].

In the phylogenetic tree constructed with integrated markers, *L. maculate* is occupied the basal position of the Stomatopoda species. Meanwhile, divergence time estimation suggests that *L. maculata* split from the clade comprising *O. oratoria*, *O. japonicus*, and *O. scyllarus* at approximately 287.02 Ma (late Permian), predating the end-Permian mass extinction and coinciding with major marine faunal turnovers [16, 71]. The divergence between *O. oratoria* and *Odontodactylus* was inferred to occur around 259.14 Ma (Middle Triassic), a period characterized by warm seas and rapid ecological expansion [64]. The split between *O. japonicus* and *O. scyllarus* was dated to ~ 169.08 Ma (Middle Jurassic), preceding the morphological differentiation suggested by fossil evidence, which may reflect heterogeneity in mitochondrial evolutionary rates or calibration bias [15, 66]. From the ecological perspective, the differentiation events of the four Stomatopoda species studied here may be associated with substantial differences in benthic substrates, trophic niches, visual adaptations, and other ecological factors [9, 72]. This also implies that habitat type, food availability, and light conditions are likely key drivers of Stomatopoda diversification [65, 70]. However, the relevant evidence still needs to be further refined.

Overall, this study demonstrates that a multi-marker integrative systematics approach is applicable to resolving the phylogenetic conundrum of Stomatopoda [18, 33, 61]. However, the evolutionary complexity of different genetic markers cannot be ignored. The future integration of molecular, paleontological, and ecological evidences for Stomatopoda is also worthy of consideration, which will help to precisely elucidate the evolutionary history of Stomatopoda.

Conclusions

This study was the first to perform whole-genome survey sequencing of three stomatopod species, and subsequently analyzed the characteristics of their exceptionally large genomes and mitochondrial genomes. Meanwhile, we reconstructed the phylogenetic relationships of Stomatopoda from multiple aspects, which provided new evidence to resolve the controversy over the phylogenetic position of Stomatopoda. Nevertheless, the relatively small number of stomatopod species included still limits the strength of our conclusions to some extent. Future research should expand taxon sampling to enable a more accurate resolution of stomatopod phylogenetic relationships.

Supplementary Information

The online version contains supplementary material available at https://doi.org/10.1186/s12864-025-12296-0.

Supplementary Material 1: Supplementary Table 1. Summary of partition testing for the ten CDSs.

Supplementary Material 2: Supplementary Table 2. Relative Synonymous Codon Usage (RSCU) and the Codon Counts of protein-coding genes *Odontodactylus japonicus*, *Odontodactylus scyllarus*, and *Lysiosquillina maculata* mitochondrial genomes.

Supplementary Material 3: Supplementary Fig. 1. Maximum likelihood

phylogenetic trees based on coding sequences (CDS) of ten mitochondrial protein-coding genes: *ATP6*, *COX1*, *COX2*, *COX3*, *CYTB*, *ND3*, *ND4*, *ND4L*, *ND5*, and *ND6* (panels A–J, respectively). Taxa are color-coded as follows: pink for Stomatopoda, MediumTurquoise for Decapoda, purple for Amphipoda, and magenta for Isopoda.

Supplementary Material 4: Supplementary Fig. 2. Maximum likelihood phylogenetic trees based on protein sequences (PEP) of ten mitochondrial protein-coding genes: *ATP6*, *COX1*, *COX2*, *COX3*, *ND5*, *ND3*, *ND4*, *ND4L*, *CYTB*, and *ND6* (panels A–J, respectively). Taxa are color-coded as follows: MediumVioletRed for Stomatopoda, Purple for Decapoda, LightPink for Amphipoda, and Crimson for Isopoda.

Authors' contributions

Xiaowen Duan: Writing—original draft, Formal analysis, Data curation. Yuan Li: Project administration, Supervision, Funding acquisition. Yue Ding: Formal analysis. Xinxian Huang, Jiaguang Xiao and Yifei Wang: Investigation, Sampling. Fangrui Lou: Writing—original draft, Writing—review & editing, Project administration, Supervision, Investigation, Funding acquisition, Conceptualization.

Funding

This work was supported by the Natural Science Foundation of Shandong Province (ZR2024MD076) and National Natural Science Foundation of China (32370565).

Data availability

The data underlying this article are stored in the SRA database of NCBI under BioProject PRJNA1205873.

Declarations

Ethics approval and consent to participate

O. japonicus, *O. scyllarus*, and *L. maculata* are not an endangered or protected animal in any country, and no special ethical approval are required. All animals used in subsequent experiments were placed on ice for frozen anesthesia before dissection to minimize the suffering. Meanwhile, we rigorously conducted the experiments under the guidance of the National Institutes of Health Guidelines for the Care and Use of Laboratory Animals (NIH Publication No. 8023, revised in 1978) and Institutional Animal Care and Use Committee of Yantai University.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

[1]School of Ocean, Yantai University, Yantai, Shandong 264003, China
[2]Third Institute of Oceanography, Ministry of Natural Resources, Xiamen, Fujian 361005, China

Received: 28 August 2025 / Accepted: 3 November 2025
Published online: 08 December 2025

References

1. Cheng J, Wang YL, Sha ZL. Progress on the systematics of Stomatopoda (Crustacea: Malacostraca). Marine Science. 2015;39(12):173–7. https://doi.org/10.11759/hykx20150121001. (in Chinese).
2. Brooks WK. Report on the Stomatopoda collected by H.M.S. Challenger during the years 1873–1876. The voyage of the H.M.S. challenger, zoology. London; 1886. p. 1–116. https://doi.org/10.5962/bhl.title.9891.
3. Ahyong ST. Revision of the Australian stomatopod crustacea. Rec Aust Mus Suppl. 2001;26:1–326. https://doi.org/10.3853/j.0812-7387.26.2001.1333.
4. Sha ZL, Wang YL, Cheng J. Application of mitochondrial COI-based DNA barcoding for the identification of stomatopod species (Crustacea: Stomatopoda) in the China seas. J Fish Sci China. 2018;25(4):858–66. https://doi.org/10.3724/SP.J.1118.2018.17338. (in Chinese).
5. Manning RB. Stomatopod Crustacea of Vietnam: the legacy of Raoul Serène. Crustacean Research, Special No. 4. Carcinological Society of Japan, Tokyo. 1995. p. 1–339. https://doi.org/10.18353/crustacea.special1995.4_1.
6. Ahyong ST. A phylogenetic analysis of the Stomatopoda (Crustacea: Malacostraca). J Crustacean Biol. 1997;17(4):695–715. https://doi.org/10.1163/193724097X00134.
7. Ahyong ST, Harling C. The phylogeny of the Stomatopoda (Crustacea). Aust J Zool. 2000;48(6):607–42. https://doi.org/10.1071/ZO00042.
8. Hof CHJ. Fossil stomatopods (Crustacea: Malacostraca) and their phylogenetic impact. J Nat Hist. 1998;32(10–11):1567–76. https://doi.org/10.1080/00222939800771101.
9. Porter ML, Zhang Y, Desai S, Caldwell RL, Cronin TW. Evolution of anatomical and physiological specialization in the compound eyes of stomatopod crustaceans. J Exp Biol. 2010;213(20):3473–86. https://doi.org/10.1242/jeb.046508.
10. Marshall NJ, Land MF, King CA, Cronin TW. The compound eyes of mantis shrimps (Crustacea: Hoplocarida: Stomatopoda). I. Compound eye structure: the detection of polarized light. Philos Trans R Soc B Biol Sci. 1991;334(1269):33–56. https://doi.org/10.1098/rstb.1991.0096.
11. Ahyong ST, Jarman SN. Stomatopod interrelationships: preliminary results based on analysis of three molecular loci. Arthropod Syst Phylogeny. 2009;67(1):91–8. https://doi.org/10.3897/asp.67.e31690.
12. Barber PH, Erdmann MV. Molecular systematics of the Gonodactylidae (Stomatopoda) using mitochondrial cytochrome oxidase subunit 1 DNA sequence data. J Crustacean Biol. 2000;20(3):20–36. https://doi.org/10.1163/1937240X-90000004.
13. Ji DP, Sun ZC, Song N, Gao TX, Xu SY. The complete mitochondrial genome of *Jaydia lineata* (Perciformes: Apogonidae) obtained by next-generation sequencing. Mitochondr DNA B. 2020;5(3):2507–8. https://doi.org/10.1080/23802359.2020.1780970.
14. Renner SS, Wu S, Pérez-Escobar OA, Silber MV, Fei ZJ, Chomicki G. A chromosome-level genome of a Kordofan melon illuminates the origin of domesticated watermelons. Proc Natl Acad Sci U S A. 2021;118(23):e2101486118. https://doi.org/10.1073/pnas.2101486118.
15. Colbourne JK, Pfrender ME, Gilbert D, et al. The ecoresponsive genome of *Daphnia pulex*. Science. 2011;331(6017):555–61. https://doi.org/10.1126/science.1197761.
16. Zhang XJ, Yuan JB, Sun YM, et al. Penaeid shrimp genome provides insights into benthic adaptation and frequent molting. Nat Commun. 2019;10:356. https://doi.org/10.1038/s41467-018-08197-4.
17. Han ZQ, Gao TX, Lou FR, Liu Q. *Charybdis japonica* genome provides insights into desiccation adaptation and sex-determining region. Zool Res. 2022;43(6):927–30. https://doi.org/10.24272/j.issn.2095-8137.2022.155.
18. Zhang DZ, Sun XL, Chen LF, et al. The chromosome-level genome provides insights into the adaptive evolution of the visual system in *Oratosquilla oratoria*. BMC Biol. 2025;23:38. https://doi.org/10.1186/s12915-025-02146-6.
19. Xiong Y, Lei X, Bai S, Xiong Y, Liu W, Wu W, et al. Genomic survey sequencing, development and characterization of single- and multi-locus genomic SSR markers of *Elymus sibiricus* L. BMC Plant Biol. 2021;21:3. https://doi.org/10.1186/s12870-020-02770-0.
20. Roy D, Tomo S, Modi A, Purohit P, Sharma P. Optimising total RNA quality and quantity by phenol-chloroform extraction method from human visceral adipose tissue: a standardisation study. MethodsX. 2020;7:101113. https://doi.org/10.1016/j.mex.2020.101113.
21. Chen SF, Zhou YQ, Chen YR, Gu J. Fastp: an ultra-fast all-in-one FASTQ preprocessor. Bioinformatics. 2018;34(24):i884–90. https://doi.org/10.1093/bioinformatics/bty560.
22. Tammi MT, Arner E, Kindlund E, Andersson B. Correcting errors in shotgun sequences. Nucleic Acids Res. 2003;31(15):4663–72. https://doi.org/10.1093/nar/gkg653.
23. Liu BH, Shi YJ, Yuan JY, et al. Estimation of genomic characteristics by analyzing k-mer frequency in de novo genome projects. arXiv [Preprint]. 2013. https://doi.org/10.48550/arXiv.1308.2012.
24. Dierckxsens N, Mardulyn P, Smits G. NovoPlasty: de novo assembly of organelle genomes from whole genome data. Nucleic Acids Res. 2017;45(4):e18. https://doi.org/10.1093/nar/gkw955.
25. Huang W, Zhao T, Fan M, et al. Phylogenetic relationships and divergence times of Odonata inferred from mitochondrial genome data. iScience. 2025;28(2):111806. https://doi.org/10.1016/j.isci.2025.111806.

26. Benjamin M, Singh J, Pandey AK, et al. Integrative genetic and transcriptomic subtyping improves prognosis prediction in B-lineage acute lymphoblastic leukemia. Lab Invest. 2025;105(9):104201. https://doi.org/10.1016/j.labinv.2025.104201.
27. Rice P, Longden I, Bleasby A. EMBOSS: the European Molecular Biology Open Software Suite. Trends Genet. 2000;16(6):276–7. https://doi.org/10.1016/S0168-9525(00)02024-2.
28. Manon MRK, Hoque MS, Alam MA, et al. Microplastic pollution in aquaculture species: quantification and characterization in Tilapia and Pangas. Aquaculture. 2026;610:742922. https://doi.org/10.1016/j.aquaculture.2025.742922.
29. Wang W, Feng C, Han R, et al. TrrosettaRNA: automated prediction of RNA 3D structure with transformer network. Nat Commun. 2023;14(1):7266. https://doi.org/10.1038/s41467-023-42528-4.
30. Zhao Y, Meng J, Wang Y, et al. Research progress of β-xylosidase in green synthesis. Int J Biol Macromol. 2025;306:141404. https://doi.org/10.1016/j.ijbiomac.2025.141404.
31. Becking T, Chebbi MA, Giraud I, et al. Sex chromosomes control vertical transmission of feminizing *Wolbachia* symbionts in an isopod. PLoS Biol. 2019;17(10):e3000438. https://doi.org/10.1371/journal.pbio.3000438.
32. Chebbi MA, Becking T, Moumen B, et al. The genome of *Armadillidium vulgare* (Crustacea: Isopoda) provides insights into sex chromosome evolution in the context of cytoplasmic sex determination. Mol Biol Evol. 2019;36(4):727–41. https://doi.org/10.1093/molbev/msz010.
33. Austin CM, Croft LJ, Grandjean F, et al. The NGS magic pudding: a nanopore-led long-read genome assembly for the commercial Australian freshwater crayfish, *Cherax destructor*. Front Genet. 2022;12:695763. https://doi.org/10.3389/fgene.2021.695763.
34. Tang BP, Wang ZK, Liu QN, et al. High-quality genome assembly of *Eriocheir japonica sinensis* reveals its unique genome evolution. Front Genet. 2020;10:1340. https://doi.org/10.3389/fgene.2019.01340.
35. Wang JJ, Lv JJ, Shi M, et al. Chromosome-level genome assembly of ridgetail white shrimp *Exopalaemon carinicauda*. Sci Data. 2024;11:576. https://doi.org/10.1038/s41597-024-03423-9.
36. Katneni VK, Shekhar MS, Jangam AK, et al. A superior contiguous whole genome assembly for shrimp (*Penaeus indicus*). Front Mar Sci. 2022;8:808354. https://doi.org/10.3389/fmars.2021.808354.
37. Cormier A, Chebbi MA, Giraud I, et al. Comparative genomics of strictly vertically transmitted, feminizing microsporidia endosymbionts of amphipod crustaceans. Genome Biol Evol. 2021;13(1):evaa245. https://doi.org/10.1093/gbe/evaa245.
38. Polinski JM, Zimin AV, Clark KF, et al. The American lobster genome reveals insights on longevity, neural, and immune adaptations. Sci Adv. 2021;7(26):eabe8290.
39. Poynton HC, Hasenbein S, Benoit JB, et al. The toxicogenome of *Hyalella azteca*: a model for sediment ecotoxicology and evolutionary toxicology. Environ Sci Technol. 2018;52(10):6009–22.
40. Ren XY, Sun DF, Lv JJ, et al. Chromosome-level genome of the long-tailed marine-living ornate spiny lobster, *Panulirus ornatus*. Sci Data. 2024;11(1):662. https://doi.org/10.1038/s41597-024-03512-9.
41. Veldsman WP, Ma KY, Hui JHL, et al. Comparative genomics of the coconut crab and other decapod crustaceans: exploring the molecular basis of terrestrial adaptation. BMC Genomics. 2021;22(1):313. https://doi.org/10.1186/s12864-021-07636-9.
42. Xu Z, Gao T, Xu Y, et al. A chromosome-level reference genome of red swamp crayfish *Procambarus clarkii* provides insights into the gene families regarding growth or development in crustaceans. Genomics. 2021;113(5):3274–84. https://doi.org/10.1016/j.ygeno.2021.07.017.
43. Russell A, Borrelli S, Fontana R, et al. Evolutionary transition to XY sex chromosomes associated with Y-linked duplication of a male hormone gene in a terrestrial isopod. Heredity. 2021;127(3):266–77. https://doi.org/10.1038/s41437-021-00457-2.
44. Patra AK, Chung O, Yoo JY, et al. First draft genome for the sand-hopper *Trinorchestia longiramus*. Sci Data. 2020;7(1):85. https://doi.org/10.1038/s41597-020-0424-8.
45. Katoh K, Standley DM. MAFFT multiple sequence alignment software version 7: improvements in performance and usability. Mol Biol Evol. 2013;30(4):772–80. https://doi.org/10.1093/molbev/mst010.
46. Edgar RC. Muscle5: high-accuracy alignment ensembles enable unbiased assessments of sequence homology and phylogeny. Nat Commun. 2022;13(1):6968. https://doi.org/10.1038/s41467-022-34630-w.
47. Capella-Gutierrez S, Silla-Martínez J, Gabaldón T. TrimAl: a tool for automated alignment trimming in large-scale phylogenetic analyses. Bioinformatics. 2009;25:1972–3. https://doi.org/10.1093/bioinformatics/btp348.
48. Bierenbroodspot MJ, Darienko T, de Vries S, et al. Phylogenomics unveil a recent origin of morphological complexity in Coleochaetophyceae. Curr Biol. 2025. https://doi.org/10.1016/j.cub.2025.08.046.
49. Stamatakis A. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics. 2014;30(9):1312–3. https://doi.org/10.1093/bioinformatics/btu033.
50. Mirarab S, Reaz R, Bayzid MS, et al. ASTRAL: genome-scale coalescent-based species tree estimation. Bioinformatics. 2014;30(17):i541–8. https://doi.org/10.1093/bioinformatics/btu462.
51. Hedges SB, Marin J, Suleski M, et al. Tree of life reveals clock-like speciation and diversification. Mol Biol Evol. 2015;32:835–45. https://doi.org/10.1093/molbev/msv037.
52. Sanderson MJ. R8s: inferring absolute rates of molecular evolution and divergence times in the absence of a molecular clock. Bioinformatics. 2003;19:301–2. https://doi.org/10.1093/bioinformatics/19.2.301.
53. Chen H, Zhang R, Liu F, et al. The chromosome-level genome of *Cherax quadricarinatus*. Sci Data. 2023;10(1):215. https://doi.org/10.1038/s41597-023-02124-z.
54. Rutz C, Bonassin L, Kress A, et al. Abundance and diversification of repetitive elements in decapoda genomes. Genes. 2023;14(8):1627. https://doi.org/10.3390/genes14081627.
55. Jeffery NW, Gregory TR. Genome size estimates for crustaceans using Feulgen image analysis densitometry of ethanol-preserved tissues. Cytometry A. 2014;85(10):862–8. https://doi.org/10.1002/cyto.a.22516.
56. Shao CW, Sun S, Liu KQ, et al. The enormous repetitive Antarctic krill genome reveals environmental adaptations and population insights. Cell. 2023;186:1279–94. https://doi.org/10.1016/j.cell.2023.02.005.
57. Manning RB. Stomatopods. In: Carpenter KE, Niem VH, editors. FAO species identification guide for fishery purposes. The living marine resources of the Western Central Pacific. Vol. 2. Cephalopods, crustaceans, holothurians and sharks. Rome: FAO; 1998. p. 827–849.
58. Yuan Z, Liu S, Zhou T, et al. Comparative genome analysis of 52 fish species suggests differential associations of repetitive elements with their living aquatic environments. BMC Genomics. 2018;19:141. https://doi.org/10.1186/s12864-018-4516-1.
59. Heckenhauer J, Frandsen PB, Sproul JS, et al. Genome size evolution in the diverse insect order Trichoptera. Gigascience. 2022;11:giac011. https://doi.org/10.1093/gigascience/giac011.
60. Qin QL, Li Y, Sun LL, et al. Trophic specialization results in genomic reduction in free-living marine *Idiomarina bacteria*. MBio. 2019;10(1):e02545-e2618. https://doi.org/10.1128/mBio.02545-18.
61. Bickhart DM, Rosen BD, Koren S, et al. Single-molecule sequencing and chromatin conformation capture enable de novo reference assembly of the domestic goat genome. Nat Genet. 2017;49(4):643–50. https://doi.org/10.1038/ng.3802.
62. Lyu T, Zhou S, Fang J, et al. Convergent genomic signatures of high-altitude adaptation among six independently evolved mammals. Animals. 2022;12(24):3572. https://doi.org/10.3390/ani12243572.
63. Ramos-Vicente D, Ji J, Gratacòs-Batlle E, et al. Metazoan evolution of glutamate receptors reveals unreported phylogenetic groups and divergent lineage-specific events. Genome Biol Evol. 2018;10(4):927–43. https://doi.org/10.1093/gbe/evy059.
64. Policarpo M, Fogg LG, Cortesi F, et al. Evolution of the nonvisual and visual opsin gene repertoire in ray-finned fishes. Genome Biol Evol. 2025;17(7):evaf129. https://doi.org/10.1093/gbe/evaf129.
65. Perna NT, Kocher TD. Patterns of nucleotide composition at fourfold degenerate sites of animal mitochondrial genomes. J Mol Evol. 1995;41(3):353–8. https://doi.org/10.1007/BF00186547.
66. Reyes A, Gissi C, Pesole G, et al. Asymmetrical directional mutation pressure in the mitochondrial genome of mammals. Mol Biol Evol. 1998;15(8):957–66. https://doi.org/10.1093/oxfordjournals.molbev.a025999.
67. Cameron SL. Insect mitochondrial genomics: implications for evolution and phylogeny. Annu Rev Entomol. 2014;59:95–117. https://doi.org/10.1146/annurev-ento-011613-162007.
68. Kilpert C, Podsiadlowski L. The complete mitochondrial genome of the common sea slater *Ligia oceanica* (Crustacea: Isopoda) bears a novel gene order and unusual control region features. BMC Genomics. 2006;7:241. https://doi.org/10.1186/1471-2164-7-241.

69. Zhao BJ, Gao ST, Zhao MY, et al. Mitochondrial genomic analyses provide new insights into the "missing" atp8 and adaptive evolution of Mytilidae. BMC Genomics. 2022;23(1):738. https://doi.org/10.1186/s12864-022-08940-8.
70. Ma H, Ma C, Li X, et al. The complete mitochondrial genome sequence and gene organization of the mud crab (*Scylla paramamosain*) with phylogenetic consideration. Gene. 2013;519(1):120–7. https://doi.org/10.1016/j.gene.2013.01.028.
71. Gissi C, Iannelli F, Pesole G. Evolution of the mitochondrial genome of Metazoa as exemplified by comparison of congeneric species. Heredity. 2008;101(4):301–20. https://doi.org/10.1038/hdy.2008.62.
72. Ohtsuki T, Kawai G, Watanabe K. The minimal tRNA: unique structure of *Ascaris suum* mitochondrial tRNA-Ser(UCN) having a short T-arm and lacking the entire D-arm. Proc Natl Acad Sci U S A. 2002;99(21):13602–7. https://doi.org/10.1073/pnas.212336399.

Publisher's Note