

RESEARCH

Open Access



Chromosome-level genome assembly of the deep-sea snail *Phymorhynchus buccinoides* provides insights into the adaptation to the cold seep habitat

Zhaoqun Liu^{1,2,3,4†}, Yuting Huang^{5,6†}, Hao Chen^{7†}, Chang Liu^{1,2,3,4}, Minxiao Wang⁷, Chao Bian^{5*},
Lingling Wang^{1,2,3,4*} and Linsheng Song^{1,2,3,4*}

Abstract

Background The deep-sea snail *Phymorhynchus buccinoides* belongs to the genus *Phymorhynchus* (Neogastropoda: Raphitomidae), and it is a dominant specie in the cold seep habitat. As the environment of the cold seep is characterized by darkness, hypoxia and high concentrations of toxic substances such as hydrogen sulfide (H₂S), exploration of the diverse fauna living around cold seeps will help to uncover the adaptive mechanisms to this unique habitat. In the present study, a chromosome-level genome of *P. buccinoides* was constructed and a series of genomic and transcriptomic analyses were conducted to explore its molecular adaptation mechanisms to the cold seep environments.

Results The assembled genome size of the *P. buccinoides* was approximately 2.1 Gb, which is larger than most of the reported snail genomes, possibly due to the high proportion of repetitive elements. About 92.0% of the assembled base pairs of contigs were anchored to 34 pseudo-chromosomes with a scaffold N50 size of 60.0 Mb. Compared with relative specie in the shallow water, the glutamate regulative and related genes were expanded in *P. buccinoides*, which contributes to the acclimation to hypoxia and coldness. Besides, the relatively high mRNA expression levels of the olfactory/chemosensory genes in osphradium indicate that *P. buccinoides* might have evolved a highly developed and sensitive olfactory organ for its orientation and predation. Moreover, the genome and transcriptome analyses demonstrate that *P. buccinoides* has evolved a sulfite-tolerance mechanism by performing H₂S detoxification. Many genes involved in H₂S detoxification were highly expressed in ctenidium and hepatopancreas, suggesting that these tissues might be critical for H₂S detoxification and sulfite tolerance.

Conclusions In summary, our report of this chromosome-level deep-sea snail genome provides a comprehensive genomic basis for the understanding of the adaptation strategy of *P. buccinoides* to the extreme environment at the deep-sea cold seeps.

[†]Zhaoqun Liu, Yuting Huang and Hao Chen contributed equally to this work.

*Correspondence:

Chao Bian

bianchao@genomics.cn

Lingling Wang

wanglingling@dlou.edu.cn

Linsheng Song

lshsong@dlou.edu.cn

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



Keywords Cold seep, *Phymorhynchus buccinoides*, Chromosome-level genome, Glutamate regulation, Olfactory/chemosensory, H₂S detoxification

Background

Deep-sea cold seeps are submarine springs where fluids emanate from the sea floor through the sediments by pressure gradients [1, 2]. The environment of deep-sea cold seeps is characterized by darkness, coldness, hypoxia, lack of photosynthesis-derived nutrients but rich in heavy metals and toxic substances [3–5]. There is no penetration of light in the 1000 m depth below the sea surface [6]. Low-oxygen zones also occur in the deeper waters of tropical and temperate oceans, usually between 100 and 1000 m [7]. Temperature in the oceans decreases with increasing depth. At tropical latitudes, the temperature range extends from 26 °C at the sea surface to 5 °C at 1500 m depth [8, 9]. In particular, the anaerobic oxidation of methane (AOM) via sulfate reduction is considered as the most important biogeochemical process at cold seeps [10]. During this process, carbonates are formed through anaerobic methane oxidation to produce extremely high concentrations of H₂S in pore waters [11–13]. Therefore, an in-depth exploration of communities living in the deep-sea cold seeps will contribute to reveal the adaptation mechanism of this unique ecosystem.

In recent years, the benthic communities at the cold seeps have attracted increasing attention with the development of technologies for deep-sea research [14, 15]. A typical cold seep environment supports various communities of metazoans containing chemoautotrophic bacteria, tubeworms [16] and mussels [17]. As light disappears below 1000 m, deep-sea organisms living in the dark communicate and interact through chemical signals [18]. The temperatures of cold seeps are 2 °C to 5 °C and the known adaptations of low temperature for other invertebrates are mainly about freezing below 0 °C [19, 20]. The cold seep habitat is characterized by chronic hypoxia, sometimes reaching complete anoxia. Organisms inhabiting these environments often adopt morphological adaptations of gills [21]. In particular, recent research has shown that the H₂S level in the bottom water of the chemosynthetic communities is remarkably high (~1940 μM) at the active cold seep at Formosa Ridge (Site F) on the continental slope of the northern South China Sea [22]. The high concentrations of H₂S will induce oxidative damage to the marine invertebrates such as molluscs [23]. Invertebrates achieve sulfide detoxification by oxidation of sulfide and thiosulfate is the main detoxification product [24–26]. Meanwhile, more and more high-throughput technologies and comparative genomics are shedding light on the researches of extreme

environments adaptation [27, 28]. Since seep animals need to evolve unique adaptation mechanism to survive in the threatening environment, studying seep organisms at the genomic levels, especially the acclimation to the cold, dark, hypoxia, and H₂S-rich environment, will help to discover novel physiological and biochemical capabilities, and provide clues to understand their adaptation and evolution strategies to thrive in the extreme environments [3, 29].

The snail *Phymorhynchus buccinoides*, which belongs to the order Neogastropoda was sampled from the active deep-sea cold seep of the South Sea of China in this research. Neogastropods are often dominant species of the benthic community at the top of the food chains due to their amazing predatory specializations [30]. Previous researches have also shown that as a secondary consumer, *P. buccinoides* feeds on mytilid mussel *Bathymodiolus platifrons*, biological carcass, and organic debris sinking down from the upper layer. It is the important predator of the food chains in the cold seeps and it contributes to the balances of the cold seep fauna communities, energy flows and other interactions among the communities living around the deep-sea cold seeps [31, 32]. In the present study, we report the first chromosome-level reference genome of the deep-sea snail *P. buccinoides*. Comparative genomic analyses of gene expansion, contraction, and identification of genes related to sulfate metabolism, chemical sensing, and glutamate regulative genes which contribute to the acclimation to hypoxia and coldness were also conducted, helping to elucidate the molecular basis of the adaptation to the deep-sea cold seep habitats.

Results

Sequencing, assembly, annotation of chromosome-scale genome

The genome of *P. buccinoides* (Fig. 1a) was sequenced using a hybrid approach. A total of raw reads and clean data including 52.2 Gb and 42.0 Gb for Illumina reads with an insert size of 350 bp (Table S1), 392.9 Gb and 327.1 Gb for Hi-C reads with insert sizes of 300–500 bp (Table S3), 47.5 Gb and 45.7 Gb for transcriptomic Illumina reads (Table S4), polymerase and subreads: 130.4 Gb and 130.1 Gb for genome (Table S2), 21.3 and 20.4 for transcriptome (Table S5) of Pacific Biosciences (PacBio) reads with a long insert size of 20 kb, were obtained using the NovaSeq 6000 platform and PacBio

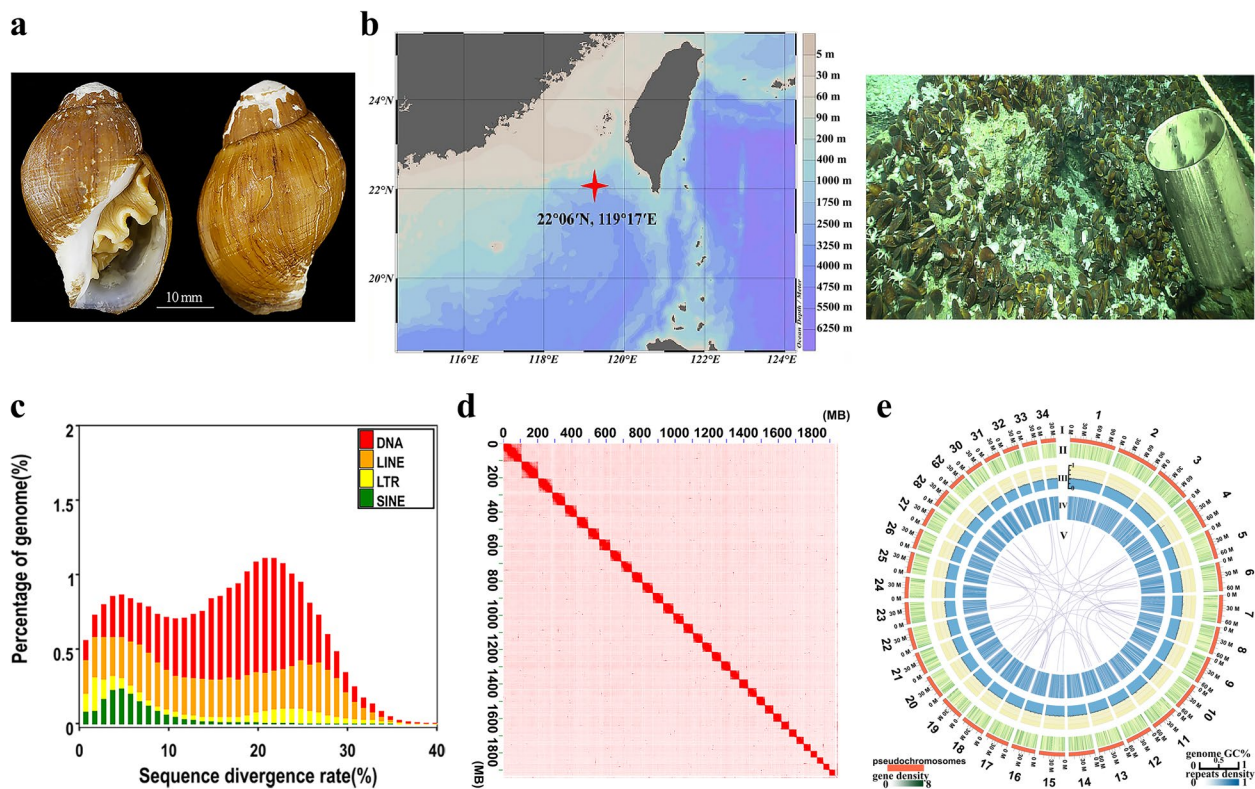


Fig. 1 Sampling site, morphology, divergence distribution of transposable elements, chromosomal contact map and genomic landscape of *P. buccinoides*. **a** Morphology of *P. buccinoides*. **b** Sampling site of *P. buccinoides*. *P. buccinoides* used in the present study were collected from cold spring district. This specie is also common in cold spring vents. The mussels associated with *P. buccinoides* are also in photo. **c** Divergence distribution of transposable elements (TEs) in the *P. buccinoides* genome. De novo prediction. The distribution of sequence divergence rates of TEs as percentages of the genome size is shown. The y-axis shows the percentage of the genome that is annotated as TEs (TE contents). The x-axis shows sequence divergence rate. DNA transposon shown as DNA is indicated with red color, long interspersed nuclear element (LINE) is indicated with orange color, long terminal repeat (LTR) is indicated with yellow color, and short interspersed nuclear element (SINE) is indicated with green color. The percentage of genome and the sequence divergence rate show the cumulative proportion and inconsistency of repeat elements, respectively. **d** Chromosomal contact map of *P. buccinoides*. Based on Hi-C data, the chromosomal contact map was built. The contacts between one location and another are referred by blocks. The blocks correspond to 34 Chrs of *P. buccinoides*. The color reflects the intensity of each contact, and it represents the interaction density from high (red) to low (white) in the plot. In the x-axis and y-axis, each number means the genomic length (Mb). **e** Diagram and genomic landscape of the gastropod *P. buccinoides*. Circos atlas represents the Chr information of *P. buccinoides*. From outside to inside of the concentric circles, (I) Chr length (Mb) and numbers, (II) Density of gene distribution in each 100 kb genomic interval, (III) GC content of 100 kb genomic intervals, and (IV) distribution of genomic repeats density in 100 kb non-overlapping windows. Deep blue color indicates higher repeat density. (V) Major interchromosomal relationships of *P. buccinoides* Chrs are presented with purple lines, and each line indicates one pair of paralog genes

Sequel instrument (Table S7). The PLATANUS v1.2.4 and DBG2OLC were employed to assemble sequence reads into contig level [33, 34]. The total length of the assembled contigs of *P. buccinoides* was 2.1 Gb, and the contig N50 value was 308.7 kb (Table S8). The Benchmarking Universal Single-Copy Orthologs (BUSCO) value was 86.0% (total BUSCO groups for searching is 5295), indicating the completeness of the assembly. The technology of pseudo-chromosomes (Chrs) construction for assistant assembly was employed to produce the final chromosome-level genome. A total of 18,751 contigs were broken, clustered, ordered and mounted successfully in

34 Chrs (Fig. 1d, e). Finally, the chromosome-level assembly of *P. buccinoides* was obtained and the Hi-C contact map was also produced (Table S9 and Fig. 1d). The longest Chr was 105.8 Mb and the shortest was 34.0 Mb. The total length of the *P. buccinoides* final chromosomal-level assembled genome and Chrs sequences were 2.1 Gb and 1.9 Gb, respectively. About 92.0% of the assembled base pairs of contigs were anchored to Chrs, and the N50 size of Chrs was 60.0 Mb (Table S9). There were numerous links between Chrs of *P. buccinoides* and that of its relative species (Figure S2). Detailed distributions of gene density, GC content, repeat sequence content of each

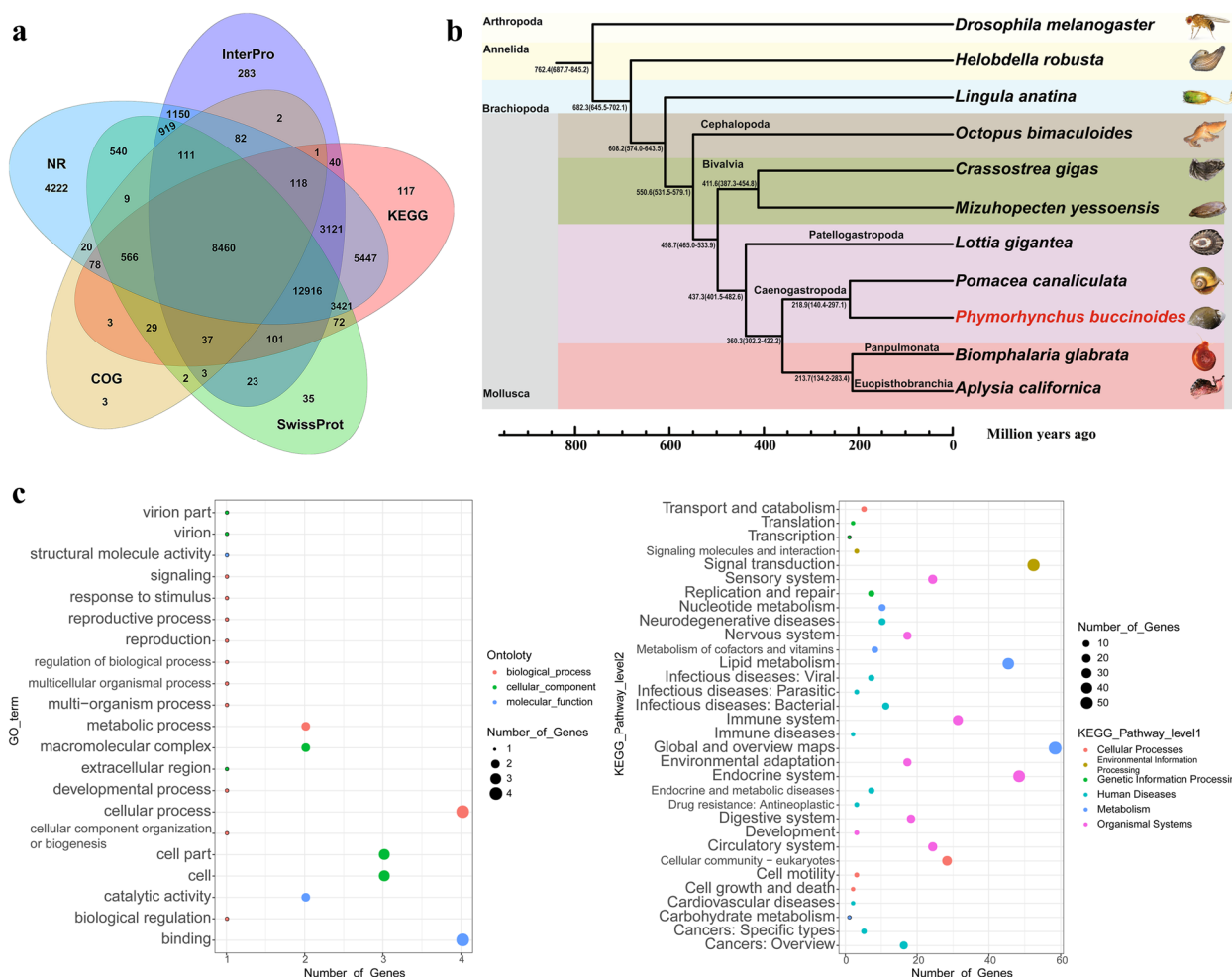


Fig. 2 Venn, phylogeny and clusters of Gene Ontology (GO) and Kyoto Encyclopedia of Genes and Genomes (KEGG) enrichment analysis. **a** Venn of the annotation in *P. buccinoides*. **b** Phylogenetic tree of *P. buccinoides*. The phylogenetic tree was based on the genome sequences. The reference divergence times for calibrations were retrieved from the TimeTree database. The black numbers on the branches represent the estimated diverge times. The *P. buccinoides* is marked in red color. **c** GO and KEGG enrichment analysis of gene families specific to *P. buccinoides*. The x-axis shows the number of genes and the y-axis shows the annotation terms. Different sizes and colors of bubbles exhibit different number and terms

Chr, and the major inner connections in *P. buccinoides* Chrs were illustrated in Fig. 1e.

In the assembled genome, the repeat content accounted for about 73.4% (Table S10). The transposable elements (TEs) including DNA transposons (24.19%) and retrotransposons (24.77%) accounted for 48.96% of the genome (Table S11), which showed high divergence (Fig. 1c, Figure S1). The retrotransposons were composed of 15.97% long interspersed nuclear elements (LINE), 1.74% short interspersed nuclear elements (SINE), and 7.06% long terminal repeats (LTR) (Table S11). A final nonredundant consensus gene set of *P. buccinoides* was obtained with the gene prediction and functional annotation (Fig. 2a). In the assembled genome, 45,545

protein-coding genes were predicted (Table S12), and 42,162 (92.6%) of them were functionally annotated (Table S13).

Gene families and phylogeny of *P. buccinoides*

In the present study, the gene family cluster analysis of *P. buccinoides* and other 10 selected species (including 7 molluscan species) were performed. In total, 33,107 gene families and 100 single-copy orthologs were identified across *P. buccinoides* and the other 10 species (Figure S3, Table S14). Comparisons of the genes of 7 molluscan species including *Octopus bimaculoides*, *Crassostrea gigas*, *Mizuhopecten yessoensis*, *Lottia gigantea*, *Pomacea canaliculate*, *Biomphalaria glabrata*, and *Aplysia californica* were summarized in Figure S4. The *P.*

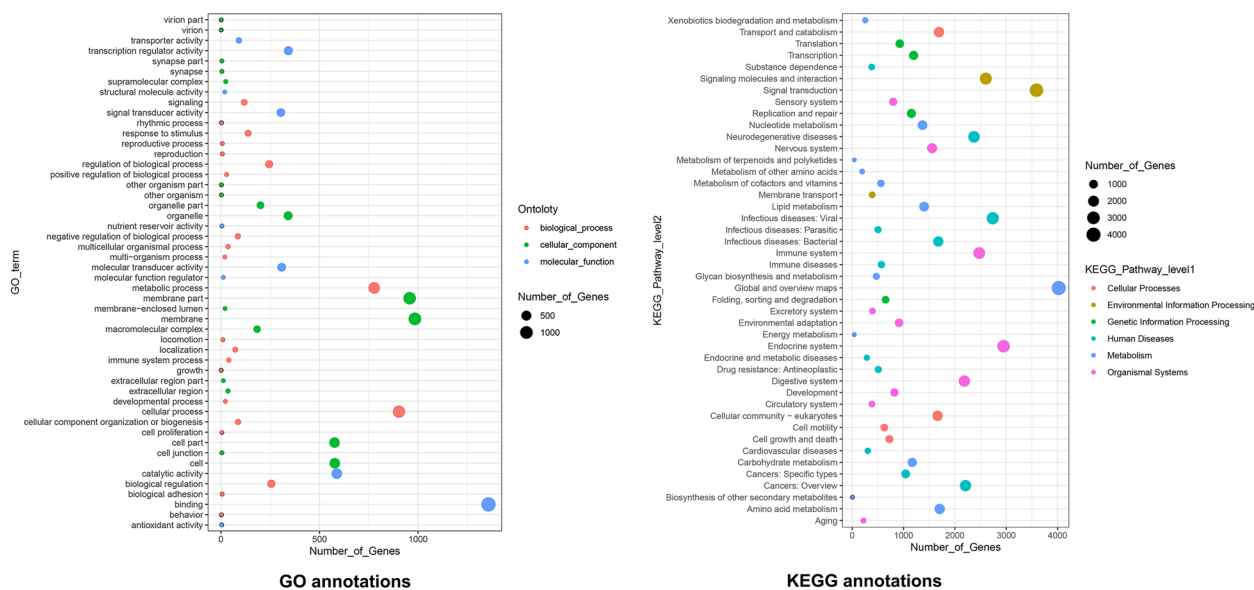


Fig. 3 GO and KEGG enrichment analysis of expanded gene families between deep-sea gastropod *P. buccinoides* and shallow sea gastropod *L. gigantea*

buccinoides-unique gene families were annotated to 21 GO terms and 32 KEGG pathways, mainly including cellular process, cellular community, and signal transduction (Fig. 2c, Table S15).

Compared with the shallow sea gastropod *L. gigantea*, 562 expanded and 15 contracted gene families were detected in *P. buccinoides*. The expanded genes in *P. buccinoides* were mainly enriched in the items such as environmental adaptation, sensory system, immune system process and antioxidant activity (Fig. 3, Table S16), and most of them were associated with olfactory and chemosensory, glutamate regulation, sulfur metabolism. Besides, GO and KEGG enriched in the contracted gene families of *P. buccinoides* compared with *L. gigantea*, mainly in association with metabolism, catalysis, and transport were showed in Figure S5, Table S17.

A phylogenetic tree was constructed using the single-copy gene families to investigate the phylogenetic evolutionary relationships among *P. buccinoides* and the other species. Phylogenetic analysis suggested that *P. buccinoides* diverged from *Pomacea canaliculata* approximately 218.9 million years ago (Mya). The estimated divergence time between *L. gigantea* and the branch group including *P. buccinoides* was 437.3 Mya. The gastropod showed an estimated divergence time of approximately 498.7 Mya from its sister group bivalve, and the time that cephalopod diverged from other lineages of mollusca was about 550.6 Mya. The mollusca were separated from brachiopoda, annelida and arthropoda about 608.2, 682.3 and 762.4 Mya (Fig. 2b).

Genomic basics of deep-sea dark, cold and toxic adaptation

Results from the genomic comparative analysis showed that the glutamate regulative and related genes including glutamine synthetase, glutamine γ -glutamyltransferase, and γ -aminobutyric acid receptor (GABA(A)) were expanded in *P. buccinoides* genome comparing with that in its relative specie living in shallow water (Fig. 3, Table S16). The olfactory/chemosensory gene families, such as olfactory specific medium-chain acyl CoA synthetase, serpentine type seven transmembrane GPCR chemoreceptor srw, G-protein coupled receptor were also significantly expanded (Fig. 3, Table S16). RNA-seqs of six tissues including hepatopancreas, foot, mantle, ctenidium, gonad and osphradium were also conducted to identify differentially expressed genes (DEGs) corresponding to olfactory and chemo/chemosensory receptor and sulfur metabolism related genes (Fig. 4, Table S18). A total of 17 olfactory/chemosensory genes, including ionotropic receptor 25a, zinc finger protein 62 homolog, transmembrane protein 256 homolog, voltage-gated hydrogen channel 1-like, G-protein coupled receptor, N-formyl peptide receptor 2-like, probable G-protein coupled receptor 139, and BTB/POZ domain containing protein KCTD7-like, were highly expressed in osphradium. These genes were enriched in five protein families including seven transmembrane odorant receptor, seven transmembrane chemosensory receptor, serpentine type seven transmembrane GPCR chemoreceptor srw, srg family chemoreceptor, and olfactory marker protein (Fig. 4a).

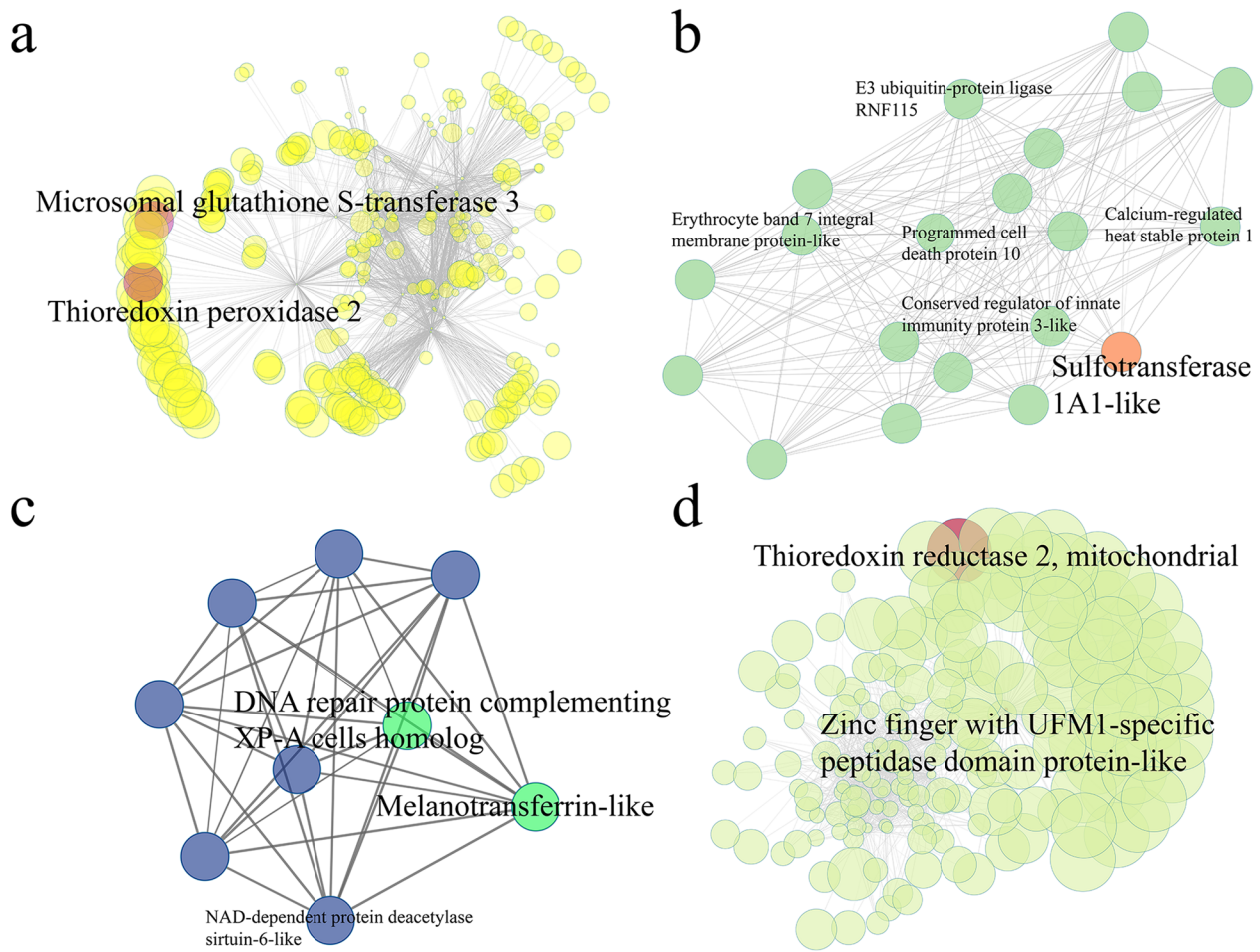


Fig. 5 The co-expression network in *P. buccinoides*. Node size represents the intramodular connectivity of a given gene. Gene names are showed for the top hub and key genes. **a** Hepatopancreas-related module (yellow, see Figure S9a). Key genes are labeled in red. **b** Hepatopancreas-related module (lightgreen, see Figure S9a). Sulfur metabolism-related gene is labeled in orange. **c** Hepatopancreas-related module (royalblue, see Figure S9a). Key genes are labeled in green. **d** Ctenidium-related module (greenyellow, see Figure S9b). Key genes are labeled in red

sirtuin-6-like. The zinc finger with UFM1-specific peptidase domain protein-like and thioredoxin reductase 2 were identified in Fig. 5d.

Discussion

The present study provides the first chromosome-level genome assembly of *P. buccinoides* and the results (Table S8, 9 and Fig. 1d, e) show that the assembled genome of *P. buccinoides* has high integrity and accuracy. The phylogeny of *P. buccinoides* (Fig. 2b) is consistent with previous researches [35–37], which confirms that other species of caenogastropoda are the most closely related species and they evolved together with other gastropod species, which diverged from other molluscan species including cephalopoda and bivalves a long time ago. In gastropod, the caenogastropoda diverged from panpulmonata and euopisthobranchia, and their ancestors diverged from patellogastropoda.

Compared with the reported genomes of gastropods, most of which are less than 2 Gb [35, 36, 38–40], the genome size of *P. buccinoides* is relatively large (Table S9). It might be due to the high proportion of repetitive elements (Table S10, 11). For example, the repeats of *Pomacea maculate* (genome size is 432 Mb), *P. canaliculate* (448 Mb), *Lanistes nyassanus* (510 Mb), and *Marisa cornuarietis* (536 Mb) are 20.5–30.8%, and that of *B. glabrata* (916 Mb), *Sinotaia purificata* (984 Mb), *Achatina fulica* (1.85 Gb), *Cepaea nemoralis* (3.5 Gb) and *Oreohelix idahoensis* (5.4 Gb) are 44.8%, 47.93%, 71%, 77% and 85.74%, respectively [36, 39, 41–44]. Overall, the chromosome-level genome assembly of *P. buccinoides* is a valuable resource for studying the adaptation to deep-sea cold seep environments.

Compared with the shallow sea gastropod *L. gigantea*, GO and KEGG enriched in the contracted gene families of *P. buccinoides* are mainly in association with

metabolism, catalysis, and transport (Figure S5, Table S17), which is consistent with previous reports that the metabolic rates of deep-sea life are orders of magnitude lower than those of life on Earth's surface [45]. Expansion of gene families play important roles in adaptation to environment including deep sea [46]. It is worth noting that the expanded genes in *P. buccinoides* are mainly associated with glutamate regulation, olfactory and chemosensory and sulfur metabolism (Table S16c).

The glutamate regulative and related genes (glutamine synthetase, glutamine γ -glutamyltransferase, and γ -aminobutyric acid receptor (GABA(A)) are noteworthy since they play important roles in hypoxia acclimation [47]. Glutamate uptake and transport by astrocytes is fundamentally important in the regulation of nervous system function, and hypoxia can suppress the glutamate transport in astrocytes [48]. Neurons and astrocytes can increase the release of glutamate, as well as improve glutamine and/or glutamate utilization to acclimate to the hypoxia condition [49]. Previous studies reveal that glutamate may be a conditionally essential amino acid resulting in enhanced tolerance to hypoxia, cold and amelioration of hypoxia-induced oxidative stress in rats [50, 51]. The γ -aminobutyric acid (GABA) exerts its inhibitory effects by binding to the GABA(A), and cDNA sequences encoding GABA(A) subunit has been cloned in mollusca [52]. GABA causes stress tolerance in plant cells and can be found in most eukaryotic organisms. Under hypoxia/anoxia stress, dual effects of GABA on both pH and TCA pathways play an important role in diminishing injury [53]. Increase of glutamate concentration can promote GABA synthesis in germinated soybean under hypoxia stress [54]. When exposure to hypoxia, GABA is neuroprotective to mature neurons of rat [55]. In the present study, the expansion of the glutamate regulating and related genes implies that the *P. buccinoides* has evolved to be more effective in glutamate production to ensure the physiological activities in a low-oxygen environment, which is of vital importance for its prosperity at the seep habitats. In addition, glutamate regulative genes are also responsive in cold acclimation. It is reported that the release of glutamate evoked by capsaicin is enhanced in spinal dorsal horn slices of repeated cold stress and adjuvant arthritic rats [56], and cold-induced glutamate release in vivo from the magnocellular region of the paraventricular nucleus is involved in ovarian sympathetic activation [57]. Similar to the response to hypoxia, the hosts intend to secrete more glutamate upon cold stress to sustain the physiological activities in mammals [58]. A substantial conversion of glutamate to GABA is proportional to the severity of cold stress and GABA accumulates to a higher extent when exposed to lower temperature. GABA is suspected to involve in

tolerance to low temperature [59]. Collectively, results in the present study indicate that the deep-sea snail *P. buccinoides* might be able to produce more glutamate and GABA to adapt to the relatively hypoxic and cold environment at the deep-sea cold seep.

Throughout the animal kingdom, chemical senses are one of the primary means by which organisms make sense of their environment [60]. The size and diversity of chemoreceptors that mediate the transduction of chemical signals can reflect the niche inhabited by the organisms. For example, *Caenorhabditis elegans* requires an abundance of chemoreceptors to navigate and interpret its nutrient-rich living environments, because they spend more foraging time compared to parasitic nematodes [61]. Especially, previous researches show that expansions of G-protein coupled receptor are correlated with environmental adaptations by enabling the evolution of sensory functions in some invertebrate species [62–64]. Therefore, the expansion of olfactory/chemosensory genes in the present study suggests that *P. buccinoides* might rely mainly on the sense of odorants or chemicals for predation and orientation to survive at the dark seep environment. Moreover, the transcriptomic survey revealed that olfactory/chemosensory genes also showed obviously high expression levels in osphradium (Fig. 4a, Table S18). Osphradium is a single or paired chemosensory organ connected with one of the visceral ganglia and situated near the gill of most aquatic molluscs [65, 66]. The olfactory organ is extremely important for molluscan environmental adaptation since it helps them to locate food, nest sites, and escape dangers. For example, nautilus, a well-known “living fossil”, possesses a pair of olfactory organs called rhinophores that are similar to the olfactory organs in octopus and other cephalopods [66]. The rhinophores serve primarily in distance chemoreception during tracking [67]. Odor on a variety of spatial scales is an important information source to nautilus in their complex coral-reef environment. Olfactory memory for predator and prey is also of great importance to their survival in the wild [68]. The osphradium also possess carrion (or prey) locating function in some gastropods [65]. And, the rhodopsin G-protein coupled receptors are highly expressed in sensory epithelia microdissected from *Aplysia* rhinophore, which are involved in its chemical detection [69]. Thus, the present study indicates that osphradium might be the olfactory/chemosensory organ of *P. buccinoides*. Since the cold seeps are usually characterized as sulfate-rich, hypoxic, dark, organic enrichments such as whale skeletons are released into the ocean, the expansion of olfactory/chemosensory genes in the genome and the relatively high mRNA expression in osphradium suggest that *P. buccinoides* has evolved a highly developed and sensitive olfactory organ

comparing with their relative specie living in the shallow water, which might contribute to their orientation, predation to adapt to the cold seep environment.

H₂S is toxic and creates extreme environmental conditions [70–72]. The toxicity of H₂S is often exhibited through inhibition of cytochrome c oxidase (COX) in the mitochondrial respiratory chain to inhibit ATP production [73]. Exposure to environmental H₂S will impact the capability of survival and reproduction of an organism [74]. Deep-sea cold seeps are rich in sulfides and heavy metals [2, 11–13, 75]. The tubeworms, deep-sea mussels, and deep-sea snails have thrived at the cold-seep environment by evolving strategies to acclimate to the toxic conditions [75]. Based on recent studies, three mechanisms of H₂S tolerance have been revealed in different organisms. First, some organisms can minimize the flux rich in H₂S into the body [76]. Second, the H₂S tolerance can be achieved through the modifications of toxicity targets that make them less sensitive to adverse consequences caused by elevated endogenous concentration in the face of continuous influx from the environment [77]. Third, the organisms can tolerate to H₂S by performing detoxification mediated by a series of enzymes such as SQR, ETHE1 and TST [78, 79]. In the present study, SQR, ETHE1, and thiosulfate sulfurtransferase (TST) were highly expressed in ctenidium and hepatopancreas (Fig. 4b, Table S18), and they are responsible for detoxification of H₂S [79]. Ctenidium is a respiratory organ found in many molluscs, which is responsible for gas (O₂ and H₂S) exchange [80]. Hepatopancreas is regarded as a critical organ for metabolism and detoxification in molluscs [81]. Therefore, the current results suggest that *P. buccinoides* might have the strong capability in tolerating sulfide by conducting H₂S detoxification, and its ctenidium and hepatopancreas are critical tissues for such process. In addition, glutathione S-transferase is a well-known antioxidant and detoxification enzyme, and it contribute to the metabolism of drugs, pesticides and other xenobiotics [82, 83]. The glutathione S-transferase and thioredoxin peroxidases also take part in response to onslaught of oxidants and function in maintaining efficient antioxidant defense in invertebrates [84–86]. In the present study, as the hub genes of hepatopancreas and ctenidium-related modules in gene co-expression networks (Fig. 5), microsomal glutathione S-transferase 3 and thioredoxin peroxidase 2 are presumed to be key genes of antioxidant and detoxification in *P. buccinoides*. Besides, sulfotransferase in the network utilizes 3'-phospho-5'-adenylyl sulfate (PAPS) as a sulfonate donor [87] to participate in sulfur metabolism, while the zinc finger with UFM1-specific peptidase domain protein-like (deubiquitylating enzyme) and thioredoxin reductase 2 are speculated to play important roles in

DNA damage responses and repair [88]. In summary, it is quite possible that *P. buccinoides* adapts to the sulfide-rich cold seep environment by enhancing the H₂S detoxification activities in ctenidium and hepatopancreas. It has been reported that there are no symbiotic bacteria in ctenidium of *P. buccinoides* [32], however, whether the sulfide detoxification is performed by endosymbiotic bacteria in other tissues deserves further investigation in the future.

Conclusions

The first chromosome-level genome assembly of the deep-sea snail *P. buccinoides* was constructed in the present study. The genome size of *P. buccinoides* is relatively large (about 2.1 Gb, scaffold N50=60.0 Mb) compared with the other known snail genomes, which might be due to the high proportion of repetitive elements. The glutamate regulative and related gene family was found to be expanded, which might contribute to the acclimation to hypoxia and coldness. The relatively high mRNA expression of the olfactory and chemosensory related genes in osphradium indicates that *P. buccinoides* might have evolved a highly developed and sensitive olfactory organ for its orientation and predation. More importantly, results of the transcriptomic and network analysis showed that *P. buccinoides* might have evolved sulfite tolerance mechanism by performing H₂S detoxification in ctenidium and hepatopancreas. The present study provides insights into the mechanisms of adaptation of gastropod to the dark, hypoxic and H₂S-rich environment in the deep-sea cold seeps.

Methods

Sample collection and sequencing

The deep-sea snails *P. buccinoides* (Fig. 1a) were collected at cold seep site F (22°06'N, 119°17'E), which was located on the continental slope of the South Sea of China during the expedition cruise of the R/V Kexue in 2018 (Fig. 1b). These snails were found 1,119 m beneath the sea surface with temperatures of 3.35–3.89 °C, salinity of 34.53–35.54 psu, and dissolved oxygen of 3.01–3.18 mg/L. In the bottom water of the chemosynthetic communities, the H₂S level increased remarkably and the highest H₂S level (~1940 μM) among all seawater samples was in the bottom water above the reduced sediments [22]. The samples were preserved at -80 °C. All of the experiments were performed following the animal ethics guidelines approved by the Ethics Committee of Dalian Ocean University.

Genomic DNA was extracted from hepatopancreas and muscle of snails using the modified phenol/chloroform method [89]. One sequencing library with insert size of 350 bp was generated using Truseq Nano DNA HT

Sample Preparation Kit (Illumina, San Diego, CA, USA) following manufacturer's recommendations. PacBio library with insert size of 20 kb was constructed using PacBio single molecule, real-time long reads sequencing technology (SMRT) SMRTbell Template Prep Kits (PacBio, Menlo Park, CA, USA). One short insert size (350 bp) library was sequenced on NovaSeq 6000 platform (Illumina, San Diego, CA, USA) using whole-genome shotgun sequencing (WGS) strategy. The raw data were generated and filtered by SOAPFILTER v2.2, a software in the SOAPdenovo package [90]. The long insert size (20 kb) library was sequenced on PacBio Sequel instrument (PacBio, Menlo Park, CA, USA) to obtain long reads (polymerase reads) data. After removing the adapters, the polymerase reads were partitioned to form subreads (Pacific Biosciences Terminology).

A total of 1000 μL (10.3 $\text{ng}/\mu\text{L}$) hepatopancreas sample of *P. buccinoides* from the same collection lot was treated [91] and the Hi-C libraries were constructed with NEBNext Ultra II DNA library Prep Kit for Illumina (NEB, Ipswich, MA, USA). The target fragments were captured with Dynabeads MyOne Streptavidin C1 (Thermo Fisher Scientific, Inc., Waltham, MA, USA). After that, the chimeric fragments were amplified with NEBNext Ultra II DNA library Prep Kit (NEB, Ipswich, MA, USA). Two replicates of libraries were generated and the libraries were sequenced on Illumina Novaseq 6000 instrument (Illumina, San Diego, CA, USA).

To perform the single-molecule long-read transcriptome sequencing with SMRT, the hepatopancreas, foot, mantle, ctenidium, gonad, and osphradium tissues were harvested. The sample is precious in the present study and three biological and technical replicates should be performed generally. RNA from different tissue samples was isolated using Trizol reagent (Sangon, Shanghai, China). By using NEBNext Ultra RNA Library Prep Kit for Illumina (NEB, MA, USA), short read RNA-Seq libraries were prepared and then sequenced on NovaSeq 6000 platform (Illumina, San Diego, CA, USA). With the Clontech SMARTer PCR cDNA Synthesis Kit (Clontech, CA, USA), one SMART bell library was constructed. The SMRT sequencing was performed on Pacific Bioscience Sequel System (Pacific Biosciences, CA, USA).

Assemblies of genome and transcriptome

All cleaned reads from short insert library were assembled using PLATANUS v1.2.4 [33] with parameter “-k 27” to obtain a de Bruijn graph assembly. Subsequently, the DBG2OLC was employed to align the de Bruijn graph assembly upon the PacBio reads for further construction of contigs [34]. Three rounds of mapping were performed with MINIMAP v2.1 [92] and polished with RACON v1.3.1 [93] to construct consensus contigs. Then, BWA

v0.7.15 [94] and PILON v1.22 [95] were used to polish the assembly one round with 350 bp library Illumina paired-end reads. Completeness of the final assembly at contig level was assessed using BUSCO v3.1.0 [96]. The mollusca_odb10 [97] orthologues gene set was used as the BUSCO reference.

The raw data generated from Hi-C library were filtered with TRIMMOMATIC v0.39 [98], and the clean data were aligned against the draft genome using JUICER v1.6.2 [99] with the default parameters. The Hi-C contacts without duplicates were used to assist genomic assembly by 3D-DNA v180114 [100]. The heatmap of chromosome interactions was constructed with 3D-DNA v180114 [100] to visualize the contact intensity among Chrs. The scaffolds were assembled and the obtained genome at contig-level was located onto the Chrs. By using JUICEBOX v1.8.8 [101], the Hi-C contact map was visualized and the extensive manual curation was performed to ensure the scaffolds within the same pseudo-chromosomal linkage group to meet the Hi-C linkage characteristics. The clustered contigs and mis-joins were ordered, oriented and fixed.

By using SMRTLINK v6.0 software (Pacific Biosciences, CA, USA) (<https://www.pacb.com/support/software-downloads>), the non-chimeric circular consensus sequences (CCSs) were generated from subread BAM files and a ccs.bam file was obtained. By performing the isoform level Iterative Clustering for Error Correction (ICE), consensus isoforms were identified from Full-length non-chimeric (FLNC) then and they were further polished with QUIVER v2.2.2 [102]. To correct nucleotide indels and mismatches in consensus reads resulting in corrected isoforms, the Illumina RNA-Seq data with same samples were used by LoRDEC v0.7 [103]. By using CD-HIT v4.6.8 [104], the redundancies in corrected consensus reads were removed and final non-redundancy transcripts for the subsequent analysis were obtained.

By using SOAPnuke v1.5.6 [105], the low-quality reads (quality score ≤ 20) were removed. The clean transcriptome reads were assembled using TRINITY v2.5.1 [106]. By using BOWTIE v1.1.1 [107], sequences obtained by Illumina-Seq were aligned to the transcript obtained by SMRT sequencing, regarded as a reference. The gene expression levels of each sample were estimated with Expectation–Maximization (RSEM) v1.3.0 [108]. By using the GSeq R package v1.10.0 [109], GO enrichment analysis of DEGs were performed. GO terms of DEGs with corrected P -value < 0.05 were considered enriched significantly. The statistical significant enrichments of DEGs in KEGG pathways were determined by KOBAS v3.0 [110] with the P -value < 0.05 .

Structure and functional annotation

Prior to gene prediction using the assembled genome, de novo and homology-based prediction were used to annotate repeat elements. The local de novo repeat reference library was generated using LTR FINDER v1.0.6 [111], REPEATMASKER v4.0.6 [112] and REPEATMOD-ELER v1.08 [113]. Subsequently, the assembled genome was aligned against this reference to produce the de novo predicted repeat elements. For the homology-based prediction, REPEATPROTEINMASK v4.06 [112], REPEATMASKER v4.0.6 [112] and TANDEM REPEATS FINDER (TRF) v4.07 [114] were run to identify, classify and mask repeats with REPEATMASKER v21.01 [115] in the *P. buccinoides* genome. Finally, the non-redundant results were generated by integrating the data from two predictions.

The assembly was annotated with three different strategies. AUGUSTUS v2.5 [116], GENSCAN v1.0 [117] and SNAP v2.0 [118] were employed for the first ab initio gene prediction method and the genome assembly was masked to exclude the repetitive elements firstly. Homologous-gene-based annotation was the second method. The protein sequences of California sea hare (*A. californica*), a freshwater snail (*B. glabrata*) [119], roundworm (*C. elegans*) [120], Eastern oyster (*Crassostrea virginica*) [121], Pacific oyster (*C. gigas*) [122], fruit fly (*Drosophila melanogaster*) [123], owl limpet (*L. gigantea*) [124], Yesso scallop (*M. yessoensis*) [125], California two-spot octopus *O. bimaculoides* [126] and golden apple snail *P. canaliculata* [35] were downloaded from the NCBI database. The TBLASTN in Basic local alignment search tool (BLAST) v2.2.26 program [127] was used to search for best-hit alignments of these proteins in the assembled *P. buccinoides* genome with E-value cutoff of 10^{-5} . Then the potential gene structure of each best-hit alignment was identified with GENEWISE v2.4.1 [128]. The transcriptomic data generated from mantle, ctenidium and 6 tissues were mapped onto the assembly to aid gene annotation. The final resultant was obtained using MAKER v2.31.8 [129].

By using BLAST v2.2.26 [127], the functional motifs and domains were identified by searching the predicted genes of *P. buccinoides* in NCBI non-redundant protein sequences (NCBI-Nr) [130], Swiss-Prot [131], Interpro [132], Clusters of Orthologous Groups (COG) [133], TrEMBL, KEGG [134] and GO [135] public functional databases. By using BLAST software v2.7.1 [127] under a threshold E-value $\leq 1e-5$, corrected isoforms of long read transcripts were searched against NCBI-Nr [130], NCBI-Nt, Swiss-Prot [131], KOG/COG [136, 137] and KEGG v2015_10_10 [134]. The Protein family (Pfam) database [138] was searched by HMMER v3.1 [139], and the Pfam

accession numbers were converted to GO terms by using 'pfam2go' mapping [135].

Phylogenetic analysis of the genome

The complete gene set of *P. buccinoides* and other 10 representative species including *A. californica* (GCF_000002075.1), *B. glabrata* (GCF_000457365.1) [140], *C. gigas* (GCF_000297895.1) [122], *D. melanogaster* (GCF_000001215.4) [123], *Helobdella robusta* (GCF_000326865.1) [124], *Lingula anatina* (GCF_001039355.2) [37], *L. gigantea* (GCF_000327385.1) [124], *M. yessoensis* (GCF_002113885.1) [125], *O. bimaculoides* (GCF_001194135.1) [126], and *P. canaliculata* (GCF_003073045.1) [35] were downloaded from NCBI. To check the homology and generate a sequence similarity matrix, the whole-genome gene sets were aligned with BLAST v2.6.0 [127]. ORTHOMCL v1.4 [141] with 1.5 inflation index was employed to distinguish gene families from the sequence similarity matrix. MUSCLE v3.8.31 [142] was used to determine homologous genes and identified single-copy orthologs. By using PHYML v3.0 [143], the phylogenetic topology with the maximum likelihood (ML) method was estimated with gamma distribution across aligned sites and HKY85 substitution model to construct the phylogenetic tree. *D. melanogaster* was used as the outgroup. To estimate divergence times among the *P. buccinoides* and the other molluscan species, the MCMCTREE in PAML v4.4 [144] was employed. The neutral evolutionary rate and species divergence time were estimated by adopting the Bayesian molecular dating [145]. Five reference divergence time points retrieved from the TimeTree database [146] were used to calibrate the phylogenetic tree [147–151].

Expansion and contraction of gene families

The program CAFÉ v2.1 [152] was adopted by determining the evolutionary dynamics of gene families to identify gene family changes between the deep-sea gastropod *P. buccinoides* and shallow sea gastropod *L. gigantea*, especially expansion and contraction of gene ortholog clusters. The gene families presented uniquely in *P. buccinoides* were also screened. Venn diagram was drawn with VENNPAINTEER v1.2.0 [153]. KEGG and GO analysis of the gene families exclusively presented and specifically expanded and contracted in the *P. buccinoides* were conducted as that in functional annotation [127, 134, 135].

Network and other analysis

By using 6 transcriptome datasets of different tissues (hepatopancreas, foot, mantle, ctenidium, gonad, and osphradium) with WGCNA v1.70–3 [154], co-expression gene networks were constructed. Cytoscape

software v3.9.1 was used to visualize the networks [155]. Block-wise network construction and consensus module detection methods were adopted, with the parameters of soft-thresholding power = 14, maximum block size = 2000 and minimum module size = 30. Module eigengene E was calculated to identify the tissue-related modules. The hub genes in a given module was measured by its connection strength with other genes in the module, and was determined by intramodular connectivity [154].

By using MicroSatellite identification tool (MISA) v1.0 [156], the microsatellites as well as compound microsatellites were identified and localized. The Animal Transcription Factor Data Base v2.0 (animalTFDB) [157] were used to predict the transcription factors. The analysis methods of long non-coding RNA (lncRNA) were performed with Coding Potential Calculator (CPC) v0.9 [158], Coding-Non-Coding Index (CNCI) v2.0 [159], predictor of long non-coding RNAs and messenger RNAs based on an improved k-mer scheme (PLEK) v1.2 [160] and Pfam v1.6 [138].

Supplementary Information

The online version contains supplementary material available at <https://doi.org/10.1186/s12864-023-09760-0>.

Additional file 1: Figure S1. Divergence distribution of transposable elements (TEs) in the *P. buccinoides* genome. **Figure S2.** Chromosomal syntenic relationships. **Figure S3.** Distribution of genes in 11 different species. **Figure S4.** Venn diagram of gene families specific to *P. buccinoides*. **Figure S5.** GO and KEGG enrichment analysis of contracted gene families between deep-sea gastropod *P. buccinoides* and shallow sea gastropod *L. gigantea*. **Figure S6.** Length distribution of unigenes in transcriptome. **Figure S7.** Length distribution of coding DNA sequence (CDS) in transcriptome. **Figure S8.** Expression of sulfur metabolism related genes in different tissues. **Figure S9.** Module eigengene E of gene co-expression networks. **Figure S10.** Gene dendrograms and module colors of gene co-expression networks. **Figure S11.** The distribution of transcript length of lncRNAs and mRNAs. **Figure S12.** The distribution of SSR motifs in transcriptome. **Figure S13.** The numbers of transcription factors involved in the top transcription factor families of transcriptome. **Table S1.** Illumina statistics of the genome sequencing data of *P. buccinoides*. **Table S2.** PacBio statistics of the genome sequencing data of *P. buccinoides*. **Table S3.** Hi-C statistics of the genome sequencing data of *P. buccinoides*. **Table S4a.** Statistics of *P. buccinoides* Illumina transcriptome reads (Raw data). **Table S4b.** Statistics of *P. buccinoides* Illumina transcriptome reads (Clean data). **Table S5.** Statistics of *P. buccinoides* Iso transcriptome reads. **Table S6.** Transcriptome sequencing data of *P. buccinoides* (for aiding gene annotation). **Table S7.** Summary statistics of the genome sequencing data of *P. buccinoides*. **Table S8.** Contig assembly of the *P. buccinoides* genome using Illumina and PacBio reads. Related to Figure 1e. **Table S9.** Summary statistics of the *P. buccinoides* chromosomal-level genome assembly. Related to Figure 1d, e. **Table S10.** Prediction of repeat elements in the *P. buccinoides* genome. Related to Figure 1c, S1. **Table S11.** Categories of repeat elements predicted in the *P. buccinoides* genome. Related to Figure 1c, S1. **Table S12.** Prediction of gene structure in *P. buccinoides* genomes. Related to Figure 1e. **Table S13.** Functional annotation of the predicted protein-coding genes in the *P. buccinoides*. Related to Figure 2a. **Table S14.** Statistics of gene families in 11 examined species.

Additional file 2: Table S15a. GO enrichment of unique gene families in *P. buccinoides* compared with seven other molluscan species. Related to Figure 2c. **Table S15b.** KEGG enrichment of unique gene families in

P. buccinoides compared with seven other molluscan species. Related to Figure 2c.

Additional file 3: Table S16a. Enriched GO terms of expanded genes in the deep-sea gastropod *P. buccinoides* compared to shallow sea *L. gigantea*. Related to Figure 3. **Table S16b.** Enriched KEGG pathways of expanded genes in the deep-sea gastropod *P. buccinoides* compared to shallow sea *L. gigantea*. Related to Figure 3. **Table S16c.** Targeted expanded genes in the deep-sea gastropod *P. buccinoides* compared to shallow sea *L. gigantea*. Related to Figure 3.

Additional file 4: Table S17a. Enriched GO terms of contracted genes in the deep-sea gastropod *P. buccinoides* compared to shallow sea *L. gigantea*. Related to Figure S5. **Table S17b.** Enriched KEGG pathways of contracted genes in the deep-sea gastropod *P. buccinoides* compared to shallow sea *L. gigantea*. Related to Figure S5.

Additional file 5: Table S18. RNA-seq differentially expressed genes (DEGs) in different tissues of *P. buccinoides*. Related to Figure 4, 5.

Additional file 6: Table S19. SSRs in RNA-seq of *P. buccinoides*. Related to Figure S12.

Additional file 7: Table S20a. Transcription factor in RNA-seq of *P. buccinoides*. Related to Figure S13. **Table S20b.** Statistics of transcription factor in *P. buccinoides* RNA-seq. Related to Figure S13.

Additional file 8: Table S21. Functional annotation of *P. buccinoides*.

Acknowledgements

We are grateful to all the laboratory members for their technical advice and helpful discussions.

Authors' contributions

Conceive and design of the study: Z.Q.L., L.S.S., L.L.W., H.C. Data analyses: Y.T.H., C.B. Samples collecting: H.C., M.X.W., C.L. Manuscript writing: Z.Q.L., Y.T.H. Manuscript revision: Z.Q.L., L.L.W., L.S.S. All authors reviewed the manuscript.

Funding

This research is supported by National Key R&D Program (2018YFC0310802), earmarked fund for CARS-49, the fund for Outstanding Talents and Innovative Team of MARA, Distinguished Professor in Liaoning (XLYC1902012), the innovation team of Aquaculture Environment Safety from Liaoning Province (LT202009), and Dalian High Level Talent Innovation Support Program (2022RG14).

Availability of data and materials

The datasets generated during the current study are available in the NCBI under accession number PRJNA700822 (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA700822>).

Declarations

Ethics approval and consent to participate

All of the experiments were performed following the animal ethics guidelines approved by the Ethics Committee of Dalian Ocean University.

Consent for publication

Not applicable.

Competing interests

The authors declare no competing interests.

Author details

¹Liaoning Key Laboratory of Marine Animal Immunology, Dalian Ocean University, Dalian 116023, China. ²Functional Laboratory of Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266235, China. ³Liaoning Key Laboratory of Marine Animal Immunology and Disease Control, Dalian Ocean

University, Dalian 116023, China. ⁴Dalian Key Laboratory of Aquatic Animal Disease Prevention and Control, Dalian Ocean University, Dalian 116023, China. ⁵Laboratory of Aquatic Genomics, College of Life Sciences and Oceanography, Shenzhen University, Shenzhen 518060, China. ⁶College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China. ⁷Center of Deep Sea Research, and CAS Key Laboratory of Marine Ecology and Environmental Sciences, Institute of Oceanology, Chinese Academy of Sciences, Qingdao 266071, China.

Received: 9 February 2023 Accepted: 22 October 2023

Published online: 10 November 2023

References

- Levin LA. Ecology of cold seep sediments: interactions of fauna with flow, chemistry and microbes. In: *Oceanogr Mar Biol. Florida: CRC Press; 2005. p. 11–56.*
- Van Dover CL, German CR, Speer KG, Parson LM, Vrijenhoek RC. Evolution and biogeography of deep-sea vent and seep invertebrates. *Science*. 2002;295(5558):1253–7. <https://doi.org/10.1126/science.1067361>.
- Hourdez S, Lallier FH. Adaptations to hypoxia in hydrothermal-vent and cold-seep invertebrates. *Rev Environ Sci Biotechnol*. 2007;6(1–3):143.
- Dong X, Rattray JE, Campbell DC, Webb J, Chakraborty A, Adebayo O, et al. Thermogenic hydrocarbon biodegradation by diverse depth-stratified microbial populations at a Scotian Basin cold seep. *Nat Commun*. 2020;11(1):1–14.
- Orphan VJ, House CH, Hinrichs KU, McKeegan KD, DeLong EF. Multiple archaeal groups mediate methane oxidation in anoxic cold seep sediments. *Proc Natl Acad Sci*. 2002;99(11):7663–8. <https://doi.org/10.1073/pnas.072210299>.
- Huang Z, Brooke BP, Harris PT. A new approach to mapping marine benthic habitats using physical environmental data. *Cont Shelf Res*. 2011;31(2):4–516.
- Gobler CJ, Baumann H. Hypoxia and acidification in ocean ecosystems: coupled dynamics and effects on marine life. *Biol Lett*. 2016;12(5):20150976. <https://doi.org/10.1098/rsbl.2015.0976>.
- Brown A, Thatje S. Explaining bathymetric diversity patterns in marine benthic invertebrates and demersal fishes: physiological contributions to adaptation of life at depth. *Biol Rev*. 2014;89(2):406–26.
- Burton EA, Walter LM. Relative precipitation rates of aragonite and mg calcite from seawater: temperature or carbonate ion control? *Geology*. 1987;15(2):111–4.
- Boetius A, Ravensschlag K, Schubert CJ, Rickert D, Widdel F, Gieseke A, Amann R, Jørgensen BB, Witte U, Pfannkuche O. A marine microbial consortium apparently mediating anaerobic oxidation of methane. *Nature*. 2000;407(6804):623–6. <https://doi.org/10.1038/35036572>.
- Maignien L, Parkes RJ, Cragg B, Niemann H, Knittel K, Coulon S, Akhmetzhanov A, Boon N. Anaerobic oxidation of methane in hypersaline cold seep sediments. *FEMS Microbiol Ecol*. 2013;83(1):214–31. <https://doi.org/10.1111/j.1574-6941.2012.01466.x>.
- Arvidson RS, Morse JW, Joye SB. The sulfur biogeochemistry of chemosynthetic cold seep communities, Gulf of Mexico, USA. *Mar Chem*. 2004;87(3–4):97–119.
- Li WL, Dong X, Lu R, Zhou YL, Zheng PF, Feng D, et al. Microbial ecology of sulfur cycling near the sulfate-methane transition of deep-sea cold seep sediments. *Environ Microbiol*. 2021;23(11):6844–58.
- Sun QL, Zhang J, Wang MX, Cao L, Du ZF, Sun YY, Sun Q-L, Zhang J, Wang M-X, Cao L, Du Z-F, Sun Y-Y, Liu S-Q, Li C-L, Sun Li. High-throughput sequencing reveals a potentially novel *Sulfurovum* species dominating the microbial communities of the seawater-sediment interface of a deep-sea cold seep in South China Sea. *Microorganisms*. 2020;8(5):687. <https://doi.org/10.3390/microorganisms8050687>.
- Lazar CS, Dinasquet J, Pignet P, Prieur D, Toffin L. Active archaeal communities at cold seep sediments populated by Siboglinidae tubeworms from the Storegga Slide. *Microb Ecol*. 2010;60(3):516–27. <https://doi.org/10.1007/s00248-010-9654-1>.
- Gardiner SL, McMullin E, Fisher CR. Seepiophila Jonesi, a new genus and species of vestimentiferan tube worm (Annelida: Pogonophora) from hydrocarbon seep communities in the Gulf of Mexico. *Proc Biol Soc Wash*. 2001;114(3):694–707.
- MacAvoy S, Carney R, Morgan E, Macko S. Stable isotope variation among the mussel *Bathymodiolus Childressi* and associated heterotrophic fauna at four cold-seep communities in the Gulf of Mexico. *J Shellfish Res*. 2008;27(1):147–51.
- Danovaro R, Corinaldesi C, Dell'Anno A, Snelgrove PV. The deep-sea under global change. *Curr Biol*. 2017;27(11):R461–465.
- Wharton DA. Cold tolerance. In: *Molecular and physiological basis of nematode survival*. Wallingford UK: CAB International; 2011. p. 182–204.
- Loomis SH. Freezing tolerance of marine invertebrates. *Oceanogr Mar Biol*. 1995;33:337–50.
- Decelle J, Andersen AC, Hourdez S. Morphological adaptations to chronic hypoxia in deep-sea decapod crustaceans from hydrothermal vents and cold seeps. *Mar Biol*. 2010;157(6):1259–69. <https://doi.org/10.1007/s00227-010-1406-8>.
- Cao L, Lian C, Zhang X, Zhang H, Wang H, Zhou L, Cao L, Lian C, Zhang X, Zhang H, Wang H, Zhou Li, Wang M, Chen H, Luan Z, Li C. In situ detection of the fine scale heterogeneity of active cold seep environment of the Formosa Ridge, the South China Sea. *J Mar Syst*. 2021;218:103530 <https://doi.org/10.1016/j.jmarsys.2021.103530>
- Joyner-Matos J, Predmore BL, Stein JR, Leeuwenburgh C, Julian D. Hydrogen sulfide induces oxidative damage to RNA and DNA in a sulfide-tolerant marine invertebrate. *Physiol Biochem Zool*. 2010;83(2):356–65.
- Vetter R, Wells M, Kurtsman AL, Somero G. Sulfide detoxification by the hydrothermal vent crab *Bythograea thermophilum* and other decapod crustaceans. *Physiol Zool*. 1987;60(1):121–37.
- Oeschger R, Vetter RD. Sulfide detoxification and tolerance in *Halicryptus spinulosus* (Priapulida): a multiple strategy. *Mar Ecol Prog Ser*. 1992;86:167–79.
- Chou PH, Hu MY, Guh YJ, Wu GC, Yang SH, Tandon K, et al. Cellular mechanisms underlying extraordinary sulfide tolerance in a crustacean holobiont from hydrothermal vents. *Proc Royal Soc B*. 2023;290(1990):20221973.
- Oh DH, Dassanayake M, Bohnert HJ, Cheeseman JM. Life at the extreme: lessons from the genome. *Genome Biol*. 2013;13(3):1–9.
- Shaikhutdinov N, Gusev O. Chironomid midges (Diptera) provide insights into genome evolution in extreme environments. *Curr Opin Insect Sci*. 2022;49:101–7.
- Zhou L, Cao L, Wang X, Wang M, Wang H, Zhong Z, et al. Metal adaptation strategies of deep-sea *Bathymodiolus* mussels from a cold seep and three hydrothermal vents in the West Pacific. *Sci Total Environ*. 2020;707:136046. <https://doi.org/10.1016/j.scitotenv.2019.136046>.
- Modica MV, Holford M. The Neogastropoda: evolutionary innovations of predatory marine snails with remarkable pharmacological potential. In: *Evolutionary Biology-Concepts, Molecular and Morphological Evolution*. Berlin: Springer; 2010. p. 249–70.
- Wang X. Nutritional sources analysis and the heavy-metal enrichment of the macrofauna from the deep-sea chemotrophic ecosystem. Beijing: University of Chinese academy of sciences; 2018.
- Fujikura K, Sasaki T, Yamanaka T, Yoshida T. Turrids whelk, phymorhynchus buccinoides feeds on *Bathymodiolus* mussels at a seep site in Sagami Bay, Japan. *Plankton Benthos Res*. 2009;4(1):23–30. <https://doi.org/10.3800/pbr.4.23>.
- Kajitani R, Toshimoto K, Noguchi H, Toyoda A, Ogura Y, Okuno M, et al. Efficient de novo assembly of highly heterozygous genomes from whole-genome shotgun short reads. *Genome Res*. 2014;24(8):1384–95. <https://doi.org/10.1101/gr.170720.113>.
- Ye C, Hill CM, Wu S, Ruan J, Ma ZS. DBG2OLC: efficient assembly of large genomes using long erroneous reads of the third generation sequencing technologies. *Sci Rep*. 2016;6:31900.
- Liu C, Zhang Y, Ren Y, Wang H, Li S, Jiang F, et al. The genome of the golden apple snail *Pomacea canaliculata* provides insight into stress tolerance and invasive adaptation. *Gigascience*. 2018;7(9):gij101.
- Jin W, Cao XJ, Ma XY, Lv GH, Xu GC, Xu P, et al. Chromosome-level genome assembly of the freshwater snail *Bellamya purificata* (Caenogastropoda). *Zool Res*. 2022;43(4):683.

37. Luo YJ, Takeuchi T, Koyanagi R, Yamada L, Kanda M, Khalturina M, Luo Y-J, Takeuchi T, Koyanagi R, Yamada L, Kanda M, Khalturina M, Fujie M, Yamasaki S-I, Endo K, Satoh N. The Lingula genome provides insights into brachiopod evolution and the origin of phosphate biomineralization. *Nat Commun.* 2015;6(1):8301. <https://doi.org/10.1038/ncomms9301>.
38. Gomes-dos-Santos A, Lopes-Lima M, Castro LFC, Froufe E. Molluscan genomics: the road so far and the way forward. *Hydrobiologia.* 2020;847(7):1705–26. <https://doi.org/10.1007/s10750-019-04111-1>.
39. Guo Y, Zhang Y, Liu Q, Huang Y, Mao G, Yue Z, et al. A chromosomal-level genome assembly for the giant African snail *Achatina fulica*. *Gigascience.* 2019;8(10):giz124. <https://doi.org/10.1093/gigascience/giz124>.
40. Masonbrink RE, Purcell CM, Boles SE, Whitehead A, Hyde JR, Seetharam AS, Masonbrink RE, Purcell CM, Boles SE, Whitehead A, Hyde JR, Seetharam AS, Severin AJ. An annotated genome for *Haliotis rufescens* (red abalone) and resequenced green, pink, pinto, black, and white abalone species. *Genome Biol Evol.* 2019;11(2):431–8. <https://doi.org/10.1093/gbe/evz006>.
41. Sun J, Mu H, Ip JC, Li R, Xu T, Accorsi A, et al. Signatures of divergence, invasiveness, and terrestrialization revealed by four apple snail genomes. *Mol Biol Evol.* 2019;36(7):1507–20.
42. Adema CM, Hillier LW, Jones CS, Loker ES, Knight M, Minx P, et al. Whole genome analysis of a schistosomiasis-transmitting freshwater snail. *Nat Commun.* 2017;8(1):15451. <https://doi.org/10.1038/ncomms15451>.
43. Linscott TM, González-González A, Hirano T, Parent CE. De novo genome assembly and genome skims reveal LTRs dominate the genome of a limestone endemic Mountainsnail (*Oreohelix idahoensis*). *BMC Genomics.* 2022;23(1):1–17.
44. Saenko SV, Groenenberg DS, Davison A, Schilthuizen M. The draft genome sequence of the grove snail *Cepaea nemoralis*. *G3.* 2021;11(2):jkaa071.
45. D'Hondt S, Rutherford S, Spivack AJ. Metabolic activity of subsurface life in deep-sea sediments. *Science.* 2002;295(5562):2067–70.
46. Takeuchi T. Molluscan genomics: implications for biology and aquaculture. *Curr Mol Biol Rep.* 2017;3(4):297–305.
47. Giesbrecht GG. Cold stress, near drowning and accidental hypothermia: a review. *Aviat Space Environ Med.* 2000;71(7):733–52.
48. Dallas M, Boycott HE, Atkinson L, Miller A, Boyle JP, Pearson HA, Dallas M, Boycott HE, Atkinson L, Miller A, Boyle JP, Pearson HA, Peers C. Hypoxia suppresses glutamate transport in astrocytes. *J Neurosci.* 2007;27(15):3946–55. <https://doi.org/10.1523/JNEUROSCI.5030-06.2007>.
49. Brose SA, Marquardt AL, Golovko MY. Fatty acid biosynthesis from glutamate and glutamine is specifically induced in neuronal cells under hypoxia. *J Neurochem.* 2014;129(3):400–12. <https://doi.org/10.1111/jnc.12617>.
50. Kumar D, Bansal A, Thomas P, Mongia S, Sharma S, Sairam M, et al. Improved high altitude hypoxic tolerance and amelioration of Anorexia and hypophagia in rats on oral glutamate supplementation. *Aviat Space Environ Med.* 1999;70(5):475–9.
51. Kumar D, Bansal A, Thomas P, Sairam M, Sharma S, Mongia S, et al. Biochemical and immunological changes on oral glutamate feeding in male albino rats. *Int J Biometeorol.* 1999;42(4):201–4.
52. Harvey RJ, Vreugdenhil E, Barnard EA, Darlison MG. Cloning of genomic and cDNA sequences encoding an invertebrate γ -aminobutyric acid_A receptor subunit. *Biochem Soc Trans.* 1990;18(3):438–9. <https://doi.org/10.1042/bst0180438>.
53. Seifkhalhor M, Aliniaefard S, Hassani B, Niknam V, Lastochkina O. Diverse role of γ -aminobutyric acid in dynamic plant cell responses. *Plant Cell Rep.* 2019;38(8):847–67. <https://doi.org/10.1007/s00299-019-02396-z>.
54. Guo Y, Yang R, Chen H, Song Y, Gu Z. Accumulation of γ -aminobutyric acid in germinated soybean (*Glycine max* L.) in relation to glutamate decarboxylase and diamine oxidase activity induced by additives under hypoxia. *Eur Food Res Technol.* 2012;234(4):679–87. <https://doi.org/10.1007/s00217-012-1678-y>.
55. Zhao P, Qian H, Xia Y. GABA and glycine are protective to mature but toxic to immature rat cortical neurons under hypoxia. *Eur J Neurosci.* 2005;22(2):289–300. <https://doi.org/10.1111/j.1460-9568.2005.04222.x>.
56. Okano K, Ueda M, Kuraishi Y, Satoh M. Effect of repeated cold stress on capsaicin-evoked release of glutamate from rat spinal dorsal horn slices. *Neurosci Res.* 1997;29(4):319–24.
57. Jara P, Rage F, Dorfman M, Grouselle D, Barra R, Arancibia S, et al. Cold-induced glutamate release in vivo from the magnocellular region of the paraventricular nucleus is involved in ovarian sympathetic activation. *J Neuroendocrinol.* 2010;22(9):979–86.
58. Gilad GM, Gilad VH, Wyatt RJ, Tizabi Y. Region-selective stress-induced increase of glutamate uptake and release in rat forebrain. *Brain Res.* 1990;525(2):335–8. [https://doi.org/10.1016/0006-8993\(90\)90886-G](https://doi.org/10.1016/0006-8993(90)90886-G).
59. Mazzucotelli E, Tartari A, Cattivelli L, Forlani G. Metabolism of γ -aminobutyric acid during cold acclimation and freezing and its relationship to frost tolerance in barley and wheat. *J Exp Bot.* 2006;57(14):3755–66. <https://doi.org/10.1093/jxb/erl141>.
60. Sharma K, Syed AS, Ferrando S, Mazan S, Korsching SI. The chemosensory receptor repertoire of a true shark is dominated by a single olfactory receptor family. *Genome Biol Evol.* 2019;11(2):398–405. <https://doi.org/10.1093/gbe/evz002>.
61. Srinivasan J, Dillman AR, Macchietto MG, Heikkinen L, Lakso M, Fracchia KM, Srinivasan J, Dillman AR, Macchietto MG, Heikkinen L, Lakso M, Fracchia KM, Antoshechkin I, Mortazavi A, Wong G, Sternberg PW. The draft genome and transcriptome of panagrellus redivivus are shaped by the harsh demands of a free-living lifestyle. *Genetics.* 2013;193(4):1279–95. <https://doi.org/10.1534/genetics.112.148809>.
62. Strotmann R, Schröck K, Bösel T, Stäubert C, Russ A, Schöneberg T. Evolution of GPCR: change and continuity. *Mol Cell Endocrinol.* 2011;331(2):170–8. <https://doi.org/10.1016/j.mce.2010.07.012>.
63. Hill CA, Fox AN, Pitts RJ, Kent LB, Tan PL, Chrystal MA, Hill CA, Fox AN, Pitts RJ, Kent LB, Tan PL, Chrystal MA, Cravchik A, Collins FH, Robertson HM, Zwiebel LJ. G protein-coupled receptors in *Anopheles gambiae*. *Science.* 2002;298(5591):176–8. <https://doi.org/10.1126/science.1076196>.
64. Thomas JH, Robertson HM. The caenorhabditis chemoreceptor gene families. *BMC Biol.* 2008;6(1):1–17.
65. Lindberg DR, Sigwart JD. What is the molluscan osphradium? A reconsideration of homology. *Zool Anz.* 2015;256:14–21. <https://doi.org/10.1016/j.jcz.2015.04.001>.
66. Young JZ. The central nervous system of nautilus. *Philos Trans R Soc Lond B Biol Sci.* 1965;249(754):1–25. <https://doi.org/10.1098/rstb.1965.0006>.
67. Ruth P, Schmidtberg H, Westermann B, Schipp R. The sensory epithelium of the tentacles and the rhinophore of *Nautilus pompilius* L. (cephalopoda, nautiloidea). *J Morphol.* 2002;251(3):239–55. <https://doi.org/10.1002/jmor.1086>.
68. Basil J, Bahctinova I, Kuroiwa K, Lee N, Mims D, Preis M, Basil J, Bahctinova I, Kuroiwa K, Lee N, Mims D, Preis M, Soucier C. The function of the rhinophore and the tentacles of nautilus pompilius L. (Cephalopoda, Nautiloidea) in orientation to odor. *Mar Freshw Behav Physiol.* 2005;38(3):209–21. <https://doi.org/10.1080/10236240500310096>.
69. Cummins SF, Erpenbeck D, Zou Z, Claudianos C, Moroz LL, Nagle GT, et al. Candidate chemoreceptor subfamilies differentially expressed in the chemosensory organs of the mollusc *Aplysia*. *BMC Biol.* 2009;7(1):1–20.
70. Kelley JL, Arias-Rodriguez L, Patacsil Martin D, Yee MC, Bustamante CD, Tobler M. Mechanisms underlying adaptation to life in hydrogen sulfide-rich environments. *Mol Biol Evol.* 2016;33(6):1419–34. <https://doi.org/10.1093/molbev/msw020>.
71. Beauchamp RO Jr, Bus JS, Popp JA, Boreiko CJ, Andjelkovich DA, Beauchamp RO, Bus JS, Popp JA, Boreiko CJ, Andjelkovich DA, Leber P. A critical review of the literature on hydrogen sulfide toxicity. *Crit Rev Toxicol.* 1984;13(1):25–97. <https://doi.org/10.3109/10408448409029321>.
72. Reiffenstein RJ, Hulbert WC, Roth SH. Toxicology of hydrogen sulfide. *Annu Rev Pharmacol Toxicol.* 1992;32(1):109–34. <https://doi.org/10.1146/annurev.pa.32.040192.000545>.
73. Cooper CE, Brown GC. The inhibition of mitochondrial cytochrome oxidase by the gases carbon monoxide, nitric oxide, hydrogen cyanide and hydrogen sulfide: chemical mechanism and physiological significance. *J Bioenerg Biomembr.* 2008;40(5):533–9. <https://doi.org/10.1007/s10863-008-9166-6>.

74. Bagarinao T. Sulfide as an environmental factor and toxicant: tolerance and adaptations in aquatic organisms. *Aquat Toxicol.* 1992;24(1–2):21–62. [https://doi.org/10.1016/0166-445X\(92\)90015-F](https://doi.org/10.1016/0166-445X(92)90015-F).
75. Wong YH, Sun J, He LS, Chen LG, Qiu JW, Qian PY. High-throughput transcriptome sequencing of the cold seep mussel *bathymodiolus platifrons*. *Sci Rep.* 2015;5(1):1–15.
76. Vismann B. Sulfide tolerance: physiological mechanisms and ecological implications. *Ophelia.* 1991;34(1):1–27. <https://doi.org/10.1080/00785326.1991.10429703>.
77. Pfenninger M, Lerp H, Tobler M, Passow C, Kelley JL, Funke E, Pfenninger M, Lerp H, Tobler M, Passow C, Kelley JL, Funke E, Greshake B, Erkoc UK, Berberich T, Plath M. Parallel evolution of cox genes in H₂S-tolerant fish as key adaptation to a toxic environment. *Nat Commun.* 2014;5(1):3873. <https://doi.org/10.1038/ncomms4873>.
78. Hildebrandt TM, Grieshaber MK. Three enzymatic activities catalyze the oxidation of sulfide to thiosulfate in mammalian and invertebrate mitochondria. *FEBS J.* 2008;275(13):3352–61. <https://doi.org/10.1111/j.1742-4658.2008.06482.x>.
79. Lagoutte E, Mimoun S, Andriamihaja M, Chaumontet C, Blachier F, Bouillaud F. Oxidation of hydrogen sulfide remains a priority in mammalian cells and causes reverse electron transfer in colonocytes. *Biochim Biophys Acta.* 2010;1797(8):1500–11. <https://doi.org/10.1016/j.bbabi.2010.04.004>.
80. McMahon RF. Respiratory response to periodic emergence in intertidal molluscs. *Am Zool.* 1988;28(1):97–114. <https://doi.org/10.1093/icb/28.1.97>.
81. Roméo M, Gnassia-Barelli M. Effect of heavy metals on lipid peroxidation in the Mediterranean clam *ruditapes decussatus*. *Comp Biochem Physiol C Toxicol Pharmacol.* 1997;118(1):33–7.
82. Oakley AJ. Glutathione transferases: new functions. *Curr Opin Struct Biol.* 2005;15(6):716–23. <https://doi.org/10.1016/j.sbi.2005.10.005>.
83. Sharma R, Yang Y, Sharma A, Awasthi S, Awasthi YC. Antioxidant role of glutathione S-transferases: protection against oxidant toxicity and regulation of stress-mediated apoptosis. *Antioxid Redox Signal.* 2004;6(2):289–300.
84. Felton GW, Summers CB. Antioxidant systems in insects. *Arch Insect Biochem Physiol.* 1995;29(2):187–97. <https://doi.org/10.1002/arch.94029.0208>.
85. Lee YM, Lee KW, Park H, Park HG, Raisuddin S, Ahn IY, Lee YM, Lee KW, Park H, Park HG, Raisuddin S, Ahn IY, Lee JS. Sequence, biochemical characteristics and expression of a novel sigma-class of glutathione S-transferase from the intertidal copepod, *tigriopus japonicus* with a possible role in antioxidant defense. *Chemosphere.* 2007;69(6):893–902. <https://doi.org/10.1016/j.chemosphere.2007.05.087>.
86. Pushpamali WA, De Zoysa M, Kang HS, Oh CH, Whang I, Kim SJ, Pushpamali WA, De Zoysa M, Kang H-S, Oh CH, Whang I, Kim SJ, Lee J. Comparative study of two thioredoxin peroxidases from disk abalone (*Haliotis discus discus*): cloning, recombinant protein purification, characterization of antioxidant activities and expression analysis. *Fish Shellfish Immunol.* 2008;24(3):294–307. <https://doi.org/10.1016/j.fsi.2007.11.016>.
87. Petkowski JJ, Bains W, Seager S. Natural products containing a nitrogen-sulfur bond. *J Nat Prod.* 2018;81(2):423–46. <https://doi.org/10.1021/acs.jnatprod.7b00921>.
88. Coleman KE, Huang TT. In a class of its own: a new family of deubiquitinases promotes genome stability. *Mol Cell.* 2018;70(1):1–3. <https://doi.org/10.1016/j.molcel.2018.03.022>.
89. Dyson N. *Essential molecular biology: a practical approach.* Oxford: IRL Press; 1991.
90. Luo R, Liu B, Xie Y, Li Z, Huang W, Yuan J, et al. SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler. *Gigascience.* 2012;1(1):2047–217 X-2041-2018.
91. Belaghzal H, Dekker J, Gibcus JH. Hi-C 2.0: an optimized Hi-C procedure for high-resolution genome-wide mapping of chromosome conformation. *Methods.* 2017;123:56–65.
92. Li H. Minimap and minimap: fast mapping and de novo assembly for noisy long sequences. *Bioinformatics.* 2016;32(14):2103–10. <https://doi.org/10.1093/bioinformatics/btw152>.
93. Vaser R, Sović I, Nagarajan N, Šikić M. Fast and accurate de novo genome assembly from long uncorrected reads. *Genome Res.* 2017;27(5):737–46. <https://doi.org/10.1101/gr.214270.116>.
94. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics.* 2009;25(14):1754–60. <https://doi.org/10.1093/bioinformatics/btp324>.
95. Walker BJ, Abeel T, Shea T, Priest M, Abouelliel A, Sakthikumar S, et al. Pilon: an integrated tool for comprehensive microbial variant detection and genome assembly improvement. *PLoS One.* 2014;9(11):e112963. <https://doi.org/10.1371/journal.pone.0112963>.
96. Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. *Bioinformatics.* 2015;31(19):3210–2. <https://doi.org/10.1093/bioinformatics/btv351>.
97. Kriventseva EV, Tegenfeldt F, Petty TJ, Waterhouse RM, Simao FA, Pozdnyakov IA, et al. OrthoDB v8: update of the hierarchical catalog of orthologs and the underlying free software. *Nucleic Acids Res.* 2015;43(D1):D250–256.
98. Bolger AM, Lohse M, Usadel B. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics.* 2014;30(15):2114–20. <https://doi.org/10.1093/bioinformatics/btu170>.
99. Durand NC, Shamim MS, Machol I, Rao SS, Huntley MH, Lander ES, et al. Juicer provides a one-click system for analyzing loop-resolution Hi-C experiments. *Cell Syst.* 2016;3(1):95–8.
100. Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Dudchenko O, Batra SS, Omer AD, Nyquist SK, Hoeger M, Durand NC, Shamim MS, Machol I, Lander ES, Aiden AP, Aiden EL. De novo assembly of the *Aedes aegypti* genome using Hi-C yields chromosome-length scaffolds. *Science.* 2017;356(6333):92–5. <https://doi.org/10.1126/science.aal3327>.
101. Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Durand NC, Robinson JT, Shamim MS, Machol I, Mesirov JP, Lander ES, Aiden EL. Juicebox provides a visualization system for Hi-C contact maps with unlimited zoom. *Cell Syst.* 2016;3(1):99–101. <https://doi.org/10.1016/j.cels.2015.07.012>.
102. Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Chin C-S, Alexander DH, Marks P, Klammer AA, Drake J, Heiner C, Clum A, Copeland A, Huddleston J, Eichler EE, Turner SW, Korlach J. Nonhybrid, finished microbial genome assemblies from long-read SMRT sequencing data. *Nat Methods.* 2013;10(6):563–9. <https://doi.org/10.1038/nmeth.2474>.
103. Salmela L, Rivals E. LoRDEC: accurate and efficient long read error correction. *Bioinformatics.* 2014;30(24):3506–14. <https://doi.org/10.1093/bioinformatics/btu538>.
104. Li W, Godzik A. Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. *Bioinformatics.* 2006;22(13):1658–9. <https://doi.org/10.1093/bioinformatics/btl158>.
105. Chen Y, Chen Y, Shi C, Huang Z, Zhang Y, Li S, et al. SOAPnuke: a MapReduce acceleration-supported software for integrated quality control and preprocessing of high-throughput sequencing data. *Gigascience.* 2017;7(1):gix120.
106. Haas BJ, Papanicolaou A, Yassour M, Grabherr M, Blood PD, Bowden J, et al. De novo transcript sequence reconstruction from RNA-seq using the Trinity platform for reference generation and analysis. *Nat Protoc.* 2013;8(8):1494–512. <https://doi.org/10.1038/nprot.2013.084>.
107. Langmead B, Trapnell C, Pop M, Salzberg SL. Ultrafast and memory-efficient alignment of short DNA sequences to the human genome. *Genome Biol.* 2009;10(3):R25. <https://doi.org/10.1186/gb-2009-10-3-r25>.
108. Li B, Dewey CN. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinform.* 2011;12(1):323.
109. Young MD, Wakefield MJ, Smyth GK, Oshlack A. Goseq: gene ontology testing for RNA-seq datasets. *R Bioconductor.* 2012;8:1–25.
110. Wu J, Mao X, Cai T, Luo J, Wei L. KOBAS server: a web-based platform for automated annotation and pathway identification. *Nucleic Acids Res.* 2006;34:W720–724.
111. Xu Z, Wang H. LTR_FINDER: an efficient tool for the prediction of full-length LTR retrotransposons. *Nucleic Acids Res.* 2007;35:W265–268.
112. Maja TG, Chen N. Using RepeatMasker to identify repetitive elements in genomic sequences. *Curr Protoc Bioinformatics.* 2009;25(1):41011–141014.
113. Smit AF, Hubley R. RepeatModeler Open-1.0. 2008–2015. Available from: <http://www.repeatmasker.org>.

114. Benson G. Tandem repeats finder: a program to analyze DNA sequences. *Nucleic Acids Res.* 1999;27(2):573–80. <https://doi.org/10.1093/nar/27.2.573>.
115. Jurka J, Kapitonov VV, Pavlicek A, Klonowski P, Kohany O, Walichiewicz J. Repbase Update, a database of eukaryotic repetitive elements. *Cytogenet Genome Res.* 2005;110(1–4):462–7.
116. Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. AUGUSTUS: ab initio prediction of alternative transcripts. *Nucleic Acids Res.* 2006;34:W435–439.
117. Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA. *J Mol Biol.* 1997;268(1):78–94. <https://doi.org/10.1006/jmbi.1997.0951>.
118. Korf I. Gene finding in novel genomes. *BMC Bioinform.* 2004;5(1): 59.
119. Rao SS, Huntley MH, Durand NC, Stamenova EK, Bochkov ID, Robinson JT, et al. A 3D map of the human genome at kilobase resolution reveals principles of chromatin looping. *Cell.* 2014;159(7):1665–80.
120. Consortium TCeS. Genome sequence of the nematode *C. Elegans*: a platform for investigating biology. *Science.* 1998;282(5396):2012–8. <https://doi.org/10.1126/science.282.5396.2012>.
121. Gómez-Chiari M, Warren WC, Guo X, Proestou D. Developing tools for the study of molluscan immunity: the sequencing of the genome of the eastern oyster, *Crassostrea virginica*. *Fish Shellfish Immunol.* 2015;46(1):2–4.
122. Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, et al. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature.* 2012;490(7418):49–54. <https://doi.org/10.1038/nature11413>.
123. Adams MD, Celniker SE, Holt RA, Evans CA, Gocayne JD, Amanatides PG, et al. The genome sequence of *Drosophila melanogaster*. *Science.* 2000;287(5461):2185–95.
124. Simakov O, Marletaz F, Cho SJ, Edsinger Gonzales E, Havlak P, Hellsten U, et al. Insights into bilaterian evolution from three spiralian genomes. *Nature.* 2013;493(7433):526–31.
125. Wang S, Zhang J, Jiao W, Li J, Xun X, Sun Y, et al. Scallop genome provides insights into evolution of bilaterian karyotype and development. *Nat Ecol Evol.* 2017;1(5):1–12.
126. Albertin CB, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, Albertin CB, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, Brenner S, Ragsdale CW, Rokhsar DS. The octopus genome and the evolution of cephalopod neural and morphological novelties. *Nature.* 2015;524(7564):220–4. <https://doi.org/10.1038/nature14668>.
127. Pevsner J. Basic local alignment search tool (BLAST). In: *Bioinformatics and functional genomics*. 2nd ed. Hoboken: John Wiley & Sons Inc; 2009. p. 121–66.
128. Birney E, Clamp M, Durbin R. GeneWise and genomewise. *Genome Res.* 2004;14(5):988–95. <https://doi.org/10.1101/gr.1865504>.
129. Campbell MS, Holt C, Moore B, Yandell M. Genome annotation and curation using MAKER and MAKER-P. *Curr Protoc Bioinformatics.* 2014;48(1):4.11.11–14.11.39.
130. Benson DA, Karsch-Mizrachi I, Lipman DJ, Ostell J, Wheeler DL. GenBank. *Nucleic Acids Res.* 2005;33:D34–8.
131. Bairoch A, Apweiler R. The SWISS-PROT protein sequence database and its supplement TrEMBL in 2000. *Nucleic Acids Res.* 2000;28(1):45–8.
132. Hunter S, Apweiler R, Attwood TK, Bairoch A, Bateman A, Binns D, et al. InterPro: the integrative protein signature database. *Nucleic Acids Res.* 2009;37:D211–215.
133. Tatusov RL, Koonin EV, Lipman DJ. A genomic perspective on protein families. *Science.* 1997;278(5338):631–7. <https://doi.org/10.1126/science.278.5338.631>.
134. Kanehisa M, Goto S. KEGG: kyoto encyclopedia of genes and genomes. *Nucleic Acids Res.* 2000;28(1):27–30.
135. Consortium GO. The Gene Ontology (GO) database and informatics resource. *Nucleic Acids Res.* 2004;32:D258–261.
136. Tatusov RL, Galperin MY, Natale DA, Koonin EV. The COG database: a tool for genome-scale analysis of protein functions and evolution. *Nucleic Acids Res.* 2000;28(1):33–6.
137. Tatusov RL, Fedorova ND, Jackson JD, Jacobs AR, Kiryutin B, Koonin EV, et al. The COG database: an updated version includes eukaryotes. *BMC Bioinform.* 2003;4(1):1–14.
138. Bateman A, Coin L, Durbin R, Finn RD, Hollich V, Griffiths-Jones S, et al. The pfam protein families database. *Nucleic Acids Res.* 2004;32:D138–141.
139. Eddy SR. Accelerated profile HMM searches. *PLoS Comput Biol.* 2011;7(10):e1002195. <https://doi.org/10.1371/journal.pcbi.1002195>.
140. Adema CM, Hillier LW, Jones CS, Loker ES, Knight M, Minx P, et al. Whole genome analysis of a schistosomiasis-transmitting freshwater snail. *Nat Commun.* 2017;8(1):1–12.
141. Li L, Stoeckert CJ, Roos DS. OrthoMCL: identification of ortholog groups for eukaryotic genomes. *Genome Res.* 2003;13(9):2178–89.
142. Edgar RC. MUSCLE: multiple sequence alignment with high accuracy and high throughput. *Nucleic Acids Res.* 2004;32(5):1792–7. <https://doi.org/10.1093/nar/gkh340>.
143. Guindon S, Dufayard JF, Lefort V, Anisimova M, Hordijk W, Gascuel O. New algorithms and methods to estimate maximum-likelihood phylogenies: assessing the performance of PhyML 3.0. *Syst Biol.* 2010;59(3):307–21. <https://doi.org/10.1093/sysbio/syq010>.
144. Yang Z. PAML 4: phylogenetic analysis by maximum likelihood. *Mol Biol Evol.* 2007;24(8):1586–91. <https://doi.org/10.1093/molbev/msm088>.
145. Yang Z, Rannala B. Bayesian estimation of species divergence times under a molecular clock using multiple fossil calibrations with soft bounds. *Mol Biol Evol.* 2006;23(1):212–26. <https://doi.org/10.1093/molbev/msj024>.
146. Kumar S, Stecher G, Suleski M, Hedges SB. TimeTree: a resource for time-lines, timetrees, and divergence times. *Mol Biol Evol.* 2017;34(7):1812–9. <https://doi.org/10.1093/molbev/msx116>.
147. Ren J, Liu X, Jiang F, Guo X, Liu B. Unusual conservation of mitochondrial gene order in *Crassostrea* oysters: evidence for recent speciation in Asia. *BMC Evol Biol.* 2010;10(1):394. <https://doi.org/10.1186/1471-2148-10-394>.
148. Zapata F, Wilson NG, Howison M, Andrade SC, Jörger KM, Schrödl M, et al. Phylogenomic analyses of deep gastropod relationships reject Orthogastropoda. *Proc Royal Soc B.* 2014;281(1794):20141739.
149. Vinther J, Sperling EA, Briggs DE, Peterson KJ. A molecular palaeobiological hypothesis for the origin of aplousobranch molluscs and their derivation from chiton-like ancestors. *Proc Royal Soc B.* 2012;279(1732):1259–68.
150. Gold DA, Runnegar B, Gehling JG, Jacobs DK. Ancestral state reconstruction of ontogeny supports a bilaterian affinity for Dickinsonia. *Evol Dev.* 2015;17(6):315–24. <https://doi.org/10.1111/ede.12168>.
151. Parfrey LW, Lahr DJ, Knoll AH, Katz LA. Estimating the timing of early eukaryotic diversification with multigene molecular clocks. *Proc Natl Acad Sci.* 2011;108(33):13624–9.
152. De Bie T, Cristianini N, Demuth JP, Hahn MW. CAFE: a computational tool for the study of gene family evolution. *Bioinformatics.* 2006;22(10):1269–71. <https://doi.org/10.1093/bioinformatics/btl097>.
153. Lin G, Chai J, Yuan S, Mai C, Cai L, Murphy RW, et al. VennPainter: a tool for the comparison and identification of candidate genes based on venn diagrams. *PLoS One.* 2016;11(4):e0154315. <https://doi.org/10.1371/journal.pone.0154315>.
154. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinform.* 2008;9(1):1–13.
155. Su G, Morris JH, Demchak B, Bader GD. Biological network exploration with cytoscape 3. *Curr Protoc Bioinformatics.* 2014;47(1):8.11–18.13.24.
156. Beier S, Thiel T, Münch T, Scholz U, Mascher M. MISA-web: a web server for microsatellite prediction. *Bioinformatics.* 2017;33(16):2583–5. <https://doi.org/10.1093/bioinformatics/btx198>.
157. Zhang HM, Chen H, Liu W, Liu H, Gong J, Wang H, et al. AnimalTFDB: a comprehensive animal transcription factor database. *Nucleic Acids Res.* 2012;40(D1):D144–149.
158. Kong L, Zhang Y, Ye ZQ, Liu XQ, Zhao SQ, Wei L, et al. CPC: assess the protein-coding potential of transcripts using sequence features and support vector machine. *Nucleic Acids Res.* 2007;35:W345–349.
159. Sun L, Luo H, Bu D, Zhao G, Yu K, Zhang C, et al. Utilizing sequence intrinsic composition to classify protein-coding and long non-coding transcripts. *Nucleic Acids Res.* 2013;41(17):e166–166.
160. Li A, Zhang J, Zhou Z. PLEK: a tool for predicting long non-coding RNAs and messenger RNAs based on an improved k-mer scheme. *BMC Bioinform.* 2014;15(1):311.

Publisher's Note

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