Scallop genome reveals molecular adaptations to semi-sessile life and neurotoxins

Yuli Li1, Xiaoqing Sun2, Xiaoli Hu1,3, Xiaogang Xun1, Jinbo Zhang2, Ximing Guo4, Wenqian Jiao1, Lingling Zhang1,3, Weizhi Liu1,5, Jing Wang1, Ji Li2, Yan Sun1, Yan Miao1, Xiaokang Zhang1, Taoran Cheng1, Guoliang Xu2, Xiaoteng Fu1, Yangfan Wang1, Xinran Yu2, Xiaoting Huang1, Wei Lu1, Jia Lv1, Chuang Mu1, Dawei Wang2, Xu Li1, Yu Xia1, Yajuan Li1, Zhihui Yang1, Fengliang Wang1, Lu Zhang1, Qiang Xing1, Huaiqian Dou1, Xianhui Ning1, Jinzhuan Dou1, Yangping Li1, Dexu Kong1, Yaran Liu1, Zhi Jiang2, Ruiqiang Li2, Shi Wang1,5 & Zhenmin Bao1,3

Bivalve molluscs are descendants of an early-Cambrian lineage superbly adapted to benthic filter feeding. Adaptations in form and behavior are well recognized, but the underlying molecular mechanisms are largely unknown. Here, we investigate the genome, various transcriptomes, and proteomes of the scallop *Chlamys farrelli*, a semi-sessile bivalve with well-developed adductor muscle, sophisticated eyes, and remarkable neurotoxin resistance. The scallop’s large striated muscle is energy-dynamic but not fully differentiated from smooth muscle. Its eyes are supported by highly diverse, intronless opsins expanded by retroposition for broadened spectral sensitivity. Rapid byssal secretion is enabled by a specialized foot and multiple proteins including expanded tyrosinases. The scallop uses hepatopancreas to accumulate neurotoxins and kidney to transform to high-toxicity forms through expanded sulfotransferases, probably as deterrence against predation, while it achieves neurotoxin resistance through point mutations in sodium channels. These findings suggest that expansion and mutation of those genes may have profound effects on scallop’s phenotype and adaptation.
bivalve molluscs, which first appeared in the early-Cambrian over 500 million years ago (MYA), represent an ancient lineage of bilaterians that has survived several mass extinction events. Yet, extant bivalves with ~9600 species remain abundant and thriving in world oceans and freshwater environments ranging from tropical to polar regions and from intertidal zones to deep seas. Bivalves are well adapted to benthic life as sessile, semi-sessile, or free-living filter feeders and play critical roles in benthic ecology. Many bivalves are important fishery and aquaculture species providing significant economic benefits to humans. Despite their biological, ecological, and economic significance, their genomes are poorly sampled for whole-genome studies, limiting our understanding of bilaterian evolution, especially molecular adaptations in this ancient but successful lineage.

Scallops are among the best-known bivalves recognized for their beautiful shells of elegant shapes, sophisticated patterns, and diverse colors. Scallops have some unique characteristics making them good models to study development, adaptation, and early animal evolution, as indicated by a recent genome analysis of *P. yessoensis* 

Scallops have a large adductor muscle, probably as an adaptation to swimming by clapping valves for avoiding predation and seeking favorable habitats. They are rare among lophotrochozoans in having numerous image-forming eyes along the edges of their mantles that perform vital functions in detecting predators and guiding swimming.

Scallops can attach to substrates as juveniles by rapidly producing adhesive and strong byssal threads that are either retained or lost in adults. As filter feeders that may feed on toxic dinoflagellates, scallops can accumulate and tolerate high levels of neurotoxins such as paralytic shellfish toxins (PSTs) that are among the most potent natural toxins for humans. These adaptive features are remarkable, and understanding the molecular innovations underlying these remarkable features may provide insights into how organisms adapt to their environments and evolve, which is a fundamental question in evolutionary biology.

The Zhihong scallop *Chlamys farreri* (Jones et Preston, 1904, also known as Chinese scallop) is a subtropical Western Pacific bivalve with wide water-temperature tolerance (−1.5 to 30 °C) and is naturally distributed along the coasts of Northern China, Korea, Japan, and Eastern Russia. *C. farreri* is epibenthic and semi-sessile. It usually attaches itself to rocks and other hard surfaces with byssal threads, but can detach under adverse conditions and swim away to new habitats. It has an outstanding ability to accumulate PSTs (up to 40,241 μg saxitoxin (STX) eq. per 100 g compared to the 80 μg STX eq. per 100 g safety level for human) and therefore is widely used for studying PST accumulation and transformation.

*C. farreri* is also a commercially important bivalve with aquaculture production once reaching ~1 million metric tons. It is among the best genetically important bivalve with aquaculture production once contig N50: 19–38 kb and scaffold N50: 167–804 kb. K-mer analysis (Supplementary Fig. 1) provides an estimate of genome size of ~1 Gb, which is similar to ~1.2 Gb estimated by flow cytometry.

The integrity and high quality of the assembly is demonstrated by the mapping of 95.8% of sequencing reads, 97.6–100% of Sanger-sequenced BAC clones, and 99.6–100% of various transcriptome datasets (Supplementary Figs. 2 and Supplementary Tables 4–6), and by the Benchmarking Universal Single-Copy Orthologs (BUSCO)-based completeness assessment (Supplementary Table 7). The assembly was anchored to chromosomes by assigning 949 scaffolds (covering 66.9% of the assembly) to 19 linkage groups (Supplementary Table 8 and Fig. 1a) of a high-density genetic linkage map.

The *C. farreri* genome contains 28,602 protein-coding genes, of which 93.3% encode proteins of at least 100 amino acid residues (aa), and 94.3% are supported by known protein sequences and/or transcriptomic data (Supplementary Fig. 3). Functional analysis via comparison with various public protein databases annotated 24,817 genes, accounting for 86.8% of all the predicted genes (Supplementary Table 9). The *C. farreri* genome contains 261.8 Mb of repetitive sequences accounting for 32.1% of the genome. This percentage is lower than that in most of existing molluscan genomes (Supplementary Table 10). Tandem repeats represent the most abundant repeat type (11.3%), followed by DNA transposons (6.7%), and long interspersed elements (4.4%). Transposable elements (TEs) show higher divergence in *C. farreri* than in other bivalves (Supplementary Figs. 4 and 5), suggesting that they are relatively old.

Polymorphism analysis identified 4.9 million single-nucleotide polymorphisms (SNPs) in the assembled individual (Supplementary Table 11), yielding an intra-individual polymorphism rate of 0.81%. A genome-wide scan of polymorphism based on the assembled and five additionally resequenced individuals identified 108 highly polymorphic genomic regions (~2500 kb) in the genome (Fig. 1a), among which six are longer than 5 Mb (approximately one-fifth of a single chromosome). SNP density in coding sequences (CDSs) varies dramatically among genes, ranging from 0 to ~117 SNPs per kb (Fig. 1a). Particularly, scanning the CDS regions identified a continuous and long SNP-scarce region (~1.74 Mb) on chromosome 1 (Fig. 1b), which harbors an intact cluster of 11 *Hox* genes (3 anterior, 6 central, and 2 posterior): key regulators of bilaterian body plan development.

Similarly, low polymorphism was also observed for the *Hox* genes of the scallop *P. yessoensis*, fruit fly, and mouse (Supplementary Fig. 6). The finding of scallop *Hox* genes largely devoid of polymorphism despite high SNP diversity in genomic background (Supplementary Table 12) suggests that the scallop’s body plan formation may be subject to rigid developmental control and its regulators are under strong purifying selection.
Phylogenetic analysis based on 1310 highly conserved orthologous genes (Supplementary Fig. 7) suggests that the scallop lineage diverged from the lineage leading to Pinctada fucata and Crassostrea gigas ~457 MYA, and Bivalvia diverged from its sister group Gastropoda ~500 MYA. Gene family analysis (Supplementary Fig. 8) revealed that C. farreri has preserved the highest number (7604) of ancestral bilaterian gene families among bivalves. This number is comparable to that in Lingula anatina, a “living fossil” lophotrochozoan.

Fig. 1 Genome landscape and polymorphism analysis of the scallop C. farreri. a From outer to inner circles: I, marker distribution on 19 chromosomes at the Mb scale; II and III, single-nucleotide polymorphism (SNP) density across genome (II) or CDS (III) was drawn in 1 Mb sliding windows with a 50 kb step (yellow columns), and polymorphism hotspot regions (500 kb–5.7 Mb) are colored red; IV and V, gene density and repeat density across the genome, respectively, drawn in 0.1 Mb nonoverlapping windows. b SNP density of the Hox gene cluster and their genomic background on chromosome 1, showing coding regions of Hox genes with extremely low SNP density, compared to high SNP density in the genomic background. The statistical significance for each Hox gene is shown in Supplementary Table 12. c Gene family expansion/contraction analysis based on 12 representative species. A total of 270 gene families are expanded in the scallop C. farreri relative to other bivalves, with the most significant expanded families involved in neurotransmission, immune response, signal transduction, and xenobiotic metabolism (Supplementary Table 13). Expanded and contracted gene families are highlighted in magenta and turquoise, respectively.

Muscle regulation and evolution. Scallops have a remarkably large adductor muscle (Supplementary Fig. 10) compared to most of sessile and endobenthic species such as oysters, mussels, and clams, probably as adaptation to swimming (Supplementary Movie 1) and the semi-sessile lifestyle. Swimming is an energy-intensive activity, and it is not surprising that the adductor muscle in scallops also serves as the primary organ of energy and glycogen storage and mobilization. We found that arginine kinase, the key enzyme responsible for producing ~70% of the ATP needed for phasic contractions (using arginine phosphate as substrate), shows extremely high levels of transcription in the adductor muscle of C. farreri, especially the striated portion (transcripts per million (TPM) = 34,704; ranked sixth among all genes; Fig. 2a). Further analysis of energy-producing pathways (glycolysis, the tricarboxylic acid (TCA) cycle, and oxidative phosphorylation) suggests that most genes related to energy production show higher expression in the C. farreri’s striated muscle than in its smooth muscle (Fig. 2a, Supplementary Fig. 11, and Supplementary Table 14), pointing to higher energy dynamics in striated muscle than in smooth muscle. These findings may reflect differences in function, with the large striated muscle responsible for fast, repetitive clapping of valves during swimming and the small smooth muscle responsible for keeping valves closed for long periods at a relatively low energy cost.

Interestingly, enzymes participating in energy (ATP/energy-rich H⁺) production (e.g., glyceraldehyde 3-phosphate dehydrogenase, pyruvate dehydrogenase, dihydrolipoamide acetyltransferase, isocitrate dehydrogenase, succinyl-CoA synthetase β-subunit in glycolysis and in the TCA cycle) generally show higher expression in the scallop C. farreri than in the oyster C. gigas, but the reverse is true for the enzymes related to energy consumption (HK and FBP; Fig. 2a and Supplementary Table 14). This finding may reflect adaptations to lifestyles with different levels of energy demand: high in semi-sessile scallop and low in sessile oyster.

To understand transcriptomic regulation in the adductor muscle, we constructed gene co-expression networks from 35 adult transcriptome datasets, and identified M3 as the only adductor muscle-related module (significantly enriched in both striated muscle- and smooth muscle-related genes;
Twist2, E75 and FBP show lower expression in generally show higher expression in the scallop oxoglutarate dehydrogenase; LSC2, succinyl-CoA synthetase dehydrogenase; DLAT, dihydrolipoamide acetyltransferase; CS, citrate synthase; ACO, aconitate hydratase; IDH, isocitrate dehydrogenase; OGDH, sucA 2-PGK, phosphoglycerate kinase; PGAM, phosphoglycerate mutase; ENO, enolase; PK, pyruvate kinase; OCDH, octopine dehydrogenase; PDH, pyruvate dehydrogenase with the former two known as key TFs in animal myogenesis,

are among the top-ranked TFs with the highest intramodular connectivity (Fig. 2b and Supplementary Data 3), enzymes participating in energy (ATP/energy-rich H+) production reactions (GAPDH, PGK, PK, PDH, DLAT, IDH, OGDH, LSC2, SDH, and MDH) phosphoryl arginine, glycolysis, and TCA cycle. The sum of TPM values of genes encoding the same enzyme was determinate for possibly related to the different modes of adult muscle growth: 

with the former two known as key TFs in animal myogenesis, are among the top-ranked TFs with the highest intramodular connectivity (Fig. 2b and Supplementary Data 3), enzymes participating in energy (ATP/energy-rich H+) production reactions (GAPDH, PGK, PK, PDH, DLAT, IDH, OGDH, LSC2, SDH, and MDH) phosphoryl arginine, glycolysis, and TCA cycle. The sum of TPM values of genes encoding the same enzyme was determinate for possibly related to the different modes of adult muscle growth: 

It is also interesting that key marker genes36 that distinguish vertebrate striated and smooth muscles that are not as distinctive as in vertebrates. The expression of the same fast contractility components in smooth as well as striated muscle has been reported for other scallop species,37, 38, ascidians39, 40, and flatworms41–43. Our findings together with those from other studies suggest that smooth and striated muscles in at least some invertebrates are not as differentiated as in vertebrates, thereby probably representing a plesiomorphic state, and still use shared basic building blocks (i.e., gene components) but in different organizations.

Opsi diversity and retina evolution. Scallops possess a large number of sophisticated non-cephalic eyes along the edge of their mantle (Fig. 3a) and thus are thought to have the best vision system among bivalve molluscs. Opsins of the G-protein-coupled receptor (GPCR) family are key light-sensing proteins
**Fig. 3** The evolution of opsin diversity and photoreception tuning in *C. farreri*. **a** Morphology of scallop eyes scattered along the edge of mantles (left, scale bar: 1 cm), schematic structure of a typical scallop eye (middle), and expression of diverse opsin genes in scallop eyes (right). Eye samples from three individuals were used in expression evaluation with standard error shown for eye and mantle groups. The asterisks indicate genes showing significantly higher expression in eyes than in the mantle (*p*-value < 0.05, the exact test by edgeR). **b** R-opsin gene structures of Mollusca and Branchiopoda. Exons rather than introns were plotted in proportion, with scale bar representing 500 bp. NETR (neurotrypsin) and FPV246 (putative ankyrin repeat protein) are the conserved neighboring genes. These data were calculated based on full protein sequences. The black line inside the box indicates the median value, and the whiskers extend from the first or third quartiles to the minimum or maximum values. **c** Sequence similarity and Ka/Ks values of all opsin genes between the scallops of *C. farreri* (%). **d** Bivalve opsin phylogeny and variation at key functional sites sensitive to various light ranges. Species abbreviations: *Chlamys farreri* (CF), *Pinctada fucata* (PF), *Crassostrea gigas* (CG), *Argopecten irradians* (AI), *Patinopecten caurinus* (PC) and *Argopecten purpuratus* (AP). Site combinations of “SFA,” “GFA” and “AFA” above the major branches are the putative ancestral bivalve types deduced from extant species. Colors of the sites correspond to the colors or wavelengths of opsin sensitivity in human opsins.

responsible for visual signal transduction[^15]. On the basis of sequence alignment with known opsins and GPCR-domain searches, we identified eight vision-related opsin genes in the scallop genome, including four *r-opsin*, two *Go-opsin* and two *c*-like-opsin (*cl-opsin*) genes (Supplementary Fig. 15). As the characteristic opsin type in invertebrates[^46], *r-opsin* (also known as Gq-coupled opsin) genes are significantly expanded (four copies) in the *C. farreri* genome, compared to a single gene copy found in other molluscs (Fig. 3b). In *C. farreri*, *r-opsin1* is likely the ancestral copy, because it shows relatively conserved gene structure and neighboring genes with the single-copy *r-opsin* genes in other molluscs (Fig. 3b). Interestingly, *r-opsin2*, *r-opsin3*, and *r-opsin4* are all intronless (Fig. 3b) and are likely retrogenes generated via retroposition of a messenger RNA (mRNA) transcript[^47]. Phylogenetic analysis suggests that these intronless *r-opsins* were generated by stepwise duplications from the original intron-containing gene (i.e., *r-opsin1*), with *r-opsin2* generated by retroposition first, followed by tandem duplications that produced *r-opsin3* and *r-opsin4* (Fig. 3d). The latter, *r-opsin4*, is the favored gene copy in the eyes of *C. farreri*, because its expression (average TPM = 2415.0) greatly exceeds that (average TPM = 2.3–46.4) of other *r-opsins* (Fig. 3a). Intronless genes are more efficient to transcribe as no post-transcriptional splicing is needed. For example, approximately 70% of early zygotic genes of *Drosophila* are intronless[^48] due to a need for efficient transcription during rapid cell divisions in early development[^48, 49]. The utilization of intronless opsins in *C. farreri* may represent an adaptive change for enhancement of transcription efficiency in support of the scallop’s unusual and advanced multi-eye visual system.

Scallop eyes possess a unique double-layered retina (Fig. 3a), which are equipped with different photoreceptors sensitive to light of different wavelengths and play different roles in visual behaviors[^50]. The proximal retina consists of rhabdomeric photoreceptor cells (also found in most of invertebrate eyes), whereas the distal retina consists of ciliary photoreceptor cells.
Neurotoxin accumulation and transformation. Bivalves can tolerate and accumulate potent neurotoxins such as PSTs, although the molecular mechanism of toxin resistance in bivalves is not well understood. Like tetrodotoxin (TTX) of puffer fish, PSTs attack the nervous system by blocking sodium channels on nerve cell membranes and by inhibiting transduction of an action potential. We identified two sodium channel genes, Nav1 and Nav2, in the C. farreri genome. Nav1 is the primary sodium channel in animal nervous systems and is targeted by PSTs. We found that the scallop’s Nav1 has a potentially toxin-resistant T mutation at position 1425 (in reference to rat sodium channel IIA69, Fig. 5a): the corresponding mutation in rat Nav1 yields a 15-fold increase in resistance to STX (the most potent PST) and a 15-fold increase in resistance to TTX71. This mutation is also present in the Nav1 genes of two puffer fish species, Tetraodon nigroviridis and Takifugu rubripes (Fig. 5a), which have strong toxin resistance, pointing to convergent evolution of toxin resistance in the scallop and puffer fish. Furthermore, our analysis revealed a Q mutation at position 945 in C. gigas and Atlantic awning clam Solenya velum (Fig. 5a), which has been shown to increase STX resistance up to 19,880-fold in a rat sodium
channel. The presence of these two novel mutations that are known to increase toxin resistance in other organisms may explain bivalves’ amazing ability to tolerate neurotoxins. Moreover, gene expression analysis in combination with PST quantification showed that toxin-rich organs (hepatopancreas and kidney) are largely devoid of sodium channel gene expression (Fig. 5b and Supplementary Table 18), which may represent another novel adaptation of the scallop for tolerance of high concentrations of PSTs via down-regulation of their targets. This situation is consistent with the hepatopancreas and kidney being the most toxin-rich or toxin-tolerant organs in the scallop.

Accumulation and transformation of PSTs in bivalves are well documented, but detailed processes and mechanisms remain obscure. To gain a deeper understanding of PST accumulation and transformation in scallops, we comprehensively studied PST accumulation and transformation in C. farreri by qualitatively and quantitatively analyzing a variety of PSTs in six scallop organs across five time points after exposure to PST-producing microalgae *Alexandrium minutum*. We found that the hepatopancreas and kidney are the two organs with the highest concentrations of PSTs, but the kidney is more toxic than hepatopancreas (Fig. 5b). This is a new and significant finding...
because the kidney has been ignored in previous PST studies. The hepatopancreas maintains a similar PST profile as the input microalgae over time; however, the PST profile of the kidney gradually changes with time and the most dramatic toxin change (from input GTXs to highly toxic STX; Supplementary Fig. 20) takes place after 5 days of *A. minutum* exposure (Fig. 5c, Supplementary Fig. 21, and Supplementary Datas 5 and 6). This finding suggests that the scallop hepatopancreas and kidney function differently, with the former mostly accumulating the incoming toxins, whereas the latter in toxin transforming and/or eliminating them. To study the molecular mechanisms of PST accumulation and transformation, we sequenced 36 transcriptomes (Supplementary Table 19) of the hepatopancreas and kidney after *A. minutum* exposure (across six time points, each represented by three individuals) and constructed a gene co-expression network for both organs (Supplementary Fig. 22). Nine and five modules were identified as toxin-responsive (TR) modules in the kidney and hepatopancreas, respectively (Fig. 5c and Supplementary Data 7). Kidney TR modules were found to be involved in diverse molecular functions (e.g., RNA/ion/carbohydrate binding, transferase activity, peptidase activity, and kinase activity), whereas hepatopancreas TR modules primarily participated in a variety of “binding” activities (Supplementary Datas 8 and 9). Notably, genes in the green module were highly expressed on day 5 after *A. minutum* exposure, coinciding with the highest transformation of PSTs in the kidney (Fig. 5c). The green module is significantly enriched with cytosolic sulfotransferase (*Sult*) genes (enrichment *p*-value = 5.9e–3; Fig. 5c and Supplementary Data 8), which may mediate the transfer of a sulfate group from a donor molecule (such as GTXs) to various acceptor molecules, endogenous metabolites, and xenobiotics. The *Sult* family is significantly expanded in the *C. farreri* genome (83 genes versus 26 in the oyster, 31 in the pearl oyster, 13 in humans, and 8 in the fly), and the *Sult* genes showing significant up-regulation in the kidney during *A. minutum* exposure all belong to the clade that is expanded in *C. farreri* or bivalves (Supplementary Fig. 23 and Supplementary Data 10). Collectively, our results suggest that the scallop hepatopancreas and kidney act as two major “centers” for toxin accumulation and transformation, respectively. The expanded *Sult* genes likely participate in conversion of GTXs to more toxic STX (Fig. 5d), which may give the scallop a powerful deterrent against predation, while the novel mutations in *Nav1* revealed in this study may provide scallops and other bivalves with the ability to tolerate those neurotoxins. It seems that dinoflagellates produce neurotoxins to inhibit grazing by filter feeders, but bivalve molluscs have adopted novel sodium channel variants to tolerate neurotoxins and converted the toxins to even more toxic forms for their own defense against predation. Our findings highlight how simple mutations and expansion in one or two key genes may have profound implications for an organism’s adaptation to the environment and the complex interactions with other organisms.

**Discussion**

Bivalves are a fascinating group of animals, which, despite long evolutionary history dating back to the early-Cambrian, are still...
and is not fully differentiated from its smooth muscle. The Methods
adaptation. may have profound effects on an organism expanded family of sulfotransferases, probably as deterrence family of tyrosinases. The scallop uses the hepatopancreas to foot and multiple proteins/enzymes including the expanded (derived from retroposition and tandem duplication) and by ported by the predominant use of novel intronless (generated by means of amino acid combinations at key positions (164, 261 and 269). The whole protein sample as well as major sodium dodecyl sulfate-polyacrylamide gel electrophoresis fractions extracted from byssal adhesive plaques by using the method of Miao et al. were subject to mass spectrometric analysis. The mass spectrometry raw data were searched against the predicted proteins from the C. farreri genome using Mascot v.2.3.0. To be stringent, the identified proteins with ≥2 unique matching peptide in both datasets and with expression ratio[foot/ave_nonfoot_organ] > 2 were excluded from further analysis. Functional annotation of scallop candidate BRPs was performed by searching against SwissProt, Pfam, IntePro, SMART, and SignalP databases. Microstructures of the byssal thread were examined by scanning electron microscopy (Hitachi S-3400N). Forty-five RNA-seq libraries covering three foot regions (proximal, middle, and distal) and five time points after the removal of byssal threads (0, 1, 1.5, 12 and 24h) were subject to Illumina sequencing and the overrepresented genes in each foot region were identified by DEG analysis using the edger package. A nitroblue tetrazolium staining assay was performed on the whole byssal threads, and a catechol oxidative assay for in situ detection of tyrosinase activity. For phylogenetic analysis of tyrosinases, a maximum likelihood (ML) tree was constructed using RAxML and the robustness of the tree was tested by reanalysis of 1000 bootstrap replicates. The voltage-gated sodium channel protein (Nav) sequences of C. farreri and other bivalves were identified via homology-based searches with an e-value threshold of 1e−5. Amino acids positions putatively conferring PST and TTX resistance were identified based on conservation of previously reported sites. Thirty-six RNA-seq libraries of the hepatopancreas and kidney from scallops fed with toxic A. minutum were subject to Illumina sequencing, and DEGs were identified using R package edgeR. The co-expression gene networks for the hepatopancreas and kidney were constructed using the R package WGCNA and over-representation analysis of the TR genes was performed for each module by a hypergeometric test (p < 0.05) to identify TR modules. GO enrichment analysis of each TR module in the networks was conducted using the EnrichPipeline. The cytosolic sulphotransferase (Sult) genes were identified in the genomes of three bivalves, H. sapiens and D. melanogaster using BLAST with an e-value threshold of 1e−5. The ML tree of SULTs and other bivalves was constructed using the PhyloTreePruner and the robustness of the tree was tested by reanalysis of 1000 bootstrap replicates. More detailed description of the above methods can be found in the Supplementary Information.

Data availability. This genome project has been registered in NCBI under the BioProject accession PRJNA185465. The sequencing data of C. farreri have been deposited in NCBI Sequence Read Archive under the accession numbers of SRX1301075, SRX2486272, SRX2486273, SRX2486281, SRX2486284, SRX2486300, SRX2913253-SRX2913260 for genomic data; and SRX2448484-SRX2448476, SRX2508197-SRX2508199, SRX2444668-SRX2444668, SRX2444950-SRX2444979 and SRX2445045-SRX2445440 for transcriptomic data. The proteomic data have been deposited in PRIDE Archive database under the accession numbers PXD007932 and PXD007987. The C. farreri genome assemblies (including an updated version improved by the addition of ~ 26 Gb PacBio data), gene sequences, and annotation data are available at the scallop genome website (http://mgb.ouc.edu.cn/cfbase/html/).

Received: 8 March 2017 Accepted: 26 October 2017
Published online: 23 November 2017

References
1. Barnosky, A. D. et al. Has the Earth’s sixth mass extinction already arrived? Nature 471, 51–57 (2011).
2. Appelans, W. et al. The magnitude of global marine species diversity.Curr. Biol. 19, 2010–2011 (2009).
3. Zhang, G. et al. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490, 49–54 (2012).
4. Takeuchi, T. et al. Bivalve-specific gene expansion in the pearl oyster genome: shell implications of adaptation to a sessile lifestyle. Zool. Lett. 2, 3 (2016).
5. Wang, S. et al. Scallop genome provides insights into evolution of bilaterian karyotype and development. Nat. Ecol. Evol. 1, 0120 (2017).
6. Sun, J. et al. Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. Nat. Ecol. Evol. 1, 0121 (2017).
7. Du, X. et al. The pearl oyster Pinctada fucata Martensii genome and multi-omic analyses provide insights into bimolecularization. GigaScience 6, 1–12 (2017).
8. Guderley, H. E. & Tremblay, I. In Scallop: Biology, Ecology, Aquaculture and Fisheries (eds Shumway, S. E. & Parsons, G. J.) Ch. 12, 535–566 (Elsevier, Amsterdam, 2016).
9. Stemmler, D. I. & Alten, L. A. In Scallop: Biology, Ecology, Aquaculture and Fisheries (eds Shumway, S. E. & Parsons, G. J.) Ch. 5, 219–252 (Elsevier, Amsterdam, 2016).
10. Shumway, S. E. A review of the effects of algal blooms on shellfish and aquaculture. J. World Aquacult. Soc. 21, 65–104 (1990).
11. Bates, H. A. & Rapoport, H. Chemical assay for saxitoxin, the paralytic shellfish poisons. J. Agr. Food Chem. 23, 237–238 (1975).
12. Guo, X. & Luo, Y. In Scallop: Biology, Ecology, Aquaculture and Fisheries (eds Shumway, S. E. & Parsons, G. J.) Ch. 22, 937–952 (Elsevier, Amsterdam, 2016).
13. Zhu, M., Zou, Y., Wu, R. & Hall, S. Accumulation and depuration of paralytic shellfish poisons (PSP) in Chinese scallop Chlamys farrei. Acta Oceanol. Sin. 25, 137–144 (2003).
14. Takatani, T., Akaeda, H., Arakawa, O. & Noguchi, T. Occurrence of paralytic shellfish poisoning (PSP) toxins in different shellfish species. J. Biochem. 151, 154–160 (2012).
15. Tian, H. et al. Comparative study on in vitro transformation of paralytic shellfish poisoning (PSP) toxins in different shellfish species. J. Biochem. 151, 161–167 (2012).
16. Guo, X., Ford, S. E. & Zhang, F. Molluscan aquaculture in China. Bull. Mich. Acad. Sci. 95, 29–35 (2006).
17. Baylies, M. K. & Bate, M. Twist: a myogenic switch in the domestication of sweetness/acidity taste in fruit trees. Int. J. Dev. Biol. 50, 259–266 (2006).
18. Huang, J. et al. The Jujube genome provides insights into genome evolution and the evolution of sweet taste in fruit trees. Nat. Commun. 6, 1721 (2015).
19. Obinata, T., Ooi, A. & Takano-Ohmuro, H. Myosin and actin from ascidian skeletal muscle. J. Biochem. 151, 2275–2281 (2002).
20. Baynes, B. L. Phenotypic flexibility and physiological tradeoffs in the feeding and growth of marine bivalve mussels. Integr. Comp. Biol. 44, 425–432 (2004).
21. Brunet, T. The evolutionary origin of bilaterian smooth and striated myocytes. Elife 5, e19607 (2016).
22. Nyströy, L., Jansson, A., Ochiai, Y., Graf, L. & Szent-Györgyi, A. G. Scallop striated and smooth muscle myosin heavy-chain isoforms are produced by alternative RNA splicing from a single gene. Proc. Natl. Acad. Sci. USA 91, 12866–12869 (1994).
23. Obinata, T. & Nishita, K. Isolation of troponins from striated and smooth adductor muscles of Akazara scallop. J. Biochem. 100, 821–824 (1986).
24. Endo, T. & Obinata, T. Troponin and its components from ascidian smooth muscle. J. Biochem. 89, 1599–1608 (1981).
25. Obinata, T., Ooi, A. & Takano-Ohmuro, H. Myosin and actin from ascidian smooth muscle and their interaction. Comp. Biochem. Phys. B 87, 437–442 (1983).
26. Koyabashi, C. et al. Identification of two distinct muscles in the planarian Dugesia japonica by their expression of myosin heavy chain genes. Zool. Sci. 15, 861–869 (1998).
27. Sulbaran, G. et al. An invertebrate smooth muscle with striated muscle myosin filaments. Proc. Natl. Acad. Sci. USA 112, E5660–E5668 (2015).
28. Witchley, J. N. et al. Muscle cells provide instructions for planarian regeneration. Cell Rep. 4, 633–641 (2013).
29. Land, M. F. & Nilsson, D. E. Animal Eyes. (Oxford University Press, Oxford, 2002).
30. Serfaty, A. A. The opsins. Genome Biol. 6, 213 (2005).
31. Fain, G. L., Hardie, R. & Laughlin, S. B. Phototransduction and the evolution of photoreceptors. Curr. Biol. 20, R114–R124 (2010).
32. Gentles, A. J. & Karlin, S. Why are human G-protein-coupled receptors predominantly intronless? Trends Genet. 15, 47–49 (1999).
33. De Renzi, S., Elemento, O., Tavaozio, S. & Wieschaus, E. F. Unmasking activation of the rgg1 gene using chromosomal deletions in the Drosophila embryo. PLoS Biol. 5, e117 (2007).
34. Gulgur, L. G. et al. Requirement for highly efficient pre-mRNA splicing during Drosophila early embryonic development. eLife 3, e02184 (2014).
35. Speiser, D. I., Loew, E. R. & Johnsen, S. Spectral sensitivity of the concave mirror eyes of scallops: potential influences of habitat, self-screening and longitudinal chromatic aberration. J. Exp. Biol. 214, 422–431 (2011).
36. Barber, V. C., Evans, E. M. & Land, M. F. The fine structure of the eye of the mussels Pecten jacobaeus and Z. zehleri. Mikrosk. Mikrosch. 76, 251–312 (1987).
37. Morton, B. The evolution of eyes in the Bivalvia: new insights. Am. Malacol. Bull. 26, 35–45 (2008).
38. Salvini-Plawen, L. V. Photoreception and the phylophylogenetic evolution of photoreceptors (with special reference to Mollusca). Am. Malacol. Bull. 26, 83–100 (2008).
39. Cronly-Dillon, J. R. Spectral sensitivity of the scallop Pecten maximus. Science 151, 345–346 (1966).
40. Kuo, C. H., Yamagata, K., Moriyz, R. K., Bitensky, M. W. & Miki, N. Multiple opsin mRNA species in bovine retina. Brain Res. 387, 251–280 (1986).
41. Yokoyama, S. Molecular evolution of color vision in vertebrates. Gene 300, 69–78 (2002).
42. Hennebert, E., Maldonado, B., Ladurme, P., Flammang, P. & Santos, R. Experimental strategies for the identification and characterization of adhesive proteins in animals: a review. Interface Focus 5, 20140064 (2015).
43. Aleandrino, A., Poslednik, L. & Serb, J. M. Convergent and parallel evolution in life habit of the scallops (Bivalvia: Pectinidae). BMC Evol. Biol. 11, 164 (2011).
44. Zhao, H. & Waite, J. H. Linking adhesive and structural proteins in the attachment plate of Mytilus californianus. J. Biol. Chem. 281, 26150–26158 (2006).
45. Miao, Y. et al. Integration of transcriptomic and proteomic approaches provides a core set of genes for understanding of scallop attachment. Mar. Biotechnol. (NY) 17, 523–532 (2015).
46. Silverman, H. G. & Roberto, F. Understanding marine mussel adhesion. Mar. Biotechnol. 9, 661–681 (2007).
47. Qin, C. L. et al. In-depth proteomic analysis of the byssus from marine mussel Mytilus coruscus. J. Proteomics 144, 87–98 (2016).
48. Guiraud, L. D. The byssal and byssus glands in Chlamys islandica and other scallops (Lamellibranchia). Zoolog. Sci. 7, 277–285 (1987).
49. Guerette, P. A. et al. Accelerating the design of biomimetic materials by integrating RNA-seq with proteomics and materials science. Nat. Biotechnol. 31, 908–913 (2013).
50. McMillan, D. R., Kayes-Wandover, K. M., Richardson, J. A. & White, P. C. Very large G protein-coupled receptor-1, the largest known cell surface protein, is highly expressed in the developing central nervous system. J. Biol. Chem. 277, 785–792 (2002).
51. Blaustein, M. P. & Lederer, W. J. Sodium/calcium exchange: its physiological implications. Physiol. Rev. 79, 763–854 (1999).
52. Bricelj, V. M. et al. Sodium channel mutation leading to saxitoxin resistance in clams increases risk of PSP. Nature 434, 763–767 (2005).
Acknowledgements

We thank Rencheng Yu, Fanzhong Kong and Yang Liu for facilitating the preparation of PST-seeding scallop materials and helpful comments on the manuscript. We acknowledge grant support from the National Natural Science Foundation of China (31310054, 31322055, 31630081, 31502162 and 31472258), the National High Technology Research and Development Program of China (863 program; 2012AA92204 and 2012AA10A405), the Fundamental Research Funds for the Central Universities (201762001 and 201564009), the Taishan Scholar Project Fund of Shandong Province of China, Chang Jiang Scholars Program of Ministry of Education, and Qingdao National Laboratory for Marine Science and Technology (AoShan Talents Program: 2015ASTP-ES02; Blue Life Program: MS2017RNO01). X.G. acknowledges support from Taishan Overseas Scholar Program of Shandong and USDA-NIFA/NIAES Project 1004475/NJ32920. We thank Xunshan Group Co. Ltd for providing scallop materials and facilities.

Author contributions

Z.B. and S.W. conceived the study and designed major scientific objectives. Z.B., S.W., R. L., Z.J. and J.Z coordinated the whole project. F.W., Lu.Z., Q.X. and H.D. prepared scallop materials for genome and transcriptome sequencing. X.F. and C.M. conducted DNA extraction, and D.W. prepared large-insert genomic libraries. X.F., Xl.H., Wei.L., Jia.L., J.D. and Yp.L. participated initial genome analysis. X.S., J.Z., Jll., G.X. and X.Y. conducted genome sequencing, assembly, annotation, and gene family analysis. W.J. and Y.X. conducted library preparation for genome resequencing, and LLZ., J.W., Y.S., XX., Y.M., X.L., Yi.J. and D.K. conducted mRNA extraction and library preparation for transcriptome sequencing. YL.L., XS. and G.X. participated in genome polymorphism analysis. YLL., SW., TC., ZY. and Yr.L. participated in muscle genes and energy pathway analysis, and JW. in gene network construction. SW., Y.S., YLL., TC. and XN. participated in opsin sequence and eye transcriptome analysis. W.L., YLL., SW. and XZ. participated in byssal proteins and foot transcriptome analysis, and Wz.L. and Y.M. designed and performed byssus-related experiments. Xl.H., X.X. and S.W. participated in Nuv sequence and toxin-related transcriptome analysis. Y.W. provided computational services and technical support. X.G. participated in final data analysis and interpretation. S.W., YLL., X.G. and Z.B. did most of the writing with input from other authors.

Additional information

Supplementary Information accompanies this paper at doi:10.1038/s41467-017-01927-0.

Competing interests: The authors declare no competing financial interests.

Reprints and permission information is available online at http://npg.nature.com/reprintsandpermissions/