

ARTICLE

DOI: 10.1038/s41467-017-01927-0

OPEN

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

Yuli Li¹, Xiaoqing Sun², Xiaoli Hu^{1,3}, Xiaogang Xun¹, Jinbo Zhang², Ximing Guo^{1,4}, Wenqian Jiao¹, Lingling Zhang^{1,3}, Weizhi Liu^{1,5}, 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, Jinzhuang Dou¹, Yangping Li¹, Dexu Kong¹, Yaran Liu¹, Zhi Jiang², Ruiqiang Li², Shi Wang^{1,5} & Zhenmin Bao^{1,3}

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 farreri*, 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.

¹Key Laboratory of Marine Genetics and Breeding (Ministry of Education), Ocean University of China, Qingdao 266003, China. ²Novogene Bioinformatics Institute, Beijing 100083, China. ³Laboratory for Marine Fisheries Science and Food Production Processes, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China. ⁴Haskin Shellfish Research Laboratory, Department of Marine and Coastal Sciences, Rutgers University, Port Norris, NJ 08349, USA. ⁵Laboratory for Marine Biology and Biotechnology, Qingdao National Laboratory for Marine Science and Technology, Qingdao 266237, China. Yuli Li, Xiaoqing Sun, Xiaoli Hu, Xiaogang Xun, Jinbo Zhang and Ximing Guo contributed equally to this work. Correspondence and requests for materials should be addressed to S.W. (email: swang@ouc.edu.cn) or to Z.B. (email: zmbao@ouc.edu.cn)

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 events¹. Yet, extant bivalves with ~9600 species² 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^{3–7}, 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 Yesso scallop *Patinopecten yessoensis*⁵. Scallops have a large adductor muscle, probably as an adaptation to swimming by clapping valves for avoiding predation and seeking favorable habitats⁸. 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 swimming⁹. 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^{10,11}. 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 Zhikong 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)¹² 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 habitats¹². 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 human¹³) and therefore is widely used for studying PST accumulation and transformation^{13–15}. *C. farreri* is also a commercially important bivalve with aquaculture production once reaching ~1 million metric tons¹⁶. It is among the best genetically characterized bivalve species with available linkage, physical and cytogenetic maps^{17–21}, fosmid and bacterial artificial chromosome (BAC) libraries^{22, 23}, and a large number of expressed sequence tags^{24, 25}, making it a good candidate for whole-genome sequencing (WGS).

Here we report the sequencing and analysis of the genome of *C. farreri* along with a comprehensive set of 117 transcriptomes and proteomes covering various organs, development stages, and characteristics of scallop biology. Our multi-omic analyses and associated assays revealed novel genomic features and molecular changes that may underlie aspects of the scallop's adaptation to semi-sessile and filter-feeding life including the well-developed adductor muscle, sophisticated photoreception system, rapid byssal production, and remarkable resistance to potent neurotoxins.

Results

Genome sequencing and characterization. Genomes of bivalves including *C. farreri* are challenging to sequence and assemble due to their exceptionally high genome heterozygosity^{3, 5, 20}. To assemble the highly polymorphic scallop genome, a modified SOAPde novo approach²⁶ was used to resolve the complex bubble structures resulting from high genome heterozygosity. Deep sequencing of a 2-year-old *C. farreri* from the Penglai-Red selectively bred population produced 362.8 Gb of clean sequences, with average genome coverage of 382× (Supplementary Table 1). The genome assembly is 779.9 Mb long with a contig N50 size of 21.5 kb and a scaffold N50 size of 602 kb (Supplementary Table 2), and over 80% of the assembly is covered by the longest 1098 scaffolds (>142 kb) (Supplementary Table 3). The assembly statistics of the scallop genome are comparable to or better than those of previously published bivalve genomes (contig N50: 19–38 kb and scaffold N50: 167–804 kb)^{3–7}. 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 cytometry²³. 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 Fig. 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 map²⁰.

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 (≥500 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 development²⁷. 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.

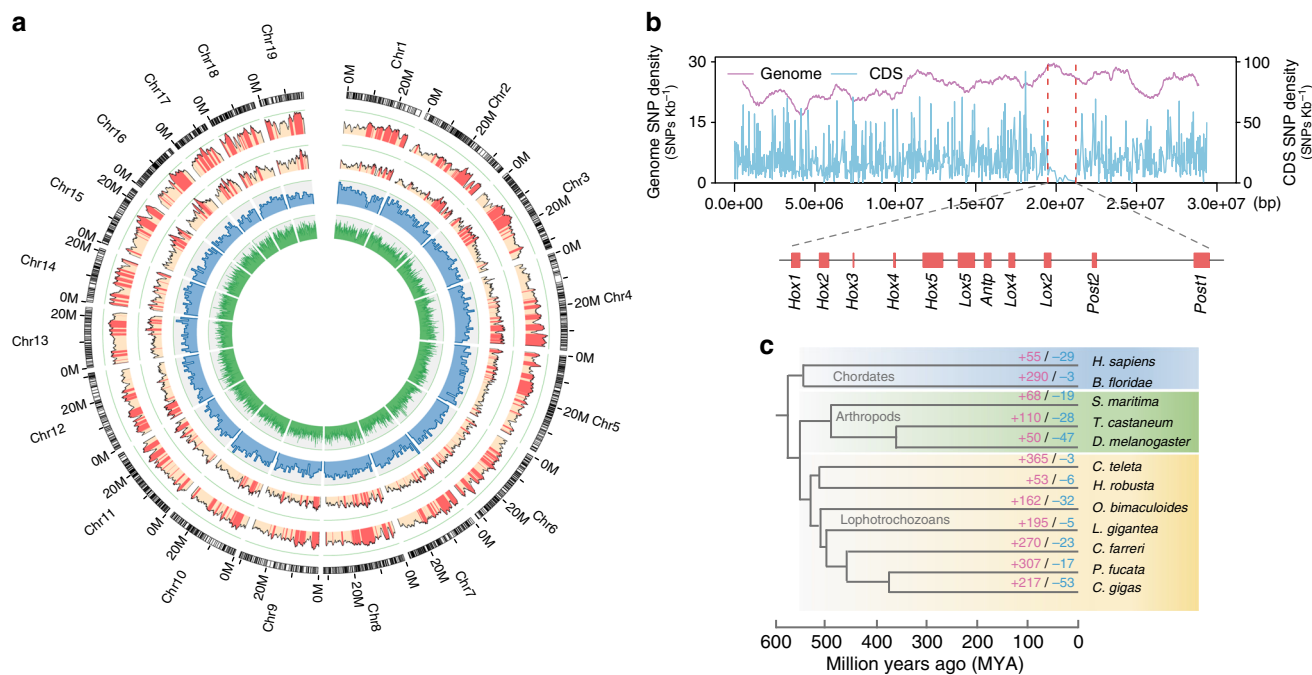


Fig. 1 Genome landscape and polymorphism analysis of the scallop *C. farreri*. **a** From outer to inner circles: I, marker distribution on 19 chromosomes at the Mb scale; II and III, single-nucleotide polymorphism (SNP) density across genome (II) or CDS (III) was drawn in 1 Mb sliding windows with a 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

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 Gastropoda²⁸ ~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 brachiopod *Lingula anatina* (7788), a “living fossil” lophotrochozoan²⁹. Compared with other bivalves, 270 gene families are significantly expanded in the scallop lineage (Fig. 1c and Supplementary Data 1) and are predominantly involved in neurotransmission, immune responses, signal transduction, and xenobiotic metabolism (Supplementary Table 13). These expanded gene families are probably important for the scallop’s lineage-specific adaptations and biology. The notable expansion of sodium- and chloride-dependent neurotransmitter transporters in *C. farreri* (61 versus 20–28 in other bivalves; Supplementary Fig. 9) may underlie the scallop’s more developed nervous and vision systems, and higher locomotion activity than other bivalves⁹.

Muscle regulation and evolution. Scallops have a remarkably large adductor muscle (Supplementary Fig. 10) compared to most of sessile and endobenthic bivalves 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)³⁰, 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 cost⁸. 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;

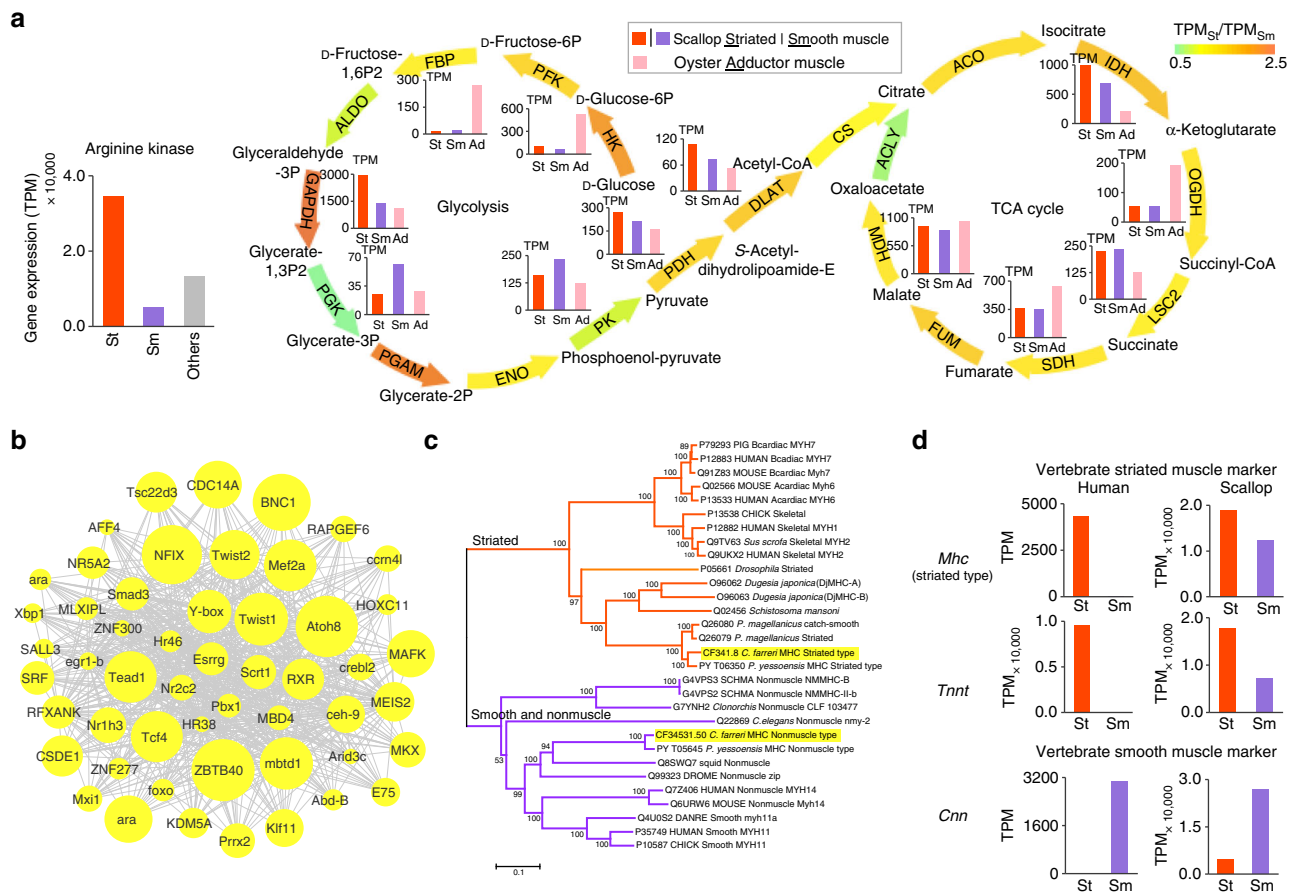


Fig. 2 Molecular characterization and evolutionary analysis of scallop adductor muscle. **a** Expression of genes involved in various energy-producing reactions or pathways in striated (ST) and smooth (SM) muscles of *C. farreri* and in the adductor muscle of oyster *C. gigas*, including hydrolysis of phosphoryl arginine, glycolysis, and TCA cycle. The sum of TPM values of genes encoding the same enzyme was defined as the expression level of a given enzyme. Enzymes participating in energy (ATP/energy-rich H^+) production reactions (GAPDH, PGK, PK, PDH, DLAT, IDH, OGDH, LSC2, SDH, and MDH) generally show higher expression in the scallop *C. farreri* than in the oyster *C. gigas*, whereas the enzymes participating in energy consumption reactions (HK and FBP) show lower expression in *C. farreri* than in *C. gigas* (Supplementary Table 14). Abbreviations of enzymes: HK, hexokinase; PFK, 6-phosphofructokinase; FBP, fructose-1,6-bisphosphatase; ALDO, fructose-bisphosphate aldolase; GAPDH, glyceraldehyde 3-phosphate dehydrogenase; PGK, phosphoglycerate kinase; PGAM, phosphoglycerate mutase; ENO, enolase; PK, pyruvate kinase; OGDH, octopine dehydrogenase; PDH, pyruvate dehydrogenase; DLAT, dihydrolipoamide acetyltransferase; CS, citrate synthase; ACO, aconitate hydratase; IDH, isocitrate dehydrogenase; OGDH, sucA 2-oxoglutarate dehydrogenase; LSC2, succinyl-CoA synthetase β subunit; SDH, succinate dehydrogenase; FUM, fumarate hydratase; MDH, malate dehydrogenase; ACLY, ATP citrate (pro-S)-lyase. **b** The co-expression network of transcription factors (TFs) in the adductor muscle-related module (M3, see Supplementary Data 3). Node size represents the intramodular connectivity of a given gene. **c** Phylogenetic analysis of scallop myosin heavy-chain (*Mhc*) genes. Genes highlighted in yellow are present in the *C. farreri* genome. The scale bar represents 0.1 amino acid substitutions per site. **d** Expression profiles of vertebrate striated muscle marker *Mhc*, *Tnnt*, and *Cnn* genes in the adductor muscle of *C. farreri*. The human expression data were obtained from the Human Protein Atlas database (<http://www.proteinatlas.org/>). Expression profiles of striated *Mhc*, *Tnnt*, and *Cnn* are shown and other marker genes (*Mrlc*, *Tnni*, *Ttn*, and *Zasp*) are available in Supplementary Fig. 14

Supplementary Data 2 and Supplementary Fig. 12). Analysis of transcription factors (TFs) in M3 suggests that *Twist*, *Nfix* and *Zbtb40* are among the top-ranked TFs with the highest intramodular connectivity (Fig. 2b and Supplementary Data 3), with the former two known as key TFs in animal myogenesis^{31, 32}. *Twist*, the master regulator of myogenesis, loses expression in adult muscles of *Drosophila*^{31, 33} and shows high expression in the adductor muscle of the adult scallop (Supplementary Fig. 13), possibly related to the different modes of adult muscle growth: determinate for *Drosophila*³⁴ and indeterminate for the scallop³⁵.

It is also interesting that key marker genes³⁶ that distinguish vertebrate striated muscles (striated *Mhc* (myosin heavy chain), *Tnnt*, *Tnni*, *Ttn*, and *Zasp*) from smooth muscles (*Cnn*) show high expression in both striated and smooth muscles of *C. farreri* (Fig. 2c, d, Supplementary Fig. 14, and Supplementary Table 15), revealing the “hybrid” nature of scallop 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}, ascidians^{39, 40}, and flatworms^{41–43}. 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 organizations⁴¹.

Opsin 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 molluscs⁴⁴. Opsins of the G-protein-coupled receptor (GPCR) family are key light-sensing proteins

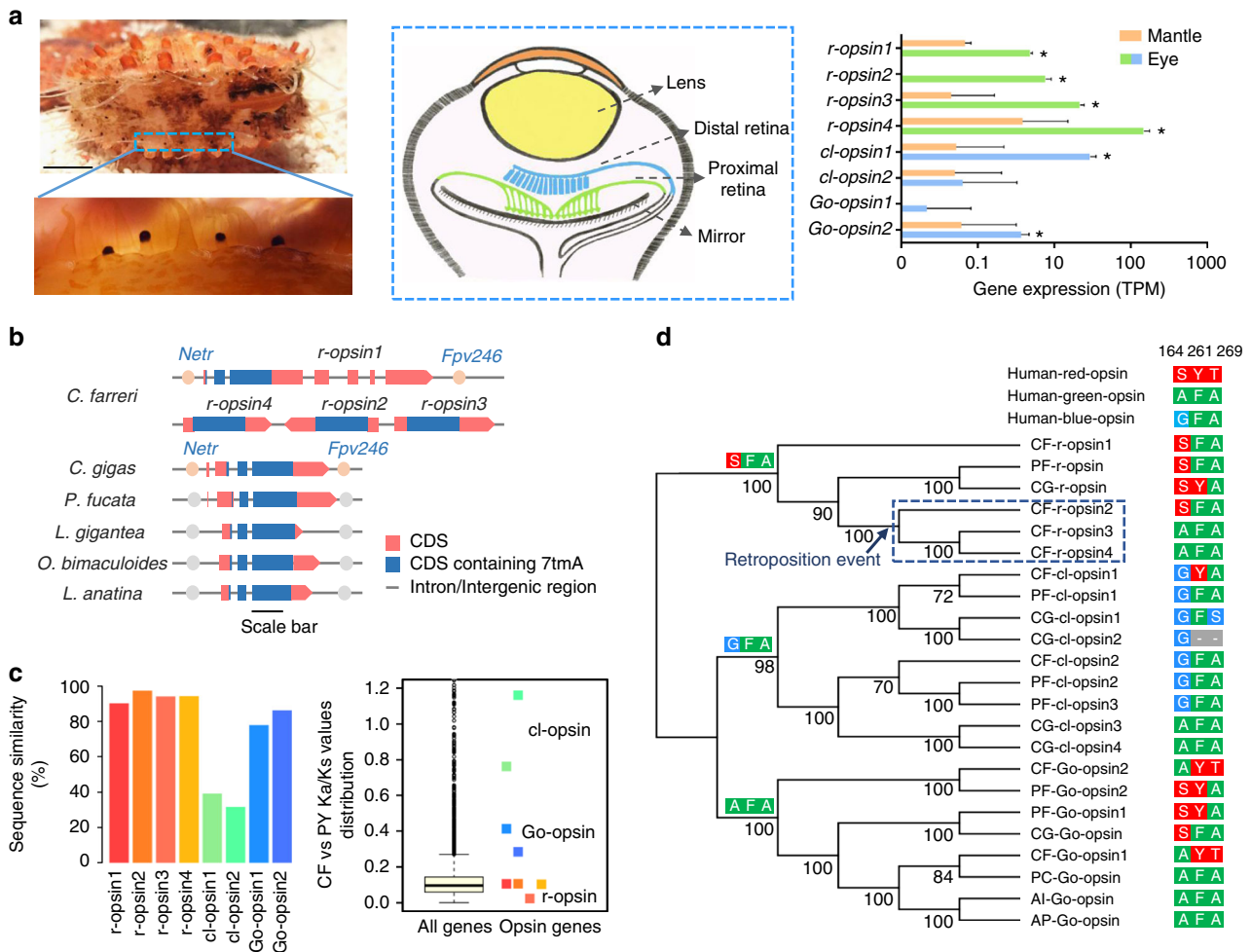


Fig. 3 The evolution of opsin diversity and photoreception tuning in *C. farreri*. **a** Morphology of scallop eyes scattered along the edge of mantles (left, scale bar: 1 cm), schematic structure of a typical scallop eye (middle), and expression of diverse opsin genes in scallop eyes (right). Eye samples 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. **c** Sequence similarity and Ka/Ks values of all opsin genes between the scallops of *C. farreri* (CF) and *P. yessoensis* (PY). 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. **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 transduction⁴⁵. 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 *cl-opsin* genes (Supplementary Fig. 15). As the characteristic opsin type in invertebrates⁴⁶, *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) transcript⁴⁷. 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 intronless⁴⁸, 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 behaviors⁵⁰. The proximal retina consists of rhabdomeric photoreceptor cells (also found in most of invertebrate eyes), whereas the distal retina consists of ciliary photoreceptor cells

(similar to rod and cone cells of vertebrates)⁵¹. Key genes participating in rhabdomeric and ciliary phototransduction pathways were identified here in the *C. farreri* genome, and higher gene expression of the Gq-coupled rhabdomeric pathway implies that the proximal retina plays prominent roles in the scallop visual system (Supplementary Table 16). The evolutionary origins of the two layers of the retina remain enigmatic. It has been suggested that the distal retina might have evolved later than the proximal retina^{51, 52} and thus is likely under relaxed selection pressure. This hypothesis is, to some extent, supported by our analysis of opsin genes from two scallop species (*C. farreri* and *P. yessoensis*): ciliary opsins expressed in the distal retina showed significantly lower sequence conservation (*t*-test *p*-value = 0.04), greater Ka/Ks values (*t*-test *p*-value = 0.03), and much weaker transcription relative to rhabdomeric opsins expressed in the proximal retina (Fig. 3a, c). The presence (addition or retention) of the otherwise vertebrate-specific distal retina may represent an evolutionary innovation giving the scallop the potential to form image and detect movement⁵³.

Scallop eyes can detect a wide spectrum of light wavelengths ($\lambda_{\text{max}} = 480\text{--}540\text{ nm}$ ⁵⁴), with the proximal retina sensitive to short wavelengths, whereas the distal retina to long wavelengths⁵⁰. Various studies have shown that three key amino acid positions in opsins (164, 261, and 269 in reference to rhodopsin of *Bos taurus*⁵⁵) are crucial determinants of spectral sensitivity to short or long wavelengths⁵⁶. *R-opsin3* and *r-opsin4*, the two opsins derived from gene duplication and most highly expressed in the proximal retina of *C. farreri*, have “AFA” at these sites (just as human green opsin does; Fig. 3d), adding a potentially valuable variant to the ancestral “SFA” type found in the scallop (*r-opsin1* and *r-opsin2*) and other bivalves. Apparently, the gene duplication increased the diversity of the r-opsin protein at core functional sites in its amino acid sequence; this change may broaden the scallop’s spectral sensitivity. Similarly, new functional site variants of Go-opsin (AYT) and cl-opsin (GYA) were observed in *C. farreri* (Fig. 3d) that may allow the scallop to sense different ranges of long-wavelength light because SYT in humans corresponds to red-light sensitivity⁵⁶. These results show that scallop eyes are not only numerous and structurally advanced but also equipped with a diverse set of opsins including novel intronless genes resulting from retroposition and tandem duplication, and new opsin variants that are polymorphic at functional sites in amino acid sequence and may broaden spectral sensitivity. These notable molecular features may provide scallops with enhanced vision or light sensitivity as part of their adaptation to epibenthic semi-sessile life.

Byssal proteins and secretion regulation. Many aquatic animals including bivalves, barnacles, and sandcastle worms have evolved effective strategies for adhesion as adaptation to turbulent currents⁵⁷. As an ancestral feature of bivalves⁵⁸, byssal attachment is an essential part of the sessile or semi-sessile lifestyle characteristic of diverse bivalve families, but, to date, has been extensively studied only in mussels^{59–61}. In contrast to mussels, scallop byssal attachment is characterized by abundant secretion of byssal proteins and temporary attachment (in concert with swimming behavior; Fig. 4a). The scallop byssus has different ultrastructure and morphology in different regions (Fig. 4a and Supplementary Fig. 16), indicating compositional and mechanical complexity. Mass spectrometric analysis identified 16 candidate byssus-related proteins (BRPs) in the whole byssal adhesive plaques of *C. farreri* (Supplementary Figs. 17 and 18 and Supplementary Data 4), including seven previously identified scallop byssal proteins (SBPs⁶²). Functional annotation of these BRPs suggests that they potentially involve in oxidative reactions (tyrosinase and

peroxidase), extracellular matrix consolidation (tenascin-X), and anti-biodegradation (serine protease inhibitor and metalloproteinase inhibitor) (Fig. 4b, Supplementary Data 4, and Supplementary Fig. 19). Of the identified scallop BRPs, none shows protein similarity to a well-known set of 11 mussel BRPs⁶³, but up to eight to an expanded set of 48 mussel BRPs recently identified by Qin et al.⁶⁴ (Supplementary Data 4). Four scallop BRPs (CF48907.12, CF47691.7, CF44339.32, and CF30077.9) that do not have either protein/domain annotations or similarity to mussel BRPs likely represent novel SBPs. Of the 16 identified BRPs, 12 show high and specific expression in the foot of *C. farreri* but nearly no expression in the foot of the adult Yesso scallop *P. yessoensis*, a species that is free living and does not produce byssi in adulthood (Fig. 4c and Supplementary Table 17). Ka/Ks analysis indicates that these BRPs diverged more rapidly than other BRPs did (Fig. 4c), reflecting differential selection in the two scallop species possibly because of different requirements for byssal attachment at the adult stage.

Byssal secretion is a complicated process involving a series of biochemical reactions occurring in various byssal and enzymatic glands located in different regions of the foot⁶⁵. To investigate the molecular mechanism of byssal secretion, we conducted comprehensive temporal and spatial gene expression profiling by sequencing 45 transcriptomes of three foot regions (proximal, middle, and distal) and at five time points during byssal secretion. Our results show that the three foot regions have different gene expression patterns, reflecting their different roles during byssogenesis (Fig. 4d). The proximal end of the foot, where the primary byssal gland resides and byssal ribbons are secreted, is characterized by high expression levels of connective proteins tenascin-X and matrilin and various related enzymes. The middle foot region where the secondary byssal and/or enzymic glands are located and the byssal ribbon sheath or envelope is formed, predominantly shows expression of a variety of tyrosinases (Fig. 4g): enzymes crucial for mussel byssogenesis that catalyze the formation of a strong adhesive, 3,4-dihydroxyphenylalanine (DOPA)⁶⁶. Tyrosinases’ participation in scallop byssogenesis is supported by the enzymatic activity in foot glands and high abundance of DOPA in byssi (Fig. 4e, f). The distal end of the foot shows dramatically increased transcription of signal transduction-related GPCR98-like proteins^{67, 68} at the initial stage (<1 h) of byssal secretion, and likely plays a role in guiding the search for suitable spots for attachment.

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 potential⁶⁹. We identified two sodium channel genes, *Nav1* and *Nav2*, in the *C. farreri* genome. *Nav1* is the primary sodium channel in animal nervous systems⁷⁰ 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^{72, 73}, *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 *Solemya velum* (Fig. 5a), which has been shown to increase STX resistance up to 19,880-fold in a rat sodium

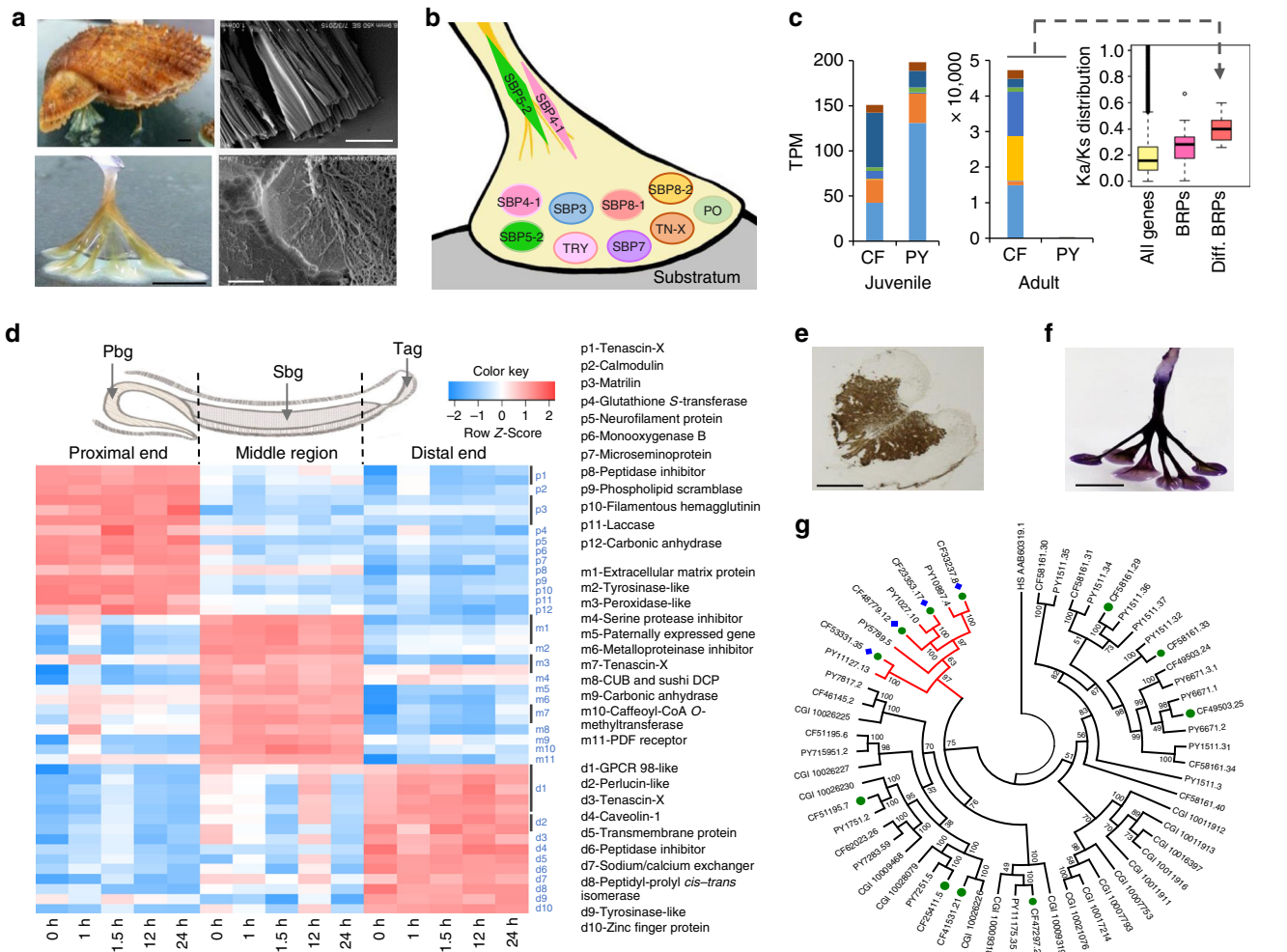


Fig. 4 Scallop byssal protein composition and molecular regulation of byssal secretion. **a** The morphology and microstructures of the byssal threads of *C. farreri*. Scale bars are 5 mm (top-left and bottom-left), 0.5 mm (top-right), and 2.5 μm (bottom-right). The full set of microstructure images across the whole byssus is presented in Supplementary Fig. 16. **b** Schematic illustration of byssal protein composition according to mass spectrometric analysis. TN-X, tenascin-X; TYR, tyrosinase; PO, peroxidase; SBP, scallop byssal proteins identified by Miao et al.⁶². **c** Gene expression and the Ka/Ks substitution rate of foot-specific byssal proteins. Expression was compared between *C. farreri* (CF) and *P. yessoensis* (PY, a species without adult byssal secretion) in the juvenile and adult foot. The Ka/Ks ratio was calculated based on homologous gene pairs from CF and PY for all proteins, all BRPs and the BRPs differentially expressed between the adult feet of CF and PY (Diff. BRPs). 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. The color code correspondence is detailed in Supplementary Table 17. **d** Spatial and temporal expression of top 15 genes that are highly and differentially expressed in proximal, middle, and distal regions of the *C. farreri* foot. The sampling time points refer to the time after removal of old byssal threads. Genes from the same gene family are grouped together (represented by a bar). Abbreviations: Pbg, primary byssal gland; Sbg, secondary byssal gland; Tag, tip attachment gland. **e** A catechol oxidase activity assay in cryosectioned foot glands reveals tyrosinase activity on L-tyrosine as a substrate. **f** Nitroblue tetrazolium staining of byssal threads reveals abundant presence of DOPA. Scale bars represent 1 mm **e** and 5 mm **f**. **g** The maximum likelihood-based phylogeny of bivalve tyrosinases (*Tyr*). The *Tyr* genes expressed in juvenile and/or adult foot (TPM ≥ 2) are indicated by green dots. Those specifically expressed in adult foot (fold change_[foot/ave_nonfoot_organ] = 206-1729) are indicated by blue diamonds, showing closer phylogenetic relationships (in the red cluster). Numbers above the branches are bootstrap values from 1000 replicates. CGI: *C. gigas*.

channel⁷³. 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^{13, 15}, 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

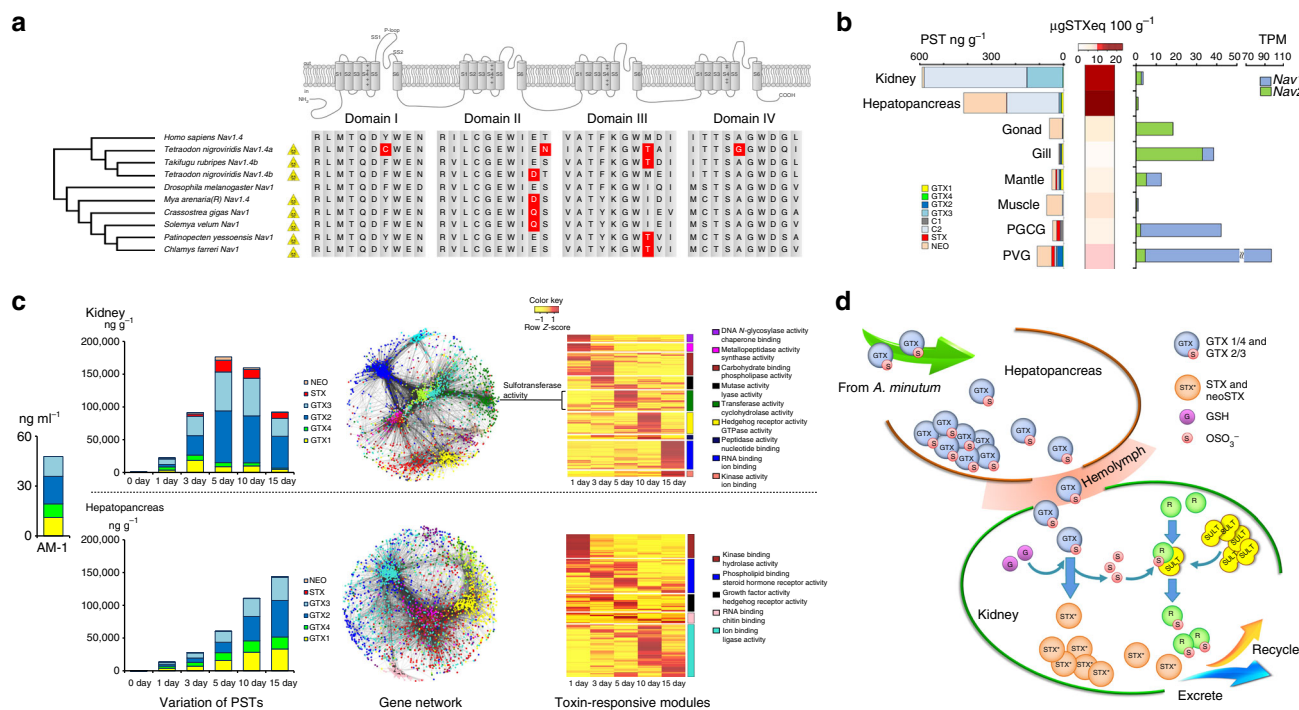


Fig. 5 Toxin resistance and response regulatory networks in the scallop kidney and hepatopancreas. **a** Amino acids conferring PST or TTX resistance on sodium channel *Nav1* (highlighted in red) identified in this and other studies^{69, 75–79}. Skull signs indicate toxin-producing and -resistant species. **b** Expression of sodium channels *Nav1* and *Nav2*, and PST concentration and toxicity levels in major organs of *C. farreri*. Toxicity (μg STX eq. per 100 g) was determined by converting total concentration of PSTs to micrograms of STX equivalents per 100 g of tissue⁸⁰. **c** Temporal variations in abundance of different PSTs and gene networks in the kidney and hepatopancreas after exposure to the toxic alga *Alexandrium minutum*. Toxin response modules were identified by enrichment analysis of DEGs during exposure to *A. minutum* and each module was annotated with the two most significantly enriched GO term(s). The green module is the largest kidney-specific response module, where cytosolic sulfotransferase (*Sult*) genes are enriched (Supplementary Data 10) and highly expressed on day 5 after *A. minutum* exposure (as indicated in the heat map). **d** A schematic diagram showing different roles of the scallop hepatopancreas and kidney in toxin metabolism, with the hepatopancreas primarily responsible for PST accumulation, whereas the kidney primarily for PST transformation mediated by SULTs

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 = 5.9×10^{-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 xenobiotics⁷⁴. 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

abundant and thriving as highly successful filter feeders dominating diverse benthic environments. Their remarkable adaptation to benthic life is not well studied at genomic levels. We sequenced the genome of the Zhikong scallop and collected extensive transcriptomes and proteomes to study molecular or genomic changes related to several of its adaptive features. Our analyses identified significant expansion in 270 gene families that may be important for *C. farreri*'s biology and adaptation. The scallop's large striated muscle shows heightened energy dynamics and is not fully differentiated from its smooth muscle. The sophisticated noncephalic multiple eyes of *C. farreri* are supported by the predominant use of novel intronless *r-opsin* genes (derived from retroposition and tandem duplication) and by diverse opsin variants for possibly broadened spectral sensitivity. The rapid secretion of byssi is enabled by a spatially differentiated foot and multiple proteins/enzymes including the expanded family of tyrosinases. The scallop uses the hepatopancreas to accumulate algae-derived neurotoxins and uses the kidney to transform them into highly toxic compounds by means of the expanded family of sulfotransferases, probably as deterrence against predation, while its own remarkable resistance to neurotoxins may be explained by mutational and expressional modulation of sodium channels. These molecular innovations may be important for the scallop's semi-sessile lifestyle as a filter feeder, suggesting that simple expansion and mutation of a few key genes may have profound effects on an organism's phenotype and adaptation.

Methods

Brief description of methods. The whole genome of a 2-year-old *C. farreri* was sequenced using the Illumina HiSeq 2000 platform through the construction and sequencing of both short-insert (180, 300 and 500 bp) and long-insert (2, 5, 10, 20 and 30 kb) DNA libraries. The genome size of *C. farreri* was estimated based on the 19-mer frequency distribution. To address the problem of high genome heterozygosity, a hierarchical strategy based on a modified version of SOAPdenovo⁴ was used for the assembly of the *C. farreri* genome. The integrity of the final assembly was assessed by means of four data sets: four BAC sequences, WGS data, transcriptome data, and an 843-BUSCO metazoan subset of genes. The assembly was further anchored to chromosomes based on a high-density genetic linkage map¹⁰, through the assignment of the scaffolds to 19 linkage groups. For repeat annotation, tandem repeats were predicted using the software Tandem Repeats Finder¹¹, and TEs were predicted via two approaches (homology-based method and de novo prediction). To predict genes in the *C. farreri* genome, three approaches (homology-based, de novo, and transcriptome-based predictions) were employed. Functional annotation of the protein-coding genes of *C. farreri* was performed by searching the SwissProt, TrEMBL, InterPro, GO (gene ontology), and KEGG (Kyoto Encyclopedia of Genes and Genomes) databases. Thirteen adult tissues/organs of the scallop were chosen for transcriptome sequencing, including striated muscle, smooth muscle, foot, hepatopancreas, kidney, female gonad, male gonad, gill, eyes, mantle, cerebral ganglion, and visceral ganglion. Differentially expressed gene (DEG) analysis was carried out using edgeR²³ with three biological replicates, and genes with a fold-change value ≥ 2 and adjusted *p*-value < 0.05 were defined as significant DEGs. To characterize the polymorphism in the *C. farreri* genome, reads from the sequenced individual and five additional resequenced individuals were aligned to the assembled genome for SNP calling using the BWA⁶ software. Genomic regions or CDs with high SNP density subjected to one-sided Fisher's exact test by comparing to the corresponding chromosomal background, and the distribution of SNP density among chromosomes or genes was visualized using the Circos software²⁶. The OrthoMCL pipeline²⁷ was used to define gene families for the selected species. For phylogenetic analysis, we selected orthologous genes using a tree-based approach PhyloTreePruner²⁸, and the phylogenetic tree was constructed using RAxML²⁹. To estimate the divergence time for *C. farreri* and other metazoans, the first and second codon positions of the orthologs were extracted for Bayesian dating using the MCMCTree program implemented in PAML³¹, with reference divergence time of selected species retrieved from the TimeTree³³ database. The evolutionary dynamics (expansion/contraction) of gene families were analyzed in the software CAFÉ³⁴, and GO enrichment analysis was performed using the EnrichPipeline³⁵. For muscle analysis, we compared the transcript abundance of various enzymes involved in glycolysis, TCA cycle, and oxidative phosphorylation pathways between different types of scallop muscles or between scallop and oyster muscles. Co-expression gene networks were constructed by means of WGCNA³⁶ using 35 transcriptomes from adult tissues/organs, and module enrichment of muscle-overrepresented genes was conducted by the hypergeometric test ($p < 0.05$). The expression profile of vertebrate muscle marker

genes³⁸ in the scallop was determined using the average TPM value of three biological replicates, and the corresponding value in human was obtained from the HPA dataset (<http://www.proteinatlas.org/>). Putative opsin genes in the scallop and other bivalves were identified by BLAST-based searching against known opsin genes of other animal species at an *e*-value threshold of $1e-5$, and only those containing seven transmembrane domains and the lysine site (296 K) were kept for subsequent analysis. The opsin phylogeny was constructed by the Bayesian method³⁹ using the sequences of seven transmembrane domains. Ka/Ks values were estimated by means of Ka_Ks_calculator 2.0⁴⁰ using the YN method. Key genes involved in rhabdomeric and ciliary phototransduction pathways were identified by homology-based search against the known genes from *Homo* and *Drosophila*, and putative light sensitivity of bivalve opsin genes was determined 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.⁴³ 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 ≤ 1 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, InterPro, 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 24 h) were subject to Illumina sequencing and the overrepresented genes in each foot region were identified by DEG analysis using the edgeR package²³. A nitroblue tetrazolium staining assay was performed on the whole byssal threads, and a catechol oxidase assay for *in situ* detection of tyrosinase activity. For phylogenetic analysis of tyrosinases, a maximum likelihood (ML) tree was constructed using RAxML²⁹ 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 $1e-10$. Amino acids positions putatively conferring PST and TTX resistance were identified based on conservation of previously reported sites⁴⁵⁻⁵⁰. 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 edgeR²³. The co-expression gene networks for the hepatopancreas and kidney were constructed using the R package WGCNA³⁶, 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 EnrichPipeline³⁵. The cytosolic sulfotransferase (*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 was constructed using RAxML²⁹ 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 SRX1305705, SRX2486272, SRX2486273, SRX2486281, SRX2486284, SRX2486300, and SRX2913253-SRX2913260 for genomic data; and SRX2444844-SRX2444876, SRX2508197-SRX2508199, SRX2444668-SRX2444682, SRX2444950-SRX2444979 and SRX2445405-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

- Barnosky, A. D. et al. Has the Earth's sixth mass extinction already arrived? *Nature* **471**, 51–57 (2011).
- Appeltans, W. et al. The magnitude of global marine species diversity. *Curr. Biol.* **22**, 2189–2202 (2012).
- Zhang, G. et al. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* **490**, 49–54 (2012).
- Takeuchi, T. et al. Bivalve-specific gene expansion in the pearl oyster genome: implications of adaptation to a sessile lifestyle. *Zool. Lett.* **2**, 3 (2016).
- 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 biomineralization. *GigaScience* **6**, 1–12 (2017).
8. Guderley, H. E. & Tremblay, I. in *Scallops: Biology, Ecology, Aquaculture and Fisheries* (eds Shumway, S. E. & Parsons, G. J.) Ch. 12, 535–566 (Elsevier, Amsterdam, 2016).
9. Speiser D. I. & Wilkens L. A. in *Scallops: 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 poison. *J. Agr. Food Chem.* **23**, 237–239 (1975).
12. Guo, X. & Luo, Y. in *Scallops: 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 farreri*. *Acta Oceanol. Sin.* **25**, 75–83 (2003).
14. Takatani, T., Akaeda, H., Arakawa, O. & Noguchi, T. Occurrence of paralytic shellfish poison (PSP) in bivalves, along with mossworm adherent to their shells, collected from Fukue Island, Nagasaki, Japan during 1995 and 1996. *Food Hyg. Safe Sci. (Shokuhin Eiseigaku Zasshi)* **38**, 430–434 (1997).
15. Tian, H. et al. Comparative study on in vitro transformation of paralytic shellfish poisoning (PSP) toxins in different shellfish tissues. *Acta Oceanol. Sin.* **29**, 120–126 (2010).
16. Guo, X., Ford, S. E. & Zhang, F. Molluscan aquaculture in China. *J. Shellfish Res.* **18**, 19–31 (1999).
17. Wang, S. et al. AFLP linkage map of an intraspecific cross in *Chlamys farreri*. *J. Shellfish Res.* **23**, 491–499 (2004).
18. Zhang, X. et al. A BAC-based physical map of Zhikong scallop (*Chlamys farreri* Jones et Preston). *PLoS ONE* **6**, e27612 (2011).
19. Zhan, A. et al. Construction of microsatellite-based linkage maps and identification of size-related quantitative trait loci for Zhikong scallop (*Chlamys farreri*). *Anim. Genet.* **40**, 821–831 (2009).
20. Jiao, W. et al. High-resolution linkage and quantitative trait locus mapping aided by genome survey sequencing: building up an integrative genomic framework for a bivalve mollusc. *DNA Res.* **21**, 85–101 (2014).
21. Feng, L. et al. An integrated genetic and cytogenetic map for Zhikong scallop, *Chlamys farreri*, based on microsatellite markers. *PLoS ONE* **9**, e2567 (2014).
22. Zhang, L., Bao, Z., Wang, S., Hu, X. & Hu, J. FISH mapping and identification of Zhikong scallop (*Chlamys farreri*) chromosomes. *Mar. Biotechnol. (NY)* **10**, 151–157 (2008).
23. Zhang, L. et al. Fosmid library construction and initial analysis of end sequences in Zhikong scallop (*Chlamys farreri*). *Mar. Biotechnol. (NY)* **9**, 606–612 (2007).
24. Wang, S. et al. Transcriptome sequencing of Zhikong scallop (*Chlamys farreri*) and comparative transcriptomic analysis with Yesso scallop (*Patinopecten yessoensis*). *PLoS ONE* **8**, e63927 (2013).
25. Wang, L. et al. Expressed sequence tags from the zhikong scallop (*Chlamys farreri*): discovery and annotation of host-defense genes. *Fish Shellfish Immunol.* **26**, 744–750 (2009).
26. Huang, J. et al. The Jujube genome provides insights into genome evolution and the domestication of sweetness/acidity taste in fruit trees. *PLoS Genet.* **12**, e1006433 (2016).
27. Krumlauf, R. Hox genes in vertebrate development. *Cell* **78**, 191–201 (1994).
28. Kocot, K. M. et al. Phylogenomics reveals deep molluscan relationships. *Nature* **477**, 452–456 (2011).
29. Luo, Y. J. et al. The Lingula genome provides insights into brachiopod evolution and the origin of phosphate biomineralization. *Nat. Commun.* **6**, 8301 (2015).
30. de Zwaan, A., Thompson, R. J. & Livingstone, D. R. Physiological and biochemical aspects of the valve snap and valve closure responses in the giant scallop *Placopecten magellanicus*. *J. Comp. Physiol.* **137**, 105–114 (1980).
31. Baylies, M. K. & Bate, M. Twist: a myogenic switch in *Drosophila*. *Science* **272**, 1481–1484 (1996).
32. Messina, M. et al. Nfix regulates fetal-specific transcription in developing skeletal muscle. *Cell* **140**, 554–566 (2010).
33. Arbeitman, M. N. et al. Gene expression during the life cycle of *Drosophila melanogaster*. *Science* **297**, 2270–2275 (2002).
34. Augustin, H. & Partridge, L. Invertebrate models of age-related muscle degeneration. *Biochim. Biophys. Acta* **1790**, 1084–1094 (2009).
35. Bayne, B. L. Phenotypic flexibility and physiological tradeoffs in the feeding and growth of marine bivalve molluscs. *Integr. Comp. Biol.* **44**, 425–432 (2004).
36. Brunet, T. The evolutionary origin of bilaterian smooth and striated myocytes. *Elife* **5**, e19607 (2016).
37. Nyitrai, L., Jancso, A., Ochiai, Y., Graf, L. & Szent-Gyorgyi, 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**, 12686–12690 (1994).
38. Ojima, T. & Nishita, K. Isolation of troponins from striated and smooth adductor muscles of Akazara scallop. *J. Biochem.* **100**, 821–824 (1986).
39. Endo, T. & Obinata, T. Troponin and its components from ascidian smooth muscle. *J. Biochem.* **89**, 1599–1608 (1981).
40. Obinata, T., Ooi, A. & Takano-Ohmuro, H. Myosin and actin from ascidian smooth muscle and their interaction. *Comp. Biochem. Phys. B* **76**, 437–442 (1983).
41. Kobayashi, 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).
42. Sulbarán, G. et al. An invertebrate smooth muscle with striated muscle myosin filaments. *Proc. Natl. Acad. Sci. USA* **112**, E5660–E5668 (2015).
43. Witchley, J. N. et al. Muscle cells provide instructions for planarian regeneration. *Cell Rep.* **4**, 633–641 (2013).
44. Land, M. F. & Nilsson, D. E. *Animal Eyes*. (Oxford University Press, Oxford, 2002).
45. Terakita, A. The opsins. *Genome Biol.* **6**, 213 (2005).
46. Fain, G. L., Hardie, R. & Laughlin, S. B. Phototransduction and the evolution of photoreceptors. *Curr. Biol.* **20**, R114–R124 (2010).
47. Gentles, A. J. & Karlin, S. Why are human G-protein-coupled receptors predominantly intronless? *Trends Genet.* **15**, 47–49 (1999).
48. De Renzi, S., Elemento, O., Tavazoie, S. & Wieschaus, E. F. Unmasking activation of the zygotic genome using chromosomal deletions in the *Drosophila* embryo. *PLoS Biol.* **5**, e117 (2007).
49. Guilgur, L. G. et al. Requirement for highly efficient pre-mRNA splicing during *Drosophila* early embryonic development. *eLife* **3**, e02181 (2014).
50. 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).
51. Barber, V. C., Evans, E. M. & Land, M. F. The fine structure of the eye of the mollusc *Pecten maximus*. *Z. Zellforsch. Mikrosk. Anat.* **76**, 25–312 (1967).
52. Morton, B. The evolution of eyes in the Bivalvia: new insights. *Am. Malacol. Bull.* **26**, 35–45 (2008).
53. Salvini-Plawen, L. V. Photoreception and the polyphyletic evolution of photoreceptors (with special reference to Mollusca). *Am. Malacol. Bull.* **26**, 83–100 (2008).
54. Cronly-Dillon, J. R. Spectral sensitivity of the scallop *Pecten maximus*. *Science* **151**, 345–346 (1966).
55. Kuo, C. H., Yamagata, K., Moyzis, R. K., Bitensky, M. W. & Miki, N. Multiple opsin mRNA species in bovine retina. *Brain Res.* **387**, 251–260 (1986).
56. Yokoyama, S. Molecular evolution of color vision in vertebrates. *Gene* **300**, 69–78 (2002).
57. Hennebert, E., Maldonado, B., Ladurner, P., Flammang, P. & Santos, R. Experimental strategies for the identification and characterization of adhesive proteins in animals: a review. *Interface Focus* **5**, 20140064 (2015).
58. Alejandrino, A., Puslednik, L. & Serb, J. M. Convergent and parallel evolution in life habit of the scallops (Bivalvia: Pectinidae). *BMC Evol. Biol.* **11**, 164 (2011).
59. Holten-Andersen, N. & Waite, J. H. Mussel-designed protective coatings for compliant substrates. *J. Dent. Res.* **87**, 701–709 (2008).
60. Mehdi-zadeh, M. & Yang, J. Design strategies and applications of tissue bioadhesives. *Macromol. Biosci.* **13**, 271–288 (2013).
61. Zhao, H. & Waite, J. H. Linking adhesive and structural proteins in the attachment plaque of *Mytilus californianus*. *J. Biol. Chem.* **281**, 26150–26158 (2006).
62. 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).
63. Silverman, H. G. & Roberto, F. Understanding marine mussel adhesion. *Mar. Biotechnol.* **9**, 661–681 (2007).
64. Qin, C. L. et al. In-depth proteomic analysis of the byssus from marine mussel *Mytilus coruscus*. *J. Proteomics* **144**, 87–98 (2016).
65. Grufflydd, L. D. The byssus and byssus glands in *Chlamys islandica* and other scallops (Lamellibranchia). *Zool. Scr.* **7**, 277–285 (1978).
66. Guerette, P. A. et al. Accelerating the design of biomimetic materials by integrating RNA-seq with proteomics and materials science. *Nat. Biotechnol.* **31**, 908–915 (2013).
67. 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).
68. Blaustein, M. P. & Lederer, W. J. Sodium/calcium exchange: its physiological implications. *Physiol. Rev.* **79**, 763–854 (1999).
69. Bricelj, V. M. et al. Sodium channel mutation leading to saxitoxin resistance in clams increases risk of PSP. *Nature* **434**, 763–767 (2005).

70. Zakon, H. H. Adaptive evolution of voltage-gated sodium channels: the first 800 million years. *Proc. Natl. Acad. Sci. USA* **109**(Suppl 1), 10619–10625 (2012).
71. Jost, M. C. et al. Toxin-resistant sodium channels: parallel adaptive evolution across a complete gene family. *Mol. Biol. Evol.* **25**, 1016–1024 (2008).
72. Choudhary, G., Yotsu-Yamashita, M., Shang, L., Yasumoto, T. & Dudley, S. C. Jr. Interactions of the C-11 hydroxyl of tetrodotoxin with the sodium channel outer vestibule. *Biophys. J.* **84**, 287–294 (2003).
73. Kontis, K. J. & Goldin, A. L. Site-directed mutagenesis of the putative pore region of the rat IIA sodium channel. *Mol. Pharmacol.* **43**, 635–644 (1993).
74. Bojarova, P. & Williams, S. J. Sulfotransferases, sulfatases and formylglycine-generating enzymes: a sulfation fascination. *Curr. Opin. Chem. Biol.* **12**, 573–581 (2008).
75. Soong, T. W. & Venkatesh, B. Adaptive evolution of tetrodotoxin resistance in animals. *Trends Genet.* **22**, 621–626 (2006).
76. Feldman, C. R., Brodie, E. D. Jr., Brodie, E. D. III & Pfrender, M. E. Constraint shapes convergence in tetrodotoxin-resistant sodium channels of snakes. *Proc. Natl. Acad. Sci. USA* **109**, 4556–4561 (2012).
77. Hanifin, C. T. & Gilly, W. F. Evolutionary history of a complex adaptation: tetrodotoxin resistance in salamanders. *Evolution* **69**, 232–244 (2015).
78. Zakon, H. H., Jost, M. C., Zwickl, D. J., Lu, Y. & Hillis, D. M. Molecular evolution of Na⁺ channels in teleost fishes. *Integr. Zool.* **4**, 64–74 (2009).
79. Terlau, H. et al. Mapping the site of block by tetrodotoxin and saxitoxin of sodium channel II. *FEBS Lett.* **293**, 93–96 (1991).
80. Shimizu, Y. & Yoshioka, M. Transformation of paralytic shellfish toxins as demonstrated in scallop homogenates. *Science* **212**, 547–549 (1981).

Acknowledgements

We thank Rencheng Yu, Fanzhou Kong and Yang Liu for facilitating the preparation of PST-feeding scallop materials and helpful comments on the manuscript. We acknowledge grant support from the National Natural Science Foundation of China (31130054, 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: MS2017NO01). X.G. acknowledges support from Taishan Oversea Scholar Program of Shandong and USDA-NIFA/NJAES 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., Xt.H., Wei.L., Jia.L., J.D. and Yp.L. participated initial genome analysis. X.S., J.Z., Ji.L., 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 Li.Z., J.W., Y.S., X.X., Y.M., X.L., Yj.L. and D.K. conducted mRNA extraction and library preparation for transcriptome sequencing. Yl.L., X.S. and G.X. participated in genome polymorphism analysis. Yl.L., S.W., T.C., Z.Y. and Yr.L. participated in muscle genes and energy pathway analysis, and J.W. in gene network construction. S.W., Y.S., Yl.L., T.C. and X.N. participated in opsin sequence and eye transcriptome analysis. W.J., Yl.L., S.W. and X.Z. 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 *Nav* 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., Yl.L., 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/>

Publisher's note: Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2017