The iron-responsive genome of the chiton *Acanthopleura granulata*

Rebecca M. Varney[1], Daniel I. Speiser[2], Carmel McDougall[3], Bernard M. Degnan[4], and Kevin M. Kocot[1,5]

[1]The University of Alabama, Department of Biological Sciences

[2]University of South Carolina, Department of Biological Sciences

[3]Griffith University, Australian Rivers Institute

[4]University of Queensland, School of Biological Sciences

[5]The University of Alabama Museum of Natural History

Corresponding author:

Kevin M. Kocot

kmkocot@ua.edu

ABSTRACT

Molluscs biomineralize structures that vary in composition, form, and function, prompting questions about the genetic mechanisms responsible for their production and the evolution of these mechanisms. Chitons (Mollusca, Polyplacophora) are a promising system for studies of biomineralization because they build a range of calcified structures including shell plates and spine- or scale-like sclerites. Chitons also harden the teeth of their rasp-like radula with a coat of iron. Here we present the genome of the West Indian fuzzy chiton *Acanthopleura granulata*, the first from any aculiferan mollusc. The *A. granulata* genome has features that may be specialized for iron biomineralization, including a high proportion of genes regulated directly by iron and two isoforms of ferritin, one iron-regulated and the other constitutively translated. The *A. granulata* genome also contains homologs of many biomineralization genes identified previously in conchiferan molluscs, suggesting the ancestral mollusc had a diverse genetic toolkit for biomineralization.

INTRODUCTION

Animals construct hardened structures by combining organic and inorganic components, a process termed biomineralization. To do so, they secrete proteins that initiate and guide the crystallization of inorganic molecules. Animals also incorporate proteins into biomineralized structures, enhancing their strength and flexibility(Cölfen 2010). Molluscs have long been models for studying the genetic mechanisms associated with biomineralization because they craft a wide range of materials into shells, spines, scales, and teeth. The ability of molluscs to produce diverse biomineralized structures likely contributes to their remarkable morphological and ecological diversity.

Chitons (Polyplacophora, Figure 1A) are a promising model for investigating mechanisms of biomineralization because they build diverse mineralized structures distinct from those of other molluscs. The shells of all molluscs are composed of calcium carbonate ($CaCO_3$), commonly in its crystal forms aragonite or calcite. Most molluscs build shells with alternating layers of aragonite and calcite, and many add an innermost layer of brick-like aragonite discs known as nacre. In contrast, chitons construct eight interlocking shell plates (Figure 1B) exclusively from aragonite and do not produce nacre. Also unlike other molluscs, chitons embed a network of sensory structures, termed aesthetes, into their shell plates. In some species, the aesthete network includes eyes with image-forming lenses made of aragonite(Speiser et al. 2011; Li et al. 2015) (Figure 1C). To protect the soft girdle tissue surrounding their shell plates, chitons produce scale- or spine-like sclerites, which are also made of aragonite(Schwabe 2010; Sigwart et al. 2014; Checa et al. 2017).

Chitons biomineralize teeth from a unique combination of materials. Most molluscs have a feeding organ, the radula, that bears rows of teeth built from chitin and, in many species, hardened with minerals such as calcium carbonate or silica. Chitons instead harden their teeth with calcium phosphate (in the form of apatite), and then cap each tooth with magnetite to reinforce its cutting edge(Lowenstam 1962) (Figure 1D). These iron coatings allow chitons to scrape algae from rocks without rapidly dulling or damaging their teeth. Chitons produce new teeth throughout their lives, making a new row every three days(Shaw et al. 2002; Joester and Brooker 2016). To make new teeth, chitons continuously sequester and transport large amounts of iron(Kim et al. 1989; Shaw et al. 2002; Shaw et al. 2010), a challenge because free iron causes oxidative stress(Dixon and Stockwell 2014).

To date, most investigations of biomineralization in molluscs have focused on species from the classes Bivalvia and Gastropoda. These, together with Monoplacophora, Cephalopoda, and Scaphopoda make up the clade Conchifera. The sister clade to Conchifera is Aculifera, made up of Polyplacophora and Aplacophora. Conchifera and Aculifera diverged approximately 550mya(Vinther, Jell, et al. 2012; Kocot et al. 2020). To make robust predictions about molluscan evolution, reconstructions of ancestral character states must include information from both conchiferans and aculiferans(Sigwart and Sutton 2007; Kocot et al. 2011; Smith et al. 2011; Vinther, Sperling, et al. 2012). Despite increasing numbers of sequenced molluscan genomes(Takeuchi et al. 2012; Zhang et al. 2012; Simakov et al. 2013; Albertin et al. 2015; Gómez-Chiarri et al. 2015; Kenny et al. 2015; Modica et al. 2015; Barghi et al. 2016; Davison et al. 2016; Murgarella et al. 2016; Adema et al. 2017; Du et al. 2017; Li et al. 2017; Nam et al. 2017; Schell et al. 2017; Sun et al. 2017; Wang et al. 2017; Calcino et al. 2018; Gerdol et al. 2018; Li et al. 2018; Liu et al. 2018; Renaut et al. 2018; Belcaid et al. 2019; Cai et al. 2019; Kijas et al. 2019; Masonbrink et al. 2019; McCartney et al. 2019; Zarrella et al. 2019), genomic resources for aculiferans remain unavailable. To

advance the study of molluscan evolution and to better understand the genetic mechanisms of biomineralization, we sequenced the genome of the West Indian fuzzy chiton *Acanthopleura granulata*. Exploring the *A. granulata* genome allowed us to: 1) identify genes chitons may use to build their shell plates, sclerites, and teeth; 2) seek genomic signatures associated with the biomineralization of iron; and 3) better understand the origin and evolution of biomineralization in molluscs.

RESULTS

A contiguous and complete chiton genome assembly

We sequenced the genome of a single specimen of *A. granulata*. We combined reads from one lane of Illumina HiSeqX paired-end sequencing (124 Gb of 2 X 150 bp reads,~204X coverage) with reads from four Oxford Nanopore flowcells run on the GridION platform (22.87 Gb, 37X coverage). Using the hybrid assembler MaSuRCA and optical mapping, we produced a haploid genome assembly for *A. granulata* that is 606.9 Mbp, slightly smaller than the 743 Mbp haploid genome size estimated by flow cytometry(Roebuck 2017). The assembled *A. granulata* genome consists of 87 scaffolds ranging in size from 50.9 Mb to 0.05 Mb, plus a single mitochondrial genome of 15,665 bp. Several of these scaffolds are similar in length to intact chromosomes from other molluscs(Sun et al. 2017; Bai et al. 2019). To verify completeness of the assembly, we mapped genomic short-read data to the genome; 85.31% of reads mapped perfectly, so we are confident the assembly encompasses a majority of sequencing data. The *A. granulata* genome has an N50 value of 23.9 Mbp and a BUSCO completeness score of 97.4%, making it more contiguous and complete than most currently available molluscan genomes (Supplementary Figure 1; Supplementary Table 1; visualized in Supplementary Figure 2).

We generated gene models by 1) aligning transcriptome data from *A. granulata* to the genome, and 2) training *de novo* gene predictors using the *A. granulata* transcriptome and protein sequences predicted from the transcriptomes of other aculiferans. Combining these two approaches, we produced a set of 81,691 gene models that is 96.9% complete according to a BUSCO transcriptomic analysis. This score is similar to the completeness score of the *A. granulata* genome, so it is likely this set of gene models missed few genes, if any, in the genome assembly. However, of the BUSCO genes expected to be single-copy in all animals, 17.2% were represented by multiple gene models. Using Markov clustering to eliminate redundant isoforms, we generated a reduced set of 20,470 gene models that is 94.7% complete. In this smaller set of gene models, only 0.5% of the BUSCO genes have multiple copies, supporting Markov clustering as an effective method for reducing the redundancy of gene models.

To provide a robust dataset for phylogenetic analysis and gene family evolution analyses, we identified homologous genes shared between *A. granulata* and other molluscs. We used the larger set of gene models from *A. granulata* to ensure a more complete input dataset, knowing that any duplicate gene models for the same locus would cluster within the same orthologous group. We compared gene models from the *A. granulata* genome to those from the genomes of nineteen other lophotrochozoans, including fourteen molluscs, two annelids, one brachiopod, one phoronid, and one nemertean. This resulted in 59,276 groups of homologous sequences including 3,379 found in all 20 genomes.

We used a tree-based approach to identify orthologous genes shared among all 20 taxa and reconstructed molluscan phylogeny using the 2,593 orthologs present in at least 17 of the 20 genomes we searched. This dataset totaled 950,322 amino acid positions with 16.2% missing data. We recovered *A. granulata* as the sister taxon of all other molluscs with sequenced genomes (Figure 1F). When we included more taxa in our phylogenetic reconstruction by using transcriptomes in addition to genomes, we recovered a position of *A. granulata* within other chitons consistent with recent phylogenetic studies(Irisarri et al. 2020) (Supplementary Figure 3).

The A. granulata *genome differs from conchiferan genomes in content and organization*

The *A. granulata* genome has a heterozygosity of 0.653%, one of the lowest heterozygosities of any molluscan genome sequenced to date (Supplementary Figure 4). High heterozygosity is often attributed to high rates of gene flow associated with broadcast spawning and far-dispersing larvae(Solé-Cava and Thorpe 1991), and it is frequently noted as an obstacle to genome assembly in molluscs (Zhang et al. 2012; Wang et al. 2017; Powell et al. 2018; Thai et al. 2019). We expected the *A. granulata* genome to have high heterozygosity because this species of chiton is a broadcast spawner with a wide geographic range(Glynn 1970). Using k-mer based analysis, we found the highest heterozygosity among the seven genomes analyzed was 3.15% in the blood clam *S. broughtonii*, and the other genomes had heterozygosities between those of *A. granulata* and *S. broughtonii*. Our findings indicate that heterozygosity may be influenced by more than an animal's reproductive mode, larval type, and geographic range (Supplementary Table 3), and that molluscan genomes should not be assumed to have high heterozygosity.

The *A. granulata* genome is arranged differently than other molluscan genomes and has fewer repetitive elements. Synteny is lower between *A. granulata* and all conchiferan molluscs we examined than it is between any two conchiferans (Supplementary Figure 5). Compared to other molluscs, *A. granulata* has fewer repetitive elements in its genome (Supplementary Table 2). Repetitive elements contribute to structural changes in genomes by providing breakpoints that increase the likelihood of chromosomal rearrangements(Weckselblatt and Rudd 2015). The low level of synteny between *A. granulata* and any conchiferan suggests the genomes of aculiferans and conchiferans became rearranged relative to each other following their evolutionary divergence, perhaps influenced by proliferations of repetitive elements in the genomes of conchiferans.

The Hox cluster is a widely conserved set of regulatory genes that contribute to the patterning of the anterior-posterior axes in bilaterian animals. In lophotrochozoans, the genes are typically collinear, beginning with *Hox1* and ending with *Post1*. Although several gastropods and bivalves possess intact Hox clusters, the clusters are dispersed in some bivalves and cephalopods(Albertin et al. 2015; Barucca et al. 2016; Belcaid et al. 2019). The Hox cluster of *A. granulata* lacks *Post1*, but is otherwise intact and collinear (Figure 2). Given current understanding of molluscan phylogeny, the order of Hox genes shared between *A. granulata* and most conchiferans likely represents the ancestral order of Hox genes in molluscs(Wanninger and Wollesen 2019) (Figure 2). *Post1* is also absent in two species of chiton from the suborder Acanthochitonina, the sister clade to that of *A. granulata* (Chitonina)(Huan et al. 2019; Wanninger and Wollesen 2019). However, *Post1* is present in aplacophorans(Iijima et al. 2006), suggesting it was lost in chitons. In conchiferan molluscs, *Post1* helps specify the posterior of an animal during development and helps pattern shell formation(Lee et al. 2003; Fröbius et al. 2008; Schiemann et al. 2017; Huan et al. 2019). In the absence of *Post1, A. granulata* and other chitons must use other sets of transcription factors to help pattern their body axes and their biomineralized structures.

Acanthopleura granulata ***shares many biomineralization genes with conchiferan molluscs***

We expected chitons to lack many genes previously identified in molluscan biomineralization pathways because their shell plates and sclerites lack both calcite and nacre. We were surprised to find homologs in the *A. granulata* genome of many biomineralization genes known from conchiferans (Supplementary

Table 4). For example, we found an ortholog in *A. granulata* for *Pif*. In pterid bivalves, the *Pif* mRNA encodes a protein that is cleaved into two peptides, PIF97 and PIF80(Suzuki et al. 2009). These peptides have different roles in biomineralization: PIF80 binds nacre and aids in nacre formation(Suzuki et al. 2009), whereas PIF97 binds to chitin and guides the growth of calcium carbonate crystals(Suzuki et al. 2013). We found that *A. granulata* possesses a *Pif* homolog, but appears to only produce PIF97 rather than two separate peptides. The expression of *Pif* mRNA was highest in girdle tissue in *A. granulata* and lowest in the radula, suggesting that PIF peptides may play a role in sclerite formation in chitons. We hypothesize that the last common ancestor of extant molluscs used PIF97 to help build mineralized structures, and that production of PIF80 is novel to bivalves.

The ancestral mollusc likely produced mineralized structures, but whether the ancestral mollusc had a single shell, multiple shell plates, or sclerites remains a matter of debate(Scherholz et al. 2013; Vinther et al. 2017; Giribet and Edgecombe 2020). Molluscs form mineralized structures by making extracellular matrices from organic components such as polysaccharides and proteins, and then hardening them with minerals(Furuhashi et al. 2009). Similarities between the extracellular matrices of different biomineralized structures suggest these structures share developmental mechanisms. The *A. granulata* genome includes genes known from conchiferan molluscs to be associated with extracellular matrices. Chitin is a major component of the extracellular matrices of all molluscan shells and radulae, and the *A. granulata* genome contains genes for chitin, chitinase, and chitin-binding proteins. We also found homologs of lustrin and dermatopontin, two proteins expressed in the extracellular matrices of conchiferans that increase the elasticity and flexibility of their shells(Gaume et al. 2014) (Supplementary Table 4).

Silk-like structural proteins are components of many biological materials, including shells(Eisoldt et al. 2011; McDougall et al. 2016; Xu et al. 2016), and several *A. granulata* genes are similar to genes known to code for silk-like proteins. These proteins are “silk-like” because they contain highly repetitive sequences of amino acids that fold into secondary structures (commonly β-pleated sheets) that impart flexibility, a phenomenon first documented in spider silk(Lewis 2006; Eisoldt et al. 2011). Silk-like domains can facilitate the precipitation and crystallization of minerals that help form structures such as bones and shells(Xu et al. 2016). We found 31 genes that code for proteins with silk-like domains in the *A. granulata* genome, 23 of which have high sequence similarity to characterized molluscan biomineralization genes (Supplementary Table 5). We found 27 of these 31 genes code for proteins with signal peptides, indicating they may be secreted as part of the extracellular matrix during biomineralization (Supplementary Table 5). Among these genes, we found three collagens, one chitinase, and one carbonic anhydrase, all possible contributors to shell formation and repair(Patel 2004) (Supplementary Table 5). Several of the genes encoding proteins with silk-like domains are highly expressed in the girdle of *A. granulata*, suggesting a role in the mineralization of sclerites (Supplementary Figure 6). Our results indicate a high number of biomineralization genes are shared between aculiferans and conchiferans, and therefore potentially among all molluscs.

The iron-responsive genome of **A. granulata**

A. granulata *has more genes with iron response elements (IREs) than other molluscs*

Chitons have more iron in their hemolymph than any other animal(Kim et al. 1988). We hypothesize that the ability of chitons to biomineralize iron requires them to respond quickly to changes in concentration of this potentially toxic metal. To assess the iron-responsiveness of the *A. granulata* genome, we searched it for iron response elements (IREs), three-dimensional hairpin structures that form in the untranslated regions (UTRs) of mRNA molecules and control translation via binding by iron regulatory protein (IRP; Supplementary Figure 7). Despite having the fewest gene models, *A. granulata* has more IREs in its genome than any other mollusc we examined. We predicted 271 IREs in the *A. granulata* genome, compared with an average of 119 IREs across other molluscan genomes (Supplementary Table 6). The highest number of predicted IREs in a conchiferan came from the genome of the blood clam *Scapharca broughtonii*, which had 201. The blood clam is so named because it is one of few molluscs that produces hemoglobin(Bai et al. 2019). We expect *A. granulata* and *S. broughtonii* have more IREs in their genomes than other molluscs because they must absorb and transport larger amounts of iron to produce iron-coated teeth and hemoglobin, respectively. We hypothesize that the evolution of increased iron use is associated with an increase in the number of genes regulated by iron, so species that produce iron-containing substances or structures will tend to have more genes with IREs than species that do not.

After identifying IREs in molluscan genomes, we asked which IREs may be associated with increased or decreased protein expression in the presence of iron. When free iron is present, IRPs do not bind to IREs. When IRPs bind to IREs in the 5' UTR of an mRNA, they block ribosomes and prevent translation; thus, mRNAs with 5' IREs will be translated in the presence of free iron (Supplementary Figure 7). When IRPs bind to IREs in the 3' UTR, they block endonucleases from degrading mRNA, thereby allowing multiple translations from a single mRNA molecule; thus, the amount of protein produced from mRNAs with 3' IREs will decrease in the presence of free iron (Supplementary Figure 7). We quantified the numbers of 5' and 3' IREs in available molluscan genomes, and found that all the genomes, including the genome of *A. granulata,* have similar proportions of each type of IRE (Figure 3A). Genes that contain a 3' IRE often contain multiple IREs in tandem to maximize protection from endonucleases, so a high number of 3' IREs in a genome does not necessarily indicate a high number of iron-responsive genes.

We next asked if the expression of IRE-containing genes in *A. granulata* differs between iron-rich and iron-poor tissues. Because IRP will not bind to an IRE in the presence of iron, IREs will not influence rates of translation in tissues with high amounts of free iron. We compared the expression of IRE-containing genes between transcriptomes sequenced from the foot, girdle, ctenidia, and four developmentally-distinct regions of the radula of *A. granulata* (Figure 3B). We found that expression of genes with IREs is lower in the anterior, iron-rich tissues of the radula than in any other tissues. The posterior portion of the radula in *A. granulata*, the region with developing teeth not yet mineralized by iron, expresses genes with IREs at levels more similar to those of non-radula tissues than to the other regions of the radula. A previous study identified a novel mineralization protein (RTMP1) in the radula of another chiton (*Cryptochiton stelleri*). RTMP1 is thought to be secreted extracellularly, and regulated via phosphorylation(Nemoto et al. 2019). We examined the mRNA of RTMP1 in *C. stelleri* and did not detect an IRE in either the 5' or 3' UTR. Given that IREs do not impact translation in the presence of free iron, we hypothesize chitons generally rely on genes that are regulated by mechanisms other than IREs to control the biomineralization of iron in the radula.

Two isoforms of ferritin may provide tissue-specific protection from oxidative stress

All metazoans require iron. However, free iron poses a threat to animals because it catalyzes the production of reactive oxygen species, which inflict damage on DNA and tissues(Dixon and Stockwell 2014). To transport iron safely, metazoans use the iron-binding protein ferritin. Previous work suggests that chitons use ferritin to transport iron to their radula(Kim et al. 1988). An iron response element (IRE) is present in the 5' UTR of the heavy chain (or soma-like) ferritin that is expressed by all metazoans(Piccinelli and Samuelsson 2007). We found two isoforms of heavy chain ferritin in our gene models for *A. granulata*: a first isoform (isoform 1) that contains the conserved 5' IRE, and a second isoform (isoform 2) that does not (Figure 4A).

Isoform 1 of ferritin from *A. granulata* contains an IRE in the 5' UTR and is thus regulated by iron(Piccinelli and Samuelsson 2007). The 5' IRE allows this isoform to be translated only in the presence of free iron. By regulating the translation of ferritin, cells can transcribe ferritin mRNA continuously so that they are primed to produce large quantities of ferritin protein rapidly if conditions require it. If no free iron is present, IRP will bind to the IRE and block translation. We found isoform 1 of ferritin is expressed at similar levels in all the transcriptomes we sequenced for *A. granulata*, including those for the foot, girdle, gonad, ctenidia, and all four regions of the radula (Figure 4B). Thus, when *A. granulata* needs to bind excess iron, it may be able to rapidly produce isoform 1 of ferritin protein throughout its body.

Isoform 2 of ferritin in *A. granulata* lacks the 5' IRE present in isoform 1. We identified an alternative transcription initiation site downstream of ferritin exon 1 in the *A. granulata* genome. Isoform 2 of ferritin, initiated at this downstream site, contains a different exon 1 than isoform 1 of ferritin, but shares exons 2-4 with isoform 1. We found transcripts of isoform 2 are expressed at a lower level than isoform 1 throughout all body tissues (foot, girdle, gonad, ctenidia) and in the posterior region of the radula that lacks iron mineralization (Figure 4B). Expression of isoform 2 is almost undetectable in the iron-rich regions of the radula. Without the 5' IRE, translation of the mRNA of isoform 2 is not blocked in the absence of free iron. The 5' IRE in ferritin is an important regulatory mechanism for protein production. In rats, for example, the expression of ferritin mRNAs is relatively constant across tissues but protein levels vary(Rogers and Munro 1987). Further, mutations in the 5' IRE cause iron-related medical conditions in mammals due to an overproduction of ferritin protein (hyperferritinaemia)(Thomson et al. 1999). We hypothesize that chitons use isoform 2 of ferritin to produce a low level of ferritin protein constitutively in tissues outside their radula as protection from the high concentrations of iron circulating throughout their bodies.

DISCUSSION

Chitons are a valuable system for investigations of biomineralization because they produce shell plates, spines, and iron-clad teeth. The unique combination of structures produced by chitons makes the *A. granulata* genome a resource for future studies of biomineralization. Although many genes involved in molluscan shell secretion are rapidly evolving(Jackson et al. 2006; Kocot et al. 2016), we were able to identify homologs of many of these biomineralization genes in the *A. granulata* genome. The expression of several genes associated with conchiferan shell secretion in the girdle of *A. granulata* suggests these genes may function in sclerite biomineralization in chitons. This suggests a common underlying biomineralization mechanism for conchiferan shells and aculiferan sclerites, structures known to share some developmental pathways even though they arise via different cell lineages(Wollesen et al. 2017).

All metazoans require iron, but they must balance iron use against potential oxidative damage. Regulating iron is a particular concern for chitons because they mineralize their teeth with magnetite. The genome of *A. granulata* contains more genes with iron response elements (IREs) than that of any other mollusc examined to date, indicating it has a larger proportion of genes regulated directly by iron. We identified two isoforms of ferritin in *A. granulata*, one that is iron-responsive and a second that is constitutively translated. We propose the second isoform of ferritin protects tissues outside the radula from oxidative stress by binding free iron. The *A. granulata* genome is a resource for future studies of metal-based mineralization as well as iron homeostasis.

The *A. granulata* genome is the first available genome for any chiton or any aculiferan. The information it provides improves our understanding of lineage-specific innovations within chitons as well as the evolution of biomineralization across Mollusca.

METHODS

Specimen collection

We collected a single male specimen of *Acanthopleura granulata* from Harry Harris State Park in the Florida Keys (Special Activity License #SAL-17-1983-SR). We cut the majority of the foot into ~1 mm^2 cubes and froze them at -80°C. We froze additional pieces of foot, girdle, ctenidia, gonad, and radula in RNAlater and stored them at -80°C as well.

Genome and transcriptome sequencing

We extracted high molecular weight DNA from frozen samples of foot tissue from *A. granulata* using a CTAB-phenol chloroform method. We cleaned DNA for short read generation with the Zymo Clean and Concentrator Kit. For library preparation and sequencing, we sent cleaned DNA to the Genomics Services Lab at HudsonAlpha (Huntsville, AL), where it was sheared with a Covaris M220 to an average fragment size of 350 bp. These fragments were used to prepare an Illumina TruSeq DNA PCR-Free library, which was sequenced using one lane of an Illumina HiSeq X (2 X 150 bp paired-end reads).

For long-read sequencing, we cleaned DNA and enriched it for higher-molecular weight fragments by performing two sequential purifications using 0.4X AmPureXP magnetic beads. We generated long reads with four flow cells on an Oxford Nanopore Technologies GridION. We prepared two sequencing libraries with ligation kit LSK-108 and sequenced them on FloMin106 (R9.4.1) flow cells. We prepared the other two sequencing libraries with the updated ligation kit LSK-109 and sequenced them on R9.4.1RevD flow cells. We generated 2.19Gb, 4.41Gb, 7.87 Gb, and 8.4 Gb respectively across the four flow cells, for a total of 22.87 Gb, or >20x coverage with long-reads. We trimmed long reads with PoreChop(Wick 2018), which was set to remove chimeras (approximately 0.0005% of reads) and all residual adapter sequences.

To generate transcriptomes, we used the Omega Bio-tek EZNA Mollusc RNA Kit to extract RNA from girdle, ctenidia, gonad, foot, and four regions of radula (representative of visibly different stages of iron mineralization) of the same individual of *A. granulata* we used for genome sequencing. We synthesized and amplified complementary DNA (cDNA) from each tissue using the SmartSeq v4 Ultra Low-input RNA kit (Clontech) from 1 ng of input RNA with 17 cycles of PCR. We created eight dual-indexed sequencing libraries with the Illumina Nextera XT kit, using 1 ng of input cDNA. We sent the eight libraries to Macrogen (Seoul, South Korea) where they were pooled and sequenced on one lane of an Illumina HiSeq 4000 (2 x 100 bp paired-end reads).

Genome and transcriptome assembly and quality assessment

We initially assembled the chiton genome with MaSuRCA v. 3.3.5(Zimin et al. 2013), which consolidates paired-end data into super reads and then uses long-read data to scaffold and gap-fill. This produced an assembly with 2,858 contigs. We filtered and collapsed heterozygous contigs with Redundans v. 0.14a(Pryszcz and Gabaldón 2016), decreasing the assembly to 1,285 contigs. To ensure that no contigs were incorrectly removed, we verified that all pre-Redundans contigs mapped to the post-Redundans assembly with bowtie2(Langmead and Salzberg 2012); all contigs mapped and thus non-redundant data were not deleted. To help decontaminate reads and contigs, we used the Blobtools2 Interface to create

blobplots(Blaxter and Challis 2018). Because Blobtools uses the NCBI nucleotide database to determine the identity of each scaffold, and chordate sequences vastly outnumber molluscan sequences in NCBI, Blobtools identified a large proportion of scaffolds as chordate. We identified contaminants as sequences that differed from the majority of scaffolds in both GC content and coverage and used BLAST to verify these sequences as bacterial before removing them from the assembly.

We scaffolded this reduced assembly with one lane of Bionano SAPHYR optical mapping, using two enzymes (BssSI and DLE1) and Bionano Solve v3.4's scaffolding software, which resulted in 87 scaffolds. We ran REAPR v. 1.0.18(Hunt et al. 2013), which map short read data and collect mapping statistics simultaneously, to determine accuracy of the assembly overall relative to all short-read data generated, and found despite reducing heterozygosity in the final assembly, 85.31% of paired-end reads map perfectly back to the genome assembly, indicating a complete genome assembly relative to the paired-end data.

To assess our genome assembly, we ran QUAST v. 5.0.2(Gurevich et al. 2013). We assessed genome completeness with BUSCO v. 4.0.2(Simão et al. 2015), using the proportions of nuclear protein-coding genes thought to be single-copy in the genomes of diverse metazoans (Metazoa odb9 dataset) and estimating the proportion of those that were complete, duplicated, fragmented, and absent.

We assembled the eight *A. granulata* transcriptomes with Trinity v. 2.84(Grabherr et al. 2011), using the --trimmomatic and --normalize reads flags. We ran CD-Hit v. 4.8.1(Fu et al. 2012) on each transcriptome separately to cluster isoforms. We also generated a composite transcriptome of all eight tissues (eight total transcriptomes including four separate radula regions) by combining reads and then following the same process described above. We used this composite transcriptome for annotation.

Genome annotation

To annotate the *A. granulata* genome, we first generated a custom repeat library with RepeatModeler v. 2.0(Smit and Hubley 2008), which was used in all subsequent analyses. We trained MAKER v. 2.31.10(Cantarel et al. 2008) on the composite transcriptome described above as well as predicted protein sequences from several other species of chitons that were generated previously (see Supplementary File 1). Using the highest quality gene models from the first as a maker-input gff3 (AED <0.5), we ran a second round of MAKER. From these resulting gene models, we used those with an AED <0.25 to train Augustus v3.0.3(Stanke et al. 2006): we extracted gene models from the genomic scaffolds along with 1,000 bp of flanking sequence on either side to ensure complete genes, and ran them through BUSCO to produce an Augustus model (.hmm) file. Separately, we ran PASA 2.4.1(Haas et al. 2003) on our composite transcriptome to maximize mapping transcripts to the genome assembly. We combined results from PASA and a trained Augustus run using the intersect tool in BEDtools v. 2.29.2(Quinlan and Hall 2010), which removed identical sequences. This yielded a set of 81,691 gene models. When we ran a BUSCO v. 3.9 analysis (Metazoa odb9 dataset), we found a 15.2% duplication rate. To decrease duplications caused by transcripts predicted for the same locus by both Augustus and PASA that varied in length (and thus were not removed by the BEDtools intersect tool), we clustered the first set of gene models using cdhit-EST v. 4.8.1(Fu et al. 2012), which we ran with the slow-but-accurate (-g) flag and with a cluster threshold value of 0.8. This produced a set of 20,470 genes. All commands we used are available in Appendix 1.

352

353 *Hox gene annotation and genomic comparisons*

354 We located the Hox cluster of *A. granulata* by first creating a BLAST database of the *A. granulata*
355 scaffolds and then querying this database with available chiton Hox sequences(Wanninger and Wollesen
356 2019). We marked *A. granulata* sequences with a BLAST hit at e-value 1e-8 as potential Hox sequences.
357 We found one clear match for each previously identified chiton Hox gene, all in a single cluster within
358 one scaffold. To verify the absence of *Post1*, we queried the *A. granulata* database with *Post1* sequences
359 from five other molluscs(Wanninger and Wollesen 2019). All matched with low support to the existing *A.*
360 *granulata Post2* sequence, so we concluded that *Post1* is absent from the *A. granulata* genome
361 assembly.

362 To graphically examine synteny between *A. granulata* and other molluscan genome assemblies, we
363 loaded each assembly and annotation into the online COGE SynMap2(Haug-Baltzell et al. 2017) server
364 and compared *A. granulata* to eight other annotated genomes with default SynMap2 settings. We
365 exported dotplots for each pair of genomes to visualize syntenic regions (or lack thereof). Scaffolds in
366 each dotplot were sorted by length, but differing assembly qualities made some dotplots difficult to
367 read due to a high number of very small scaffolds.

368 To permit direct comparisons of repeat content within *A. granulata* and other molluscs, we ran
369 RepeatModeler(Smit and Hubley 2008) on the scaffolds of a subset of genome assemblies and *A.*
370 *granulata.* We used the same default parameters for each run and quantified the number of elements in
371 each repeat family identified by RepeatModeler for each genome assembly we analyzed (LINEs, SINEs,
372 etc.).

373 *Orthology inference*

374 To identify orthologous genes shared between *A. granulata* and other molluscs, we used OrthoFinder v.
375 2.3.7(Emms and Kelly 2015). We analyzed three separate sets of data: 1) *A. granulata* and genomes of
376 nineteen other lophotrochozoans, including fourteen other molluscs, two annelids, one brachiopod, one
377 phoronid, and one nemertean 2) *A. granulata* and a subset of molluscan genomes for detailed
378 comparisons of biomineralization genes and; 3) *A. granulata* and an expanded set of data including both
379 genomes and transcriptomes, including several transcriptomes from aculiferans other than *A. granulata*.
380 For all three analyses we used the unclustered 81,691 gene set for *A. granulata*, knowing that duplicated
381 gene models would cluster together. We removed sequences from our orthogroups that were identical
382 to longer sequences where they overlapped, as well as fragmented sequences shorter than 100 amino
383 acids, using uniqHaplo(Anon). We retained orthogroups that had a minimum of four taxa, aligned the
384 sequences within them with MAFFT(Katoh et al. 2002), and cleaned mistranslated regions with
385 HmmCleaner(Di Franco et al. 2019). We used AlignmentCompare
386 (https://github.com/kmkocot/basal_metazoan_phylogenomics_scripts_01-2015) to delete sequences
387 that did not overlap with all other sequences by at least 20 AAs (starting with the shortest sequence
388 meeting this criterion).

389 *Phylogenetic analyses*

For species tree reconstruction, in cases where two or more sequences were present for any taxon in a single-gene alignment, we used PhyloPyPruner 0.9.5 (https://pypi.org/project/phylopypruner/) to reduce the alignment to a set of strict orthologs. This tool uses single-gene trees to screen putative orthogroups for paralogy. To build single-gene trees based on orthologs, we trimmed alignments with BMGE v1.12.2(Criscuolo and Gribaldo 2010) and constructed approximately maximum likelihood trees for each alignment with FastTree2(Price et al. 2010) using the "slow" and "gamma" options. We then used these alignments in PhyloPyPruner with the following settings: --min-len 100 --min-support 0.75 --mask pdist --trim-lb 3 --trim-divergent 0.75 --min-pdist 0.01 --trim-freq-paralogs 3 --prune MI. For datasets 1 ("genomes") and 3 ("all_taxa"), only orthogroups sampled for at least 85% of the total number of taxa were retained for concatenation. For dataset 2 ("biomin_subset"), only orthogroups sampled for all eight taxa were retained. Phylogenetic analyses were conducted on the supermatrix produced by PhyloPyPruner v. 1.0 in IQ-TREE v. 1.6.12(Nguyen et al. 2015) using the LG+PMSF model(Wang et al. 2018) with 1000 rapid bootstraps.

Screening for known biomineralization genes

We identified known molluscan biomineralization genes of interest in the chiton genome by first making a BLAST protein database of protein sequences for the 81,691-gene model set of *A. granulata* annotations, translated by Transdecoder(Grabherr et al. 2011; Douglas 2018). We then used the highest-quality protein sequence of that gene available on NCBI (complete where available, longest if only incomplete protein sequences existed) as a query for each biomineralization gene of interest with an initial e-value cutoff of 1e-8. In cases where multiple hits of similar support resulted, we selected the correct hit by constructing a phylogeny in RAxML v. 8.2.12(Stamatakis 2014) under the GTRGAMMA model with rapid bootstrapping and a best-scoring maximum-likelihood tree search in one run, with the number of bootstrap replicates determined by majority-rule consensus (autoMRE). This produced a list of sequences from *A. granulata* that matched the biomineralization genes of interest, and allowed us to narrow down our list of potential biomineralization genes present in *A. granulata*.

We used the above set of gene queries from other molluscs to identify the ortholog group from the above OrthoFinder2 on the subset of genomes selected as biomineralization representatives across Mollusca. We used the complete CDS or longest mRNA for each gene as a nucleotide query to search our orthogroups, again with an e-value cutoff of 1e-8 to identify the orthogroup(s) likely contained that particular biomineralization gene of interest. This produced a list of orthogroups that contained sequences with high similarity to the query, often multiple orthogroups per gene (Supplementary Table 4). This was expected due to clustering within OrthoFinder2. We used NCBI BLAST to verify the identity of the orthologous gene sequences by verifying that the top hits for each in BLAST matched the biomineralization gene of interest. We then examined these orthogroups to locate the previously identified *A. granulata* gene model that matched to each biomineralization protein. The query sequences for each gene sought in *A. granulata* are available in Supplementary Table 7.

Silk-like proteins share similar amino acid composition throughout Metazoa, but the genes that code for them are difficult to identify in genomes because their highly-repetitive sequences are often missed by traditional gene annotation tools(McDougall et al. 2016). We looked for silk-like proteins with SilkSlider(McDougall et al. 2016), run with default settings but using SignalP v. 4.01(Nielsen 2017), which identifies potential silk-like proteins by locating low-complexity repetitive domains and signal peptides.

The 31 proteins identified as silk-like by SilkSlider were then uploaded to the SignalP 5.0 webserver(Almagro Armenteros et al. 2019) for further predictions of signal peptides associated with extracellular localization.

To locate and quantify iron response elements (IREs), we screened the 20,470-gene *A. granulata* gene model set using the SIREs 2.0(Campillos et al. 2010) web server. We also ran SIRE on the subset of genomes used for biomineralization analyses (see OrthoFinder above) for comparison. We compensated for differences in annotation methods by first clustering all coding sequences from each genome with CD-Hit-EST(Fu et al. 2012) with a cluster threshold of 0.8 (to match the threshold value we used earlier to reduce redundancy in the annotations of the *A. granulata* genome). We then ran SIRE on each of these sets of predicted transcripts. We only accepted predicted IREs scored as "high quality" according to the SIRE metric (indicating both sequence and structural characteristics of a functional IRE). We pulled chiton genes containing a high quality IRE from the eight different tissue transcriptomes generated for genome annotation and assessed expression by mapping each back to the genome with Salmon v. 0.11.3(Patro et al. 2017) to generate quantifications of reads per transcript, and running these quantifications through edgeR(Robinson et al. 2010) to account for transcript length (TPM) and permit direct comparisons of gene expression. We made heatmaps with log-transformed data to compensate for outliers in expression levels with R package prettyheatmap(Kolde 2012).

Supplementary Material

The West Indian Fuzzy Chiton *Acanthopleura granulata* genome and transcriptomes from the same individuals have been deposited in the NCBI database as BioProject PRJNA578131. The genome project is registered in NCBI as JABBOT000000000. All raw reads for both the genome and transcriptomes are available online at the NCBI Sequence Read Archive under the same BioProject PRJNA578131. Transcriptome data of other species of chitons used for genome annotation are available online at the NCBI Sequence Read Archive under BioProjects PRJNA626693 and PRJNA629039. The genome assembly, all sets of gene models discussed in the manuscript, and supporting documentation for the biomineralization genes described are available in Dryad with the identifier doi:X

All code used in this study is available in Supplementary Material, Appendix 1.

Acknowledgements

We thank Ken Halanych and the crew and scientists of the Icy Inverts cruises aboard R/V Lawrence M. Gould and R/V Nathaniel B. Palmer for collecting *Callochiton* sp., Christoph Held, the Alfred Wegener Institute, and the scientists and crew of the PS96 cruise aboard R/V Polarstern for facilitating collection of *Nuttallochiton* sp., Julia Sigwart and Lauren Sumner-Rooney for collecting *Leptochiton asellus*, and Alexandra Kingston and Daniel Chappell for collecting *Chiton marmoratus and Chiton tuberculatus*, all used for transcriptome sequencing. CM was supported by a Malacological Society of London grant for transcriptome sequencing of *Acanthopleura gemmata.* We thank Kerry Roper for RNA QC of the *Acanthopleura gemmata* transcriptome samples. We thank the staff of the Smithsonian Marine Station at Ft. Pierce and the Keys Marine Lab for providing housing during collections. RMV thanks John Sutton, Michael McKain, and Michelle Lewis for assistance in optimizing nanopore library preparation protocols.

REFERENCES

Adema CM, Hillier LW, Jones CS, Loker ES, Knight M, Minx P, Oliveira G, Raghavan N, Shedlock A, do Amaral LR, et al. 2017. Whole genome analysis of a schistosomiasis-transmitting freshwater snail. Nat Commun [Internet] 8. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5440852/

Albertin CB, Simakov O, Mitros T, Wang ZY, Pungor JR, Edsinger-Gonzales E, Brenner S, Ragsdale CW, Rokhsar DS. 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature 524:220–224.

Almagro Armenteros JJ, Tsirigos KD, Sønderby CK, Petersen TN, Winther O, Brunak S, von Heijne G, Nielsen H. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nature Biotechnology 37:420–423.

Anon. Perl scripts for sequence analysis - Takebayashi. Available from: http://raven.wrrb.uaf.edu/~ntakebay/teaching/programming/perl-scripts/perl-scripts.html

Bai C-M, Xin L-S, Rosani U, Wu B, Wang Q-C, Duan X-K, Liu Z-H, Wang C-M. 2019. Chromosomal-level assembly of the blood clam, Scapharca (Anadara) broughtonii, using long sequence reads and Hi-C. Gigascience [Internet] 8. Available from: https://academic.oup.com/gigascience/article/8/7/giz067/5530322

Barghi N, Concepcion GP, Olivera BM, Lluisma AO. 2016. Structural features of conopeptide genes inferred from partial sequences of the Conus tribblei genome. Mol. Genet. Genomics 291:411–422.

Barucca M, Canapa A, Biscotti MA. 2016. An Overview of Hox Genes in Lophotrochozoa: Evolution and Functionality. Journal of Developmental Biology 4:12.

Belcaid M, Casaburi G, McAnulty SJ, Schmidbaur H, Suria AM, Moriano-Gutierrez S, Pankey MS, Oakley TH, Kremer N, Koch EJ, et al. 2019. Symbiotic organs shaped by distinct modes of genome evolution in cephalopods. PNAS:201817322.

Blaxter M, Challis R. 2018. BlobToolkit. BlobToolKit [Internet]. Available from: https://github.com/blobtoolkit

Cai H, Li Q, Fang X, Li J, Curtis NE, Altenburger A, Shibata T, Feng M, Maeda T, Schwartz JA, et al. 2019. A draft genome assembly of the solar-powered sea slug *Elysia chlorotica*. Scientific Data 6:190022.

Calcino AD, de Oliveira AL, Simakov O, Schwaha T, Zieger E, Wollesen T, Wanninger A. 2018. The quagga mussel genome and the evolution of freshwater tolerance: Supplementary Material. bioRxiv [Internet]. Available from: http://biorxiv.org/lookup/doi/10.1101/505305

Campillos M, Cases I, Hentze MW, Sanchez M. 2010. SIREs: searching for iron-responsive elements. Nucleic Acids Res 38:W360–W367.

Cantarel BL, Korf I, Robb SMC, Parra G, Ross E, Moore B, Holt C, Sánchez Alvarado A, Yandell M. 2008. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res 18:188–196.

Checa AG, Vendrasco MJ, Salas C. 2017. Cuticle of Polyplacophora: structure, secretion, and homology with the periostracum of conchiferans. Mar Biol 164:64.

Cölfen H. 2010. A crystal-clear view. Nature Mater 9:960–961.

Criscuolo A, Gribaldo S. 2010. BMGE (Block Mapping and Gathering with Entropy): a new software for selection of phylogenetic informative regions from multiple sequence alignments. BMC Evolutionary Biology 10:210.

Davison A, McDowell GS, Holden JM, Johnson HF, Koutsovoulos GD, Liu MM, Hulpiau P, Van Roy F, Wade CM, Banerjee R, et al. 2016. Formin Is Associated with Left-Right Asymmetry in the Pond Snail and the Frog. Curr Biol 26:654–660.

Di Franco A, Poujol R, Baurain D, Philippe H. 2019. Evaluating the usefulness of alignment filtering methods to reduce the impact of errors on evolutionary inferences. BMC Evolutionary Biology 19:21.

Dixon SJ, Stockwell BR. 2014. The role of iron and reactive oxygen species in cell death. Nature Chemical Biology 10:9–17.

Douglas P. 2018. TransDecoder/TransDecoder. GitHub [Internet]. Available from: https://github.com/TransDecoder/TransDecoder

Du X, Fan G, Jiao Y, Zhang H, Guo X, Huang R, Zheng Z, Bian C, Deng Y, Wang Q, et al. 2017. The pearl oyster Pinctada fucata martensii genome and multi-omic analyses provide insights into biomineralization. Gigascience 6:1–12.

Eisoldt L, Smith A, Scheibel T. 2011. Decoding the secrets of spider silk. Materials Today 14:80–86.

Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons dramatically improves orthogroup inference accuracy. Genome Biology 16:157.

Fröbius AC, Matus DQ, Seaver EC. 2008. Genomic Organization and Expression Demonstrate Spatial and Temporal Hox Gene Colinearity in the Lophotrochozoan Capitella sp. I. PLoS One [Internet] 3. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC2603591/

Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing data. Bioinformatics 28:3150–3152.

Furuhashi T, Schwarzinger C, Miksik I, Smrz M, Beran A. 2009. Molluscan shell evolution with review of shell calcification hypothesis. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 154:351–371.

Gaume B, Denis F, Van Wormhoudt A, Huchette S, Jackson DJ, Avignon S, Auzoux-Bordenave S. 2014. Characterisation and expression of the biomineralising gene Lustrin A during shell formation of

the European abalone Haliotis tuberculata. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 169:1–8.

Gerdol M, Luo Y-J, Satoh N, Pallavicini A. 2018. Genetic and molecular basis of the immune system in the brachiopod Lingula anatina. Dev. Comp. Immunol. 82:7–30.

Giribet G, Edgecombe GD. 2020. Mollusca. In: The Invertebrate Tree of Life. Princeton University Press. p. 358–390.

Glynn PW. 1970. On the ecology of the Caribbean chitons Acanthopleura granulata Gmelin and Chiton tuberculatus Linné: density, mortality, feeding, reproduction, and growth. Smithsonian Contributions to Zoology:1–21.

Gómez-Chiarri M, Warren WC, Guo X, Proestou D. 2015. Developing tools for the study of molluscan immunity: The sequencing of the genome of the eastern oyster, Crassostrea virginica. Fish Shellfish Immunol. 46:2–4.

Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng Q, et al. 2011. Trinity: reconstructing a full-length transcriptome without a genome from RNA-Seq data. Nat Biotechnol 29:644–652.

Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome assemblies. Bioinformatics 29:1072–1075.

Haas BJ, Delcher AL, Mount SM, Wortman JR, Smith RK, Hannick LI, Maiti R, Ronning CM, Rusch DB, Town CD, et al. 2003. Improving the Arabidopsis genome annotation using maximal transcript alignment assemblies. Nucleic Acids Res 31:5654–5666.

Haug-Baltzell A, Stephens SA, Davey S, Scheidegger CE, Lyons E. 2017. SynMap2 and SynMap3D: web-based whole-genome synteny browsers. Bioinformatics 33:2197–2198.

Huan P, Wang Q, Tan S, Liu B. 2019. Dorsoventral dissociation of Hox gene expression underpins the diversification of molluscs. bioRxiv:603092.

Hunt M, Kikuchi T, Sanders M, Newbold C, Berriman M, Otto TD. 2013. REAPR: a universal tool for genome assembly evaluation. Genome Biology 14:R47.

Iijima M, Akiba N, Sarashina I, Kuratani S, Endo K. 2006. Evolution of Hox genes in molluscs: a comparison among seven morphologically diverse classes. J Molluscan Stud 72:259–266.

Irisarri I, Uribe JE, Eernisse DJ, Zardoya R. 2020. A mitogenomic phylogeny of chitons (Mollusca: Polyplacophora). BMC Evol Biol 20:22.

Jackson DJ, McDougall C, Green K, Simpson F, Wörheide G, Degnan BM. 2006. A rapidly evolving secretome builds and patterns a sea shell. BMC Biology 4:40.

Joester D, Brooker LR. 2016. The Chiton Radula: A Model System for Versatile Use of Iron Oxides. Iron Oxides: From Nature to Applications:177–205.

Katoh K, Misawa K, Kuma K, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence alignment based on fast Fourier transform. Nucleic Acids Res 30:3059–3066.

Kenny NJ, Namigai EKO, Marlétaz F, Hui JHL, Shimeld SM. 2015. Draft genome assemblies and predicted microRNA complements of the intertidal lophotrochozoans Patella vulgata (Mollusca, Patellogastropoda) and Spirobranchus (Pomatoceros) lamarcki (Annelida, Serpulida). Mar Genomics 24 Pt 2:139–146.

Kijas J, Hamilton M, Botwright N, King H, McPherson L, Krsinich A, McWilliam S. 2019. Genome Sequencing of Blacklip and Greenlip Abalone for Development and Validation of a SNP Based Genotyping Tool. Front. Genet. [Internet] 9. Available from: https://www.frontiersin.org/articles/10.3389/fgene.2018.00687/full

Kim K-S, Burford MA, Macey DJ, Webb J. 1988. Iron concentrations and characterisation of the major iron binding proteins in the tissues of the chiton Clavarizona hirtosa. Comparative Biochemistry and Physiology Part B: Comparative Biochemistry 91:159–164.

Kim K-S, Macey DJ, Webb J, Mann S. 1989. Iron Mineralization in the Radula Teeth of the Chiton *Acanthopleura hirtosa*. Proceedings of the Royal Society of London B: Biological Sciences 237:335–346.

Kocot KM, Aguilera F, McDougall C, Jackson DJ, Degnan BM. 2016. Sea shell diversity and rapidly evolving secretomes: insights into the evolution of biomineralization. Frontiers in Zoology 13:23.

Kocot KM, Cannon JT, Todt C, Citarella MR, Kohn AB, Meyer A, Santos SR, Schander C, Moroz LL, Lieb B, et al. 2011. Phylogenomics reveals deep molluscan relationships. Nature 477:452.

Kocot KM, Poustka AJ, Stöger I, Halanych KM, Schrödl M. 2020. New data from Monoplacophora and a carefully-curated dataset resolve molluscan relationships. Sci Rep 10:1–8.

Kolde R. 2012. Pheatmap: pretty heatmaps. Available from: https://cran.r-project.org/web/packages/pheatmap/pheatmap.pdf

Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat Methods 9:357–359.

Lee PN, Callaerts P, De Couet HG, Martindale MQ. 2003. Cephalopod Hox genes and the origin of morphological novelties. Nature 424:1061–1065.

Lewis RV. 2006. Spider Silk: Ancient Ideas for New Biomaterials. Chem. Rev. 106:3762–3774.

Li C, Liu X, Liu B, Ma B, Liu F, Liu G, Shi Q, Wang C. 2018. Draft genome of the Peruvian scallop Argopecten purpuratus. Gigascience 7.

Li L, Connors MJ, Kolle M, England GT, Speiser DI, Xiao X, Aizenberg J, Ortiz C. 2015. Multifunctionality of chiton biomineralized armor with an integrated visual system. Science 350:952–956.

Li Yuli, Sun X, Hu X, Xun X, Zhang J, Guo X, Jiao W, Zhang Lingling, Liu W, Wang J, et al. 2017. Scallop genome reveals molecular adaptations to semi-sessile life and neurotoxins. Nat Commun 8:1–11.

Liu C, Zhang Y, Ren Y, Wang H, Li S, Jiang F, Yin L, Qiao X, Zhang G, Qian W, et al. 2018. The genome of the golden apple snail Pomacea canaliculata provides insight into stress tolerance and invasive adaptation. Gigascience [Internet] 7. Available from: https://academic.oup.com/gigascience/article/7/9/giy101/5069392

Lowenstam HA. 1962. Magnetite in Denticle Capping in Recent Chitons (Polyplacophora). Geological Society of America Bulletin 73:435–438.

Masonbrink RE, Purcell CM, Boles SE, Whitehead A, Hyde JR, Seetharam AS, Severin AJ. 2019. An annotated genome for *Haliotis rufescens* (red abalone) and resequenced green, pink, pinto, black, and white abalone species. Genome Biol Evol 11:431–438.

McCartney MA, Auch B, Kono T, Mallez S, Zhang Y, Obille A, Becker A, Abrahante JE, Garbe J, Badalamenti JP, et al. 2019. The genome of the zebra mussel, *Dreissena polymorpha*: a resource for invasive species research. bioRxiv:696732.

McDougall C, Woodcroft BJ, Degnan BM. 2016. The Widespread Prevalence and Functional Significance of Silk-Like Structural Proteins in Metazoan Biological Materials.de Brevern AG, editor. PLoS ONE 11:e0159128.

Modica MV, Lombardo F, Franchini P, Oliverio M. 2015. The venomous cocktail of the vampire snail Colubraria reticulata (Mollusca, Gastropoda). BMC Genomics 16:441.

Murgarella M, Puiu D, Novoa B, Figueras A, Posada D, Canchaya C. 2016. A First Insight into the Genome of the Filter-Feeder Mussel *Mytilus galloprovincialis*. PLoS One [Internet] 11. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC4792442/

Nam B-H, Kwak W, Kim Y-O, Kim D-G, Kong HJ, Kim W-J, Kang J-H, Park JY, An CM, Moon J-Y, et al. 2017. Genome sequence of pacific abalone (Haliotis discus hannai): the first draft genome in family Haliotidae. Gigascience 6:1–8.

Nemoto M, Ren D, Herrera S, Pan S, Tamura T, Inagaki K, Kisailus D. 2019. Integrated transcriptomic and proteomic analyses of a molecular mechanism of radular teeth biomineralization in Cryptochiton stelleri. Sci Rep 9:856.

Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A Fast and Effective Stochastic Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol Biol Evol 32:268–274.

Nielsen H. 2017. Predicting Secretory Proteins with SignalP. In: Kihara D, editor. Protein Function Prediction: Methods and Protocols. Methods in Molecular Biology. New York, NY: Springer. p. 59–73. Available from: https://doi.org/10.1007/978-1-4939-7015-5_6

Patel SV. 2004. A Novel Function of Invertebrate Collagen in the Biomineralization Process During the Shell Repair of Eastern Oyster, Crassostrea Virginica.

Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017. Salmon: fast and bias-aware quantification of transcript expression using dual-phase inference. Nat Methods 14:417–419.

Piccinelli P, Samuelsson T. 2007. Evolution of the iron-responsive element. RNA 13:952–966.

Powell D, Subramanian S, Suwansa-ard S, Zhao M, O'Connor W, Raftos D, Elizur A. 2018. The genome of the oyster Saccostrea offers insight into the environmental resilience of bivalves. DNA Res 25:655–665.

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLOS ONE 5:e9490.

Pryszcz LP, Gabaldón T. 2016. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Res. 44:e113.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics 26:841–842.

Renaut S, Guerra D, Hoeh WR, Stewart DT, Bogan AE, Ghiselli F, Milani L, Passamonti M, Breton S. 2018. Genome Survey of the Freshwater Mussel Venustaconcha ellipsiformis (Bivalvia: Unionida) Using a Hybrid De Novo Assembly Approach. Genome Biol Evol 10:1637–1646.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics 26:139–140.

Roebuck K. 2017. Nuclear Genome Size Diversity Of Marine Invertebrate Taxa Using Flow Cytometric Analysis.

Rogers J, Munro H. 1987. Translation of ferritin light and heavy subunit mRNAs is regulated by intracellular chelatable iron levels in rat hepatoma cells. Proceedings of the National Academy of Sciences 84:2277–2281.

Schell T, Feldmeyer B, Schmidt H, Greshake B, Tills O, Truebano M, Rundle SD, Paule J, Ebersberger I, Pfenninger M. 2017. An Annotated Draft Genome for Radix auricularia (Gastropoda, Mollusca). Genome Biol Evol 9:0.

Scherholz M, Redl E, Wollesen T, Todt C, Wanninger A. 2013. Aplacophoran Mollusks Evolved from Ancestors with Polyplacophoran-like Features. Curr Biol 23:2130–2134.

Schiemann SM, Martín-Durán JM, Børve A, Vellutini BC, Passamaneck YJ, Hejnol A. 2017. Clustered brachiopod Hox genes are not expressed collinearly and are associated with lophotrochozoan novelties. PNAS 114:E1913–E1922.

Schwabe E. 2010. Illustrated summary of chiton terminology. SPIXIANA 33:171–194.

Shaw JA, Brooker LR, Macey DJ. 2002. Radula tooth turnover in the chiton, *Acanthopleura hirtosa* (Blainville, 1825) (Mollusca : Polyplacophora). Molluscan Res. 22:93–99.

Shaw JA, Macey DJ, Brooker LR, Clode PL. 2010. Tooth Use and Wear in Three Iron-Biomineralizing Mollusc Species. The Biological Bulletin 218:132–144.

Sigwart JD, Sutton MD. 2007. Deep molluscan phylogeny: synthesis of palaeontological and neontological data. Proceedings of the Royal Society of London B: Biological Sciences 274:2413–2419.

Sigwart JD, Todt C, Scheltema AH. 2014. Who are the 'Aculifera'? Journal of Natural History 48:2733–2737.

Simakov O, Marletaz F, Cho SJ, Edsinger-Gonzales E, Havlak P, Hellsten U, Kuo DH, Larsson T, Lv J, Arendt D, et al. 2013. Insights into bilaterian evolution from three spiralian genomes. Nature 493:526–531.

Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome assembly and annotation completeness with single-copy orthologs. Bioinformatics 31:3210–3212.

Smit A, Hubley R. 2008. RepeatModeler Open-1.0. Available from: http://www.repeatmasker.org

Smith SA, Wilson NG, Goetz FE, Feehery C, Andrade SCS, Rouse GW, Giribet G, Dunn CW. 2011. Resolving the evolutionary relationships of molluscs with phylogenomic tools. Nature 480:364–367.

Solé-Cava AM, Thorpe JP. 1991. High levels of genetic variation in natural populations of marine lower invertebrates. Biol J Linn Soc 44:65–80.

Speiser DI, Eernisse DJ, Johnsen S. 2011. A Chiton Uses Aragonite Lenses to Form Images. Current Biology 21:665–670.

Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large phylogenies. Bioinformatics 30:1312–1313.

Stanke M, Keller O, Gunduz I, Hayes A, Waack S, Morgenstern B. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res 34:W435–W439.

Sun J, Zhang Yu, Xu T, Zhang Yang, Mu H, Zhang Yanjie, Lan Y, Fields CJ, Hui JHL, Zhang W, et al. 2017. Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes. Nat Ecol Evol 1:1–7.

Suzuki M, Iwashima A, Kimura M, Kogure T, Nagasawa H. 2013. The molecular evolution of the pif family proteins in various species of mollusks. Mar. Biotechnol. 15:145–158.

Suzuki M, Saruwatari K, Kogure T, Yamamoto Y, Nishimura T, Kato T, Nagasawa H. 2009. An acidic matrix protein, Pif, is a key macromolecule for nacre formation. Science 325:1388–1390.

Takeuchi T, Kawashima T, Koyanagi R, Gyoja F, Tanaka M, Ikuta T, Shoguchi E, Fujiwara M, Shinzato C, Hisata K, et al. 2012. Draft genome of the pearl oyster Pinctada fucata: a platform for understanding bivalve biology. DNA Res. 19:117–130.

Thai BT, Lee YP, Gan HM, Austin CM, Croft LJ, Trieu TA, Tan MH. 2019. Whole Genome Assembly of the Snout Otter Clam, Lutraria rhynchaena, Using Nanopore and Illumina Data, Benchmarked Against Bivalve Genome Assemblies. Front Genet [Internet] 10. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6880199/

Thomson AM, Rogers JT, Leedman PJ. 1999. Iron-regulatory proteins, iron-responsive elements and ferritin mRNA translation. The International Journal of Biochemistry & Cell Biology 31:1139–1152.

Vinther J, Jell P, Kampouris G, Carney R, Racicot RA, Briggs DEG. 2012. The origin of multiplacophorans – convergent evolution in Aculiferan molluscs. Palaeontology 55:1007–1019.

Vinther J, Parry L, Briggs DEG, Van Roy P. 2017. Ancestral morphology of crown-group molluscs revealed by a new Ordovician stem aculiferan. Nature 542:471–474.

Vinther J, Sperling EA, Briggs DEG, Peterson KJ. 2012. A molecular palaeobiological hypothesis for the origin of aplacophoran molluscs and their derivation from chiton-like ancestors. Proc. R. Soc. B 279:1259–1268.

Wang H-C, Minh BQ, Susko E, Roger AJ. 2018. Modeling Site Heterogeneity with Posterior Mean Site Frequency Profiles Accelerates Accurate Phylogenomic Estimation. Syst Biol 67:216–235.

Wang S, Zhang J, Jiao W, Li J, Xun X, Sun Y, Guo X, Huan P, Dong B, Zhang L, et al. 2017. Scallop genome provides insights into evolution of bilaterian karyotype and development. Nat Ecol Evol 1:1–12.

Wanninger A, Wollesen T. 2019. The evolution of molluscs. Biological Reviews 94:102–115.

Weckselblatt B, Rudd MK. 2015. Human structural variation: mechanisms of chromosome rearrangements. Trends Genet 31:587–599.

Wick R. 2018. Porechop. Available from: https://github.com/rrwick/Porechop

Wollesen T, Scherholz M, Rodríguez Monje SV, Redl E, Todt C, Wanninger A. 2017. Brain regionalization genes are co-opted into shell field patterning in Mollusca. Sci Rep [Internet] 7. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5511173/

Xu Z, Shi L, Hu D, Hu B, Yang M, Zhu L. 2016. Formation of hierarchical bone-like apatites on silk microfiber templates via biomineralization. RSC Adv. 6:76426–76433.

Zarrella I, Herten K, Maes GE, Tai S, Yang M, Seuntjens E, Ritschard EA, Zach M, Styfhals R, Sanges R, et al. 2019. The survey and reference assisted assembly of the Octopus vulgaris genome. Sci Data [Internet] 6. Available from: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC6472339/

Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, Yang P, Zhang L, Wang X, Qi H, et al. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. Nature 490:49–54.

Zimin AV, Marçais G, Puiu D, Roberts M, Salzberg SL, Yorke JA. 2013. The MaSuRCA genome assembler. Bioinformatics 29:2669–2677.

FIGURE LEGENDS:

Figure 1: (a) The West Indian Fuzzy Chiton Acanthopleura granulata. Photograph by David Liittschwager. (b) A single shell plate from A. granulata. Scale bar indicates 5mm. (c) The eyes (white arrow) and aesthetes (black arrow) of A. granulata. Scale bar indicates 200 µM. Photograph by David Liittschwager. (d) Teeth from the anterior-most region of the radula of A. granulata. The larger teeth, used for feeding, are mineralized with iron oxide (orange) and capped with magnetite (black). Scale bar indicates 300 µM (e) A genome-based phylogeny of Mollusca showing chitons as sister to all other molluscs with available genomes.

Figure 2: Synteny of Hox genes between A. granulata and other taxa. The presence of a gene is indicated by a box of the corresponding color. Continuous black lines indicate that the species has an available genome and Hox genes were located on a contiguous scaffold. Broken black lines indicate that gene(s) are located on multiple genomic scaffolds. A double slash indicates genes are located on a single contiguous scaffold but separated by greater distances than those in most other taxa.

Figure 3: Iron response elements (IREs) in the A. granulata genome. (a) The number of IREs in several molluscan gene model sets, and relative proportions of 5' and 3' IREs. A. granulata has more IREs than all other molluscs examined, but the relative proportion of 5' and 3' IREs appears consistent across molluscan genomes. (b) The relative expression [log10(TPM)] of transcripts containing IREs in the different tissues of A. granulata. The radula is divided into four developmentally distinct regions: R1,the most anterior region, contains teeth used for feeding; R2 contains teeth that are developed but are not yet used for feeding; R3 contains developing teeth that contain iron oxide; and R4, the most posterior region, contains developing teeth that have yet to be coated with iron. We found lower expression of most IRE-containing genes in the anterior regions of the radula.

Figure 4: The two isoforms of heavy-chain ferritin recovered in A. granulata. (a) The locations of the transcription initiation sites and exons of isoform 1 of ferritin (orange, above) and isoform 2 of ferritin (blue, below). A 5' IRE (red) is present in the 5' untranslated region of isoform 1, but not in isoform 2. (b) Relative expression of both isoforms of ferritin across A. granulata tissues. The radula is divided into four developmentally distinct regions as in Figure 3. Isoform 1 is transcribed more highly throughout the body than isoform 2. Isoform 2 is transcribed at lower levels in the anterior (iron-rich) regions of the radula than in other tissues.

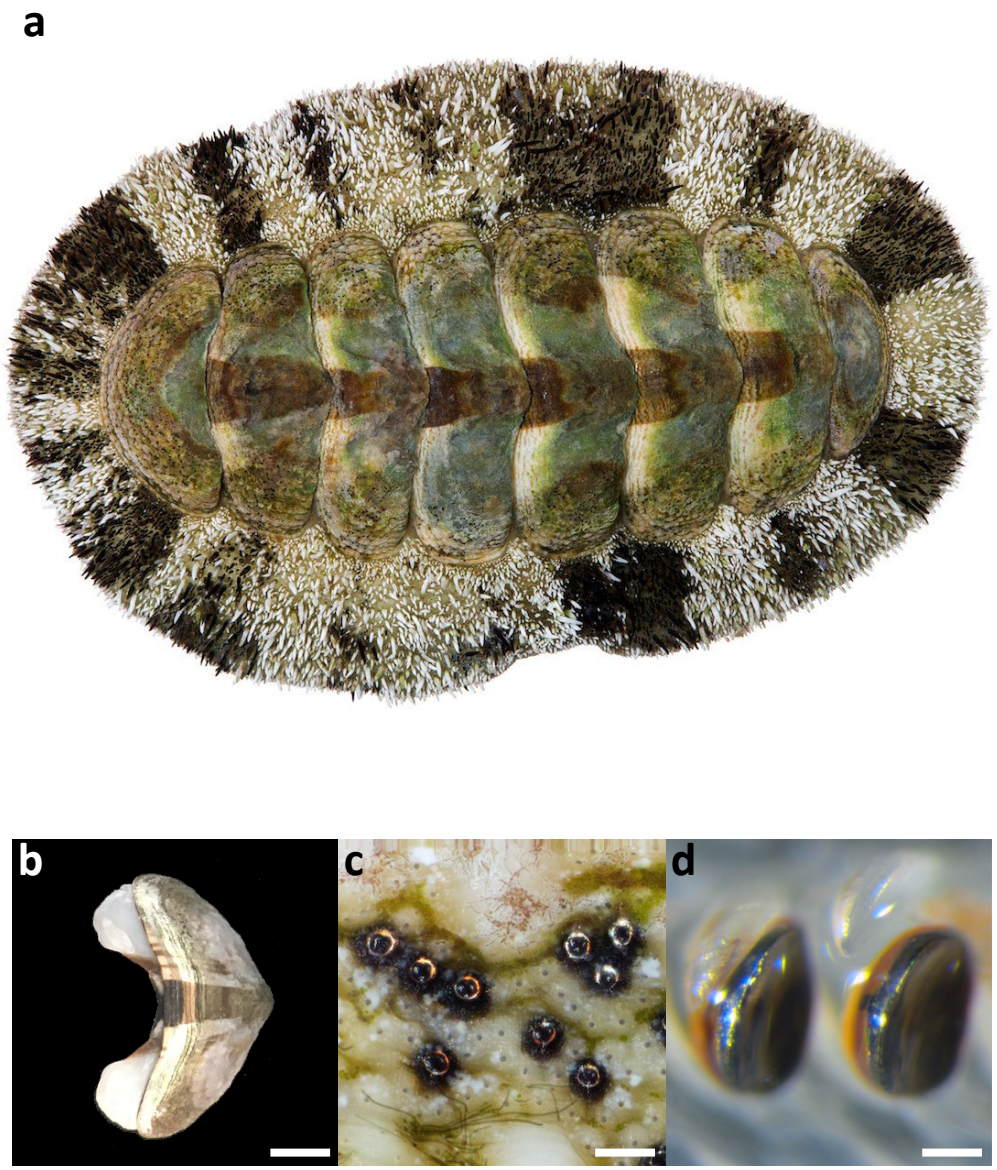
a
b
c
d
e
Acanthopleura granulata
Octopus bimaculoides
Euprymna scolopes
Dreissena rostriformis
Pinctada fucata
Crassostrea virginica
Bathymodiolus platifrons
Scapharca broughtonii
Mizuhopecten yessoensis
Haliotis rufescens
Lottia gigantea
Pomacea canaliculata
Aplysia californica
Elysia chlorotica
Biomphalaria glabrata
Notospermus cf geniculatus
Capitella teleta
Helobdella robusta
Lingula anatina
Phoronis australis
100
99
0.050
Polyplacophora
Cephalopoda
Bivalvia
Gastropoda

Figure 1

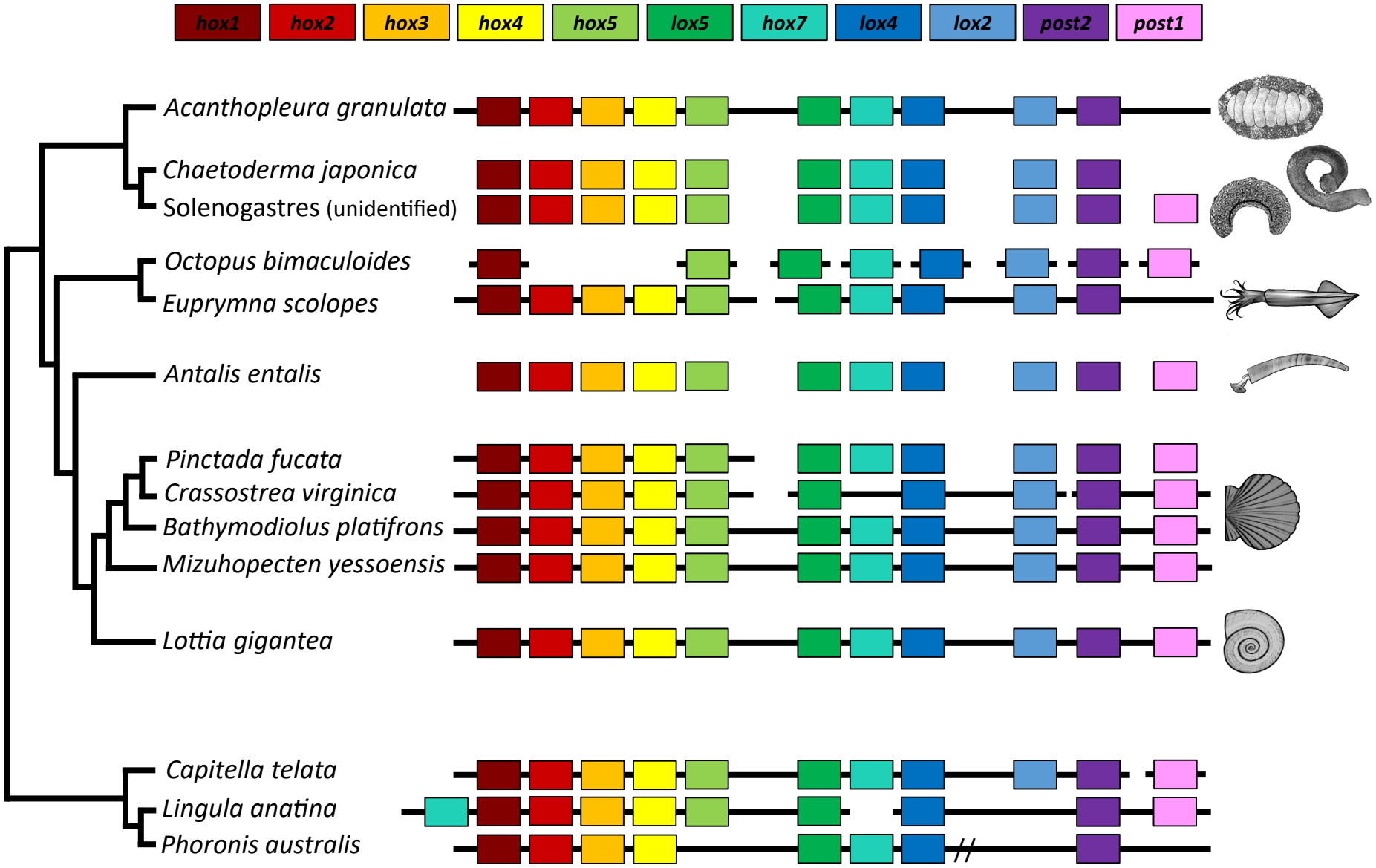
hox1
hox2
hox3
hox4
hox5
lox5
hox7
lox4
lox2
post2
post1
Acanthopleura granulata
Chaetoderma japonica
Solenogastres (unidentified)
Octopus bimaculoides
Euprymna scolopes
Antalis entalis
Pinctada fucata
Crassostrea virginica
Bathymodiolus platifrons
Mizuhopecten yessoensis
Lottia gigantea
Capitella telata
Lingula anatina
Phoronis australis

Figure 2

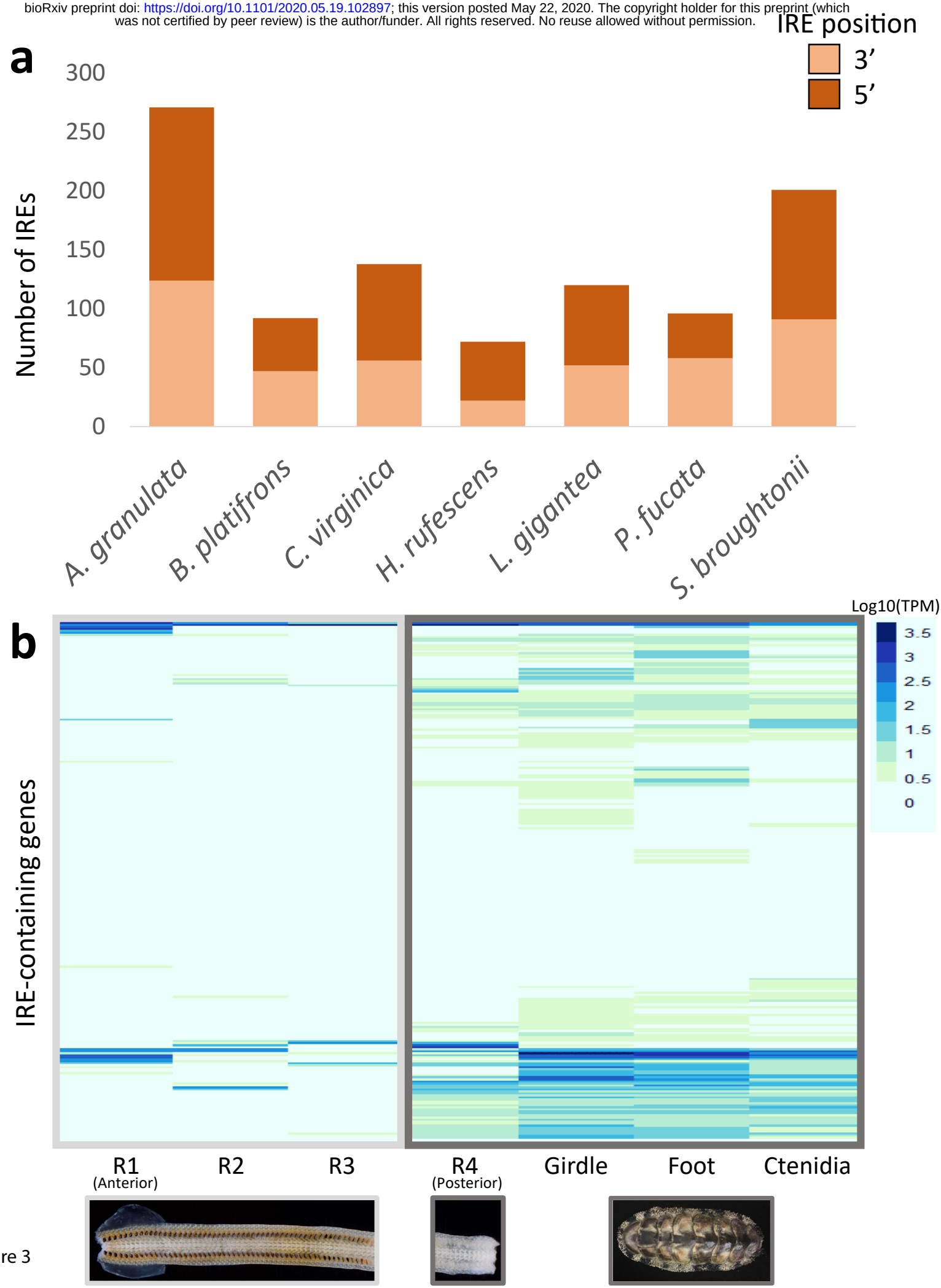

Figure 3

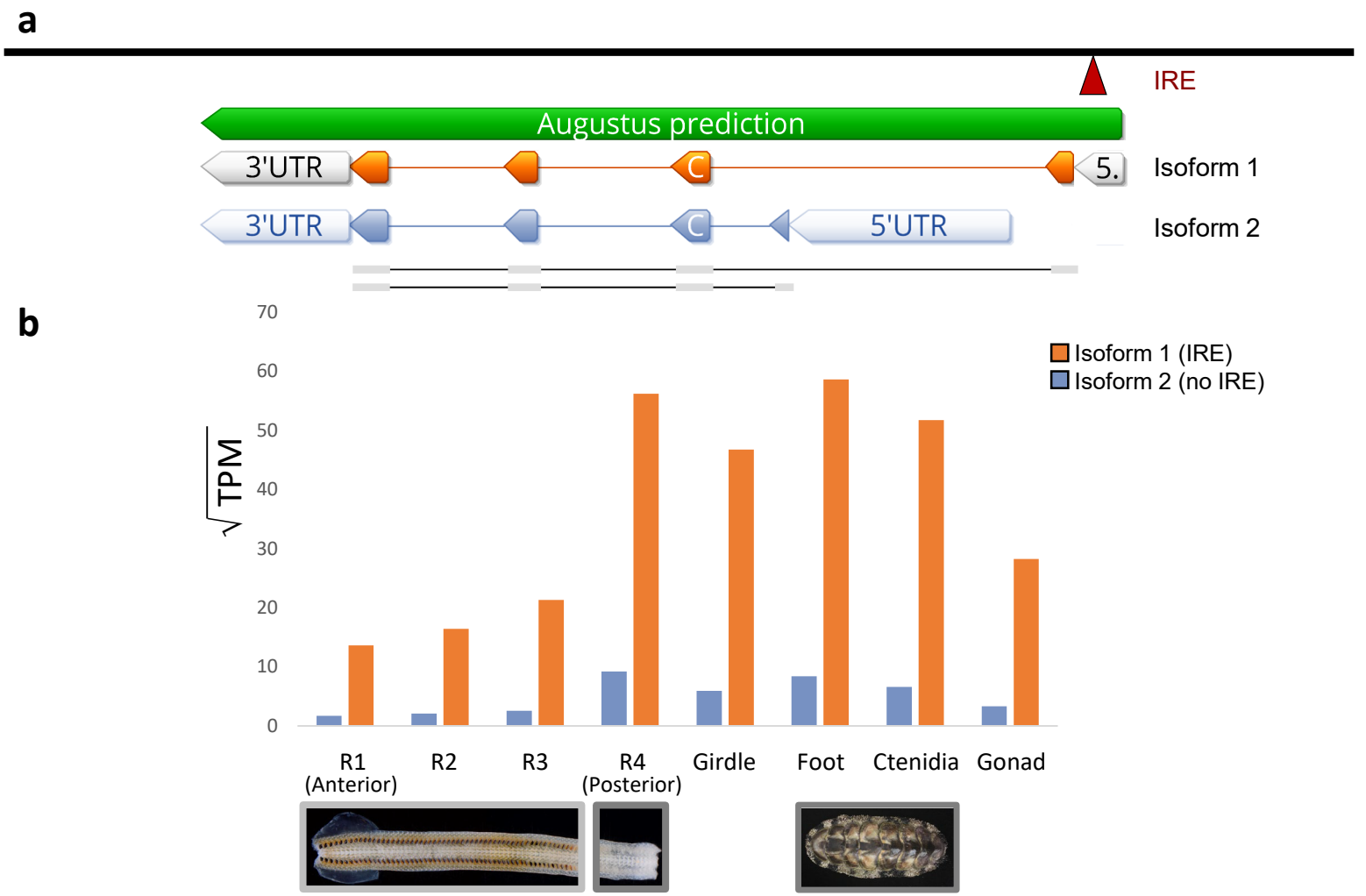
a
IRE
Augustus prediction
3'UTR
C
5.
Isoform 1
3'UTR
C
5'UTR
Isoform 2
b
70
60
50
40
30
20
10
0
√TPM
Isoform 1 (IRE)
Isoform 2 (no IRE)
R1 (Anterior)
R2
R3
R4 (Posterior)
Girdle
Foot
Ctenidia
Gonad

Figure 4