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

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]

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 calcified teeth of their rasp-like radula with a coat of iron (as magnetite). Here we present the genome of the West Indian fuzzy chiton *Acanthopleura granulata*, the first from any aculiferan mollusc. The *A. granulata* genome contains homologs of many biomineralization genes identified previously in conchiferan molluscs. We expected chitons to lack genes previously identified from pathways conchiferans use to make biominerals like calcite and nacre because chitons do not use these materials in their shells. Surprisingly, the *A. granulata* genome has homologs of many of these genes, suggesting that the ancestral mollusc had a more diverse biomineralization toolkit than expected. The *A. granulata* genome has features that may be specialized for iron biomineralization, including a higher proportion of genes regulated directly by iron than other molluscs. *A. granulata* also produces two isoforms of soma-like ferritin: one is regulated by iron and similar in sequence to the soma-like ferritins of other molluscs, and the other is constitutively translated and is not found in other molluscs. The *A. granulata* genome is a resource for future studies of molluscan evolution and biomineralization.

KEY WORDS

biomineralization, biomaterials, iron, ferritin, Iron response element (IRE), mollusc

SIGNIFICANCE STATEMENT

Chitons are molluscs that make shell plates, spine- or scale-like sclerites, and iron-coated teeth. Currently, all molluscs with sequenced genomes lie within one major clade (Conchifera). Sequencing the genome of a representative from the other major clade (Aculifera) helps us learn about the origins and evolution of molluscan traits. The genome of the West Indian Fuzzy Chiton, *Acanthopleura granulata,* reveals chitons have homologs of many genes other molluscs use to make shells, suggesting all molluscs share some shell-making pathways. The genome of *A. granulata* has more genes that may be regulated directly by iron than other molluscs, and chitons produce a unique isoform of a major iron-transport protein (ferritin), suggesting that chitons have genomic specializations that contribute to their production of iron-coated teeth.

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 (McDougall & Degnan 2018). 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 their teeth (Figure 1D) 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 the cores of their teeth with calcium phosphate (as apatite), and then reinforce their cutting edges with iron (as magnetite) (Lowenstam 1962). 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 new rows within days (Shaw et al. 2002; Joester & Brooker 2016). To make new teeth, chitons continuously sequester and transport high concentrations of iron (Kim et al. 1989; Shaw et al. 2002, 2010). Continuous iron biomineralization may thus present a physiological challenge to chitons because free iron causes oxidative stress (Dixon & 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, which is made up of Polyplacophora and Aplacophora. Conchifera and Aculifera diverged approximately 550mya (Vinther 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 & Sutton 2007; Kocot et al. 2011; Smith et al. 2011; Vinther et al. 2012). Despite increasing numbers of sequenced molluscan genomes (e.g., Takeuchi et al. 2012; Zhang et al. 2012; Simakov et al. 2013; Albertin et al. 2015; Gómez-Chiarri et al. 2015; Modica et al. 2015; Kenny et al. 2015; Barghi et al. 2016; Davison et al. 2016; Murgarella et al. 2016; Adema et al. 2017; Du et al. 08 01, 2017; Nam et al. 05 01, 2017; Schell et al. 2017; Sun et al. 2017; Wang et al. 2017; Calcino et al. 2018; Gerdol et al. 05 2018; Li et al. 04 01, 2018; Liu et al. 2018; Renaut et al. 07 01, 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; Sun et al. 2020), 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 the mitigation of iron-induced oxidative stress; and 3) better understand the origin and evolution of biomineralization in molluscs.

RESULTS AND DISCUSSION

We sequenced the genome of a single specimen of *A. granulata*. We combined reads from one lane of Illumina HiSeq X 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, 2020; 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 for *A. granulata* by 1) sequencing transcriptomes from eight different tissues from the same specimen used for genome sequencing, 2) combining these transcriptomes into a single assembly and aligning the combined transcriptome to the genome, and 3) training *de novo* gene predictors using both our combined transcriptome and protein sequences predicted from the transcriptomes of other aculiferans. Following these steps, 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 characterize proteins based on shared functional domains and sequence similarity, we analyzed the set of 20,470 gene models with InterProScan. We identified at least one GO term for 12,301 genes and a Pfam match for 15,710 genes. We also conducted a KEGG analysis and identified 7,341 proteins that could be assigned to putative molecular pathways.

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). We conducted an additional phylogenetic analysis that included more taxa by using transcriptomes in addition to genomes and recovered *Acanthopleura* within the family Chitonidae in the order Chitonida, consistent with recent phylogenetic studies of chitons based on fewer loci (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%, making it one of the least heterozygous molluscan genomes 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 & 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 genome of *A. granulata* to have high heterozygosity because this species is a broadcast spawner with a wide geographic range (Glynn 1970). To compare heterozygosity across molluscs, we selected a set of high-quality molluscan genomes for which short-read data are available (Supplementary Table 1). Using k-mer based analysis, we found the highest heterozygosity among the seven genomes we 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 2), 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. Compared to a non-molluscan lophotrochozoan, *Lingula anatina* (a brachiopod), *A. granulata* has more repetitive elements of certain types in its genome. Conversely, *A. granulata* has fewer repetitive elements in its genome than any conchiferan mollusc (Supplementary Table 2). This suggests multiple proliferations of repetitive elements during molluscan evolution. Repetitive elements contribute to structural changes in genomes by providing breakpoints that increase the likelihood of chromosomal rearrangements (Weckselblatt & Rudd 2015). Consistent with this prediction, synteny is lower between *A. granulata* and all conchiferan molluscs we examined than it is between any two of these conchiferans, and the genomes of conchiferans and *A. granulata* have little synteny with the genome of *L. anatina* (Supplementary Figure 5). A recent study found greater conservation of bilaterian ancestral linkage groups (ALGs) between a scallop and non-molluscs than between other bivalves and non-molluscs, suggesting substantial genomic rearrangements occurred within bivalves (Wang et al. 2017). Molluscan genomes appear to rearrange frequently across evolutionary time, and perhaps rearrange more frequently in conchiferans due to the proliferation of repetitive elements. However, limited synteny between *A. granulata* and the scallop genome suggests that the chiton genome has also

undergone significant rearrangement relative to the hypothetical ancestral bilaterian genome organization.

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, this cluster is dispersed in some bivalves and cephalopods (Albertin et al. 2015; Barucca et al. 2016; Belcaid et al. 2019; Wang et al. 2017). 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 & Wollesen 2019); Figure 2). *Post1* is also absent in two species of chitons from the suborder Acanthochitonina, the sister clade to Chitonina, to which *A. granulata* belongs (Huan et al. 2019; Wanninger & 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 transcription factors to help pattern their body axes and 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). A recent study in the gastropod *Lymnaea stagnalis* identified a PIF-like protein that does not promote nacre production, but is expressed in tissues that suggest this PIF protein still plays a role in shell biomineralization (Ishikawa et al. 2020). 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 (Supplementary Table 4). 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 & Edgecombe 2020; Kocot 01/2013). 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 synthase, 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 5).

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 6). 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 6). 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 6). 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).

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

Chitons have more iron in their hemolymph than any other animal studied to date (Kim et al. 1988). Iron presents physiological challenges to animals because iron can cause oxidative stress. 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 either the 3' or 5' untranslated regions (UTRs) of mRNA molecules and control translation via binding by iron regulatory protein (IRP; Supplementary Figure 7). We also examined IREs in several high-quality molluscan genomes that include UTRs as part of their available annotation data. All of the molluscan genomes we examined had similar proportions of 3' to 5' IREs (Figure 3A). Despite having the fewest gene models, the genome of *A. granulata* has more IREs than the genomes of 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 7). 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 relatively few molluscs that produces hemoglobin for use as a respiratory pigment (Kawamoto 1928; Manwell 1963; Read 1966; Collett & O'Gower 1972; 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 next asked if the expression of genes with IREs in *A. granulata* differs between iron-rich and iron-poor tissues. In the absence of iron, IRPs bind IREs. When this occurs in the 5' UTR of an mRNA, ribosomes are blocked from translating the protein; thus, mRNAs with 5' IREs will not be translated in

the absence of free iron. When IRPs bind IREs in the 3' UTR of an mRNA, they block endonucleases from degrading the mRNA, allowing multiple translations from a single mRNA molecule; thus, the translation of mRNAs with 3' IREs is higher in the absence of free iron (Supplementary Figure 7). We compared the expression of genes with IREs between transcriptomes sequenced from the foot, girdle, ctenidia, and four developmentally-distinct regions of the radula from the same specimen of *A. granulata* we used for genome sequencing, and found expression of genes with IREs in all tissue types (Figure 3B). We quantified the expression of genes that contained 3' IREs and 5' IREs separately and found many genes with 5' IREs had higher expression in the three anterior, iron-rich regions of the radula than the rest of the body (Supplemental Figure 8). Genes with 5' IREs that are expressed in iron-rich tissues are likely transcribed into proteins in only those tissues because elsewhere in the body IRPs will bind to IREs and block translation. Consequently, the proteins encoded by these genes may be unique to the radula.

After identifying genes in *A. granulata* that contain 5' IREs and are expressed at relatively high levels in the iron-rich anterior regions of the radula, we asked if these genes might have roles in the biomineralization of the radula. We used GO analysis to compare the molecular functions of the protein sequences coded by the genes with 5' IREs to the protein sequences coded by the full set of genes from the *A. granulata* genome. We found that genes with a 5' IRE that are highly expressed in the anterior of the radula are more likely to be associated with the molecular functions 'response to inorganic substance', 'response to calcium ion', and 'response to metal ion' (Supplementary Figure 9). This suggests that genes with a 5' IRE that are highly expressed in the radula may be involved in the biomineralization of the apatite (calcium) cores of teeth and their magnetite (iron) caps. A previous study by Nemoto et al. identified a novel biomineralization protein (RTMP1) in the radula of another chiton (*Cryptochiton stelleri*), and proposed that RTMP1 played a role in iron biomineralization (Nemoto et al. 2019). We examined the mRNA of RTMP1 in *C. stelleri* and did not detect an IRE in either its 5' or 3' UTR. Thus there are genes that may be important to biomineralization in the chiton radula whose expression levels are not influenced by IREs.

Two isoforms of ferritin may provide chitons with 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 & 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 & Samuelsson 2007-7). 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, allowing 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 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. We examined other mollusc genomes and transcriptomes and found a ferritin isoform present in all of them that is similar to *A. granulata* isoform 1 and contains the 5' IRE .

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 (Figure 4). We examined other molluscan genomes and transcriptomes and did not find evidence for expression of a ferritin isoform similar to *A. granulata* isoform 2 (data available on Dryad). In *A. granulata*, 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 & Munro 1987). Further, mutations in the 5' IRE of ferritin cause hyperferritinaemia in mammals, an iron-related medical condition caused by an overproduction of ferritin protein (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 through their bodies.

Conclusions

The *A. granulata* genome is the first available genome for any chiton or any aculiferan. The information it provides improves our understanding of the evolution of biomineralization across Mollusca as well as lineage-specific innovations within chitons. Chitons are a valuable system for investigating biomineralization because they produce shell plates, sclerites, 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 conchiferan 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 biomineralize their teeth with magnetite. The genome of *A. granulata* contains more genes with iron response elements (IREs) than the genome 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. Chitons are an emerging model for studies of both biomineralization and iron homeostasis. The *A. granulata* genome will aid

319 future studies by suggesting specific proteins and pathways to target with comparative studies of gene
320 expression and gene manipulation.

METHODS

Specimen collection and preservation

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 (dissected such that the tissue sample would not contain shell secretory tissue), 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. Reads were basecalled with Guppy 4.0. 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 & Gabaldón 07 08, 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 & 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 blob plots (Blaxter & 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 & Hubley 2008-2015), which was used in all subsequent analyses. We trained MAKER v. 2.31.10 (Cantarel et al. 2008-1) 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 & 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.

To identify annotated proteins in *A. granulata,* we first used Transdecoder (Douglas 2018) to produce peptide files of predicted proteins. We ran Interproscan on the set of 20,470 genes referenced above to identify GO terms and Pfam matches for proteins where possible. We used GHOSTX in the Kaas pipeline (Moriya et al. 2007) to identify KEGG pathways via comparisons to all the available molluscan taxa (*Lottia gigantea, Pomacea canaliculata, Crassostrea gigas, Mizuhopecten yessoensis,* and *Octopus bimaculoides*). Finally, we looked for shared GO terms between specific taxa with OrthoVenn(Xu et al. 2019), comparing *A. granulata* to *Lottia gigantea, Chrysomallon squaminiferon, Octopus bimaculoides,* and *Crassostrea gigas* (Supplementary Figure 10).

Hox gene annotation and genomic comparisons

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

To graphically examine synteny between *A. granulata* and other molluscan genome assemblies, we loaded each assembly and annotation into the online COGE SynMap2 (Haug-Baltzell et al. 2017) server and compared *A. granulata* to eight other annotated genomes with default SynMap2 settings. We exported dotplots for each pair of genomes to visualize syntenic regions (or lack thereof). Scaffolds in each dotplot were sorted by length, but differing assembly qualities made some dotplots difficult to read due to a high number of very small scaffolds. We assessed heterozygosity of several molluscan genomes and *A. granulata* by downloading raw paired-end data when possible and using GenomeScope2 online (Vurture et al. 2017).

To permit direct comparisons of repeat content within *A. granulata* and other molluscs, we ran RepeatModeler (Smit & Hubley 2008-2015) on the scaffolds of a subset of genome assemblies and *A. granulata.* We used the same default parameters for each run and quantified the number of elements in each repeat family identified by RepeatModeler for each genome assembly we analyzed (LINEs, SINEs, etc.; Supplementary Table 7).

Orthology inference

To identify orthologous genes shared between *A. granulata* and other molluscs, we used OrthoFinder v. 2.3.7 (Emms & Kelly 2015). We analyzed three separate sets of data: 1) *A. granulata* and genomes of nineteen other lophotrochozoans, including fourteen other molluscs, two annelids, one brachiopod, one phoronid, and one nemertean 2) *A. granulata* and a subset of molluscan genomes for detailed comparisons of biomineralization genes and; 3) *A. granulata* and an expanded set of data including both

436 genomes and transcriptomes, including several transcriptomes from aculiferans other than *A. granulata*.
437 For all three analyses we used the unclustered 81,691 gene set for *A. granulata*, knowing that duplicated
438 gene models would cluster together. We removed sequences from our orthogroups that were identical
439 to longer sequences where they overlapped, as well as fragmented sequences shorter than 100 amino
440 acids, using uniqHaplo(Anon). We retained orthogroups that had a minimum of four taxa, aligned the
441 sequences within them with MAFFT (Katoh et al. 2002), and cleaned mistranslated regions with
442 HmmCleaner (Di Franco et al. 2019). We used AlignmentCompare
443 (https://github.com/kmkocot/basal_metazoan_phylogenomics_scripts_01-2015) to delete sequences
444 that did not overlap with all other sequences by at least 20 AAs (starting with the shortest sequence
445 meeting this criterion).

446 *Phylogenetic analyses*

447 For species tree reconstruction, in cases where two or more sequences were present for any taxon in a
448 single-gene alignment, we used PhyloPyPruner 0.9.5 (https://pypi.org/project/phylopypruner/) to
449 reduce the alignment to a set of strict orthologs. This tool uses single-gene trees to screen putative
450 orthogroups for paralogy. To build single-gene trees based on orthologs, we trimmed alignments with
451 BMGE v1.12.2 (Criscuolo & Gribaldo 2010) and constructed approximately maximum likelihood trees for
452 each alignment with FastTree2(Price et al. 2010) using the "slow" and "gamma" options. We then used
453 these alignments in PhyloPyPruner with the following settings: --min-len 100 --min-support 0.75 --mask
454 pdist --trim-lb 3 --trim-divergent 0.75 --min-pdist 0.01 --trim-freq-paralogs 3 --prune MI. For datasets 1
455 ("genomes") and 3 ("all_taxa"), only orthogroups sampled for at least 85% of the total number of taxa
456 were retained for concatenation. For dataset 2 ("biomin_subset"), only orthogroups sampled for all
457 eight taxa were retained. Phylogenetic analyses were conducted on the supermatrix produced by
458 PhyloPyPruner v. 1.0 in IQ-TREE v. 1.6.12 (Nguyen et al. 2015) using the PMSF model (Wang et al. 2018)
459 with a guide tree based on the LG model. Topological support was assessed with 1,000 rapid bootstraps.

460 *Screening for known biomineralization genes*

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

472 We used the above set of gene queries from other molluscs to identify the ortholog group from the
473 above OrthoFinder2 on the subset of genomes selected as biomineralization representatives across
474 Mollusca. We used the complete CDS or longest mRNA for each gene as a nucleotide query to search
475 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 8.

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. 2017-4) 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 also separated 3’ and 5’ IREs by subsetting the high quality IREs based on whether the IRE was located at the beginning or end of the sequence. We made heatmaps with log-transformed data to compensate for outliers in expression levels with R package prettyheatmap (Kolde 2012). We then analyzed the GO terms enriched in the separate sets of 5’- and 3’-containing genes that were highly expressed in the radula with GOrilla (Eden et al. 2009), using the complete protein set as a background dataset and the sets of IRE-containing genes as the target list.

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, functional annotations, and supporting

documentation for the biomineralization genes described are available in Dryad with the identifier doi:X [For the purposes of peer review, data are currently available via Box: https://alabama.box.com/s/1hsryfff61i01qrljyasrjnu8j7qg2nj]. 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 et al. 2017. Whole genome analysis of a schistosomiasis-transmitting freshwater snail. Nat. Commun. 8. doi: 10.1038/ncomms15451.

Albertin CB et al. 2015. The octopus genome and the evolution of cephalopod neural and morphological novelties. Nature. 524:220–224. doi: 10.1038/nature14668.

Almagro Armenteros JJ et al. 2019. SignalP 5.0 improves signal peptide predictions using deep neural networks. Nat. Biotechnol. 37:420–423. doi: 10.1038/s41587-019-0036-z.

Bai C-M et al. 2019. Chromosomal-level assembly of the blood clam, Scapharca (Anadara) broughtonii, using long sequence reads and Hi-C. Gigascience. 8. doi: 10.1093/gigascience/giz067.

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. doi: 10.1007/s00438-015-1119-2.

Barucca M, Canapa A, Biscotti MA. 2016. An Overview of Hox Genes in Lophotrochozoa: Evolution and Functionality. Journal of Developmental Biology. 4:12. doi: 10.3390/jdb4010012.

Belcaid M et al. 2019. Symbiotic organs shaped by distinct modes of genome evolution in cephalopods. Proc. Natl. Acad. Sci. U. S. A. 201817322. doi: 10.1073/pnas.1817322116.

Blaxter M, Challis R. 2018. BlobToolkit. BlobToolKit. https://github.com/blobtoolkit.

Cai H et al. 2019. A draft genome assembly of the solar-powered sea slug *Elysia chlorotica*. Scientific Data. 6:190022. doi: 10.1038/sdata.2019.22.

Calcino AD et al. 2018. The quagga mussel genome and the evolution of freshwater tolerance: Supplementary Material. bioRxiv. 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. doi: 10.1093/nar/gkq371.

Cantarel BL et al. 2008-1. MAKER: An easy-to-use annotation pipeline designed for emerging model organism genomes. Genome Res. 18:188–196. doi: 10.1101/gr.6743907.

Checa AG, Vendrasco MJ, Salas C. 2017. Cuticle of Polyplacophora: structure, secretion, and homology with the periostracum of conchiferans. Mar. Biol. 164:64. doi: 10.1007/s00227-017-3100-6.

Cölfen H. 2010. A crystal-clear view. Nat. Mater. 9:960–961. doi: 10.1038/nmat2911.

Collett LC, O'Gower AK. 1972. Molluscan hemoglobins with unusual temperature-dependent characteristics. Comp. Biochem. Physiol. A Comp. Physiol. 41:843–850. doi: 10.1016/0300-9629(72)90346-5.

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 Evol. Biol. 10:210. doi: 10.1186/1471-2148-10-210.

566 Davison A et al. 2016. Formin Is Associated with Left-Right Asymmetry in the Pond Snail and the Frog.
567 Curr. Biol. 26:654–660. doi: 10.1016/j.cub.2015.12.071.

568 Di Franco A, Poujol R, Baurain D, Philippe H. 2019. Evaluating the usefulness of alignment filtering
569 methods to reduce the impact of errors on evolutionary inferences. BMC Evol. Biol. 19:21. doi:
570 10.1186/s12862-019-1350-2.

571 Dixon SJ, Stockwell BR. 2014. The role of iron and reactive oxygen species in cell death. Nat. Chem. Biol.
572 10:9–17. doi: 10.1038/nchembio.1416.

573 Douglas P. 2018. TransDecoder/TransDecoder. GitHub. https://github.com/TransDecoder/TransDecoder
574 (Accessed March 23, 2020).

575 Du X et al. 08 01, 2017. The pearl oyster Pinctada fucata martensii genome and multi-omic analyses
576 provide insights into biomineralization. Gigascience. 6:1–12. doi: 10.1093/gigascience/gix059.

577 Eden E, Navon R, Steinfeld I, Lipson D, Yakhini Z. 2009. GOrilla: a tool for discovery and visualization of
578 enriched GO terms in ranked gene lists. BMC Bioinformatics. 10:48. doi: 10.1186/1471-2105-10-48.

579 Eisoldt L, Smith A, Scheibel T. 2011. Decoding the secrets of spider silk. Mater. Today. 14:80–86. doi:
580 10.1016/S1369-7021(11)70057-8.

581 Emms DM, Kelly S. 2015. OrthoFinder: solving fundamental biases in whole genome comparisons
582 dramatically improves orthogroup inference accuracy. Genome Biol. 16:157. doi: 10.1186/s13059-015-
583 0721-2.

584 Fröbius AC, Matus DQ, Seaver EC. 2008. Genomic Organization and Expression Demonstrate Spatial and
585 Temporal Hox Gene Colinearity in the Lophotrochozoan Capitella sp. I. PLoS One. 3. doi:
586 10.1371/journal.pone.0004004.

587 Fu L, Niu B, Zhu Z, Wu S, Li W. 2012. CD-HIT: accelerated for clustering the next-generation sequencing
588 data. Bioinformatics. 28:3150–3152. doi: 10.1093/bioinformatics/bts565.

589 Furuhashi T, Schwarzinger C, Miksik I, Smrz M, Beran A. 2009. Molluscan shell evolution with review of
590 shell calcification hypothesis. Comp. Biochem. Physiol. B Biochem. Mol. Biol. 154:351–371. doi:
591 10.1016/j.cbpb.2009.07.011.

592 Gaume B et al. 2014. Characterisation and expression of the biomineralising gene Lustrin A during shell
593 formation of the European abalone Haliotis tuberculata. Comp. Biochem. Physiol. B Biochem. Mol. Biol.
594 169:1–8. doi: 10.1016/j.cbpb.2013.11.010.

595 Gerdol M, Luo Y-J, Satoh N, Pallavicini A. 05 2018. Genetic and molecular basis of the immune system in
596 the brachiopod Lingula anatina. Dev. Comp. Immunol. 82:7–30. doi: 10.1016/j.dci.2017.12.021.

597 Giribet G, Edgecombe GD. 2020. Mollusca. In: The Invertebrate Tree of Life. Princeton University Press
598 pp. 358–390. https://books.google.com/books?id=anetDwAAQBAJ.

599 Glynn P. 1970. On the Ecology of the Carribean Chitons *Acanthopleura granulata* Gmelin and *Chiton*
600 *tuberculatus* Linee: Density, Mortality, Feeding, Reproduction, and Growth. Smithson. Contrib. Zool.

601 Gómez-Chiarri M, Warren WC, Guo X, Proestou D. 2015. Developing tools for the study of molluscan
602 immunity: The sequencing of the genome of the eastern oyster, Crassostrea virginica. Fish Shellfish
603 Immunol. 46:2–4. doi: 10.1016/j.fsi.2015.05.004.

604 Grabherr MG et al. 2011. Trinity: reconstructing a full-length transcriptome without a genome from
605 RNA-Seq data. Nat. Biotechnol. 29:644–652. doi: 10.1038/nbt.1883.

606 Gurevich A, Saveliev V, Vyahhi N, Tesler G. 2013. QUAST: quality assessment tool for genome
607 assemblies. Bioinformatics. 29:1072–1075. doi: 10.1093/bioinformatics/btt086.

608 Haas BJ et al. 2003. Improving the Arabidopsis genome annotation using maximal transcript alignment
609 assemblies. Nucleic Acids Res. 31:5654–5666. doi: 10.1093/nar/gkg770.

610 Haug-Baltzell A, Stephens SA, Davey S, Scheidegger CE, Lyons E. 2017. SynMap2 and SynMap3D: web-
611 based whole-genome synteny browsers. Bioinformatics. 33:2197–2198. doi:
612 10.1093/bioinformatics/btx144.

613 Huan P, Wang Q, Tan S, Liu B. 2019. Dorsoventral dissociation of Hox gene expression underpins the
614 diversification of molluscs. bioRxiv. 603092. doi: 10.1101/603092.

615 Hunt M et al. 2013. REAPR: a universal tool for genome assembly evaluation. Genome Biol. 14:R47. doi:
616 10.1186/gb-2013-14-5-r47.

617 Iijima M, Akiba N, Sarashina I, Kuratani S, Endo K. 2006. Evolution of Hox genes in molluscs: a
618 comparison among seven morphologically diverse classes. J. Molluscan Stud. 72:259–266. doi:
619 10.1093/mollus/eyl001.

620 Irisarri I, Uribe JE, Eernisse DJ, Zardoya R. 2020. A mitogenomic phylogeny of chitons (Mollusca:
621 Polyplacophora). BMC Evol. Biol. 20:22. doi: 10.1186/s12862-019-1573-2.

622 Ishikawa A et al. 2020. Functional shell matrix proteins tentatively identified by asymmetric snail shell
623 morphology. Sci. Rep. 10:9768. doi: 10.1038/s41598-020-66021-w.

624 Jackson DJ et al. 2006. A rapidly evolving secretome builds and patterns a sea shell. BMC Biol. 4:40. doi:
625 10.1186/1741-7007-4-40.

626 Joester D, Brooker LR. 2016. The Chiton Radula: A Model System for Versatile Use of Iron Oxides. Iron
627 Oxides: From Nature to Applications. 177–205. doi: 10.1002/9783527691395.ch8.

628 Katoh K, Misawa K, Kuma K-I, Miyata T. 2002. MAFFT: a novel method for rapid multiple sequence
629 alignment based on fast Fourier transform. Nucleic Acids Res. 30:3059–3066.
630 https://www.ncbi.nlm.nih.gov/pmc/articles/PMC135756/ (Accessed February 22, 2019).

631 Kawamoto N. 1928. Oxygen capacity of the blood of certain invertebrates which contains hemoglobin.
632 Sci. Rep. Tohoku Univ. . 3:561–575.

633 Kenny NJ, Namigai EKO, Marlétaz F, Hui JHL, Shimeld SM. 2015. Draft genome assemblies and predicted
634 microRNA complements of the intertidal lophotrochozoans Patella vulgata (Mollusca,
635 Patellogastropoda) and Spirobranchus (Pomatoceros) lamarcki (Annelida, Serpulida). Mar. Genomics. 24
636 Pt 2:139–146. doi: 10.1016/j.margen.2015.07.004.

637 Kijas J et al. 2019. Genome Sequencing of Blacklip and Greenlip Abalone for Development and Validation
638 of a SNP Based Genotyping Tool. Front. Genet. 9. doi: 10.3389/fgene.2018.00687.

639 Kim K-S, Burford MA, Macey DJ, Webb J. 1988. Iron concentrations and characterisation of the major
640 iron binding proteins in the tissues of the chiton Clavarizona hirtosa. Comparative Biochemistry and
641 Physiology Part B: Comparative Biochemistry. 91:159–164. doi: 10.1016/0305-0491(88)90129-0.

642 Kim K-S, Macey DJ, Webb J, Mann S. 1989. Iron Mineralization in the Radula Teeth of the Chiton
643 *Acanthopleura hirtosa*. Proceedings of the Royal Society of London B: Biological Sciences. 237:335–346.
644 doi: 10.1098/rspb.1989.0052.

645 Kocot KM et al. 2011. Phylogenomics reveals deep molluscan relationships. Nature. 477:452. doi:
646 10.1038/nature10382.

647 Kocot KM. 01/2013. Recent advances and unanswered questions in deep molluscan phylogenetics *. Am.
648 Malacol. Bull. 31:195–208. doi: 10.4003/006.031.0112.

649 Kocot KM, Aguilera F, McDougall C, Jackson DJ, Degnan BM. 2016. Sea shell diversity and rapidly
650 evolving secretomes: insights into the evolution of biomineralization. Front. Zool. 13:23. doi:
651 10.1186/s12983-016-0155-z.

652 Kocot KM, Poustka AJ, Stöger I, Halanych KM, Schrödl M. 2020. New data from Monoplacophora and a
653 carefully-curated dataset resolve molluscan relationships. Sci. Rep. 10:1–8. doi: 10.1038/s41598-019-
654 56728-w.

655 Kolde R. 2012. *Pheatmap: pretty heatmaps*. https://cran.r-
656 project.org/web/packages/pheatmap/pheatmap.pdf.

657 Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. Nat. Methods. 9:357–359.
658 doi: 10.1038/nmeth.1923.

659 Lee PN, Callaerts P, De Couet HG, Martindale MQ. 2003. Cephalopod Hox genes and the origin of
660 morphological novelties. Nature. 424:1061–1065. doi: 10.1038/nature01872.

661 Lewis RV. 2006. Spider Silk: Ancient Ideas for New Biomaterials. Chem. Rev. 106:3762–3774. doi:
662 10.1021/cr010194g.

663 Li C et al. 04 01, 2018. Draft genome of the Peruvian scallop Argopecten purpuratus. Gigascience. 7. doi:
664 10.1093/gigascience/giy031.

665 Li L et al. 2015. Multifunctionality of chiton biomineralized armor with an integrated visual system.
666 Science. 350:952–956. doi: 10.1126/science.aad1246.

667 Liu C et al. 2018. The genome of the golden apple snail Pomacea canaliculata provides insight into stress
668 tolerance and invasive adaptation. Gigascience. 7. doi: 10.1093/gigascience/giy101.

669 Lowenstam HA. 1962. Magnetite in Denticle Capping in Recent Chitons (Polyplacophora). Geol. Soc. Am.
670 Bull. 73:435–438. doi: 10.1130/0016-7606(1962)73[435:MIDCIR]2.0.CO;2.

671 Manwell C. 1963. The chemistry and biology of hemoglobin in some marine clams—I. Distribution of the

672 pigment and properties of the oxygen equilibrium. Comp. Biochem. Physiol. 8:209–218. doi:
673 10.1016/0010-406X(63)90125-7.

674 Masonbrink RE et al. 2019. An annotated genome for *Haliotis rufescens* (red abalone) and resequenced
675 green, pink, pinto, black, and white abalone species. Genome Biol. Evol. 11:431–438. doi:
676 10.1093/gbe/evz006.

677 McCartney MA et al. 2019. The genome of the zebra mussel, *Dreissena polymorpha*: a resource for
678 invasive species research. bioRxiv. 696732. doi: 10.1101/696732.

679 McDougall C, Degnan BM. 2018. The evolution of mollusc shells. Wiley Interdiscip. Rev. Dev. Biol.
680 0:e313. doi: 10.1002/wdev.313.

681 McDougall C, Woodcroft BJ, Degnan BM. 2016. The Widespread Prevalence and Functional Significance
682 of Silk-Like Structural Proteins in Metazoan Biological Materials De Brevern, AG, editor. PLoS One.
683 11:e0159128. doi: 10.1371/journal.pone.0159128.

684 Modica MV, Lombardo F, Franchini P, Oliverio M. 2015. The venomous cocktail of the vampire snail
685 Colubraria reticulata (Mollusca, Gastropoda). BMC Genomics. 16:441. doi: 10.1186/s12864-015-1648-4.

686 Moriya Y, Itoh M, Okuda S, Yoshizawa AC, Kanehisa M. 2007. KAAS: an automatic genome annotation
687 and pathway reconstruction server. Nucleic Acids Res. 35:W182–5. doi: 10.1093/nar/gkm321.

688 Murgarella M et al. 2016. A First Insight into the Genome of the Filter-Feeder Mussel *Mytilus*
689 *galloprovincialis*. PLoS One. 11. doi: 10.1371/journal.pone.0151561.

690 Nam B-H et al. 05 01, 2017. Genome sequence of pacific abalone (Haliotis discus hannai): the first draft
691 genome in family Haliotidae. Gigascience. 6:1–8. doi: 10.1093/gigascience/gix014.

692 Nemoto M et al. 2019. Integrated transcriptomic and proteomic analyses of a molecular mechanism of
693 radular teeth biomineralization in Cryptochiton stelleri. Sci. Rep. 9:856. doi: 10.1038/s41598-018-37839-
694 2.

695 Nguyen L-T, Schmidt HA, von Haeseler A, Minh BQ. 2015. IQ-TREE: A Fast and Effective Stochastic
696 Algorithm for Estimating Maximum-Likelihood Phylogenies. Mol. Biol. Evol. 32:268–274. doi:
697 10.1093/molbev/msu300.

698 Nielsen H. 2017. Predicting Secretory Proteins with SignalP. In: Protein Function Prediction: Methods
699 and Protocols. Kihara, D, editor. Methods in Molecular Biology Springer: New York, NY pp. 59–73. doi:
700 10.1007/978-1-4939-7015-5_6.

701 Patel SV. 2004. A Novel Function of Invertebrate Collagen in the Biomineralization Process During the
702 Shell Repair of Eastern Oyster, Crassostrea Virginica. Clemson University
703 https://books.google.com/books?id=1Ew2twAACAAJ.

704 Patro R, Duggal G, Love MI, Irizarry RA, Kingsford C. 2017-4. Salmon: fast and bias-aware quantification
705 of transcript expression using dual-phase inference. Nat. Methods. 14:417–419. doi:
706 10.1038/nmeth.4197.

707 Piccinelli P, Samuelsson T. 2007-7. Evolution of the iron-responsive element. RNA. 13:952–966. doi:

10.1261/rna.464807.

Powell D et al. 2018. The genome of the oyster Saccostrea offers insight into the environmental resilience of bivalves. DNA Res. 25:655–665. doi: 10.1093/dnares/dsy032.

Price MN, Dehal PS, Arkin AP. 2010. FastTree 2 – Approximately Maximum-Likelihood Trees for Large Alignments. PLoS One. 5:e9490. doi: 10.1371/journal.pone.0009490.

Pryszcz LP, Gabaldón T. 07 08, 2016. Redundans: an assembly pipeline for highly heterozygous genomes. Nucleic Acids Res. 44:e113. doi: 10.1093/nar/gkw294.

Quinlan AR, Hall IM. 2010. BEDTools: a flexible suite of utilities for comparing genomic features. Bioinformatics. 26:841–842. doi: 10.1093/bioinformatics/btq033.

Read KRH. 1966. Molluscan hemoglobin and myoglobin. In 'Physiology of Mollusca'(KM Wilbur and CM Yonge, eds.), Vol. 2.

Renaut S et al. 07 01, 2018. Genome Survey of the Freshwater Mussel Venustaconcha ellipsiformis (Bivalvia: Unionida) Using a Hybrid De Novo Assembly Approach. Genome Biol. Evol. 10:1637–1646. doi: 10.1093/gbe/evy117.

Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. Bioinformatics. 26:139–140. doi: 10.1093/bioinformatics/btp616.

Roebuck K. 2017. Nuclear Genome Size Diversity Of Marine Invertebrate Taxa Using Flow Cytometric Analysis. https://api.semanticscholar.org/b7c1cd8822ff2c02c54e2654469ea4d1278a71e9.

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. doi: 10.1073/pnas.84.8.2277.

Schell T et al. 2017. An Annotated Draft Genome for Radix auricularia (Gastropoda, Mollusca). Genome Biol. Evol. 9:0. doi: 10.1093/gbe/evx032.

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. doi: 10.1016/j.cub.2013.08.056.

Schiemann SM et al. 2017. Clustered brachiopod Hox genes are not expressed collinearly and are associated with lophotrochozoan novelties. Proc. Natl. Acad. Sci. U. S. A. 114:E1913–E1922. doi: 10.1073/pnas.1614501114.

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. doi: 10.1071/mr02005.

Shaw JA, Macey DJ, Brooker LR, Clode PL. 2010. Tooth Use and Wear in Three Iron-Biomineralizing Mollusc Species. Biol. Bull. 218:132–144. doi: 10.1086/BBLv218n2p132.

743 Sigwart JD, Sutton MD. 2007. Deep molluscan phylogeny: synthesis of palaeontological and
744 neontological data. Proceedings of the Royal Society of London B: Biological Sciences. 274:2413–2419.
745 doi: 10.1098/rspb.2007.0701.

746 Sigwart JD, Todt C, Scheltema AH. 2014. Who are the 'Aculifera'? J. Nat. Hist. 48:2733–2737. doi:
747 10.1080/00222933.2014.964788.

748 Simakov O et al. 2013. Insights into bilaterian evolution from three spiralian genomes. Nature. 493:526–
749 531. doi: 10.1038/nature11696.

750 Simão FA, Waterhouse RM, Ioannidis P, Kriventseva EV, Zdobnov EM. 2015. BUSCO: assessing genome
751 assembly and annotation completeness with single-copy orthologs. Bioinformatics. 31:3210–3212. doi:
752 10.1093/bioinformatics/btv351.

753 Smit AFA, Hubley R. 2008-2015. RepeatModeler Open-1.0. http://www.repeatmasker.org.

754 Smith SA et al. 2011. Resolving the evolutionary relationships of molluscs with phylogenomic tools.
755 Nature. 480:364–367. doi: 10.1038/nature10526.

756 Solé-Cava AM, Thorpe JP. 1991. High levels of genetic variation in natural populations of marine lower
757 invertebrates. Biol. J. Linn. Soc. Lond. 44:65–80. doi: 10.1111/j.1095-8312.1991.tb00607.x.

758 Speiser DI, Eernisse DJ, Johnsen S. 2011. A Chiton Uses Aragonite Lenses to Form Images. Curr. Biol.
759 21:665–670. doi: 10.1016/j.cub.2011.03.033.

760 Stamatakis A. 2014. RAxML version 8: a tool for phylogenetic analysis and post-analysis of large
761 phylogenies. Bioinformatics. 30:1312–1313. doi: 10.1093/bioinformatics/btu033.

762 Stanke M et al. 2006. AUGUSTUS: ab initio prediction of alternative transcripts. Nucleic Acids Res.
763 34:W435–W439. doi: 10.1093/nar/gkl200.

764 Sun J et al. 2017. Adaptation to deep-sea chemosynthetic environments as revealed by mussel genomes.
765 Nat Ecol Evol. 1:1–7. doi: 10.1038/s41559-017-0121.

766 Sun J et al. 2020. The Scaly-foot Snail genome and implications for the origins of biomineralised armour.
767 Nat. Commun. 11:1–12. doi: 10.1038/s41467-020-15522-3.

768 Suzuki M et al. 2009. An acidic matrix protein, Pif, is a key macromolecule for nacre formation. Science.
769 325:1388–1390. doi: 10.1126/science.1173793.

770 Suzuki M, Iwashima A, Kimura M, Kogure T, Nagasawa H. 2013. The molecular evolution of the pif family
771 proteins in various species of mollusks. Mar. Biotechnol. 15:145–158. doi: 10.1007/s10126-012-9471-2.

772 Takeuchi T et al. 2012. Draft genome of the pearl oyster Pinctada fucata: a platform for understanding
773 bivalve biology. DNA Res. 19:117–130. doi: 10.1093/dnares/dss005.

774 Thai BT et al. 2019. Whole Genome Assembly of the Snout Otter Clam, Lutraria rhynchaena, Using
775 Nanopore and Illumina Data, Benchmarked Against Bivalve Genome Assemblies. Front. Genet. 10. doi:
776 10.3389/fgene.2019.01158.

777 Thomson AM, Rogers JT, Leedman PJ. 1999. Iron-regulatory proteins, iron-responsive elements and

778 ferritin mRNA translation. Int. J. Biochem. Cell Biol. 31:1139–1152. doi: 10.1016/S1357-2725(99)00080-
779 1.

780 Vinther J, Parry L, Briggs DEG, Van Roy P. 2017. Ancestral morphology of crown-group molluscs revealed
781 by a new Ordovician stem aculiferan. Nature. 542:471–474. doi: 10.1038/nature21055.

782 Vinther J, Sperling EA, Briggs DEG, Peterson KJ. 2012. A molecular palaeobiological hypothesis for the
783 origin of aplacophoran molluscs and their derivation from chiton-like ancestors. Proc. R. Soc. B.
784 279:1259–1268. doi: 10.1098/rspb.2011.1773.

785 Vurture GW et al. 2017. GenomeScope: fast reference-free genome profiling from short reads.
786 Bioinformatics. 33:2202–2204. doi: 10.1093/bioinformatics/btx153.

787 Wang H-C, Minh BQ, Susko E, Roger AJ. 2018. Modeling Site Heterogeneity with Posterior Mean Site
788 Frequency Profiles Accelerates Accurate Phylogenomic Estimation. Syst. Biol. 67:216–235. doi:
789 10.1093/sysbio/syx068.

790 Wang S et al. 2017. Scallop genome provides insights into evolution of bilaterian karyotype and
791 development. Nat Ecol Evol. 1:1–12. doi: 10.1038/s41559-017-0120.

792 Wanninger A, Wollesen T. 2019. The evolution of molluscs. Biol. Rev. Camb. Philos. Soc. 94:102–115.
793 doi: 10.1111/brv.12439.

794 Weckselblatt B, Rudd MK. 2015. Human structural variation: mechanisms of chromosome
795 rearrangements. Trends Genet. 31:587–599. doi: 10.1016/j.tig.2015.05.010.

796 Wick RR. 2018. *Porechop*. https://github.com/rrwick/Porechop.

797 Wollesen T et al. 2017. Brain regionalization genes are co-opted into shell field patterning in Mollusca.
798 Sci. Rep. 7. doi: 10.1038/s41598-017-05605-5.

799 Xu L et al. 2019. OrthoVenn2: a web server for whole-genome comparison and annotation of
800 orthologous clusters across multiple species. Nucleic Acids Res. 47:W52–W58. doi: 10.1093/nar/gkz333.

801 Xu Z et al. 2016. Formation of hierarchical bone-like apatites on silk microfiber templates via
802 biomineralization. RSC Adv. 6:76426–76433. doi: 10.1039/C6RA17199K.

803 Zarrella I et al. 2019. The survey and reference assisted assembly of the Octopus vulgaris genome. Sci
804 Data. 6. doi: 10.1038/s41597-019-0017-6.

805 Zhang G et al. 2012. The oyster genome reveals stress adaptation and complexity of shell formation.
806 Nature. 490:49–54. doi: 10.1038/nature11413.

807 Zimin AV et al. 2013. The MaSuRCA genome assembler. Bioinformatics. 29:2669–2677. doi:
808 10.1093/bioinformatics/btt476.

809

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. (LogTPM used to allow both data ranges to appear legibly on the same graph)

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.

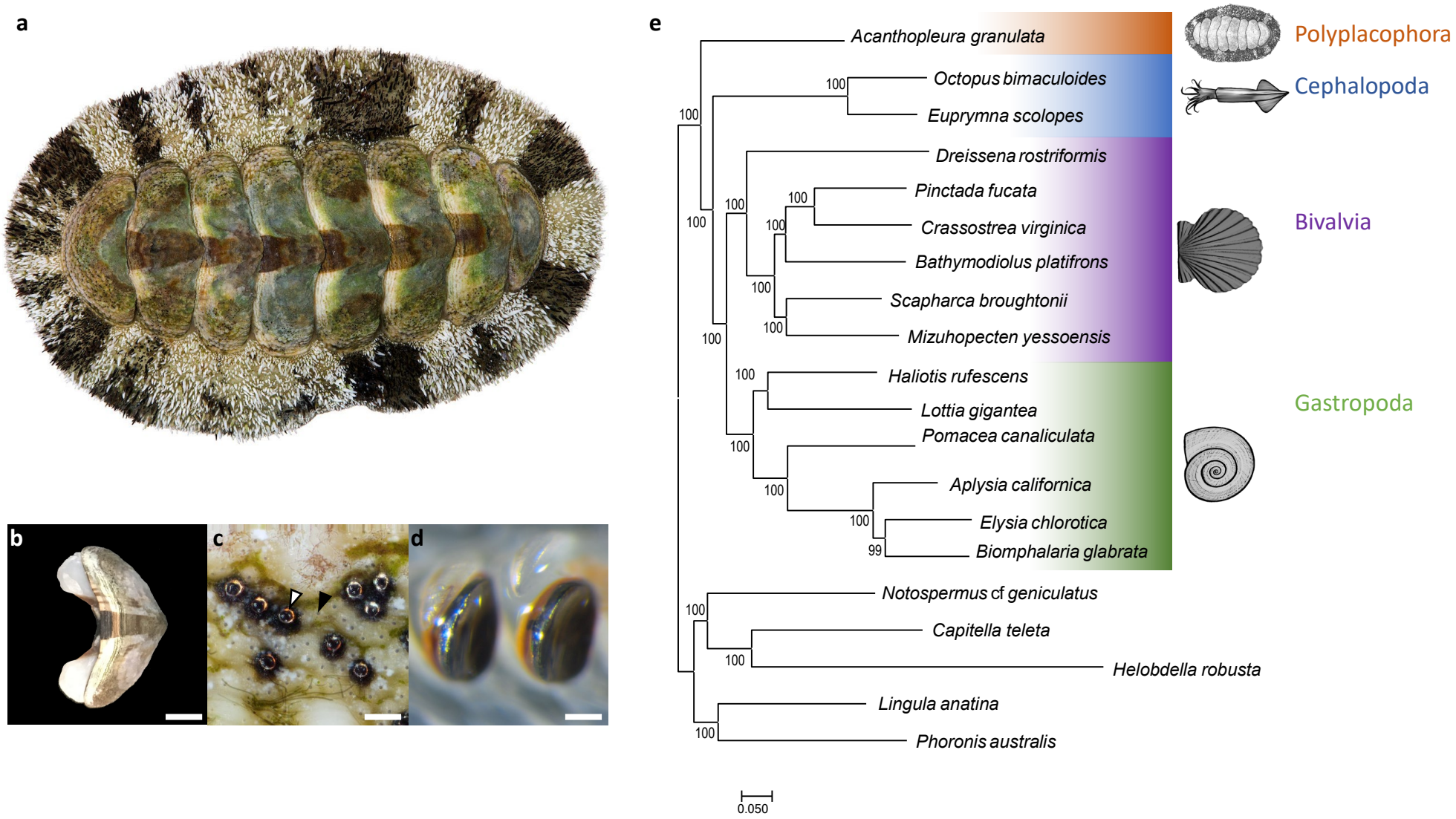

Figure 1: (a) The West Indian Fuzzy Chiton *Acanthopleura granulata*. Photograph by David Liittschwager. **(b)** A single shell plate from *A. granulata.* Scale bar indicates 5mm. **(c)** The eyes (white arrow) and aesthetes (black arrow) of *A. granulata*. Scale bar indicates 200 µM. Photograph by David 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.

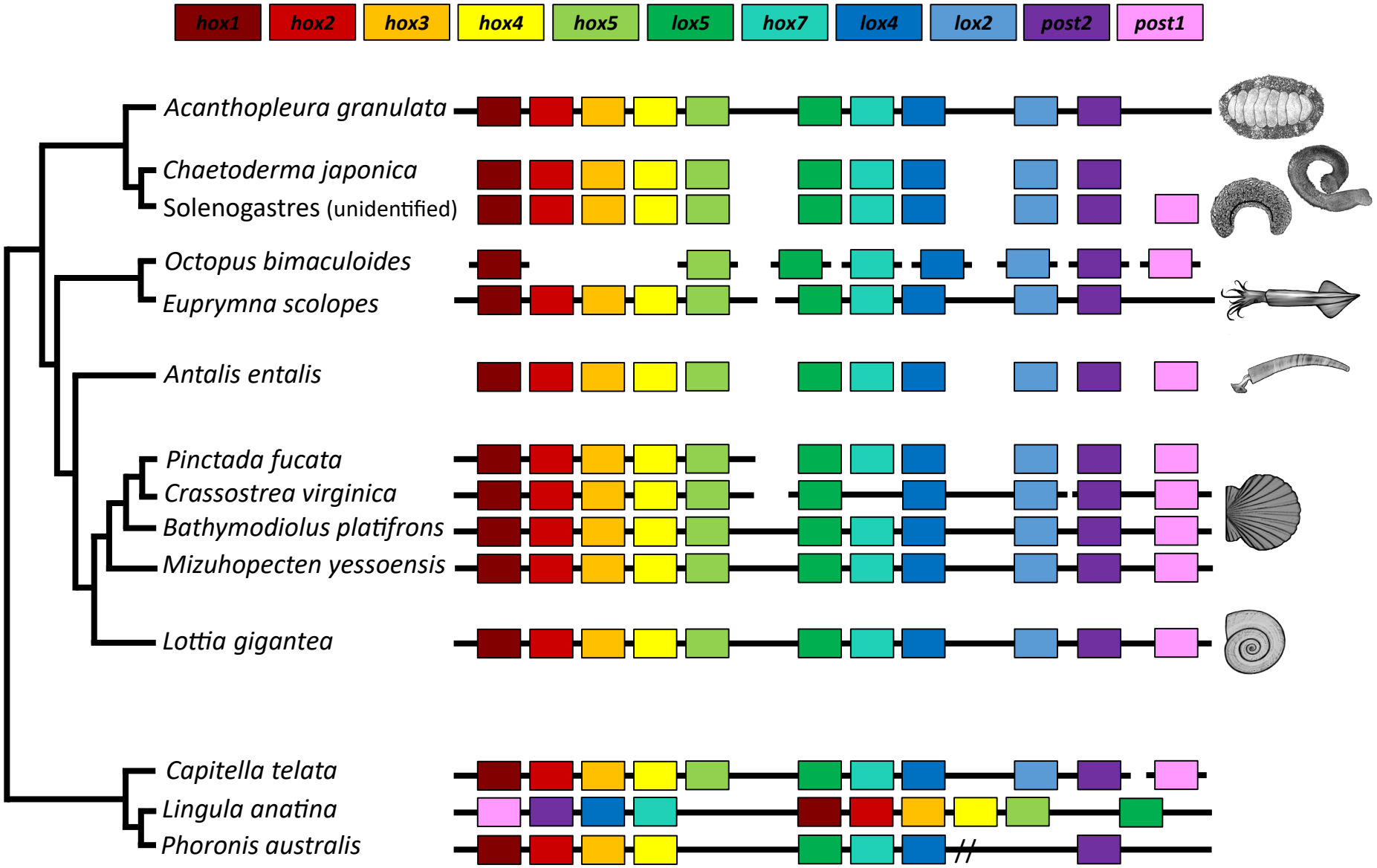

Figure 2: Synteny of Hox genes between A. granulata and other taxa. The presence of a gene is indicated by a box of the corresponding color. Continuous black lines indicate that the species has an available genome and Hox genes were located on a contiguous scaffold. Broken black lines indicate that gene(s) 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.

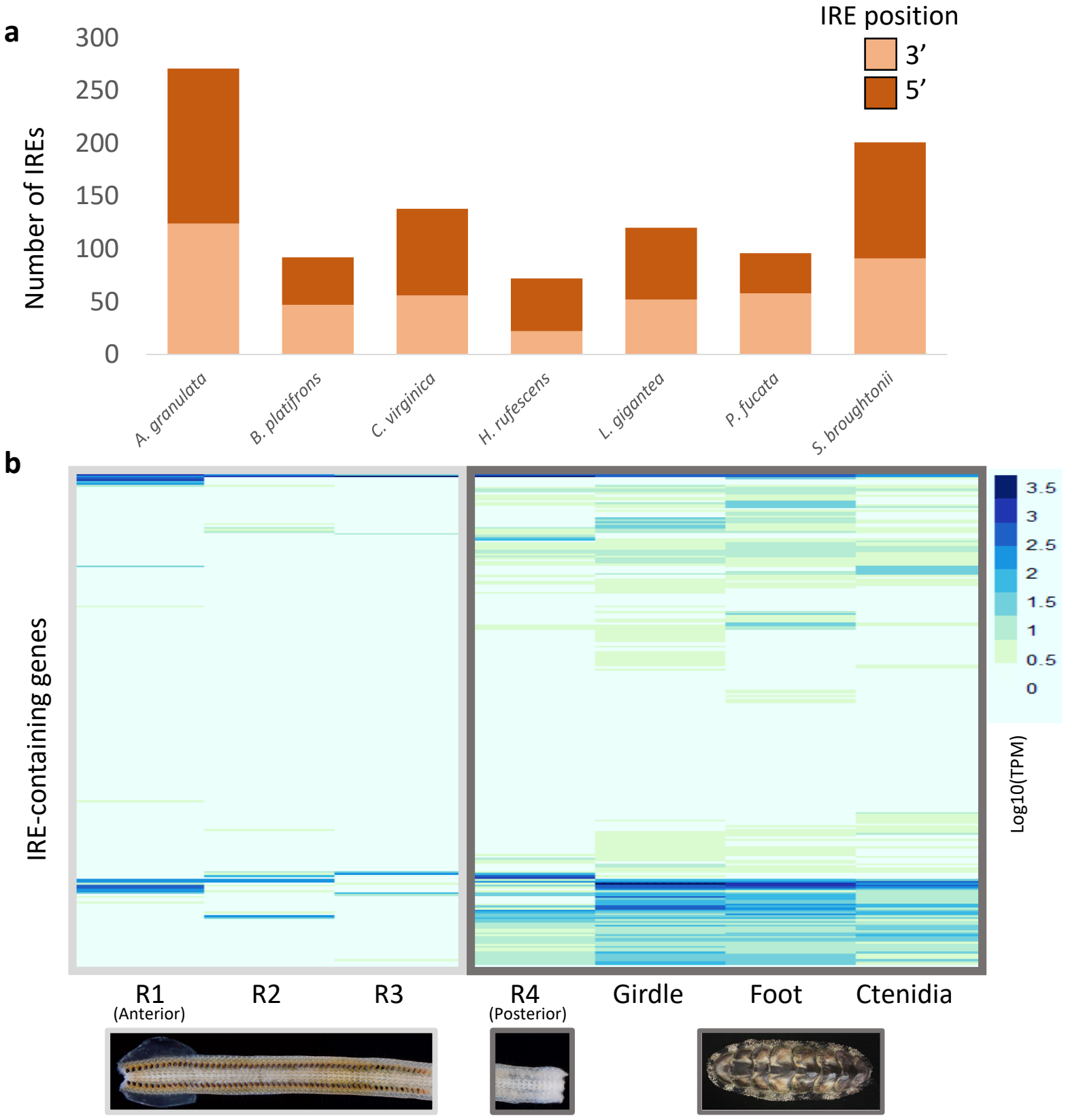

Figure 3: Iron response elements (IREs) in the *A. granulata* genome. **(a)** The number of IREs in several molluscan gene model sets, and relative proportions of 5' and 3' IREs. *A. granulata* has more IREs than all other molluscs examined, but the relative proportion of 5' and 3' IREs appears consistent 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.

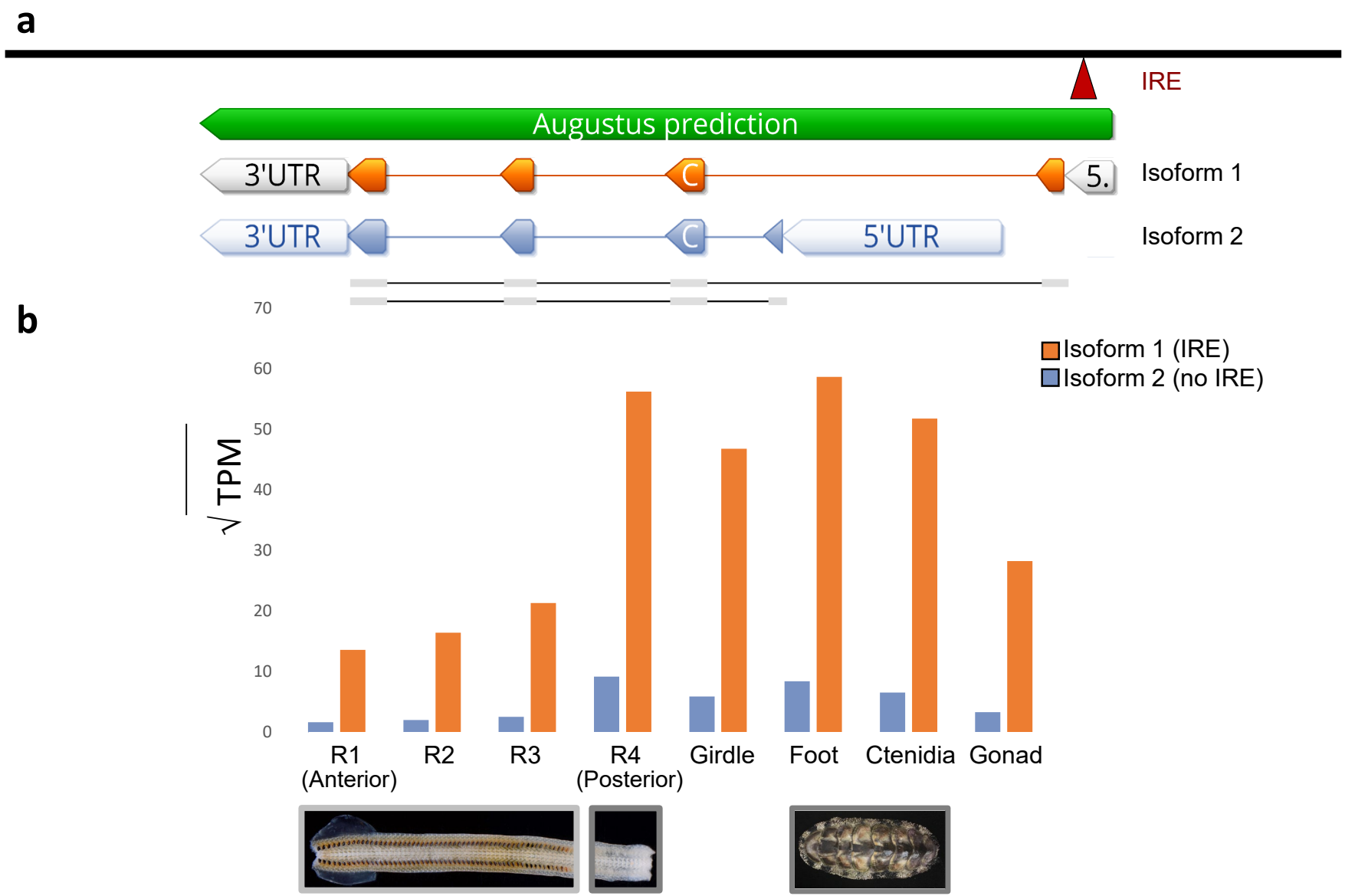

Figure 4: The two isoforms of heavy-chain ferritin recovered in *A. granulata*. **(a)** The locations of the transcription initiation sites and exons of isoform 1 of ferritin (orange, above) and isoform 2 of ferritin (blue, below). A 5' IRE (red) is present in the 5' untranslated region of isoform 1, 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.