



THE UNIVERSITY *of* EDINBURGH

Edinburgh Research Explorer

## A bivalve biomineralization toolbox

### Citation for published version:

Yarra, T, Blaxter, M & Clark, MS 2021, 'A bivalve biomineralization toolbox', *Molecular Biology and Evolution*, vol. 38, no. 9, pp. 4043-4055. <https://doi.org/10.1093/molbev/msab153>

### Digital Object Identifier (DOI):

[10.1093/molbev/msab153](https://doi.org/10.1093/molbev/msab153)

### Link:

[Link to publication record in Edinburgh Research Explorer](#)

### Document Version:

Publisher's PDF, also known as Version of record

### Published In:

Molecular Biology and Evolution

### General rights

Copyright for the publications made accessible via the Edinburgh Research Explorer is retained by the author(s) and / or other copyright owners and it is a condition of accessing these publications that users recognise and abide by the legal requirements associated with these rights.

### Take down policy

The University of Edinburgh has made every reasonable effort to ensure that Edinburgh Research Explorer content complies with UK legislation. If you believe that the public display of this file breaches copyright please contact [openaccess@ed.ac.uk](mailto:openaccess@ed.ac.uk) providing details, and we will remove access to the work immediately and investigate your claim.



# A Bivalve Biomineralization Toolbox

Tejaswi Yarra,<sup>1,2</sup> Mark Blaxter,<sup>3</sup> and Melody S. Clark<sup>\*2</sup>

<sup>1</sup>Ashworth Laboratories, Institute of Evolutionary Biology, University of Edinburgh, Edinburgh, United Kingdom

<sup>2</sup>British Antarctic Survey, Natural Environment Research Council, Cambridge, United Kingdom

<sup>3</sup>Sanger Institute, Wellcome Genome Campus, Hinxton, Cambridgeshire, United Kingdom

\*Corresponding author: E-mail: mscl@bas.ac.uk

Associate editor: John True

## Abstract

Mollusc shells are a result of the deposition of crystalline and amorphous calcite catalyzed by enzymes and shell matrix proteins (SMP). Developing a detailed understanding of bivalve mollusc biomineralization pathways is complicated not only by the multiplicity of shell forms and microstructures in this class, but also by the evolution of associated proteins by domain co-option and domain shuffling. In spite of this, a minimal biomineralization toolbox comprising proteins and protein domains critical for shell production across species has been identified. Using a matched pair design to reduce experimental noise from inter-individual variation, combined with damage-repair experiments and a database of biomineralization SMPs derived from published works, proteins were identified that are likely to be involved in shell calcification. Eighteen new, shared proteins likely to be involved in the processes related to the calcification of shells were identified by the analysis of genes expressed during repair in *Crassostrea gigas*, *Mytilus edulis*, and *Pecten maximus*. Genes involved in ion transport were also identified as potentially involved in calcification either via the maintenance of cell acid–base balance or transport of critical ions to the extrapallial space, the site of shell assembly. These data expand the number of candidate biomineralization proteins in bivalve molluscs for future functional studies and define a minimal functional protein domain set required to produce solid microstructures from soluble calcium carbonate. This is important for understanding molluscan shell evolution, the likely impacts of environmental change on biomineralization processes, materials science, and biomimicry research.

**Key words:** damage-repair, *Crassostrea gigas*, *Mytilus edulis*, *Pecten maximus*, transcriptomics, biomineralization.

## Introduction

Calcium carbonate is a major skeletal component in the natural world. Although the chemical reactions to produce solid calcium carbonate for marine skeletons are relatively simple, the structure of the biologically synthesized end products varies enormously (Clark 2020). This is particularly true within the Mollusca, which comprise over 70,000 extant species, most of which have biomineralized shells, exhibiting a vast array of shapes, sizes, and colors (Rosenberg 2014; Williams 2017). Molluscan shells are complex layered biocomposites. Calcium carbonate crystals, present in either the calcite or aragonite form, are embedded in a softer organic matrix comprising proteins, acidic polysaccharides and often, chitin (Marin 2020). While comprising a minimal proportion of the total dry shell weight (generally up to 5%), these proteins are critical factors in crystal nucleation events and crystal growth (Marin et al. 2005; Zhang et al. 2006; Kong et al. 2009; Suzuki et al. 2009; Fang et al. 2011). The exact mix of these proteins with calcium carbonate crystals significantly influences the particular shell microstructure (e.g., prism, nacre, foliae, and crossed-lamellar structures [reviewed in Clark et al. 2020]).

Although different shell microstructures appear to share common patterns of nucleation and growth, they have evolved independently several times across or within different molluscan classes (e.g., Vendrasco et al. 2011). The wide multiplicity of microstructures and evolutionary origins is reflected at the molecular level (Marin 2020). A large percentage of genes that are putatively involved in biomineralization have no functional annotation and identical shell microstructures in different species are generated by very few shared genes (Jackson et al. 2006; 2010; Marin 2020). Transcriptomic analyses of mantle tissue, the shell-secreting organ, have shown that biomineralization gene sets are characterized by rapid evolution with lineage-restricted proteins and unique combinations of co-opted ancient genes, often resulting from domain shuffling (Kocot et al. 2016; Aguilera et al. 2017). Some genes and proteins or domains have long evolutionary histories associated with biomineralization, such as carbonic anhydrase and tyrosinase (Aguilera et al. 2014; Le Roy et al. 2014). A minimal biomineralization toolbox, a critical set of functional protein domains required to produce basic calcium carbonate microstructures in molluscs, has been proposed (Arivalagan et al. 2017). This toolbox, based on shell

© The Author(s) 2021. Published by Oxford University Press on behalf of the Society for Molecular Biology and Evolution.

This is an Open Access article distributed under the terms of the Creative Commons Attribution License (<http://creativecommons.org/licenses/by/4.0/>), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.

Open Access

proteomic studies, is currently limited to four domains shared between evolutionarily distant species whose shells have different microstructures (Arivalagan et al. 2017). Therefore at the molecular level, there are still many unknowns. These range from how calcium is transported within tissues and cells to the site of manufacture, the mantle tissue and extrapallial space, and how the basic chemical reaction to make calcium carbonate crystals is linked to the co-ordinated production and secretion of the organic component to form the complex microstructures in mollusc shells.

Transcriptomic identification of genes, or more precisely, functional protein domains, involved in biomineralization has previously been derived from comparisons between tissues, or from developmental time courses encompassing initial shell formation in larvae (Zhang et al. 2012; Sleight et al. 2016; Herlitze et al. 2018; Zhao et al. 2018). In this study, shell damage was used to stimulate biomineralization, and repair followed using transcriptomics. Through this approach functional domains were identified that were involved not only in the biomineralization process but also the cellular transport of calcium, with the identification of putative biomineralization domains enhanced by the use of a curated biomineralization database (<https://doi.org/10/cz2w>). Wild populations of molluscs have high genetic variability, and between-individual variation can dwarf responses to experimental treatments. To overcome this, a matched-pair design was used with the untreated valve used as the control for the damaged valve from the same individual. The species investigated were the great scallop *Pecten maximus*, the Pacific oyster *Crassostrea gigas* and the blue mussel *Mytilus edulis*. These species are evolutionary distant, with *P. maximus* and *C. gigas* separating 420 Ma and *M. edulis* diverging from the other two around 480 Ma (Plazzi and Passamonti 2010; Ren et al. 2010), and have distinct shell microstructures. The *C. gigas* shell is formed purely of calcite, with inner layers of foliated calcite, a middle layer of chalky calcite and an outer layer of calcite prisms (Marie, Zanella-Cleon, et al. 2011). The *P. maximus* shell has a middle layer of cross-lamellar aragonite surrounded by outer and inner layers of foliated calcite (Grefsrud et al. 2008), whereas the *M. edulis* shell has inner and middle layers of aragonite nacre and an outer layer of calcite prisms (Marie, Le Roy, et al. 2011). Given the phylogenetic and structural difference between our study species, functional protein domains found to be expressed during shell repair in all three are likely to be part of the evolutionary conserved bivalve biomineralization toolbox, encompassing not only proteins within the shell itself, but also those involved in calcium transport and shell fabrication. Understanding the minimal biochemical processes used in molluscs to convert soluble calcium carbonate into complex hard skeletons will provide a valuable framework for understanding how variations of this basic process lead to the vast array of shell structural morphology seen in extant molluscs.

## Results

In this study, shell damage-repair experiments were conducted on three species of bivalve molluscs: the blue mussel

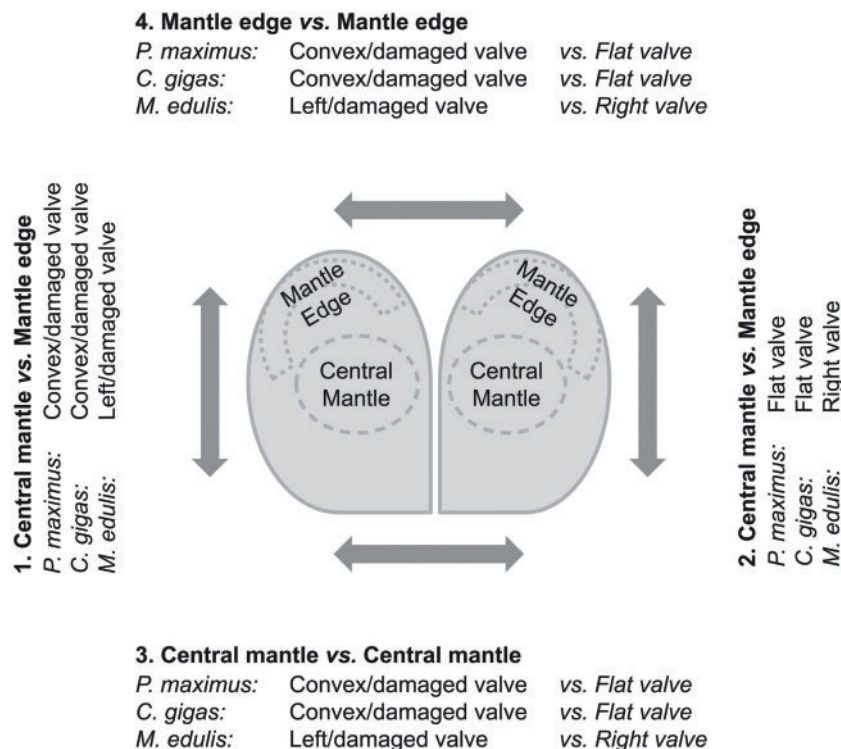
(*M. edulis*), the Pacific oyster (*C. gigas*), and the great scallop (*P. maximus*). Gene transcription profiles were analyzed from the mantle edge and the central mantle tissue sections from damaged and control valves of the same individual using a matched pair design as previously described in Hüning et al. (2016) and Yarra et al. (in review), with holes drilled in the centers of the shells above the central mantle zone in cohorts of wild-sampled, live bivalves (fig. 1). All three species were successful in initiating and completing shell repair of their damaged valves. RNA sequencing of the mantle tissue generated 191 million paired reads for *P. maximus*, 299 million paired reads for *M. edulis*, and 217 million paired reads for *C. gigas*. These data were assembled using the Trinity pipeline (Grabherr et al. 2011), inferring from 234,020 (*C. gigas*) to 560,776 (*M. edulis*) transcript fragments (“Trinity genes”) (table 1). Filtering of these assemblies to eliminate expression noise was carried out by removing all contigs with low expression levels (less than one mapped read per million mapped reads in at least half of the libraries). This filtering was based on at least 16 libraries in *P. maximus* and at least 10 libraries in *C. gigas* and *M. edulis*. This reduced the number of putative transcripts from one fifth to one third of the original total, deriving from 26,077 to 30,822 putative genes (table 1). The common dispersion values (as a measure of biological variability [McCarthy et al. 2012]) between the four mantle tissue sections were high when the matched pair nature of the experiment was not included as a factor. However, much lower dispersion values were observed under the matched pair model, which permitted an accurate quantification of differential gene expression (table 1).

Four pairwise comparisons of differential gene expression were carried out to identify genes associated with damage-repair (fig. 1). Samples were taken at a single time point in each species, with sampling conducted when there was clear evidence of shell calcification in the damaged area. Therefore, genes identified as differentially expressed in the repairing tissue were highly likely involved in biomineralization. To evaluate any effects of shell asymmetry (in *P. maximus* and *C. gigas*) transcriptomic analysis was conducted on undrilled animals. Analysis of differential gene expression of mantle edge and central mantle tissues from the convex and flat valves of undamaged *P. maximus* revealed very few differences, suggesting that these have a similar gene repertoire and that shell asymmetry would not bias further analyses (supplementary information S1, Supplementary Material online).

## Differential Gene Expression Associated with Shell Synthesis

A database has been collated of shell matrix proteins (SMP) identified in a range of mollusc species and the sequences annotated with protein domains and putative functional roles (<https://doi.org/10/cz2w>). This database represents a snapshot of the current state of understanding of gene and protein involvement in shell biology and it was used to both understand and extend the catalogue of SMPs.

In all three species, MDS plots showed separation in gene expression between the edge and central mantle sections,



**Fig. 1.** The four pairwise comparisons carried out in the damage-repair experiment.

**Table 1.** Mantle Transcriptome Assembly Metrics.

	<i>Pecten maximus</i>	<i>Crassostrea gigas</i>	<i>Mytilus edulis</i>
<b>Main assembly</b>			
Trinity transcripts	561,732	387,722	874,699
Trinity genes	320,962	234,020	560,776
% GC	38.01	38.83	32.97
N50 (bp)	1,177	909	784
Minimum length (bp)	201	201	201
Maximum length (bp)	26,479	27,343	26,467
Total assembled bases (Mb)	401	235	496
<b>Filtered assembly</b>			
Trinity transcripts	95,276	122,479	158,889
Trinity genes	26,077	29,869	30,822
<b>Common dispersion</b>			
Nonmatched pair model	0.35	0.32	0.55
Matched pair model	0.16	0.04	0.12

treatment, and between individuals ([supplementary information S2, Supplementary Material](#) online). Many contigs were differentially expressed between the central mantle sections of the damaged and undamaged valves, indicative of differential expression caused by active shell repair ([table 2](#) and [supplementary information S3, Supplementary Material](#) online). Very few contigs were observed to be differentially expressed between the mantle edge sections of the damaged and undamaged valves ([table 2](#)) implying no organism-wide effects of the damage and repair process (for example that might have arisen from infection or stress). Hence, further analyses focused on the genes differentially expressed between the central mantle sections of the damaged and undamaged valves. There was no similarity of enriched GO terms for differentially expressed sequences when damaged

central mantle was compared with undamaged central mantle in the matched pair analysis, although some terms identified in *M. edulis* differentially expressed genes, such as chitin-binding and peptidase inhibitor activity, are common functional terms associated with some SMPs ([supplementary information S4, Supplementary Material](#) online).

SMPs identified as differentially expressed between the damaged and undamaged central mantle sections in all the targets species often had different expression patterns (identified as either up-regulated, down-regulated, or variable) between the three species ([figs. 2](#) and [3](#)). In both *P. maximus* and *C. gigas*, some contigs with annotations as SMPs were also down-regulated in the central mantle of the damaged valve ([fig. 2a and b](#)), while no contigs were down-regulated in *M. edulis* ([fig. 2c](#)). Several different contigs that were annotated



**Table 2.** Number of Differentially Expressed (DE) Contigs in the Different Mantle Tissue Sections.

	<i>Pecten maximus</i>		<i>Crassostrea gigas</i>		<i>Mytilus edulis</i>	
	DE	SMP	DE	SMP	DE	SMP
<b>Central mantle vs. Mantle edge:</b>						
<b>Damaged (convex/left valve)</b>						
Mantle edge	2,303	17	1,924	39	7,340	155
Central mantle	1,305	34	1,957	39	6,229	28
<b>Central mantle vs. Mantle edge:</b>						
<b>Undamaged (flat/right valve)</b>						
Mantle edge	1,669	19	1,086	39	8,855	220
Central mantle	1,831	36	397	10	7,221	23
<b>Convex/left vs. flat/right valve:</b>						
<b>Central mantle</b>						
Mantle edge	635	19	2,323	70	653	54
Central mantle	1,066	16	1,802	24	0	0
<b>Convex/left vs. flat/right valve:</b>						
<b>Mantle edge</b>						
Mantle edge	1	0	16	2	0	0
Central mantle	4	0	2	0	0	0

NOTE.—The gray shading highlights differential expression associated with shell damage and the number of SMPs indicates the number of contigs with strong sequence similarity to known SMPs.

to the same SMP or contained an SMP-associated domain, but which responded differently to shell repair were categorized as variable (fig. 3). These contigs were further assessed to determine whether overall expression levels or presence of other functional domains could explain the variability, but no pattern was apparent. Most of the shell proteins previously identified in undamaged mature individuals of the same three species (Arivalagan et al. 2017) were found in the current data. In the Arivalagan study most of the functional domains identified were species-specific (Arivalagan et al. 2017). However, analysis of transcriptome data in this study showed that many of these species-specific contigs were expressed, and differentially regulated during shell repair, in all three species in our current study. For example, the immunoglobulin domain was originally only found in the shell matrix proteome of *C. gigas* (Arivalagan et al. 2017), but the domain was found to be present in contigs involved in shell damage responses in all three studied species here (fig. 4).

The transcriptome data were also surveyed for contigs encoding transmembrane transporters, signaling domains and secretory proteins, which could potentially be involved in calcium transport to the shell production area. A range of transmembrane transporters were identified with different expression patterns depending on the species. However, bicarbonate transporters, sodium neurotransmitter symporters and inward rectifying potassium channels were up-regulated in all three species (fig. 5). Signaling domains were found in contigs implicated in the response to shell damage in the central mantle. Rhodopsin-like G protein-coupled receptors and serine-threonine kinases were found in all three species. Additionally, frizzled domains were found in *C. gigas* and *M. edulis*. Many contigs had no functional annotation associated with them (table 3). Some of these unannotated contigs may be previously undescribed secretory proteins, as they contained signal peptides, but no detectable transmembrane domains. Coils and disordered regions were also detected in

a small number of unannotated contigs (table 3). Sequence similarity searches based on six-frame translation (tblastx) of the unknown differentially expressed contigs showed no similar sequences between the three species (table 3).

Finally, to assess the potential contribution of hemocytes to the differential expression patterns identified in mantle tissue, contigs were compared with published hemocyte transcriptome data. For *M. edulis*, 2,194 hemocyte expressed sequence tags (ESTs) were retrieved from MytiBase (Venier et al. 2009), and for *C. gigas*, 10,293 hemocyte ESTs were retrieved from GigasDatabase (Fleury et al. 2009). No hemocyte-specific data were available for *P. maximus*. Several SMP-designated contigs from *C. gigas* and a single contig from *M. edulis* that were differentially expressed during shell repair in our study showed strong sequence similarity to hemocyte ESTs (fig. 3).

## Discussion

These data derived from shell damage-repair experiments with a matched pair design considerably expand our understanding of the basic biomineralization toolbox used in the production of bivalve shells, not just for SMPs, but now also includes proteins involved in calcium transport. The mantle tissue of bivalves is multifunctional and identifying genes specifically involved in biomineralization via expression profiling can be difficult to disentangle from general metabolism (e.g., Zhang et al. 2012). However, this study optimized the detection of biomineralization genes in a number of ways. Firstly, the expression profiles were obtained after experimental shell damage to specifically invoke repair calcification, with careful note taken of when crystal structures appeared in the damaged region. In all three species, an initial organic layer formed, occluding the hole prior to calcification. Calcification proceeded at very different rates between species, hence, sampling times differed between species (10 days for *C. gigas*, 29 days for *M. edulis*, and 70 days for *P. maximus*). In each species, holes were drilled above the central mantle region,

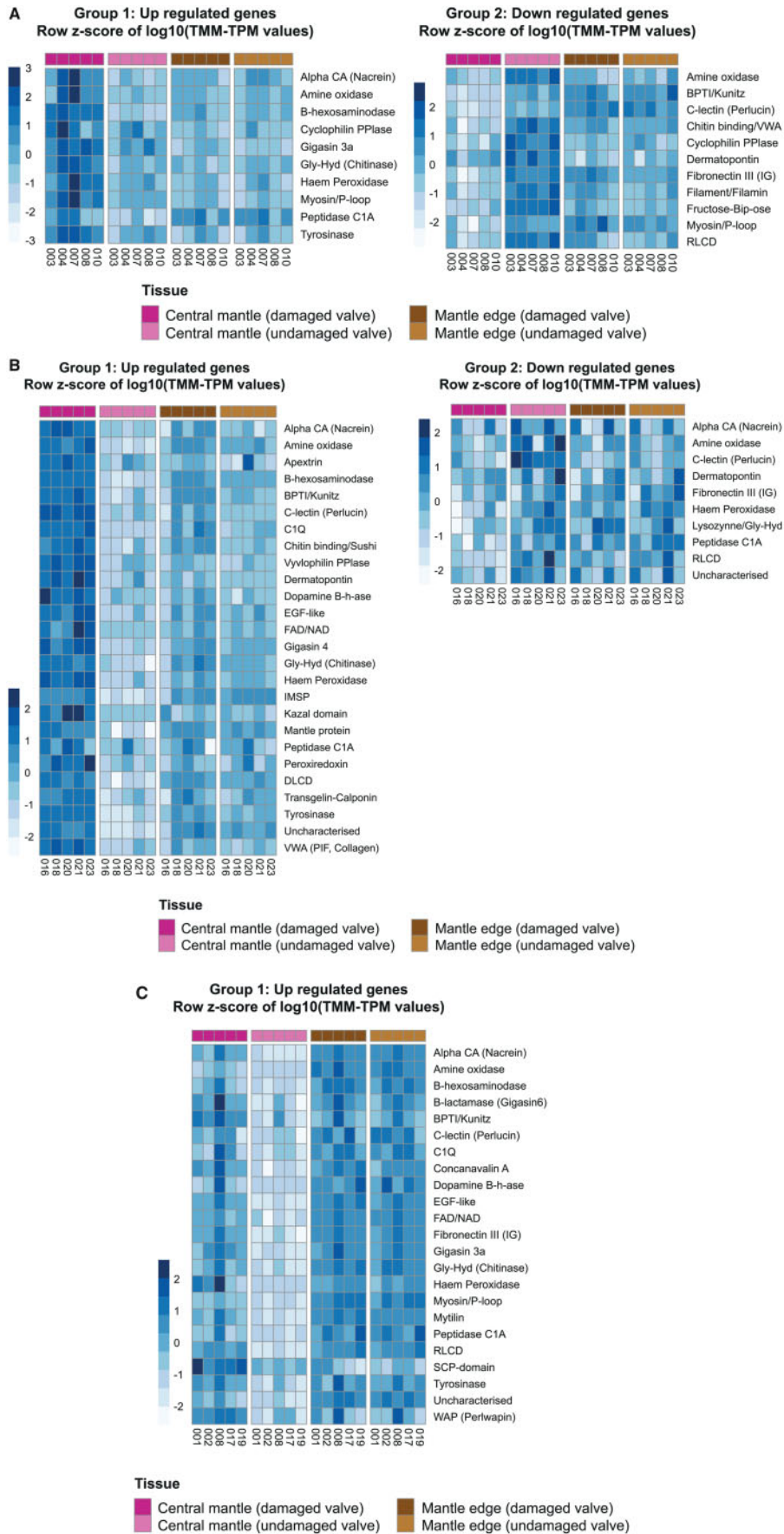
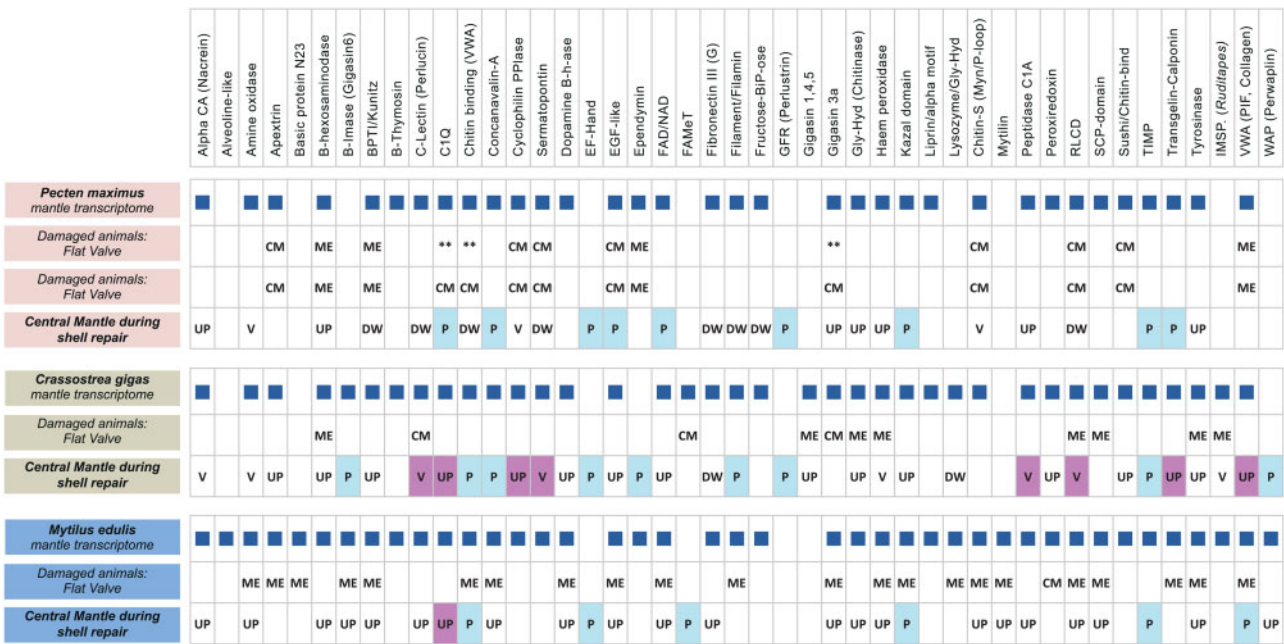
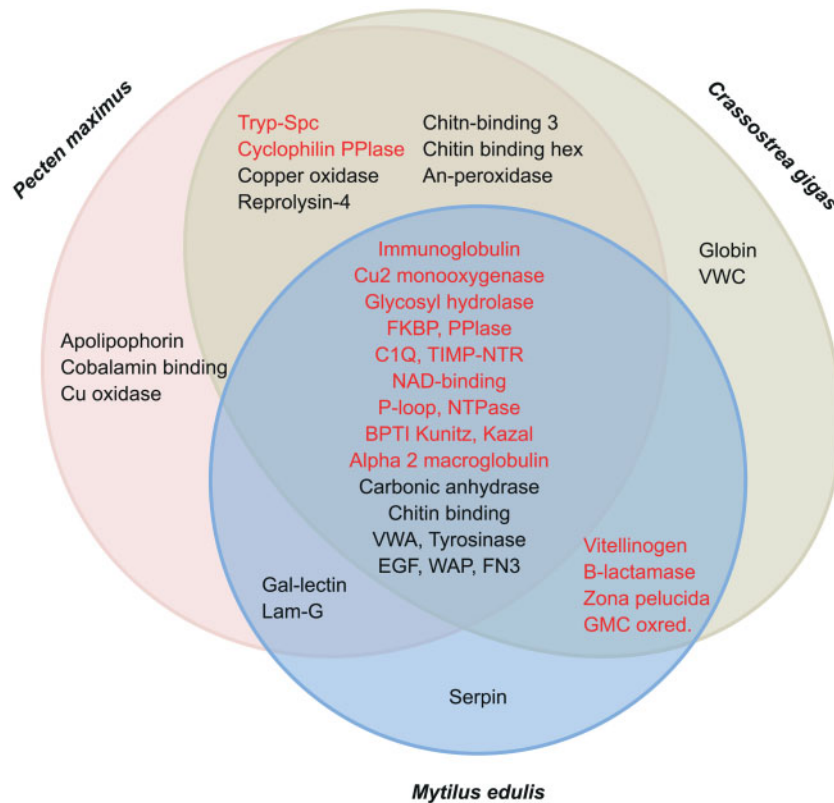


FIG. 2. Expression profiles of contigs with strong sequence similarity to SMPs. (a) *Pecten maximus*, (b) *Crassostrea gigas*, and (c) *Mytilus edulis*.

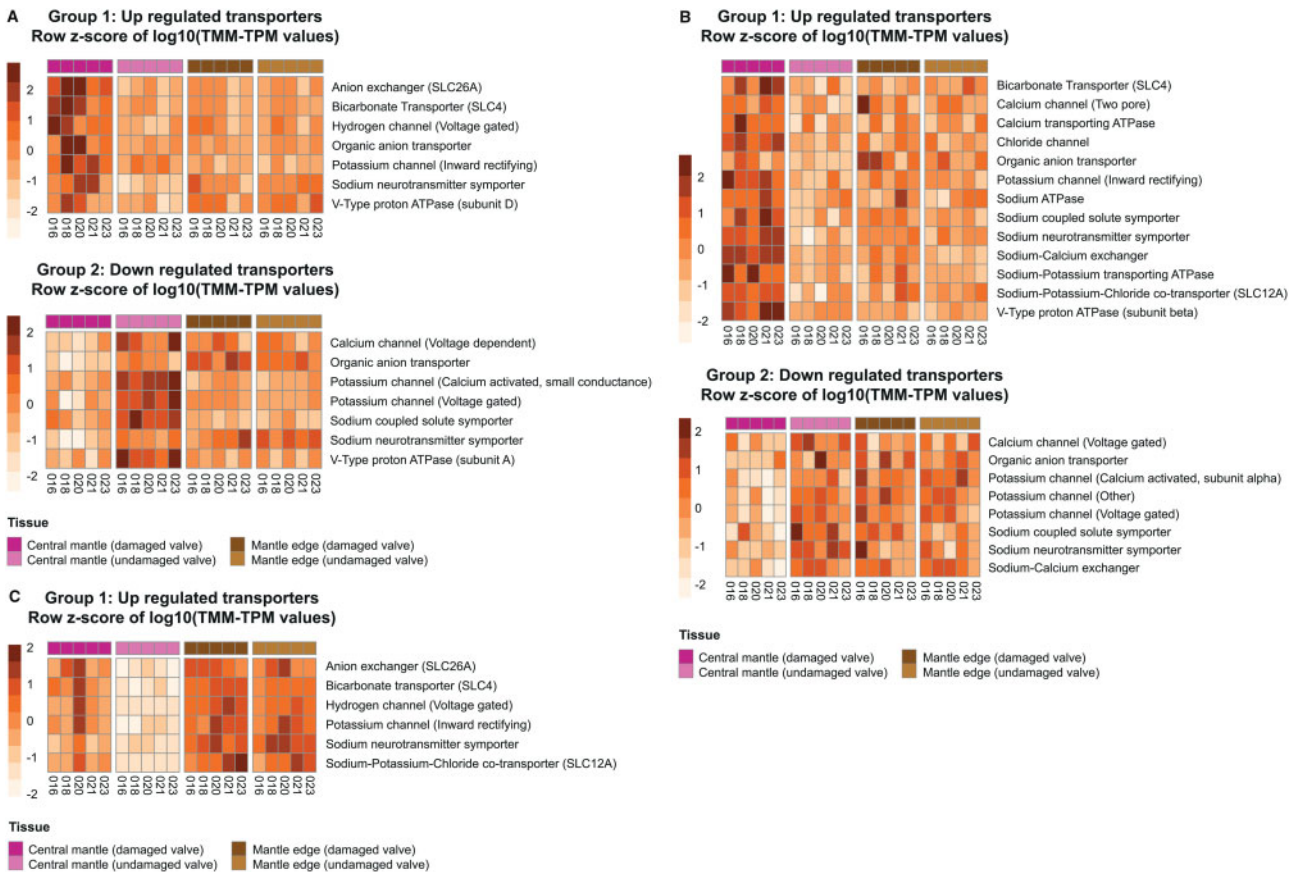


**FIG. 3.** Expression patterns of SMPs and domains found in the mantle transcriptomes. Key: CM/ME: SMPs highly expressed in the central mantle (CM) or mantle edge (ME); \*\*: FDR of at least 0.05; UP/DW/V: SMPs up-regulated (UP), down-regulated (DW), or variable (V) during repair in the central mantle of the damaged valve; P and blue shading: Putative shell proteins with no sequence similarity to, but with similar functional domains to known SMPs; Pink shading: differentially expressed contigs with sequence similarity to both SMPs and hemocyte ESTs.



**FIG. 4.** Domains identified in the proteins predicted from differentially expressed contigs following damage-repair in the central mantle. Domains in black were previously identified in shell proteomes (Arivalagan et al. 2017); those in red were common among more species than was detected in the shell proteome alone.





**Fig. 5.** Expression profiles of contigs with strong sequence similarity to transmembrane transporters. (a) *Pecten maximus*, (b) *Crassostrea gigas*, and (c) *Mytilus edulis*.

rather than at the mantle edge, as previous *Mytilus* experiments had shown that this approach induced a strong repair response (Hüning et al. 2016; Yarra et al. in review). Under normal conditions, the physiological functions of the mantle edge (i.e., growth) and the mantle center (i.e., maintenance) are very different. Damage to the central mantle results in this tissue switching from maintenance to active repair and reconstruction and cross comparison of damaged and undamaged central mantle tissue will thus emphasize repair and calcification genes.

Secondly, a matched-pair design was used to reduce noise due to biological variability, which resulted in much lower common dispersion values in all three species (table 1). The drawback of using a matched pair model is that experimental conditions affect the entire organism. However, the use of contralateral controls is common in other animal models, such as comparing a treated right limb to the control left limb in mice, or comparing cancerous tissue to normal tissue in humans (Tuch et al. 2010; Aguilar et al. 2015). Differential gene expression during central shell repair was localized to the area of damage in the central mantle and no differentially expressed contigs were identified at the mantle edge between damaged and control valves. In addition, non-drilled *P. maximus* animals showed little difference in expression profile between the edge and central mantle tissue (supplementary information S1, Supplementary Material online). Thus, the

undamaged valve in the drilled individuals was an appropriate control in this experiment, even when the valves are asymmetrical, as for *P. maximus* and *C. gigas*.

Even with this matched pair model, the variability of individuals affected detection of differentially expressed genes, as only genes differentially expressed in all individuals were designated as confirmed differentially expressed genes (fig. 3). Functional annotations did not simply correlate with expression patterns, as multiple distinct contigs that were functionally annotated as having similarity to the same shell proteins or functional domains were found to be variable in expression. Many SMPs are found as gene families with significant gene expansions in some families (e.g., tyrosinase, carbonic anhydrase, and chitinase-like) (Aguilera et al. 2014; Le Roy et al. 2014; Takeuchi et al. 2016). Furthermore SMP-associated domains are present in genes with a range of biological functions, which would invoke different expression profiles. For example, carbonic anhydrase is important in acid–base physiology as well as biomineralization (Le Roy et al. 2014).

Finally, the multifunctional aspect of the mantle transcriptome was further addressed by screening the current data for potential contamination from hemocytes. Molluscs have an open circulatory system, and hemocytes are thus present in all tissues. They have been proposed to play a role in shell formation (Mount et al. 2004; Kadar et al. 2009; Li et al. 2016; Ivanina et al. 2018). However, the evidence is equivocal, and



**Table 3.** Differentially Expressed Contigs during Shell Repair in the Central Mantle.

	<i>Pecten maximus</i>	<i>Crassostrea gigas</i>	<i>Mytilus edulis</i>
Number of DE contigs	1,701	4,125	653
Unknown DE contigs	844	424	328
Unknown DE contigs with:			
InterProscan domains	37	29	32
Putative secretory domains	10	21	39
Coils	23	27	16
Disordered regions	46	52	91

NOTE.—Investigation of unknown proteins for secretory domains, coiled and disordered regions.

involvement may be species-specific. For example, hemocytes were associated with immune processes in *C. gigas*, but with ion regulation and calcium transport in *Crassostrea virginica* (Ivanina et al. 2018). Screening of this current data set with ESTs from hemocyte-specific libraries (available for *M. edulis* and *C. gigas* only) confirmed the potential species-specific nature of this involvement, with only a single domain from the *M. edulis* repair transcriptome identified in the hemocyte ESTs, whereas eight repair-associated contigs were identified in the *C. gigas* hemocyte transcriptome. Overall, these three strategies applied to experimental design and analysis increased the confidence level of genes identified in this study as playing major roles in bivalve biomineralization.

As thousands of contigs were found to be differentially expressed in the mantle tissue sections, GO term enrichment was used to summarize the functional implications of differentially expressed contigs in each group. These analyses were generally uninformative, which is likely due to the human-centric nature of the GO annotations. The exception was *M. edulis*, where GO terms could be specifically related to known mollusc biomineralization gene functions. For example, the GO term “peptidase inhibitor activity” is associated with SMPs, such as perlwapin, BPTI/kunitz, alpha-2-macroglobulin, and kazal domain containing shell proteins. Similarly, “chitin binding” is associated with chitin-binding domain containing shell proteins, Pif, and BMSP.

Mollusc secretomes are known to contain a high proportion of novel lineage-restricted genes (Jackson et al. 2006; 2010; Kocot et al. 2016; Marin 2020). These genes often comprise unique combinations of co-opted ancient genes, which have evolved via the expansion and contraction of specific domains and domain shuffling (Kocot et al. 2016; Aguilera et al. 2017). Indeed, a recent survey of shell proteomes (“shellomes”) across the various classes of Mollusca and analyses of the first polyplacophoran genome indicate that some of these domains are present in the shellomes of highly divergent species (Marin 2020; Varney et al. 2021). Although some of these domains almost certainly represent key elements that regulate mineral disposition in the shell, their exact role within any one species is unclear and complicated by the mosaic and multiple convergent evolution of mollusc microstructures and their constituent protein families (Marie, Le Roy, et al. 2011; Vendrasco et al. 2011; Kocot et al. 2016; Marin 2020). Given this level of complexity, the main analytical approach of this current study was to curate the differentially expressed transcripts based on an in-house shell matrix database, many entries of which were domain-

specific. To expand the analysis to include calcium transport, transmembrane transporters and signaling proteins were also investigated in detail.

Shell proteome analyses have previously identified seven domains (carbonic anhydrase, chitin-binding, vWA, tyrosinase, EGF, WAP, and FN3) as part of a bivalve biomineralization toolbox (fig. 4). Most of these domains have largely well characterized functions in the protein matrix of the shell (Arivalagan et al. 2017) and are now joined by fifteen further protein domains. Some of these have previously been partially characterized for biomineralization functions. Beta-hexosaminidase, chitinase and glycoside hydrolase domains are subtypes of glycosyl hydrolases, which participate in the restructuring of the glycoside bonds in carbohydrates and potentially represent crystal framework domains. In situ hybridization and shell damage-repair experiments suggest that chitinase has an important biomineralization function (Li et al. 2017). In addition, the NAD-binding domain was implicated in shell formation in the brachiopod *Laqueus rubellus* (Isowa et al. 2015). This domain has a hydrogen carrier function that may be used to sequester H<sup>+</sup> from the crystallization milieu, thus controlling the ion concentration in the space where crystallization occurs (Isowa et al. 2015).

However, although most of these additional fifteen biomineralization domains have been previously found in shell proteomes and transcriptomes (e.g., the NTPase domain was the most common protein domain found in the date mussel *Lithophaga lithophaga* [Sivka et al. 2018]), how most of them interact and contribute to shell formation is unknown. Some domains, such as FKBP and PPLase (peptidyl-prolyl cis-trans isomerase) are known to be involved in protein folding, which may be important in the production of the extracellular matrix, but they may have other undiscovered functions. For example, the identification of protease and protease inhibitors containing kazal, BPTI/kunitz, alpha-2-macroglobulin and tissue inhibitor of metalloproteinase (TIMP) domains in mollusc shell proteomes suggested a previously unidentified biochemical defence mechanism in mollusc shells (Arivalagan et al. 2017). Experiments showed that TIMP in *C. gigas* hemocytes was associated with wound healing and defence (Montagnani et al. 2001). However, more recently TIMP has been associated with the fibrous organic matrix between aragonite crystals in the pearl oyster *Pinctada fucata* and experimental evidence suggested that PFTIMP played an important role of fibrous ligament structure (Kubota et al. 2017). Similarly, although C1q domains are traditionally associated with immune functioning in molluscs, this domain has

been subject to massive gene expansions within molluscs (Gerdol et al. 2015). Although current experimental data support the role of this domain in immune functions, it is entirely possible that some C1q domain proteins are involved in functions associated with biomineralization (Gerdol et al. 2015). This is supported by data from other species. The C1q domain in Otolin-1 (a collagen-like protein found in the inner ear of vertebrates and organic matrix of fish otoliths) is stabilized by  $\text{Ca}^{2+}$  ions and is thought to be important for the assembly of the otolith organic matrix in zebrafish (Hołubowicz et al. 2017). Furthermore, serine proteases enable mineralization in vertebrates and bacteria, while serine protease inhibitors inhibit mineralization (Tladen et al. 2012; Hershey et al. 2016). Also, metallopeptidases have been shown to promote enamel formation in humans (Prajapati et al. 2016). In molluscs, the SMP Perlwapin contains the protease inhibitor domain whey acidic protein (WAP), which inhibits growth of nacre crystals (Treccani et al. 2006).

However, there is still much to discover, as many mantle transcripts have little sequence similarity to known and functionally characterized genes (table 3). One set of proteins, where core domains are relatively conserved between species, with good annotation are the transmembrane transporters, which potentially play key roles in intercellular transport. To date, there is relatively little understanding of how calcium is transported in mollusc cells, particularly to the extrapallial space, the site of shell microstructure assembly, and how pH is controlled in this highly regulated microenvironment. Furthermore, there are considerable gaps in knowledge on the mollusc proteins, which effect transmembrane ion transfer and maintain pH homeostasis in this space. Ion transporters are highly likely to play critical roles in these processes.

In this study, 7, 13, and 6 transmembrane ion transporters were upregulated in the repairing mantle in *P. maximus*, *C. gigas*, and *M. edulis*, respectively. Three transporters were identified in all three bivalve species in the central mantle during shell repair; a bicarbonate transporter (SLC4 family), a voltage-gated hydrogen channel and a potassium channel (inward rectifying), with a further five up-regulated in two of the three species (fig. 5). Of the three “universal” ion transporters, one was previously identified in molluscs, the bicarbonate transporter (Zhao et al. 2018; Ramesh et al. 2019; 2020). Bicarbonate transporters, belonging to the solute carrier 4 (SLC4) family, are involved in the transport of bicarbonate ions across the membrane and have been shown to be important in enamel development in mice (Lacruz et al. 2010) and calcification in corals (Furla et al. 2000; Bhattacharya et al. 2016). The up-regulation of bicarbonate transporters in this study suggests that bivalves actively regulate the availability of bicarbonate ions to the site of calcification, rather than depend on bicarbonate ion availability from the surrounding environment. What is surprising is the lack of calcium transporters, such as calcium ATPase and sodium/calcium exchangers (NCX), as described in physiological studies investigating calcium transport across mantle epithelia using pharmacological inhibitors (Sillanpää et al. 2016; 2018). This may be due to the single time point sampling in this study, however, the other transporters identified almost certainly play at

least a partial role in pH homeostasis in membrane epithelia, in the cytosol and in the extrapallial space, an essential prerequisite for efficient production of mineralized microstructures. For example, a number of transporters identified in this study belong to solute carrier (SLC) families, which are passive transporters with general roles as gatekeepers of the cellular milieu (Pizzagalli et al. 2021).

Of the two other “universal” transporters identified in this study, it is likely that potassium inwardly rectifying channels play a key role in ion regulation. Potassium inwardly rectifying channels selectively mediate the uptake of potassium ions into the cell from the extracellular space, against an ion gradient (Miller 2000) and participate in bone formation in humans (Sacco et al. 2015). Furthermore, physiological studies have demonstrated higher potassium levels in the hemolymph of *M. edulis* and *C. gigas* than in the environment (Thomsen et al. 2010; Sillanpää et al. 2016). Thus, indicating that these species have the ability to regulate the extracellular levels of  $\text{K}^{+}$  to produce very specific microenvironments. The third “universal” transporter, a sodium neurotransmitter symporter is a member of the SLC6 family (Joseph et al. 2019; Pizzagalli et al. 2021). Although many of the SLC6 family members are associated with neuronal function in humans, of more relevance for mollusc shell production is the fact that certain members of the SLC6 family are actively upregulated during the maturation of dental enamel in humans (Lacruz et al. 2012).

The remaining five transporters (fig. 5) were only identified in two out of the three species (fig. 5). The  $\text{Na}^{+}\text{-K}^{+}\text{-Cl}^{-}$  cotransporter (NKCC) is a member of the SLC12 family. This transporter uses the sodium electrochemical gradient in the cell to passively drive the electrically silent influx of  $1\text{Na}^{+}:1\text{K}^{+}:2\text{Cl}^{-}$ , regulating cell volume and osmolarity. In humans, NKCC transporters mediate cell volume increases in chondrocytes, which underpin matrix synthesis and bone growth (Bush et al. 2010; Qusous et al. 2011). NKCC also has a volume regulatory role in sea urchins and is critical for the calcification of developing spicules (Basse et al. 2015). The anion exchanger is a member of the SLC 26 family. These transport a diverse range of substrates including  $\text{HCO}_3^{-}$  and therefore, in molluscs, they may perform a role similar to the previously identified SLC4 bicarbonate transporter (Alper and Sharma 2013). The final SLC family member identified in this study belonged to the organic anion transporter (OAT) SLC22 subfamily. These transporters regulate signaling molecules and key metabolites in tissues and extracellular fluids. They passively transport anionic substrates into the cell via the exchange of decarboxylates and play potential roles in perturbation of homeostasis or cell injury via remote sensing (Nigam et al. 2015). The final two transporters identified are involved in proton ( $\text{H}^{+}$ ) pumping. The V-type proton ATPase plays a critical role in establishing acidic pH in the lumens of intracellular compartments or organelles, including bone resorption lacuna in humans (Marshansky and Futai 2008; Casey et al. 2010) which hints that they may well play a role in shell remodeling in molluscs. Finally, the voltage gated hydrogen channel plays a major role in pH homeostasis, mediating  $\text{H}^{+}$  efflux, and has been shown to play a critical role

regulating pH in the calcifying cells of coccolithophores (Taylor et al. 2011). Hence, these eight ion transporters are strong candidates for future work using pharmacological inhibitors and in situ localization studies to characterize their role in molluscan biomineralization and expand the current bivalve biomineralization toolbox to include proteins involved in intracellular calcium transport.

## Conclusions

In this study, the use of a matched pair design, and induction of shell damage in the central mantle of the three bivalve species successfully detected differential gene expression during shell repair. Multiple SMPs and functional domains were regulated during repair, reflecting their importance in bivalve biomineralization processes. The previous bivalve biomineralization toolbox of four core protein domains (Arivalagan et al. 2017) has been expanded to 22 protein domains, with the further addition of several ion transporters with important roles in calcification, either via the maintenance of the cell acid–base balance or transport of critical ions to the extracellular space.

## Materials and Methods

### Experimental Design

Gene transcription profiles were analyzed from the mantle edge and the central mantle tissue sections from damaged and control valves of the same individual using a matched pair design as previously described in Hüning et al. (2016) and Yarra et al. (forthcoming) (fig. 1). Briefly, holes were drilled in the centers of the shells above the central mantle zone in cohorts of wild-sampled, live bivalves. Although *M. edulis* has symmetric valves, *P. maximus* and *C. gigas* have asymmetric valves. Shell asymmetry is related to the orientation and movement along the seabed in *P. maximus*, while no such relation is present in *C. gigas*, as they attach themselves to solid substrates and are immobile. Hence, the left valves were drilled in *M. edulis*, with the right valves acting as controls. In *P. maximus* and *C. gigas* holes were drilled in the convex valves with the opposite flatter valves acting as controls. Furthermore, to control for the potential asymmetry in *P. maximus*, gene expression in mantle tissues from nondrilled individuals was also assessed. There were no mortalities during the course of the experiment. All individuals were successful in initiating repair of the damaged valve. To capture gene expression at the onset of calcification, mantle tissue sections were sampled from individuals after 10 weeks of repair in *P. maximus*, 10 days in *C. gigas*, and 29 days in *M. edulis*. These time points coincided with the onset of calcification in the repair zones. Sequencing data for *M. edulis* were obtained in collaboration with Frank Melzner from the Helmholtz Centre for Ocean Research Kiel, Germany as described in Yarra et al. (forthcoming).

### RNA Extraction and Sequencing of *P. maximus* and *C. gigas*

Total RNA from the mantle tissue sections of *P. maximus* and *C. gigas* was extracted using the SV Total RNA Isolation

System kit (Promega) according to manufacturer's instructions. For the *P. maximus* damage repair study, four tissue sections were sampled from drilled individuals ( $n = 5$ ) (20 libraries total). An additional study to evaluate the effect of shell asymmetry was undertaken, sampling 4 tissues sections from nondrilled individuals ( $n = 3$ ) (12 libraries total). For *C. gigas*, four tissue sections were sampled from each of five drilled individuals (20 libraries in total). Isolated RNAs were assessed for concentration using a NanoDrop ND-100 Spectrometer (NanoDrop Technologies), and for quality using an Agilent 2200 TapeStation (Agilent Technologies). Strand specific libraries were prepared and the indexed libraries were pooled at equimolar concentrations and sequenced on two HiSeq2000 lanes (Illumina, USA) for *P. maximus* and one HiSeq2000 lane for *C. gigas* at Edinburgh Genomics.

### Bioinformatics Analysis

All bioinformatics analyses were carried out using default parameters unless otherwise specified. Adapter contamination was removed from the raw reads using Trimmomatic (Bolger et al. 2014) using the ILLUMINACLIP setting, with options as follows: Illumina adapter Sequences obtained from document #1000000002694 v.00 published by Illumina, seedMismatches = 2, palindromeClipThreshold = 30, simpleClipThreshold = 10, minAdapterLength = 8, keepBothReads = TRUE. Additionally, the PE setting was used to specify paired end reads. The reads were further trimmed for quality and length using Fasq-mcf (Aronesty 2011), using the settings: -q for minimum Phred score at each base, -l for minimum length after quality trimming, and -qual-min for minimum Phred score for the entire read (minimum Phred score 30, minimum length of 80 bp). After filtering for quality and length, only paired reads were retained. FastQC (Andrews 2010) was used to assess sequencing quality before and after adapter trimming and filtering for quality and length.

De novo transcriptome assemblies were constructed for all three species using Trinity v.2.2.0 (Grabherr et al. 2011). At the time of the analysis carried out in this study, only the *C. gigas* genome was available (Zhang et al. 2012). In a previous study, analysis of *C. gigas* transcriptomes was carried out using both de novo transcriptome and genome guided approaches (Yarra 2018). Although the oyster genome guided assemblies showed better N50 and median lengths, they performed poorly, with mapping rates ~20% lower than the mapping rate of the reads to the de novo transcriptome assemblies from Trinity (Yarra 2018). Hence, Trinity de novo assemblies were used for all three species in this study. The 32 libraries for *P. maximus*, and 20 libraries for *C. gigas* and *M. edulis* were normalized using the in silico normalization script (insilico\_read\_normalization.pl) with a coverage of 30 (-max cov). Trinity assemblies were constructed using the normalized reads with a minimum kmer coverage value of 2 (-min kmer cov), and the strand specific option (-SS lib type RF) for *P. maximus* and *C. gigas*. Using the Trinity pipeline, non-normalized reads were aligned to transcript assemblies using Bowtie (Langmead and Salzberg 2012) and abundance was estimated using RSEM (RNA-Seq by Expectation



Maximization) (Li and Dewey 2011). Both raw counts, and normalized counts using “Trimmed Mean of Maximum-values” (TMM) and “Transcripts Per Million mapped reads” (TPM), were generated. All differential expression analysis was performed using edgeR (Robinson et al. 2010). Raw counts were used as input for edgeR, as the software conducts its own sample normalization. Genes below 1 Counts Per Million (CPM) in at least half the libraries (16 for *P. maximus*, 10 for *C. gigas*, and *M. edulis*) were removed prior to analysis. Differential gene expression between the different mantle tissues was assessed using the paired experimental model (Individual + Tissue) and only results with a false discovery rate (FDR) of at least 0.001 were considered.

Sequence similarity searches of the transcript sequences were performed using BLAST (Altschul et al. 1997) (BlastX) with an *E*-value cutoff of  $1e^{-10}$  against SwissProt, TrEMBL, and an in-house Shell Matrix Protein database (<https://doi.org/10/cz2w> [Yarra 2018]). The SMP database contains 327 SMPs from molluscan genera, which were downloaded from Uniprot (<http://www.uniprot.org/>) using keywords related to molluscan biomineralization (molluscs, shell, bivalve, aragonite, calcite, prismatic, foliated, mantle, mantle edge, central mantle, pallial mantle). The SMP data set was manually curated by reviewing the publication related to each protein entry, and only entries that were validated to be present in molluscan shell matrices were retained. Domains found in the proteins in the SMP database were annotated Interproscan v.5.25-64.0 (Jones et al. 2014) with SMP database entries grouped by functional domain, to reconcile differing naming conventions in previous studies. Sequence similarity results were filtered to exclude matches that covered <40% of the database sequence entry. Transcripts were translated into putative protein sequences of at least 20 codons using Transdecoder (part of the Trinity pipeline) and mined for domain and family information using Interproscan (Jones et al. 2014). Gene Ontology (GO) terms for contigs were assigned based on the Interpro database (Finn et al. 2017). Enrichment of GO terms was performed using downstream Trinity pipelines for Trinotate and GSeq, with a FDR value of 0.05. TMM normalized TPM values were used to generate heatmaps for contigs of interest.

ESTs from hemocytes in *Mytilus galloprovincialis* were obtained from Mytibase (Venier et al. 2009). ESTs from *C. gigas*, uploaded as part of Gigas-Database (Fleury et al. 2009), were obtained from NCBI with the search term “hemocyte (and haemocyte) AND crassostrea,” and sequences with metadata containing “hemocyte (and haemocyte) subtracted,” and “unsure sequence” were removed before further analysis. Sequence similarity searches of the transcripts (tblastx) were performed against the hemocyte EST database and the in-house SMP database using an *E*-value cutoff of  $1e^{-10}$ .

## Supplementary Material

Supplementary data are available at *Molecular Biology and Evolution* online.

## Acknowledgments

The authors would like to thank Gordon Goldsworthy from LochFyne Farms (Tarbert, UK) for providing the *Pecten maximus* and *Crassostrea gigas*; Kim Last and Christine Beveridge from The Scottish Association of Marine Sciences (Oban, UK) who assisted with experimental set-up and animal husbandry of *P. maximus*, and Rebecca Smith from the British Antarctic Survey (Cambridge, UK) for assistance with experimental set-up and animal husbandry of *C. gigas*; Edinburgh genomics for sequencing. We also thank Kirti Ramesh, Anne Hüning, and Frank Melzner for their collaboration on the *Mytilus* experiment. Carlos Caurcel from the University of Edinburgh, UK for assistance with Interproscan domain detection and Jamie Oliver (BAS) for producing with high definition figures. This study was supported by the European Union Seventh Framework Programme [FP7] ITN project “CACHE: Calcium in a Changing Environment” under REA grant agreement (605051).

## Data Availability

The sequence data set supporting the conclusions of this article is available in SRA under accession numbers PRJNA509489 (*Pecten maximus*), PRJNA510211 (*Crassostrea gigas*), and SRP108359 (*Mytilus edulis*). The SMP database is publicly available from the NERC Polar Data Centre repository: GB/NERC/BAS/PDC/01132 with <https://doi.org/10/cz2w>.

## References

- Aguilar CA, Shcherbina A, Rieke DO, Pop R, Carrigan CT, Gifford CA, Urso ML, Kottke MA, Meissner A. 2015. *In vivo* monitoring of transcriptional dynamics after lower-limb muscle injury enables quantitative classification of healing. *Sci Rep*. 5:13885.
- Aguilera F, McDougall C, Degnan BM. 2014. Evolution of the tyrosinase gene family in bivalve molluscs: independent expansion of the mantle gene repertoire. *Acta Biomater*. 10(9):3855–3865.
- Aguilera F, McDougall C, Degnan BM. 2017. Co-option and *de novo* gene evolution underlie molluscan shell diversity. *Mol Biol Evol*. 34(4):779–792.
- Alper SL, Sharma AK. 2013. The SLC26 gene family of anion transporters and channels. *Mol Aspects Med*. 34(2–3):494–515.
- Altschul SF, Madden TL, Schaffer AA, Zhang JH, Zhang Z, Miller W, Lipman DJ. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res*. 25(17):3389–3402.
- Andrews S. 2010. FastQC: a quality control tool for high throughput sequence data. <http://www.bioinformatics.babraham.ac.uk/projects/fastqc>
- Arivalagan J, Marie B, Sleight VA, Clark MS, Berland S, Marie A. 2016. Shell matrix proteins of the clam, *Mya truncata*: roles beyond shell formation through proteomic study. *Mar Genomics*. 27:69–74.
- Arivalagan J, Yarra T, Marie B, Sleight VA, Duvernois-Berthet E, Clark MS, Marie A, Berland S. 2017. Insights from the shell proteome: biomineralization to adaptation. *Mol Biol Evol*. 34(1):66–77.
- Aronesty E. 2011. ea-utils: command-line tools for processing biological sequencing data. <https://github.com/ExpressionAnalysis/ea-utils>.
- Basse WC, Gutowska MA, Findeisen U, Stumpp M, Dupont S, Jackson DJ, Himmerkus N, Melzner F, Bleich M. 2015. A sea urchin  $\text{Na}^{(+)}\text{K}^{(+)}2\text{Cl}^{(-)}$  cotransporter is involved in the maintenance of calcification-relevant cytoplasmic cords in *Strongylocentrotus droebachiensis* larvae. *Comp Biochem Physiol A Mol Integr Physiol*. 187:184–192.
- Bhattacharya D, Agrawal S, Aranda M, Baumgarten S, Belcaid M, Drake JL, Erwin D, Foret S, Gates RD, Gruber DF, et al. 2016. Comparative



- genomics explains the evolutionary success of reef-forming corals. *Elife* 5:13288.
- Bolger AM, Lohse M, Usadel B. 2014. Trimmomatic: a flexible trimmer for Illumina sequence data. *Bioinformatics* 30(15):2114–2120.
- Bush PG, Pritchard M, Loqman MY, Damron TA, Hall AC. 2010. A key role for membrane transporter NKCC1 in mediating chondrocyte volume increase in the mammalian growth plate. *J Bone Miner Res* 25(7):1594–1603.
- Casey JR, Grinstein S, Orlowski J. 2010. Sensors and regulators of intracellular pH. *Nat Rev Mol Cell Biol* 11:1150–61.
- Clark MS. 2020. Molecular mechanisms of biomineralization in marine invertebrates. *J Exp Biol* 223(Pt 11):jeb206961. doi:10.1242/jeb.206961.
- Clark MS, Peck LS, Arivalagan J, Bäckeljau T, Berland S, Cardoso JCR, Caurcel C, Chapelle G, De Noia M, Dupont S, et al. 2020. Deciphering mollusc shell production: the roles of genetic mechanisms through to ecology, aquaculture and biomimetics. *Biol Rev Camb Philos Soc* 95(6):1812–1837.
- Fang D, Xu GR, Hu YL, Pan C, Xie LP, Zhang RQ. 2011. Identification of genes directly involved in shell formation and their functions in pearl oyster, *Pinctada fucata*. *PLoS One* 6(7):e21860.
- Finn RD, Attwood TK, Babbitt PC, Bateman A, Bork P, Bridge AJ, Chang HY, Dosztanyi Z, El-Gebali S, Fraser M, et al. 2017. InterPro in 2017—beyond protein family and domain annotations. *Nucleic Acids Res* 45(D1):D190–D199.
- Fleury E, Huvet A, Lelong C, de Lorgeril J, Boulo V, Gueguen Y, Bachere E, Tanguy A, Moraga D, Fabioux C, et al. 2009. Generation and analysis of a 29,745 unique Expressed Sequence Tags from the Pacific oyster (*Crassostrea gigas*) assembled into a publicly accessible database: the GigasDatabase. *BMC Genomics* 10:341.
- Furla P, Galgani I, Durand I, Allemand D. 2000. Sources and mechanisms of inorganic carbon transport for coral calcification and photosynthesis. *J Exp Biol* 203(22):3445–3457.
- Gerdol M, Venier P, Pallavicini A. 2015. The genome of the Pacific oyster *Crassostrea gigas* brings new insights on the massive expansion of the C1 q gene family in *Bivalvia*. *Dev Comp Immunol* 49(1):59–71.
- Grabherr MG, Haas BJ, Yassour M, Levin JZ, Thompson DA, Amit I, Adiconis X, Fan L, Raychowdhury R, Zeng QD, et al. 2011. Full-length transcriptome assembly from RNA-Seq data without a reference genome. *Nat Biotechnol* 29(7):644–U130.
- Grefsrud ES, Dauphin Y, Cuif J-P, Denis A, Strand O. 2008. Modifications in microstructure of cultured and wild scallop shells (*Pecten maximus*). *J Shellfish Res* 27(4):633–641.
- Herlitze I, Marie B, Marin F, Jackson DJ. 2018. Molecular modularity and asymmetry of the molluscan mantle revealed by a gene expression atlas. *Gigascience* 7:giy056.
- Hershey DM, Browne PJ, Iavarone AT, Teyra J, Lee EH, Sidhu SS, Komeili A. 2016. Magnetite biomineralization in *magnetospirillum magneticum* is regulated by a switch-like behavior in the HtrA protease MamE. *J Biol Chem* 291(34):17941–17952.
- Hołubowicz R, Wojtas M, Taube M, Kozak M, Ozyhar A, Dobryszczycki P. 2017. Effect of calcium ions on structure and stability of the C1q-like domain of otolin-1 from human and zebrafish. *FEBS J* 284(24):4278–4297.
- Hüning AK, Lange SM, Ramesh K, Jacob DE, Jackson DJ, Panknin U, Gutowska MA, Philipp EER, Rosenstiel P, Lucassen M, et al. 2016. A shell regeneration assay to identify biomineralization candidate genes in Mytilid mussels. *Mar Genomics* 27:57–67.
- Isowa Y, Sarashina I, Oshima K, Kito K, Hattori M, Endo K. 2015. Proteome analysis of shell matrix proteins in the brachiopod *Laqueus rubellus*. *Proteome Sci* 13:21.
- Ivanina AV, Borah BM, Vogts A, Malik I, Wu JY, Chin AR, Almarza AJ, Kumta P, Piontkivska H, Beniash E, et al. 2018. Potential trade-offs between biomineralization and immunity revealed by shell properties and gene expression profiles of two closely related *Crassostrea* species. *J Exp Biol* 221. doi:10.1242/jeb.183236.
- Jackson DJ, McDougall C, Green K, Simpson F, Wörheide G, Degnan BM. 2006. A rapidly evolving secretome builds and patterns a sea shell. *BMC Biol* 4:40.
- Jackson DJ, McDougall C, Woodcroft B, Moase P, Rose RA, Kube M, Reinhardt R, Rokhsar DS, Montagnani C, Joubert C, et al. 2010. Parallel evolution of nacre building gene sets in Molluscs. *Mol Biol Evol* 27(3):591–608.
- Jones P, Binns D, Chang HY, Fraser M, Li WZ, McAnulla C, McWilliam H, Maslen J, Mitchell A, Nuka G, et al. 2014. InterProScan 5: genome-scale protein function classification. *Bioinformatics* 30(9):1236–1240.
- Joseph D, Pidathala S, Mallela AK, Penmatsa A. 2019. Structure and gating dynamics of Na<sup>+</sup>/Cl<sup>-</sup> coupled neurotransmitter transporters. *Front Mol Biosci* 6. https://doi.org/10.3389/fmolb.2019.00080.
- Kadar E, Lobo-da-Cunha A, Azevedo C. 2009. Mantle-to-shell CaCO<sub>3</sub> transfer during shell repair at different hydrostatic pressures in the deep-sea vent mussel *Bathymodiolus azoricus* (Bivalvia: Yttilidae). *Mar Biol* 156(5):959–967.
- Kocot KM, Aguilera F, McDougall C, Jackson DJ, Degnan BM. 2016. Sea shell diversity and rapidly evolving secretomes: insights into the evolution of biomineralization. *Front Zool* 13:23.
- Kong Y, Jing G, Yan Z, Li C, Gong N, Zhu F, Li D, Zhang Y, Zheng G, Wang H, et al. 2009. Cloning and characterization of Prsilkin-39, a novel matrix protein serving a dual role in the prismatic layer formation from the oyster *Pinctada fucata*. *J Biol Chem* 284(16):10841–10854.
- Kubota K, Tsuchihashi Y, Kogure T, Maeyama K, Hattori F, Kinoshit S, Sakuda S, Nagasawa H, Yoshimura E, Suzuki M. 2017. Structural and functional analyses of a TIMP and MMP in the ligament of *Pinctada fucata*. *J Struct Biol* 199(3):216–224.
- Lacruz RS, Nanci A, White SN, Wen X, Wang H, Zalzal SF, Luong VQ, Schuetter VL, Conti PS, Kurtz I, et al. 2010. The sodium bicarbonate cotransporter (NBCe1) is essential for normal development of mouse dentition. *J Biol Chem* 285(32):24432–24438.
- Lacruz RS, Smith CE, Bringas P, Chen YB, Smith SM, Snead ML, Kurtz I, Hacia JG, Hubbard MJ, Paine ML. 2012. Identification of novel candidate genes involved in mineralization of dental enamel by genome-wide transcript profiling. *J Cell Physiol* 227(5):2264–2275.
- Langmead B, Salzberg SL. 2012. Fast gapped-read alignment with Bowtie 2. *Nat Methods* 9(4):357–U354.
- Le Roy N, Jackson DJ, Marie B, Ramos-Silva P, Marin F. 2014. The evolution of metazoan alpha-carbonic anhydrases and their roles in calcium carbonate biomineralization. *Front Zool* 11(1):75.
- Li B, Dewey CN. 2011. RSEM: accurate transcript quantification from RNA-Seq data with or without a reference genome. *BMC Bioinformatics* 12:323.
- Li H, Wang D, Deng Z, Huang G, Fan S, Zhou D, Liu B, Zhang B, Yu D. 2017. Molecular characterization and expression analysis of chitinase from the pearl oyster *Pinctada fucata*. *Compar Biochem Physiol B-Biochem Mol Biol* 203:141–148.
- Li SG, Liu YJ, Liu C, Huang JL, Zheng GL, Xie LP, Zhang RQ. 2016. Hemocytes participate in calcium carbonate crystal formation, transportation and shell regeneration in the pearl oyster *Pinctada fucata*. *Fish Shellfish Immunol* 51:263–270.
- Marie B, Le Roy N, Zanella-Cleon I, Becchi M, Marin F. 2011. Molecular evolution of mollusc shell proteins: insights from proteomic analysis of the edible mussel *Mytilus*. *J Mol Evol* 72(5–6):531–546.
- Marie B, Zanella-Cleon I, Guichard N, Becchi M, Marin F. 2011. Novel Proteins from the Calcifying Shell Matrix of the Pacific Oyster *Crassostrea gigas*. *Mar Biotechnol (NY)* 13(6):1159–1168.
- Marin F. 2020. Mollusc shellomes: past, present and future. *J Struct Biol* 212(1):107583.
- Marin F, Amons R, Guichard N, Stigter M, Hecker A, Luquet G, Layrolle P, Alcaraz G, Riondet C, Westbroek P. 2005. Caspartin and calprismin, two proteins of the shell calcitic prisms of the Mediterranean fan mussel *Pinna nobilis*. *J Biol Chem* 280(40):33895–33908.
- Marshansky V, Futai M. 2008. The V-type H<sup>+</sup>-ATPase in vesicular trafficking: targeting, regulation and function. *Curr Opin Cell Biol* 20(4):415–426.
- McCarthy DJ, Chen YS, Smyth GK. 2012. Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation. *Nucleic Acids Res* 40(10):4288–4297.
- Miller C. 2000. An overview of the potassium channel family. *Genome Biol* 1(4):REVIEWS0004.

- Montagnani C, Le Roux F, Berthe F, Escoubas JM. 2001. Cg-TIMP, an inducible tissue inhibitor of metalloproteinase from the Pacific oyster *Crassostrea gigas* with a potential role in wound healing and defense mechanisms. *FEBS Lett.* 500(1–2):64–70.
- Mount AS, Wheeler AP, Paradkar RP, Snider D. 2004. Hemocyte-mediated shell mineralization in the eastern oyster. *Science* 304(5668):297–300.
- Nigam SK, Bush KT, Martovetsky G, Ahn SY, Liu HC, Richard E, Bhatnagar V, Wu W. 2015. The organic anion transporter (OAT) family: a systems biology perspective. *Physiol Rev.* 95(1):83–123.
- Pizzagalli MD, Bensimon A, Superti-Furga G. 2021. A guide to plasma membrane solute carrier proteins. *FEBS J* 288(9):2784–2835.
- Plazzi F, Passamonti M. 2010. Towards a molecular phylogeny of Mollusks: bivalves' early evolution as revealed by mitochondrial genes. *Mol Phylogenet Evol.* 57(2):641–657.
- Prajapati S, Tao JH, Ruan QC, De Yoreo JJ, Moradian-Oldak J. 2016. Matrix metalloproteinase-20 mediates dental enamel biomineralization by preventing protein occlusion inside apatite crystals. *Biomaterials* 75:260–270.
- Qusous A, Geewan CSV, Greenwell P, Kerrigan MJ. 2011. siRNA-mediated inhibition of  $\text{Na}^+ - \text{K}^+ - 2\text{Cl}^{(-)}$  cotransporter (NKCC1) and regulatory volume increase in the chondrocyte cell line C-20/A4. *J Membr Biol.* 243(1–3):25–34.
- Ramesh K, Hu MY, Melzner F, Bleich M, Himmerkus N. 2020. Intracellular pH regulation in mantle epithelial cells of the Pacific oyster, *Crassostrea gigas*. *J Comp Physiol B.* 190(6):691–700.
- Ramesh K, Yarra T, Clark MS, John U, Melzner F. 2019. Expression of calcification-related ion transporters during blue mussel larval development. *Ecol Evol.* 9(12):7157–7172.
- Ren JF, Liu XA, Jiang F, Guo XM, Liu B. 2010. Unusual conservation of mitochondrial gene order in *Crassostrea* oysters: evidence for recent speciation in Asia. *BMC Evol Biol.* 10:394.
- Robinson MD, McCarthy DJ, Smyth GK. 2010. edgeR: a Bioconductor package for differential expression analysis of digital gene expression data. *Bioinformatics* 26(1):139–140.
- Rosenberg G. 2014. A new critical estimate of named species-level diversity of the recent Mollusca. *Am Malacol Bull.* 32(2):308–322.
- Sacco S, Giuliano S, Sacconi S, Desnuelle C, Barhanin J, Amri E-Z, Bendahhou S. 2015. The inward rectifier potassium channel Kir2.1 is required for osteoblastogenesis. *Hum Mol Genet.* 24(2):471–479.
- Sillanpää JK, Ramesh K, Melzner F, Sundh H, Sundell K. 2016. Calcium mobilisation following shell damage in the Pacific oyster, *Crassostrea gigas*. *Mar Genomics.* 27:75–83.
- Sillanpää JK, Sundh H, Sundell KS. 2018. Calcium transfer across the outer mantle epithelium in the Pacific oyster, *Crassostrea gigas*. *Proc R Soc B-Biol Sci.* 285. <https://doi.org/10.1098/rspb.2018.1676>.
- Sivka U, Toplak N, Koren S, Jakse J. 2018. *De novo* transcriptome of the pallial gland of the date mussel (*Lithophaga lithophaga*). *Comp Biochem Physiol D Genomics Proteomics.* 26:1–9.
- Sleight VA, Marie B, Jackson DJ, Dyrinda EA, Marie A, Clark MS. 2016. An Antarctic molluscan biomineralisation tool-kit. *Sci Rep.* 6:36978.
- 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(5946):1388–1390.
- Takeuchi T, Koyanagi R, Gyoja F, Kanda M, Hisata K, Fujie M, Goto H, Yamasaki S, Nagai K, Morino Y, et al. 2016. Bivalve-specific gene expansion in the pearl oyster genome: implications of adaptation to a sessile lifestyle. *Zool Lett.* 2:3.
- Taylor AR, Chrachri A, Wheeler G, Goddard H, Brownlee C. 2011. A voltage-gated  $\text{H}^+$  channel underlying pH homeostasis in calcifying coccolithophores. *PLoS Biol.* 9(6):e1001085.
- Thomsen J, Gutowska MA, Saphorster J, Heinemann A, Trubenbach K, Fietzke J, Hiebenthal C, Eisenhauer A, Kortzinger A, Wahl M, et al. 2010. Calcifying invertebrates succeed in a naturally  $\text{CO}_2$ -rich coastal habitat but are threatened by high levels of future acidification. *Bioessences* 7(11):3879–3891.
- Tiaden AN, Breiden M, Mirsaidi A, Weber FA, Bahrenberg G, Glanz S, Cinelli P, Ehrmann M, Richards PJ. 2012. Human serine protease HTRA1 positively regulates osteogenesis of human bone marrow-derived mesenchymal stem cells and mineralization of differentiating bone-forming cells through the modulation of extracellular matrix protein. *Stem Cells* 30(10):2271–2282.
- Treccani L, Mann K, Heinemann F, Fritz M. 2006. Perlwapin, an abalone nacre protein with three four-disulfide core (whey acidic protein) domains, inhibits the growth of calcium carbonate crystals. *Biophys J.* 91(7):2601–2608.
- Tuch BB, Laborde RR, Xu X, Gu J, Chung CB, Monighetti CK, Stanley SJ, Olsen KD, Kasperbauer JL, Moore EJ, et al. 2010. Tumor transcriptome sequencing reveals allelic expression imbalances associated with copy number alterations. *PLoS One* 5(2):e9317.
- Varney RM, Speiser DI, McDougall C, Degnan BM, Kocot KM. 2021. The iron-responsive genome of the chiton *Acanthopleura granulata*. *Genome Biol Evol.* 13(1). doi:10.1093/gbe/evaa263.
- Vendrasco MJ, Checa AG, Kouchinsky AV. 2011. Shell microstructure of the early bivalve *pojetaia* and the independent origin of nacre within the Mollusca. *Palaeontology* 54(4):825–850.
- Venier P, De Pitta C, Bernante F, Varotto L, De Nardi B, Bovo G, Roch P, Novoa B, Figueras A, Pallavicini A, et al. 2009. MytiBase: a knowledgebase of mussel (*M. galloprovincialis*) transcribed sequences. *BMC Genomics* 10:72.
- Williams ST. 2017. Molluscan shell colour. *Biol Rev Camb Philos Soc.* 92(2):1039–1058.
- Yarra T. 2018. Transcriptional profiling of shell calcification in bivalves [PhD thesis]. UK: University of Edinburgh.
- Yarra T, Ramesh K, Blaxter M, Hüning A, Melzner F, Clark MS. Forthcoming. Transcriptomic analysis of shell repair and biomineralization in the blue mussel, *Mytilus edulis*. *BMC Genomics*.
- Zhang GF, Fang XD, Guo XM, Li L, Luo RB, Xu F, Yang PC, Zhang LL, Wang XT, Qi HG, et al. 2012. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature* 490(7418):49–54.
- Zhang C, Li S, Ma Z, Xie L, Zhang R. 2006. A novel matrix protein p10 from the nacre of pearl oyster (*Pinctada fucata*) and its effects on both  $\text{CaCO}_3$  crystal formation and mineralogenic cells. *Mar Biotechnol (NY).* 8(6):624–633.
- Zhao R, Takeuchi T, Luo Y-J, Ishikawa A, Kobayashi T, Koyanagi R, Villar-Briones A, Yamada L, Sawada H, Iwanaga S, et al. 2018. Dual gene repertoires for larval and adult shells reveal molecules essential for molluscan shell formation. *Mol Biol Evol.* 35(11):2751–2761.