

RESEARCH ARTICLE

Open Access



# Gene expression correlated with delay in shell formation in larval Pacific oysters (*Crassostrea gigas*) exposed to experimental ocean acidification provides insights into shell formation mechanisms

Pierre De Wit<sup>1\*</sup> , Evan Durland<sup>2</sup>, Alexander Ventura<sup>3</sup> and Chris J. Langdon<sup>2</sup>

## Abstract

**Background:** Despite recent work to characterize gene expression changes associated with larval development in oysters, the mechanism by which the larval shell is first formed is still largely unknown. In *Crassostrea gigas*, this shell forms within the first 24 h post fertilization, and it has been demonstrated that changes in water chemistry can cause delays in shell formation, shell deformations and higher mortality rates. In this study, we use the delay in shell formation associated with exposure to CO<sub>2</sub>-acidified seawater to identify genes correlated with initial shell deposition.

**Results:** By fitting linear models to gene expression data in ambient and low aragonite saturation treatments, we are able to isolate 37 annotated genes correlated with initial larval shell formation, which can be categorized into 1) ion transporters, 2) shell matrix proteins and 3) protease inhibitors. Clustering of the gene expression data into co-expression networks further supports the result of the linear models, and also implies an important role of dynein motor proteins as transporters of cellular components during the initial shell formation process.

**Conclusions:** Using an RNA-Seq approach with high temporal resolution allows us to identify a conceptual model for how oyster larval calcification is initiated. This work provides a foundation for further studies on how genetic variation in these identified genes could affect fitness of oyster populations subjected to future environmental changes, such as ocean acidification.

**Keywords:** *Crassostrea gigas*, Gene expression, Larvae, Ocean acidification, Aragonite, Calcification

## Background

Calcium carbonate (CaCO<sub>3</sub>) particles have been reported to be initially formed in the intracellular compartments of specialized cells that transport the mineral to sites of shell formation in adult molluscs [1] and possibly in larvae. The composition and structure of a protein matrix that is also thought to be deposited by specialized cells [1], determines how CaCO<sub>3</sub> crystals become organized, forming different isoforms with distinct chemical and physical properties [2]. The most common of these isoforms in

marine molluscs are calcite and aragonite. Shells of adult oysters are primarily composed of calcite [3], but the larval oyster shell is composed of aragonite [4].

Aragonite of Pacific oyster larval shells starts to form after 14–18 h post-fertilization under standard culture conditions [5]. Embryos initially exist as unprotected trochophore larvae whereupon a shell gland forms and produces an organic pellicle or periostracum [6], allowing shell deposition to begin between the periostracum and the larval epithelium of the shell field [2]. These processes produce a shell that makes up to about 90% of larval dry body weight [7]. Prodissoconch I stage larvae are also called “D-larvae”, as the larval shell has a distinct D-shaped form. After the initial shell has formed,

\* Correspondence: [pierre.de\\_wit@marine.gu.se](mailto:pierre.de_wit@marine.gu.se)

<sup>1</sup>Department of Marine Sciences, University of Gothenburg, Strömstad, Sweden

Full list of author information is available at the end of the article



the larvae become planktotrophic and feed on microalgae for 2–3 weeks before developing into pediveliger larvae. When competent, pediveliger larvae settle on hard substrates and metamorphose into sessile juvenile oysters. At this life stage, shell mineral composition changes from aragonite to the calcite isoform of calcium carbonate [3, 8].

During these major life stage transitions, large-scale changes occur in the cellular biochemistry of the calcifying tissue (e.g. [9–12]). It is likely that these physiological and biochemical changes are linked to transcriptomic transitions, where for example transcripts coding for shell matrix proteins, ion pumps and other processes are up-regulated for shell formation, while others could be simultaneously down-regulated. Li et al. [13] showed that differential gene expression in the larvae of the pearl oyster, *Pinctada fucata*, occurs mostly during transitions between early developmental stages, such as from trochophore to D-shaped larva. Zhang et al. [10] reported that a fibronectin-like transcript and chitin synthase were highly expressed at the initiation of shell formation in larval *C. gigas*. While the bone-morphogenetic-protein (BMP) [14] signalling pathway has been hypothesized as an initiator of these changes [15–17], a number of studies have also been conducted on individual candidate genes putatively directly involved in biomineralization in oysters; for example tyrosinases have been suggested as having a function in periostracum formation and biogenesis [18, 19] and cyclases have been suggested to control intracellular calcium and bicarbonate ion concentrations [20, 21].

As several recent studies have shown a disconnect between gene expression and protein expression [22, 23], quantification of protein content provides an important link between gene expression changes and cellular physiology. In oysters, Huan et al. [9] compared proteins in non-calcifying trochophore larvae to calcifying D-shaped larvae and found 50 differentially expressed proteins, which they divided into the four categories “cytoskeletal components”, “biochemical regulators”, “cell proliferators” and “protein modification factors”.

On the United States (US) west coast, high mortalities of Pacific oyster larvae have recently occurred in conjunction with upwelling of deep water that is undersaturated in aragonite, causing a significant loss of income for oyster hatcheries and farmers [24–26]. Carbon dioxide (CO<sub>2</sub>)-acidified seawater can cause shell deformations and reductions in shell size of developing larvae (e.g. [27–29]) and delays in the initiation of shell formation (e.g. [27, 30–32]). The delay in shell formation of oyster larvae exposed to CO<sub>2</sub>-acidified seawater is affected by seawater aragonite saturation state ( $\Omega_{\text{ARAG}}$ ) [5, 33, 34] and/or possibly by the ratio of bicarbonate to hydrogen ions at sites of calcification [34, 35].

Concurrent with shell formation, larval oysters are undergoing a complex transition from trochophore to veliger larvae, a process which invokes a myriad of physiological and transcriptomic changes [10]. In this study, we have capitalized on the delay in initial shell development of oyster larvae under acidified seawater conditions in order to identify genes that are correlated with shell calcification during this early developmental phase. We have compared gene expression profiles during early shell development of larvae in ambient ( $\Omega_{\text{ARAG}} \approx 2.5$ –3.0) seawater with those of genetically similar larvae in acidified ( $\Omega_{\text{ARAG}} \approx 1.0$ –1.25) seawater during the first 18 h post-fertilization in order to identify expression of genes previously known to be involved in shell formation, but also expression of novel putative genes coding for shell matrix proteins, as well as other processes associated with shell formation. This is the first study to use a high temporally resolved sampling scheme (every two hours) while assessing global gene expression changes due to  $\Omega_{\text{ARAG}}$  stress in *C. gigas* larvae, an investigation which could provide a better understanding of how oyster populations may respond to environmental change. Additionally, this study could provide insight into potential targets of natural selection under future ocean acidification scenarios.

## Results

### Water chemistry

In replicate experiment 1, the partial pressure of carbon dioxide ( $p\text{CO}_2$ ) in ambient conditions ranged between 462.8–731.0  $\mu\text{atm}$  ( $\mu = 564.9$ , s.d. = 65.6), while in the treatment it ranged between 1325.8–1724.5  $\mu\text{atm}$  ( $\mu = 1515.4$ , s.d. = 104.1) (Table 1; Additional file 1: Table S1). Aragonite saturation state ranged between 2.08–2.78 ( $\mu = 2.53$ , s.d. = 0.18) in ambient cultures, while in the treatment replicates it ranged between 1.06–1.31 ( $\mu = 1.19$ , s.d. = 0.07) (Table 1). One sample from the ambient group (6 h post-fertilization, replicate A) was found to contain low  $\Omega_{\text{ARAG}}$  water, potentially as the result of a mistake during filling of the culturing vessel, and was thus treated as a low  $\Omega_{\text{ARAG}}$  treatment sample for the gene expression analyses. The exclusion of certain cultures resulting in lack of replication within certain time points is not of major concern from a statistical perspective due to our choice of regression analysis across all time points.

In replicate experiment 2, the  $p\text{CO}_2$  in ambient conditions ranged between 405.0–521.6  $\mu\text{atm}$  ( $\mu = 449.0$ , s.d. = 30.8), while in the treatment it ranged between 1103.0–1397.9  $\mu\text{atm}$  ( $\mu = 1241.9$ , s.d. = 109.2) (Table 1). Aragonite saturation state ranged between 2.63–3.25 ( $\mu = 3.00$ , s.d. = 0.175) for ambient conditions while in the treatment it ranged between 1.27–1.60 ( $\mu = 1.44$ , s.d. = 0.115) (Table 1). In this experiment one sample from the low

**Table 1** Mean ( $\pm$ SD) temperature, salinity, total alkalinity ( $\mu\text{eq kg}^{-1}$ ), total  $\text{CO}_2$  ( $\text{TCO}_2$ ), partial pressure  $\text{CO}_2$  ( $p\text{CO}_2$ ), bicarbonate ( $\mu\text{mol kg}^{-1}$ ), carbonate ( $\mu\text{mol kg}^{-1}$ ), pH ( $\text{pH}_T = \text{pH}$  on the total scale) and saturation state of aragonite ( $\Omega_{\text{ARAG}}$ ) for control and high  $p\text{CO}_2$  seawater treatments across two experiments rearing *C.gigas* larvae from 2 to 18 h post fertilization

		Parameter								
		Temp. ( $^{\circ}\text{C}$ )	Salinity (ppt)	Alkalinity ( $\mu\text{eq kg}^{-1}$ )	$\text{TCO}_2$ ( $\mu\text{mol kg}^{-1}$ )	$p\text{CO}_2$ ( $\mu\text{atm}$ )	$\text{HCO}_3^-$ ( $\mu\text{mol kg}^{-1}$ )	$\text{CO}_3^{2-}$ ( $\mu\text{mol kg}^{-1}$ )	$\text{pH}_T$	$\Omega_{\text{ARAG}}$
Experiment 1	Control	25.8 $\pm$ 0.49	30.4 $\pm$ 0	2259 $\pm$ 11	2053 $\pm$ 13	565 $\pm$ 65.6	1882 $\pm$ 21	154.7 $\pm$ 11.4	7.92 $\pm$ 0.04	2.53 $\pm$ 0.18
	High $p\text{CO}_2$	25.9 $\pm$ 0.28	30.4 $\pm$ 0	2285 $\pm$ 14	2225 $\pm$ 11	1515 $\pm$ 104.1	2109 $\pm$ 10	72.6 $\pm$ 4.5	7.54 $\pm$ 0.03	1.19 $\pm$ 0.07
Experiment 2	Control	26.1 $\pm$ 0.23	30.4 $\pm$ 0.04	2251 $\pm$ 18	2003 $\pm$ 11	449 $\pm$ 30.8	1808 $\pm$ 15	182.4 $\pm$ 10.6	8.01 $\pm$ 0.03	2.99 $\pm$ 0.17
	High $p\text{CO}_2$	26.3 $\pm$ 0.28	30.4 $\pm$ 0.12	2284 $\pm$ 14	2194 $\pm$ 13	1242 $\pm$ 109.2	2071 $\pm$ 14	87.5 $\pm$ 6.8	7.62 $\pm$ 0.04	1.44 $\pm$ 0.11

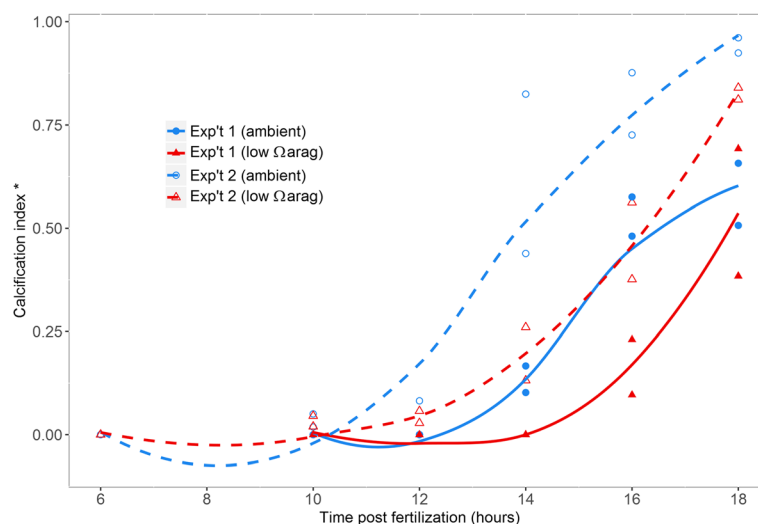
$\Omega_{\text{ARAG}}$  group (6 h post fertilization, replicate A) was found to have an abnormally low  $\Omega_{\text{ARAG}}$  (0.688). As the variance in all other replicates was very low, we interpret this as a post-experimental sterilization failure of the water sample before analysis and not a true treatment effect. Carbonate chemistry in static culture systems is strongly influenced by biologic metabolism within the culture unit(s) and, as such, inherent differences in stocking rate (biomass) result in variability of seawater carbonate chemistry. Despite notable variation in seawater  $p\text{CO}_2$  between replicates for our experiments,  $\Omega_{\text{ARAG}}$  levels were consistently and substantially different between treatments:  $> 2$  for 'ambient' conditions and  $\approx 1.0$ – $1.5$  for 'treated' seawater. No trend of decreasing  $\Omega_{\text{ARAG}}$  with time was observed (Table 1, Additional file 1: Table S1).

### Shell deposition

Both timing of onset and rate of shell calcification was significantly different between seawater treatments and

experiments ( $p < 0.05$ , Additional file 2: Table S2). In replicate experiment 1, larvae in ambient conditions began to form aragonite crystals at 14 h post-fertilization ( $\mu_{\text{CI}} = 0.13$ , s.d. = 0.05) (3% fully calcified, 20% partially calcified), while larvae in low  $\Omega_{\text{ARAG}}$  treatments did not start calcifying until 16 h post fertilization and then in much lower proportions than in ambient conditions (Fig. 1). At 18 h post fertilization, the proportion of calcified larvae in the low  $\Omega_{\text{ARAG}}$  treatment reached similar levels to that of larvae in ambient conditions ( $\mu_{\text{CI}} = 0.58 \pm 0.11$  and  $0.54 \pm 0.22$  for ambient and low  $\Omega_{\text{ARAG}}$  treatments, respectively (61% part- or fully calcified vs 62% in ambient).

In replicate experiment 2, the same general calcification pattern was seen as in experiment 1 (Fig. 1), although larvae in both ambient conditions and low  $\Omega_{\text{ARAG}}$  treatment began to partially calcify to a small extent as early as 10–12 h post fertilization ( $\approx 7\%$  partially calcified in each). At 14 h, however, only larvae in the



**Fig. 1** Calcification index of larval cultures from 6 to 18 h post fertilization. Calcification index (CI) is calculated as:  $\text{CI} = (\text{FC} + (\text{PC} * 0.5)) / \text{TL}$ , where FC, PC and TL denote the numbers of observed fully calcified, partially calcified and total larvae from each sample, respectively. Ambient (control) and low  $\Omega_{\text{ARAG}}$  (treatment) conditions are represented in blue and orange respectively, with locally estimated average (LOESS) trends represented by lines. Experiment 1 is displayed as solid lines and filled points, experiment 2 is represented by dashed lines and unfilled points. Symbols of the same type within each time point correspond to the two independent replicate cultures as specified in the "Methods" section

ambient treatment were fully calcified (48%) and at 16 h 74% were fully calcified in the ambient while only 22% were fully calcified in the low  $\Omega_{ARAG}$  treatment. Similar to the results of replicate experiment 1, this difference diminished at 18 h (97% part- or fully calcified in ambient conditions vs 88% in the low  $\Omega_{ARAG}$  treatment).

### Bioinformatic analyses of RNA transcripts

Samples were sequenced with 50 bp single-end reads, ranging from 13.2–77.4 Mreads sample<sup>-1</sup> ( $\mu = 39.0$  Mreads). After quality trimming and removal of residual adapter sequences, a mean of 97.3% of the reads were retained, with a mean quality score of 36.1 and mean length of 47.3 bases, considering data from both replicate experiments (Additional file 3: Table S3). Due to low output of the first sequencing run for replicate experiment 1, most of the samples were sequenced one more time, more than doubling the number of reads in this replicate experiment, with the exception of one 10 h replicate of the ambient treatment and both 10 h replicates of the low  $\Omega_{ARAG}$  treatment, as well as both 12 h replicates of the ambient and one 12 h replicate of the low  $\Omega_{ARAG}$  treatments, which were not re-sequenced. Unfortunately one of the ambient treatment replicates at 16 h in experiment 1 was lost during the library preparation protocol.

The fraction of reads that mapped uniquely to one position in the genome coding regions ranged from 3.1% to 38.0% ( $\mu = 25.7\%$ ). This rather low fraction was likely due to the method of extraction of RNA in bulk from seawater, which would also extract RNA from a variety of micro-organisms. The fraction of duplicate reads ranged from 6.48% to 24.79% ( $\mu = 18.72\%$ ). After removal of duplicate reads, sequencing depth ranged from 9.50–57.3 Mreads/sample ( $\mu = 30.5$  Mreads) (Additional file 4: Figure S1).

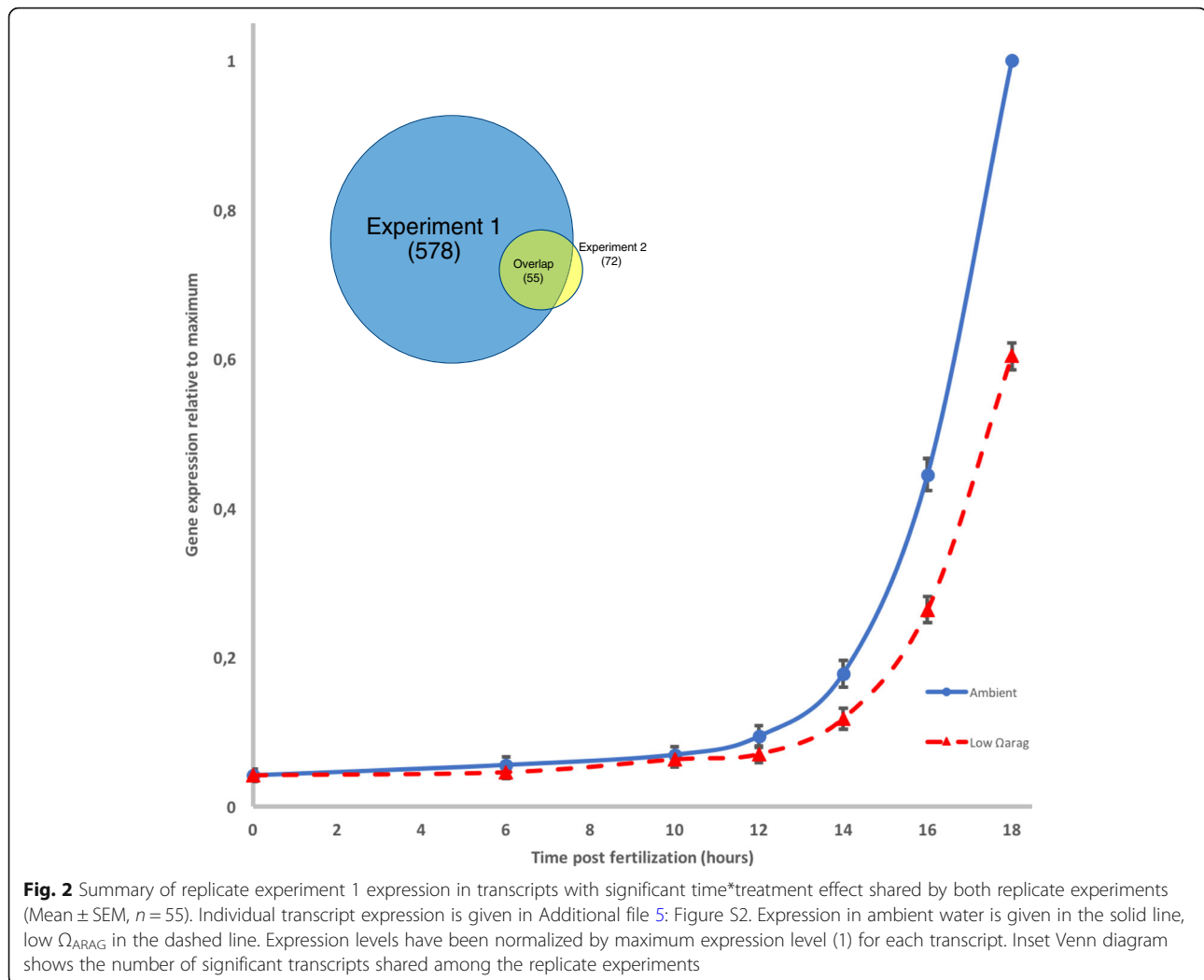
### Differentially expressed transcripts

Filtering out transcripts with low expression values, low variances and ones not showing a positive expression \* time interaction term in a generalized linear model ( $p < 0.05$ ), left 5448 transcripts in experiment 1 and 4030 transcripts in experiment 2. From these datasets, 578 transcripts showed a significant (time \* treatment) interaction effect in the log linear model:  $\log(y) = \beta_0 + \beta_1 \text{time} + \beta_2 \text{treatment} + \beta_3 (\text{time} * \text{treatment})$  in experiment 1 after a Benjamini-Hochberg false discovery rate correction (FDR = 0.05%), while there were 72 transcripts in experiment 2 ( $p < 0.05$ ). Fifty-five of the transcripts were shared between the two experiments (Inset in Fig. 2; Additional file 5: Figure S2), all of which show higher expression levels in the ambient than in the low  $\Omega_{ARAG}$  treatment (Fig. 2). Out of these, 31 had an InterPro annotation [36] through the genome sequence, and 25 had a Gene Ontology (GO) functional annotation associated

with them. In total, 37 were attached to some form of annotation, and all of these could be classified into one of four categories: Metabolic genes ( $n = 3$ ), Transmembrane Proteins (transporters) ( $n = 8$ ), Shell Matrix Proteins ( $n = 16$ ) and Protease Inhibitors ( $n = 10$ ) (Table 2). Overrepresented GO categories (within the “Molecular Function” category) in this list are: “endopeptidase inhibitor/regulator activity” (GO:0004866/GO:0061135; corrected  $p$ -value  $3.35 * 10^{-10}$ ), “serine-type endopeptidase inhibitor activity” (GO:0004867; corrected  $p$ -value  $1.68 * 10^{-10}$ ), “serine-type endopeptidase activity” (GO:0004252; corrected  $p$ -value 0.0251) and “serine-type peptidase/ hydrolase activity” (GO:0008236/GO:0017171; corrected  $p$ -value 0.0223) (Additional file 6: Figure S3). Looking closer at the transcripts within each of these categories, they code for the extracellular metalloprotease matrix protein “Papilin” (CGI\_10,020,818 / CGI\_10,021,289), as well as a variety of different protease inhibitors such as for example Antistatin (CGI\_10021371), Trypsin inhibitors (CGI\_10,015,381 / CGI\_10,012,273 / CGI\_10,020,625) and Cystatins (CGI\_10013713 / CGI\_10013715 / CGI\_10013717).

### Weighted gene correlation network analysis

Clustering the expression data from the 5448 transcripts from replicate experiment 1, after filtering out transcripts that were found not to be positively correlated with time and with low variance, rendered two major clusters, of which one showed temporal differences in expression patterns between ambient conditions and low  $\Omega_{ARAG}$  treatment (“blue” in Fig. 3a), whereas the other one did not (“turquoise” in Fig. 3a). Focusing on the “blue” cluster, it contained 2592 transcripts, with the list of genes being significantly enriched (Gene-score resampling multiple-test corrected  $p < 0.05$ ) for 47 different GO categories (10 Biological Processes, 9 Cellular Components and 28 Molecular Functions) (Additional file 6: Figure S3). In replicate experiment 2, the transcripts clustered into 4 co-expression clusters, one of which showed different expression pattern in the low  $\Omega_{ARAG}$  treatment compared to ambient (“blue” in Fig. 3b). This cluster contained 1658 transcripts, and was enriched for 12 GO categories (1 Biological Process, 4 Cellular Components and 7 Molecular Functions), all of which were also enriched in replicate experiment 1 except for tubulin-tyrosine ligase activity (GO:0004835) and serine-type peptidase activity / serine hydrolase activity (GO:0008236 / GO:0017171) (Additional file 6: Figure S3). The lower number of significant categories reflect the decrease in statistical power due to lower sequencing depth in replicate experiment 2. These categories also include the ones enriched in the dataset of transcripts showing a significant (time \* treatment) interaction described above (Additional file 6: Figure S3).



The transcripts responsible for the enriched functional categories in the “blue” clusters generated by the WGCNA analysis, together with transcripts with significant (time \* treatment) interactions, can be classified into three different main functions: 1) Extracellular Matrix Formation (Dynein chains, Myosin, Tubulin, Tektin, Integrin, Fibrillin, Cadherin, as well as Chitin binding proteins), 2) Transmembrane Ion Transport and homeostasis (Potassium, Sodium, Calcium and Copper channels, “Atrial Natriuretic Peptide” and several Lyases (Adenylate and Guanylate cyclase)), and 3) Protease Inhibitors (Papilin, Trypsin, Cystatin, Antistatin, Metalloproteinase inhibitor 3, Carboxypeptidase inhibitor SmCl).

## Discussion

### Delay in shell deposition rates

Our results indicate that shell formation of *C. gigas* larvae is affected by OA conditions. Shell development rates were reduced at aragonite conditions of 1.06–1.31

and 1.27–1.60 in experiments 1 and 2, respectively. This finding is in agreement with reports by others who have shown that shell formation of Pacific oyster larvae is impacted at  $\Omega_{ARAG}$  below 1.5 [33]. In both of our replicate experiments, the low  $\Omega_{ARAG}$  treated larvae started forming their shells at a later time than in ambient conditions. There seems to be a particularly large difference at the 14 and 16 h time points, indicating a developmental delay for larvae exposed to low  $\Omega_{ARAG}$  conditions. This is consistent with the results for larvae of bivalve mollusks (e.g. [27, 30, 37]) and of purple sea urchins [38], and suggests that the gene expression patterns correlated with shell formation have shifted as a result of exposure to low  $\Omega_{ARAG}$  conditions.

### Differentially expressed transcripts

There are many more significantly differentially expressed transcripts between low  $\Omega_{ARAG}$  and ambient treatments in the first replicate experiment than in the second. This is most likely a result of the fact that in the



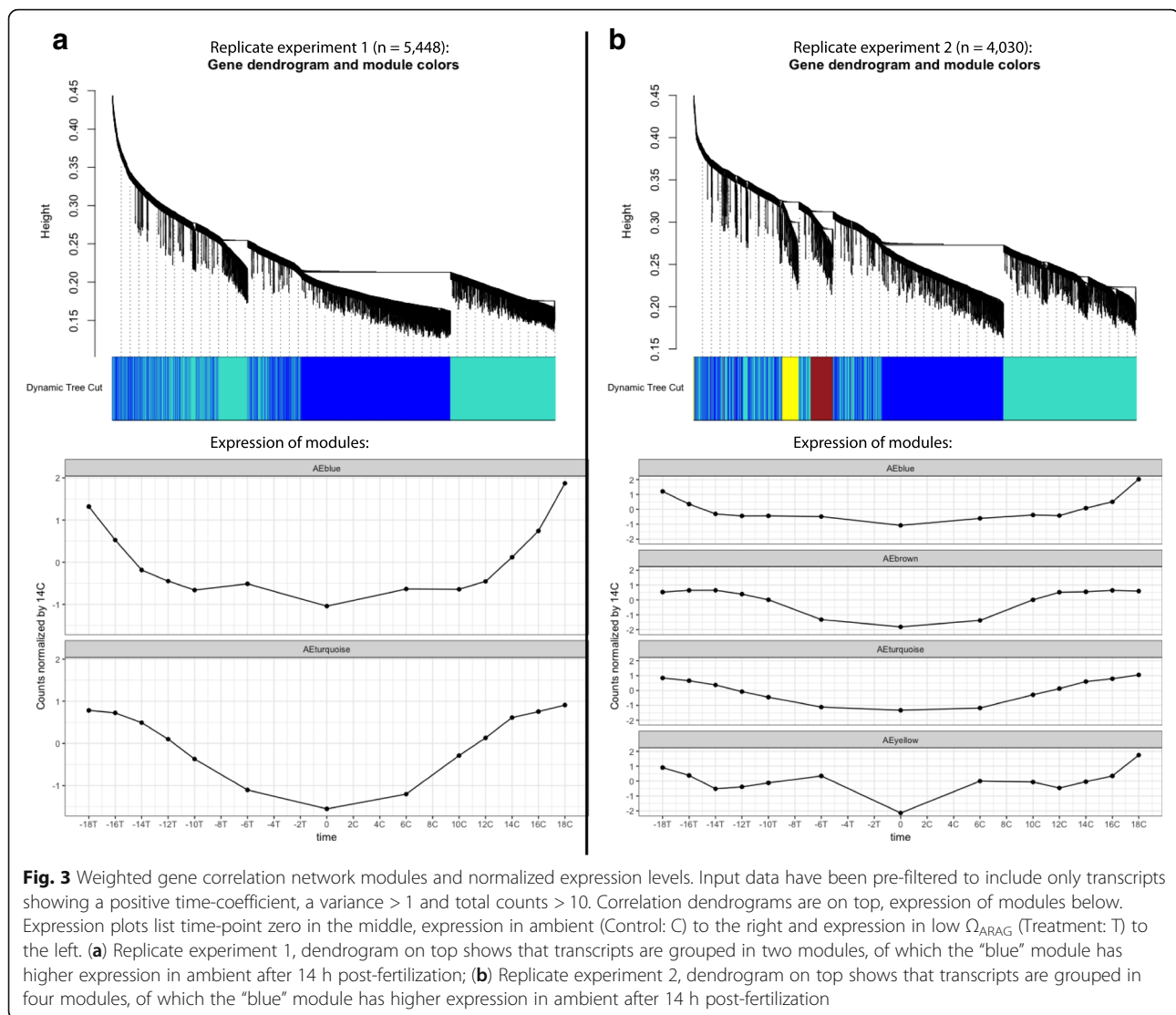
**Table 2** Thirty-seven annotated transcripts with significant time \* treatment effects in both replicate experiments, divided into four main functional categories. All are more highly expressed in ambient conditions

Metabolic transcripts	Annotation
CGI_10022578	Cytochrome P450
CGI_10025516	Sulfotransferase
CGI_10011094	Lipase
Transmembrane proteins	Annotation
CGI_10005173	Prominin
CGI_10024903	Transmembrane ion-channel
CGI_10009289	DEATH-like
CGI_10012122	Ganglioside activator protein
CGI_10007940	Caveolin
CGI_10011750	Transmembrane protein of unknown function
CGI_10012368	Pedal peptide
CGI_10022868	Leucine-rich glioma-inactivated protein
Shell matrix proteins	Annotation
CGI_10016584	C-type lectin
CGI_10024633	Chitin-binding protein
CGI_10024602	Fibrinogen, alpha/beta/gamma chain
CGI_10000698	Carbonic anhydrase / Nacrein-like protein
CGI_10027654	C-type lectin
CGI_10027048	Low-density lipoprotein receptor -like
CGI_10001361	EGF-like calcium-binding protein
CGI_10005422	Beta-lactamase-type transpeptidase
CGI_10016583	C-type lectin
CGI_10022862	Toll-like receptor
CGI_10024126	Thrombospondin
CGI_10025037	Calcium-binding EF-hand
CGI_10020619	EF-HAND 2
CGI_10013619	Temptin
CGI_10010907	Galactose-binding protein
CGI_10007447	Collagen alpha-6(VI) chain-like
Protease inhibitors	Annotation
CGI_10016790	Metalloproteinase inhibitor I3 5
CGI_10015381	Peptidase S1A, chymotrypsin-type
CGI_10018666	Cystatin-A2
CGI_10012273	Peptidase S1A, chymotrypsin-type
CGI_10010153	Protease inhibitor, Kazal-type
CGI_10020625	Peptidase S1A, chymotrypsin-type
CGI_10010888	Protease inhibitor, Kazal-type
CGI_10025096	Proteinase inhibitor I14/I15, hirudin/antistatin
CGI_10020818	Proteinase inhibitor I2, Kunitz metazoa
CGI_10021289	Proteinase inhibitor I2, Kunitz metazoa

first replicate experiment sequencing depth is twice as high for most time points. Despite this difference, there remains a remarkable overlap between the two replicate experiments: 55 out of the 72 significantly differentially expressed transcripts from replicate experiment 2 are also significantly differentially expressed in replicate experiment 1. Interestingly, all of the annotated genes from this list can be divided into only four functional categories: Metabolic Functions, Transmembrane Proteins (transporters), Shell Matrix Proteins and Protease Inhibitors. The metabolic genes are too few to result in significant enrichments for any metabolic GO category, and are restricted to specific types of metabolism, especially lipid breakdown. This could be associated with faster calcification rates in ambient seawater, as is also shown by increased expression of ion transporters and matrix protein transcripts, or by a switch in energy allocation as reported in sea urchin larvae [22]. It is somewhat unexpected to observe such a high number of protease inhibitors in this list; however, this type of inhibitor plays a very important role in preventing proteins from being hydrolysed by endopeptidases, and could be involved in shell formation as a way of protecting shell matrix proteins as they are secreted to form the extracellular matrix for mineral deposition [39]. This would especially be the case for the metallo-proteinase inhibitors, such as papilin [40], which could potentially protect important protein – CaCO<sub>3</sub> bonds. As aragonite formation is highly sensitive to the organisation of the shell protein matrix [41], degradation of some of the matrix proteins could cause the aragonite crystals to become deposited in a sub-optimal manner which would, in turn, affect the integrity of the shell. This could be a cause for the high numbers of deformed shells observed in low  $\Omega_{ARAG}$  treatments (e.g. [27, 33]).

Several of the differentially expressed transcripts are from genes known to code for parts of the shell matrix, such as nacrein [42], papilin (also a metalloprotease inhibitor; [40]), chitin-binding protein [43] and a protein with a beta-lactamase domain that is known to be part of the shell matrix, but with a currently unknown function [44]. Nacrein has a carbonic anhydrase domain [42], and has previously been shown to be strongly expressed prior to the initiation of shell formation in blue mussels [45]. Furthermore, several calcium binding proteins are represented here, and quite a few proteins involved in extracellular matrix agglutination, such as lectin, collagen, EF-Hand, thrombospondin and fibrinogen.

Transmembrane proteins are also found in this list, such as the ion channel protein caveolin, that is known to be involved in subcellular compartmentalization and vacuolar organization [46] as well as prominin that is involved in the organization of plasma membranes [47]. Interestingly, this list also includes transmembrane



proteins mostly known as neurotransmitters, such as a ganglioside activator protein, a leucine-rich glioma inactivated protein [48], the *Aplysia* “pedal peptide” gene [49] and a protein containing a DEATH-like domain [50]. These proteins could be involved in transport of ions across the plasma membrane, which in some cases can be associated with neuronal activity, but also in shell formation [51]. Altogether, this list paints an image of a machinery which binds calcium ions and synthesizes matrix proteins, then transports these components across the plasma membrane to the external environment while ensuring that the matrix proteins are not hydrolysed by proteases before calcium carbonate crystals have been deposited.

#### Weighted gene correlation network analysis

The clustering approach allows us to “zoom out” from investigations of individual transcripts to instead get an

overview of cellular functioning. The large number of transcripts falling into the co-expression cluster more highly expressed in the ambient than in the low  $\Omega_{ARAG}$  treatment highlights that the cellular response to aragonite saturation stress involves many genes/pathways. While being more sensitive and thus causing more GO categories to be enriched, the clustering result agrees strongly with the result of the individual transcript analysis (significant time\*treatment transcripts) in that the categories can be grouped into ion transport, protein synthesis, extracellular matrix proteins, and a few metabolic pathways (lipid breakdown). One type of transcript that is highly overrepresented in many of these categories is dynein (both cytoplasmic and axonemal). Dynein motor proteins are well-known major transporters of cellular components, and it is possible that these have a role in the transport of the building blocks of the shell matrix to the location where it should be deposited.

The GO category “viral capsid” is also enriched in replicate experiment 1. This could seem surprising at first sight, but these are all chitin-binding proteins and thus likely to be part of the shell formation machinery [43].

Transmembrane serine protease and antistasin are also commonly found in the enriched GO categories. These are genes known to have a part in immune defence and regulation of coagulation in other organisms [52, 53], and is possible that these genes also have a role in the control of the deposition of shell matrix proteins in larval oysters.

### Comparison with previous work

As part of the work describing the *C. gigas* genome, gene expression profiles of developing oyster larvae were produced [10]. Mining this resource (<https://www.ncbi.nlm.nih.gov/bioproject/PRJNA146329>) for expression levels in the 55 transcripts exhibiting significant time\*treatment interactions in our data, we see consistent patterns with those of Zhang et al. [10] who found that all of these transcripts also started being expressed during the transition between “trochophore” and “early D shaped larva” stages, some of these to extremely high expression levels. Interestingly, however, Zhang et al. [10] found that some of these transcripts only reached their maximum expression levels at later developmental stages, hours or even days after the end of our experiment, while others were highly expressed only in D-larvae. It is reasonable to expect that there are different mechanisms involved in the initial process of shell nucleation, compared to later shell formation stages where aragonite layers are being deposited on pre-existing layers, thus it makes sense to pay special attention to genes expressed in only early D-larvae. Unfortunately, however, the transcripts which Zhang et al. [10] found to be most highly expressed at this stage are not annotated (e.g. CGI\_10010907, CGI\_10022681, CGI\_10014978, CGI\_10022681), thus their functions are not known. Further experiments using molecular biological tools to elucidate the functions of these genes could be an interesting subject of future research.

Using a proteomics approach, Huan et al. [9] identified a number of proteins differentially expressed between trochophore (11 h post-fertilization) and D-larvae (21 h post-fertilization). Some of the expressed proteins could also be categorized as matrix proteins (e.g. tubulin, tropomyosin) and protein modification factors. Differences between their results and those of our study could have been due to a number of factors, for example the timing of sampling points, differences in oyster population used, or a lack of correlation between transcript and protein expression [22].

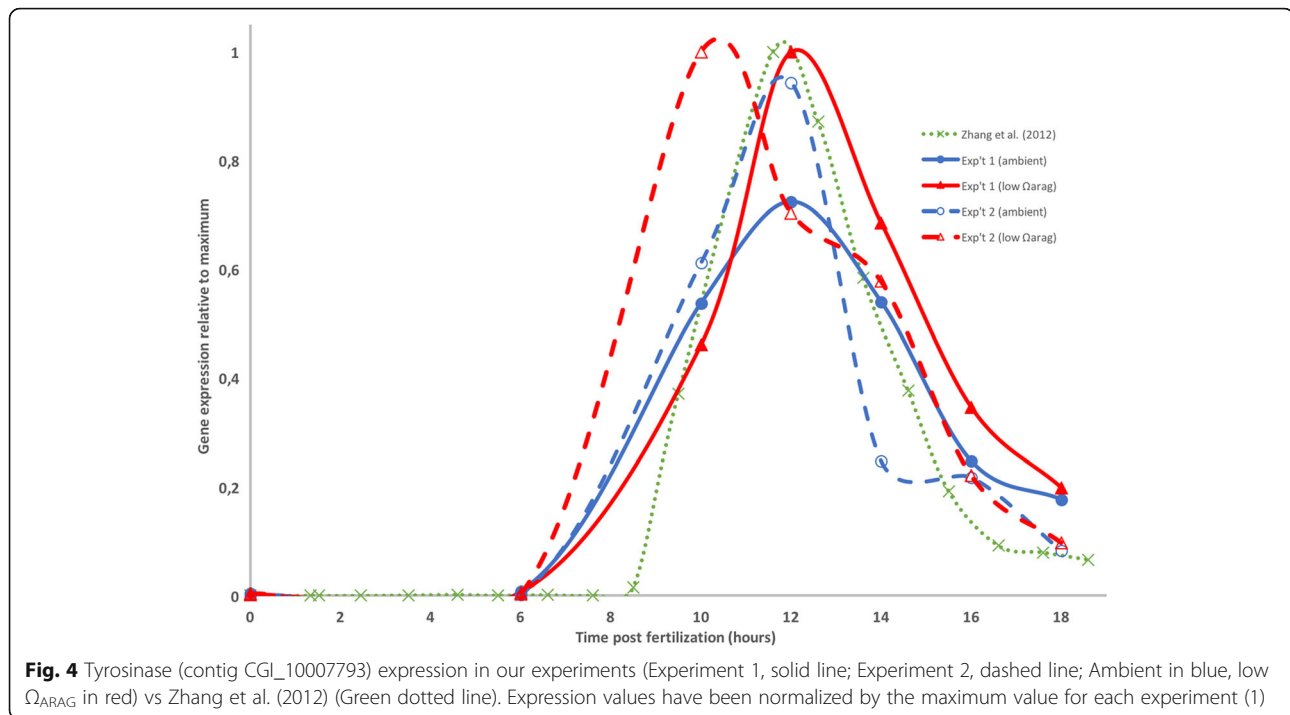
In a recent article, Wang et al. [21] found that adenylyl cyclase was an important mediator of bicarbonate ion

concentrations and intracellular pH in low  $\Omega_{ARAG}$  conditions in adult *C. gigas*. Interestingly, it seems that the same mechanism is also important in larval oysters, as shown by both adenylyl and guanylyl cyclases being commonly found in our enriched GO categories. Guanylyl cyclase is also known to mediate the import of calcium ions into cells, through production of cyclic guanosine monophosphate (cGMP), which can act to keep cGMP-mediated calcium ion channels open [20]. Alkaline phosphatases, such as adenylyl cyclase, have also been found to be highly expressed in the gastropod *Biomphalaria glabrata* just prior to the onset of shell formation [54].

Huan et al. [19] identified a tyrosinase gene (cgi-tyr1) in *C. gigas* which was highly expressed in the trochophore and D-larvae stages but not later in larval development. More recently, Yang et al. [55], found that the homolog in *C. angulata* (Ca-tyrA1) was expressed especially in trochophore larvae, but was also significantly upregulated in a high  $pCO_2$  treatment (3000 ppm). The authors concluded that this was likely in response to high  $pCO_2$  induced shell damage, and that Ca-tyrA1 was involved in larval shell repair. Wang et al. [56] proposed the involvement of a tyrosinase gene (CGI\_10017214) in the shell formation process following shell damage in adult *C. gigas*. An important role of tyrosinases in the first stages of shell repair in the blue mussel *Mytilus edulis*, has also recently been discussed [57]. Our data support previous work in that the tyrosinase gene (gene ID CGI\_10007793; available at <http://www.uniprot.org/uniprot/U5U0P0>) is highly expressed before shell formation starts (10–14 h post-fertilization), perhaps in association with formation of the larval pellicle or periostracum [18, 55, 58], and then the expression level decreases rapidly. In addition, the transcripts are expressed more in the low  $\Omega_{ARAG}$  treatment than in ambient during these hours and the larvae in low  $\Omega_{ARAG}$  in replicate experiment 2 peak in tyrosinase expression several hours earlier than remaining treatments (Fig. 4). Examining the expression data from Zhang et al. [10] shows that the tyrosinase contig (CGI\_10007793) was highly expressed at 11.5–13.5 h (Fig. 4). In addition, this contig clusters (using the WGCNA approach as described above) in a co-expression cluster together with 11 other contigs, out of which several are cation channel proteins and potential matrix proteins (Table 3). These transcripts would be interesting candidates for further study of their role in the onset of shell formation.

In search of regulatory proteins controlling larval shell formation, Liu et al. [17] found that the transcription factor GATA2/3 was expressed in the edge of the shell during trochophore and D-larval stages in *C. gigas*. In our dataset, GATA3 (GeneID: CGI\_10013217) is most highly expressed at 6 h post fertilization, then decreases





in expression with time, but there is no difference between levels under ambient and low  $\Omega_{ARAG}$  conditions. Liu et al. [17] also concluded that the GATA genes might be involved in other functions rather than in shell formation, most notably they discussed the possibility of them being involved in haemocyte formation, as has been shown in scallops [59]. This could potentially be an interesting area of further research, as it has been shown that calcification is initiated inside haemocytes before the constituents are transported to the location of shell deposition [1]. Liu et al. [17] also proposed that GATA-3 could be part of the BMP signalling pathway, which has

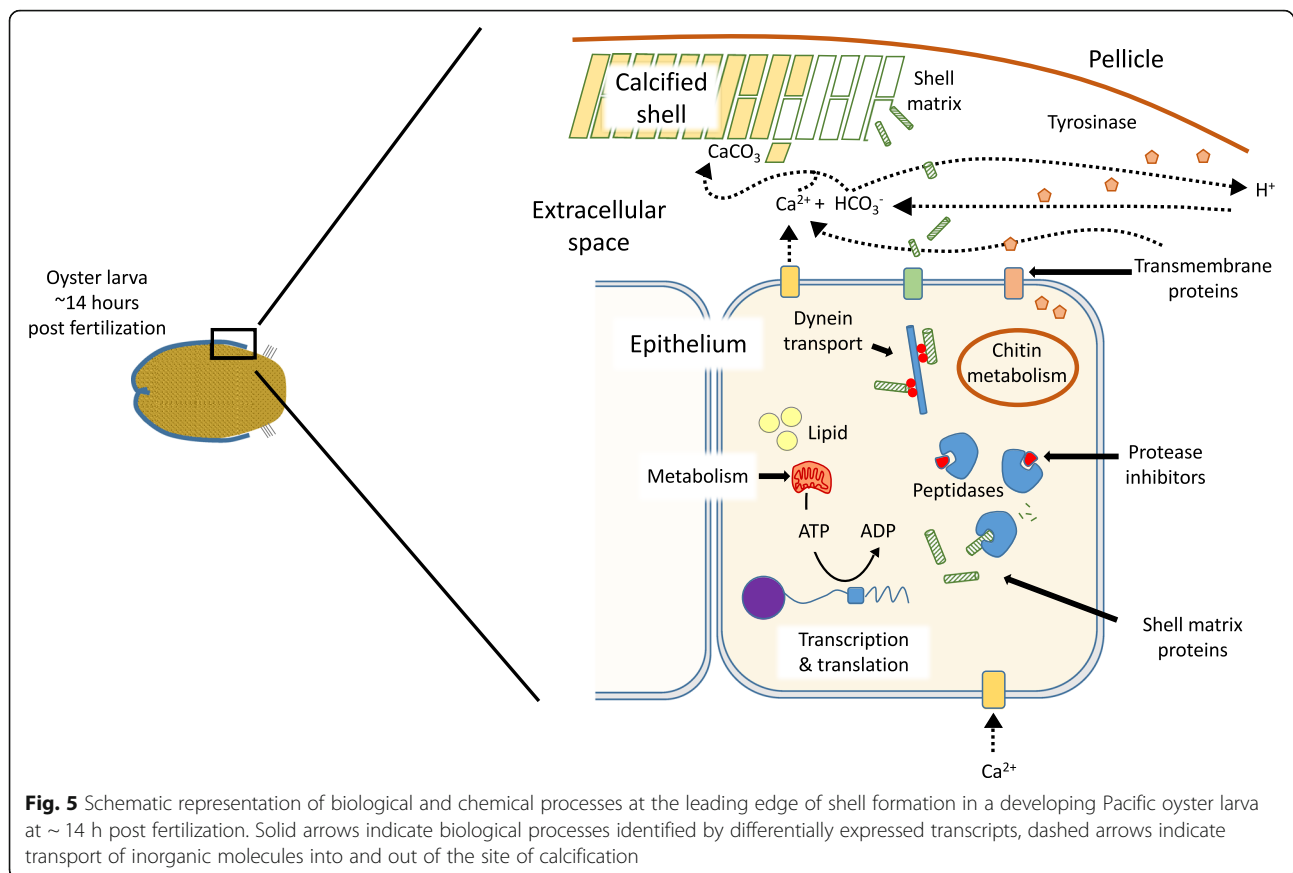
**Table 3** Contigs co-expressed with the Tyrosinase contig CGI\_10007793

Contig ID	Annotation
CGI_10003647	EGF-like calcium-binding protein
CGI_10005120	NHL repeat containing protein-binding
CGI_10011377	Glucose/ribitol dehydrogenase
CGI_10012546	No annotation
CGI_10018671	Strictosidine synthase
CGI_10018819	Transient receptor potential cation channel
CGI_10018820	Similar to transient receptor potential cation channel
CGI_10022617	No annotation
CGI_10024194	Follistatin-like, protease inhibitor, Kazal-type
CGI_10028176	No annotation
CGI_10028233	Transient receptor potential cation channel

been implied as critical for shell formation in the gastropod *Lymnaea stagnalis* [60]. Other potential parts of the BMP pathway could consist of different types of growth factors. Liu et al. [16] investigated the expression patterns of the two transforming growth factors (TGF) *cgi-smad1/5/8* and *cgi-smad4*, and found that the former was more highly expressed within the shell field than in other parts of the larvae during early shell formation, suggesting an involvement in shell formation through regulating transitions between different developmental stages in the early development of oyster larvae. In our data, *cgi-smad1/5/8* (Gene ID: CGI\_10014747) is highly expressed at all time points except in eggs, with no difference between ambient and treatment (Additional file 7: Figure S4), while *cgi-smad4* (Gene ID: CGI\_10000594) is highly expressed in eggs, then decreases in expression with time in replicate experiment 2, while in replicate experiment 1 it shows a spike in activity at 6 h post fertilization, after which expression levels decrease with time (Additional file 7: Figure S4).

## Conclusions

By examining differences in gene expression in oyster larvae during the period when shell formation is delayed by low  $\Omega_{ARAG}$  (14–16 h post fertilization), together with previous findings of other studies, we can start obtaining a better understanding of the mechanism by which the prodissoconch I larval shell is initially deposited (Fig. 5). Although it is possible that this delay in shell formation



is due to a general developmental delay, almost all of the differentially expressed transcripts observed in this study seem to be related to shell deposition. Our data suggest that the shell formation mechanism can be divided into three different main parts, namely 1) transport of ions across plasma membranes, 2) secretion of shell matrix proteins, and 3) production of protease inhibitors (Fig. 5). Increased expression of transcripts for calcium ion pumps suggests increased release of calcium ions into the epithelial space that would result in elevated aragonite saturation states. Dynein motor proteins may transport shell matrix proteins to this external space, facilitating and perhaps controlling the deposition rate and organization of shell aragonite crystals, although these observed expression changes could potentially also be due to other cellular processes. We also observed increased transcription of genes for chitin catabolism – chitin has been reported to be present in larval shells [43]. The role of the significant increase in transcripts for protease inhibitors is unknown, but one may hypothesize that they are important in preventing breakdown of proteins involved in shell formation (Fig. 5).

At later stages of development, it has been shown that metabolism is affected by aragonite saturation conditions, but this effect seems to be minor and associated

with breakdown of lipid energy reserves during the period from 14 to 16 h post fertilization when larvae are starting the process of shell formation. Thus, production of shell matrix and other proteins would seem to be fueled by higher rates of lipid breakdown, based on increased transcription of genes coding for lipases.

Interestingly, we do not see any gene expression differences in genes previously hypothesized to be involved in the shell formation signalling pathway. This could be due to a lack of power due to low replication in this study, or it could actually be that the signalling pathway remains unaffected by the water chemistry, and that the detrimental effects of low  $\Omega_{ARAG}$  on shell formation occurs downstream at the shell matrix level. We do, however, see a spike in expression of genes coding for tyrosinase prior to the onset of shell formation (10–12 h post fertilization), which is stronger in low  $\Omega_{ARAG}$ , as was recently also described in *C. angulata* [55]. This spike probably results in cross-linking of proteins and formation of the early larval pellicle. It is unlikely, however, that the pellicle can form a completely sealed space between itself and the outer shell-forming, epithelial layer of the developing larva by 14–16 h post fertilization. In support of this contention, Waldbusser et al. [7] reported that carbon isotopes of the early oyster

larval shell were dominated by those from the external seawater medium and not from respired carbon. The lack of a sealed external epithelial space also suggests that hydrogen ions, formed by formation of carbonate, as well as calcium ions can also be exchanged by diffusion with the external seawater medium [7] (Fig. 5).

This study describes how low  $\Omega_{\text{ARAG}}$  affects gene expression and shell formation of early oyster larvae. The identified genes, which show changes in expression as a result of OA conditions may play an important role in determining the capacity of oyster larvae to respond to future OA stress under natural and hatchery conditions.

## Methods

### Larval culture

Commercial broodstock of *Crassostrea gigas* was obtained from Netarts Bay, Oregon, USA, in August 2014. Recent publications have suggested that broodstock exposure to high  $p\text{CO}_2$  seawater may result, alternatively, in more robust larvae [61] or compromised ones [62]. In order to simplify our experiments, we chose to maintain all broodstock in 'optimal' conditions with  $\text{pH} \approx 8.0\text{--}8.3$  and  $\Omega_{\text{ARAG}} > 2$  (buffered with sodium carbonate) at a temperature of 19–20 °C prior to experimentation at the Hatfield Marine Sciences Center (HMSC), Oregon State University, Newport, Oregon, USA. Eggs from two females were collected through strip spawning and each sample was equally divided into four aliquots in seawater at ambient  $\Omega_{\text{ARAG}} (> 2)$  and 25 °C. Each aliquot from each female was fertilized with sperm from one of four males. After one hour, fertilization was confirmed and excess sperm were removed by washing the eggs on a 20  $\mu\text{m}$  mesh sieve, then equal numbers of fertilized eggs from each cross were combined. After pooling, the eggs were counted and  $\approx 50,000$  eggs were sampled for RNA by preserving them in RNAlater. Eggs were stocked in sealed 1 L glass jars approximately 2 h post fertilization at a density of 40 eggs  $\text{ml}^{-1}$ , as prior trials hatching *C. gigas* larvae up to 80 eggs  $\text{ml}^{-1}$  had shown no adverse effects to density at this age. The larvae were reared in either ambient ( $\Omega_{\text{ARAG}} \approx 2.5\text{--}3.0$ ) or low  $\Omega_{\text{ARAG}}$  ( $\Omega_{\text{ARAG}} \approx 1.0\text{--}1.25$ ) treatment conditions. The level of seawater acidification selected for the low  $\Omega_{\text{ARAG}}$  treatment was chosen in order to increase the likelihood of generating sufficient larval developmental delay. In full salinity seawater ( $\approx 33$  ppt), a  $p\text{CO}_2$  level of  $\approx 1500$  ppm results in  $\Omega_{\text{ARAG}}$  of 1–1.5, a range that impedes shell formation [33] and is also routinely observed during periods of upwelling along the West Coast USA [25].

Seawater types for each treatment were prepared by filling two 200 L tanks with full strength seawater ( $\approx 32$  ppt)  $\approx 18$  h before the start of each experiment and aerating them vigorously with outside (ambient) air to allow the seawater to equilibrate with ambient  $\text{CO}_2$

concentrations overnight. Next, a mixture of outside air and pure  $\text{CO}_2$  moderated by mass flow controllers (Alicat Scientific, Tuscon, AZ), was bubbled into the low  $\Omega_{\text{ARAG}}$  seawater tank for  $\approx 2$  h until a pH of 7.5 was reached. Nominal pH values were measured using an Orion Star A211 pH meter with a Ross Ultra pH/ATC triode probe (Thermo Fisher scientific) calibrated with NBS buffers and standardized with a certified seawater reference (Batch 22, A.G. Dickson, Scripps Institution of Oceanography, US.). Seawater for both ambient and low  $\Omega_{\text{ARAG}}$  conditions were transferred from corresponding storage tanks to sealed 1 L glass jars (VWR scientific, part no. 89094–014) filled with 800 mL seawater immediately prior to stocking with eggs and were sealed with a screw top lid. Cultures were treated prophylactically with 2 ppm of the antibiotic chloramphenicol to reduce bacterial respiration. Previous experiments had indicated that this antibiotic treatment had no effect on larval growth and survival [63]. Larvae were sampled at 6, 10, 12, 14, 16 and 18 h post fertilization from two independent replicate glass jars for both ambient and low  $\Omega_{\text{ARAG}}$  conditions which results in a total of 24 samples.

At each sampling point, temperature, salinity and pH of the culture medium were measured, then sample water volumes from each duplicate vessel per treatment were siphoned through a screen (attached to the end of a silicone tube immersed in the culture vessel to prevent larval removal) into a 350 ml amber glass bottle. Each sample was preserved by the addition of 30  $\mu\text{l}$  saturated  $\text{HgCl}_2$  and sealed with a polyurethane-lined metal crimp cap, until subsequent carbonate chemistry analysis was performed (see below). The remaining culture volume was sieved onto a 25  $\mu\text{m}$  mesh screen to retain larvae, which were then re-suspended in 25 ml of ambient or low  $\Omega_{\text{ARAG}}$  seawater, depending on their exposure treatment. Larval concentrations were estimated by aliquoting 30  $\mu\text{l}$  ( $n = 3$ ) volumes of vigorously agitated suspensions and counting larvae in the samples under a microscope.

At all sampling points, except at 6 h post fertilization, two to three hundred developing larvae were transferred to 20 ml shell vials with 10 ml of seawater, to which 250  $\mu\text{l}$  of buffered formalin (pH 8.2) was added, for later analyses of shell deposition by means of cross-polarized light (CPL) microscopy (see below). Remaining larvae were concentrated by centrifugation (2 min at 4500 rpm) after addition of 10 ml distilled water to reduce larval buoyancy to facilitate pellet formation, after which all supernatant seawater was immediately removed. Larvae (ca. 10,000/sample) were then flash frozen in liquid nitrogen in order to break larval shells to allow RNAlater to quickly infuse into larval tissues. The time period from addition of freshwater to flash freezing of samples was less than 5 min to avoid effects of sampling on gene expression. Samples were stored in

RNA later and divided between two replicate 1.5 ml Eppendorf tubes for later transcriptomic analyses (see below). All tubes were maintained at 4 °C for 24 h, then stored at -20 °C until RNA extraction.

The entire experiment was repeated a few days later using the same methodology but different broodstock (from the same source) was used and sperm from five males instead of four were added to batches of eggs from each of two females. In addition, eggs were stocked at a density of 20 eggs/ml (rather than 40 in experiment 1), resulting in ca. 5000 larvae per replicate culture sampled for transcriptomic analyses.

### Water chemistry analyses

Preserved water samples were analysed by the lab of Dr. Burke Hales at Oregon State University following the procedure outlined by Hales et al. [64] and Bandstra et al. [65] to obtain values for sample total dissolved carbon dioxide (TCO<sub>2</sub>), pCO<sub>2</sub>, and seawater pH, from which  $\Omega_{ARAG}$  values were calculated. This method has been shown to be highly accurate, providing TCO<sub>2</sub> and pCO<sub>2</sub> estimates with <0.2% and <5% uncertainty, respectively [7].

### Shell deposition measurements

Larvae from each time point were analysed for calcification using cross-polarized light (CPL) microscopy as described by Waldbusser et al. [5]. Patterns of refractive light viewed under CPL indicated deposition of crystalline calcium carbonate (aragonite) in the process of shell formation. A fully formed shell is curved and the secondary refraction of the polarized light creates a “Maltese cross” pattern: a dark cross area over the center of the otherwise illuminated larval shell. In this way, we were able to classify larvae into three categories: non-calcified, partially calcified and fully calcified (shelled) larvae for each sample. From these data, a “calcification index” (CI) was calculated as:  $CI = (FC + (PC * 0.5)) / TL$ , where FC, PC and TL denote the numbers of observed fully calcified, partially calcified and total larvae from each sample, respectively (Fig. 1). In order to determine the statistical significance of difference in timing of calcification onset and rate of shell formation, calcification index data from hour 10–18 (period over which calcification occurred) was arcsine transformed and fit with a generalized linear model (GLM) with the formula:  $CI \sim \beta_{Time} * \beta_{Trit} * \beta_{Exp}$ , where  $\beta_{Time}$ ,  $\beta_{Trit}$  and  $\beta_{Exp}$  represent estimates for time (hours post fertilization), treatment (seawater pCO<sub>2</sub>) and experiment number, respectively. The “full” model (with all interactions) was subsequently submitted to backwards stepwise AIC selection (“stepAIC” in R) to generate a final optimal fit model with the formula  $CI \sim \beta_{Time} + \beta_{Trit} + \beta_{Exp} + (\beta_{Time} \times \beta_{Trit}) + (\beta_{Time} \times \beta_{Exp})$ . Associated *p*-values for each parameter

were calculated on Type I sum of squares based on a chi-squared distribution.

### Transcriptomic analyses

RNA later-fixed larvae were transported to the Sven Lovén Centre for Marine Sciences, Tjärnö, Sweden, where total RNA was extracted using a Qiagen RNeasy kit (Qiagen, Hilden, Germany), following the standard protocol. RNA concentrations were measured using a QuBit 2.0 RNA fluorometric assay (Thermo Fisher Scientific, Waltham, MA, USA) and integrity was assessed with a 3-(N-morpholino)propanesulfonic acid (MOPS) denaturing agarose gel. Complementary DNA (cDNA) libraries were prepared using the Illumina TruSeq v2 mRNA sample prep kit (Illumina, San Diego, CA, USA), following a standard protocol. Briefly, mRNA was isolated with poly-A selection, followed by cDNA synthesis, Illumina standard index adapter ligation and a brief PCR reaction. Concentrations of the cDNA libraries were measured using a QuBit DNA High-sensitivity assay (Thermo Fisher Scientific, Waltham, MA, USA) and fragment length distributions were assessed using an Agilent TapeStation with a D1000 tape (Agilent, Santa Clara, CA, USA). cDNA libraries were multiplexed by equimolar pooling (6 or 7 samples/pool), and were then sent to the Swedish National Genomics Infrastructure’s SNP & SEQ platform in Uppsala for Illumina HiSeq 2500 sequencing (8 lanes; 50 bp Single-End sequencing; Illumina, San Diego, CA, USA).

Bioinformatic analyses were performed on the University of Gothenburg computer cluster Albiorix (<http://albiorix.bioenv.gu.se>) where raw reads were trimmed of low-quality (Q < 20) ends and Illumina adapter sequences were removed using the fastx toolkit ([http://hannonlab.cshl.edu/fastx\\_toolkit/](http://hannonlab.cshl.edu/fastx_toolkit/)). All reads were mapped to the oyster genome v9 coding regions (available at “[ftp://ensemblgenomes.org/pub/metazoa/release-34/fasta/crassostrea\\_gigas/cds/Crassostrea\\_gigas.G-CA\\_000297895.1.31.cds.all.fa.gz](ftp://ensemblgenomes.org/pub/metazoa/release-34/fasta/crassostrea_gigas/cds/Crassostrea_gigas.G-CA_000297895.1.31.cds.all.fa.gz)”) using the Burrows-Wheeler Aligner (bwa) ([http://bio-bwa.sourceforge.net/\[66\]](http://bio-bwa.sourceforge.net/[66])), allowing for four mismatches, after which count data for each contig was extracted using a custom script (all scripts available at [https://github.com/DeWitP/Bioinformatic\\_Pipelines/tree/master/RNA-Seq\\_materials/scripts](https://github.com/DeWitP/Bioinformatic_Pipelines/tree/master/RNA-Seq_materials/scripts)). Only reads mapping uniquely to one genome contig were considered. Duplicates were initially removed from alignment files in order to assess the effects of duplicate reads. After determining that the proportion of duplicates was low and correlated to sequencing depth, we decided to keep the duplicate reads, as is customary. To be able to compare count data across samples with differences in sequencing depth, the count data were scale-normalized using the estimateSizeFactors function in the DESeq package [67] in R.



Scale-normalized count data were then analysed in R independently for each replicate experiment ( $n = 2$ ) in several different ways. Firstly, transcripts with low variance ( $< 1$ ) and with low counts (total counts  $< 10$ ) were filtered out, after which exponential curves were fitted to the count data, keeping only transcripts showing increased expression with time in the ambient treatment. This was done in order to filter out transcripts with peak expression levels in the newly fertilized eggs or prior to the onset of shell formation, as well as transcripts showing no variation in expression over the experimental period.

With the filtered datasets, we then fitted log-linear curves of the type:  $\log(y) = \beta_0 + \beta_1 \text{time} + \beta_2 \text{treatment} + \beta_3(\text{time} * \text{treatment})$  to the transcripts of larvae from the ambient and low  $\Omega_{\text{ARAG}}$  treatments, while forcing expression to be 0 at time point 0 (time of fertilization) (i.e.  $\beta_0 = 0$ ), and keeping transcripts with a significant ( $p < 0.05$ ) time\*treatment interaction after a Benjamini-Hochberg false-discovery rate correction (Fig. 2). The lists of significant transcripts were annotated when possible using the oyster genome annotation and compared across replicate experiments for overlap (Additional file 5: Figure S2).

Transcripts with significant effects of time\*treatment in both replicated experiments were examined for non-random distributions of function using a functional enrichment Gene Score Resampling (GSR) analysis in ErmineJ [68]. Briefly, GSR analysis examines the distribution of functions (given by the Gene Ontology (GO) annotation from Zhang et al. [10]) in the transcripts of interest, and compares that to what would be expected from a random draw of all the gene models in the genome. Significant deviations from the expected random distribution can then be interpreted as evidence for a biological importance of function with regard to experimental treatments. The advantage of this approach is that it can identify patterns of biological relevance in lists of transcripts, while overcoming the issue of noise at transcript-level in gene expression data.

The filtered datasets were also analysed for gene regulatory co-expression clusters, using a weighted gene correlation network analysis (WGCNA package in R [69]) with Pearson correlation scores. This type of analysis is useful in that it summarizes all of the variation in a large dataset into correlation clusters, assigns transcripts to the different clusters and allows for visual examination for clusters showing expression patterns of interest. In order to do this, counts were first normalized across transcripts by dividing by the mean count level at 14 h post fertilization in the ambient treatment (point chosen arbitrarily as mean counts were never 0). Transcripts belonging to clusters showing different expression profiles between ambient and low  $\Omega_{\text{ARAG}}$  treatments were extracted and analysed for GO enrichment in ErmineJ as described above.

## Additional files

**Additional file 1: Table S1.** Aragonite saturation states in ambient and low aragonite saturation state treatments in the two replicate experiments. (PDF 516 kb)

**Additional file 2: Table S2.** Results of Type I sum of squares for GLM modelling the response variable Calcification Index (CI) as a function of the independent variables  $\beta_{\text{time}}$ ,  $\beta_{\text{Tt}}$ , and  $\beta_{\text{Exp}}$  representing time (hours post fertilization), seawater treatment and experiment number, respectively, along with two 2-way interactions between time and treatment ( $\beta_{\text{time}} \times \beta_{\text{Tt}}$ ) and time and experiment ( $\beta_{\text{time}} \times \beta_{\text{Exp}}$ ). All parameters are deemed significant on a  $p = 0.05$  threshold. (DOCX 13 kb)

**Additional file 3: Table S3.** Quality statistics for all samples. Sample names include: Time post-fertilization (in hours), treatment (C = ambient, T = low  $\Omega_{\text{ARAG}}$ ), and Replicate number (A or B). (XLSX 12 kb)

**Additional file 4: Figure S1.** Mapping statistics for all samples, with and without duplicate reads. Sample names include: Time post-fertilization (in hours), treatment (C = ambient, T = low  $\Omega_{\text{ARAG}}$ ), and Replicate number (A or B). (XLSX 17 kb)

**Additional file 5: Figure S2.** Expression of transcripts with significant time x treatment effect in both replicate experiments. (PDF 1564 kb)

**Additional file 6: Figure S3.** Summary of functional enrichment results of the WGCNA clusters showing differences in expression across treatments ("blue" clusters in Fig. 3), as well as the 55 transcripts showing significant time\*treatment effects (Fig. 2). (XLSX 12 kb)

**Additional file 7: Figure S4.** Relative (to maximum) expression of (A) *cgj-smad1/5/8* and (B) *cgj-smad4* in the two replicate experiments. (PDF 576 kb)

## Abbreviations

AIC: Akaike Information Criterion; BMP: bone-morphogenetic-protein; bwa: Burrows-Wheeler aligner;  $\text{CaCO}_3$ : Calcium carbonate; cDNA: complementary DNA; cGMP: cyclic guanosine monophosphate; CI: Calcification index;  $\text{CO}_2$ : Carbon dioxide; CPL: Cross-polarized light; FDR: False discovery rate; GLM: Generalized linear model; GO: Gene ontology; GSR: Gene score resampling; HMSC: Hatfield Marine Sciences Center; MOPS: 3-(N-morpholino)propanesulfonic acid; OA: Ocean Acidification;  $p\text{CO}_2$ : partial pressure of carbon dioxide in seawater;  $\text{TCO}_2$ : total dissolved carbon dioxide; TGF: transforming growth factor; US: United States of America; WGCNA: Weighted Gene Correlation Network Analysis;  $\Omega_{\text{ARAG}}$ : Aragonite saturation state

## Acknowledgements

The authors would like to acknowledge support from Science for Life Laboratory, the National Genomics Infrastructure, NGI, and UPPMAX for providing assistance in massive parallel sequencing and computational infrastructure. We would also like to thank NBIS Bioinformatics support for assistance in data analysis. All bioinformatics analyses were run on the Albiorex computer cluster (<http://albiorex.bioenv.gu.se/>) at the Department of Marine Sciences, University of Gothenburg. In addition, we acknowledge Burke Hales, the College of Earth, Ocean and Atmospheric Sciences at Oregon State University, for seawater carbonate chemistry analysis, the advice from George Waldbusser, College of Earth, Ocean and Atmospheric Sciences, OSU in adjusting carbonate chemistry parameters of the seawater treatments, and the Oregon Oyster Company for providing oysters. Travel support was provided by the Lavern Weber Fellowship for visiting scientists to the Hatfield Marine Science Center, Oregon State University. Support for sequencing was provided also by the Magnus Bergwall foundation, Helge Ax:son Johnson foundation, Adlerbertska foundation, Långmanska foundation and the Royal Society of Arts and Sciences in Sweden. PDW was supported by the Marcus and Amalia Wallenberg foundation. AV was supported by CACHE, a Marie Curie Initial Training Network (ITN) funded by the People Programme (Marie Curie Actions) of the European Union's Seventh Framework Programme FP7/2007-2013/ under REA grant agreement n° [605051]13. ED was partially funded by USDA-ARS (CRIS Project Number 2072-31000-004-00D).

This report was prepared by Oregon Sea Grant under award number NA14OAR4170064 (project number R/SAQ-20) from the National Oceanic and Atmospheric Administration (NOAA), U.S. Department of Commerce,



and by appropriations made by the Oregon State Legislature. The statements, findings, conclusions, and recommendations are those of the authors and do not necessarily reflect the views of these funders.

#### Funding

All experimental work was funded by Oregon Sea Grant under award number NA14OAR4170064. All DNA sequencing was funded by the Magnus Bergwall foundation, Helge Ax:son Johnson foundation, Adlerbertska foundation, Långmanska foundation and the Royal Society of Arts and Sciences in Sweden. The funding bodies had no role in the design of the study and collection, analysis, and interpretation of data, or in writing the manuscript.

#### Availability of data and materials

All raw Illumina data was submitted to the NCBI Short Read Sequence Archive (SRA) under the BioProject accession code PRJNA306408. Raw read count data with and without duplicate reads and the full output of the functional enrichment tests have been deposited on the Dryad digital repository <http://datadryad.org>, DOI:<https://doi.org/10.5061/dryad.8m5v4>. All custom scripts are available on GitHub at: [https://github.com/DeWitP/Bioinformatic\\_Pipelines/tree/master/RNA-Seq\\_materials/scripts](https://github.com/DeWitP/Bioinformatic_Pipelines/tree/master/RNA-Seq_materials/scripts).

#### Authors' contributions

PDW, ED, AV & CJL are responsible for project planning. PDW, ED & AV performed all experimental work. PDW is responsible for all transcriptomic analyses and manuscript writing. ED is responsible for water chemistry and microscopic analyses. CJL is responsible for funding and providing oyster growing facilities. PDW, ED, AV & CJL all contributed to the final version of the manuscript. All authors read and approved the final manuscript.

#### Ethics approval and consent to participate

Not applicable as oysters do not require ethical permits. No permits were required for the collection of oysters from the hatchery.

#### Consent for publication

Not applicable.

#### Competing interests

The authors declare that they have no competing interests.

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

#### Author details

<sup>1</sup>Department of Marine Sciences, University of Gothenburg, Strömstad, Sweden. <sup>2</sup>Department of Fisheries and Wildlife and Coastal Oregon Marine Experiment Station, Oregon State University, Newport, Oregon, USA. <sup>3</sup>Department of Biological and Environmental Sciences, University of Gothenburg, Fiskebäckskil, Sweden.

Received: 14 November 2017 Accepted: 31 January 2018

Published online: 22 February 2018

#### References

- Mount AS, Wheeler AP, Paradkar RP, Snider D. Hemocyte-mediated shell mineralization in the eastern oyster. *Science*. 2004;304:297–300.
- Marin F, Luquet G, Marie B, Medakovic D. Molluscan shell proteins: primary structure, origin, and evolution. *Curr Top Dev Biol*. 2007;80:209–76.
- Stenzel HB. Aragonite and calcite as constituents of adult oyster shells. *Science*. 1963;142:232–3.
- Stenzel HB. Oysters: composition of the larval shell. *Science*. 1964;145:155–6.
- Waldbusser GG, Gray MW, Hales B, Langdon CJ, Haley BA, Gimenez I, et al. Slow shell building, a possible trait for resistance to the effects of acute ocean acidification. *Limnol Oceanogr*. 2016;61:1969–83.
- Kniprath E. Ontogeny of the molluscan shell field: a review. *Zool Scr*. 1981;10:61–79.
- Waldbusser GG, Brunner EL, Haley BA, Hales B, Langdon CJ, Prah FG. A developmental and energetic basis linking larval oyster shell formation to acidification sensitivity. *Geophys Res Lett*. 2013;40:2171–6.
- Weiss IM, Tuross N, Addadi L, Weiner S. Mollusc larval shell formation: amorphous calcium carbonate is a precursor phase for aragonite. *J Exp Zool*. 2002;293:478–91.
- Huan P, Wang H, Dong B, Liu B. Identification of differentially expressed proteins involved in the early larval development of the Pacific oyster *Crassostrea Gigas*. *J Proteome*. 2012;75:3855–65.
- Zhang G, Fang X, Guo X, Li L, Luo R, Xu F, et al. The oyster genome reveals stress adaptation and complexity of shell formation. *Nature*. 2012;490:49–54.
- Dineshram R, Chandramouli K, Ko GWK, Zhang H, Qian P-Y, Ravasi T, et al. Quantitative analysis of oyster larval proteome provides new insights into the effects of multiple climate change stressors. *Glob Chang Biol*. 2016;22:2054–68.
- Epelboin Y, Quintric L, Guévelou E, Boudry P, Pichereau V, Corporeau C. The kinome of Pacific oyster *Crassostrea Gigas*, its expression during development and in response to environmental factors. *PLoS One*. 2016;11:e0155435.
- Li H, Zhang B, Huang G, Liu B, Fan S, Zhang D, et al. Differential gene expression during larval metamorphic development in the pearl oyster, *Pinctada Fucata*, based on transcriptome analysis. *Int J Genomics*. 2016;2016:2895303.
- Wang RN, Green J, Wang Z, Deng Y, Qiao M, Peabody M, et al. Bone morphogenetic protein (BMP) signaling in development and human diseases. *Genes Dis*. 2014;1:87–105.
- Kin K, Kakoi S, Wada H. A novel role for dpp in the shaping of bivalve shells revealed in a conserved molluscan developmental program. *Dev Biol*. 2009;329:152–66.
- Liu G, Huan P, Liu B. Cloning and expression patterns of two Smad genes during embryonic development and shell formation of the Pacific oyster *Crassostrea Gigas*. *Chin J Oceanol Limnol*. 2014;32:1224–31.
- Liu G, Huan P, Liu B. A GATA2/3 gene potentially involved in larval shell formation of the Pacific oyster *Crassostrea Gigas*. *Dev Genes Evol*. 2015;225:253–7.
- Zhang C, Xie L, Huang J, Chen L, Zhang R. A novel putative tyrosinase involved in periostracum formation from the pearl oyster (*Pinctada Fucata*). *Biochem Biophys Res Commun*. 2006;342:632–9.
- Huan P, Liu G, Wang H, Liu B. Identification of a tyrosinase gene potentially involved in early larval shell biogenesis of the Pacific oyster *Crassostrea Gigas*. *Dev Genes Evol*. 2013;223:389–94.
- Sakurai K, Chen J, Kefalov VJ. Role of guanylyl cyclase modulation in mouse cone phototransduction. *J Neurosci*. 2011;31:7991–8000.
- Wang X, Wang M, Jia Z, Wang H, Jiang S, Chen H, et al. Ocean acidification stimulates alkali signal pathway: a bicarbonate sensing soluble adenylyl cyclase from oyster *Crassostrea Gigas* mediates physiological changes induced by CO<sub>2</sub> exposure. *Aquat Toxicol*. 2016;181:124–35.
- Pan T-CF, Applebaum SL, Manahan DT. Experimental ocean acidification alters the allocation of metabolic energy. *Proc Natl Acad Sci U S A*. 2015;112:4696–701.
- Pan TCF, Applebaum SL, Lentz BA, Manahan DT. Predicting phenotypic variation in growth and metabolism of marine invertebrate larvae. *J Exp Mar Biol Ecol*. 2016;483:64–73.
- Feely RA, Sabine CL, Hernandez-Ayon JM, Janson D, Hales B. Evidence for upwelling of corrosive "acidified" water onto the continental shelf. *Science*. 2008;320:1490–2.
- Barton A, Hales B, Waldbusser GG, Langdon C, Feely RA. The Pacific oyster, *Crassostrea Gigas*, shows negative correlation to naturally elevated carbon dioxide levels: implications for near-term ocean acidification effects. *Limnol Oceanogr*. 2012;57:698–710.
- Barton A, Waldbusser GG, Feely RA, Weisberg SB, Newton JA, Hales B, et al. Impacts of coastal acidification on the Pacific northwest shellfish industry and adaptation strategies implemented in response. *Oceanography*. 2015;28:146–59.
- Kurihara H, Kato S, Ishimatsu A. Effects of increased seawater pCO<sub>2</sub> on early development of the oyster *Crassostrea Gigas*. *Aquat Biol*. 2007;1:91–8.
- Parker LM, Ross PM, O'Connor WA. The effect of ocean acidification and temperature on the fertilization and embryonic development of the Sydney rock oyster *Saccostrea Glomerata* (Gould 1850). *Glob Chang Biol*. 2009;15:2123–36.
- Welladsen HM, Southgate PC, Heimann K. The effects of exposure to near-future levels of ocean acidification on shell characteristics of *Pinctada Fucata* (Bivalvia: Pteriidae). *Molluscan Research*. 2010;30:125–30.

30. Miller AW, Reynolds AC, Sobrino C, Riedel GF. Shellfish face uncertain future in high CO<sub>2</sub> world: influence of acidification on oyster larvae calcification and growth in estuaries. *PLoS One*. 2009;4:e5661.
31. Talmage SC, Gobler CJ. The effects of elevated carbon dioxide concentrations on the metamorphosis, size, and survival of larval hard clams (*Mercenaria mercenaria*), bay scallops (*Argopecten irradians*), and eastern oysters (*Crassostrea virginica*). *Limnol Oceanogr*. 2009;54:2072–80.
32. Frieder CA, Applebaum SL, Pan T-C F, Hedgecock D, Manahan DT. Metabolic cost of calcification in bivalve larvae under experimental ocean acidification. *ICES J Mar Sci*. 2017;74:941–54.
33. Waldbusser GG, Hales B, Langdon CJ, Haley BA, Schrader P, Brunner EL, et al. Saturation-state sensitivity of marine bivalve larvae to ocean acidification. *Nat Clim Chang*. 2015;5:273–80.
34. Waldbusser GG, Hales B, Haley BA. Calcium carbonate saturation state: on myths and this or that stories. *ICES J Mar Sci*. 2016;73:563–8.
35. Cyronak T, Schulz KG, Jokiel PL. The omega myth : what really drives lower calcification rates in an acidifying ocean. *ICES J Mar Sci*. 2016;73:558–62.
36. Apweiler R, Attwood TK, Bairoch A, Bateman E, Birney E, Biswas M, et al. The InterPro database, an integrated documentation resource for protein families, domains and functional sites. *Nucleic Acids Res*. 2001;29:37–40.
37. Timmins-Schiffman E, O'Donnell MJ, Friedman CS, Roberts SB. Elevated pCO<sub>2</sub> causes developmental delay in early larval Pacific oysters, *Crassostrea gigas*. *Mar Biol*. 2013;160:1973–82.
38. Todgham AE, Hoffman GE. Transcriptomic response of sea urchin larvae *Strongylocentrotus purpuratus* to CO<sub>2</sub>-driven seawater acidification. *J Exp Biol*. 2009;212:2579–94.
39. Berezney R. Effect of protease inhibitors on matrix proteins and the association of replicating DNA. *Exp Cell Res*. 1979;123:411–4.
40. Fessler JH, Kramerova I, Kramerov A, Chen Y, Fessler LI. Papilin, a novel component of basement membranes, in relation to ADAMTS metalloproteases and ECM development. *Int J Biochem Cell Biol*. 2004;36:1079–84.
41. Falini G, Albeck S, Weiner S, Addadi L. Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science*. 1996;271:67–9.
42. Miyamoto H, Miyoshi F, Kohno J. The carbonic anhydrase domain protein nacrein is expressed in the epithelial cells of the mantle and acts as a negative regulator in calcification in the mollusk *Pinctada fucata*. *Zool Sci*. 2005;22:311–5.
43. Weiss IM, Schönitzer V, Eichner N, Sumper M. The chitin synthase involved in marine bivalve mollusk shell formation contains a myosin domain. *FEBS Lett*. 2006;580:1846–52.
44. Marie B, Zanella-Cléon I, Guichard N, Becchi M, Marin F. Novel proteins from the calcifying shell matrix of the Pacific oyster *Crassostrea gigas*. *Mar Biotechnol*. 2011;13:1159–68.
45. Medakovic D. Carbonic anhydrase activity and biomineralization process in embryos, larvae and adult blue mussels *Mytilus edulis* L. *Helgol Mar Res*. 2000;54:1–6.
46. Williams TM, Lisanti MP. The caveolin proteins. *Genome Biol*. 2004;5:214.
47. Florek M, Bauer N, Janich P, Wilsch-Braeuninger M, Fargeas CA, Marzesco A-M, et al. Prominin-2 is a cholesterol-binding protein associated with apical and basolateral plasmalemmal protrusions in polarized epithelial cells and released into urine. *Cell Tissue Res*. 2007;328:31–47.
48. Xie Y-J, Zhou L, Jiang N, Zhang N, Zou N, Zhou L, et al. Essential roles of leucine-rich glioma inactivated 1 in the development of embryonic and postnatal cerebellum. *Sci Rep*. 2015;5:7827.
49. Hall JD, Lloyd PE. Involvement of pedal peptide in locomotion in *Aplysia*: modulation of foot muscle contractions. *J Neurobiol*. 1990;21:858–68.
50. Bridgman J, Wilder J, Hollocher H, Johnson A. All in the family: evolutionary and functional relationships among death receptors. *Cell Death Differ*. 2003;10:19–25.
51. Boettiger A, Ermentrout B, Oster G. The neural origins of shell structure and pattern in aquatic mollusks. *Proc Natl Acad Sci U S A*. 2009;106:6837–42.
52. Mittl PR, Di Marco S, Fendrich G, Pohlig G, Heim J, Sommerhoff C, et al. A new structural class of serine protease inhibitors revealed by the structure of the hirustasin–kallikrein complex. *Structure*. 1997;5:253–64.
53. Hooper JD, Clements JA, Quigley JP, Antalis TM. Type II transmembrane serine proteases. Insights into an emerging class of cell surface proteolytic enzymes. *J Biol Chem*. 2001;276:857–60.
54. Marxen JC, Witten PE, Finke D, Reelsen O, Rezgaoui M, Becker W. A light- and electron-microscopic study of enzymes in the embryonic shell-forming tissue of the freshwater snail, *Biomphalaria glabrata*. *Invertebr Biol*. 2003;122:313–25.
55. Yang B, Pu F, Li L, You W, Ke C, Feng D. Functional analysis of a tyrosinase gene involved in early larval shell biogenesis in *Crassostrea angulata* and its response to ocean acidification. *Comp Biochem Physiol B Biochem Mol Biol*. 2017;206:8–15.
56. Wang X, Li L, Zhu Y, Du Y, Song X, Chen Y, et al. Oyster shell proteins originate from multiple organs and their probable transport pathway to the shell formation front. *PLoS One*. 2013;8:e66522.
57. Hüning AK, Lange SM, Ramesh K, Jacob DE, Jackson DJ, Panknin U, et al. A shell regeneration assay to identify biomineralization candidate genes in mytilid mussels. *Mar Genomics*. 2016;27:57–67.
58. Yu X, Yu H, Kong L. Molecular cloning and differential expression in tissues of a tyrosinase gene in the Pacific oyster *Crassostrea gigas*. *Mol Biol Rep*. 2014;41:5403–11.
59. Yue F, Zhou Z, Wang L, Wang M, Song L. A conserved zinc finger transcription factor GATA involving in the hemocyte production of scallop *Chlamys farreri*. *Fish Shellfish Immunol*. 2014;39:125–35.
60. Shimizu K, Sarashina I, Kagi H, Endo K. Possible functions of Dpp in gastropod shell formation and shell coiling. *Dev Genes Evol*. 2011;221:59–68.
61. Parker LM, Ross PM, O'Connor WA, Borysko L, Raftos DA, Pörtner H-O. Adult exposure influences offspring response to ocean acidification in oysters. *Glob Chang Biol*. 2012;18:82–92.
62. Griffith AW, Gobler CJ. Transgenerational exposure of North Atlantic bivalves to ocean acidification renders offspring more vulnerable to low pH and additional stressors. *Sci Rep*. 2017;7:11394.
63. Stefansson ES, Langdon CJ, Pargee SM, Blunt SM, Gage SJ, Stubblefield WA. Acute effects of non-weathered and weathered crude oil and dispersant associated with the Deepwater horizon incident on the development of marine bivalve and echinoderm larvae. *Environ Toxicol Chem*. 2016;35:2016–28.
64. Hales B, Takahashi T, Bandstra L. Atmospheric CO<sub>2</sub> uptake by a coastal upwelling system. *Glob Biogeochem Cycles*. 2005;19:GB1009.
65. Bandstra L, Hales B, Takahashi T. High-frequency measurements of total CO<sub>2</sub> : method development and first oceanographic observations. *Mar Chem*. 2006;100:24–38.
66. Li H, Durbin R. Fast and accurate short read alignment with burrows-wheeler transform. *Bioinformatics*. 2009;25:1754–60.
67. Anders S, Huber W. Differential expression analysis for sequence count data. *Genome Biol*. 2010;11:R106.
68. Gillis J, Mistry M, Pavlidis P. Gene function analysis in complex data sets using *ErmineJ*. *Nat Protoc*. 2010;5:1148–59.
69. Langfelder P, Horvath S. WGCNA: an R package for weighted correlation network analysis. *BMC Bioinformatics*. 2008;9:559.

Submit your next manuscript to BioMed Central and we will help you at every step:

- We accept pre-submission inquiries
- Our selector tool helps you to find the most relevant journal
- We provide round the clock customer support
- Convenient online submission
- Thorough peer review
- Inclusion in PubMed and all major indexing services
- Maximum visibility for your research

Submit your manuscript at  
[www.biomedcentral.com/submit](http://www.biomedcentral.com/submit)

