The Molluscan Shell Secretome: Unlocking Calcium Pathways in a Changing World

Victoria Anne Sleight

Submitted for the degree of Doctor of Philosophy

Heriot-Watt University

School of Life Sciences

September 2017

Abstract

How do molluscs build their shells? Despite hundreds of years of human fascination, the processes underpinning mollusc shell production are still considered a black box. We know molluscs can alter their shell thickness in response to environmental factors, but we do not have a mechanistic understanding of how the shell is produced and regulated. In this thesis I used a combination of methodologies - from traditional histology, to shell damage-repair experiments and 'omics technologies - to better understand the molecular mechanisms which control shell secretion in two species, the Antarctic clam *Laternula elliptica* and the temperate blunt-gaper clam *Mya truncata.* The integration of different methods was particularly useful for assigning putative biomineralisation functions to genes with no previous annotation. Each chapter of this thesis found reoccurring evidence for the involvement of vesicles in biomineralisation and for the duplication and subfunctionalisation of *tyrosinase* paralogues. Shell damage-repair experiments revealed biomineralisation in *L. elliptica* was variable, transcriptionally dynamic, significantly affected by age and inherently entwined with immune processes. The high amount of transcriptional variation across 78 individual animals was captured in a single mantle regulatory gene network, which was used to predict the regulation of "classic" biomineralisation genes, and identify novel biomineralisation genes. There were some general shared patterns in the molecular control of biomineralisation between the two species investigated in this thesis, but overall, the comparative work in this thesis, coupled to the growing body of literature on the evolution of molluscan biomineralisation, suggests that biomineralisation mechanisms are surprisingly divergent.

Dedication

To my loving family,

who have always encouraged my curiosity and adventures,

and instilled in me the courage and confidence

to pursue a career in science.

Acknowledgements

The work in this thesis would not have been possible without help, support and advice from many colleagues.

Firstly, animals were collected by dedicated members of the SCUBA diving teams at Rothera in Antarctica, and the UK National Facility for Scientific Diving in Oban, Scotland. Thank you to everybody involved in the great clam hunt saga that ensued during the last three years. Animal husbandry during holding, acclimation and transportation was also carried out by colleagues at Rothera, Cambridge and the Scottish Association for Marine Science. Thank you to everybody who has been involved in looking after the clams.

Thank you to Deborah Power and Daniel Jackson for hosting me in their labs and for training and advice on *in situ* hybridisations. Thank you to Arul Marie, Sophie Berland Benjamin Marie and Jaison Arivalagan for providing the shell proteomes for both species in this thesis. Thank you to Michael Thorne for his help with various bioinformatic analyses in this thesis. Thank you to Philipp Antczak for hosting me at the University of Liverpool Computational Biology Facility and for training and guidance on differential gene expression analysis and regulatory gene networks. Thank you to Jeremy Skepper and Lyn Carter at the Cambridge Advanced Imaging Centre for sample preparation and processing for electron microscopy. Thank you to Liz Harper for advice on shell microstructure, mantle anatomy and mollusc evolution. Thank you to all the members of the CACHE network for stimulating conversations and being a fantastic network of mollusc shell colleagues.

Finally, I am extremely grateful to my supervisors, Melody Clark, Lloyd Peck, Liz Dyrynda and Val Smith for their unwavering support and enthusiasm. Thank you for the scientific guidance and for making me think, and a special thank you to Melody for all the chocolate biscuits and encouragement

Declaration Statement

ACADEMIC REGISTRY

Research Thesis Submission

Name*:*	Victoria Anne Sleight		
School/PGI:	was School of Life Sciences, now School of Energy, Geoscience, Infrastructure & Society (EGIS)		
Version: *(i.e. First, Resubmission, Final)*	Final	Degree Sought (Award **and** Subject area)	Doctor of Philosophy – Marine Biology

Declaration

In accordance with the appropriate regulations I hereby submit my thesis and I declare that:

1) the thesis embodies the results of my own work and has been composed by myself

2) where appropriate, I have made acknowledgement of the work of others and have made reference to work carried out in collaboration with other persons

3) the thesis is the correct version of the thesis for submission and is the same version as any electronic versions submitted*.

4) my thesis for the award referred to, deposited in the Heriot-Watt University Library, should be made available for loan or photocopying and be available via the Institutional Repository, subject to such conditions as the Librarian may require

5) I understand that as a student of the University I am required to abide by the Regulations of the University and to conform to its discipline.

* *Please note that it is the responsibility of the candidate to ensure that the correct version of the thesis is submitted.*

Signature of Candidate*:*		Date:	

Submission

Submitted By *(name in capitals):*	
Signature of Individual Submitting:	
Date Submitted:	

For Completion in the Student Service Centre (SSC)

Received in the SSC by *(name in capitals):*			
Method of Submission *(Handed in to SSC; posted through internal/external mail):*			
E-thesis Submitted (mandatory for final theses)			
Signature:		Date:	

Contents

List of Figures

List of Tables

Chapter 1 INTRODUCTION – THE MOLECULAR MECHANISMS UNDERPINNING SHELL SECRETION: A METHODOLOGICAL REVIEW

1.1 Introduction

From the ancient nautilus to the humble clam, thousands of mollusc species share a common feature – a shell. For millennia humankind has been fascinated by shells, their mathematical architecture and alluring texture (Marin and Luquet 2004). In the present day, interest in mollusc shell continues. The industries of biotechnology and materials science study biomineralisation for multiple reasons: low energy and temperature requirements for producing solid calcium carbonate structures from sea water, the high strength of shell compared to other composite materials and the prospect of engineering material, such as pearl, without the need for a live animal (Meldrum 2003). On the other hand, environmental biologists are concerned about the ecological fate of calcified structures as the oceans become more acidic and aim to understand the calcification pathway in order to predict how ecosystems might alter with climate change (Gazeau et al. 2013). Mollusc species are also important in aquaculture and represent a growing food source around the world. Aquaculture is the fastest growing food production industry in the world with mollusc aquaculture making-up approximately 22% (by volume) of global production (Corbin 2007). Dominant aquaculture genera (*Mytilus*, *Crassostrea* and *Pinctada)* have therefore received primary research focus over the last 50 years, resulting in basic physiological understanding of shell growth (Bayne 1976, Palmer 1992, Harper 1997, Day et al. 2000).

The molecular methods molluscs use to extract, transport and subsequently lay-down minerals however, are largely unknown. Despite thousands of years of curiosity across multiple disciplines, the mechanisms that molluscs use to build their shells - on a molecular level - remain unclear.

The molluscan shell is the result of a controlled biological process producing a composite biomaterial containing 95-99% calcium carbonate ($CaCO_3$) and 1-5% organic matrix. A shell provides: critical protection to the soft-bodied animal within, support for internal anatomy, and a means by which the animal can seal itself off from the external environment (Figure 1-1A. Wilbur and Saleuddin, 1983, Vermeij 2005). The organ primarily responsible for shell secretion in molluscs is the mantle; although the organ does have other functions not discussed here (Galtsoff 1964, Bayne 1976, Wilbur and Saleuddin 1983). The mantle can be described as a soft, fleshy cloak which surrounds

the entire body cavity and inner organs (Figure 1-1A). Molluscs construct their shell in a process termed biomineralisation; minerals are transported across the mantle and laid down as organised crystals onto an organic matrix. A large portion of early shell literature focussed on characterising shell structure (de Paula and Silveira 2009).

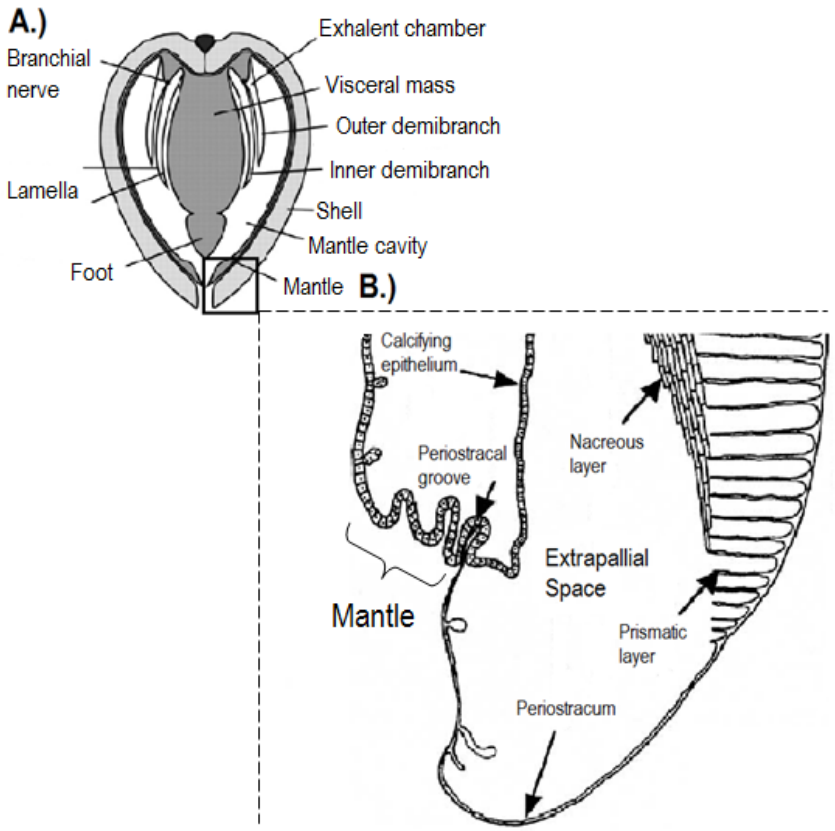

Figure 1-1. Schematic diagram of bivalve anatomy, adapted from Gainey, (2007). A.) Sagittal cross-section through hypothetical bivalve, solid box and dashed lines indicates region zoomed in. B.) Cross-section through hypothetical mantle-shell interface.

There are two major phases of calcium carbonate in mollusc shell, calcite and aragonite. The classic shell model states that shell is organized into three layers; an outermost layer is usually prismatic and made of either aragonite or calcite, a central crossed-lamellar layer made of obliquely crossed crystals is frequently present and an inner most nacreous layer made of aragonite represents the typical, desirable, smooth texture on the inside of shells (Figure 1-1B; Marin & Luquet, 2004). The outer calcite prisms have high crack propagation and puncture-resistance and hence act as the primary barrier (Harper 2000). Whereas, the internal aragonite nacreous layer has extreme high fracture resistance,

which is thought to dissipate energy and stop cracks from forming (Jackson et al. 1988). Often an uncalcified, proteinaceous periostracum layer covers the outside of the shell which provides further protection from the environment (Figure 1-1B; Harper,1997; Marin & Luquet, 2004). The described layers and diagram (Figure 1-1B) represent a basic shell model however, there are diverse combinations of structures evident across the mollusc genera (Trueman et al. 1988, de Paula and Silveira 2009).

For the past 60 years molecular biologists have facilitated tantalising glimpses into the mechanisms of biomineralisation. By studying molluscan genes, the patterns in which they are expressed and the proteins they code for, some knowledge gaps have been filled. Early work focussed on dissolving shells to identify different protein fractions with chromatography (Bevelander and Benzer 1948, Bevelander 1951, Bevelander 1952). The first shell protein - nacrein - was characterised in 1996 (Miyamoto et al. 1996), and in the following decade shell genes and proteins continued to be individually characterised one-by-one (Sudo et al. 1997, Medakovic 2000, Miyamoto et al. 2005). The development of massively parallel, Next Generation Sequencing (NGS) in 2004 marked a paradigm shift in molecular biology, instead of characterizing one gene or protein at a time, analysis became global. Four years later NGS technology was first applied to non-model species (Vera et al. 2008), and in the following years the ease and cost of application has dramatically reduced. Given the right funding, technology now exists to enable entire genome sequencing of any organisms, although challenges still arise in NGS assembly and analysis. Scientists interested in mollusc biomineralisation are beginning develop databases of molecular information.

New methods and technologies are advancing our understanding of how molluscs build their shell, here I will set out the state-of-the-art understanding of the molecular control of marine mollusc shell production at the beginning of the project, March 2014. A methodological approach will be taken and discussion will focus on what has been learnt from different techniques, moving from single gene characterisations to global ‘omics techniques. This chapter will end by outlining the aims and objective of the thesis and explaining what the project will contribute to the understanding of genes and gene networks underpinning shell biomineralisation processes.

1.2 Early work on singular protein characterisations

Prior to the introduction of genetic techniques into the study of biomineralisation, biochemical work concentrated on dissolving and resolving proteins using crystallography, chromatography or electrophoresis. Initial work in the mid-nineteenth century identified an insoluble fraction of the shell matrix (Frémy 1855), which biochemical analysis a century later simply described as a "heterogeneous mixture of proteinaceous substances" (Gregoire 1967). A break-through occurred when the soluble fraction of shell was isolated. Biochemical analysis demonstrated the soluble protein fraction had multiple negative charges (polyanionic) which was thought to be important for regulating the growth of $CaCO_3$ crystals. This discovery of a soluble, polyanionic fraction, rich in aspartic acid formulated the basis of the current shell formation hypothesis. The insoluble protein matrix acts as a framework which soluble polyanionic proteins bind to. In turn, the soluble polyanionic proteins apply antagonistic forces to either nucleate or inhibit crystal growth, resulting in highly controlled regular crystal formation (Meenakshi et al. 1971, Weiner and Hood 1975). Following the development of this simple model, research began to identify and understand the functions of different proteins contained in shell matrices with the aim of characterising a more detailed and mechanistic pathway. Molecular investigations typically focussed on one gene/protein characterisation at a time using *in vivo* histochemical and immunohistochemical localisation, protein or gene sequencing and *in vitro* crystallization experiments. The following section will highlight some of the major advances provided by such methods.

In the mid-twentieth century, histochemical studies which localised minerals and enzymes provided early insights into their spatial extent and function during biomineralisation. Classic experiments were elegantly conducted and began to pose questions which are still important in the study of biomineralisation today. For example, Bevelander et al. (1948, 1951, 1952) examined calcification in five mollusc species; they used standard histological techniques to localise mucin, phosphatase and calcium (using Hoyer's thionin, Gomori and Feigl methods respectively) in growing shells. Results indicated that calcium is extracted solely from the water column, and not from the diet, as had been previously speculated. In addition they carried out shell damage/repair studies under various different conditions, such as reduced calcium. Phosphatase was present in the shell secretion region, highlighting the importance of active enzymes in

biomineralisation, as opposed to secretion being entirely passive. In the reduced calcium concentration experimental water, a protein matrix was normally secreted - but not calcified (Bevelander, 1951, 1952, Bevelander and Benzer, 1948). Gradually researchers began to recognise the importance of organic macromolecules in the formation of shell, and in particular, the importance of the organic matrix (Gregoire, 1967).

Moving into the mid-1970s, the ability to extract soluble proteins from the organic matrix marked a significant progression in the understanding of biomineralisation. The purification and analysis of shell peptides revealed a high abundance of aspartic acid residues (Weiner and Hood 1975), which were later spectrophotometrically classified as likely calcium binding domains (Cariolou and Morse 1988). Focussing on the soluble, polyanionic component of the shell matrix, *in vitro* studies initiated crystallisation of calcium carbonate in the presence of purified, extracted proteins alone (Belcher et al. 1996). This demonstrates the soluble component of shell exerts a strong control on mineral growth and reiterating its importance in biomineralisation. Moving towards the end of the 20th century, protein classification became more detailed, thanks to mass spectrometry, which allowed researchers to speculate biomineralisation models. For example, Bowen and Tang (1996) characterised "protein H" and hypothesised it acted as an anchor molecule in the matrix complex. Valine and Leucine from protein H form hydrophobic bonds with acidic, soluble proteins, which in turn control crystal growth. Specifically, Bowen & Tang (1996) suggested that the acidic proteins are likely to have a surplus of non-bonded, ionized carboxyl groups which act as a template for crystal growth along a geometric axis.

In 1996 genetics firmly made its mark in the study of biomineralisation. Advances in technology made it possible to obtain N-terminal protein sequences and use degenerate oligonucleotide probes to find corresponding genes. Miyamoto et al. (1996) characterized the first shell protein, nacrein-a, which was isolated from the soluble fraction of the nacreous layer. Nacrein-a is a multi-functional protein, acting as a calcium binding template and an enzyme. Specifically, nacrein-a contains a carbonic anhydrase domain which is able to catalyse the rapid conversion of carbon dioxide and water to bicarbonate and protons (and vice versa), as well as acting as a structural matrix protein which binds calcium (Medakovic 2000). As the name suggests, nacrein-a was initially thought to be exclusively involved in the production of the desirable nacreous layer, but

has since also been isolated from the prismatic layer (Miyashita et al. 2002). Proteins, and the genes which code for them, have continued to be characterized in the same detailed and step-wise fashion [Table 1-1; (Shen et al. 1997, Asakura et al. 2006, Nogawa et al. 2012)]. Historically, one of the major limitations to the progression of biomineralisation research was the lack of integration of such characterisations; a complex biochemical pathway was difficult to understand by studying a single gene or protein. As technology progressed, research has become more focussed on trying to understand how genes are regulated, and how proteins interact together to control the process of biomineralisation.

Although much of the early research in the field of biomineralisation addressed important questions, methodological limitations constrained progress. How far can characterising single genes and proteins take us towards understanding how molluscs build their shells? What has the large collection of characterisations contributed to our understanding and what are the overall patterns? As described by Marin & Luquet, (2004), a glance over the information we have so far amassed can pull out some patterns. Shell matrix proteins can be divided into two broad categories: insoluble structural proteins, which determine the framework of each shell layer; and soluble polyanionic proteins, which bind to the insoluble structure and determine the type of $CaCO_3$ crystal that grows. Shell matrix proteins are typically modular (made up of different independently functioning domains) and multi-functional. Aspartic acid is potentially important in molluscan calcification, specifically in the prismatic, calcite layer. The continued characterisation of novel biomineralisation genes and proteins is important and will provide the building-blocks of future understanding. To move understanding forward from describing structures, techniques which empirically address function are important.

Table 1-1. Selection of past and present single molecular characterisations of biomineralisation proteins.

Protein Name	Putative Function
Nacrein	Soluble extracellular matrix protein involved mainly in regulation of the nacreous layer, but also prismatic layer of some species (Miyamoto et al. 1996).

Lustrin A	Highly modular and likely multifunctional insoluble extracellular matrix protein which is collagen-like, involved in regulation of the nacreous layer, protection of the matrix from degradation and elasticity (Shen et al. 1997).
Perlustrin	Soluble extracellular matrix protein involved in regulation of the nacreous layer, it is similar to collagen-like protein, Lustrin A and consists of a single insulin-like growth factor binding protein N-terminal domain (Weiss et al. 2000).
Perlucin	Soluble extracellular matrix protein involved in regulation of the nacreous layer. May also have a protective/immune role in non-self-antigen recognition (Weiss et al. 2000).
Prismalin- 14	Insoluble extracellular matrix protein involved in the regulation of the prismatic layer with chitin-binding activity (Suzuki et al. 2004).
Sarco/endoplasmic reticulum Ca2+-ATPase (SERCA)	Membrane spanning protein involved in the active transport of calcium ions across membranes, mainly for general calcium homeostasis, and potentially for aiding transport of calcium to secretory mantle cells (Fan et al. 2007).
Calmodulin	Soluble, multifunctional, calcium-binding messenger protein expressed in all eukaryotic cells, specifically involved in the extracellular matrix of the prismatic and nacreous layers and in intracellular vesicles(Fang et al. 2008).
Prisilkin- 39	Soluble extracellular matrix protein involved in the regulation of the prismatic layer, binds chitin, crystal surfaces and other acidic extracellular matrix proteins (Kong et al. 2009).
Pif	Insoluble extracellular matrix protein involved in the regulation on the nacreous layer, interacts with chitin and nucleates aragonite crystals (Suzuki et al. 2009).
Nautilin- 63	Soluble extracellular matrix glycoprotein involved in the regulation of the nacreous layers, strongly interacts with chitin and finely interacts with crystal surfaces (Marie et al. 2011).

PfN23	Soluble extracellular matrix protein involved in the regulation of the nacreous layer, catalyses $CaCO_3$ precipitation (Fang et al. 2012).
MRNP34	Insoluble extracellular matrix protein involved in the regulation of the nacreous layer (Marie et al. 2012a).
Amorphous calcium carbonate binding protein (ACCBP)	Involved in the formation of amorphous calcium carbonate - $CaCO_3$ precursor (Su et al. 2013).

1.3 Singular Gene Expression Analysis

At the most basic level, to understand the role of a gene in biomineralisation it is important to know when it is switched on or off. Therefore, once the primary structure of a gene has been described and characterised, it can be further investigated using various expression analysis techniques. *In situ* hybridisations can spatially localise expression to specific cells within the mantle tissue - which is often a first step towards speculating function. In addition, quantitative PCR (qPCR) analysis provides information on the relative expression level of different genes, from which inferences about relative importance in biomineralisation can be made. For example, the up-regulation of a gene in response to shell damage would indicate the protein it encodes is likely to play a role in the process of shell repair.

In situ hybridisations have significantly contributed towards the functional understanding of the mantle. Specifically, it has facilitated the division of the mantle into specialised zones (Joubert et al. 2010). Early morphological studies recognised different mantle folds and secretary regions (Figure 1-1); mapping gene expression patterns onto such regions presents a powerful bridge into the functional understanding of how genes control biomineralisation (Table 1-2). For example, Sudo et al. (1997) demonstrated that *MS160* is expressed in the outer epithelia of the pallial mantle, a region which secretes the nacreous layer. Whereas *MS131* is expressed in the outer epithelia of the mantle edge, a region close to the periostracal groove and therefore *MS131* is more likely to be involved

in periostracal and prismatic secretion. Similarly, *prismalin-14* is expressed only in the inner side of the outer fold of the mantle, and therefore thought to be involved in secreting the prismatic layer (Suzuki et al. 2004). More recently, *in situ* hybridisations remain a valuable method, Sato et al. (2013) revisited the expression of *MS160* and *MS131* in pearl sac implant experiments. During the formation of nacre, there was a weak expression of *MSI31* and strong expression of *MSI60*. In contrast, when the prismatic shell was formed *MSI31* was strongly expressed and *MSI60* was weakly expressed. Sato et al. (2013) recognised the limitations of single gene characterisations whilst speculating about regulatory regions. They suggested that, in addition to the *MSI* genes, there must be upstream master regulatory genes determining the type (nacreous aragonite or prismatic calcite) of calcium carbonate layer secreted. In summary, *in situ* hybridisation results suggest that the mantle outer epithelia can be divided into at least two functional regions, one responsible for nacre secretion (pallial mantle) and the other prisms (mantle edge).

Quantitative PCR (qPCR) gene expression analysis provides information on the relative expression level of different genes. Researchers often use this to track the expression pattern of specific genes through development, or over experimental time in response to a manipulation. Quantifying the relative level of gene expression can suggest relative importance in a particular process. Furthermore expression in different regions can be compared, providing understanding on functional divisions (similar to *in situ* hybridisations). A handful of studies have tracked the expression of what can be called "classic" biomineralisation genes through mollusc development. Typically biomineralisation genes are significantly up-regulated when shell formation commences, further confirming their hypothesised function (Miyazaki et al. 2010, Werner et al. 2013, Fang et al. 2011). However, qPCR has revealed some biomineralisation genes, such as *MSI* and *prismalin-14*, are expressed before a calcite shell is formed (in the veliger and pre-veliger stages), they are therefore unlikely to control crystal polymorph (Miyazaki et al. 2010). Miyazaki et al. (2010) also addressed how biomineralisation gene expression patterns synchronise with environmental variables; tides, seasons and wild vs aquarium (so called "tank effects"). Biomineralisation gene expression patterns were significantly different in aquaria compared to the wild; this result doesn't provide any insight on how molluscs build their shells, it acts an important reminder to researchers on the danger of bringing animals into the laboratory. While laboratory experiments are essential to try to reduce the confounding variables of the changing environment, it also is important to

ensure that observations are true biomineralisation signal, and not a measurement of aquarium-induced stress. Ideally the development of a best practise baseline protocol should be introduced to facilitate comparisons between experiments.

Table 1-2. Summary of shell protein and gene localisation literature, dashed line separates immune protein localisations from gene expression localisations.

Species	Name of localised product	Method of localisation	Region localised	Reference
Pinctada fucata	Prisilkin-39	Immunostaining with polyclonal antibody	In sheaths around prisms	(Kong et al. 2009)
"	PfN23	"	Nacre layer of shell	(Fang et al. 2012)
Pinctada margaritifera	Nacre proteins	"	Epithelia cells of the mantle pallium and in nacreous layer of shell	(Marie et al. 2012b)
P. fucata	*Prismalin-14*	*In situ* hybridisation	Epithelia cells of the inner side of the outer fold of the mantle	(Suzuki et al. 2004)
"	*Nacrein*	"	Outer epithelia cells of the mantle	(Miyamoto et al. 2005)
"	*Pf-ALR1*	"	Epithelia cells of the outer and middle fold	(Zhou et al. 2010)
"	*Pf-Smad3*	"	"	
"	*Cluster236*	"	All mantle epithelia cells except in the bottom of the periostracal groove	(Fang et al. 2011)
	Cluster524	"	"	
	Cluster252	"	All mantle epithelia cells including the periostracal groove	

	Cluster94	"	Outer and inner epithelia cells on inner and middle fold	
	Cluster32	"	All mantle epithelia cells but predominantly in the periostracal groove	
"	*MSI31*	“	A pearl sac which yielded a nacreous pearl showed weak expression of MSI31 A pearl sac which yielded a prismatic pearl showed strong expression of MSI31	(Sato et al. 2013)
"	*MSI60*	"	A pearl sac which yielded a nacreous pearl showed strong expression of MSI60 A pearl sac which yielded a prismatic pearl showed weak expression of MSI60	
Pinctada maxima	*PM077*	"	Outer epithelia of the dorsal mantle, terminating immediately at the ventral mantle region	(Gardner et al. 2011)
	PM037	"	"	
	PM041	"	"	
	PM316	"	Inner epithelia of the mantle and the distal outer epithelia of the middle fold	
	PM317	"	Outer epithelia of the middle fold and the inner epithelia of the outer fold	
	PM315	"	Discontinuously expressed in the	

			epithelia of the inner middle fold	
	PM233	"	Inner epithelia of the outer fold	
	PM234	"	"	
	PM235	"	"	
	PM241	"	"	
	PM238	"	Mid-way along the inner epithelia of the outer fold	
	PM239	"	"	
	PM273	"	Throughout the outer epithelia of the outer fold and ventral mantle,terminating immediately at dorsal mantle region	
	PM268	"	"	
	PM280	"	"	
	PM281	"	"	
	PM265	"	Throughout the outer epithelium of the outer fold and ventral mantle, terminating immediately at the dorsal mantle region and in the inner epithelia of the outer fold.	
P. maxima	*MP10*	"	The epithelia along the mantle edge	(Marie et al. 2012b)
and	*Clp-1*	"	"	
P. margaritifera	*Fibronectin-1*	"	"	
	Pearlin	"	The epithelia along the mantle pallium	
	MRNP34	"	"	
	NUSP-1	"	"	
Patella vulgata	*SPARC*	"	In a ring around the larval shell field	(Werner et al. 2013)

"	*Periotrohpic matrix protein (PM)*	"	"	
Hyriopsis cumingii	*Perlucin*	"	Epithelial cells of the dorsal mantle pallial	(Lin et al. 2013)

qPCR results also address some major biomineralisation themes, for example, expression patterns fluctuating with tides and season demonstrates that shell secretion is not continuous. The ability for molluscs to change their shells has been observed for decades by ecologists and physiologists (Bayne 1976), but until now understanding of how they change their shell has been limited. Molluscs can control shell secretion via up or down regulation of genes in response to the environment (Miyazaki et al. 2008). Huning et al. (2013) also found gene expression patterns alter with environment, specifically hypercapnic stress. *Tyrosinase,* a gene thought to be involved in periostracum formation, was massively up-regulated in response to reduced pH, further confirming that the periostracum plays an important role in protecting the shell from the environment. Whilst the study was primarily interested in predicting responses to ocean acidification, some more general information on biomineralisation can be interpreted. For example, the outer mantle fold had higher relative transcript abundance, in 20 out of 33 genes investigated, than the inner fold and was therefore more active. Thus providing more evidence that the mantle is functionally specialised (similar to *in situ* hybridisation results).

The gene expression analyses described above have allowed focussed study on specific biomineralisation mechanisms. It has provided understanding of shell ontology, functional mapping of the mantle and responses to stress and the environment. These techniques are useful for the study of biomineralisation and have, and will continue to, provide mechanistic understanding on how molluscs build their shells. Inherent natural variation is a factor which limits the interpretation of expression analyses and presents a constant battle to determine “signal” from “noise” (McCarthy et al. 2012). It is also difficult to establish concrete cause and effect as only a handful of genes are under investigation; one method which overcomes such limitations is gene knock-out.

1.4 Empirical Investigation of Function

Inducing mutations, gene knock-out, or gene knock-down, allows researchers to stop, or alter, gene expression and measure the effects; a technique which has been vitally important across biology to empirically confirm gene function (Arts et al. 2003). In the case of knock-down or knock-out, gene-specific interfering RNA molecules are injected into animals, or added to cell-culture solutions, where they penetrate cells and attach to the complementary RNA transcripts in the cytoplasm. Hence, interfering RNA stops the complementary mRNA from being translated into proteins or carrying-out other post-translation functions. In 2009, researchers in Japan became the first to apply gene knock-out to the study of biomineralisation, and since then only three studies have taken advantage of this technique (Suzuki et al. 2009, Fang et al. 2012, Jiao et al. 2012, Wang et al. 2013).

Suzuki et al. (2009) knocked-down *pif* in pearl oysters (*Pinctada fucata*). Pif is an aspartic acid-rich matrix protein. It is predicted to be involved in the formation of the nacreous shell layer because it shows aragonite specific binding in calcium carbonate assays. In order to observe growth rate post *pif* gene knock-down, oysters were incubated in seawater containing strontium, which was incorporated into the shell and visible during scanning electron microscopy analysis. Knocking-down the expression of *pif* resulted in a disordered crystal structure and arrested growth. Using this empirical approach, authors were able to confirm the role of *pif* in nacreous growth. Other studies have also concentrated on related *pif* genes (*pfN23* and *pif97)* with congruent results; knocking-out *pif* genes results in disordered nacreous growth (Fang et al. 2012, Wang et al. 2013). When *pif* is knocked-down in 4-cell embryo, development is impaired and eventually arrests before reaching D-shape morphology; larvae are unable to produce a shell without *pif* and therefore die (Fang et al. 2012). The way *pif* interacts with other genes as part of a network or pathway is not yet understood. When *pif* is knocked-down in oysters (*P. fucata*), the expression of other genes (*aspein, N19, pearlin, prismalin-14, and prisilkin-39*) remain unchanged (Fang et al. 2012). Considering *pif* is thought to be a dominant biomineralisation gene, it is surprising that removing *pif* expression does not affect other genes, perhaps suggesting it is disconnected from hierarchical regulatory pathways.

The sections above have reviewed the contribution of analysing single genes and proteins to our understanding of biomineralisation; they have addressed the question - what can measuring one gene, or one protein tell us about how molluscs build their shells? In summary, from working with a single gene/protein focus we have discovered, unsurprisingly, focussed and singular mechanisms. With the advancement of technology it has now become possible to sequence entire genomes, transcriptomes and proteomes - the so called, "Next-Gen" era.

1.5 Global `Omics Approaches

The `omics revolution is now embedded into biological research. Next-Generation Sequencing (NGS) of DNA and RNA allows the simultaneous study of a significant proportion of genes or transcripts in chosen organisms - producing vast amounts of data. Next-Generation technology was initially limited by low quality databases with little functional gene annotation, however, improved computational power and rigorous statistical algorithms are now facilitating the meaningful bioinformatic analysis of nucleotide sequence data. In 2010 next-generation technology was first used to study mollusc biomineralisation, and in the following years molecular resources have increased (Table 1-3). The following section will review the contribution of 'omics, and so-called "global" analyses of sequence data, to our understanding of biomineralisation.

The first sequenced and assembled mollusc genome was that of the marine gastropod, *Lottia gigantea*, by the Joint Genome Institute (Joint Genome Institute, 2007). Two years later, Takeuchi et al. (2012) published a draft genome of the first bivalve – pearl oyster, *Pinctada fucata,* and most recently the draft genome of an oyster, *Crassostrea gigas,* was published (Zhang et al. 2012). Genomic resources provide a fundamental backbone for the molecular investigation of biomineralisation, both *in vivo* and *in silico*. An organism's genome however, can only provide information on biological potential; genomes represent the scripts of life, but cannot be used to decipherer how they are read. In order to understand how animals work, and specifically how molluscs build their shells, researchers can analyse which genes are transcribed (transcriptomics) and translated into proteins (proteomics).

Table 1-3. Summary of major molecular resources available to study shell production.

Type of Sequence Data	Taxa	Species	Reference
Genome	*Gastopodia*	*Lottia gigantea*	(Joint Genome Institute, 2007)
	Bivalvia	*Pinctada fucata*	(Takeuchi et al. 2012)
	Bivalvia	*Crassostrea gigas*	(Zhang et al. 2012)
Transcriptome	*Bivalvia*	*Mytilus galloprovincialis*	(Craft et al. 2010)
	Bivalvia	*Laternula elliptica*	(Clark et al. 2010)
	Bivalvia	*Pinctada margaritifera*	(Joubert et al. 2010)
	Bivalvia	*Crassostrea gigas*	(Zhang et al., 2012)
	Gastopodia	*Patella vulgata*	(Werner et al. 2013)
	Bivalvia	*Pinctada martensii*	(Shi et al. 2013)
Proteome	*Bivalvia*	*Pinctada margaritifera*	(Joubert et al. 2010)
	Bivalvia	*Saccostrea glomerata*	(Parker et al. 2011)
	Bivalvia	*Crassostrea gigas*	(Dineshram et al. 2012)
	Bivalvia	*Crassostrea hongkongensis*	(Dineshram et al. 2013)
	Gastopodia	*Lottia gigantea*	(Marie et al. 2013)

There are various next-generation, high-throughput methods to sequence RNA. Initially 454 pyrosequencing was used to study molluscs in 2010, although the focus of study bypassed biomineralisation (Craft et al. 2010). Homing-in on the study of shell deposition, Clark et al. (2010) used 454 pyrosequencing to obtain the mantle transcriptome of the Antarctic calm, *Laternula elliptica*. Gene Ontology analysis (GO analysis) revealed approximately 40% of (annotated) transcripts were likely to be involved in binding and shell secretion. Genes of interest, which were both highly expressed and biomineralisation candidates included: *carbonic anhydrase, SPARC, tyrosinase, tenacin* and *calponin.* The *L. elliptica* biomineralisation transcript repertoire did not show significant overlap with other mollusc biomineralisation genes in database

searches. Thus indicating either: different mollusc species use different mechanisms to build their shells and therefore biomineralisation genes evolve rapidly and diverge (Jackson et al. 2010); or, a combination of inherent methodological artefacts (lack of mollusc representation in databases, low transcript abundance, timing of sampling in relation to secretion activity, individual variation, statistical methods) prevents the resolution of conserved biomineralisation mechanisms. More recently, Shi et al. (2013) characterized the Pearl Oyster (*Pinctada martensii*) mantle transcriptome; 49 biomineralisation genes were annotated, of which 20 were novel. Annotated biomineralisation genes were grouped into functional categories revealing *calponin-like* genes represented 49% of all biomineralisation transcripts. One question which has arisen from mantle transcriptome analyses regards the relative abundance of biomineralisation genes. Clark et al. (2010) found that 40% of (annotated) transcripts were likely to be involved in binding and shell secretion, similarly Joubert et al. (2010) found a high proportion of genes involved in binding, and specifically ion binding. In contrast, Shi et al. (2013) state "biomineralisation genes were of low abundance but high importance" and Werner et al. (2013) used GO analysis to show terms associated with extracellular components were significantly overrepresented in the "predicted protein secretome", compared to the transcriptome (in the marine gastropod *Patella vulgata).* Neither study speculates further on these patterns, but one explanation could be that shell secretion is a slow, controlled background process. The transcripts in the mantle are therefore likely to be low in abundance whereas the proteins secreted as a matrix into the shell build-up overtime and do not lose structural integrity.

Proteomic analyses allow molecular biologists to track the final stages of the biomineralisation process - from genes in DNA, to transcripts in mantle tissue, to the end, post-translational product of proteins within shell matrices. To be able to conduct such analyses both nucleic and amino acid sequence data are required, which enables small peptide fragments to be mapped onto longer transcripts. Joubert et al. (2010) were the first group to apply global proteomic analyses of shell matrix proteins mapped onto mantle transcriptomic data; this integrative approach facilitated the investigation of both biomineralising tissue and calcified structures. Using proteomic mass spectrometry, Joubert et al. (2010) demonstrated that mantle transcripts of *Pinctada margaritifera* encode proteins which are present in its shell, thus confirming their role in biomineralisation. Almost all of the main peptides from the mass spectrometry led to

gene identification and 13 proteins showed sequence similarity to previously characterised *P. margaritifera* shell genes. Joubert et al. (2010) concluded the level of sequence similarity in the identified proteins suggested protein repertories are conserved, at least within the pearl oyster genus *P. margaritifera.* Continuing work within the *Pinctada* genus, Marie et al. (2012b) used proteomic and transcriptomic tools to investigate the molecular origins of nacre and prism layers in two species, *Pinctada margaritifera* and *Pinctada maxima.* Over 60 unique proteins were identified and data clearly, and perhaps unsurprisingly, demonstrated the protein repertoire associated with prisms and nacre is different. Previous paleontological and ontological evidence suggests that prisms arose before nacre in the Cambrian period (Kouchinsky 2000) which led to the speculation that "nacre evolved through simple horizontal partitioning of vertical prisms" (Carter and Clark 1985). By demonstrating that matrix proteins from one layer are not a precursor for the other layer, Marie et al. (2012b) provided evidence that each layer is the result of "evolutionary innovation". Using more experimental approaches, other researchers have exposed molluscs to stress, such as shell damage or pH change, and measured the proteomic response. For example, the proteomes of oyster larvae (*Crassostrea hongkongensis* and *Crassostrea gigas)* were compared for low pH and control treatments. In both cases the number of expressed proteins was significantly reduced in low pH which was correlated with a reduction in growth rate (Dineshram et al. 2012, Dineshram et al. 2013). Specifically, two proteins calmodulin and a troponin C were down-regulated, which the authors argued highlights them as candidates for controlling the rate of calcification.

Integrating omics analyses – genomics, transcriptomics and proteomics – is one of the most powerful global methods available to understand the molecular control underpinning biomineralisation. Zhang et al. (2012) assembled a genome of the oyster, *Crassostrea gigas*, in addition to transcriptomes from different organs, developmental stages and experimental stress treatments and the proteome of the shell. Representing perhaps the single most comprehensive biomineralisation study at the time, Zhang et al. (2012) found evidence to challenge the previous paradigm matrix-model. Self-assembling silk-like proteins were absent from the oyster genome, instead, fibronectin-like proteins are present in both the genome and highly abundant in the shell proteome. The formation of fibronectin fibrils is cell-mediated via integrin binding-domains, interestingly, the haemocyte transcriptome revealed the expression of genes encoding

integrins was four times higher than other organs. One hypothesis regarding the transport of calcium from the gills to the shell which has recently gathered support is use of exosome-vesicle-like structures during transportation. Molecular investigations of the oyster revealed 23.5% of proteins identified in the mantle, matched proteins in an exome database (Zhang et al. 2012). In summary, integrating genome, transcriptome and proteome data indicated haemocyte cells may play a role in organising fibronectin-like proteins in shell matrices, thus altering the way we think about the classic matrix-model described in sections above.

In the past three years global 'omics analyses have revolutionised the study of biomineralisation. Mollusc-specific molecular resources have improved beyond recognition, facilitating meaningful bioinformatic investigation and revealing major insights. Importantly, mollusc specific resources and databases are now growing (Table 1-4).

Table 1-4. Summary of reported mollusc-specific, molecular resource searches using public NCBI databases.

Date of search	Search term	Number of nucleotide sequences	Number of genes	Number of proteins	Reference
12.07.2004	Shell protein primary sequences	n/a	n/a	16	(Marin and Luquet 2004)
18.07.2005	*Mollusca* + *Bivalvia*	3,200	n/a	n/a	(Saavedra and Bachere 2006)
25.01.2010	*Bivalvia*	25,032	356	14, 507	(Clark et al. 2010)
04.02.2014	*Bivalvia*	344,453	1,292	65,084	6 month PhD progress report – current study
28.03.2015	*Bivalvia*	714,407	33,947	115,308	18 month PhD progress report-current study
04.04.2017	*Bivalvia*	858,005	39,052	134,300	Final PhD thesis-current study

Broadly speaking we now know: there may be a core set of conserved molecular mechanisms used in mollusc biomineralisation, in addition to a rapidly evolving and diverse subset; the two separate layers of shell, nacre and prisms, are the consequence of true evolutionary innovation; and, secreted shell matrix proteins may be organised by cells and vesicles, rather than self-assembling.

1.6 Conclusions and knowledge gaps

How do molluscs build their shells? The answer is still unclear. This review has synthesised and critically analysed the molecular understanding of molluscan calcification and has found some major patterns. In summary:

- Proteins which control biomineralisation can be divided into at least two categories. Insoluble structural proteins, which provide the framework of each shell layer; and soluble polyanionic proteins, which bind to the insoluble structure and determine the type of $CaCO_3$ crystal that grows.
- Proteins which control biomineralisation are diverse but typically share the properties of being modular and multifunctional.
- The production of an external proteinacious matrix to support crystal growth may be vesicle and cell-mediated (haemocyctes) rather than self-assembling.
- The mantle organ responsible for secreting shell can be functionally divided into specialised zones responsible for the production of nacre, prisms or periostracum.
- Biomineralisation is a controlled and non-continuous process which can respond to environmental cues.
- Molluscs may share a single, core homologous biomineralisation mechanism in addition to rapidly evolving and divergent mechanisms.
- The different layers of shell, prisms and nacre, are produced using different molecular mechanisms which are the result of evolutionary innovation rather.

The use of NGS and global ‘omics analyses has significantly progressed understanding and, as such, the typical biomineralisation model requires modification. This review has identified some major knowledge gaps, some of which this thesis will aim to address. It is acknowledged that many of these limitations were due to lack of resources and technology at the time. Firstly, and most significantly, understanding of how molecules

interact to produce a coherent calcification pathway is still lacking. Secondly, there is little understanding of the major modes of transport of calcium in molluscs; leading to uncertainty such as self-assembled versus cell-mediated shell production. Thirdly, while knowledge is developing on the sequence and structure of shell genes and proteins, the empirical functions of such molecules have yet to be tested. In addition, the tissue and cellular localisation of genes and proteins has only been characterised for a handful of genes in five species - predominantly those species which produce pearl. Finally, normal variation within and between species is unknown presenting problems when trying to decipher signal from noise in large data sets.

In the last decade, the development of new technology in molecular biology and the improvement of mollusc-specific resources have rapidly increased our understanding of how molluscs build their shells. In general, to build on current work and continue progress in the field, a more integrative approach is required. For example, the integration of molecular investigations into the study of whole-organism physiology may provide understanding of energetic constraints and the way the calcification processes fit into mollusc biology. More specifically, the implementation of new techniques such as gene network analysis, in combination with empirical techniques, such as *in situ* hybridisations, gene knock-down and Clustered Regularly Interspaced Short Palindromic Repeats (CRISPR) gene editing should be used to provide a more functional understanding. In addition, in order to understand subjects such as the conservation of a core biomineralisation mechanism versus divergent mechanisms, a wider range of mollusc species should be studied. Finally, whilst the study of diverse species will answer questions regarding divergent mechanisms, the development of a general model system for biomineralisation could facilitate the characterisation the calcification pathway. For example, the development of cultured secretary mantle cell lines may facilitate rigorous, controlled and statistically robust testing (high replication and power) of hypotheses regarding the calcification pathway (Gong et al. 2008).

1.7 Thesis overview

Our climate is changing and as a result, the marine environment is changing too. The Southern Ocean for example, is warming and becoming more acidic; it is predicted to be

undersaturated in calcium carbonate ($CaCO_3$) within the current century (Hauri et al. 2016). How will life in the oceans cope with changing conditions?

There is concern for calcified organisms, such as molluscs, as the ability to extract carbonate ions (to incorporate into shell) will be compromised due to decreased pH. Although there is sophisticated understanding of calcium regulation in mammals, biomineralisation processes in other species are different and poorly characterised. Data describing calcium pathways in non-model species are lacking, and such fundamental information is required to predict how species will respond to climate change.

Climate change aside, understanding the molluscan calcification pathway is important for economic and societal benefit. Shellfish aquaculture has the potential to meet the protein demands of the growing human population. In order to provide food security and to make the shellfish industry efficient and viable, reducing shell disease and damage is vital. In addition, knowledge about how molluscs build their shells can be applied to materials science to help design smart new materials, from abundant calcium carbonate, for society.

This project will use two bivalve molluscs, the Antarctic clam, *Laternula elliptica*, and the temperate clam, *Mya truncata*, as model systems. The project aims to characterise calcification pathways in two ecologically and morphologically similar species, living in dissimilar geographical and physical environments. Shell damage-repair, as well as normal growth investigations will be used to identify the genes involved in calcification with the aim of characterising currently known and novel genes in calcium mobilisation pathways.

1.8 Thesis Species

1.8.1 Laternula elliptica

1.8.1.1 Distribution

L. elliptica are sediment-burying bivalves with a circum-Antarctic distribution and are an important infaunal species in the Southern Ocean. *L. elliptica* are one of the most abundant animals in the Southern Ocean ecosystem (in terms of biomass) and they play a key role in coupling the pelagic and benthic systems through nutrient cycling (Arntz et

al. 1994). *L. elliptica* distribution is thought to be governed by its thermal capacity, it has a short-term critical maximum temperature limit between 6 °C and 9 °C (Peck et al. 2002), and a long-term critical maximum temperature limit between 3 °C and 4 °C (Morley et al. 2012).

1.8.1.2 Biology and Ecology

L. elliptica are polar marine ectotherms, which are characterised by being highly stenothermal due to evolving in a constantly cold environment for around 22-25 million years. Fossils of *L. elliptica* from the Pilocene are found in sedimentary rocks around the Antarctic Peninsula (Jonkers 1999).

L. elliptica are subtidal filter feeders [although very recently have been reported in the intertidal zone for the first time by Waller et al. (2016)] , they bury deep into the sediment (>50 cm) and extrude their siphons into the water column above. There are no crushing predators in the Southern Ocean, and *L. elliptica* bury deep enough to avoid drilling predation by gastropods. Predation can however occur in the form of siphon nipping from notothenoid fish (Ahn 1997). A major abiotic disturbance which affects *L. elliptica* populations is ice berg scouring. Ice bergs can wipe out patches of sea bed, but if scouring is survived, *L. elliptica* have been reported to repair shell damage (Harper et al. 2012). Populations which endure frequent scouring events have thicker shells than those in more sheltered sites.

As a species, *L. elliptica* faces major environmental challenges under future climate change. The Southern Ocean is predicted to soon become under-saturated in calcium carbonate in ocean acidification models and in addition, the Western Antarctic Peninsula region has experienced one of the fastest rates of warming on the planet (Turner et al. 2014). A consequence of the latter, aside from the temperature rise *per se*, is an increase in ice calving events and iceberg scouring which can decimate seas beds and the associated infaunal communities (Cook et al. 2005, Barnes and Souster 2011).

1.8.1.3 Shell structure

The shell microstructure of *L. elliptica* is composed of 3 structurally distinct layers: (i) the outer periostracum (approximately 10 µm and relatively thick); (ii) an aragonitic outer

layer of granular prisms; and (iii) an aragonitic layer of sheet nacre [Figure 1-2, (Bieler et al. 2014)].

1.8.2 Mya truncata

1.8.2.1 Distribution

M. truncata are also sediment-burying bivalves, they have a large northern hemisphere distribution spanning the whole of the Arctic and extending down to the Bay of Biscay in Europe. Like *L. elliptica*, *M. truncata* plays an important role in bentho-pelagic coupling in the regions they inhabit (Queiros et al. 2013). Unlike *L. elliptica* however, *M. truncata* have a moderate window of physiological capability, and hence a broad circumboreal and temperate distribution (Amaro et al. 2005).

1.8.2.2 Biology and ecology

M. truncata evolved in late Eocene (34 million years ago), and later spread to reach its current large subborreal and temperate extent (MacNeil 1965). Similar to *L. elliptica*, *M. truncata* inhabit the very low intertidal stretching into the subtidal to approximately 70 m depth. Unlike its invasive and economically important relative - *Mya arenaria*, *M. truncata* has received relatively little research attention and hence, detailed data on its basic biology and ecology are lacking.

1.8.2.3 Shell structure

The shell microstructure of *M. truncata* is identical to its commercially important relative *M. arenaria*, it is composed of 4 structurally distinct layers: (i) the outer periostracum (approximately 3 µm and relatively thin); (ii) an outer shell layer of aragonitic granular prisms; (iii) a middle layer of aragonitic crossed lamellar; and (iv) an aragonitic inner layer of complex crossed lamellar [Figure 1-3, (Bieler et al. 2014)].

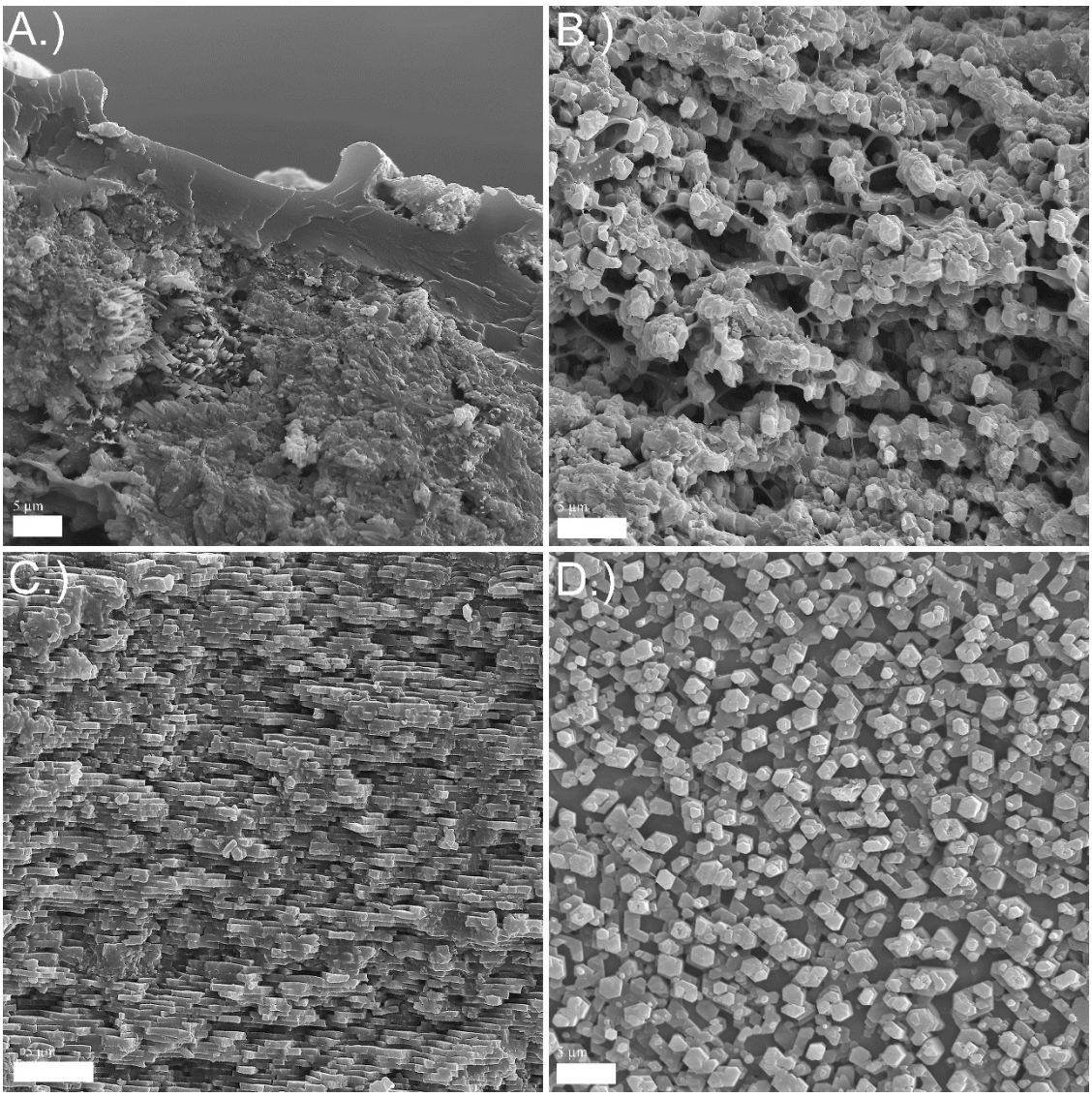

Figure 1-2. Microstructural layers of *Laternula elliptica* shell observed by Scanning Electron Microscopy (SEM) analysis. A.) Periostracum, B.) Outer shell layer made up of granular prisms, C.) Cross section of sheet nacre, D.) Plan view of sheet nacre. Microstructure nomenclature used as per Bieler et al. (2014). Images courtesy of Elizabeth M. Harper, University of Cambridge.

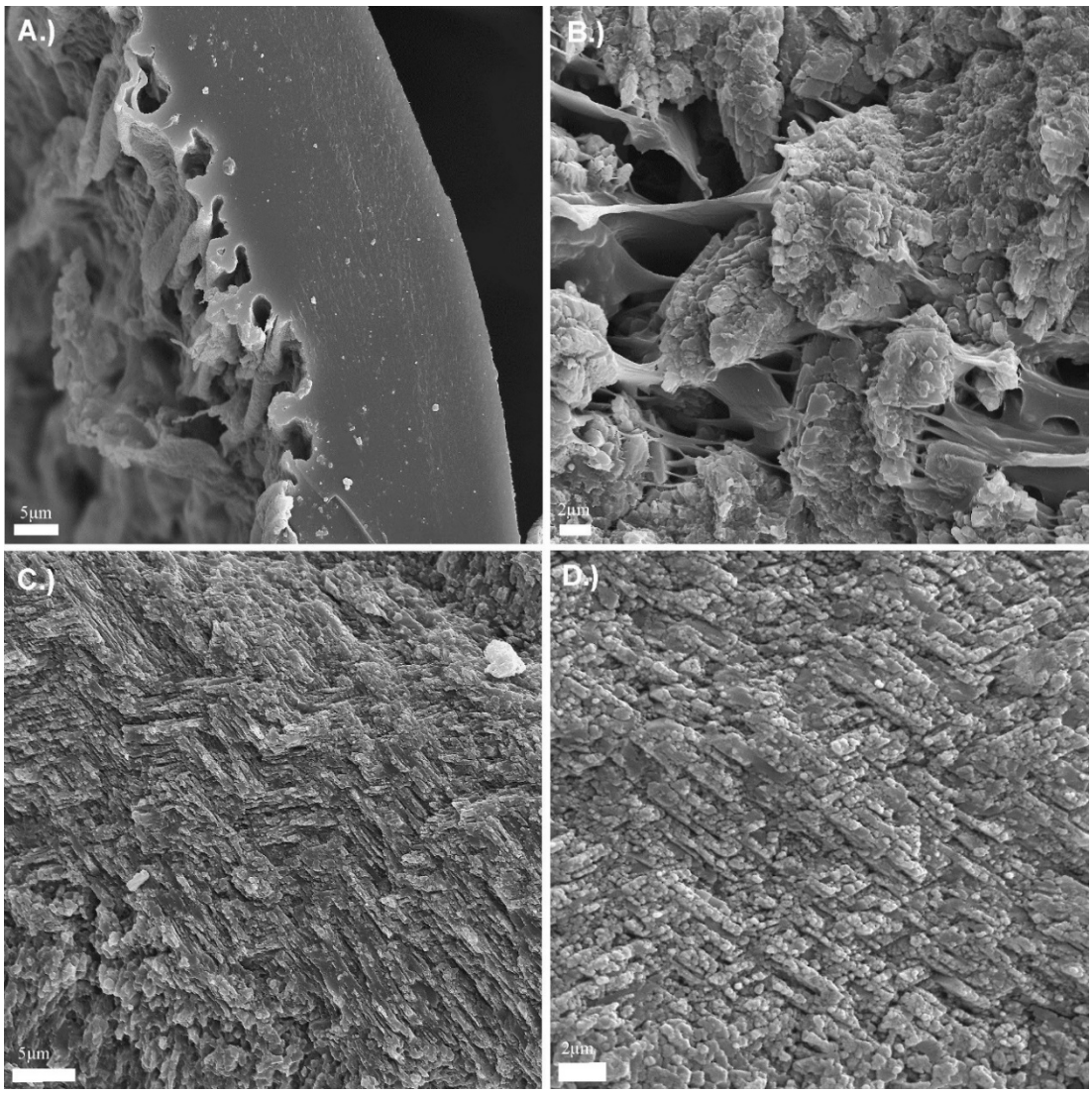

Figure 1-3. Microstructural layers of *Mya truncata* shell observed by Scanning Electron Microscopy (SEM). A.) Periostracum, B.) Outer aragonitic granular prism shell layer, C.) Middle aragonitic layer of crossed lamellar and D.) Inner aragonitic layer of complex crossed lamellar. Microstructure nomenclature used as per Bieler et al. (2014) Images courtesy of Elizabeth M. Harper, University of Cambridge.

1.8.3 Why compare *L. elliptica* and *M. truncata*?

This thesis will primarily focus on the Antarctic species *L. elliptica;* M. *truncata* has been chosen as a northern hemisphere, temperate comparison. *M. truncata* is often reported as an Arctic bivalve (Camus et al., 2002; Gillis and Ballantyne, 1999), however animals in this thesis were sampled from a more southerly, temperate, region of their distribution on

the West coast of Scotland. *M. truncata* and *L. elliptica* are ecologically and morphologically very similar, but their physical environments (Arctic to temperate versus Antarctic), geographical extent (ranging from Arctic through subboreal to temperate versus Southern Ocean exclusively) and evolutionary history (phylogenetically distant relatives) differ significantly. In addition, the two species have different shell microstructures. By comparing two species which are very similar in some ways, but different in others, this thesis aims to investigate functional questions about how species in different environments build their shells.

1.9 Thesis aims and objectives

The overall objective of this project is to better understand calcification pathways in two non-model bivalve species, the Antarctic clam (*L. elliptica*) and the temperate clam (*M. truncata*). In general, this thesis will bring together different disciplines (primarily molecular biology, bioinformatics, physiology, ecology and immunology) in order to gain a holistic, functional, understanding of biomineralisation in two species living in dissimilar habitats. More specifically the project will address the following core aims:

1.) Provide a detailed characterisation of the *L. elliptica* mantle anatomy and ultrastructure.

2.) Identify proteins in the *L. elliptica* nacreous shell layer using proteomics.

3.) Understand the function of *L. elliptica* biomineralisation candidates by localising their gene expression to tissue, cellular and subcellular resolution.

4.) Estimate the timing of the *L. elliptica* transcriptional response to shell damage.

5.) Identify candidate *L. elliptica* biomineralisation genes involved in the response to shell damage.

6.) Understand the spatial location of molecular mechanisms in the response to shell damage in *L. elliptica*.

7.) Increase the temporal resolution of the understanding of the *L. elliptica* transcriptional response to shell damage.

8.) Understand the effect of age on the *L. elliptica* transcriptional response to shell damage.

9.) Provide and apply a regulatory gene network of the *L. elliptica* mantle to aid the understanding of biomineralisation pathways.

10.) Provide a molecular resource to aid the study of biomineralisation in *M. truncata* by sequencing, assembling and putatively annotating the adult mantle transcriptome.

11.) Identify and investigate the expression and phylogeny of candidate biomineralisation genes in the *M. truncata* mantle transcriptome

12.) Compare the mantle transcriptome and biomineralisation genes of *M. truncata* to that of the Antarctic clam, *L. elliptica.*

Chapter 2 MATERIALS AND METHODS

2.1 Overview

This chapter details the core methodologies used in this thesis, which are mainly molecular methods. In each of the subsequent data chapters there is a chapter-specific methods section, detailing procedures such as experimental design, bioinformatics and statistics.

2.2 Animal collection and husbandry

2.2.1 L. elliptica

L. elliptica specimens used in this thesis were collected by SCUBA divers from Hangar Cove near Rothera Research Station, Adelaide Island, Antarctic Peninsula (67° 34' 07" S, 68° 07' 30" W) at depths of 10-15 m. Animals were immediately returned to the Rothera aquarium where they were maintained in a flow-through system with temperature of 0.6 ± 0.3 °C, under a 12 h:12 h simulated light:dark cycle. At the end of the austral summer animals were transferred to the British Antarctic Survey aquarium facilities (in the UK) where they were habituated to aquarium conditions for at least four weeks (closed recirculating system at water temperature and salinity of 0 ± 0.5 °C and 35-38 psu respectively, a 12 h:12 h light:dark regime, and fed a mircoalgal mix of *Isochrysis* and *Nannochloropsis* culture weekly) prior to any sampling or experimentation.

2.2.2 *M. truncata*

M. truncata were collected by myself and the SCUBA divers from the NERC UK National Facility for Scientific Diving from Dunstaffnage Bay, North West Scotland (56° 27' 06"N, 5° 26' 01"W) at depths of 5-15m. Animals were immediately transferred to the Scottish Association of Marine Science (SAMS) aquarium where they were maintained in a flow-though system with temperature and salinity of 15 ± 1 °C and 35-38 psu respectively, under a 12 h:12 h simulated light:dark cycle and fed a pre mixed microalgal blend (Shellfish Diet, Varicon Aqua solution Ltd., UK). Animals were habituated to aquarium conditions for at least four weeks prior to experimentation.

2.3 RNA extraction

Total RNA was extracted from tissue samples on ice using Tri-Reagent (Sigma-Aldrich, UK). In a 1.5 ml microcentrifuge tube, samples were homogenized in 600 µL of Tri-Reagent using a manual glass homogenizer and sand, 120 µL of chloroform was added to the homogenate before centrifugation at 13,000 rpm for 15 min at 4 °C. The top aqueous layer was then transferred to a new 1.5 ml microcentrifuge tube and 600 µL of isopropanol was added to precipitate out the RNA. A pellet of RNA was collected at the bottom of the 1.5 ml microcentrifuge tube via centrifugation at 13, 000 rpm for 10 mins at 4 °C. The RNA pellet was washed with 75% ethanol, dried via vacuum centrifugation and re-suspended in 30 µL of RNase free water. RNA was then purified using the RNeasy clean-up kit (QIAGEN, UK), which included a DNase step. All RNA samples were analysed for concentration and quality by spectrophotometer (NanoDrop, ND-1000) and tape station analyses (Agilent 2200 TapeStation). Only RNA samples with a 260/230 and 260/280 ratio of 1.8-2.2 and RNA integrity number (RIN) of 5 or above were used in further analysis. RNA was stored at -80 °C until reverse transcription or sequencing was performed.

2.4 Reverse Transcription

All samples were diluted to 30 ng μl^{-1} total RNA prior to reverse transcription. cDNA was synthesized from 1 µg RNA using the QuantiTect Reverse Transcription Kit (Qiagen, U.K) according to manufacturer's protocol, which included another DNase step. cDNA was synthesized under the following conditions: 42 °C for 30 minutes followed by heat-inactivation of transcriptase at 95 °C for 3 minutes. cDNA was stored at -20 °C until gene expression analysis was performed.

2.5 Candidate gene expression analyses

2.5.1 Candidate gene selection

2.5.1.1 Criteria

Candidate biomineralisation genes were selected from the *L. elliptica* and *M. truncata* mantle transcriptomes for further study. Primers for unique regions of these genes were designed using Primer 3 software (to produce single amplicons with a size of approximately 100-1000 bp depending on application, an annealing temperature of 56-66 °C and a GC content between 55–60 %, Appendix A - Primer tables). The criteria for candidate gene selection spanned four broad categories:

1. **Previous literature.** Some genes had been previously characterised as likely biomineralisation candidates in mollusc and non-mollusc species, for example a variety of the early characterisations discussed in sections 1.2 and 1.3. Genes in this category = *tyrosinase, pif, chitinase, chitin-binding domains, cartilage matrix protein, calponin, tyrosine, astacin*.

2. **Proteome.** In addition to using previous literature as a selection criteria, five candidates were selected for both species from their respective shell proteomes (the *L. elliptica* proteome was published alongside work from the present thesis [Chapter 3] and the *M. truncata* proteome was published by collaborators Arivalagan et al. (2016) in a stand-alone publication separate to work in this thesis). The proteomes provided confirmation that genes which were being expressed in the mantle tissue, were also being translated into proteins which were incorporated into the shell matrix. For *M. truncata*, the mantle tissue proteome was also available and four of the candidates selected were present in the mantle proteome (Arivalagan et al. 2016). Genes in this category = *tyrosinase, pif, chitin-binding domains, mytilin, complement control protein, astacin, calponin, cartilage matrix protein*.

3. **Mantle expression.** The third criteria used for selection was high expression in the mantle transcriptome; this criteria led to the identification of genes which had

no annotation but were extremely abundant in the mantle transcriptome. Further characterisations therefore, could reveal these genes to be involved in biomineralisation, hence providing functional information on previously unknown, unannotated genes. Genes in this category = *tyrosinase, contig 01043, contig 2930, contig 3456, tyrosine.*

4. **Immune function.** Some candidate genes were selected because they were likely to be involved in the immune response. The mollusc shell and mantle combined, act as a barrier to the external environment, and as such biomineralisation and immune processes are likely to be entwined. Unravelling immune and biomineralisation mechanisms represents a significant challenge for researchers trying to understand how molluscs build their shells. The challenge is partly due to the dual role of haemocytes both as immune cells and hypothesised calcium carbonate chaperones (Mount et al. 2004). Studying immune genes could therefore provide insight on the interaction of immune and biomineralisation functions. Genes in this category = *mytilin, complement control protein and astacin*.

2.5.1.2 Genes

Tyrosinase was selected for further study. Both species had a tyrosinase which was present in the shell proteome and one that was absent from the proteome, and both species had a *tyrosinase* in the top 50 most highly expressed genes in the mantle transcriptome. Tyrosinase is a multifunctional shell-associated protein which has been shown to have a functional role in the cross-linking of the soluble periostracum precursor (the periostracin) to form an insoluble periostracum (Waite et al. 1979). It has also been localised in the prismatic layer of shell (Kouhei Nagai et al. 2007) in addition to being expressed in the pallial mantle. Tyrosinase is therefore potentially also involved in both periostracum formation, and shell formation in the prismatic and nacreous shell layers (Takgi and Miyashita 2014).

For both species, *pif* was selected as a candidate gene for further study as it is present in the shell proteome of both species and is also well described in previous literature as a

shell matrix protein (M. Suzuki et al. 2013, Suzuki et al. 2009). Suzuki et al. (2009) used a calcium carbonate-binding assay to identify an acidic matrix protein, pif, in the pearl oyster *Pinctada fucata* that specifically binds to aragonite crystals. In addition, pif was shown to interact with chitin (Suzuki et al. 2009). The presence of chitin in mollusc shells is well documented and is hypothesised to form a major strucural component of the shell matrix, responsible for guiding crystal growth (Weiss 2012, Weiss et al. 2013). Chitin synthases and chitinases are responsible for the production and modification of chitin respectively, and therefore various chitin-related genes (some of which were present in the shell proteome and some which were absent from the shell proteome, in both species) were selected for further study. Some genes contained a single chitin-binding domain, whilst others showed sequence similarity to complete chitinase sequences.

In order to study the interaction of immune and biomineralisation functions, immune genes were selected. Mytilin for example, is an antimicrobial peptide (Mitta et al. 2000) and genes containing complement control domains, are innvolved in the identification of pathogens (Ferreira et al. 2010). Astacin has also recently been hypothesised to be involved in both immune response and biomineralisation in molluscs, and other species (Xiong et al. 2006). All of the immune candidates were present in the shell proteome.

Other genes were selected for further study based on their sequence similarity to non-mollusc biomineralisation matrix proteins (considered under the previous literature criteria), for example *cartilage matrix protein*. Cartilage matrix proteins are involved in calcium phosphate biomineralisation in vertebrates where they bind to calcium phosphate crystals and form part of an extracellular matrix (Acharya et al. 2014). Molluscs do not produce cartilage yet, the cartilage matrix protein was found in the *L. elliptica* shell proteome. In addition, calponin was selected, it has been demonstrated to be involved in the biomineralisation of bone in humans (Ueda et al. 2002) and the shell matrix protein, PFMG8, has been shown (*in silico*) to contain a calponin domain which has a calcium binding site (Evans 2012). In addition Shi et al. (2013) highlight the potential importance of calponins in molluscan biomineralisation when they characterised the Pearl Oyster (*Pinctada martensii*) transcriptome; 49 biomineralisation genes were annotated, of which 20 were novel. Annotated biomineralisation genes were grouped into functional categories revealing *calponin-like* genes represented 49% of all biomineralisation transcripts. Calponin belongs to a family of actin-binding proteins. The actin-binding

properties of calponin are regulated by either, interaction with calcium binding proteins, or interaction with tyrosine (Abouzaglou et al. 2004) and therefore tyrosine was also investigated.

2.5.2 PCR and sequencing

In order to confirm all primers in this thesis (Appendix A - Primer tables) were working efficiently and amplifying the correct product, PCR was carried out on a pool of cDNA and the products were then sequenced to confirm identity. Lyophilised primers (Invitrogen, UK) were reconstituted to 100 µmol with RNase-free water and PCR was carried out as per the Bioline, BIOTAQ™ DNA Polymerase kit (Bioline, UK). Briefly, PCR was carried out in a 43 µL reaction volume containing 0.5 µL cDNA, 23.5 µL H_20, 5 µl NH_4 10x buffer (Bioline, UK), 1.5 µL $MgCl_2$ (50mM Bioline, UK), 5 µL deoxynucleoside triphosphates mix (100 mM, Bioline, UK), 5 µL of each primer (10 mM, Invitrogen, UK) and 0.25 µL BIOTAQ DNA Polymerase (BIOTAQ™, 0.25 units [one unit is described as the amount of enzyme that incorporate 10nmoles of dNTPs into acid-insoluble form in 30 min at 72 °C], Bioline, UK). Following initial denaturation at 95 °C for 5 min, forty PCR cycles were carried out as follows: denaturation at 95 °C for 15 s, primer specific annealing temperature for 20 s, and extension at 72 °C for 30 s. To verify amplicon size, and to check for non-specific amplification, amplicons were ran on a 1.5% agarose gel.

PCR amplicons were purified using QIAquick®PCR purification kit (Qiagen, U.K.) according to manufactures instructions and sent to Source BioScience, Cambridge, for Sanger Sequencing. Amplicon sequences were aligned to their expected sequence using JEMBOSS Needle Software (www.ebi.ac.uk/Tools/psa/emboss_needle/).

2.5.3 Semi-quantitative PCR (semi-qPCR)

PCR was carried out as per the Bioline, BIOTAQ™ DNA Polymerase kit (Bioline, UK) and recommendation of Souze et al. (2009). PCR was carried out using the primers previously diluted and tested in section 2.5.2. Briefly, PCR was carried out in a 43 µL reaction volume containing 0.5 µL cDNA, 23.5 µL H_20, 5 µl NH_4 10x buffer (Bioline, UK), 1.5 µL $MgCl_2$ (50mM Bioline, UK), 5 µL deoxynucleoside triphosphates mix (100 mM, Bioline, UK), 5 µL of each primer (10 mM, Invitrogen, UK) and 0.25 µL BIOTAQ

DNA Polymerase (BIOTAQ™, 0.25 units, Bioline, UK). Following initial denaturation at 95 °C for 5 min, twenty six PCR cycles were carried out as follows: denaturation at 95 °C for 15 s, primer specific annealing temperature for 20 s, and extension at 72 °C for 30 s. PCR amplicons were sequenced to confirm identity. Amplified products were analysed by electrophoresis on 1.5 % agarose gel (Bioline, UK) containing GelRed™ and visualized under UV illumination (U:genus 3; Syngene, Cambridge, UK). The relative intensities of amplified PCR products were determined using Syngene GeneTools software (version 3.06) and expressed as an integrated density value (IDV). Primer annealing temperature was optimised across a temperature gradient to find the best temperature for each primer set, and in addition serial dilutions were used to check that semi-quantification was achievable (5 point, 2 fold dilution which resulted in a linear decrease in IDV). Each round of PCR reactions included a no template control to check for contamination.

2.5.4 Quantitative PCR (qPCR)

For quantitative PCR (qPCR), lyophilised primers (Invitrogen, UK) were reconstituted to 100 µmol with RNase-free water and mixed with Brilliant II SYBR® Green (Agilent, UK) following manufacturers guidelines (Appendix A - Primer tables). Briefly, for each reaction, 10 µL of 2x SYBR Green was added to 1.2 µL of each of the forward and reverse primers (10 mM, Invitrogen, UK), 6.6 µL was then added followed by 1 µL of cDNA template. Fluorescence was detected (Stratagene, Mx3000P) over 40 cycles with cycling conditions of 95 °C for denaturing, primer-specific annealing 62-66 °C, and extension at 72 °C. All samples and standards were run in triplicate and each plate included triplicate H_2O and no template controls. Standard curves of each gene were generated on each qPCR plate using four point, 2-fold serial dilutions of cDNA (from pooled cDNA). The efficiencies of the qPCR reactions were 90-110%, as determined using the slope of the standard curve [Equation 2-1, (Schmittgen and Livak 2008)]:

Equation 2-1

$$Efficiency\ (\%) = [10^{\left(\frac{slope}{-1}\right)} - 1]$$

Quantification of gene expression was conducted using the comparative C_T method that normalises gene expression of each sample in relation to an internal housekeeping gene as decribed by Schmittgen and Livak (2008). Evaluation of C_T values for an internal housekeeping gene across samples was performed to check there was no significant difference in expression across the different tissue types. Normalized C_T values were obtained by subtracting the C_T value of the internal housekeeping gene from that of the candidate gene in the same sample (ΔC_T). Differences between the average ΔC_T and ΔC_T of each sample were expressed as $\Delta\Delta C_T$. The fold differences ($2^{-\Delta\Delta C_T}$) of candidate gene expression were compared.

2.5.5 *In situ* hybridisations

2.5.5.1 Fixation, dehydration, embedding and sectioning

Adult *L. elliptica* mantle tissues were fixed for 12 h in freshly prepared Davidson fixative (22 % formalin, 33 % ethanol, 12 % glacial acetic acid and 33 % sterile sea water) and transferred to 70 % (room temperature [RT]) ethanol for storage. Sampled material was processed through an ethanol and xylene series to dehydrate (Shandon Duplex Tissue Processor: 70 % ethanol 2 h, 90 % ethanol 2 h, 100 % ethanol 1 h, 100 % ethanol 1 h, Histo-clear® [xylene substitute, National Diagnostics, Manville] 1 h, Histo-clear® 2 h, Histo-clear® 1 h, molten wax 2 h, 65 °C) and embedded in paraffin wax (Merck, Germany). Paraffin blocks were serially sectioned at 8 µm on a hand-rotated microtome using disposable steel blades and sections were mounted onto poly-L-lysine coated slides and dried overnight at 50 °C.

2.5.5.2 Subcloning and riboprobe synthesis

Riboprobes (all approximately 1 Kbp) were designed for unique regions of each candidate gene (Appendix A - Primer tables). PCR amplicons were purified using QIAquick®PCR purification kit (Qiagen, U.K.) according to manufacturer's instructions. Purified DNA was ligated into a pGEM-T Easy vector (Promega, U.K.). Ligations were performed overnight at 4°C in a reaction volume of 10 µL containing: 1 µL vector (Promega, U.K.), 1 µL ligation buffer (New England Biolabs, U.K.), 1 µL T_4 DNA ligase (Promega, U.K.) and 7 µL DNA sample.

Ligated vectors were used to transform *Escherichia coli* (Stratagene XL2-Blue MRF, Ultracompetent cells) as per manufacturer's instructions. Cells were thawed on ice, 1 µL of ligation mix was added to 15 µL of cells. Cells were heat shocked at 42 °C for 30 secs and returned to ice. Transformed cells were incubated for at least 1 hr at 37 °C in 500 µL SOC. Prior to plating, IPTG (25 µL, 200 mM,) and X-Gal (15 µL, 16 %) were added to facilitate the colour detection system. Cultured cells were plated onto LB/ampicillin agar plates (agar 15 g L^{-1}, Amp 50 mg L^{-1}). The pGEM-T Easy vector (Promega) provides a white/blue colour detection system (vectors contain T7 and SP6 RNA polymerase promoters flanking a multiple cloning region within the alpha-peptide coding region of the enzyme beta-galactosidase, and therefore insertional inactivation of the alpha-peptide allows recombinant clones to be directly identified by blue/white screening on indicator plates). When grown in the presence of IPTG (inducer) and X-Gal (chromogenic substrate), cells containing a vector with an insert remain white and can be selected as positive clones. Cells which contain a vector without an insert turn blue and are discarded. Individual positive clones were selected and grown overnight in 5 ml of LB broth in a 15 ml falcon tube with ampicillin (50 mg L^{-1}, 37 °C with shaking).

In order to obtain purified plasmid DNA for sequencing, bacterial cells were harvested from growth media by centrifugation and plasmids were isolated using a colour aided alkaline-lysis method (QIAGEN Plasmid Mini-Kit, Qiagen, U.K.). Pelleted bacterial cells were resuspended in a Tris-HCl-EDTA buffer containing RNase. Cells were lysed using an alkaline lysate buffer containing SDS and sodium hydroxide (NaOH), samples were mixed by inversion and kept on ice for 5 min. Lysate was neutralised with potassium acetate and centrifuged to remove cell debris (RNA, genomic DNA and proteins). Supernatent, containing plasmid DNA, was removed and bound to a resin membrane. Plasmid DNA was washed twice with a washing buffer containing NaCl, 3-(N-Morpholino) propanesulfonic acid (MOPS) and isopropanol. Clean plasmid DNA was eluted and precipitated using isopropanol. Precipitated plasmid DNA was pelleted by centrifugation (13000 rpm, 30 min), washed with 70 % ethanol and subsequently air-dried. The resulting ultra clean plasmid DNA was resuspended in sterile water and concentration and purity was determined by spectrophotometer (NanoDrop, ND-1000). Plasmid DNA was sequenced and amplicon identity was confirmed as per section 2.5.2.

Plasmid DNA was used as a template for riboprobe synthesis. Primers were designed to amplify a region of the plasmid from the M13 sites in order to include T7 and SP6 RNA polymerase promoters. Antisense riboprobes were synthesised using Promega reagents in a 10 µl reaction containing 1x reverse transcription buffer, 10 mM Dithiothreitol, 1x Digoxigenin (DIG) RNA labelling Mix (Roche, Germany), 0.25-0.5 volume PCR template and 20 Units of the appropriate RNA polymerase (SP6 or T7; Promega, Germany). Probe synthesis reactions were carried out at 37 °C for 2-4 h. All riboprobes were purified by precipitation using 0.1 volume of 3 M sodium acetate pH 5.2 and 3 volumes of absolute ethanol for 15 min, and subsequently centrifuged for 15 min at 12,000 rpm. All precipitation steps were carried out at room temperature. The resulting pellets were washed once in 75 % ethanol, air-dried and dissolved in 10 µl water at 55 °C. After quantification using a Nanodrop, 500 ng of riboprobe was denatured in 95 % deionised formamide at 75 °C for 10 min and qualitatively assessed by agarose gel electrophoresis. The remaining riboprobe solution was adjusted to a final concentration of 300 ng/µl using deionised formamide.

2.5.5.3 Tissue rehydration, proteinase K and probe hybridisation

Dried tissue sections were rehydrated through a graded ethanol series (5 min in each of the following: xylene x3, 100 % ethanol x2, 90 % ethanol, 80 % ethanol, 70 % ethanol, water, 1x phosphate buffered saline with 0.1 % Tween20 [PBTw]) before being transferred to an Invatis in situ-Pro robot for all subsequent treatments.

Tissue sections were treated with proteinase K (50 µg mL^{-1}, 10 min, RT), which was then stopped with 0.2 % glycine (5 min, RT), washed with PBTw (5 min, RT) re-fixed with 4 % paraformaldehyde (20 min, RT), incubated with hybridisation buffer (2 h, 55 °C), incubated with specific riboprobe in hybridisation buffer (500 ng μL^{-1}, 26 h, 55 °C, hybridisation buffer made as per Jackson et al. (2016)) and washed with a series of saline-sodium citrate (SSC) buffers (4x, 2x, 1x, 1x with 0.01 % Tween20, 15 min each, 55 °C).

2.5.5.4 Antibody binding, colour development and post-processing

Maleic acid buffer (MAB) was then added to the tissues (10 min, RT), followed by 2 % blocking solution (2.5 h, RT) and finally primary anti-DIG antibody conjugated to alkaline phosphatase in 2 % blocking solution (Roche, 1:10,000, 12 h, RT, 2% blocking

solution made as per Jackson et al. (2016)). Unbound antibody was removed with 15 washes in PBTw (20 min, RT) and tissue sections were removed from the robot. For colour development, tissue sections were washed twice with alkaline phosphatase (20 min, RT) before colour detection buffer was added (in the dark, time optimised for each riboprobe to obtain best signal to background ratio, RT). Tissue sections received two final washes with PBTw (5 min, RT) and were post-fixed with 3.7 % formamide in phosphate buffered saline (PBS, 2 h, RT) before being dehydrated at RT through a graded ethanol series (5 min in each of the following: 1x PBTw, water, 70 % ethanol, 80 % ethanol, 90 % ethanol, 100 % ethanol x2, xylene x2) and mounted with DPX (DPX Mountant for histology, Sigma, Germany). A list of solutions and full protocol is available in Jackson et al. (2016).

2.5.5.5 Imaging

For obtaining a broad overview of the tissue, sections were imaged under a Zeiss stereo Discovery V8 microscope, objective x0.63, running Zeiss camera software Axio Vision Rel. 4.7 for more detailed observations, x10–x100 objective, using a Zeiss Axio Imager Z1 microscope running Zeiss camera software Axio Vision Rel. 4.8. Images of all samples were captured using automatic settings for exposure and white balance. Each image was linearly adjusted for brightness and contrast and using ImageJ (Schneider et al. 2012).

2.6 Proteomic analysis of nacre shell proteins

Proteomic analyses were carried out by Benjamin Marie at the Natural History Museum in Paris and I then carried out data analysis and interpretation. The methods used to generate proteomic data are provided for completeness and repeatability.

2.6.1 Shell preparations and protein extraction

Superficial organic contaminants and the periostracum were removed by incubating intact adult *L. elliptica* shells (n = 6) in sodium hypochlorite (5 %, vol/vol) for 24-48 h followed by rinsing with water. The external prismatic layer was mechanically removed and the nacre was broken into 1-mm large fragments before being ground to fine powder (> 200

µm) and decalcified in acetic acid overnight (10 %; 4 °C). The acid-insoluble matrix (AIM) was collected by centrifugation (15,000 g; 10 min; 4 °C) and rinsed six times with MilliQ water by a series of resuspension-centrifugation steps before being freeze-dried and weighed.

Digestions of insoluble nacre matrix samples were performed in a solution of 100 µL of 10 mM dithiothreitol (Sigma-Aldrich, France) in 500 mM triethylammonium bicarbonate (pH 8.0; 30 min; 57 °C). Iodoacetamine (15 µL; 50 mM, final concentration) was added and alkylation was performed (30 min; RT in the dark). Digestion was performed by adding 10 µg of trypsin (T6567, proteomics grade, Sigma). Samples were incubated overnight (37 °C), centrifuged (30 min; 14,000 g) and injected (5 µL of acidified supernatants) into the tandem mass spectrometer with an electrospray source (LTQ Orbitrap XL Thermo Fisher Scientific, France) coupled to a liquid chromatography system (LC-ESI-MS/MS, Dionex Ultimate 3000, France).

2.6.2 High performance liquid chromatography (HPLC)

HPLC of the tryptic peptides was performed on a C18 micro-column at a flow rate of 50 µL min^{-1} with a linear gradient (10 to 80 % in 60 min) of acetonitrile and 0.1 % formic acid. Fractionated peptides were analysed in triplicate with an electrospray ionisation quadripole time-of-flight (ESI-QqTOF) hybrid mass spectrometer (pulsar i, Applied Biosystems) using information dependent acquisition (IDA), which allowed switching between MS and MS/MS experiments. Data were acquired and analysed with Analyst QS software (Version 1.1). After 1 s acquisition of the MS spectrum, the two most intense multiple charged precursor ions (+2 to +4) were selected for 2 s-MS/MS spectral acquisitions. The mass-to-charge ratios of the precursor ions selected were excluded for 60 s to avoid re-analysis. The minimum threshold intensity of the ion was set to 10 counts. The ion-spray potential and declustering potential were 5200 V and 50 V, respectively. The collision energy for the gas phase fragmentation of the precursor ions was determined automatically by the IDA based on their mass-to-charge ratio (m/z) values.

2.6.3 Nucleotide and amino acid sequence identity

MS/MS data were pooled (from triplicates) and used for database searches using an in house version of Mascot (Matrix Science, London, UK; version 2.1) and PEAKS

(Bioinformatics solutions Inc., Waterloo, Canada; version 7.0) search engines against the previously published *L. elliptica* transcriptome (Clark et al. 2010). LC-MS/MS data was searched using carbamido-methylation as a fixed modification and methionine oxidation as a variable modification. The peptide MS tolerance was set to 0.5 Da and the MS/MS tolerance was set to 0.5 Da.

Protein sequence characterisation was carried out using BLAST sequence similarity searches against the UniProtKB/Swiss-Prot database (www.uniprot.org). Signal peptides were predicted using SignalP 4.0 (www.cbs.dtu.dk/services/SignalP) and conserved domains from database models were predicted using external source database SMART (smart.embl-heidelberg.de).

2.7 Shell damage-repair

Many experiments in this thesis employed a shell-damage treatment. Non-lethal damage was inflicted to the shells in order to stimulate calcification pathways and hence repair the damaged shell.

Damage was inflicted by drilling a hole through the shell (either just inside the pallial line for very small juveniles or just outside the pallial line for larger adults close to the ventral edge of the shell and so within reach of the mantle margin), using a 10.8 V Lithium-Ion Dremel cordless modelling drill (model 800, variable speed, 5,000-35,000 rpm) fitted with a round-tipped bur to minimise any trauma to the underlying soft tissue. Damage was inflicted by a maximum of two experimenters to reduce experimenter variability. To inflict damage, animals were lifted out of the water, transferred to crushed ice on a bench in the aquarium where they were drilled, and returned immediately to their aquarium tank (holding conditions as per section 2.2). Animals were sampled at different time points during experiments (Figure 2-1). Category 1 included any hole covered by a clear mucous film, category 2 included holes covered by a brown, translucent, more proteinaceous film, and category 3 included holes covered by a brown film which shows signs of partial calcification and the final category 4 included holes which had a fully calcified layer.

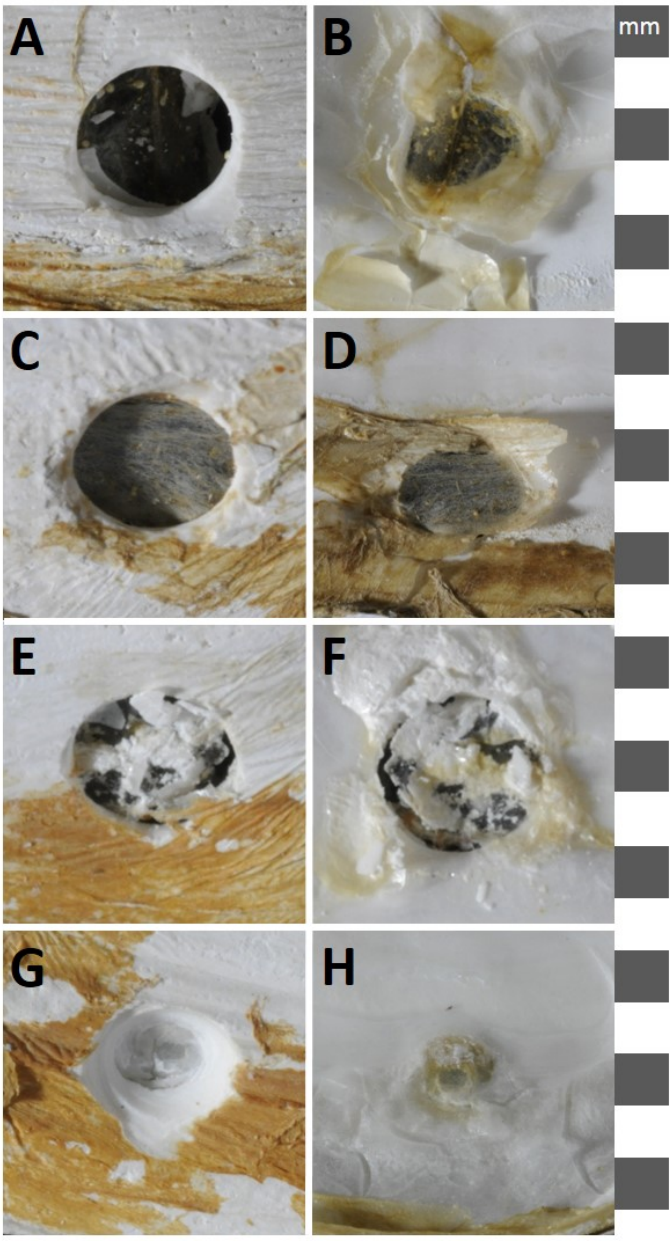

Figure 2-1. Photographs of shell repair categories, scale the same for all images as indicated by mm strip on the right-hand side. A.) Category 1, thin clear film viewed from outside of shell. B.) Category 1 viewed from inside of shell. C.) Category 2, translucent brown proteinacious film viewed from outside of shell. D.) Category 2, viewed from inside of shell. E.) Category 3, partially calcified brown film viewed from outside of shell. F.) Category 3 viewed from inside of shell. G.) Category 4, fully calcified layer, viewed from outside of shell. H.) Category 4 viewed from inside of shell.

2.8 Histology

2.8.1 Fixation, dehydration, embedding and sectioning

Adult *L. elliptica* tissues were processed as in section 2.5.5.1, the only modification being that sections were mounted onto standard glass microscope slides and dried overnight at 37 °C.

2.8.2 Tissue rehydration and staining

Adult *L. elliptica* sections were rehydrated as in section 2.5.5.3 and stained with a standard histology stain, haematoxylin and eosin (H&E). Deparaffinised sections were stained with Mayer's haematoxylin (1 %, VWR, U.K.) for 5 min and "blued" for 1 min in running tap water. Sections were rinsed in acid/alcohol (3 % v/v, Hydrochloric acid/ethanol) to clean the tissue without removing nuclear staining and subsequently stained with Eosin Y (1 %, VWR, U.K) for 30 sec. Sections were rinsed in tap water and dehydrated through a graded ethanol xylene series and mounted under a glass cover slip as per section 2.5.5.4.

2.8.3 Imaging

H&E stained sections were imaged under a Zeiss Microscope equipped with an AxioCam MR™ and captured with the software AxioVision™. For all histology analyses, three tissue sections were examined per individual and n = 4 individuals per treatment.

Images were analysed using ImageJ software (Schneider et al. 2012). Five metrics were measured: the total area of the outer epithelium on the fused inner mantle fold; the thickness of the outer epithelium on the fused inner mantle fold; the number of haemocytes in the fused inner mantle fold; the number of haemocytes in the field of view directly below the periostracal groove and the "mean redness index" of cellular organisation (Figure 2-2).

Using the objective magnification, lens magnification, c-mount, camera pixel size and camera binning settings, image pixels were used to measure length (Equation 2-2). In ImageJ, the known pixel length was set for each image and the thickness and total area of

the outer epithelia on the fused inner mantle fold were measures using the straight ruler tool and free hand tool respectively.

Equation 2-2

$$Image\ pixel\ size = camera\ pixel\ size\left(\frac{camera\ binning}{obj.mag.\times\ lens.mag.\times\ cmount}\right)$$

Roaming haemocytes (blood cells which roam within the body cavity to fulfill functions involved with the immune response) in the mantle tissue were counted using the cell counter in ImageJ . Two regions of the mantle tissue were chosen for counts: inside the fused inner mantle fold and the field of view directly below the periostracal groove as it was hypothesised these regions may be involved in the damage-repair process. The "mean redness index" of cellular organisation was developed in an attempt to quantify cellular organisation of the contractile fibres in the mantle tissue. For this index three regions of the mantle tissue were sampled, the left and right mantle fold and a region below the periostracal grooves. As the contractile fibres are stained homogenously pink by the eosin stain a histogram of the red channel was analysed in ImageJ and the mean value taken of each sample, the three samples were then averaged to give a mean average for the image. The higher the eosin index, the more red the image was and therefore the more fibres present. Densely packed fibres were disorganised, whereas neat and regular fibres were less densely packed and less intensely stained.

Statistics were performed using MiniTab 15, all histology metric data was tested for normality using the Kolmogorov–Smirnov test, where data met the assumptions of normality, a parametric Student T-Test was used to assess difference between damage and control treatments at each time point, where normality could not be achieved a nonparametric Mann Whitney-U test was used.

2.9 Transmission Electron Microscopy (TEM)

2.9.1 Fixation

Adult *L. elliptica* mantle tissues (n = 3) were fixed by immersion in 2 % formaldehyde (made from paraformaldehyde) and 2 % vacuum distilled glutaraldehyde, containing 2 mlL^{-1} calcium chloride, in 0.05 M sodium cacodylate buffer at 4 °C and pH 7.4. The fixative was made isotonic with sea water by the addition of sucrose. Tissues were then removed, sliced to 2-4 mm in one dimension and fixed for an additional 4-6 h at 4 °C.

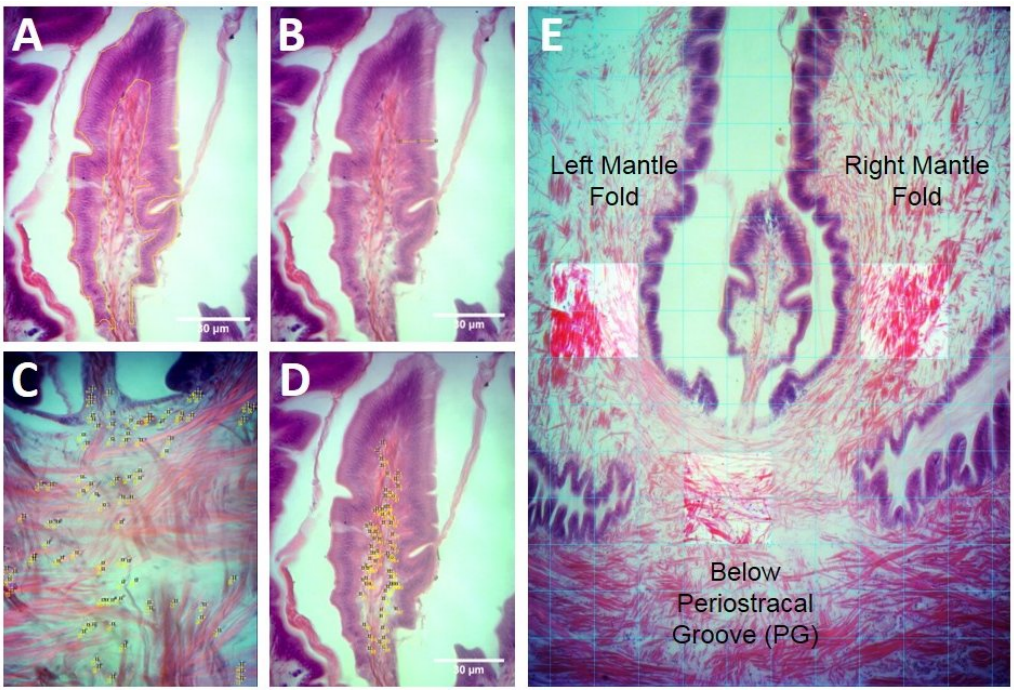

Figure 2-2. Demonstration of histology metrics. A.) The surface area of the outer epithelium on the fused inner mantle fold. B.) The thickness of the outer epithelium on the fused inner mantle fold. C.) Counting the number of haemocytes in the field of view below the periostracal grooves (black/yellow dots are haemocytes which have been counted). D.) Counting the haemocytes inside the fussed inner mantle fold. E.) Highlighted squares are the regions used to calculate the "mean redness index" of cellular organisation.

2.9.2 Processing and imaging

Tissues were then rinsed for 5 x 3 min in cold cacodylate buffer containing 2 mM calcium chloride and incubated in this solution with 1 % osmium ferricyanide for 18 h, at 4 °C and rinsed 5 times in deionised water (DIW). This was followed by 30 min in 1 % thiocarbohydrazide at room temperature (RT) followed by 5 more rinses in DIW. They were then incubated in 1 % uranyl acetate in 0.05 maleate buffer at pH 5.5 and at 4 °C overnight and rinsed with DIW at RT 5 x 3 min, with subsequent dehydration through twice each of 50 %, 70 %, 90 %, and 100 % ethanol, followed by twice each in dry ethanol, dry acetone and dry acetonitrile. Samples were then infiltrated with Quetol 651 epoxy resin over a period of 5 d. The resin was cured for 48 h at 65 °C. Thin sections, prepared with a Leica Ultracut S mounted on 200 mesh copper grids, were viewed with a TEM (Tecnai™ G2 Spirit, U.K.) operated at 200kV.

2.10 Immunology

Haemolymph (fluid equivalent to blood in most invertebrates) samples were extracted prior to sacrificing the animals. Animals were removed from the aquarium and transferred to the bench. A pair of forceps were inserted between the valves to force them open by approximately 5 mm and animals were inverted to drain any seawater from the mantle cavity. Haemolymph was withdrawn from the animals' adductor muscles via 25 gauge hypodermic needles, into an equal volume of 0.22 µm filtered sterile seawater containing 4% formaldehyde. The needle was then removed from the syringe and fixed cells were expelled into a 1.5 ml microcentrifuge tube and kept at 4 °C prior to counting.

2.10.1 Total cell counts

The total number of haemocytes per millilitre of haemolymph were counted using an improved Neubauer haemocytometer under a light microscope.

Total cell count data was tested for normality using the Kolmogorov–Smirnov test and a parametric Student T-Test was used to assess difference between damage and control treatments at each time point using MiniTab 15

Chapter 3 MANTLE AND SHELL CHARACTERISATIONS OF THE ANTARCTIC CLAM, *LATERNULA ELLIPTICA*

3.1 Introduction

The *Laternula elliptica* mantle transcriptome was first characterised in 2010 (Clark et al. 2010); a pooled RNA sample was subject to 454 pyrosequencing resulting in 18,209 assembled contigs, of which 17% were annotated. The most highly expressed genes in the transcriptome included several candidates which were "almost certainly involved in shell deposition". In addition, using Gene Ontology (GO) cellular component, Clark et al. (2010) suggested that 40% of annotated transcripts were likely to be secreted proteins, and hence potential candidates for shell matrix proteins (SMPs).

In order to develop an understanding of the expression and potential function of biomineralisation candidates sequenced by Clark et al. (2010) further work is required. Firstly, more knowledge is required on the biological structures involved. There is currently very little literature describing the *L. elliptica* mantle, the only study including *L. elliptica* mantle histology focussed on taxonomy, and arenophilic mantle glands, which are located near the siphon (Sartori et al. 2006). Once a more detailed understanding of the *L. elliptica* mantle tissue anatomy and ultrastructure is established, gene expression can be mapped more specifically to cell types and structure which could aid interpretation of putative function (Sasaki and Hogan 1993). In addition, there is currently no shell proteome available for *L. elliptica*, and it is therefore difficult to be certain that genes which are transcribed in the mantle are also translated into proteins and secreted into the shell matrix.

The objectives of this chapter are:

1.) To provide a detailed characterisation of the *L. elliptica* mantle anatomy and ultrastructure. 2.) Identify proteins in the *L. elliptica* nacreous shell layer using proteomics. 3.) To increase the understanding of the function of biomineralisation candidates by localising their gene expression to tissue, cellular and subcellular resolution.

3.2 Methods

Animal collection for the work in this chapter was carried out by SCUBA divers in Rothera. All proteomic analysis of the shell was carried out by Benjamin Marie, and sample fixation and processing for TEM was carried out by Jeremy Skepper. I carried out animal husbandry in Cambridge, all other laboratory work, data analysis and interpretation.

This chapter was published in Scientific Reports, 2016:

Sleight, V. A., Marie, B., Jackson, D. J., Dyrynda, E. A., Marie, A., Clark, M. S., (2016). An Antarctic molluscan biomineralisation tool-kit. Scientific Reports, **6**. 13 pp. 10.1038/srep36978.
http://www.nature.com/articles/srep36978

All methods for this work can be found in Chapter 2 (mantle anatomy and cellular ultrastructure characterisations via histology and TEM, candidate gene selection and gene expression analyses via semi-qPCR and *in situ* hybridisations and proteomic analysis of the nacreous shell layer via HPLC). This chapter will therefore include only a brief overview of experimental design and statistical analyses.

3.2.1 Quantification and localisation of putative biomineralisation transcripts

Seven putative biomineralisation genes were selected for tissue distribution expression profiling and *in situ* localisation as per Chapter 2 (Appendix A - Primer tables). Briefly, five were selected as they were present in the nacreous shell proteome and two were highly expressed in the previously published mantle transcriptome (Clark et al. 2010). Well-characterised biomineralisation candidates such as *pif*, *tyrosinaseA* & *tyrosinaseB* were selected, as well as less well-characterised biomineralisation candidates such as *Zn metalloendopeptidase*, *mytilin* and also two completely novel genes which had either no annotation (*contig 01043*), or only showed sequence similarity with two domains (concavalin-A and Lam-G , *contig 01311*).

3.2.1.1 Data availability

Contig numbers refer to the previously published *L. elliptica* mantle transcriptome, and the assembled contig set is available at: http://bit.ly/2cdR1eO and raw reads for assembly are available from the NCBI Short Read Archive Accession PRJNA79569.

3.2.1.2 Gene expression tissue profile

Reproductively mature animals (n = 5, mean shell length = 50 mm +/- 10 mm S.E.) were dissected into six different tissues (mantle, siphon, gill, foot, digestive gland and gonad). Gene-specific primers were designed and the *L. elliptica* 18s gene was used as a positive control and reference housekeeping gene for expression normalisation. Semi quantitative PCR (semi-qPCR) and normalised Integrated Density Value (IDV) calculations were carried out as per Chapter 2.

Gene expression data were checked for homogeneity of variance and normality using Levene's and Kolmogorov-Smirnov's tests respectively; all data met assumptions of homogeneity of variance and violated the assumption of normality. Data were transformed ($Log_{10}[X+2]$) but a normal distribution could not be achieved. Despite non-normal distribution of the transformed data, each tissue was compared using a General Linear Model Analysis of Variance (GLM-ANOVA) followed by post-hoc Tukey test. GLM-ANOVA can handle departures from normality and for added stringency, non-transformed data were also compared using a non-parametric Kruskal-Wallis (K-W) test. Differences between tissues were only considered significantly different if $P<0.05$ in both the GLM-ANOVA and the K-W tests.

3.3 Results

3.3.1 Mantle anatomy and cellular ultrastructure characterisation

The anatomy and ultrastructure of *L. elliptica* mantle tissue was characterised using standard histological staining, LM and TEM techniques and, based on the histological characterisations (Figure 3-1), a schematic illustration of the *L. elliptica* mantle was drawn to aid interpretation of the tissue (Figure 3-2). At the mantle edge *L. elliptica* have fused inner mantle folds, a periostracal groove, and what appear to be two outer mantle folds (Figure 3-1 & Figure 3-2). The mantle edge is responsible for producing the growing front of the shell, the two periostracum layers (an outer protective layer of proteinacious material which is secreted from the periostracal groove) and two shell layers – outer prisms and inner nacre. The enclosed space between the mantle and the shell is the extrapallial space. The mantle attaches to the shell at the pallial line; the mantle edge epithelial cells end and the contractile fibres of the mantle form an attachment in a line around the edge of the shell, this line continually moves with the growing front of the shell. On the dorsal side of the pallial line, the pallial mantle epithelial cells lay down nacre on the inside of the shell and control the shell thickness (Leonard et al. 1999). Inside the mantle tissue there are roaming haemocyte cells, contractile fibres and blood sinuses.

The epithelial cells of the mantle edge are columnar, with an elongate nucleus whereas the epithelial cells of the pallial mantle are more cuboidal with a large basal nucleus (Figure 3-1 & Figure 3-3). Both the mantle edge and pallial mantle epithelial cells have electron-dense vesicles (Figure 3-3).

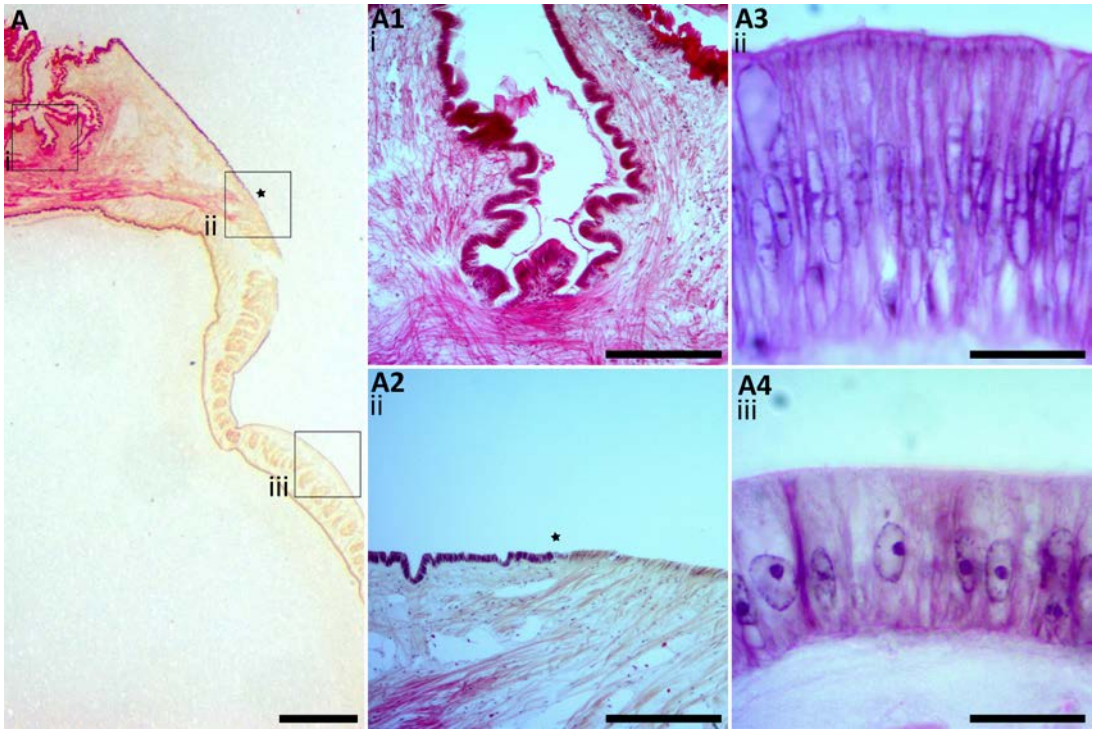

Figure 3-1. ***Laternula elliptica*** **mantle tissue stained with H&E. A = an overview of tissue anatomy, x0.63 objective, scale bar = 1.7 mm. A1 = fused in mantle fold (FIM) and outer mantle folds, x10 objective, Scale bar = 110 µm. A2 = a region of the right mantle edge, x10 objective, star indicates pallial attachment, scale bar = 110 µm. A3 = a region of mantle edge epithelium, x100 objective, scale bar = 11 µm. A4 = a region of the pallial mantle epithelium, x100 objective, scale bar = 11 µm.**

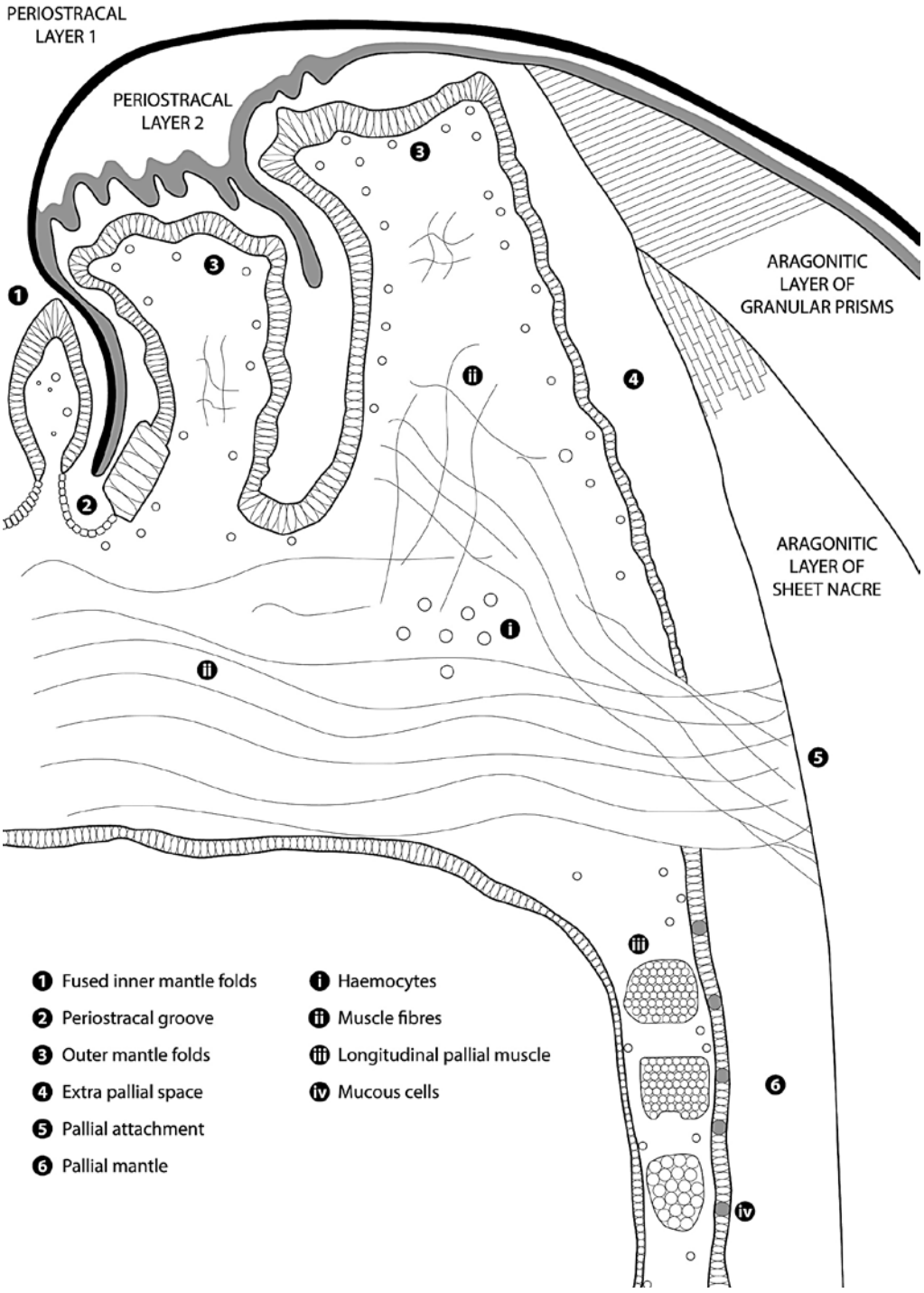

Figure 3-2. Schematic diagram representing *Laternula elliptica* mantle tissue anatomy, derived from H&E images in Figure 3-1. For illustrative purpose only, not to scale.

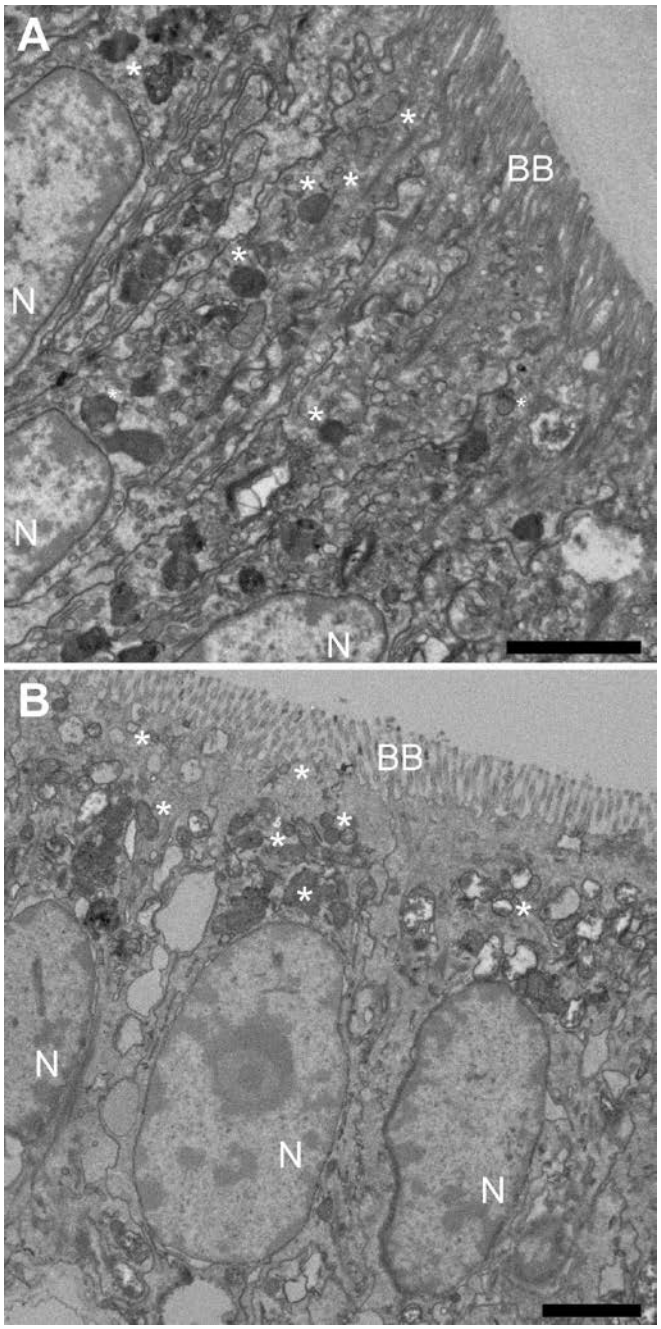

Figure 3-3. TEM of *Laternula elliptica* mantle epithelial cells. A = mantle edge epithelium cells, B = pallial mantle epithelium cells. Asterisk symbols show examples of secretory vesicles. N denotes nucleus and BB denotes brush border and the shell producing edge. Scale bars = 2 µm.

3.3.2 Nacre shell matrix proteins (SMPs)

Thirty seven proteins were identified in the nacreous layer of the *L. elliptica* shell (Table 3-1). Five conceptually translated contigs had a complete N-terminus and a signal peptide, suggesting that they were secreted by the mantle epithelia through a classic cellular secretion pathway. Twenty six proteins were detected with high confidence, either because they were detected independently by the two search engines or because they were identified by more than one peptide. The eleven other proteins were identified via one peptide.

Table 3-1. Identification of the nacre matrix proteins of *Laternula elliptica* by MS/MS analysis.

Contig reference	Proteomic identification score(nb of unique peptides)		Complete sequence / signalP	Sequence similarity		Protein feature
	Mascot	Peaks		Blast (species)	Smart domain search	
Contig01300	519 (9)	145 (4)	N/-	-	-	Novel uncharacterised protein fragment
Contig03872	496 (10)	147 (7)	Y/Y	-	-	Novel mucin-like protein
Contig02265	464 (7)	130 (3)	N/-	CA (*H. cumingii)*	CA	CA fragment
Contig01663	459 (7)	146 (7)	Y/Y	Tyrosinase 3 (*Crassostrea gigas*)	Tyrosinase	Tyrosinase
Contig01302	335 (6)	137 (4)	N/-	-	-	Novel uncharacterised protein fragment
Contig01311	301 (6)	120 (3)	N/-	Shell matrix protein(*M. californianus*)	2 concavalin-A + LamG-like	SMP-like
Contig17957	252 (4)	112 (4)	N/Y	CA (*C. midas*)	CA	CA fragment
Contig08650	223 (3)	103 (2)	N/-	-	-	LCD protein fragment
Contig01826	203 (4)	100 (3)	N/-	-	-	Novel uncharacterised protein fragment
Contig02037	196 (4)	92 (1)	N/-	Zn metalloendopeptidase(*L. gigantea*)	ZnMC metalloprotease	Zn metalloprotease fragment
Contig01312	194 (3)	104 (3)	N/-	-	-	Novel uncharacterised protein fragment
Contig01301	161 (3)	79 (1)	N/-	-	-	Novel uncharacterised protein fragment
Contig13708	161 (3)	-	N/-	-	-	Novel uncharacterised protein fragment
Contig03798	153 (2)	79 (2)	N/-	-	-	S-rich LCD protein fragment

Contig00830	147 (3)	76 (1)	N/-	Beta-hexosaminidase(*C. gigas*)	Glyco_hydro_20	Chitobiase fragment
Contig02085	135 (2)	76 (1)	N/-	CA	CA	CA fragment
Contig03967	131 (3)	-	N/-	-	-	Q-rich LCD protein fragment
Contig01785	91 (2)	60 (1)	Y/Y	Mytilin-3(*M. galloprovincialis*)	-	Mytilin-3
Contig02135	82 (2)	-	N/-	-	5 concavalin-A	Novel uncharacterised protein fragment
Contig02086	67 (2)	60 (1)	N/-	-	-	Novel uncharacterised protein fragment
Contig04690	61 (2)	-	N/Y	Insoluble matrix protein (*P. fucata)*	-	MSI60-like fragment
Contig06741	78 (1)	60 (1)	N/-	-	-	Novel uncharacterised protein fragment
Contig05798	70 (1)	54 (1)	N/-	-	-	Novel uncharacterised protein fragment
Contig01291	65 (1)	51 (1)	N/-	Papilin (*H. saltator*)	2 kunitz-like	Serine protease inhibitor fragment
Contig01036	61 (1)	46 (1)	N/-	P-U8 (*P. fucata*)	3 CCP + 3 concavalin-A	Novel uncharacterised protein fragment
Contig02084	-	104 (3)	N/-	CA	CA	CA fragment
Contig02500	62 (1)	-	N/-	Papilin (*S. mimosarum*)	2 kunitz-like	Serine protease inhibitor fragment
Contig05574	62 (1)	-	N/-	Chitin-binding protein (*P. martensii*)	Chitin-binding_3	Chitin binding protein fragment
Contig01703	-	84 (1)	N/-	Alpha-2 macroglobulin (*P. fucata*)	Alpha-2 macroglobulin	Macroglobulin fragment
Contig13709	-	68 (1)	N/-	-	-	Novel uncharacterised protein fragment
Contig01288	-	56 (1)	N/-	-	-	V-rich LCD protein fragment
Contig03070	-	50 (1)	N/-	Alpha-2 macroglobulin (*P. martensii*)	Alpha-2 macroglobulin	Macroglobulin fragment
Contig00332[#]	-	49 (1)	N/-	Uncharacterised protein (*L. gigantea*)	3 trombospondin + 3 concavalin-A + LamG	Novel uncharacterised protein fragment
Contig01639	-	47 (1)	Y/Y	Uncharacterised protein(*C. gigas*)	2 peritrophin-like	Novel uncharacterised protein
Contig00435	-	47 (1)	N/-	Uncharacterised protein(*C. gigas*)	2 VWA	Novel uncharacterised protein fragment
Contig18289	-	44 (1)	N/-	-	2 trombospondin	Novel uncharacterised protein fragment
Contig00456	-	44 (1)	Y/Y	-	-	Q-rich LCD protein

3.3.3 Tissue distribution of biomineralisation gene expression

Using semi-quantitative PCR, a mantle/siphon specific expression pattern was observed for four of the seven candidates [Figure 3-4, (*mytilin contig 1785* $f = 4.56_{29}$, $P = 0.005$,

chitin-binding contig 1311 f = 4.78_{29}, P = 0.004, *tyrosinaseB contig1359* f = 3.76_{29}, P = 0.012 and *unknown contig 1043* f = 8.75_{29}, P = 0.001)]. The remaining three contigs (*pif, Zn metalloendopeptidase and tyrosinaseA*) showed a low-level of expression across tissues and a mantle-specific signal was absent.

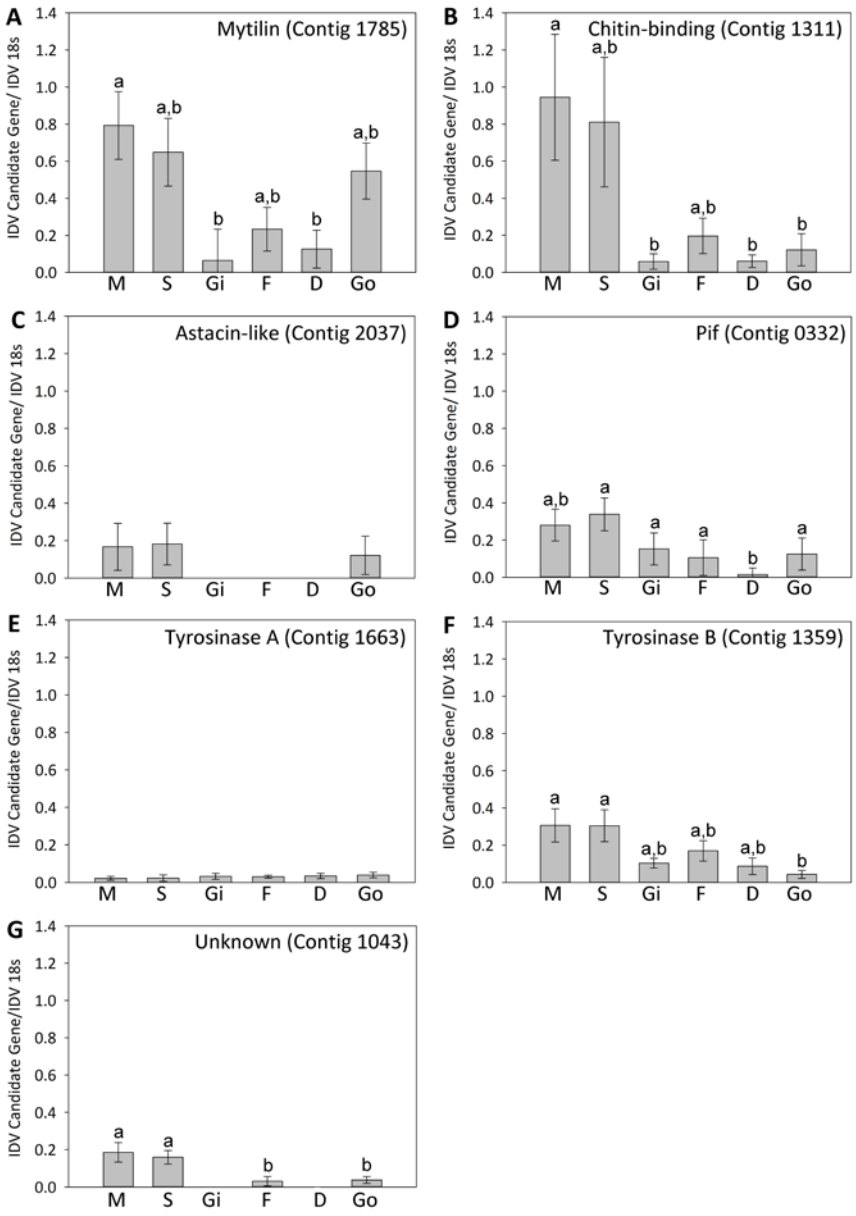

Figure 3-4. Expression of seven putative biomineralisation genes across six different *Laternula elliptica* tissues as determined by semi-quantitative PCR (mean ± S.E. n = 5). Statistically significant differences indicated by different letters above bars. M = Mantle, S = Siphon, Gi = Gill, F = Foot, D = Digestive Gland, Go = Gonad.

3.3.4 Cellular localisation of biomineralisation gene expression

In situ hybridisation of putative biomineralisation transcripts in adult mantle tissue sections revealed that different genes were expressed in different and discrete regions of the calcifying outer epithelium (Figure 3-5). Overall five major patterns were resolved:

A) Ubiquitous and continuous expression in the outer epithelial cells of the entire mantle edge and pallial mantle (*tyrosinaseA* & *tyrosinaseB*);

B) Continuous expression in the outer epithelial cells of the entire mantle edge and pallial mantle, which is much weaker in the fused inner mantle fold (FIM), and outer mantle folds (OM, *chitin-binding domain*);

C) Discrete expression in the outer epithelial cells at the edge of the OM next to the pallial attachment and continuous expression in the pallial mantle outer epithelial cells (*mytilin* and *Zn metalloendopeptidase*);

D) Discrete expression in the outer epithelial cells on the outer edge of the OM which becomes punctate and stops half way along the OM with no expression in the pallial mantle outer epithelial cells, the outer epithelial cells of the OM next to the pallial attachment, or the outer epithelial cell in the FIM (*pif*);

E) Continuous expression in the outer epithelial cells on the outer edge of the OM, with no expression in the pallial mantle outer epithelial cells, the outer epithelial cells of the OM next to the pallial attachment, or the outer epithelial cells in the FIM (*contig 01043* with no annotation).

3.4 Discussion

3.4.1 Mantle anatomy and cellular ultrastructure characterisation provides map for *in situ* hybridisation data

The mantle anatomy and cellular ultrastructure of *L.elliptica* was characterised to enable the accurate mapping of candidate biomineralisation genes to cell types using *in situ* hybridisation (Figure 3-1 & Figure 3-2). Knowing precisely where a gene is expressed - at a cellular and subcellular level – contributes to the interpretation of gene function (Sasaki and Hogan 1993).

The cellular ultrastructure described in the present chapter is similar to other mollusc species (Morse and Zardus 1996). Both the mantle edge and pallial mantle epithelial cells have electron-dense vesicles, some of which could contain calcium carbonate, that appear to be progressing to the cell apex to be deposited into the extrapallial space, where it is hypothesised that the calcium carbonate moves onto the extracellular protein shell matrix (Figure 3-3). Calcium carbonate containing vesicles progressing towards the biomineralisation site have been reported in the mantle epithelial cells of many mollusc species (Morse and Zardus 1996, Westermann et al. 2005, McDougall et al. 2011), as well as non-mollusc biomineral-producing species (Addadi and Weiner 2014, Vidavsky et al. 2014, Vidavsky et al. 2015). An important question regards the form in which calcium carbonate is carried inside the vesicles of biomineral-producing species: is it amorphous, organised, disorganised, solid, liquid, gel or indeed crystalline? Addadi and Weiner (2014) recently reviewed biomineral research and discussed the importance of fixation in determining how biominerals are observed, they recommended cryo-fixation as chemical fixation can change the state of biominerals in cells. For example, if calcium carbonate is present *in vivo* as unstable amorphous calcium carbonate (ACC), fixation can cause it to dissolve or crystallise. Due to logistical constraints the present chapter used gluteraldehyde fixations for TEM observations and therefore conclusions on the state or species of mineral inside the vesicles could not be made.

3.4.2 Analogous and unique nacre shell matrix proteins (SMPs)

The nacreous layer of the *L. elliptica* shell was subject to proteomic analysis and 37 SMPs were identified (Table 3-1). From the list of identified proteins, most of them share high sequence similarity with previously described mollusc shell proteins such as, carbonic anhydrase, tyrosinase, shell matrix protein, mytilin-3, MSI60, serine protease inhibitor, chitin-binding protein, macroglobulin, together with Q-, V-, and S-rich LCD, VWA, trombospodin, and CBD-2 bearing proteins (Marin et al. 2012, Gao et al. 2015, Mann and Edsinger 2014, Mann and Jackson 2014, Marie et al. 2013, Marie et al. 2012b). Additional *de novo* sequencing analyses of MS/MS peptides that were not involved in protein identification showed also the presence of M- and G-rich peptides (Chapter 3 Supplementary Table 1 – Appendix B Supplementary files). Previous reports have also observed that a M- and G-rich protein, called MRNP34, was present in the shell nacre of the pearl oysters (Marie et al. 2012a), but to date the function of such domains in SMP remains enigmatic. Taken together, this nacre SMP list supports the existence of a set of deeply conserved SMPs in bivalve nacre.

Whilst the majority of the proteins identified in this chapter were very similar to previously reported nacre proteins, two unique features were found in *L. elliptica* nacre. Firstly, one of the identified proteins contains a zinc-dependent metalloprotease domain which has not been reported in any shell matrix proteins to date. Secondly, *L. elliptica* nacre contained a novel mucin-like protein with remarkable T-rich composition that is not usually associated with nacre. Mucins are usually heavily glycosylated and sometimes sulfated proteins that are able to form multimeric insoluble hydrogels through cross-linking, in which nacre is believed to nucleate and develop, due to calcium and carbonate ion saturation (Marin et al. 2000). Another mucin-like protein (mucoperlin) was previously reported from the shell nacre of the fan mussel *Pinna nobilis* (Marin et al. 2000), but it has little reliable sequence similarity with the present *L. elliptica* mucin-like SMP.

3.4.3 Mantle-specific expression for some, but not all, candidate biomineralisation genes

A mantle/siphon specific expression pattern was observed for four of the seven candidate biomineralisation genes (*mytilin contig 1785*, *chitin-binding contig 1311*, *tyrosinaseB contig 1359* and unknown *contig 1043*), providing further evidence for their hypothesised role in shell deposition (Figure 3-4). Mytilin is an antimicrobial peptide which is produced by haemocytes and its high gene expression in the mantle is likely due to roaming haemocytes in the tissue (Mitta et al. 2000). *L. elliptica* has two copies of *tyrosinase* in its genome (*tyrosinaseA* & *tyrosinaseB*) which has been suggested to be the result of a duplication event followed by sub-functionalisation (Force et al. 1999, Sleight et al. 2015). Many mollusc species have multiple *tyrosinase* paralogues (Aguilera et al. 2014) and another bivalve, *Mytilus edulis*, contains at least two copies which respond differently to acidification stress (Huning et al. 2013). Tissue distribution expression profiling revealed *L. elliptica tyrosinaseB* had a mantle/siphon-specific expression pattern whereas *tyrosinaseA* showed no difference in expression across tissues and generally had a very low level of expression. Curiously, the proteome of *L. elliptica* shell nacre contained tyrosinaseA but not tyrosinaseB. Previous work on tyrosinases in *L. elliptica* using phylogenetic analysis of amino acid sequences showed that the two tyrosinases group in distant clusters (Aguilera et al. 2014). Phylogenetic differences and the tissue distribution expression patterns and shell proteome in the present chapter supports the hypothesis that the two copies of *L. elliptica tyrosinase* are carrying out different functions in the mantle.

The remaining three contigs (*pif, Zn metalloendopeptidase* and *tyrosinaseA*) showed a low-level of expression across tissues and a mantle-specific signal was absent. A peak of expression in the mantle is frequently a characteristic of biomineralisation genes (O'Neill et al. 2013) and it was surprising this pattern was absent for three relatively well-characterised biomineralisation candidates. Antarctic invertebrates such as *L. elliptica* have a low metabolism and grow slowly (Ahn et al. 2003), one explanation for the low expression of the biomineralisation candidates could be that the animals were not laying down shell at the time they were sampled, or that the rate of shell secretion is unusually slow and therefore difficult to detect at the transcript level (compared to other temperate molluscs which current characterisations are based on). The low-level of expression of

three biomineralisation candidates across tissues indicates these genes, and the proteins they encode, could be multi-functional and also demonstrates the need for higher spatial resolution in gene expression data, such as cellular localisation via *in situ* hybridisation, to contribute towards characterisation of molecular function.

3.4.4 Conserved mantle modularity provides the blueprint for a diverse array of mollusc shells

The mollusc mantle is anatomically modular in design and can be split into different regions which are thought to be responsible for secreting different layers of the shells, periostracum, prisms or nacre (Wilbur and Saleuddin 1983, Trueman et al. 1988, Morse and Zardus 1996). *In situ* hybridisation of putative biomineralisation transcripts in adult mantle tissue sections revealed that different genes were expressed in different and discrete regions of the calcifying outer epithelium (Figure 3-5). Overall five major patterns were resolved. The different and discrete gene expression patterns observed (regardless of their specific details and putative functions) provide further support for the hypothesis that the mollusc mantle is modular in design at the molecular level, as well as the anatomical level (Jackson et al. 2006). This modularity, observed at an anatomical and molecular level, acts as a "blueprint" or framework for molluscan shell production and gives rise to a huge diversity of architecture, microstructure and colour. Despite reports of rapidly evolving and diverse mollusc secretomes at the nucleotide and amino acid sequence level (Jackson et al. 2006, Jackson et al. 2007, Aguilera et al. 2014), the modularity of the mollusc mantle observed in this chapter is a conserved feature in many different shelled molluscs such as fresh water and marine gastropods and bivalves (Suzuki et al. 2004, Jackson et al. 2006, Fang et al. 2011, Gardner et al. 2011).

3.4.5 Functional understanding of SMPs

Spatial gene expression patterns can be used to infer putative gene function (Sasaki and Hogan 1993). Mapping gene expression onto different secretory regions of the mantle provides a powerful bridge into the functional understanding of how different genes control the production of specific features of the shell. *In situ* hybridisation revealed that both *L. elliptica tyrosinase* genes had intense expression in the entire mantle outer epithelium (mantle edge and pallium). Tyrosinase is involved in cross-linking of the

soluble periostracum precursor (the periostracin) to form an insoluble periostracum (Waite et al. 1979) and has previously been localised in the prismatic layer of shell (Nagai et al. 2007). Both of the *L. elliptica tyrosinase* paralogues were the only candidate genes to be expressed in the fused inner mantle fold, periostracal groove and entire mantle edge, which agrees with the previously described role of tyrosinase in the periostracum and prismatic shell matrix. Despite *tyrosinaseA* & *B* showing different tissue specificity and only *tyrosinaseA* being present in the nacre shell proteome (discussed above), both paralogues have the same spatial expression pattern within the mantle. Curiously, both genes are expressed in the pallial mantle (the region responsible for nacreous shell deposition) yet only tyrosinaseA is present in the nacreous shell proteome. TyrosinaseB could be involved in nacre formation in the extrapallial space but not become entrapped in the nacre matrix. The weight of evidence strongly suggests *tyrosinaseA* & *B* are carrying out different functions and *tyrosinaseB* could have many different roles in the pallial mantle without being a nacreous SMP (Esposito et al. 2012).

The *chitin-binding domain, mytilin* and *Zn metalloendopeptidase* genes showed similar *in situ* expression patterns and were all present in the nacre proteome, with *chitin-binding domain* and *mytilin* also having a mantle-specific tissue expression profile. All three genes were expressed along the entire pallial mantle epithelium and the outer side of the outer mantle fold; however *mytilin* and *Zn metalloendopeptidase* were restricted to a much smaller region, close to the pallial attachment, than the *chitin-binding domain*. Mytilin is hypothesised to be multi-functional and involved in both the shell matrix structure and as an anti-microbial peptide which forms part of the innate immune response (Mitta et al. 2000). It is possible the pallial attachment is a region vulnerable to the external environment and hence has a requirement for increased anti-microbials or, the region of the shell requires extra reinforcement to accommodate the pallial attachment and hence requires more shell matrix proteins.

Pif and *contig 01043* with no annotation showed no expression in the pallial mantle, this is particularly surprising for pif as it is classically thought of as a nacre protein (Suzuki et al. 2009, Suzuki et al. 2013) and indeed it was found in the nacre shell proteome. There are some possible explanations for the lack of expression of *pif* in the pallial mantle. Firstly, as previously suggested *L. elliptica* could be secreting shell much slower than other molluscs, or not at all at the point of sampling and previous work on *pif* expression

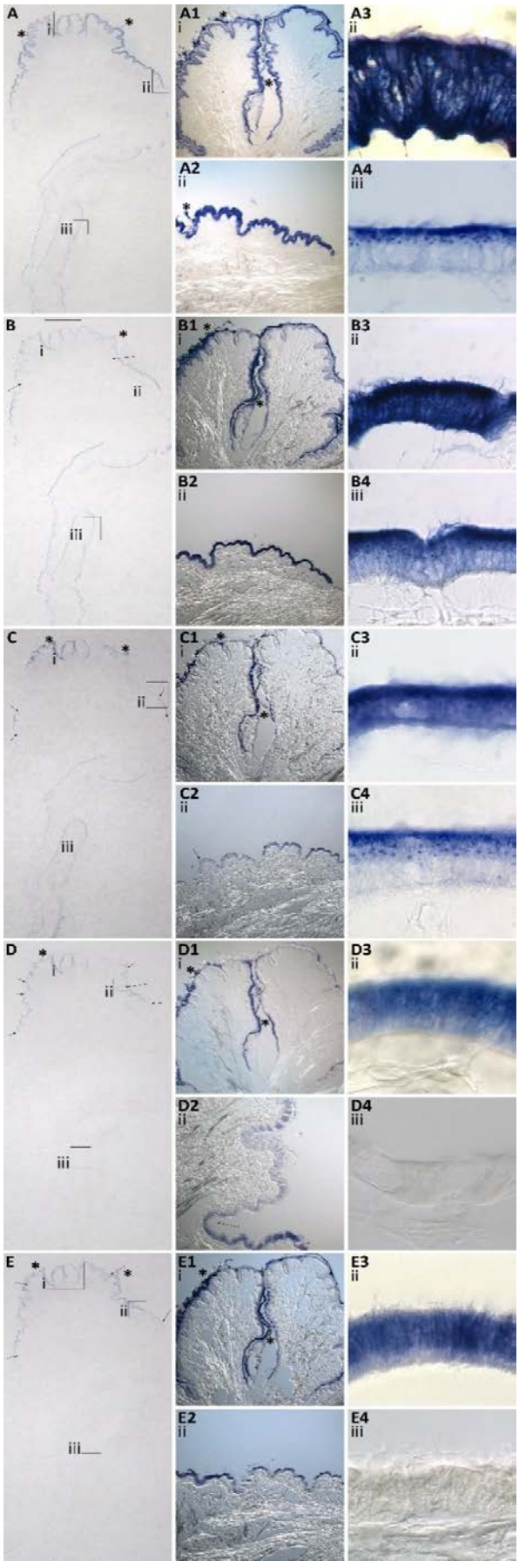

Figure 3-5. Modular mantle spatial expression patterns of *Laternula elliptica* putative biomineralisation genes. A = an overview of tyrosinaseA. B = an overview of chitin-binding domain. C= an overview of Zn metalloendopeptidase. D = an overview of pif. F= an overview of contig 01043 with unknown annotation. Boxes and roman numerals (i, ii, iii) indicate zoomed-in regions, arrows indicate expression boundaries and asterisk symbols denote extracellular organic material which is not expression signal. The following corresponding numbers match to the original overview: 1 = fused inner mantle folds (FIM) and periostracal grooves, x10 objective. 2 = a region of the right outer mantle fold of the mantle edge, x10 objective. 3 = a region of mantle edge outer mantle fold epithelium, x 100 objective and 4 = a region of the pallial mantle epithelium, x100 objective. For scale refer to Figure 3-1, for schematic representation of the tissue refer to Figure 3-2.

has shown it to be highly variable (Marie et al. 2013). Secondly, *pif* may only be involved in nacre deposition at the growing front of the shell, rather than increasing the thickness of the nacre layers in the pallial mantle. In addition, pif has been shown to interact with chitin (Suzuki et al. 2009) and it is surprising that its expression only co-localised with the *chitin-binding domain* expression at the outer edge of the outer mantle fold. *Contig 01043* has no similarity to previously characterised biomineralisation proteins and is not a known SMP, but given it is expressed in the same discrete set of cells in the mantle edge epithelial as *pif*, and that it shares a very similar tissue expression profile with *pif*, it could have similar cellular functions, thus ascribing a putative function to a previously "unknown" transcript.

Two subcellular localisation patterns were observed: ubiquitous strong expression signal in the entire epithelial cell with some vesicle staining (Figure 3-5 A3, B3, C3, D3 and E3), and expression only in the apical portion of the epithelial cell making vesicle staining easier to visualise (Figure 3-5 A4, B4 and C4). Cells in the mantle edge epithelia typically showed the ubiquitous subcellular pattern (with the exception of the punctate pattern of *pif*), whereas cells in the pallial mantle epithelium showed the apical subcellular pattern. All of the biomineralisation candidates showed subcellular expression signal in secretory vesicles. The H&E and TEM mantle characterisations of the *L. elliptica* mantle epithelium (Figure 3-1 & Figure 3-2) clearly show large basal nuclei in the epithelial cells with vesicles becoming more concentrated towards the cell apex, in addition many other molluscs (including various bivalves, gastropods and even shell-producing cephalopods such as *Nautilus pompilius*) show the same cellular ultrastructure in the mantle epithelium (Morse and Zardus 1996, Westermann et al. 2005, McDougall et al. 2011). Only one other study has investigated the subcellular localisation of biomineralisation proteins; Fang et al. (2008) used antibody protein labelling to observe calmodulin in the nucleus, endoplasmic reticulum and secretory vesicles of *Pinctada fucata* mantle epithelial cells. In the present chapter seven biomineralisation candidates localised to vesicles, and therefore they could be involved calcium carbonate transport via vesicle production, chaperoning or secretion.

3.5 Conclusions

Presented in this chapter is a detailed description of the tissues, proteins and genes involved in biomineralisation in the Antarctic clam *Laternula elliptica.* The mantle tissue anatomy and mantle epithelial cell ultrastructure were described revealing many conserved features with other shell producing molluscs, including secretory vesicles (which could contain calcium carbonate) that progress towards the shell. The proteome of the nacreous shell layer was characterised, 37 SMPs were identified, many of which corresponded to previously identified mollusc nacre SMPs, and two unique features were revealed; the presence of a zinc-dependent metalloprotease and a novel T-rich mucin-like protein. The expression patterns of seven candidate biomineralisation genes were further investigated to increase understanding of their potential functions. Four genes showed increased expression in the mantle and siphon tissues, and all seven genes had some expression in other tissues indicating they have multi-functional roles aside from biomineralisation. *In situ* hybridisation of the same transcripts revealed five different and discrete cellular expression patterns which corresponded to different secretory regions of the mantle, providing further evidence that the mollusc mantle is modular on a molecular as well as anatomical level. The different expression patterns also aided the putative interpretation of gene function in different aspects of biomineralisation. The subcellular patterns suggested that all seven biomineralisation candidates were associated with vesicles, the exact function of which is unknown, but they may be involved in calcium carbonate transport and secretion. Provided in this chapter is a multi-disciplinary analysis of the tools Antarctic clams use to produce their shells. In combination, analyses suggest that SMPs not only form the structural matrix required for calcium carbonate crystals to nucleate and grow in a highly organised and regular manner, but may also be important in the vesicular transport of biominerals and immunity.

This chapter used a mixture of traditional histology and microscopy techniques in combination with more modern ‘omics methods. Critically, it was this combination of approaches, which lead to the progressed understanding of biomineralisation and mantle functionality in *L. elliptica*. The methods used in this chapter have been especially useful for studying previously uncharacterised genes and proteins. Future work will use the characterisations described in the present chapter, and investigate how they change during shell damage-repair experiments.

Chapter 4 TRANSCRIPTOMIC RESPONSE TO SHELL DAMAGE IN THE ANTARCTIC CLAM, *LATERNULA ELLIPTICA*: PRELIMINARY TIME SCALES AND SPATIAL LOCALISATIONS

4.1 Introduction

In order to further study the molecular processes that control biomineralisation in *Laternula elliptica*, it is essential to know when and where they occur in the animal. Normal shell growth is thought to be a costly metabolic process (Palmer 1992); shell production can be a continual background process with gene (and protein) expression levels in the mantle likely to be low in abundance and therefore difficult to detect (Shi et al. 2013). One tractable method to highlight biomineralisation molecular mechanisms are shell damage-repair experiments that stimulate calcification pathways (Mount et al. 2004, Clark et al. 2013b, Huning et al. 2016). In one of the first shell damage experiments in recent history, Mount et al. (2004) notched the shells of *Crassostrea virginica* and investigated haemocyte-mediated repair 48 h later. More recently, Clark et al. (2013b) sampled *Crassostrea gigas* 1 week after damaging shells via drilling "a small hole". Using a single time point only allows for a snapshot window to capture the biomineralisation process. It can be difficult to ensure the timing of sampling is in line with any biological attribute being investigated, especially for molecular studies. For example, in a microarray experiment investigating scale regeneration (a type of biomineralisation) in sea bream, Vieira et al. (2011) found 769 differentially expressed genes after 3 days but only 21 differentially expressed after 7 days. If they had only sampled at 7 days, they would have missed the molecular processes which underpin the biological response they were investigating. Previous shell damage-repair studies were almost exclusively single time point or short-term experiments that provide a proof of concept that damage-repair can be a useful method to understand biomineralisation. Currently, the temporal and spatial components of shell damage-repair experiments are lacking.

The present chapter therefore aimed to better understand when and where molecular biomineralisation events occur in response to shell damage.

The objectives of this chapter are:

1.) To better understand the timing of biomineralisation in response to shell damage, in an Antarctic species. 2.) To identify candidate biomineralisation genes involved in response to shell damage. 3.) To understand the spatial location of molecular mechanisms in response to shell damage.

4.2 Methods

Animal collection for the work in this chapter was carried out by SCUBA divers at Rothera, aquarium experiments were designed and carried out by Melody S. Clark and Laura J. Weir and the transcriptome assembly and BLAST sequence similarity search was carried out by Michael A. S. Thorne. I carried out all of the laboratory work, data analysis (except assembly and BLAST) and interpretation.

This chapter was published in Marine Genomics, 2015:

Sleight, V. A, Thorne, M. A. S, Peck, L. S., Clark, M. S., (2015). Transcriptomic response to shell damage in the Antarctic clam, *Laternula elliptica*: time scales and spatial localisation. Marine Genomics, **20**. 45-55. 10.1016/j.margen.2015.01.009.

http://www.sciencedirect.com/science/article/pii/S1874778715000100

4.2.1 Experiment one: transcriptional profiling of damage response during a time series

4.2.1.1 Experimental design

Experiment one investigated the timing of molecular biomineralisation damage responses and was carried out over a two month time course. All of the animals were juvenile and not yet reproductively active. Half of the animals in the experiment were damaged (n = 14, mean shell length = 27.2 mm ± 1.2 S.E.), while the other half were left undamaged (n = 14, mean shell length = 23.97 mm ± 0.56 S.E.). Damage was inflicted by drilling a hole through the shell (as per 2.7). Samples of mantle tissue were subsequently taken at three time points: 1 week, 1 month and 2 months (n = 4-6 for each of the control and damaged treatments). There was no mortality in the experiment. Mantle tissues were dissected from each individual from directly under the drilled hole (or equivalent location in control animals), across the three mantle folds, snap frozen in liquid nitrogen and stored at – 80 °C prior to RNA extraction.

4.2.1.2 Sequencing

Total RNA was extracted from the mantle tissue of each animal as per Chapter 2. Equal volumes and concentrations of RNA from 4-6 individuals were pooled for each treatment and time point to make a total of 6 libraries (1 week control and damaged; 1 month control and damaged and 2 months control and damaged). The RNA was sent to the Earlham Institute, Norwich (formerly The Genome Analysis Centre [TGAC]) where TruSeq libraries were produced and subjected to 100 bp paired-end read sequencing on an Illumina Hi-Seq 2000.

4.2.1.3 Bioinformatics and statistics

A bioinformatic pipeline was developed to analyse the RNA–Seq data. The paired-end Illumina reads were assembled into a *de novo* transcriptome using Soap Denovo on default parameters (Luo et al. 2012). The newly assembled experimental transcriptome was combined with the previously published *L. elliptica* transcriptome (Clark et al. 2010), to produce a total of 42,807 contigs. All contigs were compared to the NCBI non-redundant (nr) database using Basic Local Alignment Search Tool (BLAST) to search for sequence similarity and putative gene annotation [$1e^{-10}$ cut-off, (Altschul et al. 1990). The paired-end reads from each experimental treatment were aligned to the transcriptome using Maq on default parameters (Li et al. 2008).

The aligned read counts produced in Maq were then analysed for differential expression. DEXUS (a Bioconductor package available in R that uses read counts as a finite mixture of negative binomial distributions) was used on default parameters (including a default normalisation step) to detect differential expression between control and damaged treatments at each time point (Klambauer et al. 2013). DEXUS uses a Bayesian approach to provide evidence of differential expression measured by the informative/noninformative (*I/NI*) value. The top 50 annotated, differentially expressed contigs for each time point were used for further analysis. Putative annotations were produced via sequence similarity searching against the NCBI nr database and confirmed manually using Blastx against both the Uniprot human database. In addition, seven "classic" candidate biomineralisation genes were selected using keyword searches and tracked through the DEXUS analysis. These "classic" biomineralisation candidates have

been previously characterised as shell matrix proteins (pif, nacrein, PfN44, perlucin, insoluble shell matrix protein) and information on their previous characterisations can be found in Chapter 1 - Table 1-1, and Chapter 2. Chitin synthases produce chitin, and chitin has been found in shells. More specifically, chitin synthase has been shown to be involved in shell biomineralisation in many species (Badariotti et al. 2007b, Schoenitzer and Weiss 2007, Ehrlich 2010, Weiss 2012). And the final candidate, dentin, is a protein involved in tooth biomineralisation (Arany et al. 2009).

To provide a visual qualitative assessment of the differential biological processes at each time point, the STRING v9.1 program was used (Franceschini et al. 2013). The tBlastx UniProt human annotations for the top 50 annotated differentially expressed contigs were entered into the STRING program. Using the protein-protein interactive network mode in STRING, all non-interacting proteins were removed from analysis and the remaining interactions were clustered using the Markov cluster algorithm (MCL).

4.2.1.4 Data availability

The RNA-Seq Illumina reads from the current chapter were submitted to the NCBI SRA (Sequence Read Archive), BioProject accession number: PRJNA268918. The updated assembled contig dataset – a combination of the previously published and the current experimental *Laternula elliptica* RNA-Seq data - is available for download from the Polar Data Centre (http://tinyurl.com/l5uvczh).

4.2.2 Experiment two: localisation of damage response

4.2.2.1 Experimental design

Experiment two investigated the spatial location of molecular events and was carried out over 1 week using two treatments; damaged, with a single hole half-way along the shell edge (method as per 2.7, n = 5, mean shell length = 68.2 mm ± 1.07 S.E.) and undamaged control (n = 5, mean shell length = 66.4 mm ± 3.06 S.E.). There was no mortality in the experiment. Mature, reproductively active animals were used in experiment two, which were larger and older than those in experiment one, in order to maximise the spatial distance between the areas of the mantle sampled. After 1 week, animals were sacrificed and mantle tissues were dissected (as above) into four regions, with two samples taken

from each region (Figure 4-1). Tissue samples were snap frozen in liquid nitrogen and stored at – 80 °C prior to RNA extraction.

4.2.2.2 Semi-quantitative PCR (semi-qPCR)

Total RNA was extracted from each section of mantle tissue (Figure 4-1) of individual animals (n = 5) and semi-qPCR was performed as per Chapter 2.

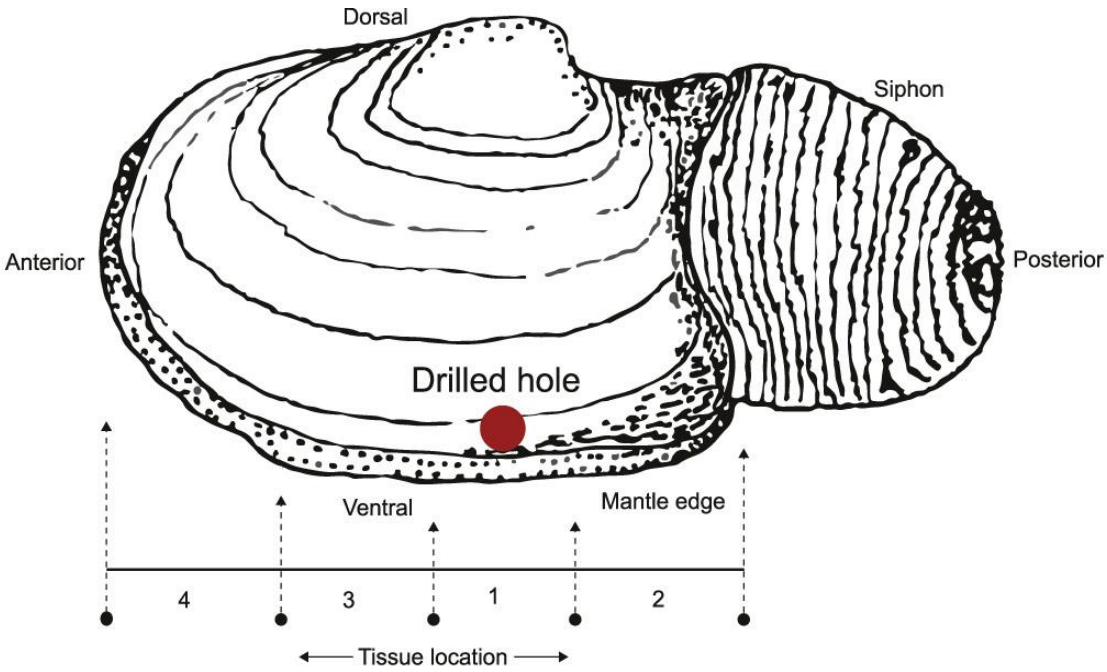

Figure 4-1. Schematic diagram showing the site of damage and the localisation sampling protocol in *Laternula elliptica* (experiment two). Two tissue samples were taken from each tissue location. Image not to scale.

A total of twelve candidate biomineralisation genes were selected to investigate the localisation of the repair response within the mantle tissue (Appendix A - Primer tables). More information on the candidate genes, and the reasons for their selection can be found in 2.5.1. Four additional transcripts, which were highly up-regulated in response to damage after 1 week using DEXUS in experiment one, were also included. Candidate genes were also isolated from the DEXUS analysis in experiment one to investigate their expression over time, thus providing data on both timing and localisation for a key set of transcripts, as well as linking experiment one with experiment two.

4.2.2.3 Statistical and transcriptomic analyses

Integrated density values (IDVs) from the GeneTools analysis were manipulated into an index for gene expression, as well as normalised, by division using a housekeeping reference 18s IDV per Chapter 2. 18s was confirmed as an appropriate housekeeping reference gene as there was no significant difference in IDV between samples. Gene expression data were not normal and could not be transformed to reach normality therefore, non-parametric statistical analyses were used. To test if the candidate gene was up or down-regulated in response to damage, data from control and damaged treatments were compared using a Kruskal–Wallis one-way analysis of variance by ranks. To test if there was a spatial pattern in expression across the four mantle regions in response to damage, data within the control and damaged treatments were compared using a Kruskal–Wallis one-way analysis of variance by ranks. Where statistically significant differences were found ($p < 0.05$) a non-parametric, ranked-sum, post-hoc HSD test was used to identify differences between regions. Statistical analyses for experiment two were carried out using MiniTab 15 and MATLAB.

4.3 Results

4.3.1 Experiment one: transcriptional profiling of damage response during a time series

The previously published, and new experimental transcriptome from the current chapter, were combined to make a total of 42,807 contigs. Thirty two percent of contigs were assigned putative functions using BLAST sequence similarity searching (below an e-value of $1e^{-10}$).

At the first time point in the experiment, one week after damage, only half of the animals had healed their shells. After one month however, all shell holes were occluded with at least an organic layer and after 2 months all organic layers appeared to be mineralised (Figure 4-2).

DEXUS analysis revealed differential expression between control and damaged groups, which varied over the time series (Table 4-1). 6046 transcripts (14 % of total) were differentially expressed at week 1, the highest number of differentially expressed transcripts was at 1 month (7402; 17 %) and the lowest was at the 2 month time point (2998; 7 %).

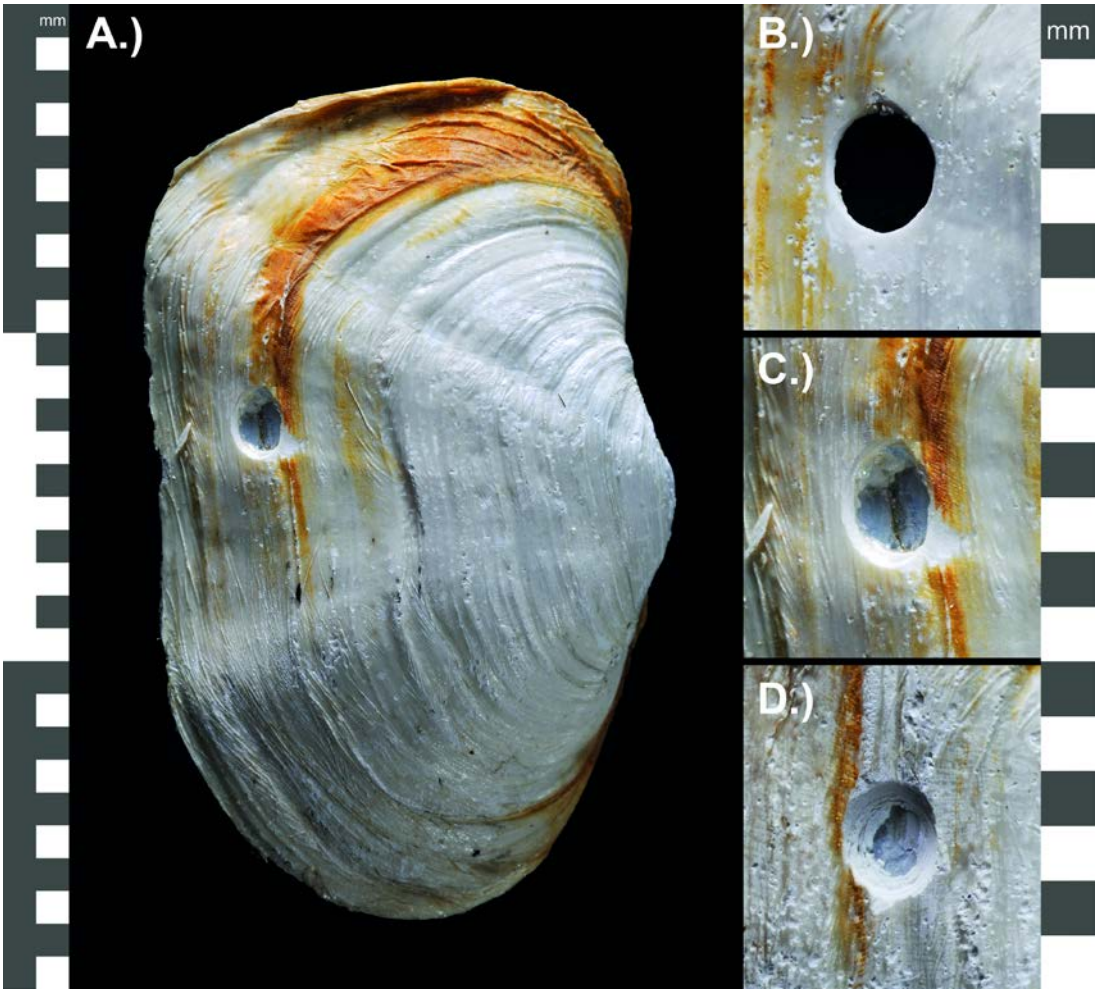

Figure 4-2. Photographs of representative damage-repair in *Laternula elliptica* at each time point during time series. A.) Example of a damaged shell, scale bar to left of image. Zoomed in photos of damage-repair B.) 1 week, C.) 1 month, and D.) 2 months after damage, scale bar on the right of images.

Table 4-1. Overall numbers of differentially expressed transcripts between control and damaged treatments at each time point according to DEXUS analysis of transcriptomic data (experiment one). Numbers in (bold) refer to annotated genes.

Time point	Number of contigs differentially expressed	Percentage of contigs differentially expressed
1 week	6,046 **(1,453)**	14.1% **(3.39)**
1 month	7,402 **(1,493)**	17.3% **(3.48)**
2 months	2,998 **(676)**	7% **(1.58)**

At each time point, the top (annotated) differentially expressed contigs between control and damaged treatments were different (Chapter 4 Supplementary Tables 1, 2 and 3 – Appendix B Supplementary files). Qualitative analysis using STRING indicated different biological processes were important at different time points (Figure 4-3). Specifically, after 1 week, DNA repair, immune response, RNA processing, cytoskeleton and mitosis were the dominant differential processes. After 1 month, respiratory electron transport chain and cell cycle regulation were most prominent and after 2 months apoptosis, protein folding, mRNA splicing and protein regulation dominated the differentially expressed processes.

The "classic" biomineralisation genes remained generally unchanged over the time series (Table 4-2; Chapter 4 Supplementary Tables 4 – Appendix B Supplementary files). After 1 week, only the *insoluble matrix shell protein* showed differential expression between control and damaged treatments. After 1 month, only *pfN44* was differentially expressed and after 2 months, *pif* and *chitin synthase* were differentially expressed.

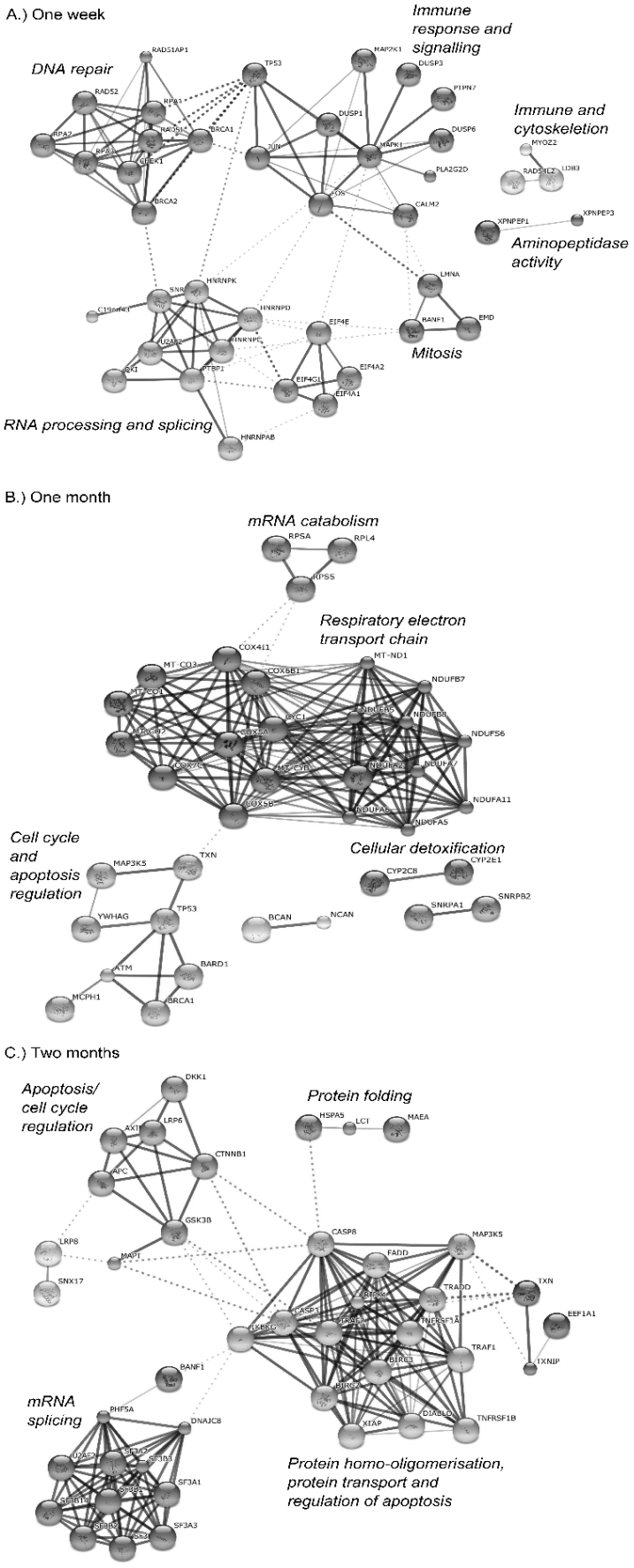

Figure 4-3. STRING Database predicted protein-protein interactions for top 50 annotated (<1e^{-10}) differentially expressed contigs between control and damaged treatments in experiment one after A.) 1 week B.) 1 month and C.) 2 months. Clustered using Markov cluster algorithm (MCL), thickness of line indicates strength of interaction, italicized labels indicate most likely biological processes or pathways.

Table 4-2. The differential expression status between control and damaged treatments of "classic" biomineralisation genes at each time point according to DEXUS analysis of transcriptomic data (experiment one). Contig Id of *pif* different to Table 4-3 due to multiple copies of *pif-like* genes in the *Laternula elliptica* transcriptome.

	1 week		1 month		2 months	
Contig I.D. / description	**Comparison**	***I/NI* value**	**Comparison**	***I/NI* value**	**Comparison**	***I/NI* value**
18958277 / Pif	Unchanged	$3.18e^{-14}$	Unchanged	$1.17e^{-06}$	**Different**	**1.16**
CL767 / Nacrein-like 3	Unchanged	$4.83e^{-08}$	Unchanged	$6.49e^{-08}$	Unchanged	$1.03e^{-07}$
18941698 / Perlucin	Unchanged	$3.10e^{-05}$	Unchanged	$7.15e^{-06}$	Unchanged	0.001
18977631 / PfN44	Unchanged	$7.93e^{-06}$	**Different**	**0.42**	Unchanged	$2.16e^{-14}$
18956919 / Chitin synthase	Unchanged	$1.94e^{-14}$	Unchanged	$2.23e^{-07}$	**Different**	**1.14**
18934508 / Dentin matrix protein	Unchanged	0.0003	Unchanged	$1.74e^{-06}$	Unchanged	0.0014
18972263 / Insoluble matrix shell protein	**Different**	**0.74**	Unchanged	$8.27e^{-15}$	Unchanged	$8.88e^{-05}$

The twelve candidate genes of interest identified in experiment two were also specifically investigated in the DEXUS analysis (Table 4-3). Ten out of the twelve genes showed a time-dependant expression pattern. After 1 week all genes of interest, except a *chitin-binding* gene, showed differential expression between control and damaged treatments. After 1 month, six of the genes were differentially expressed and after 2 months only four on the genes were differentially expressed.

Table 4-3. The differential expression status (between control and damaged treatments) of twelve candidate genes of interest, over time, according to DEXUS analysis of transcriptomic data (experiment one). Genes highlighted in bold failed to show time-dependant expression patterns.

	1 Week		1 Month		2 Months	
Contig I.D. ; description	**Comparison**	***I/NI* value**	**Comparison**	***I/NI* value**	**Comparison**	***I/NI* value**
Contig 02037; Astacin-like	Different	0.331	Different	0.250723	Unchanged	0.059
Contig 00332; Pif	Different	0.138	Different	0.252471	Unchanged	0.099
Contig 01359; Tyrosinase B	Different	0.335	Different	0.165317	Unchanged	$2.50e^{-13}$
Contig 01311; Shell matrix protein	Different	0.379	Unchanged	0.029	Different	0.115
Contig 01785; Mytilin	Different	0.133	Different	0.170	Unchanged	$5.43e^{-13}$
Contig 00041; Chitin-binding peritrophin	**Unchanged**	**0.062**	**Unchanged**	**$1.53e^{-13}$**	**Unchanged**	**$4.61e^{-13}$**
Contig 01043; Unknown	Different	0.127	Different	0.307	Unchanged	0.08
Contig 01663; Tyrosinase A	Different	0.251	Unchanged	0.078	Unchanged	0.025
Contig 18937102; Chitinase	Different	0.946	Unchanged	$4.29e^{-06}$	Different	0.693
Contig 1586; Tyrosine	Different	0.866	Unchanged	$1.06e^{-09}$	Unchanged	$2.86e^{-05}$
Contig 2930; Unknown	**Different**	**2.409**	**Different**	**0.420**	**Different**	**1.528**
Contig 3459; Unknown	Different	2.329	Unchanged	$7.92e^{-08}$	Different	1.344
Number of differentially expressed	11		6		4	

4.3.2 Experiment two: localisation of damage response

The response of twelve candidate genes to shell damage was investigated by comparing control and damaged treatments in experiment two by semi-quantitative PCR (Table 4-4). Three genes were unresponsive to damage, three genes were significantly up-regulated and the remaining six genes were significantly down-regulated.

Table 4-4. Expression response of twelve candidate genes to damage during spatial localisation experiment (experiment two). Comparison between control and damaged treatments using Kruskal–Wallis one-way analysis of variance by ranks to test for up or down-regulation.

Contig I.D. ; description	H	*df*	P	Response to damage across whole mantle
Contig 02037; Astacin-like	0.71	1	0.4	No change
Contig 00332; Pif	30.3	1	<0.001	↓ Down-regulated
Contig 01359; Tyrosinase B	25.04	1	<0.001	↑ Up-regulated
Contig 01311; Shell matrix protein	0.08	1	0.772	No change
Contig 01785; Mytilin	6.39	1	0.011	↑ Up-regulated
Contig 00041; Chitin-binding peritrophin	6.35	1	0.012	↓ Down-regulated
Contig 01043; Unknown	41.32	1	<0.001	↓ Down-regulated
Contig 01663; Tyrosinase A	25.2	1	<0.001	↓ Down- regulated
Contig 18937102; Chitinase	NA	NA	NA	No change
Contig 1586; Tyrosine	34.93	1	<0.001	↓ Down-regulated
Contig 2930; Unknown	13.82	1	<0.001	↑ Up-regulated
Contig 3459; Unknown	52.23	1	<0.001	↓ Down-regulated

To test if each candidate gene showed a spatial expression pattern across the four mantle regions in response to damage, data within the control and damaged treatments were compared (Figure 4-4). In every instance the control treatment showed no significant difference in expression across the four mantle regions. In contrast, in the damaged treatment, seven genes showed a significant spatial difference in expression across the four mantle regions. Each of the seven differentially expressed genes showed the same spatial pattern; a decrease in expression towards the siphon/posterior mantle edge (tissue

location 2, Figure 4-1) and an increase in expression away from the drilled hole at the foot/anterior edge of the mantle (tissue location 4, Figure 4-1).

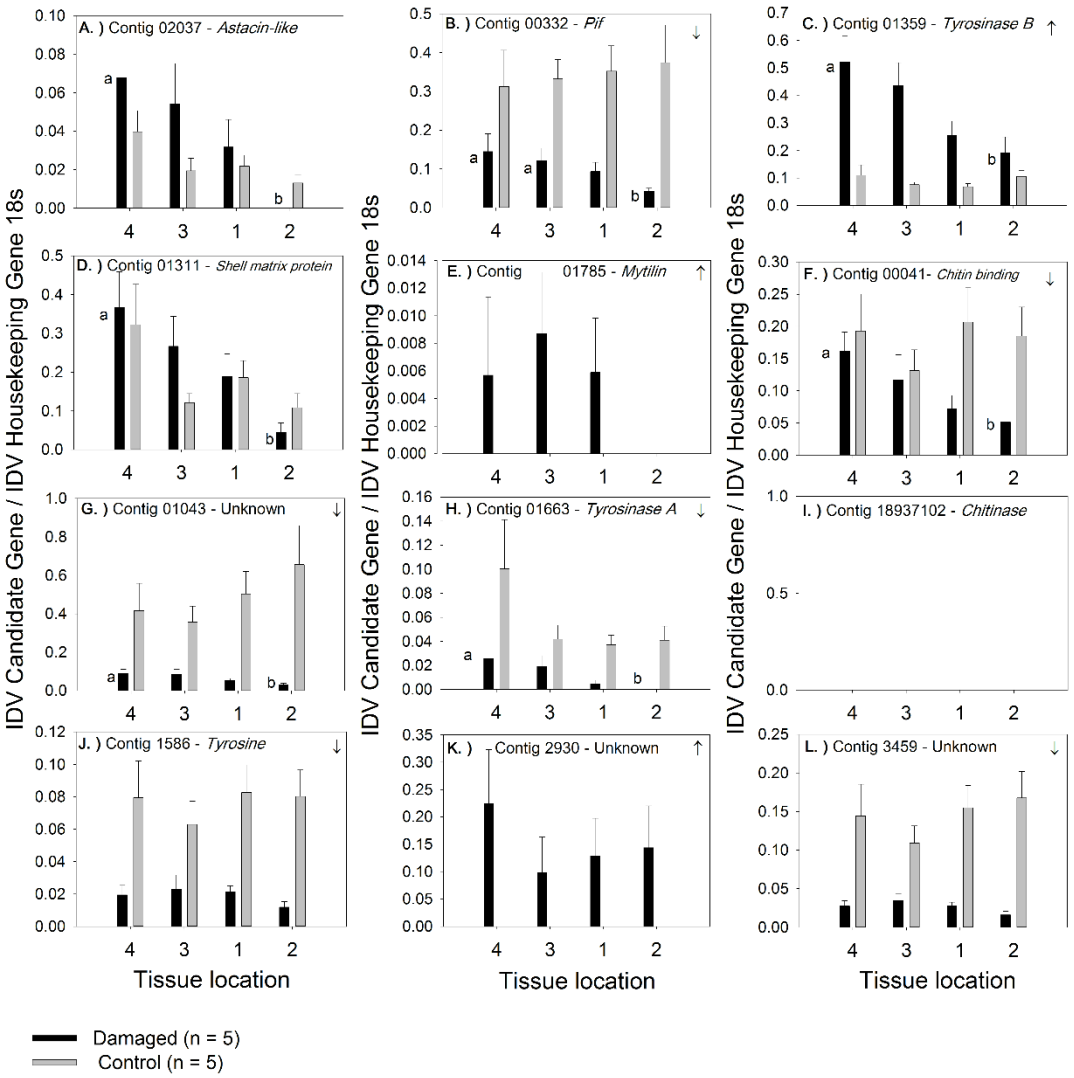

Figure 4-4. Expression of twelve candidate genes across four tissue locations in experiment two as determined by semi-quantitative PCR (mean ± S.E. n = 5). Statistically significant differences within treatments (across locations) indicated by different letters above bars. Statistically significant differences between treatments (up↑ or down↓ regulation between damaged vs control) indicated by arrow in top of plot.

4.4 Discussion

4.4.1 Experiment one: transcriptional profiling of damage response during a time series

Antarctic marine ectotherms are characterised by slow rates of growth, development and metabolism (Peck et al. 2006, Peck 2016) therefore it was unsurprising to discover that the transcriptional response of *L. elliptica* to shell damage was a slow process lasting at least 2 months. In a previous pilot study using reproductively mature clams of approximately 10 years age (60 mm ± 2.85 S.E.) and damage inflicted using the same drilling technique, hole occlusion took 4 months (Clark, pers comm). Animals in the present chapter were smaller (younger at approximately 4-5 years and immature) with thinner shells, and therefore were expected to repair faster. Qualitative analysis of differentially expressed transcripts over the time course showed the highest number occurred after 1 month, and the lowest after 2 months (Table 4-1). When the twelve candidate transcripts (from experiment two) were highlighted in the DEXUS analysis (Table 4-3), a decrease in differential expression over time was also observed. The reduction in differential expression over time is consistent with observations of shell occlusion, where the hole had occluded in 2 months in all damaged individuals (Figure 4-2). The pattern of differential expression over time was likely due to an initial immune/shock response to the damage, followed by rapid hole-filling, and then a slower response as the shell matrix above the closed hole was filled-in and strengthened. Hence, although there was a qualitative decrease in differential gene expression over time for both the overall differentially expressed transcripts (Table 4-1) and the twelve transcripts of interest (Table 4-3), substantial differential expression was still present after 2 months (7% of transcripts were differentially expressed between control and damaged treatments,Table 4-1), which indicated that shell repair was likely on-going.

Other studies have investigated shell repair in temperate bivalves over various, relatively short, time scales. Clark et al. (2013b) studied *Crassostrea gigas* 1 week after damage, Cho and Jeong (2011) also studied *C. gigas* with three sampling times, 7, 14 and 21 days after injury and, in one of the first documented shell damage-repair experiments, Mount et al. (2004) studied *Crassostrea virginica* only 48 hours after injury. Clark et al. (2013b) observed shell occlusion in all animals after 7 days. Cho and Jeong (2011) also observed

occlusion (as a white semi-transparent membrane) which was visible between 24 hours and 4 days after damage. Mount et al. (2004) observed haemocyte mediated hole-filling after 48 hours. Most recently Huning et al. (2016) drilled 9 small 1 mm holes into the shells of Baltic Sea *Mytilus edulis*×*Mytilus trossulus* hybrids and studied the shell-repair 20, 29 and 36 days after damage. The small holes were covered by a "fragile organic sheet approx. 0.5 mm thick with a periostracum-like light brown appearance" after 20 days and by 36 days aragonitic calcite structures were beginning to form over the holes. The shell repair of other mollusc groups has also been investigated. In a slightly longer-term study, Fleury et al. (2008) damaged *Haliotis tuberculata* shells and studied mantle histology 7, 14, 30 and 60 days after injury. They concluded the hole was occluded by a brownish organic lamella, which resembled a periostracal layer, after 1 day. In the present chapter, however, only half the damaged animals showed partial occlusion after 1 week and it took 1 month before all animals sealed the hole (Figure 4-2). As the present chapter used three time points spread over 2 months, data on the exact time and rate of occlusion were unattainable, but the occlusion process took a minimum of twice as long in the Antarctic clam compared to temperate molluscs.

Similar to the present chapter, Husmann et al. (2011) investigated transcriptional responses to shell damage in *L. elliptica* by collecting siphon and haemocyte samples 2 and 21 days post-injury. Shell damage was inflicted using a wrench and was a blunt, crushing motion. Husmann and colleagues did not investigate transcriptional differences between control and damaged groups over time, but instead their main findings were regarding effects of age and starvation on the ability for animals to repair. The present chapter however, demonstrates that 21 days is perhaps insufficient time capture Antarctic shell repair. Transcriptional responses to damage have also been studied in non-mollusc species. In a microarray experiment investigating scale regeneration (a type of biomineralisation) in sea bream, Vieira et al. (2011) found 769 differentially expressed genes after 3 days but only 21 differentially expressed after 7 days. Taken together with the results of the present chapter, damage-repair experiments emphasise the importance of timing in biomineralisation studies.

Qualitative analysis of the genes differentially expressed between control and damaged treatments highlights the different biological processes that are likely to be important at different time scales (Figure 4-3). The qualitative STRING analysis provides a visual

tool to consider processes which are important at different stages of the damage-repair process. The tool however, has some limitations. Firstly, the annotations used have to be from a model organism database, Uniprot human was chosen as it has the best annotation, therefore mollusc specific proteins, such as shell biomineralisation proteins will not be identified. Secondly, only proteins which interact are included in analysis; it is possible that some key processes are only represented by a single, non-interacting match in the top 50 annotated differentially expressed contains. To supplement the STRING analyses Blastx against the entire NCBI nr database (Chapter 4 Supplementary Tables 1, 2 and 3 – Appendix B Supplementary files) was also carried out. Although many of the annotations were conserved using both methods, some notable differences were identified, such as *contig 18955293* in Chapter 4 Supplementary Table 1(Appendix B Supplementary files) was annotated as hematopoietic prostaglandin D synthase which is involved in lipid metabolism when UniProt Human was used as the reference database, but when the NCBI nr was used it was annotated as glutathione S-transferase sigma 3 which is involved in detoxification.

4.4.1.1 One week after damage

In the immediate short-term after damage (1 week) STRING analysis highlighted differences between control and damaged treatments in typical stress responses such as: immune system activation, cytoskeleton processes, DNA repair and signalling (Tomanek 2011). The immune response was unsurprising as the external barrier of the animals was breached, allowing ingress of pathogens and potential infection, and indicated that it is likely the primary response was to combat disease and seal the hole (Jeffroy et al. 2013). The cytoskeleton has a role not only in morphological restructuring (such as cell proliferation) and intracellular protein transport (Porat-Shliom et al. 2013, Vindin and Gunning 2013, Mishra et al. 2014), but also potentially in the stress response at the 1 week time point (Tomanek 2011). Stress has long been linked to DNA damage, and a causal mechanism between stress and DNA damage has been found (Hara et al. 2011). Signalling molecules initiate biochemical pathways, such as stress responses, and this process is often referred to as a signalling cascade (Barolo and Posakony 2002). In the present chapter the primary initiator of signalling cascades was likely to be the damage inflicted. Damage is likely to activate the production of first signal molecules that trigger the production of further signalling molecules that serve to amplify the original stimulus

and result in the activation of effector molecules, which carry out the cellular response to the initial signal. The cellular responses observed were typical stress signals, as outlined above, and also more specific shell-repair signals, discussed below. Signalling is therefore, a critical early stress response mechanism that is important in the shell repair response in *L. elliptica*, and has been documented in the response to many different stressors in many organisms (Gitter et al. 2013).

The full NCBI nr annotation of differentially expressed transcripts highlighted additional transcripts of interest which were unresolved in the STRING analysis (Chapter 4 Supplementary Table 1 – Appendix B Supplementary files). Notably two transcripts that were likely to represent the activation of biomineralisation processes to occlude the hole: *calmodulin* and *fibulin-2*. Yan et al. (2007) demonstrated that calmodulin and calmodulin-like proteins modify the morphology of calcite and aragonite crystals *in vitro*. Calmodulin proteins have also been localised in the extracellular matrix of both prismatic and nacreous shell layers. Fibulins are a diverse family of extracellular matrix proteins, and fibulin-2 specifically contributes to elasticity of matrices (Timpl et al. 2003). Fibulins, so far, are uncharacterised in terms of any potential role in mollusc biomineralisation. Arany et al. (2009) however, showed that fibulin is involved in the biomineralisation of mouse odontoblast cells. Only one candidate "classic" biomineralisation gene *(insoluble shell matrix protein*) was differentially expressed after 1 week (Table 4-2), highlighting that the "classic" genes are not an exclusive set of molecules involved in biomineralisation, but are rather likely to be part of a more complex pathway or network. In order to understand the more complex system, of which the "classic" biomineralisation genes are a part, a discovery-led approach is required.

4.4.1.2 One month after damage

One month after damage, STRING analysis detected a major difference in respiratory electron transport chain processes between damaged and control treatments (Figure 4-3). Biomineralisation, and specifically the production of extracellular matrix proteins, is costly (Palmer 1983, Day et al. 2000). The difference in respiration between treatments is likely due to the energetic requirements for shell repair. Other differentially expressed processes included: mRNA catabolism, cell cycle/apoptosis regulation and cellular detoxification. The differences observed in mRNA catabolism and cell cycle/apoptosis

regulation could represent the re-configuring of cellular processes, from stress response to energy production required for repair. The differences in cellular detoxification between treatments could be linked to oxidative stress. Damaged animals were likely to be experiencing oxidative stress for two reasons: increased aerobic respiration and the macrophages oxidative burst during the immune response. STRING identified changes in the cytochrome P450 system, which is known to be involved in mediating oxidative stress. In addition, the full NCBI nr annotation of the differentially expressed transcripts (Chapter 4 Supplementary Table 2 – Appendix B Supplementary files) included *thioredoxin,* which is involved in cellular redox homeostasis (Jones and Go 2010).

At the 1 month time point, STRING analysis and the full NCBI nr annotations failed to highlight a biomineralisation signal; (Figure 4-3; Table 4-2 and Chapter 4 Supplementary Tables 2 – Appendix B Supplementary files). Only one candidate "classic" biomineralisation gene (*pfN44*) was differentially expressed. At this time point all the damaged shells had occluded with an organic layer and the initial stress of a breach to the external barrier was likely resolved. It is possible that biomineralisation to fill in the hole (mineralise the occluded layer) had not started, and instead animals were re-allocating cellular energy before biomineralisation could commence.

4.4.1.3 Two months after damage

After 2 months, the STRING analysis revealed differences in protein transport and turn-over, protein folding, mRNA splicing and cell cycle/apoptosis regulation (Figure 4-3). Protein turn-over indicated that the mantle tissue was metabolically active, with increased cell turnover potentially due to tissue remodelling in response to the damage.

Although the STRING analysis was unable to detect a biomineralisation signal at 2 months, the full NCBI nr annotation of the differentially expressed transcripts (Chapter 4 Supplementary Table 3 – Appendix B Supplementary files) highlighted four genes potentially involved in the production of the proteinaceaous extracellular matrix component of the shell: *collagen* (Benson et al. 1986), *calmodulin* (Fang et al. 2008), *fibrinogen* (Fang et al. 2011) and *low-density lipoprotein receptor* (Leupin et al. 2011). In addition, two candidate classic biomineralisation transcripts (*pif* and *chitin synthase*) were differentially expressed (Table 4-2). The differential expression of

biomineralisation transcripts between control and damaged treatments strongly supported the hypothesis that shell repair processes were on-going after 2 months.

4.4.2 Experiment two: localisation of damage response

Presented in this chapter is the first investigation into the spatial molecular response of molluscs to shell damage in mantle tissue. Genes differed in response to shell damage on a spatial scale and surprisingly, there was no peak in expression directly below the drilled hole (Figure 4-4). A consistent spatial expression pattern emerged for seven of the twelve genes investigated; a decrease in expression towards the siphon/posterior mantle edge and an increase in expression away from the drilled hole at the foot/anterior edge of the mantle. Currently, the reason for the observed spatial pattern can only be speculated. Mantle tissue can be functionally divided into at least three zones responsible for secreting nacre, prisms or periostracum (Jolly et al. 2004, Suzuki et al. 2004, Fang et al. 2011, Gardner et al. 2011, Sato et al. 2013); the expression of certain genes in each zone may be pre-programmed despite the influence of injury. Or, the expression pattern may be linked to the close proximity of the foot which protrudes outside the mantle when in use. One hypothesis could be that the anterior region of the mantle edge is more sensitive to damage and therefore has a higher induced response than other regions. The anterior region of the shell is the leading edge which buries into sediment and is where the foot protrudes. The movement of the foot in and out of the pallial space, coupled with burying into the sediment, could make the anterior region of the shell more susceptible to damage and therefore more sensitive.

Three candidate genes were up-regulated in response to shell damage (*contig 01359 – tyrosinaseB, contig 01785 - mytilin* and *contig 2930* – unknown; Table 4-4). Two of the up-regulated genes, *mytilin* and an unknown gene, showed no spatial pattern and were ubiquitously increased across the mantle edge. Mytilin is an anti-microbial peptide isolated from a marine mussel which is synthesised in, and transported by, haemocytes (Mitta et al. 2000). Mitta et al. (2000) found no expression of *mytilin* in the mantle tissue of control *Mytilus galloprovincialis*. It is likely the up-regulation and ubiquitous expression of *mytilin* detected in the present chapter was due to an infiltration of haemocytes into the mantle tissue as part of the immune response to injury. The spatial expression pattern shows that the *mytilin* immune response is unspecific to the injury site.

On the other hand, *tyrosinase B - contig 01359*, was up-regulated with a specific spatial pattern in relation to the injury site. Tyrosinase is involved in cross-linking of the periostracum, forming an insoluble outer layer protecting the shell; it is also involved in the pigmentation and formation of prismatic shell layers (Waite et al. 1979, Zhang et al. 2006). Similar to the present chapter Huning et al. (2016) and found one of the *Mytilus edulis tyrosinase* paralogues was hugely responsive to shell damage (damage via acidification and drilling respectively). Three candidate genes were unresponsive to damage (*contig 02037 – astacin-like*, *contig 01331 – shell matrix protein* and *contig 183937 – chitinase*; Table 4-4). *Chitinase* expression was not detected in the control or damaged treatment at any mantle region (Figure 4-4, I). Similarly Huning et al. (2016) found a chitin related gene (*chitin synthase*) was not responsive to shell damage in gene expression analyses. The negative *chitinase* and *chitin synthase* results in both experiment two of the present chapter and in recent shell damage experiments on *Mytilus sp* (Huning et al. 2016), could infer that chitin is absent from shell repair processes after 1 week; however, in experiment one, at 1 week in the time-series, *chitinase* was significantly up-regulated and hence was selected for further localisation in experiment two (Table 4-3). The function of chitinase and chitinase-like proteins have been investigated in arthropods with chitinous exoskeletons; they are typically involved in moult-cycles, wound healing and tissue repair (Chen et al. 2004, Bonneh-Barkay et al. 2010). Previous work showed *chitinase* expression is up-regulated in response to injury after 21 days in young, but not old, *L. elliptica* (Husmann et al. 2014), as well as in the oyster *Crassostrea gigas* (Badariotti et al. 2007a). Clark et al. (2013a) also found age to be a significant factor when investigating responses to hypoxia in *L. elliptica*. It is possible the present chapter found differential expression of certain genes, such as *chitinase*, in experiment one, but not two, because the size and age of the animals were different in the two experiments. The genes coding for an astacin-like and shell matrix protein were neither up nor down-regulated in response to damage, however both showed a significant spatial expression pattern in the damaged treatment, which was absent in the controls (Figure 4-4, A & D). Despite the lack of up or down-regulation, the presence of a spatial pattern in the damaged but not control treatment represents the likely re-configuration of molecular events across the mantle.

Six candidate genes were down-regulated in response to damage (*contig 1586 – tyrosine*, *contig 3459* – unknown, *contig 00332 – pif*, *contig 01043* – unknown, *contig 00041* -

chitin-binding peritrophin and *contig 01663 – tyrosinaseA*; Table 4-4). Two of the down-regulated genes, *tyrosine* and an unknown gene, showed no spatial pattern and were ubiquitously decreased across the mantle edge. The remaining four genes – *pif, chitin-binding peritrophin*, an unknown gene and *tyrosinaseA* – were down-regulated with a specific spatial pattern in relation to the damage site. Similar to the results discussed above, the spatial patterns observed are likely to represent the movement of cellular processes and functional zonation in the mantle.

4.5 Conclusions

In experiment one *Laternula elliptica,* shell repair was revealed to be a slow process lasting at least 2 months. Results highlighted different biological processes were important at different time scales during repair, thus enabling more targeted analyses in the future. For example, more sampling time points around the immediate short-term (e.g. 1-7 days) would provide a better understanding of the interaction between the immune, stress and biomineralisation responses. In addition, around 1 month after damage there appears to be a switch from energy production to biomineralisation, which is a critical step. Longer experiments, with a higher resolution of time-points, are also needed to fully characterise the biomineralisation component of the shell repair. As the time-scale of shell repair is unknown for most species, preliminary studies such as the one detailed here, provide a cost-effective approach to discover the most important molecular time-scales for the species being investigated. Knowledge on time-scales maximises sampling efficiency, as well as minimising sequencing costs; a particularly important factor for slow growing animals with enhanced longevity, such as those inhabiting the Southern Ocean, where biological processes are slow and age is a significant factor in response timings.

Although some biomineralisation signal was detected in the transcriptional profiling experiment, the signal was not dominant over other biological processes and the "classical" biomineralisation candidate genes were largely unchanged overtime. In order to better characterise and understand the molluscan biomineralisation pathway, the regulation of genes with the same expression pattern as "classic" biomineralisation genes should be investigated. Methods such as co-regulation and gene network analysis could

lead to the identification of novel genes and a better understanding on how biomineralisation genes interact forming a coherent pathway.

Experiment two highlighted the spatial localisation of repair mechanisms, different genes showed different spatial patterns in relation to a single drilled hole. The spatial pattern revealed is important for all future damage-repair experiments as the location and number of holes, in relation to where the underlying tissue is sampled from, could bias transcriptomic results. Future work investigating damage-repair mechanisms in large adults should drill several holes around the edge of the shell in order to stimulate expression across the entire mantle edge and maximise the tissue samples that can be taken for analysis.

There were some surprising differences in the gene expression response to damage between experiment one and two, for example, *chitinase* was up-regulated in response to damage in experiment one but not in experiment two. One explanation for this could be the different ages of animals in the experiments, immature in experiment one versus mature in experiment two. Future experiments therefore, should consider age as a factor when assessing the molecular control of shell production.

The pilot experiments in this chapter generated valuable data on the temporal and spatial response of shell damage-repair and therefore provide a baseline not only for future studies in *L. elliptica*, but also other molluscs.

Chapter 5 TIME SCALES IN THE *LATERNULA ELLIPTICA* RESPONSE TO SHELL DAMAGE: FROM PHYSIOLOGY TO A REGULATORY GENE NETWORK

5.1 Introduction

The previous chapters in this thesis investigated the tissues, cells, proteins and genes used for biomineralisation in the Antarctic clam and presented pilot experiments that generated data on the temporal and spatial response of shell damage-repair. Shell repair in *Laternula elliptica* was slow, lasting at least 2 months, and different biological processes were important at varying time scales. A spatial pattern in relation to a single drilled hole was revealed for some candidate biomineralisation genes, suggesting the mantle may be spilt into functionally specialised regions.

Previous work on *L. elliptica* has highlighted the importance of age when assessing population wide changes and responses to climate change (Clark et al. 2013a). *L. elliptica* can live for 36 years and, like many long-lived species, their muscle mass decreases with age as they become increasingly sedentary (Clark et al. 2013a). As a result of aging, the physiological flexibility of young clams diminishes as they develop. In addition age has been shown to be a significant factor in transcriptional response to injury and starvation (Husmann et al. 2014), as well as in the physiological response to disturbance, such as burying (Peck et al. 2004). It is therefore hypothesised that responses to shell-damage are also likely to be different in juvenile and adult animals.

To further understand the molecular control of shell damage-repair in the Antarctic clam, Chapter 4 suggests that experiments including more time points over a longer period of time are required. In addition, Chapter 4 suggests individual animals should be sampled for RNA-Seq libraries, rather than pooled, to allow for more sophisticated statistical analyses and computational modelling. Collecting data on physiological processes, such as shell repair, histology and immune cell counts, are also required to provide a framework in which to embed the molecular gene expression data.

The present chapter therefore aims to better understand the molecular response to shell damage in *L. ellipitca*, with a deeper, more comprehensive focus on time-scales and age.

The objectives of this chapter are to better understand: 1.) The timing of biomineralisation in response to shell damage in a Antarctic species 2.) The effect of age on biomineralisation and shell damage-repair mechanisms and 3.) The regulatory gene network underpinning biomineralisation in *L. elliptica.*

5.2 Materials and Methods

Animal collection for the work in this chapter was carried out by British Antarctic Survey SCUBA divers in Rothera, Antarctica. I carried out all of the experiments (except for the juvenile experiment which was extracted from previous work, more information below), laboratory work, data analysis and interpretation.

Detailed methods on animal collection and husbandry, experimental shell-damage, RNA extraction, shell-repair assessment, histology and immunology can be found in Chapter 2.

5.2.1 Experimental design

Three independent shell damage-repair experiments were conducted on three different age categories: juvenile (~ 1-2 years old), adolescent (~ 3-5 years old) and adult (> 6 years old), ages were calculated as per Philipp et al. (2008). The shell length of animals in the "juvenile" category were all ≤ 34 mm, "adolescent" shell lengths were > 34 mm - ≤ 45 mm and "adult" shell lengths > 45 mm (Figure 5-1). Juveniles were not reproductively active, adolescents were starting to develop gonads and adults were fully reproductively mature. Previous work on the *L. elliptica* population studied in this chapter showed that animals less than 38 mm in length have no gonad development (Powell 2001) and similar age categories defined by reproductive activity were used by Clark et al. (2013a).

The juvenile experiment was carried out June-August 2012 and the experimental design is outlined in Section 4.2.1.1. Briefly, animals were reproductively immature (n = 28, mean shell length = 25.6 mm ± 0.88 SE), half of the animals were damaged by drilling a single hole and the other half were controls. Samples of mantle tissue were subsequently taken at three time points: 1 week, 1 month and 2 months (n = 4-6 for each of the control and damaged treatments). In the previous chapter, the RNA from these samples were pooled to make 6 libraries (n = 1), in the present chapter however, each individual was used to make a new individual library (n = 3 for each of the control and damaged treatments), and a total of 18 libraries were re-sequenced.

The adolescent and adult experiments were carried out three years later in January-April 2015. Adolescent animals were developing gonads (n=32, mean shell length = 39.2 mm ± 0.56 SE), half of the animals were damaged with two shell holes and the other half were controls. Samples of mantle tissue were subsequently taken at four time points: 1 month, 2 months, 3 months and 4 months. Finally, adult animals were all sexually mature (n=48, mean shell length = 59.1 mm ± 1.9 SE), half of the animals were damaged with three shell holes and the other half were controls. Samples of mantle tissue were subsequently taken at six time points: 5 days, 1 week, 1 month, 2 months, 3 months and 4 months. For the adult animals, tissue and haemolymph samples were also fixed as per Section 2.5.5.1. The number of holes drilled in the shell was proportional to the shell length of the animal, for example only one hole would fit on the shells of the juvenile animals. In addition, results from Chapter 4 showed a spatial pattern in response to a single hole in large

Figure 5-1. *Laternula elliptica* age categories and corresponding shell lengths used in experiments, aproximate ages calculated using published growth curve (Philipp et al. 2008).

animals and therefore multiple holes were used to reduce the heterogeneity of the molecular response across the tissue.

5.2.2 Sequencing

To provide a buffer in case of mortality or RNA extraction failure, four individuals were sampled in each category (time x treatment) and three of those individual RNA samples per category were sent for library preparation. cDNA libraries were made for each individual in each of the age experiments (n = 78), library preparation was conducted by the Earlham Institute, Norwich, UK (formerly known as The Genome Analysis Centre). Stranded libraries were prepared using the NEXTflex™ Rapid Illumina Directional RNA-Seq Library Prep Kit and sequencing was carried out over 5 lanes on a Hi-Seq 2000 generating 125 base paired-end reads (with an exception of one library from the juvenile experiment which was sequenced on a MiSeq due to human error at the Earlham Institute, Norwich, UK).

5.2.3 Bioinformatics and statistics

All raw reads went through an initial quality control process conducted by the Earlham Institute that removed Illumina adaptor sequences and ribosomal RNA reads. The rest of the bioinformatic pipe-line was conducted in-house at the British Antarctic Survey, by myself and is summarised in Figure 5-2. The only exception being clustering and gene network algorithms run on servers at the Liverpool University Computational Biology Facility. Reads were further cleaned for quality (Phred score 30) and minimum read length (80 bp) using the ea-utils tool (v1.1.2) fastq-mcf (https://github.com/ExpressionAnalysis/ea-utils/blob/wiki/FastqMcf.md). The cleaned reads were normalised using Trinity's (v2.2.0) *In silico* Read Normalisation tool (Haas et al. 2013), with default parameters. The left and right reads for each library were normalised, all of the left and right reads were then concatenated and the resulting concatenated left and right read file was normalised a second time with default parameters. The concatenated, normalised reads were then *de novo* assembled using Trinity (v2.2.0) with default parameters (Grabherr et al. 2011). Trinity genes less than 300 bp in length were filtered from the assembly (Lenz et al. 2014).

The *de novo* transcriptome was assembled using reads from all 78 libraries, the downstream differential expression analysis however, was performed on each age experiment independently. Transcript abundance was estimated by alignment-based

quantification using Trinity (v2.2.0) utilities (Grabherr et al. 2011, Haas et al. 2013). Transcripts were aligned to the *de novo* transcriptome using bowtie with default parameters and transcript abundance estimates were calculated using RNA-Seq by Expectation-Maximization (RSEM). The gene-level abundance estimates (raw counts) for each of the libraries were constructed into a matrix for downstream expression analyses (using the Trinity abundance_estimates_to_matrix.pl script). Transcript abundance estimation was quality checked using the Trinity Perl-to-R 'PtR' toolkit and principal component analysis (PCA) was used to check for batch effects and outliers (Haas et al. 2013). For PCA analysis, Trimmed Mean of M-values [TMM] normalised Fragments Per Kilobase Of Exon Per Million Fragments Mapped [FPKM] values were used. Differentially expressed genes, between control and damaged animals at each time point, were identified using the Bioconductor (v3.4) edgeR package in R (v3.1.1) with a false discovery rate (FDR) of 5% (Robinson et al. 2010, McCarthy et al. 2012). A Log Fold-Change (LogFC) cut-off was not set as many of the time points had very few detected differentially expressed genes using the FDR < 5% cutoff alone. EdgeR outputs an estimator of relative gene expression called LogFC which is the unit used in many of the plots and tables in this chapter, this FC is set, as a default, to Log_2 of the ratio between treatments, in this experiment between control and damaged groups. For example, a LogFC value of 10 is a raw ratio value of 1024 [$Log_2(1024) = 10$] meaning there were 1024 times more normalised read counts in the damaged group compared to the control. In the same analysis, genes which showed significant temporal changes in response to damage, termed here, time-dependant damage-response genes, were identified (again using a FDR < %5 cutoff). The full R scripts for all downstream analyses are available in Appendix B Supplementary files (Chapter 5 Supplementary files 1, 2 and 3.).

The longest isoform of each gene was extracted from the transcriptome for annotation as per An *et al.* (2014) using the Trinity utility script "get_longest_isoform_seq_per". The longest isoforms of each gene were compared to a local NCBI non-redundant (nr) database (updated 01 June 2016) using Basic Local Alignment Search Tool (blastx, cut-off $<1e^{-10}$) to search for sequence similarity and putative gene annotation (Altschul et al. 1990).

A gene network was created from differential expression data from all 78 libraries. All libraries were aligned to the transcriptome, as described above, and the gene-level

abundance estimates (TMM normalised FPKM values) for each of the 78 libraries were constructed into one total mantle gene expression matrix, where each row was a trinity gene and each column a library. The data matrix was loaded into TM4 MultiExperiment Viewer (termed here TMeV - an open-source system originally designed for microarray data, which has recently been updated to handle RNA-Seq data) and clustered based on expression profiles for each gene (Howe et al. 2011). Within TMeV, SOTA (Self Organizing Tree Algorithm) was used for clustering (Herrero et al. 2001). SOTA is an unsupervised machine learning algorithm which constructs a dendrogram based on Euclidian distance, the terminal nodes in the dendrogram are the resulting clusters; i.e. genes which showed very similar expression patterns, or are "co-regulated", across the

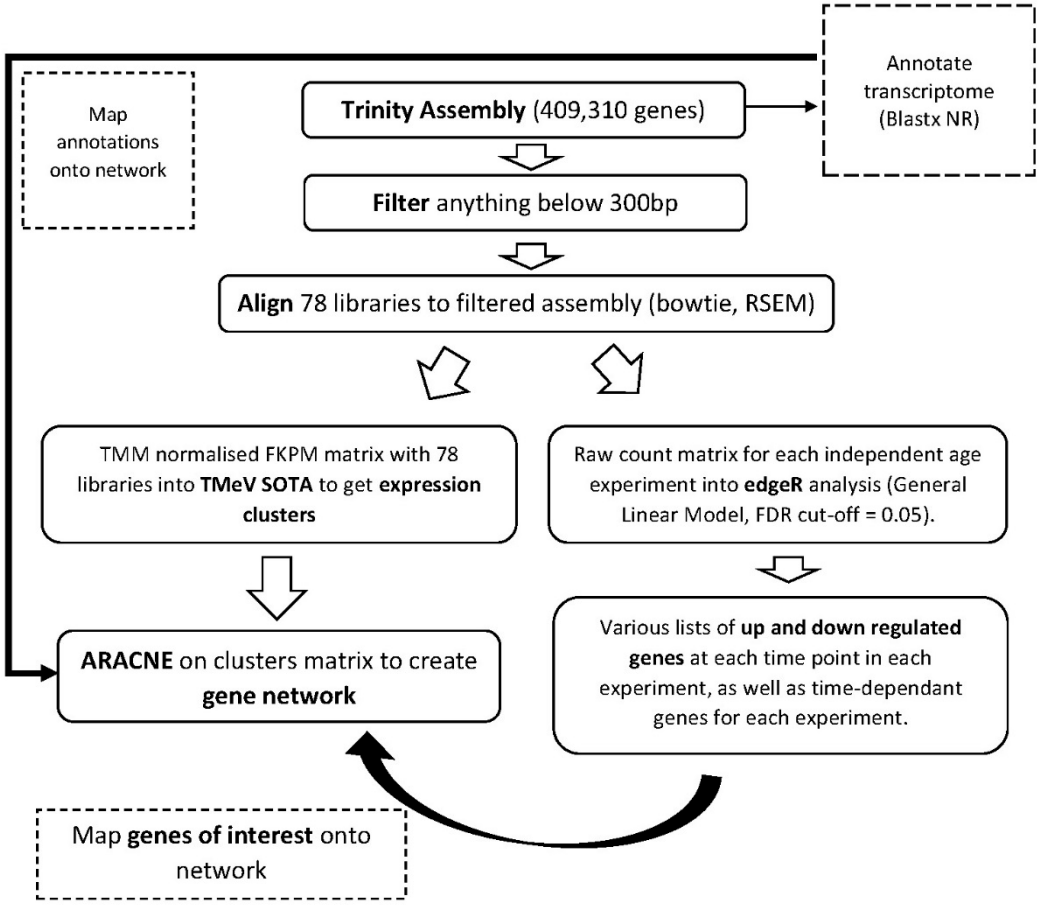

Figure 5-2. Flow diagram summarising the total bioinformatic pipline for analyses in this chapter.

78 libraries are collapsed into a single cluster. Clusters based on co-regulation/correlation (grouping together genes that exhibit similar transcriptional responses to various conditions) are a very coarse representation of a cellular network but cannot separate statistical interactions that are direct, from those arising from correlations between many non-interacting genes

To construct a more "data-informed" gene regulatory network, ARACNE (Algorithm for the Reconstruction of Accurate Cellular Networks) was used with a p-value cut-off of $1e^{-7}$ (Margolin et al. 2006). ARACNE is an information-theoretic algorithm, the edges which connect nodes in the network represent irreducible (i.e. direct) statistical dependency between gene expression profiles that cannot be explained as an artefact of other statistical dependencies in the network. The ARACNE algorithm uses mutual information to measure the degree of statistical dependency between two variables. Mutual information is estimated using a computationally efficient Gaussian Kernel estimator. Direct statistical dependencies in the network are therefore likely to identify biologically relevant regulatory interactions. The ARACNE network output was loaded in Cytoscape v3.4.0 for visualisation and exploration (Shannon et al. 2003). To identify highly interconnected sub-networks, GLay (clusterMaker) was applied (Morris et al. 2011). Gene annotations (Blastx nr, as above) were mapped onto nodes in the regulatory gene network.

Genes of interest were highlighted using edgeR (above) and mapped onto the network nodes. To identify direct interactions with these genes of interest, the first-neighbours of node, connected by edges or "direct statistical dependency" were pulled-out

5.2.3.1 Data availability

The RNA-Seq Illumina reads from the current chapter have been submitted to the NCBI SRA (Sequence Read Archive) and be made publicly available upon publication of the manuscript arising from this chapter (SRP115712). In addition, the assembled transcriptome, RSEM matrix dataset and R script will be made available for download from the Polar Data Centre.

5.3 Results

5.3.1 Physiological measurements

5.3.1.1 Shell repair

As per Chapter 4, in the juvenile experiment half of the damaged animals showed holes which were occluded and had a partially calcified layer after 1 week and it took 1 month before all animals occluded and partially calcified their holes. By 2 months 60 % of the animals had occluded and partially calcified holes, and the other 40 % were fully calcified. In the adolescent experiment, 1 month after damage, half of the holes were occluded and partially calcified, this rate of healing stayed consistent throughout damaged individuals at 2 and 3 months, with a slightly reduced rate in individual animals sampled at 4 months (Figure 5-3). In the adult experiment 5 days after damage 75 % of holes were occluded with a transparent mucus. One week and 1 month after damage, all of the holes were occluded and had a partially calcified layer. At 2 months after damage 40 % of holes were covered by a brown membrane and the remaining 60 % of holes were either partially or fully calcified. At the 3 month time point, all holes were partially calcified and at the final 4 month time point 15 % of holes were covered by a brown membrane and the remaining 85 % were either partially or fully calcified (Figure 5-3).

5.3.1.2 Histology and Immunology

Samples from the adult experiment were used for histology and immunology measurements (Figure 5-4). At the 5 day time point damaged animals had a significantly higher fused inner mantle fold surface area ($T_4 = 4.03$, $P = 0.016$) however, at all other time points there was no significant difference in fused inner mantle fold surface area. There was no significant difference in the thickness of the fused inner mantle fold between damaged and control animals at any time point. There were significantly more roaming haemocyte cells in the fused inner mantle fold of damaged animals at 3 months ($T_4 = 6.25$, $P = 0.025$), but there was no significant difference at any other time point. There was no significant difference between damaged and control animals in the number of roaming haemocytes below the periostracal groove at any time point.

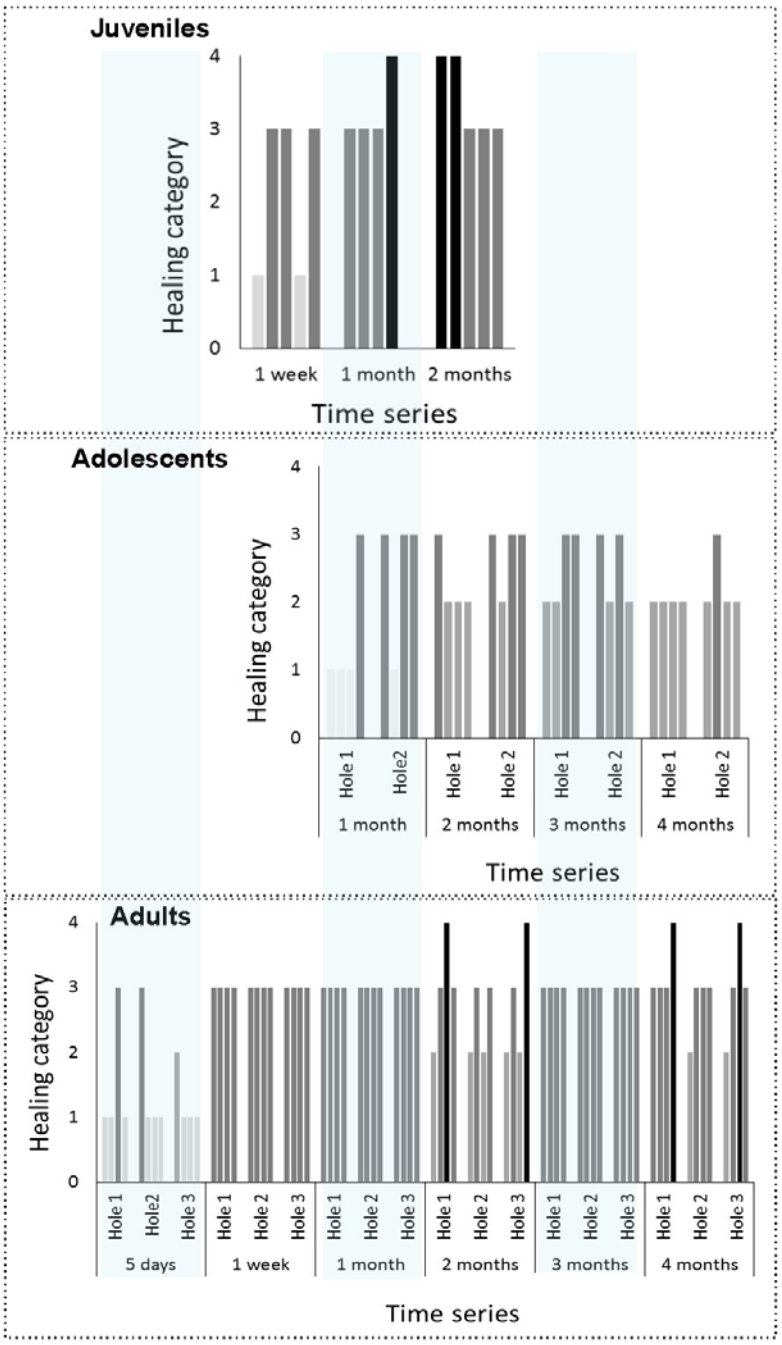

Figure 5-3. Shell repair observed in all age experiments, at each hole and time point, for each damaged individual. Healing categories described in Chapter 2.

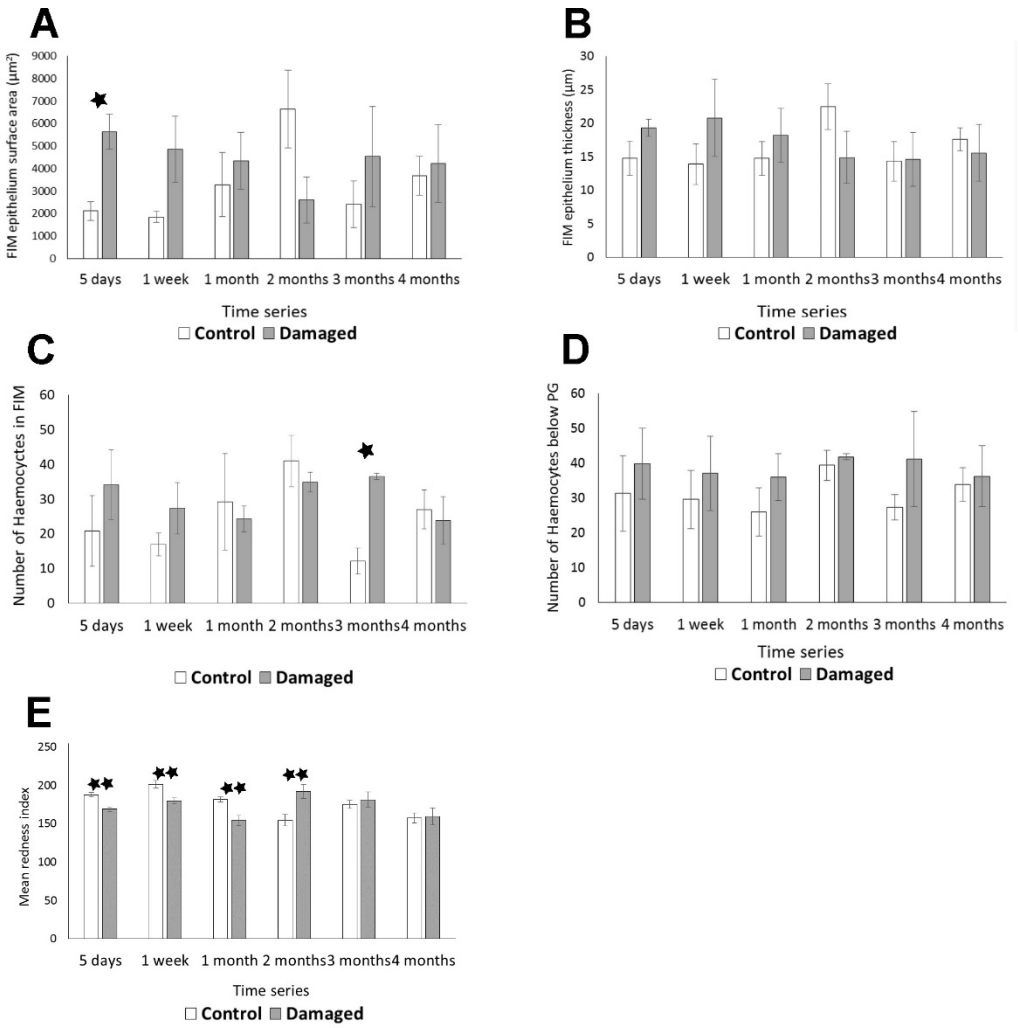

Figure 5-4. Histology metrics observed for adult animals (n = 4 ± S.E. per time point). A.) The surface area of the outer epithelium on the fused inner mantle fold. B.) The thickness of the outer epithelium on the fused inner mantle fold. C.) The number of haemocytes in the field of view below the periostracal grooves. D.) The number of haemocytes inside the fused inner mantle fold. E.) The "mean redness index" of cellular organisation. Single star = P<0.05, double star = P<0.01, histology metrics defined in Chapter 2.

At the 5 day (T_{21} = 4.47, P < 0.001), 1 week (T_{20} = 3.49, P < 0.002), and 1 month (T_{16} = 3.7, P < 0.001), time points, the mean redness index was higher in control animals than damaged animals. At the 2 month time point this pattern reversed and the mean redness index was higher in damaged animals (T_{11} = 3.17, P < 0.009), and for the remaining 3 and 4 month time points there was no significant difference.

There was no difference in total circulating haemocyte cell counts (TCC) between control and damaged individuals at any time point (Figure 5-5).

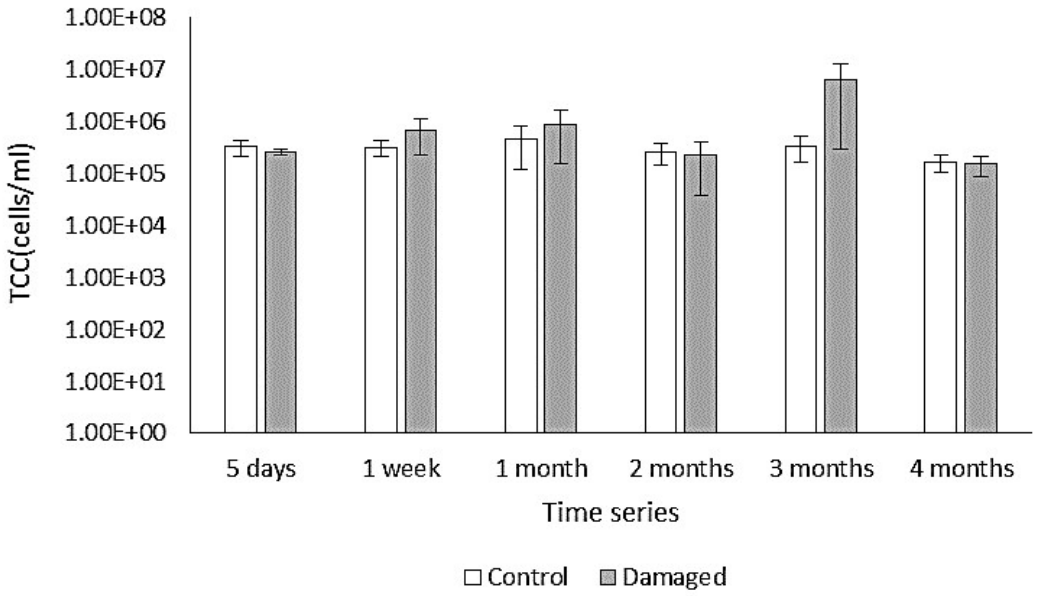

Figure 5-5. Total circulating haemocyte cells per ml of haemolymph (mean ±S.E., n = 4) in adult damaged and control animals at each time point.

5.3.2 Molecular measurements

5.3.2.1 Transcriptome

All 78 RNA libraries were cleaned, normalised and assembled into a *de novo* transcriptome, which was then filtered to remove very short transcripts (< 300 bp ,Table 5-1). Twenty four percent of Trinity genes (48, 236) were assigned putative annotation using Blastx sequence similarity searching against the NCBI nr database (below an e-value of $1e^{-10}$).

Table 5-1. Assembly statistics for the *Laternula elliptica* mantle *de novo* transcriptome

Reads:	
Raw reads	916,012,670
Clean reads (q30, l80)	842,890,608
Normalised reads (K25, C50, pctSD200)	58,610,296
Filtered Assembly:	
Total trinity transcripts	324,119
Total trinity genes	199,321
% GC	38.7
Statistics based on longest isoform per gene:	
N50 (bp)	892
Median length (bp)	457
Mean average length (bp)	746

5.3.2.2 Differential gene expression

Transcript abundance estimates were quality checked using PCA analysis and deemed suitable for further analysis (Figure 5-6, Figure 5-7, Figure 5-8) because; there was no clear clustering for time (Figure 5-7) or time×treatment (Figure 5-8), and there were also no clear outliers or batch effects, which simply indicates high levels of biological variation between individuals. Clustering for treatment was evident and was most prominent in the adult experiment (Figure 5-6).

EdgeR was used to identify differentially expressed genes (FDR < 0.05) between control and damaged animals at each time point for each age experiment (Figure 5-9, Chapter 5 Supplementary Tables 1-11 – Appendix B Supplementary files). Overall, there was very little differential gene expression in juvenile and adolescent animals compared to adults. Juveniles showed no response to damage at any time point, adolescents had a peak in response to damage at 2 months (116 significantly up-regulated genes and 108 significantly down-regulated genes, FDR < 0.05) and the adult response increased over time, reaching a peak in response 4 months after damage (3,513 significantly up-regulated genes and 3,212 significantly down-regulated genes, FDR < 0.05).

Three key time points were selected for further study and comparison: adolescent 2 months and adult 3 & 4 months. The significantly up-regulated genes were compared between these time points to find shared and unique genes in the damage response at different times and ages (Figure 5-10). A core set of 42 up-regulated damage-response genes were identified between the three selected time points, of these 42 genes, 21 were assigned putative annotation via sequence similarity searching (Blastx against nr, Chapter 5 Supplementary Table 12 – Appendix B Supplementary files). In addition to a core conserved set of significantly up-regulated damage response genes, each time point investigated also had a large suite of unique genes. For the adolescents at 2 months 34 % of genes were unique, for adults and 3 and 4 months 18 % and 72 % respectively were unique, again these genes were assigned putative annotation via sequence similarity searches (Chapter 5 Supplementary Tables 13-15 – Appendix B Supplementary files).

In addition to investigating differential expression within each time point, genes that showed a significant temporal change across time points, in response to damage, termed

time-dependant damage-response genes, were also identified in each of the three age experiments. Juvenile, adolescent and adults had 89, 1,278 and 8,417 significant (FDR<0.05) time-dependant damage response genes, respectively (the time-dependant genes which could be annotated can be found in Appendix B Supplementary files [Chapter 5 Supplementary Tables 16-18]).

The time-dependant damage-response genes for each age experiment were compared to find conserved and unique genes. Only 4 time-dependant damage-response genes were shared between the three experiments, 1 of which was assigned a putative function via sequence similarity searching (TRINITY_DN252582_c0_g3, cAMP-dependent protein kinase type II regulatory subunit, Chapter 5 Supplementary Table 19 – Appendix B Supplementary files). For the juvenile, adolescent and adult experiments, 54 %, 47 % and 92 % of time-dependant damage response genes, respectively, were unique, again these genes were assigned putative annotation via sequence similarity searches (Figure 5-11, Chapter 5 Supplementary Tables 20, 21 and 22 – Appendix B Supplementary files).

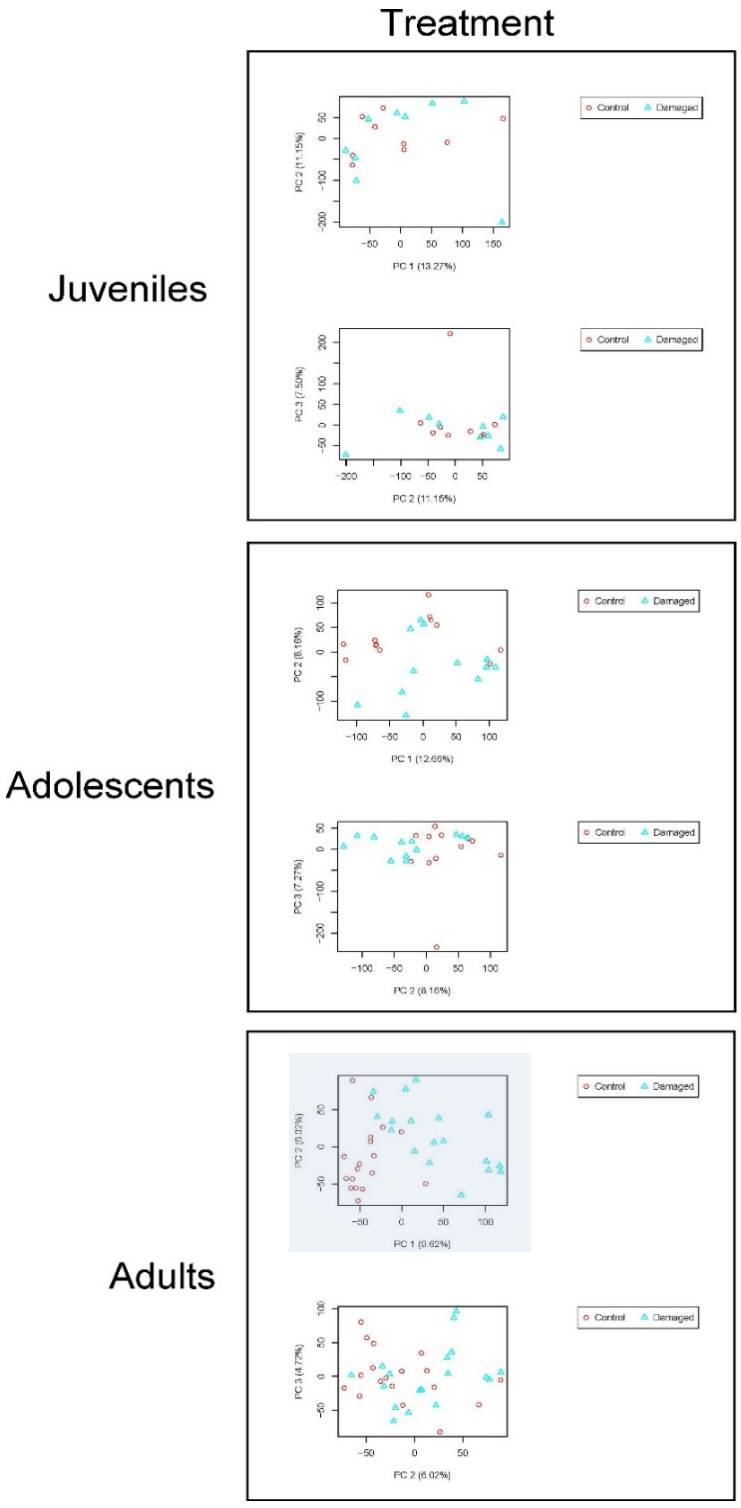

Figure 5-6. Principal component analysis of transcript abundance estimates for each library (TMM normalised FPKM values). Labelled to highlight just the effect of treatment, which was most prominent in adult animals.

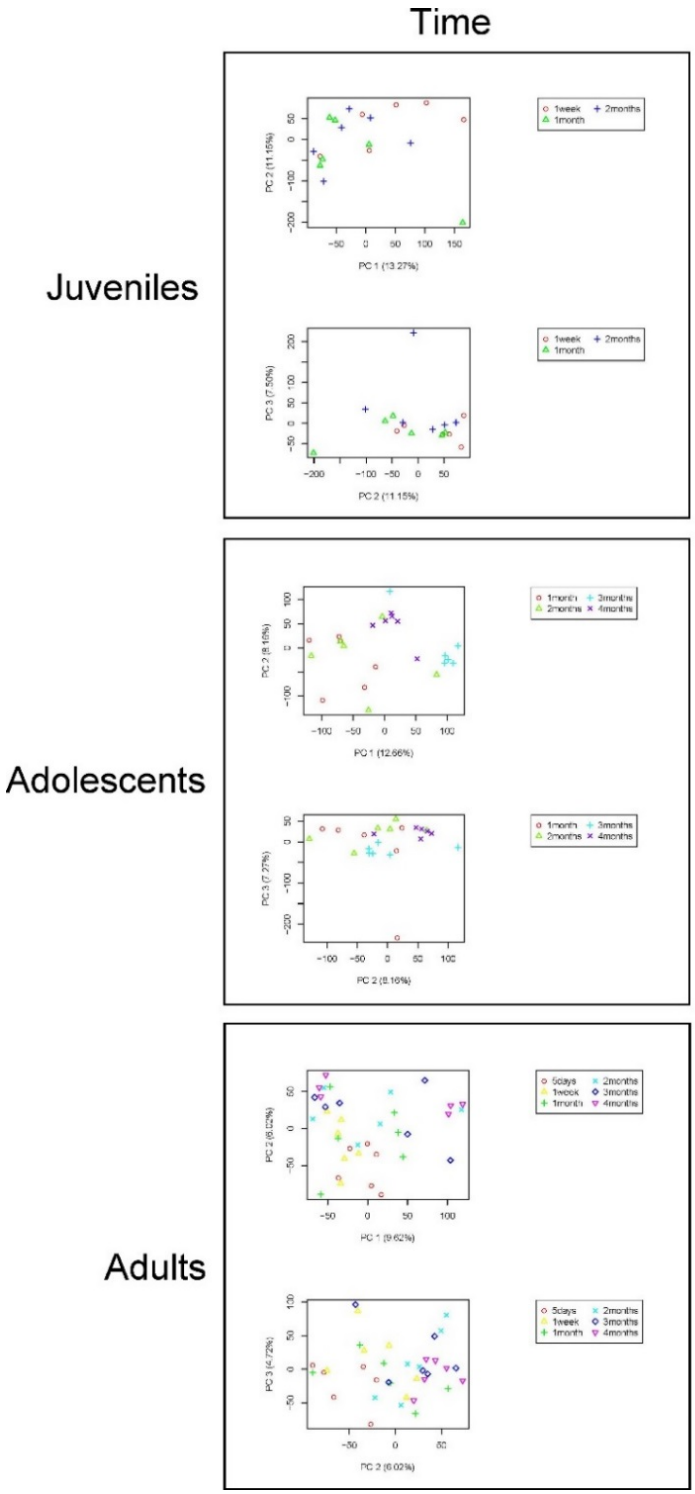

Figure 5-7. Principal component analysis of transcript abundance estimates for each library (TMM normalised FPKM values). Labelled to highlight just the effect of time- there was no clear clustering.

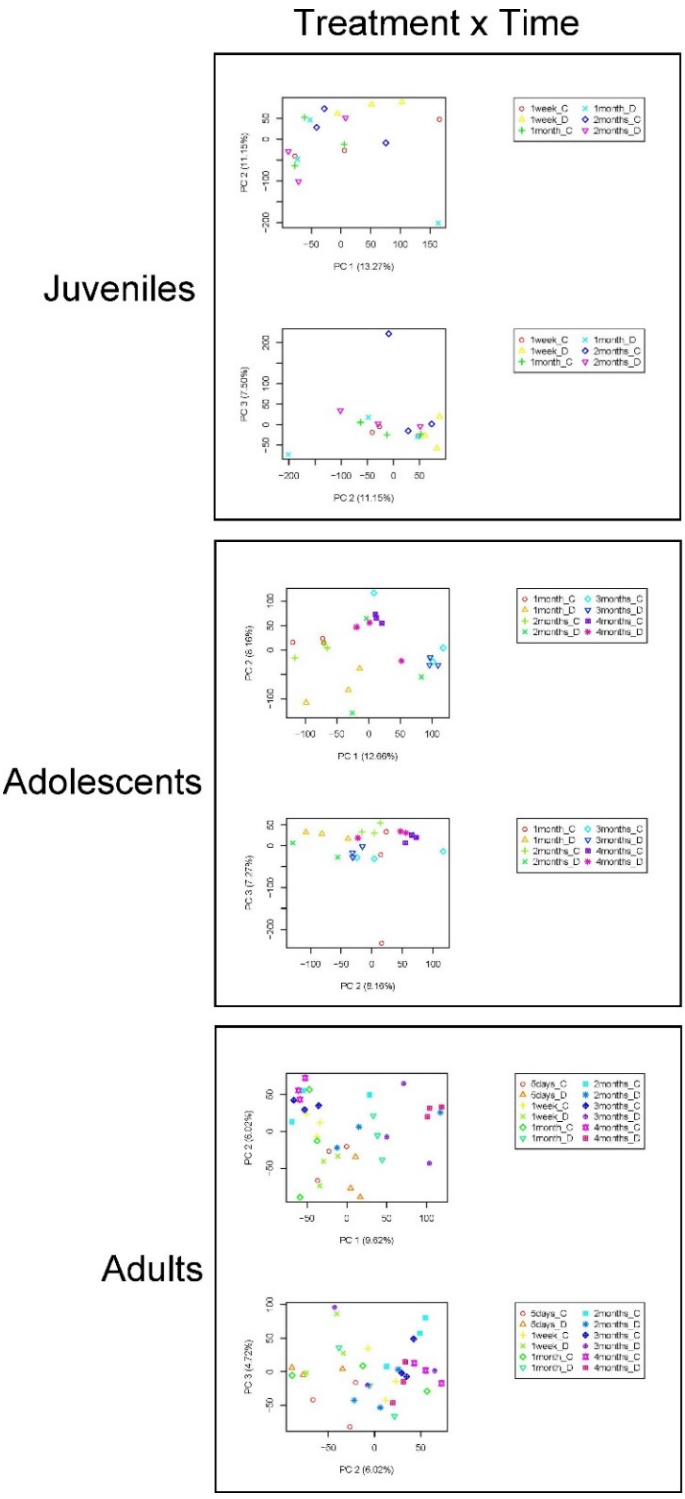

Figure 5-8. Principal component analysis of transcript abundance estimates for each library (TMM normalised FPKM values). Labelled to highlight the effects of treatment x time – there was no clear clustering.

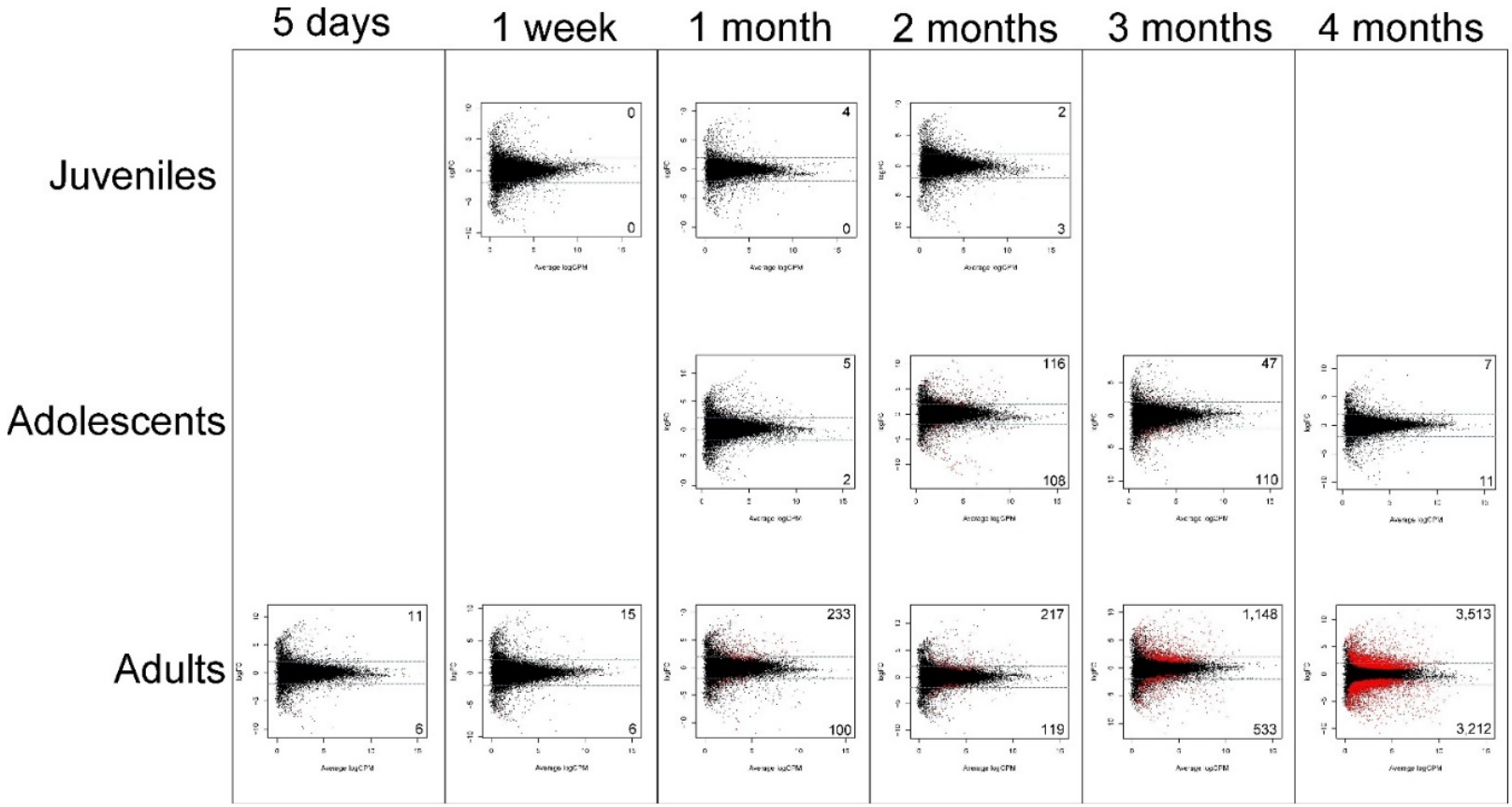

Figure 5-9. Smear plots of up and down-regulated genes between damaged and control animals at each time point in each age experiment. Grey dashed line indicates log_2-fold and red indicates genes which were found to be significantly different (< FDR 0.05).

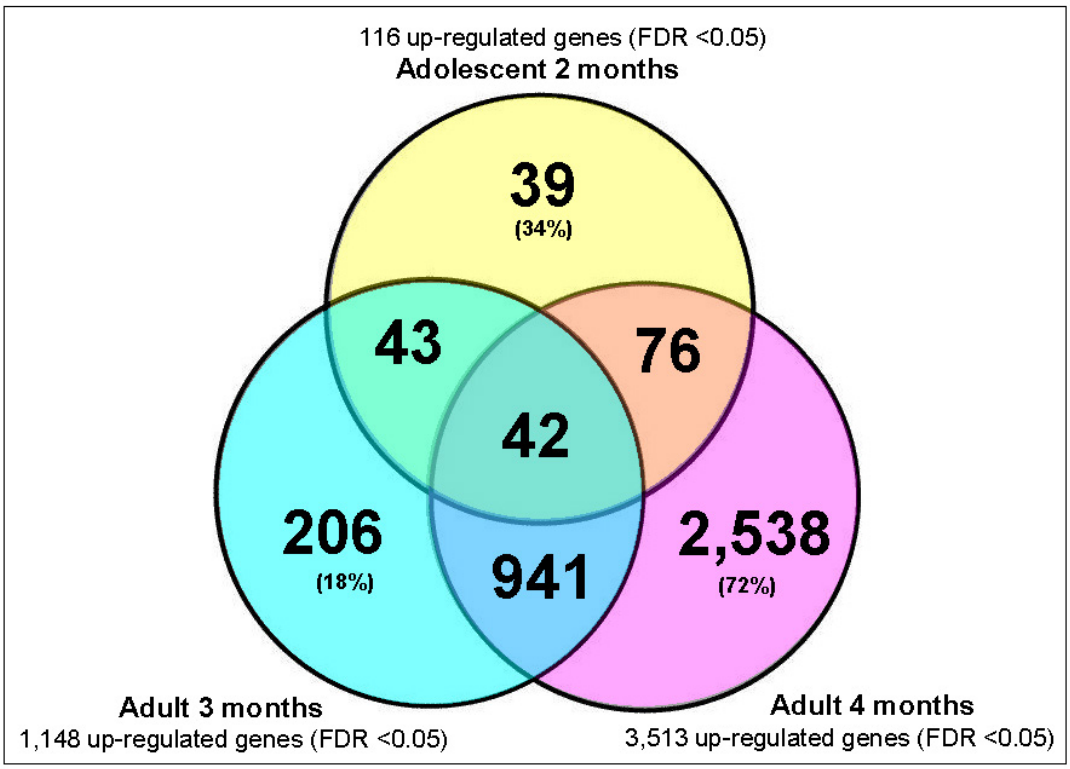

Figure 5-10. A comparison of the up-regulated genes between three key experimental time points. For annotations please see Appendix B - Supplementary files.

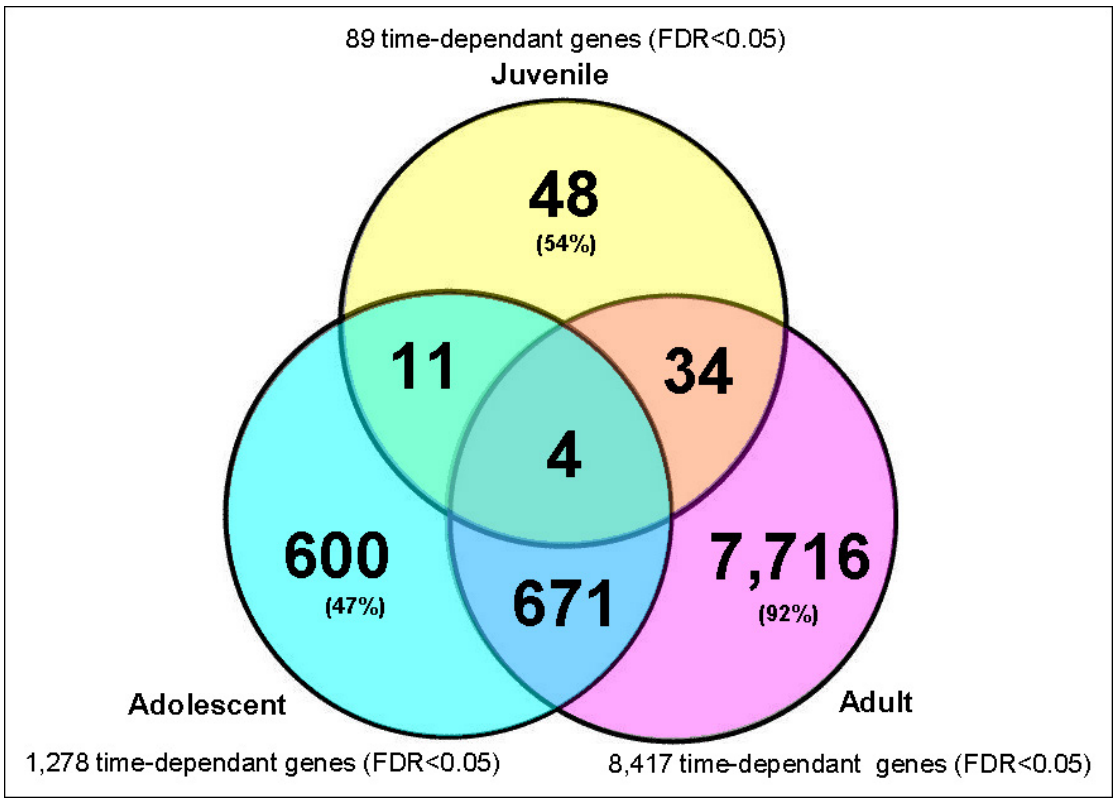

Figure 5-11. A comparison of the time-dependant damage-response genes between the three age experiments. For annotations please see Appendix B - Supplementary files.

5.3.2.3 Regulatory gene network

Gene expression profiles for the 199,321 Trinity genes were collapsed into 18,862 representative clusters using SOTA. To create a regulatory gene network from the SOTA gene expression profiles, ARACNE was used (p-value cut-off $1e^{-7}$) and the resulting network consisted of 13,577 nodes. To aid interpretation of this network, highly interconnected regions were identified using GLay yielding three large and fifteen small expression modules (Figure 5-12).

Genes of interest, highlighted from edgeR analyses, were mapped onto the regulatory gene network to identify nodes which directly interacted with well-characterised candidates, representing potential novel biomineralisation genes. Differential gene expression analysis using edgeR highlighted some well-known biomineralisation candidates for example three Trinity genes showed sequence similarity to a shell matrix protein, pif. *Pif* was significantly up-regulated in both adolescents at 2 months and adults at 3 and 4 months as well as showing a significant time-dependant damage response in adults. To investigate the possibility that these three Trinity genes were gene duplicates and potential paralogues, they were translated into amino acid sequences and aligned to two full length pif sequences extracted from NCBI, although all three were well aligned to the full length sequences, there were no overlapping regions and hence further phylogenetic analyses could not be pursued. The *pif* with the longest alignment (TRINITY_DN250713_c6_g1, Cluster-171) was selected for gene-regulation analysis using the gene network. In addition, two other shell matrix proteins, perlucin (TRINITY_DN252420_c3_g2, Cluster-11530) – which showed significant time-dependant damage response in adolescents - and a cartilage matrix protein (TRINITY_DN251420_c3_g1, Cluster-13458) - which was significantly up-regulated in adults at 4 months and overall had a time-dependant damage response in adults - were mapped onto the network. *Pif* had 34 first-neighbours connected by 79 interacting network edges (Figure 5-13), 22 of the first-neighbours were annotated (Table 5-2). *Perlucin* had 16 first-neighbours connected by 26 interacting network edges (Figure 5-13), 9 of the first-neighbours were annotated (Table 5-3). *Cartilage matrix protein* had 7 first-neighbours connected by 9 interacting network edges (Figure 5-13), and 2 of the first-neighbours were annotated (Table 5-4).

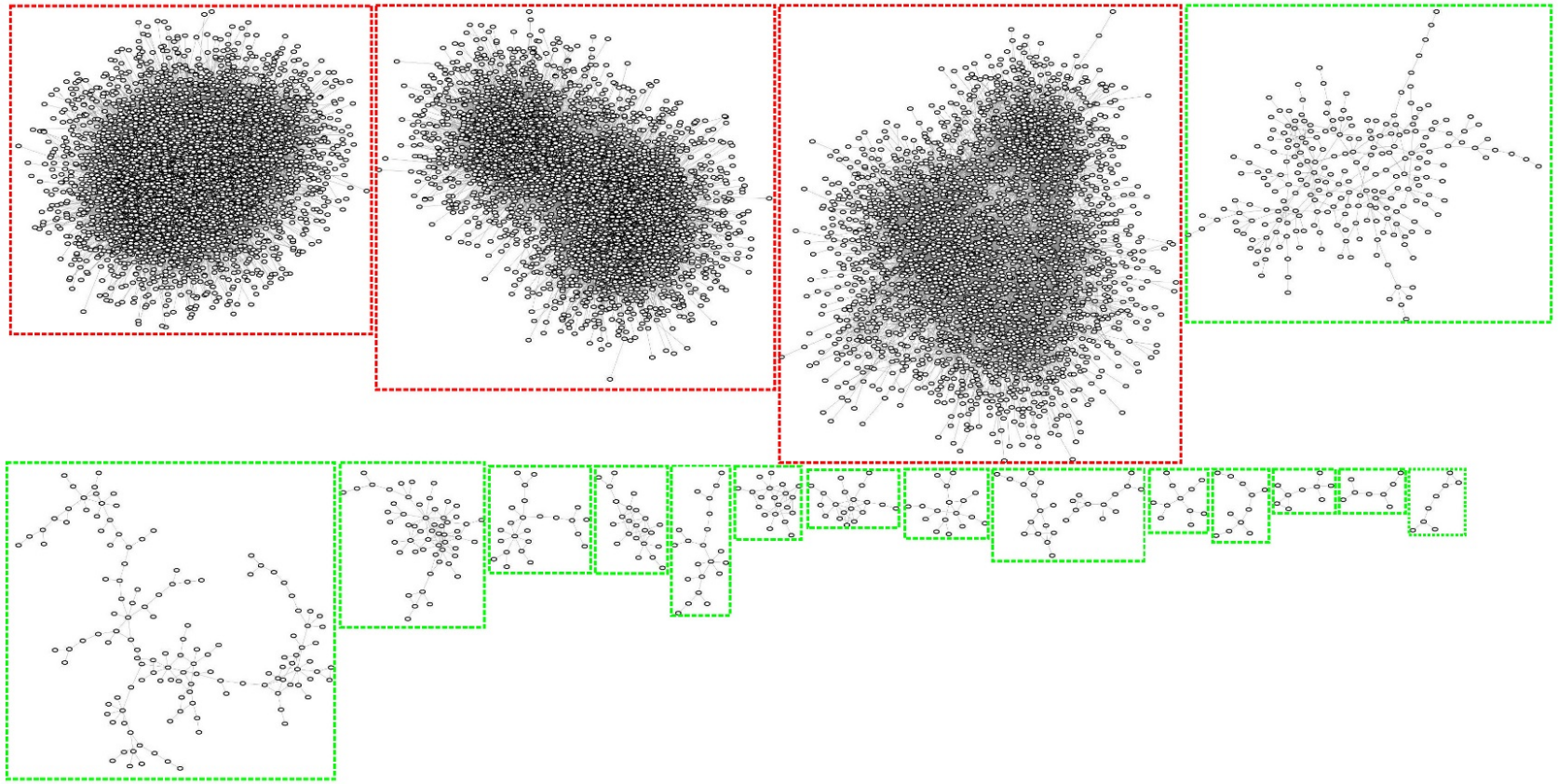

Figure 5-12. ***Laternula elliptica*** **mantle tissue gene regulatory network created using ARACNE. Eighteen highly interconnected transcriptional modules, "large" modules highlighted in red and "small" highlighted in green.**

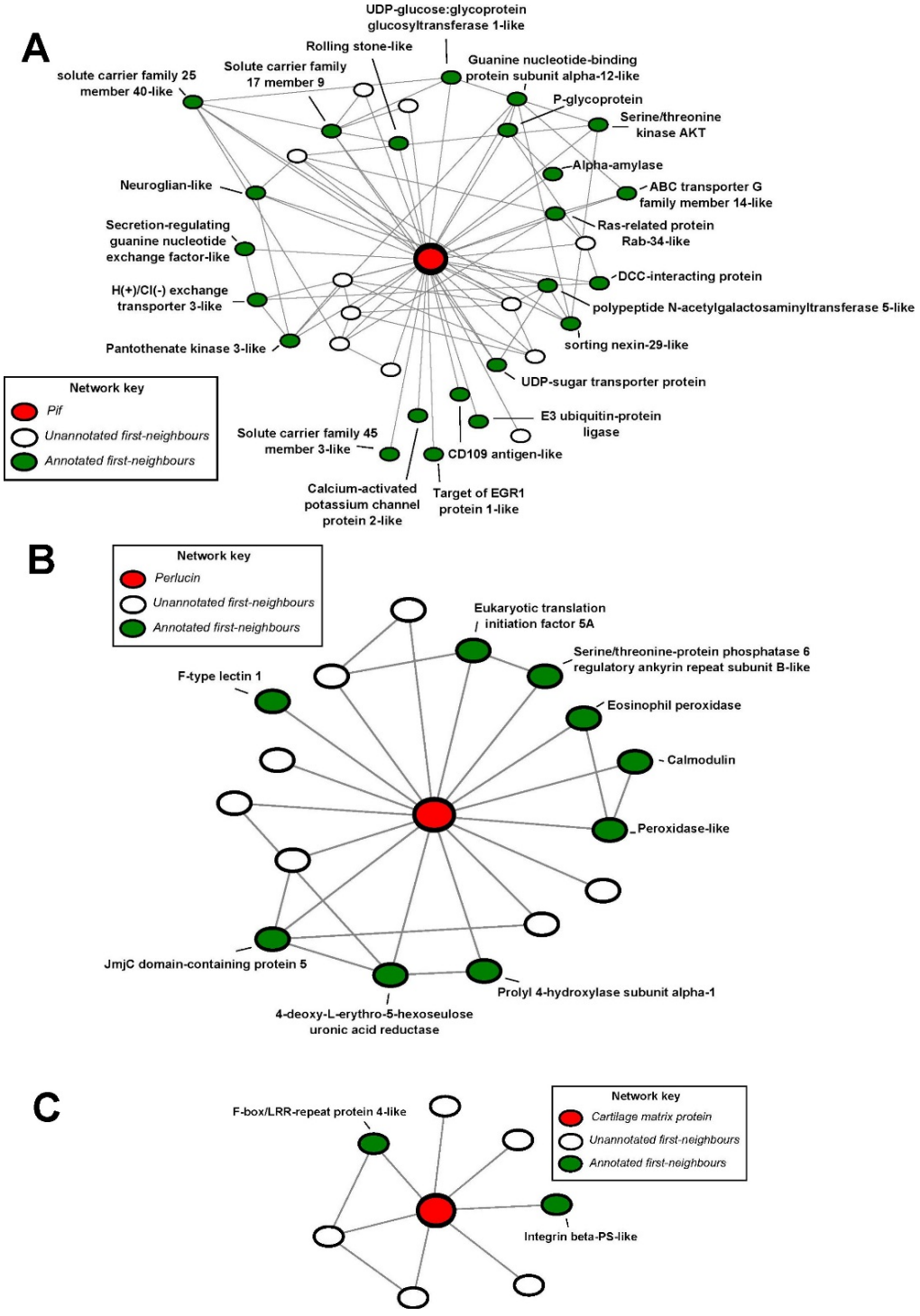

Figure 5-13. The first-neighbour networks of three significantly up-regulated biomineralisation candidates, < $1e^{-10}$ cut-off used for annotations.

Table 5-2. The annotated (blastx, cut-off < $1e^{-10}$) first-neighbours of up-regulated biomineralisation candidate *pif*.

Trinity_ID	Putative annotation	Species name	E-value	Putative function
TRINITY_DN250833_c3_g4	Alpha-amylase	*Hyriopsis cumingii*	0	Carbohydrate metabolism
TRINITY_DN259102_c1_g1	Neuroglian-like isoform X1	*Crassostrea gigas*	$1.65E^{-32}$	Cell adhesion and transmembrane transport
TRINITY_DN242704_c0_g1	Target of EGR1 protein 1-like	*Lingula anatina*	$1.62E^{-121}$	Cell cycle
TRINITY_DN226389_c0_g1	Pantothenate kinase 3-like isoform X1	*Octopus bimaculoides*	$2.38E^{-176}$	Coenzyme A biosynthesis
TRINITY_DN259983_c3_g4	PCD109 antigen-like isoform X5	*Crassostrea gigas*	$5.41E^{-153}$	Immune
TRINITY_DN253187_c4_g1	Small conductance calcium-activated potassium channel protein 2-like isoform X1	*Crassostrea gigas*	0	Ion transport
TRINITY_DN257781_c0_g2	H(+)/Cl(-) exchange transporter 3-like isoform X1	*Crassostrea gigas*	0	Ion transport and vescile trafficking
TRINITY_DN258797_c2_g1	UDP-glucose:glycoprotein glucosyltransferase 1-like	*Crassostrea gigas*	0	Misfolded protein tag
TRINITY_DN247112_c1_g2	Protein rolling stone-like	*Crassostrea gigas*	$1.63E^{-55}$	Muscle fibre fusion
TRINITY_DN252937_c0_g3	Polypeptide N-acetylgalactosaminyltransferase 5-like	*Aplysia californica*	0	Protein modification
TRINITY_DN254087_c1_g1	Sorting nexin-29-like isoform X2	*Crassostrea gigas*	0	Protein transport and vesicle trafficking
TRINITY_DN252773_c1_g1	Ras-related protein Rab-34-like	*Crassostrea gigas*	$2.11E^{-117}$	Protein transport and immune
TRINITY_DN233162_c0_g1	Secretion-regulating guanine nucleotide exchange factor-like isoform X1	*Lingula anatina*	$9.27E^{-91}$	Secretion
TRINITY_DN241857_c3_g1	Serine/threonine kinase AKT	*Crassostrea hongkongensis*	0	Signalling
TRINITY_DN236405_c0_g1	Guanine nucleotide-binding protein subunit alpha-12-like	*Crassostrea gigas*	0	Signalling
TRINITY_DN252858_c0_g4	DCC-interacting protein 13-alpha	*Crassostrea gigas*	0	Signalling
TRINITY_DN238563_c0_g1	Probable E3 ubiquitin-protein ligase DTX3 isoform X1	*Crassostrea gigas*	$8.69E^{-76}$	Signalling and protein modification
TRINITY_DN253024_c2_g1	Solute carrier family 25 member 40-like	*Aplysia californica*	$1.07E^{-135}$	Transmembrane transport
TRINITY_DN255659_c0_g1	Solute carrier family 17 member 9	*Crassostrea gigas*	0	Transmembrane transport
TRINITY_DN252539_c1_g3	P-glycoprotein	*Tegillarca granosa*	0	Transmembrane transport
TRINITY_DN253347_c5_g1	ABC transporter G family member 14-like isoform X3	*Lingula anatina*	$4.34E^{-177}$	Transmembrane transport
TRINITY_DN245947_c2_g1	Solute carrier family 35 member A4	*Lingula anatina*	$3.34E^{-77}$	Transmembrane transport
TRINITY_DN256578_c4_g3	Solute carrier family 45 member 3-like isoform X6	*Crassostrea gigas*	0	Transmembrane transport

Table 5-3. The annotated (blastx, cut-off <1e^{-10}) first-neighbours of up-regulated biomineralisation candidate *perlucin*.

Trinity_ID	Putative annotation	Species name	E-value	Putative function
TRINITY_DN253595_c4_g1	Calmodulin	*Myotis brandtii*	1.23E^{-48}	Calcium binding
TRINITY_DN256444_c0_g1	Prolyl 4-hydroxylase subunit alpha-1	*Crassostrea gigas*	1.87E^{-66}	Collagen fibril organisation
TRINITY_DN254850_c3_g4	Eosinophil peroxidase	*Crassostrea gigas*	3.46E^{-103}	Immune
TRINITY_DN255243_c1_g2	F-type lectin 1	*Pinctada fucata*	1.54E^{-24}	Immune
TRINITY_DN242538_c7_g2	Peroxidase-like protein	*Crassostrea gigas*	8.02E^{-33}	Oxidative Stress
TRINITY_DN250855_c1_g1	4-deoxy-L-erythro-5-hexoseulose uronic acid reductase	*Haliotis discus hannai*	3.27E^{-124}	Oxidoreductase activity
TRINITY_DN253914_c5_g2	Eukaryotic translation initiation factor 5A	*Athalia rosae*	1E^{-51}	Positive regulation of translation
TRINITY_DN227098_c0_g1	Serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit B-like	*Octopus bimaculoides*	8.22E^{-93}	Recognition
TRINITY_DN240723_c1_g1	JmjC domain-containing protein 5	*Crassostrea gigas*	5.86E^{-113}	Transcription factor

Table 5-4. The annotated (blastx, cut-off <1e^{-10}) first-neighbours of up-regulated biomineralisation candidate *cartilage matrix protein*.

Trinity_ID	**Putative annotation**	**Species name**	**E-value**	**Putative function**
TRINITY_DN259474_c3_g1	Integrin beta-PS-like	*Priapulus caudatus*	2.53E^{-44}	Calcium dependant matrix adhesion
TRINITY_DN253434_c6_g2	F-box/LRR-repeat protein 4-like	*Crassostrea gigas*	0	Ubiquitin-protein transferase activity

5.4 Discussion

5.4.1 Age-dependant shell damage-repair response

5.4.1.1 Ageing delays the molecular shell damage-repair response

The Antarctic clam, *L. elliptica,* can live for up to 36 years and, like many long-lived species, their ageing process includes a slow loss in energetic function coupled to a similarly slow increase in oxidative stress, culminating in the deterioration of cellular health, and eventually a decline in physiological function and death (Philipp et al. 2005). Over this long lifetime age significantly reduces the ability of *L. elliptica* to respond to stressors, such as injury, starvation, hypoxia and physical disturbance requiring re-burying (Peck et al. 2004, Philipp and Abele 2010, Clark et al. 2013a, Husmann et al. 2014). The present chapter therefore aimed to better understand the effect of age on shell damage-repair processes in *L. elliptica* and it was hypothesised that older animals would find shell-repair more difficult. At the physiological level, there was high inter- and intra-individual variability in the rate of the shell-repair and, given the qualitative categorical nature of the shell-repair measurements, a significant difference in the rate of shell-repair was not observed between the three age categories (Figure 5-3). Previous shell damage-repair studies have also found very variable healing rates, for example in a study on *Mytilus edulis* Huning et al. (2016) stated that there was "high variation in the regeneration speed between individuals". At the molecular level however, there was a marked difference in the timing of differential expression (between control and damaged animals) for juvenile, adolescent and adult animals (Figure 5-9). There was no peak in differential expression at any time during the experiment for juvenile animals, whereas adolescent animals showed a peak in differential expression after 2 months and the differential expression in adults had the highest differential expression at the end of the experiment, 4 months after damage was inflicted (Figure 5-9). Hence, the molecular response was progressively delayed with age and significantly extended in adult animals. In medical research, it is well established that age causes delays in wound repair and the effect of ageing has been documented at the physiological and molecular level in models such as rats, mice and humans (Khodr and Khalil 2001, Sgonc and Gruber 2013). In

mollusc ageing research however, the temporal component of damage-repair has not been studied and therefore results in the present chapter represent the first report of age-delayed repair, at the cellular level - for molluscs.

The analytical pipeline in this chapter was unable to detect a large number of differentially expressed genes in juvenile animals. One possible explanation for the absence of a detectable molecular response in juveniles could be that these animals were pre-reproductive maturity and could have already been investing all their available free energy into somatic growth, including shell growth. Hence, even when the shell was damaged, all available energy was already invested in somatic growth and shell production and either, it was not possible to increase the rate of growth, or, no shunting in process, or changes to gene expression profiles, were required. In contrast, mature animals were likely to be investing much of their available resources into reproductive output and diverting resources away from reproduction towards shell-repair could therefore be delayed. Fitting in with this possible explanation for delayed repair in adults, the "disposable soma theory" of ageing predicts that, in long-lived animals, mature specimens will divert energy away from tissue maintenance to reproduction (Philipp and Abele 2010). Previous work has suggested *L. elliptica* fits the "disposable soma theory" of aging (Clark et al. 2013a), and the delayed response to damage in older animals in the present study also supports this hypothesis.

5.4.1.2 Older animals find shell damage more difficult

The molecular response in adult animals was not just slower, it was also much more intense and, in general, included more stress-response genes (Chapter 5 Supplementary Tables 1-11 – Appendix B Supplementary files). For the adults, the number of differentially expressed genes was an order of magnitude higher than the adolescents, and likewise, the adolescents were an order of magnitude higher than the juveniles (Figure 5-9). At the peak of the adolescent response, 2 months after damage, there were only three up-regulated transcripts related to the stress response (TRINITY_DN259186_c1_g5 *heatshock 70kDa protein13-like isoformX2*, TRINITY_DN251698_c2_g1 *heatshock protein 60* and TRINITY_DN258017_c0_g5 *glutathioneS-transferase sigma2*). At the

peak of the adult response, 4 months after damage, however, many of the up-regulated genes encoded stress relates genes (too many to list in text), including oxidative stress, for example, seven different *HSP*s and three *glutathione-transferase* family members (Chapter 5 Supplementary Table 11 – Appendix B Supplementary files). It was unsurprising to find shell damage was more stressful for older animals as many previous bivalve experiments have demonstrated similar patterns. Clark et al. (2013a) found older *L. elliptica* exhibited more signs of oxidative stress after exposure to hypoxia, Ivanina et al. (2008) detected age related changes in the expression of oxidative stress genes and chaperones in *Crassostrea virginica* and Sukhotin et al. (2002) found lipid peroxidation (which is caused by free radicals) increases exponentially with age in *Mytilus edulis.*

5.4.1.3 Very few shared shell damage-response genes across ages

After documenting that the molecular response to shell damage was generally slower, more intense and included more stress response genes in adult animals, the next step was to directly compare the molecular pathways and specific genes which were differentially expressed between the ages. In order to identify shared and unique damage-response genes between age groups the three time points with the most differential gene expression were compared (up-regulated genes only, Figure 5-10), and in addition, the time-dependant damage-response genes for each age group were also compared (Figure 5-11). For each comparison remarkably few transcripts were shared between the age groups.

5.4.1.3.1 Shared genes from the highly up-regulated comparison

When the up-regulated genes were compared between the three key experimental time points, only 42 transcripts were shared between all three groups and of those, 21 could be assigned a putative function via sequence similarity searching (Chapter 5 Supplementary Table 12 – Appendix B Supplementary files, Figure 5-10). Even more surprising was that only one of those shared genes, *pif* (TRINITY_DN244470_c1_g2), was a likely biomineralisation candidate. As explained in previous chapters, pif is a well characterised extracellular matrix protein. It was found in the *L. elliptica* nacreous shell proteome (Chapter 3) and its gene expression was localised to specific cells in the *L. elliptica* mantle epithelium (Chapter 3). In addition it has been described in many other mollusc shell

proteomes and mantle transcriptomes (Suzuki et al. 2009, Werner et al. 2013, Joubert et al. 2010, Shi et al. 2013, Michio Suzuki et al. 2013, Bahn et al. 2015). Whilst it was surprising to find only one shared biomineralisation gene between the three experimental time points, this finding does demonstrate that *pif* is likely to be central to shell-repair, across age ranges, in this species.

5.4.1.3.2 Shared genes from the time-dependent comparison

When the time-dependant damage-response genes for each age group were compared, even fewer genes were shared between the three age categories. Only 4 time-dependant genes were shared, and of those only 1 was assigned a putative function (TRINITY_DN252582_c0_g3, *cAMP-dependent protein kinase type II regulatory subunit – PRKAR2A,* Chapter 5 Supplementary Table 19 – Appendix B Supplementary files). *cAMP-dependent protein kinase type II regulatory subunit - PRKAR2A*- is involved in the G protein-coupled receptor-triggered signalling cascade used in cell communication, and has not previously been reported in mollusc biomineralisation. From a general perspective, cAMP signalling cascades are important in a huge suite of biological processes (Tasken and Aandahl 2004), and due to their involvement in so many different pathways perhaps it is not surprising *PRKAR2A* had a time-dependant damage-response in all age groups. Taking a more biomineralisation specific perspective, the importance of protein kinase and cAMP signalling in bone and cartilage development has been documented in the mammalian medical literature for many years (Karsenty et al. 2009). Recently a gene knock-out study in mice found that *PRKAR2A* improves the bone phenotype of mice (Liu et al. 2015). Given the importance of *PRKAR2A* in the healthy production of extracellular matrices and biominerals in mammals, and it being the only shared time-dependant gene, out of several thousand, it could be possible it also plays an important role in the modulation of molluscan biomineralisation.

5.4.1.4 High proportion of unique shell damage-response genes

For all of the compared sets of genes (Figure 5-10, Figure 5-11), a high proportion of the differentially expressed transcripts were unique to a time-point or age group, suggesting that, at a gene expression level, the shell damage-repair response was different for each

age category. For the adolescents at 2 months 34 % of genes were unique, for the adults at 3 and 4 months 18 % and 72 % respectively were unique and for the time-dependant genes in the juvenile, adolescent and adult experiments, 54 %, 47 % and 92 % of time-dependant damage response genes, respectively, were unique. Another possible reason the transcriptional analyses in this chapter were unable to resolve a core set of damage-response genes shared amongst ages and time points, and instead found large proportions of genes uniquely up-regulated, could be because the animals used in the experiments were wild invertebrates, and gene expression is a very dynamic and variable process. Gene expression in wild animals (compared to inbred laboratory model animals) is known to be highly variable (Golden and Melov 2004, Garfield et al. 2013). More specifically, studies focussing on molluscan biomineralisation in wild populations frequently report high degrees of variability and the very dynamic nature of these expression profiles (Jackson et al. 2007). The study in this chapter aimed to assess shell damage-repair over time and age and therefore the number of experimental time points were high, and the biological replication was relatively low (n=3 for transcriptomics, which is still higher than most other biomineralisation studies). In addition there was the added logistical restraint on the number of animals that could be collected from Antarctica. In future studies however, as the cost of sequencing decreases, it could be interesting to increase replication from n=3, and use power analyses to achieve the best number of replicates required to resolves patterns which may currently be hidden in these highly variable data sets.

5.4.1.4.1 Unique genes from the highly up-regulated comparison

When comparing the annotations of the unique genes in the different categories, more evidence suggested adults found shell-repair more difficult. The comparison of the three most highly up-regulated time points (adolescents at 2 months and adults at 3 and 4 months) showed that the adolescent unique genes were dominated by biomineralisation processes, whereas the adults genes had more stress and signalling genes than biomineralisation. For example, the most highly up-regulated, annotated, unique gene for adolescents at 2 months was *carbonic anhydrase* (*CA*, LogFC = 10.2, Chapter 5 Supplementary Table 13 – Appendix B Supplementary files). CAs are a group of metallo-

enzymes that catalyse the reversible hydration of carbon dioxide to form one bicarbonate ion and one proton, this reaction is essential for many physiological functions such as pH regulation, respiration, photosynthesis, and biomineralisation (Le Roy et al. 2014). CA was one of the first biomineralisation domains to be characterised and as such has become a favourite amongst many researchers interested in the biological effects of ocean acidification on marine calcifiers, and more generally, in shell production (Miyamoto et al. 1996, Medakovic 2000, Miyamoto et al. 2005). Indeed CA domains were present in the nacreous shell proteome of *L. elliptica* (Chapter 3) and in the shell proteome of many other bivalves (Marie et al. 2010, Berland et al. 2013, Arivalagan et al. 2016). In addition to *CA*, seven other up-regulated, annotated, unique transcripts for adolescents at 2 months, were vesicle trafficking or membrane transport proteins (such as *sorting nexin-20*, *solute carrier family 22* and *vesicle-fusing ATPase 1*, Chapter 5 Supplementary Table 13 – Appendix B Supplementary files). The presence of vesicles and vesicle-associated gene expression was documented in the *L. elliptica* mantle epithelium in Chapter 3. In addition, these proteins are classically associated with "secretomes" (Jackson et al. 2006, Aguilera et al. 2017), and were therefore likely to be involved in the transportation of calcium carbonate - and various other components required for the extracellular matrix - to the shell deposition site. In contrast, the most highly up-regulated, annotated, unique transcripts for adults at 3 and 4 months, were stress proteins (Chapter 5 Supplementarys Table 14 and 15 – Appendix B Supplementary files). For example, the second and third highest unique up-regulated transcripts for adults at 3 months were *HSP*s (TRINITY_DN246078_c0_g1 & TRINITY_DN257198_c0_g2, *HSP70*, LogFC = 10.3 & 8.8 respectively) and the most highly up-regulated annotated transcript for adults at 4 months was *thioredoxin* (TRINITY_DN245544_c1_g4, LogFC = 10.4).

5.4.1.4.2 Unique genes from the time-dependent comparison

The annotated, unique, time-dependant damage-response genes were investigated to look for age differences, and specifically age differences in biomineralisation genes. The juveniles only had 4 annotated unique time-dependant transcripts, none of which were likely to be involved in biomineralisation (Chapter 5 Supplementary Table 20 – Appendix B Supplementary files). The adolescents had 168 annotated unique time-dependant

transcripts, including some shell matrix proteins (two *perlucins* and a *Blue Mussel Shell Protein [BMSP]*) along with vesicle trafficking and membrane transport proteins (Chapter 5 Supplementary Table 21 – Appendix B Supplementary files). And finally the adults had 2265 unique time-dependant transcripts (Chapter 5 Supplementary Table 22 – Appendix B Supplementary files). The number of unique adult transcripts was too high for manual assessment, the transcripts were therefore sorted in descending order of LogFC in the 4 month time point to highlight those which were increasingly up-regulated over time in the damaged animals. This filtering method revealed that many of the adult time-dependant transcripts that were increasingly up-regulated over time were likely biomineralisation candidates, for example the shell matrix genes *pif* (n.b. different isoform, TRINITY_DN250713_c6_g1, to the *pif* found as a shared gene above), *pfN44* and *cartilage matrix protein* (Chapter 5 Supplementary Table 22 – Appendix B Supplementary files). In addition, chitin remodelling genes were also found such as *acidic mammalian chitinase*. The unique time-dependent shell matrix proteins in adolescents (perlucin and BMSP) and adults (pif, pfN44 and cartilage matrix protein) are all involved in the production of the nacreous layer, suggesting that *L. ellipica,* of various ages, repair shell damage using nacre, rather than prisms. Nacre is the layer which covers the inside of the shell and is closest to the majority of the calcifying epithelium, most of the cells in the mantle therefore are designed to secrete nacre, whereas only a relatively small proportion of mantle cells, at the growing front of the shell where the mantle edge and periostracal groove are, secrete prismatic shell (Chapter 1, Chapter 3). Although the biomineralisation genes in both ages are related to the nacreous layer, the role these shell matrix protein play in nacre biomineralisation are subtly different. Perlucins and BMSP found in the adolescents both interact with aragonite, perlucin reportedly functions to nucleate aragonite crystals (Dodenhof et al. 2014) and BMSP functions to bind to both aragonite crystals and chitin (Suzuki et al. 2011). Pif and pfN44 found in adults also interact with aragonite but in a slightly different way. Similar to perlucin, pif binds to both chitin and aragonite (Suzuki et al. 2009), however pfN44 interacts with magnesium calcite to inhibit the crystallisation of aragonite and therefore contributes to the control of aragonite crystallisation (Pan et al. 2014). The difference in the putative functional roles of the unique time-dependent shell matrix proteins in adolescent and adult animals was

only very subtle, but could indicate that the way *L. elliptica* repairs its shell using nacre differs with age.

5.4.2 Physiological measurements as a framework for transcriptomics

Various physiological measurements were taken during shell-damage repair experiments with the aim of using them as a framework for which to compare the molecular measurement to. Previous experiments have observed a tight interaction between gene expression patterns and structural changes to repairing shells (Huning et al. 2016), and hence the present chapter aimed to use a range of physiological observations, not just shell-repair, to help interpret some of the molecular data. Adult animals were chosen for histology and immunology measurements as that experiment included the most time points, and in addition, it was only from the larger animals that sufficient quantities of haemolymph could be extracted for analyses. Overall there was a high degree of variation across all of the histology and immunology metrics used and these data were comparable with similar measurements of other wild mollusc species (Bodin et al. 2004, Ford and Paillard 2007). Despite the high degree of variation some significant patterns indicated increased epithelium inflammation, haemocyte infiltration into the mantle edge and tissue remodelling in damaged animals, which, in some cases, tightly coincided with large changes in gene expression profiles.

After 5 days, damaged animals had a significantly higher fused inner mantle fold surface area than controls (Figure 5-4). At this time point however, the gene expression profiles of control and damaged animals were very similar (Figure 5-9). Pro-inflammatory gene expression responses that were likely to have caused the observed increase in epithelium surface area probably occurred on a much shorter time scale than the observed histological changes. The temporal separation in the early onset of pro-inflammation gene expression and residing physiological signs of inflammation has been well documented in medical models (Bonneh-Barkay et al. 2010, Yang et al. 2012, Gyoneva and Ransohoff 2015, Tomida et al. 2015). In contrast, the genetic control of inflammation, in general, is not well understood in bivalves, and results in the present

chapter highlight the highly dynamic nature of physiological inflammation in response to shell-damage, and gene expression.

After 3 months, damaged animals had significantly more roaming haemocytes in the fused inner mantle fold than controls (Figure 5-4) and, at the same time point, over a thousand genes were up-regulated (compared to only 217 in the previous time point, Figure 5-9). Some of the most highly up-regulated genes at this time point included both immune and biomineralisation genes [TRINITY_DN254110_c1_g1, logFC = 8.2, annotation = *CD109 antigen* (Solomon et al. 2004); TRINITY_DN256879_c1_g2, logFC = 7.5, annotation = *T-cell activation inhibitor,mitochondrial* (Keeren et al. 2009); TRINITY_DN235682_c0_g1, logFC = 4.8, annotation = *chitotriosidase-1-like* (Boot et al. 1995); TRINITY_DN255711_c1_g2, logFC = 3.8, annotation = *chit3* (Gao et al. 2017); TRINITY_DN245203_c1_g1, logFC = 3.8, annotation = *pfN44* (Pan et al. 2014)]. Haemocyte-mediated shell-repair has been documented in many bivalve species (Mount et al. 2004, Kadar 2008, Li et al. 2016), and haemocytes have been reported to transport calcium carbonate and shell matrix proteins to the extrapallial space where shell is secreted (Li et al. 2016). The exact role of haemocyctes in both normal biomineralisation and shell-repair is still unclear; in both scenarios these cells are likely to be playing a dual role in the immunoprotection of the animal at its external barrier, and in shell production. Increased haemocyte infiltration into the fused inner mantle fold, coupled to a large increase in differential gene expression – including both biomineralisation and immune genes - further supports the hypothesis of haemocycte-mediated biomineralisation.

During the first three time points (5 days, 1 week and 1 month) damaged animals had a consistently lower mean redness index, and then at the 2 month time point, this pattern reversed and damaged animals had a significantly higher mean redness index than controls. Interestingly in the next two time points, 3 and 4 months, when the most differential gene expression was seen, the mean redness index in control and damaged animals was the same. Similar to the epithelium inflammation, this histological measurement was temporally de-coupled from the molecular results. The “mean redness index” of cellular organisation was developed in an attempt to quantify cellular organisation of the contractile fibres in the mantle tissue. Densely packed fibres were

disorganised, whereas ordered and regular fibres were less densely packed and less intensely stained. The histology results therefore showed that damaged animals had less densely packed contractile fibres, followed by a large increase in fibres and finally a levelling off, indicating that cellular remodelling of the mantle tissue occurs over time in response to damage.

5.4.3 Regulatory gene network provides insights into the molecular control of biomineralisation: investigation of candidate gene connections

Regulatory gene networks generated from RNA-Seq gene expression profiles are a powerful tool for inferring gene function and understanding the molecular pathways that a gene may be operating in (Schaefer et al. 2017). Presented in this chapter is the first comprehensive regulatory gene network constructed from mollusc mantle gene expression profiles. Mantle gene expression profiles (FPKM values generated by Trinity) from 78 individual *L. elliptica*, of different ages (juveniles, adolescents and adults with shell length ranging from 19 to 84 mm), in different temporal stages of shell damage-repair (39 control animals plus 39 damaged animals 5 days, 1 week and 1, 2, 3 and 4 months after shell damage) were used to construct the network. Regions of the network that were highly interconnected were identified, resulting in three large and fifteen small expression modules (Figure 5-12). This network will be extremely valuable for the study of biomineralisation, and potentially many other processes such as Antarctic ageing and stress-responses. One application of the network is to highlight a node with well-characterised functional annotation and investigate which other nodes in the network it is connected to, with each significant connection representing a likely biological regulatory relationship (Wolfe et al. 2005, Saito et al. 2008, Tenyi et al. 2016, Schaefer et al. 2017). Three well-characterised biomineralisation genes– *pif*, *perlucin* and *cartilage matrix protein*- highlighted from the edgeR analytical pipeline described above, were mapped onto the regulatory gene network and their connected first-neighbour nodes were investigated (Figure 5-13).

Pif was connected to 34 first-neighbours (Figure 5-13), 22 of which were annotated (Table 5-2). Of the 22 annotated first-neighbours, 4 were involved in signalling and 12 were involved in protein transport, vesicle trafficking, ion transport and general transmembrane transport. The signalling genes connected to *pif* could represent candidates for the signalling cascade which controls biomineralisation in *L. elliptica.* Serine/threonine kinase Akt for example, is a signal transduction molecule that modifies other proteins via phosphorylation. It has a well-documented function in the mediation of extracellular matrices, for example Krepinsky et al. (2005) demonstrated that mechanical strain leads to activation of the PI3K/Akt signalling pathway leading to the up-regulation and secretion of the extracellular matrix protein collagen I. The gene regulatory network in the present chapter suggests that serine/threonine kinase Akt could play a similar role in controlling the secretion of pif. Pif is a shell matrix protein, it – and all the other components of the shell such as, calcium carbonate, other proteins and glycoproteins - have to somehow travel across the membrane of secretory mantle epithelium cells to the shell. These shell components could be using the various transport channels connected to *pif* in the regulatory gene network. For example, one of *pif*'s first-neighbours was a P-glyoprotein, which is an integral membrane protein that acts as an ATP-dependent efflux pump in many tissues (Warren et al. 1995, Leader and O'Donnell 2005).

Perlucin was connected to 16 first-neighbours (Figure 5-13), 9 of which were annotated (Table 5-3). Similar to *pif*, three of the *perlucin* annotated first-neighbours were involved in signal transduction; *eukaryotic translation initiation factor 5*A, *serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit B-like* and *jmjc domain-containing protein 5*. Two of the three genes were interconnected in the gene regulatory network, indicating there are regulatory processes both between the signal transduction genes and between each of the genes with *perlucin*. Serine/threonine-protein phosphatase 6 regulatory ankyrin repeat subunit B-like is involved in recognition (Kwiek et al. 2006, Stefansson et al. 2008), whereas jmjc domain-containing protein 5 is a transcription factor (Klose et al. 2006), and eukaryotic translation initiation factor 5A regulates translation (Li et al. 2004). Together these genes could be part of a signalling cascade which controls the expression of *perlucin*, and therefore could be involved in the regulation of

biomineralisation. Another *perlucin* first-neighbour, *prolyl 4-hydroxylase subunit alpha-1*, is involved in the organisation of collagen fibres and (Lamberg et al. 1996), given collagen and collagen-like matrix proteins have been found in many shells (Fernandez et al. 2001, Blank et al. 2003, Yin et al. 2005, Tanur et al. 2010), it is possibly also involved in the regulation of the extracellular components such as *perlucin*. Three of *perlucin*'s first neighbours formed an interconnected group involved in immunity, *eosinophil peroxidase, F-type lectin 1* and *peroxidase-like protein* (Wang and Slungaard 2006, Anju et al. 2013). Increasingly, the immunoprotective properties of shell matrix proteins are being revealed (Arivalagan et al. 2017, Arivalagan et al. 2016). The gene regulatory network suggests that structural matrix proteins, such as perlucin, may be tied into regulatory processes with immune genes, therefore further demonstrating the immune role of the shell both as a physical and biochemical barrier to pathogens.

Cartilage matrix protein was connected to 7 first-neighbours (Figure 5-13), 2 of which were annotated (Table 5-4). Although it was connected to fewer first-neighbours than the other two biomineralisation genes investigated, the small network was still valuable. Of particular interest was the connection to *integrin beta-PS-like*, which is involved in calcium dependent matrix adhesion. Integrin transmembrane receptors function as a link for the extracellular matrix and the intracellular actin cytoskeleton and is crucial for the controlled organisation of matrices (Brower et al. 1995). Hence, given the regulatory gene network predicts governing processes between the *cartilage matrix protein* and *integrin beta-PS-like*, it could be possible *integrin beta-PS-like* is involved in the controlled regulation of cartilage matrix proteins within the shell matrix.

Pif, *perlucin* and *cartilage matrix protein* were also connected to many unannotated nodes. These unannotated nodes however, are not to be dismissed, their lack of annotation simply means they showed no sequence similarity to anything currently in the NCBI nr database, which does not mean they are unimportant. Whilst these nodes cannot currently be assigned a functional annotation, the regulatory gene network putatively suggests that they may be involved in biomineralisation due to their association with known genes. The bioinformatic analyses in the present chapter were able to assign

putative function to a range of novel biomineralisation candidates, some of which were previously unannotated, that will be further characterised in future research.

5.5 Conclusions

The shell damage-repair experiments in this chapter found that age causes a transcriptional delay in the response to shell damage and, in addition, older animals showed more signs of stress, and hence found shell repair more difficult. It was hypothesised that juvenile animals were already investing all available resources into somatic growth, including shell production, whereas older adults had to divert resources away from reproduction to repair the shell.

Surprisingly there were very few shared differently expressed genes between time points, or ages, indicating that animals at different ages, and stages in repair, were transcriptionally different. Only one biomineralisation gene, *pif*, was shared between key experimental time points and different age groups up-regulated different nacre-related biomineralisation genes, indicating subtle differences in the repair process.

Time points with major differential expression, were tightly coupled to some, but not all, histological metrics. Overall the transcriptional response to shell-damage, in all ages, was a highly dynamic event, which in some cases involved a time-lag between up-regulation of transcripts and observable changes at the histological and shell-repair level.

Gene expression profiles from the shell damage-repair experiments were used to create the first comprehensive regulatory gene network for mollusc mantle tissue. The network was used to identify new biomineralisation candidates, which should be further investigates in future research.

Chapter 6 CHARACTERISATION OF THE MANTLE TRANSCRIPTOME AND BIOMINERALISATION GENES IN THE BLUNT-GAPER CLAM, *MYA TRUNCATA*

6.1 Introduction

Mya truncata is a marine bivalve which is part of the Myidae family of soft-shelled clams; there is currently very little knowledge of how *M. truncata* build their shells and no molecular data on this species exist in public databases. This chapter will characterise the *M. truncata* transcriptome and biomineralisation genes to provide a preliminary understanding of how *M. truncata* build their shells in comparison to other bivalves, especially the Antarctic clam, *Laternula elliptica*.

More specifically, the objectives of the present chapter are:

1.) To develop a molecular resource to aid the study of biomineralisation in *M. truncata* by sequencing, assembling and putatively annotating the adult mantle transcriptome. 2.) To identify and investigate the expression and phylogeny of candidate biomineralisation genes in the newly-assembled *M. truncata* mantle transcriptome and 3.) To compare the mantle transcriptome of *M. truncata* to that of the Antarctic clam, *L. elliptica*.

6.2 Methods

In this chapter animals were collected by SCUBA divers from the NERC UK National Facility for Scientific Diving, Oban. Animal dissections were carried out by Laura J. Weir, the mantle transcriptome was assemble by Michael A.S. Thorne and the main mollusc tyrosinase alignment was provided by Felipe Aguilera. I carried out all of the laboratory work, data analysis and interpretation.

Work contained in this chapter was published in the following papers:

Arivalagan, J., Yarra, T., Marie, B., Sleight, V. A., Duvernois-Berthet, E., Clark, M. S., Marie, A., Berland, S. (2017). Insights from the shell proteome: biomineralization to adaptation. *Molecular Biology and Evolution*, **34**. 66-77. doi:10.1093/molbev/msw219.

https://academic.oup.com/mbe/article-lookup/doi/10.1093/molbev/msw219

Arivalagan, J., Marie, B., Sleight, V. A., Clark, M. S., Berland, S., Marie, A. (2016). Shell matrix proteins of the clam, *Mya truncata*: Roles beyond shell formation through proteomic study. Marine Genomics, **27**. 69-74. doi:10.1016/j.margen.2016.03.005.

http://www.sciencedirect.com/science/article/pii/S1874778716300186

Sleight, V. A., Thorne, M. A. S., Peck, L. S., Arivalagan, J., Berland, S., Marie, A., Clark, M. S. (2016). Characterisation of the mantle transcriptome and biomineralisation genes in the blunt-gaper clam, *Mya truncata*. Marine Genomics, **27**. 47-55. doi:10.1016/j.margen.2016.01.003.

http://www.sciencedirect.com/science/article/pii/S1874778716300034

6.2.1 *Mya truncata* mantle transcriptome

6.2.1.1 Animals

Mya truncata (n = 9, mean shell length = 64.5 mm ± 1.38 S.E.) were collected by the NERC UK National Facility for Scientific Diving from Dunstaffnage Bay, North West Scotland in August 2011 as per Chapter 2. Mantle tissue was dissected from each animal and RNA was extracted as per Chapter 2.

6.2.1.2 Sequencing and Bioinformatics

The total mantle RNA from nine individuals was pooled prior to sequencing. Pooled RNA was subject to 454 GS FLX Titanium sequencing at Cambridge University Department of Biochemistry Sequencing Centre.

454 reads were assembled into a de novo transcriptome with GS Data Analysis Software (454.com/products/analysis-software/) on default genomic-style parameters, resulting in a total of 20,106 contigs with an average read length of 675 bp. All contigs were compared to the NCBI non-redundant (nr) database (downloaded for in-house use January 2015) using the Basic Local Alignment Search Tool (Blast) to search for sequence similarity and putative gene annotation [Blastx, cut-off $<1e^{-10}$, (Altschul et al. 1990)]. The most highly expressed annotated contigs were identified to highlight dominant processes in the mantle at the transcript level. The mantle transcriptomes of *M. truncata* and *L. elliptica* were compared using tBlastx with default parameters.

Putative biomineralisation genes were identified using keyword searches for candidates which have previously been shown to be associated with shell production and calcification (Appendix A - Primer tables). In addition, contigs which were present in the mantle and shell proteome with sequence similarity to biomineralisation domains were also included (Arivalagan et al. 2016), see 2.5.1 for more inormation on candidate gene selection criteria.

6.2.1.3 Data availability

All reads generated in this chapter are available from NCBI SRA accession number: SRP064949, the assembled contigs are also available at: http://bit.ly/1QcFiVH.

6.2.2 Tyrosinase bivalve phylogeny

Fifteen of the *M. truncata* mantle contigs showed high sequence similarity to an important biomineralisation candidate, *tyrosinase* (see 2.5.1 for more information). To understand the evolution of tyrosinase proteins in *M. truncata*, phylogenetic analyses were carried out.

The fifteen *tyrosinase* transcripts were translated into amino acid sequences, mapped to a reference tyrosinase domain (PF00264) and preliminarily aligned using Clustal-W with default parameters (Larkin et al. 2007). The alignment indicated that most contigs represented gene fragments that mapped to different, non-overlapping regions (Chapter 6 Supplementary Table 1 – Appendix B Supplementary files), only those fragments which overlapped were selected for further analysis. From the initial alignment of fifteen contigs, five contigs were identified as potential paralogues, the remaining ten contigs could still be paralogues but could not be included in analysis as they did not overlap with conserved regions of the tyrosinase reference domain.

The derived amino acid sequences of the five selected contigs were added to a previously published bivalve tyrosinase alignment (Aguilera et al. 2014) and the phylogeny was determined using the method of Aguilera et al. (2014). Briefly, alignments were created using the MAFFT algorithm (Katoh et al. 2005), refined using the RASCAL webserver (Thompson et al. 2003) and analysed with Gblocks 9.1b (Castresana 2000) to select conserved regions. The final alignment was used to run three phylogenetic models: Neighbor-joining (NJ) reconstructions were performed using MEGA 5.2.2 (Tamura et al. 2011), Maximum-likelihood (ML) trees were constructed using RAxMLGUI v. 1.3 (Silvestro and Michalak 2012) and Bayesian inferences (BIs) were performed using MrBayes v. 3.2 (Ronquist et al. 2012). Data from the three models were manually combined to produce a consensus tree (Figure 6-1).

6.2.3 Tissue distribution of biomineralisation gene expression

6.2.3.1 Experimental Design

Animals (same collection as above, n = 5, mean shell length = 60 mm ± 2.06 S.E.) were dissected into six different tissues: mantle, siphon, gill, foot, digestive gland and gonad to investigate the expression of biomineralisation genes in each tisssue.

6.2.3.2 RNA extraction, Quantitative-PCR and Statistical Analysis

RNA was extracted from each sample and reverse transcribed into cDNA as per Chapter 2. A total of seven candidate biomineralisation genes were selected for tissue distribution gene expression analysis (Appendix A - Primer tables). The *ribosomal 18s* gene was selected as a housekeeping reference gene as recommended by previous work on a *M. truncata*'s sister species – *Mya arenaria* (Siah et al. 2008). qPCR was performed as per 2.5.4.

Data were non-normally distributed (due to the number of zeros or very low values) and could not be transformed to reach normality. Given the non-normal distribution and unbalanced design due to RNA degradation in some samples (mantle n = 5, siphon n = 4, gill n = 5, foot n = 3, digestive gland n = 5, gonad n = 5), data were compared using 95% confidence intervals around the mean average fold difference and compared to zero, ie if the confidence interval overlapped with zero (indicating the fold difference was equal to zero) or not. For additional stringency, mantle gene expression data were tested against a set median of zero using the non-parametric Wilcoxon Signed Rank Test.

6.3 Results

6.3.1 *Mya truncata* mantle transcriptome

RNA-Seq reads from the mantle tissue of nine animals were assembled to create 20,106 contigs in the final *de novo* transcriptome. 19 % of contigs were assigned putative functions using BLAST sequence similarity searching (below an e-value of 1e $^{-10}$).

The top 50 most highly expressed, annotated, transcripts included many putatively involved in muscle contraction (40 %) such as: ,*myosin, paramyosin, tubulin, tropomyosin* and *actin*. Energy production was also a dominant process, with annotation in twelve transcripts (24 %), eg *NADH dehydrogenase* and *cytochrome c oxidase*. Other notable transcripts included three encoding putative biomineralisation genes– *calponin-3, calponin-2* and *tyrosinase*; and two chaperone genes – *heat shock protein 90-alpha 1* and *heat shock protein 70* (Table 6-1).

When the mantle transcriptomes of *M. truncata* and *L. elliptica* were compared using tBLASTx, 17 % of the *M. truncata* contigs showed similarity to an *L. elliptica* contig (below an e-value of $1e^{-10}$; Chapter 6 Supplementary Table 2 – Appendix B Supplementary files). The top 50 most similar contigs included one notable biomineralisation gene – *tyrosinase* (Chapter 6 Supplementary Table 3 – Appendix B Supplementary files). All of the candidate biomineralisation genes selected for qPCR also showed strong sequence similarity to (below an e-value of $1e^{-15}$) a *L. elliptica* contig (Chapter 6 Supplementary Table 2 – Appendix B Supplementary files).

Table 6-1. The top 50 most highly expressed annotated (blastx, cut-off <1e^{-10}) contigs in the *Mya truncata* mantle transcriptome.

Contig ID	Contig length	No. of reads	Description	Species	Common name	E-value
2653	1206	5190	Paramyosin	*Crassostrea gigas*	Pacific oyster	5.22E^{-111}

270	2488	4806	Myosin heavy chain, striated muscle-like	*Aplysia californica*	Californian sea hare	0
9876	569	3952	Paramyosin	*Mytilus galloprovincialis*	Mediterranean mussel	$8.82E^{-034}$
1535	1479	3676	Myosin heavy chain	*Placopecten magellanicus*	Atlantic deep-sea scallop	$5.00E^{-141}$
14488	307	3308	Alpha-L1 nicotinic acetyl choline receptor	*Brugia malayi*	Elephantiasis nematode	$5.66E^{-015}$
13639	350	3165	Actin	*Marsupenaeus japonicus*	Japanese tiger prawn	$9.48E^{-055}$
11333	482	2504	Paramyosin	*Mytilus galloprovincialis*	Mediterranean mussel	$7.35E^{-063}$
8180	661	2106	Actin	*Drosophila persimilis*	Fruit fly	$2.14E^{-071}$
12700	401	2079	Myosin, regulatory light chain	*Mercenaria*	Hard-shell clam	$1.66E^{-030}$
15612	255	1993	Myosin heavy chain, striated muscle-like	*Aplysia californica*	Californian sea hare	$9.36E^{-031}$
18456	149	1981	Actin subfamily protein	*Acanthamoeba castellanii*	Soil amoebae	$3.23E^{-013}$
2367	1265	1673	Elongation factor 1 alpha	*Mytilus galloprovincialis*	Mediterranean mussel	0
15826	245	1541	Transcript Antisense to Ribosomal RNA (Tar1p)	*Medicago truncatula*	Barrel clover	$3.48E^{-017}$
433	2216	1533	NADH dehydrogenase subunit 2 (mitochondrion)	*Mya arenaria*	Soft-shelled clam	$7.77155e^{-94}$
16924	203	1424	Calponin-2	*Crassostrea gigas*	Pacific oyster	$9.27E^{-013}$
209	2646	1354	Protein disulfide-isomerase	*Crassostrea gigas*	Pacific oyster	$1.23E^{-173}$
619	1969	1341	Arginine kinase	*Pholas orientalis*	Oriental angel wing	0
12154	439	1280	Ribosomal protein rps12	*Eurythoe complanata*	Fire worm	$1.81E^{-056}$
190	2707	1243	Myosin heavy chain, striated muscle-like	*Aplysia californica*	Californian sea hare	0
19872	107	1227	Actin-4	*Toxocara canis*	Dog roundworm	$3.03E^{-011}$
3040	1140	1200	Myosin heavy chain	*Pecten maximus*	King scallop	$2.55E^{-096}$
13067	381	1197	Transcript Antisense to Ribosomal RNA (Tar1p)	*Medicago truncatula*	Barrel clover	$7.13255e^{-18}$

4048	1004	1161	Myosin heavy chain, striated muscle	*Crassostrea gigas*	Pacific oyster	1.09E^{-121}
12702	401	1042	Ligand-gated ion channel 4-like	*Aplysia californica*	Californian sea hare	6.99E^{-013}
19227	124	1041	Myosin, regulatory light chain	*Macrocallista nimbosa*	Sunray venus clam	1.21E^{-012}
187	2730	1030	Heat shock protein HSP 90-alpha 1	*Crassostrea gigas*	Pacific oyster	0
1953	1363	1002	Fructose-1, 6-bisphosphate aldolase	*Meretrix meretrix*	Orient clam	1.19E^{-077}
5581	846	993	60S ribosomal protein L4	*Crassostrea gigas*	Pacific oyster	3.00803e^{-129}
15404	264	979	Cytochrome c oxidase subunit III (mitochondrion)	*Mya arenaria*	Soft-shelled clam	1.14E^{-031}
17523	181	963	Cytochrome c oxidase subunit III (mitochondrion)	*Mya arenaria*	Soft-shelled clam	2.93134e^{-19}
16470	221	960	Calponin-3	*Pinctada fucata*	Pearl oyster	8.72E^{-013}
5696	836	950	Cytochrome c oxidase subunit I (mitochondrion)	*Mya arenaria*	Soft-shelled clam	1.93028e^{-120}
15046	280	944	Myosin heavy chain	*Argopecten irradians*	Atlantic bay scallop	3.17082e^{-18}
15886	243	943	Myosin, essential light chain	*Mercenaria mercenaria*	Hard-shell clam	1.12192e^{-15}
7264	717	919	Beta-actin	*Meretrix meretrix*	Orient clam	7.06472e^{-133}
1680	1435	903	Tubulin alpha-1 chain	*Harpegnathos saltator*	Indian jumping ant	0
13124	378	870	Tropomyosin	*Tresus keenae*	Horse clam	5.07894e^{-48}
495	2111	826	Polyadenylate-binding protein 4	*Crassostrea gigas*	Pacific oyster	0
11402	478	812	NADH dehydrogenase subunit 1 (mitochondrion)	*Mya arenaria*	Soft-shelled clam	1.27497e^{-35}
162	2807	787	Phosphoenolpyruvate carboxykinase [GTP]	*Crassostrea gigas*	Pacific oyster	0
4251	980	775	NADH dehydrogenase subunit 5 (mitochondrion)	*Mya arenaria*	Soft-shelled clam	9.23E^{-110}
211	2639	771	Tyrosinase (tyr-3)	*Crassostrea gigas*	Pacific oyster	4.33E^{-086}
268	2492	741	Heat shock protein 70	*Corbicula fluminea*	Golden clam	0

7748	686	734	60S ribosomal protein	*Aplysia californica*	Californian sea hare	$2.16E^{-088}$
1122	1641	729	Voltage-dependent anion channel 2	*Haliotis diversicolor*	Many-coloured abalone	$1.26E^{-095}$
10131	550	713	60S ribosomal protein L4	*Crassostrea gigas*	Pacific oyster	$6.82E^{-022}$
2645	1207	694	Receptor for activated protein kinase	*Scrobicularia plana*	Peppery furrow shell	$1.44E^{-169}$
10039	555	689	NADH dehydrogenase subunit 1 (mitochondrion)	*Mya arenaria*	Soft-shelled clam	$1.81E^{-033}$
2970	1152	687	Cytochrome b (mitochondrion)	*Mya arenaria*	Soft-shelled clam	$2.93E^{-140}$
3020	1144	672	ADP,ATP carrier protein	*Crassostrea gigas*	Pacific oyster	$7.38E^{-136}$

6.3.2 Tyrosinase bivalve phylogeny

At least five putative *tyrosinase* paralogues were identified in the *M. truncata* mantle transcriptome. The derived amino acid sequences were added to a previously published molluscan tyrosinase phylogenetic analysis (Aguilera et al. 2014). Many of the nodes had low support values (Figure 6-1, Chapter 6 Supplementary Figures 1, 2 & 3 – Appendix B Supplementary files), however some reoccurring patterns were observed across the three phylogenetic models used. Two major clades (A&B) were resolved (although it should be noted that it is likely there is some confusion in the literature with regards to nomenclature which I believe explains clade discrepancies) in addition to two large, independent expansions in the taxa *Crassostrea* and *Pinctada*. In general the *M. truncata* transcripts clustered most closely with the A-type tyrosinases. One of the *M. truncata* transcripts (contig00553) clustered loosely with the *L. elliptica A* genes and *TyrA3* genes from *Crossostrea gigas* and *Pinctada fucata*, whilst all other copies clustered with the other *L. elliptica* gene, *TyrB*, and showed evidence of early expansion in the *M. truncata* genome (Figure 6-1).

6.3.3 Tissue distribution of biomineralisation gene expression

Seven candidate biomineralisation transcripts were identified for further analysis based on sequence similarity to known biomineralisation genes, some of which were also present in the top 50 most highly expressed annotated transcripts (Table 6-1), the shell or mantle proteome (Arivalagan et al. 2016), or any combination of the three (Appendix A - Primer tables).

A mantle-specific signal was detected for all candidates (Figure 6-2). None of the mantle 95% confidence intervals overlapped zero and a non-parametric Wilcoxon Signed Rank test showed that all of the mantle gene expression values were above zero ($P<0.05$). In contrast, all of the remaining tissues tested (except siphon for 16470 and 395) were not significantly different to zero.

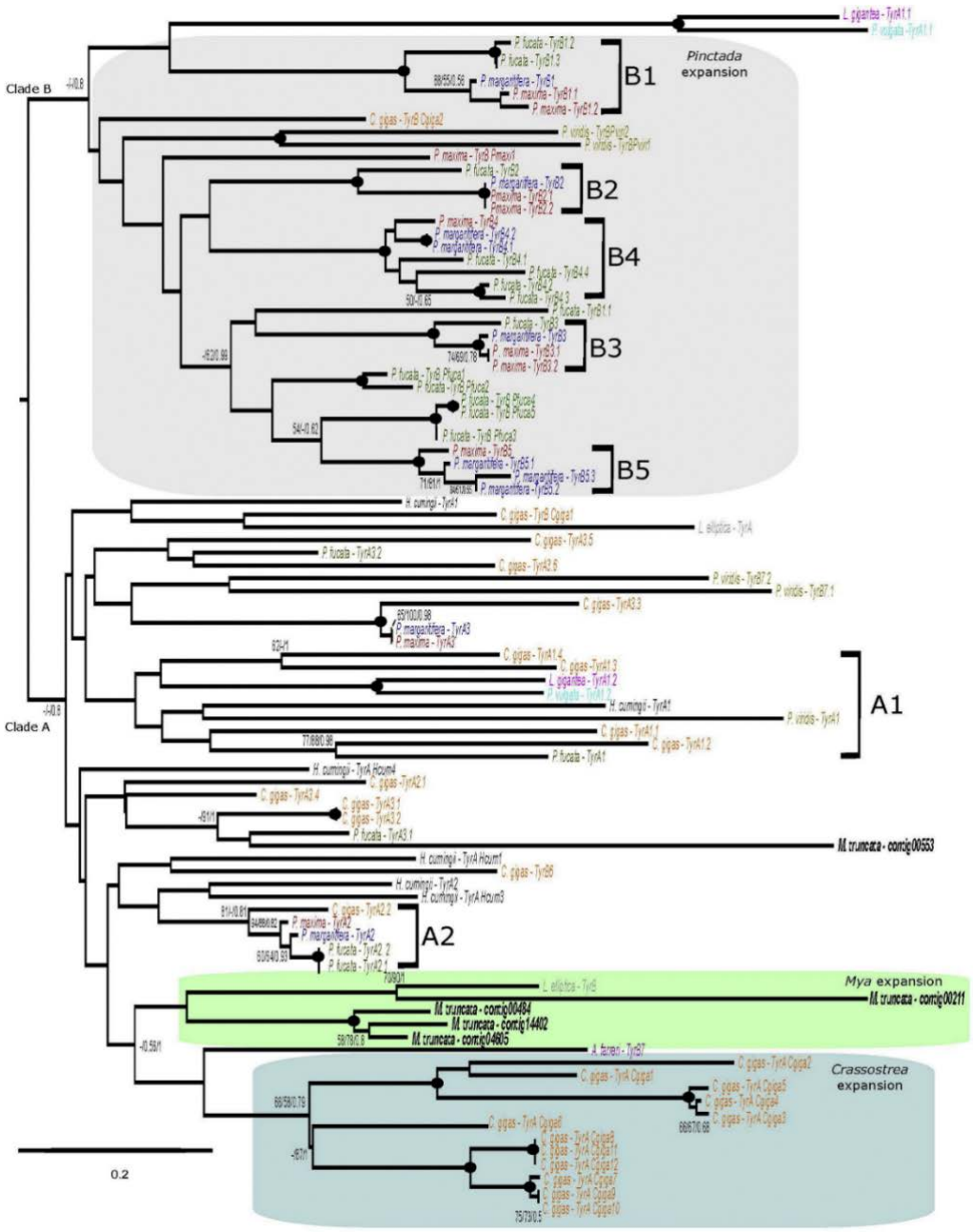

Figure 6-1. Phylogenetic analysis of tyrosinase proteins in shell-building molluscs. A consensus midpoint-rooted tree based on Neighbor-Joining (NJ) topology. Only bootstrap support values > 50% and posterior probabilities > 0.50, from three different phylogenetic models, are shown at the nodes as follows: NJ bootstrap support/Maximum Likelihood (ML) bootstrap support/Bayesian Posterior Probabilities (BPP). A black dot at the node represents NJ and ML bootstrap > 90% and BPP > 0.9. Tree labels and nomenclature are consistent with Aguilera et al. (2014) in order to provide an easy visual comparison between the two studies. See Chapter 6 Supplementary Figures 1, 2 & 3 (Appendix B Supplementary files) for trees generated from each model.

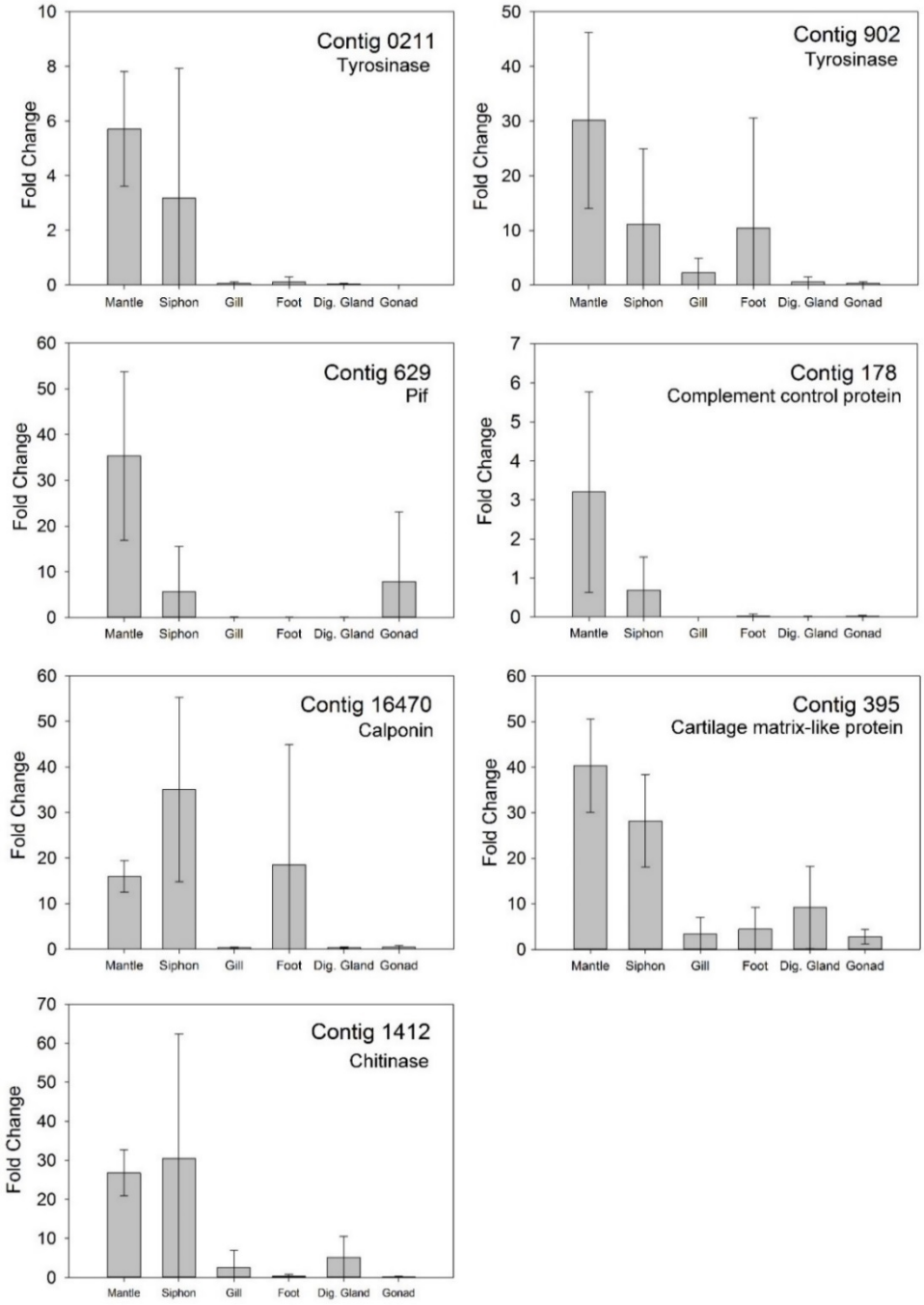

Figure 6-2. Tissue distribution expression patterns of *Mya truncata* candidate biomineralisation genes determined via qPCR (mean average ± 95% confidence intervals). Fold change calculated as $2^{-\Delta\Delta CT}$ using Ribosomal 18s as an internal housekeeping gene.

6.4 Discussion

6.4.1 *Mya truncata* mantle transcriptome

Presented in this chapter is the first substantial molecular resource for *M. truncata*, which is valuable data for biomineralisation and comparative studies. The *M. truncata* mantle transcriptome was similar in size (~20,000 contigs), composition of most highly expressed transcripts and percentage annotation to other previously characterised bivalve mantle transcriptomes (Clark et al. 2010, Joubert et al. 2010, Niu et al. 2013, Shi et al. 2013, Freer et al. 2014); in particular it shares specific features with that of the Antarctic clam *L. elliptica*. When the *M. truncata* and *L. elliptica* mantle transcriptomes were compared using tBlastx, 17% of contigs were shared (below an e-value of $1e^{-10}$) representing a highly conserved core set of genes (Clark et al. 2010). Both the *M. truncata* and *L. elliptica* mantle transcriptomes were heavily dominated by muscle related genes (Table 6-1; *myosin, actin* etc.), reflecting the contractile nature of this organ. In addition, putative mitochondrial respiratory chain genes (*NADH dehygrogenase, cytochrome c, arginine kinase* etc) were highly expressed in both, demonstrating that the mantle is a metabolically and transcriptionally active tissue. The primary species in this thesis is *L. elliptica* and *M. truncata* has been chosen as a northern hemisphere, temperate comparison. *M. truncata* is often reported as an Arctic bivalve (Gillis and Ballantyne 1999, Camus et al. 2002). Animals in this chapter however, were sampled from a more southerly, and hence much warmer, region of their distribution on the West coast of Scotland. As explained in Chapter 1, *M. truncata* and *L. elliptica* are ecologically and morphologically very similar, but their physical environments (Arctic to temperate versus Antarctic), geographical extent (ranging from Arctic through subboreal to temperate versus Southern Ocean exclusively) and evolutionary history (phylogenetically distant relatives) differ significantly. In addition, the two species have also different shell microstructure (see figures in Chapter 1). As a result of their independent evolutionary trajectories in dissimilar physical conditions, with diverse selection pressures, *M. truncata* which inhabit temperate regions have a higher metabolic rate and shorter

lifespan than *L. elliptica* (Peck et al. 2002, Camus et al. 2003, Camus et al. 2005, Philipp and Abele 2010)

As well as sharing a core set of highly conserved genes, the *M. truncata* and *L. elliptica* most highly expressed mantle transcripts share more specific similarities at the individual gene level. Here the present chapter focussed on the genes likely to be involved in biomineralisation. Both transcriptomes have a single *tyrosinase* gene in the most highly expressed set of transcripts, *L. elliptica* has *Tyrosinase B* and *M. truncata* has *Tyrosinase A3*. Tyrosinase is a biomineralisation protein involved in the formation of the shell matrix and periostracum (Sanchez-Ferrer et al. 1995, Zhang et al. 2006, Huning et al. 2013), its extremely high expression in the mantle organ of both clam species provides further evidence for its important role in shell deposition. A less well characterised biomineralisation candidate is calponin; both species have two *calponin* genes in the mostly highly expressed transcripts. Calponin proteins are typically involved in muscle contraction and interact closely with other muscle action proteins such as actin and myosin (Matthew et al. 2000). Its role in muscle contraction however, is thought to be primarily cross-linking and stabilisation of the muscle fibres (Jensen et al. 2014), and it could be hypothesised to carry out a similar role in the stabilisation of the shell protein matrix. Several pieces of evidence support this idea: 1.) It is highly expressed at the transcript level in the mantle of the two clam species and a species of pearl oyster, *Pinctada martensii* (Shi et al. 2013). 2.) It has been demonstrated to be involved in the biomineralisation of bone in humans (Ueda et al. 2002). 3.) The shell matrix protein, PFMG8, has been shown (*in silico*) to contain a calponin domain which has a calcium binding site (Evans 2012). 4.) It has a mantle specific expression pattern comparable with other tissue (Figure 6-2) and finally, 5.) Calponin proteins were found in both the shell and mantle proteome of *M. truncata* where it may interact with myosin and contribute to shell elasticity (Arivalagan et al. 2016).

One noticeable difference between the *M. truncata* and *L. elliptica* mantle transcriptomes concerns the constitutive expression of heat shock proteins (HSPs). HSPs are involved in protein-folding and chaperoning and are either constitutively expressed, or induced in response to stress (Hartl 1996). Both *heat shock protein 70* (*Hsp70*) and *heat shock*

protein 90-alpha (*Hsp90*) were highly expressed in the *M. truncata* mantle, whereas in *L. elliptica* there was no high expression of any HSP family members [above 300 reads – which was the cut-off used by Clark et al. (2010)]. One form of *Hsp70* is classically regarded as inducible, rather than constitutive, and forms part of a classic "stress" response in many organisms (Clark and Peck 2009, Clark et al. 2008). One possible explanation for the high background expression of inducible *Hsp70* in *M. truncata* is that the animals in the present chapter were sampled from a southerly, and hence warmer, point in their overall distribution (summer sea surface temperatures around Oban have been recorded >14°C) where they are close to their upper thermal tolerance limits (Amaro et al. 2005). Work by Amaro et al. (2005) demonstrated that *M. truncata* at its southern distribution limit (the Frisian Front in the North Sea experiencing water temperatures >14°C) have low numbers of ripe oocytes and frequent years of poor recruitment, indicating the southern populations could be in a chronic state of low-level thermal stress. Hsp90 on the other hand, is thought of as a constitutively expressed protein, and its high expression in the *M. truncata* mantle transcriptome is likely to represent normal cellular processes in the mantle, similar to other marine invertebrate species (Huang et al. 2013).

6.4.2 Tyrosinase bivalve phylogeny

Tyrosinase is a multifunctional well-characterised, shell-associated protein which has been shown to have a functional role in the cross-linking of the soluble periostracum precursor (the periostracin) to form an insoluble periostracum (Waite et al. 1979); to be localised in the prismatic layer of shell (Nagai et al. 2007); and to be expressed in the pallial mantle and hence potentially also involved in the nacreous layer of the shell (Takgi and Miyashita 2014). In addition, tyrosinase is involved in numerous other biological processes such as innate immunity, pigmentation and wound healing (Zhang et al. 2006, Aguilera et al. 2014). This chapter has identified at least five gene copies of *tyrosinase* in the *M. truncata* mantle transcriptome which are likely to be the result of several gene duplication events followed by sub-functionalisation (Force et al. 1999). It is possible that an expansion and subsequent sub-functionalisation has produced *tyrosinase* paralogues which have completely new functions besides shell formation, for example in the immune system, as suggested by Wang et al. (2009) and Esposito et al. (2012). Given

the well-characterised and clearly important nature of tyrosinase in molluscan shell formation however, it is important to investigate its evolution in order to better understand how mollusc shell is produced between species. Much work into the molecular control of shell production has searched for a "conserved" molluscan shell tool-kit, and, if such a conserved mechanism exists, it has been hypothesised tyrosinase could be central to its function (Aguilera et al. 2013, Aguilera et al. 2014, Aguilera et al. 2017, Arivalagan et al. 2017). *M. truncata* tyrosinase amino acid sequences were therefore investigated within a phylogenetic context. Similar to Aguilera et al. (2014) and other previous tyrosinase phylogenies (Aguilera et al. 2013), many of the nodes had low support values (Figure 6-1, Chapter 6 Supplementary Figures 1, 2 & 3 – Appendix B Supplementary files). Adding *M. truncata* sequence data to three independent phylogenetic analyses did not alter the overall tree topology and the major patterns described by Aguilera and colleagues were largely still resolved – further validating their work.

Four of the five *M. truncata* tyrosinase amino acid sequences formed a well-supported cluster with *L. elliptica* tyrosinaseB. It is possible this cluster represents an expansion within the *M. truncata* genome. The fifth *M. truncata* sequence grouped with tyrosinaseA3 sequences from *C. gigas* and *P. fucata*, and *L. elliptica* tryrosinaseA1. Given that both of the *L. elliptica* sequences group with the *M. truncata* sequences it is possible they have evolved under at least some similar selection pressures with regards to shell growth. *M. truncata* however, show evidence for a tyrosinase expansion, which is absent in *L. elliptica.* Antarctic marine invertebrates are largely stenothermal due to evolution at constant cold temperatures for long periods of geological time (Rogers 2007). In this chapter I find that, for the tyrosinase gene family, Antarctic clams are less diverse than their temperate counterparts *M. truncata*, and other marine shelled-molluscs. An explanation for higher tyrosinase diversity in *M. truncata* than *L. elliptica* could be due to *M. truncata's* wider geographical distribution and spread over environmental gradients. As a species, *M. truncata* could require more diverse molecular machinery to cope with the diverse environments they inhabit; supporting the hypothesis that diversity is positively correlated to environmental heterogeneity and stress (Van Valen 1965, Nevo 2001).

6.4.3 Tissue distribution of biomineralisation gene expression

As expected, all of the putative biomineralisation candidate genes selected for tissue distribution analysis (Appendix A - Primer tables) showed a mantle/siphon-specific expression pattern (Figure 6-2). Tyrosinase and pif are well characterised shell proteins (Kouhei Nagai et al. 2007, Suzuki et al. 2009); they have been shown to respond to shell damage in *L. elliptica* (Chapter 4, Chapter 5), and the mantle/siphon specific expression patterns found in *M. truncata* provide further evidence to support their hypothesised functional role in shell deposition in the two clam species. Calponin is likely to be involved in biomineralisation (as discussed at length above) however, like many biomineralisation proteins, it is multi-functional and also involved in muscle contraction where it interacts with myosin and hence it also showed variable expression in the foot (a muscle). Cartilage matrix proteins are involved in calcium phosphate biomineralisation in vertebrates where they bind to calcium phosphate crystals and form part of an extracellular matrix (Acharya et al. 2014). Arivalagan et al. (2016) identified cartilage matrix protein in the *M. truncata* shell proteome, and taken together with the mantle-specific gene expression pattern found in the present chapter, it is likely to play a similar matrix-like role in mollusc biomineralisation. Chitinase is an enzyme hypothesised to be involved in mollusc shell matrix construction, as well as immunity (Badariotti et al. 2007a). As explained in Chapter 4, the function of chitinase and chitinase-like proteins have been investigated in arthropods with chitinous exoskeletons; they are typically involved in moult-cycles, wound healing and tissue repair (Chen et al. 2004, Bonneh-Barkay et al. 2010), however their exact function in the mollusc shell matrix is still unclear. Previous work shows chitinase expression is up-regulated in response to injury after 21 days in young, but not old, *L.elliptica* (Husmann et al. 2014). In Chapter 4 of this thesis there was also variable expression in *L. elliptica* over time in response to shell damage and similarly Huning et al. (2016) also found high variability in a chitin related gene in adult mussel mantle tissue. The mantle specific expression pattern shown in the present chapter, as well as chitinase presence in the *M. truncata* shell proteome (Arivalagan et al. 2016), provides further support for its active involvement in shell deposition. More research however, is required to understand its hugely variable

expression in young versus old animals and its exact function during matrix formation and shell secretion.

Complement control proteins are multifunctional and involved in both the immune system (Ferreira et al. 2010) and possibly biomineralisation (Arivalagan et al. 2016). Bivalve shell and mantle combined act as a barrier to the external environment, and as such are likely to be entwined with immune processes. Disentangling immune and biomineralisation mechanisms represents a significant challenge for researchers trying to understand how molluscs build their shells. The challenge is partly due to the dual role of haemocytes both as immune cells and hypothesised calcium carbonate chaperones (Mount et al. 2004). In addition to the observed mantle specific expression of complement control protein in the present chapter and previous reports of immune genes such as *mytilin* responding to damage in *L. elliptica* (Chapter 4), Arivalagan et al. (2016) found immune proteins in the *M. truncata* shell proteome (and verified their presence was not due to contamination). The expression of immune genes in the *M. truncata* mantle, as well as the immune proteins found in the shell, could be explained in several ways: 1.) General haemocyte circulation in the mantle could result in coincidental incorporation into the shell as an accidental bi-product of their immune function (as per the mantle's role as a barrier). 2.) Whilst haemocytes actively deposit calcium carbonate to the shell secretion site they could be coincidentally trapped in the shell matrix space, as an accidental bi-product of their role in biomineralisation, as proposed by Arivalagan et al. (2016); or 3.) Immune proteins could serve a genuine dual-function role, both in aiding biomineralisation during calcium carbonate secretion from haemocytes, and possibly also at a structural matrix-level, as well as fighting infection as an anti-microbial peptide in the shell, and in the circulating haemocytes in the mantle.

6.5 Conclusions

This chapter presents the first substantial molecular resource for *M. truncata*. The mantle transcriptome was 454-sequenced, *de novo*-assembled and BLAST sequence similarity-annotated to produce a total of 20,106 contigs, of which approximately 19 % were

assigned putative functions. The mantle transcriptomes of *M. truncata* and the Antarctic clam (*L. elliptica*) were compared using tBLASTx and overall, shared a core complement (17 %) of highly conserved transcripts. Looking at the most highly expressed genes in the two species showed that many of the dominant biological functions (contraction, energy production, biomineralisation) were conserved. The tyrosinase proteins from *M. truncata* were analysed phylogenetically and showed a small expansion which was closely related to *L. elliptica*, however *M. truncata* had more diversity in tyrosinase proteins. The tissue distribution expression pattern of candidate biomineralisation genes was investigated using qPCR; all genes showed a mantle specific expression pattern supporting their hypothesised role in shell secretion. This chapter provides preliminary insights on how clams in different environments (temperate versus polar) build their shells, a topic which will be explored further in the final chapter of this thesis.

Chapter 7 SYNTHESIS AND FUTURE DIRECTIONS

The overall objective of this project was to better understand calcification pathways in two non-model bivalve species, the Antarctic clam (*Laternula elliptica*) and the temperate clam (*Mya truncata*). This general objective was broken down into 12 more specific aims, all of which were completed in Chapters 3 - 6 of this thesis. In this chapter I will briefly recap each of the aims and summarise the related main findings. The core emerging conclusions will then be synthesised and recommendations made for future work.

7.1 Chapter aims and the corresponding summaries of findings

Chapter 3:

1.) Provide a detailed characterisation of the *L. elliptica* mantle anatomy and ultrastructure.

Using histology, light and electron microscopy, the *L. elliptica* adult mantle tissue anatomy and mantle epithelium cell ultrastructure were described and a schematic illustration was produced. The detailed observations of the mantle revealed many conserved features with other shell producing molluscs, including secretory vesicles (which could contain calcium carbonate) that migrate towards the shell.

2.) Identify proteins in the *L. elliptica* nacreous shell layer using proteomics.

The proteome of the nacreous shell layer was characterised, 37 shell matrix proteins were identified, many of which corresponded to previously identified mollusc nacre shell matrix proteins, and two unique features were revealed; the presence of a zinc-dependent metalloprotease and a novel T-rich mucin-like protein.

3.) Understand the function of *L. elliptica* biomineralisation candidates by localising their gene expression to tissue, cellular and subcellular resolution.

The expression patterns of seven candidate biomineralisation genes were investigated to increase understanding of their potential functions. Four genes showed increased expression in the mantle and siphon tissues, and all seven genes had some expression in other tissues indicating they have multi-functional roles

aside from biomineralisation. *In situ* hybridisation of the same transcripts revealed five different and discrete cellular expression patterns which corresponded to different secretory regions of the mantle, providing further evidence that the mollusc mantle is modular on a molecular as well as anatomical level. The subcellular expression patterns suggested that all seven biomineralisation candidates were associated with vesicles, the exact function of which is unknown, but they may be involved in calcium carbonate transport and secretion.

Chapter 4:

4.) Estimate the timing of the *L. elliptica* transcriptional response to shell damage.

The preliminary transcriptional profiling of *L. elliptica* during a shell damage-repair experiment revealed shell repair was a slow process, lasting at least 2 months. Different biological processes were important at different time scales during repair, thus enabling more targeted analyses in future experiments.

5.) Identify candidate *L. elliptica* biomineralisation genes involved in the response to shell damage.

A low level of biomineralisation signal was detected in the transcriptional profiling experiment and these candidates were taken forward for further characterisation (in Chapter 5), however, in general, the "classic" biomineralisation candidate genes were largely unchanged over time. In order to better characterise and understand the molluscan biomineralisation pathway, the regulation of genes with the same expression pattern as "classic" biomineralisation genes, using gene –network analysis was recommended.

6.) Understand the spatial location of molecular mechanisms in the response to shell damage in *L. elliptica*.

When investigating the spatial localisation of repair mechanisms, different genes showed different spatial patterns in relation to a single drilled hole. The spatial pattern revealed is important for all future damage-repair experiments as the location and number of holes, in relation to where the underlying tissue is sampled from, could bias transcriptomic results. It was recommended that future work investigating damage-repair mechanisms in large adults should therefore drill several holes around the edge of the shell in order to stimulate expression across the entire mantle edge and maximise the tissue samples that can be taken for analysis.

Chapter 5:

7**.) Increase the temporal resolution of understanding of the *L. elliptica* transcriptional response to shell damage**.

An age-related shell damage-repair experiment was conducted over 4 months with six different sampling time points at three different life history stages. Shell repair mechanisms were still on-going at 4 months and in general, animals did not fully heal and calcify their holes within the experiment. Time points with major differential expression were tightly coupled to some, but not all, histological metrics. Overall the transcriptional response to shell-damage was highly dynamic, which in some cases involved a time-lag between up-regulation of transcripts and observable changes at the histological and shell-repair level.

8.) Understand the effect of age on the *L. elliptica* transcriptional response to shell damage.

The damage-repair experiments in this chapter found that age causes a transcriptional delay in the response to shell damage and in addition, older animals showed more signs of stress, and hence found shell repair more difficult. It was hypothesised that juvenile animals were already investing all available energy into somatic growth, including shell production, whereas older adults had to divert energy away from reproduction to repair the shell.

Surprisingly there were very few shared differentially expressed genes between time points, or ages, indicating that animals at different ages, and stages in repair, were transcriptionally different. Only one biomineralisation gene, *pif*, was shared between key experimental time points. In contrast, different age groups up-regulated different nacre-related biomineralisation genes, indicating subtle differences in the repair process.

9.) Provide and apply a regulatory gene network of the *L. elliptica* mantle to aid the understanding of biomineralisation pathways

Gene expression profiles from the shell damage-repair experiments were used to create the first comprehensive regulatory gene network for mollusc mantle tissue. The network was used to identify putative new biomineralisation candidates, which should be further investigated in future work.

Chapter 6:

10.) Provide a molecular resource to aid the study of biomineralisation in *M. truncata* by sequencing, assembling and putatively annotating the adult mantle transcriptome.

The first substantial molecular resource for *M. truncata* was produced. The mantle transcriptome was 454-sequenced, *de novo*-assembled and BLAST sequence similarity-annotated to produce a total of 20,106 contigs, of which approximately 19 % were assigned putative functions.

11.) Identify and investigate the expression and phylogeny of candidate biomineralisation genes in the *M. truncata* mantle transcriptome

The tyrosinase proteins from *M. truncata* were analysed phylogenetically and showed a small expansion in this species. The tissue distribution expression pattern of candidate biomineralisation genes was investigated using qPCR; all genes showed a mantle specific expression pattern supporting their hypothesised role in shell secretion.

12.) Compare the mantle transcriptome and biomineralisation genes of *M. truncata* to that of the Antarctic clam, *L. elliptica*

The mantle transcriptomes of *M. truncata* and the Antarctic clam (*L. elliptica*) were compared using tBLASTx and overall, shared a core complement (17 %) of highly conserved transcripts. Looking at the most highly expressed genes in the two species showed that many of the dominant biological functions (contraction, energy production, biomineralisation) were conserved. The *tyrosinase* expansion in *M. truncata* genome was closely related to the *L. elliptica tyrosinases*, however *M. truncata* had more diversity in *tyrosinase* transcripts, possibly reflecting the broad range of environmental conditions they inhabit.

7.2 Two species comparison

Most of the work in this thesis focused on the Antarctic species, *L. elliptica*, and *M. truncata* was initially chosen as a temperate comparison. As work progressed however, the complexity of the *L. elliptica* shell damage-response was revealed, especially in relation to the different life stages, and hence the comparative approach was scaled-down in order to focus on *L. elliptca.* To directly compare the results from each species was difficult as the transcriptomes, proteomes and tissue distribution expression profiles were carried out using subtly different methods. A qualitative comparison of some of the key findings however, can be useful to draw out broad patterns (Figure 7-1).

In general, both species had a copy of *tyrosinaseA* and *pif* in the shell proteome (*L. elliptica* was just a nacreous shell proteome and *M. truncata* was a whole shell proteome). In addition, both species had a copy of *tyrosinaseB* in the most highly expressed annotated transcripts in the mantle transcriptome which, interestingly, despite the high expression in the mantle, was not present in either shell proteome. For both species, the fact that one *tyrosinase* paralogue was in the proteome (A), and one was absent (B), and likewise, one paralogue was in the most highly expressed mantle transcripts (B), and one was absent (A), strongly suggests each paralogue was carrying out a different function. The presence of shared proteins in both of the shell proteomes, and shared highly expressed transcripts

in the mantle transcriptomes, indicates that - at a molecular level – these two species use at least some similar pathways to build their shells.

Tissue distribution gene expression profiles of candidate genes in *M. truncata* (qPCR) all showed clear mantle-specific expression patterns, whereas candidate genes in *L. elliptica* (semi-qPCR) tended to show expression in other tissues, in addition to the mantle. In general this could indicate that the putative biomineralisation genes in *L. elliptica* are more multifunctional, compared to *M. truncata.* The qualitative differences found between the two species in this thesis provides support for the growing consensus in the literature that molluscan biomineralisation pathways are surprisingly divergent (Jackson et al. 2010, McDougall et al. 2013, Aguilera et al. 2017). For example, even when methods were carefully standardised, there are only a handful of shared proteins between four divergent bivalve shell proteomes [four proteins = carbonic anhydrase, chitin-binding 2, vonwillebrand factor type-A and tyrosinase (Arivalagan et al. 2017)].

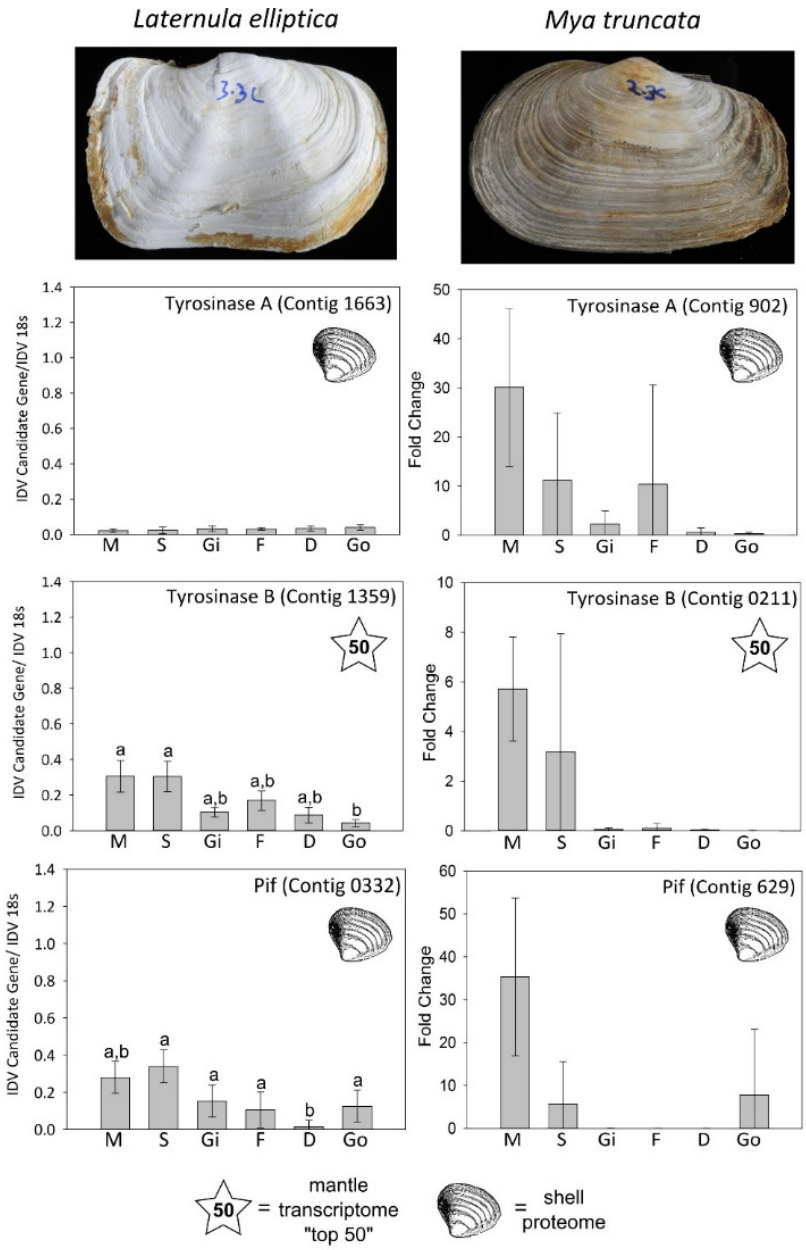

Figure 7-1. Summary of *Laternula elliptica* and *Mya truncata* comparison. Expression of putative biomineralisation genes across six different tissues as determined by semi-quantitative PCR (*L. elliptica,* mean ± S.E. n = 5, statistically significant differences indicated by different letters above bars), and qPCR (*M. truncata,* mean average ± 95% confidence intervals, n = 5). M = Mantle, S = Siphon, Gi = Gill, F = Foot, D = Digestive Gland, Go = Gonad.

During the time-scale of the project, I did conduct a shell damage-repair experiment on *M. truncata*. Due to unexpected difficulties in animal collection, followed by delays at the sequencing centre, data were only received towards the end of the project and hence the results are not included in this thesis. Briefly, preliminary data analysis indicates that adult *M. truncata* found shell damage much more stressful than adult *L. elliptica* and therefore it may be difficult to use the data to learn about biomineralisation as almost all of the up-regulated genes in damaged animals are involved in stress responses and apoptosis. Previous work has shown that animals living at high latitudes, with a longer lifespan, have better wound and shell-healing capacity (Ziuganov et al. 2000). The difference in repair capacity between high and low latitude populations has been hypothesised to be due to the reduced energy expenditure for growth, and greater utilisation for shell and tissue regeneration to sustain homeostasis. In the case of *L. ellipitca* and *M. truncata*, I hypothesise that *L. ellipitca*'s longevity and adaptation to cold environments is only part of their reason for enhanced biomineralisation repair. Their evolutionary history, entwined with damage from iceberg scouring and without heavy predation, has likely resulted in their enhanced biomineralisation repair, compared to *M. truncata*. Previous work has shown that shell-repair in *L. ellipitca* is a phenotypically plastic in response to iceberg scouring (Harper et al. 2012). I predict *L. elliptica*'s phenotypically plastic biomineralisation pathway has been selected for, and is hence an adaptive trait (although this hypothesis has not been explicitly tested).

When reflecting on the comparative aspect of this project it is useful to consider the value and limitations of this kind of approach. As discussed above, directly comparing data from one species to another is difficult as transcriptomes, proteomes, tissue distribution expression profiles etc are often carried out using different methods. In addition, factors such as life stage, season, tank effects, longevity and general capacity to repair have all been shown to affect transcriptional profiles [Chapters 3-6 in addition to Jackson et al. (2007) and Philipp and Abele (2010)]. Wild molluscs are inherently very variable and in addition, gene expression is a dynamic process; it is therefore hard to be certain that transcriptional profiles from a handful of animals are representative of a species in terms of their genome and more specifically, biomineralisation pathway. Overall the comparative approach was useful for qualitatively comparing the molecular control of

biomineralisation in two species. Using a broad approach when comparing mantle transcriptomes and shell proteomes allowed me to look for general patterns in each species. In particular, the approach has generated interesting hypotheses about the role of longevity, ageing and biomineralisation in species living in constantly cold sub-zero waters compared to ones living in very variable temperature regions. To take the comparison further, and really explicitly test these hypotheses, a comprehensive genome for each species would be useful, in addition to a better defined molluscan phylogeny. Finally, methodology should be tightly controlled (same library prep, sequencing depth, sequencing methodology, assembly and downstream analysis) and, in order to do a like-for-like comparison, life stage and other factors such as season should be considered. It could also be interesting to compare the embryonic development of the shell in these two species. Examining the evolutionary development of biomineralisation and increasing understanding on how these organisms create a shell from scratch, will help to assess if there is an evolutionarily conserved molluscan shell secretion mechanism and how the first ever molluscan shell field evolved.

7.3 Conclusion

The combination of methodologies, from traditional histology, to shell damage-repair experiments and ‘omics technologies, provided a holistic understanding of shell production, which was particularly useful for assigning putative biomineralisation functions to genes with no previous annotation. There was reoccurring evidence, across all three *L. elliptica* chapters, for the involvement of vesicles in biomineralisation and in addition for the duplication and subfunctionalisation of *tyrosinase* paralogues. Shell damage-repair experiments revealed biomineralisation in *L. elliptica* was variable, transcriptionally dynamic, significantly affected by age and inherently entwined with immune processes. The high amount of variation in gene expression profiles across 78 individual mantle samples was captured in a single regulatory gene network, which was used to predict the regulation of “classic” biomineralisation genes. There were some general shared patterns in the molecular control of biomineralisation between the two species investigated, but overall, the comparative work in this thesis, coupled to the

growing body of literature on the evolution of molluscan biomineralisation, suggests that biomineralisation mechanisms are surprisingly divergent. This thesis significantly increased publicly available data resources for two species of bivalve, which will feed into both future comparative work, as well as provide the foundation of biomineralisation understanding in each species.

7.4 Future work

The *de novo* transcriptome of a non-model organism was first published only nine years ago (Vera et al. 2008), and since then rapidly emerging sequence technologies have been increasingly applied to the field of biomineralisation (Clark et al. 2010, Joubert et al. 2010, Freer et al. 2014, Clark et al. 2016). To date, a considerable amount of the work on the molecular control of shell production represents the collection and description of much-needed sequence data and initial characterisations of gene and protein expression patterns. Biomineralisation has been revealed to be an incredibly complex process involving the regulation of potentially thousands of genes and tens or hundreds of proteins. Such a complex system is evidently hard to comprehend due to the sheer number of interacting biological variables, each of which being vital for the precise control of shell production. Detangling this immensely complex problem, to understand how molluscs build their shells, and to usefully apply knowledge on the molecular mechanisms underpinning shell production to materials science, aquaculture and ecosystem resilience predictions, requires the continued co-ordination and integration of research efforts.

Studying of non-model species, especially transcriptional profiling of wild invertebrates, can be very challenging. One of the challenges to overcome is the high degree of individual variability and dynamic nature of gene expression. Future work should consider using an experimental design to reduce or overcome this variability, for example matched-pairs where a control sample is taken from the same animal, could help. Matched-pairs would also decrease the number of animals required for the experiment, which is especially useful for *L. ellipitca* considering the logistical difficulties of

collection from Antarctica. If financial constraints were also taken into consideration, another amendment to experimental design could be to prioritise higher replication of fewer experimental time points, which would aid the statistical interpretation of the data. If financial constraints were not an issue however, having a higher replication and more fine scale resolution of time-scales would provide even more detailed understanding of the transcriptional response to shell damage. In addition to a more detailed, fine-scale, time resolution, finer scale spatial resolution would also be useful. One interesting experiment to conduct would be to drill a single hole near the ventral edge of the shell in small juvenile animals, similar to Chapter 4, but instead of sampling broad regions of tissue, the whole animal could be fixed and whole mount *in situ* hybridisations could be used to understand the finer scale spatial response of different biomineralisation and immune candidates to damage. In addition, electron microscopy could be used, with cryo-fixation in order to assess calcium carbonate in vesicles, to investigate the epithelium of both damaged and control animals. Animals from the same experiment could be used for transcriptional profiling, but with a higher spatial resolution, for example 10 regions instead of 4, and also carefully separating the mantle edge from mantle pallial. With the rapid progress of single-cell sequencing technology, the amount of tissue required to gain a good quality assembly is reducing, and hence the spatial precision of expression data can therefore increase. Again these types of experiments would ideally be done with much higher replication when using wild animals.

The argument for a finer scale resolution in both temporal and spatial molecular data is fundamental across all of biology. The more we learn about a plethora of cellular mechanisms, including biomineralisation, the more complicated the picture gets. If we are to gain a truly functional understanding of molluscan biomineralisation, the complete signalling pathway that controls it, or indeed how it first evolved, we urgently need to understand the function of the vast myriad of novel, unannotated, genes and proteins that have been sequenced from the mantle transcriptomes and shell proteomes in various molluscs in the last decade. These unannotated genes are so often overlooked, with a preference to search for the “classic” biomineralisation genes. Work in this thesis has shown how various methods can be used to increase our understanding of gene function and a dominant feature of this work has been to assess the location and timing of gene

expression and protein levels in relation to certain known secretory regions of the mantle, known layers of the shell or in response to a known stimulus, i.e. shell damage. Clearly, the finer the resolution of the data we can obtain, the more detailed our understanding becomes of how these genes and proteins might function. Continuing to seek more molecular resources and experimental evidence on gene and protein function, which is not a trivial task, will certainly continue to be a major contribution to the field of molluscan biomineralisation in the next decade.

In reference to the clear need for more descriptive and experimental molecular data in the field of biomineralisation, much of the work in this thesis culminated in the production of a regulatory gene network for mantle tissue in *L. elliptica*, this resource will be hugely useful for many fields, and in particular it could provide the foundation of future biomineralisation studies in this species, as well as acting as a proof-of-concept, methodologically, for other species. Further characterisation (*in situ* hybridisations, tissue distribution profiles etc) of the first-neighbor connections (highlighted in Chapter 5) in order to test their hypothesised role in biomineralisation, and essentially understand their function, would be a good starting point for future work.

For the first time in molluscan biomineralisation research, Chapter 5 of this thesis found that age was significant factor in the transcriptional regulation of biomineralisation. The animals used in this thesis spanned many life stages, from pre-reproductive juveniles to old and fecund adults. For reasons explained in the sections above, future work should also include the investigation of embryonic and larval life stages.

Comparative, multi-species, multi-life stage, experiments can be very powerful for testing evolutionary questions (as explained in the sections above). Many of these questions make up the contemporary core of biomineralisation research, for example: is there an evolutionarily conserved molluscan shell secretion mechanism? And how did the first ever molluscan shell field evolve? Have species living in extreme environments, for example the sub-zero polar regions, adapted their biomineralisation mechanisms? And if so, how? Will molluscs around the globe be able to adapt to climate change? To move these questions forward, more comprehensive genomes and phylogenies are certainly

required. Again it is important to stress that this is no simple task and is the life work of many scientists around the world. It can be tempting to trivialize the collection and description of data on gene function, but for non-model organisms such as molluscs, this work is essential in order to understand how biomineralisation works. Chapter 1 highlighted how the molecular resources for studying molluscs are rapidly increasing, and the production of comprehensive genomes for more mollusc species, in combination with emerging molecular technologies, is likely to underpin the progress of questions stated above. For example, it would be very exciting to apply emerging technologies, such as gene knock-out and genome editing with CRISPR (on the candidates highlighted in this thesis for example, including those in the regulatory gene network without annotation), to the question of how molluscs build their shells. Given very few groups in the world have either of these methods convincingly working on molluscs yet, their development could represent an opportunity for significant progress in the field of biomineralisation as they would help us to understand the function of genes and proteins which, as explained above, is urgently needed. Mantle and haemoctye cell cultures could also aid the advance of these technologies especially for species where collection, long-term husbandry, and variation are significant challenges.

Continuing on the theme of emerging technologies progressing the understanding of how molluscs build their shells, there are many lessons the field of molluscan biomineralisation can learn from medical literature and the use of model organisms. The investigation of how healthy and pathological bones are formed is a huge field of science. Recently, solid-state nuclear magnetic resonance (NMR) spectroscopy has been used to identify the atomic finger print of bone extracellular matrices, in animal tissues and in *in vitro* tissues. NMR showed that sugars, and specifically the glycosylation of collagen fibrils by PolyADPribose, are structurally important in bone; a finding that fundamentally changed the traditional bone biomineralisation model that will have major clinical implications (Chow et al. 2014, Duer 2015). NMR atomic finger printing was a key technique which lead to a paradigm shift in bone research, its application to molluscan shells will surely resolve more components of the shell matrix and lead to a better understanding of how molluscs build their shells.

Another consideration for future research is which species should be studied? Is *L. elliptica* a good model for biomineralisation questions? Or even, could a single mollusc species help us to answer all the questions we want to ask about biomineralisation? In general I propose that future research should continue to study the less economically important species, that are fascinating and useful in their own right, as well as the economically important and more tractable species. *L. elliptica* for example, is probably the best mollusc species to use for shell damage-repair experiments owing to its superior capacity to heal damage (due to its evolutionary history with ice-berg scouring). Building on the theme of ageing and longevity presented in Chapter 5, a future project could involve a range of species, both long-lived living at high latitudes (examples = Antarctic clam, *Laternula elliptica*; ocean quahog, *Arctica islandica*) and long-lived living at low latitudes (examples = giant clam *Tridacna gigas*, geodruck clam *Panopea abrupta*), and also short-lived for each geographic location – perhaps short-lived economically important animals. For each species a comprehensive genome, as well as adult mantle transcriptome could be developed and a shell-damage repair experiment could be employed (using the same methodology throughout in order to compare like-for-like), in addition, the phylogenetic relationships between the species would have to be constrained. With these data researchers could ask questions such as: Does biomineralisation, and shell damage-repair, play a role in mollusc longevity? Does reduced energy expenditure for growth, and greater utilisation for shell and tissue regeneration lead to increased longevity in high latitude bivalves? Does temperature fundamentally influence biomineralisation and longevity, or are organisms more constrained by their evolutionary history? Is phenotypic plasticity in the biomineralisation pathway an adaptive trait?

Many of the questions outlined in this chapter represent the genuine future of molluscan biomineralisation research, while others simply represent my own personal musings and curiosity. Having spent the last 3.5 years totally obsessed with clam shells, one thing is certain: they're much more complicated than I could have ever anticipated. Every question I adressed in this work inspired the formation of deeper and more detailed questions for further research. These are questions that myself, and many of my colleuges, will continue to persue for years to come.

Appendix A

Primer tables

Appendix A contains information on all the primers used in this thesis.

Table A-1. Primer and transcript information for *Laternula elliptica* candidate genes used in this thesis. 18s = housekeeping gene for normalisation. M13 primers used to amplify cloned plasmid insert for riboprobe synthesis

Transcript information			Primers for Semi-qPCR				Primers for *in situ* localisation			
Contig I.D	**Annotation** (Blastx)	**Present in nacre shell layer**	**Forward**	**Reverse**	**Product Size**	**Annealing Temp. (°C)**	**Forward**	**Reverse**	**Product Size**	**Annealing Temp. (°C)**
N/A	*L. elliptica 18s*	N/A	GGCCGTTCTTAGTTGGTGGA	TCATTAACTGGGGCGATCGG	445	60	N/A	N/A	N/A	N/A
N/A	M13 plasmid region	N/A	N/A	N/A	N/A	N/A	CGCCAGGGTTTTCCCAGTCACGAC	TCACACAGGAAACAGCTATGAC	Insert Specific	58
Contig01785	Mytilin	Yes	CGCCGATATGGATTTACGTG	GGCGAGACAGAATTCGGATA	413	60	TTATGCCGCCAATGTTCCCT	AAGGCGAGACAGAATTCGGA	871	60

Contig 01311	Chitin-binding domain	Yes	CACTGTGCTTCGCTGAACAT	GATCTACGCCCTCGTCAGAG	414	58	AAGCCACTTCCCTCATTCGG	ACTCTGTTGGGTATGCAGGC	1096	60
Contig 02037	Zn Metallo-endopeptidase	Yes	GAAAACGCCCGACTTGAATA	TCGGTGTGACATCGTTTGAT	496	58	TGCATGACGGAGGAAAAGCT	TGTTACGCGTGACAGAGGAG	912	60
Contig 00332	Pif	Yes	GGACAACACATTGGGGTAGG	GGATGTATCGGGGAAAAGGT	350	58	CAATCCGTCCACTTCTGCCT	CCACTGAACTGTTCGACCGA	1002	60
Contig 01663	Tyrosinase A	Yes	CCATCCGCTATCTGTGGTCC	TCTTTTGCGACCTCAGACCC	399	60	GGTAACCAGGCATGACGGAA	TACCGCGCCTATCAGAACAC	892	60
Contig 01359	Tyrosinase B	No	CGGCCTCATCGTGATAATCT	GGGAAGATTTTCGAATGCAA	479	58	TGTCACCCAATGTCCTGTCG	ATGACTTCCTGGCCAGCTTC	930	60
Contig 01043	Unknown	No	GGGTCAGCTGGTATCCTTGA	AGCGCTTGCAAAATTGTCTT	456	60	CCAAGCAGTCCATCGTCCTT	GGTGAATACGGACCCAGGAC	1097	60

Contig 00041	Chitin binding Peritrophin-A domain	No	TTTCGCCGCTAGAAGTGTGT	GGGTTCGTCAACATCCAGGT	407	60	N/A	N/A	N/A	N/A
Contig 18937102	Chitinase	No	CGTTGGTGTAGGTTCCATTC	GGCATGACCGGTATTAACAG	232	56	N/A	N/A	N/A	N/A
Contig 1586	Tyrosine	No	AGTTTTGTGCTGCCTTCAAC	CACGTGTTTGTGCTCTTCAC	210	57	N/A	N/A	N/A	N/A
Contig 2930	Unknown	No	TTGAAGTGTTACGCCTCCTC	AGCACTGCCTGGTGATAGAG	175	56	N/A	N/A	N/A	N/A
Contig 3459	Unknown	No	AAAATCAACGAGGAGGAGGA	GGGAATCCAGGTCTCTCTGT	150	57	N/A	N/A	N/A	N/A

Table A-2. Primer and transcript information for *Mya truncata* candidate genes used for qPCR in this thesis. 18s = housekeeping gene for normalisation. "Top 50" = The top 50 most highly expressed annotated (blastx, cut-off <1e^{-10}) contigs in the *Mya truncata* mantle transcriptome

Transcript information					Primers for qPCR			
Contig I.D.	**Annotation (Blastx)**	**Present in shell proteome**	**Present in mantle proteome**	**Present in "Top 50"**	**Forward**	**Reverse**	**Amplicon size**	**Annealing temp. (°C)**
M. truncata 18s	*M. truncata* 18s	N/A			GCTCGTAGTTGGATCTCGGG	ATCAAGAGCACCAAGGGACG	102	62
Contig 16470	Calponin	Yes	Yes	Yes	CGTACCCAGTCATACCCTTCT	GGCAAAGATATCAAAGCCGATG	106	64
Contig 629	Pif	Yes	Yes	No	CAGTCAGTGTCTGCCAGGTA	ACTACATCCACCACAGAGCC	107	64
Contig 1412	Chitin-binding domain	Yes	Yes	No	TTTACTCCCGATGCCAGTGT	CTTCGTACCTCCGCAATTGG	222	64
Contig 178	Complement control protein domain	Yes	No	No	CTTGCGATCCTGTTCCGAAG	TTGCAGGGtTTACACGTGTG	187	64
Contig 902	Tyrosinase A	Yes	No	No	CACCCTAATGCGTCAATGGG	GACATGAAGGTACCGGGTCA	124	64

Contig 0211	Tyrosinase B	No	No	Yes	CCCGGGCCTTCTAAATGTGT	ACACAACCTTTGTTAACCGGC	103	64
Contig 395	Cartilage matrix-like protein/ Von Willebrand factor type A domain	No	Yes	No	CCTCGTTCTTTGCCTCATCG	CAGGAATGTTAAGCTCGGCC	236	64

Appendix B

Supplementary files

Appendix B contains information on all the supplementary files for this thesis. Many of the chapters in this thesis includes the analysis of large, next-generation, sequence data sets, which produces many long lists of genes. These gene lists, and associated files, are an integral part of analysis and are included on the attached CD. Each file is described below.

Chapter 1: There are no supplementary files associated with this chapter.

Chapter 2: There are no supplementary files associated with this chapter.

Chapter 3:

Chapter3_Supplementary_table1.docx = *De novo* sequencing of the nacre tryptic peptides that do not match with *Laternulla elliptica* transcriptome according to Mascor search using Peaks software.

Chapter 4:

Chapter4_Supplementary_table1.xlsx = The putative *Laternula elliptica* gene annotations for the top 50 (annotated) differentially expressed genes, between control and damaged animals, at 1 week.

Chapter4_Supplementary_table2.xlsx = The putative *Laternula elliptica* gene annotations for the top 50 (annotated) differentially expressed genes, between control and damaged animals, at 1 month.

Chapter4_Supplementary_table3.xlsx = The putative *Laternula elliptica* gene annotations for the top 50 (annotated) differentially expressed genes, between control and damaged animals, at 2 months.

Chapter4_Supplementary_table4.xlsx = The putative *Laternula elliptica* gene annotations for selected “classic” biomineralisation genes.

Chapter 5:

Chapter5_Supplementary_file1_juveniles.R = The edgeR script used for differential expression analysis on the juvenile RNA-Seq data.

Chapter5_Supplementary_file2_adolescents.R = The edgeR script used for differential expression analysis on the adolescent RNA-Seq data.

Chapter5_Supplementary_file3_adults.R = The edgeR script used for differential expression analysis on the adult RNA-Seq data.

Chapter5_Supplementary_table1.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the juvenile age category, at 2 months. N.b. There were no annotated differentially expressed genes at 1 week or 1 month,

Chapter5_Supplementary_table2.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adolescent age category, at 1 month.

Chapter5_Supplementary_table3.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adolescent age category, at 2 months.

Chapter5_Supplementary_table4.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adolescent age category, at 3 months.

Chapter5_Supplementary_table5.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adolescent age category, at 4 months.

Chapter5_Supplementary_table6.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adult age category, at 5 days.

Chapter5_Supplementary_table7.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adult age category, at 1 week.

Chapter5_Supplementary_table8.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adult age category, at 1 month.

Chapter5_Supplementary_table9.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adult age category, at 2 months.

Chapter5_Supplementary_table10.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adult age category, at 3 months.

Chapter5_Supplementary_table11.xlsx = The putative *Laternula elliptica* gene annotations for the differentially expressed genes, between control and damaged animals, in the adult age category, at 4 months.

Chapter5_Supplementary_table12.xlsx = The shared, annotated, differentially expressed, genes between three highlight experimental time points, adolescents at 2 months, adults at 3 months and adults at 4 months.

Chapter5_Supplementary_table13.xlsx = The unique, annotated, differentially expressed, genes to adolescents at 2 months compared to, adults at 3 months and adults at 4 months.

Chapter5_Supplementary_table14.xlsx = The unique, annotated, differentially expressed, genes to adults at 3 months compared to, adolescents at 2months and adults at 4 months.

Chapter5_Supplementary_table15.xlsx = The unique, annotated, differentially expressed, genes to adults at 4 months compared to, adolescents at 2months and adults at 3 months.

Chapter5_Supplementary_table16.xlsx = The annotated time-dependant damage-response genes from the juvenile time category.

Chapter5_Supplementary_table17.xlsx = The annotated time-dependant damage-response genes from the adolescent time category.

Chapter5_Supplementary_table18.xlsx = The annotated time-dependant damage-response genes from the adult time category.

Chapter5_Supplementary_table19.xlsx = The shared, annotated, time-dependant damage-response genes between the juvenile, adolescent and adult age categories.

Chapter5_Supplementary_table20.xlsx = The unique, annotated, time-dependant damage-response genes to juveniles compared to, adolescents and adults.

Chapter5_Supplementary_table21.xlsx = The unique, annotated, time-dependant damage-response genes to adolescents compared to, juveniles and adults.

Chapter5_Supplementary_table22.xlsx = The unique, annotated, time-dependant damage-response genes to adults compared to, adolescents and juveniles.

Chapter 6:

Chapter6_Supplementary_table1.doc = Summary information of an amino acid alignment of all *Mya truncata* contigs annotated as Pif, to a reference full-length Pif amino acid sequence.

Chapter6_Supplementary_table2.xlsx = Output table of the tBLASTx results using *Mya truncata* transcriptome as subject and *Laternula elliptica* transcriptome as query.

Chapter6_Supplementary_table3.xlsx = Putative annotations for the top 50 most similar contigs between the *Mya truncata* and *Laternula elliptica* mantle transcriptomes.

Chapter6_Supplementary_Figure1.pdf = Neighbor-Joining (NJ) phylogenetic tree of the shell-building molluscan Tyrosinase proteins obtained by MEGA5.2.2 under the JTT + G substitution model. The tree is rooted using the midpoint-rooted option. Statistical support for each node is indicated as percentage of 1000 replicates. Nomenclature used as per Aguilera et al. (2014).

Chapter6_Supplementary_Figure2.pdf = Maximum Likelihood (ML) phylogenetic tree of shell-building molluscan Tyrosinase proteins obtained by RAxMLGUI v1.3 under the PROTGAMMAWAG substitution model (final ML optimization likelihood: −7786.291168). The tree is rooted using the midpoint-rooted option. Statistical support for each node is indicated as percentage of 1000 replicates. Nomenclature used as per Aguilera et al. (2014).

Chapter6_Supplementary_Figure3.pdf = Bayesian Inference (BI) phylogenetic tree of the molluscan Tyrosinase proteins obtained by MrBayes v3.2 under the WAG + G substitution model. The tree is rooted using the midpoint-rooted option. Statistical support for each node is indicated as posterior probabilities after 1,500,000 generations. Nomenclature used as per Aguilera et al. (2014).

References

Abouzaglou, J., Benistant, C., Gimona, M., Roustan, C., Kassab, R. and Fattoum, A. (2004) 'Tyrosine phosphorylation of calponins - Inhibition of the interaction with F-actin', *European Journal of Biochemistry*, 271(13), 2615-2623.

Acharya, C., Yik, J. H. N., Kishore, A., Dinh, V. V., Di Cesare, P. E. and Haudenschild, D. R. (2014) 'Cartilage oligomeric matrix protein and its binding partners in the cartilage extracellular matrix: Interaction, regulation and role in chondrogenesis', *Matrix Biology*, 37, 102-111.

Addadi, L. and Weiner, S. (2014) 'Biomineralization: mineral formation by organisms', *Physica Scripta*, 89(9), 13.

Aguilera, F., McDougall, C. and Degnan, B. (2013) 'Origin, evolution and classification of type-3 copper proteins: lineage-specific gene expansions and losses across the Metazoa', *BMC Evolutionary Biology*, 13(96), 1471-2148.

Aguilera, F., McDougall, C. and Degnan, B. M. (2014) 'Evolution of the tyrosinase gene family in bivalve molluscs: Independent expansion of the mantle gene repertoire', *Acta Biomaterialia*, 10(9), 3855-65.

Aguilera, F., McDougall, C. and Degnan, B. M. (2017) 'Co-option and de novo gene evolution underlie molluscan shell diversity', *Molecular Biology and Evolution*, 34(4), 779-792.

Ahn, I.-Y. (1997) 'Feeding ecology of the Antarctic lamellibranch *Laternula elliptica* (Laternulidae) in Marian Cove and vicinity, King George Island, during one austral summer', *Antarctic communities: species, structure, and survival*, 142.

Ahn, I. Y., Surh, J., Park, Y. G., Kwon, H., Choi, K. S., Kang, S. H., Choi, H. J., Kim, K. W. and Chung, H. (2003) 'Growth and seasonal energetics of the Antarctic bivalve *Laternula elliptica* from King George Island, Antarctica', *Marine Ecology Progress Series*, 257, 99-110.

Altschul, S. F., Gish, W., Miller, W., Myers, E. W. and Lipman, D. J. (1990) 'Basic local alignment search tool', *Journal of Molecular Biology*, 215(3), 403-10.

Amaro, T., Duineveld, G. and Tyler, P. (2005) 'Does *Mya truncata* reproduce at its southern distribution limit? Preliminary information', *Journal of Shellfish Research*, 24(1), 25-28.

An, J., Shen, X. F., Ma, Q. B., Yang, C. Y., Liu, S. M. and Chen, Y. (2014) 'Transcriptome profiling to discover putative genes associated with paraquat resistance in goosegrass (*Eleusine indica L.*)', *Plos One*, 9(6), e99940.

Anju, A., Jeswin, J., Thomas, P. C. and Vijayan, K. K. (2013) 'Molecular cloning, characterization and expression analysis of F-type lectin from pearl oyster *Pinctada fucata'*, *Fish & Shellfish Immunology*, 35(1), 170-174.

Arany, S., Koyota, S. and Sugiyama, T. (2009) 'Nerve growth factor promotes differentiation of odontoblast-like cells', *Journal of Cellular Biochemistry*, 106(4), 539-545.

Arivalagan, J., Marie, B., Sleight, V. A., Clark, M. S., Berland, S. and Marie, A. (2016) 'Shell matrix proteins of the clam, *Mya truncata*: Roles beyond shell formation through proteomic study', *Marine Genomics*, 27, 69-74.

Arivalagan, J., Yarra, T., Marie, B., Sleight, V. A., Duvernois-Berthet, E., Clark, M. S., Marie, A. and Berland, S. (2017) 'Insights from the shell proteome: Biomineralization to adaptation', *Molecular Bioliogy and Evolution*, 34(1), 66-77.

Arntz, W. E., Brey, T. and Gallardo, V. A. (1994) 'Antarctic Zoobenthos', *Oceanography and Marine Biology, Vol 32: an Annual Review*, 32, 241-304.

Arts, G. J., Langemeijer, E., Tissingh, R., Ma, L. B., Pavliska, H., Dokic, K., Dooijes, R., Misic, E., Clasen, R., Michiels, F., van der Schueren, J., Lambrecht, M., Herman, S., Brys, R., Thys, K., Hoffmann, M., Tomme, P. and van Es, H. (2003)

'Adenoviral vectors expressing siRNAs for discovery and validation of gene function', *Genome Research,* 13(10), 2325-2332.

Asakura, T., Hamada, M., Nakazawa, Y., Ha, S. W. and Knight, D. P. (2006) 'Conformational study of silk-like peptides containing the calcium-binding sequence from calbindin D-9k using C-13 CP/MAS NMR spectroscopy', *Biomacromolecules,* 7(2), 627-634.

Badariotti, F., Lelong, C., Dubos, M. P. and Favrel, P. (2007a) 'Characterization of chitinase-like proteins (Cg-Clp1 and Cg-Clp2) involved in immune defence of the mollusc *Crassostrea gigas*', *The FEBS journal,* 274(14), 3646-3654.

Badariotti, F., Thuau, R., Lelong, C., Dubos, M. P. and Favrel, P. (2007b) 'Characterization of an atypical family 18 chitinase from the oyster *Crassostrea gigas*: Evidence for a role in early development and immunity', *Developmental and Comparative Immunology,* 31(6), 559-570.

Bahn, S. Y., Jo, B. H., Hwang, B. H., Choi, Y. S. and Cha, H. J. (2015) 'Role of Pif97 in nacre biomineralization: *in vitro* characterization of recombinant Pif97 as a framework protein for the association of organic-inorganic layers in nacre', *Crystal Growth & Design,* 15(8), 3666-3673.

Barnes, D. K. A. and Souster, T. (2011) 'Reduced survival of Antarctic benthos linked to climate-induced iceberg scouring', *Nature Climate Change,* 1(7), 365-368.

Barolo, S. and Posakony, J. W. (2002) 'Three habits of highly effective signaling pathways: principles of transcriptional control by developmental cell signaling', *Genes & Development,* 16(10), 1167-1181.

Bayne, B. L. (1976) *Marine mussels : their ecology and physiology* (Vol. 10). Cambridge University Press.

Belcher, A. M., Wu, X. H., Christensen, R. J., Hansma, P. K., Stucky, G. D. and Morse, D. E. (1996) 'Control of crystal phase switching and orientation by soluble mollusc-shell proteins', *Nature,* 381(6577), 56-58.

Benson, S. C., Benson, N. C. and Wilt, F. (1986) 'The organic matrix of the skeletal spicule of sea-urchin embryos', *Journal of Cell Biology,* 102(5), 1878-1886.

Berland, S., Ma, Y. F., Marie, A., Andrieu, J. P., Bedouet, L. and Feng, Q. L. (2013) 'Proteomic and profile analysis of the proteins laced with aragonite and vaterite in the freshwater mussel *Hyriopsis cumingii* shell biominerals', *Protein and Peptide Letters,* 20(10), 1170-1180.

Bevelander, G. (1951) 'Calcification in molluscs - the localization of ca-95 and p-32 in the mantle and regenerating shell', *Biological Bulletin,* 101(2), 197-198.

Bevelander, G. (1952) 'Calcification in molluscs .3. Intake and deposition of ca-45 and p-32 in relation to shell formation', *Biological Bulletin,* 102(1), 9-15.

Bevelander, G. and Benzer, P. (1948) 'Calcification in marine molluscs', *Biological Bulletin,* 94(3), 176-183.

Bieler, R., Mikkelsen, P. M., Collins, T. M., Glover, E. A., Gonzalez, V. L., Graf, D. L., Harper, E. M., Healy, J., Kawauchi, G. Y., Sharma, P. P., Staubach, S., Strong, E. E., Taylor, J. D., Temkin, I., Zardus, J. D., Clark, S., Guzman, A., McIntyre, E., Sharp, P. and Giribet, G. (2014) 'Investigating the Bivalve Tree of Life - an exemplar-based approach combining molecular and novel morphological characters', *Invertebrate Systematics,* 28(1), 32-115.

Blank, S., Arnoldi, M., Khoshnavaz, S., Treccani, L., Kuntz, M., Mann, K., Grathwohl, G. and Fritz, M. (2003) 'The nacre protein perlucin nucleates growth of calcium carbonate crystals', *Journal of Microscopy-Oxford,* 212, 280-291.

Bodin, N., Burgeot, T., Stanisiere, J. Y., Bocquene, G., Menard, D., Minier, C., Boutet, I., Amat, A., Cherel, Y. and Budzinski, H. (2004) 'Seasonal variations of a battery of biomarkers and physiological indices for the mussel *Mytilus galloprovincialis* transplanted into the northwest Mediterranean Sea', *Comparative Biochemistry and Physiology C-Toxicology & Pharmacology*, 138(4), 411-427.

Bonneh-Barkay, D., Zagadailov, P., Zou, H. C., Niyonkuru, C., Figley, M., Starkey, A., Wang, G. J., Bissel, S. J., Wiley, C. A. and Wagner, A. K. (2010) 'YKL-40 Expression in Traumatic Brain Injury: An Initial Analysis', *Journal of Neurotrauma*, 27(7), 1215-1223.

Boot, R. G., Renkema, G. H., Strijland, A., Vanzonneveld, A. J. and Aerts, J. (1995) 'Cloning of a cDNA-encoding chitotriosidase, a human chitinase produced by macrophages', *Journal of Biological Chemistry*, 270(44), 26252-26256.

Bowen, C. E. and Tang, H. (1996) 'Conchiolin-protein in aragonite shells of mollusks', *Comparative Biochemistry and Physiology A-Physiology*, 115(4), 269-275.

Brower, D. L., Brabant, M. C. and Bunch, T. A. (1995) 'Role of the PS integrins in Drosophila development', *Immunology and Cell Biology*, 73(6), 558-564.

Camus, L., Birkely, S. R., Jones, M. B., Borseth, J. F., Grosvik, B. E., Gulliksen, B., Lonne, O. J., Regoli, F. and Depledge, M. H. (2003) 'Biomarker responses and PAH uptake in *Mya truncata* following exposure to oil-contaminated sediment in an Arctic fjord (Svalbard)', *Scieince of the Total Environment*, 308(1-3), 221-34.

Camus, L., Gulliksen, B., Depledge, M. H. and Jones, M. B. (2005) 'Polar bivalves are characterized by high antioxidant defences', *Polar Research*, 24(1-2), 111-118.

Camus, L., Richardsen, S. R., Borseth, J. F., Grosvik, B. E., Gulliksen, B., Jones, M. B., Lonne, O. J., Regoli, F. and Depledge, M. H. (2002) 'Biomarkers in the soft shell Arctic clam *Mya truncata*: Seasonal variability and impact of PAH', *Marine Environmental Research*, 54(3-5), 830-830.

Cariolou, M. A. and Morse, D. E. (1988) 'Purification and characterization of calcium-binding conchiolin shell peptides from the mollusk, *Haliotis rufescens*, as a function of development', *Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology,* 157(6), 717-729.

Carter, J. G. and Clark, G. R. (1985) 'Classification and phylogenetic significance of mollusc shell microstructure' in Broadhead, T. W., ed. *Classification and Phylogenetic Significance of Mollusk Shell Microstructures*, Knoxville, Tennessee: Tennessee Press, 67.

Castresana, J. (2000) 'Selection of conserved blocks from multiple alignments for their use in phylogenetic analysis', *Molecular Biology and Evolution,* 17(4), 540-552.

Chen, L., Wu, W., Dentchev, T., Zeng, Y., Wang, J. H., Tsui, I., Tobias, J. W., Bennett, J., Baldwin, D. and Dunaief, J. L. (2004) 'Light damage induced changes in mouse retinal gene expression', *Experimental Eye Research,* 79(2), 239-247.

Cho, S.-M. and Jeong, W.-G. (2011) 'Prismatic shell repairs by hemoctyes in the extrapallial fluid of the Pacific Oyster, *Crassostrea gigas'*, *Korean Journal of Malacology,* 27(3), 223-228.

Chow, W. Y., Rajan, R., Muller, K. H., Reid, D. G., Skepper, J. N., Wong, W. C., Brooks, R. A., Green, M., Bihan, D., Farndale, R. W., Slatter, D. A., Shanahan, C. M. and Duer, M. J. (2014) 'NMR spectroscopy of native and in vitro tissues implicates polyADP ribose in biomineralization', *Science*, 344(6185), 742-746.

Clark, M., Fraser, K. P. and Peck, L. (2008) 'Antarctic marine molluscs do have an HSP70 heat shock response', *Cell Stress and Chaperones*, 13(1), 39-49.

Clark, M. S., Husmann, G., Thorne, M. A. S., Burns, G., Truebano, M., Peck, L. S., Abele, D. and Philipp, E. E. R. (2013a) 'Hypoxia impacts large adults first: consequences in a warming world', *Global Change Biology*, 19(7), 2251-2263.

Clark, M. S. and Peck, L. S. (2009) 'HSP70 heat shock proteins and environmental stress in Antarctic marine organisms: A mini-review', *Marine Genomics*, 2(1), 11-18.

Clark, M. S., Power, D. M. and Sundell, K. (2016) 'Cells to shells: The genomics of mollusc exoskeletons', *Marine Genomics*, 27, 1-2.

Clark, M. S., Thorne, M. A. S., Amaral, A., Vieira, F., Batista, F. M., Reis, J. and Power, D. M. (2013b) 'Identification of molecular and physiological responses to chronic environmental challenge in an invasive species: the Pacific oyster, *Crassostrea gigas*', *Ecology and Evolution*, 3(10), 3283-3297.

Clark, M. S., Thorne, M. A. S., Vieira, F. A., Cardoso, J. C. R., Power, D. M. and Peck, L. S. (2010) 'Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyrosequencing', *BMB Genomics*, 11.

Cook, A. J., Fox, A. J., Vaughan, D. G. and Ferrigno, J. G. (2005) 'Retreating glacier fronts on the Antarctic Peninsula over the past half-century', *Science*, 308(5721), 541-544.

Corbin, J. S. (2007) 'Marine aquaculture: Today's necessity for tomorrow's seafood', *Marine Technology Society Journal*, 41(3), 16-23.

Craft, J. A., Gilbert, J. A., Temperton, B., Dempsey, K. E., Ashelford, K., Tiwari, B., Hutchinson, T. H. and Chipman, J. K. (2010) 'Pyrosequencing of *Mytilus galloprovincialis* cDNAs: Tissue-Specific Expression Patterns', *Plos One*, 5(1), e8875.

Day, E. G., Branch, G. M. and Viljoen, C. (2000) 'How costly is molluscan shell erosion? A comparison of two patellid limpets with contrasting shell structures', *Journal of Experimental Marine Biology and Ecology*, 243(2), 185-208.

de Paula, S. M. and Silveira, M. (2009) 'Studies on molluscan shells: Contributions from microscopic and analytical methods', *Micron*, 40(7), 669-690.

Dineshram, R., Thiyagarajan, V., Lane, A., Yu, Z., Shu, X. and Leung, P. T. Y. (2013) 'Elevated CO2 alters larval proteome and its phosphorylation status in the commercial oyster, *Crassostrea hongkongensis*', *Marine Biology*, 160(8), 2189-2205.

Dineshram, R., Wong, K. K. W., Xiao, S., Yu, Z. N., Qian, P. Y. and Thiyagarajan, V. (2012) 'Analysis of Pacific oyster larval proteome and its response to high-CO2', *Marine Pollution Bulletin*, 64(10), 2160-2167.

Dodenhof, T., Dietz, F., Franken, S., Grunwald, I. and Kelm, S. (2014) 'Splice variants of perlucin from *Haliotis laevigata* modulate the vrystallisation of CaCO3', *Plos One*, 9(5), e97126.

Duer, M. J. (2015) 'The contribution of solid-state NMR spectroscopy to understanding biomineralization: Atomic and molecular structure of bone', *Journal of Magnetic Resonance*, 253, 98-110.

Ehrlich, H. (2010) 'Biomineralization-Demineralization-Remineralization Phenomena in Nature', *Biological Materials of Marine Origin: Invertbrates*, 1, 59-101.

Esposito, R., D'Aniello, S., Squarzoni, P., Pezzotti, M. R., Ristoratore, F. and Spagnuolo, A. (2012) 'New insights into the evolution of metazoan tyrosinase gene family', *PLoS One*, 7(4), e35731.

Evans, J. S. (2012) 'Aragonite-associated biomineralization proteins are disordered and contain interactive motifs', *Bioinformatics*, 28(24), 3182-5.

Fan, W., Li, C., Li, S., Feng, Q., Xie, L. and Zhang, R. (2007) 'Cloning, characterization, and expression patterns of three sarco/endoplasmic reticulum Ca2+-ATPase

isoforms from pearl oyster (*Pinctada fucata*)', *Acta Biochimica et Biophysica Sinica,* 39(9), 722-730.

Fang, D., Pan, C., Lin, H. J., Lin, Y., Zhang, G. Y., Wang, H. Z., He, M. X., Xie, L. P. and Zhang, R. Q. (2012) 'Novel basic protein, PfN23, functions as key macromolecule during nacre formation', *Journal of Biological Chemistry,* 287(19), 15776-15785.

Fang, D., Xu, G., Hu, Y., Pan, C., Xie, L. and Zhang, R. (2011) 'Identification of genes directly involved in shell formation and their functions in pearl oyster, *Pinctada fucata*', *Plos One,* 6(7), e21860.

Fang, Z., Yan, Z., Li, S., Wang, Q., Cao, W., Xu, G., Xiong, X., Xie, L. and Zhang, R. (2008) 'Localization of calmodulin and calmodulin-like protein and their functions in biomineralization in *P. fucata*', *Progress in Natural Science-Materials International,* 18(4), 405-412.

Fernandez, M. S., Moya, A., Lopez, L. and Arias, J. L. (2001) 'Secretion pattern, ultrastructural localization and function of extracellular matrix molecules involved in eggshell formation', *Matrix Biology,* 19(8), 793-803.

Ferreira, V. P., Pangburn, M. K. and Cortes, C. (2010) 'Complement control protein factor H: The good, the bad, and the inadequate', *Molecular Immunology,* 47(13), 2187-2197.

Fleury, C., Marin, F., Marie, B., Luquet, G., Thomas, J., Josse, C., Serpentini, A. and Lebel, J. M. (2008) 'Shell repair process in the green ormer *Haliotis tuberculata*: A histological and microstructural study', *Tissue & Cell,* 40(3), 207-218.

Force, A., Lynch, M., Pickett, F. B., Amores, A., Yan, Y. L. and Postlethwait, J. (1999) 'Preservation of duplicate genes by complementary, degenerative mutations', *Genetics,* 151(4), 1531-1545.

Ford, S. E. and Paillard, C. (2007) 'Repeated sampling of individual bivalve mollusks I: Intraindividual variability and consequences for haemolymph constituents of the Manila clam, *Ruditapes philippinarum*', *Fish & Shellfish Immunology*, 23(2), 280-291.

Franceschini, A., Szklarczyk, D., Frankild, S., Kuhn, M., Simonovic, M., Roth, A., Lin, J. Y., Minguez, P., Bork, P., von Mering, C. and Jensen, L. J. (2013) 'STRING v9.1: protein-protein interaction networks, with increased coverage and integration', *Nucleic Acids Research*, 41(D1), D808-D815.

Freer, A., Bridgett, S., Jiang, J. H. and Cusack, M. (2014) 'Biomineral proteins from *Mytilus edulis* mantle tissue transcriptome', *Marine Biotechnology*, 16(1), 34-45.

Frémy, E. (1855) 'Recherches chimiques sur les os', *Annales de Chimie et de Physique*, 43, 47-107.

Galtsoff, P. S. (1964) 'The american oyster *Crassostrea virginica* Gmelin', *Fisheries. Bulletin.*, 64, 421-425.

Gao, C., Cai, X., Zhang, Y., Su, B., Song, H. and Li, C. (2017) 'Characterization and expression analysis of chitinase genes (CHIT1, CHIT2 and CHIT3) in turbot (*Scophthalmus maximus* L.) following bacterial challenge', *Fish & shellfish immunology*, 64, 357-366.

Gao, P., Liao, Z., Wang, X. X., Bao, L. F., Fan, M. H., Li, X. M., Wu, C. W. and Xia, S. W. (2015) 'Layer-by-layer proteomic analysis of *Mytilus galloprovincialis* shell', *Plos One*, 10(7), e0133913.

Gardner, L. D., Mills, D., Wiegand, A., Leavesley, D. and Elizur, A. (2011) 'Spatial analysis of biomineralization associated gene expression from the mantle organ of the pearl oyster *Pinctada maxima*', *BMC Genomics*, 12(1), 455.

Garfield, D. A., Runcie, D. E., Babbitt, C. C., Haygood, R., Nielsen, W. J. and Wray, G. A. (2013) 'The impact of gene expression variation on the robustness and evolvability of a developmental gene regulatory network', *Plos Biology*, 11(10), e1001696.

Gazeau, F., Parker, L. M., Comeau, S., Gattuso, J.-P., O'Connor, W. A., Martin, S., Poertner, H.-O. and Ross, P. M. (2013) 'Impacts of ocean acidification on marine shelled molluscs', *Marine Biology*, 160(8), 2207-2245.

Gillis, T. E. and Ballantyne, J. S. (1999) 'Mitochondrial membrane composition of two arctic marine bivalve mollusks, *Serripes groenlandicus* and *Mya truncata'*, *Lipids*, 34(1), 53-57.

Gitter, A., Carmi, M., Barkai, N. and Bar-Joseph, Z. (2013) 'Linking the signaling cascades and dynamic regulatory networks controlling stress responses', *Genome Research*, 23(2), 365-376.

Golden, T. R. and Melov, S. (2004) 'Microarray analysis of gene expression with age in individual nematodes', *Aging Cell*, 3(3), 111-124.

Gong, N., Li, Q., Huang, J., Fang, Z., Zhang, G., Xie, L. and Zhang, R. (2008) 'Culture of outer epithelial cells from mantle tissue to study shell matrix protein secretion for biomineralization', *Cell and Tissue Research*, 333(3), 493-501.

Grabherr, M. G., Haas, B. J., Yassour, M., Levin, J. Z., Thompson, D. A., Amit, I., Adiconis, X., Fan, L., Raychowdhury, R., Zeng, Q. D., Chen, Z. H., Mauceli, E., Hacohen, N., Gnirke, A., Rhind, N., di Palma, F., Birren, B. W., Nusbaum, C., Lindblad-Toh, K., Friedman, N. and Regev, A. (2011) 'Full-length transcriptome assembly from RNA-Seq data without a reference genome', *Nature Biotechnology*, 29(7), 644-652.

Gregoire, C. (1967) 'Structure of organic matrices of mollusc shells', *Biological Reviews of the Cambridge Philosophical Society*, 42(4), 653-684.

Gyoneva, S. and Ransohoff, R. M. (2015) 'Inflammatory reaction after traumatic brain injury: therapeutic potential of targeting cell-cell communication by chemokines', *Trends in Pharmacological Sciences*, 36(7), 471-480.

Haas, B. J., Papanicolaou, A., Yassour, M., Grabherr, M., Blood, P. D., Bowden, J., Couger, M. B., Eccles, D., Li, B., Lieber, M., MacManes, M. D., Ott, M., Orvis, J., Pochet, N., Strozzi, F., Weeks, N., Westerman, R., William, T., Dewey, C. N., Henschel, R., LeDuc, R. D., Friedman, N. and Regev, A. (2013) 'De novo transcript sequence reconstruction from RNA-Seq: reference generation and analysis with Trinity', *Nature protocols*, 8(8).

Hara, M. R., Kovacs, J. J., Whalen, E. J., Rajagopal, S., Strachan, R. T., Grant, W., Towers, A. J., Williams, B., Lam, C. M., Xiao, K. H., Shenoy, S. K., Gregory, S. G., Ahn, S., Duckett, D. R. and Lefkowitz, R. J. (2011) 'A stress response pathway regulates DNA damage through beta(2)-adrenoreceptors and beta-arrestin-1', *Nature*, 477(7364), 349-353.

Harper, E. M. (1997) 'The molluscan periostracum: An important constraint in bivalve evolution', *Palaeontology*, 40, 71-97.

Harper, E. M. (2000) 'Are calcitic layers an effective adaptation against shell dissolution in the Bivalvia?', *Journal of Zoology*, 251, 179-186.

Harper, E. M., Clark, M. S., Hoffman, J. I., Philipp, E. E. R., Peck, L. S. and Morley, S. A. (2012) 'Iceberg scour and shell damage in the Antarctic bivalve *Laternula elliptica*', *Plos One*, 7(9), e46341.

Hartl, F. U. (1996) 'Molecular chaperones in cellular protein folding', *Nature*, 381(6583), 571-580.

Hauri, C., Friedrich, T. and Timmermann, A. (2016) 'Abrupt onset and prolongation of aragonite undersaturation events in the Southern Ocean', *Nature Climate Change,* 6(2), 172-176.

Herrero, J., Valencia, A. and Dopazo, J. (2001) 'A hierarchical unsupervised growing neural network for clustering gene expression patterns', *Bioinformatics,* 17(2), 126-136.

Howe, E. A., Sinha, R., Schlauch, D. and Quackenbush, J. (2011) 'RNA-Seq analysis in MeV', *Bioinformatics,* 27(22), 3209-3210.

Huang, A.-M., Geng, Y., Wang, K.-Y., Zeng, F., Liu, Q., Wang, Y., Sun, Y., Liu, X.-X. and Zhou, Y. (2013) 'Molecular cloning and expression analysis of heat shock protein 90 (Hsp90) of the mud crab, *Scylla Paramamosain'*, *Journal of Agricultural Science;* (5)7, 1.

Huning, A., Melzner, F., Thomsen, J., Gutowska, M. A., Kramer, L., Frickenhaus, S., Rosenstiel, P., Portner, H. O., Philipp, E. E. R. and Lucassen, M. (2013) 'Impacts of seawater acidification on mantle gene expression patterns of the Baltic Sea blue mussel: implications for shell formation and energy metabolism', *Marine Biology,* 160(8), 1845-1861.

Huning, A. K., Lange, S. M., Ramesh, K., Jacob, D. E., Jackson, D. J., Panknin, U., Gutowska, M. A., Philipp, E. E. R., Rosenstiel, P., Lucassen, M. and Melzner, F. (2016) 'A shell regeneration assay to identify biomineralization candidate genes in mytilid mussels', *Marine Genomics,* 27, 57-67.

Husmann, G., Abele, D., Rosenstiel, P., Clark, M. S., Kraemer, L. and Philipp, E. E. R. (2014) 'Age-dependent expression of stress and antimicrobial genes in the hemocytes and siphon tissue of the Antarctic bivalve, *Laternula elliptica*, exposed to injury and starvation', *Cell Stress & Chaperones,* 19(1), 15-32.

Husmann, G., Philipp, E. E. R., Rosenstiel, P., Vazquez, S. and Abele, D. (2011) 'Immune response of the Antarctic bivalve *Laternula elliptica* to physical stress and

microbial exposure', *Journal of Experimental Marine Biology and Ecology,* 398(1-2), 83-90.

Ivanina, A. V., Sokolova, I. M. and Sukhotin, A. A. (2008) 'Oxidative stress and expression of chaperones in aging mollusks', *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology,* 150(1), 53-61.

Jackson, A. P., Vincent, J. F. V. and Turner, R. M. (1988) 'The mechanical design of nacre', *Proceedings of the Royal Society Series B-Biological Sciences,* 234(1277), 415-+.

Jackson, D. J., Herlitze, I. and Hohagen, J. (2016) 'A whole mount in situ hybridization method for the gastropod mollusc *Lymnaea stagnalis*', *Jove-Journal of Visualized Experiments,* (109).

Jackson, D. J., McDougall, C., Green, K., Simpson, F., Woerheide, G. and Degnan, B. M. (2006) 'A rapidly evolving secretome builds and patterns a sea shell', *BMC Biology,* 4(1). 40.

Jackson, D. J., McDougall, C., Woodcroft, B., Moase, P., Rose, R. A., Kube, M., Reinhardt, R., Rokhsar, D. S., Montagnani, C., Joubert, C., Piquemal, D. and Degnan, B. M. (2010) 'Parallel evolution of nacre building gene sets in molluscs', *Molecular Biology and Evolution,* 27(3), 591-608.

Jackson, D. J., Worheide, G. and Degnan, B. M. (2007) 'Dynamic expression of ancient and novel molluscan shell genes during ecological transitions', *BMC Evolutionary Biology,* 7(1), 160.

Jeffroy, F., Brulle, F. and Paillard, C. (2013) 'Differential expression of genes involved in immunity and biomineralization during Brown Ring Disease development and shell repair in the Manila clam, *Ruditapes philippinarum', Journal of Invertebrate Pathology,* 113(2), 129-136.

Jensen, M. H., Morris, E. J., Gallant, C. M., Morgan, K. G., Weitz, D. A. and Moore, J. R. (2014) 'Mechanism of calponin stabilization of cross-linked actin networks', *Biophysical Journal*, 106(4), 793-800.

Jiao, Y., Wang, H., Du, X., Zhao, X., Wang, Q., Huang, R. and Deng, Y. (2012) 'Dermatopontin, a shell matrix protein gene from pearl oyster *Pinctada martensii*, participates in nacre formation', *Biochemical and Biophysical Research Communications*, 425(3), 679-683.

Jolly, C., Berland, S., Milet, C., Borzeix, S., Lopez, E. and Doumenc, D. (2004) 'Zonal localization of shell matrix proteins in mantle of Haliotis tuberculata (Mollusca, Gastropoda)', *Marine Biotechnology*, 6(6), 541-551.

Jones, D. P. and Go, Y. M. (2010) 'Redox compartmentalization and cellular stress', *Diabetes Obesity & Metabolism*, 12, 116-125.

Jonkers, H. A. (1999) 'Aligned growth positions in Pliocene *Laternula elliptica* (King & Broderip) (Bivalvia : Anomalodesmata : Laternulidae)', *Antarctic Science*, 11(4), 463-464.

Joubert, C., Piquemal, D., Marie, B., Manchon, L., Pierrat, F., Zanella-Cleon, I., Cochennec-Laureau, N., Gueguen, Y. and Montagnani, C. (2010) 'Transcriptome and proteome analysis of *Pinctada margaritifera* calcifying mantle and shell: focus on biomineralization', *BMC Genomics*, 11(1), 613.

Kadar, E. (2008) 'Haemocyte response associated with induction of shell regeneration in the deep-sea vent mussel *Bathymodiolus azoricus* (Bivalvia : Mytilidae)', *Journal of Experimental Marine Biology and Ecology*, 362(2), 71-78.

Karsenty, G., Kronenberg, H. M. and Settembre, C. (2009) 'Genetic control of bone formation', *Annual Review of Cell and Developmental Biology*, 25, 629-648.

Katoh, K., Kuma, K. i., Toh, H. and Miyata, T. (2005) 'MAFFT version 5: improvement in accuracy of multiple sequence alignment', *Nucleic Acids Research,* 33(2), 511-518.

Keeren, K., Friedrich, M., Gebuhr, I., Philipp, S., Sabat, R., Sterry, W., Brandt, C., Meisel, C., Grutz, G., Volk, H. D. and Sawitzki, B. (2009) 'Expression of tolerance associated gene-1, a mitochondrial protein inhibiting T cell activation, can be used to predict response to immune modulating therapies', *Journal of Immunology,* 183(6), 4077-4087.

Khodr, B. and Khalil, Z. (2001) 'Modulation of inflammation by reactive oxygen species: Implications for aging and tissue repair', *Free Radical Biology and Medicine,* 30(1), 1-8.

Klambauer, G., Unterthiner, T. and Hochreiter, S. (2013) 'DEXUS: identifying differential expression in RNA-Seq studies with unknown conditions', *Nucleic Acids Research,* 41(21).

Klose, R. J., Kallin, E. M. and Zhang, Y. (2006) 'JmjC-domain-containing proteins and histone demethylation', *Nature Reviews Genetics,* 7(9), 715-727.

Kong, Y., Jing, G., Yan, Z., Li, C., Gong, N., Zhu, F., Li, D., Zhang, Y., Zheng, G., Wang, H., Xie, L. and Zhang, R. (2009) 'Cloning and characterization of prisilkin-39, a novel matrix protein serving a dual role in the prismatic layer formation from the oyster *Pinctada fucata', Journal of Biological Chemistry,* 284(16), 10841-10854.

Kouchinsky, A. (2000) 'Shell microstructures in Early Cambrian molluscs', *Acta Palaeontologica Polonica,* 45(2), 119-150.

Krepinsky, J. C., Li, Y. X., Chang, Y. F., Liu, L. Q., Peng, F. D., Wu, D. C., Tang, D. M., Scholey, J. and Ingram, A. J. (2005) 'Akt mediates mechanical strain-induced collagen production by mesangial cells', *Journal of the American Society of Nephrology,* 16(6), 1661-1672.

Kwiek, N. C., Thacker, D. F., Datto, M. B., Megosh, H. B. and Haystead, T. A. J. (2006) 'PITK, a PP1 targeting subunit that modulates the phosphorylation of the transcriptional regulator hnRNP K', *Cellular Signalling,* 18(10), 1769-1778.

Lamberg, A., Helaakoski, T., Myllyharju, J., Peltonen, S., Notbohm, H., Pihlajaniemi, T. and Kivirikko, K. I. (1996) 'Characterization of human type III collagen expressed in a baculovirus system - Production of a protein with a stable triple helix requires coexpression with the two types of recombinant prolyl 4-hydroxylase subunit', *Journal of Biological Chemistry,* 271(20), 11988-11995.

Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. and Higgins, D. G. (2007) 'Clustal W and clustal X version 2.0', *Bioinformatics,* 23(21), 2947-2948.

Le Roy, N., Jackson, D. J., Marie, B., Ramos-Silva, P. and Marin, F. (2014) 'The evolution of metazoan alpha-carbonic anhydrases and their roles in calcium carbonate biomineralization', *Frontiers in Zoology,* 11(1), 75.

Leader, J. P. and O'Donnell, M. J. (2005) 'Transepithelial transport of fluorescent p-glycoprotein and MRP2 substrates by insect Malpighian tubules: confocal microscopic analysis of secreted fluid droplets', *Journal of Experimental Biology,* 208(23), 4363-4376.

Lenz, P. H., Roncalli, V., Hassett, R. P., Wu, L. S., Cieslak, M. C., Hartline, D. K. and Christie, A. E. (2014) 'De Novo Assembly of a Transcriptome for *Calanus finmarchicus* (Crustacea, Copepoda) - The Dominant Zooplankter of the North Atlantic Ocean', *Plos One,* 9(2), e88589.

Leonard, G. H., Bertness, M. D. and Yund, P. O. (1999) 'Crab predation, waterborne cues, and inducible defenses in the blue mussel, *Mytilus edulis*', *Ecology,* 80(1), 1-14.

Leupin, O., Piters, E., Halleux, C., Hu, S., Kramer, I., Morvan, F., Bouwmeester, T., Schirle, M., Bueno-Lozano, M., Fuentes, F. J. R., Itin, P. H., Boudin, E., de Freitas, F., Jennes, K., Brannetti, B., Charara, N., Ebersbach, H., Geisse, S., Lu, C. X., Bauer, A., Van Hul, W. and Kneissel, M. (2011) 'Bone Overgrowth-associated Mutations in the LRP4 Gene Impair Sclerostin Facilitator Function', *Journal of Biological Chemistry,* 286(22), 19489-19500.

Li, A. L., Li, H. Y., Jin, B. F., Ye, Q. N., Zhou, T., Yu, X. D., Pan, X., Man, J. H., He, K., Yu, M., Hu, M. R., Wang, J., Yang, S. C., Shen, B. F. and Zhang, X. M. (2004) 'A novel eIF5A complex functions as a regulator of p53 and p53-dependent apoptosis', *Journal of Biological Chemistry,* 279(47), 49251-49258.

Li, H., Ruan, J. and Durbin, R. (2008) 'Mapping short DNA sequencing reads and calling variants using mapping quality scores', *Genome Research,* 18(11), 1851-1858.

Li, S. G., Liu, Y. J., Liu, C., Huang, J. L., Zheng, G. L., Xie, L. P. and Zhang, R. Q. (2016) 'Hemocytes participate in calcium carbonate crystal formation, transportation and shell regeneration in the pearl oyster *Pinctada fucata'*, *Fish & Shellfish Immunology,* 51, 263-270.

Lin, J. Y., Ma, K. Y., Bai, Z. Y. and Li, J. L. (2013) 'Molecular cloning and characterization of perlucin from the freshwater pearl mussel, *Hyriopsis cumingii*', *Gene,* 526(2), 210-6.

Liu, S., Saloustros, E., Mertz, E. L., Tsang, K., Starost, M. F., Salpea, P., Faucz, F. R., Szarek, E., Nesterova, M., Leikin, S. and Stratakis, C. A. (2015) 'Haploinsufficiency for either one of the type-II regulatory subunits of protein kinase A improves the bone phenotype of Prkar1a(+/-) mice', *Human Molecular Genetics,* 24(21), 6080-6092.

Luo, R., Liu, B., Xie, Y., Li, Z., Huang, W., Yuan, J., He, G., Chen, Y., Pan, Q., Liu, Y., Tang, J., Wu, G., Zhang, H., Shi, Y., Liu, Y., Yu, C., Wang, B., Lu, Y., Han, C., Cheung, D. W., Yiu, S. M., Peng, S., Xiaoqian, Z., Liu, G., Liao, X., Li, Y., Yang, H., Wang, J., Lam, T. W. and Wang, J. (2012) 'SOAPdenovo2: an empirically improved memory-efficient short-read de novo assembler', *Gigascience,* 1(1), 18.

MacNeil, F. S. (1965) 'Evolution and distribution of the genus *Mya*, and Tertiary migrations of Mollusca', *US Geological Survey*, 483(1).

Mann, K. and Edsinger, E. (2014) 'The *Lottia gigantea* shell matrix proteome: re-analysis including MaxQuant iBAQ quantitation and phosphoproteome analysis', *Proteome Science*, 12, 28.

Mann, K. and Jackson, D. J. (2014) 'Characterization of the pigmented shell-forming proteome of the common grove snail *Cepaea nemoralis'*, *BMC Genomics*, 15, 249.

Margolin, A. A., Nemenman, I., Basso, K., Wiggins, C., Stolovitzky, G., Dalla Favera, R. and Califano, A. (2006) 'ARACNE: An algorithm for the reconstruction of gene regulatory networks in a mammalian cellular context', *BMC Bioinformatics*, 7(1), S7.

Marie, B., Jackson, D. J., Ramos-Silva, P., Zanella-Cleon, I., Guichard, N. and Marin, F. (2013) 'The shell-forming proteome of *Lottia gigantea* reveals both deep conservations and lineage-specific novelties', *The FEBS Journal*, 280(1), 214-32.

Marie, B., Joubert, C., Belliard, C., Tayale, A., Zanella-Cleon, I., Marin, F., Gueguen, Y. and Montagnani, C. (2012a) 'Characterization of MRNP34, a novel methionine-rich nacre protein from the pearl oysters', *Amino Acids*, 42(5), 2009-2017.

Marie, B., Joubert, C., Tayale, A., Zanella-Cleon, I., Belliard, C., Piquemal, D., Cochennec-Laureau, N., Marin, F., Gueguen, Y. and Montagnani, C. (2012b) 'Different secretory repertoires control the biomineralization processes of prism and nacre deposition of the pearl oyster shell', *Proceedings of the National Academy of Sciences of the United States of America*, 109(51), 20986-20991.

Marie, B., Zanella-Cleon, I., Corneillat, M., Becchi, M., Alcaraz, G., Plasseraud, L., Luquet, G. and Marin, F. (2011) 'Nautilin-63, a novel acidic glycoprotein from

the shell nacre of *Nautilus macromphalus*', The *FEBS Journal,* 278(12), 2117-2130.

Marie, B., Zanella-Cleon, I., Le Roy, N., Becchi, M., Luquet, G. and Marin, F. (2010) 'Proteomic analysis of the acid-soluble nacre matrix of the bivalve unio pictorum: detection of novel carbonic snhydrase and putative protease inhibitor proteins', *ChemBioChem,* 11(15), 2138-2147.

Marin, F., Corstjens, P., de Gaulejac, B., Vrind-De Jong, E. D. and Westbroek, P. (2000) 'Mucins and molluscan calcification - Molecular characterization of mucoperlin, a novel mucin-like protein from the nacreous shell layer of the fan mussel *Pinna nobilis* (Bivalvia, Pteriomorphia)', *Journal of Biological Chemistry,* 275(27), 20667-20675.

Marin, F., Le Roy, N. and Marie, B. (2012) 'The formation and mineralization of mollusk shell', *Frontiers in Bioscience,* 4, 1099-125.

Marin, F. and Luquet, G. (2004) 'Molluscan shell proteins', *Comptes Rendus Palevol,* 3(6-7), 469-492.

Matthew, J. D., Khromov, A. S., McDuffie, M. J., Somlyo, A. V., Somlyo, A. P., Taniguchi, S. and Takahashi, K. (2000) 'Contractile properties and proteins of smooth muscles of a calponin knockout mouse', *Journal of Physiology,* 529(3), 811-24.

McCarthy, D. J., Chen, Y. S. and Smyth, G. K. (2012) 'Differential expression analysis of multifactor RNA-Seq experiments with respect to biological variation', *Nucleic Acids Research,* 40(10), 4288-4297.

McDougall, C., Aguilera, F. and Degnan, B. M. (2013) 'Rapid evolution of pearl oyster shell matrix proteins with repetitive, low-complexity domains', *Journal of the Royal Society Interface,* 10(82), 20130041.

McDougall, C., Green, K., Jackson, D. J. and Degnan, B. M. (2011) 'Ultrastructure of the mantle of the gastropod *Haliotis asinina* and mechanisms of shell regionalization', *Cells Tissues Organs,* 194(2-4), 103-107.

Medakovic, D. (2000) 'Carbonic anhydrase activity and biomineralization process in embryos, larvae and adult blue mussels *Mytilus edulis* L', *Helgoland Marine Research,* 54(1), 1-6.

Meenakshi, V. R., Hare, P. E. and Wilbur, K. M. (1971) 'Amino acids of the organic matrix of neogastropod shells', *Comparative Biochemistry and Physiology Part B- Comparative Biochemistry,* 40(4), 1037-1043.

Meldrum, F. C. (2003) 'Calcium carbonate in biomineralisation and biomimetic chemistry', *International Materials Reviews,* 48(3), 187-224.

Mishra, M., Huang, J. Q. and Balasubramanian, M. K. (2014) 'The yeast actin cytoskeleton', *FEMS Microbiology Reviews,* 38(2), 213-227.

Mitta, G., Vandenbulcke, F., Hubert, F., Salzet, M. and Roch, P. (2000) 'Involvement of mytilins in mussel antimicrobial defense', *Journal of Biological Chemistry,* 275(17), 12954-12962.

Miyamoto, H., Miyashita, T., Okushima, M., Nakano, S., Morita, T. and Matsushiro, A. (1996) 'A carbonic anhydrase from the nacreous layer in oyster pearls', *Proceedings of the National Academy of Sciences of the United States of America,* 93(18), 9657-9660.

Miyamoto, H., Miyoshi, F. and Kohno, J. (2005) '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 mollusc *Pinctada fucata', Zoological Science,* 22(3), 311-315.

Miyashita, T., Takagi, R., Miyamoto, H. and Matsushiro, A. (2002) 'Identical carbonic anhydrase contributes to nacreous or prismatic layer formation in *Pinctada fucata* (Mollusca : Bivalvia)', *Veliger*, 45(3), 250-255.

Miyazaki, Y., Nishida, T., Aoki, H. and Samata, T. (2010) 'Expression of genes responsible for biomineralization of *Pinctada fucata* during development', *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, 155(3), 241-248.

Miyazaki, Y., Usui, T., Kajikawa, A., Hishiyama, H., Matsuzawa, N., Nishida, T., Machii, A. and Samata, T. (2008) 'Daily oscillation of gene expression associated with nacreous layer formation', *Frontiers of Materials Science in China*, 2(2), 162-166.

Morley, S. A., Hirse, T., Thorne, M. A. S., Portner, H. O. and Peck, L. S. (2012) 'Physiological plasticity, long term resistance or acclimation to temperature, in the Antarctic bivalve, *Laternula elliptica'*, *Comparative Biochemistry and Physiology A-Molecular & Integrative Physiology*, 162(1), 16-21.

Morris, J. H., Apeltsin, L., Newman, A. M., Baumbach, J., Wittkop, T., Su, G., Bader, G. D. and Ferrin, T. E. (2011) 'clusterMaker: a multi-algorithm clustering plugin for Cytoscape', *BMC Bioinformatics*, 12(1), 436.

Morse, M. P. and Zardus, J. D. (1996) 'Mollusca II. Bivalvia' in Harrison, F. W. and Kohn, A. J., eds., *Microscopic anatomy of invertebrates*, Wiley, 7-118.

Mount, A. S., Wheeler, A. P., Paradkar, R. P. and Snider, D. (2004) 'Hemocyte-mediated shell mineralization in the eastern oyster', *Science*, 304(5668), 297-300.

Nagai, K., Yano, M., Morimoto, K. and Miyamoto, H. (2007) 'Tyrosinase localization in mollusc shells', *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, 146(2), 207-214.

Nevo, E. (2001) 'Evolution of genome–phenome diversity under environmental stress', *Proceedings of the National Academy of Sciences*, 98(11), 6233-6240.

Niu, D. H., Wang, L., Sun, F. Y., Liu, Z. J. and Li, J. L. (2013) 'Development of molecular resources for an intertidal clam, *Sinonovacula constricta*, using 454 transcriptome sequencing', *Plos One*, 8(7), e67456.

Nogawa, C., Baba, H., Masaoka, T., Aoki, H. and Samata, T. (2012) 'Genetic structure and polymorphisms of the N16 gene in *Pinctada fucata'*, *Gene*, 504(1), 84-91.

O'Neill, M., Gaume, B., Denis, F. and Auzoux-Bordenave, S. (2013) 'Expression of biomineralisation genes in tissues and cultured cells of the abalone, *Haliotis tuberculata*', *Cytotechnology*, 65(5), 681-681.

Palmer, A. R. (1983) 'Relative cost of producing skeletal organic matrix versus calcification: Evidence from marine gastropods', *Marine Biology*, 75(2-3), 287-292.

Palmer, A. R. (1992) 'Calcification in marine molluscs: how costly is it?', *Proceedings of the National Academy of Sciences*, 89(4), 1379-1382.

Pan, C., Fang, D., Xu, G. R., Liang, J., Zhang, G. Y., Wang, H. Z., Xie, L. P. and Zhang, R. Q. (2014) 'A novel acidic matrix protein, PfN44, stabilizes magnesium calcite to inhibit the crystallization of aragonite', *Journal of Biological Chemistry*, 289(5), 2776-2787.

Parker, L. M., Ross, P. M., Raftos, D., Thompson, E. and O'Connor, W. A. (2011) 'The proteomic response of larvae of the Sydney rock oyster, *Saccostrea glomerata* to elevated pCO(2)', *Australian Zoologist*, 35(4), 1011-1023.

Peck, L. S. (2016) 'A Cold Limit to Adaptation in the Sea', *Trends in Ecology and Evolution*, 31(1), 13-26.

Peck, L. S., Convey, P. and Barnes, D. K. A. (2006) 'Environmental constraints on life histories in Antarctic ecosystems: tempos, timings and predictability', *Biological Reviews,* 81(1), 75-109.

Peck, L. S., Portner, H. O. and Hardewig, I. (2002) 'Metabolic demand, oxygen supply, and critical temperatures in the antarctic bivalve *Laternula elliptica*', *Physiological and Biochemical Zoology,* 75(2), 123-133.

Peck, L. S., Webb, K. E. and Bailey, D. M. (2004) 'Extreme sensitivity of biological function to temperature in Antarctic marine species', *Functional Ecology,* 18(5), 625-630.

Philipp, E., Brey, T., Portner, H. O. and Abele, D. (2005) 'Chronological and physiological ageing in a polar and a temperate mud clam', *Mechanisms of Ageing and Development,* 126(5), 598-609.

Philipp, E., Brey, T., Voigt, M. and Abele, D. (2008) 'Growth and age of *Laternula elliptica* in Potter Cove, King-George Island', *The Antarctic ecosystem of Potter Cove, King-George Island (Isla 25 de Mayo) Synopsis of research performed 1999-2006 at the Dallmann Laboratory and Jubany Station*, 216.

Philipp, E. and Abele, D. (2010) 'Masters of Longevity: Lessons from Long-Lived Bivalves - A Mini-Review', *Gerontology*, 56(1), 55-65.

Porat-Shliom, N., Milberg, O., Masedunskas, A. and Weigert, R. (2013) 'Multiple roles for the actin cytoskeleton during regulated exocytosis', *Cellular and Molecular Life Sciences,* 70(12), 2099-2121.

Powell, D. K. (2001) *The Reproductive Ecology of Antarctic Free-spawning Molluscs*, Doctoral dissertation, University of Southampton.

Queiros, A. M., Birchenough, S. N. R., Bremner, J., Godbold, J. A., Parker, R. E., Romero-Ramirez, A., Reiss, H., Solan, M., Somerfield, P. J., Van Colen, C., Van Hoey, G. and Widdicombe, S. (2013) 'A bioturbation classification of European marine infaunal invertebrates', *Ecology and Evolution,* 3(11), 3958-3985.

Robinson, M. D., McCarthy, D. J. and Smyth, G. K. (2010) 'edgeR: a Bioconductor package for differential expression analysis of digital gene expression data', *Bioinformatics,* 26(1), 139-140.

Rogers, A. D. (2007) 'Evolution and biodiversity of Antarctic organisms: a molecular perspective', *Philosophical Transactions of the Royal Society B: Biological Sciences,* 362(1488), 2191-2214.

Ronquist, F., Teslenko, M., van der Mark, P., Ayres, D. L., Darling, A., Hohna, S., Larget, B., Liu, L., Suchard, M. A. and Huelsenbeck, J. P. (2012) 'MrBayes 3.2: Efficient bayesian phylogenetic inference and model choice across a large model space', *Systematic Biology,* 61(3), 539-542.

Saavedra, C. and Bachere, E. (2006) 'Bivalve genomics', *Aquaculture,* 256(1-4), 1-14.

Saito, K., Hirai, M. Y. and Yonekura-Sakakibara, K. (2008) 'Decoding genes with coexpression networks and metabolomics - 'majority report by precogs'', *Trends in Plant Science,* 13(1), 36-43.

Sanchez-Ferrer, A., Rodriguez-Lopez, J. N., Garcia-Canovas, F. and Garcia-Carmona, F. (1995) 'Tyrosinase: a comprehensive review of its mechanism', *Biochimica et Biophysica Acta (BBA)-Protein Structure and Molecular Enzymology,* 1247(1), 1-11.

Sartori, A. F., Passos, F. D. and Domaneschi, O. (2006) 'Arenophilic mantle glands in the Laternulidae (Bivalvia : Anomalodesmata) and their evolutionary significance', *Acta Zoologica,* 87(4), 265-272.

Sasaki, H. and Hogan, B. L. M. (1993) 'Differential expression of multiple fork head-related genes during gastrulation and axaial pattern-formation in the mouse embryo', *Development,* 118(1), 47-59.

Sato, Y., Inoue, N., Ishikawa, T., Ishibashi, R., Obata, M., Aoki, H., Atsumi, T. and Komaru, A. (2013) 'Pearl microstructure and expression of shell matrix protein genes MSI31 and MSI60 in the pearl sac epithelium of *Pinctada fucata* by in situ hybridization', *Plos One,* 8(1), e52372.

Schaefer, R. J., Michno, J. M. and Myers, C. L. (2017) 'Unraveling gene function in agricultural species using gene co-expression networks', *Biochimica et Biophysica Acta-Gene Regulatory Mechanisms,* 1860(1), 53-63.

Schmittgen, T. D. and Livak, K. J. (2008) 'Analyzing real-time PCR data by the comparative C-T method', *Nature Protocols,* 3(6), 1101-1108.

Schneider, C. A., Rasband, W. S. and Eliceiri, K. W. (2012) 'NIH Image to ImageJ: 25 years of image analysis', *Nature Methods,* 9(7), 671-675.

Schoenitzer, V. and Weiss, I. M. (2007) 'The structure of mollusc larval shells formed in the presence of the chitin synthase inhibitor Nikkomycin Z', *BMC Structural Biology,* 7(71), 1-24.

Sgonc, R. and Gruber, J. (2013) 'Age-related aspects of cutaneous wound healing: A mini-review', *Gerontology,* 59(2), 159-164.

Shannon, P., Markiel, A., Ozier, O., Baliga, N. S., Wang, J. T., Ramage, D., Amin, N., Schwikowski, B. and Ideker, T. (2003) 'Cytoscape: A software environment for integrated models of biomolecular interaction networks', *Genome Research,* 13(11), 2498-2504.

Shen, X. Y., Belcher, A. M., Hansma, P. K., Stucky, G. D. and Morse, D. E. (1997) 'Molecular cloning and characterization of lustrin A, a matrix protein from shell

and pearl nacre of *Haliotis rufescens*', *Journal of Biological Chemistry*, 272(51), 32472-32481.

Shi, Y., Yu, C., Gu, Z., Zhan, X., Wang, Y. and Wang, A. (2013) 'Characterization of the pearl oyster (*Pinctada martensii*) mantle transcriptome unravels biomineralization genes', *Marine Biotechnology*, 15(2), 175-187.

Siah, A., Dohoo, C., McKenna, P., Delaporte, M. and Berthe, F. C. J. (2008) 'Selecting a set of housekeeping genes for quantitative real-time PCR in normal and tetraploid haemocytes of soft-shell clams, *Mya arenaria*', *Fish & Shellfish Immunology*, 25(3), 202-207.

Silvestro, D. and Michalak, I. (2012) 'raxmlGUI: a graphical front-end for RAxML', *Organisms Diversity & Evolution*, 12(4), 335-337.

Sleight, V. A., Thorne, M. A., Peck, L. S. and Clark, M. S. (2015) 'Transcriptomic response to shell damage in the Antarctic clam, *Laternula elliptica*: time scales and spatial localisation', *Marine Genomics*, 20, 45-55.

Sleight, V. A., Thorne, M. A. S., Peck, L. S., Arivalagan, J., Berland, S., Marie, A. and Clark, M. S. 'Characterisation of the mantle transcriptome and biomineralisation genes in the blunt-gaper clam, Mya truncata', *Marine Genomics*, 27, 47-55.

Solomon, K. R., Sharma, P., Chan, M., Morrison, P. T. and Finberg, R. W. (2004) 'CD109 represents a novel branch of the alpha 2-macroglobulin/complement gene family', *Gene*, 327(2), 171-183.

Souza, A., Souza, D. R. V., Sanabani, S. S., Giorgi, R. R. and Bendit, I. (2009) 'The performance of semi-quantitative differential PCR is similar to that of real-time PCR for the detection of the MYCN gene in neuroblastomas', *Brazilian Journal of Medical and Biological Research*, 42(9), 791-795.

Stefansson, B., Ohama, T., Daugherty, A. E. and Brautigan, D. L. (2008) 'Protein phosphatase 6 regulatory subunits composed of ankyrin repeat domains', *Biochemistry,* 47(5), 1442-1451.

Su, J. T., Liang, X., Zhou, Q., Zhang, G. Y., Wang, H. Z., Xie, L. P. and Zhang, R. Q. (2013) 'Structural characterization of amorphous calcium carbonate-binding protein: an insight into the mechanism of amorphous calcium carbonate formation', *Biochemical Journal,* 453, 179-186.

Sudo, S., Fujikawa, T., Nagakura, T., Ohkubo, T., Sakaguchi, K., Tanaka, M., Nakashima, K. and Takahashi, T. (1997) 'Structures of mollusc shell framework proteins', *Nature,* 387(6633), 563-564.

Sukhotin, A. A., Abele, D. and Portner, H. O. (2002) 'Growth, metabolism and lipid peroxidation in *Mytilus edulis*: age and size effects', *Marine Ecology Progress Series,* 226, 223-234.

Suzuki, M., Iwashima, A., Kimura, M., Kogure, T. and Nagasawa, H. (2013) 'The molecular evolution of the pif family proteins in various species of mollusks', *Marine Biotechnol,* 15(2), 145-58.

Suzuki, M., Iwashima, A., Tsutsui, N., Ohira, T., Kogure, T. and Nagasawa, H. (2011) 'Identification and characterisation of a calcium carbonate-binding protein, blue mussel shell protein (BMSP), from the nacreous layer', *ChemBioChem,* 12(16), 2478-2487.

Suzuki, M., Murayama, E., Inoue, H., Ozaki, N., Tohse, H., Kogure, T. and Nagasawa, H. (2004) 'Characterization of prismalin-14, a novel matrix protein from the prismatic layer of the Japanese pearl oyster (*Pinctada fucata*)', *Biochemical Journal,* 382, 205-213.

Suzuki, M., Saruwatari, K., Kogure, T., Yamamoto, Y., Nishimura, T., Kato, T. and Nagasawa, H. (2009) 'An acidic matrix protein, pif, is a key macromolecule for nacre formation', *Science,* 325(5946), 1388-1390.

Suzuki, M., Iwashima, A., Kimura, M., Kogure, T., & Nagasawa, H. (2013) 'The molecular evolution of the pif family proteins in various species of mollusks', *Marine Biotechnology,* 15(2), 145-158.

Takeuchi, T., Kawashima, T., Koyanagi, R., Gyoja, F., Tanaka, M., Ikuta, T., Shoguchi, E., Fujiwara, M., Shinzato, C., Hisata, K., Fujie, M., Usami, T., Nagai, K., Maeyama, K., Okamoto, K., Aoki, H., Ishikawa, T., Masaoka, T., Fujiwara, A., Endo, K., Endo, H., Nagasawa, H., Kinoshita, S., Asakawa, S., Watabe, S. and Satoh, N. (2012) 'Draft genome of the pearl oyster *Pinctada fucata*: A platform for understanding bivalve biology', *DNA Research,* 19(2), 117-130.

Takgi, R. and Miyashita, T. (2014) 'A cDNA cloning of a novel alpha-class tyrosinase of *Pinctada fucata*: Its expression analysis and characterization of the expressed protein', *Enzyme research,* 2014, 780549-780549.

Tamura, K., Peterson, D., Peterson, N., Stecher, G., Nei, M. and Kumar, S. (2011) 'MEGA5: Molecular evolutionary genetics analysis using maximum likelihood, evolutionary distance, and maximum parsimony methods', *Molecular Biology and Evolution,* 28(10), 2731-2739.

Tanur, A. E., Gunari, N., Sullan, R. M. A., Kavanagh, C. J. and Walker, G. C. (2010) 'Insights into the composition, morphology, and formation of the calcareous shell of the serpulid *Hydroides dianthus', Journal of Structural Biology*, 169(2), 145-160.

Tasken, K. and Aandahl, E. M. (2004) 'Localized effects of cAMP mediated by distinct routes of protein kinase A', *Physiological Reviews,* 84(1), 137-167.

Tenyi, A., de Atauri, P., Gomez-Cabrero, D., Cano, I., Clarke, K., Falciani, F., Cascante, M., Roca, J. and Maier, D. (2016) 'ChainRank, a chain prioritisation method for contextualisation of biological networks', *BMC Bioinformatics*, 17(1), 17.

Thompson, J. D., Thierry, J. C. and Poch, O. (2003) 'RASCAL: rapid scanning and correction of multiple sequence alignments', *Bioinformatics*, 19(9), 1155-1161.

Timpl, R., Sasaki, T., Kostka, G. and Chu, M. L. (2003) 'Fibulins: A versatile family of extracellular matrix proteins', *Nature Reviews Molecular Cell Biology*, 4(6), 479-489.

Tomanek, L. (2011) 'Environmental proteomics: Changes in the proteome of marine organisms in response to environmental stress, pollutants, infection, symbiosis, and development', *Annual Review of Marine Science*, 3, 373-399.

Tomida, T., Takekawa, M. and Saito, H. (2015) 'Oscillation of p38 activity controls efficient pro-inflammatory gene expression', *Nature Communications*, 6.

Trueman, E. R., Wilbur, K. M. and Clarke, R. (1988) *The Mollusca: Form and function*, Academic Press.

Turner, J., Barrand, N. E., Bracegirdle, T. J., Convey, P., Hodgson, D. A., Jarvis, M., Jenkins, A., Marshall, G., Meredith, M. P., Roscoe, H., Shanklin, J., French, J., Goosse, H., Guglielmin, M., Gutt, J., Jacobs, S., Kennicutt, M. C., Masson-Delmotte, V., Mayewski, P., Navarro, F., Robinson, S., Scambos, T., Sparrow, M., Summerhayes, C., Speer, K. and Klepikov, A. (2014) 'Antarctic climate change and the environment: an update', *Polar Record*, 50(3), 237-259.

Ueda, T., Araki, N., Mano, M., Myoui, A., Joyama, S., Ishiguro, S., Yamamura, H., Takahashi, K., Kudawara, I. and Yoshikawa, H. (2002) 'Frequent expression of smooth muscle markers in malignant fibrous histiocytoma of bone', *Journal of Clinical Pathology*, 55(11), 853-858.

Van Valen, L. (1965) 'Morphological variation and width of ecological niche', *American Naturalist*, 377-390.

Vera, J. C., Wheat, C. W., Fescemyer, H. W., Frilander, M. J., Crawford, D. L., Hanski, I. and Marden, J. H. (2008) 'Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing', *Molecular Ecology*, 17(7), 1636-1647.

Vermeij, G. J. (2005) 'Shells inside out: The architecture, evolution and function of shell envelopment in molluscs', *Evolving Form and Function: Fossils and Development*, 197-221.

Vidavsky, N., Addadi, S., Mahamid, J., Shimoni, E., Ben-Ezra, D., Shpigel, M., Weiner, S. and Addadi, L. (2014) 'Initial stages of calcium uptake and mineral deposition in sea urchin embryos', *Proceedings of the National Academy of Sciences of the United States of America*, 111(1), 39-44.

Vidavsky, N., Masic, A., Schertel, A., Weiner, S. and Addadi, L. (2015) 'Mineral-bearing vesicle transport in sea urchin embryos', *Journal of Structural Biology*, 192(3), 358-365.

Vieira, F. A., Gregorio, S. F., Ferraresso, S., Thorne, M. A. S., Costa, R., Milan, M., Bargelloni, L., Clark, M. S., Canario, A. V. M. and Power, D. M. (2011) 'Skin healing and scale regeneration in fed and unfed sea bream, *Sparus auratus*', *BMC Genomics*, 12(1), 490.

Vindin, H. and Gunning, P. (2013) 'Cytoskeletal tropomyosins: choreographers of actin filament functional diversity', *Journal of Muscle Research and Cell Motility*, 34(3-4), 261-274.

Waite, J. H., Saleuddin, A. S. M. and Andersen, S. O. (1979) 'Periostracin - soluble precursor of sclerotized periostracum in *Mytilus edulis*-L', *Journal of Comparative Physiology*, 130(4), 301-307.

Waller, C. L., Overall, A., Fitzcharles, E. M. and Griffiths, H. (2016) 'First report of *Laternula elliptica* in the Antarctic intertidal zone', *Polar Biology*, 1-4.

Wang, J. G. and Slungaard, A. (2006) 'Role of eosinophil peroxidase in host defense and disease pathology', *Archives of Biochemistry and Biophysics*, 445(2), 256-260.

Wang, P., Bouwman, F. G. and Mariman, E. C. M. (2009) 'Generally detected proteins in comparative proteomics - A matter of cellular stress response?', *Proteomics*, 9(11), 2955-2966.

Wang, X. T., Song, X. R., Wang, T., Zhu, Q. H., Miao, G. Y., Chen, Y. X., Fang, X. D., Que, H. Y., Li, L. and Zhang, G. F. (2013) 'Evolution and functional analysis of the Pif97 gene of the Pacific oyster *Crassostrea gigas*', *Current Zoology*, 59(1), 109-115.

Warren, L., Malarska, A. and Jardillier, J. C. (1995) 'The structure of p-glycoprotein and the secretion of lysosomal-enzymes in multidrug-resistant cells', *Cancer Chemotherapy and Pharmacology*, 35(3), 267-269.

Weiner, S. and Hood, L. (1975) 'Soluble protein of the organic matrix of mollusk shells: a potential template for shell formation', *Science*, 190(4218), 987-989.

Weiss, I. M. (2012) 'Species-specific shells: Chitin synthases and cell mechanics in molluscs', *Zeitschrift Fur Kristallographie*, 227(11), 723-738.

Weiss, I. M., Kaufmann, S., Mann, K. and Fritz, M. (2000) 'Purification and characterization of perlucin and perlustrin, two new proteins from the shell of the mollusc *Haliotis laevigata*', *Biochemical and Biophysical Research Communications*, 267(1), 17-21.

Weiss, I. M., Lueke, F., Eichner, N., Guth, C. and Clausen-Schaumann, H. (2013) 'On the function of chitin synthase extracellular domains in biomineralization', *Journal of Structural Biology*, 183(2), 216-225.

Werner, G. D. A., Gemmell, P., Grosser, S., Hamer, R. and Shimeld, S. M. (2013) 'Analysis of a deep transcriptome from the mantle tissue of *Patella vulgata*

Linnaeus (Mollusca: Gastropoda: Patellidae) reveals candidate biomineralising genes', *Marine Biotechnology,* 15(2), 230-243.

Westermann, B., Schmidtberg, H. and Beuerlein, K. (2005) 'Functional morphology of the mantle of *Nautilus pompilius* (Mollusca, Cephalopoda)', *J Morphol,* 264(3), 277-85.

Wilbur, K. M. and Saleuddin, A. S. M. (1983) *The Mollusca: Physiology, part 2 / edited by A.S.M. Saleuddin, Karl M. Wilbur,* Academic Press.

Wolfe, C. J., Kohane, I. S. and Butte, A. J. (2005) 'Systematic survey reveals general applicability of "guilt-by-association" within gene coexpression networks', *BMC Bioinformatics,* 6(1), 277.

Xiong, X., Chen, L., Li, Y., Xie, L. and Zhang, R. (2006) 'Pf-ALMP, a novel astacin-like metalloproteinase with cysteine arrays, is abundant in hemocytes of pearl oyster *Pinctada fucata*', *Biochimica et Biophysica Acta - Gene Structure and Expression,* 1759(11–12), 526-534.

Yan, Z. G., Fang, Z., Ma, Z. J., Deng, J. Y., Shuo, L. A., Xie, L. and Zhang, R. Q. (2007) 'Biomineralization: Functions of calmodulin-like protein in the shell formation of pearl oyster', *Biochimica et Biophysica Acta-General Subjects,* 1770(9), 1338-1344.

Yang, Q., Orman, M. A., Berthiaume, F., Ierapetritou, M. G. and Androulakis, I. P. (2012) 'Dynamics of short-term gene expression profiling in liver following thermal injury', *Journal of Surgical Research,* 176(2), 549-558.

Yin, Y., Huang, J., Paine, M. L., Reinhold, V. N. and Chasteen, N. D. (2005) 'Structural characterization of the major extrapallial fluid protein of the mollusc *Mytilus edulis*: Implications for functions', *Biochemistry,* 44(31), 10720-10731.

Zhang, C., Xie, L. P., Huang, J., Chen, L. and Zhang, R. Q. (2006) 'A novel putative tyrosinase involved in periostracum formation from the pearl oyster (*Pinctada fucata*)', *Biochemical and Biophysical Research Communications*, 342(2), 632-639.

Zhang, G., Fang, X., Guo, X., Li, L., Luo, R., Xu, F., Yang, P., Zhang, L., Wang, X., Qi, H., Xiong, Z., Que, H., Xie, Y., Holland, P. W. H., Paps, J., Zhu, Y., Wu, F., Chen, Y., Wang, J., Peng, C., Meng, J., Yang, L., Liu, J., Wen, B., Zhang, N., Huang, Z., Zhu, Q., Feng, Y., Mount, A., Hedgecock, D., Xu, Z., Liu, Y., Domazet-Loso, T., Du, Y., Sun, X., Zhang, S., Liu, B., Cheng, P., Jiang, X., Li, J., Fan, D., Wang, W., Fu, W., Wang, T., Wang, B., Zhang, J., Peng, Z., Li, Y., Li, N., Wang, J., Chen, M., He, Y., Tan, F., Song, X., Zheng, Q., Huang, R., Yang, H., Du, X., Chen, L., Yang, M., Gaffney, P. M., Wang, S., Luo, L., She, Z., Ming, Y., Huang, W., Zhang, S., Huang, B., Zhang, Y., Qu, T., Ni, P., Miao, G., Wang, J., Wang, Q., Steinberg, C. E. W., Wang, H., Li, N., Qian, L., Zhang, G., Li, Y., Yang, H., Liu, X., Wang, J., Yin, Y. and Wang, J. (2012) 'The oyster genome reveals stress adaptation and complexity of shell formation', *Nature*, 490(7418), 49-54.

Zhou, Y. J., He, Z. X., Huang, J., Gong, N. P., Yan, Z. G., Liu, X. J., Sun, J. A., Wang, H. Z., Zhang, G. Y., Xie, L. P. and Zhang, R. Q. (2010) ' Cloning and characterization of the activin like receptor 1 homolog (Pf-ALR1) in the pearl oyster, *Pinctada fucata*' *Comparative Biochemistry and Physiology B-Biochemistry & Molecular Biology*, 156(3), 158-167.

Ziuganov, V., San Miguel, E., Neves, R. J., Longa, A., Fernandez, C., Amaro, R., Beletsky, V., Popkovitch, E., Kaliuzhin, S. and Johnson, T. (2000) 'Life span variation of the freshwater pearl shell: A model species for testing longevity mechanisms in animals', *AMBIO: A Journal of the Human Environment*, 29(2), 102-105.