GROWTH-RELATED GENE EXPRESSION IN *HALIOTIS MIDAE*

Mathilde van der Merwe

Dissertation presented for the degree of Doctor of Philosophy (Genetics) at Stellenbosch University

Promoter: Dr Rouvay Roodt-Wilding
Co-promoters: Dr Stéphanie Auzoux-Bordenave and Dr Carola Niesler

December 2010

Declaration

By submitting this dissertation, I declare that the entirety of the work contained therein is my own, original work, that I am the authorship owner thereof (unless to the extent explicitly otherwise stated) and that I have not previously in its entirety or in part submitted it for obtaining any qualification.

Date: 09/11/2010

Acknowledgements

I would like to express my sincere gratitude and appreciation to the following persons for their contribution towards the successful completion of this study:

Dr **Rouvay Roodt-Wilding** for her continued encouragement, careful attention to detail and excellent facilitation throughout the past years;

Dr **Stéphanie Auzoux-Bordenave** for valuable lessons in abalone cell culture and suggestions during completion of the manuscript;

Dr **Carola Niesler** for setting an example and providing guidance that already started preparing me for a PhD several years ago;

Dr **Paolo Franchini** for his patience and greatly valued assistance with bioinformatics;

Dr **Aletta van der Merwe** and my fellow lab-colleagues for their technical and moral support;

My dear husband **Willem** for his love, support and enthusiasm, for sitting with me during late nights in the lab and for making me hundreds of cups of tea;

My **parents** for their love and encouragement and for instilling the determination in me to complete my studies;

All my family and friends for their sincere interest.

Abstract

The slow growth rate of *Haliotis midae* impedes the optimal commercial production of this most profitable South African aquaculture species. To date, no comprehensive effort has been made to identify genes associated with growth variation in farmed *H. midae*. The aim of this study was therefore to investigate growth variation in *H. midae* and to identify and quantify the expression of selected growth-related genes. Towards this aim, molecular methodologies and cell cultures were combined as a time-efficient and economical way of studying abalone transcriptomics and cell biology.

Modern Illumina sequencing-by-synthesis technology and subsequent sequence annotation were used to elucidate differential gene expression between two sibling groups of abalone demonstrating significant growth variation. Following transcriptome sequencing, genes involved in growth and metabolism, previously unknown in *H. midae*, were identified. The expression of selected target genes involved in growth was subsequently analyzed by quantitative real-time PCR (qPCR).

The feasibility of primary cell cultures for *H. midae* was furthermore investigated by targeting embryo, larval and haemolymph tissues for the initiation of primary cell culture. Larval cells and haemocytes could be successfully maintained *in vitro* for limited periods. Primary haemocyte cultures demonstrated to be a suitable *in vitro* system for studying gene expression and were subsequently used for RNA extraction and qPCR, to evaluate differential growth induced by bovine insulin and epidermal growth factor (EGF).

Gene expression was thus quantified in fast and slow growing abalone and in *in vitro* primary haemocyte cultures treated with different growth stimulating factors. The results obtained from transcriptome analysis and qPCR revealed significant differences in gene expression between large and small abalone, and between treated and untreated haemocyte cell cultures. Throughout *in vivo* and *in vitro* qPCR experiments, the up-regulation of genes involved in the insulin signalling pathway provides evidence for the involvement of insulin in enhanced growth rate for various *H. midae* tissues.

Besides the regulation of target genes, valuable knowledge was also gained in terms of reference genes, during qPCR experimentation. By quantifying the stable expression of two genes (8629, ribosomal protein S9 and 12621, ornithine decarboxylase) in various tissues and under various conditions, suitable reference genes, that can also be used in future *H. midae* qPCR studies, were identified.

By providing evidence at the transcriptional level for the involvement of insulin, insulin-like growth factors (IGFs) and insulin-like growth factor binding proteins (IGFBPs) in improved growth rate of *H. midae*, the relevance of investigating ways to stimulate insulin/IGF release in aquaculture species was again emphasized. As nutritional administration remains the most probable route of introducing agents

that can stimulate the release of insulin-related peptides, continuous endeavours to stimulate abalone growth through a nutritional approach is encouraged.

This is the first time next generation sequencing is used towards the large scale transcriptome sequencing of any haliotid species and also the first time a comprehensive investigation is launched towards the establishment of primary cell cultures for *H. midae*. A considerable amount of sequence data was furthermore annotated for the first time in *H. midae*. The results obtained here provide a foundation for future genetic studies exploring ways to optimise the commercial production of *H. midae*.

Opsomming

Die stadige groeitempo van *Haliotis midae* belemmer die optimale kommersiële produksie van hierdie mees winsgewende Suid-Afrikaanse akwakultuur spesie. Tot op hede is geen omvattende poging aangewend om gene verwant aan groeivariasie in *H. midae* te identifiseer nie. Die doel van hierdie studie was dus om groeivariasie in *H. midae* te ondersoek en om spesifieke groei-gekoppelde gene te identifiseer en hul uitdrukking te kwantifiseer. Ter bereiking van hierdie doel is molekulêre metodes en selkulture gekombineer as 'n tydsbesparende en ekonomiese manier om perlemoen transkriptomika en selbiologie te bestudeer.

Moderne Illumina volgordebepaling-deur-sintese tegnologie en daaropvolgende annotasie is gebruik om verskille in geenuitdrukking tussen naby-verwante groepe perlemoen, wat noemenswaardige groeivariasie vertoon, toe te lig. Na afloop van die transkriptoom volgordebepaling is gene betrokke by groei en metabolisme, vantevore onbekend in *H. midae*, geïdentifiseer. Die uitdrukking van uitgesoekte teikengene betrokke by groei is vervolgens ge-analiseer deur kwantitatiewe "real-time PCR" (qPCR).

Die lewensvatbaarheid van primêre selkulture vir *H. midae* is ook ondersoek deur embrio, larwe en hemolimf weefsels te teiken vir die daarstelling van primêre selkulture. Larweselle en hemosiete kon *in vitro* suksesvol onderhou word vir beperkte periodes. Primêre hemosietkulture het geblyk 'n gepaste *in vitro* sisteem te wees om geenuitdrukking te bestudeer en dit is vervolgens gebruik vir RNS ekstraksie en qPCR, om differensiële groei, geïnduseer deur insulien en epidermale groeifaktor (EGF), te evalueer.

Geenuitdrukking is dus gekwantifiseer in vinnig- en stadiggroeiende perlemoen en in *in vitro* primêre hemosiet selkulture wat behandel is met verskillende groei stimulante. Die resultate wat verkry is van transkriptoomanalise en qPCR het noemenswaardige verskille in geenuitdrukking tussen groot en klein perlemoen, en tussen behandelde en onbehandelde hemosiet selkulture uitgelig. Die op-regulering van gene betrokke by die insulien sein-padweg, tydens *in vivo* en *in vitro* qPCR eksperimente, bied getuienis vir die betrokkenheid van insulien in die verhoogde groeitempo van verskeie *H. midae* weefsels.

Benewens die regulering van teikengene is waardevolle kennis ook ingewin in terme van verwysingsgene tydens qPCR eksperimentering. Deur die stabiele uitdrukking van twee gene (8629, ribosomale proteïen S9 en 12621, ornitien dekarboksilase) te kwantifiseer in verskeie weefsels en onder verskeie kondisies is gepaste verwysingsgene, wat ook in toekomstige *H. midae* qPCR eksperimente aangewend kan word, geïdentifiseer.

Deur getuienis vir die betrokkenheid van insulien, insuliensoortige groeifaktor en insuliensoortige groeifaktor-bindingsproteïene by verbeterde groei van *H. midae* op transkripsievlak te bied, is die toepaslikheid van bestudering van maniere om insulienvrystelling in akwakultuurspesies te stimuleer,

beklemtoon. Aangesien voeding die mees waarskynlike roete is om middele wat insuliensoortige peptiedvrystelling stimuleer daar te stel, word vogehoue pogings om perlemoengroei deur die regte voeding te stimuleer, aangemoedig.

Hierdie is die eerste studie wat volgende generasie volgordebepaling ("next generation sequencing") gebruik vir die grootskaalse transkriptoom volgordebepaling van enige haliotied spesie. Dit is ook die eerste keer dat 'n omvattende ondersoek geloods word na die daarstelling van primêre selkulture vir *H. midae*. 'n Aansienlike hoeveelheid volgorde data is ook vir die eerste keer geannoteer in *H. midae*. Die resultate wat hier verkry is bied 'n basis vir toekomstige genetiese studies wat maniere ondersoek om die kommersiële produksie van perlemoen te optimiseer.

Publications and conference proceedings resulting from this PhD

An article reporting a large portion of the content of **chapter three** was accepted for publication:

Van der Merwe, M., Auzoux-Bordenave, S., Niesler, C and Roodt-Wilding, R. 2010. Investigating the establishment of primary cell culture from different abalone (*Haliotis midae*) tissues. *Cytotechnology* 62 (3): 265-277.

Results from **chapter two** were presented at the 10th International Symposium on Genetics in Aquaculture, Bangkok, Thailand, 22 - 26 June 2009:

Van der Merwe, M., Franchini, P. and Roodt-Wilding, R. Growth-related Gene Expression in *Haliotis midae*: Study of Transcriptome Sequence Data using Next Generation Sequence Technology. 2009. Proceedings of the ISGA X: 55

Results from **chapters two and four** were presented at the 17th World Congress of Malacology, Phuket, Thailand, 18 - 24 July 2010:

Van der Merwe, M and Roodt-Wilding, R. Growth-related gene expression in *Haliotis midae*: Analysis of transcriptome sequence data using next generation sequence technology and quantitative real-time PCR. 2010. Tropical Natural History, Supplement 3: 12

Table of Contents

Declaration I
Acknowledgements II
Abstract III
Opsomming V
Publications and conference proceedings resulting from this PhD VII
Table of Contents VIII
List of Tables XI
List of Figures XII
List of Abbreviations XIV

1 LITERATURE REVIEW, BACKGROUND AND AIM 1
1.1 *Haliotis midae* 1
1.1.1 Biology 2
1.1.1.1 Anatomy 2
1.1.1.2 Development 4
1.1.1.3 Feeding, metabolism and growth 5
1.1.2 As aquaculture species 6
1.2 Growth as a desirable trait in animal husbandry 7
1.2.1 Genes within the somatotropic axis and central nervous system 9
1.2.2 Genes from the muscle tissue and haemolymph 14
1.2.3 Genes involved in shell deposition and growth 16
1.2.4 Miscellaneous genes and proteins that could play a role in growth regulation 19
1.3 Methods to study growth variation 21
1.3.1 Molecular biology approach to study gene expression 21
1.3.1.1 Transcriptome analysis 22
1.3.1.2 Real-time PCR 24
1.3.2 *In vitro* investigation of growth using cell culture 27
1.3.2.1 Overview 27
1.3.2.2 Applications of cell culture 28
1.4 Aim of this study 29
1.5 References 31
2 NEXT-GENERATION SEQUENCING OF THE *H. MIDAE* TRANSCRIPTOME TO IDENTIFY DIFFERENTIALLY EXPRESSED GENES 46
2.1 Introduction 46
2.2 Materials and Methods 49
2.2.1 Sampling 49
2.2.2 RNA extractions 50
2.2.3 cDNA Library preparation and sequencing 51

2.2.4 Bioinformatics 53
2.2.4.1 Sequence assembly 53
2.2.4.2 Differential expression analysis 55
2.2.4.3 Annotation 55
2.3 Results 59
2.3.1 Sampling 59
2.3.2 RNA extraction 59
2.3.3 cDNA Library preparation, sequencing and sequence assembly 60
2.3.4 Differential expression analysis 61
2.3.5 Annotation 61
2.4 Discussion 71
2.5 References 80
3 INVESTIGATING THE ESTABLISHMENT OF PRIMARY CELL CULTURES FROM *H. MIDAE* TISSUES 87
3.1 Introduction 87
3.1.1 Cell culture in marine molluscs 87
3.1.2 Considerations for primary culture initiation 88
3.1.2.1 Choice of tissue 88
3.1.2.2 Primary culture initiation methods 89
3.1.2.3 Medium and maintenance 92
3.1.2.4 Contamination 95
3.1.3 Cell characterization 95
3.1.4 Tissues of origin for *H. midae* (embryos, larvae, haemocytes) 96
3.1.5 Initiative for establishing *H. midae* primary cell cultures 99
3.2 Materials and Methods 100
3.2.1 Cell collection, dissociation and culture initiation 100
3.2.2 Conditions for cell maintenance 103
3.2.3 Viability assessment 106
3.2.4 Statistical analysis 107
3.3 Results 108
3.3.1 Embryo cell cultures 108
3.3.2 Larval cell cultures 109
3.3.3 Haemocyte cell cultures 114
3.3.4 Contamination 117
3.4 Concluding remarks 118
3.4.1 Suitability of various tissues for primary cell culture 118
3.4.2 Applications 123

3.5 References 125
4 APPLICATION OF CELL CULTURE AND QUANTITATIVE REAL-TIME PCR TOWARDS IDENTIFICATION OF GROWTH AND METABOLISM GENES IN *H.MIDAE* 130
4.1 Introduction 130
4.2 Materials and Methods 133
4.2.1 Haemocyte cell cultures 133
4.2.2 RNA extraction and reverse transcription 133
4.2.3 Primer design 135
4.2.4 PCR optimization 138
4.2.5 Reference gene validation 139
4.2.6 Quantitative real-time PCR 139
4.2.7 Data analysis and bioinformatics 141
4.3 Results 142
4.3.1 Haemocyte cell cultures 142
4.3.2 RNA extraction and reverse transcription 142
4.3.3 Primer optimization and confirmation of primer specificity 148
4.3.4 Reference gene validation 152
4.3.5 Quantitative real-time PCR 152
4.3.5.1 Standard curves 152
4.3.5.2 Data analysis and bioinformatics (Delta delta CT results, REST) 156
4.4 Discussion 162
4.5 References 170
5 CONCLUSIONS AND RECOMMENDATIONS 177
5.1 Conclusions and recommendations 177
5.2 References 186
6 APPENDIX 190

List of Tables

Table 1.1 Taxonomic classification of *Haliotis* (The Uniprot Consortium, 2010) 1

Table 2.1 Previously identified molecular markers in *Haliotis midae* 46

Table 2.2 Databases used during dCAS annotation 57

Table 2.3 Results of ANOVA for differences between large (L) and small (S) groups 59

Table 2.4 Concentrations and absorbance ratios for RNA after extraction and cleanup 59

Table 2.5 Functional classification of contigs from the *H. midae* transcriptome (R) with a BLAST E-value of ≤ 10^{-10}, based on the KOG database 64

Table 2.6 Relative representation of contigs in selected subcategories of KOG annotation 66

Table 2.7 Functional classification of contigs from the *H. midae* transcriptome (R) with a BLAST E-value of ≤ 10^{-10}, based on the KEGG database 67

Table 2.8 Relative representation of contigs in selected subcategories of KEGG annotation 69

Table 2.9 Annotation results for differentially expressed L and S *H. midae* sequences, across all databases 70

Table 3.1 Dissociation methods reported for other mollusc primary tissue culture trials 90

Table 3.2 Media and supplements reported for other mollusc primary tissue culture trials 93

Table 3.3 Dissociation methods used for *H. midae* embryos and larvae 102

Table 3.4 Culture media formulations used for *H. midae* primary cell cultures 104

Table 3.5 Media supplementation to study the effect on viability of cultured *H. midae* larval cells 105

Table 3.6 Medium D supplementation to study the effect on viability of cultured *H. midae* haemocytes 106

Table 3.7 Summary of the effects of different combinations of dissociation and maintenance protocols on viability of *H. midae* larval cell cultures 110

Table 4.1 Culture medium and supplementations used for *H. midae* haemocyte cell culture 133

Table 4.2 Sequences chosen for verification of differential expression by quantitative real-time PCR 136

Table 4.3 Primer sequences for amplification of target and reference genes 138

Table 4.4 Concentration and 260/280 absorbance ratios for RNA samples (presented as an average for three replicates per sample) 142

Table 4.5 PCR optimization for the first five primer pairs 148

Table 4.6 Expected sizes of amplification products 148

Table 4.7 Relative fold change calculated by the $2^{-\Delta\Delta CT}$ method 157

Table 4.8 Relative expression values of genes 752, 2380 and 13596 in large and small abalone, with associated P-values and 95 % confidence intervals (CI) 160

Table 4.9 Relative expression values of genes 809, 54 and 3309 in haemocytes in response to growth factor treatments, with associated P-values and 95 % confidence intervals (CI) 161

List of Figures

Figure 1.1 Dorsal and ventral views of *H. midae*.....2
Figure 1.2 Ventral view of organs and soft body parts of the abalone (Henry, 1995).....4
Figure 2.1 Abalone sampled for transcriptome sequencing.....49
Figure 2.2 Example of 2 % agarose gel after DNA excision.....53
Figure 2.3 Example of 2 % agarose gel for verification of correct amplified product.....53
Figure 2.4 Denaturing formaldehyde agarose gel with RNA extracted from small (S1, S2, S3) and large (L1, L2, L3) tissue samples.....60
Figure 2.5 Summary of the sequencing and assembly of the *H. midae* reference transcriptome.....60
Figure 2.6 Volcano plot displaying the $-\log_{10}$ of the P-values from Kal's statistical test in terms of different group means.....61
Figure 2.7 Categorization of *H.midae* contigs with significant BLAST hits (E-value cutoff $\leq 10^{-10}$) to the GO database, using three main categories.....62
Figure 2.8 Species distribution of top BLAST hits for annotation of the *H. midae* transcriptome.....63
Figure 2.9 Categorization of *H. midae* contigs with significant BLAST hits (E-value cutoff $\leq 10^{-10}$) to the KOG database, using four main categories.....65
Figure 2.10 Categorization of *H. midae* contigs with significant BLAST hits (E-value cutoff $\leq 10^{-10}$) to the KEGG database, using six main categories.....68
Figure 3.1 Early developmental stages of abalone (Shallow Seafarming Research Institute, 1990).....97
Figure 3.2 Embryo and larval stages of development in *H. midae*.....98
Figure 3.3 *H. midae* embryo cells in culture medium A (Table 3.4), four days after dissociation with abalone sperm (1×10^8 sperm/ml for 30 minutes) in a poly-D-lysine coated six-well tissue culture plate.....108
Figure 3.4 *H. midae* larval cells at day 4 of culture in culture medium C showing epithelial-like (E), fibroblast-like (F) and large round (R) morphologies.....110
Figure 3.5 Change in cell viability in cultured *H. midae* larval cells over twelve days.....111
Figure 3.6 Change in cell viability in cultured *H. midae* larval cells over nine days.....112
Figure 3.7 Change in cell viability in cultured *H. midae* larval cells over eight days.....112
Figure 3.8 Linear relationship of increasing absorbance with increase in *H. midae* larval cell number dertermined by XTT assay over seven days of culture.....113
Figure 3.9 Change in metabolic activity of cultured *H. midae* larval cells over ten days.....113
Figure 3.10 *H. midae* haemocytes in medium D (Table 3.4) attached to the surface of a six-well plate at day five of culture: A = Amoeboid-like cells, F = Fibroblast-like cells.....114
Figure 3.11 Linear relationship between *H. midae* haemocyte density and absorbances determined by XTT assay over nine days of culture for increasing cell densities.....115
Figure 3.12 Change in cell viability of cultured *H. midae* haemocytes (60 hours). Significant difference from control at $P < 0.05$ (*) and $P < 0.01$ (**).....116
Figure 3.13 Change in cell viability of cultured *H. midae* haemocytes (60 hours). Significant difference from control and D-1 and D-7 at $P < 0.01$ (**).....116

Figure 3.14 Summary of the effect of different concentrations of bovine insulin on *H. midae* haemocyte cell viability 117
Figure 4.1 Quantitative real-time PCR graph of *H. midae* cDNA, depicting the four phases; background, exponential, linear and plateau 131
Figure 4.2 Denaturing 2 % agarose gel of RNA isolated from haemocyte cell cultures 142
Figure 4.3 Second derivative amplification plots for all primers 147
Figure 4.4 2 % Agarose gels of PCR products amplified from cDNA of large (L) and small (S) animals using the initial five primer pairs 149
Figure 4.5 2 % Agarose gels of PCR products amplified from cDNA of cultured haemocytes using four subsequent primer pairs 149
Figure 4.6 Melt curve analysis following real-time PCR on triplicate samples 151
Figure 4.7 Standard curves for determining amplification efficiencies of all primers 155
Figure 4.8 Relative expression of genes 752, 2380 and 13596 160
Figure 4.9 Relative expression of genes 809, 54 and 3309 161
Figure 4.10a One route in the MAP-kinase signaling pathway 167
Figure 4.11a One route in the insulin signaling pathway 167

List of Abbreviations

ΔG	Free energy
A	Adenine
AB	Antibiotic
ADA	Adenosine deaminase
ADGF	Adenosine deaminase-related growth factor
AGSA	Atrial gland granule-specific antigen
ANOVA	Analysis of variance
ASW	Artificial seawater
ATP	Adenosine triphosphate
BLAST	Basic local alignment search tool
BMP	Bone morphogenic protein
bp	Base pairs
BrdU	5-bromo-2'-deoxyuridine
C	Cytosine CI Confidence Interval
CI	Confidence interval
CA	Carbonic anhydrase
C_b	Calibrator
CDD	Conserved domain databas
cDNA	Complementary DNA
CECR	Cat eye syndrome critical region
CGRP	Calcitonin gene related peptide
CLP	Chitinase-like protein
CMFSS	Calcium and magnesium free artificial seawater solution
C_n	Control
CoA	Coenzyme A
COG	Clusters of orthologous groups
cov	Coverage
C_q	Quantification cycle
CREB	cAMP response element-binding
CRP	Cysteine-rich polypeptide
dCAS	cDNA annotation software
DART-PCR	Data Analysis for Real-Time PCR
DD	Differential display
DMEM	Dulbecco's modified eagle's medium
DNA	Deoxyribonucleic acid
DNase	Deoxyribonuclease
dNTP	Deoxyribonucleotide triphosphate
DS	Dissociation solution
DTT	Di- thiothreitiol
E	PCR efficiency
EDTA	Ethylenediaminetetraacetic acid
EGF	Epidermal growth factor
EGFR	Epidermal growth facto receptor
EIF4E	Elongation initiation factor 4E
JAK2	Janus kinase 2
JNK	c-Jun N-terminal kinase
KAAS	Kegg automatic annotation server
KEGG	Kyoto encyclopedia of genes and genomes
KO	KEGG orthology
KOG	Eukaryotic orthologous groups
L	Large
L-EGRF	*Lymnaea stagnalis* EGFR
LGC	Light green cells
M	DNA marker
M	Average expression stability
M	Slope
MAP	Mitogen activated protein
MAPK	Mitogen activated protein kinase
MAS	Marker assisted selection
MDGF	Mollusc derived growth factor
MEM	Minimum essential medium
mGDF	Molluscan growth and differentiation factor
MIP	Molluscan insulin-related peptide
MIQE	Minimum information for publication of quantitative real-time PCR experiments
MKKK	MAP kinase kinase kinase
M-MLV	Moloney murine leukemia virus
Mnk	MAP kinase interacting serine/threonine kinase
MOPS	3-(N-morpholino)propanesulfonic acid
MPSS	Massively parallel signature sequencing
mRNA	Messenger RNA
MSTN	Myostatin
MTT	3-(4,5-Dimethylthiazol-2-yl)-2,5-diphenyltetrazolium bromide
N/C	Nucleus to cytoplasm
NCBI	National center for biotechnology information
NF-κB	Nuclear factor kappa-light-chain-enhancer of activated B cells
NG	Next generation
NLR	NOD-like receptor
NOD	Nucleotide-binding oligomerization domain
NRG	Neuregulin
NTC	No template control
ODC	Ornithine decarboxylase
P:C:I	Phenol:Chloroform:Isoamylalcohol
PCR	Polymerase chain reaction
PDGF	Platelet derived growth factor
PG	Proteglycan

eIF4E	Eukaryotic initiation factor 4E
ERK	Extracellular regulated kinase
ErbB	Receptor tyrosine kinases
EST	Expressed sequence tag
FASL	Fas ligand
FBS	Fetal bovine serum
FCS	Fetal calf serum
FDR	False discovery rate
FDD-RT-PCR	Fluorescent differential display RT-PCR
FGF	Fibroblast growth factor
FGFR	Fibroblast growth factor receptor
FSW	Filtered seawater
G	Guanine
G6PDH	Glucose-6-phosphate dehydrogenase
GA	Genome analyser
GAG	Glucosaminoglycan
GAPD	Glyceraldehyde-3-phosphate dehydrogenase
GDF	Growth and differentiation factor
GH	Growth hormone
GHR	Growth hormone receptor
GHRH	Growth hormone releasing hormone
GMEM	Glasgow MEM
GO	Gene ontology
GRP	Glucose-regulated protein
GST	Glutathione S-transferase
HBSS	Hanks balanced salt solution
Hdcols	*Haliotis discus* collagens
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hnRNA	Heterogenous nuclear RNA
HSP	Heat shock protein
IAP	Inhibitor of apoptosis
ICES	International council for the exploration of the sea
IDGF	Insect derived growth factor
IDT	Integrated DNA technologies
IGF	Insulin-like growth factor
IGFBP	Insulin-like growth factor binding protein
IGFR	Insulin-like growth factor receptor
IL-1	Interleukin-1
IU	International unit
PMS	N-methyl dibenzopyrazine methyl sulfate
PS	Penicillin-streptomycin
qPCR	Quantitative real-time PCR
QTL	Quantitative trait loci
R	Reference
r^2	Coefficient of determination
REST	Relative expression software tool
RNA	Ribonucleic acid
RNase	Ribonuclease
RPMI	Roswell park memorial institute
rRNA	Ribosomal RNA
RT	Room temperature
-RT	Minus reverse transcriptase
RT-PCR	Reverse transcriptase PCR
S	Small
SAGE	Serial analysis of gene expression
SBS	Sequencing-by-synthesis
SDS	Sodium dodecyl sulfate
SFK	Src family kinase
SNP	Single nucleotide polymorphism
SSIII	Superscript III
Stats	Signal transducers and transcription activators
T	Thymine
TB	Trypan blue
TBE	Tris-borate-EDTA
TGF	Transforming growth factor
TGF- α	Transforming growth factor-α
TGF-B	Transforming growth factor beta
T_m	Melting temperature
TNF	Tumor necrosis factor
Topo	Topoisomerase
TRAMP	Tyrosine-rich acidic matrix protein
TSGF	Tsetse salivary growth factor
TSP1	Thrombospondin-1 precursor
UBQ	Ubiquitin
UV	Ultraviolet
VEGF	Vascular endothelial growth factor
XTT	Sodium 3´-[1- (phenylaminocarbonyl)- 3,4-tetrazolium]-bis (4-methoxy-6-nitro) benzene sulfonic acid hydrate

1 LITERATURE REVIEW, BACKGROUND AND AIM

1.1 *Haliotis midae*

Abalone (*Haliotis*) belong to the phylum Mollusca which is, after Arthropoda, the second largest phylum in the animal kingdom. The Mollusca comprise of between 50 000 and 200 000 living species and 35 000 fossil species. It is a widespread phylum, with species present in marine, freshwater and terrestrial environments. Species include chitons, snails, abalone, oysters and octopuses, amongst others (Hickman and Roberts, 1994; Bourquin, 2009; Bunje, 2010). Table 1.1 presents the taxonomic classification of *Haliotis.*

Table 1.1 Taxonomic classification of *Haliotis* (The Uniprot Consortium, 2010)

Phylum	Mollusca
Class	Gastropoda
Subclass	Orthogastropoda
Superorder	Vetigastropoda
Family	Haliotidae
Genus	*Haliotis*

Haliotids belong to the order Vetigastropoda, which is the oldest and most "primitive" group of gastropods (Latin: "stomach foot") (Purchon, 1977; Muller, 1986; Bourquin, 2009). There are six haliotid species that occur in Southern African waters, namely *Haliotis midae* (Linnaeus*), H. parva* (Linnaeus)*, H. pustulata* (Reeve), *H. queketti* (Smith), *H. spadicea* (Donovan) and *H. speciosa* (Reeve) (Muller, 1986; Hecht, 1994, Geiger, 2000). *Haliotis midae*, known locally as 'perlemoen', occurs along the Western and Eastern Cape shores of South Africa, and is the only abalone species with importance to aquaculture in South Africa. The other five species of abalone are relatively small and not harvested commercially (Henry, 1995).

Strict conservation measures were implemented since 1965 to prevent overfishing of *H. midae* (Genade *et al*., 1988). In that year, the highest abalone harvest ever was reported at an annual catch of 2800 tonnes. In 1968 a maximum production quota of 386 tons was imposed and this was reduced to 227 tons in 1970 (Tarr, 1992). Due to continued concern over the state of the resource, the production quota was reduced to 181 ton in 1971. From 1979 to 1982, it was even further reduced by 10 percent to 163 tons. After this, the control system was changed to a whole mass quota and continuous efforts were made to manage this resource (Tarr, 1992). Years of uncontrolled commercial fishing and poaching however brought the South African abalone, *H. midae*, to the brink of extinction. In February 2008 a complete ban on abalone fishing was issued by the Department of environmental affairs and

tourism of the South African government (Department Environmental Affairs and Tourism, 2008). The conditional lifting of this ban was however approved by cabinet in June 2010 (GCIS, 2010).

1.1.1 Biology

1.1.1.1 *Anatomy*

All haliotids, including *H. midae*, are large, herbivorous, marine gastropods with a depressed shell, enlarged body whorl and reduced spire near the back of the shell. The round or ear-shaped shell is characteristically perforated by a line of small respiratory pores located along the left margin of the shell. The older pores close successively as growth proceeds (Figure 1.1) (Muller, 1986; Genade *et al*., 1988; Hahn, 1989). The flat shell, which reduces resistance to waves, and the wide shell-mouth, which enables the animal to attach firmly to the substratum, reflects adaptation of *Haliotis* to conditions of strong wave action (The South African Institute for Aquatic Biodiversity, 2004).

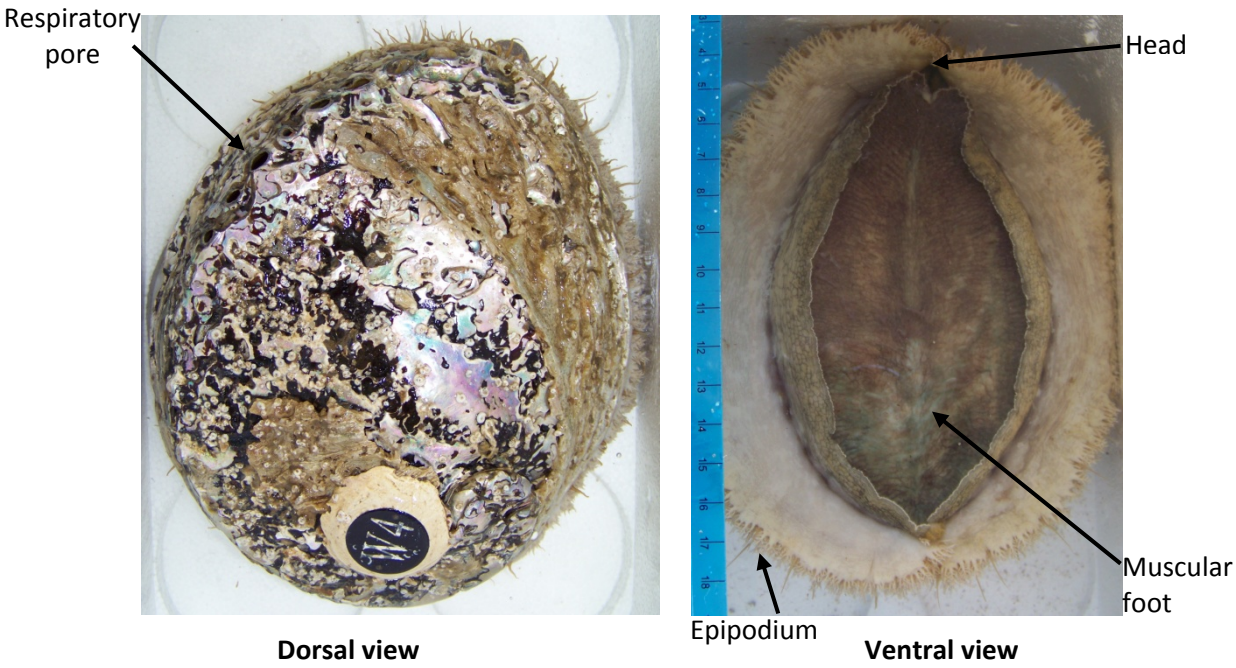


Figure 1.1 Dorsal and ventral views of *H. midae* (A. Roux, 2008). The respiratory pores are visible on the left side of the shell in the dorsal view. The ventral view shows the head, muscular foot and epipodia.

Underneath the shell lie the anterior head, a large muscular foot and the soft body that is attached to the shell by a column of shell muscles (adductor muscle). The muscular foot is encircled by the mantle as well as the epipodium – a sensory structure bearing the tentacles (Figure 1.1). The epipodium, which projects beyond the shell edge has a smooth or pebbly surface with a frilly or scalloped edge and is a reliable structure for identifying abalone species (Fishtech Inc., 2010). The foot is the edible part of the animal and can account for more than a third of the animal's weight. It is used by the animal to attach tightly to rocky surfaces by suction (Department of Fisheries, Government of Western Australia, 2005). The organs arranged around the foot and under the shell comprise of a pair of eyes, a mouth with a

long tongue called the radula, an enlarged pair of tentacles and the crescent-shaped gonad. Next to the mouth and under the respiratory pores is the pallial cavity where water is drawn in under the edge of the shell and flows over the gills and out the pores, carrying waste and reproductive products out in the exhalant water (Fishtech Inc., 2010). The abalone has a heart on its left side and blood, called hemolymph, flows through the arteries, veins and sinuses (open circulatory system). The central nervous system lacks concentration of ganglia into complex organs, although distinctive ganglia do occur in the head (Hahn, 1992). Because it has no obvious organized brain structure, the abalone is considered a “primitive” animal (Fishtech Inc., 2010).

Abalone are gonochoric animals and have a single gonad, either ovary or testis, enveloping the digestive gland, which forms the bulk of the visceral mass (Newmann, 1967; Purchon, 1977; Henry, 1995). The gonad constitutes 15 to 20 percent of the soft body mass during the breeding season and remains this size until spawning, after which it rapidly decreases in size (Hahn, 1989; Henry, 1995). The combined structure of gut enveloped by gonad is called the conical appendage (Hahn, 1989; Fallu, 1991; Henry, 1995; Hooker and Creese, 1995). This structure is developed extensively to the right side of the body and around the right posterior margin of the adductor muscle (Henry, 1995). The gonad consists of a large lumen, bounded by germinal epithelium with a connective tissue base, which is well supplied with blood vessels (Newmann, 1967). Figure 1.2 depicts the various organs and other soft body parts of the abalone.

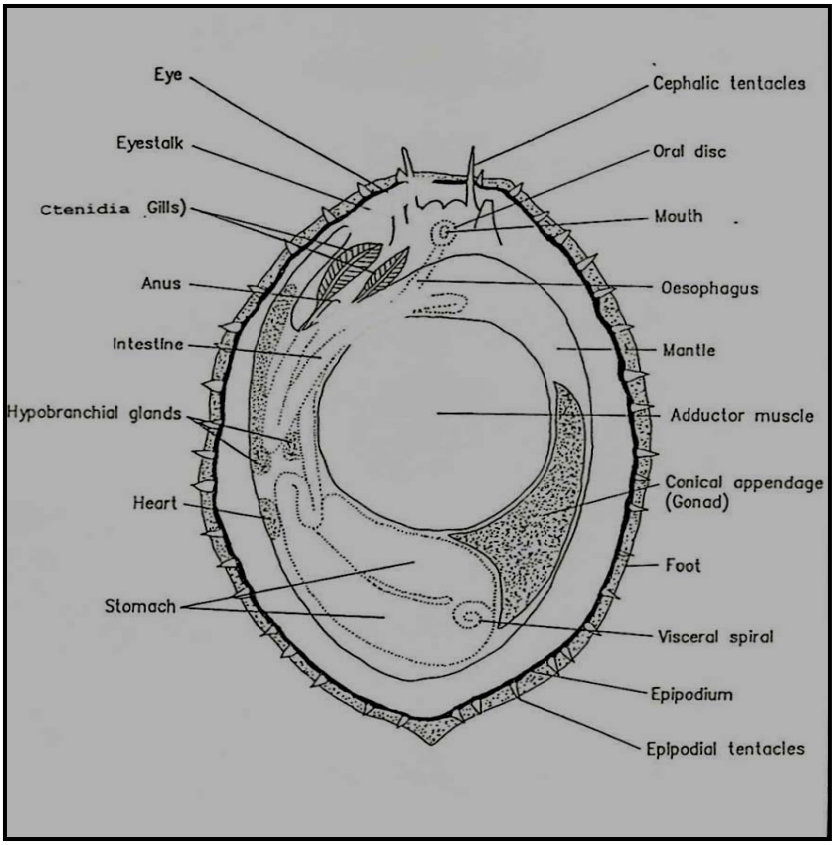


Figure 1.2 Ventral view of organs and soft body parts of the abalone (Henry, 1995)

H. midae is the largest of the South African abalone species and have a grayish shell colour, usually covered with epibiota (other, smaller organisms attaching to the shell). The foot is pale cream to mottled light brown colour and tentacles and gills are yellow to gray (Muller, 1986). The gonad of the female is green and that of the male cream coloured when ripe. *H. midae* can reach a size of 90 mm shell length in six years and a maximum size of about 200 mm shell length at an age of over 30 years in the wild (Hahn, 1989; Sales and Britz, 2001).

1.1.1.2 *Development*

Embryo and early larval stages

The developmental process from fertilization to settlement takes four to ten days, depending on the species and water temperature (Hahn, 1989; Henry, 1995). The time that the larvae spend in a floating state is called the pelagic/planktonic stage. During this time, the larvae undergo successive changes from trochophore to veliger stages. An investigation by Genade *et al*. (1988) confirmed that the planktonic larval stage of *H. midae* is more or less within the time confines of that of other abalone species.

Larval to post larval stages

Settlement, metamorphosis and deposition of the peristomal shell characterize the transition from larval to post-larval development (Hahn, 1989). Settlement occurs at a week to a month after the veliger stage depending on the species and conditions. For *H. midae*, settlement occurs about five days (at 20 ˚C) or seven days (at 17.5 ˚C) after fertilization (Genade *et al*., 1988). This is when the larvae sink to the bottom and start crawling in search of a suitable substratum. Crawling continues until the larvae attach to the substratum. This is followed by metamorphosis, which is characterized by development of the mouth and radula, digestive tract, circulatory system with a beating heart, sensory organs and adult form (Hahn, 1989). Larvae are now called "spat" and feed on micro-algae (Fallu, 1991). Twenty-four hours after metamorphosis they start feeding on benthic diatoms and this will remain their principal food source until individuals are 7 to 10 mm in length, when they change to a diet of macro-algae (Hahn, 1989; Henry, 1995).

Juvenile stage to sexual maturity

The post larval period continues until formation of the first respiratory pore, which announces the notch stage. This stage occurs at an age of about one to three months (Hahn, 1989). *Haliotis midae* reach the notch stage at a size of 2.1 - 2.2 mm and formation of the first respiratory pore is completed at 2.3 mm (Genade *et al*., 1988). Growth rates of juveniles sharply increase after the notch stage is reached as this is also when weaning begins and the abalone starts feeding on macro-algae (seaweed) (Hahn, 1989; Fallu, 1991). Juveniles of about 10 mm in length will consume 10 - 30 percent of their whole wet body weight in macro-algae each day. The abalone will slowly increase in size until sexual maturity is reached (at around 80 - 105 mm shell length in *H. midae*) and beyond (Barkai and Griffiths, 1988; Henry, 1995; Tarr, 1995).

1.1.1.3 *Feeding, metabolism and growth*

Haliotis midae are strictly herbivorous gastropods. The main source of energy of the adult abalone is kelp (*Ecklonia maxima*) which is ingested from late afternoon to early morning (Barkai and Griffiths, 1988). Large, mature individuals usually aggregate on an outcrop of reef, extending from 0.5 to 2 m above the seabed, facing the incoming swell in the midst of dense kelp forests. Food availability in such aggregations is probably enhanced because of individuals trapping large drift-kelp fronds (Tarr, 1995). Feed intake (wet weight) in wild *H. midae* is estimated at 8.1 % of soft body weight per day at 14 °C and 11.4 % at 19 °C (Sales and Britz, 2001) and the absorption efficiency of *H. midae* feeding on a natural diet of kelp is estimated at 37.25 % (Barkai and Griffiths, 1988). Further studies on *H. midae* also reported that about 63 % of the energy consumed in food is lost as faeces in wild animals and a further 32 % is expended on respiration. It is suggested that some energy may also be used for mucus

production during locomotion (Barkai and Griffiths, 1988; Farias *et al*., 2003). This leaves about 5 % of energy intake available for growth and reproductive output with an increasing proportion of this energy utilized for reproduction in older animals. The assumption is made that energy expended on reproduction is similar for both male and female abalone (Barkai and Griffiths, 1988).

During the reproductive cycle, gametogenesis takes place. The production of gametes requires a large amount of nutrients for metabolic requirements and synthesis of vitellogenin, which serves as fuel for larval development (Hahn, 1989). The foot and the digestive gland are indicated as sources of metabolic energy for gametogenesis. During gamete development, the size of the foot decreases and glycogen levels drop significantly. It is proposed that glycogen is converted to lipids and transferred to the ovary where it is incorporated in vitellogenesis. Additionally lipids are supplied to the ovaries by the digestive gland (Hahn, 1989). Consequently, more metabolic energy is diverted towards gametogenesis and less energy and resources are available for somatic growth of the animal during reproduction (Boudry *et al*., 1998; Garnier-Géré *et al*., 2002). As the abalone reaches sexual maturity, somatic growth is thus significantly reduced (Yang *et al*., 1998).

1.1.2 As aquaculture species

The South African abalone fishery has been in existence since 1949, but the first attempts to cultivate *H. midae* were made in 1981 when captured specimens were successfully spawned to produce spat and juveniles (Genade *et al*., 1988; Sales and Britz, 2001). *Haliotis midae* was however only confirmed as a suitable marine species for aquaculture in 1990. Since then, concerted research and development efforts towards the establishment of commercial abalone farming progressed (Henry, 1995; Sales and Britz, 2001). In the past two decades, 18 abalone farms have been established, ranging in distribution from Port Nolloth on the Atlantic/West Coast to East London on the Indian/East Coast. Abalone farming currently contributes a commercial production of 934 tons with a production value of R268 million to the South African aquaculture industry. This makes abalone aquaculture the largest contributing sector, representing 81 % of the total rand value of the local aquaculture sector (Britz *et al*., 2009).

Abalone have a very slow growth rate, typically two to three centimeters per year. At this rate, two to five years is required for an abalone to reach market size (Hahn, 1989). Like most other commercially important abalone species, the slow growth rate of *H. midae* is an obstacle in the profitable farming and global competitiveness of this species (Stepto, 1997; Elliott, 2000). Continuous research efforts focusing on optimal culture conditions, nutrition and genetic improvement in abalone are implemented to address this problem of slow growth.

1.2 Growth as a desirable trait in animal husbandry

In *H. midae*, as in most farmed species, understanding physiological traits that are of economic value, like growth rate, disease resistance and meat quality is of utmost importance for the market competency and profitability of the product. In terrestrial livestock production, genetic information has been harnessed with the aim of improving production traits. Initial livestock improvement practices were mainly based on selection and resulted in increased productivity in some livestock species, such as a six times weight gain in broiler chickens (Havenstein *et al*., 2003). Abalone aquaculture researchers have also reported that genetic improvement of desirable traits (like weight and length) can be achieved through selective breeding and that abalone breeding programs can deliver substantial economic benefits to the abalone aquaculture industry (Li, 2008). Selection practices however, do not account for all observed patterns of genetic variation since for many traits allelic variation at individual genes may disproportionately influence overall phenotypic expression (De Santis and Jerry, 2007).

Quantitative traits are those that represent phenotypic characteristics that are usually controlled by many genes at one or more loci. A few studies of quantitative traits in aquaculture species have been conducted (Abalone: Hayes *et al*., 2007; Trout: Haidle *et al*., 2008; Salmon: Houston *et al*., 2008). The International Council for the Exploration of the Sea (ICES, 2008) made a comprehensive summary of quantitative trait loci (QTL) studies in aquaculture species. Particular traits that are relevant for production are of interest in aquaculture. These include growth rate, body weight at marketing, feed conversion efficiency, disease resistance, flesh quality and age at maturation (Beaumont and Hoare, 2003; ICES, 2008). Growth rate is the primary trait of interest that is intrinsically linked to productivity and profitability of aquaculture enterprises.

In recent decades, the genomic revolution allowed researchers to acknowledge the contribution of candidate genes – genes of which the physiological function is known and its direct effect on the expression of a trait is quantifiable (De Santis and Jerry, 2007). Candidate genes are often targeted in growth variation studies based on prior knowledge of their specific regulatory roles in metabolic pathways influencing growth rate.

If prior information in the species being studied is not available, comparative genomic approaches are followed whereby sequence information from other already sequenced and identified genomes can be utilized to attain best-match information and identify similar genes in the species of interest (De Santis and Jerry, 2007). Such genes that are functionally conserved in different species and that have branched from a common ancestor by speciation are called orthologous genes (Moriya *et al.,* 2007). Such genes can be sourced from collections of identified and annotated genomes and their associated genes, proteins and metabolic pathways, that are available in public databases. KEGG (Kyoto

Encyclopedia of Genes and Genomes), GO (Gene Ontology database), KOG (EuKaryotic Orthologous Groups), NCBI GenBank (a genetic sequence database), NCBI RefSeq (a curated non-redundant collection of genome, transcript and protein sequences) and Uniprot (a comprehensive resource for protein sequences and functional information) are all examples of databases where large collections of nucleotide and protein sequences are made available for public use.

Several candidate genes influencing growth in livestock and finfish have already been isolated and their effects quantified. Amongst these, many play roles in somatotropic governing pathways that are similar in livestock and finfish. Somatogenesis is a polygenic trait that regulates energy metabolism and muscle growth, making genes within the somatotropic axis and transforming growth factor superfamily the most targeted candidate genes in livestock and finfish (De Santis and Jerry, 2007). Similar genes could be potential candidate genes in molluscs. No investigation has been made towards identification of specific growth-related genes associated with growth and observed size differentiation in *H. midae.*

Already in the 1970s however the idea of molluscs having endocrine factors involved in growth control emerged, when it was shown that a factor produced by the cerebral ganglion of the prosobranch, *Crepidula fornicata*, stimulated somatic growth (Lubet, 1971). In 1976, studies on the gastropod *Lymnaea stagnalis* followed suit with evidence of growth factors from the neurosecretory ganglion that stimulate tissue growth and shell formation (Geraerts, 1976).

Cytoskeleton structure, myogenesis and shell growth are other physiological characteristics that are also closely linked with commercially important traits like meat quality and growth. Since all farmed mollusc species, including *H. midae,* have external shells, growth is an intricate coordination between somatic/soft body growth and shell growth. Shell formation involves organic matrix components (proteins, glycoproteins, lipids, chitin and acidic polysaccharides) driving crystal nucleation. These components control the growth and spatial arrangement of minerals during calcium carbonate polymorph formation (Wilbur and Saleuddin, 1983; Falini *et al*., 1996; Levi-Kalisman *et al*., 2001). Proteins and peptides of the shell organic matrix have only been studied in a few abalone species e.g. in *H. asinina* (Jackson *et al.*, 2007), *H. laevigata* (Weiss *et al.*, 2001), and *H. tuberculata* (Jolly *et al*., 2004) compared to the numerous data available on the pearl oyster, *Pinctada fucata* (Samata *et al*., 1999; Zhang *et al*., 2003; Takeuchi and Endo, 2006).

Somatic growth and shell growth must be regulated in a parallel fashion and occur simultaneously in order to ensure the correct functional relationship for the animal (Duvail *et al.*, 1998). Several genes and peptides have been suggested to play roles in such coordinated growth regulation. Examples of such genes in abalone include actin, identified in several abalone species (Bryant *et al.,* 2006; Sin *et al.*, 2007); tropomyosin identified in *Haliotis rufescens* (Degnan *et al.*, 1997); proteoglycan and collagen

identified in *H. tuberculata* (Poncet *et al.*, 2000) and the transcription factor genes Pou, Sox and Pax identified in *H. asinina* (O'Brien and Degnan, 2000). These genes and others that are considered as important in growth regulation of molluscs will be discussed in more detail below and are summarized in Table 1.2 of the Appendix. Genes from this table may be used as possible orthologous genes when conducting an investigation towards identification of genes associated with growth and observed size differentiation in *H. midae.*

1.2.1 Genes within the somatotropic axis and central nervous system

Growth hormone and associated growth factors

A review by De Santis and Jerry (2007), focused on the identification of candidate growth genes for finfish, by using available information on genes influencing growth rate in terrestrial vertebrates. Since physiological growth-pathways are fundamentally conserved among vertebrates, it is suggested that researchers focus on the GH-IGF-I cascade and myogenic transforming growth factors as candidates for initial investigations. Genes encoding for somatotropic axis hormones already identified for terrestrial livestock include Growth hormone (GH) gene, Growth hormone receptor (GHR) gene, Insulin-like growth factor I (IGF-I) gene, Growth hormone-releasing hormone (GHRH) and Leptin. In vertebrates, insulins and insulin-related peptides form a superfamily of regulatory peptides that control growth, metabolism, reproduction and development (Smit *et al.*, 1991; De Santis and Jerry, 2007).

In mammals, the growth-promoting action of GH is mediated through mainly hepatic synthesis and secretion of IGF-1, which stimulates maturation and cell division of chondrocytes in the epiphyseal plates of long bones. Growth hormone (GH) and its associated factors of the somatotropic axis have been studied extensively and it has been shown to perform various functions in addition to growth promotion. These functions include the regulation of lipid and carbohydrate metabolism, normal reproductive function, beneficial physiological actions on cardiac and immune function, modulation of gut function and enhanced uptake of macro- and micronutrients. GH can also act as an agent of neural repair and promotes neural stem cell proliferation (Lanning and Carter-Su, 2006; Lichanska and Waters, 2007).

Mammalian GH is secreted by the anterior pituitary and exerts its effects on target tissues expressing GHR, a cytokine receptor for GH. Upon binding, the tyrosine kinase Janus kinase 2 (JAK2) and to some extent a Src Family kinase (SFK) are activated by phosphorylation and various signaling cascades are activated, resulting in a variety of biological responses including cellular proliferation, differentiation and migration, prevention of apoptosis and regulation of metabolic pathways. Some signaling proteins involved in pathways activated by GH that have been identified include Stats (signal transducers and transcription activators), MAPKs (Mitogen activated protein kinases) and ERKs (extracellular-regulated

kinases). Other putative GH-responsive pathways that may be responsible for elevating cytosolic Ca^{2+} in response to GH include phospholipase Cg- and protein kinase C-pathways (Lanning and Carter-Su, 2006; Lichanska and Waters, 2007). Signaling events initiated by GHR activation when GH binds to it are diverse, complex and intricately related and the mechanisms by which downstream responses are activated in different cell types and tissues are continually being elucidated (Lanning and Carter-Su, 2006).

Although not much progress has been made in identifying growth hormone in molluscs, a conserved function of vertebrate growth hormones in terms of physiological responses in molluscs has been suggested (Lucas, 2007). This follows the observation of a few studies that investigated the effect of vertebrate growth factors on mollusc growth. After exposure to recombinant salmon growth hormone, by immersion and intramuscular injection, accelerated growth rate was observed in *Haliotis discus hannai* (Moriyama and Kawauchi, 2004). Also for the Eastern oyster, *Crassostrea virginica*, immersion of juveniles in biosynthetic rainbow trout growth hormone resulted in increased growth (Paynter and Chen, 1991). Bovine growth hormone has also been shown to result in increased growth rate in young *Haliotis rufescens* postlarvae (Morse, 1984). However, following a comprehensive search for genes involved in growth-control in *H. asinina,* no haliotid homologs of vertebrate growth hormone has been reported to date (Lucas, 2007).

Insulin-like factors (IGFs, IGFBPs and MIPs)

Factors that constitute the somatotropic axis appear to be highly conserved in vertebrates and evidence for similar expression patterns and functions in fish have been reported (Li *et al.*, 2006). Components and mechanisms concerned with the IGF signaling pathway are evolutionarily conserved amongst vertebrates and invertebrates and research on the interaction between these mechanisms and the aetiology of human age-related diseases has been promoted greatly by work with invertebrates. Model invertebrate organisms like the nematode worm, *Caenorhabditis elegans* and the fruit-fly *Drosophila melanogaster* often provide simpler study material because of their short lifespan, ease of handling and genetic accessibility and consequently, they are used to make more rapid progress in elucidating molecular mechanisms underlying the somatotropic axis. To date 39 *C. elegans* and seven *Drosophila* insulin-like genes have been identified (Piper *et al.*, 2008). In invertebrates, insulin and insulin-related molecules are secreted by neuroendocrine cells of the central nervous system, instead of cells associated with the intestinal tract, as is the case in vertebrates (Smit *et al.*, 1991).

The freshwater gastropod *Lymnaea stagnalis* and marine gastropod *Aplysia californica* are two molluscs that are attractive models for neurobiological research and have as a result also contributed to the identification of insulin-like growth related peptides. As both of these molluscs' genomes are being

sequenced at present they, together with the gastropod *Lottia gigantea* whose genome has already been sequenced (U.S. Department of Energy Joint Genome Institute, 2007), pose the best model organisms for mollusc genome and transcriptome research to date.

The Light Green Cells (LGC) are neuroendocrine cells present as two paired clusters totaling approximately 150 neurons in both cerebral ganglia of *L. stagnalis.* These cells produce hormones, collectively called molluscan insulin-related peptides (MIPs), that control somatic growth, shell formation and protein and glycogen metabolism. The peripheries of the median lip nerves are used as the storage and release site of the peptides derived from the MIP precursors (Dogterom, 1980; Li *et al*., 1992; Smit *et al*., 1991). The MIPs, which resemble vertebrate insulin-related peptides, were first shown to stimulate somatic growth and enlargement of the shell in *L. stagnalis* by Dogterom (1980). Seven MIP peptides and their coding genes have since been isolated and characterized and MIPs I and II has been described as possibly being the most complexly folded molecules of the insulin superfamily. A scheme similar to the processing of preproinsulin in the B-cells of pancreatic islets and related tissues in vertebrates has also been proposed for the processing of the MIP II precursor in the neuroendocrine LGC system of *Lymnaea* (Li *et al*., 1992; Roovers *et al*., 1995).

In the initial study by Dogterom (1980), it was also concluded that the size of the shell plays a determining role in the extent of somatic growth that occurs and that the size of the animal is ultimately determined by the effect of the MIPs on the mantle edge. Since the characterization of MIPs in *Lymnaea*, insulin-like peptides have been identified in the nervous ganglia neurosecretory cells in various molluscan species, including the gastropods *Aplysia californica, Helisoma duryi, Planorbarius corneus, Otala lactea* and *Helix aspersa* and the bivalve *Mytilus edulis.* The same substances have also specifically been localized in the digestive tract, an organ undergoing continuous regeneration, of *M. edulis, H. duryi, O. lactea* and *H. aspersa,* amongst others, and the haemolymph of *H. duryi* and *O. lactea*. Similar insulin-like peptides involved in growth regulation by stimulating protein synthesis in mantle edge cells involved in shell and soft tissue growth is reported for the oyster *Crassostrea gigas* (Gricourt *et al.*, 2003). This growth-promoting effect is mediated by specific receptors belonging to class II of the receptor protein-tyrosine kinase family. These receptors that are similar to insulin-like receptor sequences in mammals, specifically in the tyrosine-kinase catalytic domain, are expressed in the mantle edge cells, labial palps and gonad, and contribute to the suggestion that a high level of conservation of this receptor family is maintained during evolution (Gricourt *et al.*, 2003).

Adenosine deaminase-related growth factors

Adenosine deaminase-related growth factors (ADGFs; known as CECR1 in vertebrates) belong to the adenosine deaminase (ADA) subfamily. They have mitogenic properties and possess adenosine

deaminase enzymatic activity in invertebrates. Adenosine deaminase (ADA) is an enzyme involved in the depletion of adenosine and deoxyadenosine levels by catalyzing the irreversible, hydrolytic deamination of adenosine and 2-deoxyadenosine to inosine and 2-deoxyinosine. It thus prevents the accumulation of the toxic, sometimes lethal levels of these substrates in the blood or haemolymph (of *Drosophila* and other invertebrates). The growth stimulating properties of ADGFs in invertebrates are suggested to be due to the destruction of adenosine by ADA activity, thereby protecting growing tissues, or due to the unique amino-terminal region of this gene family that could have a classic growth hormone function by binding to a yet unknown receptor. The tissues where the highest ADA enzymatic activity was observed in vertebrates and invertebrates studied are the gut and lymphoid organs and/or blood cells (Maier *et al.,* 2001; Akalal *et al.*, 2003; Dolezelova *et al.*, 2005; Maier *et al.,* 2005).

The initial member that was described in invertebrates for this subfamily of secreted growth factors was IDGF (insect-derived growth factor) in the flesh fly, *Sarcophaga peregrina* (Homma *et al.*, 1996). Since then, members that are similar to IDGF have been described in other invertebrates, including MDGF (mollusc derived growth factor) in *Aplysia californica* (Akalal and Nagle, 2001); TSGF-1 and -2 (tsetse salivary growth factor) in the tsetse fly, *Glossina morsitans morsitans* (Li and Aksoy, 2000); LuloADA (*L. longipalpis* salivary gland ADA) in the sand fly, *Lutzomyia longipalpis* (Charlab *et al*., 2000, 2001); salivary ADA in the mosquito, *Aedes aegypti* (Valenzuela *et al.,* 2002) and six *Drosophila* ADGF homologues (Maier *et al.*, 2001). In humans CECR1, a candidate gene for the rare duplication disorder of cat eye syndrome, show significant amino acid similarity to invertebrate ADGF. This, together with additional vertebrate homologues (CECR1 in pig, cow, frog and zebrafish) suggests that ADGFs are a widely dispersed growth factor gene family of which catalytic residues involved in ADA activity are conserved (Maier *et al.*, 2001; Akalal *et al.*, 2004; Maier *et al.*, 2005).

MDGF, characterized in *Aplysia*, has ADA activity and stimulates cell proliferation *in vitro*. Initially, MDGF was called AGSA (atrial gland granule-specific antigen) until the significant homology to IDGF was observed after characterizing atrial gland cDNA for AGSA. In *Aplysia*, MDGF protein is expressed in the atrial gland and reproductive tract and also transiently in the central nervous system during development. This supports the suggestion that MDGF may play a growth factor role during periods of cell proliferation, including neuronal reorganization and restoration after injury (Akalal and Nagle, 2001; Akalal *et al.*, 2003).

The stimulation of embryonic fly NIH-Sape-4 cell proliferation *in vitro* by addition of 1 ng/ml MDGF agrees with similar cell proliferation observed with IDGF and *Drosophila* ADGFs at the same concentration. Broad cross-species reactivity of the conserved catalytic domain has also been illustrated as *Aplysia* MDGF and calf ADA both stimulate NIH-Sape-4 cell proliferation (Akalal *et al.*, 2003, 2004).

EGF and granulin

Another growth factor that has been studied in molluscs and other invertebrates is epidermal growth factor (EGF). In vertebrates EGF belongs to a peptide family consisting of ten known members: EGF; transforming growth factor-α (TGF- α); heparin-binding EGF; epiregulin; betacellulin; amphiregulin; and four neuregulins (NRG1–4). These peptides interact with one of four receptor tyrosine kinases called ErbB1–4. Upon activation, ErbB receptor signaling supports various functions during nervous system development. In invertebrates, EGF-like peptides and receptors have been identified in *Drosophila* and *C. elegans*. Recent experimental research on *Lymnaea stagnalis* supports the hypothesis that EGF ligands in invertebrates have functions similar to those observed for vertebrate neurotrophins, factors that induce cellular signaling pathways underlying neurite outgrowth and synapse formation. The authors also suggest that receptor binding and activation are conserved between EGF family members based on the observation that human EGF and TGF-α promoted synapse formation through *L. stagnalis* EGF-receptor (L-EGFR) (Van Kesteren *et al.*, 2008). Human EGF was also shown to stimulate protein and DNA synthesis of abalone (*Haliotis tuberculata*) haemocytes *in vitro* (Lebel *et al.*, 1996).

Granulin/epithelin-like peptides are putative growth modulating factors that show qualitative similarities to the physiological activities of EGF and TGF-α, namely functions related to cell proliferation and growth. Several granulin/epithelin-like peptides have been identified in invertebrates including a molluscan granulin/epithelin-like peptide, CRP1, which acts as a growth-modulating factor in the Japanese scallop *Patinopecten yessoensis* (Nara *et al*., 2004).

Developmental genes and transcription factors

Various genes and transcription factors associated with development of the molluscan central nervous system have been identified. Proteins encoded by homeobox genes (like Hox cluster genes, Mox, POU, Sox and Pax) comprise a large family of conserved transcription factors that control development and spatial patterning mechanisms in vertebrates and invertebrates (Degnan *et al*., 1997; Giusti *et al*., 2000). Amongst haliotids, Mox homologues with great sequence identity to evolutionary distant taxa have been described for *Haliotis rubra* larvae (Hru-Mox) (Degnan *et al.*, 1997) and for *H. asinina* (Has-Mox) where it is expressed in larval mesodermal cells destined to form adult foot muscle (Hinman and Degnan, 2002). Members of POU, Sox, and Pax multigene transcription factor families that are expressed in the adult brain ganglia have been described for *H. asinina* (O'Brien and Degnan, 2000). Apart from its neuro-developmental functions, members of the POU family are also involved in the activation of growth genes (Managalam *et al*., 1989) and in the regulation of neurotransmitter synthesis (O'Brien and Degnan, 2000). Hox homologues have altogether been identified in seven molluscan classes (Iijima *et al.*, 2006). The expression of the SoxB gene (PvuSoxB) was characterized when, during early development it was observed in neurogenesis regions in the mollusc *Patella vulgata* (Le Gouar *et*

al., 2004). Although it has been suggested that the expression of POU, Sox and Pax genes in the adult brain may have a role in regulating the expression of genes directly responsible for controlling growth and reproduction in the abalone, it still needs to be elucidated how exactly the genes of this diverse gene family play their developmental and physiological roles (O'Brien and Degnan, 2000).

β-thymosins

Other genes that are possibly involved in growth regulation are those transcribing β-thymosin. These bioactive proteins were recently isolated in *Aplysia californica* and are believed to have a diverse range of intracellular and extracellular physiological functions depending on their site of expression. Proteins homologous to vertebrate β-thymosins are secreted from whole ganglia in the *Aplysia* central nervous system and have been shown to positively affect neurite extension and branching in cultured *Aplysia* neurons. Together with the observation that mRNA encoding *Aplysia* β-thymosin is over-expressed during nerve injury, it is suggested that this protein may be involved in the regulation of structural plasticity in the central nervous system of *Aplysia* (Romanova *et al*., 2006).

1.2.2 Genes from muscle tissue and haemolymph

Different genes are involved in muscular activity and since most vertebrate muscle genes and proteins have invertebrate homologs (Hooper and Thuma, 2005), databases with annotated vertebrate muscle growth genes are of great value in the search for candidate growth genes in molluscs. Muscle proteins that have been identified in molluscs include the thin filament proteins actin, tropomyosin and troponin and calponin/caldesmon. The thick filament proteins include myosin heavy chain, catchin/myorod, coded by alternative splice of the myosin heavy chain gene, myosin light chains, kinases that phosphorylate and thus modulate myosin, paramyosins/miniparamyosins, filagenins and myonin (Hooper and Thuma, 2005).

Actin

Amongst the thin filament proteins identified in molluscs, the most completely identified and covered in literature is actin. Actin, a ubiquitous conserved protein found in eukaryotic cells, is abundant in muscle cells as a major component of myofibrils involved in muscle contraction and in non-muscle cells as microfilaments involved in cytoskeleton morphology, cell division, organelle movements, cell motility, locomotion, phagocytosis, endocytosis and exocytosis (Hennessey *et al.*, 1993; Gomez *et al.*, 1999; Wang *et al.*, 2008). The functional importance of actin is evident in the high level of sequence conservation displayed by actin, which implies evolutionary constraint (Hennessey *et al.*, 1993).

Actins are encoded by a multigene family with a variable number of members in different species: six genes in humans and *Drosophila melanogaster*; nine genes in the pufferfish, *Fugu rubripes* and eight

genes in the abalone, *Haliotis iris* (Gomez-Chiarri *et al.*, 1999; Hooper and Thuma, 2005; Bryant *et al*., 2006). Other molluscs in which actin genes have been identified include *Aplysia californica* (Des Groseillers *et al*., 1990, 1994), the Pacific oyster, *Crassostrea gigas* (Cadoret *et al.,* 1999; Miyamoto *et al*., 2002), the scallop, *Placopecten magellanicus* (Patwary *et al*., 1996), the limpet, *Patella vulgata* (van Loon *et al*., 1993), the red abalone, *Haliotis rufescens* (Gomez-Chiarri *et al.*, 1999) and the European abalone, *Haliotis tuberculata* (Kissinger, 1998). It has been suggested that the actin protein sequence may be used as an indicator of meat quality in commercially valuable molluscs (Bryant *et al*., 2006).

Being a multigene family, actin is able to perform a wide range of specialized biological functions. Multiplicity allows tissue and developmental stage specific differential expression (by changes in regulatory sequences) and the provision of a higher copy number to cope with the varying demands for actin in different tissues and at different developmental stages (Rubenstein, 1990; Herman, 1993; Bryant *et al.*, 2006). While muscle actin has been observed to be highly conserved in gene sequence, copy number, tissue-specific expression patterns and function, regardless of developmental mode, cytoplasmic actin varies in gene number, gene sequence and expression patterns between orthologous genes (Kissinger, 1998). Due to its ubiquitous transcription, actin is often used as an internal control in gene expression studies (VanGuilder *et al*., 2008; Wang *et al*., 2008; Li *et al*., 2010).

Transforming growth factor β (TGF-β) superfamily

The transforming growth factor β (TGF-β) superfamily includes a number of factors that can be divided into three groups: bone morphogenic proteins (BMPs), activins and TGF-β *sensu stricto* (s.s.). A variety of TGF-β members has been characterized at the molecular level and all the proteins of the superfamily share characteristic features. One such important feature is the ability to act as ligands that bind to type I and II serine/threonine kinase receptors to activate signal transduction pathways involved in tissue growth and development (Lelong *et al*., 2001; Kim *et al*., 2004; Herpin *et al*., 2005). One member of this superfamily, originally identified as growth and differentiation factor-8 (GDF-8) but renamed as myostatin (MSTN), was first characterized in mice where null mutations at the MSTN locus resulted in increased muscle mass and growth. This observation lead to the hypothesis that MSTN functions as a negative regulator of muscle development (Kim *et al*., 2004; De Santis and Jerry, 2007).

MSTN has been associated with growth differences in several vertebrate species, specifically those of agricultural importance. In terrestrial vertebrates, MSTN is predominantly expressed in skeletal muscle (Kim *et al*., 2004; De Santis and Jerry, 2007). MSTN has also been characterized in a number of phylogenetically diverse teleost species where it is present as a double copy gene and expressed in several tissues, including muscle, gill, renal and gonad tissue (Kim *et al*., 2004; De Santis and Jerry, 2007).

For invertebrates, less homologues for MSTN has been reported. Probably the first true molluscan MSTN homologue was identified in the scallop *Argopecten irradians* and named sMSTN. Although sMSTN RNA was detected in mantle, gonad, heart, digestive gland and gill tissues, by far the highest mRNA expression was observed in the skeletal (adductor) muscle of adult scallops. It is proposed that sMSTN is involved in respiration, muscle growth and development, and the relationship between somatic and shell growth (Kim *et al*., 2004).

Another member of the TGF-β superfamily that may function as a regulator of body size in bivalve molluscs is molluscan growth and differentiation factor (mGDF) which was identified in the oyster, *Crassostrea gigas*. mGDF showed slightly higher expression in digestive gland, the mantle and the gills than in other tissues tested, but the overall widespread expression of this factor suggests that it has various physiological functions (Lelong *et al*., 2001).

1.2.3 Genes involved in shell deposition and growth

Shell growth consists of enlargement at the shell edge orchestrated by the epithelial mantle cells in close contact with the shell edge; and thickening of the inner surface, a process regulated by the whole mantle (Dogterom, 1980; Wilbur and Saleuddin, 1983). Recent studies of biomineralization processes confirm the involvement of the mantle epithelium in organic matrix deposition and shell formation and suggest that these growth processes are under control of calcium-binding proteins and hormonal factors like calcitonin-gene related peptides that are also known to affect calcium metabolism in vertebrates (Auzoux-Bordenave *et al*., 2007). Amongst the many proteins and genes involved in molluscan shell growth, quite a few display remarkable similarities to growth factors and proteins involved in calcification in a diverse range of species (Sarashina, 2006).

Shell formation and growth

The molluscan shell is a composite biomaterial made of calcium carbonate intimately associated with organic matrix components, secreted by mantle epithelial cells. The shell shows incremental growth throughout the life of the organism (Bøggild, 1930; Wilbur and Saleuddin, 1983; Watabe, 1988). Growth takes place on the distal border of the shell where minute layers are added to adjacent layers, resulting in shell growth in length more than in thickness (Marin and Luquet, 2004). The mantle epithelium from the anterior edge of the mantle tissue secretes the organic matrix (proteins, glycoproteins, proteoglycans, chitin, lipids and polysaccharides) and inorganic substances (Ca^{2+} and CO_3^{2-} ions) that are responsible for $CaCO_3$ precipitation and further shell growth (Wilbur and Saleuddin, 1983; Marin and Luquet, 2004; Lin and Meyers, 2005). Calcium and bicarbonate ions are absorbed from the body surface, inner mantle epithelium, gills and gut and transported via the haemolymph to the epithelial cells where it is stored. From here, it is actively pumped from the cytosol to the extrapallial space by

putative calcium and bicarbonate channels in the membranes of the epithelial cells. The extrapallial fluid that forms in the space between the internal shell surface and the mantle is a supersaturated fluid where the organic matrix is also secreted. Interaction between organic matrix components and mineral ions result in $CaCO_3$ calcification (Marin and Luquet, 2004).

Most of the matrix proteins are rich in aspartic acid and have calcium-binding properties. These proteins control crystal nucleation, select the $CaCO_3$ crystal-morph (aragonite or calcite) and control the growth and spatial arrangement of minerals (Falini *et al*., 1996; Levi-Kalisman *et al*., 2001). Functional enzymes like carbonic anhydrase contribute to the overall control of biomineralization. The aragonitic structures are visible in growth bands (mesolayers) that correspond to periods of growth arrest interspersed with periods of tiled aragonite growth. It has been proposed that aspartic acid rich proteins attach itself to the plane of the aragonite in a periodic fashion that is mediated by biological mechanisms. There appears to be increasing substantiation that various shell matrix proteins are expressed in a coordinated way to produce the elaborate and highly functional structures that constitute the molluscan shell (Shen *et al*., 1997; Weiss *et al*., 2000; Marin and Luquet, 2004; Lin and Meyers, 2005; Takeuchi and Endo, 2006).

Proteins involved in calcium binding and $CaCO_3$ deposition

A number of transcripts and proteins involved in mollusc shell formation have been isolated from the shell of the calcyfing mantle. Some of these are expressed in various regions of the mantle.

LustrinA was first characterized and described in mantle pallial cells of the abalone, *Haliotis rufescens*. It contains protease inhibitor-like domain, suggesting a capacity to bind proteases, thereby protecting the secreted proteins of the shell-forming matrix from degradation. Such mechanisms that rely on protease inhibitors in the outer protective layer are common in a variety of organisms (Shen *et al.*, 1997). A protein similar to lustrinA, called perlustrin was isolated and identified in the abalone, *H. laevigata* (Weiss *et al.*, 2000, 2001). Perlustrin showed homology to the N-terminal domain of mammalian insulin-like growth factor binding proteins (IGFBPs). In mammals, IGFBPs are involved in bone formation and resorption (Ueland, 2004). This sequence similarity, supported by growth factor binding experimental data showing perlustrin's notable binding affinity for IGFs and insulin, confirms perlustrin's identity as a member of the IGFBP family. It is possible that the presence of this binding protein could be an indication for the existence of growth factor-like substances in abalone nacre (Weiss *et al.*, 2001).

Dermatopontins

Dermatopontin, also known as tyrosine-rich acidic matrix protein (TRAMP), is a matrix protein found in arthropod and sponge exoskeleton and also in mammalian skin, muscle, lung, kidney, cartilage and bone tissue (Sarashina *et al*., 2006). In molluscs it was first isolated from the shell of the gastropod *Biomphalaria glabrata* (as *Bg*Derm1). Fourteen homologues of Dermatopontin have been identified since. Of these, eight are considered to be involved in mollusc shell formation as their expression level was highest in the mantle tissue. The other proteins show expression that is distributed amongst mantle, kidney, foot and hepatopancreas, probably indicating a more general function (Sarashina *et al*., 2006).

Peptides and enzymes involved in shell formation control

The importance of calcium ions in the process of growth and shell formation suggests that molecules functionally related to calciotropic hormones may be involved in the control of these processes (Duvail *et al.*, 1998). Calcium-binding proteins and hormonal factors like calcitonin-gene related peptides are known to affect calcium metabolism in vertebrates and it is postulated that similar proteins control organic matrix deposition and shell formation in molluscs (Auzoux-Bordenave *et al.*, 2007).

Calcitonin gene related peptide (CGRP) is encoded by the same gene as calcitonin, but produced by alternative splicing of a common primary transcript. Sequences of this neuropeptide that shows a high level of conservation with human CGRP has been identified in trout, crustaceans, scallop and abalone. CGRP stimulates osteoblast proliferation, suggesting that it may play an anabolic role in bone metabolism (Duvail *et al*., 1997). Detection of CGRP in abalone mantle tissue led to the suggestion that this protein may play a major role as a calciotropic hormone in invertebrates, especially molluscs where it may control calcium metabolism associated with the synthesis of nacre (Duvail *et al.*, 1997). *In vitro* cell culture experiments confirmed the involvement of CGRP in abalone mantle and haemocyte cellular activity as cell metabolism was stimulated when high concentrations (50 ng/ml for mantle cells and 500 ng/ml for haemocytes) of human CGRP were added to the cells (Auzoux-Bordenave *et al.*, 2007).

Carbonic anhydrase (CA), the enzyme that catalyzes the conversion of $C0_2$ to HCO_3^- is present in the mantle and gills of molluscs. It is intimately linked to the calcium reserves of the mantle where it maintains the ionized calcium concentration (Istin and Girard, 1970). The high CA concentration in the branchial membranes of scallops also suggests that it is mainly involved in respiration, acid-base regulation and ionic transport (Duvail *et al.*, 1998). Gill carbonic anhydrase activity displays maximum activity in periods of rapid growth. This positive correlation between gill CA activity and tissue weight as well as between CGRP related molecules and tissue weight suggests that CA and CGRP-like molecules may affect tissue growth to a greater extent than it does shell formation in the scallop, *Pecten*

maximus. These observations lead to the proposal of a possible interaction between carbonic anhydrase activity and CGRP (Duvail *et al.*, 1998). A calcitonin receptor Cg CT-R, expressed in the gills, mantle edge, muscles, digestive gland, heart and labial palps, has been cloned and characterized in the bivalve mollusc *Crassostrea gigas* (Dubos *et al*., 2003). This is the only available sequence for CGRP in molluscs.

Has-genes

A study by Jackson *et al.* (2006) shed some further light on the complexity of the mantle transcriptome. By using an expressed sequence tag (EST) approach, 15 genes were identified that encode proteins involved in the supply of shell building components. Their findings suggest that the proteins involved in shell construction are encoded primarily by rapidly evolving genes. In addition, it appears that there are patterns of complex spatial expression within the mantle tissue of *H. asinina* as specific mantle zones influence the crystal morphology of discrete layers of the mature shell (Jackson *et al*., 2006).

A subsequent study on ontogenic gene expression in the mantle of *H. asinina* highlighted nine genes that display dynamic spatial and temporal expression profiles within the larval shell field and juvenile mantle. *Has-ubfm* shares homology with ubiquitin-like fold modifying proteins from various organisms. *Has-ferrt* shows a high degree of sequence conservation with ferritin proteins from a variety of metazoan taxa. *Has-calmbp1* encodes a protein with sequence similarity to proteins with Ca^{2+} binding sites and *Has-tsfgr1* shares similarity with glycine rich sequences like spidroin, an elastomeric spider silk protein. *Has-cam1* displays a high level of sequence similarity to calmodulin proteins across a broad range of taxa. These proteins are known to have calcium-binding properties and to be involved in muscle contraction. Alignment of *Has-lustA* and LustrinA from *H. rufescens* reveals a high level of sequence conservation, indicating a similar involvement in nacre synthesis in both of these abalone species. *Has-Som* shares some primary structural features with a family of rapidly evolving extracellular glycoproteins, the ependymin proteins (Jackson *et al*., 2007).

1.2.4 Miscellaneous genes and proteins that could play a role in growth regulation

Topoisomerase

Topoisomerase II (TopoII) which is known as a specific marker of cell proliferation has been identified in gill, mantle, foot and digestive gland tissue of the mussel, *Mytilus edulis*. TopoII is a ubiquitously expressed nuclear protein required for cell division in all living organisms. Its primary role is to disentangle intertwined chromosomes during anaphase to allow chromosome segregation to occur prior to cell division. It also plays an important role in transcription and it is suggested to be involved in DNA repair (Withoff *et al*., 1996).

Chitinase-like proteins

Chitinase-like proteins (CLPs) are a class of proteins that belong to the greater family of O-glycosyl hydrolases, GH18. CLPs are implicated in diverse biological processes, like embryonic development, mitogenesis and stimulation of cell proliferation in coordination with insulin in insects. They also act as cytokines during tissue remodeling in humans and in the oyster. It is suggested that CLPs are highly conserved, both structurally and functionally. *Cg*-CLP1 and *Cg*-CLP2 identified in the oyster, *Crassostrea gigas* are highly expressed in oocytes and early stages of embryonic development as well as in adult haemocytes. Weaker expression also occurs in the adult mantle edge, digestive tract and post-spawning gonads and it is postulated that *Cg*-CLP1 and *Cg*-CLP2 are involved in tissue growth, remodeling and repair, regulation of immunity and embryonic development (Badariotti *et al.*, 2007a, b).

Proteoglycans and Collagen

Proteoglycans (PG) /glucosaminoglycans (GAG) and collagen are components of the extracellular matrix that participate in molecular events that regulate cell adhesion, migration, and proliferation (Poncet *et al.*, 2000). Proteoglycans, which are produced by most eukaryotic cells, perform various functions: from the formation of multimolecular aggregates in cartilage to heparan sulfate proteoglycans that provide matrix binding sites and cell-surface receptors for growth factors such as fibroblast growth factor (FGF) (Hardingham and Fosang, 1992). Proteoglycans have been implicated as fundamental components of mineralized tissues and has been identified in the sea urchin shell as well as the organic shell matrix of the snail *Biomphalaria glabrata* (Manouras *et al*., 1991; Marxen *et al*., 1998). The extracellular matrix has the ability to bind growth factors such as TGF-β or FGF (Taipale and Keski-Oja, 1997). This growth factor binding is an important mechanism regulating growth factor activities and subsequent cellular events. Likewise, growth factors can modulate the synthesis of proteoglycans/glucosaminoglycans and collagen of the extracellular matrix. Collagen and PG/GAG synthesis in the mantle cells of the abalone, *Haliotis tuberculata*, was confirmed by *in vitro* cell culture studies (Poncet *et al*., 2000).

Collagen, the most abundant structural protein in the animal body is located in the extracellular matrix of connective tissues where it provides an insoluble scaffold for the attachment of macromolecules, glycoproteins, hydrated polymers and inorganic ions, and for cell attachment. Nineteen types of collagen have been characterized in vertebrates alone and comparable heterogeneity is believed to exist in invertebrates. Fibril-forming collagen, which belongs to the vertebrate type I collagen family, is present in large amounts in the foot muscle of gastropods. Two full-length cDNAs coding for fibril-forming collagen proα-chains were isolated from the foot muscle of the abalone, *H. discus*, and these proteins were named Hdcols (*H. discus* collagens) 1α and 2α. Thick layers of collagen fibrils surround the myofibrils in the abalone foot muscle, causing the foot muscle to be very solid, yet morphologically complex, with high behavioral plasticity. Besides the foot muscle, mRNAs encoding abalone collagen

proα -chains were also identified in mantle and in the adductor muscle. It has been suggested that collagen may represent a key component of molluscan biomineralization (Yoneda *et al*., 1999; Poncet *et al.*, 2000).

Ferritin and metallothionein

In a study by Lucas (2007) investigating genetic aspects of growth of the tropical abalone *Haliotis asinina*, ferritin and metallothionein were found to be up-regulated in fast growing animals and were consequently suggested as markers of fast growth in abalone. Ferritin is involved in metal binding and is essential for detoxification and metabolism of oxygen and iron (Da Costa *et al*., 2001). In the gastropod mollusc, *Littorina littorea,* ferritin is also suggested to play a protective role against oxygen free radicals (Larade and Storey, 2004). Metallothioneins are known to bind metals with cysteine residues and are as a result used as bioindicators for pollution studies (Livingstone, 2007; Lemoine *et al*., 2000; Lucas, 2007). Their up-regulation in fast-growing *Haliotis asinina* was speculated to either reflect a different metabolic rate in these abalone, or differential ingestion of copper-containing feed (Lucas, 2007).

1.3 Methods to study growth variation

In order to produce aquaculture animals that carry a trait or traits of interest, the different approaches that are followed include selection and breeding, ploidy manipulation and genetic engineering (gene transfer). While breeding entails a long term approach that may only result in the expected outcome after several generations, ploidy manipulation and genetic engineering offer a solution that is more feasible in terms of commercial gain (Beaumont and Hoare, 2003). Genetic management practices are being implemented on some South African abalone farms and research on genetic diversity, parentage assignment, linkage mapping, marker assisted selection and triploidy is ongoing (for example De Beer, 2004; Roodt-Wilding and Slabbert, 2006; Slabbert *et al*., 2009 a, b). Although a study of differential gene expression may offer valuable insight into growth variation, no comprehensive effort has been made to identify and quantify the expression of growth-related genes in *H. midae*.

1.3.1 Molecular biology approach to study gene expression

The mRNA population, which translates into the protein products of an organism, represents how its genes are expressed under a particular set of conditions and the analysis of this RNA and proteins can provide the investigator with a reflection of the organism's gene expression profile. Some gene expression experiments are designed to elucidate temporal or spatial patterns of expression. Mostly, gene expression experiments take on a comparative nature, where the gene expression profiles of two or more different tissues or conditions are compared (Lorkowski and Cullen, 2003). Because RNA is transiently expressed and its levels vary depending on cell type, developmental stage, physiology and

pathology, RNA quantification is context dependent and inherently variable. For this reason gene expression data should always be generated and assessed meticulously to ensure its technical accuracy, before biological relevance can be deduced (Adams, 2006).

Differential gene expression analyses have been delivering successful results from the 1970's onwards and have developed dramatically over the past few decades. Early days of gene expression analysis relied on northern blots to detect and quantify the amount of RNA in a sample. The northern blot technique involves the electrophoresis, blotting (onto a nitrocellulose membrane) and hybridization of single stranded RNA to radioactive labeled DNA. Visualization of the membrane results in semi-quantitative analysis of differential gene expression. Although these assays only satisfy low-throughput applications and do not lend itself to statistical analysis, they are still used to confirm differential expression reported by using other methods (Lorkowski and Cullen, 2003). In the 1980's, differential screening and subtractive hybridization methodologies were developed to focus more specifically on mRNA expression and lead to gene discovery in addition to gene expression analysis. Amongst the more sophisticated methods that were developed subsequently to investigate gene expression at the mRNA level, are Differential Display (DD), Serial Analysis of Gene Expression (SAGE), DNA microarrays, Real-Time PCR and recent applications of next generation sequencing, like mRNA Seq (Lorkowski and Cullen, 2003; Marioni *et al*., 2008; Mortazavi *et al*., 2008). By carefully selecting and combining two or more methods to analyse gene expression to answer the specific research question at hand, and by meticulous statistical and bioinformatic analyses of the data, these molecular biology approaches can reveal more about gene expression today than was possible ever before.

1.3.1.1 *Transcriptome analysis*

Techniques employed to study transcriptomics present a snapshot of the current situation in which the RNA was harvested and should be integrated with genomic, proteomic and metabolomic information to attain a comprehensive understanding of gene expression. There are currently two strategies used to assess global transcriptome activity: hybridization-based and DNA sequencing-based methods. Hybridization-based methods were first developed as cDNA arrays utilizing cDNA prepared from whole organs as hybridization probes and have expanded with the advent of large-scale EST generation into oligonucleotide-based expression profiling platforms that operate by spotting oligonucleotides onto a glass surface or by *in situ* probe synthesis (Rosenkranz *et al*., 2008).

Sequencing can be used to quantify gene expression by mapping sequence reads to transcripts. The larger the collection of reads is, the more confidence one can have in the information generated (Rosenkranz *et al*., 2008). The sequencing of ESTs has been used extensively since the 1990's to accomplish gene discovery, complement genome annotation, establish the viability of alternative

transcripts and to aid proteome analysis. Large collections of ESTs have been generated for various organisms and these are made available in public databases (Nagaraj *et al*., 2007). Other methods used to attain improved information content of reads include SAGE and massively parallel signature sequencing (MPSS) (Rosenkranz *et al*., 2008).

MPSS makes EST studies more robust and efficient, because the need for the cloning of ESTs is eliminated and a much larger quantity of data can be generated in a short time, and at reduced cost (Elmer *et al*., 2010). Other advantages of MPSS include its resolution to detect uncharacterized transcripts, rare messages and small differences in expression, its ability to assess gene expression in a comprehensive quantitative manner and its applicability to the study of previously uncharacterized genomes (Hedgecock *et al.*, 2007). The capabilities of MPSS have been drastically improved over the past five years and the technique has been applied successfully in studies characterizing, amongst others; transcripts in the oyster, *Crassostrea gigas* (Hedgecock *et al.*, 2007), clam, *Laternula elliptica* (Clark *et al*., 2010), mouse (Rosenkranz *et al*., 2008), yeast, *Schizosaccharomyces pombe* (Wilhelm *et al*., 2008), butterfly, *Melitaea cinxia* (Vera *et al*., 2008), grapevine, *Vitis vinifera* (Iandolino *et al*., 2008) and hookworm, *Necator americanus* (Cantacessi *et al*., 2010). No MPSS studies have been reported for abalone to date.

The several next generation (NG) sequencing platforms available that can perform MPSS include Roche 454's Genome Sequencer, Applied Biosystems' SOLiD and Illumina's Genome Analyzer (Asmann *et al*., 2008). The Genome analysis system process of Illumina, which is used during this study, is briefly discussed below:

For whole transcriptome analysis, poly-A containing mRNA molecules are purified and fragmented before being copied into first strand cDNA. After second strand cDNA synthesis cDNA fragments are ligated to adapters and ligation products are purified and enriched with PCR to create the final cDNA library (Illumina Inc., 2009). Samples are hybridized onto a flow cell where cDNA anneal to primers acting as capture probes, and templates are amplified by bridge amplification. At the outset of sequencing on the Genome Analyzer, the flow cell contains millions of clusters, each containing ~ 1000 copies of template. On this clustered template sequencing is performed through robust four-color DNA Sequencing-By-Synthesis (SBS) technology. Sequencing employs reversible terminators with removable fluorescence. Single-base incorporation occurs at each cycle and one of four chemically cleavable fluorescent labels corresponds to the identity of each nucleotide. Image acquisition at the end of each cycle relies on fluorescence detection, which is achieved using laser excitation and total internal reflection optics (Illumina Inc., 2008).

The short sequence reads generated in this way can be used for *de novo* transcriptome assembly or can be aligned against a reference genome/transcriptome, which allows for the quantification of transcript abundance. New promoters, exons, untranscribed regions and candidate microRNA precursors can also be identified and RNA splice events can be detected directly by mapping splice-crossing sequence reads (Mortazavi *et al*., 2008).

1.3.1.2 *Real-time PCR*

Developed in the early 1990s, real-time PCR is an improved, refined RNA quantification method based on the conventional polymerase chain reaction (PCR) (Kubista *et al*., 2006). While real-time PCR can be used for pathogen detection, single nucleotide polymorphism (SNP) analysis, analysis of chromosome aberrations and protein detection by real-time immuno PCR, the application of interest in this study is the use of real-time PCR to investigate gene expression i.e. quantitative real-time PCR (qPCR). Quantitative real-time PCR is considered a sensitive and flexible quantification method and can be used on very little starting material to compare mRNA levels in different sample populations simultaneously, characterize patterns of mRNA expression, differentiate between closely related mRNAs and analyse RNA structure (Bustin, 2000; Kubista *et al*., 2006).

Quantitative real-time PCR combines the transcription of a RNA template into cDNA, the subsequent exponential amplification by PCR and quantification of the amplified product. Fluorescent reporters (DNA dyes or probes) which are introduced into the reaction, have a fluorescence that is proportional to the amount of amplified DNA product formed. Monitoring of the fluorescence of these reporters can thus be interpreted as an indication of the amount of amplified product during the course of the reaction (Bustin, 2000; Kubista *et al.*, 2006). When more template is present at the beginning of the reaction, less cycles will be needed to reach a point in which the fluorescent signal is first recorded as statistically significant above the background (Gibson *et al*., 1996). By recording the number of amplification cycles required to obtain a particular amount of DNA while assuming a certain amplification efficiency, it is possible to calculate the number of DNA molecules originally present in the sample (Kubista *et al*., 2006).

Following isolation and purification of the RNA template, the next and very important step in qPCR is the copying of mRNA into cDNA by reverse transcription. This step is critical for accurate quantification, because the amount of cDNA produced must correctly reflect the amount of RNA. As this is the qPCR step where most variation is introduced, it is recommended that replicates should already be used at the reverse transcription step in order to obtain higher experimental accuracy (Kubista *et al*., 2006).

One-step or two-step qPCR can be performed. In the former, all reactants for reverse transcription and subsequent PCR are included in one tube and all the available template will be used during the reaction.

In two-step qPCR, reverse transcription is first performed in one tube and then PCR is performed in another, using the cDNA formed during the first step as template. Two-step qPCR has the advantage that cDNA may be stored and used in portions. It also allows for a minus reverse transcription control to be introduced (Shipley, 2006). This control consists of a RNA sample that was treated in the same way as all other samples, but where reverse transcriptase was omitted from the reverse transcription reaction. No cDNA formation should ideally result from this control (Adams, 2006). If the control is positive, it indicates the presence of genomic DNA in the RNA sample. In order to minimise such residual genomic DNA that might have been carried over from RNA extraction, RNA is routinely treated with DNase prior to reverse transcription (Schmittgen, 2006).

Three priming strategies may be used during reverse transcription. These are random sequence (or hexamer) primers, gene specific primers and oligo(dT) primers (Bustin and Nolan, 2004; Kubista *et al*., 2006; Shipley, 2006). Random primers prime at multiple origins along the RNA template. This means that cDNA is made from mRNA, ribosomal RNA and transfer RNA and that more than one cDNA can be made from each mRNA target. This leads to subsequent amplification of the gene of interest often being over-estimated (Bustin and Nolan, 2004; Shipley, 2006). Gene specific primers results in synthesis of specific cDNA and allows for sensitive quantitative assays, but by using them one is limited to implementing one-step qPCR, which may again be less sensitive than the two-step approach (Bustin and Nolan, 2004; Kubista *et al*., 2006). Oligo(dT) priming offers a more specific method than random priming and allows the user to perform two-step qPCR. Oligo(dT) primers hybridise to the poly(A) tail of mRNA and will thus initiate reverse transcription from the 3'-end of the mRNA. Although this seems to be the most popular method of priming, it also has disadvantages. If the RNA is fragmented or reverse transcription is terminated before the PCR target sequence is reached, the resulting cDNA will not be a suitable template for qPCR (Bustin and Nolan, 2004; Kubista *et al*., 2006; Shipley, 2006).

After reverse transcription, gene specific primers are used during the PCR step when performing two-step qPCR. These primers should facilitate maximum amplification efficiency in order to estimate quantification accurately (Bustin and Nolan, 2004). In an optimal assay the amount of product will double with each cycle (100 % amplification efficiency). Standard curves are used to determine the amplification efficiency of an assay (Adams, 2006; Shipley, 2006; Bustin *et al*., 2009). A standard/calibration curve is constructed for each quantified target by plotting the logarithm of the initial template concentration (of a serial dilution of a standard) on the x-axis to the quantification cycle (C_q: the point where fluorescence is first observed as significantly above background) on the y-axis. A slope of -3.32, y intercept between 33 and 37 and a coefficient of determination (r^2) of 1.00 depict a perfect assay. Amplification efficiency is particularly important when performing relative quantification

assays, where the reference and target gene should have similar efficiencies, to facilitate accurate comparison (Adams, 2006; Bustin *et al*., 2009).

Estimation of the input copy number by relating the PCR signal (C_q) generated by the target sequence to a standard curve is known as absolute quantification (Livak and Schmittgen, 2001). Relative quantification, on the other hand, is when the mRNA concentration of a target gene is reported relative to that of a reference gene. Relative quantification relates the C_q value of the target transcript in a treatment group to that of another sample, such as the untreated/control group, while normalising to the C_q of a reference transcript (Livak and Schmittgen, 2001; Bustin *et al*., 2009). Normalisation not only controls for efficiency of amplification, but also for variations in extraction yield, reverse transcription yield as well as enabling comparisons of mRNA concentrations across different samples (Bustin *et al*., 2009).

The choice of reference gene for normalisation is critical, as its expression should be stable across all samples and should not be affected by the treatment (Pfaffl *et al*., 2002; Vandesompele *et al*., 2002; Pfaffl, 2006). Various popular reference genes called "housekeeping" genes are used in qPCR studies. These include tubulins, actins, glyceraldehyde-3-phosphate dehydrogenase (GAPD), albumins, cyclophilin, micro-globulins, ribosomal units (18S or 28S rRNA) and ubiquitin (UBQ) (Pfaffl *et al*., 2004). However, it has been shown that reference gene expression can vary considerably. Therefore it is important to validate reference genes for stability of expression for each specific study and the use of more than one reference gene is also recommended (Vandesompele *et al*., 2002; Bustin *et al*., 2009). Different software solutions are available to assess stability of gene expression and selection of proper reference genes (GeNorm: Vandesompele *et al*., 2002; Bestkeeper: Pfaffl *et al*., 2004).

After the amplification efficiencies of target and reference genes were confirmed as being similar (within 10 % of each other - Schmittgen and Livak, 2008), quantification runs can be performed. Proper controls should be included in each run. In addition to the minus reverse transcription control, a no template control and positive control are also introduced. A no template control (where template cDNA is substituted by water) testifies that it is indeed the sample being measured and not a contaminant. A positive control (where a template of known amplification efficiency is used) is included in order to assist with troubleshooting in the case where no amplification was measured for test samples (Adams, 2006).

Various data analysis software applications have been developed to group and standardize raw data generated by qPCR. Most of these use algorithms to statistically analyse C_q values on the basis of classical standard parametric tests (analyses of variance or t-tests). Examples include the LightCycler relative quantification software (2001), Relative Expression Software Tool (REST: Pfaffl *et al*., 2002), Q-

Gene (Muller *et al*., 2002), qBASE (Hellemans *et al*., 2007), SoFAR (Metralabs), Data Analysis for Real-Time PCR (DART-PCR: Peirson *et al*., 2003) and Rotor-Gene 6000 Series Software (2006) (Pfaffl, 2006).

Although other researchers have reported gene expression results for *Haliotis midae* generated by microarrays and fluorescent differential display RT-PCR, no gene expression studies, using MPSS and qPCR have been reported for *H. midae*. The previous gene expression studies focused on juvenile abalone's response to changing temperature and oxygen levels and gene expression in immune-stimulated *H. midae* (Janse van Rensburg and Coyne, 2009; Vosloo *et al.,* 2009). No research concerning gene expression involved in differential growth of *H. midae* have been reported to date.

1.3.2 *In vitro* investigation of growth using cell culture

Cell culture enables the researcher to characterize normal dividing cells in a controlled environment using a vast array of techniques, including morphology-, viability-, metabolism-, proliferation- and gene expression assays. Investigation of cell biology on this scale allows controlled assessment of cell health and metabolic state by minimizing variability of *in vivo* responses caused by effects of stress, environmental influences and differences in genetic background (Villena, 2003).

1.3.2.1 *Overview*

Cell culture is a general term referring to the maintenance of cells outside the organism in laboratory controlled conditions. Less often, the term tissue culture is used when the cultured cells originate from pieces of tissue (explants) excised from the source organism (Lackie and Dow, 1999). Cell culture originated in early 1900's in the field of neurobiology when amphibian spinal cord explants were incubated in tissue fluid in a warm, moist environment and monitored at regular intervals under the microscope (Alberts *et al.*, 2002). In modern laboratories, cell cultures are more often made from suspensions of cells dissociated from tissues, instead of tissue explants. Most cells require a substrate on which to grow and divide. This substrate is usually a plastic or glass tissue-culture dish, which is often treated to provide a hydrophilic surface that facilitates cell attachment. Additional coating with extracellular matrix components that promote adhesion, for example collagen, fibronectin, laminin and poly-D-Lysine is sometimes added depending on the specific cell type's requirements (Buchanan *et al*., 1999; ECACC, 2001; Alberts *et al*., 2002).

To create conditions in which cells can live, multiply, and express differentiated properties in a tissue-culture dish, the environment from which the cell originated is simulated. Generally cells require a sterile, pH balanced, osmotically balanced and temperature stable environment with a fresh supply of nutrients, which include inorganic salts, carbohydrates, amino acids, vitamins, fatty acids and lipids, proteins and peptides. Specific media suitable for supporting the growth of a wide range of cell types

have been formulated over the past few decades and are available commercially. Examples include the basal media MEM (Minimum Essential Medium), DMEM (Dulbecco's Modified Eagle's Medium) and GMEM (Glasgow MEM); and the complex media RPMI (Roswell Park Memorial Institute) 1640, Iscove's DMEM and Leibovitz L-15 (ECACC, 2001).

Similar to the first cell cultures initiated in the early 1900's, cultures that are prepared directly from the tissues of an organism are known as primary cell culture. Such cells can be stimulated to proliferate and can be transferred to additional tissue culture dishes to generate a large number of secondary cultures. Subculturing of cells in this fashion can be repeated for weeks or months, enabling the researcher to study the differentiated characteristics of cells *in vitro*, however, the use of primary cultures is restricted because of their limited proliferation potential (Littlewood, 1998; Alberts, 2002). Cells in primary cultures are often described according to their morphology and behaviour, such as "small round" cells, "epithelial-like" cells, "fibroblast-like" cells or "spindle-shaped" cells. These cells are also described as "cells adapted to culture, but with finite division potential" (Brooks and Kurtti, 1971; Lackie and Dow, 1999).

Primary cell culture is advantageous in that a diversity of cell types isolated from the host organism may be maintained with morphology close to *in vivo* conditions. The specific functionalities of different cells are well-preserved in primary cell culture. The disadvantages of primary cell culture include the introduction of contamination from the host organism, the limited number of cells available and the limited lifetime of culture (Freshney, 2005).

Conversely, a cell line is a permanently established cell culture that will proliferate indefinitely given the appropriate conditions (Lackie and Dow, 1999). Some primary cell cultures, like rodent cultures, can form continuous cell lines either spontaneously or following exposure to a mutagenic agent (e.g. the tumourigenic Epstein Barr Virus). Others require additional genetic manipulation, which can result in altered or abnormal cellular behaviour *in vitro*. The disadvantage of such cell cultures is that they cannot categorically be regarded as representative of the *in vivo* condition, but are nevertheless very useful in cell research as they can generate almost unlimited numbers of cells of a known morphology and can be stored in liquid nitrogen for an indefinite period while retaining their viability when thawed (Alberts, 2002; Qiagen, 2002). Also, a continuous cell line has a limited risk of being contaminated by exogenous infective agents like parasites, bacteria, fungi, rickettsia and viruses (which are prevalent in the tissue of origin and thus a greater risk for primary cell cultures) (Littlewood, 1998).

1.3.2.2 *Applications of cell culture*

In aquaculture, the application of cell culture technology is greatly driven by the need for knowledge concerning health protection and disease management. The progress in fish cell culture has also been

driven by work done on important research models like the zebrafish (*Danio rerio*), a fundamental animal model for developmental biology (Villena, 2003). Researchers have developed *in vitro* assays for fish toxicological, pathological, and immunological studies based both on primary cell cultures (hepatocytes) and established cell lines (*Sparus aurata* – Bejar *et al.*, 2002; *Scophthalmus maximus* – Chen *et al.*, 2005; *Lates calcarifer* – Parameswaran *et al.*, 2006).

Mollusc cell cultures have been used to identify transcripts (Badariotti *et al*., 2007a, b), assess cytotoxicity (Domart-Coulon *et al.*, 2000; Le Pennec and Le Pennec, 2001), investigate myogenesis (Naganuma *et al*., 1994; Odintsova *et al.*, 2001, 2010) and biomineralisation (Auzoux-Bordenave *et al*., 2007) and to study cellular responses to vertebrate growth factors (Lebel *et al*., 1996; Serpentini *et al*., 2000). In these studies primary cultures originating from larvae, haemolymph, mantle, heart, gill and digestive gland were used. The only report on *Haliotis midae* cell maintenance is from Janse van Rensburg and Coyne (2009) where haemocytes were cultured for 120 minutes to evaluate the effect of Antimycin A on cell viability. No other comprehensive investigation towards the establishment of primary cell cultures for *H. midae* has been reported to date.

1.4 Aim of this study

The slow growth rate of *H. midae* impedes the optimal commercial production of this most profitable South African aquaculture species. Although continuous research strive towards optimizing abalone culture conditions, nutrition and genetic improvement, no comprehensive effort has been made to identify genes associated with observed size differentiation in farmed *H. midae*.

The aim of this study is therefore to investigate growth variation in *H. midae* at the cellular and molecular level and to identify and quantify the expression of selected growth-related genes. The objectives set out to reach this aim are three-fold and will be reported in separate chapters.

Firstly, the *H. midae* transcriptome will be sequenced by MPSS and genes involved in or regulating growth and metabolism will be identified on the basis of differential expression analysis by mRNASeq (CLCbio). Sequence annotation of differentially expressed sequences to general and custom databases will be used to identify growth-related sequences. Chapter two will thus describe the first attempt to characterize the transcriptome of a haliotid species by using MPSS.

In chapter three, the feasibility of primary cell cultures for *H. midae* will be investigated. Since no comprehensive research towards the establishment of primary cell cultures for *H. midae* has been reported, different tissues of origin will be targeted for initiation of primary cell culture. These cell cultures can provide a useful platform for performing *in vitro* assays. *In vitro* cellular responses to exogenous growth factors will be evaluated by measuring changes in cell viability.

Subsequently, in chapter four qPCR will be employed to quantify the expression of selected target genes, normalized with two reference genes. Target and reference genes will be selected based on differential expression analyses performed in chapter two. Identified genes will be studied for levels of gene expression i) in fast and slow growing abalone and ii) in *in vitro* primary haemocyte cultures treated with different growth stimulating factors. The differential expression of specific target genes evident from these qPCR experiments will be discussed in terms of their possible role in metabolic pathways related to cell viability and proliferation.

By addressing the above objectives it is anticipated that genes involved in growth regulation and metabolism previously unknown in *H. midae,* but homologous to known genes in related species, will be identified. The evaluation of expression of these genes will represent a step towards the genetic improvement of *H. midae* aquaculture.

1.5 References

Adams, P.S. 2006. Data analysis and reporting. *In:* Real-time PCR. M. T. Dorak, editor. Abingdon, UK, Taylor and Francis Group. 39-62

Akalal, D.B.G., and G.T. Nagle. 2001. Mollusk-derived growth factor: cloning and developmental expression in the central nervous system and reproductive tract of *Aplysia*. *Molecular Brain Research.* 91:163-168.

Akalal, D.B.G., J.E. Bottenstein, S.H. Lee, J.H. Han, D.J. Chang, B.K. Kaang, and G.T. Nagle. 2003. *Aplysia* mollusk-derived growth factor is a mitogen with adenosine deaminase activity and is expressed in the developing central nervous system. *Molecular Brain Research.* 117:228-236.

Akalal, D.B.G., C.H. Schein, and G.T. Nagle. 2004. Mollusk-derived growth factor and the new subfamily of adenosine deaminase-related growth factors. *Current Pharmaceutical Design.* 10:3893-3900.

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 2002. Molecular Biology of the Cell, 4th ed. New York, Garland Science. 448p.

Asmann, Y.W., M.B. Wallace, and E.A. Thompson. 2008. Transcriptome profiling using next-generation sequencing. *Gastroenterology.* 135:1466-1468.

Auzoux-Bordenave, S., M. Fouchereau-Peron, M.N. Helléouet, and D. Doumenc. 2007. Activity of mantle cells and hemocytes in abalone primary cell cultures (*Haliotis tuberculata*). *Journal of Shellfish Research.* 26:887-894.

Badariotti, F., M. Kypriotou, C. Lelong, M.P. Dubos, E. Renard, P. Galera, and P. Favrel. 2006. The phylogenetically conserved molluscan chitinase-like protein 1 (Cg-Clp1), homologue of human HC-gp39, stimulates proliferation and regulates synthesis of extracellular matrix components of mammalian chondrocytes. *The Journal of Biological Chemistry.* 281:29583-29596.

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

Badariotti, F., R. Thuau, C. Lelong, M.P. Dubos, and P. Favrel. 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.* 31:559-–570.

Barkai, R., and C.L. Griffiths. 1988. An energy budget for the South African abalone *Haliotis midae* Linnaeus. *Journal of Molluscan Studies.* 54:43-51.

Beaumont, A.R., and K. Hoare. 2003. Biotechnology and genetics in fisheries and aquaculture, 1st ed. Kent, Great Brittain, Blackwell Science Ltd. 158p.

Béjar, J., Y. Hong, and M.C. Alvarez. 2002. An ES-like cell line from the marine fish *Sparus aurata*: Characterization and chimaera production. *Transgenic Research.* 11:279-289.

Bøggild, O.B. 1930. The shell structure of the mollusks. *Det Kongelige Danske Videnskabernes Selskab Skrifter-Naturvidenskabelig og Mathematisk Afdeling.* 9:233–326.

Boudry, P., M. Barré, and A. Gérard. 1998. Genetic improvement and selection in shellfish: A review based on oyster research and production. *Cahiers Options Mediterraneennes (CIHEAM).* 34:61-75.

Bourquin, A. 2009. Welcome to Man and Mollusc/Mollusk. 2009. [Online]. Available: http://www.manandmollusc.net/ [Accessed 29 May 2010].

Britz, P.J., B. Lee, and L. Botes. 2009. AISA 2009 Aquaculture Benchmarking Survey: Primary Production and Markets. A report for the Aquaculture Institute of South Africa and Swisscontact, produced by Enviro-Fish Africa (Pty.) Ltd. 130p.

Brooks, M.A., and T.J. Kurtti. 1971. Insect cell and tissue culture. *Annual Review of Entomology.* 16:27-52.

Bryant, M.J., H.J. Flint, and F.Y.T. Sin. 2006. Isolation, characterization and expression analysis of three actin genes in the New Zealand black-footed abalone, *Haliotis iris*. *Marine Biotechnology.* 8:110-119.

Buchanan, J.T., J.F. La Peyre, R.K. Cooper, and T.R. Tiersch. 1999. Improved attachment and spreading in primary cell cultures of the eastern oyster, *Crassostrea virginica*. *In Vitro Cellular and Developmental Biology.* 35:593-598.

Bunje, P. 2010. The Mollusca: Sea slugs, squid, snails, and scallops. 2010. [Online]. Available: http://www.ucmp.berkeley.edu/taxa/inverts/mollusca/mollusca.php [Accessed 29 May 2010].

Bustin, S.A. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology.* 25:169-193.

Bustin, S.A., and T. Nolan. 2004. Pitfalls of quantitative real-time reverse-transcription polymerase chain reaction. *Journal of Biomolecular Techniques.* 15:155-166.

Bustin, S.A., V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, and C.T. Wittwer. 2009. The MIQE Guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry.* 55:611-622.

Cadoret, J.P., R. Debón, L. Cornudella, V.V. Lardans, A. Morvan, P. Roch, and V. Boulo. 1999. Transient expression assays with the proximal promoter of a newly characterized actin gene from the oyster *Crassostrea gigas*. *FEBS Letters.* 460:81-85.

Cantacessi, C., M. Mitreva, A.R. Jex, N.D. Young, B.E. Campbell, R.S. Hall, M.A. Doyle, S.A. Ralph, E.M. Rabelo, S. Ranganathan, P.W. Sternberg, A. Loukas, and R.B. Gasser. 2010. Massively parallel sequencing and analysis of the *Necator americanus* transcriptome. *PLoS Neglected Tropical Diseases.* 4:e684.

Charlab, R., E.D. Rowton, and J.M. Ribeiro. 2000. The salivary adenosine deaminase from the sand fly *Lutzomyia longipalpis*. *Experimental Parasitology.* 95:45-53.

Charlab, R., J.G. Valenzuela, J. Andersen, and J.M.C. Ribeiro. 2001. The invertebrate growth factor/CECR1 subfamily of adenosine deaminase proteins. *Gene.* 267:13-22.

Chen, S.L., G.C. Ren , Z.X. Sha, and Y. Hong. 2005. Development and characterization of a continuous embryonic cell line from turbot (*Scophthalmus maximus*). *Aquaculture.* 249:63-68.

Clark, M.S., M.A.S. Thorne, F.A. Vieira, J.C.R. Cardoso, D.M. Power, and L.S. Peck. 2010. Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyroseqeuencing. *BMC Genomics.* 11:362.

Colby, G.P., Y.J. Sung, and R.T. Ambron. 2005. mRNAs encoding the *Aplysia* homologues of fasciclin-I and beta-thymosin are expressed only in the second phase of nerve injury and are differentially segregated in axons regenerating *in vitro* and *in vivo*. *Journal of Neuroscience.* 82:484-498.

Da Costa, P.N., C.V. Romão, J. LeGall, A.V. Xavier, E. Melo, M. Teixeira, and L.M. Saraiva. 2001. The genetic organization of *Desulfovibrio desulphuricans* ATCC 27774 bacterioferritin and rubredoxin-2 genes: involvement of rubredoxin in iron metabolism. *Molecular Microbiology.* 41:217-227.

De Beer, M. 2004. Induction of Triploidy in the South African abalone, *Haliotis midae*, by the use of hydrostatic pressure. Unpublished MSc thesis, University of Stellenbosch, Stellenbosch, South Africa.

De Zoysa, M., and J. Lee. 2007. Two ferritin subunits from disk abalone (*Haliotis discus discus*): cloning, characterization and expression analysis. *Fish and Shellfish Immunology.* 23:624-635.

Degnan, B.M., S.M. Degnan, and D.E. Morse. 1997. Muscle-specific regulation of tropomyosin gene expression and myofibrillogenesis differs among muscle systems examined at metamorphosis of the gastropod *Haliotis rufescens*. *Development Genes and Evolution.* 206:464-471.

Department Environmental Affairs and Tourism. 2008. Government notice: Regulations for the protection of wild abalone (*Haliotis*). No.R.62, South Africa.

Department of Fisheries, Government of Western Australia. 2005. Farming Abalone. [Online]. Available: http://www.fish.wa.gov.au/docs/aqwa/Abalone/FarmingAbalonePage01.php?0308 [Accessed 24 June 2010].

De-Santis, C., and D.R. Jerry. 2007. Candidate growth genes in finfish — Where should we be looking? *Aquaculture.* 272:22-38.

DesGroseillers, L., D. Auclair, and L. Wickham. 1990. Nucleotide sequence of an actin cDNA gene from *Aplysia californica*. *Nucleic Acids Research.* 18:3654.

DesGroseillers, L., D. Auclair, L. Wickham, and M. Maalouf. 1994. A novel actin cDNA is expressed in the neurons of *Aplysia californica*. *Biochimica et Biophysica Acta (BBA) - Gene Structure and Expression.* 1217:322-324.

Dogterom, A.A. 1980. Growth hormone, metabolism and shell formation in the pond snail *Lymnaea stagnalis*. Unpublished PhD dissertation, Vrije Universiteit te Amsterdam, The Netherlands.

Dolezelova, E., M. Zurovec, T. Dolezal, P. Simek, and P.J. Bryant. 2005. The emerging role of adenosine deaminases in insects. *Insect Biochemistry and Molecular Biology.* 35:381-389.

Domart-Coulon, I., S. Auzoux-Bordenave, D. Doumenc, and M. Khalanski. 2000. Cytotoxicity assessment of antibiofouling compounds and by-products in marine bivalve cell cultures. *Toxicology in Vitro.* 14:245-251.

Dubos, M.P., F. Badariotti, F. Rodet, C. Lelong, and P. Favrel. 2003. Molecular and physiological characterization of an invertebrate homologue of a calcitonin-related receptor. *Biochemical and Biophysical Research Communications.* 310:972-978.

Duvail, L., E. Lopez, and M. Fouchereau-Peron. 1997. Characterization of a calcitonin gene related peptide-like molecule in the abalone, *Haliotis tuberculata*. *Comparative Biochemistry and Physiology.* 116C:155-159.

Duvail, L., J. Moal, and M. Fouchereau-Peron. 1998. CGRP-like molecules and carbonic anhydrase activity during the growth of *Pecten maximus*. *Comparative Biochemistry and Physiology Part C.* 120:475-480.

Elliott, N.G. 2000. Genetic improvement programmes in abalone: what is the future? *Aquaculture Research.* 31:51-59.

Elmer, K.R., S. Fan, H.M. Gunter, J.C. Jones, S. Boekhoff, S. Kuraku, and A. Meyer. 2010. Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. *Molecular Ecology.* 19:197-211.

European Collection of Cell Cultures (ECACC) and Sigma-Aldrich. 2001. Fundamental techniques in cell culture. Salisbury, UK. 68p.

Falini, G., S. Albeck, S. Weiner, and L. Addadi. 1996. Control of aragonite or calcite polymorphism by mollusk shell macromolecules. *Science.* 271:67-69.

Fallu, R. 1991. Abalone farming, 1st ed., Oxford, Fishing News Books. 195 p.

Farias, A., Z. Garcia-Esquivel, and M.T. Viana. 2003. Physiological energetics of the green abalone, *Haliotis fulgens*, fed on a balanced diet. *Journal of Experimental Marine Biology and Ecology.* 289:263-276.

Fishtech Inc. 2010. Facts about abalone. 2010. [Online]. Available: http://www.fishtech.com/facts.html [Accessed 24 June 2010].

Franchini, P., R. Slabbert, M. van der Merwe, A. Roux, and R. Roodt-Wilding. 2010. Karyotype and genome size estimation of *Haliotis midae*: estimators to assist future studies on the evolutionary history of Haliotidae. *Journal of Shellfish Research*. In press.

Freshney, R.I. 2005. Culture of animal cells: a manual of basic techniques, 5th ed. Hoboken, New Jersey, John Wiley and sons Inc. 642p.

Garnier-Géré, P.H., Y. Naciri-Graven, S. Bougrier, A. Magoulas, M. Héral, G. Kotoulas, A. Hawkins, and A. Gérard. 2002. Influences of triploidy, parentage and genetic diversity on growth of the Pacific oyster, *Crassostrea gigas* reared in contrasting natural environments. *Molecular Ecology.* 11:1499-1514.

Genade, A.B., A.L. Hirst, and C.J. Smit. 1988. Observations on the spawning, development and rearing of the South African abalone *Haliotis midae* Linn.. *South African Journal of Marine Science.* 6:3-12.

Geraerts, W.P.M. 1976. Control of growth by the neurosecretory hormone of the Light Green Cells in the freshwater snail *Lymnaea stagnalis*. *General and Comparative Endocrinology.* 29:61-71.

Geiger, D.L. 2000. Distribution and biogeography of the Haliotidae (Gastropoda: Vetigastropoda) world-wide. *Bolletino Malacologica.* 35:57-120.

Gibson, U.E.M., C.A. Heid, and P.M. Williams. 1996. A novel method for real-time quantitative RT-PCR. *Genome Research.* 6:995-1001.

Giusti, A.F., V.F. Hinman, S.M. Degnan, B.M. Degnan, and D.E. Morse. 2000. Expression of a *Src/Hox5* gene in the larval central nervous system of the gastropod *Haliotis*, a non-segmented spirilian lophotrochozoan. *Evolution and Development.* 2:294-302.

Gómez-Chiarri, M., V.L. Kirby, and D.A. Powers. 1999. Isolation and characterization of an actin promoter from the red abalone (*Haliotis rufescens*). *Marine Biotechnology.* 1:269-278.

Government Communications (GCIS). 2010. Statement on the Cabinet Meeting, 15 June 2010. [Homepage of the Department of Foreign Affairs, Republic of South Africa], [Online]. Available: http://www.dfa.gov.za/docs/2010/cabinet0618.html [Accessed 01 July 2010].

Gricourt, L., G. Bonnec, D. Boujard, M. Mathieu, and K. Kellner. 2003. Insulin-like system and growth regulation in the Pacific oyster *Crassostrea gigas*: hrIGF-1 effect on protein synthesis of mantle edge cells and expression of an homologous insulin receptor-related receptor. *General and Comparative Endocrinology.* 134:44-56.

Gricourt, L., M. Mathieu, and K. Kellner. 2006. An insulin-like system involved in the control of Pacific oyster *Crassostrea gigas* reproduction: hrIGF-1 effect on germinal cell proliferation and maturation associated with expression of an homologous insulin receptor-related receptor. *Aquaculture.* 251:85-98.

Hahn, K.O. 1989. CRC Handbook of culture of abalone and other marine gastropods, 1st ed. Florida, CRC Press. 348 p.

Hahn, K.O. 1992. Review of endocrine regulation of reproduction in abalone *Haliotis* spp. *In:* Abalone of the world: Biology, Fisheries and Culture. S.A. Shepherd, M.J. Tegner and S.A. Guzman del Proo, editors. Oxford, UK, Blackwell Scientific Publications: 49-58.

Haidle, L., J.E. Janssen, K. Gharbi, H.K. Moghadam, M.M. Ferguson, and R.G. Danzmann. 2008. Determination of Quantitative Trait Loci (QTL) for early maturation in rainbow trout (*Oncorhynchus mykiss*). *Marine Biotechnology.* 10:579-592.

Hardingham, T.E., and A.J. Fosang. 1992. Proteoglycans: many forms and many functions. *The FASEB Journal.* 6:861-870.

Havenstein, G.B., P.R. Ferket, and M.A. Qureshi. 2003. Carcass composition and yield of 1957 versus 2001 broilers when fed representative 1957 and 2001 broiler diets. *Poultry Science.* 82:1509-1518.

Hayes, B., M. Baranski, M.E. Goddard, and N. Robinson. 2007. Optimisation of marker assisted selection for abalone breeding programs. *Aquaculture.* 265:61-69.

Hecht, T. 1994. Behavioural thermoregulation of the abalone, *Haliotis midae,* and implications for intensive culture. *Aquaculture.* 126:171-181.

Hedgecock, D., J. Lin, S. DeCola, C.D. Haudenschild, E. Meyer, D.T. Manahan, and B. Bowen. 2007. Transcriptomic analysis of growth heterosis in larval Pacific oysters (*Crassostrea gigas*). *Proceedings of the National Academy of Sciences USA.* 104:2313-2318.

Hellemans, J., G. Mortier, A. De Paepe, F. Speleman, and J. Vandesompele. 2007. qBase relative quantification framework and software for management and automated analysis of real-time quantitative PCR data. *Genome Biology.* 8:R19.

Hennessey, A.S., D.R. Drummond, and J.C. Sparrow. 1993. Molecular genetics of actin function. *Biochemical Journal.* 282:657-671.

Henry, N.R. 1995. Studies related to the artificial spawning and culture of the abalone, *Haliotis midae* Linne, 1785. Unpublished MSc thesis, University of Cape Town, Cape Town, South Africa.

Herman, I.M. 1993. Actin isoforms. *Current Opinion in Cell Biology.* 5:48-55.

Herpin, A., C. Lelong, T. Becker, F. Rosa, P. Favrel, and C. Cunningham. 2005. Structural and functional evidence for a singular repertoire of BMP receptor signal transducing proteins in the lophotrochozoan *Crassostrea gigas* suggests a shared ancestral BMP/activin pathway. *The FEBS Journal.* 272:3424-3440.

Hickman, C.P., and L.S. Roberts, editors. 1994. Biology of Animals, 6th ed. Dubuque, Iowa, Wm. C. Brown Publishers. 764p.

Hinman, V.F., and B.M. Degnan. 2002. Mox homeobox expression in muscle lineage of the gastropod *Haliotis asinina*: evidence for a conserved role in bilaterian myogenesis. *Development Genes and Evolution.* 212:141-144.

Homma, K.I., T. Matsushita, and S. Natori. 1996. Purification, characterization, and cDNA cloning of a novel growth factor from the conditioned medium of NIH-Sape-4, an embryonic cell line of *Sarcophaga peregrina* (Flesh fly). *The Journal of Biological Chemistry.* 271:13770-13775.

Hooker, S.H., and R.G. Creese. 1995. Reproduction of Paua, *Haliotis iris* Gmelin 1791 (Mollusca: Gastropoda), in North-eastern New Zealand. *Marine and Freshwater Research*. 46:617-622.

Hooper, S.L., and J.B. Thuma. 2005. Invertebrate muscles: muscle specific genes and proteins. *Physiological Reviews.* 85:1001-1060.

Hoskins, R.A., J.W. Carlson, C. Kennedy, D. Acevedo, M. Evans-Holm, E. Frise, K.H. Wan, S. Park, M. Mendez-Lago, F. Rossi, A. Villasante, P. Dimitri, G.H. Karpen, and S.E. Celniker. 2007. Sequence finishing and mapping of *Drosophila melanogaster* heterochromatin. *Science.* 316:1625-1628.

Houston, R.D., C.S. Haley, A. Hamilton, D.R. Guy, A.E. Tinch, J.B. Taggart, B.J. McAndrew, and S.C. Bishop. 2008. Major quantitative trait loci affect resistance to infectious pancreatic necrosis in Atlantic salmon (*Salmo salar*). *Genetics.* 178:1109-1115.

Iandolino, A., K. Nobuta, F. Goes da Silva, D.R. Cook, and B.C. Meyers. 2008. Comparative expression profiling in grape (*Vitis vinifera*) berries derived from frequency analysis of ESTs and MPSS signatures. *BMC Plant Biology.* 8:53-69.

ICES. 2008. Report of the working group on the application of genetics in fisheries and mariculture (WGAGFM). Pitlochry, Scotland, UK. 77p.

Iijima, M., N. Akiba, I. Sarashina, S. Kuratani, and K. Endo. 2006. Evolution of Hox genes in molluscs: A comparison among seven morphologically diverse classes. *Journal of Molluscan Studies.* 72:259-266.

Illumina Inc. 2008. Paired-end sequencing user guide for cluster station and genome analyzer. San Diego, USA.

Illumina Inc. 2009. mRNA Sequencing sample preparation guide. San Diego, USA.

Istin, M., and J.P. Girard. 1970. Carbonic anhydrase and mobilisation of calcium reserves in the mantle of lamellibranchs. *Calcified Tissue International.* 5:247-260.

Jackson, D.J., C. McDougall, K. Green, F. Simpson, G. Wörheide, and B.M. Degnan. 2006. A rapidly evolving secretome builds and patterns a sea shell. *BMC Biology.* 4:40-50.

Jackson, D.J., G. Wörheide, and B.M. Degnan. 2007. Dynamic expression of ancient and novel molluscan shell genes during ecological transitions. *BMC Evolutionary Biology.* 7:160-177.

Janse van Rensburg, M., and V.E. Coyne. 2009. The role of electron transport in the defence response of the South African abalone, *Haliotis midae*. *Fish and Shellfish Immunology.* 26:171-176.

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

Kim, H.W., D.L. Mykles, F.W. Goetz, and S.B. Roberts. 2004. Characterization of a myostatin-like gene from the bay scallop, *Argopecten irradians*. *Biochimica et Biophysica Acta.* 1679:174-179.

Kissinger, J.C., and R.A. Raff. 1998. Evolutionary changes in sites and timing of actin gene expression in embryos of the direct- and indirect-developing sea urchins, *Heliocidaris erythrogramma* and *H. tuberculata*. *Development Genes and Evolution.* 208:82-93.

Kono, M., N. Hayashi, and T. Samata. 2000. Molecular mechanism of the nacreous layer formation in *Pinctada maxima*. *Biochemical and Biophysical Research Communications.* 269:213-218.

Kubista, M., J.M. Andrade, M. Bengtsson, A. Forootan, J. Jona, K. Lind, R. Sindelka, R. Sjöback, B. Sjögreen, Stroömbom, L, Ståhlberg. A., and N. Zoric. 2006. The real-time polymerase chain reaction. *Molecular Aspects of Medicine.* 27:95-125.

Lackie, J., and J.A.T. Dow. 1999. The Dictionary of Cell and Molecular Biology., 3rd ed. London, Academic Press.

Lanning, N.J., and C. Carter-Su. 2006. Recent advances in growth hormone signaling. *Reviews in Endocrine and Metabolic Disorders.* 7:225-235.

Larade, K., and K.B. Storey. 2004. Accumulation and translation of ferritin heavy chain transcripts following anoxia exposure in a marine invertebrate. *Journal of Experimental Biology.* 207:1353-1360.

Le Gouar, M., A. Guillou, and M. Vervoort. 2004. Expression of a SoxB and a Wnt2/13 gene during the development of the mollusc *Patella vulgata*. *Development Genes and Evolution.* 214:250-256.

Le Pennec, G., and M. Le Pennec. 2001. Acinar primary cell culture from the digestive gland of *Pecten maximus* (L.): an original model for ecotoxicological purposes. *Journal of Experimental Marine Biology and Ecology.* 259:171-187.

Lebel, J.M., W. Giard, P. Favrel, and E. Boucaud-Camou. 1996. Effects of different vertebrate growth factors on primary cultures of hemocytes from the gastropod mollusc, *Haliotis tuberculata. Biology of the Cell.* 86:67-72.

Lelong, C., M. Mathieu, and P. Favrel. 2000. Structure and expression of mGDF, a new member of the transforming growth factor-beta superfamily in the bivalve mollusc *Crassostrea gigas*. *European Journal of Biochemistry.* 267:3986-3993.

Lelong, C., M. Mathieu, and P. Favrel. 2001. Identification of new bone morphogenetic protein-related members in invertebrates. *Biochimie.* 83:423-426.

Lemoine, S., Y. Bigot, D. Sellos, R.P. Cosson, and M. Laulier. 2000. Metallothionein isoforms in *Mytilus edulis* (Mollusca, Bivalvia): Complementary DNA characterization and quantification of expression in different organs after exposure to cadmium, zinc, and copper. *Marine Biotechnology.* 2:195-203.

Levi-Kalisman, Y., G. Falini, L. Addadi, and S. Weiner. 2001. Structure of the nacreous organic matrix of a bivalve mollusk shell examined in the hydrated state using cryo-TEM. *Journal of Structural Biology.* 135:8-17.

Li, K.W., W.P.M. Geraerts, R.H.M. Ebberink, and J. Joosse. 1992. Purification and sequencing of molluscan insulin-related peptide I (MIP I) from the neuroendocrine light green cells of *Lymnaea stagnalis*. *Moleculur and Cellular Endocrinology.* 85:141-150.

Li, M., J. Greenaway, J. Raine, J. Petrik, A. Hahnel, and J. Leatherland. 2006. Growth hormone and insulin-like growth factor gene expression prior to the development of the pituitary gland in rainbow trout (*Oncorhynchus mykiss*) embryos reared at two temperatures. *Comparative Biochemistry and Physiology, Part A.* 143:514-522.

Li, S., and S. Aksoy. 2000. A family of genes with growth factor and adenosine deaminase similarity are preferentially expressed in the salivary glands of *Glossina m. morsitans*. *Gene.* 252:83-93.

Li, X. 2008. Abalone aquaculture subprogram: Selective breeding of farmed abalone to enhance growth rates (II). FRDC Final Report. 2001/254. West Beach, South Australia, South Australian Research and Development Institute (SARDI).

Li, Z., L. Yang, J. Wang, W. Shi, R.A. Pawar, Y. Liu, C. Xu, W. Cong, Q. Hu, T. Lu, F. Xia, W. Guo, M. Zhao, and Y. Zhang. 2010. ß-Actin is a useful internal control for tissue-specific gene expression studies using quantitative real-time PCR in the half-smooth tongue sole *Cynoglossus semilaevis* challenged with LPS or *Vibrio anguillarum*. *Fish and Shellfish Immunology.* 29:89-93.

Lichanska, A.M., and M.J. Waters. 2007. How growth hormone controls growth, obesity and sexual dimorphism. *Trends in Genetics.* 24:41-47.

Lin, A., and M.A. Meyers. 2005. Growth and structure in abalone shell. *Materials Science and Engineering A.* 390:27-41.

Littlewood, T. 1998. Cell biology aspects of safety in cell culture. *In:* Safety in cell and tissue culture. G. Stacey, A. Doyle and P. Hambleton, editors. Dordrecht, Kluwer: 26-46.

Livak, K.J., and T.D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the 2DDCT method. *Methods.* 25:402-408.

Livingstone, D.R. 2007. Biotechnology and pollution monitoring: Use of molecular biomarkers in the aquatic environment. *Journal of Chemical Technology and Biotechnology.* 57:195-211.

Lorkowski, S., and P. Cullen. 2003. Analysing gene expression Vol. 1, 1st ed. Weinheim, Germany, Wiley-VCH Verlag GmbH and Co. 407p.

Lubet, P. 1971. Influence des ganglions cerebroides sur la croissance de *Crepidula fornicata* Phil. (mollusque mesogasteropode). *Comptes Rendus de l'Académie des Sciences.* 273:2309-2311.

Lucas, T. 2007. Investigating genetic and molecular aspects of growth of the tropical abalone *Haliotis asinina*. Unpublished PhD dissertation, University of Queensland, Queensland, Australia.

Maier, S.A., L. Podemski, S.W. Graham, H.E. McDermid, and J. Locke. 2001. Characterization of the adenosine deaminase-related growth factor (ADGF) gene family in *Drosophila*. *Gene.* 280:27-36.

Maier, S.A., J.R. Galellis, and H.E. McDermid. 2005. Phylogenetic analysis reveals a novel protein family closely related to adenosine deaminase. *Journal of Molecular Evolution.* 61:776-794.

Mangalam, H.J., V.R. Albert, H.A. Ingraham, A. Kapiloff, L. Wilson, C. Nelson, H. Elsholtz, and M.G. Rosenfeld. 1989. A pituitary POU domain protein, Pit-1, activates both growth hormone and prolactin promoters transcriptionally**.** *Genes and Development.* 3:946-958.

Mann, K., I.M. Weiss, S. Andre, H.J. Gabius, and M. Fritz. 2000. The amino-acid sequence of the abalone (*Haliotis laevigata*) nacre protein perlucin: Detection of a functional C-type lectin domain with galactose/mannose specificity. *European Journal of Biochemistry.* 267:5257-5264.

Manouras, A., N.K. Karamanos, T. Tsegenidis, and C.A. Antonopoulos. 1991. Isolation and chemical characterization of two acid carbohydrates from the sea urchin shell, extraction and fractionation of their protein complexes. *Comparative Biochemistry and Physiology, Part B*:119-124.

Marin, F., and G. Luquet. 2004. Molluscan shell proteins. *Comptes Rendus Palevol.* 3:469-492.

Marin, F., P. Corstjens, B. De Gaulejac, E. De Vrind-De Jong, and P. Westbroek. 2000. Mucins and molluscan calcification. *The Journal of Biological Chemistry.* 275:20667-20675.

Marioni, J.C., C.E. Mason, S.M. Mane, M. Stephens, and Y. Gilad. 2008. RNA-seq: An assessment of technical reproducibility and comparison with gene expression arrays. *Genome Research.* 18:1509-1517.

Marxen, J.C., M. Hammer, T. Gehrke, and W. Becker. 1998. Carbohydrates of the organic shell matrix and the shell-forming tissue of the snail *Biomphalaria glabrata* (Say). *The Biological Bulletin.* 194:231-240.

Miyamoto, H., T. Miyashita, M. Okushima, S. Nakanoi, T. Morita, and A. Matsushiro. 1996. A carbonic anhydrase from the nacreous layer in oyster pearls. *Proceedings of the National Academy of Sciences USA.* 93:9657-9660.

Miyamoto, H., M. Hamaguchi, and K. Okoshi. 2002. Analysis of genes expressed in the mantle of oyster *Crassostrea gigas*. *Fisheries Science.* 68:651-658.

Moccia, R., D. Chen, V. Lyles, E. Kapuya, Y. E, S. Kalachikov, C.M. Spahn, J. Frank, E.R. Kandel, M. Barad, and K.C. Martin. 2003. An unbiased cDNA library prepared from isolated *Aplysia* sensory neuron processes is enriched for cytoskeletal and translational mRNAs. *Journal of Neuroscience.* 23:9409-9417.

Moriya, Y., M. Itoh, S. Okuda, A.C. Yoshizawa, and M. Kanehisa. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research.* 35:W182–W185.

Moriyama, S., and H. Kawauchi. 2004. Somatic growth acceleration of juvenile abalone, *Haliotis discus hannai*, by immersion in and intramuscular injection of recombinant salmon growth hormone. *Aquaculture.* 229:469-478.

Morse, D.E. 1984. Biochemical and genetic engineering for improved production of abalones and other valuable molluscs. *Aquaculture.* 39:263-282.

Mortazavi, A., B.A. Williams, K. McCue, L. Schaeffer, and B. Wold. 2008. Mapping and quantifying mammalian transcriptomes by RNA-Seq. *Nature Methods.* 5:621-628.

Muller, P.Y., H. Janovjak, A.R. Miserez, and Z. Dobbie. 2002. Processing of gene expression data generated by quantitative real-time RT-PCR. *Biotechniques.* 32:1372-1374.

Muller, S. 1986. Taxonomy of the genus *Haliotis* in South Africa. *Transactions of the Royal Society of South Africa.* 46:69-77.

Naganuma, T., B.M. Degnan, K. Horikoshi, and D.E. Morse. 1994. Myogenesis in primary cell culture from larvae of the abalone, *Haliotis rufescens* . *Molecular Marine Biology and Biotechnology.* 3:131-140.

Nagaraj, S.H., N. Deshpande, R.B. Gasser, and S. Ranganathan. 2007. ESTExplorer: an expressed sequence tag (EST) assembly and annotation platform. *Nucleic Acids Research.* 35:W143-W147.

Nara, K., H. Matsue, and T. Naraoka. 2004. Granulin-like peptide in the mid-gut gland of the bivalve mollusk, *Patinopecten yessoensis*. *Biochimica et Biophysica Acta.* 1675:147-154.

Newmann, G.G. 1967. Reproduction of the South African abalone, *Haliotis midae*. 64. South Africa, Division of Sea Fisheries, 24p.

O'Brien, E.K., and B.M. Degnan. 2000. Expression of POU, Sox, and Pax genes in the brain ganglia of the tropical abalone *Haliotis asinina*. *Marine Biotechnology.* 2:545-557.

Odintsova, N.A., S.V. Plotnikov, A.A. Karpenko, and M.G. Eliseikina. 2001. Myogenic differentiation of *Mytilus* larval cells *in vitro*. *Russian Journal of Developmental Biology.* 32:304-309.

Odintsova, N.A., V.A. Dyachuk, and L.P. Nezlin. 2010. Muscle and neuronal differentiation in primary cell culture of larval *Mytilus trossulus* (Mollusca: Bivalvia). *Cell and Tissue Research.* 339:625-637.

Pancer, Z., J.P. Rast, and E.H. Davidson. 1999. Origins of immunity: transcription factors and homologues of effector genes of the vertebrate immune system expressed in sea urchin coelomocytes. *Immunogenetics.* 49:773-786.

Parameswaran, V., R. Shukla, R.R. Bhonde, and A.S. Sahul Hameed. 2006. Splenic cell line from sea bass, *Lates calcarifer*: Establishment and characterization. *Aquaculture.* 261:43-53.

Patwary, M.U., M. Reith, and E.L. Kenchington. 1996. Isolation and characterization of cDNA encoding an actin gene from sea scallop (*Placopecten magellanicus*). *Journal of Shellfish Research.* 15:265-270.

Paynter, K.T., and T.T. Chen. 1991. Biological activity of biosynthetic rainbow trout growth hormone in the Eastern oyster, *Crassostrea virginica*. *Biological Bulletin.* 181:459-462.

Peirson, S.N., J.N. Butler, and R.G. Foster. 2003. Experimental validation of novel and conventional approaches to quantitative real-time PCR data analysis. *Nucleic Acids Research.* 31:e73.

Pfaffl, M.W. 2006. Relative quantification. *In:* Real-time PCR. M. T. Dorak, editor. Abingdon, UK, Taylor and Francis Group: 63-82.

Pfaffl, M.W., G.W. Horgan, and L. Dempfle. 2002. Relative expression software tool (REST(c)) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research.* 30:e36-e46.

Pfaffl, M.W., A. Tichopad, C. Prgomet, and T.P. Neuvians. 2004. Determination of stable housekeeping genes, differentially regulated target genes and sample integrity: BestKeeper – Excel-based tool using pair-wise correlations. *Biotechnology Letters.* 26:509-515.

Piper, M.D.W., C. Selman, J.J. McElwee, and L. Partridge. 2008. Separating cause from effect: how does insulin/IGF signalling control lifespan in worms, flies and mice? *Journal of Internal Medicine.* 263:179-191.

Poncet, J.M., A. Serpentini, B. Thiébot, C. Villers, J. Bocquet, E. Boucaud-Camou, and J.M. Lebel. 2000. *In vitro* synthesis of proteoglycans and collagen in primary cultures of mantle cells from the nacreous mollusk, *Haliotis tuberculata*: A new model for study of molluscan extracellular matrix. *Marine Biotechnology.* 2:387-398.

Purchon, R.D. 1977. The biology of the Mollusca, 2nd ed. Oxford, UK, Pergamon press: 269-332.

Qiagen. 2002. The Qiagen transfection resource book, 2nd ed. Qiagen Germany.

Romanova, E.V., M.J. Roth, S.S. Rubakhin, J.A. Jakubowski, W.P. Kelley, M.D. Kirk, N.L. Kelleher, and J.V. Sweedler. 2006. Identification and characterization of homologues of vertebrate ß-thymosin in the marine mollusk *Aplysia californica*. *Journal of Mass Spectrometry.* 41:1030-1040.

Roodt-Wilding, R., and R. Slabbert. 2006. Molecular markers to assist the South African abalone industry. *South African Journal of Science.* 102:99-102.

Roovers, E., M.E. Vincent, E. Van Kesteren, W.P.M. Geraerts, R.J. Planta, E. Vreugdenhil, and H. Van Heerikhuizen. 1995. Characterization of a putative molluscan insulin-related peptide receptor. *Gene.* 162:181-188.

Rosenkranz, R., T. Borodina, H. Lehrach, and H. Himmelbauer. 2008. Characterizing the mouse ES cell transcriptome with Illumina sequencing. *Genomics.* 92:187-194.

Rubenstein, P.A. 1990. The functional importance of multiple actin isoforms. *Bioessays.* 12:309-315.

Sales, J., and P.J. Britz. 2001. Research on abalone (*Haliotis midae* L.) cultivation in South Africa. *Aquaculture Research.* 32:863-874.

Samata, T., N. Hayashi, M. Kono, K. Hasegawa, C. Horita, and S. Akera. 1999. A new matrix protein family related to the nacreous layer formation of *Pinctada fucata*. *FEBS Letters.* 462:225-229.

Sarashina, I., H. Yamaguchi, T. Haga, M. Iijima, S. Chiba, and K. Endo. 2006. Molecular evolution and functionally important structures of molluscan Dermatopontin: Implications for the origins of molluscan shell matrix proteins. *Journal of Molecular Evolution.* 62:307-318.

Schmittgen, T.D. 2006. Quantitative gene expression by real-time PCR: a complete protocol. *In:* Real-time PCR. M. T. Dorak, editor. Abingdon, UK, Taylor and Francis Group: 127-137.

Schmittgen, T.D., and K.J. Livak. 2008. Analyzing real-time PCR data by the comparative CT method. *Nature Protocols.* 3:1101-1108.

Serpentini, A., C. Ghayor, J.M. Poncet, V. Hebert, P. Galera, J.P. Pujol, E. Boucaud-Camou, and J.M. Lebel. 2000. Collagen study and regulation of the *de novo* synthesis by IGF-I in hemocytes from the gastropod mollusc, *Haliotis tuberculata*. *Journal of Experimental Zoology.* 287:275-284.

Shen, X., A.M. Belcher, P.K. Hansma, G.D. Stucky, and D.E. Morse. 1997. Molecular cloning and characterization of lustrin A, a matrix protein from shell and pearl nacre of *Haliotis rufescens*. *The Journal of Biological Chemistry.* 272:32472-32481.

Shipley, G.L. 2006. An introduction to real-time PCR. *In:* Real-time PCR. M. T. Dorak, editor. Abingdon, UK, Taylor and Francis Group: 1-31.

Sin, F.Y.T., M.J. Bryant, and A. Johnstone. 2007. Molecular evolution and phylogeny of actin genes in *Haliotis* species (Mollusca: Gastropoda). *Zoological Studies.* 46:734-745.

Slabbert, R., A.E. Bester, and M.E. D'Amato. 2009a. Analysis of genetic diversity and parentage within a South African hatchery of the abalone *Haliotis midae* Linnaeus using microsatellite markers. *Journal of Shellfish Research.* 28:369-375.

Slabbert, R., A. Roux, M. Van der Merwe, L. Watson, P. Wang, D. Brink, and R. Roodt-Wilding. 2009b. Genetic enhancement in the South African abalone (*Haliotis midae*): current status. The 10th International Symposium on Genetics in Aquaculture, Bangkok, Thailand:76.

Smit, A.B., E. Vreugdenhil, R.H.M. Ebberink, W.P.M. Geraerts, J. Klootwijk, and J. Joosse. 1988. Growth-controlling molluscan neurons produce the precursor of an insulin-related peptide. *Nature.* 331:535-538.

Smit, A.B., W.P.M. Geraerts, I. Meester, H. Van Heerikhuizen, and J. Joosse. 1991. Characterization of a cDNA clone encoding molluscan insulin-related peptide II of *Lymnaea stagnalis*. *European Journal of Biochemistry.* 199:699-703.

Smit, A.B., A. Van Marle, R. Van Elk, J. Bogerd, H. Van Heerikhuizen, and W.P.M. Geraerts. 1993. Evolutionary conservation of the insulin gene structure in invertebrates: cloning of the gene encoding molluscan insulin-related peptide III from *Lymnaea stagnalis*. *Journal of Molecular Endocrinology.* 11:103-113.

Sossin, W.S., T. Kreiner, M. Barinaga, J. Schilling, and R.H. Scheller. 1989. A dense core vesicle protein is restricted to the cortex of granules in the exocrine atrial gland of *Aplysia california*. *The Journal of Biological Chemistry.* 264:16933-16940.

Stepto, N.K. 1997. Triploid induction in the South African Abalone, *Haliotis midae*. Unpublished MSc thesis, University of Cape Town, Cape Town, South Africa.

Sudo, S., T. Fujikawa, T. Nagakura, T. Ohkubo, K. Sakaguchi, M. Tanaka, and K. Nakashima. 1997. Structures of mollusc shell framework proteins. *Nature.* 387:563-564.

Suzuki, M., E. Murayama, H. Inoue, N. Ozaki, H. Tohse, T. Kogure, and H. Nagasawa. 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.

Taipale, J., and J. Keski-Oja. 1997. Growth factors in the extracellular matrix. *The FASEB Journal.* 11:51-59.

Takeuchi, T., and K. Endo. 2006. Biphasic and dually coordinated expression of the genes encoding major shell matrix proteins in the pearl oyster *Pinctada fucata*. *Marine Biotechnology.* 8:52-61.

Tarr, R.J.Q. 1992. The abalone fishery of South Africa. *In:* Abalone of the world: Biology, Fisheries and Culture. S. A. Shepherd, M. J. Tegner and S. A. Guzman del Proo, editors. Oxford, UK, Blackwell Scientific Publications: 438-447.

Tarr, R.J.Q. 1995. Growth and movement of the South African abalone *Haliotis midae*: a reassessment. *Marine and Freshwater Research.* 46:583-590.

The South African Institute for Aquatic Biodiversity. 2004. Coastal fishery resources: an easy guide, 1st ed. Grahamstown, South Africa, 42p.

The UniProt Consortium. 2010. The Universal Protein Resource (UniProt) in 2010. *Nucleic Acids Research.* D142-D148.

Tsuzuki, S., M. Iwami, and S. Sakurai. 2001. Ecdysteroid-inducible genes in the programmed cell death during insect metamorphosis. *Insect Biochemistry and Molecular Biology.* 31:321-331.

U.S. Department of Energy Joint Genome Institute (DOE JGI). 2007. *Lottia gigantea* v1.0 [Homepage of U.S. Department of Energy Joint Genome Institute (DOE JGI)], [Online]. Available: http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html [Accessed 27 May 2010].

Ueland, T. 2004. Bone metabolism in relation to alterations in systemic growth hormone. *Growth Hormone and IGF Research.* 14:404-417.

Valenzuela, J.G., V.M. Pham, M.K. Garfield, I.M.B. Francischetti, and J.M.C. Ribeiro. 2002. Toward a description of the sialome of the adult female mosquito *Aedes aegypti*. *Insect Biochemistry and Molecular Biology.* 32:1101-1122.

Van Kesteren, R.E., J.S. Gagatek, A. Hagendorf, Y. Gouwenberg, A.B. Smit, and N.I. Syed. 2008. Postsynaptic expression of an epidermal growth factor receptor regulates cholinergic synapse formation between identified molluscan neurons. *European Journal of Neuroscience.* 27:2043-2056.

Van Loon, A.E., H.J. Goedemans, A.J.J.M. Daemen, A.J. Van de Kamp, and J.A.M. Van den Biggelaar. 1993. Actin genes expressed during early development of *Patella vulgata*. *Development Genes and Evolution.* 202:77-84.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology.* 3:1-12.

VanGuilder, H.D., K.E. Vrana, and W.M. Freeman. 2008. Twenty-five years of quantitative PCR for gene expression analysis. *BioTechniques.* 44:619-626.

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

Villena, A.J. 2003. Applications and needs of fish and shellfish cell culture for disease control in aquaculture. *Reviews in Fish Biology and Fisheries.* 13:111-140.

Vosloo, D., A. Vosloo, P. Sommer, A. Laas, and J. Samuels. 2009. Differential gene expression in South African Abalone (*Haliotis midae*) in response to oxidative and thermal stress. *Comparative Biochemistry and Physiology - Part A: Molecular and Integrative Physiology.* 153:S75.

Wang, K.J., H.L. Ren, D.D. Xu, L. Cai, and M. Yang. 2008. Identification of the up-regulated expression genes in hemocytes of variously colored abalone (*Haliotis diversicolor* Reeve, 1846) challenged with bacteria. *Developmental and Comparative Immunology.* 32:1326-1347.

Watabe, N. 1988. Shell structure. *In:* The Mollusca. E. R. Trueman and M. R. Clark, editors. New York, Academic Press. 69-104.

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

Weiss, I.M., W. Göhring, M. Fritz, and K. Mann. 2001. Perlustrin, a *Haliotis laevigata* (Abalone) nacre protein, is homologous to the insulin-like growth factor binding protein N-terminal module of verrtebrates. *Biochemical and Biophysical Research Communications.* 285:244-249.

Wilbur, K.M., and A.S.M. Saleuddin. 1983. Shell formation. *In:* The Mollusca. A. S. M. Saleuddin and K. M. Wilbur, editors. New York, Academic Press. 235-287.

Wilhelm, B.T., S. Marguerat, S. Watt, F. Schubert, V. Wood, I. Goodhead, C.J. Penkett, J. Rogers, and J. Bähler. 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature.* 453:1239-1245.

Withoff, S., S. De Jong, E.G. De Vries, and N.H. Mulder. 1996. Human DNA topoisomerase II: biochemistry and role in chemotherapy resistance (review). *Anticancer Research.* 16:1867-1880.

Yang, H.S., Y.Y. Ting, and H.C. Chen. 1998. Blocking polar body with cytochalasin B in the fertilized eggs of the small abalone, *Haliotis diversicolor supertexta* (Lischke), and the development and ploidy of the resultant embryos. *Aquaculture Research.* 29:775-783.

Yoneda, C., Y. Hirayama, M. Nakaya, Y. Matsubara, S. Irie, K. Hatae, and S. Watabe. 1999. The occurrence of two types of collagen proα-chain in the abalone *Haliotis discus* muscle. *European Journal of Biochemistry.* 261:714-721.

Zhang, Y., L. Xie, Q. Meng, T. Jiang, R. Pu, L. Chen, and R. Zhang. 2003. A novel matrix protein participating in the nacre framework formation of pearl oyster, *Pinctada fucata*. *Comparative Biochemistry and Physiology, Part B.* 135:565-573.

2 NEXT-GENERATION SEQUENCING OF THE *H. MIDAE* TRANSCRIPTOME TO IDENTIFY DIFFERENTIALLY EXPRESSED GENES

2.1 Introduction

To date, the characterization of genes and DNA sequences in *Haliotis midae* has been driven mainly by a need for knowledge to improve this agriculturally important species and better understand the organism in its natural environment. *H. midae* has been the subject of several population studies to determine the genetic diversity and evaluate its status in terms of conservation requirements. Allozyme electrophoresis, mitochondrial DNA restriction fragment length polymorphism, microsatellite DNA variation and single nucleotide polymorphism (SNP) methods have been used to assess genetic diversity in *H. midae* (Table 2.1).

Table 2.1 Previously identified molecular markers in *Haliotis midae*

Type of marker	Number identified	Reference
Microsatellite	11	Bester *et al*., 2004
Allozyme loci	7	Evans *et al*., 2004
Mitochondrial	3	
Microsatellite	3	
SNP	12	Rhode *et al*., 2008
SNP	20	Bester *et al*., 2008
Microsatellite	63	Slabbert *et al*., 2008
Microsatellite	44	Slabbert *et al.,* 2010

The aim of marker assisted selection (MAS) is to identify genetic markers surrounding quantitative trait loci (QTL). Quantitative variation, controlled by QTL and environmental influences, characterizes economically important traits in farmed animals, such as growth, meat quality and disease resistance. This information can be utilized to maximise the rate of genetic gain from selective breeding programs for abalone (Hayes *et al*., 2007; Massault *et al*., 2008). Extensive research on *H. rubra* (Hayes *et al*., 2007) and *H. discus hannai* (Li *et al*., 2007; Qi and Akihiro, 2007; Sekino and Hara, 2007; Li *et al*., 2010) have been done towards MAS and a similar approach, through linkage mapping and QTL identification, is envisaged for *H. midae* (Slabbert, 2010).

At the chromosomal level, the diploid chromosome number of *H. midae* has been determined as 36, with 6 metacentric, 10 submetacentric and 2 subtelocentric chromosome pairs. The haploid genome size (C-value) for *H. midae* was recently determined as 1.43 (±0.02) pg, using flow cytometry (Van der Merwe and Roodt-Wilding, 2008; Franchini *et al*., 2010).

Not many studies have focused exclusively on targeting expressed genes in *H. midae,* but a recent report from Vosloo *et al.* (2009) confirmed the differential expression of 39 genes involved in juvenile abalone's response to changing temperature and oxygen levels. Amongst these, detected by

fluorescent differential display RT-PCR (FDD-RT-PCR), are malate dehydrogenase, cytochrome C oxidase, ubiquitin and arginine-N-methyltransferase. Janse van Rensburg and Coyne (2009) also reported two differentially expressed electron transport chain genes, namely cytochrome b and cytochrome c oxidase III, which are up-regulated in haemocytes from immune-stimulated *H. midae*, in response to probiotic treatment.

There is still a vast amount of knowledge to be gained concerning the genome and transcriptome of *H. midae*. Specifically knowledge of QTL's and genes of interest, can greatly improve the productivity of farmed animals. An example is the sequencing of the partial transcriptome of the gonad in the marine scallop (*Argopecten purpuratus*), which revealed specific genes involved in reproduction mechanisms (Boutet *et al*., 2008).

The study of gene expression enables the researcher to understand the mechanisms that contribute to specific phenotypic traits to a greater extent. According to the central dogma, DNA is transcribed into RNA and RNA translated into polypeptides. These are processed into proteins, which affect phenotypic characteristics of cells, tissues, organs and ultimately the whole organism. As not all genes in the genome are encoded simultaneously in all tissues, the study of the transcriptome (the complement of genes encoded at any specific moment or in any specific tissue) offers a directed approach to the study of a specific phenomenon or phenotype. Although identification and characterization of such transcripts can provide invaluable insight, this information should always be regarded within the context of gene regulation. Gene expression is regulated on transcriptional, post-transcriptional, translational and post-translational levels. Techniques employed to study transcriptomics present a snapshot of the situation in which the RNA was harvested and should be integrated with genomic, proteomic and metabolomic information to attain a more comprehensive understanding.

Massively parallel signature sequencing (MPSS) has the advantages of improved resolution for detection of uncharacterized and rare transcripts and for small differences in expression. It also has the ability to assess gene expression in a comprehensive quantitative manner and is applicable to the study of previously uncharacterized genomes. Overall, it offers an attractive method to study gene expression (Hedgecock *et al*., 2007). In the Pacific oyster (*Crassostrea gigas)*, the application of MPSS revealed patterns of gene expression in larvae, that provided novel information about the molecular and physiological causes of heterosis (Hedgecock *et al*., 2007). Although the technique has been successfully applied in various organisms including a bacteria, *Pseudomonas syringae* (Filiatrault *et al*., 2010) a yeast, *Schizosaccharomyces pombe*, (Wilhelm *et al.,* 2008), a fungus, *Grosmannia clavigera* (DiGuistini *et al*., 2009), the hookworm, *Necator americanus* (Cantacessi *et al*., 2010), a butterfly, *Melitaea cinxia*

(Vera *et al*., 2008) and the mouse, *Mus musculus* (Rozenkranz *et al*., 2008), this chapter will describe the first attempt to characterize the transcriptome of a haliotid species, using MPSS.

As the MPSS approach allows one to view a large number of transcripts, the challenge lies in knowing what it is you are looking for. This study will focus on identifying genes of interest involved in differential growth in *Haliotis midae*. Sequences that were previously identified as involved in growth in related mollusc species will be used to guide the identification of similar genes in *H. midae*.

2.2 Materials and Methods

2.2.1 Sampling

Sample animals designated for RNA extraction and downstream transcriptome sequencing originated from a commercial abalone farm on the Southwestern coast of South Africa. Animals were all two-year old siblings from a specific family (7B) of which the parents are known. Sample animals were starved for two days before selection to ensure efficient anesthesia, reduce stress after anesthesia and to prevent injury during the selection process. Anesthesia was achieved by dissolving $MgSO_4$ in seawater at a concentration of 46 g/L, to a final volume of 200 Liters and immersing the basket containing the animals therein until animals detached easily. Initially, three animals were randomly sampled for construction of a reference transcriptome. For investigating differential growth, animals were selected on the basis of differences in weight, shell length and shell width. From a family of 1000 animals, eight of the largest and eight of the smallest animals were selected and allowed to recover from anesthesia in fresh oxygenated seawater (Figure 2.1). The number of animals available for sampling was limited due to two factors. Firstly, tag loss resulted in a reduced number of identifyable animals within the family. Secondly, because the majority of animals in the family were used for continuous growth data collection, as few as possible animals had to be used for destructive sampling. The weights, shell widths and shell lengths were recorded and Analyses of Variance (ANOVAs) confirmed significant differences between the two groups of animals ($P < 0.001$), henceforth named groups L and S. Large animals ranged from 49.05 mm to 63.52 mm in length, 30.11 mm to 36.68 mm in width and 21.13 g to 29.28 g in wet weight. Small animals ranged from 26.07 mm to 37.73 mm in length, 16.48 mm to 24.56 mm in width and 2.4 g to 10.12 g in weight.

Figure 2.1 Abalone sampled for transcriptome sequencing Top row = group L (large); Bottom row = group S (small)

Animals were kept in oxygenated seawater for 3 hours before they were sacrificed. For tissue collection, animals were taken from the water one by one and placed on ice for ten minutes, shell-side

down, to allow muscle contraction to slow down. This was necessary to assist fast and effective dissection. All mucus and water was wiped away with tissue paper and subsequently all soft tissue was dissected away from the shell and placed in a 90 mm petri dish. The tissue was cut into 5 mm strips and transferred to a tube containing RNALater solution (Ambion). RNALater is an aqueous tissue storage reagent that rapidly permeates the tissue to stabilize and protect RNA, until RNA extractions can be completed.

The wet weight of each animal's soft tissue was calculated by subtracting the weight of the tube with its RNALater contents from the final weight of the tube containing the tissue. After the tissue from all 16 animals were collected in RNALater, each tube was opened again to cut the tissue into smaller pieces (2-3 mm^2) to ensure efficient penetration of the RNALater solution. From each individual L/S tube, two grams of tissue was transferred to a communal L/S tube. Samples were pooled to due to limited sequencing space on the Illumina Genome Analyzer, which did not allow separate sequencing of each individual. Each communal tube with pooled tissue was split in three and these tubes were filled with RNALater and stored for one day at 4 ˚C (as suggested by the RNALater product literature) before being transferred to -20 ˚C for long term storage.

2.2.2 RNA extractions

All glassware used during RNA extraction procedures were baked at 145 ˚C for 6-7 hours and plastic ware were soaked in a solution of 0.1 % SDS and 0.1 M NaOH at 37 ˚C overnight. A protocol for extraction of cytosolic RNA adapted from Carninci *et al.* (2002) and Dos Reis Falcão *et al*. (2008) was followed. RNA extractions were performed on the selected samples (three randomly sampled animals, for the reference transcriptome and eight large and eight small animals, for differential expression analysis).

The starting material for RNA extractions consisted of 2.5 g of tissue from each group. While working on ice, 11.25 ml of extraction buffer (50 mM Tris-HCL, 5 mM $MgCl_2$,150 mM KCL, 10 mM β-Mercaptoethanol, 0.8 M Sucrose, 10 mM Ribonucleoside Vanadyl Complexes (SIGMA)) was added to each tube and the tissue was ground to a fine consistency with a polytron homogeniser. Between samples, the homogeniser was cleaned with 70 % EtOH and RNAseZap (Ambion). The liquefied tissue was transferred to 15 ml tubes and centrifuged at 15 500 x g, 4 ˚C for 15 minutes. The supernatant was transferred to 50 ml tubes and 750 µl of H_2O added per 1 g of starting material. An equal volume, to the volume in the tube, of SDS buffer (20 mM Tris-HCL, 1 %SDS, 200 mM NaCl, 40 mM EDTA, 20 ng/ml Proteinase K) was added to each tube before incubating at 45 ˚C for 40 minutes. An equal volume, to the volume in the tube, of Phenol:Chloroform:Isoamylalcohol (P:C:I; 50:49:1) was added to each tube, vortexed well and placed on ice for 15 minutes for the phases to separate. After centrifuging at 10 000 x

g, 20 ˚C for 10 minutes, the upper phase of each tube was transferred to a new tube and an equal volume P:C:I was added. This was repeated for a total of three P:C:I additions and the upper phase of the last centrifugation was transferred to clean tubes. To facilitate precipitation, 2 M of LiCl and an equal volume of 2-propanol was added to each tube before incubating at -20 °C overnight. All samples were subsequently centrifuged at 15 000 x g, 4 ˚C for 60 minutes to obtain RNA pellets. Pellets were dissolved completely in 500 µl sterile, RNase free ddH_2O by vortex. Absorbance measurements were performed on a NanoDrop® ND-1000 spectrophotometer to determine the RNA concentration and purity. To improve purity, RNA samples were cleaned using the RNEasy Midi Kit from Qiagen according to kit instructions. RNA integrity was determined by denaturing agarose gel electrophoresis (2 % agarose; 1 X 3-(N-morpholino)propanesulfonic acid (MOPS) buffer). The gel was run for 90 minutes and photographed under UV light.

2.2.3 cDNA Library preparation and sequencing

To convert the total RNA into a library of template molecules suitable for high throughput DNA sequencing, poly-A containing mRNA molecules were isolated, fragmented and copied into cDNA. The short cDNA fragments were then prepared for sequencing on an Illumina Genome Analyzer II (GA II). A more comprehensive description of library preparation follows:

A starting amount of 11 µg of total RNA was used for all samples. After incubation at 65 °C for five minutes, during which RNA secondary structures were disrupted, the RNA samples were added, together with Binding Buffer (20 mM Tris-HCl pH 7.5, 1.0 M LiCl and 2 mM EDTA; Invitrogen, #610-06) to washed Dynal oligo(dT) magnetic beads (from Dynabeads mRNA Purification Kit, Invitrogen, #610-06), according to the kit instructions. After washing the beads with Washing Buffer B (10 mM Tris-HCl PH 7.5, 0.15 M LiCl, 1 mM EDTA; Invitrogen, #610-06), elution of mRNA by addition of 10 mM Tris-HCl (Invitrogen, #610-06) and heating at 80 ˚C for 2 minutes, a second round of oligo-dT purification was performed. The final elution resulted in a volume of ~9 µl mRNA.

mRNA was fragmented by adding 1 µl of 10 x Fragmentation Buffer (Ambion, #AM8740) to 9 µl of mRNA and incubating at 70 °C for 5 minutes. A volume of 1 µl Stop Buffer (Ambion, #AM8740), was added to terminate fragmentation. These RNA fragmentation reagents are designed to fragment RNA to sizes between 60-200 nucleotides (Ambion #AM8740 product insert). For precipitation of the RNA, 1 µl of 3 M NaOAC, pH 5.2, 2 µl of glycogen (5ug/µl, Ambion, #AM9510), and 30 µl of 100 % EtOH was added and tubes were incubated at -80 °C for 30 minutes. Tubes were centrifuged in a microcentrifuge at 20 200 x g for 25 minutes at 4 °C. Pellets were washed with 70 % EtOH, air-dried and resuspended in 10.5 µl sterile, RNase free ddH_2O.

The cleaved RNA fragments were then copied into first strand cDNA using SuperScript II reverse transcriptase (200U/µl, Invitrogen, #18064-014) and a high concentration of random hexamer primers (3 µg/µl, Invitrogen, #48190-011). This was followed by second strand cDNA synthesis using DNA Polymerase I (10U/µl, Invitrogen, #18010-025) and RNaseH (2 U/µl, Invitrogen, #18021-014). DNA was purified using QIAquick PCR spin columns (Qiagen, #28106) and eluted in 30 µl of Elution Buffer solution.

The following part of the library preparation was performed using the Illumina Genomic DNA Sample Prep Kit #FC-102-1001 according to manufacturer's instructions. To convert overhangs that might have formed during fragmentation into blunt ends, T4 DNA polymerase and Klenow enzyme was used. After the 3' to 5' exonuclease activity of the enzymes removed 3' overhangs, the polymerase activity filled in the 5' overhangs (Illumina Inc., 2008). Following end repair, an 'A' base was added to the 3' end of the blunt phosphorylated DNA fragments, using the polymerase activity of Klenow fragment (3' to 5' exo minus). This prepared the DNA fragments to be ligated to the adapters, which had a single 'T' base overhang at their 3' end (Illumina Inc., 2008). After addition of the 'A' base, DNA was purified using the MinElute PCR Purification Kit (QIAGEN, part # 28004) and eluted in 6 µl of elution buffer. Next, adapters (each 45 bp in size) were ligated to the ends of the DNA fragments, preparing them to be hybridized to the flow cell of the GA II. Paired end adapter oligo mix and DNA ligase were used for adapter ligation and DNA was purified once more and eluted in 15 µl. Subsequently DNA products of the ligation reaction were purified on a gel to remove all unligated adapters, remove adapters that may have ligated to one another, and select a size-range of templates to be transferred to the cluster generation platform of the GA II. Different DNA samples were loaded onto separate agarose gels (2 % agarose; 1 X Tris-borate-EDTA (TBE) buffer) together with a low molecular weight DNA marker (NEB, part # N3233L). After running at 120 V for 120 minutes, the gels were visualized on a UV transilluminator and templates sized ±300 base pairs were excised carefully with a scalpel (Figure 2.2). As RNA fragmentation resulted in RNA of sizes between 60-200 nucleotides and the combined flanking sequence after adapter ligation was approximately 90 bp, fragments in this size range should include the desired 200 bp inserts for sequencing. Additional bands of ±600 base pairs were excised and stored at -20 °C as a back-up RNA source. The template was purified using the QIAquick Gel Extraction Kit (QIAGEN cat# 28706) and eluted in 15 µl of elution buffer.

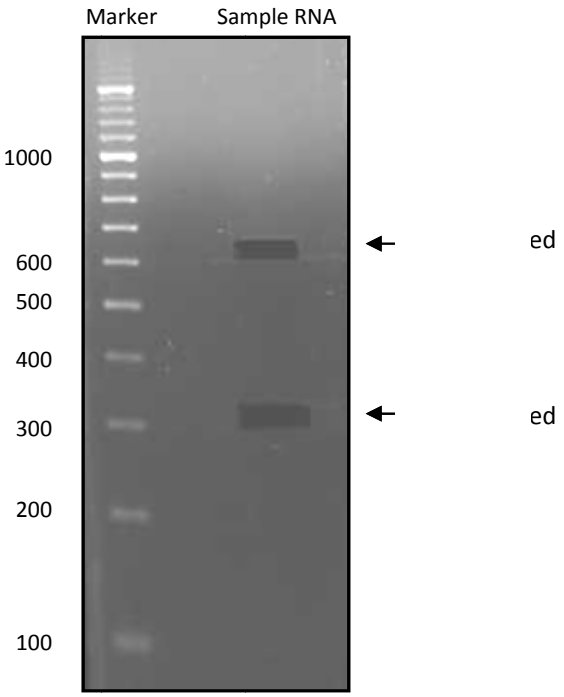


Figure 2.2 Example of 2 % agarose gel after DNA excision

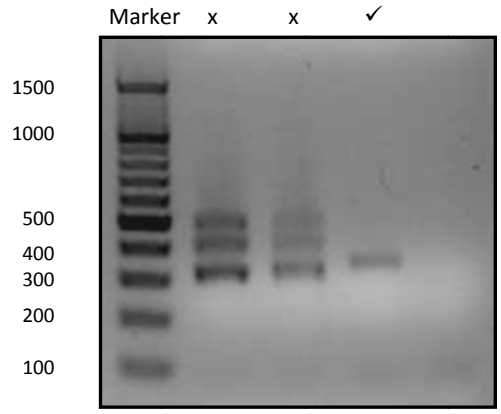


Figure 2.3 Example of 2 % agarose gel for verification of correct amplified product. Lanes 1 and 2 display more than one band indicating poor sample preparation. Lane 3 is a positive verification that the correct template (±300 base pairs) was amplified

PCR was done to enrich for those DNA fragments that had adapter molecules on both ends, and to amplify the amount of DNA in the library. Two primers that annealed to the ends of the adapters were used for the PCR amplification. Only 18 cycles of PCR was employed, to avoid any skewing of the representation of the library (Illumina Inc., 2008). DNA was purified using the QIAquick PCR Purification Kit (QIAGEN, #28104) and eluted in 20 µl of elution buffer. The amplified product was loaded onto a 2 % agarose gel to verify that the correct sized template (±300 bp) was amplified (Figure 2.3).

Each DNA sample was hybridized to a lane on a flow cell on the cluster station of the Illumina GA II at a density of approximately 100 000 clusters per tile according to the instrument's clustering protocol. Sequencing of clustered template DNA on the GA II was performed using four-colour DNA SBS technology. Paired end short reads of 45 and 40 bases were sequenced for the reference transcriptome sample, consisting of three animals. Single end reads of 45 bases were sequenced for the L and S samples, each consisting of eight animals.

2.2.4 Bioinformatics

2.2.4.1 Sequence assembly

Sequence assembly was done with Velvet v0.7.52 (Zerbino and Birney, 2008). The output consisted of contiguous sequences (contigs) stored in a multifasta file. The headline of each sequence is structured as in the following example: **NODE_3_length_4577_cov_21.957**. "**NODE_3"** is the unique name of the contig, "**length_4577"** refers to the length of the contig and "**cov_21.957"** refers to coverage. Due to

the lengths measured by Velvet that are in overlapping k-mers (k-letter words), the actual length in base pair (bp) is calculated by (N + k - 1), where N is the number of overlapping k-mers in a row and k is the size of a k-mer (set as 21 in this case). The actual length of the relevant sequence is thus 4597 bp (= 4577 + 21 - 1). The coverage is the average amount of times each k-mer is observed in the dataset. In other words, after aligning all the reads onto the consensus contig sequence, how many reads would 'cover' a given k-mer. This measure can be interpreted as an indication of certainty of sequencing accuracy and can also be related to expression in an mRNA sample.

Reference transcriptome

Initially RNA from three animals from the same family was extracted and prepared for sequencing as described above. The raw paired end sequences (forward and reverse reads) from the GA were converted from fastq to fasta format using a Python script. Using the Perl script included in the Velvet package (shuffleSequences.pl), the forward and reverse reads were ordered one beneath the other and a hash table was built with the Velvet module "velveth", using a k-mer size of 21. Then, the module "velvetg" was run to build contigs according to the following parameters: insert length 270; expected coverage 50; cutoff 2; minimum length 80 nucleotides. Insert length refers to the maximum estimated length of the insert between the forward and reverse read. In order to reduce errors in the assembly, which will be represented by short, low-coverage nodes, the cutoff value is set. The lower this value is set, the more low coverage nodes, which may or may not be informative, will be included. The minimum length refers to the contig length. When set at 80, only contigs of 80 bp and longer will be included in the output file (Celton, 2008 pers. comm.; Zerbino and Birney, 2008).

Reads that only appeared once in the assembly or that could not be merged in a contig consistently were regarded as singletons (Green, 1994; Wall *et al*., 2009). After the sequencing of L and S sequences was completed, all short reads (original reference transcriptome and L and S transcriptomes) were pooled together to create a more comprehensive reference transcriptome, henceforth designated R. Parameters used for assembly of this final reference transcriptome were: k-mer size 23, insert length 250; expected coverage 100; cutoff 5; minimum length 80 nucleotides. This assembly, thus consisting of the combined transcriptome of 19 individual animals, was used as reference transcriptome (R) for differential expression analysis.

L and S transcriptomes

For the L and S sequences, contig assembly was performed similarly, as for the reference transcriptome, according to the following parameters: k-mer size 23, cutoff 3, and minimum length 80 nucleotides. These transcriptome assemblies were used to BLAST to general databases to confirm the quality of sequencing and assembly.

2.2.4.2 Differential expression analysis

CLC Genomics Workbench v3.5.1 (CLCbio, 2009) was used to analyse differential expression between the two samples (L and S) of DNA. Short read sequence files from L and S were imported into CLC and the mRNASeq module was then used to align them against the velvet-assembled reference transcriptome (R). Separate mRNASeq files for L and S were created, using *unique gene reads* as expression measure. As the genes in this dataset do not have an mRNA annotation, a count parameter (unique gene reads) was used. The subsequent statistical test corrects for the total read count in the experiment (Dekker, 2009 pers. comm.). An *Expression Analysis Experiment* was created in CLC, using the S and L mRNASeq files as input. With the *Statistical Analysis Tool*, Kal's statistical test, a proportion based Z-test originally developed to compare SAGE libraries, was used to compare L and S in terms of expression (Kal *et al*., 1999). By comparing the expression levels at the level of proportions, the data is corrected for sample size, allowing one to compare libraries of different size. The most significantly differential expression values ($P < 0.0005$, False discovery rate (FDR) corrected) were selected to create a sub-experiment.

Three sequence clusters resulted from the expression analysis experiment. Cluster L contained sequences with significantly higher expression in L than in S, cluster S contained sequences with significantly higher expression in S than in L and cluster E contained sequences with equal expression in S and L. These three clusters were extracted from the original reference transcriptome contig file in fasta format, using the online sequence extractor, FaBox (1.35) (Villesen, 2009) before proceeding to sequence annotation.

2.2.4.3 Annotation

The whole reference transcriptome, as well as the three clusters of sequence (L, S and E) resulting from the differential expression analysis were annotated against various databases. Sequence annotation was accomplished by performing sequential alignment searches using cDNA annotation software (dCAS v1.4: http://exon.niaid.nih.gov/cas/). The software is a standalone package that performs raw sequence cleaning, sequence clustering and consensus sequence building before using Basic Local Alignment Search Tool (BLAST) algorithms (Altschul *et al.*, 1997) to compare input sequence with selected sequence databases (Guo *et al*., 2009). Two general databases implemented in the default dCAS installation package were used - the eukaryote clusters of genes (KOG: Tatusov *et al*., 2003) in the Conserved Domain Database (CDD) and Gene Ontology (GO: Ashburner *et al*., 2000). One imported database, the Kyoto Encyclopedia of Genes and Genomes (KEGG: Kanehisa and Goto, 2000; Kanehisa *et al*., 2006; 2010), was also utilised. Throughout the annotation pipeline, BLAST hits with E-values equal to or smaller than 10^{-10} were regarded as significant (Moroz *et al.*, 2006; Wang *et al*., 2008).

The Conserved Domains Database (CDD) is a combination of several collections of multiple sequence alignments representing conserved domains i.e. functional units within a protein that have been used as building blocks in molecular evolution and recombined in various arrangements, resulting in proteins with different functions. CDD is available on the National Center for Biotechnology Information (NCBI; http://www.ncbi.nlm.nih.gov/) and includes the external source databases COG (Clusters of Orthologous Groups), an NCBI-curated protein classification resource. EuKaryotic Orthologous Groups (KOG) is the eukaryote-specific version of COG for identifying ortholog and paralog proteins. KOG consists of 4852 clusters of orthologs, including 59 838 proteins from seven eukaryotic genomes: three animal (*Homo sapiens*, *Drosophila melanogaster* and *Caenorhabditis elegans*), one plant (*Arabidopsis thaliana*), two fungi (*Saccharomyces cerevisiae* and *Schizosaccharomyces pombe*) and one parasite (*Encephalitozoon cuniculi*) genome. KOG provides four functional groups, each divided into KOG classifications identified by letters. The functional groups are "Information storage and processing", "Cellular processes and signaling", "Metabolism", and "Poorly characterized". The classifications by letter for each of these functional groups can be viewed at ftp://ftp.ncbi.nih.gov/pub/COG/KOG/fun.txt. The rpsBLAST algorithm is used to search KOG (Tatusov *et al*., 2003; Anderson, 2008).

Gene Ontology (GO) is a database that has been designed with the aim of providing consistent descriptions of gene products in terms of their biological processes, cellular components and molecular functions in a species-independent manner (The Gene Ontology, 2010). The BLASTx algorithm is used to search GO (Anderson, 2008).

The Kyoto Encyclopedia of Genes and Genomes (KEGG) provides a reference knowledge base for linking genomes to biological systems and environments by metabolic pathway mapping and functional hierarchical classification (Kanehisa and Goto, 2000). KEGG is an integrated database resource consisting of 16 main databases that are categorized into systems information, genomic information, and chemical information. KEGG is widely used as a reference knowledge base for biological interpretation of large-scale datasets generated by sequencing (Kanehisa Laboratories, 2010a). The database is searched using the BLASTx algorithm.

Custom databases can be constructed by collecting the desired sequences in fasta format, and formatting and storing it as a database in dCAS using the "Format Blast Database" utility. From the annotation output, in tab delimited format, biological functions of input sequences can be derived. Three of the customized databases used for annotation included collections of *Haliotis*-specific nucleotide sequences, ESTs and proteins available on the NCBI. The fourth customized database consisted of a list of growth-related sequences, mostly from molluscs, that was regarded as possible

target sequences for differential expression in fast and slow-growing abalone. All these sequences, collected from literature, were reported to contribute to faster growth in molluscs and other invertebrates (see Table 1.2 of the Appendix for the list of sequences and accession numbers). The last customized database was a collection of gene models from the sequenced genome of the gastropod mollusc *Lottia gigantea*, produced by the US Department of Energy Joint Genome Institute in collaboration with the user community (U.S. Department of Energy Joint Genome Institute, 2007). At the time of annotation, this was the only available sequenced mollusc genome. Table 2.2 provides a summary of databases used for annotation.

Table 2.2 Databases used during dCAS annotation

Database name	BLAST function	Database Description
General		
GO	BLASTx	Gene Ontology
KOG	rpsBLAST	Clusters of euKaryotic Orthologous Groups (part of the NCBI's Conserved Domain Database (CDD))
KEGG	BLASTx	Kyoto Encyclopedia of Genes and Genomes
Specific/Customised		
HALIOTIS_EST_NCBI	BLASTn	Collection of EST sequences for all *Haliotis* species available on NCBI
HALIOTIS_NT_NCBI	BLASTn	Collection of nucleotide sequences for all *Haliotis* species available on NCBI
HALIOTIS_PROT_NCBI	BLASTx	Collection of protein sequences for all *Haliotis* species available on NCBI
LOTTIA	BLASTn	Collection of gene models from the sequenced genome of the gastropod mollusc *Lottia gigantea*
SEQ_LIT_NT	BLASTn	Collection of relevant growth related nucleotide sequences reported in literature, available on NCBI
SEQ_LIT_PROT	BLASTx	Collection of relevant growth related protein sequences reported in literature, available on NCBI

From the annotation output, transcripts were categorised into various biological classes. With the aim of identifying sequences that had a higher expression in L than S, results were exported in Excel spreadsheets and sorted according to E–value. Further criteria for selecting sequences from L included an expression ratio of ≤ 0.1 (translating to L having expression values of ≥ 10 times than that of S) and a coverage of more than 30 as well as sequences that showed significant alignment with known growth-related sequences during dCAS annotation. In general, in databases that were not *Haliotis* specific, best hits with molluscs and invertebrates were considered as priority hits before looking at hits within the vertebrate group.

Significant hits from databases used during dCAS annotation, that were considered as sequences that might contribute to faster growth based on the abovementioned criteria, were furthermore BLASTed to online databases available for molluscs. MolluscDB, a collection of mollusc ESTs (Blaxter, 2010) was

searched using tBLASTx. Likewise, these sequences were BLASTed to SnailDB, a collection of ESTs available for the gastropod mollusc *Biomphalaria glabrata* (De Cássia Ruy *et al*., 2010).

2.3 Results

2.3.1 Sampling

Difference between the groups L and S, each consisting of eight animals, was confirmed with Analyses of Variance (ANOVAs) for shell length, shell width and wet weight, respectively. Results of these ANOVAs are summarized in Table 2.3 and confirm significant differences ($P < 0.0001$) for all three parameters.

Table 2.3 Results of ANOVA for differences between large (L) and small (S) groups

Parameter	P-value (ANOVA)	Difference between L and S
Length	1.74×10^{-07}	Significant ($P < 0.0001$)
Width	3.4×10^{-07}	Significant ($P < 0.0001$)
Weight	6.3×10^{-09}	Significant ($P < 0.0001$)

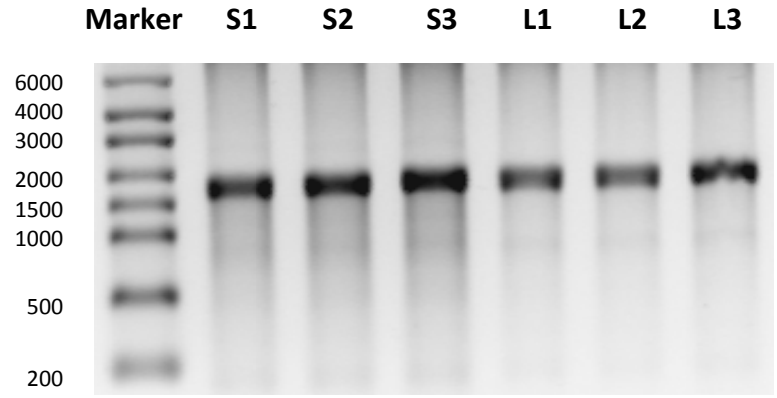
2.3.2 RNA extraction

The 260/280 and 260/230 ratios of absorbance are used to assess the purity of RNA. A 260/280 ratio of ~2.0 and a 260/230 ratio of 1.8 – 2.2 is generally accepted as pure for RNA (Nanodrop Technical Support Bulletin T009, 2007). Concentrations and absorbance ratios for RNA from L and S after RNA extraction and after an additional RNA cleanup are presented in Table 2.4. The improved absorbance ratios indicate that the RNA cleanup with the RNEasy kit (Qiagen) removed proteins and polysaccharides. Although a much lower RNA concentration is present after cleanup, this is desirable to RNA with impurities, which can inhibit downstream reactions like reverse transcription, PCR and sequencing.

Table 2.4 Concentrations and absorbance ratios for RNA after extraction and cleanup

Before RNA cleanup		**Concentration**	**260/280**	**260/230**
	L	2152.4 ng/µl	1.79	1.6
	S	2426.2 ng/µl	1.83	1.68
After RNA cleanup				
	L	354.4 ng/µl	1.93	1.89
	S	446.5 ng/µl	1.98	1.90

Following gel electrophoresis, as presented in Figure 2.4, a single 28S RNA band between 1500 and 2000 base pairs for L and S RNA samples confirmed that the RNA was intact and similar to other molluscan RNA that displayed only a single RNA band instead of the three bands often seen in other eukaryotes (Gambacurta *et al.,* 1993). The mean concentration for RNA extracted from the initial three abalone was 550 ng/µl, but details of absorbance ratios and gel electrophoresis are not available as this RNA was prepared by a previous worker in the same laboratory.



with RNA
L3) tissue
base pairs

2.3.3 cDNA Library preparation, sequencing and sequence assembly

Reference transcriptome

Assembly of the reference transcriptome (R), using the combined short reads of 19 individual animals resulted in 127687 nodes (from 8312245 k-mers of good quality). These were used to construct 30689 contigs with a median length of 296 and a maximum length of 5740. The number of contigs with lengths over 500 base pairs was 3336 and the nucleotide-wise coverage, i.e. the average number of supporting reads per base pair, was estimated at 30 times (Figure 2.5). Sixty-three percent (15982020 from 25261244) of the total reads used were singletons, meaning that they could not be combined into contigs, due to insufficient overlap. Singletons were excluded from annotation and further analysis. This assembly was used henceforth as reference transcriptome (R) for *H. midae*.

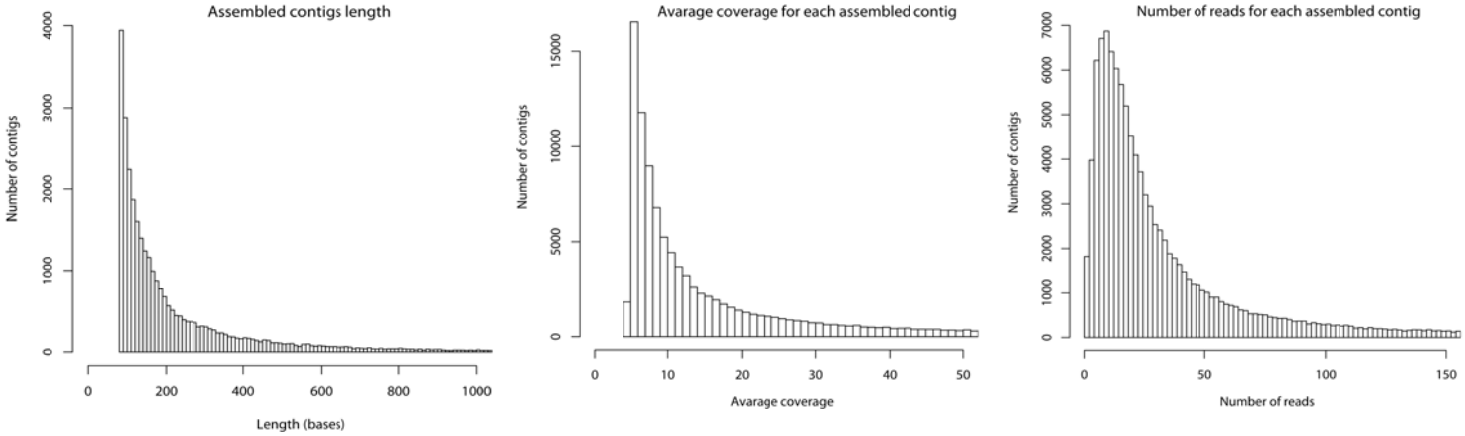


Figure 2.5 Summary of the sequencing and assembly of the *H. midae* reference transcriptome. The three histograms show the distribution of the length (A), the average coverage (B) and the number of reads of the assembled contigs (C)

2.3.4 Differential expression analysis

For L, the mRNASeq procedure resulted in 1904042 assembled reads and 6583312 non-assembled reads. For S, the mRNASeq procedure resulted in 1386015 assembled reads and 4589541 non-assembled reads. After performing Kal's statistical test, the most significant differential expression values (P < 0.0005, FDR corrected) returned 1174 sequences that were highly expressed in L ([Expression Small/Expression Large] < 1), 1324 sequences which were highly expressed in S ([Expression Small/Expression Large] > 1) and 8 sequences with equal expression, named E ([Expression Small/Expression Large] = 1) (Figure 2.6). These sequences are represented by the volcano plot created with CLC Genomics Workbench v3.5.1 (CLCbio) in Figure 2.6. L sequences were selected from the boxed area on the right-hand side of the graph, while S sequences were selected from the boxed area on the left-hand side of the graph. E sequences are represented by the circled area.

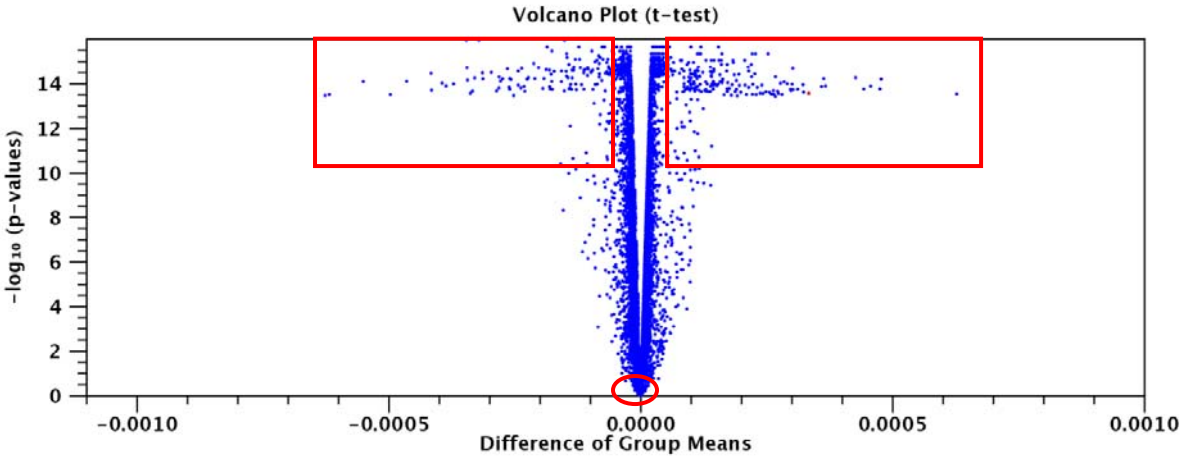


Figure 2.6 Volcano plot displaying the $-\log_{10}$ of the P-values from Kal's statistical test in terms of different group means. The boxed areas represent sequences that have significantly different expression values ($P \leq 10^{-15} - P \leq 10^{-10}$). The circled area represents sequences with similar expression

2.3.5 Annotation

The annotation results (after an E-value cutoff of 10^{-10} was applied) of the reference (R), large (L) and small (S) transcriptomes to the general databases GO, KOG and KEGG are presented in the following paragraphs. Annotation hits were classified into functional categories according to the categories available for the respective databases.

Gene Ontology (GO) annotation of R returned 2987 contigs that had BLAST hits with an E-value of $\leq 10^{-10}$. From these, a subsection of 1095 was identified using "BLAST2GO 2.4.2" (Conesa and Götz, 2009). BLAST2GO uses BLAST searches to find sequences that are similar to input sequences. GO terms associated with each of the obtained hits are extracted and an evaluated GO annotation for the query sequence(s) is returned (Conesa and Götz, 2009). These contigs were divided into three functional

categories: “Cellular Component”, “Molecular Function” and “Biological Process". In each of these, the contigs were classified into various subcategories according to the Go Slim classification system, using the Generic Go Slim available online (GO Consortium, 2009). The number of unique contigs of each main category represents the number of unique sequences contained in that category, regardless of the number of subcategories in which it appears (Table 2.10, Appendix). Using these numbers, the pie charts of the reference transcriptome (R) as well as the large (L) and small (S) transcriptomes (for which annotation tables are not shown here) were constructed (Figure 2.7).

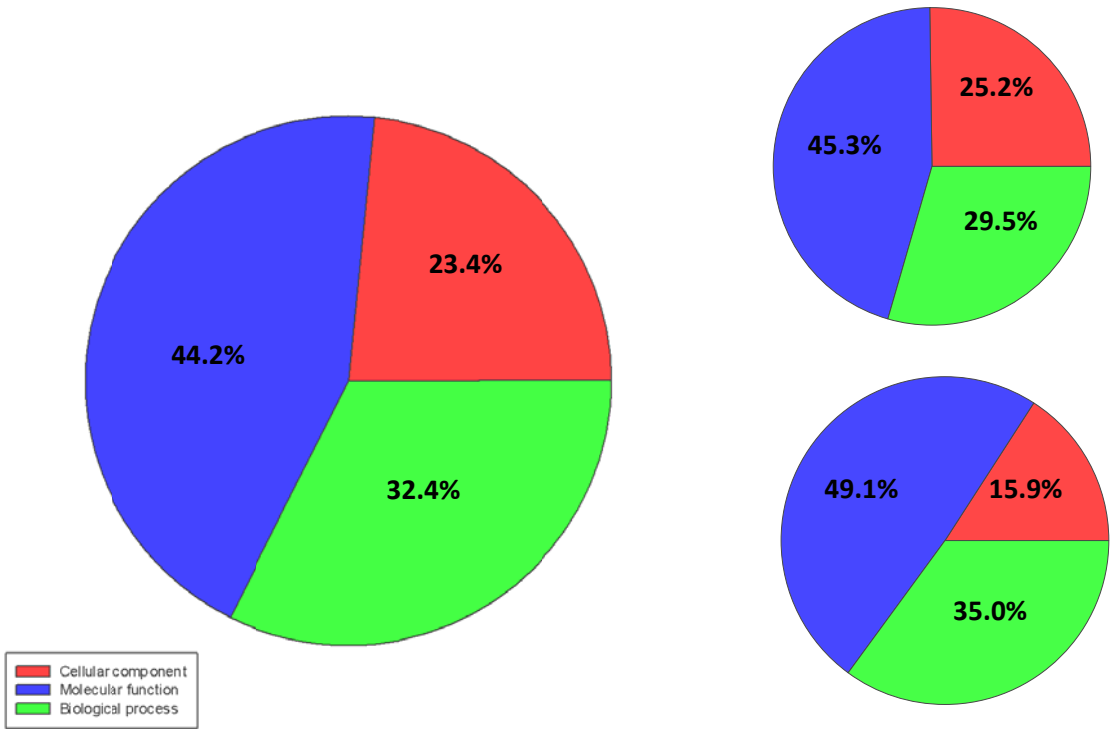


Figure 2.7 Categorization of *H.midae* contigs with significant BLAST hits (E-value cutoff ≤ 10^{-10}) to the GO database, using three main categories

For the reference *H. midae* transcriptome, 1095 sequences could be assigned to GO categories. Assigned sequences had the greatest representation in the “Molecular function” category, followed by the “Biological process” and “Cellular component” categories (Figure 2.7). The annotation of the L and S contigs resulted in a similar proportional distribution of assigned sequences in the three categories.

After GO annotation of the reference *H. midae* transcriptome, the distribution of species to which *H. midae* contigs had the best BLAST hits was assembled. These are presented in Figure 2.8. The distribution of BLAST hits (with an E-value ≤ 10^{-10}) to known sequences in a variety of organisms, including molluscs (*Haliotis*, *Crassostrea* and *Aplysia*) attests to the accuracy of the contig assembly. For *de novo* assembly such as this, where no annotated reference is available, the matching of contigs to known proteins gives an indication of the quality of assembly (Parchman *et al*., 2010).

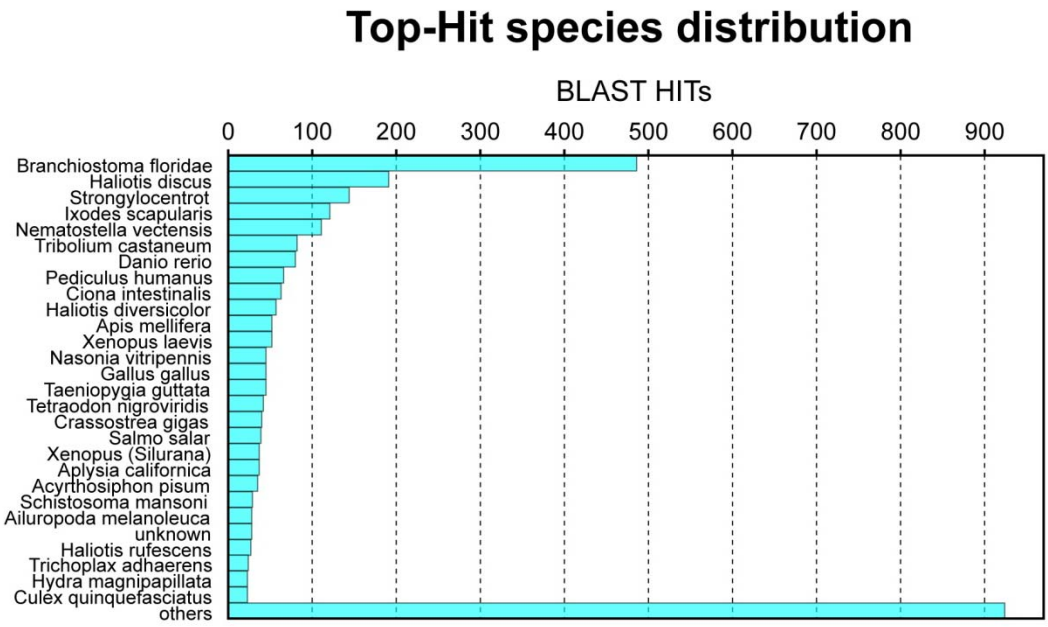


Figure 2.8 Species distribution of top BLAST hits for annotation of the *H. midae* transcriptome

Annotation of R to the KOG database returned 2654 contigs that had BLAST hits with an E-value of ≤ 10^{-10}. These were classified into four functional groups, "Cellular processes and signaling", "Information storage and processing", "Metabolism" and "Poorly categorized" (Table 2.5, Figure 2.9).

Table 2.5 Functional classification of contigs from the *H. midae* transcriptome (R) with a BLAST E-value of ≤ 10^{-10}, based on the KOG database

FUNCTIONAL CLASSIFICATION	Number of contigs	Percentage of total
Cellular processes and signalling	**1058**	
Cell wall/membrane/envelope biogenesis	29	1.09%
Cell motility	6	0.22%
Posttranslational modification, protein turnover, chaperones	284	10.70%
Signal transduction mechanisms	339	12.77%
Intracellular trafficking, secretion, and vesicular transport	131	4.93%
Defense mechanisms	17	0.64%
Extracellular structures	52	1.95%
Nuclear structure	13	0.48%
Cytoskeleton	187	7.04%
Information storage and processing	**540**	
RNA processing and modification	131	4.93%
Chromatin structure and dynamics	31	1.16%
Translation, ribosomal structure and biogenesis	252	9.49%
Transcription	113	4.25%
Replication, recombination and repair	13	0.48%
Metabolism	**752**	
Energy production and conversion	163	6.14%
Cell cycle control, cell division, chromosome partitioning	39	1.46%
Amino acid transport and metabolism	123	4.63%
Nucleotide transport and metabolism	38	1.43%
Carbohydrate transport and metabolism	148	5.57%
Coenzyme transport and metabolism	21	0.79%
Lipid transport and metabolism	87	3.27%
Inorganic ion transport and metabolism	69	2.59%
Secondary metabolites biosynthesis, transport and catabolism	64	2.41%
Poorly characterized	**304**	
General function prediction only	203	7.64%
Function unknown	101	3.80%
Total	**2654**	

The pie charts in Figure 2.9 are graphical presentations of the number of contigs of each main category in the annotation table. The whole transcriptome (R) as well as the large (L) and small (S) transcriptomes (for which annotation tables are not shown here) were constructed.

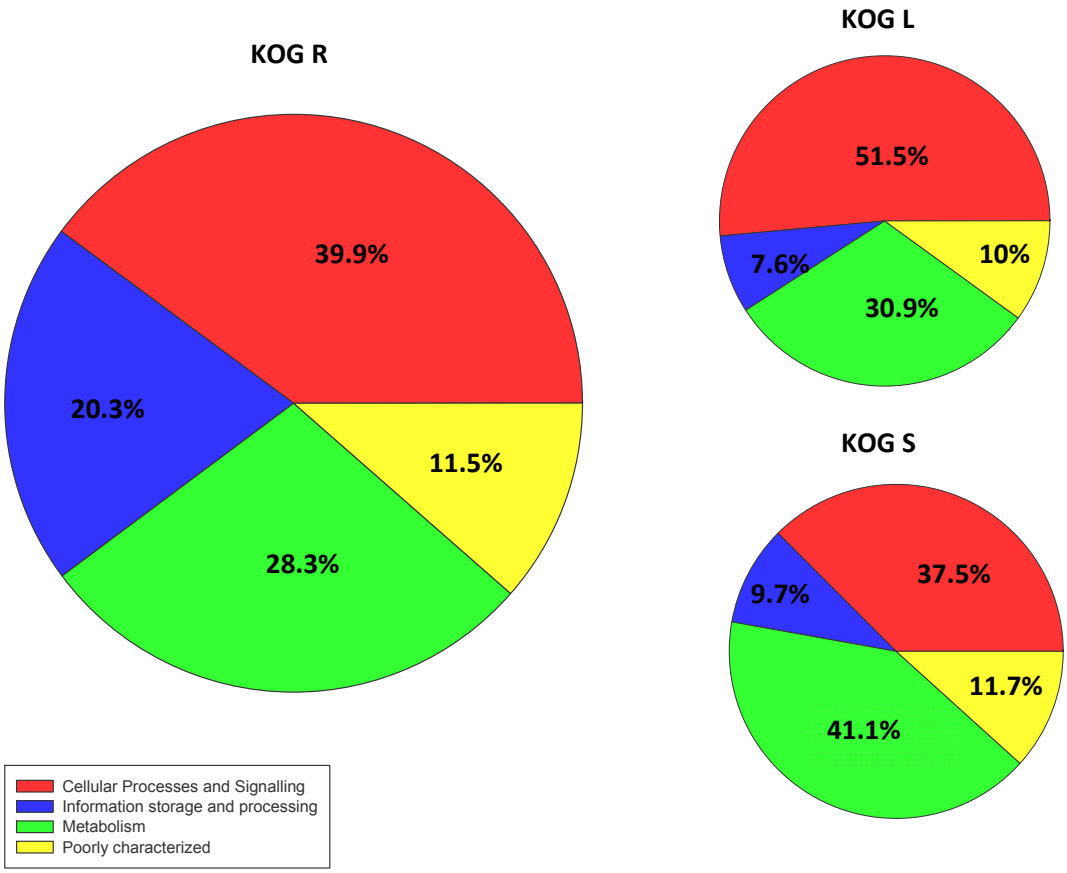


Figure 2.9 Categorization of *H. midae* contigs with significant BLAST hits (E-value cutoff ≤ 10^{-10}) to the KOG database, using four main categories

Assigned sequences in R had the greatest representation in the "Cellular processes and signaling" category, followed by the "Metabolism" and "Information storage and processing" categories (Table 2.5, Figure 2.9). The annotation of the complement of contigs with significantly higher expression in large animals resulted in a distribution of assigned sequences with the greatest representation of sequences in the "Cellular processes and signaling" category, followed by the "Metabolism" and "Poorly characterized" categories. The annotation of the complement of contigs with significantly higher expression in small animals resulted in a similar distribution of assigned sequences in the four categories as for L (Figure 2.9).

When comparing L and S relative to the R annotation at the level of subcategories, different distributions become apparent. Categories where a higher representation of L contigs is found include "Signal transduction mechanisms", "Extracellular structures" and "Inorganic ion transport and metabolism". Categories where a higher representation of S contigs are found, include "Carbohydrate transport and metabolism"' "Lipid transport and metabolism", and "Secondary metabolites biosynthesis, transport and catabolism". All these subcategories, which differ by a factor of more than two between L and S, are presented in Table 2.6.

Table 2.6 Relative representation of contigs in selected subcategories of KOG annotation

Main category	**Functional classification subcategory**	**Percentage representation**			
		R	**Large**	**Small**	**L>S**
Cellular processes and signaling	Signal transduction mechanisms	12.77%	25.29%	11.69%	2.16 X
	Extracellular structures	1.95%	8.23%	0%	8.23 X
Metabolism	Inorganic ion transport and metabolism	2.59%	7.35%	2.82%	2.61 X
					S>L
Metabolism	Carbohydrate transport and metabolism	5.57%	3.82%	10.88%	2.85 X
Metabolism	Lipid transport and metabolism	3.27%	0.88%	7.25%	8.24 X
Metabolism	Secondary metabolites biosynthesis, transport, catabolism	2.41%	2.35%	5.64%	2.4 X

Annotation of R to KEGG Metabolic Pathways returned 1212 contigs that had BLAST hits with a bit score of > 60. These contigs were annotated by the KEGG Automatic Annotation Server, which automatically removes BLAST hits with bit scores less than 60 (KAAS; Moriya *et al*., 2007). This server provides functional annotation of genes against the manually curated KEGG GENES database, by using BLAST comparisons. The resulting annotation contains KEGG Orthology (KO) assignments and automatically generates KEGG pathways. The six functional categories used by KAAS includes: “Metabolism”, “Genetic Information Processing”, “Environmental Information Processing", “Cellular Processes”, “Organismal Systems” and “Human Diseases” (Table 2.7). Each category is subdivided into subcategories according to the KO classification. Since a particular sequence can belong to more than one subcategory the sum of the “Number of unique contigs” is lower than the combined number of contigs in subcategories. Likewise, a particular sequence can belong to more than one main category and the combined number of sequences in main categories will be higher than 1212.

Table 2.7 Functional classification of contigs from the *H. midae* transcriptome (R) with a BLAST E-value of ≤ 10^{-10}, based on the KEGG database

FUNCTIONAL CLASSIFICATION	Number of unique contigs
Metabolism	**527**
Carbohydrate metabolism	214
Energy metabolism	137
Lipid metabolism	95
Nucleotide metabolism	48
Amino acid metabolism	155
Metabolism of other amino acids	61
Glycan biosynthesis and metabolism	55
Biosynthesis of polyketides and nonribosomal peptides	1
Metabolism of cofactors and vitamins	41
Biosynthesis of secondary metabolites	32
Xenobiotics biodegradation and metabolism	72
Enzyme families	108
Genetic information processing	**499**
Transcription	200
Translation	250
Folding, sorting and degradation	222
Replication and repair	104
Environmental information processing	**174**
Membrane transport	31
Signal transduction	103
Signaling molecules and interaction	126
Cellular processes	**221**
Transport and catabolism	102
Cell motility	76
Cell growth and death	68
Cell communication	63
Organismal systems	**145**
Immune system	82
Endocrine system	78
Circulatory system	31
Excretory system	2
Nervous system	29
Sensory system	8
Development	14
Environmental adaptation	4
Human diseases	**210**
Cancers	107
Immune system diseases	13
Neurodegenerative diseases	241
Circulatory diseases	28
Metabolic diseases	7
Infectious diseases	58

The pie charts in Figure 2.10 are graphical presentations of the number of contigs of each main category in the annotation table. The whole transcriptome (R) as well as the large (L) and small (S) transcriptomes (for which annotation tables are not shown here) were constructed.

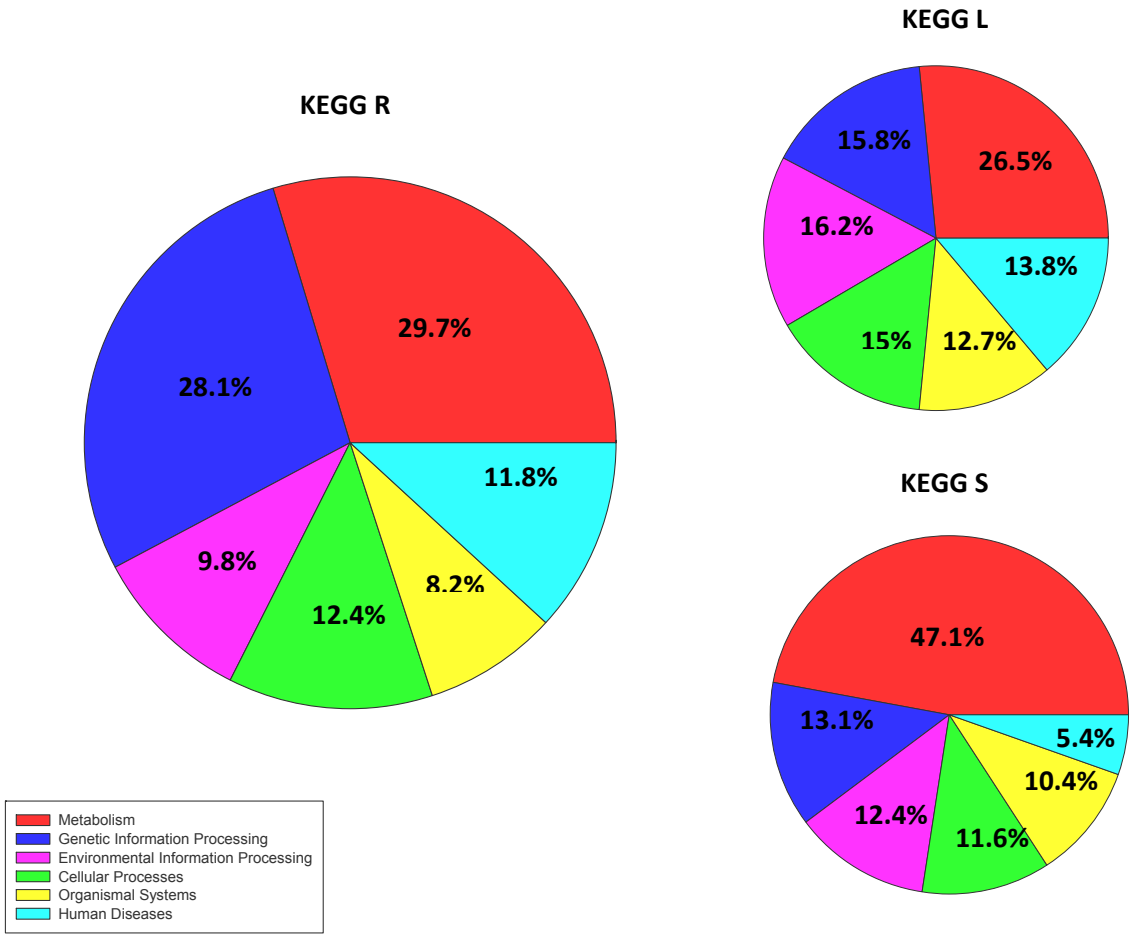


Figure 2.10 Categorization of *H. midae* contigs with significant BLAST hits (E-value cutoff ≤ 10^{-10}) to the KEGG database, using six main categories

For the complete (R) *H. midae* transcriptome, 1212 sequences could be assigned to KEGG categories. Assigned sequences had the greatest representation in the "Metabolism" category, followed by the "Genetic information processing" and "Cellular processes" categories (Table 2.7, Figure 2.10). The annotation of the complement of contigs with significantly higher expression in large animals as well as those with significantly higher expression in small animals resulted in a similar distribution of assigned sequences in the six categories to that found in R. Contigs with significantly higher expression in small animals had an even greater representation in the Metabolism category, than for R and L (Figure 2.10). Similar to the KOG annotation, when comparing L and S relative to the R annotation at the level of KEGG subcategories, different distributions become apparent. Categories where a higher representation of L contigs is found include "Signal transduction", "Cell Motility", "Cell Communication" and "Immune System". Categories where a higher representation of S contigs is found include

"Carbohydrate Metabolism", "Lipid Metabolism", "Amino Acid Metabolism", "Metabolism of Other Amino Acids" and "Xenobiotics Biodegradation and Metabolism". All these subcategories, which differ by a factor of more than two between L and S, are presented in Table 2.8.

Table 2.8 Relative representation of contigs in selected subcategories of KEGG annotation

Main category	**Functional classification subcategory**	**Percentage representation**			
		R	**Large**	**Small**	**L>S**
Environmental information processing	Signal transduction	8.49%	13.55%	4.89%	2.77 X
Cellular processes	Cell motility	6.27%	12.26%	2.72%	4.5 X
Cellular processes	Cell communication	5.2%	15.48%	4.89%	3.17 X
Organismal systems	Immune system	6.77%	17.42%	4.89%	3.56 X
					S>L
Metabolism	Carbohydrate metabolism	17.66%	12.9%	26.09%	2.02 X
Metabolism	Lipid metabolism	7.84%	0%	21.74%	21.74 X
Metabolism	Amino acid metabolism	12.79%	14.19%	28.8%	2.03 X
Metabolism	Metabolism of other amino acids	5.03%	4.52%	11.96%	2.65 X
Metabolism	Xenobiotics biodegradation and metabolism	5.94%	1.3%	14.67%	11.29 X

In both the KOG and KEGG annotations (Tables 2.6 and 2.8), it appears that pathways related to signal transduction mechanisms have a higher representation amongst contigs form the L group, than from the S group. Also, from KEGG annotation, it is evident that more contigs from the L group were assigned to pathways involved in immunity, cell motility and cell communication. A higher percentage of contigs from the S group are assigned to the processes of carbohydrate, lipid and amino acid metabolism. In addition, the metabolism and biodegradation of xenobiotics and secondary metabolite biosynthesis, transport and catabolism are categories where S group contigs have a higher representation.

Annotation of differentially expressed sequences to general databases was complemented by database searches that were mollusc and species specific. Within the *Lottia* database, 63 of the 1174 sequences highly expressed in L and 42 of the 1324 sequences highly expressed in S BLASTed with a significant E-value ($< 10^{-10}$). These sequences and their associated functions in *Lottia gigantea* are summarized in tables 2.11 and 2.12 of the Appendix. None of the 8 sequences in group E, where equal expression between L and S was shown, had known similarities in the *Lottia* database. Annotation results of general and specified databases are summarised in Table 2.9.

Table 2.9 Annotation results for differentially expressed L and S *H. midae* sequences, across all databases

	L			S		
Database	**Hits to known sequence**	**Hits to known sequence with E<10^{-10}**	**Effective % high stringency hits**	**Hits to known sequence**	**Hits to known sequence with E<10^{-10}**	**Effective % high stringency hits**
GO	57.3 %	65.4 %	37.5 %	52.0 %	57.9 %	30.1 %
KOG	92.3 %	33.4 %	30.8 %	85.5 %	24.2 %	20.7 %
KEGG	47.5 %	85.6 %	40.7 %	36.7 %	44.3 %	16.3 %
LOTTIA	5.3 %	100 %	5.3 %	3.3 %	100 %	3.3 %
SEQ_LIT_NT	100 %	2.4 %%	2.4 %	100 %	0.5 %	0.5 %
SEQ_LIT_PROT	98.4 %	6.1 %	6.0 %	98 %	2.6 %	2.6 %
HALIOTIS_EST_NCBI	20.2 %	100 %	20.2 %	28.0 %	100 %	28.0 %
HALIOTIS_NT_NCBI	8.7 %	100 %	8.7 %	7.0 %	100 %	7.0 %
HALIOTIS_PROT_NCBI	91.7 %	15.1 %	13.8 %	87.1 %	14.0 %	12.2 %

From Table 2.9 it is apparent that, for L and S transcripts, similar percentages matched known sequences, throughout all annotation databases. Of these, varying percentages of "high-stringency hits" - with an E-value of ≤ 10^{-10} - were found: For L transcripts, less than 50 % high-stringency annotation matches were returned for the KOG, Seq_Lit and Haliotis_Prot databases. For S transcripts, less than 50% high-stringency annotation matches were returned for the KOG, KEGG, Seq_Lit and Haliotis_Prot databases. The low percentage of matches to known sequences within the Lottia and Haliotis_NT databases is noteworthy.

2.4 Discussion

This chapter describes the first comprehensive sequencing effort for *Haliotis midae*. Illumina sequencing-by-synthesis and assembly of the complete *H. midae* transcriptome resulted in the generation of 127 687 contigs from the reference transcriptome (from RNA isolated from 19 animals). From these, only contigs consisting of 80 or more base pairs (30 689 contigs) were used for further differential expression and annotation analyses. These contigs represent the first large collection of expressed sequence tags (ESTs) in *H. midae*.

EST collections have been reported for various molluscs in recent years. An EST library for the gastropod mollusc *Biomphalaria glabrata*, that was constructed from various tissues (haemocytes, the haemopoietic organ, the ovotestis, the head and foot, and brain tissue) resulted in 1843 non-redundant EST sequences (Lockyer *et al*., 2007). Other mollusc EST collections that were constructed from whole body tissue include that of the clam, *Ruditapes decussatus* (1938 ESTs; Tanguy *et al.*, 2008), the deep-sea mussel *Bathymodiolus azoricus* (2280 ESTs; Tanguy *et al.*, 2008) and the zhikong scallop *Chlamys farreri* (5123 ESTs; Wang *et al*., 2009). For another marine invertebrate, the barnacle *Balanus amphitrite*, construction of a cDNA library of ESTs from adult individuals resulted in a total of 609 unique sequences (De Gregoris *et al*., 2009). Sequencing of the neuronal transcriptome of *Aplysia californica*, resulted in a collection of 175 000 ESTs, which was suggested to comprise more than half of the protein coding genes in this gastropod mollusc (Moroz *et al*., 2006). Sequencing of the digestive gland of the abalone *Haliotis discus discus* resulted in the identification of 841 ESTs (Munasinghe *et al.*, 2006). All of these EST collections were however constructed using traditional capillary sequencing methods.

Conversely, increasing reports where next generation (NG) sequencing is used for construction of EST libraries, which usually include more ESTs than libraries constructed by traditional methods, are published. *De novo* assembly of the Glanville fritillary butterfly (*Melitaea cinxia*) transcriptome using 454 pyrosequencing data, resulted in 48 354 contigs with an average length of 197 bp, and 59 943 singletons (Vera *et al*., 2008). The neuronal transcriptome of the gastropod mollusc, *Lymnea stagnalis*, consisted of 7 712 distinct EST sequences, with a mean length of 847 nucleotides per EST (Feng *et al*., 2009). The liver transcriptome of the eelpout (*Zoarces viviparus*), a marine fish, was also sequenced with 454 pyrosequencing and the 36 110 contigs (from ~50 000 transcripts) contained in it is reported to represent 40 % of the total transcriptome of the organism (Kristiansson *et al.*, 2009).

While difficult to compare the size of the *H. midae* transcriptome with that of other organisms, where specific life stage or organ-level transcriptomes were sequenced, it is important to note the improved sensitivity of NG sequence technology, when compared to traditional Sanger-based sequencing (Huse

et al., 2007). When the transcriptome of floral tissue from *Arabidopsis*, the genome of which has been sequenced since 2000 by conventional capillary sequencing (Arabidopsis Genome Initiative, 2000), was sequenced using 454 technology, new transcript boundaries of 8 662 genic regions were mapped (Wall *et al*., 2009). Similarly, by using high-throughput 454 sequencing, 298 838 ESTs were identified in the transcriptome of the first larval stage of *C. elegans*. Almost 22 % of these ESTs are thought to represent previously unidentified genetic structures (Shin *et al*., 2008).

NG sequencing is particularly suitable when sequencing a transcriptome of an organism for the first time. However, a current shortcoming of NG sequencing is that it presents significant challenges in assembly and sequence accuracy due to short read lengths, especially for *de novo* assembly (Kristiansson *et al*., 2009; Wall *et al*., 2009). In expression studies, short read lengths also increase the potential for incorrect mapping to a known transcriptome, if available (Grigoriades *et al.*, 2009). New improved assembly approaches that may overcome these limitations are, however, becoming increasingly available (SHARCGS: Dohm *et al*., 2007; SSAKE: Warren *et al.,* 2007; ABySS: Simpson *et al*., 2009; SOAPdenovo: Li *et al*., 2010; OASES: Schulz and Zerbino, 2010). The advantage that an increased number of transcripts, including ones that represent novel transcribed elements, can be identified, and that this technology is attainable at a fraction of the cost of previous large-scale sequencing efforts, makes NG sequencing a preferred modern technology.

Thirty-seven percent of the reference *H. midae* transcriptome was assembled into contigs that were used for further analyses. The other 63 % consisted of singletons. These transcripts were reads that could not be matched to any other reads. Although singletons are still useful for annotation purposes, they were excluded from further analysis in this study, in order to streamline the annotation process. In other words, only transcripts that were at least constructed of two reads (contigs) were included in further analyses.

When compared to the eelpout liver transcriptome that has a distribution of 67.5 % contigs and 32.5 % singletons (Kristiansson *et al*., 2009), the assembly of the *H. midae* transcriptome returned a rather large proportion of singletons. Conversely, the assembly of the *Arabidopsis* floral tissue transcriptome returned a similar distribution with ~10 % contigs and ~90 % singletons (Wall *et al*., 2009). Also following *de novo* assembly of the transcriptome of pine (*Pinus contorta*) needles and conelets, a relatively small portion of reads (48 %) was assembled into contigs (Parchman *et al*., 2010). The same phenomenon was noted after assembly of the tick, *Ixodes scapularis,* genome where 62% of the total trace reads were singletons (NIAID Bioinformatics Resource Center, 2010). Such a large proportion of singletons may be ascribed to a high level of polymorphism within the sequenced population of individuals. Singletons could also result from contaminants or insufficiently trimmed reads from the

sequencing or from mis-assemblies and repeats present during assembly (Salzberg *et al*., 2006; Miller *et al*., 2010; Parchman *et al*., 2010). Transcriptome assembly is also encumbered by alternative splicing and *de novo* assembly is complicated by long repeat regions (Parchman *et al*., 2010). The assembly program therefore may be instrumental in the distribution of sequences between contigs and singletons. However, while more stringent filtering can improve the quality of assembly; important data may be lost in this way (Huse *et al.,* 2007; Parchman *et al*., 2010).

In the present study, the high proportion of singletons may be explained by a possible high level of polymorphism and difficulties with assembly. As assembly software continues to improve in terms of power and efficiency, it will likely be able to handle large data sets, more repetitive sequences and less homogeneous samples (Miller *et al*., 2010). The sequences assembled into contigs in this study were used for annotation and differential expression analyses.

Following KOG annotation of the reference transcriptome, 2350 (~89 %) of the 2654 annotated contigs could be assigned a function. The number of predicted genes that have only general function prediction or an unknown function (~11 %) is comparable to that reported for the sea urchin, *Strongylocentrotus purpuratus* (8 %) (Goel and Mushegian, 2006) and the tick, *Ixodes scapularis* (15 %) (Rebeiro *et al*., 2006). The largest percentage (~13 %) of assigned contigs was classified into the "Signal transduction mechanisms" subcategory, followed by "Posttranslational modification, protein turnover, chaperones" (10.77 %), "Translation, ribosomal structure and biogenesis" (9.49 %) and "Cytoskeleton" (7.04 %). The "Metabolism" subcategories with greatest representation included: "Energy production and conversion" (6.14 %), "Carbohydrate transport and metabolism" (5.57 %) and "Amino acid transport and metabolism" (4.63 %). The complement of contigs with metabolism-related functions comprised 28.33 %. When compared to KOG annotation for the sea urchin genome, with its ~8 % metabolism related sequences (Goel and Mushegian, 2006), many more contigs were assigned metabolism-related functions in *H. midae*.

Following KEGG annotation of the reference transcriptome, the majority of contigs were assigned to the main categories "Metabolism" (43.48 %) "Genetic information processing" (41.17 %) and "Cellular processes" (18.23 %). Interestingly, the next best represented category was "Human diseases" (17.33 %). A similar observation was made after annotation of the neuronal transcriptome of *Aplysia californica* when 104 genes orthologous to human neurological disease genes were identified. In *Aplysia*, a model organism for cellular and systems neural science, this presents a unique opportunity to study molecular and cellular functions of these genes and develop models for neurodegenerative diseases (Moroz *et al.*, 2006).

Using NG sequencing, genome scale expression information can be obtained in a direct and cost-effective way from organisms that do not have a genome sequence and comprehensive microarray platform available (Wall *et al*., 2009). The millions of sequence reads that is made available by Illumina sequencing-by-synthesis have a large dynamic range similar to traditional microarray experiments. These sequence reads can be used to construct microarrays or to design qPCR experiments in order to verify gene expression. One example of such verification of assembled transcripts with corresponding oligonucleotide probes in a microarray is the construction of the eelpout microarray based on 40 % of the transcriptome, generated by NG sequencing. A high correlation between the developed eelpout microarray and the number of reads from the sequencing attested to consensus between the microarray and the transcriptome sequence data (Kristiansson *et al*., 2009). Another study that also reports validation of NG sequence quality and assembly by using microarrays is that of Vera *et al*. (2008), where 60-mer oligonucleotide microarray probe sequences were constructed from contigs and singletons generated by 454 sequencing and *de novo* assembly. An example where qPCR was used to validate differential expression observed in both MPSS and microarray expression experiments is the study of Grigoriades *et al.* (2009), where expression of a subset of sense-antisense transcript pairs were confirmed in breast cancer malignant cell-lines and solid tumours. Similarly, 454 sequencing was used to sequence the wasp brain transcriptome, in order to identify test genes for quantitative RT-PCR analysis (Toth *et al.*, 2007).

In a similar approach, *H. midae* transcriptome data was explored in terms of differential gene expression and annotation of differentially expressed transcripts, with the aim of later validating differential expression of selected genes with qPCR. Differential expression of the sequenced transcriptomes of L and S was investigated with mRNASeq analysis in the CLC Genomics Workbench v3.5.1 (CLCbio, 2009). Following differential expression analysis and annotation of differentially expressed sequences in the L and S *H. midae* groups, a distinct distribution of contig sequences amongst the annotation subcategories is apparent. On the basis of KEGG annotation, categories that are better represented by contigs from the L group include “Signal transduction mechanisms”, “Immune System”, “Cell Motility” and “Cell Communication”. The metabolic pathways associated with these categories are discussed in more detail in the following paragraphs.

“Signal transduction mechanism” pathways include the Wnt signaling pathway (involved in basic developmental processes), the Notch signaling pathway (involved in embryonic development), the Hedgehog signaling pathway (involved in morphogenesis and development), the VEGF signaling pathway (involved in angiogenesis) and the Calcium signaling pathway. The specific pathway, into which the most sequences grouped and that will be focused on in the remainder of this study, is the mitogen-activated protein kinase (MAPK) signaling pathway.

The MAPK signaling pathway is highly conserved and involved in various cellular functions such as cell proliferation, -differentiation, -migration and -death. The name mitogen-activated protein kinase acknowledges the fact that these enzymes had first been detected as mitogen-stimulated tyrosine phosphoproteins. MAPKs are directly activated by phosphorylation and have two activating phosphorylation sites, a tyrosine and a threonine (Pearson *et al*., 2001). MAPK pathways commonly occur downstream of ErbB receptors. The ErbB family of receptor tyrosine kinases couples binding of extracellular growth factors to intracellular signaling pathways that regulate biologic responses like proliferation, differentiation and cell motility. One example of such a receptor tyrosine kinase is epidermal growth factor receptor (EGFR) (Kanehisa Laboratories, 2010b). The factors in this pathway that matched to contigs from the L group include protein phosphatase 3, fibroblast growth factor receptor 2, MAP kinase interacting serine/threonine kinase, Ras-related C3 botulinum toxin substrate 1, filamin and heat shock 70kDa (HSP70) protein. MAP kinase interacting serine/threonine kinase will be investigated in more detail in Chapter 4.

The “Immune System” pathways that are regarded as relevant in this study and into which the most sequences grouped are the NOD-like receptor-signaling pathway and the Antigen processing and presentation pathway. Factors identified from the complement of L contigs involved in these pathways include caspase 8, baculoviral IAP repeat-containing, molecular chaperone HtpG, heat shock protein 90kDa (HSP90), heat shock protein70kDa (HSP70), protein disulfide isomerase family A and cathepsin L.

The intracellular NOD-like receptor (NLR) family is a family of pattern recognition receptors responsible for detecting pathogens and initiating immune responses. It plays an important role in the recognition of intracellular ligands. Bacterial peptidoglycan fragments may escape from endosomal compartments and become involved in driving the activation of MAPK, NF-κB (nuclear factor of kappa light polypeptide gene enhancer in B-cells), cytokine production and apoptosis. Some NLRs detect the presence of these bacterial fragments in the cytosol and initiate an immune response. Other NLRs induce caspase-1 activation through the assembly of inflammasomes (multiprotein complexes) which are critical for generating mature proinflammatory cytokines (Kanehisa Laboratories, 2010 c, d). From these pathways, HSP90 will be investigated further in Chapter 4.

The relevant “Cell Motility” pathway into which the most sequences grouped was the Regulation of actin cytoskeleton pathway. Factors involved in this pathway include fibroblast growth factor receptors, myosin-light-chain kinase and Ras-related C3 botulinum toxin substrate. Many sequences were also classified as being cytoskeletal proteins (actin, myosin, actinin alpha, filamin, spectrin beta, gelsolin, tropomyosins and tubulin beta) (Kanehisa Laboratories, 2010e).

The focal adhesion pathway is the “Cell Communication” pathway that contained the most sequences. Focal adhesion/cell-matrix adhesions are involved in various important biological processes like cell motility, cell proliferation, cell differentiation, regulation of gene expression and cell survival. Some focal adhesion constituents are signaling molecules, including different protein kinases. Cell-cell adherens junctions have the potential to limit cell movement and proliferation. The integrity cell adhesion molecules involved in adherens junctions is regulated by receptor tyrosine kinases and cytoplasmic tyrosine kinases. Gap junctions have intercellular channels that enable communication between neighbouring cells. Direct transfer of small molecules (ions, amino acids, nucleotides, second messengers and other metabolites) that are involved in metabolic transport, apoptosis, and tissue homeostasis is facilitated by these gap junctions. Factors involved in this pathway that matched to sequences from the L complement of contigs include various collagens, thrombospondin, tenascin, myosin-light-chain kinase, actin, actinin alpha, filamin, Ras-related C3 botulinum toxin substrate 1 and apoptosis regulator BCL-2 (Kanehisa Laboratories, 2010f). Of these, collagen and thrombospondin will be investigated in further detail in Chapter 4.

Selected sequences that belong to the abovementioned categories were targeted for further differential expression analysis using qPCR and the respective pathways will be discussed in more detail in that chapter (Chapter 4).

On the basis of KOG and KEGG annotation, categories that were better represented by contigs from the S group include “Carbohydrate transport and metabolism”, “Amino Acid Metabolism”, “Lipid transport and metabolism”, and “Xenobiotics Biodegradation and Metabolism”. For the latter two categories, there was a marked difference between the L and S, with a much greater representation in S (Table 2.9). The following paragraphs will elaborate on the interaction of metabolites between these two categories.

“Xenobiotics Biodegradation and Metabolism” includes the degradation of caprolactam, gamma-Hexachlorocyclohexane, Chloroacrylic acid, 1,2-Dichloroethane, 1,4-Dichlorobenzene, Benzoate, Geraniol, Fluorobenzoate, Naphthalene and anthracene and xenobiotic and drug metabolism by cytochrome P450. One of the organelles involved in sequestration and metabolism of natural toxins is the lysosome. Cathepsins are proteases that reside in the lysosome and that are involved in lysosomal proteolysis of endocytosed proteins (Jia *et al*., 2009). The majority of contigs (44 %) that were assigned functions in the “Immune system” category of small abalone could be identified as cathepsins.

The lysosomal proteases cathepsin B and cathepsin L were previously reported to be responsible for large variation in maintenance energy expenditure in mussels suffering partial starvation. The activities of lysosomal enzymes can be connected to 44 % of whole-body protein turnover. A study of mussel

susceptibility to summer mortality confirmed that cathepsin L levels were increased in mussels susceptible to summer mortality when compared to those resistant to it (Tremblay *et al*., 1998). Similar patterns of stunted growth and mortality during summer months have been reported for abalone (Winstanley, 1972; Prince *et al.,* 1987; Steinbeck *et al*., 1992). Periods during which summer mortality occurs are characterised by physiological stress and subsequent elevated levels of lysosomal enzymes (Tremblay *et al.,* 1998), which could ultimately lead to gamete degradation via lysosomal processes (Bayne *et al.*, 1987). Lysosomal membrane destabilisation during these conditions is also associated with increased intralysosomal protein catabolism and consequently increased concentrations of amino acids and small peptides as products of this protein breakdown (Bayne *et al*., 1981). Another result of destabilization of the lysosomal membrane is atrophy if the digestive cells (Moore, 1988), which impedes effective digestion and food conversion. This atrophy may be explained by a recent study confirming the involvement of cathepsins in apoptosis in scallops (Wang *et al*., 2008).

It is also interesting to note the presence of glutatione peroxidase amongst contigs highly expressed in small abalone and involved in lipid metabolism and xenobiotic degradation. Certain industrial chemicals, such as the pesticide gamma-Hexachlorocyclohexane, a lipophilic compound, are known to be stimulators of oxyradical production. Oxidative damage in molluscs may result from bioaccumulation of such xenobiotics due to lipid peroxidation, the oxidative degradation of lipids (Livingstone, 1990). During this process, free radicals remove electrons from the lipids in cell membranes, resulting in cell damage. Glutatione peroxidase plays an important role in protecting membranes from this damage (Livingstone, 1990; Jai *et al.*, 2009).

Another intracellular enzyme involved in xenobiotic metabolism that has a similar activity as glutathione peroxidase, is glutathione S-transferase (GST). GSTs are also among the contigs highly expressed in small abalone and involved in xenobiotic degradation, specifically the cytochrome P450 pathways. GST has the ability to detoxify organic lipoperoxides and provide protection against oxidative damage (Slatinská *et al*., 2008; Jai *et al*., 2009). In fish, GST is used as a biomarker to indicate aquatic environment pollution with municipal, industrial, agricultural or mining wastewater (Slatinská *et al.*, 2008). Both glutatione peroxidase and GST have also been identified in disk abalone, *Haliotis discus discus*, and led to the suggestion that further studies on abalone GSTs may elucidate their biological significance in cellular detoxification. This is proposed to be important to future improvement of fisheries by limiting accumulation of toxic wastes (Munasinghe *et al*., 2006).

The majority of contigs that were grouped into “Lipid transport and metabolism” belonged to the processes Fatty acid metabolism and Glycerolipid metabolism. The enzymes involved in fatty acid metabolism include 3-hydroxyacyl-CoA dehydrogenase, enoyl-CoA hydratase and acyl-CoA

dehydrogenase. These are all involved in β -oxidation, the process by which fatty acids are broken down to generate Acetyl-CoA, which enters the tricarboxylic acid cycle (or Krebs) cycle, where it is further oxidized to CO_2 and usable energy (adenosine triphosphate; ATP). The oxidation of fatty acids yields significantly more energy than the oxidation of carbohydrates (King, 2010). Periods of immune challenge is characterised by a reduction in energy intake and an increase in fatty acid oxidation to provide energy (Spurlock, 1997).

The elevated levels of cathepsins in the small abalone group may indicate a response to physiological stress due to xenobiotic exposure and an increased activity in immune response. Because of the diversion of nutrients away from growth in support of immune-related processes, immune challenge is considered a major obstacle for animals to achieve their genetic potential for growth (Spurlock, 1997). The increased fatty acid oxidation noted in small abalone is also likely to be an effect of this increased immune activity.

This study focused on differential gene expression between two sibling groups of abalone demonstrating significant differences in body weight, shell length and shell width. After sequencing and annotation of a reference transcriptome, differential gene expression was assessed by means of mRNASeq analysis (CLCBio) and resulting sequences were annotated using various annotation databases. The procedure resulted in identification of differentially expressed sequences, of which some will be further investigated by quantitative real-time PCR (Chapter 4).

A great amount of sequence data remains unexplored following the sequencing of the *H. midae* transcriptome. Annotation can still be refined and expression analysis focusing on other genes of interest can be performed. This data is, however, of limited use in answering questions about organ-specific gene expression, as RNA originating from whole animal soft tissue was used for this transcriptome sequencing. The sequence data generated here can be useful in comparative phylogenetic studies and also to supplement existing knowledge of marker assisted selection, molecular marker detection and QTL identification for *H. midae*.

Novaes *et al.* (2008) have demonstrated that nucleotide diversity of an organism can be sampled by the high-throughput sequencing of a pool of genotypes. A comprehensive approach using molecular markers and genomics to establish functional and causal links between genotype, phenotype and natural selection is thus attainable by using NG sequencing. This has been demonstrated in lake whitefish (*Coregonus* spp. Salmonidae) where a large data set of candidate SNP markers was gathered from 454 pyrosequencing. SNP distribution within coding regions was assessed and the effect of species divergence on allele frequencies evaluated. This research aims at linking NG sequencing marker information to previous genomic studies (QTL, eQTL, genome scan and gene expression) in order to

identify genomic locations involved in adaptive divergence (Bernatchez *et al*., 2010; Renaut *et al*., 2010). A similar approach in *Haliotis* can elucidate evolutionary relationships in abalone and population genetic status of specific species like *H. midae* as well as aid the identification of genes of interest in selective breeding programs of farmed abalone.

2.5 References

Altschul, S.F., T.L. Madden, A.A. Schäffer, J. Zhang, Z. Zhang, W. Miller, and D.J. Lipman. 1997. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Research.* 25:3389-3402.

Anderson, J.M. 2008. Databases Used By CAS Lecture notes, National Institute of Allergy and Infectious Diseases, Bioinformatics and Scientific IT Program, USA.

Ashburner, M., C.A. Ball, J.A. Blake, D. Botstein, H. Butler, J.M. Cherry, A.P. Davis, K. Dolinski, S.S. Dwight, J.T. Eppig, M.A. Harris, D.P. Hill, L. Issel-Tarver, A. Kasarskis, S. Lewis, J.C. Matese, J.E. Richardson, M. Ringwald, G.M. Rubin, and G. Sherlock. 2000. Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. *Nature Genetics.* 25:25-29.

Bayne, B.L., M.N. Moore, and R.K. Koehn. 1981. Lysosomes and the response by *Mytilus edulis* L. to an increase in salinity. *Marine Biology Letters.* 2:193-204.

Bayne, B.L., A.J.S. Hawkins, and E. Navarro. 1987. Feeding and digestion by the mussel *Mytilus edulis* L. (Bivalvia: Mollusca) in mixtures of silt and algal cells at low concentrations. *Journal of Experimental Marine Biology and Ecology.* 111:1-22.

BCICC/OCICB/NIAID. 2010. Desktop cDNA Annotation System (dCAS) v1.4.3, USA. [Online]. Available: http://exon.niaid.nih.gov/cas/ [Accessed 02 July 2010].

Bernatchez, L., S. Renaut, A.R. Whiteley, N. Derome, J. Jeukens, L. Landry, G. Lu, A.W. Nolte, K. Østbye, S.M. Rogers, and J. St-Cyr. 2010. On the origin of species: insights from the ecological genomics of lake whitefish. *Philosophical Transactions of the Royal Society* B. 365:1783-1800.

Bester, A.E., R. Slabbert, and M.E. D'Amato. 2004. Isolation and characterization of microsatellite markers in the South African abalone (*Haliotis midae*). *Molecular Ecology Notes.* 4:618-619.

Bester, A.E., R. Roodt-Wilding, and H.A. Whitaker. 2008. Discovery and evaluation of single nucleotide polymorphisms (SNPs) for *Haliotis midae*: a targeted EST approach. *Animal Genetics.* 39:321-324.

Blaxter, M. 2010. MolluscDB: A web interface for the MolluscDB Partigene database. [Homepage of The University of Edinburgh], [Online]. Available: http://xyala.cap.ed.ac.uk/NeglectedGenomes/MOLLUSCA/ [Accessed 28 May 2010].

Boutet, I., D. Moraga, L. Marinovic, J. Obreque, and P. Chavez-Crooker. 2008. Characterization of reproduction-specific genes in a marine bivalve mollusc: Influence of maturation stage and sex on mRNA expression. *Gene.* 407:130-138.

Cantacessi, C., M. Mitreva, A.R. Jex, N.D. Young, B.E. Campbell, R.S. Hall, M.A. Doyle, S.A. Ralph, E.M. Rabelo, S. Ranganathan, P.W. Sternberg, A. Loukas, and R.B. Gasser. 2010. Massively parallel sequencing and analysis of the *Necator americanus* transcriptome. *PLoS Neglected Tropical Diseases.* 4:e684.

Carninci, P., M. Nakamura, K. Sato, Y. Hayashizaki, and M.J. Brownstein. 2002. Cytoplasmic RNA extraction from fresh and frozen mammalian tissues. *BioTechniques.* 33:306-309.

CLCbio. 2009. CLC Genomics Workbench 3.5.1, Denmark.

Conesa, A., and S. Gotz. 2009. Blast2GO Tutorial. Bioinformatics and Genomics Department Prince Felipe Research Center, Valencia, Spain.

De Cássia Ruy, P, A.Z. Neto, J.C. Ruiz, G.C. De Oliveira, E. Aguiar, F.S. De Oliveira and D.A.S. Cipriano. 2010. Functional analysis of *Biomphalaria glabrata* ESTs. [Homepage of Genomics and Computational Biology Group, Fiocruz], [Online]. Available: http://www.snaildb.org/ [Accessed 27 May 2010].

De Gregoris, T.B., M. Borra, E. Biffali, T. Bekel, J.G. Burgess, R.R. Kirby, and A.S. Clare. 2009. Construction of an adult barnacle (*Balanus amphitrite*) cDNA library and selection of reference genes for quantitative RT-PCR studies. *BMC Molecular Biology.* 10:62-73.

DiGuistini, S., N.Y. Liao, D. Platt, G. Robertson, M. Seidel, S.K. Chan, T.R. Docking, I. Birol, R.A. Holt, M. Hirst, E. Mardis, M.A. Marra, R.C. Hamelin, J. Bohlmann, C. Breuil, and S.J.M. Jones. 2009. *De novo* genome sequence assembly of a filamentous fungus using Sanger, 454 and Illumina sequence data. *Genome Biology.* 10:R94.

Dohm, J.C., C. Lottaz, T. Borodina, and H. Himmelbauer. 2007. SHARCGS, a fast and highly accurate short-read assembly algorithm for *de novo* genomic sequencing. *Genome Research.* 17:1697-1706.

Dos Reis Falcão, V., A.P. Tonon, M.C. Oliveira, and P. Colepicolo. 2008. RNA Isolation method for polysaccharide rich algae: agar producing *Gracilaria tenuistipitata* (Rhodophyta). *Journal of Applied Phycology.* 20:9-12.

Evans, B.S., N.A. Sweijd, R.C.K. Bowie, P.A. Cook, and N.G. Elliott. 2004. Population genetic structure of the perlemoen *Haliotis midae* in South Africa: evidence of range expansion and founder events. *Marine Ecology Progress Series.* 270:163-172.

Feng, Z., Z. Zhang, R.E. Van Kesteren, V.A. Straub, P. Van Nierop, K. Jin, N. Nejatbakhsh, J.I. Goldberg, G.E. Spencer, M.S. Yeoman, W. Wildering, J.R. Coorssen, R.P. Croll, L.T. Buck, N.I. Syed, and A.B. Smit. 2009. Transcriptome analysis of the central nervous system of the mollusc *Lymnaea stagnalis*. *BMC Genomics.* 10:451-465.

Filiatrault, M.J., P.V. Stodghill, P.A. Bronstein, S. Moll, M. Lindeberg, G. Grills, P. Schweitzer, W. Wang, G.P. Schroth, S. Luo, I. Khrebtukova, Y. Yang, T. Thannhauser, B.G. Butcher, S. Cartinhour, and D.J. Schneider. 2010. Transcriptome analysis of *Pseudomonas syringae* identifies new genes, ncRNAs, and antisense activity. *Journal of Bacteriology.* 192:2359-2372.

Gambacurta, A., M.C. Piro, and F. Ascoli. 1993. Cooperative homodimeric hemoglobin from *Scapharca inaequivalvis* cDNA cloning and expression of the fully functional protein in *E. coli*. *FEBS Letters.* 330:90-94.

GO Consortium. 2009. GO Slim and Subset Guide. [Online]. Available: http://www.geneontology.org/GO.slims.shtml [Accessed 27 May 2010].

Goel, M., and A. Mushegian. 2006. Intermediary metabolism in sea urchin: The first inferences from the genome sequence. *Developmental Biology.* 300:282-292.

Green, P. 1994. Phrap:phragment assembly program. [Online]. Available: http://www.phrap.org/ [Accessed 02 July 2010].

Grigoriadis, A., G.R. Oliver, A. Tanney, H. Kendrick, M.J. Smalley, P. Jat, and A.M. Neville. 2009. Identification of differentially expressed sense and antisense transcript pairs in breast epithelial tissues. *BMC Genomics.* 10:324-336.

Guo, Y., J.M.C. Ribeiro, J.M. Anderson, S. Bour. 2009. dCAS: a desktop application for cDNA sequence annotation. *Bioinformatics*. 25:1195-1196.

Hayes, B., M. Baranski, M.E. Goddard, and N. Robinson. 2007. Optimisation of marker assisted selection for abalone breeding programs. *Aquaculture.* 265:61-69.

Hedgecock, D., J. Lin, S. DeCola, C.D. Haudenschild, E. Meyer, D.T. Manahan, and B. Bowen. 2007. Transcriptomic analysis of growth heterosis in larval Pacific oysters (*Crassostrea gigas*). *Proceedings of the National Academy of Sciences USA.* 104:2313-2318.

Huse, S.M., J.A. Huber, H.G. Morrison, M.L. Sogin, and D.M. Welch. 2007. Accuracy and quality of massively parallel DNA pyrosequencing. *Genome Biology.* 8:R143.

Illumina, Inc. 2008. Paired-end sequencing user guide for cluster station and genome analyzer. San Diego, USA.

Janse van Rensburg, M., and V.E. Coyne. 2009. The role of electron transport in the defence response of the South African abalone, *Haliotis midae*. *Fish and Shellfish Immunology.* 26:171-176.

Jia, X., Z. Zhang, G. Wang, Z. Zou, S. Wang, B. Huang, and Y. Wang. 2009. Expressed sequence tag analysis for identification and characterization of genes related to Tributyltin (TBT) exposure in the abalone *Haliotis diversicolor supertexta*. *Comparative Biochemistry and Physiology, Part D.* 4:255-262.

Kal, A.J., A.J. Van Zonneveld, V. Benes, M. Van den Berg, M.G. Koerkamp, K. Albermann, N. Strack, J.M. Ruijter, A. Richter, B. Dujon, W. Ansorge, and H.F. Tabak. 1999. Dynamics of gene expression revealed by comparison of serial analysis of gene expression transcript profiles from yeast grown on two different carbon sources. *Molecular Biology of the Cell.* 10:1859-1872.

Kanehisa, M., and S. Goto. 2000. KEGG: Kyoto Encyclopedia of Genes and Genomes. *Nucleic Acids Research.* 28:27-30.

Kanehisa, M., S. Goto, M. Hattori, K.F. Aoki-Kinoshita, M. Itoh, S. Kawashima, T. Katayama, M. Araki, and M. Hirakawa. 2006. From genomics to chemical genomics: new developments in KEGG. *Nucleic Acids Research.* 34:D354-D357.

Kanehisa Laboratories. 2010a. KEGG Overview. [Homepage of Kanehisa Laboratories], [Online]. Available: http://www.genome.jp/kegg/kegg1a.html [Accessed 27 May 2010].

Kanehisa Laboratories. 2010b. MAPK signaling pathway - *Homo sapiens* (human). [Homepage of Kanehisa Laboratories], [Online]. Available: http://www.genome.jp/kegg/pathway/hsa/hsa04010.html [Accessed 25 May 2010].

Kanehisa Laboratories. 2010c. NOD-like receptor signaling pathway - *Homo sapiens* (human) [Homepage of Kanehisa Laboratories], [Online]. Available: http://www.genome.jp/kegg/pathway/hsa/hsa04621.html [Accessed 25 May 2010].

Kanehisa Laboratories. 2010d. Antigen processing and presentation - Reference pathway [Homepage of Kanehisa Laboratories], [Online]. Available: http://www.genome.jp/kegg-bin/show_pathway?map04612 [Accessed 25 May 2010].

Kanehisa Laboratories. 2010e. Regulation of actin cytoskeleton - *Homo sapiens* (human) [Homepage of Kanehisa Laboratories], [Online]. Available: http://www.genome.jp/kegg/pathway/hsa/hsa04810.html [Accessed 25 May 2010].

Kanehisa Laboratories. 2010f. Focal adhesion - *Homo sapiens* (human) [Homepage of Kanehisa Laboratories], [Online]. Available: http://www.genome.jp/kegg/pathway/hsa/hsa04510.html [Accessed 25 May 2010].

Kanehisa, M., S. Goto, M. Furumichi, M. Tanabe, and M. Hirakawa. 2010. KEGG for representation and analysis of molecular networks involving diseases and drugs. *Nucleic Acids Research.* 38:D355-D360.

King, M.W. 2010. Fatty acid oxidation. [Homepage of IU School of Medicine], [Online]. Available: http://themedicalbiochemistrypage.org/fatty-acid-oxidation.html [Accessed 28 May 2010].

Kristiansson, E., N. Asker, L. Förlin, and D.G.J. Larsson. 2009. Characterization of the *Zoarces viviparus* liver transcriptome using massively parallel pyrosequencing. *BMC Genomics.* 10:345-355.

Li, Q., J. Shu, R. Yu, and C. Tian. 2007. Genetic variability of cultured populations of the Pacific abalone (*Haliotis discus hannai* Ino) in China based on microsatellites. *Aquaculture Research.* 38:981-990.

Li, R., H. Zhu, J. Ruan, W. Qian, X. Fang, Z. Shi, Y. Li, S. Li, G. Shan, K. Kristiansen, S. Li, H. Yang, J. Wang, and J. Wang. 2010. *De novo* assembly of human genomes with massively parallel short read sequencing. *Genome Research.* 20:265-272.

Livingstone, D.R., P.G. Garcia Martinez, X. Michel, J.F. Narbonne, S. O'Hara, D. Ribera, and G.W. Winston. 1990. Oxyradical production as a pollution-mediated mechanism of toxicity in the common mussel, *Mytilus edulis* L., and other molluscs. *Functional Ecology.* 4:415-424.

Lockyer, A.E., J.N. Spinks, A.J. Walker, R.A. Kane, L.R. Noble, D. Rollinson, E. Dias-Neto, and C.S. Jones. 2007. *Biomphalaria glabrata* transcriptome: Identification of cell-signalling, transcriptional control and immune-related genes from open reading frame expressed sequence tags (ORESTES). *Developmental and Comparative Immunology.* 31:763-782.

Massault, C., H. Bovenhuis, C. Haley, and D. De Koning. 2008. QTL mapping designs for aquaculture. *Aquaculture.* 285:23-29.

Miller, J.R., S.K. Koren, and G. Sutton. 2010. Assembly algorithms for next-generation sequencing data. *Genomics.* 95:315-327.

Moore, M.N. 1988. Cytochemical responses of the lysosomal system and NADPH-ferrihemoprotein reductase in molluscan digestive cells to environmental and experimental exposure to xenobiotics. *Marine Ecology Progress Series.* 46:81-89.

Moriya, Y., M. Itoh, S. Okuda, A. Yoshizawa, and M. Kanehisa. 2007. KAAS: an automatic genome annotation and pathway reconstruction server. *Nucleic Acids Research.* 35:W182-W185.

Moroz, L.L., J.R. Edwards, S.V. Puthanveettil, A.B. Kohn, T. Ha, A. Heyland, B. Knudsen, A. Sahni, F. Yu, L. Liu, S. Jezzini, P. Lovell, W. Iannucculli, M. Chen, T. Nguyen, H. Sheng, R. Shaw, S. Kalachikov, Y.V. Panchin, W. Farmerie, J.J. Russo, J. Ju, and E.R. Kandel. 2006. Neuronal transcriptome of *Aplysia*: Neuronal compartments and circuitry. *Cell.* 127:1453-1467.

Munasinghe, H., H.S. Kang, and J. Lee. 2006. Analysis of digestive gland expressed sequence tag library from the disk abalone, *Haliotis discus discus*. *Journal of the World Aquaculture Society.* 37:96-106.

NIAID Bioinformatics Resource Center. 2010. VectorBase: *I. scapularis*. [Homepage of NIAID Bioinformatics Resource Center], [Online]. Available: http://iscapularis.vectorbase.org/SequenceData/ [Accessed 25 May 2010].

Novaes, E., D.R. Drost, W.G. Farmerie, G.J. Pappas Jr, D. Grattapaglia, R.R. Sederoff, and M. Kirst. 2008. High-throughput gene and SNP discovery in Eucalyptus grandis, an uncharacterized genome. *BMC Genomics*. 9:312-326.

Parchman, T.L., K.S. Geist, J.A. Grahnen, C.W. Benkman, and C.A. Buerkle. 2010. Transcriptome sequencing in an ecologically important tree species: assembly, annotation, and marker discovery. *BMC Genomics.* 11:180-195.

Pearson, G., F. Robinson, T.B. Gibson, B. Xu, M. Karandikar, K. Berman, and M.H. Cobb. 2001. Mitogen-Activated Protein (MAP) kinase pathways: Regulation and physiological functions. *Endocrine Reviews.* 22:153-183.

Prince, J.P., T. Sellers, W. Ford, and S. Talbot. 1987. A survey of the abalone stocks of the Kent and Hogan groups in Bass Straight. Technical Report No. 24. Department of Sea Fisheries, Taroona, Tasmania.

Qi, L., and K. Akihiro. 2007. Sequences characterization of microsatellite DNA sequences in Pacific abalone (*Haliotis discus hannai*). *Journal of Ocean University of China.* 6:47-52.

Renaut, S., A.W. Nolte, and L. Bernatchez. 2010. Mining transcriptome sequences towards identifying adaptive single nucleotide polymorphisms in lake whitefish species pairs (*Coregonus* spp. Salmonidae). *Molecular Ecology*. 19:115-131.

Rhode, C., R. Slabbert, and R. Roodt-Wilding. 2008. Microsatellite flanking regions: a SNP mine in South African abalone (*Haliotis midae*). *Animal Genetics.* 39:328-332.

Ribeiro, J.M.C., F. Alarcon-Chaidez, I.M.B. Francischetti, B.J. Mans, T.N. Mather, J.G. Valenzuela, and S.K. Wikel. 2006. An annotated catalog of salivary gland transcripts from *Ixodes scapularis* ticks. *Insect Biochemistry and Molecular Biology.* 36:111-129.

Rosenkranz, R., T. Borodina, H. Lehrach, and H. Himmelbauer. 2008. Characterizing the mouse ES cell transcriptome with Illumina sequencing. *Genomics.* 92:187-194.

Salzberg, S., Yorke, J., Delcher, A. and Pop, M. 2006. Genome assembly validation [Homepage of Center for Bioinformatics and Computational Biology, University of Maryland, USA], [Online]. Available: http://www.cbcb.umd.edu/research/assembly_validation.shtml [Accessed 25 May 2010].

Schulz, M., and D. Zerbino. 2010. Oases: *De novo* transcriptome assembler for very short reads. [Online]. Available: http://www.ebi.ac.uk/~zerbino/oases/ [Accessed 10 June 2010].

Sekino, M., and M. Hara. 2007. Linkage maps for the Pacific abalone (Genus *Haliotis*) based on microsatellite DNA markers. *Genetics.* 175:945-958.

Shin, H., M. Hirst, M.N. Bainbridge, V. Magrini, E. Mardis, D.G. Moerman, M.A. Marra, D.L. Baillie, and S.J.M. Jones. 2008. Transcriptome analysis for *Caenorhabditis elegans* based on novel expressed sequence tags. *BMC Biology.* 6:30-44.

Simpson, J.T., K. Wong, S.D. Jackman, J.E. Schein, S.J.M. Jones, and I. Birol. 2009. ABySS: A parallel assembler for short read sequence data. *Genome Research.* 19:1117-1123.

Slabbert, R. 2010. Identification of faster growth rate quantitative trait loci within abalone, *Haliotis midae*, using comparative microsatellite bulked segregant analysis. Unpublished PhD dissertation, Stellenbosch University, Stellenbosch, South Africa.

Slabbert, R., N.R. Ruivo, Van Den Berg, N. C., D.L. Lizamore, and R. Roodt-Wilding. 2008. Isolation and characterization of 63 microsatellite loci for the abalone, *Haliotis midae*. *Journal of the World Aquaculture Society.* 39:429-435.

Slabbert, R., J. Hepple, A. Venter, S. Nel, L. Swart, N.C. Van den Berg, and R. Roodt-Wilding. 2010. Isolation and segregation of 44 microsatellite loci in the South African abalone *Haliotis midae* L. *Animal Genetics.* 41:332-336.

Slatinská, I., M. Smutná, M. Havelková, and Z. Svobodová. 2008. Biochemical markers of aquatic pollution in fish - Glutathione-S-Transferase. *Folia Veterinaria.* 52:129-134.

Spurlock, M.E. 1997. Regulation of metabolism and growth during immune challenge: an overview of cytokine function. *Journal of Animal Science.* 75:1773-1783.

Steinbeck, J.R., J.M. Groff, C.S. Friedman, T. McDowell, and R.P. Hendrick. 1992. Investigations into a mortality among populations of the California black abalone *Haliotis cracherodii*, on the central coast of California , USA. *In:* Abalone of the world: Biology, fisheries and culture. S. A. Shepherd, M. J. Tegner and S. A. Guzman del Proo, editors. Oxford, Fishing News Press: 407-426.

Tanguy, A., N. Bierne, C. Saavedra, B. Pina, E. Bachère, M. Kube, E. Bazin, F. Bonhomme, P. Boudry, V. Boulo, I. Boutet, L. Cancela, C. Dossat, P. Favrel, A. Huvet, S. Jarque, D. Jollivet, S. Klages, S. Lapègue, R. Leite, J. Moal, D. Moraga, R. Reinhardt, J. Samain, E. Zouros, and A. Canario. 2008. Increasing genomic information in bivalves through new EST collections in four species: Development of new genetic markers for environmental studies and genome evolution. *Gene.* 408:27-36.

Tatusov, R.L., N.D. Fedorova, J.D. Jackson, A.R. Jacobs, B. Kiryutin, E.V. Koonin, D.M. Krylov, R. Mazumder, S. Mekhedov, A.N. Nikolskaya, B.S. Rao, S. Smirnov, A.V. Sverdlov, S. Vasudevan, Y.I. Wolf, J.J. Yin, and D.A. Natale. 2003. The COG database: an updated version includes eukaryotes. *BMC Bioinformatics.* 4:41-54.

The Arabidopsis Genome Initiative. 2000. Analysis of the genome sequence of the flowering plant *Arabidopsis thaliana*. *Nature.* 408:796-815.

The Gene Ontology. 2010. An Introduction to the Gene Ontology. [Online], Available: http://www.geneontology.org/GO.doc.shtml [Accessed 27 May 2010].

Toth, A.L., K. Varala, T.C. Newman, F.E. Miguez, S.K. Hutchison, D.A. Willoughby, J.F. Simons, M. Egholm, J.H. Hunt, M.E. Hudson, and G.E. Robinson. 2007. Wasp gene expression supports an evolutionary link between maternal behavior and eusociality. *Science.* 318:441-444.

Tremblay, R., B. Myrand, and H. Guderley. 1998. Temporal variation of lysosomal capacities in relation to susceptibility of mussels, *Mytilus edulis*, to summer mortality. *Marine Biology.* 132:641-649.

U.S. Department of Energy Joint Genome Institute (DOE JGI). 2007. *Lottia gigantea* v1.0 [Homepage of U.S. Department of Energy Joint Genome Institute (DOE JGI)], [Online]. Available: http://genome.jgi-psf.org/Lotgi1/Lotgi1.home.html [Accessed 27 May 2010].

Van der Merwe, M., and R. Roodt-Wilding. 2008. Chromosome number of the South African abalone *Haliotis midae*. *African Journal of Marine Science.* 30:195-198.

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

Villesen, P. 2009. Welcome to FaBox (1.35) - an online fasta sequence toolbox. [Homepage of Aarhus University, Denmark], [Online]. Available: http://www.birc.au.dk/fabox [Accessed 27 May 2010].

Vosloo, D., A. Vosloo, P. Sommer, A. Laas, and J. Samuels. 2009. Differential gene expression in South African Abalone (*Haliotis midae*) in response to oxidative and thermal stress. *Comparative Biochemistry and Physiology, Part A*. 153:S75.

Wall, P.K., J. Leebens-Mack, A.S. Chanderbali, A. Barakat, E. Wolcott, H. Liang, L. Landherr, L.P. Tomsho, Y. Hu, J.E. Carlson, H. Ma, S.C. Schuster, D.E. Soltis, P.S. Soltis, N. Altman, and C.W. dePamphilis. 2009. Comparison of next generation sequencing technologies for transcriptome characterization. *BMC Genomics.* 10:347-366.

Wang, L., L. Song, J. Zhao, L. Qiu, H. Zhang, W. Xua, H. Li, C. Li, L. Wu, and X. Guo. 2009. Expressed sequence tags from the zhikong scallop (*Chlamys farreri*): Discovery and annotation of host-defense genes. *Fish and Shellfish Immunology.* 26:744-750.

Wang, S., Z. Sha, T.S. Sonstegard, H. Liu, P. Xu, B. Somridhivej, E. Peatman, H. Kucuktas, and Z. Liu. 2008. Quality assessment parameters for EST-derived SNPs from catfish. *BMC Genomics.* 9:450-461.

Warren, R.L., G.G. Sutton, S.J.M. Jones, and R.A. Holt. 2007. Assembling millions of short DNA sequences using SSAKE. *Bioinformatics.* 23:500-501.

Wilhelm, B.T., S. Marguerat, S. Watt, F. Schubert, V. Wood, I. Goodhead, C.J. Penkett, J. Rogers, and J. Bähler. 2008. Dynamic repertoire of a eukaryotic transcriptome surveyed at single-nucleotide resolution. *Nature.* 453:1239-1245.

Winstanley, R.H. 1972. Abalone and rock lobster mortality and abnormal water temperatures. *Tasmanian Fisheries Research.* 6:21.

Zerbino, D.R., and E. Birney. 2008. Velvet: Algorithms for *de novo* short read assembly using de Bruijn graphs. *Genome Research.* 18:821-829.

3 INVESTIGATING THE ESTABLISHMENT OF PRIMARY CELL CULTURES FROM *H. MIDAE* TISSUES

3.1 Introduction

3.1.1 Cell culture in marine molluscs

The practice of cell culture in marine invertebrate species, specifically marine molluscs, is not as well developed as in human, mammalian or fish species. Despite the need for cell cultures from marine invertebrates that are important in aquaculture, toxicology, biotechnology and the pharmaceutical industry, attempts made to initiate such cell cultures have been ineffective. No established cell lines have been reported for any marine invertebrate. The lack of knowledge pertaining to invertebrate cell physiology and biochemistry, their slow rate of proliferation *in vitro* and the lack of appropriate culture media to sustain cultures remain unsolved challenges in the quest for marine invertebrate cell lines (Rosenfield, 1993; Lebel *et al*., 1996; Rinckevich, 2005).

The isolation and maintenance of primary cell cultures for limited periods, ranging from a few days to several months, is a more readily used practice. Despite their disadvantage when compared to cell lines, namely the increased risk of contamination and the limited number of cells and lifetime of cultures, primary cell cultures provide suitable models for short term *in vitro* investigations. The increasing development of aquaculture has particularly stimulated the development of cell cultures from economically important marine mollusc species.

Badariotti *et al.* (2007) cultured oyster haemocytes to assist in the characterization of a chitinase homologue and mantle explant cultures of the mussel, *Lamellidens marginalis,* were maintained for 70 days in a study to evaluate mussel cell culture as a possible route to produce tissue-cultured pearls (Barik *et al*., 2004). Primary cell cultures from oyster heart and clam gill cells were used as experimental *in vitro* models to assess cytotoxicity of aquatic pollutants (Domart-Coulon *et al.*, 2000) and the digestive gland of the scallop, *Pecten maximus,* were also maintained *in vitro* for long enough to be useful in ecotoxicological studies (Le Pennec and Le Pennec, 2001). Mussel and abalone larval cell cultures were maintained for three and 10 - 12 weeks, respectively, to investigate myogenesis (Naganuma *et al*., 1994; Odintsova *et al.*, 2001, 2010). Other abalone primary cell cultures include mantle explants and haemocyte primary cell cultures of *Haliotis tuberculata* which were maintained for up to two weeks to investigate the control of biomineralization (Auzoux-Bordenave *et al*., 2007) and the effect of vertebrate growth factors on haemocyte physiology (Lebel *et al*., 1996). Abalone haemocyte primary cell cultures were also used to determine the transcription expression pattern of heat shock proteins HSP70 and HSP90, under conditions of thermal stress (Farcy *et al*., 2007) and to investigate the effect of insulin-like growth factor-I (IGF-1) addition on haemocyte collagen synthesis (Serpentini *et al*.,

2000). *Haliotis tuberculata* mantle cell cultures were also used to test the effect of the water-soluble matrix fraction from the nacre of the silver-lipped pearl oyster, *Pinctada maxima*, on mantle cell viability and alkaline phosphatase activity (Sud *et al*., 2001).

It is clear from these studies that much knowledge pertaining to the biology and physiology of molluscs can be gained from the use of primary cell culture. The application of primary cell culture to any organism that has not been used for cell culture before should, however, be preceded by the identification of suitable tissues, optimal initiation methods, optimal medium and maintenance practices and evaluation of contamination concerns.

3.1.2 Considerations for primary culture initiation

3.1.2.1 Choice of tissue

The animal from which the tissues are harvested should be well characterised in terms of its anatomy, physiology, habitat, life cycle and behavioural responses, before embarking on primary cell culture initiation. Species that are less sensitive to varying environmental temperature and water conditions will offer greater flexibility for the initiation and maintenance of cell culture. According to Evenden (2000), two important considerations driving the choice of tissue for cell culture are i) the developmental and differentiation properties of the tissue and ii) the intended purpose of the cultures.

Embryonic stem cells are pluripotent and have infinite proliferative potential; attributes that make them attractive cells to culture, but current efforts in culturing decapsulated embryos and embryo cells from marine invertebrates have had variable success (Bulgakov *et al*., 2002; Kuang *et al.*, 2002). Although many cell types from marine invertebrates do have wide morphogenic potentialities (such as multipotency, pluripotency and neoplasia), identification and characterisation of such cell populations *in vivo* is not yet optimal (Evenden, 2000; Rinckevich, 2005). Blast-like cells from the haemolymph of *Haliotis tuberculata* have been reported to exhibit characteristics of undifferentiated cells (Travers *et al.*, 2008). Although undifferentiated cell types have greater proliferation rates and are consequently assumed to be readily propagated *in vitro*, other cell types may be just as useful for primary cultures. Culture of committed progenitor cells and mature differentiated cells of marine invertebrates (as listed above in section 3.1.1) has been more successful. Tissue-specific research questions in abalone have been answered by using mantle cells (Sud *et al*., 2001; Auzoux-Bordenave *et al*., 2007), larval cells (Naganuma *et al*., 1994) and haemocytes (Lebel *et al*., 1996; Farcy *et al*., 2006, Auzoux-Bordenave *et al*., 2007). Still, the use of these primary cell cultures is not culminating in the development of established cell cultures. This is probably due to incomplete cell characterisation and information regarding cell biology, -physiology and -biochemistry and the yet unmet challenge of harnessing unlimited proliferative potential in marine invertebrate cells.

3.1.2.2 Primary culture initiation methods

Primary cell cultures can be initiated by tissue fragmentation (explants) or by dissociating the tissue to isolate individual cells. From tissue explants (aseptically excised tissue cut into small fragments) cells are allowed to migrate and adhere to the culture dish. After cells have started forming a monolayer, within four to six days, the explants are removed and cells are collected for *in vitro* experiments. Isolation of individual cells from tissue can be accomplished in several ways including mechanical, chemical and enzymatic dissociation methods (Mulcahy, 2000; Pomponi and Willoughby, 2000). The choice of dissociation method will be determined by the application of the dissociated cells and the nature of the tissue to be dissociated. When a cell population consisting of individual cells is required, the method of choice will be one which delays re-aggregation of cells; thus chemical or enzymatic dissociation (Mulcahy, 2000; Pomponi and Willoughby, 2000). To isolate single cells, the extracellular matrix is disrupted by proteolytic enzymes such as collagenase, dispase, pronase or trypsin, that digest extracellular matrix proteins or by chemical agents such as EDTA which bind the Ca^{2+} on which cell-cell adhesion depends (Mulcahy, 2000; Alberts, 2002). During all types of dissociation, cells are vulnerable to mechanical and toxic damage and therefore care should be taken to limit the duration of the dissociation procedure as much as possible (Pomponi and Willoughby, 2000). Examples where cell dissociation and tissue fragmentation has been applied to molluscan tissues for cell culture are presented in Table 3.1.

Table 3.1 Dissociation methods reported for other mollusc primary tissue culture trials

Species (common name)	Tissue	Fragmentation/Dissociation method	Application	Reference
Haliotis rufescens (red abalone)	Trochophore and veliger larvae	Enzymatic (porcine trypsin)	Myogenesis	Naganuma *et al*., 1994
Crassostrea virginica (Eastern oyster)	Ventricle	Enzymatic (pronase)	Improvement of attachment and spreading of cells towards development of bivalve cell line	Buchanan *et al.*, 1999
Crassostrea gigas (Pacific oyster)	Heart Mantle Digestive gland Embryo	Tissue fragmentation (explants), enzymatic/chemical (trypsin-EDTA), enzymatic (collagenase) Enzymatic/chemical (trypsin-EDTA), enzymatic (collagenase)	Establishment of cell culture systems for oyster and clam	Chen and Wen, 1999
Meretrix lusoria (hard clam)	Heart Foot Ovary Arterial bulb Mantle	Tissue fragmentation (explants) and enzymatic (trypsin or collagenase) Tissue fragmentation (explants) Tissue fragmentation (explants) Tissue fragmentation (explants)		Chen and Wen, 1999
Crassostrea gigas (Pacific oyster) *Ruditapes decussatus* (grooved carpet shell clam)	Heart Gill	Tissue fragmentation (explants) Chemical (Alsever) and mechanical	Cytotoxicity assessment	Domart-Coulon *et al*., 2000
Haliotis tuberculata (European abalone)	Mantle	Tissue fragmentation (explants), enzymatic (collagenase) and mechanical (stirring)	Development of mantle cell culture	Sud *et al.*, 2001

Mytilus edulis (common mussel)	Digestive gland Gill	Tissue fragmentation (explants) and mechanical (stirring and filtering)	Develop primary culture	Faucet *et al*., 2003
Mytilus galloprovincialis (Mediterranean mussel)	Gill	Enzymatic (dispase)	Characterization of mussel gill cells	Gómez-Mendikute *et al.*, 2005
Mytilus trossulus (bay mussel)	Trochophore and veliger larvae	Enzymatic (collagenase)	Investigation of myogenic and neuronal differentiation	Odintsova *et al*., 2001, 2010

3.1.2.3 Medium and maintenance

The culture medium significantly affects the success of cell culture experiments and because different cell types have different growth requirements, the most appropriate medium should be formulated through careful experimental observation (Qiagen, 2002). Various basal and complex media types have been developed that are often used as a basis from which to construct specified media. Examples include MEM (developed for primary and diploid culture), Leibovitz-L15 (developed for CO_2 free environments), TC100 (developed for insect cells) and RPMI 1640 (developed for human leukemic cells) (ECACC and Sigma-Aldrich, 2001). These media all contain inorganic salts that maintain osmotic balance of the cells and help regulate membrane potential. A balanced pH in the culture medium is accomplished by the natural buffering system where gaseous CO_2 balances with the CO_3 / HCO_3 content of the culture medium (in a controlled gaseous environment) or by adding the zwitterion, HEPES (4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid) to the medium. The inclusion of phenol red as a pH indicator helps the researcher decide when to replace the medium. When the colour changes to yellow the medium's pH is too low and when the colour changes to purple, it is too high (ECACC and Sigma-Aldrich, 2001).

Basal and complex media are supplemented prior to use, with serum, amino acids and antibiotics according to the cells' requirements. Serum (from fetal bovine, newborn calf, porcine or equine origin) is partially defined as consisting of a complex mixture of albumins, growth factors and growth inhibitors and it is used with variable success to support cell growth, increase the buffering capacity of the medium, protect against mechanical damage and neutralise toxins. However, serum may differ from batch to batch and there is also a risk of contamination associated with its use (ECACC and Sigma-Aldrich, 2001; Qiagen, 2002).

When experimenting with new cell cultures, such as aquatic invertebrate cell culture, it is often challenging to formulate an appropriate medium as most commercial media were developed for use in mammalian cell culture. These commercial media are used as a foundation and modified in terms of osmolarity, pH and nutritional components to simulate the aquatic environment, whether it is marine or freshwater. As little is known about the nutritional requirements of aquatic invertebrates it is still a trial and error process where a similar approach to mammalian cell culture is sometimes followed. The addition of serum or haemolymph extracts from the host organism has lead to varying success in cell culture for these types of cells (Brooks and Kurtti, 1971; Domart-Coulon *et al.*, 1994; Evenden, 2000). Insect haemolymph provides a source of proteins and peptides, B vitamins, lipids and sterols but it also contains wastes that might be inhibitory to cell growth (Brooks and Kurtti, 1971). Selected examples of basal media and supplements used for primary mollusc cell cultures and reported in the literature are summarised in Table 3.2.

Table 3.2 Media and supplements reported for other mollusc primary tissue culture trials

Species (common name)	Tissue	Medium	Supplements	Reference
Crassostrea gigas (Pacific oyster)	Heart	Modified Leibovitz L -15	100 U/ml penicillin G, 100 µg/ml streptomycin, 10 % fetal calf serum, haemolymph, bovine insulin, hydrocortisone, oestradiol-β, putrescine, type IV concanavalin A, PokeWeed mitogen, *E. coli* lipopolysaccharide, Multiplication-Stimulating-Activity (Sigma), rhbFGF, rhEGF, rhPDGF BB, rhTGFβ1, aprotinin, fetuin, catalase, gluthation-GSH, Ex-Cyte VLE, Egg yolk extract, myo-inositol, prostaglandin F2α	Domart-Coulon *et al*., 1994
Haliotis rufescens (red abalone)	Larvae	Modified Leibovitz L -15	25 mg/ml amphotericin B, 5 mg/ml Ampicillin, 1000 U/ml PenicillinG, 10 mg/ml Streptomycin Sulphate, 10 % fetal calf serum	Naganuma *et al*., 1994
Haliotis tuberculata (European abalone)	Haemocytes	Modified Hanks-199 medium	100 µg/ml streptomycin sulfate; 60 µg/ml penicillin G, 50 µg/ml gentamycin sulfate; 0.20 µg/ml amphotericin B, 8 µg/ml nystatin, L-glutamine (2 mM), concanavalin A (2 mM)	Lebel *et al*., 1996
Crassostrea virginica (Eastern oyster)	Ventricle	JL-ODRP-4 (balanced salt solution with various added components)	None	Buchanan *et al.*, 1999
Crassostrea gigas (Pacific oyster)	Heart Mantle Digestive gland Embryo	2 x Leibovitz's L-15 medium (osmolarity adjusted)	10 % fetal bovine serum, 500 µg/ml streptomycin and 500 IU/ml penicillin Oyster body fluid, pituitary gland extract from rabbit, rat, cow, or common carp, extract from oyster embryo or gonad,	Chen and Wen, 1999

Meretrix lusoria (hard clam)	Heart Foot Ovary Arterial bulb Mantle		oyster soft body fluid, insulin, fibroblast growth factor, epidermal growth factor, endothelial cell growth factor, and Nutridoma serial products, 100 µg/ml collagenase or pronase	
Crassostrea gigas (Pacific oyster) *Ruditapes decussatus* (grooved carpet shell clam)	Heart Gill	Modified Leibovitz L-15 medium Modified Leibovitz L-15 medium	10 % fetal calf serum , 1 % penicillin-streptomycin	Domart-Coulon *et al*., 2000
Mytilus trossulus (bay mussel)	Larvae	Mofied Leibovitz L-15 medium	2 % embryonic calf serum, 1.75 mg/l vitamin E, 5 mg/l insulin, 40 mg/l gentamycin	Odintsova *et al*., 2001
Haliotis tuberculata (European abalone)	Mantle	Modified L15-Leibowitz medium	50 U/ml Nystatin, 1.25 µg/ml amphotericin B	Sud *et al*., 2001
Mytilus edulis (common mussel)	Digestive gland Gill	Modified Leibovitz L-15 medium	10 % fetal bovine serum, 1 ml/l tylosin solution, 1 % gentamycin sulphate	Faucet *et al*., 2003
Mytilus galloprovincialis (Mediterranean mussel)	Gill	Leibovitz L-15 medium	1 mg/ml glucose, 50 µg/ml glucosamine, 1.7 mg/ml HEPES, 5 % newborn calf serum, 100 U/ml penicillin, 100 µg/ml streptomycin, 100 µg/ml neomycin, 100 µg/ml kanamycin	Gómez-Mendikute *et al.*, 2005
Haliotis tuberculata (European abalone)	Mantle Haemocytes	Modified Leibovitz L -15	100 U/ml penicillin, 100 µg/ml streptomycin, 200 µg/ml gentamycin and 1 µg/ml amphotericin B, calcitonin-related molecules (h-CGRP)	Auzoux-Bordenave *et al.*, 2007
Mytilus trossulus (bay mussel)	Larvae	Mofied Leibovitz L-15 medium	2 % fetal bovine serum, 50 mg/l insulin, 1.75 mg/l α-tocopherol-acetate	Odintsova *et al*., 2010

3.1.2.4 Contamination

Persistent contamination in mollusc cell cultures and the difficulties to overcome it is often reported (Mulcahy, 2000). Contamination in the form of mycoplasma, protozoa, bacteria and fungi is a major challenge when initiating primary cell cultures. Antibiotics such as Fungizone, Amphotericin B, Penicillin/Streptomycin and Gentamycin may be added to the culture medium. However, it is generally recommended to discard contaminated culture dishes as the use of antibiotics may be toxic to invertebrate cells or lead to the development of chronic resistant bacteria that are difficult to eradicate (Freshney, 1992; Domart-Coulon *et al.*, 1994; Buchanan *et al.,* 1999; Mothersill *et al*., 2000; ECACC and Sigma-Aldrich, 2001). The effects of mycoplasma infection are more subtle than those of bacteria and fungi and may cause reduced growth rate, morphological changes, chromosome abnormalities and alterations in amino acid and nucleic acid metabolism over an extended period. Mycoplasma testing by using commercially available mycoplasma test kits should be part of any cell culture laboratory's routine, as mycoplasma is not eradicated by the use of antibiotics (Freshney, 1992; ECACC and Sigma-Aldrich, 2001; Qiagen, 2002).

3.1.3 Cell characterization

Cell characterization *in vitro* is essential for the determination of cellular functions and understanding of cell behaviour, physiology and biochemical patterns, and can be approached in terms of morphology, karyotype, differentiation status, cytological properties, biochemical properties, immunocytochemical properties, cell viability and proliferation (Lyons-Alcantara, 2001; Rinkevich, 2005). New primary cultures are often described in terms of their morphology, general appearance and cell behaviour. Careful microscopic and electron-microscopic examination can reveal pseudopodia and other cytoplasmic protrusions, adhesion properties, phagocytotic behaviour, cytoplasmic inclusions and the nucleus to cytoplasm (N/C) ratio. N/C ratio can be used as an indicator of differentiation status, where a high nucleus to cytoplasm ratio is a morphological characteristic of undifferentiated, blast-like cells (Hansen, 1988; Lyons-Alcantara, 2001; Travers *et al.,* 2008). Density gradient centrifugation, including Percoll and Ficoll gradients, are often adopted to separate cells by size or density. This gives an indication of the various cell types isolated from a tissue during primary cell culture initiation (Pomponi and Willoughby, 2000; Alberts, 2002; Richelle-Maurer *et al*., 2003). Karyotype analysis by mitotic arrest and G banding can be implemented to confirm a stable karyotype, which is characteristic of the normal or undifferentiated status of cells (Lyons-Alcantara, 2001; Brimble *et al.,* 2004).

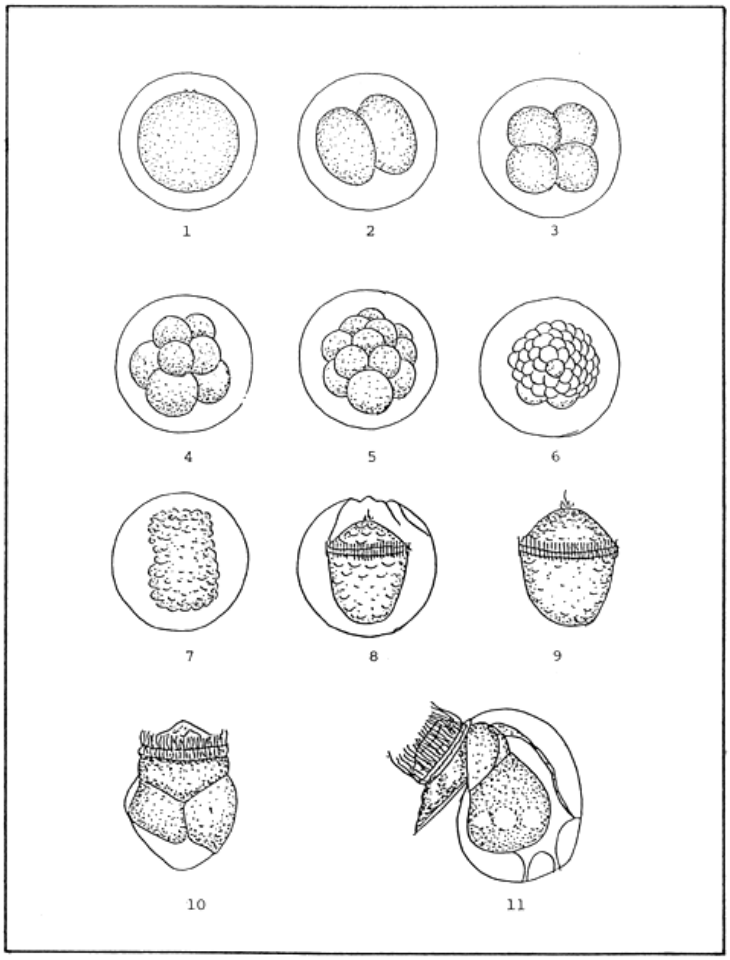
Several histological and cytological stains can be employed to identify general structure, intracellular polysaccharides, lipids, lysosomes, neutrophils, acidophils, basophilic granules and muco-polysaccharides (Cima *et al.,* 2000; Gómez-Mendikute *et al.,* 2005; Travers *et al.,* 2008).

Immunocytochemical assays such as the specific anti-CD34 antibody test are used to identify cells with stem cell-like qualities (Cima *et al.*, 2000). Viability and proliferation assays that measure vital functions associated with healthy or growing cells are routinely used to estimate the condition of a cell culture. These include the Trypan Blue exclusion test where viable cells do not take up the dye while non-viable cells do and are thus stained dark blue and the MTT or XTT assays where metabolically active cells cleave a tetrazolium salt into an orange coloured formazan product (Phillips and Terryberry, 1957; Domart-Coulon *et al.,* 1994; Roche Applied Science, 2003). Measurement of DNA synthesis is used to analyze cell proliferation *in vitro*. Labeled DNA precursors such as [^{3}H]-Thymidine or 5-bromo-2′-deoxyuridine (BrdU) or [H^3]-uridine are added to cells and their incorporation into genomic DNA during the S phase of the cell cycle is quantified (Odintsova and Khomenko, 1991; Roche Applied Science, 2003).

3.1.4 Tissues of origin for *H. midae* (embryos, larvae, haemocytes)

The abalone exhibits the most primitive condition of spawning within the phylum Mollusca. Ova and sperm are emitted in the exhalant water current and fertilization occurs at random in the surrounding seawater (Purchon, 1977). Spawning begins when the abalone's posterior end touches the substrate and the anterior is raised. The shell is lifted (approximately five to ten minutes prior to spawning in the female abalone) to the extent where the gonad is visible (Hahn, 1989). The adductor muscle contracts rapidly and these contractions compress the conical appendage between the foot and the shell, resulting in the gametes being released from the respiratory pores (Hahn, 1989; Henry, 1995). Upon spawning, the gametes are passed via a single gonad duct directly into the kidney and from there via the nephridiopore duct out through the respiratory holes in the shell (Purchon, 1977).

Abalone are poikilotherms and therefore seawater temperature plays an important role in the fertilization success, timing of hatch-out and early life stages of abalone. Larval development is faster in warmer water, but is terminated at high temperatures, while sub-optimal temperatures result in a longer planktonic larval phase (Hahn, 1989; Henry, 1995). Haliotid larvae are considered to be lecithotrophic since they do not feed before settlement, but survive on nutrients in the egg yolk (Hahn, 1989; Henry, 1995). Larval development is a dynamic process characterized by high biosynthetic activities, with up to 40 % of the whole-body protein being turned over each day (Vavra and Manahan, 1999). The larval development stages of most abalone species, including *H. midae*, as confirmed by microscopic observation, are quite similar to that described by Hahn (1989) and Courtois De Viçose *et al*. (2007) and summarized in Figure 3.1.



1: Polar bodies discharged (15 - 40 min)

2: 2-cell stage (~50 min)

3: 4-cell stage (80 - 90 min)

4: 8-cell stage (90 - 120 min)

5: 16-cell stage (120 - 160 min)

6: Morula stage (150 - 195 min)

7: Gastrula stage (4.5 - 6 hrs)

8: Trochophore in membrane (7 - 9 hrs)

9: Hatched trochophore larva (10 - 12 hrs)

10: Early veliger (15 hrs)

11: Late veliger (48 hrs)

Figure 3.1 Early developmental stages of abalone (Shallow Seafarming Research Institute, 1990)

The average size of fertilized *H. tuberculata* eggs is reported as 205 ± 8 µm and enclosed in a 1 µm thick glycoproteinaceous vitelline envelope which starts thinning after ten to 12 hours before it bursts at hatching (Shallow Seafarming Research Institute, 1990). Similar egg size is observed in *H. midae*. The size of developing gastrula larvae of *H. midae* is approximately 180 µm in length and 150 µm in width (Figure 3.2). The length of free swimming trochophore larvae of 166.6 ± 7.6 µm and width, 124.7 ± 4.75 µm as reported for *H. tuberculata* corresponds to the size in *H. midae* (Figure 3.2) (Vacquier *et al.,* 1990; Courtois de Viçose *et al*., 2007).

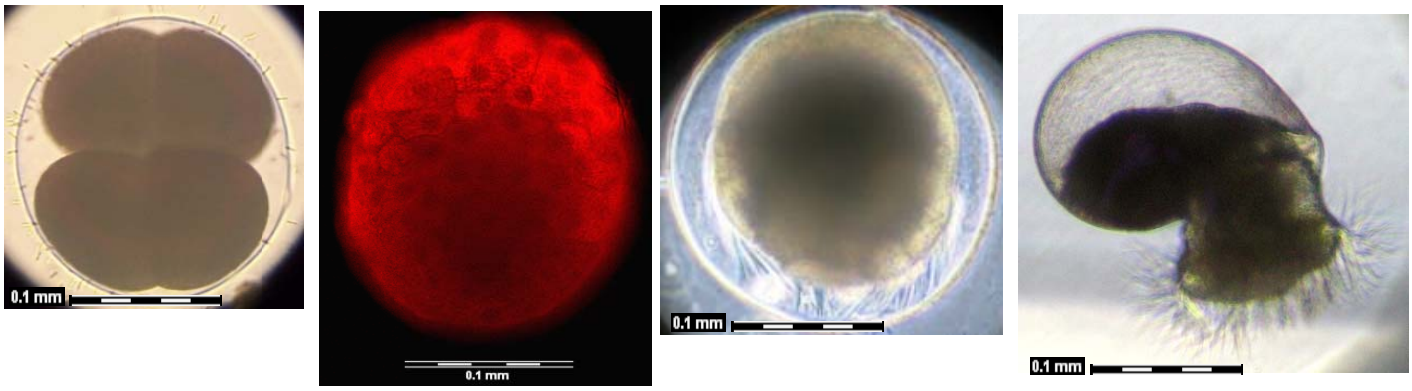


Figure 3.2 Embryo and larval stages of development in *H. midae*. A: 4 cell stage embryo (A. Roux, 2007) B: Gastrula stage abalone embryo stained with Mitotracker mitochondrial dye (M. van der Merwe, 2007). C: Trochophore larva in membrane and D: Veliger larva (A. Roux)

Following the account of Crofts (1937), the gastrula stage of embryogenesis is where macromeres and micromeres form and give rise to primitive endo-, meso- and ectoderm cells. The next stage is where cilia become visible and starts beating at the top of the embryo within the egg membrane (Courtois de Viçose *et al.*, 2007). At the veliger larval stage (Figure 3.2 D), cell differentiation into various cell types, including those comprising the larval retractor muscle, larval shell, foot mass, eye spot and velum with cilia has occurred (Courtois de Viçose *et al.*, 2007).

Mollusc haemocytes represent a heterogeneous population both in morphological and functional terms. They have been suggested to play roles in defense mechanisms (like cell-mediated cytotoxicity, encapsulation and phagocytosis), production of antimicrobial peptides, digestion, metabolite transport, wound and shell repair and secretion of extracellular matrix components (Lebel *et al*., 1996; Cima *et al.*, 2000; Humphries and Yoshino, 2003; Travers *et al.,* 2008). Abalone have an open circulatory system, in which haemolymph, consisting of haemocytes in plasma, circulates freely (Jorgensen *et al.*, 1984). To date no site (or sites) of haemocyte production in abalone have been identified (Travers *et al.*, 2008). Three types of abalone (*H. tuberculata*) haemocytes exist: large hyalinocytes, basophilic granulocytes and small blast-like cells (with a high nuclear-to-cytoplasm ratio). This ratio, together with the acidic cytoplasm implies that these blast-like cells could be undifferentiated cells (Travers *et al.*, 2008). It has been suggested that haemocytes may provide a suitable model to investigate the effects of factors involved in processes of proliferation and differentiation and various reports on haemocyte primary culture are available (Lebel *et al*., 1996; Serpentini *et al.*, 2000; Terahara *et al.*, 2005; Gagnaire *et al.*, 2006; Auzoux-Bordenave *et al.*, 2007). Small amounts of haemolymph can be collected from the pallial sinuses of adult abalone using a thin needle or maximal amounts can be collected by sacrificing the animal, incising the foot and allowing haemolymph to bleed out.

3.1.5 Initiative for establishing *H. midae* primary cell cultures

Due to the long generation time of *H. midae* and the cost and time involved in conventional genetic improvement practices (such as selection, crossbreeding and hybridization), biotechnological avenues such as gene transfer or transfection, where desired exogenous genetic material is inserted into a cell by viral or physical methods, may be a faster and more directed approach. Embryonic stem cell technology presents the possibility of combining genetic modifications *in vitro* and embryonic stem cell-mediated germline transmission, whereby transfected cells are successfully integrated into developing blastocyst embryos, retained in the animal through adulthood and passed on to their offspring (Shui and Tan, 2004). As the development of a pluripotent embryonic cell line stands central to effective germline transmission it was envisaged to investigate the likelihood of establishing such a cell line for *H. midae*. Other useful applications of *in vitro* cell cultures such as to further the understanding of shellfish pathologies and to study nutritional, toxicological and immune metabolic responses, also make investigations of this kind worthwhile (Villena, 2003). Towards identification of *H. midae* cells suitable for such applications, the establishment of larval and haemocyte primary cell cultures will be investigated in addition to embryonic primary cell cultures.

3.2 Materials and Methods

3.2.1 Cell collection, dissociation and culture initiation

Collection

All materials originated from a commercial abalone hatchery situated on the Southwestern coast of South Africa. Care was taken throughout tissue collection to sterilize all equipment and surfaces with 70 % ethanol and 1 % Virkon disinfectant. As far as possible, all procedures were carried out in a laminar flow cabinet. For embryos and larvae, normal embryo division was verified after fertilization by light microscopic observation and healthy developing embryos at the four to eight cell stage were collected at a temperature of 18 ˚C in filtered (0.2 µm) seawater supplemented with antibiotics (AB wash solution, Table 3.8, Appendix).

Embryos were transported within two hours from the abalone farm to the laboratory in this solution on ice. The solution was replaced with fresh AB wash solution in which embryos were allowed to develop to the gastrula or veliger stages respectively. The gastrula stage is the last stage before any visible functional differentiation takes place and therefore the gastrula embryo was considered a suitable target tissue for embryo cell culture. Embryos and larvae were dissociated using a variety of methods; chemical, enzymatic and mechanical (see Table 3.3).

For larval cell cultures, gastrula stage embryos were transferred to sterile glass bottles with filtered seawater and incubated at 18 °C overnight. The next morning hatched veliger larvae were collected, washed and dissociated according to the methods listed in Table 3.3. The various wash solutions for larvae consisted of the following:

- Hanks Balanced Salt Solution (HBSS) containing 1 % Penicillin/Streptomycin with or without 1 % Amphotericin B
- Filtered seawater (FSW) containing 1 % Penicillin/Streptomycin and 1 % Amphotericin B
- Calcium and Magnesium Free artificial Seawater Solution (CMFSS, see Table 3.8, Appendix for composition) with 15 mM EDTA and 1 % Penicillin/Streptomycin
- FSW containing 1 % Penicillin/Streptomycin, 2.5 µg/ml Amphotericin B and 50 µg/ml Gentamycin

The majority of larvae were hatched at the laboratory, as described above, with the only exception being the first trial, which was performed on the farm. During this trial fertilized embryos were transferred to buckets containing 5 L of 1 µm filtered, UV sterilized seawater at 18 °C to which 0.5 % Penicillin/Streptomycin was added, and allowed to hatch overnight. Larvae were transported to the laboratory in seawater with 0.5 % Penicillin/Streptomycin, collected, and washed using the same

procedures as for laboratory-hatched larvae. All larvae were collected by pouring the seawater in which they hatched through 100 µm mesh and then rinsing them with a small amount of FSW into sterile 50 ml conical tubes (Greiner Bio-One). After centrifuging at 300 x g, 17 °C for five minutes, the supernatant was aspirated carefully and the larval pellet was used for subsequent dissociation steps.

Before haemolymph collection, adult abalone of size 70 to 80 mm were starved and kept in clean ultra-violet sterilized seawater for three days. Dissection, haemolymph collection and culture initiation were performed in a laminar flow cabinet. The foot of each animal was wiped with 70 % ethanol before dissection, to clean the surface of mucous and debris. The foot was quickly dissected away from the shell and organs and placed, cut side down, in a 90 mm Petri dish where it was allowed to bleed for ten minutes. Care was taken not to sever the gut or gonad, as this would release contaminating material. The volume of haemolymph collected ranged from six to 12 ml per abalone.

Dissociation

All dissociation procedures were preceded by various washes in antibiotic solution and conducted at room temperature. The methods that returned the most promising results for embryo and larvae primary cell culture are summarized in Table 3.3. For embryo cell culture, these methods are the ones that delivered material where the capsule around the embryonic cell mass was partially or completely dismantled and where cell viability in the form of adherent cells or continued differentiation and development was apparent after initiation of culture. Since individually isolated cells were seldom observed after embryo dissociation, cell number was generally not quantified. For larval cell culture, the methods summarized in Table 3.3 delivered material where the larvae were dissociated into a mixture of cell clumps and single cells. Cell viability in the form of adhering cells or ciliated movement was apparent after initiation of culture. Where individual cells could be distinguished, cell number was estimated by counting on a haemocytometer before seeding into tissue culture plates.

Table 3.3 Dissociation methods used for *H. midae* embryos and larvae

Dissociation methods for embryos	**Dissociation methods for larvae**
1. <u>Enzymatic: different concentrations sperm</u> Embryos were exposed to 1 x 10^7 – 1 x 10^8 sperm per ml in FSW for 20 to 40 minutes; this was followed by washing the embryos seven times in FSW	1. <u>Mechanical: maceration</u> Larvae were macerated with a pestle in an eppendorf tube
2. <u>Enzymatic: sperm lysin</u> Different dilutions sperm lysin[a] was added to embryos for 30 to 40 minutes; this was followed by washing the embryos four times in FSW	2. <u>Enzymatic : trypsin</u> **2A**: Larvae were treated with 0.25 % trypsin at 17°C for three hours **2B**: Larvae were treated with 0.25 % trypsin at RT for one hour **2C**: Larvae were treated with 0.15 % trypsin at 17°C for three hours **2D**: Larvae were treated with 0.15 % trypsin at RT for one hour
3. <u>Enzymatic: hatch medium[b] + sperm</u> Embryos were treated with hatch medium for one hour and then exposed to 1 x 10^8 sperm per ml in FSW for 20 to 30 minutes; this was followed by washing the embryos four times in FSW	3. <u>Enzymatic: collagenase</u> Larvae were treated with 0.125 % collagenase in CMFSS for one hour with gentle agitation at RT
4. <u>Enzymatic: hatch medium + lysin</u> Embryos were treated with hatch medium for one hour and then exposed to undiluted sperm lysin for 30 minutes; this was followed by washing the embryos four times in FSW	4. <u>Chemical and mechanical</u> Larvae were treated with AB wash solution for 45 minutes followed by [CMFSS +15 mM EDTA +1 %PS] for ten minutes at RT. After collecting the tissue pellet by centrifuging at 300 x g for three minutes, it was resuspended in 3 ml [CMFSS +15 mM EDTA +1 %PS] and gently pipetted ten times
5. <u>Enzymatic: collagenase 1</u> Embryos were treated with 1 mg/ml collagenase in DS for 50 minutes with gentle magnetic stirring followed by 50 minutes of gentle agitation	
6. <u>Enzymatic : collagenase 2</u> Embryos were treated with 1 mg/ml collagenase in CMFSS +1 % PS for one hour with gentle magnetic stirring; cells were separated from clumps and debris by filtering through a 70 µm cell strainer; this was followed by washing the embryo cells once in CMFSS + 1 mM EDTA + 1 % PS	

FSW = Filtered seawater; DS = Dissociation Solution; CMFSS = Calcium and Magnesium free artificial seawater solution (pers. comm. Odintsova, 2007, see Table 3.8, Appendix); PS = Penicillin/Streptomycin; RT = Room Temperature

[a] Sperm lysin was prepared according to the method for inducing the acrosome reaction proposed by Lewis *et al.* (1982): Freshly spawned sperm from two mature abalone males was collected in seawater. After centrifuging at 1000 x g for ten minutes at 4 °C the pellet was resuspended in 20 to 50 volumes filtered seawater (4 °C). Calcium chloride ($CaCl_2$) was added to a final concentration of 0.3 M and incubated at 4 °C for 30 minutes after which it was centrifuged for 30 minutes at 10 000 x g at 4 °C. The

supernatant was centrifuged for another 30 minutes at the same conditions and stored at 4 °C. The resulting supernatant was used as crude lysin and various dilutions (in FSW) thereof. Lysin is a non-enzymatic acrosomal protein that, upon contact with the abalone egg, causes its vitelline envelope to unravel. It has been reported that increasing lysin concentrations leads to dissolving of the vitelline layers in a stoichiometric fashion (Lewis *et al.*, 1982; Shaw, 1995; Kresge *et al.*, 2001).

[b] The solution called "hatch medium" is the filtered seawater in which larvae hatched, re-filtered at 0.2 µm and stored at -20 °C. The suggestion that a natural hatching enzyme may promote dismantling of the vitelline layer was obtained from Sakai (1961) who reported the dissolving of the inner layer of chorions of Medaka fish (*Oryzias latipes*), after exposure to a mixture of hatching enzyme (prepared from pre-hatch embryos) and pancreatin. This procedure of dechorionation did not have any detrimental effects on the subsequent development of the Medaka embryos.

Culture initiation

Following dissociation (larvae and embryos) and quantification of cell number (where possible), cells were seeded in six-well or 12-well plastic tissue culture dishes (CELLSTAR, Greiner Bio-One), with or without Poly-D-Lysine coating, and specified culture medium was added (Table 3.4).

For haemocyte culture initiation, haemolymph was collected into a sterile 50 ml conical tube (Greiner Bio-One) and an equal volume of AB wash solution was added. The mixture was passed through a 70 µm cell strainer into a clean 50 ml tube, to eliminate debris, before two volumes (to one volume of haemolymph) of Alsever Solution (see Table 3.8, Appendix) was added to act as anticoagulation agent. A sample from the haemolymph mixture was counted using a haemocytometer and cells were seeded in the required densities in 96-well tissue culture plates. The cells were incubated at 18 °C and allowed to adhere to the surface of the tissue culture dish for one to two hours. Following this incubation, the solution was removed and replaced with a mixture of 50 % culture medium and 50 % AB wash solution. After another one to two hours in the incubator, the medium was changed to 100 % culture medium (100 µl per well).

3.2.2 Conditions for cell maintenance

Cells were incubated at 18 °C in a Hotpack low temperature incubator in specified culture medium. Tissue culture plates were sealed with parafilm and no artificial humidity- or CO_2 environment was simulated in the incubator.

The pH of seawater in the hatchery was verified as 7.6 - 7.8 and the osmolarity as 1030 - 1070 mmol/kg. This was taken into consideration when formulating culture media. Various formulations of culture media were tested for both embryo and larval cell culture maintenance (Media A to D, Table 3.4) before optimal cell culture media were adopted (medium C, following Domart-Coulon *et al.*, 1994). Haemocyte cell cultures were maintained under the same atmospheric conditions as embryo and larval cell cultures, in culture media C and D only (Table 3.4). After preparation, cell culture media were

filtered at 0.2 µm and stored at 4 °C for no longer than three weeks. Glutamine was only added prior to use.

Cultures were observed under light and phase contrast on an Olympus CKX31 inverted microscope. Cultures were inspected regularly for contamination. A rapid change in pH (visible by a colour change from red to orange/yellow in the medium), cloudiness of the medium, extracellular granularity and visible floating or moving material were all regarded as indicators of contamination (Freshney, 1992).

Table 3.4 Culture media formulations used for *H. midae* primary cell cultures

Culture medium A		**Culture medium B**	
1.5 % NaCl in Leibovitz-L15		Filtered, autoclaved seawater	
Add		**Add**	
Glutamine	2 mM	Fetal Calf Serum	10 %
Penicillin/Streptomycin	1 %	Penicillin/Streptomycin	1 %
Fetal Calf Serum	10 %	L-Glutamine	1 %
Osmolarity ±800 mmol/kg		Gentamycin	0.1 %
		+ or - 20 µg/ml bovine insulin	
Culture medium C (Domart-Coulon *et al.*, 1994)		**Culture medium D (ASW) (Stanford University, 1997: Artificial Seawater)**	
NaCl	10.1 g	NaCl	6.15 g
KCL	0.27 g	KCL	0.1675 g
$CaCl_2$	0.3 g	$CaCl_2.2H_2O$	0.34 g
$MgSO_4.7H_2O$	0.5 g	$MgSO_4.7H_2O$	1.5725 g
$MgCl_2.6H_2O$	1.95 g	$MgCl_2.6H_2O$	1.165 g
		$NaHCO_3$	0.045 g
Dissolve in 500 ml Leibovitz-L15			
		Dissolve in 250 ml dH_2O	
Add			
Penicillin/Streptomycin	1 %	**Adjust** pH to 7.6. Filter through	
Gentamycin	0.5 %	0.2 µm into sterilized bottle	
Amphotericin B	0.4 %		
		Add	
Adjust pH to 7.6. Filter through 0.2 µm into sterilized bottle		Penicillin/Streptomycin	1 %
		Gentamycin	0.5 %
		Glutamax	2 %
Add Glutamax (or L-Glutamine)	2 %		
Osmolarity ~ 1000 mmol/kg			

ASW = Artificial Seawater

For larval cell cultures, culture media C and D (which supported cell viability best of all tested media) were supplemented with different factors shown to enhance viability and proliferation (Table 3.5). These include: Bovine insulin - reported to stimulate cell viability in *Haliotis tuberculata* (Lebel *et al*., 1996) and *Crassostrea gigas* (Domart-Coulon *et al.*, 1994); bovine fibroblast growth factor (bFGF) and collagenase - shown to enhance growth in *C. gigas* and *Meretrix lusoria* cells (Chen and Wen, 1999) and

Fetal Calf Serum - reported to improve cell adhesion for *H. rufescens* larval cells (Naganuma *et al*., 1994).

Table 3.5 Media supplementation to study the effect on viability of cultured *H. midae* larval cells

Medium	**Added factor**	**Concentration**	**Presentation of results**
Medium A-1	Bovine fibroblast growth factor (bFGF)	10 ng/ml	Table 3.5
Medium C-1	FCS + *Bovine insulin*	10 % + *5 µg/ml*	Figure 3.6
Medium C-2	FCS + *Bovine insulin*	10 % + *50 µg/ml*	-
Medium C-3	FCS + *Bovine insulin* + *bFGF*	10 % + *5 µg/ml* + *2 ng/ml*	-
Medium C-4	FCS + *Bovine insulin* + *bFGF*	10 % + *50 µg/ml* + *2 ng/ml*	Figure 3.6
Medium C-5	Collagenase	100 µg/ml	Figure 3.7
Medium C-6	Fetal Calf serum (FCS)	5 %	Figure 3.7
Medium C-7	Fetal Calf serum (FCS)	10 %	Figure 3.7
Medium C-8	Collagenase + *FCS*	100 µg/ml + *5 %*	Figure 3.7
Medium C-9	Collagenase + *FCS*	100 µg/ml + *10 %*	Figure 3.7
Medium C-10	Amino acids	1 %	Figure 3.9
Medium D-11	Amino acids + *FCS*	1 % *+ 5 %*	Figure 3.9

The effect of Poly-D-Lysine coating on adherence behaviour of larval cells was also tested as it had been reported to enhance adherence and spreading in *Crassostrea virginica* cells over five days of culture (Buchanan *et al.*, 1999). For wells where cells were viable, half of the medium was replaced with fresh medium every three to four days. This was done carefully to avoid disturbing the cells.

For haemocytes, culture medium D was supplemented with the different factors summarized in Table 3.6 to investigate the effect of these known growth stimulants on haemocyte viability. These putative factors are all reported to result in an increase in metabolic activity in cultured mollusc cells (Domart-Coulon *et al*., 1994; Lebel *et al.,* 1996; Chen and Wen, 1999).

Table 3.6 Medium D supplementation to study the effect on viability of cultured *H. midae* haemocytes

Medium name	**Supplement/supplement combination**	**Concentration**
D-1	Bovine insulin	50 µg/ml
D-2	Epidermal Growth Factor (EGF)	25 ng/ml
D-3	Catalase	20 µg/ml
D-4	Collagenase	100 µg/ml
D-5	50 % Insulin + *50 % EGF*	25 µg/ml + 1*2.5 ng/ml*
D-6	50 % Insulin + *50 % Catalase*	25 µg/ml + *10 µg/ml*
D-7	50 % Insulin + *50 % Collagenase*	25 µg/ml + *50 µg/ml*
D-8	50 % EGF + *50 % Catalase*	12.5 ng/ml + *10 µg/ml*
D-9	50 % EGF + *50 % Collagenase*	12.5 ng/ml + *50 µg/ml*
D-10	50 % Catalase + *50 % Collagenase*	10 µg/ml + *50 µg/ml*

Medium D1 was identified as the medium affecting haemocyte growth the most and Medium D (Table 3.4) was subsequently supplemented with different bovine insulin concentrations of 0.25, 0.5, 1, 2, 4, 8, 16, 32, 64, 128 and 256 µg/ml (Media D-12 to D-22), to determine at which concentration it had an optimal growth stimulating effect on cultured haemocytes.

3.2.3 Viability assessment

The Trypan Blue (TB) exclusion test was employed to verify embryo cell viability after culture initiation. TB (0.4 %) was added to the cells and incubated for five minutes before microscopic examination. Viability was evident in cells with a brown colour on the interior and a radiant ring surrounding the cell membrane, while dead cells stained blue on the inside and lacked the radiant ring. Cultures where the majority of cells were dead were discarded.

Where factors were added to cell culture media (larval and haemocyte cultures), their effect on cell viability was monitored over time by haemocytometer counts using Trypan Blue, or by interpreting the results of XTT assays on specified days of culture. The XTT assay relies on the cleavage of a yellow tetrazolium salt into an orange formazan product by metabolic active cells (Roche Applied Science, 2003). Measurement of such metabolic activity over time can be regarded as indicative of cell viability or cell density variations. Five milliliter XTT (sodium 3´-[1- (phenylaminocarbonyl)- 3,4-tetrazolium]-bis (4-methoxy-6-nitro) benzene sulfonic acid hydrate) and 100 µl PMS (N-methyl dibenzopyrazine methyl sulfate) were thawed, combined and mixed thoroughly according to the manufacturers instructions. Fifty microliters of XTT/PMS mix was added to each well of the 96-well plate containing the cells in 100 µl of culture medium. The plate was wrapped in foil and returned to the incubator (18 °C) for six hours. After incubation, an absorbance reading was performed on the 96-well plate with a Bio-Rad

model 680 Microplate Reader using a measurement filter at 490 nm and a reference filter at 655 nm. Each plate was discarded after the measurement was taken.

3.2.4 Statistical analysis

Results of viability assays were reported by using t-tests and Analyses of Variance (ANOVAs). ANOVAs that confirmed a significant difference between treatment groups ($P < 0.05$) were followed by pairwise t-tests to evaluate the difference between individual treatment groups and the untreated control. Prior to performing a t-test, an F-test for equal variances was performed. According to the results of this F-test, the appropriate t-test (assuming equal or unequal variances) was performed. T-tests that returned P-values of < 0.05 were regarded as confirmation of significant differences between treated and control groups.

3.3 Results

3.3.1 Embryo cell cultures

Effect of dissociation protocol on cell morphology

Only enzymatic dissociation trials for embryos are reported here, as preliminary trials with mechanical and chemical dissociation seemed to be too harsh for embryos, causing severe damage to embryo cells.

Embryo dissociation methods where higher concentrations of sperm and lysin were used, rendered cells that either continued differentiating into larval cells with beating cilia (some of which were unable to hatch out) or, into cells adhering to the surface of the poly-D-lysine coated tissue culture plates – as single cells or small groups of cells (Figure 3.3). Enzymatic dissociation by collagenase treatment caused partial to complete removal of embryo capsules from the gastrula cell mass. During subsequent days of culture, these cell masses seemed to continue along the larval cell differentiation pathways. Pieces of ciliated larvae were observed moving about in the culture dish whilst other embryo/larval cells adhered to the culture dish and exhibited contractile behavior.

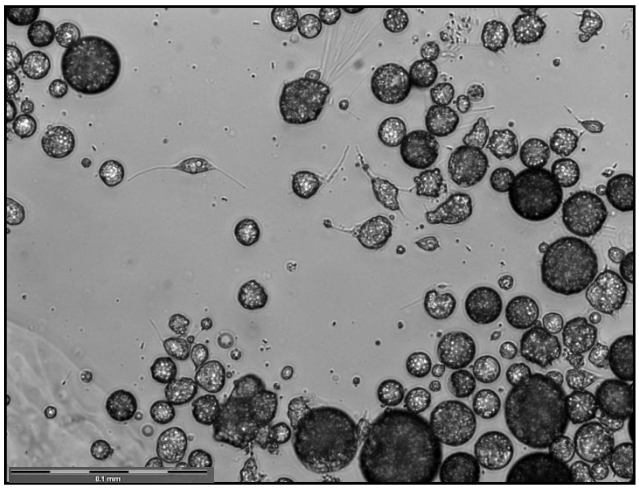


Figure 3.3 *H. midae* embryo cells in culture medium A (Table 3.4), four days after dissociation with abalone sperm (1 x 10^8 sperm/ml for 30 minutes) in a poly-D-lysine coated six-well tissue culture plate

Cells cultured in culture medium B (Table 3.4) appeared more viable and remained adhered to the tissue culture dish for longer than cells cultured in culture medium A (six days vs. four days). Adherent cells appeared fibroblast-like and some showed signs of contractile behavior. All embryo cells were subject to progressive protozoan, bacterial and fungal contamination that finally claimed the cultures after six days.

Microscopic observation confirmed the progressive development of what is here referred to as 'embryo cell culture' into a variety of differentiated larval cell types. The cultures arising from the methods described above can thus not be regarded as embryo primary cell culture exclusively. The process of

pre-programmed cellular differentiation in early development is a complex one and has not been described in detail for *Haliotis.* Although methods exist to arrest cells in an embryonic, proliferative state (Smith, 2001; Nichols and Ying, 2006), these methods were not investigated in this study as a viable, healthy collection of embryo cells could not be established due to persistent contamination. The early termination of trials to establish healthy embryo cell cultures is also the reason why other media compositions, later adapted for larval and haemocyte cell cultures (media C and D), were not tested on embryo cell cultures.

3.3.2 Larval cell cultures

Effect of dissociation protocol on cell morphology and maintenance

Larval cells in culture tended to separate into two groups: cells adhering to the tissue culture dish within the first 24 hours of culture, and cells that clustered together and remained in suspension. Both populations were characterized by movement; the adherent cells by contractile movement and the suspension cells by ciliated movement. Adherent larval cells had fibroblast-like, epithelial-like and large round morphologies. Groups of adherent epithelial-like cells seemed to communicate to neighbouring clusters through thin tube-like protrusions (Figure 3.4). Such communication was evident during contraction, when two neighbouring cell clusters would contract in succession. Suspension cells were characterized as clusters of small cells, some of which were ciliated.

Dissociation method 1 resulted in a low yield of cell number and formed cell cultures that remained viable for 19 days. Dissociation method 2 (A – D) resulted in higher cell numbers forming cell cultures that remained viable for 12 to 21 days. Larval cells dissociated with method 3 remained viable for eight days and those dissociated with method 4 remained viable for ten days. Methods 3 and 4 had similarly high yields of cell number as dissociation method 2. Viability of resultant cell cultures decreased over time, regardless of the dissociation method used. Because of the similar success of dissociation methods 2, 3 and 4, all three are suitable for use in future experiments. These methods delivered material where the larvae were dissociated into a mixture of cell clumps and single cells. Cell viability in the form of adhering cells or ciliated movement was apparent after initiation of culture.

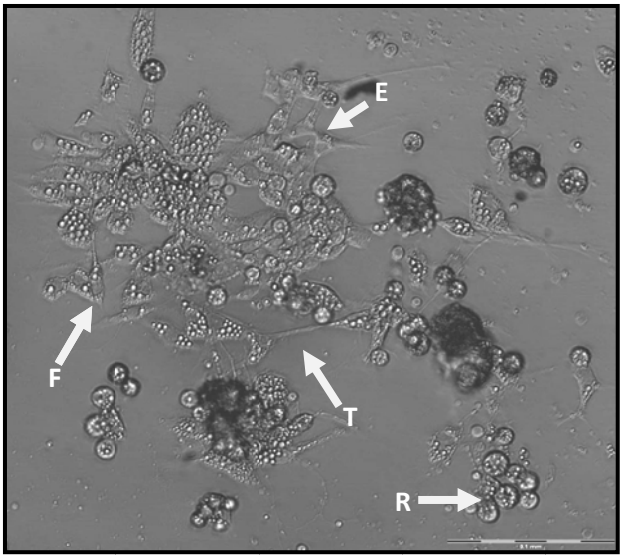


Figure 3.4 *H. midae* larval cells at day 4 of culture in culture medium C showing epithelial-like (E), fibroblast-like (F) and large round (R) morphologies. Clusters are connected by thin tube-like structures (T)

Effect of medium composition

The effects of different media compositions are summarized in Table 3.7. Effects that are mentioned in Table 3.7, but not elaborated on in the text are those that where only monitored by microscopic observation and thus not quantified in terms of viability. Where viability was quantified, it was initially done with Trypan Blue exclusion and later with XTT.

Table 3.7 Summary of the effects of different combinations of dissociation and maintenance protocols on viability of *H. midae* larval cell cultures

Dissociation method (From Table 3.3)	**Medium (From Tables 3.4, 3.5)**	**Duration, viability and behaviour of cell cultures**
2A	A (Table 3.4)	19 days survival, irregular attachment and contraction
2B	A (Table 3.4)	14 days survival, irregular attachment and contraction
2B	A-1 (Table 3.5)	12 days survival and irregular attachment, no record of contraction
2B	C-1 and C-4 (Table 3.5)	12 days survival and irregular attachment, no record of contraction
2D	A (Table 3.4)	19 days survival, 14 days irregular attachment and contraction
3	C-5 - C-9 (Table 3.5)	Eight days + survival and irregular attachment, no record of contraction
4	C-10 and D-11 (Table 3.5)	Ten days +survival and irregular attachment, no record of contraction

Larval cell culture, initiated using dissociation method 2B (enzymatic, Table 3.3) and maintained in culture medium A-1 (Table 3.5) was observed over the course of 12 days to determine the effect of bovine fibroblast growth factor (bFGF). Cell viability was quantified by trypan blue staining and counting

cells on a haemocytometer. The result of a homoscedastic t-test (F-test for equal variances, $P > 0.05$) indicated that there was no significant difference between the mean viability measured over 12 days of cultures treated with bFGF and control cultures ($P > 0.05$) (Figure 3.5). It was therefore concluded that 10 ng/ml bFGF did not have a significant effect on viability of cultured *H. midae* larval cells over a period of 12 days of culture.

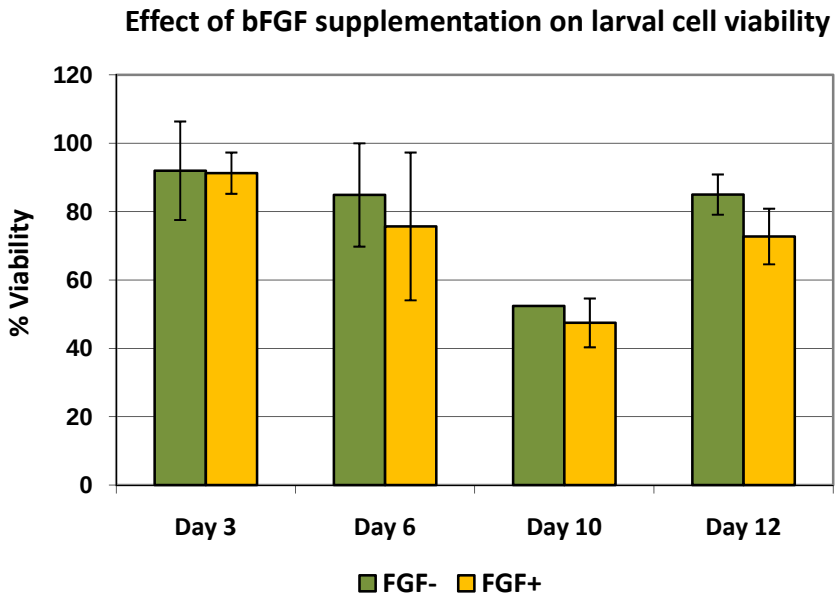


Figure 3.5 Change in cell viability in cultured *H. midae* larval cells over twelve days: FGF+ = Medium A-1 with 10 ng/ml bFGF. FGF- = Medium A-1 with no added bFGF. Error bars indicate the standard error.

Larvae dissociated with method 2B (enzymatic, Table 3.3) and maintained in culture media C-1 – C-4 (different combinations of bovine insulin and bFGF, Table 3.5), were observed over the course of 12 days to compare the different supplemented media. On day 12, too few cells remained attached to the tissue culture plates to perform a sensible count. Cells were counted on days three, six and nine of culture. The change in larval cell viability (determined with Trypan Blue exclusion), only in significantly differing culture media (media C-1 and C-4) is reported in Figure 3.6. T-tests report a significant difference in cell viability between day three and day nine with both media C-1 and C-4 demonstrating a significant decrease in viability at day nine of culture for both media ($P < 0.01$). Subsequent t-tests reported no significant difference in viability between the two media on day three, but on day nine significantly less viability was observed in medium C-4 cells than medium C-1 cells ($P < 0.01$), indicating that over time the medium that sustained cell viability the best of the two, was medium C-1.

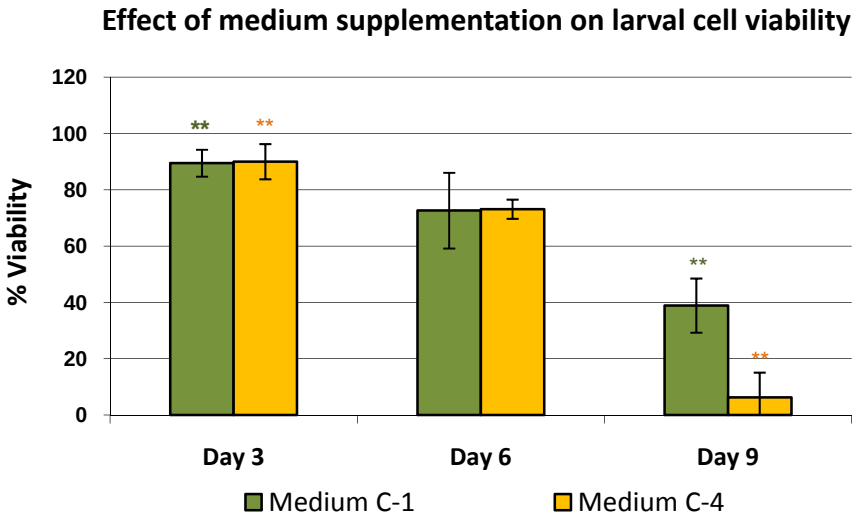


Figure 3.6 Change in cell viability in cultured *H. midae* larval cells over nine days: Medium C-1 = Culture medium C with 10 % FCS and 5 µg/ml bovine insulin. Medium C-4 = Culture medium C with 10 % FCS and 50 µg/ml bovine insulin and 2 ng/ml bFGF. Error bars indicate the standard error. Significant difference from Day 3 at $P < 0.01$(**)

Larvae that were dissociated with method 3 (enzymatic, Table 3.3) and maintained in culture media C-5 - C-9 (different combinations of added fetal calf serum and collagenase), were observed over the course of eight days. Viable cells (determined with Trypan Blue exclusion) were counted using a haemocytometer on days one, four and eight of culture. Results from a two way ANOVA reported no significant difference in cell viability between the media ($P > 0.05$), and that overall viability decreased significantly over the eight days of culture ($P < 0.01$) (Figure 3.7).

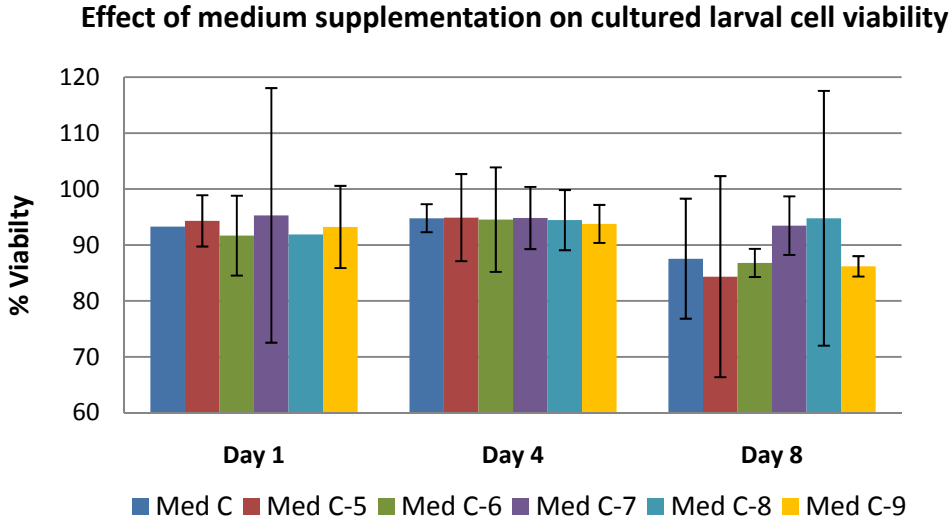


Figure 3.7 Change in cell viability in cultured *H. midae* larval cells over eight days: Med C = control; Med C-5 = Med C + 100 µg/ml collagenase; Med C-6 = Med C + 5 % FCS; Med C-7 = Med C + 5 % FCS + 100 µg/ml collagenase; Med C-8 = Med C + 10 % FCS; Med C-9 = Med C + 10 % FCS + 100 µg/ml collagenase. Error bars indicate the standard error

Primary cultures resulting from dissociation method 4 (chemical and mechanical, Table 3.3) was cultured in 96-well tissue culture plates (Greiner Bio-One) at a density of 15 500 cells per well. This cell number was determined from the linear relationship of increase in absorbance according to increased cell number in an XTT assay over seven days of culture and was regarded as a reliable mean value for optimal cell viability (Figure 3.8). Culture media C-10 and D-11 (supplemented with 1 % amino acids, Table 3.5) were compared to control media C and D, respectively (no supplements, Table 3.4) over a period of ten days, taking absorbance measurements following XTT assays on days three, five, seven and ten. The result from this experiment is represented graphically in Figure 3.9.

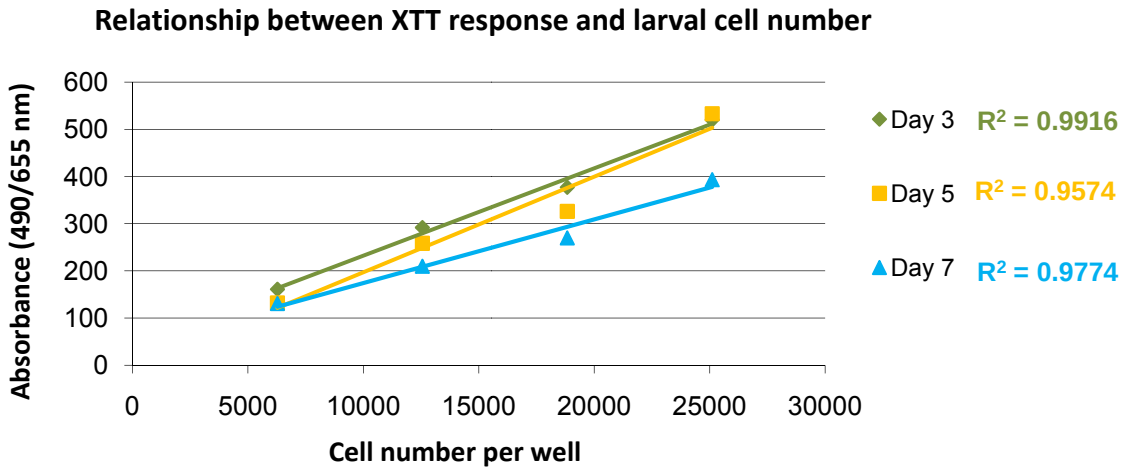


Figure 3.8 Linear relationship of increasing absorbance with increase in *H. midae* larval cell number dertermined by XTT assay over seven days of culture. Cells were cultured in 96-well plates in medium C

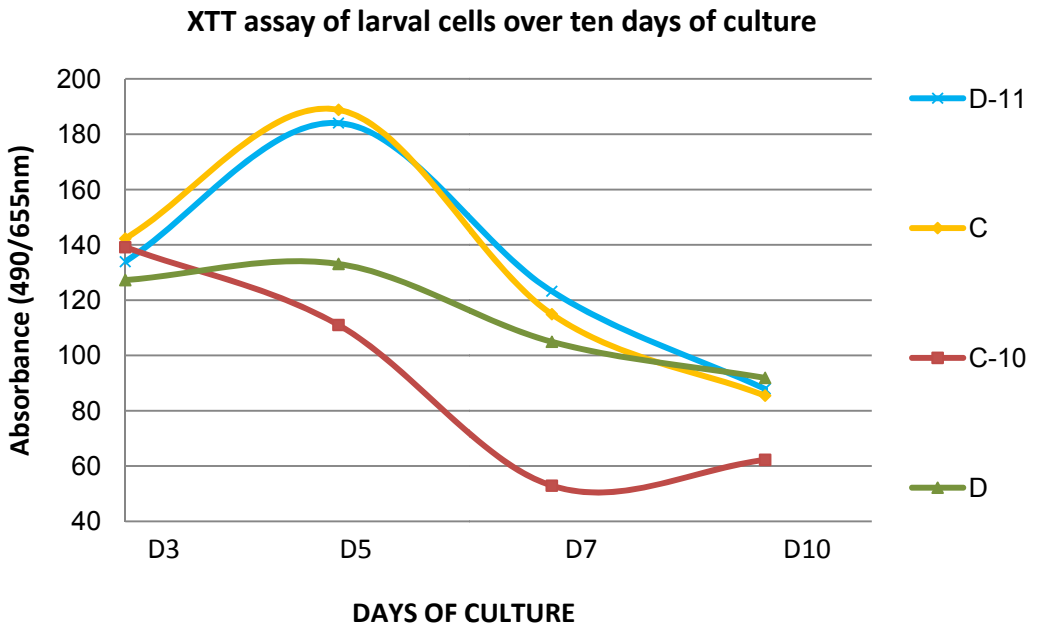


Figure 3.9 Change in metabolic activity of cultured *H. midae* larval cells (15 500 cell per well) over ten days. C = Control medium C, C-10 = Medium C + 1 % amino acids, D = Control medium D, D-11 = Medium D + 1 % amino acids + 5 % FCS (absorbance values reported at 1000 x)

An increase of cell viability was observed during the first five days of primary culture for medium C and D-11 and decreased afterwards. Cell viability remained quite stable in medium D, while a rapid decline in cell viability over time was observed for cells cultured in medium C-10, with the lowest viability at day ten of culture.

3.3.3 Haemocyte cell cultures

Cell morphology and behaviour

The majority of *Haliotis midae* haemocytes appeared as small round brightly coloured amoeboid-like cells that readily clumped together (Figure 3.10). The cells attached individually and in clusters to the surface of the culture dish or even to the surface of the glass haemocytometer within 30 minutes and started spreading out by the next day of culture, adopting a fibroblast-like morphology upon adhering. Thin pseudopodial connections between cells became visible upon adhering. Cells were observed to cover an increasing portion of the well area over the first three days of culture after which it didn't change until diminishing at day nine to ten of culture when media was not changed. Amoeboid- and fibroblast-like cells remained present throughout *in vitro* culture, with fibroblast-like cells decreasing and amoeboid-like cells increasing in number towards the second week of culture.

Contamination was less of a problem in haemocyte cultures when compared to embryo and larval cell cultures. Medium C was used for initial experiments, but Medium D was used for later haemocyte cultures described below, due to the simple nature of the medium (which would facilitate observation of the effect of added factors) and the resemblance of it to natural seawater in terms of pH and osmolarity.

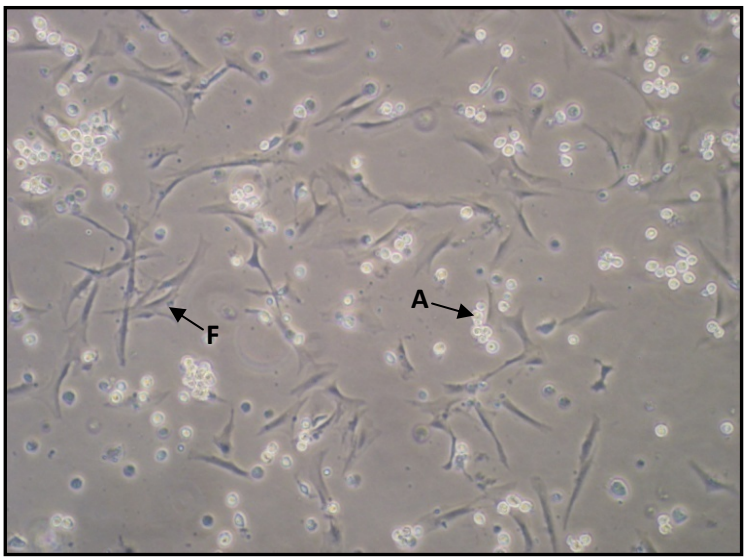


Figure 3.10 *H. midae* haemocytes in medium D (Table 3.4) attached to the surface of a six-well plate at day five of culture: A = Amoeboid-like cells, F = Fibroblast-like cells

Evaluation of cell viability

The relationship between absorbance and cell number was determined as linear in a XTT assay over nine days of culture (Figure 3.11). Cells were cultured in 96-well plates in culture medium C (Table 3.4). A consistent increase in absorbance with increased cell number per well was observed with no detrimental effects evident at the highest density of cells (109 500 cells per well). An estimated haemocyte cell number of 75 000 cells per well returned XTT absorbance values between 0.76 and 0.86 for the first seven days. At day nine the absorbance value decreased to 0.07. This may be due to medium depletion and cell deterioration. Subsequent primary haemocyte cultures were performed in 96-well tissue culture plates at a density of 50 000 cells per well. This was regarded as an adequate cell number for optimal cell viability.

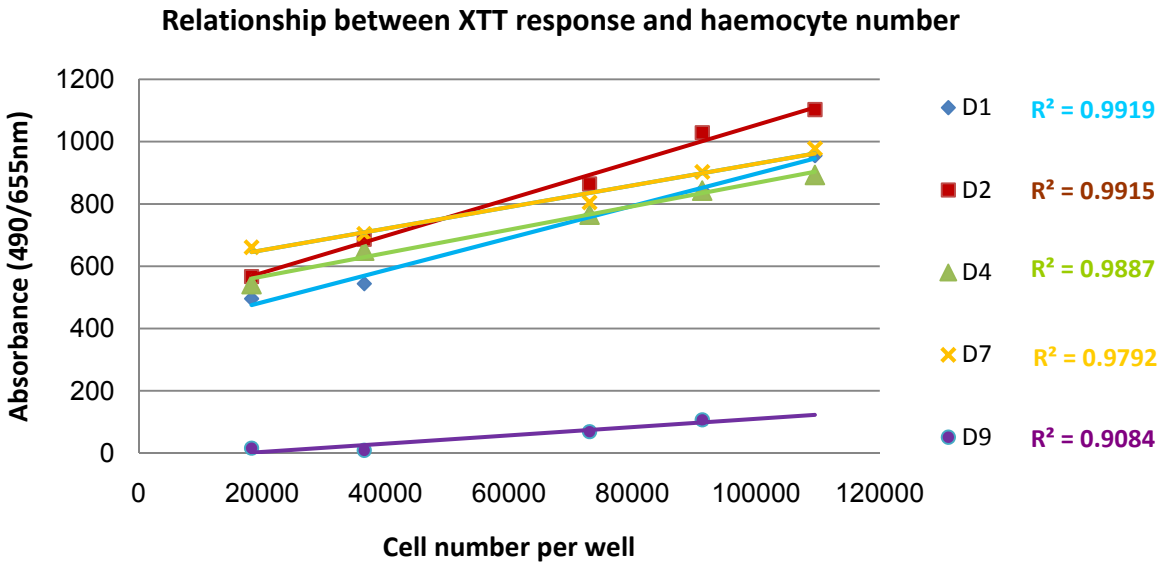


Figure 3.11 Linear relationship between *H. midae* haemocyte density and absorbances determined by XTT assay over nine days of culture for increasing cell densities. Cells were cultured in 96-well plates in medium C (Table 3.4)

To investigate the effect that media supplements had on haemocyte viability, different supplements were added to culture medium D (see Table 3.6). Cells were incubated in the respective media in 96-well plates at 50 000 cells per well for 60 hours before an XTT assay was initiated. The result of five replicates of this experiment is presented in Figure 3.12.

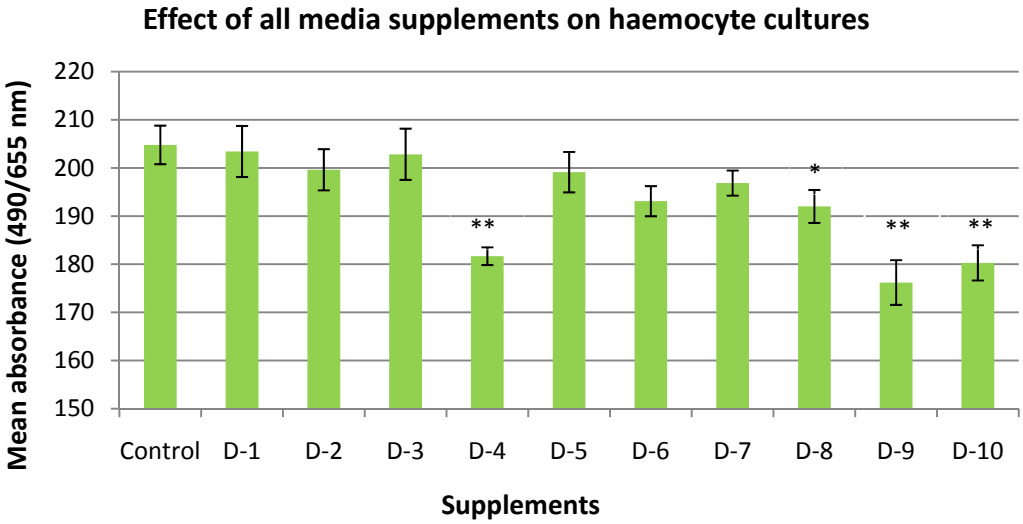


Figure 3.12 Change in cell viability of cultured *H. midae* haemocytes (60 hours): Control = Culture medium D; D-1 - D-9 = Culture medium D with different supplements (Table 3.6). Error bars indicate the standard error. Significant difference from control at $P < 0.05$ (*) and $P < 0.01$ (**)

Media that lead to a significant decrease in haemocyte viability and that was supplemented with collagenase (D-4, D-9 and D10) are contrasted with control medium and media D-1 and D-7 in Figure 3.13. Significant reduction ($P < 0.01$) of haemocyte viability is apparent for media D-4 (100 µg/ml collagenase), D-9 (12.5 ng/ml EGF + 50 µg/ml collagenase) and D-10 (10 µg/ml catalase + 50 µg/ml collagenase) when compared to the control (Figure 3.13). No significant difference in haemocyte viability was observed between medium D and media D-1 (50 µg/ml bovine insulin) and D-7 (25 µg/ml bovine insulin+ 50 µg/ml collagenase). Although medium D-7 also contains collagenase as a supplement, viability was not significantly decreased as for other collagenase-containing media.

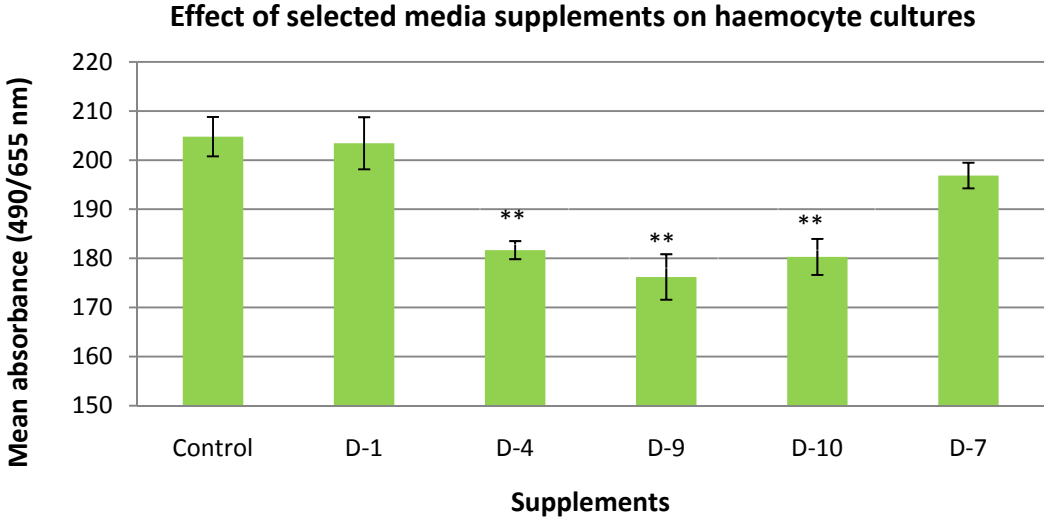


Figure 3.13 Change in cell viability of cultured *H. midae* haemocytes (60 hours): Control = Culture Medium D; Medium D with different supplements (Table 3.6): D-1 = 50 µg/ml insulin, D-4 = 100 µg/ml collagenase, D-7 = 25 µg/ml insulin + 50 µg/ml collagenase, D-9 = 12.5 ng/ml EGF + 50 µg/ml collagenase, D-10= 10 µg/ml catalase+ 50 µg/ml collagenase. Error bars indicate the standard error. Significant difference from control and D-1 and D-7 at $P < 0.01$ (**)

The most significant results from a subsequent experiment, performed in duplicate, where medium D was supplemented with bovine insulin in increasing concentrations of 0.25, 0.5, 1, 2, 4, 8, 16, 32, 64, 128 and 256 µg/ml (D-12 to D-22), are presented in Figure 3.14. Different concentrations of insulin were tested because of previous reports identifying insulin as an agent that increased haemocyte and oyster heart cell viability (Domart-Coulon *et al*., 1994; Lebel *et al*., 1996).

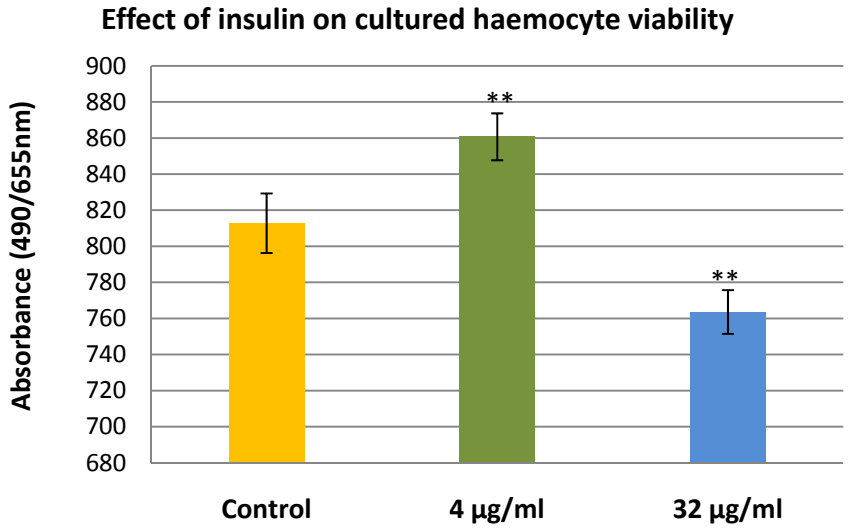


Figure 3.14 Summary of the effect of different concentrations of bovine insulin on *H. midae* haemocyte cell viability. Error bars indicate the standard error. Significant difference from control at $P < 0.01$(**)

3.3.4 Contamination

Overall, contamination by protozoa and bacteria was widespread throughout embryo and larval cell cultures despite the addition of antibiotics. Protozoa presented as oval-shaped bodies exhibiting rapid, directional movement. Bacteria presented as small round or rod-shaped bodies that had a very fast proliferation rate. After first appearing, the surface of the entire tissue culture plate would be covered in bacteria within 24 hours. Addition of Amphotericin B kept fungal contamination at bay, but the presence of Amphotericin B had a detrimental effect on cells, causing them to become shriveled and to loosen from the culture plate surface. This supported previous reports that antifungals may be toxic to molluscan cells (Mulcahy, 2000). When Amphotericin B was omitted, fungal contamination presented as cylindrical, thread-like hyphae, which branched and proliferated to a network of hyphae that would fill the culture well if allowed to grow. When such fungal contamination was apparent, cultures were discarded. Contamination was not observed frequently in haemocyte cultures.

3.4 Concluding remarks

This report of initial cell culture trials on *Haliotis midae* tissues cannot be supported by any literature for cell culture on the same species. There are, however reports of similar investigations for other abalone species (Naganuma *et al*., 1994; Lebel *et al*., 1996; Serpentini *et al.*, 2000; Auzoux-Bordenave *et al.*, 2007).

As these trials were pilot studies, they were met with various challenges. Attaining culture conditions, including optimal pH, osmolarity and temperature, that support cell viability while keeping contamination at bay with the addition of antibiotics summarize the main challenges. Contamination frequently claimed cultures within a few days. From initiation of embryo and larval cell cultures, contamination threatened culture quality and stability. Regardless of the utmost care that was taken in assuring clean and sterile working conditions, the very nature of the tissues of origin guaranteed the presence of contamination from initiation of culture. Since abalone are broadcast spawners and their gametes are released through their respiratory pores into the environment, the sterility of materials originating from gametes are jeopardized. It is recommended that future studies on material originating from gametes rather be conducted on mature excised gametes from adult animals. This is however not a feasible practice when working with farmed animals as the mature, spawning adults constitute the valuable broodstock of the farm and excising gametes cannot be performed without sacrificing the animal.

Although it is generally not recommended, antibiotics were added to wash solutions and culture media throughout trials. The trade-off when doing this is that cell integrity is challenged and that resistant bacterial strains that are difficult to eradicate may develop (Domart-Coulon *et al.*, 1994; Buchanan *et al.,* 1999; ECACC and Sigma-Aldrich, 2001). The use of antifungal agents was discontinued after it was observed that such treatment was detrimental to cultured cells. Other studies with molluscan larval cell culture also reported the use of antibiotics. Chen and Wen (1999) used streptomycin and penicillin; Odintsova *et al*. (2001) used gentamycin and Naganuma *et al*. (1994) used amphotericin B, Ampicillin, Penicillin G and Streptomycin Sulphate.

3.4.1 Suitability of various tissues for primary cell culture

Embryos and Larvae

Overall, trials to establish primary cell cultures from *Haliotis midae* embryos were met with poor success. Transportation of embryos during their early developmental stages may have been detrimental to cellular viability as slight variation in temperature as well as agitation is inevitable during transportation. Further challenges included the efficient removal of the embryonic capsule without

damaging the cell mass on the inside, effective decontamination of embryo surfaces and in cases where viable cells could be isolated from the embryo, continued differentiation into larval tissue types.

Enzymatic dissociation by collagenase treatment caused partial to complete removal of embryo capsules from the gastrula cell mass. During subsequent days of culture, these cell masses seemed to continue along the larval cell differentiation pathways. Similar continued differentiation was reported in oyster embryo cell culture. After gastrula stage embryos from oyster were dissociated with collagenase, individual cells regrouped and developed to the trochophore stage (Chen and Wen, 1999). Dissociation methods where high concentrations of sperm and lysin were used, also rendered embryo cells that either continued differentiating into larval cells with beating cilia (some of which were unable to hatch out) or, into cells adhering to the surface of the poly-D-lysine coated tissue culture plates – as single cells or small groups of cells (Figure 3.3). The use of sperm lysin as a dissociative agent of the embryo capsule with the aim to initiate cell culture has never been reported before. Although sperm and sperm lysin have been reported to dissolve the egg capsule of abalone embryos (Lewis *et al*., 1982; Shaw *et al*., 1995; Kresge *et al.*, 2001) the effect of high concentrations sperm on viability of embryonic cells has not been reported previously. According to Lewis *et al.* (1982) and Kresge *et al*. (2001), abalone sperm lysin creates a hole in the vitelline envelope of the egg by a stoichiometric, non-enzymatic interaction. The non-enzymatic nature of lysin prevents complete degradation of the vitelline envelope and also prevents degradation of other cell membrane proteins (Lewis *et al*., 1982; Kresge *et al.*, 2001). The successful removal of the embryo capsule and subsequent survival and differentiation of embryo cells in short term culture, in the present study, confirm that sperm lysin is not detrimental to embryo cells. This method of decapsulation is recommended for future studies aimed at initiating abalone embryo primary cell culture.

The processes underlying *Haliotis* embryonic development should be described and understood to a greater extent before successful embryonic stem cell culture can be achieved. Cells with stem-like qualities have not yet been identified in *Haliotis* embryos, but in another mollusc, *Patella vulgata,* macromeres at the 32-cell stage are suggested to be the stem cells of the mesoderm (De Laat *et al.*, 1980). Rigorous characterization of *Haliotis* embryonic stages and their associated cells is needed before a targeted attempt to isolate, culture and immortalize stem cells can be made.

Larval cell cultures were slightly more viable and endured for longer periods of time than embryo cell cultures, probably because more robust handling during wash and dissociation steps was tolerated by larval tissues. In addition, larvae consist of differentiated cell types that might play supportive cell-to-cell roles during *in vitro* culture.

Among the dissociation protocols used, enzymatic dissociation with trypsin (method 2) and collagenase (method 3) and chemical-mechanical dissociation with EDTA and gentle pipetting (method 4) demonstrated to be the most effective in dissociating larvae into various cell types, while retaining cell viability. Viability of ~90 % after dissociation with these methods compares well with viability of 90–95 % reported for *Mytilus trossulus* larvae after dissociation with 0.25 % collagenase (Odintsova *et al*., 2001) and 0.125 % collagenase (Odintsova *et al*., 2010). Larval dissociation with 0.25 % trypsin, reported for abalone (*H. rufescens*), also resulted in successful short-term primary cell cultures (Naganuma *et al*., 1994). Although mechanical and chemical dissociation can also be used to initiate primary larval cell cultures successfully, the yield obtained from mechanical dissociation is smaller and both mechanical and chemical treatments probably cause more damage to larval cells than enzymatic treatments.

Morphology and cellular behaviour similar to that described in other molluscan larval cell cultures were confirmed during this study: Although Naganuma *et al*. (1994) reported viable larval cell cultures for up to 12 weeks, the same pattern of decreased cell number after about day seven of culture was observed. The same authors also reported contracting myocytes after one to two days of culture, which was confirmed as parallel development to the *in vivo* situation. These authors did however not report synchronicity of contraction in neighbouring contractile cells, as was the case for the present study. The presence of filamentous actin, striated I-bands and desmin, a muscle specific intermediate filament, in *H. rufescens*, was confirmed by phalloidin and immunohistochemical staining methods and the authors concluded that mesodermal cells were already committed to their myogenic fate at the trochophore larval stage (Naganuma *et al*., 1994). Observations from the present study where dissociated gastrula embryo cells continued differentiating into ciliated, swimming cell clusters or adherent contractile cells suggest that cell commitment may even be secured earlier than the trochophore stage.

Also for cultured mussel (*Mytilus trossulus*) larval cells, the presence of actin microfilaments was confirmed by rhodamine-phalloidine staining and contractile behaviour was described as resembling smooth muscle cells (Odintsova *et al.*, 2001). Odintsova *et al*. (2010) recently reported similar cell aggregates connected with long spindle-shaped cells in mussel (*Mytilus trossulus*) larval cell cultures, as was observed here for *Haliotis midae* larval cell cultures. The mussel cell aggregates were described as cells that differentiated toward muscle, neuron-like and ciliated cells and spontaneous contraction in cell aggregates could be detected for longer than two months after culture initiation (Odintsova *et al*., 2010).

The influence of culture medium composition and supplementation on the longevity and viability of larval cell cultures was studied at the hand of various experiments. Culture media A and C (Leibovitz-based) and D (artificial seawater-based) all demonstrated to be media that has the ability to sustain

short-term larval cell cultures. Supplementation did not always have an evident effect on viability: addition of 10 ng/ml bFGF did not influence larval cell viability during 12 days and addition of different combinations of fetal calf serum and collagenase did not significantly alter cell viability during eight days of culture. There are, however, cases where supplementation did appear to enhance cell viability. Supplementation of medium C with 10 % Fetal Calf Serum and 5 µg/ml bovine insulin sustained larval cell cultures significantly better over nine days of culture than the same medium supplemented with 10 % Fetal Calf Serum, 50 µg/ml bovine insulin and 2 ng/ml bFGF. Supplementation of medium D with 1 % amino acids seemed to sustain larval cell culture better over a period of seven days, than medium D without addition of the amino acids.

The successful maintenance of primary cell cultures from *H. midae* larvae in Leibovitz-based medium is supported by reports of primary cell culture from other molluscan larvae. Examples include oyster (*Crassostrea gigas*) larval cell cultures that was maintained for 50 days in 2 x Leibovitz's L-15 medium (Chen and Wen, 1999) and mussel (*Mytilus trossulus*) larval cell cultures that was maintained for ten weeks in modified Leibovitz-L15 medium (Odintsova *et al*., 2010). Larval cell cultures from red abalone (*H. rufescens*) were also maintained in modified Leibovitz L -15 medium for up to twelve weeks (Naganuma *et al*., 1994). The use of artificial seawater for maintenance of larval cell cultures cannot be supported by literature, but from the present study it seems a feasible alternative to Leibovitz-based media. Another cell type that was successfully cultured for up to four months using artificial seawater, supplemented with amino acids, vitamins, fetal calf serum and whole egg ultrafiltrate, is cardiac tissue of the surf clam (*Spisula solidissima*) (Cecil, 1969). Sterilised (0.22 µm filtered) natural seawater, supplemented with antibiotics, is also frequently used for sponge cell cultures (Klatau *et al*., 1993; De Rosa *et al*., 2001; Cao *et al*., 2007).

Where supplementation of larval cell cultures is concerned, Chen and Wen (1999) reported that the addition of fibroblast growth factor, epidermal growth factor, endothelial cell growth factors and Nutridoma serial products to medium did not yield positive results in terms of cell growth. Fetal calf serum (10 %) was added to all their media. Addition of embryonic calf serum, vitamin E and insulin to modified Leibovitz L-15 medium for culture of larval cells of the bay mussel (*Mytilus trossulus*) supported cell viability (Odintsova *et al*., 2001). Naganuma *et al*. (1994) also supplemented their media for abalone larval cell cultures with 10 % fetal calf serum. Although it is generally recommended to omit antibiotics from cell culture, the present study, like the three abovementioned reports for mollusc larval cell culture, included antibiotics in the culture media. Chen and Wen (1999) used streptomycin and penicillin; Odintsova *et al*. (2001) used gentamycin and Naganuma *et al*. (1994) used amphotericin B, Ampicillin, Penicillin G and Streptomycin Sulphate.

Haemocytes

Abalone haemocytes adapted well to *in vitro* conditions, adhering rapidly and firmly to the surface of the tissue culture plate and adopting morphologically defined populations that remained well-defined for the duration of the experiments. Haemocyte cell cultures could be maintained sufficiently in both media C (Leibovitz-based) and D (artificial seawater-based) for up to seven days of culture. When maintained for longer periods, the condition of the cells deteriorated and increasing numbers of cells started lifting away from the surface of the tissue culture dish. The amoeboid-like cells observed in the present study possibly correspond to the blast-like cells in *H. tuberculata*, which were described as exhibiting characteristics of undifferentiated cells and poor adhesion behaviour. The fibroblast-like cells observed here may correspond to the large hyalinocytes in *H. tuberculata*, which were proposed to play roles in glycogen storage and phagocytosis (Travers *et al.*, 2008).

From XTT viability assays, it can be concluded that haemocytes have a higher metabolic activity than larval cells. An estimated haemocyte cell number of ± 20 000 cells per well returned absorbance values between 0.5 and 0.66, while for the same number of larval cells the absorbance values varied between 0.27 and 0.37. The high metabolic activity of haemocytes has been previously described as resulting from their high respiratory burst effect, an early defense mechanism to infection or contamination, of haemocytes in culture (Boulo *et al.*, 1991; Auzoux-Bordenave *et al.*, 2007). This immune response, together with the phagocytic characteristics of the fibroblast-like cells may be reasons for the lower susceptibility to contamination and improved health of haemocyte cultures when compared to larval cell cultures. The improved sterile conditions of primary culture initiation, where haemolymph is bled from the cut surface of the foot muscle may also contribute to diminished contamination, when compared to primary cultures from embryos and larvae that originate from spawning of gametes through the respiratory pores into the environment.

The effects of various media supplementations on cell viability of molluscan cell cultures in previous reports were confirmed in 60-hour-old *H. midae* primary haemocyte cultures to some extent. Bovine insulin did not have a significant stimulatory or detrimental effect at a concentration of 50 µg/ml, such as reported for haemocytes of *H. tuberculata* and cultured oyster heart cells (Domart-Coulon *et al.,* 1994; Lebel *et al*., 1996). A significant increase in metabolic activity is evident at a bovine insulin concentration of 4 µg/ml, while a significant depression of metabolic activity is evident at 32 µg/ml (Figure 3.14). At a concentration of 25 µg/ml a possible protective effect of bovine insulin against collagenase-induced cell damage is suggested. Following the supplementation experiment, collagenase and combinations with collagenase seemed to have a detrimental effect on haemocyte cell viability. Despite the suggestion of Chen and Wen (1999) that 100 µg/ml collagenase might enhance the growth of oyster and hard clam heart cell cultures, various reports on the detrimental effect of collagenase on

cultured cells are available. In a report where human fibroblast cell cultures were used to study the wound healing process, collagenase was used to reduce the number of viable fibroblast cells, in a dose-dependant manner. A concentration of 0.5 mg/ml decreased fibroblast cell number by 60 % in 72 hours and lower concentrations (0.01, 0.05 and 0.25 mg/ml) also confirmed a trend toward decreasing cell viability (Zamboni *et al*., 2004). Lo and Kim (2004) showed that collagenase caused apoptosis in primary human chondrocyte cell culture in a time and dose dependent manner (0.02 % - 0.08 % collagenase from 12 to 48 hours). While outlining dissociation procedures for initiating nerve and muscle cell cultures from *Xenopus laevis*, Gomez *et al*., (2003) also warns against prolonged exposure to collagenase, due to its detrimental effect on cell survival. In cell cultures, collagenase probably causes degradation of the extracellular matrix components, which leads to an impairment in the interaction between cells and matrix components (Zamboni *et al*., 2004).

When bovine insulin was added to haemocyte cell cultures in the present study, the detrimental effect noticed for collagenase was markedly less severe (Figure 3.13). Likewise, IGF- 1 was previously shown to be an effective inhibitor of collagenase-induced chondrocyte apoptosis, in a dose dependent manner (100 and 200 ng/ml) (Lo and Kim, 2004). Another report, where insulin increased live cell number of porcine aortic endothelial cells after damaging treatment with high glucose and heparinase, confirms the protective function of insulin (Han *et al*., 2005). Insulin is also known as an important survival factor for primary cerebellar neurons in mice and has been reported to protect cultured neuronal cells from neonatal rat retina (R28 cell line) against apoptosis via a PI 3-kinase/Akt-dependent pathway (Barber *et al*., 2001). Insulin is also present in fetal bovine serum (FBS) and Burke and Vuk-Pavlovic (1993) showed that insulin and IGF elicit a similar proliferative response in human neuroblastoma cells as when these cells are treated with 10 % FBS. The fact that FBS is also widely used during subculturing of vertebrate cell lines to inactivate proteases after cells have been detached from the flask bottom, suggests that some FBS component(s) counter the action of proteases. A similar protective action of insulin might be at play in abalone haemocyte cultures.

Catalase and EGF did not have any significant effect on metabolic activity of the haemocytes. The lack of success in reproducing the results published by other researchers may be due to differences in haemolymph composition and culture conditions. Further studies with *H. midae* haemocytes where the time of incubation and concentration of supplements are varied are needed to elucidate their possible growth stimulatory effects.

3.4.2 Applications

Primary cell culture has been applied with useful results in various studies of marine molluscs amongst which count reports on cell differentiation (myogenesis), regulation of metabolic processes

(biomineralization), enzyme characterisation, cytotoxicity assessment and ecotoxicological applications (Naganuma *et al*., 1994; Domart-Coulon *et al.*, 2000; Le Pennec and Le Pennec, 2001; Odintsova *et al.*, 2001; Faucet *et al.*, 2003; Barik *et al*., 2004; Auzoux-Bordenave, 2007; Badariotti *et al.*, 2007; Suja *et al.,* 2007; Odintsova *et al.*, 2010). The successful short term culture of larval and especially haemocyte cells from *H. midae* described in this chapter may be used as a foundation for *in vitro* assays which could facilitate advances in knowledge pertaining to cell physiology, biochemistry and behaviour of the South African abalone. One example of such an implementation that will be adressed in Chapter 4 is the measurement of growth response of cultured cells to exogenous growth stimulants. Haemocyte cultures, which can be propagated within a short time, can be used in this way for RNA extraction and subsequent quantitative analyses of gene expression, in order to gain information about the genes involved in cell viability, growth and proliferation.

3.5 References

Alberts, B., A. Johnson, J. Lewis, M. Raff, K. Roberts, and P. Walter. 2002. Molecular Biology of the Cell, 4th ed. New York, Garland Science. 448p.

Auzoux-Bordenave, S., M. Fouchereau-Peron, M.N. Helléouet, and D. Doumenc. 2007. Activity of mantle cells and hemocytes in abalone primary cell cultures (*Haliotis tuberculata*). *Journal of Shellfish Research.* 26:887-894.

Badariotti, F., R. Thuau, C. Lelong, M.P. Dubos, and P. Favrel. 2007. 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.* 31:559-570.

Barber, A.J., M. Nakamura, E.B. Wolpert, C.E.N. Reiter, G.M. Seigel, D.A. Antonetti, and T.W. Gardner. 2001. Insulin rescues retinal neurons from apoptosis by a phosphatidylinositol 3-Kinase/Akt-mediated mechanism that reduces the activation of caspase-3. *The Journal of Biological Chemistry.* 276:32814-32821.

Barik, S.K., J.K. Jena, and K. Janaki Ram. 2004. $CaCO_3$ crystallization in primary culture of mantle epithelial cells of freshwater pearl mussel. *Current Science.* 86:730-734.

Boulo, V., D. Hervio, A. Morvan, E. Bachere, and E. Mialhe. 1991. *In vitro* culture of mollusc hemocytes. Functional study of burst respiratory activity and analysis of interactions with protozoan and procaryotic pathogens. *In Vitro* 27:42A.

Brimble, S.N., X. Zeng, D.A. Weiler, Y. Luo, Y. Liu, I.G. Lyons, W.J. Freed, A.J. Robins, M.S. Rao, and T.C. Schultz. 2004. Karyotypic stability, genotyping, differentiation, feeder-free maintenance and gene expression sampling in the three human embryonic cell lines derived prior to August 9, 2001. *Stem Cells and Development.* 13:585-596.

Brooks, M.A. and T.J. Kurtti. 1971. Insect cell and tissue culture. *Annual Review of Entomology.* 16:27-52.

Buchanan, J.T., J.F. La Peyre, R.K. Cooper, and T.R. Tiersch. 1999. Improved attachment and spreading in primary cell cultures of the Eastern oyster, *Crassostrea virginica. In Vitro Cellular and Developmental Biology.* 35:593-598.

Bulgakov, V.P., N.A. Odintsova, S.V. Plotnikov, K.V. Kiselev, E.V. Zacharov, and Y.N. Zhuravlev. 2002. *Gal-4*-Gene-dependent alterations of embryo development and cell growth in primary culture of sea urchins. *Marine Biotechnology.* 4:480-486.

Burke, T.W., and S. Vuk-Pavlović. 1993. Insulin-like growth factor-1 is a serum component stimulating growth of human neuroblastoma. *In Vitro Cellular & Developmental Biology*. 29A:391-394.

Cao, X., W. Fu, X. Yu, and W. Zhang. 2007. Dynamics of spicule production in the marine sponge *Hymeniacidon perlevis* during *in vitro* cell culture and seasonal development in the field. *Cell and Tissue Research.* 329:595-608.

Cecil, J.T. 1969. Mitosis in cell cultures from cardiac tissue of the surf clam *Spisula solidissima*. *Journal of Invertebrate Pathology.* 14:407-410.

Chen, S.L., G.C. Ren , Z.X. Sha, and Y. Hong. 2005. Development and characterization of a continuous embryonic cell line from turbot (*Scophthalmus maximus*). *Aquaculture.* 249:63-68.

Chen, S.N., and C.M. Wen. 1999. Establishment of cell lines derived from oyster, *Crassostrea gigas* Thunberg and hard clam, *Meretrix lusoria* Röding. *Methods in Cell Science.* 21:183-192.

Cima, F., V. Matozzo, M.G. Marin, and L. Ballarin. 2000. Haemocytes of the clam *Tapes philippinarum* (Adams and Reeve, 1850): morphofunctional characterisation. *Fish and Shellfish Immunology.* 10:677-693.

Courtois De Viçose, G., M.P. Viera, A. Bilbao, and M.S. Izquierdo. 2007. Embryonic and larval development of *Haliotis tuberculata coccinea* Reeve: an indexed micro-photographic sequence. *Journal of Shellfish Research.* 26:847-854.

Crofts, D.R. 1937. The development of *Haliotis tuberculata*, with special reference to organogenesis during torsion. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences.* 228:219-268.

De Laat, S.W., L.G.J. Tertoolen, A.W.C. Doresteijn, and J.A.M. Van den Biggelaar. 1980. Intercellular communication patterns are involved in cell determination in early molluscan development. *Nature.* 287:546-548.

De Rosa, S., S. De Caro, G. Tommonaro, K. Slantchev, K. Stefanov, and S. Popov. 2001. Development in a primary cell culture of the marine sponge *Ircinia muscarum* and analysis of the polar compounds. *Marine Biotechnology.* 3:281-286.

Domart-Coulon, I., D. Doumenc, S. Auzoux-Bordenave, and Y.L. Fichant. 1994. Identification of media supplements that improve the viability of primarily cell cultures of *Crassostrea gigas* oysters. *Cytotechnology.* 16:109-120.

Domart-Coulon, I., S. Auzoux-Bordenave, D. Doumenc, and M. Khalanski. 2000. Cytotoxicity assessment of antibiofouling compounds and by-products in marine bivalve cell cultures. *Toxicology in Vitro.* 14:245-251.

European Collection of Cell Cultures (ECACC) and Sigma-Aldrich. 2001. Fundamental techniques in cell culture. Salisbury, UK. 68p.

Evenden, A.J. 2000. Tissue choices for aquatic invertebrate tissue culture. *In:* Aquatic Invertebrate Cell Culture. C. Mothersill and B. Austin, editors. Chichester, UK, Praxis Publishing: 15-24.

Fan, T.J., and X.F. Wang. 2002. *In vitro* culture of embryonic cells from the shrimp, *Penaeus chinensis*. *Journal of Experimental Marine Biology and Ecology.* 267:175-184.

Farcy, E., A. Serpentini, B. Fiévet, and J.M. Lebel. 2007. Identification of cDNAs encoding HSP70 and HSP90 in the abalone *Haliotis tuberculata*: Transcriptional induction in response to thermal stress in hemocyte primary culture. *Comparative Biochemistry and Physiology, Part B.* 146:540-550.

Faucet, J., M. Maurice, B. Gagnaire, T. Renault, and T. Burgeot. 2003. Isolation and primary culture of gill and digestive gland cells from the common mussel *Mytilus edulis*. *Methods in Cell Science.* 25:177-184.

Freshney, R.I. 1992. Animal Cell Culture: A Practical Approach, 2nd ed. Oxford, IRL Press. 329p.

Gagnaire, B., H. Frouin, K. Moreau, H. Thomas-Guyon, and T. Renault. 2006. Effects of temperature and salinity on haemocyte activities of the Pacific oyster, *Crassostrea gigas* (Thunberg). *Fish and Shellfish Immunology.* 20:536-547.

Gómez, T.M., D. Harrigan, J. Henley, and E. Robles. 2003. Working with *Xenopus* spinal neurons in live cell culture. *In:* Neurons: Methods and Applications for the Cell Biologist. P. J. Hollenbeck and J. R. Bamburg, editors. San Diego, USA, Academic Press, Elsevier Science. 144p.

Gómez-Mendikute, A., M. Elizondo, P. Venier, and M.P. Cajaraville. 2005. Characterization of mussel gill cells *in vivo* and *in vitro*. *Cell and Tissue Research.* 321:131-140.

Hahn, K.O. 1989. CRC Handbook of culture of abalone and other marine gastropods, 1st ed. Florida, CRC Press. 348 p.

Han, J., A.K. Mandal, and L.M. Hiebert. 2005. Endothelial cell injury by high glucose and heparanase is prevented by insulin, heparin and basic fibroblast growth factor. *Cardiovascular Diabetology.* 4:12.

Hansen, E.L. 1988. A cell line from embryos of *Biomphalaria glabrata* (Pulmonata): Establishment and characteristics. *In:* Invertebrate Tissue Culture: Research Applications. K. Maramorosch, editor. New York, Academic Press: 75-97.

Henry, N.R. 1995. Studies related to the artificial spawning and culture of the abalone, *Haliotis midae* Linne, 1785. Unpublished MSc thesis, University of Cape Town, Cape Town, South Africa.

Humphries, J.E., and T.P. Yoshino. 2003. Cellular receptors and signal transduction in molluscan hemocytes: Connections with the innate immune system of vertebrates. *Integrative and Comparative Biology.* 43:305-312.

Jorgenson, D.D., S.K. Ware, and J.R. Redmond. 1984. Cardiac output and tissue blood flow in the Abalone, *Haliotis cracherodii* (Mollusca, Gastropoda). *Journal of Experimental Zoology.* 231:309-324.

Klautau, M., M.R. Custodio, and R. Borojevic. 1993. Cell cultures of sponges *Clathrina* and *Polymastia*. *In Vitro Cellular and Developmental Biology.* 29A:97-99.

Kresge, N., V.D. Vacquier, and C.D. Stout. 2001. Abalone lysin: the dissolving and evolving sperm protein. *BioEssays.* 23:95-103.

Kuang, S., M. Regnier, and J.I. Goldberg. 2002. Long-term culture of decapsulated gastropod embryos: A transplantation study. *Biological Bulletin.* 203:278-288.

Le Pennec, G., and M.M. Le Pennec. 2001. Acinar primary cell culture from the digestive gland of *Pecten maximus* (L.): an original model for ecotoxicological purposes. *Journal of Experimental Marine Biology and Ecology.* 259:171-187.

Lebel, J.M., W. Giard, P. Favrel, and E. Boucaud-Camou. 1996. Effects of different vertebrate growth factors on primary cultures of hemocytes from the gastropod mollusc, *Haliotis tuberculata*. *Biology of the Cell.* 86:67-72.

Lewis, C.A., and Talbot, C.F. and Vacquier, V.D. 1982. A protein from abalone sperm dissolves the egg vitelline layer by a nonenzymatic mechanism. *Developmental Biology.* 92:227-239.

Littlewood, T. 1998. Cell biology aspects of safety in cell culture. *In:* Safety in cell and tissue culture. G. Stacey, A. Doyle and P. Hambleton, editors. Dordrecht, Kluwer: 26-46.

Lo, M.Y., and H.T. Kim. 2004. Chondrocyte apoptosis induced by collagen degradation: inhibition by caspase inhibitors and IGF- 1. *Journal of Orthopaedic Research.* 22:140-144.

Lyons-Alcantara, M. 2001. Characterization. *In:* Aquatic invertebrate Cell Culture. C. Mothersill and B. Austin, editors. Chichester, UK, Praxis: 25-49.

Mothersill, C., A.L. Mulford, and B. Austin. 2000. Basic methods and media. *In:* Aquatic Invertebrate Cell Culture. C. Mothersill and B. Austin, editors. Chichester, UK, Praxis: 9-23.

Mulcahy, M.F. 2000. Culture of molluscan cells. *In:* Aquatic Invertebrate Cell Culture. C. Mothersill and B. Austin, editors. Chichester, UK, Praxis: 165-177.

Naganuma, T., B.M. Degnan, K. Horikoshi, and D.E. Morse. 1994. Myogenesis in primary cell culture from larvae of the abalone, *Haliotis rufescens*. *Molecular Marine Biology and Biotechnology.* 3:131-140.

Nichols, J., and Q.L. Ying. 2006. Derivation and propagation of embryonic stem cells in serum- and feeder-free culture. *In:* Embryonic Stem Cell Protocols. K. Turksen, editor. Totowa, NJ, Humana Press: 91-98.

Odintsova, N.A., and A.V. Khomenko. 1991. Primary cell culture from embryos of the Japanese scallop *Mizuchopecten yessoensis* (Bivalvia). *Cytotechnology.* 6:49-54.

Odintsova, N.A., S.V. Plotnikov, A.A. Karpenko, and M.G. Eliseikina. 2001. Myogenic differentiation of *Mytilus* larval cells *in vitro*. *Russian Journal of Developmental Biology.* 32:304-309.

Odintsova, N.A., V.A. Dyachuk, and L.P. Nezlin. 2010. Muscle and neuronal differentiation in primary cell culture of larval *Mytilus trossulus* (Mollusca: Bivalvia). *Cell and Tissue Research.* 339:625-637.

Phillips, H.J., and J.E. Terryberry. 1957. Counting actively metabolizing tissue cultured cells. *Experimental Cell Research.* 13:341-347.

Pomponi, S.A., and R. Willoughby. 2000. Development of sponge cell cultures for biomedical application. *In:* Aquatic Invertebrate Cell Culture. C. Mothersill and B. Austin, editors. Chichester, UK, Praxis: 323-336.

Purchon, R.D. 1977. The biology of the Mollusca, 2nd ed. Oxford, UK, Pergamon press: 269-332.

Qiagen. 2002. The Qiagen Transfection Resource Book., 2nd ed. Germany, Qiagen:36-52.

Richelle-Maurer, E., R. Gomez, J.C. Braekman, G. Van de Vyver, R.W.M. Van Soest, and C. Devijver. 2003. Primary cultures from the marine sponge *Xestospongia muta* (Petrosiidae, Haplosclerida). *Journal of Biotechnology.* 100:169-176.

Rinkevich, B. 2005. Marine Invertebrate Cell Cultures: New Millennium Trends. *Marine Biotechnology.* 7:429-439.

Roche Applied Science. 2003. Cell proliferation and viability measurement. *Biochemica.* 3:26-28.

Rosenfield, A. 1993. Marine invertebrate cell culture: Breaking the barriers. NOAA Technical Memorandum NMFS-F/NEC-98, Department of Commerce, USA. 40p.

Sakai, Y.T. 1961. Method for removal of chorion and fertilization of the naked egg in *Oryzias latipes. Embryologica.* 5:357-368.

Serpentini, A., C. Ghayor, J.M. Poncet, V. Herbert, P. Galéra, J.P. Pujol, E. Boucaud-Camou, and J.M. Lebel. 2000. Collagen study and regulation of the d*e novo* synthesis by IGF-I in hemocytes from the gastropod mollusc*, Haliotis tuberculata*. *Journal of Experimental Zoology.* 287:275-284.

Shallow Seafarming Research Institute. 1990. Training manual on artificial breeding of abalone (*Haliotis discus hannai*) in Korea DPR. Kosong, Korea, Regional Seafarming Development and Demonstration Project.

Shaw, A., P.A.G. Fortes, and V.D. Vacquier. 1995. Crystal structure and subunit dynamics of the abalone sperm lysin dimer- egg envelopes dissociate dimers, the monomer is the active species. *The Journal of Cell Biology.* 130:1117-1125.

Shui, J.W., and T.H. Tan. 2004. Germline transmission and efficient DNA recombination in mouse embryonic stem cells mediated by adenoviral-cre transduction. *Genesis.* 39:217-223.

Smith, A.G. 2001. Embryo-derived stem cells: Of mice and men. *Annual Review of Cell and Developmental Biology.* 17:435-462.

Stanford University (1997) Sea urchin embryology: Artificial seawater. [Online]. Available: http://www.stanford.edu/group/Urchin/seawater [Accessed 21 May 2010].

Sud, D., D. Doumenc, E. Lopez, and C. Milet. 2001. Role of water-soluble matrix fraction, extracted from the nacre of *Pinctada maxima*, in the regulation of cell activity in abalone mantle cell culture (*Haliotis tuberculata*). *Tissue and Cell.* 33:154-160.

Suja, C.P., N. Sukumaran, and S. Dharmaraj. 2007. Effect of culture media and tissue extracts in the mantle explant culture of abalone, *Haliotis varia* Linnaeus. *Aquaculture.* 271:516-522.

Terahara, K., K.G. Takahashi, and K. Mori. 2005. Pacific oyster hemocytes undergo apoptosis following cell-adhesion mediated by integrin-like molecules. *Comparative Biochemistry and Physiology, Part A.* 141:215- 222.

Travers, M.A., P.M. Da Silva, P.M. Le Goi, D. Marie, A. Donval, S. Huchette, M. Koken, and C. Paillard. 2008. Morphologic, cytometric and functional characterisation of abalone (*Haliotis tuberculata*) haemocytes. *Fish and Shellfish Immunology.* 24:400-411.

Vacquier, V.D., K.R. Carner, and C.D. Stout. 1990. Species-specific sequences of abalone lysin, the sperm protein that creates a hole in the egg envelope. *Proceedings of the National Academy of Sciences USA.* 87:5792-5796.

Vavra, J., and D.T. Manahan. 1999. Protein metabolism in lecithotrophic larvae (Gastropoda: *Haliotis rufescens*). *Biological Bulletin.* 196:177-186.

Villena, A.J. 2003. Applications and needs of fish and shellfish cell culture for disease control in aquaculture. *Reviews in Fish Biology and Fisheries.* 13:111-140.

Zamboni, P., M. De Mattei, A. Ongaro, L. Fogato, S. Carandina, M. De Palma, S. Tognazzo, G.L. Scapoli, M.L. Serino, A. Caruso, A. Liboni, and D. Gemmati. 2004. Factor XIII contrasts the effects of metalloproteinases in human dermal fibroblast cultured cells. *Vascular and Endovascular Surgery.* 38:431-438.

4 APPLICATION OF CELL CULTURE AND QUANTITATIVE REAL-TIME PCR TOWARDS IDENTIFICATION OF GROWTH AND METABOLISM GENES IN *H. MIDAE*

4.1 Introduction

The combination of cell culture and quantitative real-time PCR (qPCR) has been used widely to elucidate gene expression mechanisms and quantify expression of genes of interest. Especially in the field of medical research, normal and cancer cell lines are used to investigate gene expression (Combes *et al.*, 2009; Ghaemmaghami *et al*., 2010; Jheon *et al*., 2009; Fu *et al*., 2010). Cell culture provides an efficient means for measuring the response of cells to various treatments in terms of gene expression, because of the fast propagation potential of cells in culture and the controlled environment of such a system.

Aquaculture and marine environmental research areas have also benefited from the combined use of cell culture and qPCR. Fish myogenic cell cultures provide an attractive system for evaluating the effects of various regulatory molecules on gene expression, under precisely controlled conditions (Bower and Johnston, 2009). Cell lines derived from the gilthead sea bream (*Sparus aurata*) was used to evaluate the expression of several mineralization-related genes by qPCR, in response to treatment with bovine and fish sera (Rosa *et al*., 2010). A primary hepatocyte cell culture from silver sea bream (*Sparus sarba*) liver was also developed in order to assess changes in gene expression of metallothionein and glucose-6-phosphate dehydrogenase (G6PDH), following cadmium exposure (Man and Woo, 2008). Oyster mantle primary cell culture was used in drug treatment experiments to investigate mRNA expression of the transcriptional modulator Pf-Smad3, the activin like receptor 1 homolog Pf-ALR1 and two oyster nuclear factor-κB (NF-kB) members (Zhou *et al*., 2010).

Quantitative real-time PCR is an accurate, sensitive and fast method of quantifying low target copy numbers or identifying minor changes in mRNA expression levels. It is especially useful when limited amounts of tissue is available, like in cell culture, and can even be used to analyse gene expression in single cells (Derveaux *et al*., 2010; Pfaffl, 2010; Ståhlberg and Bengtsson, 2010). qPCR combines the reverse transcription of a RNA template into cDNA with the subsequent exponential amplification by PCR and quantification of the amplified product (Bustin, 2000). SYBRGreen, a fluorescent reporter used in the reaction, is a minor groove binding dye that binds to double-stranded DNA. The fluorescent signal is proportional to the amount of amplified DNA product formed. Monitoring of the fluorescence can thus be interpreted as an indication of the amount of amplified product during the course of the reaction (Kubista *et al.*, 2006; Schmittgen, 2006). The number of DNA molecules originally present in the sample can be calculated by recording the number of amplification cycles required to obtain a

particular amount of DNA molecules, while assuming a certain amplification efficiency (Kubista *et al*., 2006). When more template is present at the beginning of the reaction less cycles will be needed to reach a point in which the fluorescent signal is first recorded as statistically significant above the background (Bustin, 2000; Ginzinger, 2002).

The typical qPCR reaction graph depicts four phases (Figure 4.1): The first is where the fluorescent signal is below the detection limit and not distinguishable from the background fluorescence. The second phase is the growth/exponential phase, where the amount of product, and therefore the amount of fluorescence, increases exponentially. The point where fluorescence is first observed as significantly above background is referred to as the quantification cycle (C_q). The third is a linear phase that indicates signal saturation, where the reaction starts running out of one of the critical components and PCR product ceases to increase exponentially and finally reaches a plateau. The C_q values of different response curves in the exponential phase reflect the difference in their initial amounts of template molecules (Lorkowski and Cullen, 2003; Kubista *et al*., 2006).

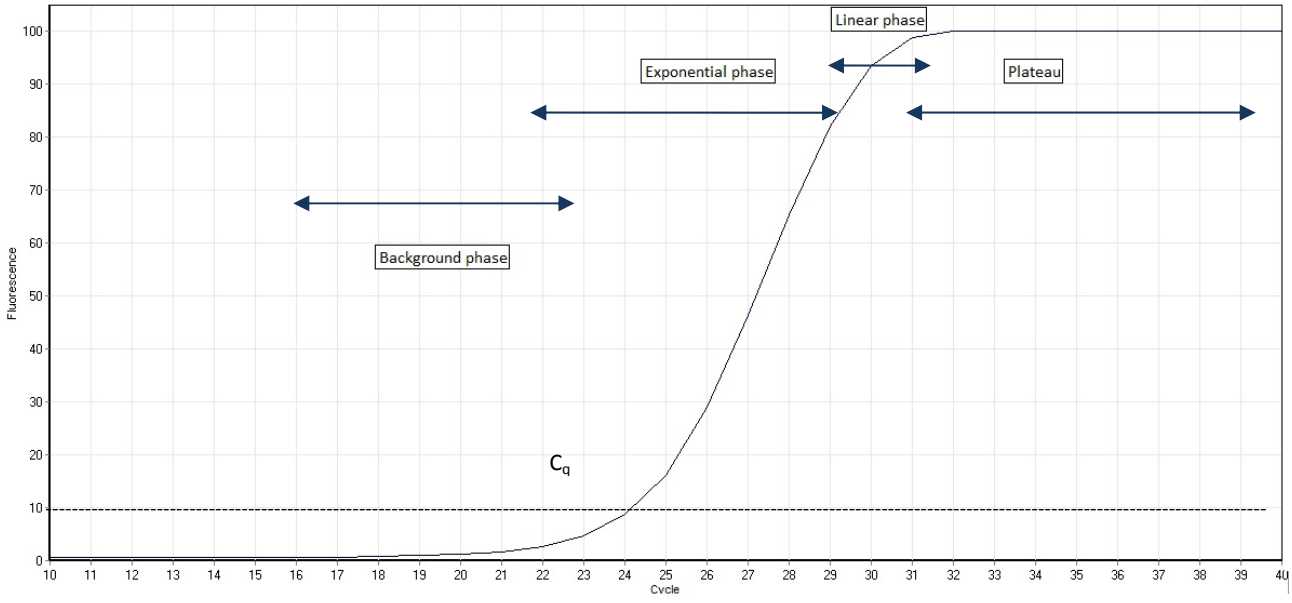


Figure 4.1 Quantitative real-time PCR graph of *H. midae* cDNA, depicting the four phases; background, exponential, linear and plateau. The dotted line represents the threshold and C_q the quantification cycle

The popularity of qPCR as a technique for quantifying gene expression differences is increasing exponentially. The large number of papers reporting the use of qPCR and the lack of consensus in terms of reagents, protocols, analysis methods and reporting formats, motivated the formation of the Minimum Information for Publication of Quantitative real-time PCR Experiments (MIQE) guidelines (Bustin *et al*., 2009; Taylor *et al*., 2010). These guidelines are aimed at establishing a framework within which researchers can conduct qPCR experiments. Such a consistent format encourages detailed

revision of experimental detail, data analysis and reporting principles (Bustin *et al*., 2010; Taylor *et al*., 2010). The suggested application of MIQE guidelines include detailed adherence to and description of the following processes: experimental design, RNA isolation procedures, RNA quality control, reverse transcription, primer and amplicon design, qPCR validation, choice of reference genes and experimental reproducibility (Taylor *et al*., 2010).

In this chapter, the application of qPCR to quantify the expression of seven target genes, normalized with two reference genes, is demonstrated. The differential expression observed for target genes in Chapter 2 is verified by qPCR in large and small abalone tissue and then tested in treated and control haemocyte cultures. Gene expression differences are expected to occur between large and small abalone and in cultured abalone haemocytes treated with two growth factors (bovine insulin and EGF) at different concentrations, in comparison with the untreated control. As shown in Chapter 3, *Haliotis midae* haemocyte primary cell culture is a suitable system for short-term experimentation. In this study cell cultures were maintained for 65 hours before harvesting cells for RNA extraction and subsequent qPCR experimentation. The process of qPCR is described, from RNA extraction, through primer design and real-time PCR, to data analysis using two software tools: Rotor-Gene 6000 Series Software and the Relative Expression Software Tool (REST: Pfaffl *et al*., 2002). The differential expression of specific target genes evident from these qPCR experiments are discussed in more detail in terms of their possible role in metabolic pathways related to cell viability and proliferation.

4.2 Materials and Methods

4.2.1 Haemocyte cell cultures

For haemolymph collection, eight adult abalone (shell length 60 to 70 mm) were sacrificed. Dissection, haemolymph collection and culture initiation were performed in a laminar flow cabinet, as described in Chapter 3. Haemolymph from all eight animals, mixed with antibiotic wash solution and Alsever solution, was pooled and amounted to ± 400 ml with 250 000 cells per milliliter. Thirty 90 mm culture dishes were seeded, each with 3.3 x 10^6 cells. After an initial period of two-and-a-half hours at 18 °C, during which time haemocytes were allowed to adhere to the surface of the culture dishes, 10 ml of culture media, supplemented according to Table 4.1, was added to each. For each medium, six culture dished were seeded. After addition of specified culture medium, tissue culture dishes were sealed with parafilm and incubated at 18 °C in a Hotpack low temperature incubator for 65 hours.

Table 4.1 Culture medium and supplementations used for *H. midae* haemocyte cell culture

Medium				
Control (C_n) (ASW medium)	**I-4**	**I-50**	**E-25**	**E-600**
Artificial Seawater (ASW) + 1 % Penicillin/Streptomycin +0.5 % Gentamycin +2 % Glutamax pH 7.45	ASW medium + 4 µg/ml bovine insulin	ASW medium + 50 µg/ml bovine insulin	ASW medium +25 ng/ml epidermal growth factor (EGF)	ASW medium + 600 ng/ml epidermal growth factor (EGF)

The supplements added to the culture medium are all reported to result in an increase in metabolic activity and act as putative growth stimulants in cultured mollusc cells (Domart-Coulon *et al*., 1994; Lebel *et al.,* 1996; Giard *et al*., 1998; Chen and Wen, 1999).

4.2.2 RNA extraction and reverse transcription

The RNeasy Midi Kit (Qiagen, USA) was used for extraction of cytoplasmic RNA from cultured haemocytes. Cytoplasmic RNA extraction limits the presence of unprocessed heterogeneous nuclear RNA (hnRNA) and immature mRNA, which can eventually lead to contamination of ESTs by cDNA sequences derived from incompletely spliced RNAs (Carnici *et al*., 2002).

In order to collect a sufficient amount of cells for RNA extraction, haemocytes from two culture dishes were regarded as one replicate of each treatment (or control), resulting in three biological replicates per treatment. For each culture dish, the top 3.5 ml medium was discarded and remaining cells were collected by scraping with the rubber tip of a 2.5 ml syringe. Cells from each replicate were pooled together in a 15 ml centrifuge tube. Tubes were centrifuged at 300 x g, 18 ˚C for seven minutes and cell pellets were resuspended in 500 µl RLN buffer [50 mM Tris·Cl, pH 8.0, 140 mM NaCl, 1.5 mM $MgCl_2$, 0.5 % (v/v) Igepal CA-630 (SIGMA), 1000 U/ml Ribolock RNase inhibitor (Fermentas, Canada), precooled

to 4 ˚C] to lyse plasma membranes. After incubating on ice for five minutes, the cell lysate was centrifuged at 500 x g, 4 °C for seven minutes and returned to the ice. Supernatants were transferred to clean 15 ml centrifuge tubes and 2 ml of guanidine isothiocyanate-containing lysis buffer (buffer RLT) was added to each tube. After vortexing, 1.4 ml of absolute ethanol was added to each tube and the contents were mixed by vigorous shaking. Samples were applied to the RNEasy-supplied columns and centrifuged at 5000 x g, 21 °C for five minutes. The flow-through was discarded and 4 ml of guanidine isothiocyanate-containing buffer RW1 (Qiagen, USA) added before centrifuging at 5000 x g, 21 °C for two minutes. After two washes in 2.5 ml buffer RPE, each column was transferred to a clean 15 ml collection tube and RNA was eluted in 150 µl RNAse-free water. RNA was quantified and purity determined using a NanoDrop® ND-1000 spectrophotometer.

Although it is suggested that DNase digestion is generally not necessary with RNeasy Kits since the RNeasy silica-membrane technology efficiently removes most of the DNA (Qiagen, USA), the RNA extraction in the present study was followed with DNase digestion. This was done to limit the amount of DNA that might have carried over from the RNA extraction to a minimum (Shipley, 2006). To accomplish DNase digestion, 50 µl of each RNA sample was transferred to a clean 500 µl eppendorf tube and incubated with 5 µl 10 x Turbo DNase buffer and 1 µl Turbo DNase (Turbo DNA-free Kit, Ambion, USA) at 37 °C for 30 minutes. The reaction was terminated by incubation with 5 µl DNase inactivation reagent for five minutes at room temperature. The supernatant, containing purified, DNase free RNA, was collected after centrifuging at 10 000 x g, 21 °C for 1.5 minutes. RNA integrity was determined by denaturing agarose gel electrophoresis (2 % agarose; 1 X MOPS buffer). The gel was run for 90 minutes and photographed under UV light.

First strand cDNA was prepared from 170 ng RNA using 1 µl of 100 µM oligo $(dT)_{18}$ primer (Fermentas, Canada) and 1 µl (200U) Superscript III (SS III) M-MLV Reverse Transcriptase (Invitrogen, USA), per 20 µl reaction. RNA, oligo primer and 1 µl dNTP Mix (10mM) were heated for five minutes at 65 °C before cooling on ice for one minute. Subsequently, 4 µl first strand buffer (5 x), 1 µl DTT (0.1M), 1 µl Ribolock (40U/µl, Fermentas, Canada) and SS III were added and the tube was incubated at 55 °C for 50 minutes, followed by 70 °C for 15 minutes. Reverse transcription was performed for each replicate of the original RNA samples. One minus reverse transcriptase control reaction (-RT) was included for each treatment, using equal amounts of RNA from each replicate, mixed together, as template. The same reagents were included in this reaction, except for SS III, which was replaced with water. Following first-strand cDNA synthesis, samples were treated with Ribonuclease H (Fermentas, Canada) (2U RnaseH per 20 µl cDNA) in order to digest the remaining RNA complementary to the cDNA. This RNaseH step after reverse transcription has been reported to increase copy numbers up to fourfold and result in earlier detection (lower crossing point) of many target genes (Mackay, 2007; Smith *et al*., 2007). It was shown that, when

RNaseH treatment is not performed, cDNA/RNA hybrids might prevent maximum primer annealing to cDNA during PCR (Mackay, 2007).

4.2.3 Primer design

Target and reference sequences for real-time PCR were carefully selected from the transcriptome data generated by Illumina sequencing-by-synthesis technology (Chapter 2). Target sequences were selected from the subset of 1174 sequences highly expressed in large animals and that had significantly differential expression when compared to small animals ($P < 0.0005$, FDR corrected). Reference sequences were selected from eight sequences with equal expression between large and small animals (according to CLC Genomics Workbench v3.5.1 differential expression analysis). All sequences used for qPCR had a coverage of more than ten and showed significant alignment with known growth related sequences during dCAS annotation (minimum E-value of 10^{-10}). These sequences and their associated putative functions are summarized in Table 4.2. The first three target sequences and the two reference sequences were initially amplified in template cDNA derived from whole animal soft body tissues (extraction described in Chapter 2). When only the reference sequences amplified in template cDNA derived from cultured haemocytes, four subsequent target sequences were selected and primers designed.

Table 4.2 Sequences chosen for verification of differential expression by quantitative real-time PCR

Sequence name	Times higher expression in L	Annotation
NODE_752_cov_15.2	2.7	**GO** ($9E^{-39}$): Generation of neurons, lung development, brain development, induction of an organ, negative regulation of apoptosis, cell maturation, blood vessel morphogenesis, branching morphogenesis of a tube, positive regulation of mesenchymal cell proliferation, extracellular space, integral to membrane, angiogenesis, FGFR activity, regulation of cell proliferation, embryonic limb morphogenesis, salivary gland morphogenesis, FGFR signaling pathway. positive regulation of cell proliferation **KOG** ($2E^{-53}$): Fibroblast/platelet-derived growth factor receptor and related receptor tyrosine kinases **KEGG** ($2E^{-37}$): FGFR1 **Haliotis_prot** ($8E^{-09}$): Map kinase [*Haliotis discus discus*] **Haliotis nt** (0E): *Haliotis midae* clone 2C3 FGFR-like mRNA, partial sequence **MolluscDB** ($2E^{-28}$): Similar to Q7YT64 (Q7YT64) Insulin-related peptide receptor (Acc: CGC01730_1) **SnailDB** ($3E^{-24}$): EST_Bglabrata-16-03-09_54304-contig_13590 similar to TEK tyrosine kinase, endothelial
NODE_13596_cov_14.5	4.0	**GO** ($1E^{-06}$): Nucleus, IGFBP complex, IGF binding, IGFR signaling pathway, inner ear morphogenesis, skeletal development, otic vesicle formation **KEGG** ($9E^{-06}$): IGFBP1 **Haliotis_prot** ($4E^{-44}$): RecName: Full=Perlustrin **Haliotis nt** ($6E^{-30}$): *Haliotis discus discus* perlustrin mRNA, complete cds **Seq_Lit** ($5E^{-31}$): *Haliotis discus discus* perlustrin mRNA, complete cds **MolluscDB** (0.16E): Similar to Q95V57 (Q95V57) Annexin **SnailDB** (0.15E): EST_Bglabrata-16-03-09_54304-contig_4975 similar to annexin
NODE_2380_cov_70.7	16.2	**GO** ($4E^{-25}$): Extracellular space, inflammatory response, negative regulation of angiogenesis **KOG** ($1E^{-22}$): Netrin transmembrane receptor unc-5 **KEGG** ($3E^{-30}$): Hypothetical protein **Haliotis_prot** ($2E^{-11}$): Beta 1,3-glucan binding protein [*Haliotis discus discus*] **MolluscDB** ($3E^{-18}$): Similar to TSP1_XENLA (P35448) Thrombospondin 1 precursor **SnailDB** ($6E^{-08}$): EST_Bglabrata-16-03-09_54304-contig_7930 similar to SCO-spondin
NODE_7048_cov_81.4	1.2	**GO** (0E): Mitochondrion **KOG** (0E): Endoplasmic reticulum glucose-regulated protein (GRP94/endoplasmin), HSP90 family **KEGG** ($3E^{-30}$): Hypothetical protein **Haliotis_prot** (0E): HSP90A [*Haliotis asinina*] **Haliotis nt** (0E): *Haliotis tuberculata* mRNA for 84kDa heat shock protein (HSP84 gene) **Haliotis_EST** (0E): ABHSL-3_G04_31_14 **Lottia** ($7E^{-85}$)**:** jgi|Lotgi1|161608|, [CHICK Heat shock protein HSP 90-alpha, *Gallus gallus*] **MolluscDB** (0E): Similar to UniRef100_Q6USB9 Heat shock protein 90 [*Chlamys farreri*] (Acc: CGC01730_1) **SnailDB** (0E): EST_Bglabrata-16-03-09_54304-contig_378, 90-kDa heat shock protein HSP83

NODE_54_cov_48.4	1.5	**GO** (0E): Collagen type IV, cellular component organization and biogenesis, extracellular matrix structural constituent **KOG** ($1E^{-112}$): Collagens (type XV) **KEGG** (0): Collagen alpha 2(IV) chain precursor **Haliotis_prot** ($2E^{-152}$): Alpha 1 type IV collagen [*Haliotis tuberculata*] **Haliotis_EST** (0E): RM035-SK *Haliotis discus* cDNA library (RM) *Haliotis discus* cDNA, mRNA sequence **Lottia** ($1E^{-39}$)**:** jgi|Lotgi1|116244|, extracellular matrix structural constituent; collagen **MolluscDB** ($1E^{-98}$): Similar to Q07265 (Q07265) 3 alpha procollagen (Acc: AIC03579_1) **SnailDB** ($3E^{-28}$): EST_Bglabrata-16-03-09_54304-contig_2680, similar to Collagen protein 2
NODE_3309__cov_19.5	1.8	**GO** ($7E^{-59}$): Cyclin-dependent protein kinase regulator activity, regulation of progression through cell cycle **KOG** ($2E^{-31}$): Cyclin B and related kinase-activating proteins **KEGG** ($3E^{-75}$): Hypothetical protein **Haliotis_prot** ($1E^{-08}$): Cyclin B [*Dreissena polymorpha*] (Acc: AAC35952.1) **MolluscDB** ($1E^{-70}$): Similar to CYCI_HUMAN (Q14094) Cyclin I **SnailDB** ($8E^{-13}$): EST_Bglabrata-16-03-09_54304-contig_13351, similar to Cyclin G1
NODE_809_cov_25.0	2.0	**GO** ($1E^{-117}$): Protein amino acid phosphorylation, protein binding, serine/threonine kinase activity, protein kinase activity **KOG** ($1E^{-135}$): MAP kinase-interacting kinase and related serine/threonine protein kinases **KEGG** (1E-^{117}): Similar to Lk6 CG17342-PA, isoform A **Haliotis_prot** ($7E^{-12}$): Map kinase [*Haliotis discus discus*] (Acc: ABO26691.1) **Haliotis_EST** ($1E^{-176}$): AHSL-2_A07_49_01 **Lottia (**$1E^{-22}$)**:** jgi|Lotgi1|138360|,Mnk [*Aplysia californica*] **MolluscDB** ($1E^{-112}$): Similar to Q803R1. Similar to MAP kinase-interacting serine/threonine kinase 2 **SnailDB** ($1E^{-112}$): EST_Bglabrata-16-03-09_54304-contig_5817 , Putative map kinase interacting kinase
NODE_8629_cov_528.7	Equal expression	**GO** ($5E^{-18}$): Intracellular, ribosome, cytosolic small ribosomal subunit (sensu Eukaryota), RNA binding, translation **KOG** ($3E^{-14}$): Ribosomal protein S4 **KEGG** ($2E^{-16}$): Similar to ribosomal protein S9 **Haliotis_prot** ($3E^{-21}$): Ribosomal protein S9 [*Haliotis discus discus*] **Haliotis_EST** ($6E^{-59}$): GOM154 *Haliotis discus* cDNA library (GOM) *Haliotis discus* cDNA, mRNA sequence **Haliotis_nt** ($3E^{-59}$): *Haliotis discus discus* ribosomal protein S9 mRNA, complete cds **MolluscDB** ($1E^{-22}$): Similar to Q7ZYU4 (Q7ZYU4) Similar to ribosomal protein S9 **SnailDB** ($1E^{-22}$): EST_Bglabrata-16-03-09_54304-contig_2050, similar to 40S ribosomal protein S9
NODE_12621_cov_30.8	Equal expression	**GO** ($1E^{-120}$): Ornithine decarboxylase activity, protein binding, polyamine biosynthetic process, cellular_component **KOG** ($1E^{-136}$): Ornithine decarboxylase **KEGG** ($1E^{-141}$): Hypothetical protein **Haliotis_prot** (0E): Ornithine decarboxylase [*Haliotis diversicolor supertexta*] **Haliotis_EST** (0E): GOF007-SK *Haliotis discus* cDNA library (GOF) *Haliotis discus* cDNA, mRNA sequence **Haliotis_nt** (0E): *Haliotis diversicolor supertexta* ornithine decarboxylase mRNA, complete cds **MolluscDB** ($1E^{-162}$): Similar to DCOR_BOVIN (P27117) Ornithine decarboxylase **SnailDB** ($1E^{-148}$): EST_Bglabrata-16-03-09_54304-contig_1352, similar to ornithine decarboxylase 1

Primers were designed in Oligo 4.1 Primer Analysis Software (Rychlik, 1992) according to the following optimal parameters: Primer length, 18 to 25 bases; 3'end consisting of a G or C; melting temperature (T_m) of 63; T_m difference between forward and reverse primer ≤ 1 °C; ΔG of ≤ -5 (in order to limit primer selfdimers and heterodimers); product to amplify, 75 to 250 bp with a GC content of 40 to 60 %. Primer pairs were further verified with PrimerExpress (Applied Biosystems, USA) and the online primer designing tools IDT Oligo Analyzer (Integrated DNA Technologies, USA) and Primer-BLAST (NCBI, 2009). Primer sequences are presented in Table 4.3. Initially only three primer pairs (752, 2380 and 13596) were designed for target sequences, together with two for reference sequences (8629 and 12621). When the target primers did not amplify products from cDNA template isolated from cultured haemocytes, another four primer pairs (7048, 54, 809, and 3309) were designed for target sequences regarded as more specifically expressed in haemocytes.

Table 4.3 Primer sequences for amplification of target and reference genes

Sequence name	Forward primer (5'→3')	Reverse primer (5'→3')
Target genes		
752 (Fgf receptor/insulin receptor)	CTTCACTGGGGATTCGGAC	AGGATGCGAATGAGAAGAGTG
2380 (Thrombospondin precursor)	GACCAATCGCTCCATCCAC	GAGGACCCGCCAGTGTAAC
13596 (Perlustrin)	TTTATACGCTTCTATCTTGTTCCC	CCCTGTATTCTGATGCCTTGT
7048 (HSP 90)	ACCTCCCCTTGCTGTATTGTC	TGCCTTCTCCTTCAGTGTCTTC
54 (Collagen)	GGTCCTACTGGCTTTCAACAAC	GGGAGCATCACATACAACACAG
809 (MAP kinase/Mnk)	AATGGAGGATAACCGCAAAG	GAGATTCAGATGTCCGCACC
3309 (Cyclin)	AGGTAAGCAGATGGTGGCAG	GAGTGTGGCTGTTCCGTTTC
Reference genes		
8629 (Ribosomal protein S9)	CATCCAACACACCAATACGC	TTGATGAGAAGGAACCCAGAC
12621 (Ornithine decarboxylase)	TCATCGTCCCCAGTCACC	GCTTTCACTCTCACTGCCAAC

4.2.4 PCR optimization

Preceding real-time PCR experiments, successful amplification of all products were verified with conventional PCR. Different annealing temperatures (55 °C, 58 °C and 60 °C) and primer concentrations (50, 100, 200 and 400 nM) were tested for the first five primer pairs (three target sequences and two reference sequences, Table 4.5) in order to arrive at optimal amplification of products. Presence of the correct-sized products was confirmed by 2 % agarose gel electrophoresis. At the optimal annealing temperature for these primer pairs, two primer concentrations (50 and 100 nM) were tested by real-time PCR. The subsequent four primer pairs were also tested at four different annealing temperatures (55 °C, 57 °C, 58 °C and 60 °C), only at the optimal primer concentration (100 nM) from the initial test.

4.2.5 Reference gene validation

The selection of proper reference genes for data normalisation is important in order to deduce true expression differences between genes in a sample. The geNorm method (Vandesompele *et al*., 2002) is considered as the standard algorithm for evaluating candidate reference genes in terms of expression stability (Derveaux *et al*., 2010). GeNorm returns an average expression stability (*M*) value that describes the variation of a gene compared to all other candidate genes. The lower the *M* value, the more stable the expression of the gene (Vandesompele *et al*., 2002). The reference genes used in the present study were selected on the basis of equal expression determined by digital differential expression analysis (Chapter 2). In order to verify that the two selected reference genes were indeed suitable for qPCR normalisation purposes, a geNorm analysis was performed on both sample types. For whole abalone tissue samples, five genes (two reference genes - 12621 and 8629 - and three target genes - 752, 13596 and 2380) were included in a geNorm analysis, using stepwise elimination of the least stable genes. Following this process, the last pair of genes remaining is recommended as the best possible pair of reference genes. The same procedure was repeated for the six genes tested in cultured abalone haemocytes (two reference genes - 12621 and 8629 - and four target genes - 7048, 54, 3309 and 809).

4.2.6 Quantitative real-time PCR

All real-time PCR runs were performed on the Rotor-Gene 6000 (Corbett Research, Australia) in the 36-well rotor, using 0.2 ml thin walled PCR tubes. Initially, qPCR was performed on cDNA constructed from RNA that was isolated from whole body tissues of small and large abalone that was used for transcriptome sequencing (Chapter 2). This was done to verify the differential expression observed by digital differential expression analyses. Reverse transcription was performed as described above for haemocyte samples, starting with 160 ng of RNA per sample, followed by Ribonuclease H treatment. cDNA from small animals (S), large animals (L) and an equal volumes mixture of S and L as calibrator (C_b) was used as template. Standard curves were constructed for all primer pairs to determine the accuracy and sensitivity of the assay and calculate amplification efficiencies (Adams, 2006). Calibrator cDNA was used for this purpose and six five-fold serial dilutions were prepared, starting with 400 ng/µl. Each 20 µl reaction contained 1 µl of first-strand cDNA, 10 µl of 2x SYBRFAST qPCR Master Mix (KapaBiosystems, USA) and 0.2 µl (100 nM) of each primer. The thermocycle program used was 95 °C for five minutes, followed by 40 cycles of 95 °C for 30 seconds, 58 °C for 30 seconds and 72 °C for 30 seconds. After cycling, a melt curve analysis was performed, consisting of an increase in temperature by 1 °C every five seconds, from 72 °C to 95 °C. Each sample was evaluated in triplicate and each PCR run included a no template control (NTC) reaction.

After verifying amplification efficiencies from the standard curves, target and reference primer pairs with comparable efficiencies were used together in subsequent quantification runs on the Rotor-Gene 6000. For quantification runs, each 20 µl reaction contained 1 µl of an 80 times dilution of first-strand cDNA (±10 ng), 10 µl of 2x SYBRFAST qPCR Master Mix (KapaBiosystems, USA) and 0.2 µl (100 nM) of each primer. The same thermocycle program described above was used, followed by a melt curve analysis. Four replicates of each cDNA (L, S and C_b) were included per run and for each primer in a run a no template control (NTC, cDNA substituted with water), a minus reverse transcriptase control (-RT, where no SS III was added during reverse transcription) and a positive control (50 ng of L cDNA) were included.

Quantitative real-time PCR was subsequently performed on cDNA constructed from RNA that was isolated from cultured abalone haemocytes (sections 4.2.1 and 4.2.2). After initial verification of amplification by conventional PCR, using this new cDNA as template, only the reference genes (8629 and 12621) amplified. Lowering the annealing temperature to 55 °C did not result in amplification of the original target sequences. Real-time amplification on the Rotor-Gene 6000, using 200 ng cDNA as template and 55 °C as annealing temperature for 40 cycles (keeping all other parameters similar to previous runs), confirmed amplification in the reference sequences, but not in the original target sequences.

For new primers (7048, 54, 3309, and 809; Table 4.3) that were designed on target sequences regarded as more specifically expressed in haemocytes, initial verification of amplification by conventional PCR was repeated. Presence of the correct-sized products was confirmed by 2 % agarose gel electrophoresis.

Standard curves were constructed for the new target sequence primers to determine the accuracy and sensitivity of the assay and calculate amplification efficiencies (Adams, 2006). Cycling conditions were kept the same as for previous qPCR runs, to ensure comparability between all runs. Equal volumes of cDNA originating from three biological replicates of control cell cultures were combined as calibrator cDNA (C_b). This was used as template for standard curve reactions. From a starting concentration of 400 ng/µl, six five-fold serial dilutions were prepared. Reaction conditions were the same as for generation of standard curves from initial primer pairs. Each sample was evaluated in triplicate and each PCR run included a NTC reaction.

Construction of standard curves was followed by quantification runs, where target and reference primer pairs with comparable efficiencies (within 10 % of each other) were combined in one run. For quantification runs, each 20 µl reaction contained 1 µl of a 30 times dilution of first-strand cDNA (±20 ng), 10 µl of 2x SYBRFAST qPCR Master Mix (KapaBiosystems, USA) and 0.2 µl (100 nM) of each

primer. The same thermocycle program as described above was used, followed by a melt curve analysis. Three replicates of each cDNA (C_n, I-4, I-50, E-25, and E-600) were included per run and for each primer pair in a run, a no template control (cDNA substituted with water), a minus reverse transcriptase control (-RT, where no SS III was added during reverse transcription) and a positive control (50 ng of L cDNA) were included.

4.2.7 Data analysis and bioinformatics

Real-time PCR data was analysed using the Rotor-Gene 6000 Series Software version 1.7.87 and REST (Pfaffl and Corbett Research, 2008). For construction of standard curves, the "Quantitation" function in the Rotor-Gene 6000 Series Software was used. Standard curves were, as far as possible, constructed on triplicate C_q-values per sample. The standard curve for a perfect assay will have a slope of -3.32 (100 % efficiency), a y-intercept of 33 to 37 cycles (the essay is sensitive and the template was quantified accurately) and an r^2 of 1.00 (a good fit of the data to the theoretical regression line) (Adams, 2006).

For relative quantification data analysis, the comparative CT method (Livak and Schmittgen, 2001) was performed using the Rotor-Gene 6000 Series Software ("Delta Delta CT" analysis). When using this method, it is essential that the amplification efficiencies of the target gene (gene of interest) and the reference gene (normaliser gene) are identical (Rotor Gene 6000 operator manual, 2006). The method to verify similarity of amplification efficiencies proposed by Livak and Schmittgen (2001) was employed. When the slope of the ΔCt plot was smaller than 0.1 (amplification efficiencies differ by less than 10 %), the amplification efficiencies were considered as similar (Schmittgen and Livak, 2008).

The RG Mode of REST (2008) was also employed to verify relative expression. REST determines whether there are significant differences between samples and controls, taking into account reference gene normalization and reaction efficiency, by employing randomization techniques (Pfaffl and Corbett Research, 2008). REST, which is based on an efficiency corrected mathematical model for data analysis, calculates relative expression ratios using the PCR efficiency (E) and crossing point deviation (delta CT) of transcripts in the sample and control groups. Constant amplification efficiencies in all compared samples are needed for reliable comparison between samples, since small efficiency differences between target and reference genes generate a false expression ratio. When amplification efficiencies are not constant, efficiency correction is important because it results in a more reliable estimation of the real expression ratio compared (Pfaffl *et al*., 2002; Pfaffl, 2006). Data resulting from a "Comparative Quantitation analysis" in the Rotor-Gene Software is entered into REST for relative expression analysis in RG Mode.

4.3 Results

4.3.1 Haemocyte cell cultures

Haemocytes that were maintained in primary cultures had normal viability and no contamination throughout 65 hours of culture. Cultures consisted of adherent fibroblast-like cells and floating amoeboid-like cells. There was no apparent difference in cell morphology for cells maintained in the different supplemented media.

4.3.2 RNA extraction and reverse transcription

Following gel electrophoresis, presence of the expected single 28S RNA band between 1500 and 2000 bases for all samples was confirmed (Figure 4.2). Concentrations and absorbance ratios for RNA following extraction are presented in Table 4.4.

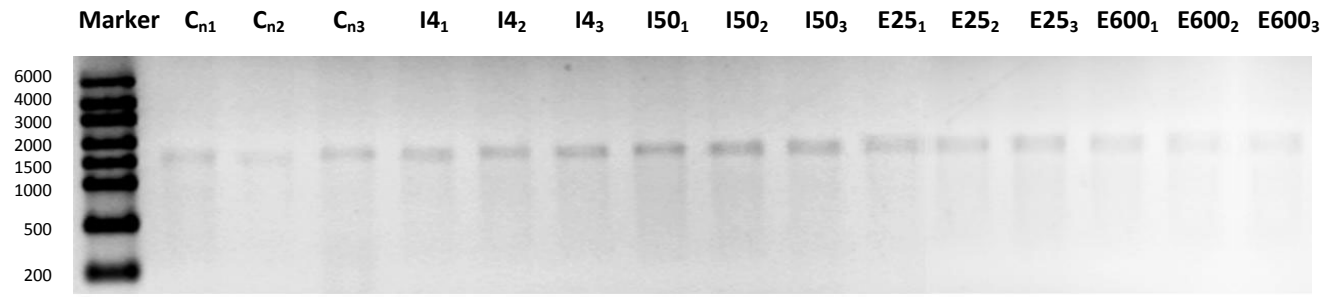


Figure 4.2 Denaturing 2 % agarose gel of RNA isolated from haemocyte cell cultures. C_n = control; I4/I50= ASW with bovine insulin supplement (4 µg/ml or 50 µg/ml); E25/600 = ASW with EGF supplement (25 ng/ml or 600 ng/ml)

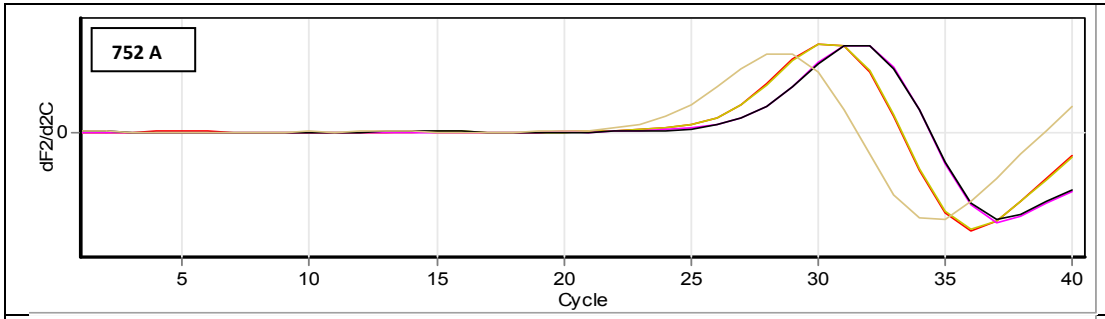
Table 4.4 Concentration and 260/280 absorbance ratios for RNA samples (presented as an average for three replicates per sample)

	C_n	I-4	I-50	E-25	E-600
Concentration (ng/µl)	31.03	30.83	36.37	31.37	28.07
260/280	2.14	1.86	1.96	1.96	1.94

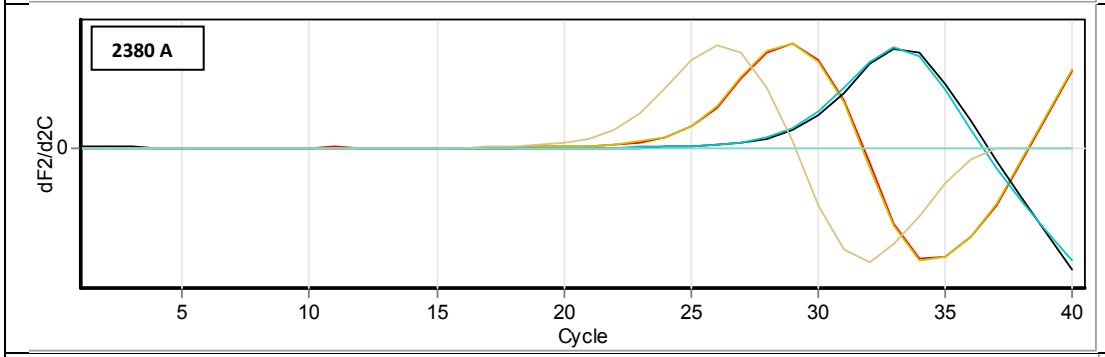
C_n = control; $I_{4/50}$ = ASW with bovine insulin supplement (4 µg/ml or 50 µg/ml); $E_{25/600}$ = ASW with EGF supplement (25 ng/ml or 600 ng/ml)

The level of genomic DNA contamination following DNase digestion and reverse transcription is reported in Figure 4.3 as second derivatives of the amplification plots from real-time PCR reactions. The peaks in the graphs correspond to the maximum rate of fluorescence increase in the reaction. The takeoff point (Cq) is defined as the cycle at which the second derivative is at 20 % of the maximum level, and indicates the transition into the exponential phase (Rotor Gene 6000 operator manual, 2006). Minus reverse transcription controls (-RT) had very low exponential amplification efficiencies (where 2.0 equals 100 % efficiency) and in the cases where some product did amplify, the Cq's appeared 7 to 10 cycles later than target amplicons. All amplification efficiencies were calculated in the Rotor-Gene

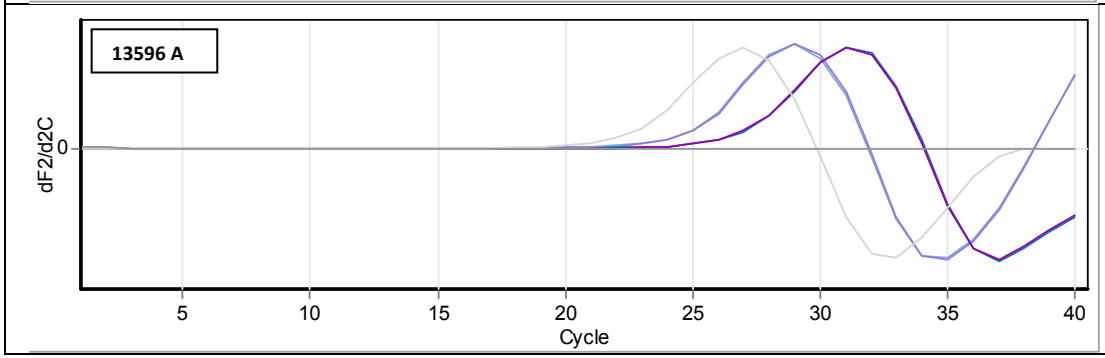
6000 Series Software. The slope op the standard curve (**M**) is used to determine the amplification efficiency where: exponential amplification = $10^{(-1/\mathbf{M})}$.



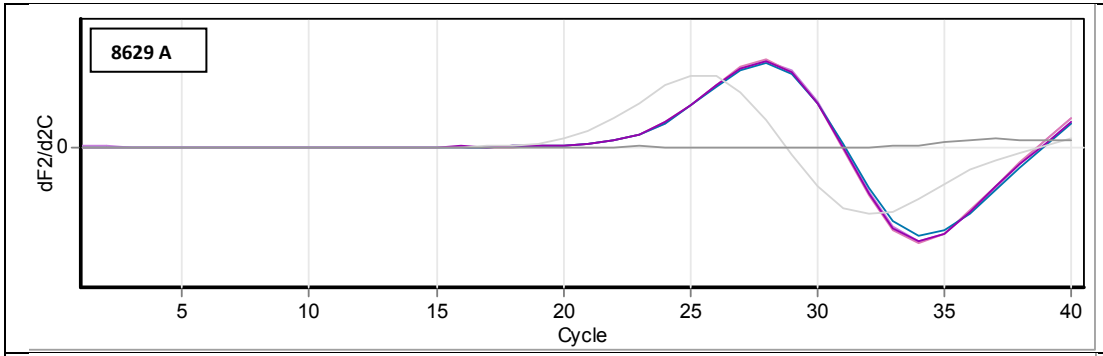
Name	C_q	Efficiency
L1	27	1.86
L2	27	1.91
S1	28.1	1.84
S2	28	1.86
P	24.7	1.76
-RT	34.6	0.56



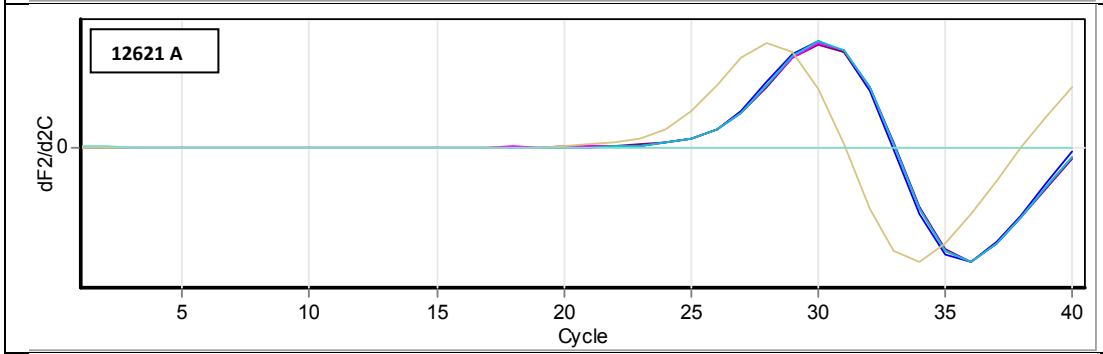
Name	C_q	Efficiency
L1	25.6	1.87
L2	25.5	1.9
S1	29.9	1.79
S2	29.7	1.81
P	22.9	1.84
-RT	9.5	-0.03



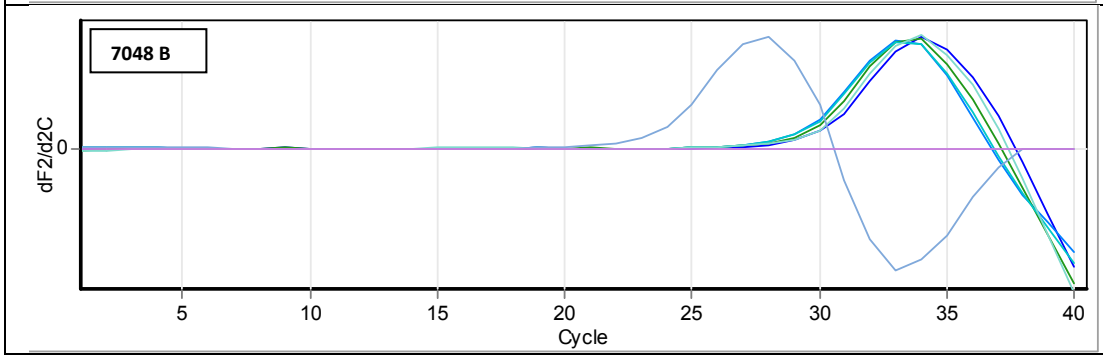
Name	C_q	Efficiency
L1	25.8	1.94
L2	25.9	1.96
S1	27.9	1.87
S2	27.9	1.88
P	23.6	1.86
-RT	21.3	0.06



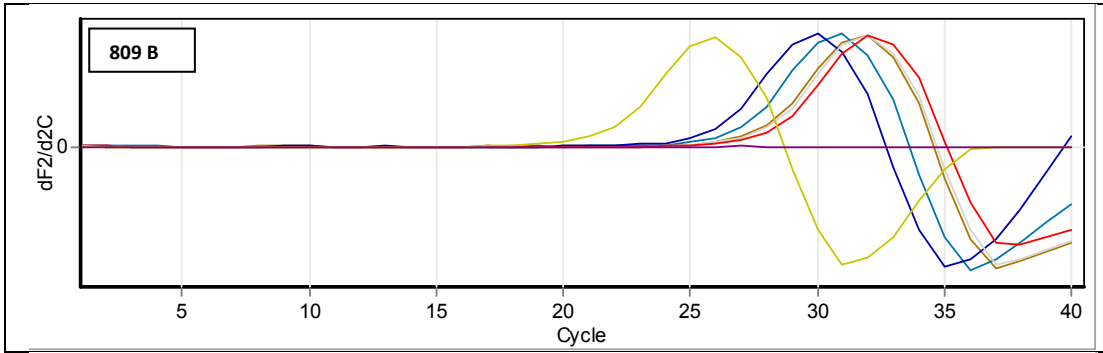
Name	C_q	Efficiency
L1	24.1	1.81
L2	24	1.82
S1	24	1.8
S2	24	1.77
P	21.5	1.85
-RT	33.9	2.32



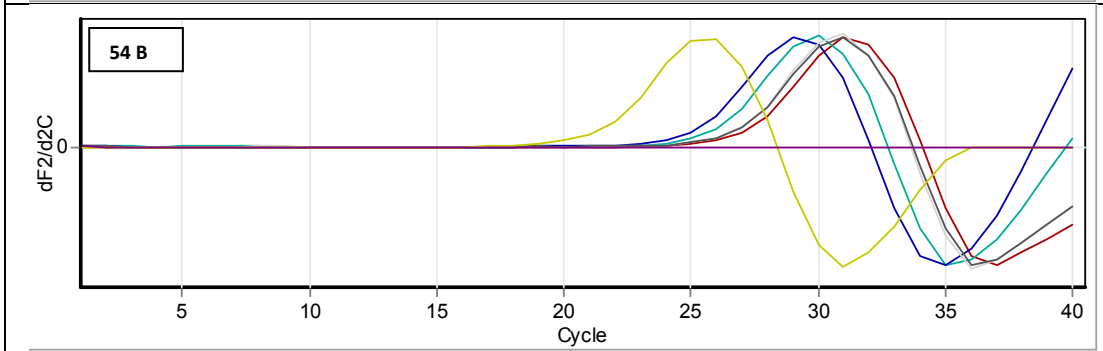
Name	C_q	Efficiency
L1	26.8	1.93
L2	26.9	1.88
S1	26.9	1.9
S2	26.9	1.92
P	24.9	1.89
-RT	7.9	0.6



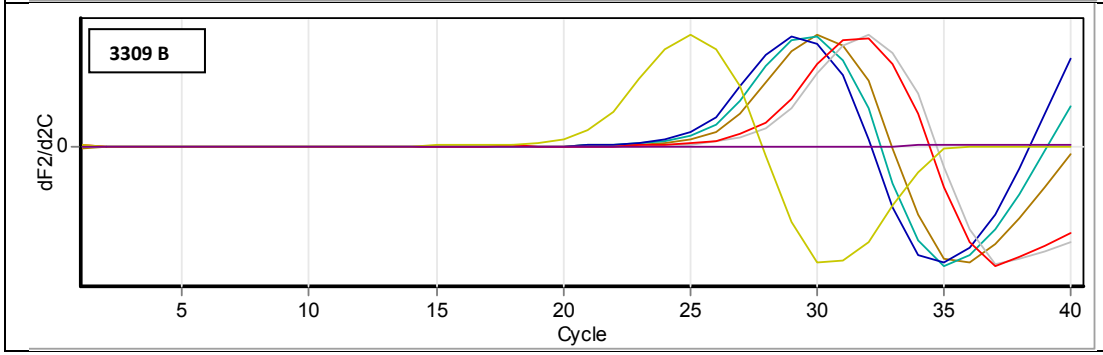
Name	C_q	Efficiency
C_n	31	2.01
I-4	30.3	1.98
I-50	30.5	2.03
E-25	30.3	1.94
E-600	30.8	2.05
P	24.7	2.00
-RT	29.1	-0.24



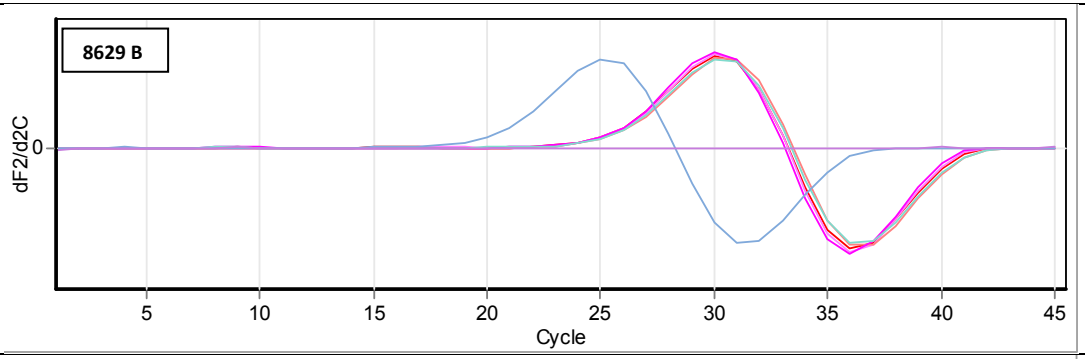
Name	C_q	Efficiency
C_n	28.7	2.02
I-4	27.8	2.04
I-50	26.9	2.06
E-25	28.8	2.03
E-600	29.2	1.99
P	22.7	1.98
-RT	25.9	-0.01



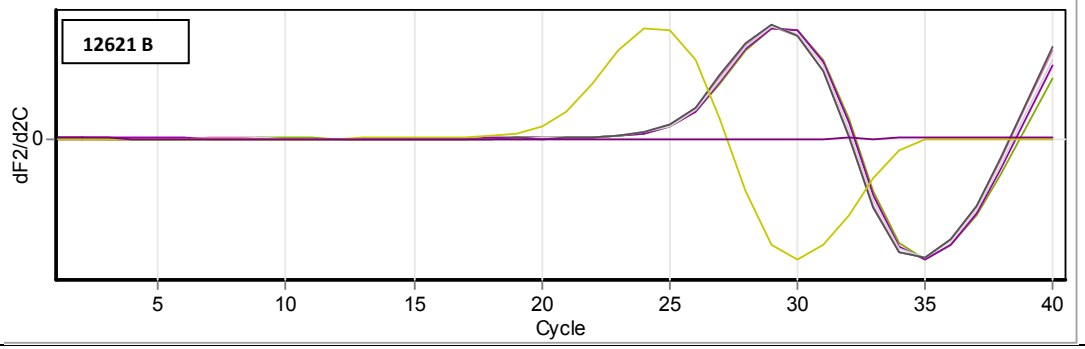
Name	C_q	Efficiency
C_n	28.2	2
I-4	26.8	2.04
I-50	26.5	1.96
E-25	27.9	2.04
E-600	27.8	1.98
P	22.5	2.02
-RT	25.6	0.28



Name	C_q	Efficiency
C_n	27.1	2
I-4	26.7	2.01
I-50	26.3	2.01
E-25	28.9	1.99
E-600	28.5	2.02
P	22.0	2.05
-RT	29.8	0.83



Name	C_q	Efficiency
C_n	26.5	1.81
I-4	26.6	1.82
I-50	26.7	1.82
E-25	26.6	1.82
E-600	26.6	1.82
P	21.5	1.79
-RT	38.1	0.15



Name	C_q	Efficiency
C_n	26.2	2
I-4	26.4	2.05
I-50	26.4	2.05
E-25	26.3	2.05
E-600	26.2	2.07
P	21.4	2.01
-RT	34.6	1.55

Figure 4.3 Second derivative amplification plots for all primers. A = PCR products amplified from cDNA of whole animals (S and L), B = PCR products amplified from cDNA of cultured haemocytes (C_n, I-4, I-50, E-25, E-600). P indicates the positive control and -RT the minus reverse transcription control.

4.3.3 Primer optimization and confirmation of primer specificity

Results of primer optimization in terms of annealing temperatures and primer concentrations for the initial five primer pairs are presented in Table 4.5. Presence of the correct-sized products was evident as clear bands on 2 % agarose gels (Table 4.6, Figure 4.4).

Table 4.5 PCR optimization for the first five primer pairs

			Genes				
Annealing temperature	**PCR method**	**Primer concentration**	**12621**	**13596**	**2380**	**752**	**8629**
55 °C	Conventional	400nM	Yes	Yes	Yes	Yes	Yes
58 °C	Real-time	50 nM	No	No	No	No	No
	Real-time	100 nM	Yes	Yes	Yes	Yes	Yes
	Conventional	100 nM	Yes	Yes	Yes	Yes	Yes
	Conventional	200 nM	Yes	Yes	Yes	Yes	Yes
	Conventional	400 nM	Yes	Yes	Yes	Yes	Yes
60 °C	Conventional	100 nM	Yes	Yes	Yes	No	No
	Conventional	200 nM	No	No	No	No	No
	Conventional	400 nM	No	No	No	No	No

Conventional: Yes = one clear band on gel, no dimers; No = smear or dimers visible. Real-time: Yes = normal sigmoidal curve, no dimers; No = incomplete curve or curve with dimers

For the subsequent four primer pairs, 2 % agarose gel electrophoresis also confirmed clear bands for the correct sized products (Table 4.6, Figure 4.5) with no dimers, at all the tested annealing temperatures (55 °C, 57 °C, 58 °C and 60 °C) using 100 nM of primer. Primer specificity for all primers was furthermore confirmed by melt curve analysis following real-time PCR on triplicate samples. Results are presented in Figure 4.6.

Table 4.6 Expected sizes of amplification products

Primer	**Product size (bp)**	**Figure**	**Primer**	**Product size (bp)**	**Figure**
752 (Fgf /insulin receptor)	206	4.4	*7048* (HSP 90)	171	4.5
2380 (Thrombospondin)	132	4.4	*809* (MAP kinase/Mnk)	172	4.5
13596 (Perlustrin)	218	4.4	*54* (Collagen)	101	4.5
8629 (Ribosomal protein)	78	4.4	*3309* (Cyclin)	235	4.5
12621 (ODC)	101	4.4			

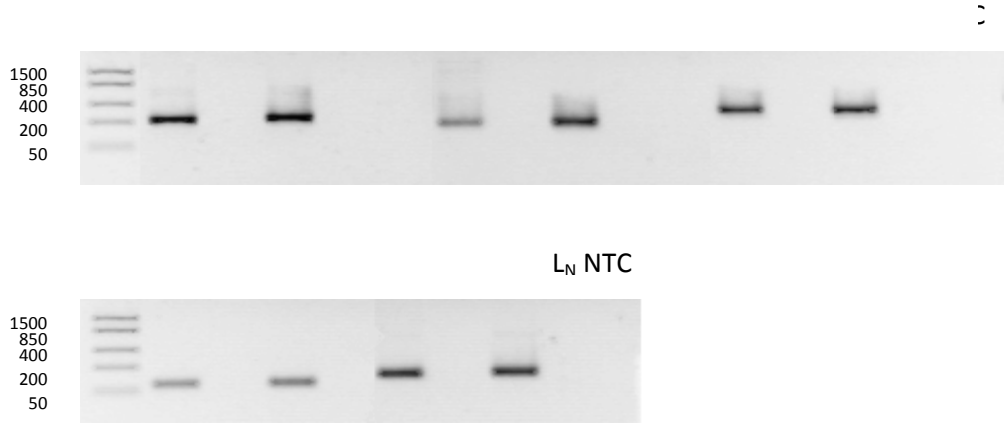


Figure 4.4 2 % Agarose gels of PCR products amplified from cDNA of large (L) and small (S) animals using the initial five primer pairs; M = DNA marker, $_N$ = minus reverse transcription control, NTC = no template control

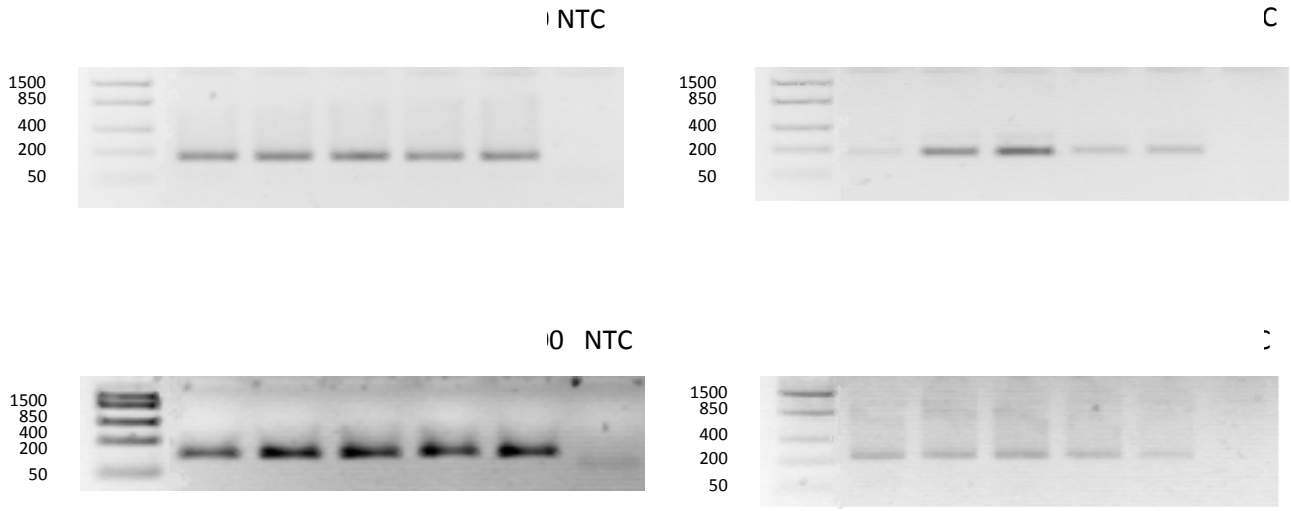
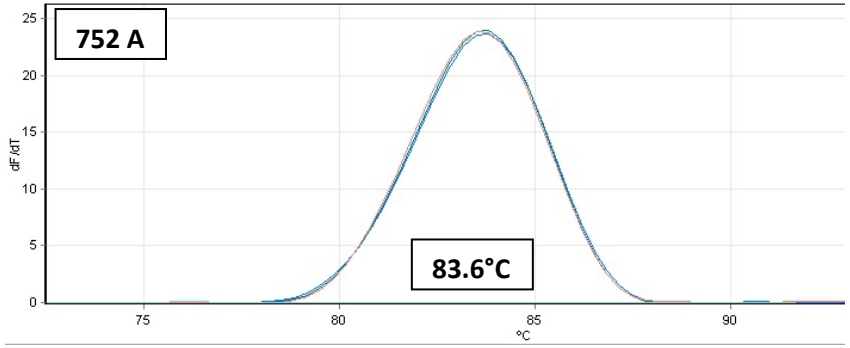
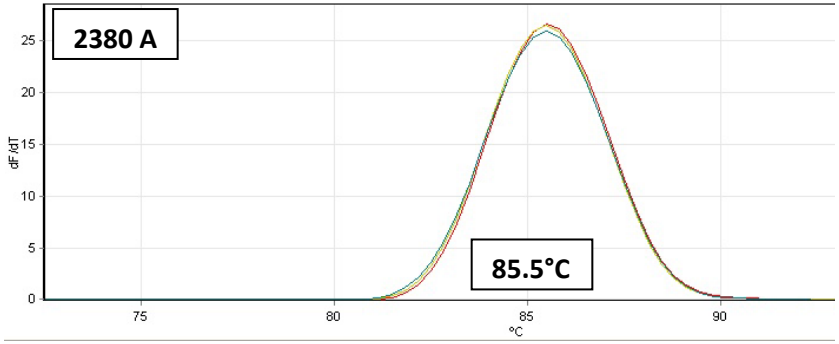
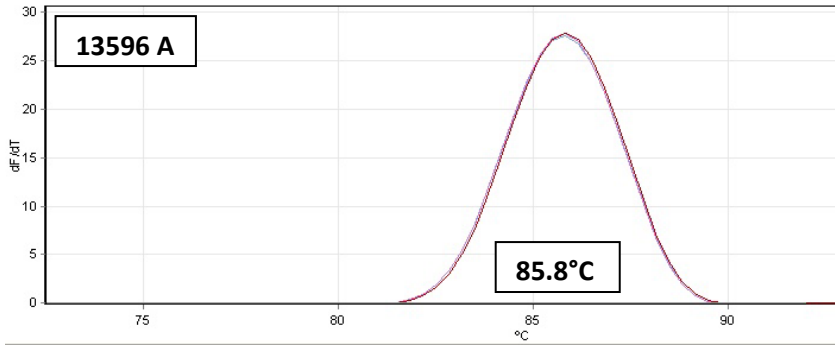
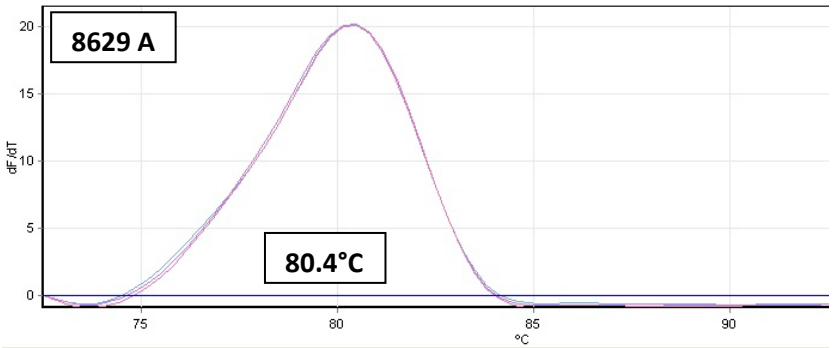


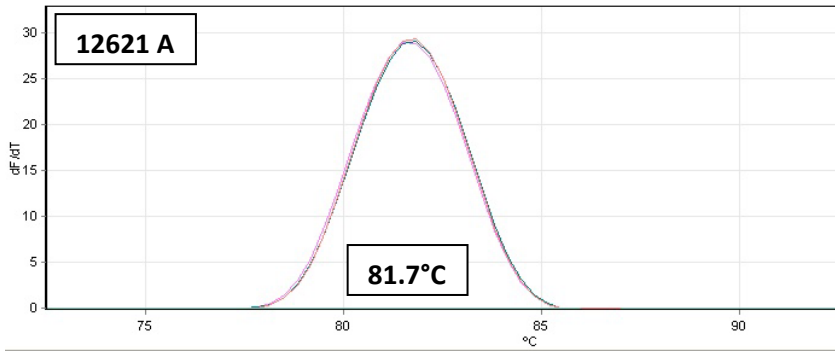
Figure 4.5 2 % Agarose gels of PCR products amplified from cDNA of cultured haemocytes using four subsequent primer pairs; M = DNA marker, C_n = untreated control, I-4 = 4 µg/ml bovine insulin-treated, I-50 = 50 µg/ml bovine insulin-treated, E-25 = 25 ng/ml EGF-treated, E-600 = 600 ng/ml EGF-treated, NTC = no template control

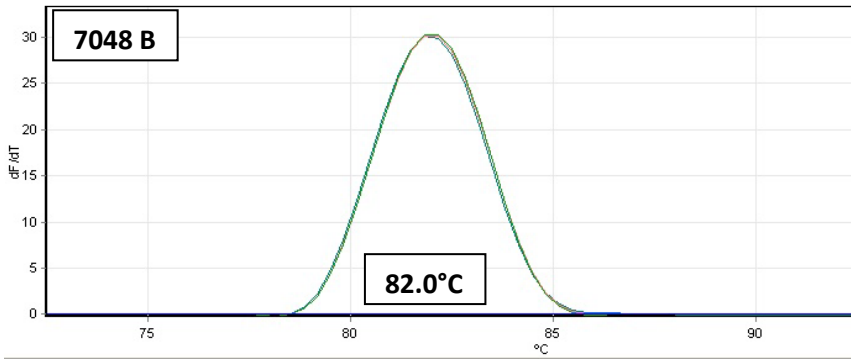

752 A
83.6°C
dF/dT
°C


2380 A
85.5°C
dF/dT
°C


13596 A
85.8°C
dF/dT
°C


8629 A
80.4°C
dF/dT
°C


12621 A
81.7°C
dF/dT
°C


7048 B
82.0°C
dF/dT
°C

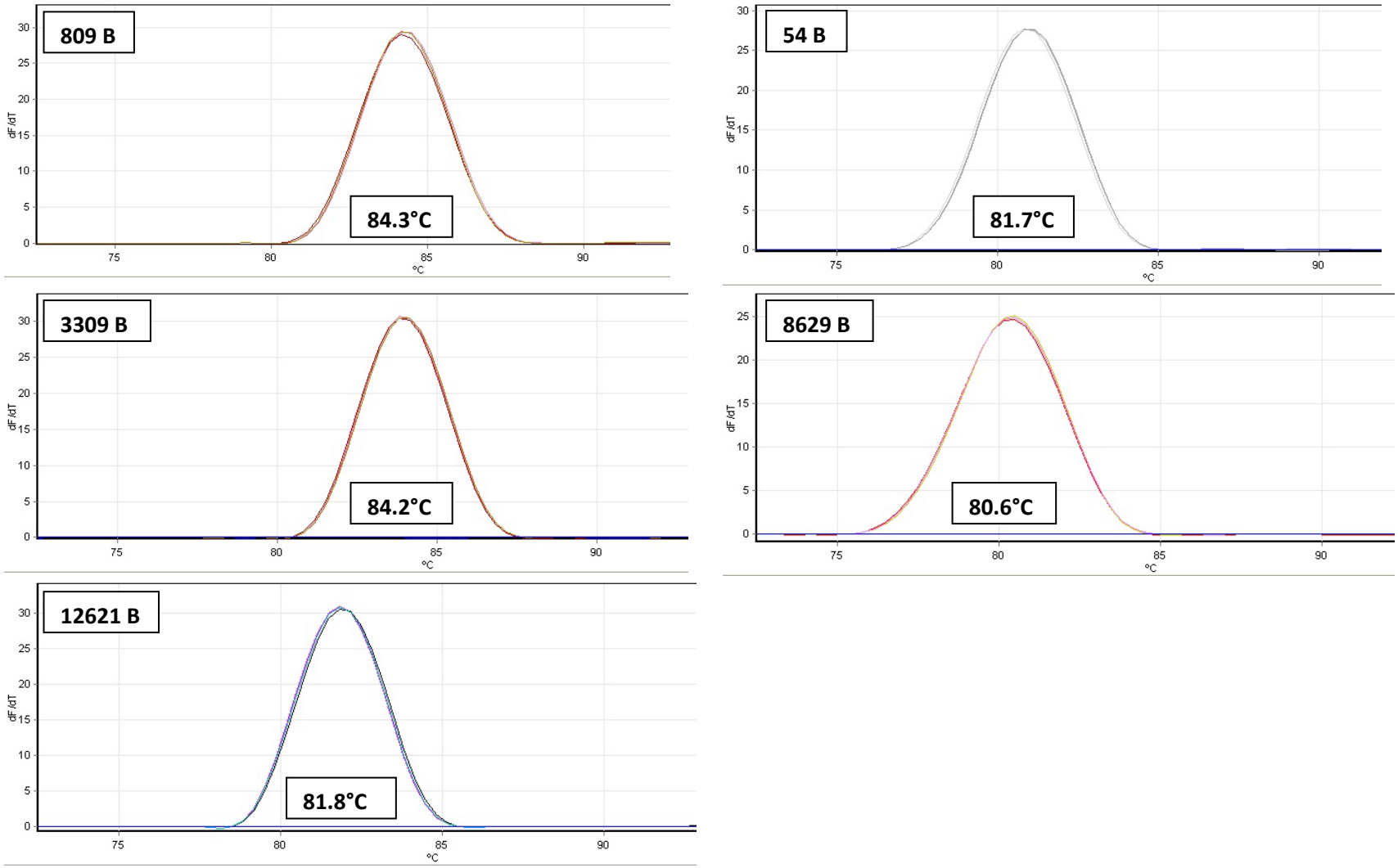


Figure 4.6 Melt curve analysis following real-time PCR on triplicate samples; A = PCR products amplified from cDNA of whole animals, B = PCR products amplified from cDNA of cultured haemocytes. The triplicate average melting temperatures are indicated in the boxed area below each graph.

The peak of each melt curve in Figure 4.6 corresponds to a single amplified fragment with a specific melting temperature. This confirms the specificity of the reactions, as no primer dimers or other non-specific amplification is observed.

4.3.4 Reference gene validation

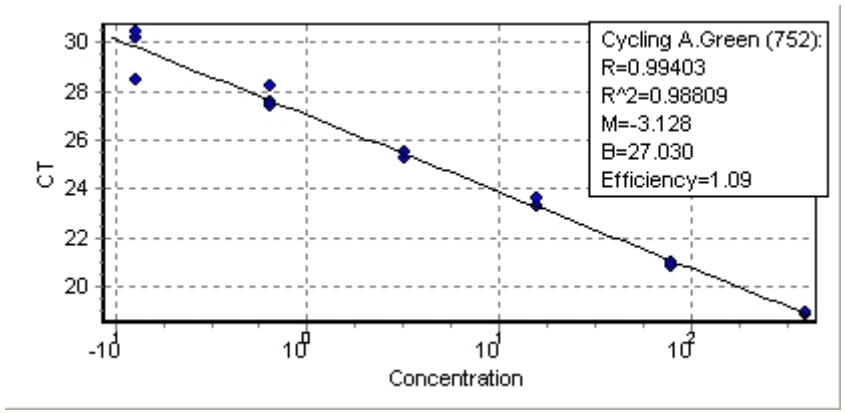
Following the geNorm procedure, the reference genes 12621 (ornithine decarboxylase) and 8629 (ribosomal protein S9) were confirmed to be valid reference genes for qPCR in both sample types (whole soft body tissue and cultured haemocytes). After stepwise elimination of the least stable genes, these two genes remained as the most stable reference genes (M = 0.131 for whole abalone tissue samples, compared to the least stable gene 2380 with M = 1.679; and M = 0.154 for cultured abalone haemocytes sample, compared to the least stable gene 3309 with M = 0.886).

4.3.5 Quantitative real-time PCR

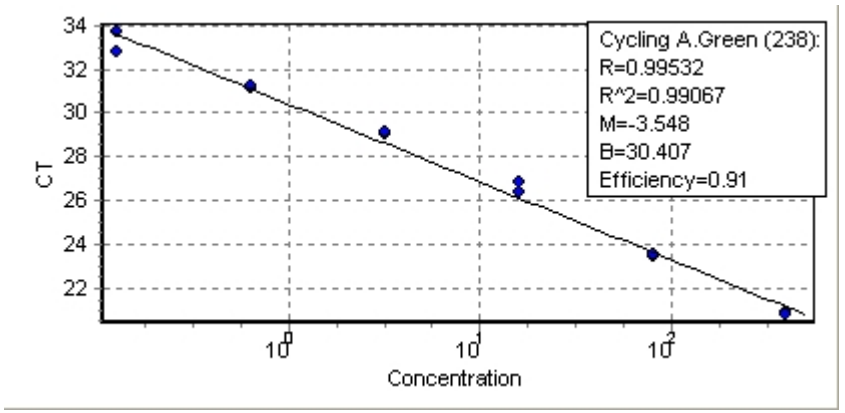
4.3.5.1 Standard curves

Standard curves were constructed to estimate amplification efficiencies of real-time PCR reactions for each primer pair (Figure 4.7). The log of the starting quantity of the template was plotted against the C_q values. In most cases, triplicate reactions of a five-fold dilution series over six points are presented. In some instances, the last point was omitted to improve the linearity of the standard curve. Optimal reaction efficiency between 90 and 110 % (0.9 and 1.1) with an r^2-value > 0.980 indicates that the data fits the linear regression line well (Taylor *et al.*, 2010).

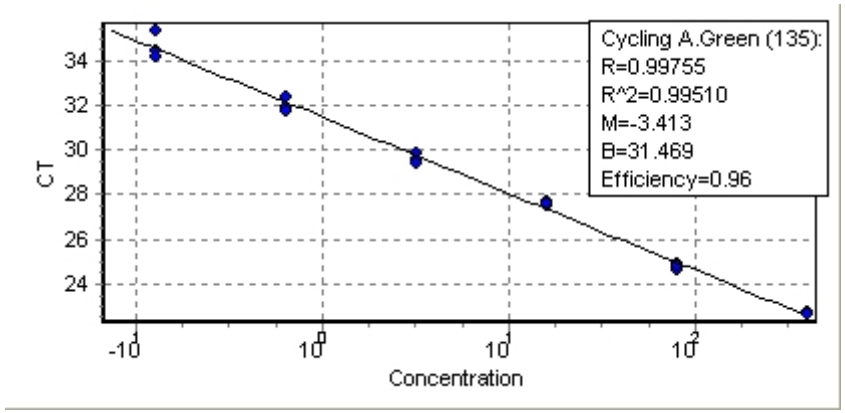
752 A



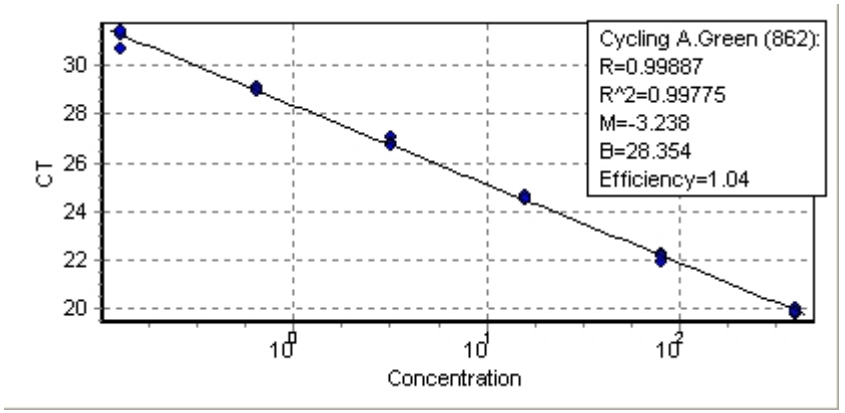
2380 A



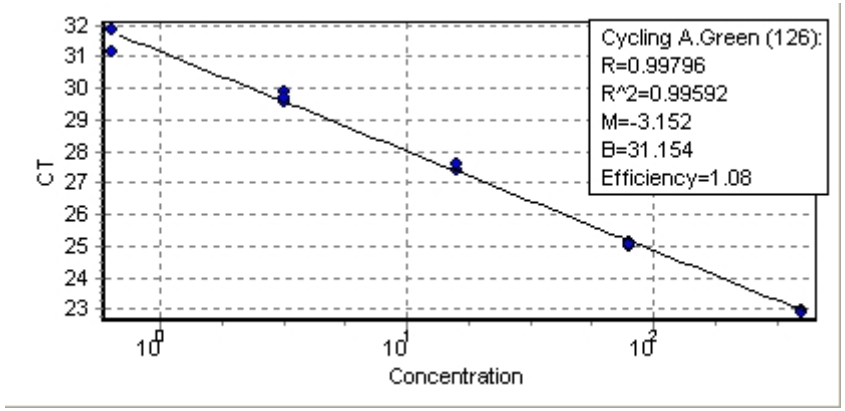
13596 A



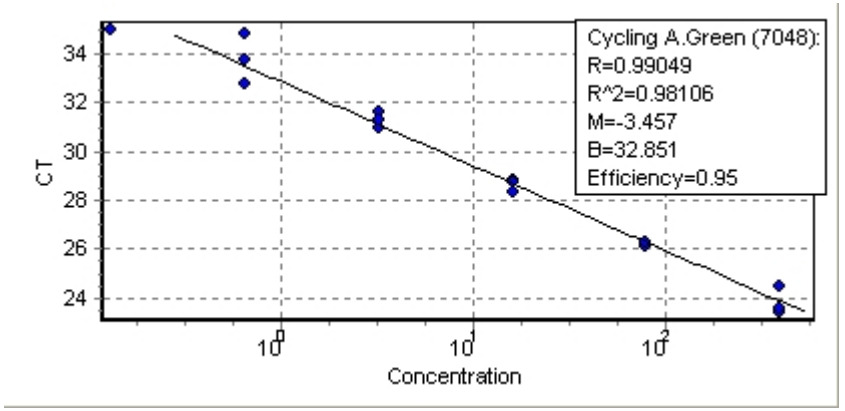
8629 A



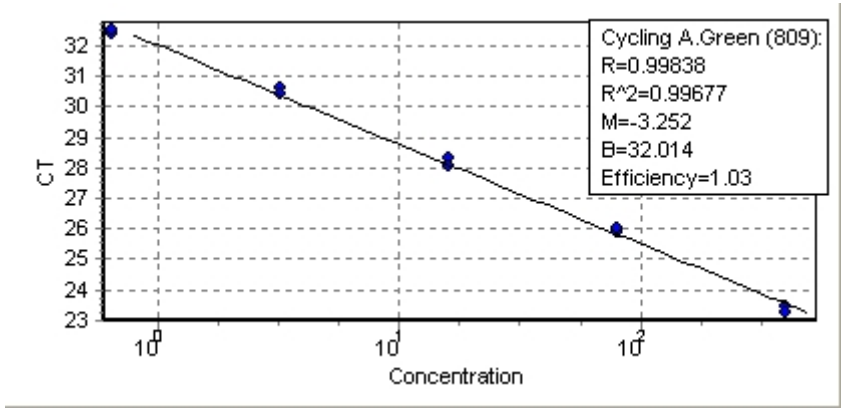
12621 A



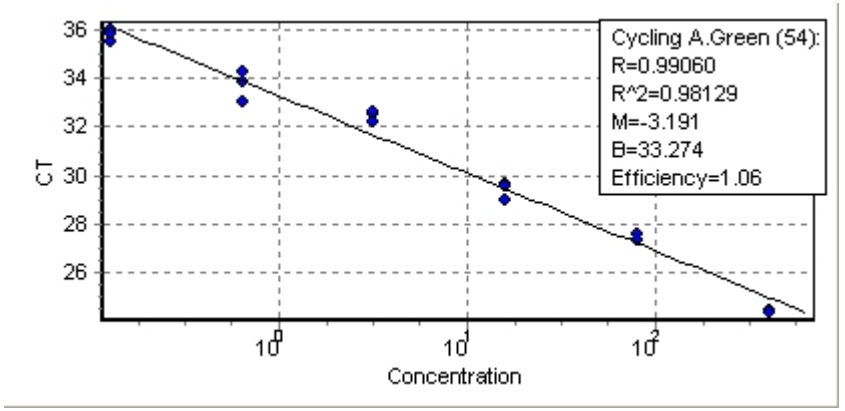
7048 B



809 B



54 B



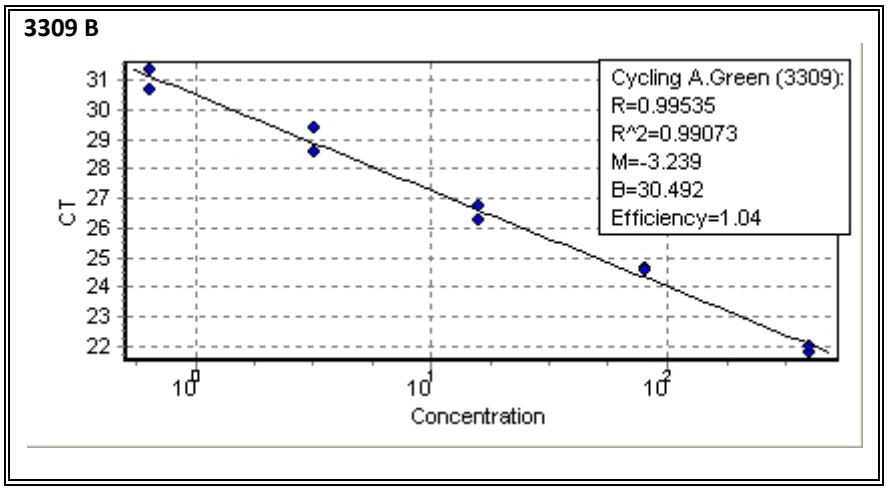


Figure 4.7 Standard curves for determining amplification efficiencies of all primers. A = PCR products amplified from cDNA of whole animals, B = PCR products amplified from cDNA of cultured haemocytes. For reference genes, only standard curves of A samples are indicated.

4.3.5.2 Data analysis and bioinformatics (Delta delta CT results, REST)

For samples that had equal amplification efficiencies (within 10 % of each other), relative changes in gene expression was analysed using the comparative CT module in the Rotor-Gene 6000 Series Software. This module is based on the $2^{-\Delta\Delta CT}$ method, a method of relative quantification without efficiency correction (Livak and Schmittgen, 2001; Rotor Gene 6000 operator manual, 2006; DNAbiotec (Pty) Ltd, 2008). Accordingly, reference gene 8629 could be combined with target genes 752, 13596, 7048, 809, 54 and 3309. Reference gene 12621 could be combined with target genes 752, 809, 54 and 3309. Target gene 2380 could not be combined with any reference gene due to unequal amplification efficiencies (slope of ΔCT plot > 0.1). ANOVAs and t-tests were performed on relative concentration values emanating from the expression analysis to determine significant differences between groups. Relative concentrations and results of ANOVAs / t-tests are summarised in Table 4.7. Using a significance level of 0.05, none of the target genes were indicated as up- or down-regulated for samples derived from whole animal soft body tissues. With the same significance level ($P < 0.05$), one gene (3309 i.e. Cyclin) was indicated as down-regulated in two samples derived from cultured haemocytes treated with EGF (25 ng/ml and 600 ng/ml) following t-tests. One gene (809 i.e. MAP kinase/Mnk) was indicated as up-regulated in one sample derived from cultured haemocytes treated with bovine insulin (50 µg/ml) following t-tests. Although some target genes appear to be up- or down-regulated, according to the raw data, this regulation cannot be regarded as statistically significant.

Table 4.7 Relative fold change calculated by the $2^{-\Delta\Delta CT}$ method. Replicates printed in bold were used as calibrators and differences between sample groups and control group are indicated with "=" (no significant difference), "↑" (up-regulated) and "↓" (down-regulated) according to ANOVA and t-test*; $P < 0.05$

752_8629				**752_12621**			**13596_8629**			
Replicate name	Relative fold change	P-value	=↑↓	Relative fold change	P-value	=↑↓	Replicate name	Relative fold change	P-value	= ↑↓
C_1	1.20			0.97			C_1	1.02		
C_2	1.00			1.00			**C_2**	1.00		
C_3	1.11			1.16			C_3	1.02		
C_4	1.34			1.14			C_4	1.03		
L_1	1.07			1.01			L_1	1.01		
L_2	1.10	0.54	=	1.05	1.00	=	L_2	1.02	1.00	=
L_3	2.00			1.20			L_3	1.02		
L_4	1.10			1.01			L_4	1.02		
S_1	1.05			0.86			S_1	1.00		
S_2	0.76	0.17	=	0.93	0.54	=	S_2	0.99	0.14	=
S_3	0.64			1.05			S_3	1.00		
S_4	1.24			1.20			S_4	1.02		

$C_{1,2,3,4}$= four replicates calibrator cDNA; $L_{1,2,3,4}$ = four replicates cDNA from large animals; $S_{1,2,3,4}$ = four replicates cDNA from small animals

54_8629				**54_12621**			**3309_8629**				**3309_12621**		
Replicate name	Relative fold change	P-value	= ↑↓	Relative fold change	P-value	= ↑↓	Replicate name	Relative fold change	P-value	= ↑↓	Relative fold change	P-value	= ↑↓
C_1	1.78			0.44			C_1	2.06			2.21		
C_2	0.97			2.08			C_2	0.13			1.85		
C_3	1.00			1.00			**C_3**	1.00			1.00		
I-4_1	0.98			1.33			I-4_1	0.70			0.88		
I-4_2	1.10	0.98	=	0.46	0.7	=	I-4_2	1.31	0.96	=	1.55	0.54	=
I-4_3	1.70			1.06			I-4_3	1.08			1.74		
I-50_1	1.50			0.61			I-50_1	1.54			1.40		
I-50_2	1.59	0.33	=	0.71	0.30	=	I-50_2	0.57	0.88	=	0.88	0.21	=
I-50_3	1.55			0.46			I-50_3	1.39			1.05		
E-25_1	0.53			1.46			E-25_1	0.18			0.33		
E-25_2	0.83	0.35	=	1.25	0.86	=	E-25_2	0.14	0.38	=	0.45	0.04*	↓
E-25_3	1.29			0.48			E-25_3	1.01			1.21		
E-600_1	0.59			1.38			E-600_1	0.25			0.21		
E-600_2	0.63	0.08	=	0.95	0.84	=	E-600_2	0.39	0.26	=	0.42	0.02*	↓
E-600_3	0.63			0.86			E-600_3	0.34			0.30		

809_8629				**809_12621**			**7048_8629**			
Replicate Name	Relative fold change	P-value	= ↑↓	Relative fold cahnge	P-value	= ↑↓	Replicate Name	Relative fold change	P-value	= ↑↓
C_1	1.95			0.32			C_1	1.28		
C_2	0.48			0.35			C_2	1.66		
C_3	1.00			1.00			**C_3**	1.00		
I-4_1	1.54			0.52			I-4_1	1.17		
I-4_2	2.16	0.20	=	0.45	0.61	=	I-4_2	1.21	0.62	=
I-4_3	1.86			1.27			I-4_3	1.25		
I-50_1	2.57			0.69			I-50_1	1.06		
I-50_2	1.76	0.04*	↑	0.49	0.97	=	I-50_2	1.10	0.55	=
I-50_3	3.33			0.52			I-50_3	1.36		
E-25_1	0.75			0.73			E-25_1	1.26		
E-25_2	0.96	0.68	=	0.51	0.98	=	E-25_2	1.29	0.90	=
E-25_3	2.73			0.41			E-25_3	1.47		
E-600_1	0.52			0.82			E-600_1	1.53		
E-600_2	0.57	0.24	=	0.35	0.80	=	E-600_2	1.17	0.69	=
E-600_3	0.54			0.27			E-600_3	1.53		

$C_{1,2,3}$ = three replicates calibrator cDNA; I-$4_{1,2,3}$ = three replicates cDNA from bovine insulin (4 µg/ml) treated haemocytes; I-$50_{1,2,3,}$ = three replicates cDNA from bovine insulin (50 µg/ml) treated haemocytes; E-$25_{1,2,3}$ = three replicates cDNA from EGF (25 ng/ml) treated haemocytes; E-$600_{1,2,3}$ = three replicates cDNA from EGF (600 ng/ml) treated haemocytes.

Results generated by the RG Mode of REST, which is a method to analyse relative gene expression, with efficiency correction, are presented in bar charts (Figures 4.8, 4.9) and tables (4.8, 4.9). REST enables the user to take into consideration multiple reference genes when determining expression. Therefore, reference genes 12621 and 8629 were both used in calculating relative expression. Normalisation to multiple reference genes allows for alternative approximations of the true expression values (Pfaffl and Corbett Research, 2008).

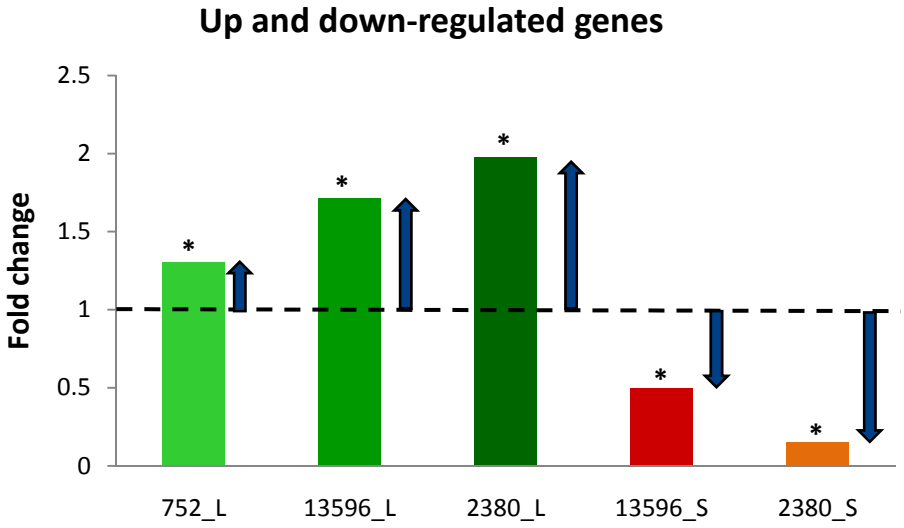


Figure 4.8 Relative expression of genes 752, 2380 and 13596. Real-time PCR was performed on template cDNA from whole animals; L = large, S = small. The dotted line indicates expression of the control group (equal amount mixture of L and S), normalized to both reference genes 12621 and 8629. Arrows indicate up- or down-regulation, $P < 0.05$ (*)

A fold change of one (dotted line in Figures 4.8 and 4.9) signifies the expression of the concerned gene under control conditions. Control conditions are i) a combination of equal amounts of L and S, for whole animal soft body tissue cDNA and ii) haemocytes that received no growth factor treatment, for haemocyte cDNA. Only genes that were significantly up-regulated (fold change > 1; $P < 0.05$) and significantly down-regulated (fold change < 1; $P < 0.05$) are shown in Figure 4.8. Significantly up-regulated genes in large abalone soft body tissues include 752 (FGF /insulin receptor), 13596 (Perlustrin) and 2380 (Thrombospondin). Significantly down-regulated genes in small abalone soft body tissues include 13596 (Perlustrin) and 2380 (Thrombospondin) (Figure 4.8, Table 4.8).

Table 4.8 Relative expression values of genes 752, 2380 and 13596 in large and small abalone, with associated P-values and 95 % confidence intervals (CI)

Gene	Fold change	P-value	CI (95%)
752_L	1.306	0.029	1.087 - 1.555
13596_L	1.712	0.007	1.248 - 2.184
2380_L	1.975	0	1.613 - 2.481
13596_S	0.497	0.01	0.393 - 0.683
2380_S	0.147	0.031	0.113 - 0.194

In addition to significantly up- and down-regulated genes, Figure 4.9 includes one gene (809_I-4; Mnk/MAP kinase) that was up-regulated in response to 4 µg/ml insulin treatment, although this up-regulation was not significant (P = 0.101).

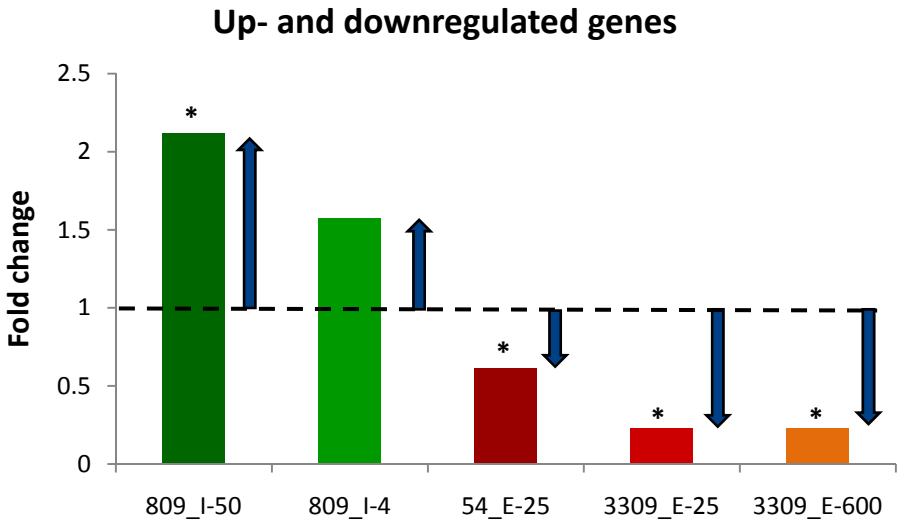


Figure 4.9 Relative expression of genes 809, 54 and 3309. Real-time PCR was performed on template cDNA from cultured haemocytes treated with 50 µg/ml (I-50) and 4 µg/ml (I-4) bovine insulin and 25 ng/ml (E-25) and 600 ng/ml (E-600) epidermal growth factor. The dotted line indicates expression of the control group (untreated cells), normalized to both reference genes 12621 and 8629. Arrows indicate up- or down-regulation, P < 0.05 (*)

The relative expression of genes 809 (Mnk/MAP-kinase), 54 (Collagen) and 3309 (Cyclin) in cultured haemocytes is shown Figure 4.9. The Mnk/MAP-kinase encoding gene (809) appears significantly up-regulated in haemocytes treated with 50 µg/ml bovine insulin while other genes (54, 7048 and 3309) were not influenced significantly by insulin treatment (data not shown). The collagen (54) and cyclin (3309) encoding genes appears to be significantly down-regulated in response to EGF treatment. Other genes (809 and 7048) were not influenced significantly by EGF treatment (data not shown).

Table 4.9 Relative expression values of genes 809, 54 and 3309 in haemocytes in response to growth factor treatments, with associated P-values and 95 % confidence intervals (CI)

Gene	Fold change	P-value	CI (95%)
809_I-5	2.118	0.000	1.203 - 4.041
809_I-4	1.576	0.101	0.984 - 2.643
54_E-2	0.614	0.038	0.412 - 0.901
3309_E-2	0.228	0.012	0.109 - 0.621
3309_E-6	0.226	0.021	0.139 - 0.343

4.4 Discussion

Whole animal tissue and short-term primary haemocyte cell cultures from *Haliotis midae* were used successfully in this study as material to investigate differential gene expression by qPCR. Seven genes of interest, selected on the basis of their differential expression following transcriptome sequencing and digital differential expression analysis, were studied.

Procedures were conducted according to the MIQE guidelines (Bustin *et al*., 2009). Samples of template cDNA free from contaminating genomic DNA were prepared from the tissues of origin and subsequent melt curves and agarose gels confirmed primer specificity and amplification of the desired product. Reference genes were verified as having stable expression under variable test conditions and therefore were suitable for normalising gene expression data in both sample types. Standard curves were used to verify amplification efficiency for all primer pairs. Amplification reactions for all samples were performed in triplicate and no-template controls included in all runs confirmed the absence of contaminating DNA and non-specific products.

Initially three target genes (752, 13596 and 2380) and two reference genes (8629 and 12621) were investigated, using whole animal tissue cDNA as template. The resulting fold-changes in gene expression (from REST analysis) corroborated the up-regulated expression of the target genes in large animals, observed during digital differential expression analysis. These five genes were subsequently targeted for amplification in haemocyte-derived template cDNA.

Sequence 752 resembles an insulin-related peptide receptor expressed in the mantle edge of the Pacific oyster, *Crassostrea gigas*. Insulin-like biological effects involve a variety of mollusc cell types, including those from the nervous ganglia, mantle, digestive gland, gonad and haemolymph (Gricourt *et al*., 2003, 2006). Amplification of sequence 752 demonstrates expression of an insulin-related peptide receptor in *Haliotis midae*. However, although insulin-related peptides were shown to stimulate protein synthesis in haemocytes of *H. tuberculata* (Lebel *et al*., 1996; Serpentini *et al*., 2000), no gene expression could be quantified in cultured haemocytes of *H. midae* in this study. The up-regulated expression of 752 in large animal whole soft body tissue can probably be ascribed to its expression in other tissues, such as the nervous ganglia, mantle or digestive gland.

Sequence 13596 resembles perlustrin from *Haliotis laevigata* and *H. discus discus*. Perlustrin, a nacre protein isolated from abalone shell, with homology to the insulin-like growth factor binding protein in vertebrates, suggests the existence of growth factor-like substances in abalone nacre (Weiss *et al*., 2001). The process of shell formation/biomineralisation, where perlustrin is mainly at play, involves the absorption of calcium and bicarbonate ions from the body surface, inner mantle epithelium, gills and gut and transportation thereof via the haemolymph to the epithelial cells (Marin and Luquet, 2004).

The close association of haemolymph and mantle tissue during biomineralisation makes investigation of the prevalence of nacre proteins in haemocytes worthwhile. Mount *et al.* (2004) indeed reported that granulocytic haemocytes of the oyster (*Crassostrea virginica*) can initiate mineral growth and may be directly involved in shell crystal production. Although amplification of 13596 in whole soft body tissue confirmed expression of perlustrin, no gene expression could be quantified for 13596 in cultured haemocytes of *Haliotis midae* in this study. The up-regulation of 13596 in large animal whole soft body tissue can thus probably be ascribed to its expression in mantle tissue. Although Weiss *et al*. (2001) isolated perlustrin directly from the nacre of abalone shell and did not investigate the involvement of any tissue, deposition of other shell matrix molecules are known to primarily involve the outer mantle epithelium (Duvail *et al*., 1997; Zhang *et al*., 2003; Jolly *et al*., 2004; Gong *et al*., 2008).

Sequence 2380 resembles Thrombospondin-1 precursor (TSP1). TSP1 is an extracellular matrix adhesive glycoprotein involved in cell adhesion, spreading, migration, and proliferation (Lawler *et al*., 1993; Urry *et al*., 1998). TSP1 influences cell function by modulating cell-matrix interactions through cell surface receptors, cytokines, growth factors and proteases. In vertebrates, TSP1 is expressed primarily during development, during growth, in response to injury and in tissues with continued turnover (Bornstein, 2001). Following a study by Lucas (2007) thrombospondin-1 precursor was up-regulated in anterior soft body tissues of fast-growing abalone (*H. asinina*), when compared to slow-growing and diseased abalone. If a similar function in abalone to that in vertebrates is assumed, it is proposed that thrombospondin can affect growth by causing up-regulation of transforming growth factor β, which interacts with insulin-like proteins (Lucas, 2007). Being an extracellular matrix protein, TSP1 is expected to be prevalent in all tissues. Although its expression in whole soft body tissues was confirmed, expression of 2380 could not be quantified in cultured haemocytes of *H. midae* in this study. Significant up-regulation in large animal whole soft body tissue, however, confirmed the observation of Lucas (2007) and suggests that TSP1 may indeed contribute to increased growth in abalone.

Sequences 8629 and 12621 were confirmed as reliable reference genes in this study, following reference gene validation with geNorm (Vandesompele *et al*., 2002). Reference gene 8629 resembles ribosomal protein S9 from other molluscs. Genes transcribing ribosomal proteins have been used effectively in previous qPCR studies as reference genes. Examples in molluscs include S-18 ribosomal RNA in clam (*Mya arenaria*) haemocytes (Araya *et al*., 2008), 28S ribosomal RNA in oyster (*Crassostrea gigas*) gill and digestive gland tissue (Choi *et al*., 2008) and ribosomal protein S9 gene in abalone (*Haliotis discus discus*) gill, mantle, gonad, foot and digestive tract tissues (Wan *et al*., 2008). The *H. discus discus* ribosomal protein S9 reference gene was reported to have more stable expression in abalone tissues than β-actin, a reference gene frequently used by other researchers (Wan *et al*., 2008). This same ribosomal protein S9 mRNA (Genbank accession EU247757.1) is nearly identical to 8629 from

H. midae (megablast, $3E^{-62}$). The successful amplification of 8629 in both soft body tissues and cultured haemocytes of *Haliotis midae* confirms expression of ribosomal protein S9 in various abalone tissues.

Reference gene 12621 resembles ornithine decarboxylase (ODC) from other molluscs. ODC has been used in previous qPCR studies as a reference gene. After Draper *et al*. (2001) suggested that ODC is expressed at near constant levels during early development of zebrafish (*Danio rerio*), it was used as a reference gene by Nogare *et al*. (2007) and White *et al*. (2009) during qPCR experiments on zebrafish (*Danio rerio*) embryos. ODC has also been used as reference gene during qPCR on various adult tissues and embryos of *Xenopus laevis* (Nanba *et al*., 2010). Šindelka *et al*. (2006) did, however point out that ODC mRNA is a suitable reference gene only at certain stages of development of the *Xenopus* embryo, while at other stages up to tenfold increase and decrease in mRNA levels was measured. No report could be found for the use of ODC as reference gene in molluscs. Amplification of 12621 confirms the expression of ODC in whole soft body tissues and cultured haemocytes of *Haliotis midae*.

Although the two reference genes (8629 and 12621) were confirmed to have stable expression across all sample types (from whole animal tissue: large and small groups; from haemocyte cultures: across all treatments and control), the initial three target genes did not show any detectable expression in cultured haemocyte cDNA. It is therefore inferred that these three genes are expressed elsewhere in abalone tissues. Four new sequences, carefully selected for their up-regulation reported by digital differential expression analysis (Chapter 2) and for their known expression in mollusc haemocytes, were subsequently targeted for their expression in cultured *H. midae* haemocytes.

Sequence 7048 resembles HSP90, a group of highly conserved stress proteins expressed in all eukaryotic cells (Pratt, 1997). HSP90 is a molecular chaperone that, as part of the HSP90-based chaperone system, facilitates the folding of a variety of proteins. This system is utilized by various signal transduction systems, one of which is signal transduction via activation of the MAP (mitogen-activated protein) family of protein kinases. As HSP90 is critical for normal signal transduction and forms stable complexes with a variety of protein kinases, it is closely involved in the binding of a variety of polypeptide ligands (insulin, epidermal growth factor, platelet-derived growth factor and nerve growth factor) through the activation of MAP kinases (Pratt, 1997). HSP90 has been identified in the haemocytes of the Zhikong scallop, *Chlamys farreri*, (Gao *et al*., 2007); the bay scallop, *Argopecten irradians*, (Gao *et al*., 2008); the soft-shell clam *Mya arenaria* (Mateo *et al*., 2010) and the abalones *Haliotis tuberculata* (Farcy *et al*., 2007) and *H. discus* (Wang *et al*., 2010) where its expression was up-regulated by heat shock, bacterial challenge and heavy metal stress. In unstressed conditions, HSP90 of *H. discus* is constitutively expressed in all the tissues, including haemolymph. Following heat shock treatment, transcription of the HSP90 gene was significantly up-regulated in the gills (Wang *et al*.,

2010). Amplification of 7048 confirms the expression of HSP90 in cultured *H. midae* haemocytes. The fact that no significant up- or down-regulation was observed across all treatments suggests that 7048 is constitutively expressed in haemocytes, regardless of the growth factor treatments. This can be supported by results from the reference gene validation with geNorm, where 7048 ranked third (of six) in terms of expression stability. The reason why 7048 was reported as up-regulated (1.2 x) in large animals by differential expression analysis (Chapter 2), may thus be ascribed to up-regulation of the HSP90 gene in another tissue, possibly as a result of bacterial challenge, heavy metal stress or temperature fluctuations experienced by the sample animals.

Sequence 54 resembles alpha 1 type IV collagen (in *Haliotis tuberculata*) and collagen pro alpha-chain (in *H. discus*). Collagen is one of the components of the extracellular matrix that participate in molecular events that regulate cell adhesion, migration, proliferation and regeneration (Poncet *et al.*, 2000). Different types of collagen have been identified in abalone. Fibril-forming collagen, Hdcols 1α and 2α, were identified in *H. discus* mantle and muscle (Yoneda *et al*., 1999) and non-fibrillar type IV collagen was identified in *H. tuberculata* haemocytes (Genbank accession: FN566840.1). Serpentini *et al.* (2000) also reported that *H. tuberculata* haemocytes synthesize type I and possibly type IV collagen *in vitro*. The 6.4-kb transcript they identified was proposed as consistent with that of the basement membrane collagen (type IV collagen) previously detected in *Drosophila* (Monson *et al*., 1982), sea urchin (*Strongylocentrotus purpuratus*) (Exposito *et al*., 1993) and sponge (*Pseudocorticium jarrei*) (Boute *et al*., 1996). The successfull amplification of sequence 54 demonstrates the expression of a related collagen gene in cultured *Haliotis midae* haemocytes. As it is a 6.5-kb transcript it may correspond well to non-fibrillar type IV collagen (6.4-kb). This confirms the idea of Serpentini *et al*. (2000) that a basement membrane type IV collagen is present in abalone haemocytes, which is involved in basement membrane deposition.

It has been shown that insulin and IGF-I stimulate protein (Sevala *et al*., 1993; Lebel *et al*., 1996; Giard *et al*., 1998) and specifically collagen (Serpentini *et al*., 2000) synthesis in mollusc haemocytes. The reason why insulin did not cause up-regulation of 54 in the present study may be because 54 is a non-fibrillar type IV collagen and not a type I collagen, which was the major fraction of collagen synthesized in response to IGF-1 treatment in the study of Serpentini *et al.* (2000). Although insulin has been shown to promote growth in various molluscan cell types (Domart-Coulon *et al*., 1994; Lebel *et al*., 1996; Giard *et al*., 1998; Chen and Wen, 1999), Weiss *et al.* (2001) pointed out that insulin bound with a twofold lower affinity to Perlustrin, an abalone nacre protein, than IGF-I did. If similar binding affinity can be assumed for type IV collagen, this may alternatively explain why the same effect that Serpentini *et al.* (2000) observed was not seen for 54 in *H. midae*. The reason why 54 was reported as up-regulated (1.5

x) in large animals by differential expression analysis (Chapter 2), may also be ascribed to up-regulation of the collagen gene in another tissue.

Sequence 54 was significantly down-regulated in response to EGF treatment (25 ng/ml). Similar down-regulation of collagen expression in mammalian astrocytes was reported following EGF-receptor activation (Heck *et al*., 2007). EGF was also reported to down-regulate *de novo* collagen synthesis and to promote degradation of type I collagen in *in vitro* cultured human fibroblast cells. These authors suggest that the MEK/ERK signaling pathway is involved in the EGF-mediated down-regulation of collagen synthesis (Mimura *et al*., 2006).

Sequence 3309 resembles cyclin. Cyclins are regulatory subunits in the progression of the mitotic cell cycle and also play a role in meiosis. Elevated levels of cyclin B expression are associated with actively proliferating cells (Scheurlen *et al*., 1996; Lamers *et al*., 1999) and cyclins A and B have been found to be expressed in mature gonads, gill, mantle, muscle and eggs of the zebra mussel, *Dreissena polymorpha* (Lamers *et al*., 1999). Proliferation of haemocytes of the prosobranch mollusc *Littorina littorea*, is also associated with increased cyclin D expression (Gorbushin and Iakovleva, 2008). Furthermore, BLASTn revealed identity between 3309 and HasCL296Contig1, a contig sequenced from the nacre-secreting mantle tissue of the tropical abalone *Haliotis asinina* (Jackson *et al*., 2010).

Amplification of 3309 evidences the expression of a cyclin in cultured *H. midae* haemocytes. Furthermore 3309 expression was significantly down-regulated in EGF-treated haemocytes. Previous studies on mammalian tissues have reported the stimulatory effect of EGF on cyclin D1 expression (Ravitz *et al*., 1996; Perry *et al*., 1998). Although the MAPK pathway, which has elevated activity in response to EGF, functions as a positive regulator of growth, it can also cause inhibition of cyclin dependent kinase activity and lead to cell cycle arrest. Pumiglia and Decker (1997) propose that transient or cyclical activation of the MAPK pathway may contribute to cell cycle progression, while sustained high activity levels may lead to cell cycle arrest.

It has been reported that human cyclins A, D, E and possibly C are synthesized during the G1 phase of the cell cycle and that cyclins D and E are rate-limiting for the G1 to S transition (Ohtsubo and Roberts, 1993; Polyak *et al*., 1994). During a normal mammalian cell cycle, a three to five-fold fluctuation in cyclin E-dependent kinase expression, which also reflects the level of cyclin E expression, occurs (Ohtsubo and Roberts, 1993). Furthermore, when apoptosis is induced in mammalian cells, mitotic cyclins A2 and B1 are down-regulated, but cylin D is not (Gentile *et al.*, 2003). Thus, depending on which cyclin is encoded by the gene sequence 3309, its expression will vary considerably at different stages of the cell cycle. Future studies where the cell cycle is carefully monitored and where 3309 expression is

quantified at different cell cycle stages, as well as in response to different treatments, is needed to elucidate the type of cyclin involved and its expression patterns in abalone haemocytes.

Sequence 809 resembles the MAP kinase-interacting serine/threonine kinase Mnk from *Aplysia californica* (Genbank accession: ABD47746.1). The MAP-Kinase pathway is highly conserved and involved in cell proliferation, -differentiation, -migration and -death. MAP kinase pathways involve MAP kinase kinase kinase (MKKK), a cascade of at least three protein kinase activities culminating in the phosphorylation and activation of a MAP kinase (The Gene Ontology, 2010). The typical MAP-Kinase pathway includes the following route of protein activation (Figure 4.10a) (the complete MAP-Kinase pathway can be found as Figure 4.10b in the Appendix):

EGF ⇨ EGFR ⇨ GRB2 ⇨ SOS ⇨ Ras ⇨ Raf1 ⇨ MEK1/MEK2 ⇨ ERK ⇨ **MNK1 / 2** ⇨ CREB ⇨ DNA ⇨ proliferation/differentiation.

Figure 4.10a One route in the MAP-kinase signaling pathway: EGF, epidermal growth factor; EGFR, epidermal growth factor receptor; GRB2, growth factor receptor-bound protein 2; SOS, son of sevenless homolog 1; Ras, related RAS viral (r-ras) oncogene homolog 2; Raf1, v-raf-1 murine leukemia viral oncogene homolog 1; MEK1/2, mitogen-activated protein kinase kinase 1/2; ERK, extracellular signal-regulated kinase 1/2 / mitogen-activated protein kinase 1; MNK1/2, MAP kinase interacting serine/threonine kinase 1/2; CREB, activating transcription factor 4

Extracellular-ligand-regulated kinase (ERK), which phosphorylates Mnk, can be activated by the binding of extracellular growth factors (like EGF, FGF and PDGF) to their cell membrane receptors, or in response to pro-inflammatory stimuli, like TNF, IL-1, TGF-B and FASL (Kanehisa Laboratories, 2010). Following activation by ERK, Mnk phosphorylates the transcriptional regulatory protein CREB, which leads to induction of immediate early gene transcription leading to cell proliferation or differentiation (Waskiewicz *et al*., 1997; Kanehisa Laboratories, 2010).

Mnk also acts as transcriptional and translational regulator of mRNA via phosphorylation of the elongation initiation factor (EIF4E), an important modulator of cell growth and proliferation (Pyronnet *et al*., 1999; Clark *et al*., 2010). Mnk plays this crucial role in the insulin signaling pathway, where it acts as a MAP kinase-activated kinase that regulates eukaryotic initiation factor-4E (eIF4E) phosphorylation. The typical insulin signaling pathway includes the following route of protein activation (Figure 4.11a) (the complete insulin signaling pathway can be found as Figure 4.11b in the Appendix):

INS ⇨ INSR ⇨ SHC ⇨ GRB2 ⇨ SOS ⇨ Ras ⇨ Raf ⇨ MEK1/2 ⇨ ERK1/2 ⇨ **MNK** ⇨ eIF4E ⇨ protein synthesis

Figure 4.11a One route in the insulin signaling pathway: INS, insulin; INSR, insulin receptor; SHC, src homology 2 domain-containing transforming protein C; GRB2, growth factor receptor-bound protein 2; SOS, son of sevenless homolog 1; Ras, Harvey rat sarcoma viral oncogene homolog; Raf, murine sarcoma 3611 viral oncogene homolog; MEK, mitogen-activated protein kinase kinase 1; ERK, extracellular signal-regulated kinase 1/2 /mitogen-activated protein kinase 1, Mnk, MAP kinase interacting serine/threonine kinase 2; eIF4E, eukaryotic translation initiation factor 4E

Two forms of Mnk exist: Mnk1 and Mnk2. Both are phosphorylated by extracellular-ligand-regulated kinases (Erk1 and Erk2). This phosphorylation stimulates the Mnks' kinase activity towards their substrate eIF-4E, a regulatory phosphoprotein that plays a vital role in the mechanism and translation of mRNA (Flynn and Proud, 1996; Waskiewicz *et al*., 1997). Phosphorylation of eIF4E is increased following insulin treatment, in an ERK-dependent manner (Flynn and Proud, 1996; Waskiewicz *et al*., 1997).

Mnk has been associated with the regulation of cap-dependent translation in *Aplysia* neurons (Ross *et al*., 2006) and is suggested to be involved in the continual growth of clam, *Laternula elliptica*, mantle tissue (Clark *et al*., 2010). Mnks have been identified in various vertebrate organisms as well as the following invertebrates, apart from *Aplysia* and *L. elliptica*: the purple sea urchin, *Strongylocentrotus purpuratus* (Oulhen *et al*., 2007), the nematodes *Caenorhabditis elegans* (The *C. elegans* sequencing consortium, 1998) and *C. briggsae* (Stein *et al*., 2003) and the fruit fly, *Drosophila melanogaster* (Brodsky *et al*., 2004). Mnk's role in increasing the efficiency of translation under limiting conditions has been suggested in *Drosophila* (Reiling *et al*., 2005). Its role in increasing translation rates in rat vascular smooth muscle cells and feline cardiomyocytes was also reported (Ishida *et al*., 2003; Tuxworth *et al*., 2004).

There is evidence to suggest that protein kinases such as MAPK are involved in the regulation of molluscan haemocyte cell spreading (Humphries and Yoshino, 2003). Lebel *et al.* (1996) showed that EGF and insulin stimulate DNA and protein synthesis in *Haliotis tuberculata* haemocytes and Humphries and Yoshino (2003) cloned insulin-like receptor cDNA from *Biomphalaria glabrata* embryonic cells. These reports support the notion of the presence of growth factor/hormone receptors and ligand-mediated intracellular signaling in molluscan haemocytes.

Amplification of 809 demonstrates the expression of Mnk, an essential component of the MAP kinase and insulin signaling pathways, in cultured *Haliotis midae* haemocytes. Furthermore, treatment of haemocytes with 50 µg/ml bovine insulin caused significant up-regulation (2 X, $P < 0.05$) of 809 gene expression. Treatment with 4 µg/ml bovine insulin also caused up-regulation (1.5 X, $P = 0.113$) of 809 expression, although not significant. Gene expression in haemocytes treated with EGF did not differ significantly from that in control cells. Activation of ERK, which phosphorylates Mnk, probably occurred more readily following insulin treatment than EGF treatment. Although both growth factors have the ability to stimulate cell proliferation through MAP kinase and insulin signaling pathways, a lesser effect of EGF was demonstrated at a physiological level during XTT assays in haemocyte cell culture trials (Chapter 3). Although viability of cells treated with 25 ng/ml EGF was not less than viability of control cells, EGF did not have a stimulating effect on cell proliferation as was shown for insulin. The ability of

insulin to stimulate cell proliferation is confirmed by the qPCR results. Admittedly, even though the effect of EGF on cell proliferation was not tested over a range of concentrations, as was insulin's, it is still unexpected to see no up-regulation in Mnk expression following EGF treatment of cells. This may be a result of sub-optimal EGF concentration used or it may be ascribed to variation hiding subtle gene expression differences. Variation is introduced during the course of a qPCR experiment at different stages, including cell harvesting, RNA extraction, reverse transcription and PCR (Adams, 2006).

The combined use of cell culture and qPCR in *Haliotis midae* is a powerful method of investigating gene expression. It allows the investigator to relate phenotypic responses to controlled treatments at the cellular level to the transient expression of genes. Meticulous selection of target and reference genes is the crucial first step towards finding answers related to gene expression patterns. In this chapter, three genes, differentially expressed in whole abalone soft body tissues and thought to be candidate genes related to growth, were shown to exhibit very low or no expression in cultured haemocytes. The expression of six carefully selected genes (two reference and four target) in cultured abalone haemocytes, was studied further. The four target genes, selected for their known expression in mollusc haemocytes and their up-regulation in fast growing abalone *in vivo,* were all confirmed to be expressed in *H. midae* haemocytes. Only one of these, a MAP-kinase/Mnk gene (809), showed significant up-regulation in response to insulin treatment. Although 54 (collagen) and 7084 (HSP90) showed up-regulation in soft body tissues of large abalone, this may be ascribed to expression in other tissues. A more comprehensive trial where different growth factor treatments are evaluated over a range of concentrations and cell cycle stages may elucidate expression of cyclin (3309), involved in regulation of the cell cycle.

Although the expression of most of the abovementioned genes has been already studied in molluscs (some of them in other abalone species), their expression in *H. midae* (*in vivo* and *in vitro*) is reported here for the first time. The system described in this chapter can be used successfully to investigate the expression of genes involved in the regulation of metabolic processes, disease resistance and cytotoxicity in *Haliotis midae* haemocytes. It provides a suitable method to evaluate factors causing a range of physiological responses in a controlled environment. Primary cell culture of other *H. midae* tissues (like mantle and gill) can also be used together with qPCR in future, to investigate tissue specific gene expression.

4.5 References

Adams, P.S. 2006. Data analysis and reporting. *In:* Real-time PCR. M. T. Dorak, editor. Abingdon, UK, Taylor and Francis Group. 39-62

Araya, M.T., A. Siah, D. Mateo, F. Markham, P. McKenna, G. Johnson, and F.C.J. Berthe. 2008. Selection and evaluation of housekeeping genes for haemocytes of soft-shell clams (*Mya arenaria*) challenged with *Vibrio splendidus*. *Journal of Invertebrate Pathology.* 99:326-331.

Bornstein, P. 2001. Thrombospondins as matricellular modulators of cell function. *The Journal of Clinical Investigation.* 107:929-934.

Boute, N., J. Exposito, N. Boury-Esnault, J. Vacelet, N. Noro, K. Miyazaki, K. Yoshizato, and R. Garrone. 1996. Type IV collagen in sponges, the missing link in basement membrane ubiquity. *Biology of the Cell.* 88:37-44.

Bower, N.I., and I.A. Johnston. 2009. Selection of reference genes for expression studies with fish myogenic cell cultures. *BMC Molecular Biology.* 10:80-91.

Brodsky, M.H., B.T. Weinert, G. Tsang, Y.S. Rong, N.M. McGinnis, K.G. Golic, D.C. Rio, and G.M. Rubin. 2004. *Drosophila melanogaster* MNK/Chk2 and p53 regulate multiple DNA repair and apoptotic pathways following DNA damage. *Molecular and Cellular Biology.* 24:1219-1231.

Bustin, S.A. 2000. Absolute quantification of mRNA using real-time reverse transcription polymerase chain reaction assays. *Journal of Molecular Endocrinology.* 25:169-193.

Bustin, S.A. 2010. Why the need for qPCR publication guidelines?—The case for MIQE. *Methods.* 50:217-226.

Bustin, S.A., V. Benes, J.A. Garson, J. Hellemans, J. Huggett, M. Kubista, R. Mueller, T. Nolan, M.W. Pfaffl, G.L. Shipley, J. Vandesompele, and C.T. Wittwer. 2009. The MIQE Guidelines: Minimum information for publication of quantitative real-time PCR experiments. *Clinical Chemistry.* 55:611-622.

Carninci, P., M. Nakamura, K. Sato, Y. Hayashizaki, and M.J. Brownstein. 2002. Cytoplasmic RNA extraction from fresh and frozen mammalian tissues. *BioTechniques.* 33:306-309.

Chen, S.N., and C.M. Wen. 1999. Establishment of cell lines derived from oyster, *Crassostrea gigas* Thunberg and hard clam, *Meretrix lusoria* Röding. *Methods in Cell Science.* 21:183-192.

Choi, Y.K., P.G. Jo, and C.Y. Choi. 2008. Cadmium affects the expression of heat shock protein 90 and metallothionein mRNA in the Pacific oyster, *Crassostrea gigas*. *Comparative Biochemistry and Physiology, Part C.* 147:286-292.

Clark, M.S., M.A.S. Thorne, F.A. Vieira, J.C.R. Cardoso, D.M. Power, and L.S. Peck. 2010. Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyroseqeuencing. *BMC Genomics.* 11:362.

Combes, J.D., G. Grelier, M. Laversanne, N. Voirin, S. Chabaud, R. Ecochard, C. Lasset, and C. Moyret-Lalle. 2009. Contribution of cell culture, RNA extraction, and reverse transcription to the

measurement error in quantitative reverse transcription polymerase chain reaction-based gene expression quantification. *Analytical Biochemistry.* 393:29-35.

Corbett Research. 2006. Rotor-Gene 6000 Operator manual 1.7.87, Corbett Research, Australia.

Derveaux, S., J. Vandesompele, and J. Hellemans. 2010. How to do successful gene expression analysis using real-time PCR. *Methods.* 50:227-230.

DNAbiotec (Pty) Ltd. 2008. Essential real-time PCR short course. 1st edn, Pretoria, South Africa.

Domart-Coulon, I., D. Doumenc, S. Auzoux-Bordenave, and Y. Le Fichant. 1994. Identification of media supplements that improve the viability of primary cell cultures of *Crassostrea gigas* oysters. *Cytotechnology.* 16:109-120.

Domart-Coulon, I., S. Auzoux-Bordenave, D. Doumenc, and M. Khalanski. 2000. Cytotoxicity assessment of antibiofouling compounds and by-products in marine bivalve cell cultures. *Toxicology in Vitro.* 14:245-251.

Draper, B.W., P.A. Morcos, and C.B. Kimmel. 2001. Inhibition of Zebrafish fgf8 Pre-mRNA splicing with morpholino oligos: A quantifiable method for gene knockdown. *Genesis.* 30:154-156.

Duvail, L., E. Lopez, and M. Fouchereau-Peron. 1997. Characterization of a calcitonin gene related peptide-like molecule in the abalone, *Haliotis tuberculata*. *Comparative Biochemistry and Physiology, Part C*:155-159.

Exposito, J., M. D'Alessio, M. Di Liberto, and F. Ramirez. 1993. Complete primary structure of a sea urchin type IV collagen α chain and analysis of the 5' end of its gene. *The Journal of Biological Chemistry.* 268:5249-5254.

Farcy, E., A. Serpentini, B. Fiévet, and J. Lebel. 2007. Identification of cDNAs encoding HSP70 and HSP90 in the abalone *Haliotis tuberculata*: Transcriptional induction in response to thermal stress in hemocyte primary culture. *Comparative Biochemistry and Physiology, Part B.* 146:540-550.

Flynn, A., and C.G. Proud. 1996. Insulin-stimulated phosphorylation of initiation factor 4E is mediated by the MAP kinase pathway. *FEBS Letters.* 389:162-166.

Fu, J., A.M. Weise, J.L. Falany, C.N. Falany, B.J. Thibodeau, F.R. Miller, T.A. Kocarek, and M. Runge-Morris. 2010. Expression of estrogenicity genes in a lineage cell culture model of human breast cancer progression. *Breast Cancer Research and Treatment.* 120:35-45.

Gao, Q., L. Song, D. Ni, L. Wu, H. Zhang, and Y. Chang. 2007. cDNA cloning and mRNA expression of heat shock protein 90 gene in the haemocytes of Zhikong scallop *Chlamys farreri*. *Comparative Biochemistry and Physiology, Part B.* 147:704-715.

Gao, Q., J. Zhao, L. Song, L. Qiu, Y. Yu, H. Zhang, and D. Ni. 2008. Molecular cloning, characterization and expression of heat shock protein 90 gene in the haemocytes of bay scallop *Argopecten irradians*. *Fish and Shellfish Immunology.* 24:379-385.

Gentile, M., L. Latonen, and M. Laiho. 2003. Cell cycle arrest and apoptosis provoked by UV radiation-induced DNA damage are transcritionally highly divergent responses. *Nucleic Acids Research.* 31:4779-4790.

Ghaemmaghami, S., J. Ullman, M. Ahn, S. St. Martin, and S.B. Prusiner. 2010. Chemical induction of misfolded prion protein conformers in cell culture. *The Journal of Biological Chemistry.* 285:10415-10423.

Giard, W., J. Lebel, E. Boucaud-Camou, and P. Favrel. 1998. Effects of vertebrate growth factors on digestive gland cells from the mollusc *Pecten maximus* L.: an *in vitro* study. *Journal of Comparative Physiology Part B.* 168:81-86.

Ginzinger, D.G. 2002. Gene quantification using real-time quantitative PCR: An emerging technology hits the mainstream. *Experimental Haematology.* 30:503-512.

Gong, N., Z. Ma, Q. Li, Q. Li, Z. Yan, L. Xie, and R. Zhang. 2008. Characterization of calcium deposition and shell matrix protein secretion in primary mantle tissue culture from the marine pearl oyster *Pinctada fucata*. *Marine Biotechnology.* 10:457-465.

Gorbushin, A.M., and N.V. Iakovleva. 2008. The enigma of the haemogram ''left-shift'' in periwinkles infected with trematodes. *Fish and Shellfish Immunology.* 24:745-751.

Gricourt, L., G. Bonnec, D. Boujard, M. Mathieu, and K. Kellner. 2003. Insulin-like system and growth regulation in the Pacific oyster *Crassostrea gigas*: hrIGF-1 effect on protein synthesis of mantle edge cells and expression of an homologous insulin receptor-related receptor. *General and Comparative Endocrinology.* 134:44-56.

Gricourt, L., M. Mathieu, and K. Kellner. 2006. An insulin-like system involved in the control of Pacific oyster *Crassostrea gigas* reproduction: hrIGF-1 effect on germinal cell proliferation and maturation associated with expression of an homologous insulin receptor-related receptor. *Aquaculture.* 251:85-98.

Heck, N., J. Garwood, A. Dobbertin, V. Calco, S. Sirko, T. Mittmann, U.T. Eysel, and A. Faissner. 2007. Evidence for distinct leptomeningeal cell-dependent paracrine and EGF-linked autocrine regulatory pathways for suppression of fibrillar collagens in astrocytes. *Molecular and Cellular Neuroscience.* 36:71-85.

Humphries, J.E., and T.P. Yoshino. 2003. Cellular receptors and signal transduction in molluscan hemocytes: Connections with the innate immune system of vertebrates. *Integrative and Comparative Biology.* 43:305-312.

Ishida, M., T. Ishida, H. Nakashima, N. Miho, K. Miyagawa, K. Chayama, T. Oshima, M. Kambe, and M. Yoshizumi. 2003. Mnk1 is required for angiotensin II–induced protein synthesis in vascular smooth muscle cells. *Circulation Research.* 93:1218-1224.

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

Jheon, A., A.K. Bansal, B. Zhu, B. Ganss, S. Cheifetz, and J. Sodek. 2009. Characterisation of the constitutive over-expression of AJ18 in a novel rat stromal bone marrow cell line (D8-SBMC). *Archives of Oral Biology.* 54:705-716.

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

Kanehisa Laboratories. 2010. MAPK signaling pathway - *Homo sapiens* (human). 2010. [Homepage of Kanehisa Laboratories], [Online]. Available: http://www.genome.jp/kegg/pathway/hsa/hsa04010.html [Accessed 17 June 2010].

Kubista, M., J.M. Andrade, M. Bengtsson, A. Forootan, J. Jona, K. Lind, R. Šindelka, R. Sjöback, B. Sjögreen, Ströbom, L, Ståhlberg. A., and N. Zoric. 2006. The real-time polymerase chain reaction. *Molecular Aspects of Medicine.* 27:95-125.

Lamers, A.E., J.P. Heiney, and J.L. Ram. 1999. Cloning and sequence analysis of two cDNAs encoding cyclin A and cyclin B in the zebra mussel *Dreissena polymorpha*. *Biochimica et Biophysica Acta.* 1448:519-524.

Lawler, J., M. Duquette, C.A. Whittaker, J.C. Adams, K. McHenry, and D.W. DeSimone. 1993. Identification and characterization of Thrombospondin-4, a new member of the thrombospondin gene family. *The Journal of Cell Biology.* 120:1059-1067.

Lebel, J., W. Giard, P. Favrel, and E. Boucaud-Camou. 1996. Effects of different vertebrate growth factors on primary cultures of hemocytes of the gastropod mollusc, *Haliotis tuberculata*. *Biology of the Cell.* 86:67-72.

Livak, K.J., and T.D. Schmittgen. 2001. Analysis of relative gene expression data using real-time quantitative PCR and the $2^{-\Delta\Delta CT}$ method. *Methods.* 25:402-408.

Lorkowski, S., and P. Cullen. 2003. Analysing gene expression Vol. 1, 1st ed. Weinheim, Germany, Wiley-VCH Verlag GmbH and Co. 407p.

Lucas, T. 2007. Investigating genetic and molecular aspects of growth of the tropical abalone *Haliotis asinina*. Unpublished PhD dissertation, University of Queensland, Queensland, Australia.

Mackay, J. 2007. Introduction to kinetic (real-time) PCR. *In:* Methods in Molecular Biology, vol. 353: Protocols for Nucleic Acid Analysis by Nonradioactive Probes. E. Hilario and J. Mackay, editors. Totowa, NJ, Humana Press Inc. 167-176.

Man, A.K.Y., and N.Y.S. Woo. 2008. Upregulation of metallothionein and glucose-6-phosphate dehydrogenase expression in silver sea bream, *Sparus sarba* exposed to sublethal levels of cadmium. *Aquatic Toxicology.* 89:214-221.

Marin, F., and G. Luquet. 2004. Molluscan shell proteins. *Comptes Rendus Palevol.* 3:469-492.

Mateo, D.R., S.J. Greenwood, M.T. Araya, F.C.J. Berthe, G.R. Johnson, and A. Siah. 2010. Differential gene expression of $-actin, Toll-like receptor 2 (TLR-2) and interleukin-1 receptor-associated kinase 4 (IRAK-4) in *Mya arenaria* haemocytes induced by *in vivo* infections with two *Vibrio splendidus* strains. *Developmental and Comparative Immunology.* 34:710-714.

Mimura, Y., H. Ihn, M. Jinnin, Y. Asano, K. Yamane, and K. Tamaki. 2006. Epidermal growth factor affects the synthesis and degradation of type I collagen in cultured human dermal fibroblasts. *Matrix Biology.* 25:202-212.

Monson, J.M., J. Natzle, J. Friedman, and B.J. McCarthy. 1982. Expression and novel structure of a collagen gene in *Drosophila*. *Proceedings of the National Academy of Sciences USA.* 79:1761-1765.

Mount, A.S., A.P. Wheeler, R.P. Paradkar, and D. Snider. 2004. Hemocyte-mediated shell mineralization in the Eastern oyster. *Science.* 304:297-300.

Nanba, R., N. Fujita, and S. Nagata. 2010. Structure and expression of myelin basic protein gene products in *Xenopus laevis*. *Gene.* 459:32-38.

Nogare, D.E.D., A. Arguello, S. Sazer, and M.E. Lane. 2007. Zebrafish cdc25a is expressed during early development and limiting for post-blastoderm cell cycle progression. *Developmental Dynamics.* 236:3427-3435.

Ohtsubo, M., and J.M. Roberts. 1993. Cyclin-dependent regulation of G1 in mammalian fibroblasts. *Science.* 259:1908-1912.

Oulhen, N., P. Salaün, B. Cosson, P. Cormier, and J. Morales. 2007. After fertilization of sea urchin eggs, eIF4G is post-translationally modified and associated with the cap-binding protein eIF4E. *Journal of Cell Science.* 120:425-434.

Perry, J.E., M.E. Grossmann, and D.J. Tindall. 1998. Epidermal growth factor induces cyclin D_1 in a human prostate cancer cell line. *The Prostate.* 35:117-124.

Pfaffl, M.W. 2006. Relative quantification. *In:* Real-time PCR. M. T. Dorak, editor. Abingdon, UK, Taylor and Francis Group. 63-82.

Pfaffl, M.W. 2010. The ongoing evolution of qPCR. *Methods.* 50:215-216.

Pfaffl, M.W. and Corbett Research. 2008. Relative expression software tool. Australia.

Pfaffl, M.W., G.W. Horgan, and L. Dempfle. 2002. Relative expression software tool (REST©) for group-wise comparison and statistical analysis of relative expression results in real-time PCR. *Nucleic Acids Research.* 30:e36-e46.

Polyak, K., J. Kato, M.J. Solomon, C.J. Sherr, J. Massague, J.M. Roberts, and A. Koff. 1994. p27 ^{Kip1}, a cyclin-Cdk inhibitor, links transforming growth factor-B and contact inhibition to cell cycle arrest. *Genes and Development.* 8:9-22.

Poncet, J., A. Serpentini, B. Thiebot, C. Villers, J. Bocquet, E. Boucaud-Camou, and J. Lebel. 2000. In vitro synthesis of proteoglycans and collagen in primary cultures of mantle cells from the nacreous mollusk, *Haliotis tuberculata*: A new model for study of molluscan extracellular matrix. *Marine Biotechnology.* 2:387-398.

Pratt, W.B. 1997. The role of the hsp90-based chaperone system in signal transduction by nuclear receptors and receptors signaling via MAP kinase. *Annual Review of Pharmacology and Toxicology.* 37:297-326.

Pumiglia, K.M., and S.J. Decker. 1997. Cell cycle arrest mediated by the MEK/mitogen-activated protein kinase pathway. *Proceedings of the National Academy of Sciences USA.* 94:448-452.

Pyronnet, S., H. Imataka, A.C. Gingras, R. Fukunaga, T. Hunter, and N. Sonenberg. 1999. Human eukaryotic translation initiation factor 4G (eIF4G) recruits Mnk1 to phosphorylate eIF4E. *The EMBO Journal.* 18:270-279.

Ravitz, M.J., S. Yan, C. Dolce, A.J. Kinniburch, and C.E. Wenner. 1996. Differential regulation of p27 and Cyclin D1 by TGF-B and EGF in C3H 10T1/2 mouse fibroblasts. *Journal of Cellular Physiology.* 168:510-520.

Reiling, J.H., K.T. Doepfner, E. Hafen, and H. Stocker. 2005. Diet-dependent effects of the *Drosophila* Mnk1/Mnk2 homolog Lk6 on growth via eIF4E. *Current Biology.* 15:24-30.

Rosa, J., D.M. Tiago, J. Dias, M.L. Cancela, and V. Laizé. 2010. Serum-specific stimulation of proliferation and mineralization of fish bone-derived cells. *Journal of Applied Ichthyology.* 26:251-256.

Ross, G., J.R. Dyer, V.F. Castellucci, and W.S. Sossin. 2006. Mnk is a negative regulator of cap-dependent translation in *Aplysia* neurons. *Journal of Neurochemistry.* 97:79-91.

Rychlik, W. 1992. OLIGO 4.1 Primer Analysis Software. Plymouth, MN, National Bioscience Inc.

Scheurlen, I., S.A.H. Hoffmeister, and H.C. Schaller. 1996. Presence and expression of G2 cyclins in the coelenterate hydra. *Journal of Cell Science.* 109:1063-1069.

Schmittgen, T.D. 2006. Quantitative gene expression by real-time PCR: a complete protocol. *In:* Real-time PCR. M. T. Dorak, editor. Abingdon, UK, Taylor and Francis Group: 127-137.

Schmittgen, T.D., and K.J. Livak. 2008. Analyzing real-time PCR data by the comparative CT method. *Nature Protocols.* 3:1101-1108.

Serpentini, A., C. Ghayor, J.M. Poncet, V. Hebert, P. Galera, J.P. Pujol, E. Boucaud-Camou, and J.M. Lebel. 2000. Collagen study and regulation of the de novo synthesis by IGF-I in hemocytes from the gastropod mollusc, *Haliotis tuberculata*. *Journal of Experimental Zoology.* 287:275-284.

Sevala, V.M., V.L. Sevala, and A.S.M. Saleuddin. 1993. Hemolymph insulin-like peptides (ILP) titers and the influence of ILP and mammalian insulin on the amino acid incorporation in the mantle collar *in vitro* in *Helisoma* (mollusca). *Biological Bulletin.* 185:140-148.

Shipley, G.L. 2006. An introduction to real-time PCR. *In:* Real-time PCR. M. T. Dorak, editor. Abingdon, UK, Taylor and Francis Group: 1-31.

Šindelka, R., Z. Ferjentsik, and J. Jonák. 2006. Developmental expression profiles of *Xenopus laevis* reference genes. *Developmental Dynamics.* 235:754-758.

Smith, C., D. Berg, S. Beaumont, N.T. Standley, D.N. Wells, and P.L. Pfeffer. 2007. Simultaneous gene quantitation of multiple genes in individual bovine nuclear transfer blastocysts. *Reproduction.* 133:231-242.

Ståhlberg, A., and M. Bengtsson. 2010. Single-cell gene expression profiling using reverse transcription quantitative real-time PCR. *Methods.* 50:282-288.

Stein, L.D., Z. Bao, D. Blasiar, T. Blumenthal, M.R. Brent, N. Chen, A. Chinwalla, L. Clarke, C. Clee, A. Coghlan, Coulson A, P. D'Eustachio, D.H.A. Fitch, L.A. Fulton, R.E. Fulton, S. Griffiths-Jones, T.W. Harris, L.W. Hillier, R. Kamath, P.E. Kuwabara, E.R. Mardis, M.A. Marra, T.L. Miner, P. Minx, J.C. Mullikin, R.W. Plumb, J. Rogers, J.E. Schein, M. Sohrmann, J. Spieth, J.E. Stajich , C. Wei, D. Willey, R.K. Wilson, R.M. Durbin, and R.H. Waterston. 2003. The genome sequence of *Caenorhabditis briggsae*: a platform for comparative genomics. *PLoS Biology.* 1:166-192.

Taylor, S., M. Wakem, G. Dijkman, M. Alsarraj, and M. Nguyen. 2010. A practical approach to RT-qPCR—Publishing data that conform to the MIQE guidelines. *Methods.* 50:S1–S5.

The *C. elegans* sequencing consortium. 1998. Genome sequence of the nematode *C. elegans*: a platform for investigating biology. *Science.* 282:2012-2018.

The Gene Ontology. 2010. MAPKKK cascade. 2010. [Homepage of The Gene Ontology], [Online]. Available: http://amigo.geneontology.org/cgi-bin/amigo/term-details.cgi?term=GO:0000165 [Accessed 17 June 2010].

Tuxworth, W.J., A.N. Saghir, L.S. Spruill, D.R. Menick, and P.J. McDermott. 2004. Regulation of protein synthesis by eIF4E phosphorylation in adult cardiocytes: the consequence of secondary structure in the 5'-untranslated region of mRNA. *Biochemical Journal.* 378:73-82.

Urry, L.A., C.A. Whittaker, M. Duquette, J. Lawler, and D.W. DeSimone. 1998. Thrombospondins in early *Xenopus* embryos: Dynamic patterns of expression suggest diverse roles in nervous system, notochord, and muscle development. *Developmental Dynamics.* 211:390-407.

Vandesompele, J., K. De Preter, F. Pattyn, B. Poppe, N. Van Roy, A. De Paepe, and F. Speleman. 2002. Accurate normalization of real-time quantitative RT-PCR data by geometric averaging of multiple internal control genes. *Genome Biology.* 3:1-12.

Wan, Q., I. Whang, and J. Lee. 2008. Molecular characterization of mu class glutathione-S-transferase from disk abalone (*Haliotis discus discus*), a potential biomarker of endocrine-disrupting chemicals. *Comparative Biochemistry and Physiology, Part B.* 150:187-199.

Wang, N., I. Whang, J.S. Lee, and J. Lee. 2010. Molecular characterization and expression analysis of a heat shock protein 90 gene from disk abalone (*Haliotis discus*). *Molecular Biology Reports.* in press, doi: 10.1007/s11033-010-9972-x.

Waskiewicz, A.J., A. Flynn, C.G. Proud, and J.A. Cooper. 1997. Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. *The EMBO Journal.* 16:1909-1920.

Weiss, I.M., W. Göhring, M. Fritz, and K. Mann. 2001. Perlustrin, a *Haliotis laevigata* (abalone) nacre protein, is homologous to the insulin-like growth factor binding protein N-terminal module of vertebrates. *Biochemical and Biophysical Research Communications.* 285:244-249.

White, Y.A.R., J.T. Kyle, and A.W. Wood. 2009. Targeted gene knockdown in Zebrafish reveals distinct intraembryonic functions for insulin-like growthfFactor II signaling. *Endocrinology.* 150:4366-4375.

Yoneda, C., Y. Hirayama, M. Nakaya, Y. Matsubara, S. Irie, K. Hatae, and S. Watabe. 1999. The occurrence of two types of collagen proa-chain in the abalone *Haliotis discus* muscle. *European Journal of Biochemistry.* 261:714-721.

Zhang, Y., L. Xie, Q. Meng, T. Jiang, R. Pu, L. Chen, and R. Zhang. 2003. A novel matrix protein participating in the nacre framework formation of pearl oyster, *Pinctada fucata*. *Comparative Biochemistry and Physiology, Part B.* 135:565-573.

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

5 CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions and recommendations

This research combined the use of molecular methodologies and cell culture as a time-efficient and economical way of studying abalone transcriptomics and cell biology, specifically to investigate growth variation in *Haliotis midae*. This is the first time next generation sequencing is used towards the large scale sequencing and annotation of any haliotid species and also the first time a comprehensive investigation is launched towards the establishment of primary cell cultures for *H. midae*. Genes involved in growth regulation and metabolism, previously unknown in *H. midae,* were identified and the expression of selected target genes involved in growth was quantified by qPCR.

The *H. midae* transcriptome data can be useful for future novel gene identification and can also supplement existing knowledge of marker detection, QTL identification and gene expression. The data can reveal more genes associated with economically important traits like growth, disease resistance and meat quality. By combining QTL and gene expression information, these traits can be harnessed in farmed abalone during selective breeding programs and in future, possibly through genetic engineering. The large amount of data generated by sequencing the *H. midae* transcriptome can also be useful in comparative phylogenetic studies towards understanding genetic variation within wild populations. The increased precision of such an approach was recently reported during a phylogeographic study of a North American butterfly genus *Lycaeides* (Gompert *et al*., 2010).

As accurate sequencing and assembly of short reads into contigs still pose challenges and sophisticated assembly algorithms are continuously being developed, different assemblies can be produced from the same raw sequence data in future. Improved assembly algorithms may result in a more comprehensive contig assembly, with a reduction in singletons. One such program, specifically developed for *de novo* transcriptome assembly, is Oases (Schulz and Zerbino, 2010). This assembler uses a preliminary assembly produced by Velvet, and clusters the contigs into groups. By exploiting the available paired-end sequence read and/or long sequence read information it constructs transcript isoforms. Its added ability to detect and report alternative splicing events makes Oases a superior *de novo* assembler. Oases is therefore the recommended program for further refinement of the transcriptome assembly produced by Velvet during this study.

Following annotation, many of the contigs matched to transcripts with known functions. From the reference transcriptome, 10 % (2987) of contigs could be annotated with high stringency ($E \leq 10^{-10}$) to an inferred biological function based on the Gene Ontology (GO). This percentage is similar to that found in other non-model organisms sequenced by next generation sequencing (~13 % for both cichlid

teleosts *Amphilophus astorquii* and *A. zaliosus*, E ≤ 10^{-05} – Elmer *et al*., 2010; 17 % for the Antarctic clam, *Laternula elliptica,* E ≤ 10^{-10} – Clark *et al*., 2010). A large amount of sequence data was thus annotated for the first time in *H. midae*. Despite this advance, annotation can still be refined and many sequences with lower BLAST scores that may also be of interest can be mined from the data.

Modern Illumina sequencing-by-synthesis technology and subsequent sequence annotation were used to elucidate differential gene expression between two sibling groups of abalone demonstrating significant growth variation. Over the three months before sample animals were sacrificed, differences in growth rate was evident between the two groups. Large animals showed a significant gain in length, width and weight ($P < 0.005$) over three months, while small animals did not increase significantly in any of the measured parameters during this time. Global assessment of the transcriptomes of large and small animals allowed for the discrimination of trends in differential gene expression. The wealth of EST information produced by next generation sequencing for *H. midae* was only explored to a limited extent in this study. For differential expression analysis, only the most differentially expressed ($P < 0.0005$) sequences were used. This accounts for a mere ~8 % of the total amount of contigs from the reference transcriptome. Sequences that was significantly differentially expressed but that did not show any homology to known sequences were not analaysed. These may provide a future source for novel gene identification in *H. midae*.

As a result of the constitution of the sample material (whole soft body tissue pooled together for all animals in a specific group), this data cannot be used to make inferences about organ- or tissue-specific gene expression. Also, because all animals were of the same age, this study is not suitable for making life-stage specific inferences regarding gene expression. It is assumed that growth variation was triggered by external factors because the animals are all siblings and therefore minimal inherent genetic variation is expected. Although this study is not able to predict a specific time-point or event that could cause growth variation, it may speculate on the nature of such a trigger, given the gene expression data obtained.

The pattern of up-regulated expression of genes concerned with xenobiotic biodegradation and metabolism observed in small animals leads to speculation regarding their relative size. It is possible that the small animals were small in relation to the “large” animals, which were in fact more representative of the normal size, because of a rerouting of energy towards xenobiotic degradation and metabolism, and away from growth. Various events can lead to an increased energy allocation towards xenobiotic degradation and metabolism.

Toxic accumulation in the water used on the abalone farm could be one causative agent. The farm where sample animals originated from conforms to a permit requirement of yearly testing for heavy

metals and pesticides, and no abnormalities were found during the time of this study. This route of xenobiotic contamination is therefore unlikely. However, the chemical analysis performed on water used by abalone farms only measure a fraction of contaminating substances in polluted water and it is possible that low levels of contamination may be overlooked (Slatinská *et al*., 2008). Xenobiotic/toxic substances in the water may originate from municipal, industrial and agricultural activities in the vicinity of the abalone farm, where sites of wastewater deposition into the sea are in close proximity to the inflowing water of the farm. In the present case, possible sources of xenobiotic substances in the inflowing water include sewage works, approximately 1.2 km away from the farm and the farm's own outlet for wastewater, approximately 150 m from the source of inflowing water. The sewage works are reportedly not depositing any wastewater into the ocean. The close proximity of the farm's own wastewater outlet to the inflowing water may however result in low levels of waste substances being pumped back to the farm.

One source of toxic pollution, present on all abalone farms as a product of protein degradation and microbial breakdown of organic nitrogen, is ammonia (Reddy-Lopata *et al*., 2006; Batley and Simpson, 2009). The growth rate of abalone is slower when kept in baskets close to the outflow of tanks, where a higher ammonia concentration is measured, when compared to baskets near the inflow. Concentrations of toxic, free ammonia is also higher in tanks when sludge, consisting of faeces, uneaten feed and organic and inorganic particulate matter entering the tank with the incoming seawater, accumulates at the bottom (Yearsley, 2007). From a previous study on *H. midae*, it is known that small abalone are more sensitive to ammonia concentration in the water than larger abalone. Although abalone can adapt to sub-lethal levels of ammonia, continued exposure causes a substantial reduction in growth (Reddy-Lopata *et al*., 2006).

Although retarded growth in small animals may be partially explained by their difficulty to cope with ammonia levels in the water it is also true that all animals, small and large, were kept in the same housing throughout their two years, and thus probably received similar ammonia exposure. Whether small animals resided closer to the bottom of the basket and therefore closer to the ammonia-forming sludge, which was cleaned out once a week, remains unknown as animal movement within the basket was not monitored during the two years. If small animals stayed close to the bottom of the basket during a crucial growth period, for instance following weaning and transfer to the baskets, their growth might have been handicapped from an early stage. In addition to continued ammonia exposure, there might also be other causes of retarded growth in small animals. Examples include fluctuations in water temperature and periods of immune challenge experienced during the occasional pathogen outbreaks that may occur on abalone farms.

The inability to satisfactorily answer these questions on the possible causes of retarded growth in small abalone is largely due to the many variables associated with the *in vivo* situation. Although farms adhere to strict operational protocols, variable water temperature and ammonia levels, pathogen outbreaks and trace amounts of chemicals and heavy metals are inevitable. In principal, using an *in vitro* system eliminates such environmental variables because a controlled environment can be maintained.

The feasibility of *in vitro* cell culture for *Haliotis midae* is demonstrated by the successful short term culture of larval and haemocyte cells in this study. The effect of exogenous growth factors on larval and haemocyte cell viability could be assessed and subsequent use of haemocytes in qPCR analyses attested to the applicability of such a system for investigating gene expression. In this way, *in vitro* cell culture allows one to relate phenotypic responses to controlled treatments, at the cellular level, to the transient expression of genes. Cell cultures were used here mainly to investigate the effect of growth stimulatory factors on cell proliferation.

Even though haemocyte cell cultures could be successfully maintained and haemocytes originate from the haemolymph, a ubiquitous tissue, this cell culture system was of limited use for quantifying the expression of the first three selected target genes using qPCR. These genes, showing significant differential expression between large and small animals could not be amplified in cultured haemocytes, leading to the conclusion that they are not expressed in haemocytes. It is therefore pertinent to have knowledge of where genes are expected to be expressed when aiming to quantify their expression. Other tissues may be more useful for studying expression of these specific genes. For example, mantle cell culture will probably be more suitable when studying expression of insulin-related peptide receptor (sequence 752) and perlustrin (sequence 13596), a nacre protein with homology to the insulin-like growth factor binding protein in vertebrates. Insulin-related peptide receptor closely resembles a similar sequence expressed in the mantle edge of the Pacific oyster, *Crassostrea gigas* (Gricourt *et al*., 2003, 2006) and perlustrin plays a role in biomineralisation, a process centered in the mantle tissue (Marin and Luquet, 2004). Primary mantle cell culture has been accomplished with success in other abalone species, including *H. tuberculata* (Auzoux-Bordenave *et al*., 2007) and *H. varia* (Suja *et al*., 2007). It is therefore recommended that the regulation of the concerned genes, in response to insulin and IGF treatment, be investigated in a mantle primary cell culture system of *H. midae* in future studies.

Throughout *in vivo* and *in vitro* qPCR experiments, the up-regulation of genes involved in the insulin signaling pathway (sequences 752 and 13596 and 809) provides evidence for the involvement of insulin in enhanced growth rate for various *H. midae* tissues. Insulin is an anabolic hormone which promotes protein synthesis in muscle and enhances incorporation of nutrients into cells (Machado *et al*., 1988; Haruta *et al*., 2000; Montserrat *et al*., 2007). It has been shown to promote growth acceleration in

mammals and fish via the mTOR (target of rapamycin) cell signaling pathway, which is involved in the regulation of initiation of mRNA translation (Bower and Johnston, 2010; Kaushik and Seiliez, 2010). The mTOR signaling pathway is intricately linked with both the MAP kinase and insulin signaling pathways (see Figure 5.1, Appendix) and insulin action is mediated through all three pathways (Haruta *et al*., 2000). Expression of genes specifically involved in the latter two pathways was quantified in this study.

In addition to insulin, IGFs also promote growth by stimulating the proliferation and differentiation of muscle cells (Bower, 2008). Cellular responses to IGF stimulation is mediated through the IGF receptor, a member of the tyrosine kinase family, and by the availability of the IGF, which is regulated by the IGFBPs. Binding of the receptor leads to the activation of multiple signal transduction cascades including the mTOR and MAP kinase pathways (Engert *et al*., 1996).

The growth stimulating properties associated with insulin and IGFs have lead to various attempts and successes at harnessing its action in farmed animals. It is known that the dietary protein arginine can stimulate the release of insulin and arginine supplementation has successfully been used to enhance growth in terrestrial livestock (Flynn *et al*., 2002; Kim *et al*., 2004; Yao *et al*., 2008; Wu *et al*., 2009). Large scale arginine supplementation also holds great promise for growth and health management in aquaculture (Li *et al*., 2009). It has been used with success in aquacultured fish, like rainbow trout (*Oncorhynchus mykiss*; Gutiérrez and Plisetskaya, 1991; Riley *et al*., 1996), giant gouramy (*Osphronemus gouramy*; Suprayudi *et al*. 2000) and cobia (*Rachycentron canadum*; Zhao *et al*., 2007). It has also been demonstrated that IGFBPs in Atlantic salmon (*Salmo salar*) are differentially regulated with nutritional status (Bower *et al*., 2008), but specific amino acid composition of the formulated diet used was not reported.

Also for abalone (*H. tuberculata* and *H. discus hannai*) the increase of arginine concentrations in formulated diets is suggested to be beneficial in terms of improved growth (Mai *et al*., 1994). On the contrary, arginine supplementation trials in *Haliotis midae* reported that arginine does not promote growth in this species (Britz *et al*., 1997; Shipton, 1999). This finding was however due in part to technical challenges to correctly quantify abalone amino acid requirements. These challenges are concerned with the apparent inability of abalone to utilize whole proteins and crystalline amino acids used in feeding trials (Britz *et al*., 1997; Shipton, 1999). A later study by Sales and Britz (2003) concluded that proteins and amino acids in oil seeds and legumes are however highly digestible by *H. midae* and that one such oil seed, cottonseed, contains high levels of arginine. Although protein digestibility is an indicator of mean amino acid availability, complete information on availability of amino acids from the diet is still lacking. Further studies are needed to determine the correct proportions of essential amino acids and the extent of nutrient leaching in formulated diets (Sales and Britz, 2003; Sales *et al.*, 2003).

Another method used in vertebrates to manipulate insulin levels is growth hormone treatment, as growth hormone is the primary stimulus for synthesis and release of IGF-I (Feng *et al*., 2009; Reinecke, 2010). Although growth hormone has not been identified in abalone, it is implied that it also plays a role in the growth stimulation of molluscs. After immersion and intramuscular injection of recombinant salmon growth hormone, an increase in shell length and body mass was observed in juvenile abalone (*H. discus hanai*) (Moriyama and Kawauchi, 2004; Moriyama *et al*., 2008). Recent studies where salmon growth hormone was entrapped in sodium alginate gel and fed to abalone, reports the efficiency of this approach to accelerate the somatic growth of juvenile abalone (Moriyama *et al*., 2009).

Although sodium alginate gel has also been reported to be a good binder for feed consistency (with minimal leaching of nutrients) in sea urchin (*Strongylocentrotus droebachiensis*) feeding trials, it is not recommended for use on a commercial scale because of cost implications. These authors recommended gelatin as the best binder in terms of feed stability and cost-effectiveness (Pearce *et al*., 2002). Both gelatin and sodium alginate have been used as binders in abalone feeding trials before but despite their inclusion, large amounts of essential amino acids (53.5 % of arginine) were still reported to leach from the food within 24 hours. Also, as gelatin was found to be rejected by abalone, it was suggested that further research is needed to evaluate protein loss and develop strategies aimed at reducing it (Coote *et al*., 2000).

By providing evidence at the transcriptional level for the involvement of insulin, IGFs and IGFBPs in improved growth rate of *Haliotis midae* with this study, the relevance of investigating ways to stimulate insulin/IGF release is again emphasized. Although other factors, like temperature and photoperiod, environmental salinity and endocrine disrupting compounds may influence levels of peptides involved in the growth hormone/insulin-like growth peptide system, nutritional status is believed to be the principal environmental regulator of circulating IGF (Reinecke, 2010). Therefore, nutritional administration remains the most probable route of introducing agents that can stimulate the release of insulin-related peptides. Both arginine supplementation and growth hormone as agents of stimulating insulin and IGF release in abalone holds promise for significant improvement in growth rate. Continuous endeavours to stimulate abalone growth through a nutritional approach are therefore encouraged. It is specifically recommended to continue optimizing arginine availability and to initiate research to incorporate growth hormone in the *H. midae* diet.

Confirmation of the beneficial properties of arginine may be evident from the *H. midae* larval cell culture trial performed during this study. Although it was not intended to investigate the effect of arginine, supplementation of medium with 1 % amino acids (which contains 1.05 g/L arginine) seemed to sustain larval cell culture better over a period of seven days, than medium without addition of the

amino acids. This is an indication of the *in vitro* growth stimulatory effect of amino acids. Further trials, also using *in vitro* culture of other tissues, may shed light on this matter. As the digestive gland cells are the main site of digestion and nutrient assimilation, *in vitro* culture of these cells may be a suitable approach to study the effect of arginine supplementation. Digestive gland primary cell cultures have been successfully performed in other molluscs, like the scallop, *Pecten maximus* (Le Pennec and Le Pennec, 2001) and the mussel, *Mytilus edulis* (Faucet *et al*., 2003).

Haliotis midae haemocyte cell cultures could be used as a suitable *in vitro* system for studying gene expression after new genes were carefully selected for their possible involvement in growth metabolic pathways and their known expression in mollusc haemocytes. The up-regulation of sequence 809, or Mnk, a MAP kinase-interacting serine/threonine kinase involved in the MAPK- and insulin signaling pathways, in response to insulin treatment corroborates the notion of a growth hormone-IGF axis in abalone, similar to that established for mammals and fish (Reinecke, 2010).

The haemocytes represent the main cellular component of the molluscan immune system (Humphries and Yoshino, 2003) and was as such not the ideal tissue for investigating growth-related gene expression. Nevertheless, as more work has been reported on immune function and immune-regulating systems in molluscan haemocytes, this study could draw from that knowledge. Involvement of a variety of receptor types (including receptors for growth factors, hormones and cytokines) in common signaling elements suggest a conservation of signaling pathways within moluscan haemocytes (Humphries and Yoshino, 2003). Some of these signaling pathways considered to be involved in immunity also play a role in growth. Examples include pathways initiated by binding of insulin-like receptors and MAPK-mediated signaling pathways (Humphries and Yoshino, 2003; Iakovleva *et al*., 2006a, b).

MAPK homologues have been identified in molluscan haemocytes before. In the haemocytes of the periwinkle (*Littorina littorea*), MAPKs of all three MAPK subfamilies (ERK, p38 and JNK) was identified when investigating the contribution of different MAPKs to the responses triggered by immune stimuli (Iakovleva *et al.* 2006a, b). Since the earliest identification of MAPKs, in response to stimuli with growth factors (including EGF and insulin), it has been established that MAPKs are expressed in all eukaryotic cells and that basic assembly of MAPK pathways is conserved. Activation of MAPKs occurs in response to a diverse array of stimuli, ranging from growth factors to irradiation or mechanical stress (Widmann *et al*., 1999). It has been shown that even cell adhesion to cell culture plates causes MAPK activation in insect (*Ceratitis capitata*) haemocytes (Foukas, 1998). The large number of MKKKs that allows for various inputs from numerous stimuli and the intricate web of subfamilies that have been identified for the different components involved in MAPK pathways, accentuate the challenge of placing a MAPK component identified for the first time in a species, within the known framework.

The fact that sequence 809 (Mnk) was up-regulated in insulin-treated cells when compared to control cells maintained under the same conditions and in the same cell culture plates, lead to the conclusion that this MAPK component was stimulated by the binding of insulin to its receptor in *H. midae* haemocytes. If the MAPK pathway is involved, it is curious why treatment with EGF did not elicit the same up-regulation of Mnk. It is possible that the concentration of EGF used in the treatments was not optimal, or that it had a toxic or inhibitory effect. It has been found that some naturally occurring halogen compounds in the marine environment can elicit inhibitory activity against EGF receptor kinase. These include bromotyrosine alkaloids and tauroacidins isolated from sponges (*Hymeniacidon* sp. and *Psammaplysilla purea*) (Kobayashi *et al.*, 1997; Laus, 2001). The possibility that similar compounds to those involved in marine sponge defense mechanisms could be present in the vicinity of farmed abalone, and thus the animals used during haemolymph collection, cannot be excluded.

It is more likely that it was not via the MAPK pathway, but rather via the insulin signaling pathway that sequence 809 (Mnk) was up-regulated in response to insulin treatment. Within the insulin signaling pathway Mnk also acts as a MAP kinase-activated kinase - one that regulates eIF4E phosphorylation. It is known that phosphorylation of eIF4E is increased following insulin treatment, in an ERK-dependent manner (Flynn and Proud, 1996; Waskiewicz *et al*., 1997).

Besides the regulation of target genes, valuable knowledge was also gained in terms of reference genes, during qPCR experimentation. By quantifying the stable expression of two genes (8629, ribosomal protein S9 and 12621, ornithine decarboxylase) in various tissues and under various conditions (whole animal and haemocyte cultures receiving different growth factor treatments), suitable reference genes were identified. These were used successfully in this study to calculate relative gene expression and they can also be used in future qPCR studies on *H. midae* tissues.

In a similar approach as was followed here, genes identified as up-regulated in small animals may also be targeted in tissue-specific *in vitro* assays in future. Speculations about possible toxins and their effects on various abalone tissues may be investigated in a controlled *in vitro* environment. Many aquatic invertebrate species, including farmed molluscs are used as bioindicators of environmental pollution. *In vitro* assays, especially cell culture from fish tissues, have been developed for evaluating environmental toxicology (Villena, 2003). Also oyster (*Crassostrea gigas*) heart and clam (*Ruditapes decussates*) gill primary cell cultures have been used to evaluate the cytotoxicity of an organic molluscicide and several organohalogenated compounds (Domart-Coulon *et al*., 2000). Oyster haemocytes have also been used in order to investigate the effects of acute cadmium and mercury toxicity (Gagnaire *et al*., 2004; Gagnaire *et al*., 2006). Recently, *Haliotis tuberculata* primary haemocyte cultures were also successfully used to investigate the effects of zinc pollution *in vitro* (Mottin *et al*.,

2010). *Haliotis midae* haemocyte cell cultures can be used in a similar way as a suitable system for detection of cytotoxicity.

Optimal commercial production of *H. midae* can be attained by using an interdisciplinary approach. The combined application of improved on-farm culture practices, nutrition, *in vitro* assays and genetic enhancement steered by advanced molecular techniques can overcome obstacles ranging from slow growth rate to disease resistance and detrimental water quality. Close cooperation between farms and research institutions provide the best opportunity towards improving the local production of this most valuable of South African aquaculture species. With the advent of next-generation sequencing, not only can we answer questions pertaining to improved farming of *H. midae* more thoroughly, but also understand the phylogeny and evolution of one of the oldest classes of living molluscs to a greater extent. This research hopes to contribute to this journey of understanding, improvement and appreciation of haliotids.

5.2 References

Auzoux-Bordenave, S., M. Fouchereau-Peron, M.N. Helléouet, and D. Doumenc. 2007. Activity of mantle cells and hemocytes in abalone primary cell cultures (*Haliotis tuberculata*). *Journal of Shellfish Research.* 26:887-894.

Batley, G.E., and S.L. Simpson. 2009. Development of guidelines for ammonia in estuarine and marine water systems. *Marine Pollution Bulletin.* 58:1472-1476.

Bower, N.I., and I.A. Johnston. 2010. Transcriptional regulation of the IGF signaling pathway by amino acids and insulin-like growth factors during myogenesis in Atlantic salmon. *PLoS ONE.* 5:e11100.

Bower, N.I., X. Li, R. Taylor, and I.A. Johnston. 2008. Switching to fast growth: the insulin-like growth factor (IGF) system in skeletal muscle of Atlantic salmon. *The Journal of Experimental Biology.* 211:3859-3870.

Britz, P.J., N. Bacela, and T. Hecht. 1997. Can crystalline arginine be used to quantify the arginine requirement of abalone? *Aquaculture.* 157:95-105.

Clark, M.S., M.A.S. Thorne, F.A. Vieira, J.C.R. Cardoso, D.M. Power, and L.S. Peck. 2010. Insights into shell deposition in the Antarctic bivalve *Laternula elliptica*: gene discovery in the mantle transcriptome using 454 pyroseqeuencing. *BMC Genomics.* 11:362.

Coote, T.A., P.W. Hone, R.J. Van Barneveld, and G.B. Maguire. 2000. Optimal protein level in a semipurified diet for juvenile greenlip abalone *Haliotis laevigata*. *Aquaculture Nutrition.* 6:213-220.

Domart-Coulon, I., S. Auzoux-Bordenave, D. Doumenc, and M. Khalanski. 2000. Cytotoxicity assessment of antibiofouling compounds and by-products in marine bivalve cell cultures. *Toxicology in Vitro.* 14:245-251.

Elmer, K.R., S. Fan, H.M. Gunter, J.C. Jones, S. Boekhoff, S. Kuraku, and A. Meyer. 2010. Rapid evolution and selection inferred from the transcriptomes of sympatric crater lake cichlid fishes. *Molecular Ecology.* 19:197-211.

Engert, J.C., E.B. Berglund, and N. Rosenthal. 1996. Proliferation precedes differentiation in IGF-I-stimulated myogenesis. *Journal of Cell Biology.* 135:431-440.

Faucet, J., M. Maurice, B. Gagnaire, T. Renault, and T. Burgeot. 2003. Isolation and primary culture of gill and digestive gland cells from the common mussel *Mytilus edulis*. *Methods in Cell Science.* 25:177-184.

Feng, J., Z. Gu, M. Wu, F.C. Gwazdauskas, and H. Jiang. 2009. Growth hormone stimulation of serum insulin concentration in cattle: Nutritional dependency and potential mechanisms. *Domestic Animal Endocrinology.* 37:84-92.

Flynn, A., and C.G. Proud . 1996. Insulin-stimulated phosphorylation of initiation factor 4E is mediated by the MAP kinase pathway. *FEBS Letters.* 389:162-166.

Flynn, N.E., C.J. Meininger, T.E. Haynes, and G. Wu. 2002. The metabolic basis of arginine nutrition and pharmacotherapy. *Biomedicine and Pharmacotherapy.* 56:427-438.

Foukas, L.C., H.K. Katsoulas, N. Paraskevopoulou, A. Metheniti, M. Lambropoulou, and V.J. Marmaras. 1998. Phagocytosis of Escherichia coli by insect hemocytes requires both activation of the Ras/Mitogen-activated protein kinase signal transduction pathway for attachment and b3 integrin for internalization. *The Journal of Biological Chemistry.* 273:14813-14818.

Gagnaire, B., H. Thomas-Guyon, and T. Renault. 2004. *In vitro* effects of cadmium and mercury on Pacific oyster, *Crassostrea gigas* (Thunberg), haemocytes. *Fish and Shellfish Immunology.* 16:501-512.

Gagnaire, B., H. Frouin, K. Moreau, H. Thomas-Guyon, and T. Renault. 2006. Effects of temperature and salinity on haemocyte activities of the Pacific oyster, *Crassostrea gigas* (Thunberg). *Fish and Shellfish Immunology.* 20:536-547.

Gompert, Z., M.L. Forister, J.A. Fordyce, C.C. Nice, R.J. Williamson, and C.A. Buerkle. 2010. Bayesian analysis of molecular variance in pyrosequences quantifies population genetic structure across the genome of Lycaeides butterflies. *Molecular Ecology.* 19:2455-2473.

Gricourt, L., G. Bonnec, D. Boujard, M. Mathieu, and K. Kellner. 2003. Insulin-like system and growth regulation in the Pacific oyster *Crassostrea gigas*: hrIGF-1 effect on protein synthesis of mantle edge cells and expression of an homologous insulin receptor-related receptor. *General and Comparative Endocrinology.* 134:44-56.

Gricourt, L., M. Mathieu, and K. Kellner. 2006. An insulin-like system involved in the control of Pacific oyster *Crassostrea gigas* reproduction: hrIGF-1 effect on germinal cell proliferation and maturation associated with expression of an homologous insulin receptor-related receptor. *Aquaculture.* 251:85-98.

Gutiérrez, J., and E.M. Plisetskaya. 1991. Insulin binding to liver plasma membranes in salmonids with modified plasma insulin levels. *Canadian Journal of Zoology.* 69:2745-2750.

Haruta, T., T. Uno, J. Kawahara, A. Takano, K. Egawa, P.M. Sharma, J.M. Olefsky, and M. Kobayashi. 2000. A rapamycin-sensitive pathway down-regulates insulin signaling via phosphorylation and proteasomal degradation of insulin receptor substrate-1. *Molecular Endocrinology.* 14:783-794.

Humphries, J.E., and T.P. Yoshino. 2003. Cellular receptors and signal transduction in molluscan hemocytes: Connections with the innate immune system of vertebrates. *Integrative and Comparative Biology.* 43:305-312.

Iakovleva, N.V., A.M. Gorbushin, and K.B. Storey. 2006a. Modulation of mitogen-activated protein kinases (MAPK) activity in response to different immune stimuli in haemocytes of the common periwinkle *Littorina littorea*. *Fish and Shellfish Immunology.* 21:315-324.

Iakovleva, N.V., A.M. Gorbushin, and U.E. Zelck. 2006b. Partial characterization of mitogen-activated protein kinases (MAPK) from haemocytes of the common periwinkle, *Littorina littorea* (Gastropoda: Prosobranchia). *Fish and Shellfish Immunology.* 20:665-668.

Kaushik, S., and I. Seiliez. 2010. Protein and amino acid nutrition and metabolism in fish: current knowledge and future needs. *Aquaculture Research.* 41:322-332.

Kim, H.W., D.L. Mykles, F.W. Goetz, and S.B. Roberts. 2004. Characterization of a myostatin-like gene from the bay scallop, *Argopecten irradians*. *Biochimica et Biophysica Acta.* 1679:174-179.

Kobayashi, J., K. Inaba, and M. Tsuda. 1997. Tauroacidins A and B, new bromopyrrole alkaloids possessing a taurine residue from *Hymeniacidon* sponge. *Tetrahedron.* 53:16679-16682.

Laus, G. 2001. Biological activities of natural halogen compounds. *Studies in Natural Products Chemistry.* 25:757-810.

Le Pennec, G., and M. Le Pennec. 2001. Acinar primary cell culture from the digestive gland of *Pecten maximus* (L.): an original model for ecotoxicological purposes. *Journal of Experimental Marine Biology and Ecology.* 259:171-187.

Li, P., K. Mai, J. Trushenski, and G. Wu. 2009. New developments in fish amino acid nutrition: towards functional and environmentally oriented aquafeeds. *Amino Acids.* 37:43-53.

Machado, C.R., M.A.R. Garofalo, J.E.S. Roselino, I.C. Kettlehut, and R.H. R. H. Migliorini. 1988. Effects of starvation, refeeding, and insulin on energy-linked metabolic processes in catfish (*Rhamdia hilarii*) adapted to a carbohydrate-rich diet. *General and Comparative Endocrinology.* 71:429-437.

Mai, K., J.P. Mercer, and J. Donlon. 1994. Comparative studies on the nutrition of two species of abalone, *Haliotis tuberculata* L. and *Haliotis discus hannai* Ino II. Amino acid composition of abalone and six species of macroalgae with an assessment of their nutritional value. *Aquaculture.* 128:115-130.

Marin, F., and G. Luquet. 2004. Molluscan shell proteins. *Comptes Rendus Palevol.* 3:469-492.

Montserrat, N., P. Gómez-Requeni, G. Bellini, E. Capilla, J. Pérez-Sánchez, I. Navarro, and J. Gutiérrez. 2007. Distinct role of insulin and IGF-I and its receptors in white skeletal muscle during the compensatory growth of gilthead sea bream (*Sparus aurata*). *Aquaculture.* 267:188-198.

Moriyama, S., and H. Kawauchi. 2004. Somatic growth acceleration of juvenile abalone, *Haliotis discus hannai*, by immersion in and intramuscular injection of recombinant salmon growth hormone. *Aquaculture.* 229:469-478.

Moriyama, S., K. Tashiro, S. Furukawa, and H. Kawauchi. 2008. Ability of salmon growth hormone to accelerate somatic growth of juvenile abalone *Haliotis discus hannai*. *Fisheries Science.* 74:860-866.

Moriyama, S., S. Furukawa, and H. Kawauchi. 2009. Growth stimulation of juvenile abalone *Haliotis discus hannai* by feeding with salmon growth hormone in sodium alginate gel. *Fisheries Science.* 75:689-695.

Mottin, E., C. Caplat, M. Mahaut, K. Costil, D. Barillier, J. Lebel, and A. Serpentini. 2010. Effect of *in vitro* exposure to zinc on immunological parameters of haemocytes from the marine gastropod *Haliotis tuberculata*. *Fish and Shellfish Immunology.* In press, doi: 10.1016/j.fsi.2010.07.022.

Pearce, C.M., T.L. Daggett, and S.M.C. Robinson. 2002. Effect of binder type and concentration on prepared feed stability and gonad yield and quality of the green sea urchin, *Strongylocentrotus droebachiensis*. *Aquaculture.* 205:301-323.

Reddy-Lopata, K., L. Auerswald, and P. Cook. 2006. Ammonia toxicity and its effect on the growth of the South African abalone *Haliotis midae* Linnaeus. *Aquaculture.* 261:678-687.

Reinecke, M. 2010. Influences of the environment on the endocrine and paracrine fish growth hormone – insulin-like growth factor-I system. *Journal of Fish Biology.* 76:1233-1254.

Riley, W.W., D.A. Higgs, B.S. Dosanjh, and J.G. Eales. 1996. Influence of dietary arginine and glycine content on thyroid function and growth of juvenile rainbow trout, *Oncorhynchus mykiss* (Walbaum). *Aquaculture Nutrition.* 2:235-242.

Sales, J., and P.J. Britz. 2003. Apparent and true availability of amino acids from common feed ingredients for South African abalone (*Haliotis midae* L.). *Aquaculture Nutrition.* 9:55-64.

Sales, J., P.J. Truter, and P.J. Britz. 2003. Optimum dietary crude protein level for growth in South African abalone (*Haliotis midae* L.). *Aquaculture Nutrition.* 9:85-89.

Schulz, M., and D. Zerbino. 2010. Oases: *De novo* transcriptome assembler for very short reads. 0.1.8 [Online]. Available: http://www.ebi.ac.uk/~zerbino/oases/ [Accessed 10 June 2010].

Shipton, T.A. 1999. The protein requirements of the South African abalone, *Haliotis midae*. Unpublished PhD dissertation, Rhodes University, Grahamstown, South Africa.

Slatinska, I., M. Smutna, M. Havelkova, and Z. Svobodova. 2008. Biochemical markers of aquatic pollution in fish - Glutathione-S-Transferase. *Folia Veterinaria.* 52:129-134.

Suja, C.P., N. Sukumaran, and S. Dharmaraj. 2007. Effect of culture media and tissue extracts in the mantle explant culture of abalone, *Haliotis varia* Linnaeus. *Aquaculture.* 271:516-522.

Suprayudi, M.A., T. Takeuchi, I. Mokoginta, and A.T. Kartikasari. 2000. The effect of additional arginine in the high defatted soybean meal diet on the growth of giant gouramy *Osphronemus gouramy* Lac. *Fisheries Science.* 66:807-811.

Villena, A.J. 2003. Applications and needs of fish and shellfish cell culture for disease control in aquaculture. *Reviews in Fish Biology and Fisheries.* 13:111-140.

Waskiewicz, A.J., A. Flynn, C.G. Proud, and J.A. Cooper. 1997. Mitogen-activated protein kinases activate the serine/threonine kinases Mnk1 and Mnk2. *The EMBO Journal.* 16:1909-1920.

Widmann, C., S. Gibson, M.B. Jarpe, and G.L. Johnson. 1999. Mitogen-activated protein kinase: conservation of a three-kinase module from yeast to human. *Physiological Reviews.* 79:143-180.

Wu, G., F.W. Bazer, T.A. Davis, S.W. Kim, P. Li, J.M. Rhoads, M.C. Satterfield, S.B. Smith, T.E. Spencer, and Y. Yin. 2009. Arginine metabolism and nutrition in growth, health and disease. *Amino Acids.* 37:153-168.

Yao, K., Y.L. Yin, W. Chu, Z. Liu, D. Deng, T. Li, R. Huang, J. Zhang, B. Tan, W. Wang, and G. Wu. 2008. Dietary arginine supplementation increases mTOR signaling activity in skeletal muscle of neonatal pigs. *The Journal of Nutrition.* 138:867-872.

Yearsley, R.D. 2007. Water quality, abalone growth and the potential for integrated mariculture on a South African abalone *Haliotis midae* L. farm. Unpublished MSc thesis, Rhodes University, Grahamstown, South Africa.

Zhao, H.X., J.M. Cao, J.K. Wu, Y.G. Tan, M. Zhou, H.O. Liang, and D.W. Yang. 2007. Studies of arginine requirement for juvenile cobia. *Journal of South China Agricultural University*: 4, doi: CNKI:SUN:HNNB.0.2007-04-022

6 APPENDIX

Table 1.2 Growth related nucleotide and protein sequences in mollusc and other model invertebrate organisms

Organism	Protein/Gene	Genbank accession	Authors
Somatotropic axis / CNS		# Nucleotide **# Protein**	
Aplysia californica	MDGF	**AAD13112**	Akalal and Nagle, 2003
Aplysia californica	MDGF	AF117336.1	Akalal *et al*., 1999 (unpubl)
Lymnaea stagnalis	Insulin-related peptide III	AH004423.1	Smit *et al*., 1993
Lymnaea stagnalis	MIP	X06983.1 **P07223.2**	Smit *et al*., 1988
Lymnaea stagnalis	MIPII	X59302.1 **P25289.1**	Smit *et al*., 1991
Lymnaea stagnalis	MIPIII (3 segments)	AH004423.1 **AAB28954.1**	Smit *et al.,* 1993
Lymnaea stagnalis	MIPIII segment 2	S69104.1	Smit *et al.,* 1993
Lymnaea stagnalis	MIPIII segment 3	S69155.1	Smit *et al.,* 1993
Lymnaea stagnalis	MIPV	S42163.1 **AAA09966.1**	Smit *et al.,* 1993
Lymnaea stagnalis	MIPVII	S82894.1 **AAB46831.1**	Smit *et al.,* 1993
Aplysia californica	MIPII	AF364181.2 **AAL99711.2**	Moroz *et al.,* 2006
Bombyx mori	Insulin receptor-like protein (BIR)	AF025542.1 **AAF21243.1**	Lindstrom-Dinnetz and Iatrou, 1997 (unpubl)
Crassostrea gigas	Partial ir gene for MIP receptor (CIR)	AJ535669.1 **CAD59674.1**	Gricourt *et al*., 2006
Bombyx mori	IDGF like protein	AB041634.1 **BAB16695.1**	Tsuzuki *et al*., 2001
Bombyx mori	IDGF	AB183872.1 **BAF73623.1**	Tsuzuki *et al*., 2001
Sarcophaga peregrina	IDGF	D83125.1 **BAA11812.1**	Homma *et al*., 1996
Aplysia californica	Atrial gland-specific antigen	J05059.1 **AAA27741.1**	Sossin *et al*., 1989
Glossina morsitans morsitans	TSGF-1	AF140521.1 **AAD52850.1**	Li and Aksoy, 2000
Glossina morsitans morsitans	TSGF-2	AF140522.1 **AAD52851.1**	Li and Aksoy, 2000
Lymnaea stagnalis	EGFR	EF492516.1 **ABQ10634.1**	Van Kesteren *et al*., 2008
Mizuhopecten yessoensis	CRP1	AB167507.1 **BAD72613.1**	Nara *et al*., 2004

Lutzomyia longipalpis	Putative ADA	AF234182 **AAF78901.1**	Charlab *et al*., 2000
Drosophila melanogaster	Male-specific insect derived growth factor (Msi)	NM_080281.3 **NP_525020.2**	Hoskins *et al*., 2007
Haliotis asinina	Class III POU	AF231942.2 **AAK26108.2**	O'Brien and Degnan, 2000
Haliotis asinina	Class IV POU transcription factor	AF231943.2 **AAK26109.2**	O'Brien and Degnan, 2000
Haliotis asinina	Transcription factor sox-B	AF231944 **AAK26110.1**	O'Brien and Degnan, 2000
Haliotis asinina	Sox group C transcription factor	AF231945 **AAK26111.1**	O'Brien and Degnan, 2000
Haliotis asinina	Transcription factor pax-258	AF231946 **AAK26112.1**	O'Brien and Degnan, 2000
Haliotis asinina	Pax-6	AF231947 **AAK26113.1**	O'Brien and Degnan, 2000
Patella vulgata	uSoxB	AJ580004.1 **Q7YUD7**	Le Gouar *et al*., 2003 (unpubl)
Aplysia californica	Beta-thymosin	AF481063.1 **AAM09681.1**	Moccia *et al*., 2003
Aplysia californica	Beta-thymosin	AF454398.1 **AAM22407.1**	Colby *et al*., 2005
Strongylocentrotus purpuratus	Thymosin beta	AF076515 **AAC26833.1**	Pancer *et al*., 1999
Muscle growth			
Argopecten irradians	sMSTN	AY553362 **AAT36326.1**	Kim *et al*., 2004
Mytilus edulis	Actin	AF172606.1 **AAD48064.1**	Luedeking *et al*., 1999 (unpubl)
Haliotis discus discus	Actin	EF103363.1 **ABO26621.1**	Kang *et al.,* 2008 (unpubl)
Haliotis diversicolor	Actin promotor	EU622901.2 **ACC91877.2**	Li *et al*., 2008 (unpubl)
Haliotis diversicolor	Actin	EF587284.1 **ABU86741.1**	Wang *et al*., 2008
Haliotis diversicolor	Beta-thymosin	EF542808.1 **ABU53029.1**	Wang *et al*., 2008
Haliotis diversicolor	Profilin	EU244332.1 **ABY87349.1**	Wang *et al*., 2008
Haliotis iris	Actin	AY921239.1 **AAX19288.1**	Bryant *et al*., 2006
Haliotis virginea	Actin	AY959327.1 **AAY00160.1**	Bryant *et al*., 2006
Haliotis iris	Actin A1a	AY961954.1 **AAY00162.1**	Sin *et al*., 2005 (unpubl)
Haliotis iris	Actin A1b	AY961955.1 **AAY00163.1**	Sin and Bryant, 2005 (unpubl)
Haliotis iris	Actin A1c	AY961956.1 **AAY00163.1**	Sin and Bryant, 2005 (unpubl)
Haliotis tuberculata	mRNA for Actin	AM236595.1	Fleury *et al*., 2006

		CAJ85786.1	(unpubl)
Haliotis iris	Actin A1 gene	AY921237.1 **AAX19286.1**	Bryant *et al*., 2006
Haliotis iris	Actin A2	AY921238.1 **AAX19287.1**	Bryant *et al*., 2006
Haliotis rufescens	Tropomyosin	X75218.1 **CAA53028.1**	Degnan *et al*., 1997
Crassostrea gigas	mGDF	AJ130967.1 **CAA10268.1**	Lelong *et al.*, 2000
Shell growth			
Pinctada fucata	Nacrein	D83523.1 **BAA11940.1**	Miyamoto *et al*., 1996
Pinctada fucata	Nacrein	**BAA11940**	Miyamoto *et al*., 1996
Haliotis rufescens	Lustrin A	AF023459 **AAB95154.1**	Shen *et al.*, 1997
Pinctada fucata	MSI31	D86073 **BAA20465.1**	Sudo *et al.*, 1997
Pinctada fucata	MSI60	D86074 **BAA20466.1**	Sudo *et al*., 1997
Pinctada fucata	MSI7	AF516712 **AAQ08227.1**	Zhang *et al*., 2003
Pinctada fucata	N16	AB023067.1 **BAA83732.1**	Samata *et al.,* 1999
Pinctada maxima	N14	AB032612.1 **BAA90539.1**	Kono *et al.,* 2000
Pinctada maxima	N66	AB032613.1 **BAA90540.1**	Kono *et al.* 2000
Pinna nobilis	Mucoperlin (mcp)	AF145215 **AAK18045.1**	Marin *et al.*, 2000
Biomphalaria glabrata	BgDerm1	AB210096 **BAD97853.1**	Sarashina *et al.,* 2006
Biomphalaria glabrata	Dermatopontin2	AB210097.1 **BAD97854.1**	Sarashina *et al.,* 2006
Lymnaea stagnalis	LsDerm1	AB210093 **BAD97850.1**	Sarashina *et al*., 2006
Lymnaea stagnalis	Dermatopontin2	AB210094.1 **BAD97851.1**	Sarashina *et al*., 2006
Lymnaea stagnalis	Dermatopontin3	AB210095.1 **BAD97852.1**	Sarashina *et al*., 2006
Euhadra herklotsi	EhDerm1	AB210100 **BAD97857.1**	Sarashina *et al.,* 2006
Euhadra herklotsi	Dermatopontin2	AB210101.1 **BAD97858.1**	Sarashina *et al.,* 2006
Euhadra brandtii	EbDerm1	AB210099.1 **BAD97856.1**	Sarashina *et al.,* 2006
Euhadra brandtii	EbDerm2	AB210098 **BAD97855.1**	Sarashina *et al.,* 2006
Euhadra amaliae	EaDerm1	AB210102 **BAD97859.1**	Sarashina *et al*., 2006
Euhadra peliomphala	Dermatopontin2	AB210103.1 **BAD97860.1**	Sarashina *et al*., 2006

Mandarina aureola	Dermatopontin1	AB210104.1 **BAD97861.1**	Sarashina *et al*., 2006
Mandarina aureola	MaDerm2	AB210105 **BAD97862.1**	Sarashina *et al*., 2006
Satsuma japonica	Dermatopontin2	AB210106.1 **BAD97863.1**	Sarashina *et al*., 2006
Haliotis asinina	Has-ubfm	DW986191	Jackson *et al*., 2006
Haliotis asinina	Has-ferrt	DW986406	Jackson *et al*., 2006
Haliotis asinina	Has-calmbp	DW986217	Jackson *et al*., 2006
Haliotis asinina	Has-cam	DW986371	Jackson *et al*., 2006
Haliotis asinina	Has-lustA	DQ298402 **ABC00197.1**	Jackson *et al*., 2006
Haliotis asinina	Has-Som	DW986219	Jackson *et al*., 2006
Haliotis asinina	Has-tfgr	DW986319	Jackson *et al*., 2006
Mytilus edulis	Topoisomerase II	AF227976 **AAF35892.1**	Luedeking *et al*., 2000 (unpubl)
Haliotis discus discus	Perlustrin	EF103331.1 **ABO26589.1**	Lee *et al*., 2006 (unpubl)
Haliotis laevigata	Perlustrin	**P82595.2**	Weiss *et al.,* 2001
Haliotis laevigata	Perlucin	**P82596.3**	Mann *et al.,* 2000
Pinctada fucata	Aspein	AB094512.1 **BAD00044.1**	Tsukamoto *et al*., 2002 (unpubl)
Pinctada fucata	Aspein	EU177508.1 **ABY84933.1**	Guan and He, 2007 (unpubl)
Pinctada fucata	Prismalin-14	AB159512.1 **BAD27406.1**	Suzuki *et al*., 2004
Crassostrea gigas	mRNA for calcitonin-like receptor precursor	AJ551182.1 **CAD82836.1**	Dubos *et al*., 2003
Ornithodoros parkeri	adrenomodulin-like protein	**ABR23448.1**	Francischetti *et al*., 2007 (unpubl)
Strongylocentrotus purpuratus	CGRP-receptor component protein	**XP_788533.2**	Predicted
Ciona intestinalis	calcitonin gene-related peptide receptor component protein	**BAB85994.1**	Inaba and Kho, 2002 (unpubl)
Nasonia vitripennis	CGRP receptor component	**NP_001136339.1**	NCBI Refseq, 2008
Nasonia vitripennis	CGRP receptor component	**NP_001136340.1**	NCBI Refseq, 2008
Lepeophtheirus salmonis	Calcitonin gene-related peptide type 1 receptor precursor	**ACO12944.1**	Yasuike *et al*., 2009 (unpubl)

Miscellaneous			
Crassostrea gigas	Clp1 protein	AJ971240.1 **CAI96028.1**	Badariotti *et al*., 2006
Crassostrea gigas	Clp2 protein	**CAI96023.1**	Badariotti, 2005 (unpubl)
Crassostrea gigas	Clp3 protein	**CAI96024.1**	Badariotti, 2005 (unpubl)
Haliotis diversicolor	Collagen type XXII alpha 1-like protein	EU244379.1 **ABY87395.1**	Wang *et al*., 2008
Haliotis diversicolor	Collagen type XXI-like protein	EU244340.1 **ABY87357.1**	Wang *et al*., 2008
Haliotis discus	Collagen pro alpha-chain	AB017601.1 **BAA75669.1**	Yoneda *et al*., 1999
Haliotis discus	Collagen pro alpha-chain	AB017600.1 **BAA75668.1**	Yoneda *et al*., 1999
Haliotis discus hannai	Ferritin subunit 1	EU660052.1 **ACD12135.1**	Liu *et al*., 2008 (unpubl)
Haliotis discus discus	Ferritin subunit 1	DQ821493.1 **ABG88845.1**	De Zoysa and Lee, 2007
Haliotis discus discus	Ferritin subunit 2	DQ821494.1 **ABG88846.1**	De Zoysa and Lee, 2007
Haliotis discus hannai	Ferritin	DQ845482.1 **ABH10672.1**	Zheng and Wang, 2006 (unpubl)
Haliotis asinina	Metallothionein ML5H2	DW986335.1	Jackson *et al*., 2006
Littorina littorea	Ferritin heavy chain	AY090096.1 **AAM10781.1**	Larade and Storey, 2004

Table 2.10 Functional classification of contigs from the *Haliotis midae* transcriptome (R) with a BLAST E-value of ≤ 10^{-10}, based on the GO database

GENE ONTOLOGY CLASSIFICATION	Numbers of ID
Cellular component	**516**
Cellular component	353
Extracellular region	18
Proteinaceous extracellular matrix	6
Extracellular space	2
Cell wall	1
Intracellular	242
Cell	345
Nucleus	71
Nuclear envelope	4
Nucleoplasm	2
Chromosome	9
Nucleolus	1
Cytoplasm	150
Mitochondrion	26
Lysosome	1
Endosome	1
Vacuole	3
Endoplasmic reticulum	19
Golgi apparatus	14
Microtubule organizing center	2
Cytosol	18
Ribosome	72
Cytoskeleton	70
Plasma membrane	26
Cilium	3
Plastid	12
Thylakoid	1
Cytoplasmic membrane-bounded vesicle	21
Organelle	140
Protein complex	154
Molecular function	**973**
Nucleotide binding	216
Molecular function	2
Nucleic acid binding	35
DNA binding	40
Chromatin binding	5
Transcription factor activity	12
RNA binding	35
Motor activity	36
Actin binding	32
Catalytic activity	254
Nuclease activity	3
Protein kinase activity	21

Phosphoprotein phosphatase activity	9
Receptor activity	14
Receptor binding	3
Structural molecule activity	87
Transporter activity	93
Ion channel activity	7
Neurotransmitter transporter activity	12
Binding	354
Calcium ion binding	37
Protein binding	102
Cytoskeletal protein binding	3
Translation factor activity, nucleic acid binding	23
Peptidase activity	49
Lipid binding	4
Electron carrier activity	26
Antioxidant activity	8
Kinase activity	17
Transferase activity	70
Hydrolase activity	147
Enzyme regulator activity	19
Carbohydrate binding	8
Transcription regulator activity	9
Biological process	**713**
Reproduction	3
Carbohydrate metabolic process	57
Generation of precursor metabolites and energy	52
Nucleobase, nucleoside, nucleotide and nucleic acid metabolic process	65
DNA metabolic process	6
Transcription	6
Translation	93
Protein modification process	42
Cellular amino acid and derivative metabolic process	43
Lipid metabolic process	13
Transport	114
Ion transport	38
Response to stress	19
Organelle organization	25
Cytoskeleton organization	4
Cell cycle	8
Cell communication	5
Signal transduction	27
Cell-cell signaling	3
Multicellular organismal development	17
Behavior	1
Biological process	454
Metabolic process	403
Cell death	4
Catabolic process	89

Biosynthetic process	80
Response to external stimulus	2
Response to biotic stimulus	2
Response to abiotic stimulus	3
Anatomical structure morphogenesis	5
Embryonic development	4
Protein transport	38
Cellular component organization	38
Protein metabolic process	77
Pellular homeostasis	13
Secondary metabolic process	6
Cell differentiation	7
Growth	1
Regulation of gene expression, epigenetic	1
Primary metabolic process	1
Regulation of biological process	62

Table 2.11 Collection of 63 sequences highly expressed in L with significant ($E<10^{-10}$) similarity sequences in the *Lottia gigantea* database

Contig Name	Best Match LOTTIA	E Value	Protein Function best hit	Transcript Function description	KOG classification
NODE_1163	jgi|Lotgi1|169374|	1E-157	Sodium/potassium-transporting ATPase alpha chain (Sodium pump) (Na(+)/K(+) ATPase) (model%: 99, hit%: 99, score: 4130, %id: 76) [*Drosophila melanogaster*]	Monovalent inorganic cation transporter activity	Inorganic ion transport and metabolism
NODE_38783	jgi|Lotgi1|204297|	1E-137	Tubulin beta chain (Beta tubulin) (model%: 99, hit%: 99, score: 2323, %id: 97) [*Paracentrotus lividus*]	GTPase activity, structural molecule activity	Cytoskeleton
NODE_25210	jgi|Lotgi1|208914|	1E-112	Calcium-ATPase [*Mizuhopecten yessoensis*] (model%: 97, hit%: 100, score: 4232, %id: 81)	ATPase activity, coupled to transmembrane movement of ions, phosphorylative mechanism	Inorganic ion transport and metabolism
NODE_1673	jgi|Lotgi1|236462|	2E-92	GBB_LOLFO Guanine nucleotide-binding protein subunit beta (model%: 99, hit%: 100, score: 1711, %id: 93) [*Loligo forbesi*]	Signal transducer activity	General function prediction only ; G-protein beta subunit
NODE_10812	jgi|Lotgi1|189716|	1E-87	PREDICTED: similar to CG4376-PB, isoform B [*Tribolium castaneum*] (model%: 99, hit%: 97, score: 3690, %id: 77) [*Tribolium castaneum*]	Actin binding, calcium ion binding	Cytoskeleton
NODE_7048	jgi|Lotgi1|161608|	7E-85	CHICK Heat shock protein HSP 90-alpha (model%: 99, hit%: 100, score: 3003, %id: 79) [*Gallus gallus*]	Chaperone activity, ATP binding	Posttranslational modification, protein turnover, chaperones
NODE_960	jgi|Lotgi1|139452|	5E-70	CG5870 gene product from transcript CG5870-RA (model%: 99, hit%: 99, score: 7359, %id: 62) [*Drosophila melanogaster*]	Actin binding	Cytoskeleton
NODE_11040	jgi|Lotgi1|105848|	7E-52	S-adenosylhomocysteine hydrolase-like 1 (model%: 99, hit%: 77, score: 2043, %id: 83) [*Xenopus tropicalis*]	Adenosylhomocysteinase activity	Coenzyme transport and metabolism
NODE_1743	jgi|Lotgi1|219544|	1E-50	Similar to TGF beta-inducible nuclear protein 1 (L-name related LNR42) (model%: 99, hit%: 100, score: 1165, %id: 81) [*Bos taurus*]	Structural constituent of ribosome	Uncharacterized conserved protein related to ribosomal protein S8E;

					General function prediction only
NODE_1440	jgi|Lotgi1|238559|	2E-42	Similar to CG14516-PA, isoform A isoform 1 [*Apis mellifera*] (model%: 97, hit%: 90, score: 2112, %id: 35) [*Apis mellifera*]	Membrane alanyl aminopeptidase activity	Amino acid transport and metabolism
NODE_54	jgi|Lotgi1|116244|	1E-39	NONE	Extracellular matrix structural constituent; collagen	Extracellular structures
NODE_1020	jgi|Lotgi1|177837|	4E-37	Similar to Heat shock protein 8 (model%: 99, hit%: 100, score: 2832, %id: 84) [*Danio rerio*]		Molecular chaperones HSP70/HSC70, HSP70 superfamily; Posttranslational modification, protein turnover, chaperones
NODE_3904	jgi|Lotgi1|223645|	6E-36	AEQIR Myosin heavy chain, striated muscle (model%: 97, hit%: 26, score: 2152, %id: 76) [*Argopecten irradians*]	Motor activity; ATP binding	Cytoskeleton; Myosin class II heavy chain
NODE_1641	jgi|Lotgi1|224185|	2E-35	ATPase, H+ transporting, V0 subunit C [EC:3.6.3.14] [KO:K02155] (model%: 95, hit%: 96, score: 619, %id: 85) [*Xenopus tropicalis*]	46933 Hydrogen-transporting ATP synthase activity, rotational mechanism	Energy production and conversion
NODE_9524	jgi|Lotgi1|230842|	4E-33	Similar to myosin regulatory light polypeptide 9 isoform b (model%: 99, hit%: 98, score: 771, %id: 85) [*Danio rerio*]	Calcium ion binding	Cytoskeleton
NODE_530	jgi|Lotgi1|120098|	2E-32	Hypothetical protein SPCC63.09 - fission yeast (*Schizosaccharomyces pombe*) (model%: 64, hit%: 54, score: 117, %id: 61) [*Schizosaccharomyces pombe*]	NONE	NONE
NODE_294	jgi|Lotgi1|59092|	4E-29	Myosin heavy chain, striated muscle (model%: 100, hit%: 6, score: 538, %id: 80) [*Argopecten irradians*]	NONE	Cytoskeleton; Myosin class II heavy chain

NODE_28554	jgi|Lotgi1|223645|	8E-29	Myosin heavy chain, striated muscle (model%: 97, hit%: 26, score: 2152, %id: 76) [*Argopecten irradians*]	Motor activity; ATP binding	Cytoskeleton
NODE_17420	jgi|Lotgi1|230301|	1E-28	Ras-related C3 botulinum toxin substrate 1 (rho family, small GTP binding protein Rac1) [KO:K04392] (model%: 99, hit%: 100, score: 928, %id: 89) [*Homo sapiens*]	GTP binding	General function prediction only ; Ras-related small GTPase, Rho type
NODE_2451	jgi|Lotgi1|149096|	6E-28	H2A_SEPOF Histone H2A (model%: 97, hit%: 99, score: 600, %id: 96) [*Sepia officinalis*]	DNA binding	Chromatin structure and dynamics ; Histone 2A
NODE_10330	jgi|Lotgi1|106937|	2E-27	Ezrin/radixin/moesin [*Aplysia californica*] (model%: 99, hit%: 53, score: 1482, %id: 86) [*Aplysia californica*]	Cytoskeletal protein binding	General function prediction only ; Radixin, moesin and related proteins of the ERM family
NODE_648	jgi|Lotgi1|229535|	4E-23	QM-like protein [*Pinctada fucata*] >gi|37723970|gb|AAN85578.1| QM protein [*Pinctada fucata*] (model%: 96, hit%: 91, score: 979, %id: 85) [*Pinctada fucata*]	Structural constituent of ribosome	Translation, ribosomal structure and biogenesis
NODE_809	jgi|Lotgi1|138360|	1E-22	Mnk [*Aplysia californica*] (model%: 95, hit%: 86, score: 1518, %id: 60) [*Aplysia californica*]	Protein kinase activity	Signal transduction mechanisms
NODE_45091	jgi|Lotgi1|139452|	1E-22	Gene product from transcript CG5870-RA (model%: 99, hit%: 99, score: 7359, %id: 62) [*Drosophila melanogaster*]	Actin binding	Cytoskeleton
NODE_70315	jgi|Lotgi1|164867|	8E-22	SJCHGC00821 protein [S*chistosoma japonicum*] (model%: 92, hit%: 67, score: 488, %id: 67) [*Schistosoma japonicum*]	Calcium ion binding	Cytoskeleton
NODE_485	jgi|Lotgi1|158607|	2E-20	PREDICTED: filamin 1 (actin-binding protein-280) isoform 1 [*Macaca mulatta*] (model%: 99, hit%: 99, score: 6036, %id: 41) [*Macaca mulatta*]	Actin binding	Cytoskeleton
NODE_10743	jgi|Lotgi1|214616|	5E-20	NONE	NONE	Energy production and conversion; Prohibitins and stomatins of the PID superfamily

NODE_5130	jgi|Lotgi1|217342|	1E-19	Fructose-bisphosphate aldolase (model%: 98, hit%: 98, score: 1314, %id: 70) [*Echinococcus multilocularis*]	Fructose-bisphosphate aldolase activity	Carbohydrate transport and metabolism
NODE_4317	jgi|Lotgi1|150024|	2E-19	Chaperonin containing TCP1, subunit 5 (epsilon) (model%: 99, hit%: 99, score: 2236, %id: 79) [*Xenopus laevis*]	NONE	Posttranslational modification, protein turnover, chaperones
NODE_19714	jgi|Lotgi1|57638|	5E-19	RPGWamide precursor [*Mytilus edulis*] >gi|1588531|prf||2208453A APGWamide-related peptide precursor (model%: 81, hit%: 58, score: 641, %id: 40) [*Mytilus edulis*]	NONE	Posttranslational modification, protein turnover, chaperones Molecular chaperone (HSP90 family)
NODE_1509	jgi|Lotgi1|148309|	6E-19	Cytochrome P450 like_TBP [*Nicotiana tabacum*] (model%: 70, hit%: 6, score: 162, %id: 88) [*Nicotiana tabacum*]	NONE	NONE
NODE_17084	jgi|Lotgi1|115468|	1E-18	PREDICTED: similar to CG5871-PA [*Apis mellifera*] (model%: 97, hit%: 97, score: 2049, %id: 33) [*Apis mellifera*]	NONE	Posttranslational modification, protein turnover, chaperones ; Hyaluronoglucosaminidase
NODE_952	jgi|Lotgi1|230064|	2E-18	Glutamine synthetase [*Crassostrea gigas*] (model%: 99, hit%: 99, score: 1478, %id: 69) [*Crassostrea gigas*]	Glutamate-ammonia ligase activity	Amino acid transport and metabolism
NODE_22223	jgi|Lotgi1|214098|	6E-18	Mvp; major vault protein (model%: 96, hit%: 98, score: 3078, %id: 69) [*Danio rerio*]	Calcium ion binding	Posttranslational modification, protein turnover, chaperones
NODE_5618	jgi|Lotgi1|200105|	4E-17	NONE	NONE	Intracellular trafficking, secretion, and vesicular transport ; Flotillins
NODE_1361	jgi|Lotgi1|235056|	4E-17	Dual oxidase 1 [*Lytechinus variegatus*] (model%: 97, hit%: 95, score: 3902, %id: 41) [*Lytechinus variegatus*]	Oxidoreductase activity	Inorganic ion transport and metabolism
NODE_3978	jgi|Lotgi1|139286|	7E-17	Na-dependent Cl/HCO3 exchanger [*Loligo pealei*] (model%: 97, hit%: 82, score: 2476, %id: 45) [*Loligo pealei*]	Inorganic anion exchanger activity	Inorganic ion transport and metabolism

NODE_36141	jgi|Lotgi1|232873|	1E-16	Gene product from transcript CG30420-RA (model%: 76, hit%: 19, score: 362, %id: 64) [*Drosophila melanogaster*]		Signal transduction mechanisms ; Neural proliferation, differentiation and control protein
NODE_11042	jgi|Lotgi1|167517|	2E-16	Gene 11-1 protein precursor (model%: 86, hit%: 52, score: 501, %id: 2) [*Plasmodium falciparum 3D7*]	NONE	Function unknown ; Uncharacterized conserved protein
NODE_247	jgi|Lotgi1|207101|	1E-15	Eukaryotic translation initiation factor 4A, isoform 2 [KO:K03257] (model%: 93, hit%: 95, score: 1612, %id: 80) [*Bos taurus*]	Nucleic acid binding	Translation, ribosomal structure and biogenesis
NODE_13490	jgi|Lotgi1|190816|	4E-15	Glucose regulated protein 58kD [*Bos taurus*] >gi|729433|sp|P38657|PDIA3_BOVIN Protein disulfide-isomerase A3 precursor (Disulfide isomerase ER-60) (ERP60) (58 kDa microsomal protein) (P58) (ERp57) >gi|303524|dbj|BAA03760.1| PLC alpha [*Bos taurus*] >gi|1585552|prf||2201353A glucose-regulated protein ERp57/GRP58 (model%: 96, hit%: 94, score: 1354, %id: 52) [*Bos taurus*]	Isomerase activity, electron transporter activity	Posttranslational modification, protein turnover, chaperones
NODE_6879	jgi|Lotgi1|118685|	1E-14	CHICK Angiotensin-converting enzyme (Dipeptidyl carboxypeptidase I) (Kininase II) >gi|994708|gb|AAA75554.1| angiotensin converting enzyme (model%: 94, hit%: 90, score: 3357, %id: 54) [*Gallus gallus*]	Peptidyl-dipeptidase A activity; metallopeptidase activity	Amino acid transport and metabolism
NODE_3052	jgi|Lotgi1|198383|	1E-14	Myosin II heavy chain [*Ilyanassa obsoleta*] (model%: 73, hit%: 40, score: 877, %id: 83) [*Ilyanassa obsoleta*]	Motor activity	Cytoskeleton
NODE_2183	jgi|Lotgi1|197084|	2E-14	Tumor rejection antigen (gp96) 1 (model%: 97, hit%: 97, score: 2819, %id: 69) [*Gallus gallus*]	Chaperone activity	Posttranslational modification, protein turnover, chaperones

NODE_1106	jgi|Lotgi1|199050|	3E-14	Elongation factor 2 (EF-2) (model%: 99, hit%: 100, score: 3249, %id: 72) [*Drosophila melanogaster*]	Translation elongation factor activity	Translation, ribosomal structure and biogenesis
NODE_25076	jgi|Lotgi1|111250|	1E-13	Taurine transporter [*Mytilus galloprovincialis*] (model%: 94, hit%: 91, score: 2230, %id: 66) [*Mytilus galloprovincialis*]	Neurotransmitter:sodium symporter activity	Signal transduction mechanisms
NODE_6066	jgi|Lotgi1|139303|	2E-13	Na-dependent Cl/HCO3 exchanger [*Loligo pealei*] (model%: 97, hit%: 77, score: 2380, %id: 44) [*Loligo pealei*]	Inorganic anion exchanger activity	Inorganic ion transport and metabolism
NODE_3580	jgi|Lotgi1|137449|	5E-13	PREDICTED: similar to poly A binding protein, cytoplasmic 1 isoform 4 [*Tribolium castaneum*] (model%: 99, hit%: 100, score: 2023, %id: 55) [*Tribolium castaneum*]	Nucleic acid binding	RNA processing and modification
NODE_19641	jgi|Lotgi1|223645|	5E-13	Myosin heavy chain, striated muscle (model%: 97, hit%: 26, score: 2152, %id: 76) [*Argopecten irradians*]	Motor activity	Cytoskeleton
NODE_358	jgi|Lotgi1|167500|	7E-13	Ribosomal protein L11 [*Argopecten irradians*] (model%: 99, hit%: 100, score: 706, %id: 78) [*Argopecten irradians*]	Structural constituent of ribosome	Translation, ribosomal structure and biogenesis
NODE_12312	jgi|Lotgi1|213568|	1E-12	PONPY ATP-dependent RNA helicase DDX3Y (DEAD box protein 3, Y-chromosomal) (model%: 99, hit%: 87, score: 1942, %id: 67) [*Pongo pygmaeus*]	Nucleic acid binding	RNA processing and modification
NODE_1180	jgi|Lotgi1|65952|	2E-12	Myosin heavy chain, striated muscle (model%: 100, hit%: 15, score: 1245, %id: 76) [*Argopecten irradians*]	Motor activity	Cytoskeleton
NODE_34129	jgi|Lotgi1|233411|	1E-11	Hypothetical LOC496448 [KO:K04646] (model%: 98, hit%: 99, score: 7368, %id: 82) [*Xenopus tropicalis*]	NONE	Intracellular trafficking, secretion, and vesicular transport

NODE_66616	jgi|Lotgi1|231565|	1E-11	Hypothetical protein MGC69492 (model%: 99, hit%: 100, score: 2165, %id: 78) [*Xenopus tropicalis*]	NONE	Posttranslational modification, protein turnover, chaperones
NODE_8171	jgi|Lotgi1|225089|	3E-11	NONE	NONE	NONE
NODE_573	jgi|Lotgi1|230588|	4E-11	Similar to solute carrier family 5 (sodium/glucose cotransporter), member 9 (model%: 98, hit%: 95, score: 1862, %id: 53) [*Bos taurus*]	Transporter activity	Inorganic ion transport and metabolism
NODE_5602	jgi|Lotgi1|219848|	1E-10	STRIE Calmodulin (CaM) (model%: 92, hit%: 97, score: 709, %id: 92) [*Strongylocentrotus intermedius*]	Calcium ion binding	Signal transduction mechanisms
NODE_26721	jgi|Lotgi1|105757|	2E-10	Gelsolin [*Suberites ficus*] (model%: 85, hit%: 63, score: 876, %id: 67) [*Suberites ficus*]	Actin binding	Cytoskeleton
NODE_2586	jgi|Lotgi1|233095|	2E-10	BIOGL Tropomyosin-1 (TMI) (Bg 39) (model%: 99, hit%: 100, score: 1173, %id: 82) [*Biomphalaria glabrata*]	NONE	Cytoskeleton
NODE_1783	jgi|Lotgi1|203016	2E-10	Putative lipoprotein [*Pseudoalteromonas tunicata* D2] >gi|88817438|gb|EAR27255.1| putative lipoprotein [*Pseudoalteromonas tunicata* D2] (model%: 89, hit%: 93, score: 852, %id: 20) [*Pseudoalteromonas tunicata* D2]	NONE	NONE
NODE_12302	jgi|Lotgi1|94313|	4E-10	PREDICTED: similar to Glycogenin CG9480-PB, isoform B [*Apis mellifera*] (model%: 97, hit%: 48, score: 967, %id: 52) [*Apis mellifera*]	Transferase activity, transferring hexosyl groups	Carbohydrate transport and metabolism
NODE_4383	jgi|Lotgi1|196756|	8E-10	Ribophorin I [EC:2.4.1.119] [KO:K00730] (model%: 94, hit%: 95, score: 1666, %id: 54) [*Xenopus tropicalis*]	DNA binding	Posttranslational modification, protein turnover, chaperones

Table 2.12 Collection of 42 sequences highly expressed in S with significant ($E<10^{-10}$) similarity sequences in the *Lottia gigantea* database

Contig Name	Best Match LOTTIA	E Value	Protein Function best hit	Transcript Function description	KOG classification
NODE_3734	jgi|Lotgi1|204198|	3E-84	PATVU Tubulin alpha-2/alpha-4 chain (model%: 99, hit%: 100, score: 2369, %id: 98) [*Patella vulgata*]	GTPase activity	Cytoskeleton
NODE_639	jgi|Lotgi1|206609|	2E-66	Calreticulin [KO:K08057] (model%: 98, hit%: 96, score: 1622, %id: 71) [*Mus musculus*]	Chaperone activity	Posttranslational modification, protein turnover, chaperones
NODE_6009	jgi|Lotgi1|141813|	1E-33	RAT S-adenosylmethionine synthetase isoform type-2 (Methionine adenosyltransferase 2) (AdoMet synthetase 2) (Methionine adenosyltransferase II) (MAT-II) (model%: 93, hit%: 96, score: 1649, %id: 77) [*Rattus norvegicus*]	Methionine adenosyltransferase activity	Coenzyme transport and metabolism
NODE_6094	jgi|Lotgi1|207422|	4E-33	Selenophosphate synthetase [*Bombyx mori*] >gi|76496248|gb|ABA43640.1| selenophosphate synthetase [*B. mori*] (model%: 97, hit%: 93, score: 1527, %id: 74) [no tax name]	Catalytic activity	Signal transduction mechanisms
NODE_47013	jgi|Lotgi1|177987|	5E-30	Aldehyde dehydrogenase 2 family (mitochondrial) (model%: 99, hit%: 95, score: 1943, %id: 72) [*Gallus gallus*]	Oxidoreductase activity	Energy production and conversion
NODE_769	jgi|Lotgi1|223645|	5E-29	MYS_AEQIR Myosin heavy chain, striated muscle (model%: 97, hit%: 26, score: 2152, %id: 76) [*Argopecten irradians*]	Motor activity	Cytoskeleton
NODE_2429	jgi|Lotgi1|148309|	4E-26	Cytochrome P450 like_TBP [*Nicotiana tabacum*] (model%: 70, hit%: 6, score: 162, %id: 88)	NONE	NONE
NODE_2567	jgi|Lotgi1|65952|	2E-24	Myosin heavy chain, striated muscle (model%: 100, hit%: 15, score: 1245, %id: 76) [*Argopecten irradians*]	Motor activity	Cytoskeleton
NODE_6312	jgi|Lotgi1|205491|	3E-22	RAT Tubulin alpha-1 chain (Alpha-tubulin 1) >gi|55977473|sp|P68362|TBA2_CRIGR Tubulin alpha-2 chain (Alpha-tubulin 2) (Alpha-tubulin II) >gi|55977479|sp|P68369|TBA1_MOUSE Tubulin alpha-1 chain (Alpha-tubulin 1) (Alpha-tubulin isotype M-alpha-1) >gi|55977864|sp|Q71U36|TBA3_HUMAN Tubulin alpha-3 chain (Alpha-tubulin 3) (Tubulin B-alpha-1) (model%: 99, hit%: 100, score: 2368, %id: 97) [*Rattus norvegicus*]	GTPase activity	Cytoskeleton

NODE_24185	jgi|Lotgi1|211667|	2E-21	Tubulin alpha-2/alpha-4 chain (model%: 99, hit%: 100, score: 2369, %id: 98) [*Patella vulgata*]	GTPase activity	Cytoskeleton
NODE_69787	jgi|Lotgi1|233830|	2E-19	Eukaryotic translation initiation factor 2, subunit 2 beta, 38kDa (model%: 99, hit%: 99, score: 889, %id: 57) [*Gallus gallus*]	Translation initiation factor activity	Translation, ribosomal structure and biogenesis
NODE_37761	jgi|Lotgi1|209056|	3E-18	Heat shock protein 70 B2 (model%: 99, hit%: 99, score: 2559, %id: 76) [*Anopheles albimanus*]	NONE	Posttranslational modification, protein turnover, chaperones
NODE_117074	jgi|Lotgi1|237425|	1E-16	Hypothetical protein [*Cleome spinosa*] (model%: 31, hit%: 42, score: 424, %id: 19) [*C. spinosa*]	NONE	Carbohydrate transport and metabolism
NODE_38112	jgi|Lotgi1|215397|	2E-16	LUMRU Calmodulin (CaM) (model%: 99, hit%: 100, score: 726, %id: 93) [*Lumbricus rubellus*]	Calcium ion binding	Signal transduction mechanisms
NODE_30051	jgi|Lotgi1|239568|	3E-16	Generic methyltransferase [*Geobacter uraniumreducens* Rf4] >gi|88917456|gb|EAR36650.1| generic methyltransferase [*G. uraniumreducens* Rf4] (model%: 77, hit%: 76, score: 1186, %id: 27) [*Geobacter uraniumreducens* Rf4]	NONE	Function unknown
NODE_56300	jgi|Lotgi1|157973|	4E-16	Hydroxysteroid (17-beta) dehydrogenase 4 (model%: 99, hit%: 100, score: 2196, %id: 58) [*Gallus gallus*]	Oxidoreductase activity	Lipid transport and metabolism
NODE_5974	jgi|Lotgi1|233095|	3E-15	BIOGL Tropomyosin-1 (TMI) (Bg 39) (model%: 99, hit%: 100, score: 1173, %id: 82) [*Biomphalaria glabrata*]	NONE	Cytoskeleton
NODE_1709	jgi|Lotgi1|198383|	1E-14	Myosin II heavy chain [*Ilyanassa obsoleta*] (model%: 73, hit%: 40, score: 877, %id: 83) [*I.obsoleta*]	Motor activity	Cytoskeleton
NODE_10526	jgi|Lotgi1|135551|	4E-14	AEQIR Myosin heavy chain, striated muscle >gi|5612|emb|CAA39247.1| myosin heavy chain [*Argopecten irradians*] (model%: 99, hit%: 43, score: 3381, %id: 70) [*A. irradians*]	Motor activity	Cytoskeleton
NODE_55030	jgi|Lotgi1|197134|	6E-14	Cellulase [*Haliotis discus hannai*] (model%: 95, hit%: 76, score: 1709, %id: 69) [*H. discus hannai*]	Hydrolase activity, hydrolyzing O-glycosyl compounds	Carbohydrate metabolism
NODE_106626	jgi|Lotgi1|234726|	1E-13	Peritrophin membrane protein 1 [*Spodoptera frugiperda*] (model%: 63, hit%: 59, score: 304, %id: 12)	Chitin binding	Carbohydrate transport and metabolism
NODE_20977	jgi|Lotgi1|212802|	2E-13	Gi|70909593|emb|CAJ17217.1| ribosomal protein S21e [*Sphaerius* sp. APV-2005] (model%: 98, hit%: 100, score: 315, %id: 69) [*Sphaerius* sp. APV-2005]	Structural constituent of ribosome	Translation, ribosomal structure and biogenesis

NODE_30588	jgi|Lotgi1|202942|	2E-13	TO71-3 [*Taraxacum officinale*] (model%: 50, hit%: 38, score: 114, %id: 88) [*Taraxacum officinale*]	NONE	NONE
NODE_110795	jgi|Lotgi1|112980|	3E-13	PREDICTED: arylsulfatase B isoform 6 [*Macaca mulatta*] (model%: 92, hit%: 87, score: 1155, %id: 36) [*Macaca mulatta*]	Sulfuric ester hydrolase activity	General function prediction only ; Sulfatase
NODE_117952	jgi|Lotgi1|158430|	5E-13	Cellulase [*Haliotis discus*] (model%: 95, hit%: 95, score: 1273, %id: 40) [*Haliotis discus*]	Hydrolase activity, hydrolyzing O-glycosyl compounds; carbohydrate metabolism	NONE
NODE_108864	jgi|Lotgi1|112980|	7E-13	PREDICTED: arylsulfatase B isoform 6 [*Macaca mulatta*] (model%: 92, hit%: 87, score: 1155, %id: 36) [*Macaca mulatta*]	Sulfuric ester hydrolase activity	General function prediction only ; Sulfatase
NODE_12520	jgi|Lotgi1|190317|	8E-13	Acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain [*Danio rerio*] >gi|42542450|gb|AAH66533.1| Acyl-Coenzyme A dehydrogenase, C-4 to C-12 straight chain [*Danio rerio*] (model%: 94, hit%: 92, score: 1610, %id: 75) [*Danio rerio*]	Oxidoreductase activity	Lipid transport and metabolism
NODE_116318	jgi|Lotgi1|217885|	9E-13	Hypothetical LOC496744 [EC:3.2.1.51] [KO:K01206] (model%: 94, hit%: 92, score: 1520, %id: 59) [*Xenopus tropicalis*]	Alpha-L-fucosidase activity	Carbohydrate transport and metabolism
NODE_20792	jgi|Lotgi1|202942|	1E-12	TO71-3 [*Taraxacum officinale*] (model%: 50, hit%: 38, score: 114, %id: 88) [*Taraxacum officinale*]	NONE	NONE
NODE_60522	jgi|Lotgi1|105466|	1E-12	Cellulase [*Haliotis discus*] (model%: 96, hit%: 97, score: 1973, %id: 61) [*Haliotis discus*]	Hydrolase activity, hydrolyzing O-glycosyl compounds	NONE
NODE_12559	jgi|Lotgi1|123611|	2E-12	Thioredoxin peroxidase [*Apis mellifera ligustica*] (model%: 98, hit%: 68, score: 676, %id: 72) [*Apis mellifera ligustica*]	NONE	Posttranslational modification, protein turnover, chaperones
NODE_84703	jgi|Lotgi1|133359|	2E-12	Mandelate racemase/muconate lactonizing enzyme:Mandelate racemase/muconate lactonizing enzyme [*Azotobacter vinelandii* AvOP] >gi|67088342|gb|EAM07808.1| Mandelate racemase/muconate lactonizing enzyme:Mandelate racemase/muconate lactonizing enzyme [*Azotobacter vinelandii* AvOP] (model%: 92, hit%: 61, score: 1126, %id: 51) [*Azotobacter vinelandii* AvOP]	Catalytic activity	NONE

NODE_11805	jgi|Lotgi1|184532|	3E-12	S-adenosylhomocysteine hydrolase [*Danio rerio*] >gi|28278417|gb|AAH44200.1| S-adenosylhomocysteine hydrolase [*Danio rerio*] >gi|37681725|gb|AAQ97740.1| S-adenosylhomocysteine hydrolase [*Danio rerio*] (model%: 99, hit%: 98, score: 1870, %id: 80) [*Danio rerio*]	Adenosylhomocysteinase activity	Coenzyme transport and metabolism
NODE_35051	jgi|Lotgi1|112613|	4E-12	Gi|61742178|gb|AAX54912.1| antiquitin [*Acanthopagrus schlegelii*] (model%: 99, hit%: 100, score: 2022, %id: 72) [*Acanthopagrus schlegelii*]	Oxidoreductase activity	Energy production and conversion
NODE_115305	jgi|Lotgi1|132347|	1E-11	Cellulase [*Haliotis discus*] (model%: 92, hit%: 93, score: 1531, %id: 45) [*Haliotis discus*]	Hydrolase activity, hydrolyzing O-glycosyl compounds	NONE
NODE_2784	jgi|Lotgi1|192810|	2E-11	Aldehyde dehydrogenase class 1 [*Xenopus laevis*] (model%: 99, hit%: 98, score: 1784, %id: 65) [*Xenopus laevis*]	Oxidoreductase activity	Energy production and conversion
NODE_79571	jgi|Lotgi1|114783|	3E-11	Beta 1,3-glucanase [*Strongylocentrotus purpuratus*] >gi|1488257|gb|AAC47235.1| beta 1,3-glucanase (model%: 97, hit%: 88, score: 1270, %id: 47) [*Strongylocentrotus purpuratus*]	NONE	NONE
NODE_13416	jgi|Lotgi1|238461|	1E-10	Unnamed protein product [*Tetraodon nigroviridis*] (model%: 99, hit%: 100, score: 942, %id: 49) [*Tetraodon nigroviridis*]	NONE	Translation, ribosomal structure and biogenesis
NODE_113634	jgi|Lotgi1|209552|	2E-10	RAT Succinate semialdehyde dehydrogenase (NAD(+)-dependent succinic semialdehyde dehydrogenase) (model%: 96, hit%: 98, score: 1559, %id: 61) [*Rattus norvegicus*]	Oxidoreductase activity	Energy production and conversion
NODE_1117	jgi|Lotgi1|178121|	3E-10	Similar to phosphoenolpyruvate carboxykinase 2 (mitochondrial) [EC:4.1.1.32] [KO:K01596] (model%: 99, hit%: 94, score: 2262, %id: 68) [*Xenopus laevis*]	Phosphoenolpyruvate carboxykinase activity	Energy production and conversion
NODE_11725	jgi|Lotgi1|67491|	5E-10	Dienelactone hydrolase family protein (model%: 100, hit%: 71, score: 489, %id: 54) [*Arabidopsis thaliana*]	Hydrolase activity	General function prediction only; Predicted hydrolase related to dienelactone hydrolase

Table 3.8 Recipes of solutions used during cell culture procedures

Alsever solution (personal communication S Auzoux-Bordenave, 2007)	
NaCl	5.6 g
Glucose	5.2 g
Sodium citrate	2 g
Penicillin/Streptomycin	2.5 ml
Dissolve salts in 250 ml milliq H_2O, add Penicillin/Streptomycin. Set pH 7.5. Filter through 0.2 µm into sterilized bottle	
Antibiotic Wash solution (AB wash) (personal communication S Auzoux-Bordenave, 2007)	
Filtered Seawater	483.5 ml
PenStrep	2 %
Gentamycin	250 µg/ml
Amphotericine B	2 µg/ml
Adjust pH t0 7.6. Filter through 0.2µm into sterilized bottle	
Dissociation solution (DS) (personal communication S Auzoux-Bordenave, 2007)	
Glucose (1 mg/ml)	250 mg
Filtered Seawater	250 ml
Penicillin/Streptomycin (1 %)	2.5 ml
Adjust pH to 7.4. Filter through 0.2 µm.	
Isotonic Invertebrate Solution (ISO) (Cima *et al*., 2000)	
Tris	2.82 g/L
NaCl	29.25 g/L
Adjust pH to 7·4	
Osmolarity±1000 mmol/kg	
Calcium Magnesium free Artificial Seawater (CMFSS) (personal communication N Odintsova, 2007)	
NaCl	25.5 g/L
KCL	0.8 g/L
Na_2HPO_4	3.0 g/L
Glucose	3.0 g/L
HEPES	2.86 g/L
Adjust pH to 7.4. Filter through 0.2 µm.	
Osmolarity ± 900 mmol/kg	

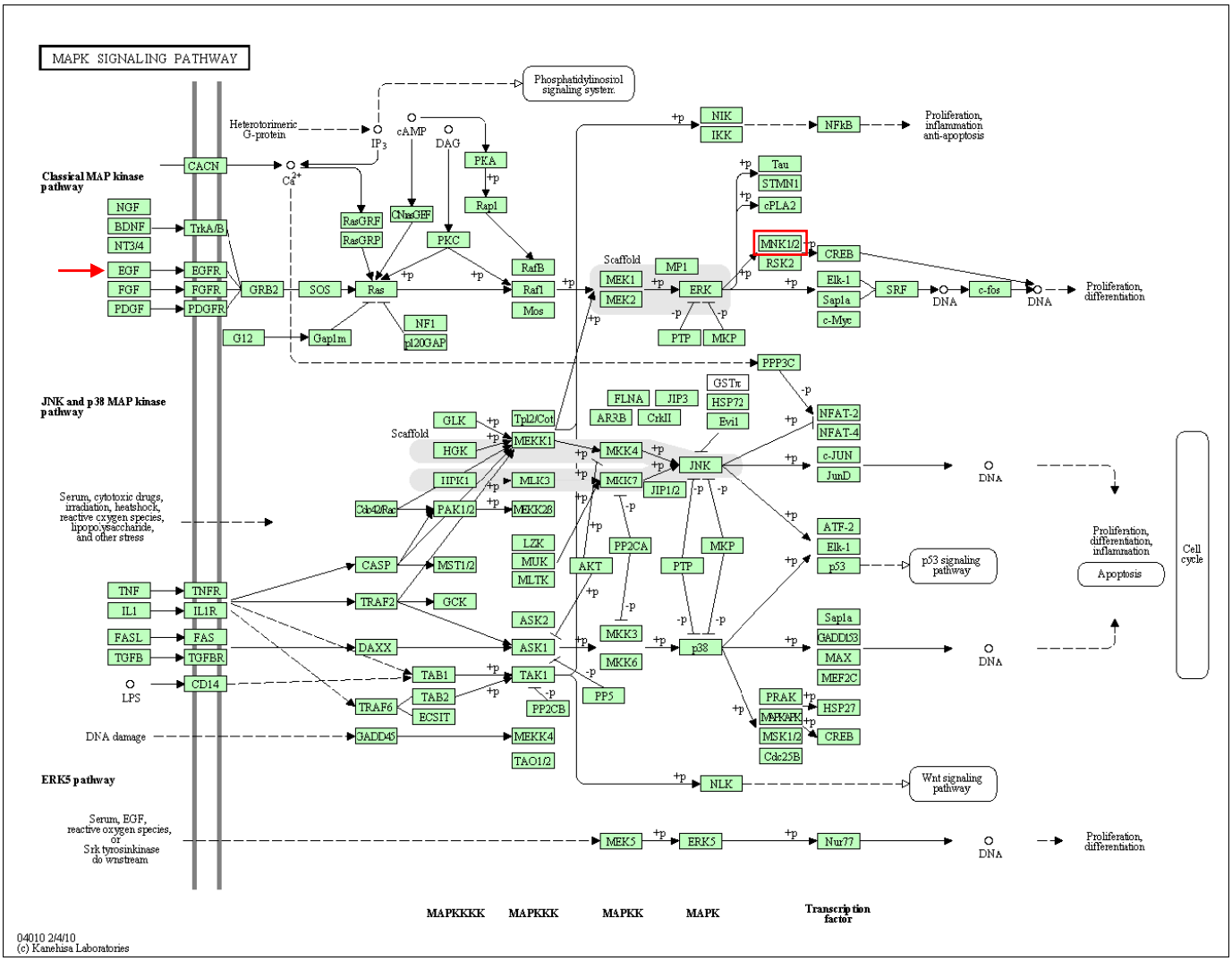


Figure 4.10b MAP-Kinase signaling pathway (Kanehisa, 2010). Activation by EGF binding is indicated with a red arrow and Mnk is boxed in red

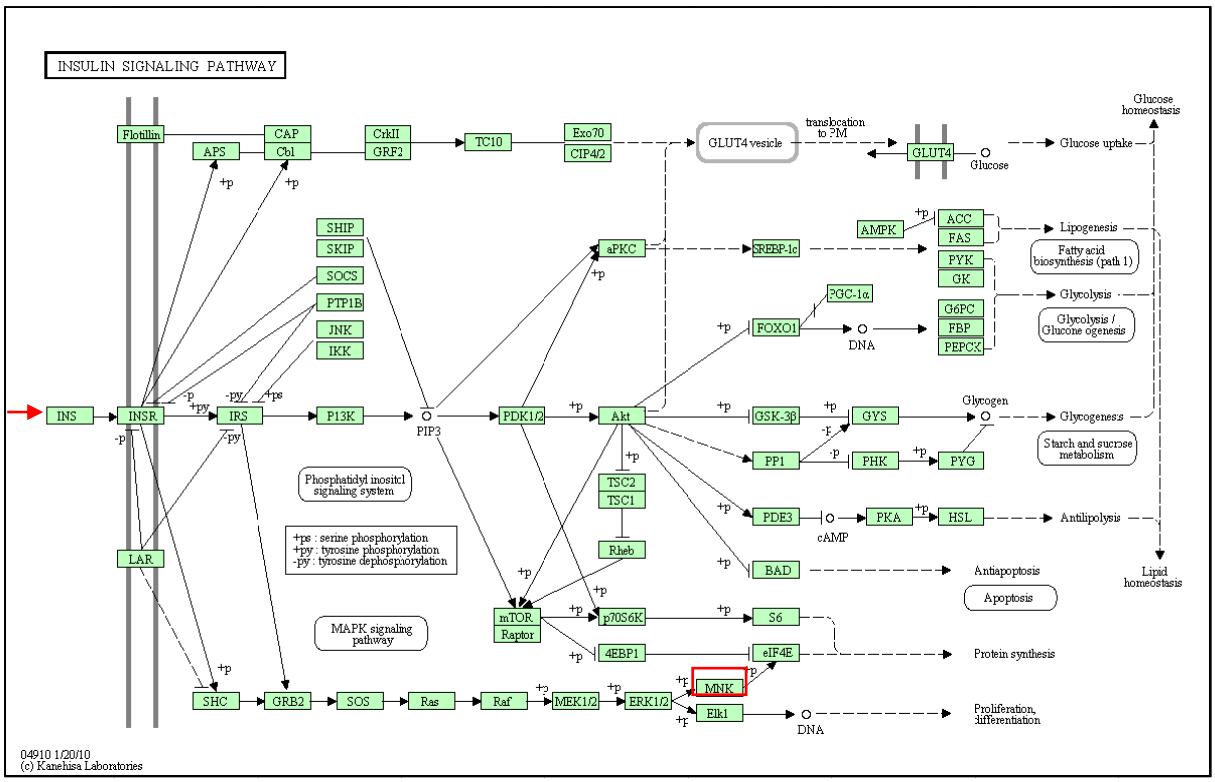


Figure 4.11b Insulin signaling pathway (Kanehisa, 2010). Activation by insulin binding is indicated with a red arrow and Mnk is boxed in red

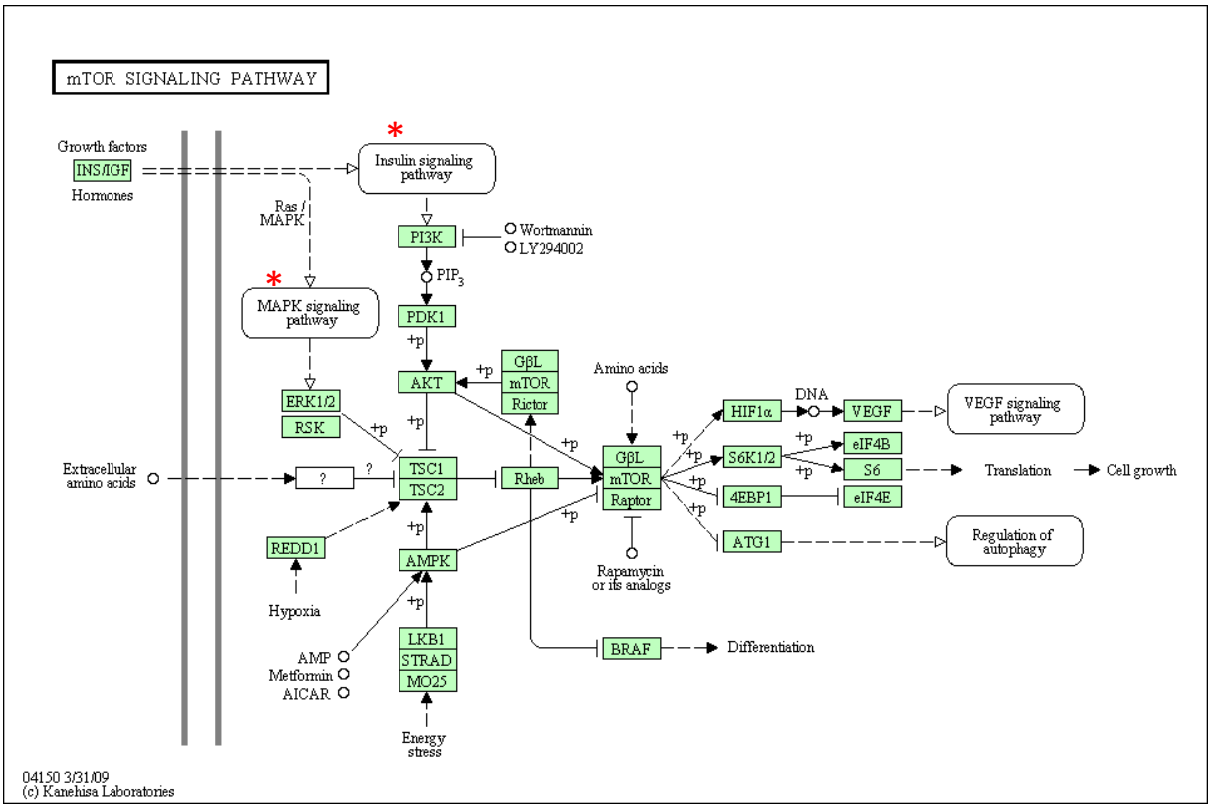


Figure 5.1 mTOR signaling pathway (Kanehisa, 2010). Interaction with MAPK and Insulin signaling pathways is indicated with a red asterisk