Prevention and control of maturation to address multiple key abalone production constraints

Natasha Botwright, Michelle Colgrave, Mathew Cook and Nick Elliott

Project No. [2010/767]

30/4/2014

National Research
FLAGSHIPS
Food Futures

CSIRO

This project was conducted by CSIRO National Food Futures Flagship

ISBN: 978-1-4863-0320-5

Copyright, 2014: The Seafood CRC Company Ltd, the Fisheries Research and Development Corporation and CSIRO Food Futures National Research Flagship.

This work is copyright. Except as permitted under the Copyright Act 1968 (Cth), no part of this publication may be reproduced by any process, electronic or otherwise, without the specific written permission of the copyright owners. Neither may information be stored electronically in any form whatsoever without such permission.

The Australian Seafood CRC is established and supported under the Australian Government's Cooperative Research Centres Program. Other investors in the CRC are the Fisheries Research and Development Corporation, Seafood CRC company members, and supporting participants.

Office Mark Oliphant Building, Laffer Drive, Bedford Park SA 5042
Postal Box 26, Mark Oliphant Building, Laffer Drive, Bedford Park SA 5042
Tollfree 1300 732 213 Phone 08 8201 7650 Facsimile 08 8201 7659
Website www.seafoodcrc.com ABN 51 126 074 048

Important Notice
Although the Australian Seafood CRC has taken all reasonable care in preparing this report, neither the Seafood CRC nor its officers accept any liability from the interpretation or use of the information set out in this document. Information contained in this document is subject to change without notice.

Australian Government
Fisheries Research and Development Corporation

An Australian Government Initiative

Non-Technical Summary

2010/767: Prevention and control of maturation to address multiple key abalone production constraints

PRINCIPAL INVESTIGATOR: Natasha Botwright

ADDRESS: **CSIRO Animal, Food and Health Sciences**
Queensland Bioscience Precinct
306 Carmody Road
St Lucia QLD 4067

PROJECT OBJECTIVES:

1. To establish molluscan specific neuropeptide databases that will underpin abalone maturation tissue studies

2. Develop reliable sample collection, preservation, laboratory processing of abalone maturation tissues to ensure optimal analyses by mass spectrometry

3. To undertake mass spectroscopy and bioinformatic analyses of samples to map neuropeptides throughout the maturation cycle in different tissues and stages of maturation

4. Select candidates that may be useful for prevention and control of maturation for validation

5. Select candidates responsible for influencing decreased gamete condition and spawning success over successive seasons

6. Undertake development of a non-destructive sampling technique for monitoring of animal responses to experimental intervention strategies

OUTCOMES ACHIEVED

Key to achieving the outcomes of this project was the development herein of significant abalone sequence resources. Consolidation of these with previously publically available genome resources to develop the molluscan theoretical protein database improved the identification success up to four fold. Sample collection, extraction and mass spectroscopy methodologies in combination with the novel pipeline developed specifically for this project enabled the core focus to be only on those peptides suspected to be involved in the gonad maturation process.

A shortlist of 25 maturation related candidates were identified and profiled in response to increasing visual gonad index. Of these, eight were identified and known to be involved in maturation related processes. Sequence information is available for a further seven peptides. The remaining 10 have mass, charge and retention time data with the mass spectral data available for reinterrogation as further genetic resources become available.

Approximately half of the peptide candidates were able to be profiled in younger and older abalone to determine those likely to be responsible for compromised gamete and spawning success. Of these, five are likely to have inhibitory effects through enhanced expression. A further two are suspected of being expressed at insufficient

levels to promote successful gamete conditioning and spawning.

The high rate of degradation of signalling peptides involved in biological processes negates the development of a non-destructive sampling method for the future monitoring of interventional strategies. However knowledge of the sequence of active biological maturation related peptides provides the opportunity to conduct functional studies to progress towards an interventional strategy to promote or inhibit gonad condition as required in commercial or breeding stock.

LIST OF OUTPUTS PRODUCED

Molluscan theoretical protein and peptide databases based on abalone and closely related mollusc species for identification, quantification and monitoring of key maturation/reproductive neuropeptides in temperate abalone.

Abalone neuropeptides homologous to invertebrate species identified.

Novel methods for analysis and interpretation of mass spectroscopy generated data for species with insufficient genetic knowledge for traditional analyses.

Short list of candidates involved in gonad maturation processes and potentially responsible for influencing gamete condition and spawning success in abalone.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge the following individuals or groups for various forms of assistance and collaboration during this and associated research, and apologise to anyone we have inadvertently omitted.

Harry King, Richard Taylor - CSIRO Marine and Atmospheric Research

Leanne Dierens, John Hoad, Sean McWilliam, Eddie Parker, Tony Reverter, Lynette Williams, Gene Wijffels - CSIRO Animal, Food and Health Sciences

Ondrej Hlinka - CSIRO Information Management and Technology

Scott Cummins - University of the Sunshine Coast

James Harris - Flinders University

Justin Fromm, Dan Machin, Nick Savva - Australian Abalone Growers Association

David Connell, Daniel Price, Justin Sanguinetti - Kangaroo Island Abalone Pty Ltd

Adam Clark, Anton Krsinich, Luke McPherson - Great Southern Waters Pty Ltd

Nick Savva - Australian Ocean Biotechnology Pty Ltd

CSIRO Food Futures National Research Flagship

1. Introduction and Background

Maturation was associated with several key priority R&D areas identified by the Australian Abalone Growers Association in 2009. These are summer mortality, live transport, controlling sex, maturation and spawning. Industry estimated the cost benefit of controlling maturation in the key production areas of health, reduced production, broodstock conditioning and product quality as 5-10%, 5-10%, 1-5% and 5-10% respectively.

At the current time there is a dearth of publically available molecular knowledge on abalone reproduction. This fundamental knowledge is essential to determine effective candidates and timing for the greatest success of interventional methods. In the absence of an abalone whole genome sequence, the use of liquid chromatography/ mass spectroscopy (LC/MS) allows for the rapid identification and quantification of peptides in maturation tissues throughout the maturation cycle. In this instance, the peptides of interest are hormones which are responsible for control of reproduction through signalling and response mechanisms.

This project will advance our understanding of maturation pathways in temperate abalone to enable:

1. Future development of strategies to prevent maturation in commercial stock resulting in improved abalone health, increased production and confidence in product quality. This will result in more consistent quality and supply of abalone to market and increased profitability.

2. Tighter control of maturation in breeding stock resulting in synchronised broodstock conditioning, improved seasonal reconditioning of aging broodstock and on-demand spawning technologies. This will result in more reliable conditioning of highly valued 'elite' selected broodstock. Combining these benefits with reliable spawning methodologies will enable the greatest gains from selective breeding programs particularly with the advent of closed breeding cycles in operation on most farms.

The primary goal of this project is to provide a foundation of knowledge on maturation. Two key outcomes of the methodologies proposed herein are to identify critical maturation time points and identify key molecules involved in and/or responsible for maturation. This will provide a basis for the future development and implementation of commercial scale technologies for use within the industry.

1.1 Need

Maturation is a complex issue with key abalone production issues in:

- Health

There is a concordance between maturation and the spawning process, immune status and pathogen susceptibility (Travers *et al.,* 2008). These factors may combine with elevated water temperatures contributing to summer mortalities.

- Reduced production

Gonad is commercially undesirable as this results in lost opportunity for conversion of feed to growth. Reduced profitability occurs as a result of increases in gonad weight where frozen or canned abalone products are based on meat weight and through

decreases in gonad weight due to pre-harvest spawning of live product sold on whole weight.

- Broodstock conditioning

Most farms now utilise closed life cycle breeding. This relies on the use of mass selected or genetically selected broodstock that are ready for synchronised spawning on demand. To make the greatest gains the industry needs to breed from select elite individuals rather than from a select group of abalone. The industry is not able to select an individual with 100% confidence for spawning.

- Product Quality

Animal health provides industry with confidence in their product all year. Currently the industry has reduced confidence in the quality of live abalone exports over the summer period. The fragile physiology of spawning and maturing animals and the stress involved in transport and holding leads to increased mortality risks during transport. Consequently some exporters reduce or cease exports during this time.

The ability to prevent maturation will provide the Australian temperate abalone industry with a competitive edge by reducing mortality, increasing growth and improving delivery of live product to market. Controlling maturation will assist in breeding. The key to achieving these outcomes is to first gain an understanding of the maturation process at the molecular level. This will provide the knowledge to initiate development of new technologies and/or interventional strategies to address these key industry issues.

1.2 Objectives

Each of the following objectives has been achieved during the course of this project.

1. To establish molluscan specific neuropeptide databases that will underpin abalone maturation tissue studies

2. Develop reliable sample collection, preservation, laboratory processing of abalone maturation tissues to ensure optimal analyses by mass spectrometry

3. To undertake mass spectroscopy and bioinformatic analyses of samples to map neuropeptides throughout the maturation cycle in different tissues and stages of maturation

4. Select candidates that may be useful for prevention and control of maturation for validation

5. Select candidates responsible for influencing decreased gamete condition and spawning success over successive seasons

6. Undertake development of a non-destructive sampling technique for monitoring of animal responses to experimental intervention strategies

2. Methods

2.1 Development of molluscan specific protein databases to underpin maturation tissue studies

A theoretical mollusan protein database consisting of nine smaller subsets based on publically and privately acquired data have been developed. In brief, greenlip ganglia cDNA libraries were constructed using In-Fusion SMARTer Directional cDNA Library Construction Kit (Clontech) as per the manufacturer's instructions. Sequencing was carried out at Macrogen Inc. (Korea). RNA was extracted using the Trizol (Life Technologies) method as per the manufacturer's instructions from greenlip ganglia and gonad samples for transcriptome sequencing by BGI Tech Solutions (Hong Kong) Co (China). Additional sequences for a custom built abalone database were obtained from closely related *Haliotis* species such as *H. rufescens*, *H. midae* and *H. asinina*. Other publically available sequence resources used for database construction include NCBI and UniProt. Bioinformatic analyses were carried out using the Galaxy bioinformatics platform (Blankenberg *et al.*, 2007) to predict and annotate theoretical proteins.

2.2 Abalone sample collection and preservation

Sampling of greenlip and hybrid abalone was undertaken at three sites in Australia beginning with juveniles pre-gonad development at ca. 15 months. Continuous sampling of the same cohort in South Australia and Victoria was undertaken every two months roughly coinciding with the end of the first spawning season. The preferred age limit was relaxed (ca. 27 months) for the Tasmanian cohorts as these were significantly behind their interstate counterparts in terms of development with approximately two times less weight gain at the same age. Phenotypic data collected at each sampling time point included water temperature, sex, weight, length and visual gonad index (VGI). Blacklip abalone of the required age were not available during the initial sampling phase therefore a subset based on VGI was collected at a later date.

Abalone were euthanized with 360 mM magnesium chloride to preserve peptide integrity. Similarly dissection of gonad, ganglia and muscle samples were carried out on freezer packs to preserve peptide integrity. Haemolymph was collected by allowing it to drain naturally out of the remaining tissue into a plastic bag on ice. All samples were stored in cryogenic vials and snap frozen in liquid nitrogen before being stored at -80°C for laboratory processing.

2.3 Abalone peptide extraction and mass spectroscopy assays

Samples were extracted individually in triplicate (biological replicates) for each sex at strategic time points in the maturation cycle as determined by the VGI. Ganglia and gonad tissue were thermally treated to denature proteolytic enzymes and peptides were extracted using a liquid-liquid acetic acid method previously described for mice (Svennson *et al.*, 2003) and cattle (Colgrave *et al.*, 2010) hypothalamus tissue. Samples were vacuum dried and stored at -20°C until analysis. The peptide extraction method used for ganglia and gonad tissues was modified for haemolymph. In brief, 1x Complete protease inhibitor cocktail (Roche) was added to 500 µl haemolymph and mixed gently. Samples were then centrifuged at 1000 g for 5 mins at 4°C to remove cellular debris. The supernatant was mixed 1:1 with 0.5% acetic acid and then centrifuged at 14 000 g for 30 mins at 4°C. Peptides smaller than 10 kDa were isolated by filtration using Microcon YM-10 filter devices (Millipore) and

centrifuged at 20 000 g for 90 min at 4°C. Samples were vacuum dried and stored at -20°C until analysis.

Peptides were chromatographically separated using a Shimadzu Prominence LC20 HPLC system with a ZORBAX 300SB-C18 column (75 µm x 15 cm, 3.5 µm) (Agilent Technologies). Samples were reconstituted in 1% formic acid at 2 µl/mg of tissue and then diluted 1:5 for a 20 µl on column injection volume. A linear gradient at a flow rate of 300 nl/min from 2-40% solvent B over 44 mins was utilised where solvent A was 0.1% formic acid and solvent B was 0.1% formic acid in 90% acetonitrile.

Eluent was directed into the nanoelectrospray ionisation source of the TripleTOF™ 5600 system (ABSciex). Data was acquired in two modes where MS analysis was performed in positive ion mode over the mass range *m/z* 300–1800 with a 0.5 s accumulation time. The ion spray voltage was set to 2400 V, the curtain gas was set to 25, the nebuliser gas to 12 and the heated interface was set to 180°C. Mass spectra were acquired over the mass range m/z 300-1800. The information-dependent acquisition method consisted of a high-resolution TOF-MS survey scan (as described above) followed by 20 MS/MS in the first stage. MS/MS spectra were analysed over the range *m/z* 300-1800 using rolling collision energy for optimum peptide fragmentation. Precursor ion masses were excluded for 8 s after two occurrences.

2.4 Bioinformatic selection of maturation related candidates

A novel bioinformatic pipeline was developed specifically for this project due to the paucity of sequence data available for this species. Differential analysis of samples was undertaken using MarkerView^{TM} 1.2.1.1 software (ABSciex) to profile and select candidate peptides predicted to be involved in maturation processes. In brief, mass spectral data was imported in experimental groups consisting of the same species, tissue source, sex and collection site. The minimum retention time for data processing was 10 mins. Peak finding was enhanced by using a subtraction offset of 10 scans, minimum spectral and retention time peak width of 5 ppm and 20 scans respectively and noise threshold of 5. Charge sites were automatically assigned by the software. Alignment of chromatograms had a retention time tolerance of 2 mins and a mass tolerance of 1 Da.

Peaks occurring in less than 2 samples were automatically removed. Samples were normalised using chromatogram total area sums (Rocchiccioli *et al.*, 2010). All peaks unassigned or with a charge state of 1 were removed from the analysis. Principle component analysis with logarithm weighting and Pareto scaling was used to identify samples which were not uniformly extracted, had less sample loaded on column and could be expected to be missing low level expression peptide peaks. Samples were compared using a t-test and profiles were then plotted of response against the visual gonad index of the sample. These were manually filtered to obtain candidate peptides of interest. Prior to data filtering a complete list prior of all peptides within a group was exported as a text file for later analyses with CrossMatch.

Peptides were identified by database searching of spectral datasets using ProteinPilot^{TM} 4.5 software (ABSciex) against the individual subsets of the theoretical molluscan protein database described previously. Search parameters were defined as no cysteine alkylation and no digestion enzyme. The ID focus was on biological modifications and thorough identifications. The resultant peptide identifications were imported into PeakView^{TM} 1.1.0.0 software (ABSciex) with the Protein QtoolDirector 1.0.0.0 plug-in for peak alignment. A complete list of peptides with confidence greater

than 80%, 50% and 1% for each group file was exported to MarkerView to be saved as a text file for manual curation against the candidate list.

A Perl 5.14.2 program, massMatchSeq was written to assign identity to maturation peptide candidates using candidate peptide lists and group database search files. In brief the process involved matching firstly to *m/z* values +/- 0.02 and then filtered based on retention time +/- 0.5 mins. The complete sequence associated with the unique database identifier is then retrieved from the original database file and added to the output. Submitting the sequence to NeuroPred (UIUC Centre for Neuroproteomics) to predict peptide cleavage sites (Hummon *et al*., 2003) enabled prediction and separation of degradation protein fragments from the active signalling peptides of interest.

A second Perl program, CrossMatch was written to check that all peptides potentially involved in maturation processes were selected in the original MarkerView candidate profiling. CrossMatch enabled automated comparison of the candidate list against each of the complete peptide lists prior to candidate filtering for each group of tissues. The output of this process provided data for graphical representation and interpretation of each of the candidates across species, sex and site. The graphs produced for each of the unknown peptides were also compared with each other to predict peptides likely to belong to the same protein. A shortlist of candidates involved in the maturation process was then extracted.

2.5 Candidate selection for decreased gamete conditioning and spawning success

A comparison of the 2009 year class greenlip abalone with the 2008 hybrids and 2007 greenlips was undertaken to select candidates from the short list that may influence decreased gamete condition and spawning success over successive seasons. Quantified response data for each of the peptides was extracted from each species complete mass spectra list if it was detected. Signal peptides are under constant degradation therefore all peptides matched to a specific protein were combined into a 'young' year class (2009) and an 'old' year class (2007 and 2008) for analysis. To observe general trends for each protein graphs were produced. The fold change for each peptide was calculated by subtracting VGI 0 data from VGI 3 and the average for young and old were used to select candidates where sufficient data was present.

3. Results

3.1 Spatial and temporal analysis

Samples were collected to assess both spatial and temporal variation to assist in producing a profile relevant to the majority of the Australian cultured abalone industry. The number of samples collected at each site per species is provided in Table 1.

Table 1 Number of samples collected to develop a proteomic/peptidomic profile of the physiological maturation process

Site	Species	Sex	Number of samples
South Australia	Greenlip	Unknown	80
		Male	146
		Female	116
Victoria	Hybrid	Unknown	60
		Male	132
		Female	130
	Blacklip	Male	25
		Female	24
Tasmania	Greenlip	Unknown	2
		Male	68
		Female	70
	Hybrid	Unknown	17
		Male	62
		Female	66

Weight (Fig. 1A) and length (Fig. 1B) data follow a similar trend to temperature (Fig. 1C) for sites in South Australia and Victoria. Weight and length data collected for Tasmanian cohorts are not shown as these were from older year classes. The visual VGI was recorded for each sample collected showing that the two hybrid collection sites exhibited peak VGIs at different times. We predict that the earlier maximum average at the Tasmanian site (Figure 1D) more closely resembles the spawning season commonly associated with blacklip abalone. The Victorian hybrid peak VGI appears to closely resemble that for greenlip samples collected from South Australia and Victoria (Figure 1E).

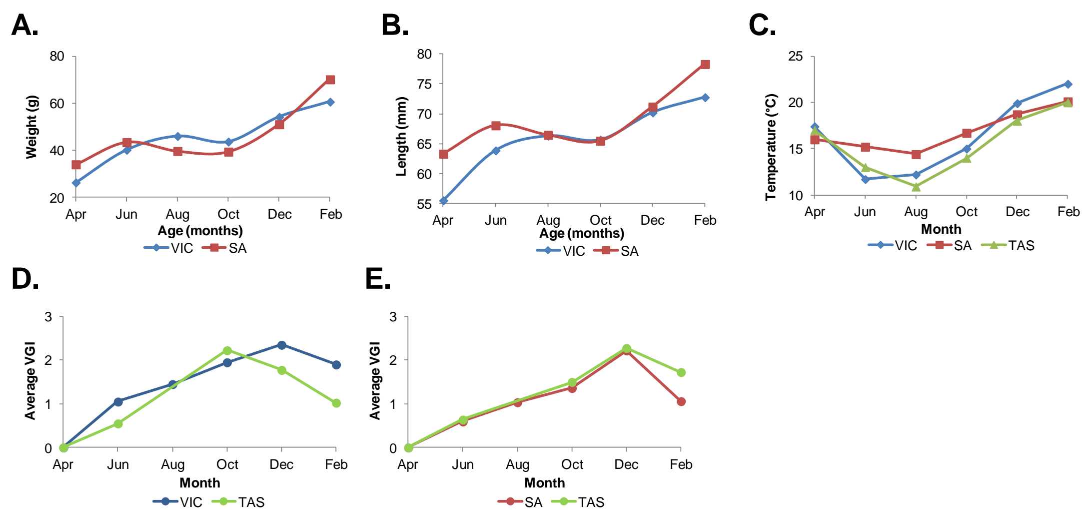

Figure 1 A – Average weight of animals sampled throughout the sampling period; B – Average length of animals sampled throughout the sampling period; C – Water temperature profile over the sample collection period; D – Comparison of the average visual gonad index (VGI) between hybrid samples collected; and E - Comparison of the average VGI between greenlip samples collected

3.2 Molluscan Peptide Databases

A total of nine molluscan peptide databases were developed specifically for this project for a total of 12,843,879 sequences from a variety of sources. Five databases were developed from greenlip ganglia and gonad tissue sourced during the collection phase incorporating annotated and unannotated sequence data contributing 64,346 sequences to the molluscan peptide databases. Annotated data allowed for sequence identification as well as providing evidence for the most likely function of proteins involved in the maturation process. This also enabled some proteins to be discarded from the final candidate list where their function was inconsistent with maturation related processes. Unannotated data enabled sequence identification however the function of these proteins remains unknown at the current time.

Two databases were developed from private sequence collections based on predicted open reading frames (gene coding protein sequence) provided a further 12,657,888 sequences. These sequence databases supplemented the 121,645 sequences sourced from publically available sequence collections. Molluscan peptide databases developed specifically for this project improved identification from 7% to 25% for ganglia, 13% to 44% for gonad and 1% to 10% for haemolymph candidates selected for their potential involvement in sexual maturation processes.

3.3 Mass spectroscopy validation of sample collection and processing techniques

Transcriptome data was searched against a dataset of molluscan neuropeptides to ascertain homology to other species. These homolog's were successfully identified in ganglia mass spectroscopy data validating the extraction method for this tissue source (Table 2). Those that were not identified may have been due to sequence dissimilarity. Alternatively homolog's may not have been present in the samples at the stage of development and therefore did not get sequenced. Fewer neuropeptides were able to be detected in gonad data negating the utility of mass spectroscopy for detection of signal molecules in this tissue. Signal molecule detection was obscured by detection of greater peptide levels of gonad tissue related molecules which are unable to be separated out during the extraction process.

Table 2 Molluscan neuropeptide homologues detected by mass spectral analysis of abalone ganglia tissue

Molluscan homologues detected		
Achatin	Enterin	NKY
Allatostatin	FCAP	Pedal
Allatropin	FFamide	PKYMDT
APGWamide	FMRFamide	Pleurin
Buccalin	Insulin	sCAP
Cerebrin	LASGL	Tachykinin
Conopressin	Lymnokinin	WWamide
Elevenin	Myomodulin	

Molluscan homologues not detected	
Bursicon	LFRYAMIDE
CCAP	LRNFVamide
ELH	Luqin
GGNG	NdWFamide
GnRH-like	NPY
GPA2	PXFVamide
GPB5	
LFRFFamide	

3.4 Maturation related preliminary candidate peptide selection

A total of 16 mass spectroscopy experiments were completed to enable maturation peptide profiles to be developed for ganglia and gonad tissues as well as haemolymph. Sample groups consisted of juveniles, males and females with a VGI of zero (no evidence of gonad development) through to three (gonad considered ripe for spawning induction). Greenlip and hybrid ganglia profile analysis results were consistent with known peptides involved in the maturation cycle which enabled

subsequent experiment numbers to be reduced (Table 3). Routine maintenance of the mass spectrometer over two and half years of experimental analysis resulted in variation in the number of peaks per sample group detected. These ranged from ~600 through to ~4000 peaks per experimental group (Table 3).

Table 3 Experiment summary for selection of maturation related peptide profiles of interest

Tissue	Spp.	Year Class	Sex	Number of samples	Peptide Profiles		
					Total	Maturation related	Identified
Ganglia	GL	2009	F	26	3061	74	25
			M	28	3074	61	24
		2007	F	9	2686	139	29
			M	9	2743	253	42
	HY	2009	F	28	3366	278	59
			M	22	4143	358	69
		2008	F	9	2877	243	39
			M	9	2912	186	42
	BL	2010	F	9	697	79	13
			M	9	521	66	15
Gonad	GL	2009	F	13	2375	133	59
			M	14	2257	142	39
	HY	2009	F	14	2543	101	41
			M	13	2478	95	31
Haemolymph	GL	2009	F	9	3676	125	7
			M	11	3747	138	22

GL – greenlip abalone; HY – hybrid abalone; BL – blacklip abalone; F – female; M - male

A novel bioinformatic pipeline developed specifically for this project outlined in Section 2.4 was used to select candidates associated with the sexual maturation process. The number of candidates in the preliminary stages of analysis varied between tissue type, species and source. For ganglia, the number of candidates varied between 61 and 358, for gonad between 95 and 142 and for haemolymph approximately 130 (Table 3). The number of peptides identified also varied dependent on the number of candidates and tissue source. The greatest percentage of candidates was identified in gonad tissue with between 27% and 44%. For ganglia it was between 16% and 40%. The lowest number was for haemolymph samples with between 5% and 15% of candidates identified.

CrossMatch analysis of the 10 ganglia experiments resulted in consolidation of a preliminary list of 90 candidates proposed to be involved in maturation processes. Of these 32 were assigned an identity using MatchMass (Appendix 3) while the remaining 58 were unidentified (Appendix 4). Graphical representations revealed the proposed method of action for each of the peptides in controlling maturation which is summarised in Table 4. A review of the literature to determine the biological function of each of these proteins resulted in a short list of 25 peptides suspected to be involved in maturation processes (Appendix 5). For gonad tissue a total of 104 potential candidates were selected however none of these were consistent with those detected in the ganglia samples. In addition the majority of the identified proteins (39) were related to gamete development and not expected to be involved in the signalling mechanisms responsible for gonad maturation processes (Appendix 6 and 7). The regulatory methods as predicted by graphical analysis are summarised in Table 4.

Table 4 Summary of proposed regulatory mechanisms for peptide candidates involved in maturation processes in ganglia and gonad tissue

	Identified			Unidentified		
Regulation	**Up**	**Down**	**Variable**	**Up**	**Down**	**Variable**
Ganglia	11	3	17	42	10	6
Gonad	25	14		27	38	

3.5 Response of maturation related candidates with increasing age

Maturation related candidates identified in Section 3.4 were further analysed for their role in decreased gonad conditioning and spawning success. Each of the peptides associated with an identified protein in the candidate list was not detected for every species or age group. This is due to the rapid degradation rate of peptides involved in signalling processes. Therefore predictions based on average fold changes for all peptides detected for a protein within a year class were used to compare changes in response to age.

Three groups emerged from the analysis (Fig. 2). The first included proteins such as FMRF1, myomodulin, pedal peptide and a predicted protein which are characterised as being involved in a wide variety of biological processes. The second group includes pedal peptide and four unidentified proteins which show between a 25% and 82% increase in the average fold change of protein expression levels in response to increased age. The third group consists of whitnin, a known maturation related protein and an unidentified protein which are expressed at 25% and 38% higher average fold change expression levels in younger abalone.

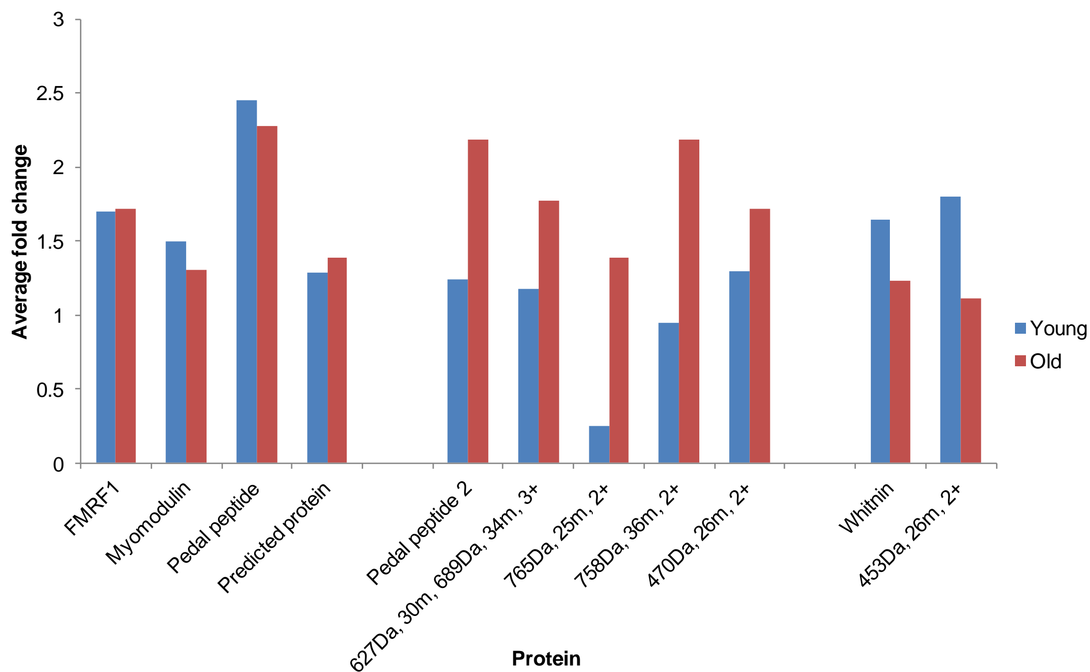

Figure 2 Changes in maturation related protein expression levels in response to age where fold change was calculated based on visual gonad indice of zero for each peptide degradation product detected by mass spectroscopy and the average for each protein calculated.

3.6 *Non-destructive sampling to monitor animal responses to experimental peptide based intervention strategies*

Two mass spectroscopy experiments comparing haemolymph from male and female greenlip abalone were conducted to assess suitability as a non-destructive sampling method. Eighteen proteins were able to be identified (Appendix 8) and a further 34 were unable to be identified (Appendix 9). Paramyosin was the only protein detected in both ganglia and haemolymph samples. Other proteins listed in the candidate short list were not present in these samples. The chromatograms associated with peptide elution off the column displayed distinctive contaminant peak/s in all samples. The sample volume injected onto the column was reduced to account for the contaminant however this also reduced peptide detection capability.

4. Discussion

4.1 Method development

Despite method development specific for abalone, the lack of an annotated proteome has reduced the expected outcomes of this project. This has resulted in a much larger number of peptides potentially involved in maturation remaining unidentified. Successful identification of proteins by mass spectrometry is reliant on the protein database searched. Species with rich genomic resources such as cattle, sheep and humans have a higher rate of identification and sequence assignment.

For species such as abalone which are less well studied at the genomic level identification success is much reduced. In the initial stages of this project publically available resources including those from closely related species were acquired. However due to limitations of search algorithms and downstream analyses as little as one amino acid change in the sequence can result in non identification and subsequent sequence assignment. The development of genetic resources specific to greenlip abalone out of necessity became a central focus of the project and as a result this improved sequence identification by approximately four fold.

The lower identification success by traditional methods resulted in the development of a novel pipeline using MarkerView for analysis. This software, while used predominately for metabolite and environmental contaminant detection provided the opportunity to concurrently examine different maturation stages. This method ensured that the majority of proteins and peptides involved in maturation stages would be selected, identified, sequenced and associated with a specific biological process where possible. Planned *de novo* sequencing of unidentified or novel peptides did not take place due to difficulty in definitively associating peptides with maturation without identification.

4.2 Maturation candidates

Graphical representations have not improved the selection process due to the rapid degradation rate of signalling peptides which results in multiple sequences of different lengths of the same peptide. Therefore there are a number of peptides which are unable to be assigned to a specific peptide and whose roles and biological function with respect to maturation remains unknown.

In general peptides with variable regulation are those which are involved in other biological processes and also contribute to maturation processes. For example buccalin is a modulatory neuropeptide, acting presynaptically on nerve terminals to inhibit acetylcholine release which results in a decrease muscle contraction. It is

predominately associated with feeding behaviour and contraction of the accessory radula closer muscle in *Aplysia*. However exogenous application of buccalin A also has modulatory effects on mechanosensory neurons (Raymond *et al.*, 1989; Rosen *et al.*, 1989). In addition, buccalin-like immunoreactivity has been detected in tissues of the cardiovascular, reproductive, and digestive systems suggesting that it has the potential to be involved in a variety of actions in these systems as well (Miller *et al.*, 1993).

One of the seasonal cues for gonad conditioning in wild abalone is increased food availability in the environment. The abalone aquaculture systems in this study are all open to the environment and are therefore subject to these same cues which in combination with the other biological actions of this neuropeptide accounts for the variable profiles. Functional studies are required to ascertain whether up regulation of this peptide in combination with other maturation related peptides would contribute to induction of conditioning. Its role in decreasing muscle contraction by inhibiting acetylcholine release at nerve terminals may also provide a means to inhibit spawning through a reduction in contraction of the foot muscle to expel gametes.

FMRF and myomodulin peptides have regulatory roles in feeding behaviour and reproduction (Veenstra, 2010). FMRF is also cardioexcitatory with diverse biological roles such as heart rate, blood pressure and gut motility (Cummins *et al.*, 2011). FMRF and myomodulin gene expression studies in *H. asinina* however indicate decreases in expression levels just prior to spawning in females while expression peaks in males just prior to spawning (York *et al.*, 2012). FMRF and myomodulin as for buccalin have variable peptide expression profiles suggesting that while they have a role to play in conjunction with other peptides in reproductive processes they may not be the major neuropeptides responsible for conditioning and spawning. The three *H. laevigata* predicted proteins with variable regulation may have similar functions to FMRF and myomodulin however functional studies with the biologically active peptides and histology are essential to determine their roles and involvement in maturation processes.

Pedal peptide 2 was up regulated in response to increasing VGI while pedal peptide precursor expression exhibited variable expression. This may be due to different splice variants which results in binding to different receptors to promote a different function. For example pedal peptides are involved in locomotion as well as reproductive behaviour in terms of spawning positioning (Moroz, 2006). Future functional studies of pedal peptide 2 will provide further insight into the involvement of this peptide in reproductive processes.

Twitchin is a regulator of muscle contraction and relaxation. It senses mechanical strain that occurs during muscle activity by unfolding in clearly resolvable steps at differing forces (Funabara et al., 2001). Up regulation of this peptide in response to increasing VGI is likely to be associated with increased locomotion activity. This is in response to seasonal increases in water temperatures which also results in increased feeding activity and growth over the warmer months of the year. Expression studies specifically associated with spawning activity are needed to verify the role of this peptide in reproductive processes. However it may act in conjunction with pedal peptide 2 to expel gametes from the gonad.

Whitnin peptide expression is up regulated in response to increasing VGI which is consistent with gene expression studies. In *H. asinina* whitnin gene expression peaks at release of oocytes from the extracellular matrix in females while in male's expression levels peak around spawning (York *et al.*, 2012). This difference in gene expression may be related to observations of abalone male gonad physiology where

sperm have been observed to be motile within the male gonad unlike other fish species (H. King pers. comm.). As for previous peptides this peptide is likely to function in conjunction with other maturation related peptides.

Functional studies and histology is required for the seven unknown, one *H. laevigata* predicted protein and one *Crassotrea gigas* uncharacterised protein which exhibit peptide up regulation in response to increased VGI. However based on their expression pattern they are likely to act in conjunction or have similar actions to twitchin and/or whitnin in terms of reproductive processes. There are also two *H. laevigata* predicted peptides and three unknown peptides which are down regulated in response to increase in response to increased VGI. From a theoretical perspective these peptides may be associated with the conversion from juvenile to mature adult. In theory they may provide a potential avenue for inhibiting or delaying reproductive processes. However as reproductive success is essential to species survival inhibition of other peptides in concert is likely to be necessary for this strategy to be effective.

4.3 Candidate influence on gamete condition and spawning success

Not all candidate proteins were detected in both the young and old year classes of abalone examined in this study. This is likely to be due to variable peptide degradation. There is also the potential for collection site effects. Gene expression studies are required to completely discount these peptides from involvement in maturation and reproductive related processes. Fold change in expression was used to compare between year classes as expression levels overall were 10 fold higher in older abalone as expected due to normal physiological differences associated related to size.

Proteins likely to have inhibitory effects on gamete or gonad conditioning and hence spawning success are those which exhibit increased fold change in expression levels in response to age and increasing VGI. The peptides in this group were mostly unidentified with the exception of pedal peptide 2 which is involved in locomotion as well as reproductive behaviour in terms of spawning positioning. This peptide however could be expected to exhibit a higher fold change due to increased effort to move a larger abalone. Water temperature also increases with VGI which promotes feeding behaviour and requires movement. In this instance it is possible that too much energy is being expended on locomotion to fuel feeding behaviour inhibiting successful gamete conditioning. In this instance nutritional interventional is likely to be more effective than peptide interventional methods.

A second hypothesis is that proteins with lower fold change levels in older abalone are required at higher expression levels to successfully induce gamete conditioning and spawning. For example insufficient whitnin may result in insufficient oocytes being released from the extracellular matrix (York *et al.,* 2012) contributing to low spawning success in females. However whitnin would not be related to gamete condition in females as it would promote release of gametes from the extracellular matrix regardless of their condition. Insufficient whitnin in males would also be counterproductive to successful spawning as this peptide peaks at spawning in *H. asinina* as discussed previously (York *et al.*, 2012).

Further functional, histological and gene expression studies of the remaining unidentified candidates are essential to ascertain whether the proposed inhibitory and excitatory effects are related to gamete condition and spawning success.

4.4 Non destructive sampling for monitoring interventional strategies

Several issues were encountered in developing a non-destructive sampling method. One of these was related to the sampling method. Mass spectrometry is extremely sensitive to contaminants. In this instance it is believed that the plastic of the bag used to hold tissue to drain haemolymph for collection inadvertently introduced a contaminant into the samples. Due to the high degradation rate of peptides post-signalling as evidenced in the ganglia data (Appendix 3) improving collection methods by drawing haemolymph directly from a sinus are unlikely to resolve this issue. Destructive sampling to measure gene or peptide expression responses is essential in combination with functional studies to develop a model of the interaction between maturation related peptides and ascertain those candidates essential for successful intervention.

5. Benefits and Adoption

The aim of this project was to generate a foundation of knowledge to support the development of interventional methods to prevent or promote conditioning and spawning in abalone in support of the Australian cultured industry. In the long term application of this knowledge to the industry is expected to have a cost benefit of 5-20% across the key production areas identified in Section 1.1. Data generated during this project provides a framework for the development of functional assays to assist in reaching this goal.

Novel tools and resources are now available to industry, abalone and proteomic research communities. These are a novel peptide analytical pipeline which is applicable to any species where there is a paucity of protein databases; and development of the molluscan databases has generated greater than 12 million sequences improving genomic and transcriptomic data resources for Australian cultured abalone species.

6. Further Development

The raw MS data produced during this project will be able to be interrogated to provide further essential knowledge of maturation processes. Rapid growth in genetic research in abalone species worldwide will see database resources continue to improve over the next few years. Knowledge gained during the course of this project suggests that molecules responsible for gonad maturation are also tightly linked with growth. Therefore with the foundation of knowledge now available for the identified peptides we are able to proceed with functional assays and determination of practical applications in the commercial environment.

7. Planned Outcomes

Public Benefit Outcomes

The public benefits associated with this project are an increase in genetic resources and novel proteomic data analysis methods which may be utilised by other molluscan and proteomic researchers.

Private Benefit Outcomes

The private benefit of this project is progress towards the goal of enhancing abalone aquaculture by preventing maturation in commercial stock resulting in improved abalone health, increased production and improved confidence in product quality.

This will result in more consistent quality and supply of abalone to market and increased profitability. Tighter control of maturation in breeding stock will result in synchronised broodstock conditioning, improved seasonal conditioning of aging broodstock and assistance with on-demand spawning technologies. This will result in more reliable conditioning of highly valued 'elite' selected broodstock. Combining these benefits with reliable spawning methodologies will enable the greatest gains from selective breeding programs particularly with the advent of closed breeding cycles in operation on most farms.

Linkages with CRC Milestone Outcomes

This project is aligned with CRC Output 1.3 which is the 'removal or reduction of key production constraints in selected aquaculture system' and Milestone 1.3.5 which is 'production efficiency gains from genetic, health management and nutritional interventions quantified to inform long term strategies and estimated commercial benefits. The primary goal of this project was to provide a foundation of knowledge on maturation. Specifically to identify critical maturation time points and key molecules involved in and/or responsible for maturation to provide a basis for the future development and implementation of commercial scale technologies for use within the industry.

The key outputs from this project were the development of a molluscan specific neuropeptide database which can be utilised routinely used for identification, quantification and monitoring of key maturation/reproductive neuropeptides in temperate abalone. Validation of sampling and collection methods provided identification of abalone neuropeptides homologous to other species. Difficulty in identification of all peptides predicted to be involved in the maturation process has hindered progress towards a clearly defined strategy for developing a commercially applicable system to control maturation and spawning in temperate abalone. However a solid foundation now exists for a reduction of the identified key production constraints in the future.

8. Conclusion

This project has highlighted both the lack of genetic knowledge available for abalone species and the difficulty in determining gene and protein interaction networks in invertebrate species. Data generated during this project is available to be reinterrogated as further genetic resources become available in the future which may lead to further insights. Candidates have been identified for involvement in reproductive maturation processes. Knowledge of the sequence of active biological maturation related peptides provides the opportunity to conduct functional studies to progress towards an interventional strategy to promote or inhibit gonad condition as required in commercial or breeding stock.

9. References

Blankenberg D, Taylor J, Schenck I, He J, Zhang Y, Ghent M, Veeraraghavan N, Albert I, Miller W, Makova KD, Hardison RC, Nekrutenko A. (2007) A framework for collaborative analysis of ENCODE data: making large-scale analyses biologist-friendly, Genome Research, 17: 960-964

Cummins, S.F., Tollenaere, A., Degnan, B.M. and Croll, R.P. (2011) Molecular analysis of two FMRFamide-encoding transcripts expressed during the development of the tropical abalone Haliotis asinina, Journal of Comparative Neurology, 519: 2043-2059

Funabara D., Kinoshita S., Watabe S., Siegman M.J., Butler T.M. and Hartshorne D.J. (2001) Phosphorylation of molluscan twitchin by the cAMP-dependent protein kinase, Biochemistry 40:2087-2095

Hummon, A.B., Hummon, N.P., Corbin, R.W., Li, L.J., Vilim, F.S., Weiss, K.R., Sweedler, J.V. (2003) From precursor to final peptides: a statistical sequence-based approach to predicting prohormone processing. Journal of Proteome Research, 2: 650-656

Miller, M.W., Beushausen, S., Cropper, E.C., Eisinger, K., Stamm, S., Vilim, F.S., Vitek, A., Zajc, A., Kupfermann, I., Brosius, J. and Weiss, K.R. (1993) The buccalin-related neuropeptides: isolation and characterization of an *Aplysia* cDNA clone encoding a family of peptide cotransmitters. The Journal of Neuroscience, 13: 3348-3357

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

Raymond, J.L., Shulman, E.E., Baxter, D.A., Cleary, L.J. and Byrne, J.H. (1989) Differential effects of the peptide buccalin and serotonin on membrane currents, action potential duration, and excitability in pleural sensory neurons of Aplysia, Society of Neuroscience Abstract 15: 1284

Rocchiccioli, S., Citti, L., Boccardi, C., Ucciferri, N., Tedeschi, L., Lande, C., Trivella, M.G. and Cecchettini, A. (2010) A gel-free approach in vascular smooth muscle cell proteome: perspectives for a better insight into activation, Proteome Science, 8: 15-23

Rosen, S.C., Susswein, A.J., Cronner, E.C., Weiss, K.R. and Kuofermann, I. (1989) Selective modulation of spike duration by serotonin and the neuropeptides FMRFamide, SCP, buccalin and myomodulin in different classes of mechanoafferent neurons in the cerebral ganglion of Aplysia, Journal of Neuroscience 9: 390-402.

Travers, M., Le Goic, N., Huchette, A., Koken, M. and Paillard, C. (2008) Summer immune depression associated with increased susceptibility of the European abalone, *Haliotis tuberculata* to *Vibrio harveyi* infection. Fish and Shellfish Immunology, 25: 800-808

Veenstra, J.A. (2010) Neurohormones and neuropeptides encoded by the genome of Lottia gigantean, with reference to other mollusks and insects, General and Comparative Endocrinology, 167: 86-103

York, P.S., Cummins, S.F., Degnan, S.M., Woodcroft, B.J. and Degnan, B.M. (2012) Marked changes in neuropeptide expression accompany broadcast spawning in the gastropod *Haliotis asinina*, Frontiers in Zoology, 9: 9

10. Appendices

Appendix 1. Intellectual property

The intellectual property and valuable information arising from this research lies with the Seafood CRC on behalf of all its participants. At this time no specific intellectual property is identified as being generated from the results of this research.

Location and format of data	Topic(s)	Author/Custodian	Access to data
Database	Abalone specific nucleotide, protein and peptide sequences	Natasha Botwright	For project participants only
Database	List of samples collected and observations	Natasha Botwright	For project participants only
Lab notebook	Preparation of ganglia cDNA library	Natasha Botwright	For project participants only
Lab notebook	Abalone neuropeptide methodologies	Natasha Botwright	For project participants only
Spreadsheets/s	Analysis of gonad and ganglia cDNA libraries	Natasha Botwright	For project participants only
Secure backed up server	Mass spectroscopy data	Natasha Botwright	For project participants only
Secure backed up server	*Haliotis laevigata* ganglia and gonad transcriptomes	Natasha Botwright	For project participants only
Lab notebook	Sample preparation for transcriptome sequencing	Natasha Botwright	For project participants only
Secure backed up server	Analysis of *Haliotis laevigata* transcriptomes	Natasha Botwright	For project participants only
Database/s	Molluscan mass spectroscopy databases	Natasha Botwright	For project participants only
Spreadsheet/s	Analysis of mass spectroscopy data	Natasha Botwright	For project participants only
Spreadsheet/s	List of abalone neuropeptides and sequences homologous to invertebrate species	Natasha Botwright	For project participants only
Spreadsheet/s	Preliminary list of maturation peptide profiles of interest	Natasha Botwright	For project participants only
Secure backed up server	Perl MS comparative programs	Natasha Botwright	For project participants only
Secure backed up server	Protein Pilot search results	Natasha Botwright	For project participants only
Spreadsheet/s	Complete list of maturation peptides/proteins of interest	Natasha Botwright	For project participants only
Spreadsheet/s	Shortlist of maturation peptides/proteins of interest	Natasha Botwright	For project participants only

Appendix 2 Staff

Staff engaged on the project:

Principal Investigator

Natasha Botwright	CSIRO Animal, Food and Health Sciences

Co-Investigators

Dr Michelle Colgrave	CSIRO Animal, Food and Health Sciences
Dr Mathew Cook	CSIRO Marine and Atmospheric Research
Dr Nick Elliott	CSIRO Marine and Atmospheric Research

Appendix 3 Identified ganglia maturation related preliminary candidates

m/z	RT	*z*	DB_ID	Peptide	Protein	Function
Up regulated with increasing VGI						
526	23	2	U27933 (2)	SANPLEIDEGLPRLRPAPR	Myomodulin neuropeptide 2 [*Crassostrea gigas*]	Feeding behaviour, potentiates accessory radula closer muscle contraction
402	28	2	CL5039.C1 (6)	ALSKAAEI	Neural-cadherin [*Crassostrea gigas*]	Cell adhesion, developmental protein
873	30	2	U24790 (2)	GSFFTQSDENTFPRI	Neuropeptide precursor [*Helix lucorum*]	Unknown
400	17	2	U13634 (2)	VTRDVKIVRI	Paramyosin [*Haliotis discus discus*]	Major structural component of invertebrate muscle thick filament
742	22	2		1Ac.MDYGDVSSQVVRS		
842	24	2		1Ac.MDYGDVSSKVVRTVS		
911	23	2		1Ac.MDYGDVSSQVVRSVTH		
530	26	2	U9697 (2)	MANGLSRFY	Pedal peptide 2 precursor [*Aplysia californica*]	Locomotion, reproductive behaviour i.e. positioning
548	27	2		GLDHGLSSFY		
400	29	2	A9QLN4 (9)	GIGGGGLGGGP	Precollagen-P [*Mytilus californianus*]	Collagen
742	28	2	S249375_4 (4)	PFDSIASGRGMAGFA	Predicted protein [*Haliotis laevigata*]	Unknown
513	22	2	U33767 (3)	LKVESYSVT	Predicted protein [*Haliotis laevigata*]	Unknown
367	26	2	Q7YT99 (9)	LAFNAPT	Twitchin [*Mytilus galloprovincialis*]	Regulator of muscle contraction and relaxation.

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; VGI – visual gonad index; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

m/z	RT	*z*	DB_ID	Peptide	Protein	Function
653	21	2	K1QED4 (9)	REINMMIQRN	Uncharacterized protein [*Crassostrea gigas*]	Unknown
552	23	2	U24770 (2)	DEGRLDEASL	Whitnin prepropeptide [*Haliotis asinina*]	Peaks at release of oocytes from the EC matrix in females; maximum in males at spawning
588	21	2		DEGRLDEASLA		
624	21	2		ADEGRLDEASLA		
637	20	2		LPADEGRLDEAS		
637	28	2		PADEGRLDEASL		
693	25	2		LPADEGRLDEASL		
729	24	2		LPADEGRLDEASLA		
764	25	2		LPADEGRLDEASLAA		
800	25	2		SGLPADEGRLDEASLA		
828	24	2		LPADEGRLDEASLAAE		
Down regulated with increasing VGI						
340	27	2	S166847_25 (4)	GPKGYSA	Predicted protein [*Haliotis laevigata*]	Unknown
516	31	2	S187249_5 (4)	LTGPFSKVVTRKLT	Predicted protein [*Haliotis laevigata*]	Unknown
592	21	2	K1QX02 (9)	GDGLVSVAGGSGHA	Tenascin-X [*Crassostrea gigas*]	Cell adhesion, cell communication
615	23	2		ASGASHGGLGGAGACG		
Variable regulation						
542	31	2	CL3199.C2 (6)	ANPDLLPEIT	CDK5 regulatory subunit-associated protein 2-like [*Strongylocentrotus purpuratus*]	Regulation of neuronal differentiation

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; VGI – visual gonad index; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

m/z	RT	*z*	DB_ID	Peptide	Protein	Function
480	32	2	U19376 (2)	FDPLANGLI	Buccalin precursor [*Aplysia californica*]	Modulatory neuropeptide, inhibits acetylcholine release to decrease muscle contraction
506	34	2		PFDPLAS[Dhy]GLI		
515	34	2		PFDPLASGLI		
531	28	2		PFDELASGLI		
541	29	2		SPLDPLASGLI		
665	33	2		APFDPLANELIE		
502	15	2	CL2930.C1 (2)	VETKAVEAGNKDIE	FMRF1 [*Haliotis asinina*]	Cardioexcitatory peptide e.g. heart rate, blood pressure, gut motility, feeding behaviour, reproduction
512	30	2		LAGDSFLRF		
563	31	2		TLAGDSFLRF		
570	19	2		ILSGGPPQNEE		
647	25	2		SDSDLDDVIRAS		
660	24	2		DSDLDDVIRASL		
773	27	2		LIKW[Kyn]ISCINGFSDFCNKP[Oxi]VN		
581	21	2	B2ZSS8 (9)	SVQEPSVASADA	FMRF2 [*Haliotis asinina*]	Cardioexcitatory peptide e.g. heart rate, blood pressure, gut motility, feeding behaviour, reproduction
646	23	2		SVQEPSVASADAM		
512	24	2	2500_20 (8)	MGQSSFVRI	FVRIamide neuropeptide precursor [*Haliotis*]	May be involved in suppressing male reproductive activity. Phototaxis
500	21	2	U27887 (2)	SDPAVSVVPK	Hypothetical protein [*Crassostrea gigas*]	Unknown
318	26	2	U47254 (2)	YGPGQN	Lipopolysaccharide-induced TNF factor like protein [*Polysphondylium pallidum*]	Transcription regulation

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; VGI – visual gonad index; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

m/z	RT	*z*	DB_ID	Peptide	Protein	Function
402	28	2	CL3106.C2 (2)	NMLRLGDGLQMLRL	Myomodulin prepropeptide [*Haliotis asinina*]	Modulation of reproduction, feeding status (peaks in males prior to spawning)
408	28	2		ALFAKAVT[Dhy]		
607	29	2	GT274785_13 (8)	AMDKFGFASQL	Predicted protein [*Haliotis asinina*]	Unknown
619	23	2	U2250 (2)	SFDKIAHGGFAS	Pedal peptide precursor [*Helix lucorum*]	Locomotion, reproductive behaviour i.e. positioning
632	24	2		PFDSIASGRGMAG		
690	30	2		PLDSISAGSGLSGFA		
406	29	2	S100857_3 (4)	LTLSTYN	Predicted protein [*Haliotis laevigata*]	Unknown
508	28	2	S49531_1 (4)	LKAVPFNNI	Predicted protein [*Haliotis laevigata*]	Unknown
562	22	2	S64767_5 (4)	SPKYGYYFV	Predicted protein [*Haliotis laevigata*]	Unknown
573	38	2		LVNPFVYYALGKNKNSGNTP		
544	29	2	S138335_1 (4)	SALGRSLIISA	Predicted protein [*Haliotis laevigata*]	Unknown
509	30	2	S153809_2 (4)	ISISGLSDLI	Predicted protein [*Haliotis laevigata*]	Unknown
651	27	2	S355084_1 (4)	LGSANEIRISQI	Predicted protein [*Haliotis laevigata*]	Unknown
641	23	2	CL2908.C2 (6)	PVYSIAFQNDGA	U4/U6 small nuclear ribonucleoprotein Prp4 [*Crassostrea gigas*]	Participates in pre-mRNA splicing. Spliceosome building block.

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; VGI – visual gonad index; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

Appendix 4 Unidentified ganglia maturation related preliminary candidates

Up regulated with increasing VGI

m/z	RT	*z*
393	28	2
398	23	2
441	21	2
466	32	2
484	30	2
490	25	2
510	33	2
513	24	2
532	30	2
539	29	2
545	29	2
546	26	2
546	29	2
557	22	2
572	31	2
589	22	2
590	26	2
593	20	2
597	29	2
608	29	2
625	18	2
639	29	2
665	25	2
666	33	2
673	23	2
674	18	2
694	30	2
716	30	2
748	21	2
751	28	2
758	36	2
765	25	2
766	30	2
786	25	2
813	16	2
830	25	2
838	26	2
872	34	2
1100	33	2
1160	33	2
1216	34	2
627*	30	3
636*	31	3
687*	28	3
689*	34	3
775*	33	3

m/z – mass/charge; RT – retention time;
z – charge; VGI – visual gonad index;

Down regulated with increasing VGI

m/z	RT	*z*
357	24	2
362	25	2
391	16	3
433	29	2
470	30	2
481	30	2
560	32	2
606	32	2
628	32	2
784	18	5

Variable regulation

m/z	RT	*z*
453	26	2
470	26	2
575	22	2
777	27	2
442*	23	2
789*	33	2
864*	32	2
501*	21	2
650*	32	2
703*	33	2
894*	34	2

Appendix 5 Maturation related candidate short list

m/z	RT	*z*	DB_ID	Peptide	Protein	Function	GLF09	GLM09	HYF09	HYM09	GLF07	GLM07	HYF08	HYM08	BLF	BLM
Variable regulation																
480 506 515 531 541 665	32 34 34 28 29 33	2 2 2 2 2 2	U19376 (2)	FDPLANGLI PFDPLAS[Dhy]GLI PFDPLASGLI PFDELASGLI SPLDPLASGLI APFDPLANELIE	Buccalin precursor [*Aplysia californica*]	Modulatory neuropeptide, inhibits acetylcholine release to decrease muscle contraction	X	X	X	X	X	X	X	X		X
502 512 563 570 647 660 773	15 30 31 19 25 24 27	2 2 2 2 2 2 2	CL2930.C1 (2)	VETKAVEAGNKDIE LAGDSFLRF TLAGDSFLRF ILSGGPPQNEE SDSDLDDVIRAS DSDLDDVIRASL LIKW[Kyn]ISCINGFSDFCNKP[Oxi]VN	FMRF1 [*Haliotis asinina*]	Cardioexcitatory peptide e.g. heart rate, blood pressure, gut motility, feeding behaviour, reproduction	X	X	X	X		X	X	X	X	X
581 646	21 23	2 2	B2ZSS8 (9)	SVQEPSVASADA SVQEPSVASADAM	FMRF2 [*Haliotis asinina*]	Cardioexcitatory peptide e.g. heart rate, blood pressure, gut motility, feeding behaviour, reproduction						X	X	X		

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; GLF – greenlip female; GLM – greenlip male; HYF – hybrid female; HYM – hybrid male; BLF – blacklip female; BLM – blacklip male; 09 – 2009 year class; 08 – 2008 year class; 07 – 2007 year class; VGI – visual gonad index; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

m/z	RT	*z*	DB_ID	Peptide	Protein	Function	GLF09	GLM09	HYF09	HYM09	GLF07	GLM07	HYF08	HYM08	BLF	BLM
402 408	28 28	2 2	CL3106.C2 (2)	NMLRLGDGLQMLRL ALFAKAVT[Dhy]	Myomodulin prepropeptide [*Haliotis asinina*]	Modulation of reproduction, feeding status (peaks in males prior to spawning)	X	X			X	X	X	X		
619 632 690	23 24 30	2 2 2	U2250 (2)	SFDKIAHGGFAS PFDSIASGRGMAG PLDSISAGSGLSGFA	Pedal peptide precursor [*Helix lucorum*]	Locomotion, reproductive behaviour i.e. positioning	X	X			X	X	X	X		
562 573	22 38	2 2	S64767_5 (4)	SPKYGYYFV LVNPFVYYALGKNKNSGNTP	Predicted protein [*Haliotis laevigata*]	Unknown			X	X					X	X
544	29	2	S138335_1 (4)	SALGRSLIISA	Predicted protein [*Haliotis laevigata*]	Unknown					X	X	X	X		
651	27	2	S355084_1 (4)	LGSANEIRISQI	Predicted protein [*Haliotis laevigata*]	Unknown					X	X	X	X		
Up regulated in response to increasing VGI																
530 548	26 27	2 2	U9697 (2)	MANGLSRFY GLDHGLSSFY	Pedal peptide 2 precursor [*Aplysia californica*]	Locomotion, reproductive behaviour i.e. positioning	X	X			X	X	X	X		

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; GLF – greenlip female; GLM – greenlip male; HYF – hybrid female; HYM – hybrid male; BLF – blacklip female; BLM – blacklip male; 09 – 2009 year class; 08 – 2008 year class; 07 – 2007 year class; VGI – visual gonad index; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

m/z	RT	*z*	DB_ID	Peptide	Protein	Function	GLF09	GLM09	HYF09	HYM09	GLF07	GLM07	HYF08	HYM08	BLF	BLM
742	28	2	S249375_4 (4)	PFDSIASGRGMAGFA	Predicted protein [*Haliotis laevigata*]	Unknown	X	X	X	X						
367	26	2	Q7YT99 (9)	LAFNAPT	Twitchin [*Mytilus galloprovincialis*]	Regulator of muscle contraction and relaxation.	X	X			X		X	X		
653	21	2	K1QED4 (9)	REINMMIQRN	Uncharacterized protein [*Crassostrea gigas*]	Unknown					X	X	X	X		
552	23	2	U24770 (2)	DEGRLDEASL	Whitnin prepropeptide [*Haliotis asinina*]	Peaks at release of oocytes from the EC matrix in females; maximum in males at spawning	X	X	X	X	X	X	X	X		X
588	21	2		DEGRLDEASLA												
624	21	2		ADEGRLDEASLA												
637	20	2		LPADEGRLDEAS												
637	28	2		PADEGRLDEASL												
693	25	2		LPADEGRLDEASL												
729	24	2		LPADEGRLDEASLA												
764	25	2		LPADEGRLDEASLAA												
800	25	2		SGLPADEGRLDEASLA												
828	24	2		LPADEGRLDEASLAAE												
627*	30	3				Unknown	X	X	X	X		X	X	X	X	
636*	31	3														
687*	28	3														
689*	34	3														
775*	33	3														

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; GLF – greenlip female; GLM – greenlip male; HYF – hybrid female; HYM – hybrid male; BLF – blacklip female; BLM – blacklip male; 09 – 2009 year class; 08 – 2008 year class; 07 – 2007 year class; VGI – visual gonad index; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence; EC – extracellular matrix

m/z	RT	*z*	DB_ID	Peptide	Protein	Function	GLF09	GLM09	HYF09	HYM09	GLF07	GLM07	HYF08	HYM08	BLF	BLM
765	25	2				Unknown	X		X	X				X		
453	26	2				Unknown	X	X	X	X	X	X	X		X	
557	22	2				Unknown			X	X	X	X	X	X		
625	18	2				Unknown			X	X	X	X	X			
575	22	2				Unknown			X	X	X	X	X	X		
758	36	2				Unknown	X	X					X	X		
Down regulated in response to increasing VGI																
340	27	2	S166847_25 (4)	GPKGYSA	Predicted protein [*Haliotis laevigata*]	Unknown					X	X	X	X		
516	31	2	S187249_5 (4)	LTGPFSKVVTRKLT	Predicted protein [*Haliotis laevigata*]	Unknown					X	X	X	X		
362	25	2				Unknown			X	X	X		X	X		
453	26	2				Unknown	X	X	X	X	X	X	X	X	X	X
470	26	2				Unknown	X	X	X	X	X	X	X	X	X	

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; GLF – greenlip female; GLM – greenlip male; HYF – hybrid female; HYM – hybrid male; BLF – blacklip female; BLM – blacklip male; 09 – 2009 year class; 08 – 2008 year class; 07 – 2007 year class; VGI – visual gonad index; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence; EC – extracellular matrix

Appendix 6 Identified gonad maturation related candidates

Down regulated with increasing VGI

GLF09	GLM09	HYF09	HYM09	*m/z*	RT	*z*	DB_ID	Peptide	Protein
X		X	X	403	30	2	D2CMM6 (9)	TNKANGSN	NR2 [*Aplysia californica*]
X	X			595	40	2	M_08-D06_1 (5)	TIKIFVDLIK	Predicted protein [*Haliotis rubra*]
X	X	X		577	34	2	U38486 (6)	HRIGRTGRSN	Predicted: probable ATP-dependent RNA helicase DDX5-like isoform 1 [*Strongylocentrotus purpuratus*]
X		X	X	531	33	2	S9878_8 (4)	TGKLMTTNPV	Predicted protein [*Haliotis laevigata*]
X		X	X	417	31	2	U17925 (3)	SKKNRSN	Predicted protein [*Haliotis laevigata*]
X		X		340	29	2	CL7729.C1 (2)	GPKGYTG	Serine/threonine-protein kinase RIO3 [*Crassostrea gigas*]
X		X		384	31	2	S231661_10 (4)	LPLPLGGT	Predicted protein [*Haliotis laevigata*]
X		X		425	31	2	S93925_2 (4)	YGGFTFSA	Predicted protein [*Haliotis laevigata*]
X		X		445	31	2	Q1MW90 (9)	GGLGGMGGLGGG	Shematrin-7 [*Pinctada fucata*]
X	X			530	32	2	DN763571_13 (8)	SASAKVLALAK	Predicted protein [*Haliotis discus discus*]
X	X	X	X	498	33	2	K1QVA4 (9)	QVQAAQAAAH	Virilizer-like protein [*Crassostrea gigas*]
X		X	X	641	32	2	C49558537_4 (4)	LCLSLFLSLALALSHRY	Predicted protein [*Haliotis laevigata*]
X		X	X	487	32	2	S297330_3 (4)	NLFLIPSVA	Predicted protein [*Haliotis laevigata*]
		X	X	479	33	2	K1PSG4 (9)	ATRRPQTQ	Uncharacterized protein [*Crassostrea gigas*]

VGI – visual gonad index; GLF09 – greenlip female 2009 year class; GLM09 – greenlip male 2009 year class; HYF09 – hybrid female 2009 year class; HYM09 – hybrid male 2009 year class; *m/z* – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

Up regulated with increasing VGI

GLF09	GLM09	HYF09	HYM09	m/z	RT	z	DB_ID	Peptide	Protein
X		X		601	33	2	CL5473.C2 (2)	VCMLVLILINA	Hypothetical protein CGI_10013002 [*Crassostrea gigas*]
X		X		673	31	2	S67859_6 (4)	FAVGGLLGDVFLH	Predicted protein [*Haliotis laevigata*]
X		X		395	31	2	CL3052.C1 (3)	SPGTLQW	Predicted protein [*Haliotis laevigata*]
	X		X	812	30	2	U18450 (7)	VEVSGYISVEIDNAK	Predicted protein [*Haliotis laevigata*]
				384	28	2		LPIPLKS	
				492	24	2		DRVRNAVDRYTPDRET	
				638	28	2		PGQ.QERVHGLNIPE	
X		X		676	27	2	U19297 (2)	SGEVDTTNPYRL	Hypothetical protein CGI_10006238 [*Crassostrea gigas*]
X		X		824	37	2	U14720 (2)	VQKVNPARLPVVVGGL	Clathrin heavy chain 1 [*Crassostrea gigas*]
X		X		727	29	2	K1RHE8 (9)	LVIMTLFTIFAAI	Polycystic kidney disease and receptor for egg jelly-related protein [*Crassostrea gigas*]
X		X		867	37	2	S100644_4 (4)	LKTVQPCIPISPLPAR	Predicted protein [*Haliotis laevigata*]
X		X		596	37	3	S105781_16 (4)	N[Dea]ELVRWSNAGVSRIAN	Predicted protein [*Haliotis laevigata*]
X		X		925	37	2	S116241_7 (4)	VLNIVTLIPNAPSIILAS	Predicted protein [*Haliotis laevigata*]
X		X		670	27	2	S19758_48 (4)	ERLRLPVEPET	Predicted protein [*Haliotis laevigata*]
X		X		878	37	2	S97414_10 (4)	VGVGGRLLALACLLACLL	Predicted protein [*Haliotis laevigata*]
X		X		714	35	2	U14036 (7)	VVIESQLNWIIN	Predicted protein [*Haliotis laevigata*]
X		X		805	30	2	U34830 (2)	GDNDYAPEFPSPVNS	Protocadherin-like wing polarity protein [*Crassostrea gigas*]
X		X		670	31	2	U915 (2)	TAHWQTQTTAPP	Splicing factor 1 [*Crassostrea gigas*]

VGI – visual gonad index; GLF09 – greenlip female 2009 year class; GLM09 – greenlip male 2009 year class; HYF09 – hybrid female 2009 year class; HYM09 – hybrid male 2009 year class; *m/z* – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

GLF09	GLM09	HYF09	HYM09	m/z	RT	z	DB_ID	Peptide	Protein
X		X		523	28	2	K1PG12 (9)	LAPLLALQLPPEQY	Uncharacterized protein [*Crassostrea gigas*]
				396	26	2		SPADVRF	
				449	34	2		VPFIPINP	
X		X		545	32	2	U35159 (6)	PLGRLSPMFA	Vitellogenin [*Haliotis discus hannai*]
	X		X	732	31	2	K1RDX3 (9)	AGAKSSGEKGEN[Dea]VDI	Uncharacterized protein [*Crassostrea gigas*]
	X		X	474	29	2	M-13-H8_1-4 (1)	N[Dea]LLCEILE	Enterin neuropeptide precursor [*Haliotis*]
	X		X	463	25	2	S192471_2 (4)	KNHVDSPE	Predicted protein [*Haliotis laevigata*]
				669	30	2		TVLEGYDEDEPQDLIRN	
				710	30	2		IRNEVMGPPTSSQE[NaX]NMWE	
				712	31	2		TVLEGYDEDEPQDLIRNE	
				730	31	2		DEPQDLIRNEVM	
	X		X	755	26	2	U10988 (3)	TVLEGYDEDEPQD	Predicted protein [*Haliotis laevigata*]
	X		X	591	32	2	U45337 (6)	DTILLGDLTSGKLSEIP	Predicted: intraflagellar transport protein 172 homolog isoform 1 [*Saccoglossus kowalevskii*]
	X		X	797	29	2	Q1MW94 (9)	GGYGRFLGGGVIGGGGVI	Shematrin-3 [*Pinctada fucata*]
X		X		522	24	2	U22974 (6)	VTYYLVGDI	PAX-interacting protein 1 [*Crassostrea gigas*]
X		X		591	37	2	S49128_2 (4)	SSMVTHLAKKLSDLLQ	Predicted protein [*Haliotis laevigata*]

VGI – visual gonad index; GLF09 – greenlip female 2009 year class; GLM09 – greenlip male 2009 year class; HYF09 – hybrid female 2009 year class; HYM09 – hybrid male 2009 year class; *m/z* – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

Appendix 7 Unidentified gonad maturation related candidates

Up regulated with increasing VGI

m/z	RT	*z*
416	31	2
523	24	2
581	28	4
602	33	2
610	22	4
615	34	2
629	30	3
650	28	4
661	27	4
671	27	2
711	34	2
712	37	3
780	37	2
892	30	2
935	34	3
941	32	3
943	34	3
953	36	2
1004	30	2
1054	30	2
1062	28	2
1066	31	5
1069	31	2
1072	32	2
1110	31	2
1346	33	2
1404	34	2

m/z – mass/charge; RT – retention time;
z – charge; VGI – visual gonad index;

Down regulated with increasing VGI

m/z	RT	*z*
455	34	3
481	32	2
486	35	3
505	35	3
538	34	2
564	33	2
586	33	2
596	40	2
636	34	2
643	37	3
665	43	2
674	39	2
683	35	2
687	40	2
689	34	2
700	36	3
700	28	2
705	35	2
727	35	2
747	36	2
772	39	2
773	39	2
784	21	5
786	38	4
791	36	2
792	28	4
793	36	2
836*	27	3
975*	27	3
847	25	5
884	34	2
914	40	2
947	44	2
1050	36	2
1056	28	3
1057	25	2
1094	28	3
1113	28	3
1174	27	5

Appendix 8 Identified haemolymph maturation related candidates

m/z	RT	*z*	DB_ID	Peptide	Protein	Function
691	26	2	M-10-F6_3-5 (1)	LDEMDTTCLPED	Hypothetical protein [*Haliotis laevigata*]	
634	25	2	E3VVP8 (9)	AGLGSLGAQGLGGLP	Molluscan prismatic and nacreous layer protein [*Pinctada margaritifera*]	Shell formation
744	28	2	U13634 (2)	CRM.MD[Oxi]YGDVSSKVVRSVTHRAY	Paramyosin [*Haliotis discus discus*]	Muscle thick filament
566	34	2	K1QAU8 (9)	PFGDILDIQI	Peptidyl-prolyl cis-trans isomerase E [*Crassostrea gigas*]	Accelerates protein folding and catalyses peptide bonds
441	18	2	CL8344.C1 (3)	IKKYSEL	Predicted protein [*Haliotis laevigata*]	
582	39	2	S91914_21 (4)	TKTPPGYPRF	Predicted protein [*Haliotis laevigata*]	
589	27	2	S104530_1 (4)	ATETLPQPTKSLLIPR	Predicted protein [*Haliotis laevigata*]	
622	26	2	S131753_3 (4)	TLTQTLTHSQI	Predicted protein [*Haliotis laevigata*]	
633	26	2	S164314_3 (4)	IVCTRTHQPIP	Predicted protein [*Haliotis laevigata*]	
634	23	2	S138278_4 (4)	PSNALESQVLLP	Predicted protein [*Haliotis laevigata*]	
640	28	2	S1945_8 (4)	GKSWVLAKEFAGNPTEAN	Predicted protein [*Haliotis laevigata*]	
646	29	2	U10988 (3)	SAYLDETD[NaX]LLE	Predicted protein [*Haliotis laevigata*]	
695	30	2	S88026_20 (4)	AGLVDMCSGVDLIP	Predicted protein [*Haliotis laevigata*]	
699	30	2	S326470_8 (4)	PSILPMPSLLKDS	Predicted protein [*Haliotis laevigata*]	
933	32	2	S77305_11 (4)	FQSGGNLDLLNDSRSEL	Predicted protein [*Haliotis laevigata*]	
624	22	2	K1QMC2 (9)	VEMIDPELGFP	Protein white [*Crassostrea gigas*]	ATP catabolism; eye pigmentation, gravitaxis, male courtship behaviour
570	21	2	U26119 (2)	LDDDEDRFD	Transcription factor ETV7 [*Crassostrea gigas*]	Transcription repressor
635	29	2	CL4233.C2 (2)	ASSKSAYHDSCI	Uncharacterized protein C2orf50 [*Crassostrea gigas*]	

m/z – mass/charge; RT – retention time; *z* – charge; DB_ID – database identification; green text – greater than 80% sequence confidence; blue text – 50-79% sequence confidence; red text – 1-49% sequence confidence

Appendix 9 Unidentified haemolymph maturation related candidates

m/z	RT	*z*
511	23	2
746	16	3
777	27	5
786	22	2
790	17	5
794	17	5
847	35	4
857	35	4
857	36	5
996	16	4
1051	32	5
1053	29	5
1076	27	2
1148	34	5
1152	32	5
1161	24	4
1161	25	4
1165	37	5
1173	37	3
1174	37	4
1184	32	5
1187	32	5
1231	33	5
1241	31	5
1314	32	4
1441	30	4
1480	32	4
1510	31	4
1534	33	4
1537	31	4
1540	31	4
1551	31	4
1555	32	4
1668	32	5

m/z – mass/charge; RT – retention time;
z – charge