Thermal stress in the Antarctic clam Laternula and the temperate mussel Mytilus.

Garcia, Manuela Trueban

THERMAL STRESS IN THE ANTARCTIC CLAM *Laternula* AND THE TEMPERATE MUSSEL *Mytilus*

BY

MANUELA TRUEBANO GARCIA

SUBMITTED TO THE UNIVERSITY OF WALES IN FULFILMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

SWANSEA UNIVERSITY

SCHOOL OF MEDICINE

2009

SUMMARY

There is ample evidence that a period of global warming is already affecting ecosystems worldwide. In order to predict the effects of a warming climate on organism physiology and biogeography, a description of the mechanisms involved in species responses to elevated temperatures is needed. Comparative studies examining species inhabiting different environments provide important information on the relative susceptibility of ecosystems to climate change. Antarctic marine ectotherms have evolved in a stable cold environment. They live within a narrow thermal window and experience stress with small elevations in temperature. In contrast, temperate intertidal species experience considerable temperature changes on a daily basis. The Antarctic clam *Laternula elliptica* and temperate mussel *Mytilus edulis* were selected as representative species for their respective environments. This thesis presents i) a description of the construction of a cDNA microarray for *L. elliptica*, ii) analysis of gene expression in *L. elliptica* upon acute exposure to 3°C, iii) a comparative study between the two species at the protein level via two dimensional electrophoresis, and iv) analysis of corticosteroid synthesis in *Mytilus*. Significant changes in the expression of 294 clones, representing 160 transcripts were observed. Of these, 33 were identified by sequence similarity searches and classified to a variety of cellular functions including protein turnover, folding and chaperoning, intracellular signalling and trafficking and cytoskeletal activity. In addition, the expression of 264 and 375 proteins in *L. elliptica* and *M. edulis* respectively was studied, 14 and 26 of which presented changes in expression between treatments. Only changes in proteins involved in energy metabolism were detected in both species. A higher level of biological variation in response to stress was observed in *M. edulis* at the protein level. The relevance of the observed results in determining the relative susceptibility of these species to climate change is discussed.

DECLARATION

This work has not previously been accepted in substance for any degree and is not being concurrently submitted in candidature for any degree.

Signed ....

Date .....3rd May 2010.....

This thesis is the result of my own investigations, except where otherwise stated. No correction services were used.

Other sources are acknowledged by footnotes giving explicit references. A bibliography is appended.

Signed ....

Date .....3rd May 2010.....

I hereby give consent for my thesis, if accepted, to be available for photocopying and for inter-library loan, and for the title and summary to be made available to outside organisations.

Signed ....

Date .....3rd May 2010.....

TABLE OF CONTENTS

TABLE OF FIGURES

TABLE OF TABLES

ACKNOWLEDGEMENTS

I would like to thank Melody Clark and Keiron Fraser for giving me this opportunity, especially Mel for always having the time to help and never getting tired. Thanks to Adelino Canario for allowing me to become part of his welcoming team. I thank David Skibinski, for his endless involvement, guidance and support throughout my PhD despite the distance, and for all I learnt during our long discussions.

Very special thanks go to Angel Diz and Gavin Burns without whose tuition, training and advice this thesis would have been an even harder task. Thanks to all those, too many to mention, whose doors were always open for advice... Dave Barnes, Lloyd Peck, Chester Sands, Will Goodall-Copestake, Guy Hillyard, Jaume Forcada and many others at BAS and Swansea, but most especially Michael Thorne and Paul Seear, thanks for giving me so much of your time and for your constant interest. To those always ready to help with a smile in their faces despite the smells, the mess, the exploding tanks... Sandy Cordiner-Lawrie, Lucy Conway, Min Gordon, Paul Geissler, Elaine Fitzcharles and Elsa Couto, thank you!

I would like to thank everyone who helped me keep my sanity during these years. Thanks to my lab colleagues at Swansea and the students at BAS for all their support and encouragement, and for sharing good and bad times. Thanks to the coffee crew at stack-a-pleb, for the much needed breaks, the exotic coffee beans and for putting up with my moods and always cheering me up when it was most needed. But above all thank you all for your company and for making everything that little bit brighter. Thanks to all my friends outside this little world of science, for always being there and giving me so much credit. But most especially, thanks to my family, to César, Nori and Elvira, for it is their unconditional support and love that has got me where I am and always keeps me going. Thank you for always understanding me, even when it all made so little sense.

This work was carried out at the School of Medicine, University of Swansea, The British Antarctic Survey and CCMAR, University of the Algarve. The work was funded by the Marine Genomics Europe (MGE) Network (EU-FP6 contract No. GOCE-CT-2004-505403) and the BAS Q4 BIOREACH / BIOFLAME core.

ABBREVIATIONS

17P	17α-Hydroxy[1,2,6,7-^{3}H]-progesterone
1DE	One dimensional gel electrophoresis
2DE	Two dimensional gel electrophoresis
aa-tRNA	Aninoacyl- transfer ribonucleic acid
ACTH	Corticotrophin
AdoHcy	S-adenosylhomocysteine
AdoMet	Adenosylmethionine
AIC	Akaike Information Criterion
ANOVA	Analysis of variance
AOGCMS	Atmosphere-Ocean General circulation models
AR	Androgen receptor
ARE	Antioxidant responsive element
ATP	Adenosine triphosphate
BH	Benjamini and Hochberg
BHMT	Betaine homocysteine S methyltransferase
bp	Base pair
BSA	Bovine serum albumin
CA	Catecholamines
CaM	Calmodulin
cDNA	Complementary deoxyribonucleic acid
CHAPS	3-[(3-Cholamidopropyl)dimethylammonio]-1-propanesulfonate
cpm	Counts per minute
CRH	Corticotrophin releasing hormone
CSR	Cellular stress response
Ct	Threshold cycle
ddH_2O	Double distilled water
ds	Double stranded
DEPC	Diethylpyrocarbonate
Df	Degrees of freedom
DIGE	Difference gel electrophoresis
DMSO	Dimethylsulfoxide
DNA	Deoxyribonucleic acid
dNTP	Deoxyribonucleotide triphosphate
DTT	Dithiothreitol
EBI	European Bioinformatics Institute
EDTA	Ethylenediaminetetraacetic acid
ELISA	Enzyme-linked immunosorbent assay
ER	Endoplasmatic reticulum
ERα/β	Estrogen receptor α/β
ESI	Electrospray ionization
EST	Expressed sequence tag
FC	Fold change
FDR	False discovery rate
GAL	Genepix array list
GO	Gene ontology
GOLD	Genomes On-Line Database
GPR	Genepix results file

GR	Glucocorticoid receptor
GSH	Glutathione
GST	Glutathione-S-transferase
HA	Hyaluronic acid
HC	Heavy chain
HEPES	4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
HIF	Hypoxia inducible factor
HMG	High mobility group
HMGB	High mobility group box
HPA	Hypothalamic-pituitary-adrenal axis
HPI	Hypothalamic-pituitary-interrenal axis
HSP	Heat shock protein
IAA	Iodoacetamide
ID	Identifier
IEF	Isoelectric focusing
IPCC	Inter-Governmental Panel on Climate Change
IPG	Immobilised pH gradient
IαI	Inter alpha trypsin inhibitor
L. ratio	Log likelihood ratio
L:D	Light:dark
LB	Luria Bertani
LC	Liquid chromatography
LD PCR	Long distance polymerase chain reaction
LDLR	Low density lipoprotein receptor
Ma	Millions of years (mega annum)
MALDI-TOF	Matrix assisted laser desorption ionization time of flight
MDH	Malate dehydrogenase
mgf	Mascot generic file
ML	Maximum Likelihood
MPSS	Massively parallel signature sequencing
MR	Mineralocorticoid receptor
mRNA	Messenger ribonucleic acid
MS	Mass spectrometry
MS/MS	Tandem mass spectrometry
$MS_{between}$	Between group mean square
MS_{within}	Within group mean square
MW	Molecular weight
Mya	Million years ago
NADH	Nicotinamide adenine dinucleotide (reduced)
NCBI	National Center for Biotechnology Information
NMR	Nuclear magnetic resonance
nts	Nucleotides
P	Probability
p.p.t.	Parts per thousand
PAGE	Polyacrylamide gel electrophoresis
PAK	P21 activated kinase
PCR	Polymerase chain reaction
pI	Isoelectric point
PMT	Photomultiplier tube
PR	Progestin receptor

PRP19	Pre-messenger ribonucleic acid processing factor 19
Prx	Peroxiredoxin
qPCR	Quantitative polymerase chain reaction
Q-Q	Quantile-quantile
APS	Ammonium persulfate
REML	Restricted Maximum Likelihood
REST	Relative expression software tool
RIA	Radioimmunoassay
RNA	Ribonucleic acid
ROS	Reactive oxygen species
rRNA	Ribosomal ribonucleic acid
RT	Reverse transcriptase
SAGE	Serial analysis of gene expression
SAP	Shrimp alkaline phosphatase
SDS	Sodium dodecyl sulphate
SE	Standard error
Sn RNPs	Small nuclear ribonucleoprotein particles
SOC	Super optimal broth with catabolite repression
SOLiD	Sequencing by oligonucleotide ligation and detection
SSC	Saline-sodium citrate
SSH	Secure shell (Unix)
Tc	Critical temperature
TCP-1	T-complex protein 1
TEMED	Tetramethylethylenediamine
TFA	Trifluoroacetic acid
TIFF	Tagged Image File Format
TLC	Thin-layer chromatography
Tp	Pejus threshold
TRiC	T-complex protein ring complex
Tris	Tris(hydroxymethyl)aminomethane
tRNA	Transfer ribonucleic acid
U	Units
UV	Ultraviolet

For César, Nori and Elvira

1 INTRODUCTION

Since the first Inter-Governmental Panel on Climate Change (IPCC) Assessment Report in 1990 and allied initiatives such as the Millennium Ecosystem Assessment (http://www.millenniumassessment.org/), the debate has moved from "Is climate change really happening?" to discussions on prediction and mitigation. This has now become the main driver behind environmental research worldwide. In the UK, the latest strategy by the Natural Environment Research Council "Next Generation Science for Planet Earth" has understanding the effects of climate change on biodiversity as one of its key strategic goals. The study of "biodiversity" has many different aspects, but there is an increasing focus on functional aspects i.e. how an organism responds to environmental challenge with an assessment of their ability to cope. It is this area of functional biodiversity that this thesis addresses using laboratory studies of thermal challenge.

Temperature is acknowledged as one of the more prominent abiotic factors and its pervading effect on individuals and species has been of major interest in ecology since the early part of the last century (Krogh, 1916). In the light of climate change and the growing concern of human impact on the planet, thermal biology has come into renewed prominence. Of special interest in this regard is the situation where a temperature change induces a stress response in individuals or species and how this should be quantified in respect of the long-term consequences for the organism and population. Analysing the effects of temperature on living systems, from molecules to ecosystems, should reveal patterns of evolutionary adaptation and help predict how climate change might affect organism physiology and distribution (Hochachka and Somero, 2002). This project was aimed at evaluating the effect of acute exposure to elevated temperatures in two non model marine bivalve species, the Antarctic clam (*Laternula elliptica*) and the euythermal inter-tidal mussel (*Mytilus edulis*) in terms of their transcriptional and proteomic responses.

1.1 CLIMATE CHANGE AND BIODIVERSITY

Over the past decade, the IPCC has evaluated an overwhelming body of scientific evidence suggesting that the planet is undergoing a period of global warming, characterized by an increase in average surface temperatures of around 0.75°C since

the 19th century (Solomon et al, 2007). The latest IPCC report introduced six potential scenario families of future climate change predictions based on Atmosphere-Ocean General Circulation models (AOGCMS), each with a range of possible outcomes associated with it. According to current models, the beginning of the century will be characterised by a surface temperature increase of 0.64-0.69°C, almost irrespective of the chosen scenario. By mid-century, predictions range between increases of 1.3 and 1.7°C with predictions for the second half of the century largely dependent on the choice of scenario. These are more dramatic, ranging between increases of 1.8 and 4°C (IPCC, 2007). Global sea surface temperatures are predicted to increase by an average 2°C in the next century (IPCC, 2007). Additional to this, is the fact that certain areas are more vulnerable to climate change. The Antarctic Peninsula has experienced an increase in surface temperatures of 3°C in the last 50 years (Vaughan et al., 2003) over 2°C higher than the global average, whilst in the marine environment, shallow seawater temperatures in the Bellingshausen Sea have risen in excess of 1°C over the last 50 years (Meredith and King 2005). Although the predictions for the future show a high degree of uncertainty, it is clear, whichever the scenario, that both surface temperatures and oceanic sea surface temperatures are likely to undergo significant increases over the next century.

There is now ample evidence that these recent climatic changes have affected a wide range of species worldwide with consequences to the phenology and physiology of organisms, distributional ranges, species composition and interactions, and ecosystems structure and dynamics (Peñuelas and Filella, 2001; Walther et al., 2002). Most notably, global meta-analysis has confirmed significant range shifts towards the poles and advancement of spring events, in line with climate change predictions (Parmesan and Yohe, 2003). If predictions hold, warming over the next 100 years could have devastating effects on biodiversity worldwide. In order to predict the effects of a warming climate on organism physiology and biogeography, a quantitative understanding of the temperatures experienced by ectothermic organisms in nature is needed (Helmuth et al., 2002).

1.2 THE STRESS RESPONSE

Organisms exist within a thermal envelope directly related to habitat temperature. When environmental conditions change and this envelope is exceeded, individuals become vulnerable and experience what is generically described as a stress response. Responses to environmental challenge can be measured at many levels from ecological to behavioural, physiological and cellular. To date, most measures of species responses to environmental change have been either ecological or physiological (c.f. Walther et al., 2002; Peck, 2005; Pörtner et al., 2007; Chown and Gaston, 2008). But molecular investigation of physiological stress has provided finer detailed insights into capacity to cope (Gardeström et al., 2007; Kültz et al., 2007) and it is generally accepted that signs of stress will first appear at the molecular level. This has lead to the proposal of a generic universal cellular stress response (CSR), which comprises a series of biochemical changes aimed at maintaining homeostasis (Kültz, 2005). These biochemical changes, or rather the expression levels of the genes and proteins underlying them can potentially be used as molecular biomarkers to forecast changes in higher levels of biological organisation (Cajaraville et al., 2000; Dondero et al., 2006). Early detection of the effects of environmental perturbation via the CSR has invaluable applications across several fields including ecology, biogeography and aquaculture.

Experiments based on acute exposure to elevated temperatures provide an excellent tool for identifying internal mechanisms that are activated in response to thermal challenge. Such experimental approaches allow the identification of the limiting processes involved in an organism or species failing to cope with external challenge and consequently, via identification of the genes underlying this response, provide molecular biomarkers for species or population vulnerability.

1.3 STRESS AS A CONCEPT

Some terms related to stress and the stress response have ambiguous definitions in the literature. Living organisms survive by maintaining an immensely complex dynamic and harmonious equilibrium, or homeostasis, that is constantly challenged

or threatened by disturbing forces (Chrousos and Gold, 1992). "Stress" will be defined here as a state of disharmony, or threatened homeostasis. The disturbing forces, or stimuli that cause stress, will be referred to as "stressors" whether they occur in nature or are laboratory induced. "Stress" and "environmental stress" will be used interchangeably to refer to the response of an animal to an environmental perturbation (this can be varied in nature, such as salinity, temperature, hypoxia etc). The consequence of temperature acting as a stressor will be referred to as temperature stress, thermal stress or heat stress. The first two terms may be caused by exposure to unusually low or high temperatures, whereas the latter involves only temperature elevations. The term "stress response" will be used to refer to any response in an organism that occurs as a consequence of a stressor acting upon that organism. When this term refers strictly to the "cellular stress response" or the "classical stress response" this will be specified.

1.4 THERMAL STRESS AND MOLECULAR CHAPERONES

Non-lethal heat shock results in the activation of the CSR. This mechanism protects cells from sudden fluctuations in the environment, and ultimately reacts to the threat of macromolecular change (Kültz, 2003). Although some responses to heat stress, such as the expression of heat shock proteins, are well understood (Feder and Hofmann, 1999) others remain to be identified; and our understanding of a system wide response in environmentally relevant organisms is still limited.

The biologically relevant function of proteins is largely dependent on the formation and dissolution of chemical bonds. Some of these are weak, and largely affected by environmental factors, such as temperature (Feder, 1999). Alterations of these weak bonds may result in departure from the functional structure of the protein, and in turn present devastating consequences for the individual (Somero, 1995). When free, weak bonds can bind to other proteins, forming aggregates that become non functional and reduce the yield of proteins available for a particular process in the cell. In addition, aggregates can become cytotoxic and difficult to eliminate (Feder, 1999). Denatured proteins must be either degraded by cellular proteases or refold into their functional structure (Wickner et al., 1999). Molecular chaperones bind to

proteins in a non-native conformation and either allow them to attain their initial conformation, or make them available for degradation and removal from the cell (Hartl, 1996). Chaperones are universal and although they play a housekeeping role in the normal functioning of the cell, they are best known for being important molecular mechanisms of stress tolerance (Feder, 1999).

Cells contain at least two major chaperone systems, the heat shock proteins and the chaperonin system:

- Chaperonins: This contains two subgroups: type I chaperonins or members of the GroEL family found in eubacteria, mitochondria and chloroplasts, and type II chaperonins, members of the TRiC family (TCP-1 ring complex) found in archaebacteria and in the eukaryotic cytosol (Llorca et al., 1999). Members of the TRiC family are thought not to be stress inducible in eukaryotes and have a restricted range of substrates (Hartl, 1996).

- Heat shock proteins (HSPs): These are the best understood family of molecular chaperones. These are highly conserved, ubiquitous proteins (Kültz, 2005) designated and classified according to their molecular weight. The most widely studied family members are the 70 kDa proteins which comprise a constitutive (permanently expressed) and an inducible (expressed in response to stress) form. The latter is of particular importance for its role in assisting mis-folded proteins and targeting proteins degraded as a result of environmental insult (Parsell and Lindquist, 1993).

1.5 HEAT SHOCK PROTEINS AND THE CLASSICAL HEAT SHOCK RESPONSE

Although heat shock proteins (HSPs) play a role in normal cell physiology, folding and transporting native polypeptides (Feder and Hofmann, 1999), they were first described as proteins associated with heat shock responses. However, they are induced in the presence of practically every environmental stress studied (Feder and

Hofmann, 1999) and when induced in response to thermal stress, result, at least temporarily, in increased thermotolerance (Sonna et al., 2002).

The classical heat shock response, consisting on the expression of the inducible HSPs in the presence of a thermal challenge, has been identified in all species studied, with the exception of some Antarctic species (reviewed in Clark and Peck, 2009a). The ubiquity of HSPs, the high level of sequence conservation between species and their sensitivity to stressors of different types has made them prime candidates for stress biomarkers. However, despite the numerous studies on the behaviour of HSPs in response to temperature, their utility as universal stress markers is under scrutiny (Iwama et al, 2004) for several reasons:

- The induction of HSPs in response to stress, in terms of temperature thresholds, HSP concentrations and diversity, differs significantly as a function of the magnitude and timing of the stress, as well as the thermal history of the species (Buckley et al., 2001; Hofmann, 2005; Tomanek, 2008).

- Some species have higher threshold induction temperatures than the temperature at which the first signs of cellular stress appear (reviewed in Clark and Peck, 2009a).

- Because HSPs are induced by almost any stress (e.g. thermal, chemical, osmotic etc), changes in HSP expression cannot generally be attributed with confidence to a single environmental stress (Feder and Hofmann, 1999).

There is thus a need for more appropriate biomarkers for environmental stress caused by different stressors. While showing that an unexamined organism shows HSP expression in response to heat no longer tells us anything new (Feder, 1999); non-HSP mechanisms of stress tolerance in non model organisms are not well understood and require further examination. There is a need for a unifying concept explaining the universality of the stress response across different species and types of stress, beyond the expression of HSPs. Indeed, this response may take two forms: stress intolerant organisms may possess a particular set of proteins especially vulnerable to

stress on which molecular chaperones act to prevent or repair damage (Feder, 1999). The identity of these proteins is poorly understood and whether a similar set of vulnerable proteins exists across taxa remains unknown. There is also the molecular response to stress, which clearly in most species includes the induction of HSPs, but there may be other proteins which may be equally important in the ability of the organism to combat the effects of environmental stress.

Both the potential for identification of vulnerable proteins, and the need to identify mechanisms involved in the stress response, provide an excellent opportunity for the discovery of alternative molecular markers for environmental stress. This type of question cannot be efficiently answered by a targeted candidate gene approach, but requires large-scale screening of the transcriptome and proteome, as detailed below.

1.6 THE ENVIRONMENTAL STRESS RESPONSE AND THE GENOMIC REVOLUTION

Genome based technologies have the potential to revolutionize our understanding of the CSR. These are often referred to as the "omics" and include genomics, transcriptomics, proteomics and metabolomics, to study the genome, transcriptome, proteome and metabolome respectively (figure 1.1). The changes that allow the cell to respond dynamically to environmental stimuli can be measured via different techniques within these fields. They allow the discovery of differentially expressed genes, proteins and metabolites between stressed and unstressed individuals (Zivy and de Vienne, 2000). Understanding gene and protein expression and the changes involved in the response to environmental stress can help characterise the functional phenotype and determine the factors responsible for setting the limits of the ecological niche (Dupont et al., 2007).

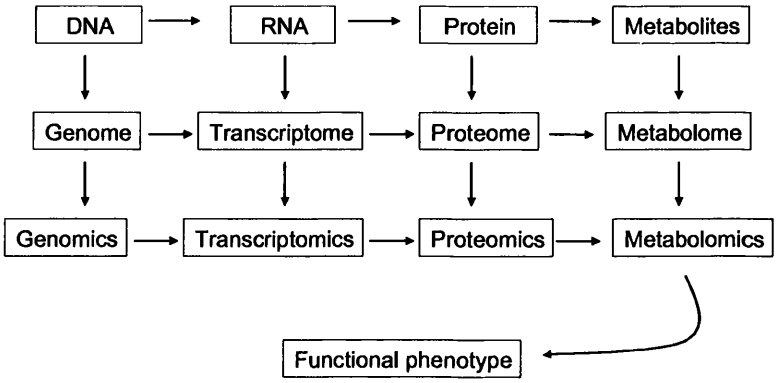


Figure 1.1 The link between genomics tools and the characterization of the functional phenotype. Horizontal arrows indicate information flow.

Some of the most widely used techniques in environmental studies are discussed below alongside examples of their application to study the stress response of non model marine organisms.

1.6.1 Transcriptomics

Transcriptomics is the term given to the study of gene expression. Which genes are expressed in particular cell types, at what levels and how these expression levels change under different circumstances is often called expression profiling. Comprehensive knowledge of the genes that change their expression pattern in response to stress can reveal much about the mechanisms that are turned on and those that become suppressed, in which tissues and over what time course. This is essential to increase our understanding of environmental plasticity (Cossins et al., 2006). The lists of genes that are identified as differentially expressed in response to stress through large scale screening, provide the basis for subsequent hypothesis testing at higher levels of biological organization by targeting specific genes and processes. Several approaches are available for the study of the transcriptome (figure 1.2) the most relevant of which are discussed below.

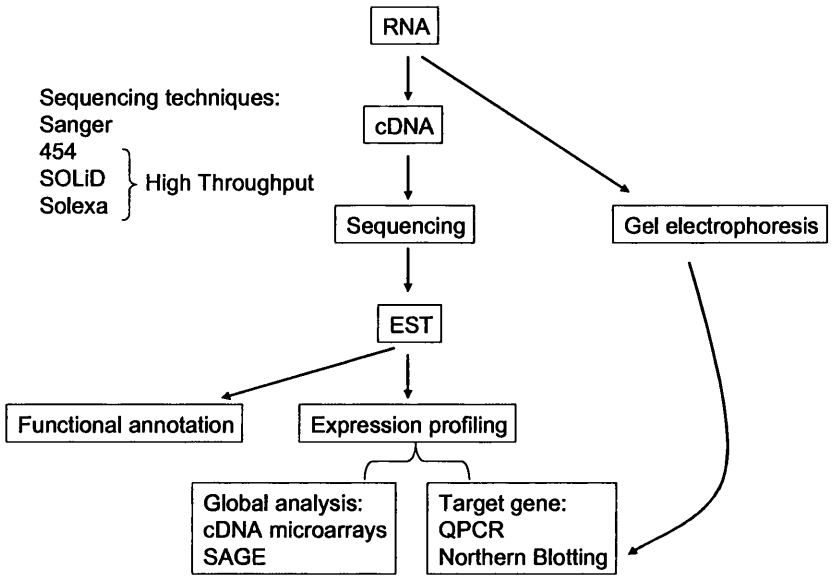


Figure 1.2 Some of the most widely used transcriptomics tools in ecological studies.

1.6.1.1 EST libraries and sequencing

An Expressed Sequence Tag (EST) is a single sequencing read from a cloned piece of cDNA. The production of a cDNA occurs when an RNA molecule corresponding to the expressed part of a gene is converted into a DNA copy to increase stability. Each clone is then usually sequenced from one end (generally the 5' end of the gene to avoid the complications of sequencing errors associated with the polyA tail). These EST sequence reads are generally short (300-800bp) and because they have only been sequenced once, often contain errors. However this technique does allow for relatively cheap gene discovery in organisms where there is very little sequence data. It also contributes to our understanding of the functionality of genes in non model species via sequence similarity searches with data from model species and provides data which can be used in further applications in taxonomic, evolutionary and systematics studies (Dupont et al., 2007).

An EST library, or collection of ESTs from the same individual, tissue or condition, can be produced via different methods depending on the characteristics required by the final application. If cDNA is made directly from the cells or tissues under investigation, the number of times a sequence is present in a library is a direct reflection of the quantity of the RNA corresponding to that gene in the cell. This

may mean that highly expressed sequences (e.g. actin in muscle) may mask the detection of rarely expressed clones and this is not efficient in terms of sequencing costs and gene discovery. Knowing the relative quantities of an RNA molecule (gene) across various tissues can be very useful, but generally techniques are used to circumvent this situation:

- Normalisation: Methods can be applied to equalize numbers of gene copies in the library and ensure the representation of genes transcribed at relatively low levels (Shcheglov et al., 2007).

- Subtraction: A subtraction library is made by removing identical genes present in two libraries from different conditions.

- Enrichment: A library is constructed from a large number of organisms from a specific condition in order to enrich transcripts induced by that particular treatment (Dupont et al., 2007).

With the reduction in the cost of sequencing and accessibility of small labs to sequencing equipment, EST sequence information in non model marine species is increasing considerably. GOLD (Genomes On-Line Database, http://genomesonline.org/index2.htm) gives a comprehensive listing of genome and EST projects. Although at the time of writing, NCBI holds approximately only 150,000 ESTs for bivalve molluscs, this information is increasing due to various European and international initiatives such as MytiBase. This database currently contains around 20000 high quality ESTs from several tissues from *Mytilus galloprovincialis* exposed to different pollutants, temperatures and bacteria (http://mussel.cribi.unipd.it/).

1.6.1.2 High through-put approaches

EST libraries are widely used for global analysis of gene expression in high throughput approaches, for which two conceptually different methodologies exist. The first group involves serial analysis of gene expression (SAGE) and massively parallel signature sequencing (MPSS), and the second group involves microarrays.

SAGE and MPSS are quantitative transcriptome profiling techniques that differ in their methodology but are essentially based on generating unique cDNA 'tags' that are sequenced to quantify transcript abundance and to identify the transcribed gene. From the relative abundance of the tags, gene expression profiles can be described (Hofmann et al., 2005a). These methods are not widespread in the analysis of species for which little sequence information is available as it is difficult to identify the tags (Thomas and Klaper, 2004). In particular, MPSS because of cost and technology requirements is restricted to some specialized laboratories (Morozova and Marra, 2008).

EST libraries (or rather the cDNA clones the ESTs are produced from) are often used for the production of microarrays, first described by Schena et al. (1995) and currently one of the most widely used protocols for transcriptome analysis. The arrays consist of a coated glass slide onto which a number of cDNA sequences are immobilised in an ordered array. Samples from a treated and a reference organism are labelled, each with a fluorescent dye, and these are then hybridised onto the array in competition with each other. Any mRNA sequence that shares enough similarity with one of the sequences on the array will bind non-covalently to that specific spot. As a result, the corresponding cDNA will be labelled a different colour, depending on whether the abundance of that particular sequence was greater in the sample or the reference. By measuring the signal of each spot via specific image analysis, the expression levels of thousands of genes is determined (Van Straalen and Roelofs, 2006). Using image analysis software and statistical tools, it is possible to determine which genes are up- or down-regulated, and which remain unchanged in relation to the expression levels in the control sample. Genes that show differences in expression are regarded as candidates for further analysis (figure 1.3). cDNA sequencing and gene identification of each of the clones on the array can be performed prior to printing. Alternatively, if costs are limited, it is possible to sequence only those clones of interest which have been shown to be differentially expressed in the experiment. Examples of the use of microarrays to elucidate changes in gene expression in response to temperature in non model marine species are given in chapter three.

The sequence data from ESTs can also be used to generate gene or clone-specific oligonucleotides to produce oligoarrays. These types of arrays offer some advantages because oligonucleotides can be selected from specific regions to discriminate between closely related genes and avoid cross hybridization. However, the selection of oligonucleotides depends on the availability of sequence data, thus this technique is generally more often used and most effective in well characterised model species (Gracey and Cossins, 2003).

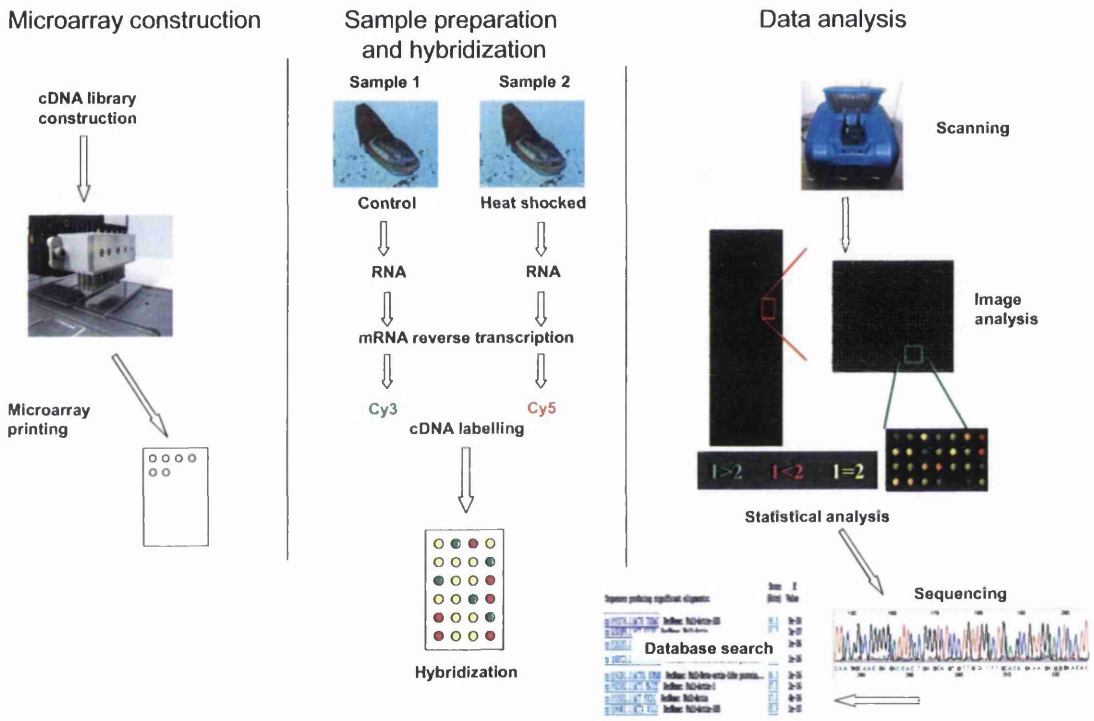


Figure 1.3 Pipeline for analysis of gene expression using microarrays. This includes microarray construction, sample preparation and hybridization and image and data analysis. Sequencing of the whole cDNA library can be performed before carrying out the experiment or alternatively, only those clones showing statistically significant changes in expression can be sequenced as represented here.

1.6.1.3 Gene quantification and expression validation

If a specific gene or set of genes is of particular interest, then targeted profiling can be carried out using those sequences. This involves the use of either Northern

blotting (Alwine et al., 1977) or the quantitative real time polymerase chain reaction (qPCR) approach (Kang et al., 2000).

Northern blotting involves the gel electrophoresis of tissue RNA extracts followed by capillary blotting onto a membrane. A radiolabelled probe is hybridised to the target and hybridization sites are visualized using an X-ray film and quantified by densitometry. Non-radioactive strategies are also available, which rely in chemiluminescence or fluorescence (Trayhurn, 1996). QPCR requires the design of primers from the gene sequences of interest. Amplifications are then performed in the presence of a DNA-binding fluorescent dye and the measured fluorescence is represented as a function of the PCR cycle number and related to the abundance of a transcript of interest. Both Northern blotting (Hofmann et al., 2005b) and qPCR (Cellura et al., 2006; Song et al., 2006; Park et al., 2007; Clark et al., 2008a; Clark et al., 2008b) have been extensively used for the analysis of targeted expressed genes, such as HSP expression at the transcript level. QPCR is also the favoured technique for microarray validation.

1.6.1.4 Next Generation Sequencing

The development of new high throughput sequencing techniques including 454 (Margulies et al., 2005), Solexa (Bennett et al., 2005) and SOLiD (Shendure et al., 2005) has minimised costs and labour associated with large scale sequencing, compared to the traditional Sanger method (Morozova and Marra, 2008). This technology has revolutionised our approach to gene expression studies and is just coming into use for environmentally relevant species (c.f. Vera et al., 2008; Hahn et al., 2009). *In vivo* cloning is substituted by clonal amplification of spatially separated molecules using either emulsion PCR (454, SOLiD) or bridge amplification on solid surface (Solexa) to produce shorter reads but increasing the number of base pairs (bp) per run (table 1.1). The next generation sequencing technologies will allow a more efficient characterization of the transcriptome and expression studies (using library tagging) to be carried out at a lower cost (Morozova and Marra, 2008). In non model species, 454 pyrosequencing has already been identified as a fast, cost effective tool for transcriptome assembly and for the development of ESTs and functional genomics tools, including the development of

oligoarrays for expression studies (Vera et al., 2008; Hahn et al., 2009). These areas are currently dominated by EST library production and Sanger sequencing, but that is changing. Although promising, these methods are still in their infancy, and robust verification and universal data analysis methods are still needed.

Table 1.1 Read lengths and length capacity per run for the traditional automated Sanger method compared to the more recent high throughput sequencing techniques. (Table adapted from Morozova and Marra, 2008)

Platform	Read length (bp)	Total length per run
Sanger	<900	96Kb*
454	200-300	80-120Mb
Solexa	30-40	1Gb
SOLiD	35	1-3Gb

* based on a 96 capillary array format

1.6.2 Proteomics

Transcriptomics, whilst identifying the mRNAs produced within a cell, produces no detail on how these are subsequently processed into proteins, the functional "end product". Proteomics is the study of proteins expressed by a genome (Williams and Hochstrasser, 1997). The proteome can be defined as the full complement of proteins expressed by the genome of a cell, a tissue or an organism at a specific time point (Williams and Hochstrasser, 1997). It can be regarded as a molecular phenotype (Biron et al., 2006), closer to the phenotype than the transcriptome. For expression studies, the study of the proteome requires the separation of complex protein mixtures, extracted from a cell, tissue or whole organism, followed by some form of relative quantification and identification of proteins of interest (summarised in figure 1.4).

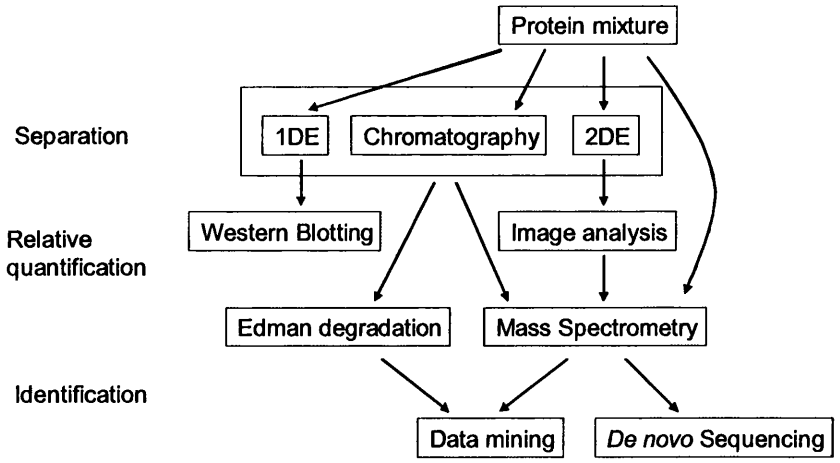


Figure 1.4 The most widely used proteomics tools in ecological studies. These include several techniques at the stages of protein separation, quantification and identification.

1.6.2.1 Separation and quantification techniques

There are several separation techniques including different formats of chromatography, of which the most widely used in proteomics is liquid chromatography (LC), one dimensional gel electrophoresis (1DE) and two-dimensional gel electrophoresis (2DE) (Twyman, 2004).

Chromatography is used to separate complex protein mixtures prior to other applications. These can then be subjected to mass spectrometric analysis for relative quantification as detailed below, or sequenced using Edman degradation, a technique that allows the sequencing of amino acids in a peptide (Edman, 1950).

1DE is often used to separate proteins on a gel matrix according to their molecular mass, prior detection of target proteins using antibodies via Western blotting (Renart et al., 1979). Western blotting has been used in non model marine species, to verify changes in mRNA expression, and determine which genes are translated and to what extent the corresponding protein is expressed (Evans and Somero, 2008). A vast number of studies have used Western blots to analyse HSP expression (Hofmann et al., 2000 Drew et al., 2001; Halpin et al., 2004; Hofmann, 2005). But perhaps the most powerful technique is 2DE which, first described by O'Farrell in 1975, allows

proteins to be separated in a gel matrix according to their isoelectric point (pI) and molecular weight (MW) (figure 1.5).

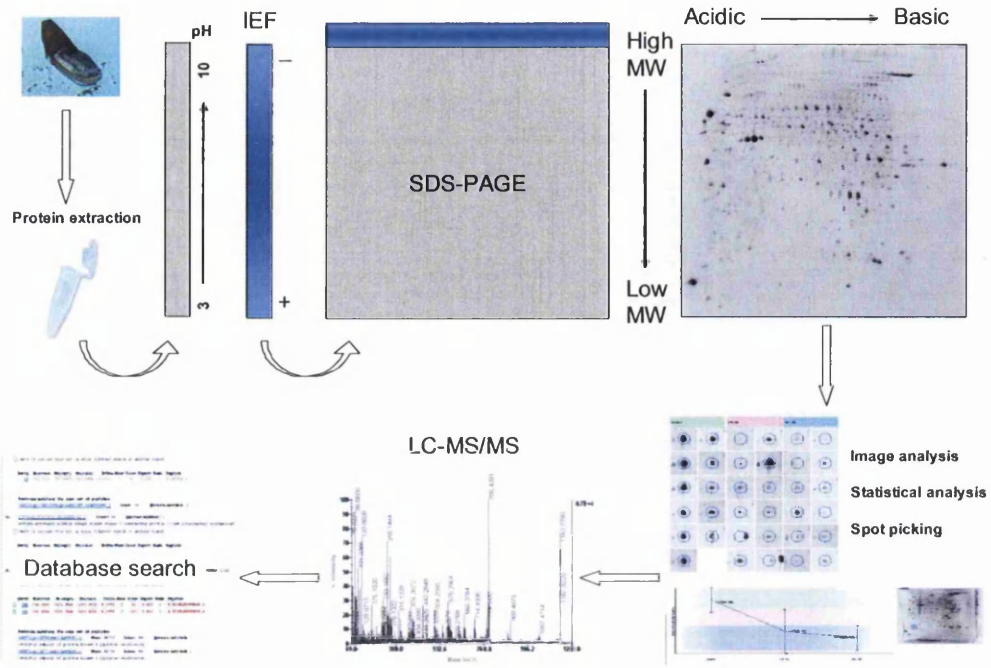


Figure 1.5 Pipeline for proteomics experiments using two dimensional gel electrophoresis, including sample preparation, first and second dimension gel electrophoresis, image and data analysis and identification of spots differentially expressed between the treatments via mass spectrometry. Alternative identification tools include Edman degradation and *de novo* sequencing.

2DE involves a combination of isoelectric focusing gel electrophoresis (IEF) and sodium dodecyl sulphate polyacrylamide gel electrophoresis (SDS-PAGE). Combined with non-specific staining, it allows the visualization of a large proportion of the expressed proteome exhibited by an individual in a particular cell or tissue, at a particular time. It is also possible to identify some post-translational modifications such as phosphorylation (Lopez, 2007) which cannot be predicted using transcriptome analysis. The quantification of each protein, represented as a spot on the gel, requires specific image analysis software. A number of packages are commercially available, which quantify each spot in terms of its intensity, directly related to the number of copies for that particular protein. Comparisons of spot intensities between treatments are then performed using carefully designed statistical

analysis. In common with microarray analysis, it is possible to determine which proteins are up- or down-regulated between the treatments, and those showing significant changes in expression levels become candidates for further analysis. Examples of studies of thermal stress in non model species using 2DE are discussed in chapter four.

Since its development in the 1970s, 2DE has undergone numerous improvements including standardization and simplification of the methodology, the ability to load larger amounts of protein with consequent identification of least abundant proteins, better resolution and reproducibility. A reduction of the cost has also made the technique more accessible to a larger number of laboratories. However, 2DE has not been used to the extent of transcriptomic analysis, particularly in environmental studies. This is probably due to its technical difficulties and time requirements (Lopez, 2007). A modification of the technique known as difference gel electrophoresis (DIGE), has overcome some of the technical limitations intrinsic to 2DE studies by introducing the possibility of analysing several samples labelled with fluorescent dyes in the same gel thus reducing the number of gels required and minimising technical variation (Alban et al., 2003).

1.6.2.2 Protein identification techniques

Regardless of the separation technique, protein identifications are currently performed using one of the available methods based on mass spectrometry (MS). In the early days of 2DE, proteins separated in the gels could only be identified by co-migration with those that had already been purified, or those for which an antibody was available for detection via immunoblotting studies. This is still used in Western blotting for the analysis of target proteins. Later developments into Edman microsequencing allowed the identification of previously known proteins, but also the *a priori* characterization of novel proteins. This method for protein identification is restricted by the amount of protein required, which is relatively large and requires several gels to be run for the identification of a single spot. As a result, spots corresponding with proteins with very low levels of expression may not be used in this type of analysis (Zivy and de Vienne, 2000). This problem was largely overcome in the 1990s with the development of several methods based on MS

(Yates, 1998). These techniques allowed the identification of peptides resulting from the digestion of a protein by a protease.

In matrix assisted laser desorption ionization time of flight mass spectrometry (MALDI-TOF), the masses of several peptides for a single protein are measured and fingerprints are produced, which are then compared with those predicted from sequences in the databases in order to find a match. This is applicable in species where a large proportion of the genome has been sequenced. For those organisms where few cDNAs have been sequenced, electrospray ionization tandem mass spectrometry (ESI-MS/MS) is usually the choice. The isolated peptides are further fragmented and the masses of the fragments are measured and a series of spectra produced, which can then be compared with those expected from database sequences. Although these techniques are very different in their outcome, they share a main limitation. The identification of their product relies on database searches, and is therefore dependent upon the availability of previously described peptides. An alternative to protein identification using theoretical spectra is *de novo* sequencing, where peptide sequences are inferred independently from the distribution of ions within the MS/MS spectrum and compared to sequences in databases using similarity search algorithms (Palagi et al., 2006). This method is limited to high quality spectra.

1.6.3 Metabolomics

One stage further on from proteomics is the study of the endogenous, low molecular weight metabolites within a cell, tissue or biofluid, defined as metabolomics (Bundy et al., 2009). As for transcriptomics and proteomics, hundreds of metabolites can be measured at once, allowing the characterization of metabolic processes affected by abiotic stressors. The most widely used techniques in large scale metabolomics are ^{1}H nuclear magnetic resonance spectroscopy (NMR) and MS (Viant, 2007). These are exploratory tools that allow the identification of unknown metabolites and are beyond the scope of this introduction. However, measurements of targeted metabolites can also be carried out on a smaller scale. These include hormones; whose concentration can be quantified using immunoassays, including radioimmunoassays (RIA) (Yalow and Berson, 1960) and enzyme-linked

immunosorbent assay (ELISA) (Engvall and Perlmann, 1972). RIA is an extremely sensitive technique, which relies on the competition between a labelled and an unlabelled antigen for the same antibody. Known amounts of labelled antigen and its specific antibody are allowed to react in the sample with the unknown amount of unlabelled antigen. The higher the antigen concentration in the sample, the less chance the labelled antigen will have of combining with the antibody. Antigen concentrations in the unknown sample are estimated by measuring the quantity of labelled antigen combined with antibody with isotope counting equipment. ELISA is a similar technique that relies on immunoreactivity, but substitutes the radiolabelled compounds with an enzyme that can be labelled with non-radioactive substances, minimising the risks of manipulating radioactive compounds.

1.7 CHOOSING A SINGLE DISCIPLINE OR AN INTEGRATED APPROACH

All of the technologies described above have their advantages and limitations. Gene expression profiling using transcriptomic tools allows the study of thousands of genes at the same time and is an invaluable tool to reveal candidate mRNAs related to stress. However, the role of the transcriptomic studies in identifying biological processes has been questioned as changes measured at the mRNA level do not necessarily reflect changes in the final gene products, the proteins (Anderson and Seilhamer, 1997; Gygi et al., 1999). The reason for this lies within the very basics of protein metabolism. Not only is there temporal variation in mRNA species expressed in a cell, but transcript translation also results in variations in protein expression. In addition, post-translational modification of proteins has the potential to greatly complicate the types of protein species produced from an individual mRNA (Chang et al., 2004). Even though the synthesis of proteins is closely related to mRNA abundance, its concentration is the net of synthesis and degradation, thus leading to variation in protein expression levels (Feder and Walser, 2005). Several studies, however, suggest that for the majority of genes, mRNA and protein abundance are highly correlated, although this is not the case for those genes whose expression is regulated post-translationally (Lu et al., 2007). Moreover, there is increasing evidence that transcriptional regulation is an essential means of biological regulation, thus all changes in transcript abundance are likely to be relevant when

evaluating potential mechanisms involved in the response to abiotic stressors (Podrabsky and Somero, 2004). Oleksiak et al. (2002) demonstrated a high degree of natural variation in gene expression, and suggested that evolution may depend more strongly on such variation than on differences between protein variants.

The study of the transcriptome therefore opens insights into one level of biological organization and is extremely useful to identify further testable hypotheses. It does not substitute, however, for the exploration of responses at the level of proteome or metabolome (Cossins et al., 2006). Parallel studies of the proteome and transcriptome overcome many of the limitations that arise from using them in isolation. They can increase our understanding of post-transcriptional and translational processes and the overall relationship between proteome and transcriptome can potentially be elucidated. Understanding how an organism interacts with its environment benefits from knowledge of both the transcriptome and proteome. Clearly, whichever the discipline, the techniques described here are most effective in species where there is a large amount of sequence data in the form of ESTs or draft genomes, however for the majority of environmentally relevant species, such resources are extremely limited. There are a number of difficulties in working with what are termed "non model" species.

1.8 ADVANTAGES AND LIMITATIONS OF WORKING WITH NON MODEL SPECIES

Non model species, in brief, are generally species for which there are often no inbred lines, no breeding stocks, little historical research, physiology, genomics or are intractable for genetic manipulation studies (Wilson et al., 2005). Indeed, some or all of these characteristics may apply and with some environmentally relevant species there are the additional problems of laboratory culture and supply (either in terms of numbers and/or the requirement for remote sampling, when in extreme environments). August Krogh, introduced the idea, currently termed Krogh's principle, that for any question a biologist asks there is an appropriate species on which it can be studied (Krebs, 1975).

Genome studies have classically been conducted on a small range of model species, typically *Homo sapiens, Drosophila melanogaster*, *Caenorhabditis elegans*, *Mus musculus*, *Arabidopsis thaliana*, *Saccharomyces cerevisiae* and *Danio Rerio.* However, this is changing as technology has moved towards a focus on draft genome coverage rather than the fully finished gold standard coverage associated with the human genome project (http://www.ensembl.org/index.html). From a genomics point of view, the criteria established to select model species do not necessarily coincide with those used in other fields, thus excellent candidates for genome analysis may not be of ecological importance (Dupont et al., 2007).

Nevertheless, model species, for which the whole genome is known, provide an excellent tool for the potential identification by database sequence similarity searching of unknown proteins or genes isolated from non model species. This works most effectively for those genes and proteins that are highly conserved between taxa. Although the lack of completed genome sequence in non model species does not necessarily prevent the study of changes in the proteome or transcriptome (Gracey and Cossins, 2003) it limits the extent to which such changes can be identified. A comparison of EST and protein sequence records available in NCBI between some multicellular model and non model species is shown in table 1.2. The non model species are represented by the bivalve genus *Mytilus*, which has some poorly characterised and some relatively well characterised species.

Table 1.2 The number of protein sequences and EST records held in NCBI in September 2009 for four multicellular model species (*D. melanogaster, C. elegans, M. musculus and A. thaliana*) and four non model species of the genus *Mytilus* (*M. californianus*, *M. edulis*, *M. galloprovincialis* and *M. trossulus*). Total numbers are shown, thus some degree of redundancy is expected.

Species	EST	Protein
D. melanogaster	847756	138911
C. elegans	325527	54079
M. musculus	4933819	262501
A. thaliana	1778434	394938
M. californianus	42364	828
M. edulis	4889	1470
M. galloprovincialis	19703	1206
M. trossulus	10	327

Despite being amongst the best characterised bivalve species, the genus *Mytilus* is poorly characterised when compared to other invertebrate model species, thus the assignment of functionality to gene sequences and proteins from this genus (using sequence similarity searches) is limited.

Notwithstanding the caveats, there is great advantage in not restricting studies of gene or protein expression to a limited number of genetically well characterised species. The development of genomic tools and resources largely relies on previously existent information, which results in the well studied organisms becoming better characterised, while non model species remain largely uncharacterised. Moving away from classic model organisms not only improves resources for a wider range of species, but allows the use of species best suited to address specific processes. With this in mind, the Antarctic clam *Laternula elliptica* and temperate mussel *Mytilus edulis* have been selected as representatives of the environments they inhabit, to understand the mechanisms involved in their response to thermal stress, and how these vary according to the characteristic of their habitat. A justification for these choices is provided below.

1.9 THERMAL HISTORY AND STRESS RESPONSES

The climate variability hypothesis suggests that the greater the magnitude of temperature variation in the environment, the wider the thermal tolerance windows of the ectothermic species that inhabit it (Stevens, 1989). Species inhabiting temperate climates should thus be able to survive a wider range of temperatures than species inhabiting the tropics or the poles. For example, a large scale study on the thermal limits of 17 bivalve species from tropical and temperate environments, demonstrated that tropical species have higher lethal thermal limits but thermal windows 7°C smaller on average than temperate species (Compton et al. 2007). In the marine environment, the extremes can be represented by the eurythermal temperate intertidal species and the highly stenothermal Antarctic subtidal species.

1.9.1 Eurythermy and the temperate intertidal

Temperate intertidal species undergo daily periods of emersion and immersion and are thus exposed to concomitant variations in temperature and other abiotic stressors (Menge and Branch, 1981). In addition, they are exposed to seasonal temperature variations, which are magnified during aerial exposure. The intertidal mussel *M. trossulus*, for example, is subjected to daily temperature changes often as high as 20°C depending on the season (Hofmann and Somero, 1995). Species of the genus *Mytilus* are abundant and ubiquitous, with a large geographical distribution. They inhabit a largely variable environment, subjected to a steep gradient in abiotic stress in the intertidal. Because of their accessibility and the characteristics of their environment, they have proven excellent models to understand the effect of thermal variation on a small spatial scale, and the effect of thermal history on HSP expression (Hofmann, 2005). Intertidal mussels have thus been the organism of choice in numerous studies on HPS expression (Hofmann and Somero, 1995; Smerdon et al., 1995; Hofmann and Somero, 1996; Chapple et al., 1997; Chapple et al., 1998; Buckley et al., 2001; Halpin et al., 2004; Cellura et al., 2006; Gonzalez-Riopedre et al., 2007). Other aspects of their heat shock response, however, have received limited attention. The organism under study, *M. edulis* (Linnaeus) lives attached to hard surfaces in the intertidal zone and has a wide geographical distribution (Gosling, 1984). It has an upper sustained thermal tolerance limit of 29°C (Almada-Villela et al, 1982; Read and Cumming, 1967) but the average lethal temperature was found to be as high as 40°C (Henderson, 1929). In Swansea Bay, in the Southwest of Wales (51°34′11″N, 3°58′49″W), minimum and maximum annual air temperatures raged between -0.8°C and 32°C during the year of collection (2006/07); whereas seawater temperatures varied seasonally from around 6°C to 18°C (data provided by Keith Naylor, Swansea University and Phil Glover, Pollution control, Swansea City Council)

1.9.2 Extreme stenothermality: Antarctic marine ectotherms

In contrast, the Southern Ocean is characterised by low but stable temperatures (Clarke et al., 2008). Antarctic marine ectotherms have evolved in a constantly cold environment for around 25–22 Ma (Clarke and Johnston, 1996) and are extremely

stenothermal (Somero and DeVries, 1967; Peck, 2002). Most species die with short term temperature rises of just 5-10°C but lose essential biological functions much earlier when seawater temperatures are raised just 2°C above current summer maxima (Peck et al., 2004). Their restricted thermal envelope, combined with the absence of other major anthropogenic influences in the environment, make them excellent candidates for the study of warming seawater temperatures. Biomarkers for heat stress in Antarctic species can thus be identified in the absence of confounding effects, such as pollution. The organism under study, the Antarctic clam *Laternula elliptica* (King and Broderip) is an abundant sediment burrowing bivalve. Around the British Antarctic Survey's Rothera Research Station on Adelaide Island, Antarctica (67°34'S, 66°8'W), where individuals for this study were collected, annual seawater temperatures at 15 m depth range from -1.8°C to 1.8°C (Clarke et al., 2008). Since marine ectotherms present body temperatures equivalent to those of their surroundings (Feder and Hofmann, 1999) the homogeneity of their environment will likely make this species highly susceptible to small temperature elevations.

1.9.3 Response to thermal challenge: eurythermy versus stenothermy

Given these data, it is expected that stenothermal species with restricted thermal windows will activate the stress response with much smaller temperature elevations than eurytherms. If we consider the induction of HSP expression as a sign of stress, this typically occurs 10-15°C above the optimum growth temperature for species that inhabit environments experiencing a broad range of temperatures. For organisms with more restricted ranges, maximum responses occur approximately 5°C above the optimum (Lindquist, 1986). Antarctic species represent a special case in their stress response, as initially revealed by studies with notothenoid fish. These have lost the ability to express the inducible form of HSP70 in response to stress (Hofmann et al., 2000). It has been suggested that this is due to a mutation in the promoter region of the gene that prevents binding of the HIF (hypoxia inducible factor) transcription factor necessary to initiate transcription (Buckley et al., 2004) and that this occurred 12-5 Mya (Hofmann et al., 2005b). However, these fish do permanently express this inducible form of HSP70 (Place et al., 2004) and this is thought to be due to the increased requirement for protein folding in response to increased denaturation at

low temperatures (Privalov, 1990). These studies, showing the lack of the classical heat shock response, raised the concern over whether the ability to respond to stress had been lost in Antarctic species. However, further analysis of the stress response in other Antarctic marine species revealed that some are able to induce HSP expression in response to elevated temperatures (Clark et al., 2008b). The induction temperature of the inducible form of HSP70 in these animals is in excess of +8°C, a temperature these animals never experience in the Antarctic marine environment. Hence the benefit of Antarctic species retaining the ability to induce HSP expression in response to heat stress is not immediately clear. Additional experiments have shown that other stressors induce HSP70 in these animals and also that alternative members of the HSP70 family are induced in response to both chronic and acute heat stress (Clark and Peck, 2009a; b). Hence, the situation with regards to heat shock proteins and the stress response is not as simple as previously thought, as exemplified in Iwama et al (2004) and Clark and Peck (2009a). These discuss examples which illustrate, even with studying three or four genes in isolation, how complex the stress response can be in different organisms and draws into question the universality of the CSR (Kültz, 2005).

In addition to experiencing stress with smaller temperature elevations, once their thermal window has been exceeded, some of the molecular mechanisms underlying the stress response of Antarctic bivalve species might differ from those observed in temperate species. It is thus possible that Antarctic marine species display mechanisms to maintain homeostasis as yet unidentified in their temperate counterparts. Comparative studies examining species inhabiting such different habitats can assist in the understanding of adaptation to stable and highly variable environments. Analysing their responses to change can provide important information on the relative susceptibility of ecosystems to climate change. The study of physiological, behavioural and metabolic responses is essential to understanding the extent of phenotypic plasticity in response to a changing environment (Helmuth, 2009).

1.10 OBJECTIVES AND THESIS STRUCTURE:

The main aim of this thesis is to analyse changes in gene and protein expression in response to elevated temperatures in the highly stenothermal Antarctic clam *L. elliptica* and the eurythermal temperate mussel *M. edulis*, and compare responses at the protein level. This will be achieved by expanding the currently limited genomic resources available for the Antarctic bivalve *L. elliptica*, by creating a custom made microarray. This could be used in future studies with other closely related species to determine the universality of the stress response by studying responses to other abiotic stresses.

Further aims were to i) identify potentially novel transcripts, proteins and metabolites associated with the heat stress response in these species, which can be proposed as candidates for biomarkers, ii) determine whether the physiological responses to elevated temperatures observed at the organism level are apparent at the level of gene and protein expression, and whether changes observed at the molecular level are accurate predictors on responses at higher levels of biological organization, iii) establish the link between the transcriptome and proteome in *L. elliptica* in response to heat stress, and iv) discuss the relative susceptibility of these species to climate change.

With these aims in mind, the thesis is divided into five further chapters: chapter two includes a description of the materials and methods used for sample collection, acclimation of animals to laboratory conditions and experimental manipulations. Chapter three presents the results from the analysis of gene expression profiles following heat stress in *L. elliptica* using a custom made microarray. The analysis of changes in protein expression in response to heat stress in *M. edulis* and *L. elliptica* using 2DE is described in chapter four. Chapter five describes the analysis of the endogenous synthesis of corticosteroids and their role as stress hormones in bivalves, and a general discussion is provided in chapter six.

1.11 REFERENCES

Alban A, David SO, Bjorkesten L, Andersson C, Sloge E, Lewis S, Currie I (2003) A novel experimental design for comparative two-dimensional gel analysis: Two-dimensional difference gel electrophoresis incorporating a pooled internal standard. Proteomics 3:36-44

Almada-Villela PC, Davenport J, Gruffydd LD (1982) The effects of temperature on the shell growth of young *Mytilus edulis* L. Journal of Experimental Marine Biology and Ecology 59(2-3):275-288

Alwine JC, Kemp DJ, Stark GR (1977) Method for detection of specific RNAs in agarose gels by transfer to diazobenzyloxymethyl-paper and hybridization with DNA probes. Proceedings of the National Academy of Sciences of the United States of America 74:5350-5354

Anderson L, Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. Electrophoresis 18:533-537

Bennett ST, Barnes C, Cox A, Davies L, Brown C (2005) Toward the $1000 human genome. Pharmacogenomics 6:373-382

Biron DG, Brun C, Lefevre T, Lebarbenchon C, Loxdale HD, Chevenet F, Brizard JP, Thomas F (2006) The pitfalls of proteomics experiments without the correct use of bioinformatics tools. Proteomics 6:5577-5596

Buckley BA, Place SP, Hofmann GE (2004) Regulation of heat shock genes in isolated hepatocytes from an Antarctic fish, *Trematomus bernacchii*. Journal of Experimental Biology 207:3649-3656

Buckley BA, Owen ME, Hofmann GE (2001) Adjusting the thermostat: the threshold induction temperature for the heatshock response in intertidal mussels (genus *Mytilus*) changes as a function of thermal history. Journal of Experimental Biology 204:3571-3579

Bundy JG, Davey MP, Viant MR (2009) Environmental metabolomics: a critical review and future perspectives. Metabolomics 5:3-21

Cajaraville MP, Bebianno MJ, Blasco J, Porte C, Sarasquete C, Viarengo A (2000) The use of biomarkers to assess the impact of pollution in coastal environments of the Iberian Peninsula: a practical approach. Science of the Total Environment 247:295-311

Cellura C, Toubiana M, Parrinello N, Roch P (2006) HSP70 gene expression in *Mytilus galloprovincialis* hemocytes is triggered by moderate heat shock and *Vibrio anguillarum*, but not by *V. splendidus* or *Micrococcus lysodeikticus*. Developmental and Comparative Immunology 30:984-997

Chang JS, Van Remmen H, Ward WF, Regnier FE, Richardson A, Cornell J (2004) Processing of data generated by 2-dimensional gel electrophoresis for statistical analysis: Missing data, normalization, and statistics. Journal of Proteome Research 3:1210-1218

Chapple JP, Smerdon GR, Hawkins AJS (1997) Stress-70 protein induction in *Mytilus edulis*: Tissue-specific responses to elevated temperature reflect relative vulnerability and physiological function. Journal of Experimental Marine Biology and Ecology 217:225-235

Chapple JP, Smerdon GR, Berry RJ, Hawkins AJS (1998) Seasonal changes in stress-70 protein levels reflect thermal tolerance in the marine bivalve *Mytilus edulis* L. Journal of Experimental Marine Biology and Ecology 229:53-68

Chrousos GP, Gold PW (1992) The concepts of stress and stress system disorders: overview of physical and behavioural homeostasis. Journal of the American Medical Association 267:1244-1252

Clark MS, Fraser KPP, Burns G, Peck LS (2008a) The HSP70 heat shock response in the Antarctic fish *Harpagifer antarcticus*. Polar Biology 31:171-180

Clark MS, Fraser KPP, Peck LS (2008b) Antarctic marine molluscs do have an HSP70 heat shock response. Cell Stress and Chaperones 13:39-49

Clark MS, Peck LS (2009a) HSP70 heat shock proteins and environmental stress in Antarctic marine organisms: A mini-review. Marine Genomics 2:11-18

Clark MS, Peck LS (2009b) Triggers of the HSP70 stress response: environmental responses and laboratory manipulation in an Antarctic marine invertebrate (*Nacella concinna*). Cell Stress and Chaperones 14:649-660

Clarke A, Johnston IA (1996) Evolution and adaptive radiation of Antarctic fishes. Trends in Ecology and Evolution 11:212-218

Clarke A, Meredith M, Wallace M, Brandon M, Thomas D (2008) Seasonal and interannual variability in temperature, chlorophyll and macronutrients in northern Marguerite Bay, Antarctica. Deep-sea research. Part II, Topical studies in oceanography 55:1988-2006

Compton TJ, Rijkenberg MJA, Drent J, Piersma T (2007) Thermal tolerance ranges and climate variability: A comparison between bivalves from differing climates. Journal of Experimental Marine Biology and Ecology 352:200-211

Cossins A, Fraser J, Hughes M, Gracey A (2006) Post-genomic approaches to understanding the mechanisms of environmentally induced phenotypic plasticity. Journal of Experimental Biology 209:2328-2336

Dondero F, Dagnino A, Jonsson H, Capri F, Gastaldi L, Viarengo A (2006) Assessing the occurrence of a stress syndrome in mussels (*Mytilus edulis*) using a combined biomarker/gene expression approach. Aquatic Toxicology 78:S13-S24

Drew B, Miller D, Toop T, Hanna P (2001) Identification of expressed HSP's in blacklip abalone (*Haliotis rubra* leach) during heat and salinity stresses. Journal of Shellfish Research 20:695-703

Dupont S, Wilson K, Obst M, Skold H, Nakano H, Thorndyke MC (2007) Marine ecological genomics: when genomics meets marine ecology. Marine Ecology-Progress Series 332:257-273

Edman P (1950) Method for determination of the amino acid sequence in peptides. Acta Chemica Scandinavica 4:277–282

Engvall E, Perlmann P (1972) Enzyme-linked immunosorbent assay. ELISA III. Quantitation of specific antibodies by enzyme-labelled anti-immunoglobulin in antigen-coated tubes. The Journal of Immunology 109:129-135

Evans TG, Somero GN (2008) A microarray-based transcriptomic time-course of hyper- and hypo-osmotic stress signalling events in the euryhaline fish *Gillichthys mirabilis*: osmosensors to effectors. The Journal of Experimental Biology 211:3636-3649

Feder ME, Hofmann GE (1999) Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annual Review of Physiology 61:243-282

Feder ME (1999) Organismal, ecological, and evolutionary aspects of heat-shock proteins and the stress response: Established conclusions and unresolved issues. American Zoologist 39:857-864

Feder ME, Walser JC (2005) The biological limitations of transcriptomics in elucidating stress and stress responses. Journal of Evolutionary Biology 18:901-910

Gardeström J, Elfwing T, Lof M, Tedengren M, Davenport J, Davenport J (2007) The effect of thermal stress on protein composition in dogwhelks (*Nucella lapillus*) under non-normoxic and hyperoxic conditions. Comparative Biochemistry and Physiology A Molecular and Integrative Physiology 148:869-875

Chown SL, Gaston KJ (2008) Macrophysiology for a changing world. Proceedings of The Royal Society B-Biological Sciences 275:1469-1478

Gonzalez-Riopedre M, Novas A, Dobano E, Ramos-Martinez JI, Barcia R (2007) Effect of thermal stress on protein expression in the mussel *Mytilus galloprovincialis* Lmk. Comparative Biochemistry and Physiology B-Biochemistry and Molecular Biology 147:531-540

Gosling EM (1984) The systematic status of *Mytilus galloprovincialis* in Western Europe: a review. Malacologia 25:551-568

Gracey AY, Cossins AR (2003) Application of microarray technology in environmental and comparative physiology. Annual Review of Physiology 65:231-259

Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. Molecular and Cellular Biology 19:1720-1730

Hahn DA, Ragland GJ, Shoemaker DD, Denlinger DL (2009) Gene discovery using massively parallel pyrosequencing to develop ESTs for the flesh fly *Sarcophaga crassipalpis*. BMC Genomics 10:234-243

Halpin PM, Menge BA, Hofmann GE (2004) Experimental demonstration of plasticity in the heat shock response of the intertidal mussel *Mytilus Californianus*. Marine Ecology Progress Series 276:137-145

Hartl FU (1996) Molecular chaperones in cellular protein folding. Nature 381:571-580

Helmuth B, Harley CDG, Haplin P, O'Donnell M, Hofmann G, Blanchette C (2002) Climate change and latitudinal patterns of intertidal thermal stress. Science 298:1015-1017

Henderson JT (1929) Lethal temperatures of *Lamellibranchiata*. Contributions to Canadian Biology, 4:399-410

Hochachka PW, Somero GN (2002) Biochemical Adaptation: Mechanism and Process in Physiological Evolution. Oxford University Press, New York. 466pp

Hofmann GE, Buckley BA, Airaksinen S, Keen J, Somero GN (2000) Heat-shock protein expression is absent in the Antarctic fish *Trematomus bernacchii* (family *Nototheniidae*). Journal of Experimental Biology 203:2331-2339

Hofmann GE, Burnaford JL, Fielman KT (2005a) Genomics-fueled approaches to current challenges in marine ecology. Trends in Ecology and Evolution 20:305-311

Hofmann GE, Lund SG, Place SP, Whitmer AC (2005b) Some like it hot, some like it cold: the heat shock response is found in New Zealand but not Antarctic notothenioid fishes. Journal of Experimental Marine Biology and Ecology 316:79-89

Hofmann GE, Somero GN (1995) Evidence for protein damage at environmental temperatures: seasonal-changes in levels of ubiquitin conjugates and hsp70 in the intertidal mussel *Mytilus trossulus*. Journal of Experimental Biology 198:1509-1518

Hofmann GE (2005) Patterns of Hsp gene expression in ectothermic marine organisms on small to large biogeographic scales. Integrative and Comparative Biology 45:247-255

Hofmann GE, Somero GN (1996) Interspecific variation in thermal denaturation of proteins in the congeneric mussels *Mytilus trossulus* and *M. galloprovincialis*:

Evidence from the heat shock response and protein ubiquitination. Marine Biology 126:65-75

IPCC (2007) Forth Assessment report. Working Group I Report (WGI): Climate change 2007: The physical science basis.

Iwama GK, Afonso LOB, Todgham A, Ackerman P, Nakano K (2004) Are hsps suitable for indicating stressed states in fish. Journal of Experimental Biology 207: 15-19

Kang JJ, Watson RM, Fisher ME, Higuchi R, Gelfand DH, Holland MJ (2000) Transcript quantitation in total yeast cellular RNA using kinetic PCR. Nucleic Acids Research 28:e2

Krebs H (1975) The August Krogh principle: "For many problems there is an animal on which it can be most conveniently studied". Journal of Experimental Zoology 194:221-226

Krogh A (1916) The Respiratory Exchange of Animals and Man. Longmans, Green and Co., London. 173pp

Kültz D (2003) Evolution of the cellular stress proteome: from monophyletic origin to ubiquitous function. Journal of Experimental Biology 206:3119-3124

Kültz D (2005) Molecular and evolutionary basis of the cellular stress response. Annual Review of Physiology 67:225-257

Kültz D, Fiol D, Valkova N, Gomez-Jimenez S, Chan SY, Lee J (2007) Functional genomics and proteomics of the cellular osmotic stress response in `non-model' organisms. Journal of Experimental Biology 210:1593-1601

Lindquist S (1986) The heat shock response. Annual Review of Biochemistry 55:1151-1191

Llorca O, McCormack E, Hynes G, Grantham J, Cordell J, Carrascosa J, Willison K, Fernandez J, Valpuesta J (1999) Eukaryotic type II chaperonin CCT interacts with actin through specific subunits. Nature 402:693-696

Lopez JL (2007) Two-dimensional electrophoresis in proteome expression analysis. Journal of Chromatography B-Analytical Technologies in the Biomedical 849:190-202

Lu P, Vogel C, Wang R, Yao X, Marcotte EM (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nature Biotechnology 25:117-124

Margulies M et al. (2005) Genome sequencing in microfabricated high-density picolitre reactors. Nature 437:376-380

Menge BA, Branch GM (2001) Rocky intertidal communities. In: Bertness MD, Gaines SD, Hay ME (Eds). Marine Community Ecology. Sinauer, Hamshire. pp221-251.

Meredith MP, King JC (2005) Rapid climate change in the ocean west of the Antarctic Peninsula during the second half of the 20th century. Geophysical Research Letters 32:L19604-L19609

Morozova O, Marra MA (2008) Applications of next-generation sequencing technologies in functional genomics. Genomics 92:255-264

O'Farrell PH (1975) High resolution 2-dimensional electrophoresis of proteins. Journal of Biological Chemistry 250:4007-4021

Oleksiak MF, Churchill GA, Crawford DL (2002) Variation in gene expression within and among natural populations. Nature Genetics 32:261-266

Palagi PM, Hernandez P, Walther D, Appel RA (2006) Proteome informatics I: Bioinformatics tools for processing experimental data. Proteomics 6:5435-5444

Park H, Ahn IY, Lee HE (2007) Expression of heat shock protein 70 in the thermally stressed Antarctic clam *Laternula elliptica*. Cell Stress and Chaperones 12:275-282

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. Nature 421:37-42

Parsell DA, Lindquist S (1993) The function of heat-shock proteins in stress tolerance: degradation and reactivation of damaged proteins. Annual Review of Genetics 27:437-496

Peck LS (2002) Ecophysiology of Antarctic marine ectotherms: limits to life. Polar Biology 25:31-40

Peck LS, Webb KE, Bailey DM (2004) Extreme sensitivity of biological function to temperature in Antarctic marine species. Functional Ecology 18:625-630

Peck LS (2005) Prospects for survival in the Southern Ocean: vulnerability of benthic species to temperature change. Antarctic Science 17:497-507

Peñuelas J, Filella I (2001) Phenology: Responses to a warming world. Science 294:793-795

Place SP, Zippay ML, Hofmann GE (2004) Constitutive roles for inducible genes: evidence for the alteration in expression of the inducible hsp70 gene in Antarctic notothenioid fishes. American Journal of Physiology. Regulatory, Integrative and Comparative Physiology 287:R429-R436

Podrabsky JE, Somero GN (2004) Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. Journal of Experimental Biology 207:2237-2254

Pörtner HO, Peck LS, Somero GN (2007) Thermal limits and adaptation in marine Antarctic ectotherms: an integrative view. Philosophical Transactions of the Royal Society B-Biological Sciences 362:2233-2258

Privalov PL (1990) Cold denaturation of proteins. Critical Reviews in Biochemistry and Molecular Biology 25:281-305

Read KRH, Cumming KB (1967) Thermal tolerance of the bivalve molluscs *Modiolus modiolus* (L.), *Mytilus edulis* (L.) and *Brachidontes demissus* (Dillwyn), Comparative Biochemistry and Physiology, 22:149-155

Renart J, Reiser J, Stark GR (1979) Transfer of proteins from gels to diazobenzyloxymethyl-paper and detection with antisera: method for studying antibody specificity and antigen structure. Proceedings of the National Academy of Science 76:3116-3120

Schena M, Shalon D, Davis RW, Brown PO (1995) Quantitative monitoring of gene-expression patterns with a complementary-DNA microarray. Science 270:467-470

Shcheglov A, Zhulidov P, Bogdanova E, Shagin D (2007) Generation of normalized cDNA libraries. In: Buzdin A, Lukyanov S (Eds.) (2007) Nucleic Acids Hybridization: Modern Applications, Springer, New York. pp. 97-124

Shendure J, Porreca GJ, Reppas NB, Lin X, McCutcheon JP, Rosenbaum AN, Wang MD, Zhang K, Mitra RD, Church GM (2005) Accurate multiplex polony sequencing of an evolved bacterial genome. Science 309:1728-1732

Smerdon GR, Chapple JP, Hawkins AJS (1995) The simultaneous immunological detection of 4 Stress-70 protein isoforms in *Mytilus edulis*. Marine Environmental Research 40:399-407

Solomon S, et al. (2007) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Somero GN (1995) Proteins and temperature. Annual Review of Physiology 57:43-68

Somero GN, DeVries AL (1967) Temperature Tolerance of Some Antarctic Fishes. Science 156:257-258

Song L, Wu L, Ni D, Chang Y, Xu W, Xing K (2006) The cDNA cloning and mRNA expression of heat shock protein 70 gene in the haemocytes of bay scallop (*Argopecten irradians*, Lamarck 1819) responding to bacteria challenge and naphthalin stress. Fish and Shellfish Immunology 21:335-345

Sonna LA, Fujita J, Gaffin SL, Lilly CM (2002) Invited Review: Effects of heat and cold stress on mammalian gene expression. Journal of Applied Physiology 92:1725-1742

Stevens GC (1989) The latitudinal gradient in geographical range: how so many species coexist in the tropics. American Naturalist 133:240-256

Thomas MA, Klaper RD (2004) Genomics for the ecological toolbox. Trends in Ecology and Evolution 19:439-445

Tomanek L (2008) The importance of physiological limits in determining biogeographical range shifts due to global climate change: the heat shock response. Physiological and Biochemical Zoology 81:709-717

Trayhurn P (1996) Northern blotting. Proceedings of the Nutrition Society 55:583-589

Twyman RM (2004) Principles of proteomics. BIOS Scienftific Publishers. Oxford. 241pp.

Van Straalen NM, Roelofs, D (2006) An introduction to ecological genomics. Oxford University Press, Oxford. 320pp.

Vaughan DG, Marshall GJ, Connolley WM, Parkinson C, Mulvaney R, Hodgson DA, King JC, Pudsey CJ, Turner J (2003) Recent rapid regional climate warming on the Antarctic Peninsula. Climatic Change 60:243-274

Vera JC, Wheat CW, Fescemyer HW, Frilander MJ, Crawford DL, Hanski I, Marden JH (2008) Rapid transcriptome characterization for a nonmodel organism using 454 pyrosequencing. Molecular Ecology 17:1636-1647

Viant MR (2007) Metabolomics of aquatic organisms: the new 'omics' on the block. Marine ecology progress series 332:301-306

Walther GR, Post E, Convey P, Menzel A, Parmesan C, Beebee TJT, Fromentin JM, Hoegh-Guldberg O, Bairlein F (2002) Ecological responses to recent climate change. Nature 416:389-395

Wickner S, Maurizi MR, Gottesman S (1999) Posttranslational quality control: Folding, refolding, and degrading proteins. Science 286:1888-1893

Williams KL, Hochstrasser DF (1997) Introduction to the proteome. In: Wilkins MR, Williams KL, Appel RD, Hochstrasser DF (Eds). Proteome research: new frontiers in functional genomics. Springer, New York, pp 1–12

Wilson K, Thorndyke M, Nilsen F, Rogers A, Martinez P (2005) Marine systems: moving into the genomics era. Marine Ecology: An Evolutionary Perspective 26:3-16

Yalow RS, Berson SA (1960) Immunoassay of endogenous plasma insulin in man. Journal of Clinical Investigation 39:1157-1175

Yates JR (1998) Mass spectrometry and the age of the proteome. Journal of Mass Spectrometry 33:1-19

Zivy M, de Vienne D (2000) Proteomics: a link between genomics, genetics and physiology. Plant Molecular Biology 44:575-580

2 General Materials and Methods

L. elliptica and *M. edulis* specimens were collected in 2006 and treated according to the methodology described below. Samples were used throughout the experiments for the analysis of gene and protein expression, and the identification of stress hormones. Methodology specific to each chapter will be discussed separately.

2.1 SAMPLE COLLECTION

2.1.1 *L. elliptica*

Forty *L. elliptica* individuals (shell length 51.10±10.4 mm (mean±SE)) were collected by scuba divers at a depth of 10-18 m in February 2006 at North Cove, Rothera Point, Adelaide Island, Antarctic Peninsula (67°34′07″S, 68°07′30″W).

a) b)

Figure 2.1 Representative example of a) the substrate inhabited by *L. elliptica* with buried individuals and visible siphons on the surface (provided by Martin White, British Antarctic Survey) and b) *L. elliptica* individuals underwater with extended siphon (provided by Alfred Wegener Institute).

L. elliptica live permanently buried in the sediment accessing the surface via the siphon. The substratum consisted of fine sediment overlaying gravel and cobbles (figure 2.1). After collection, the animals were returned to the laboratory and maintained in a flow-through aquarium under a simulated natural light regime controlled by a mechanical timer (Poltips 3, Proudman Oceanographic Laboratory, Birkenhead). Water temperatures ranged between 0 and -0.1°C. Animals were shipped to the British Antarctic Survey facilities in Cambridge, UK and allowed to acclimate to laboratory conditions for four weeks before the experiments were

carried out. During the acclimation period, animals were kept in plastic meshed baskets suspended in a 500 l tank with a circulating water system. Water was brought from the North Sea and biological, mechanical and chemical filters were used through the system in order to ensure high water quality. The water was maintained at a constant temperature of 0±0.5°C and salinity of 34±1 p.p.t. Animals were exposed to a L:D cycle of 12 h:12 h throughout the four weeks. A marine microalgae (*Nannochloropsis* species) concentrate (Reed Mariculture) was added to the water on a weekly basis according to manufacturer instructions.

2.1.2 *M. edulis*

M. edulis individuals (shell length 48.73±6.9 mm (mean±SE)) were collected at low tide in July 2006 from Mumbles, Swansea Bay, UK (51°34′11″N, 3°58′49″W). Intertidal blue mussels live attached to hard surfaces and are exposed to periods of immersion and emersion on a daily basis. The mussels selected were attached to the pier and were collected during emersion at low tide (figure 2.2).

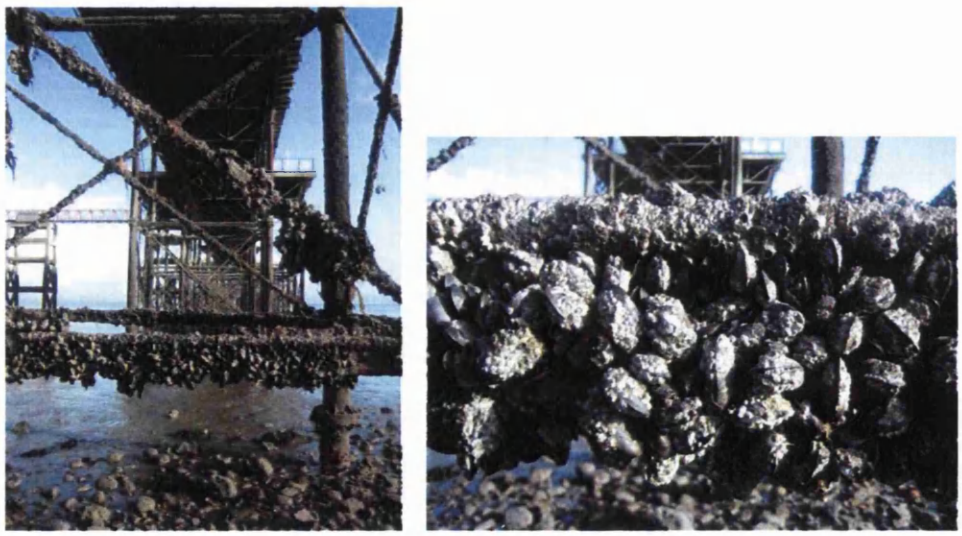


Figure 2.2 *M. edulis* attached to the pier at Mumbles, Swansea at the time of collection (July 2006) shown during tidal emersion.

After collection, the animals were transported to the laboratory and kept in a 15 l aquarium. Animals were allowed to acclimate for four weeks to laboratory conditions before further experiments were carried out. During the acclimation

period the aquarium was kept in a constant temperature room to ensure water was maintained at 12±1°C. Salinity was kept constant at 31±1 p.p.t and animals were exposed to a L:D cycle of 12 h:12 h. In order to ensure water quality, the tank was emptied weekly using suction through a pipe to avoid disturbing the animals, and the water was replaced with clean seawater, previously cooled to 12°C. Mussels were fed daily approximately 500 ml of a live *Tetraselmis suecica* culture. Typical densities were around 1 million cell ml^{-1}. Algae were harvested from batch culture at the Centre for Sustainable Aquaculture Research at Swansea University.

2.1 HEAT SHOCK EXPERIMENTS

A control and two temperature treatments were carried out for each species, corresponding to a moderate and extreme heat shock. These were carried out at different temperatures for each species to reflect habitat differences and thermal histories. The length of the exposure to elevated temperatures also varied between species. A justification for the selected experimental conditions is provided below.

2.1.3 *L. elliptica*

At the end of the acclimation period, three separate heat shock experiments in *L. elliptica* were performed at the selected temperatures of 0, 3 and 9°C (±0.5). The temperature for each treatment was selected based on previous studies (Peck et al., 2004) where *L. elliptica* was shown to suffer 50% failure in essential biological activities at 3°C and survive only a few days at 9°C.

The length of the heat shock for *L. elliptica* was chosen based on previous experiments on HSP expression, which showed no induction after 4-6h exposure (L. Peck pers. comm.) and knowledge about metabolic rates in Antarctic species. Metabolic rates of Antarctic benthic ectotherms are significantly lower from those in temperate and tropical ecosystems (Clarke, 2003); with most physiological rates two to five times slower than in temperate species (Peck, 2002). The length of heat exposure in *L. elliptica* was prolonged accordingly to ensure that a slow response was not missed.

All treatments were carried out in the same tank and conditions but at separate times. All animals within the treatment were placed in the tank at the same time and treated equally. The animals were placed in a 60 l sedimentation tank with an approximately 15 cm deep sand substratum and aerated seawater. The sand surface was covered with a plastic tray, to prevent the animals from burying until the experimental temperature was reached. A thermo-circulator (Grant LTD 20G, Grant Instruments LTD) was used to gradually increase the temperature over a 12 h period to the maximum temperature at each treatment. At the end of the 12 h period, the tray was removed and the animals where left in contact with the sand to allow them to bury. The temperature was maintained constant for 24 h, after which period the number of buried animals was noted. Approximately 90% individuals reburied at 0°C, 25% at 3°C and none at 9°C (n = 10, 20 and 10 for 0, 3 and 9°C respectively). Individuals exposed to a moderate heat shock were divided into two groups named 3°C buried and 3°C non buried. After the heat shocks, each animal was individually removed from the tank and processed.

2.1.4 *M. edulis*

Three separate heat shock experiments were carried out at the end of the acclimation period at the selected temperatures of 12, 27 and 37°C. Temperatures were selected based on previous studies on *M. edulis*. The control temperature was 12°C, the annual average seawater temperature in Swansea Bay. 27°C approximately reflects the temperature experienced at the southern limit of *M. edulis* distribution (Clay, 1965) and it is just below the temperature at which individuals become unable to acclimate, estimated at 28.5°C (Chapple et al., 1998). The upper lethal temperature in *M. edulis* has been estimated at between 31 and 41°C (Davenport and Davenport, 2005; Henderson, 1929), thus 37°C was used as the extreme heat shock.

Changes in HSP expression indicate the activation of a heat shock response, which occurs after a short period of exposure to elevated temperatures, thus exposure length was selected based on previous studies on HSP expression on *M. edulis*. These have varied in the length of exposure, rate of warming and heat shock temperature. Gonzalez-Riopedre et al. (2007), for example, showed that the expression of some

members of the HSP70 family increased significantly in *M. edulis* after 5 min of exposure at 45°C. Chapple et al. (1997), however, observed that the expression of some HSPs increased significantly after 2 h of exposure at 28°C. In the present study, ten animals were exposed to each treatment as before and the temperature in each tank was adjusted using aquarium heaters. Once the animals had been placed in the tank, water temperature was gradually increased to the maximum for each treatment over a 2 h period, at the end of which the animals were left at that constant temperature for another 2 h.

2.2 SAMPLE PREPARATION

Immediately after the heat shocks, individuals were removed from the tanks. Haemolymph samples were extracted from all individuals following a different protocol for each species. In *L. elliptica*, haemolymph was taken from the central sinus of each animal by injecting the tip (3 mm) of a syringe needle through the ligament as close to the posterior side of the hinge as possible. In the case of *M. edulis*, haemolymph was extracted from the adductor muscle of each individual by making a small incision at the edge of the shell close to the anterior adductor mussel and inserting a thin needle. Sample volumes were <1 ml varying accordingly with the size of the animal and the success of the extraction. Haemolymph samples were placed in 1.5 ml microtubes and stored at -80°C until further use.

Following haemolymph extraction in *L. elliptica*, the shell was forced open and tissue samples from the mantle, siphon, gill and digestive gland were taken. Each tissue sample was cut into four pieces, each of which was individually wrapped in foil and flash-frozen in N_2 (l) before storage at -80°C. *M. edulis* individuals were laterally bisected and the foot was excised, flash frozen and stored at -80°C until further use. All dissections were completed within 2 h from the subtraction of the first animal from the tank. All experiments were carried out at the same time of the day in order to exclude potential differences as a result of daily patterns in metabolism.

2.3 TISSUE SELECTION

Proteomic and transcriptomic responses from the same tissue in two different species might be very different for physiological reasons (Sheehan and McDonagh, 2008). Accordingly, the most appropriate tissue was selected for each species independently based on previous recommendations and preliminary work. In the case of *M. edulis*, the foot was selected for protein extraction following the rationale presented in Lopez (2005). The foot is likely to show clearer 2DE patterns, be less subjected to seasonal changes in protein expression during the gametogenic cycle and present less contaminating exogenous proteins than the mantle, branchia and digestive gland. For *L. elliptica*, preliminary experiments were carried out with all tissues and 2D gels were successfully produced with clear reproducible patterns in all cases except for proteins extracted from the digestive gland (details provided in chapter four). The mantle was selected based on preliminary work producing cDNA. The RNA extraction protocol had to be optimized in order to achieve high RNA quality, and mantle RNA produced the highest quality, with reproducible results. Mantle tissue was thus selected for further analysis.

2.4 REFERENCES

Chapple JP, Smerdon GR, Berry RJ, Hawkins AJS (1998) Seasonal changes in stress-70 protein levels reflect thermal tolerance in the marine bivalve *Mytilus edulis* L. Journal of Experimental Marine Biology and Ecology 229:53-68

Clarke A (2003) Costs and consequences of evolutionary temperature adaptation. Trends in Ecology and Evolution 18:573-581

Clay, E. (1965) Literature survey of the common fauna of estuaries. 14. *Mytilus edulis* L. Imperial Chemical Industries Ltd., Brixham Laboratory, UK.

Davenport J, Davenport JL (2005) Effects of shore height, wave exposure and geographical distance on thermal niche width of intertidal fauna. Marine Ecology-Progress Series 292:41-50

Gonzalez-Riopedre M, Novas A, Dobano E, Ramos-Martinez JI, Barcia R (2007) Effect of thermal stress on protein expression in the mussel *Mytilus galloprovincialis* Lmk. Comparative Biochemistry and Physiology B-Biochemistry and Molecular Biology 147:531-540

Henderson JT (1929) Lethal temperature of *Lamellibranchiata.* Contributions to Canadian Biology and fisheries, 4, 397-412

Lopez JL (2005) Role of proteomics in taxonomy: the *Mytilus* complex as a model of study. Journal of Chromatography B-Analytical Technologies in the Biomedical and Life Sciences 815:261-274

Peck LS (2002) Ecophysiology of Antarctic marine ectotherms: limits to life. Polar Biology 25:31-40

Peck LS, Webb KE, Bailey DM (2004) Extreme sensitivity of biological function to temperature in Antarctic marine species. Functional Ecology 18:625-630

Sheehan D, McDonagh B (2008) Bivalves and oxidative stress: A proteomics approach. Invertebrate Survival Journal 5:110-123

Chapple JP, Smerdon GR, Hawkins AJS (1997) Stress-70 protein induction in *Mytilus edulis*: Tissue-specific responses to elevated temperature reflect relative vulnerability and physiological function. Journal of Experimental Marine Biology and Ecology 217:225-235

3 Analysis of Gene Expression following Heat Stress in *L. elliptica*

3.1 INTRODUCTION

3.1.1 Limits to survival in *L. elliptica* under acute thermal stress

Temperature is one of the most prominent abiotic stressors, and the ability of an individual or species to cope with such stress has been a major interest in ecology since the early part of last century (Krogh, 1916). Current understanding of the mechanisms limiting species exposed to elevated experimental temperatures centres around limitations to physiological capacities (Pörtner, 2002; Pörtner et al., 2006). The oxygen limitation hypothesis (Pörtner, 2002) suggests that a drop in aerobic scope is the first mechanism to limit survival at the low and high ends of the thermal envelope. The loss of aerobic scope involves a mismatch between oxygen demand and supply to tissues, caused by capacity limitations of the circulatory and ventilatory systems. Under acute exposure to elevated temperatures, oxygen limitation has been demonstrated to limit thermal tolerance (Pörtner et al., 2006) and it is likely to be the mechanism dictating survival limits.

Under the oxygen limitation hypothesis, pejus (=getting worse) thresholds (Tp) indicate the limits of optimum haemolymph oxygenation and thus designate the temperature at which aerobic scope falls from optimal levels and performance starts to decline. This marks the limit of the temperature range which allows long term survival. With further warming, aerobic scope decreases and critical temperatures (Tc) are reached. These define the point at which anaerobic products begin to accumulate and have been suggested to mark the physiological limit for survival (Pörtner, 2002). After this threshold, survival is passive and relies on anaerobic metabolism. Under normoxia and acute temperature exposure, *L. elliptica* suffers 50% failure in the ability to burrow, an essential biological activity, at 2-3°C, and complete loss at 5°C (Peck et al., 2004). The upper Tp in *L. elliptica* has been estimated around 4.3°C based on analysis of haemolymph oxygen tensions (Pörtner et al., 2006). Beyond this temperature, a significant reduction in aerobic scope occurs until an upper critical temperature is reached at around 6°C, when the transition to anaerobic metabolism occurs as indicated from the level of tissue succinate accumulation (Peck et al., 2002). The upper lethal temperature has been

shown to be around 9°C (Peck et al., 2002). Hyperoxia alleviates these limitations through improved oxygen uptake into haemolymph circulation. Doubling the oxygen content of ambient seawater raises the upper temperature limit in *L. elliptica* by 2.5°C (Pörtner et al., 2006), providing support for the oxygen limitation hypothesis.

3.1.2 The molecular basis of stress in *L. elliptica*

Much of the work carried out in *L. elliptica* relates to physiological responses to elevated temperatures. Physiological studies have identified the thermal limits and changes in metabolism in response to temperature and oxygen availability (Peck et al., 2002; Pörtner et al., 2006; Peck et al., 2007). Movement and burrowing capacity (Peck et al., 2004) and seasonal energetics with consequences to activity (Morley et al., 2007) have also been examined. The factors determining maximum life span, characteristically long in Antarctic species, have also been under investigation (Philipp et al., 2005a; Philipp et al., 2005b). Metabolic studies include lipid radical (Estevez et al., 2002) and reactive oxygen species generation (Heise et al., 2003), both examined in relation to temperature.

At a molecular level, a very limited amount of work has been carried out, thus *L. elliptica* is extremely poorly characterized in genome databases. A database search in GenBank in July 2009 retrieved 18 nucleotide and 14 protein sequences, mainly from partial mRNAs which included three glutathione-S-transferase (GST) genes, two peroxiredoxin genes, actin, a number of ribosomal and heat shock protein genes and two metallothionein genes. To date, molecular studies using *L. elliptica* as the species of choice include HSP expression (Park et al., 2007; Clark et al., 2008), the study of antioxidant systems with the characterization of a GST (Kim et al., 2009) and two peroxiredoxin genes (Park et al., 2008a). At the protein level, an investigation of the effect of thermal stress on antioxidant defence systems using enzyme assays has been carried out (Park et al., 2008b).

3.1.3 Microarrays and thermal stress in non model organisms

An emergent application of microarray technology is the study of the molecular basis of physiological responses to stress. The aim is to identify genes that are differentially expressed in response to stress, which may provide the basis for subsequent hypothesis testing. As previously discussed, although some responses to heat stress are well understood, our understanding of a system wide response is still limited.

The availability of full genome sequences has allowed the construction of whole genome microarrays in model species. This has enabled the successful characterization of the basis of their heat stress response at the level of gene expression. Examples of this include studies in the fruit fly *Drosophila melanogaster* (Sorensen et al., 2005), the yeast *Saccharomyces cerevisiae* (Causton et al., 2001), the nematode *Caenorhabditis elegans* (Hahn et al., 2004) and a number of microbes (Gao et al., 2004; Koide et al., 2006) amongst others. Such comprehensive genomic resources, however, are limiting for non model species; hence there are few studies that have used microarrays to evaluate their response to heat stress. These are particularly scarce in marine species and to date include analysis of gene expression following heat stress in the killifish *Austrofundulus limnaeus* (Podrabsky and Somero, 2004) goby *Gillichthys mirabilis* (Buckley et al., 2006), the porcelain crab *Petrolisthes cinctipes* (Teranishi and Stillman, 2007) and the coral *Montastraea faveolata* (Desalvo et al., 2008) amongst others. Direct comparisons of changes in expression profiles are difficult between studies, as these differ in the length and intensity of the treatment, and responses may differ accordingly. Nonetheless, some of the changes in gene expression observed are common across these studies. These relate to maintenance of protein homeostasis through protein synthesis, degradation and chaperoning; lipid and/or carbohydrate metabolism; cytoskeletal reorganization; and signal transduction. As expected when working with non model species, these studies share limitations in the identification of a large number of transcripts (20-80% differentially expressed genes were of unknown function).

Another major limitation to the use of microarray technology in non model organisms is the elevated cost associated with the *de novo* development of resources for each organism of interest. Commercial platforms are only available for widely studied species and custom made microarrays can be labour intensive and costly. An increasingly popular alternative is to use an already developed microarray generated for one species to examine gene expression in a closely related species. Considering species A, for which the microarray platform has been developed, and species B and C, both closely related to A; heterologous microarray experiments can be used to compare relative expression levels between: species A and B/C, B and C, or two samples of B (Kassahn, 2008). The latter can be used to assess gene expression in response to a given treatment between samples of a novel species, for which microarray resources are not available. Examples of the use of heterologous microarray experiments to study stress responses in non model marine species are shown in table 3.1. Although likely to be restricted to the most highly conserved genes, the success of heterologous microarrays in identifying genes induced by different stressors is comparable to that using species specific microarrays.

Table 3.1 Examples of heterologous microarray experiments used in studies of the stress response to several stressors in non model marine species.

Species represented	Species hybridized	Type of stress	Platform	Number of features[1]	Differentially expressed clones	Reference
D. rerio (Zebrafish)	*P. moluccensis* (coral fish)	Temperature	Oligo	16399	111	(Kassahn et al., 2007)
F. rubripes (Fugu)	*D. mykiss* (Rainbow trout)	Zinc exposure	cDNA	18432	150	(Hogstrand et al., 2002)
G. mirabilis (Goby fish)	*T. bernacchii* (Antarctic fish)	Temperature	cDNA	9205	35-113	(Buckley and Somero, 2009)
G. mirabilis (Goby fish)	*T .orientalis* (Blue fin tuna)	Temperature	cDNA	9205	36-110	(Castilho et al., 2009)
P. flesus (European flounder)	*S. senegalensis* (Sole)	Aquaculture stressors	cDNA	3336	271-664	(Osuna-Jimenez et al., 2009)

[1] A feature represents a single spot of homogeneous cDNA amplified by PCR from a single cDNA clone that represents a single gene derived from the source tissue

3.1.4 Aims

In the present study, a custom made microarray was constructed to monitor the expression of several thousand genes in response to thermal stress in *L. elliptica*. This was used with the aims of i) identifying potentially novel mechanisms associated with the heat stress response in this species, which can be proposed as candidate stress biomarkers ii) determining whether the physiological responses to elevated temperatures observed at the organism level were apparent at the level of gene expression iii) test the potential of the microarray to be used with closely related species. This study provides the first gene expression profiling analysis in an Antarctic marine invertebrate using microarray technology and significantly expands the gene sequence information on this species.

3.2 MATERIALS AND METHODS

3.2.1 Experimental design

L. elliptica individuals were collected, acclimated to lab conditions and heat shocked as described in chapter two. The original design included the use of the same individuals for transcriptome and proteome analysis. However, despite flash freezing and storage at -80°C, RNA was not preserved intact over the length of the study (the protein and gene expression analyses were carried out in consecutive years) and the heat shocks had to be repeated. As a result of limitations in the number of individuals available, the experimental component previously used to assess loss of burying capacity was eliminated. No substrate was provided for individuals to bury. The experiment therefore consisted of two temperature treatments of 3°C and 9°C compared to a control (0°C). A pool of total RNA from eight individuals from the control group was used as a common reference to allow comparisons between treatments. Biological replicates, different individuals within the treatments, were individually compared with the pooled control group.

3.2.2 RNA preparation

RNA extractions from siphon, mantle and gill were carried out from individuals in all treatments using a modified tri reagent method (Tri Reagent, Sigma). Optimisation of the method was required as a high degree of degradation was observed using standard protocols. Trials included a final precipitation of the RNA pellet with 7.5 M LiCl, a final precipitation with sodium acetate and 100% ethanol, a modification of the tri reagent method using a saline solution. An RNA extraction directly from tissue using RNeasy mini kit (Qiagen) was also tested. Details for the results from these preliminary tests are provided in section 3.3.1. The optimized protocol involved a combination of the different methods, and is explained in detail below.

Total RNA was extracted on ice by homogenising ~50 mg of tissue samples in 600 µl Tri Reagent (Sigma) using glass homogenisers. Immediately after, 120 µl

chloroform was added to the homogenate, which was mixed vigorously and centrifuged at 13000*g* for 15 min at 4°C. After centrifugation, the upper aqueous phase was transferred into a fresh 1.5 ml tube to which 150 µl isopropanol and 150 µl saline solution (1:1 0.8 M sodium citrate and 1.2 M NaCl) was added. After a second 10 min centrifugation at the same speed and temperature, the supernatant was discarded and the RNA pellet washed with 100 µl 75% ethanol. Following a 1 min centrifugation, the ethanol was removed and, after drying, the pellet was resuspended in 250 µl DEPC treated water. 250 µl 7.5 M LiCl was added to the sample, which was incubated at -80°C for 30 min and centrifuged at 13000*g* for 20 min at 4°C. The supernatant was removed and the pellet was washed with ethanol, air dried and dissolved in 50 µl DEPC treated water. RNA concentrations were estimated using a Nanodrop spectrophotometer (ND-1000, Thermo Scientific). Samples for RNA amplification and labelling were cleaned using RNeasy mini kit columns (Qiagen) following manufacturer instructions in order to eliminate LiCl and salt residues. Samples were stored at -80°C until further use.

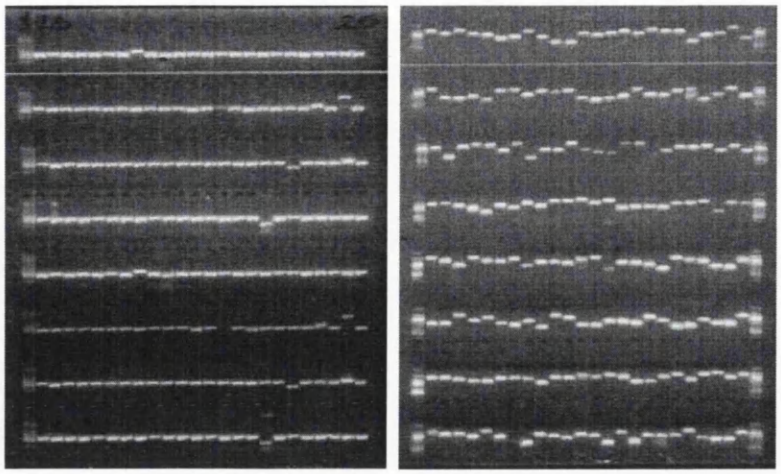
3.2.3 cDNA library construction

Three directional cDNA libraries were produced from mRNAs isolated from gill (Le_A01), mantle (Le_A02) and siphon (Le_A03). Each library was constructed from a pool of 14 individuals, in which control and temperature treatments were approximately equally represented. As part of the optimization process, two different protocols for the construction of a cDNA library were attempted with the aim of identifying the most appropriate technique. These are explained in detail below:

3.2.3.1 cDNA synthesis by primer extension

PolyA+ mRNA was prepared from total RNA using the MicroPoly(A) purist kit (Ambion) according to manufacturer's instructions. The Creator SMART cDNA library construction kit (Clontech) was then used for cDNA synthesis using the primer extension protocol as described by the manufacturer, with 2 µg of mRNA reverse transcribed using 200 U of PrimeScript reverse transcriptase (Takara). An additional denaturation step of 15 sec at 95°C was added to the recommended PCR

cycle. Once the reaction was complete, the size distribution of the ds cDNA was analysed on a 1.1% agarose gel alongside a DNA ladder (Hyperladder I, Bioline). Product sizes were smaller than expected at 400-800 bp, which may have been a result of low quality RNA. The recommended protocol was followed through cDNA size fractionation and ligation of cDNA to the pDNR-LIB vector provided with the kit. Size fractionation allows the selection of the longest cDNA sequences present. Transformations of recombinant plasmids into bacteria were performed using XL1-Blue Component Cells (Stratagene) according to manufacturer's instructions with the following modifications: the volume of cells, β-mercaptoethanol and SOC were 50% of those indicated in the protocol and 1 µl of ligation reaction was used. Chloramphenicol was used throughout the procedure (i.e. pDNR-LIB vector contains chloroamphenicol resistance genes, thus surviving cells will have the vector incorporated). After overnight growth, colonies were picked into 100 µl LB media and grown overnight at 37°C, followed by addition of 40 µl 50% glycerol. Inserts from each isolated clone were amplified via PCR in a 20 µl reaction mixture containing 0.084 µM dNTP, 1xNH_4 reaction buffer, 1.5 µM $MgCl_2$ (Bioline), 0.125 µM of each amine terminated M13 primer (forward: GTAAAACGACGGCCAGT; reverse: AACAGCTATGACCATG. (Invitrogen)) and 0.5 U Taq polymerase (Bioline). An initial denaturation step of 2 min at 96°C was followed by 30 cycles of denaturation at 96°C for 20 sec, primer annealing at 49°C for 20 sec and extension at 72°C for 3 min. In the final cycle, 5 min were added to the extension step. After PCR, products were diluted with 30 µl double distilled water after which 5 µl was analyzed on a 1.1% gel (figure 3.1a). Virtually all fragments were the same size, suggesting that the inserts were not incorporated into the plasmids.



a) b)

Figure 3.1 PCR amplification of cDNA inserts from isolated clones from *L. elliptica*. Sampled were electrophoresed on 1.1% agarose gels and correspond with two different amplification protocols a) cDNA by primer extension and b) cDNA synthesis by LD PCR. Hyperladder I (Bioline) is shown on the first and last lanes.

3.2.3.2 cDNA synthesis by Long-Distance (LD) PCR

cDNA was produced from total RNA using the LD PCR amplification protocol from the Creator SMART cDNA library construction kit according to manufacturer's instructions. 1µg total RNA and the reverse transcriptase PrimeScript RT (Takara) was used for first strand synthesis and Advantage 2 PCR kit (Clontech) was used for second strand synthesis according to the Creator SMART cDNA kit protocol. The final PCR consisted of 19 cycles.

Proteinase K digestion, size fractionation, ligation to pDNR-LIB vector, transformations of plasmids into bacteria and insert amplification were carried out as explained above. Successful PCR amplification of cDNA inserts is shown in figure 3.1b. Fragments with inserts of variable length indicate successful cloning (>1kb).

3.2.4 Microarray construction and design

A 16896 feature cDNA microarray was constructed consisting of 8448 PCR amplified cDNA clones printed in duplicate. 2784, 3456 and 2208 cDNA clones were printed from the Le_A01, Le_A02 and Le_A03 libraries respectively. The

Stratagene SpotReport Alien cDNA microarray validation system was included on the microarray resulting in a total of 17664 features. Prior to printing, 15µl of each amplified cDNA clone was transferred from 96 to 384 well plates and 5µl 4× spotting buffer (1 M sodium phosphate pH 8.5 and 0.001% sarkosyl) added to each well. A total of 23 384-well plates (22 plates of clones and one plate of spikes) were used for printing. Quality control was performed after PCR amplification, in which every individual amplified cDNA clone was visualised on an agarose gel. Those lanes with either multiple or no product were flagged as failed. There were 916 failed clones in total, 107, 450 and 359 from the Le_A01, Le_A02 and Le_A03 libraries respectively. A total of 7532 (89.16%) clones were flagged as pass and used in the final analysis. Failed clones were left on the plates and printed onto the microarray to minimise errors during the transfer from 96 to 384 plates but were excluded from the final analysis. Probes were prepared and printed onto amine binding Codelink Activated Slides (Amersham) according to manufacturer instructions with the substitution of SDS by sarkosyl, at a final concentration of 0.00025%. Printing was carried out in a Genetix QArray2 robot at 50% humidity. Each microarray consisted of 48 blocks arranged in 12 rows and four columns. Each block contained 184 clones in duplicate arranged in 16 rows and 23 columns.

3.2.5 RNA amplification, labelling and hybridization

PCR amplified ds cDNA targets were prepared from 1 µg total RNA using the protocol described in Petalidis et al (2003) with the addition of 1 µl Stratagene Alien mRNA spikes and the substitution of PowerScript with PrimeScript reverse transcriptase. Briefly, 500 ng of total RNA was used as an input into the first strand reaction along with 1µl Stratagene Alien Spike, 1 µl of 10 µM SMART IIA oligonucleotide (Metabion), and 1µl of 10 µM SMART CDS IIA (Sigma). The reaction volume was made up to 6µl with DEPC treated water which was then heated to 72°C for 2 min and then cooled on ice to allow the primers to anneal. To this, 2 µl of first strand buffer, 1 µl 10 mM dNTP and 1µl Primescript reverse transcriptase (Takara) were added. The samples were then incubated at 42°C for 1 h. A 1 µl aliquot of first strand reaction was used as a template in a second strand amplification reaction consisting of 40 µl ddH_2O, 5 µl 10x PCR Advantage buffer II (Takara), 1 µl 10 mM dNTP, 2 µl 10 µM 5' PCR Primer IIA (Clontech) and 1 µl

PCR Advantage 2 polymerase (Takara). The following cycle conditions were used for PCR amplification: 95°C for 1 min and then 19 cycles of 95°C for 5 sec, 65°C for 5 sec and 68°C for 6 min. Amplification efficiency at different numbers of cycles was tested and is shown in figure 3.2 from which 19 cycles was chosen as the most efficient. Amplified ds cDNA was purified using Illustra GFX PCR purification columns (GE Healthcare). After purification, sample concentration was measured, and 1 μg cDNA was labelled and purified according to the protocol described in Petalidis et al. (2003). Briefly, 1 μg cDNA was diluted in DEPC treated water to a volume of 22 μl, to which 20 μl of 2.5x random primer reaction buffer (Invitrogen) was added. The mixture was incubated at 95°C for 5 min and placed on ice, prior to the addition of 5 μl of Low-C dNTP mix (5 mM dATP, 5 mM dGTP, 5 mM dTTP, 2 mM dCTP), 3 μl Cy3 or Cy5-dCTP (1mM stock, Amersham) and 40 U Klenow polymerase (Invitrogen). The reaction was incubated at 37°C for 2 h in the dark, after which time 5 μl stop buffer (Invitrogen) was added to terminate the reaction. Labelled products were purified using G50 columns (Amersham) according to manufacturer's instructions. After purification, Cy3 (control) and Cy5 (treatment) labelled samples were pooled together and 4 μg poly dA and 4 μg yeast tRNA (Sigma) were added as blocking agents. Targets were ethanol precipitated, dissolved in 40 μl hybridization buffer (40% deionised formamide, 5x Denhart's solution, 1 mM sodium pyrophosphate, 50 mM Tris pH 7.4 and 0.1% SDS), denatured at 95°C for 5 min and allowed to cool to room temperature. Labelled targets for each treatment group were hybridized to the microarray with labelled control pool target as a reference. Hybridizations were performed in a humidified incubator at 49°C for approximately 16 h. After hybridization, the microarrays were washed at room temperature in 2xSSC 0.1xSDS for 15 min, followed by 2xSSC for 5 min and 0.1xSSC for 5 min. After washing, the microarrays were dried by centrifugation, scanned in a microarray scanner (Genepix 4100A, Molecular Devices) at two wavelengths (532 nm green and 635 nm red) and saved as TIFF images. Detected fluorescence was adjusted by adjusting the voltage to the photomultiplier tubes (PMT's) to ensure that the ratio of Cy5 to Cy3 signals was close to one and no features were saturated.

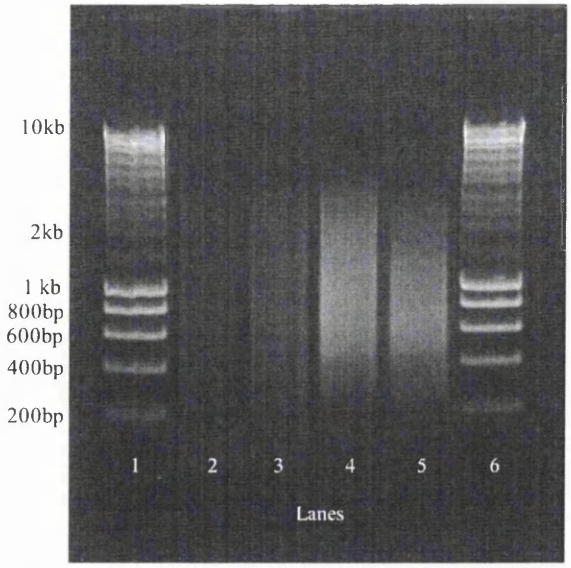


Figure 3.2 PCR amplified ds cDNA from *L. elliptica* mantle after 15, 17, 19 and 21 amplification cycles in lanes 2-5 respectively. The highest amplification efficiency was achieved after 19 cycles. Hyperladder I (Bioline) markers are shown in lanes 1 and 6.

3.2.6 Microarray data analysis

Images were analysed using the Genepix Pro software v 6.0.1 (Molecular Devices). A Genepix array list (GAL) file was generated, containing information on the size and position of the blocks, the layout of features and individual feature identifiers. The GAL file was used to grid the microarray and failed features were flagged and their position indicated. Failed features included those not found, those in areas of abnormally intense background, and those abnormally shaped. This was followed by visual inspection of the features including flagging of any additional abnormal ones and data extraction from the scanned images to GPR (Genepix results) files.

Data were analysed using the limma package (Smyth, 2005) in R (R Development Core Team, 2008). Median foreground and background intensities were extracted from the GPR files. Features flagged as abnormal in Genepix and those flagged as failed were given a zero weight. Unflagged features were given weight of one if the following criteria were met:

- Median foreground intensity for each individual feature was three times the median background intensity in both channels

- At least one duplicate of each probe was present in more than 50% of the microarrays within a treatment.

Background subtraction was applied using the Half function (Ritchie et al., 2007), which subtracts median local background from median feature intensities and gives a value of 0.5 to all values <0.5. This prevents zero and negative feature intensities, a convenience for subsequent statistical analysis. Within microarray normalisation was performed applying the printiploess function (Smyth & Speed, 2003), a robust method that normalises each print tip independently. As the design consisted of a single sample experiment with a common reference in the 532 nm wavelength, between microarray normalisation was performed using the Gquantile function. Differentially expressed genes were identified by fitting a linear model to every gene, where empirical Bayes methods were employed to use information from all genes in the microarray to moderate the standard errors of expression changes in each gene and improve power (Smyth, 2004). Intra duplicate correlation of features was integrated within the linear model (Smyth, Michaud & Scott, 2005). Differentially expressed clones were selected at a Benjamini and Hochberg (BH) (Benjamini and Hochberg, 1995) adjusted p-value $P<0.05$ and a minimum two fold change. The array design and experiment have been submitted to Array Express (accession number A-MEXP-1676, E-MEXP-2336)

3.2.7 EST sequencing

Inserts from each differentially expressed isolated clone from the cDNA library were amplified via PCR using the same reaction mix and amplification cycle stated in a previous section (cDNA synthesis by primer extension). Enzymatic clean up of the amplified cDNA was carried out in a 60 µl reaction containing 1 µl PCR product, 0.4U Shrimp Alkaline Phosphatase (SAP), 0.6 U Exonuclease I and 0.567 µl SAP dilution buffer (USB, Affymetrix). The reactions were incubated for 30 min at 37°C, followed by 10 min at 80°C to inactivate the enzymes. After incubation, 4 µl Dye terminator mix (GE Healthcare) and 0.5 µl 10 µM M13R primer was added to each

PCR template, and subjected to 26 thermal cycles of 20 sec at 95°C, and 2 min 20 sec at 60°C. Products were ethanol precipitated, rehydrated in 10 µl MegaBACE loading solution and run on a MegaBACE 1000 capillary sequencer (Amersham Biosciences, GE Healthcare).

3.2.8 Sequence data analysis

Chromatograms were imported into Geneious Pro 4.5 (Drummond et al., 2007) where sequences were individually edited to ensure high quality. The pDNR-LIB vector and 3′-polyadenylate tail were identified and removed. Low quality 3′end bases were also excluded. Edited sequences under 100bp were excluded from further analysis. The average read length of the remaining sequences was 323.5±9 (mean±SE). Contig assembly was performed using the TGI clustering tool (Pertea et al., 2003) with a minimum overlap of 30 and 94% overlap identity. Sequences were searched using blastx against the Uniprot/Swissprot database (Bairoch et al., 2007). Sequence matches were considered significant with an e value of less than $1e^{-7}$. Gene Ontology (GO) identifiers were obtained using the GO Consortium at EBI (Ashburner et al., 2000). Genes were allocated to functional categories based on GO annotations and literature searches. All sequences will be submitted to GenBank upon acceptance for publication.

3.2.9 Real-time PCR

For confirmation of microarray results, the expression levels of five candidate genes were analysed via quantitative PCR (qPCR) in a MX3000P qPCR thermocycler (Stratagene). Total RNA for real-time PCR was isolated as previously described from the same individuals used in the microarrays. cDNA was produced from 1 µg total RNA reversed transcribed using a Quantitect Reverse Transcription kit (Qiagen) according to manufacturer instructions. cDNA was diluted 2:1 in ddH_2O prior to use. A reference gene (Le_A01_P0_24D02), which showed no significant variation between individuals as indicated by the microarray results, was used for normalization of expression.

Primers were designed using Primer3 (Rozen and Skaletsky, 2000) and manufactured by Invitrogen. Aliquots of 1 µl cDNA were amplified in 20 µl reaction mixtures in the presence of Brilliant SYBR Green qPCR Master Mix (Stratagene). PCR conditions for all genes were 10 min at 95°C, followed by 40 cycles of 30 sec at 95°C, 60 sec at 58°C, and 30 sec at 72°C. Amplifications were analysed using the MxPro™ software (Agilent Technologies) and Ct values exported into Excel. Primer specificity was initially checked via agarose gel electrophoresis. Amplifications using primer pairs resulting in either several or no product on the gel were considered failures and primers were redesigned until a single product was observed. Primer sequences are shown in table 3.2. Efficiency of amplification was determined for each primer pair using serial 2-fold dilutions and was in all cases >95%. Dissociation curves were visually inspected to ensure that no primer dimers had been produced and a single product had been amplified during the PCR reaction.

QPCR reactions for each primer pair were carried out in a 96 well plate, containing three individuals per treatment, each amplified in triplicate on separate wells. The same set up was applied to both the candidate gene and the control. Negative controls for each primer pair were performed without primers (no primer control) and without cDNA template (no template control).

Table 3.2 PCR primers used for qPCR analysis of five genes. All genes were upregulated with heat stress. Primer sequence, direction, optimal annealing temperature, product size, correlation coefficient of the standard curve (Rsq) and PCR efficiency are shown.

Gene ID	Direction	Primer sequence	Annealing T (°C)	Product size (bp)	RSq	PCR Efficiency (%)
Le_A01_P0_02A04	F	GAGGGTGGTAACCAGACCAA	59.82	133	0.995	106.3
	R	CAATGCAGGAAAACCCAACT	59.97			
Le_A01_P0_03H10	F	CAATCCTGAGGCTGTCATCA	59.79	131	0.995	102.0
	R	ATGCCCTTTCATCCAGTTCA	60.46			
Le_A01_P0_09E10	F	AAAGACCCAACAAAGCCAAA	59.59	134	0.999	116.8
	R	AAGTTTCTCCACGCTTCACC	59.33			
Le_A01_P0_17H11	F	CAAGGAAAACGCCAAAGAAG	59.85	137	0.988	96.8
	R	AGGATTTCGGAAGCACAATG	60.07			
Le_A03_P0_27A01	F	ACCCAGGAAGAGCGTAACCT	60.13	102	0.999	128.7
	R	TGCAGATTTTGTTTCGACCA	60.23			

Relative expression ratios of each gene (compared to the reference gene) between treated and control samples were calculated using the relative expression software tool (REST) (Pfaffl et al., 2002) based on the Pfaffl method, which uses efficiency corrected mathematical model (Pfaffl, 2001). REST is an excel macro that incorporates a mathematical model to calculate relative expression ratios on the basis of the PCR efficiency and crossing point derivation of the unknown sample versus the control. The crossing point is defined as a threshold cycle at which a statistically significant increase in fluorescence is detected. The threshold cycle is inversely proportional to the log of the initial copy number. Expression ratios are tested for significance by a randomisation test. This test makes no distributional assumptions about the data but assumes that treatments were randomly allocated. Ten thousand randomisations were used.

3.3 RESULTS

3.3.1 RNA quality

As stated in the methodology section, problems with extracting RNA of sufficient integrity in *L.elliptica* resulted in the optimization of the RNA extraction method for this species. The RNA precipitation methods alone did not improve RNA quality substantially, whereas extractions using the RNeasy mini kit resulted in RNA that was less degraded. However, the kit produced RNA concentrations that were lower than required. An example of the products visualised in 1% agarose gels of the different extraction methods is shown in figure 3.3. Samples show the classical molluscan appearance of a single band, possibly due to a break of the 28S band which converts the 28S into components of the size of 18S during the extraction procedure (Barcia et al., 1997). A modified tri reagent method with an added saline solution was most successful and used throughout the subsequent experiments.

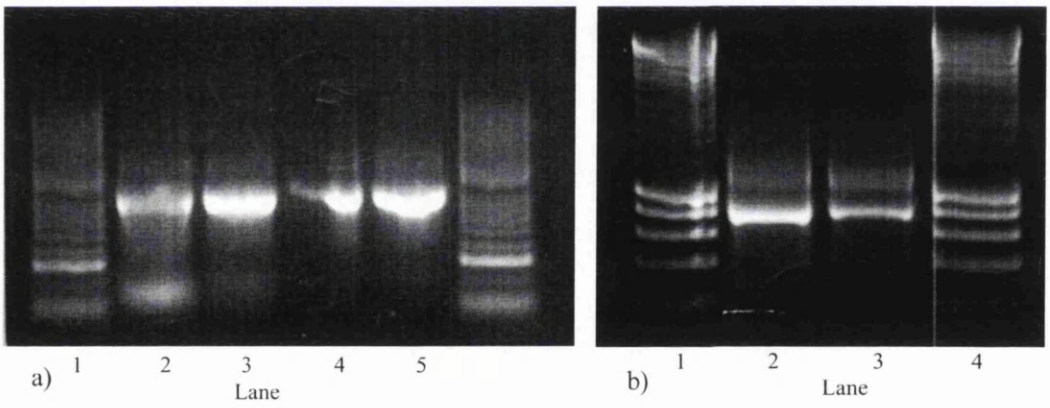


Figure 3.3 Representative samples of total RNA isolated from *L. elliptica* mantle tissue using a) tri reagent standard protocol, tri reagent method with final LiCl precipitation, modified tri reagent method using a saline solution and modified tri reagent method followed by LiCl precipitation in lanes 2-5 respectively. b) Qiagen RNeasy Mini Kit. Approximately 2 µg of each was electrophoresed on a 1% agarose gel. In both images, Hyperladder I (Bioline) is shown in the first and last lanes.

The quality of RNA extracted from individuals from all groups was assessed by visual inspection of RNA bands in 1.5% agarose gels and the ratio of absorbances at 260 and 280 nm (A_{260}:A_{280}). RNA was considered of acceptable integrity and purity when samples consisted of a high intensity band around 1kb and A_{260}:A_{280} ratio>1.8. Based on these parameters there was no indication of different RNA quality from individuals in the different treatment groups. However, following problems with labelling and microarray hybridisation of 9°C treated samples, the RNA quality of six samples extracted from control and 9°C treated animals was further assessed using the Lab901 Screen Tape R6K (Lab901 Ltd.), a newly developed tool for RNA quality control. Samples were prepared according to manufacturer instructions and 200 ng total RNA were analysed per sample. The instrument combines RNA separation via polyacrylamide electrophoresis, profile imaging and sample analysis. RNA integrity is assessed qualitatively using the gel image and quantitatively via the RNA profile where 28S and 18S ribosomal peaks are located and compared with known standards, currently limited to mammalian and model species.

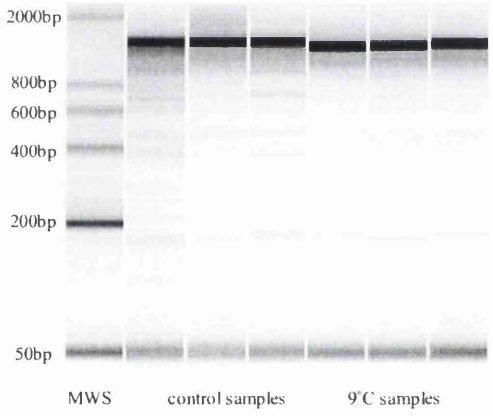


Figure 3.4 Banding patterns of RNA samples from *L. elliptica* separated on an acrylamide matrix using a Lab901 Screen Tape R6K analyser. Lane 1 shows an RNA ladder, followed by three samples from individuals from the control group (0°C), and three samples from individuals in the 9°C treated group. RNA extracted from all samples was treated in the same way.

The lack of high quality standards for *L. elliptica* and appearance of a single high intensity ribosomal band did not allow a quantitative assessment of RNA quality or

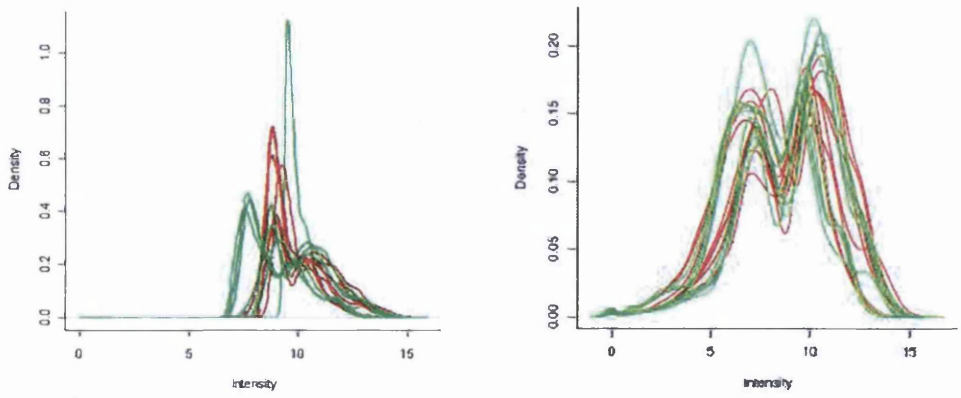
degradation. Nonetheless, visual inspection of the high resolution RNA banding patterns revealed some differences between the control (0°C) and 9°C treated samples (figure 3.4) which may relate to different quality of RNA extracted from individuals in the two groups. This could explain the limited success in RNA labelling of the 9°C treated samples resulting in poor hybridizations and extremely high intensity background in the microarrays. It was therefore not possible to conduct a sufficiently extensive microarray analysis with the 9°C samples and microarray work with these was discontinued. The experiment was conducted comparing 3°C treated animals with the control (0°C).

3.3.2 Microarray quality control

Preliminary quality screening resulted in the introduction of a series of quality control measures. Initial screening included visual inspection of the distribution of intensities across features, background intensities and MA plots before and after normalization and print tip order plots (appendix I). Uneven background intensities within arrays indicated the need for local normalization methods. Print tip order analysis revealed that no bias in the overall intensity of features had been introduced by the printing procedure. MA plots show the relationship of signal ratios to intensities (M-values are calculated $\log_2$ expression ratios, A-values are average $\log_2$ expression values). For a large chip where the expression of most genes does not change between treatments, the bulk of the data should fall close to the zero line. Normalization should thus bring outliers closer to this line. Analysis of boxplots (shown below) was carried out in raw, background corrected and normalised data.

Figure 3.5 shows density plots for signal distributions for both channels in all microarrays prior to and post filtering. Intensities are presented $\log_2$ transformed. Prior to filtering, a large proportion of features presented intensities <1000 with distributions in the red channel slightly different from those in the green with at least one chip with a very different distribution in the green channel from all of the others. Filters applied prior to data analysis resulted in a relatively high proportion of missing values (mean±SE= 63.03±2.79%) so that a large proportion of features was not used for further analysis (table 3.3). However, the use of stringent filters ensured the quality and reliability of the dataset maintained. After data filtering, signal

distributions showed a more homogeneous profile between all microarrays and channels.



a) b)

Figure 3.5 Density plots for $\log_2$ signal distributions for a) raw intensities prior to applying filtering criteria and b) raw intensities after filtering. Each line represents one microarray and channel. Green and red lines correspond to the green (control) and red (3˚C treated) channels respectively.

Table 3.3 Proportion of missing values after applying intensity based quality control filters. These include features flagged fail by Genepix, features not present in at least 50% of microarrays and those with median intensity values less than three times the median background.

Array	Fail	Pass	% Fail	% Pass
Le_01	9939	7725	56.27	43.73
Le_02	9765	7899	55.28	44.72
Le_03	9875	7789	55.90	44.10
Le_04	13262	4402	75.08	24.92
Le_05	10006	7658	56.65	43.35
Le_06	11704	5960	66.26	33.74
Le_07	12067	5597	68.31	31.69
Le_08	12447	5217	70.47	29.53

Boxplots of intensity data by channel are shown in figure 3.6. Foreground and background signal intensities are shown for both channels. There were differences in median values across the microarrays and their distributions were not similar, suggesting that data normalization was required.

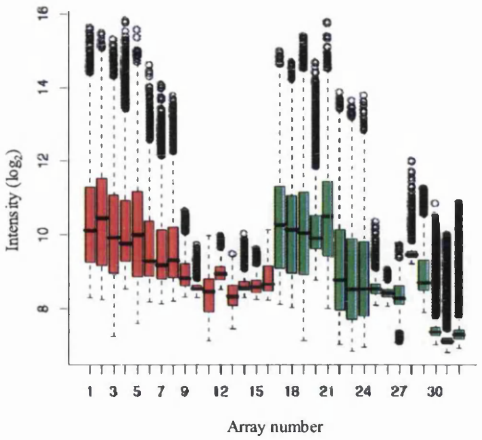


Figure 3.6 Boxplots of log_2 raw intensity by channel and microarray. Foreground red (3°C treated) intensities for microarrays 1 to 8 are labelled 1-8 in the x axis, followed by background red intensities in the same order labelled 9-16, green (control group) foreground 16-23 and green background 24-32.

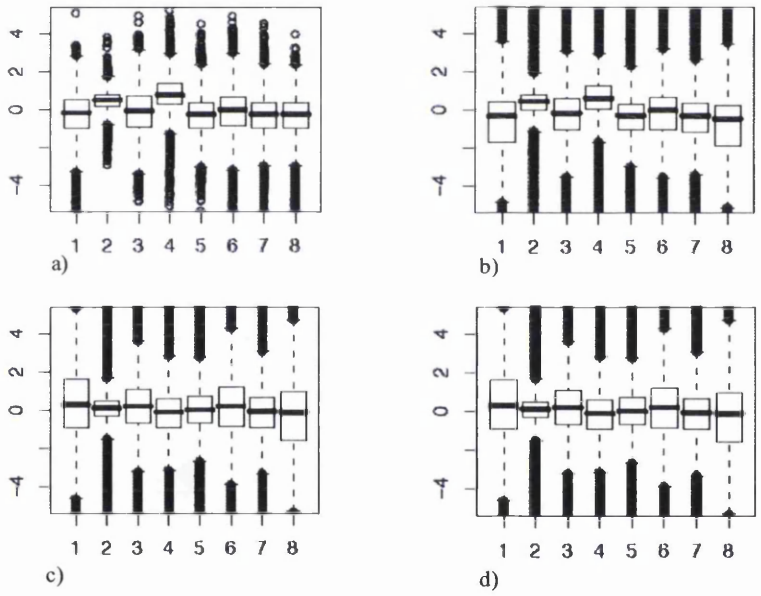


Figure 3.7 Boxplots of log_2-ratios by microarray for a) ratios of raw data, b) background subtracted data using the half method, c) data after background subtraction and within microarray normalization using printiploess and d) data after background subtraction, within and between microarray normalization using printiploess and Gquantile. Microarrays 1-8 and corresponding M-values are shown in the x and y axes respectively. Boxplots indicate median M-values, upper and lower quartiles, variability range, and individual extreme values. Images are zoomed in to show distribution of median values.

The result of background subtraction and normalization is shown in figure 3.7. Boxplots comparing the distribution of log-ratios between microarrays, before and after background subtraction and normalization indicate that the selected methods allowed further comparisons between microarrays. After normalization the median M-values are centralized around zero, making between microarray comparisons possible. The overall variability within and between arrays was reduced by normalization. The effect of between microarray normalization cannot be appreciated from the plots, as the Gquantile method ensures the same empirical distribution of the common reference across microarrays without altering the M-values. The effect of this normalization is apparent on the MA plots (appendix I). After between microarray normalization, some outliers were eliminated bringing the data closer to the straight horizontal line along zero. It must be noted that low intensity features have been excluded by filtering.

3.3.3 Microarray analysis

Eight microarrays from individual samples of the 3°C treated group were compared to a pooled control group (0°C). 4015 features passed quality screening (47.5% of all features without taking into account duplicates). Where both duplicates passed the filtering process, these were incorporated in the analysis but are referred to as a single feature for convenience. Of these, 294 features presented significantly different intensities in the treated samples when compared to the control at a BH adjusted p-value $P<0.05$, with the mean level of expression altered by at least two fold. This represented 3.5% of the total number of features printed on the microarray and 7.3% of the total number of features maintained for analysis. 173 of these were upregulated and 121 downregulated in the 3°C treated animals in relation to the control. Up- and down- regulated genes varied in expression by up to 5.89 and -4.86 fold respectively. Following sequencing and editing, 206 clones remained, of which 64 clustered into 18 contigs, and 142 were singletons, thus the 294 features represent 160 unique transcripts. The rest of the sequences were of low quality, below the minimum length of 100bp after trimming and were thus excluded from further analysis. Only 42 clones, representing 33 transcripts showed significant sequence similarity to a known protein. Results are shown in table 3.4.

Table 3.4 Genes identified as differentially expressed, with a minimum two fold change on *L. elliptica* individuals treated at 3°C in relation to controls maintained at 0°C, as determined by microarray data. Feature names, blast IDs for the most significant hit in Uniprot Swissprot, Uniprot recommended gene names, blast scores and e-values and allocated functional categories are indicated. (*) Indicates a representative clone within a contig. The library each sequence derives from is indicated in the feature name (A01, A02 and A03 for gill, mantle and siphon respectively)

cDNA feature	Description	Organism	Uniprot ID	Blast score	e-value	Functional category	Mean FC	P
Le_A02_P0_13E05*	Actin-103	*Nicotiana tabacum*	P93376	89.0	2.E-37	Cytoskeletal activity	2.69 ±0.54	0.014-0.023
Le_A02_P0_19A04*	60S acidic ribosomal protein P1	*Caenorhabditis elegans*	P91913	110.0	5.E-24	Protein turnover	-4.24 ±0.30	0.003-0.013
Le_A03_P0_01E04*	Tropomyosin	*Turbo cornutus*	Q7M3Y8	59.7	4.E-09	Cytoskeletal activity	2.39 ±0.14	0.004-0.031
Le_A01_P0_11A08*	Beta-actin	*Schizophyllum commune*	Q9Y702	83.2	4.E-16	Cytoskeletal activity	2.26 ±0.01	0.001-0.003
Le_A03_P0_01A04*	Myosin essential light chain	*Argopecten irradians*	P07291	168.0	7.E-43	Cytoskeletal activity	2.04 ±0.03	0.026-0.047
Le_A01_P0_17H11	Epithelial chloride channel protein	*Bos taurus*	P54281	56.2	5.E-08	Intracellular trafficking and signalling	3.56	<0.001
Lc_A01_P0_03A08	Bifunctional aminoacyl-tRNA synthetase	*Drosophila melanogaster*	P28668	64.3	2.E-15	Protein turnover	3.41	0.046
Le_A03_P0_06D08	40S ribosomal protein S23	*Sus scrofa*	Q6SA96	164	2.E-40	Protein turnover	3.32	0.021
Le_A01_P0_06H06	Peroxiredoxin-1	*Gekko japonicus*	Q6DV14	122	5.E-28	Antioxidant activity	3.28	0.002
Le_A02_P0_16F09	Inter-alpha-trypsin inhibitor heavy chain H4	*Bos taurus*	Q3T052	47	1.E-08	Acute phase response	3.17	<0.001
Le_A01_P0_03H10	Serine/threonine-protein kinase PAK 3	*Rattus norvegicus*	Q62829	60.08	2.E-09	Intracellular trafficking and signalling	3.12	0.000
Lc_A03_P0_18E11	Probable glutathione S-transferase 9	*Caenorhabditis elegans*	Q21743	68.6	9.E-12	Antioxidant activity	3.01	0.000
Le_A03_P0_21G06	Sarcoplasmic calcium-binding protein	*Mizuhopecten yessoensis*	P02637	68.9	7.E-12	Intracellular trafficking and signalling	2.94	0.011
Le_A01_P0_09E10	High mobility group protein B3	*Gallus gallus*	P40618	59.3	6.E-11	Regulation of global transcription	2.89	0.001
Lc_A03_P0_27A01	14-3-3 protein homolog	*Neospora caninum*	Q25538	93.2	3.E-19	Intracellular trafficking and signalling	2.85	<0.001

cDNA feature	Description	Organism	Uniprot ID	Blast score	e-value	Functional category	Mean FC	P
Le_A03_P0_19E07	BTB/POZ domain-containing protein KCTD7	*Bos Taurus*	A4IFB4	69.3	5.E-12	Protein turnover	2.79	0.024
Le_A01_P0_10F10	Centrin-1	*Bos Taurus*	Q32LE3	114	9.E-31	Intracellular trafficking and signalling	2.70	0.004
Le_A03_P0_04H03	Tropomyosin	*Haliotis diversicolor*	Q9GZ71	99.8	4.E-21	Cytoskeletal activity	2.69	0.010
Le_A03_P0_03B11	ATP synthase subunit alpha heart isoform, mitochondrial	*Bos Taurus*	P19483	84.7	1.E-16	Energy metabolism	2.63	0.010
Le_A03_P0_11F04	Betaine--homocysteine S-methyltransferase 1	*Xenopus (Silurana) tropicalis*	Q5M8Z0	54.3	7.E-14	Protein turnover	2.60	0.050
Le_A01_P0_02H01	Very low-density lipoprotein receptor	*Oryctolagus cuniculus*	P35953	52.8	7.E-07	Intracellular trafficking and signalling	2.59	0.005
Le_A02_P0_05F11	Calponin-2	*Homo sapiens*	Q99439	61.02	2.E-09	Intracellular trafficking and signalling	2.47	0.010
Le_A03_P0_03D06	Histone H2A.V	*Strongylocentrotus purpuratus*	P08991	193	2.E-49	DNA synthesis	2.38	0.021
Le_A02_P0_29A03	Probable ribosome biogenesis protein RLP24	*Drosophila melanogaster*	Q9VGN9	111	1.E-24	Protein turnover	2.30	0.016
Le_A01_P0_12A02	Calmodulin	*Ciona intestinalis*	O02367	271	2.E-72	Intracellular trafficking and signalling	2.27	0.015
Le_A01_P0_01E09	Proteasome subunit beta type-5	*Rattus norvegicus*	P28075	110	2.E-24	Protein turnover	2.12	0.006
Le_A02_P0_07A12	Prefoldin subunit 2	*Bos Taurus*	A1A4P5	86.7	3.E-17	Protein folding/chaperoning	2.09	0.007
Le_A02_P0_29G11	40S ribosomal protein S3-A	*Xenopus laevis*	P02350	270	2.E-72	Protein turnover	2.07	0.018
Le_A01_P0_09C10	Pre-mRNA-processing factor 19	*Bos Taurus*	Q08E38	76.3	9.E-14	Protein turnover	2.07	0.026
Le_A03_P0_01E09	40S ribosomal protein S17	*Bos Taurus*	A5PK63	187	1.E-47	Protein turnover	2.06	0.026
Le_A02_P0_21A08	Dolichyl-diphosphooligosaccharide--protein glycosyltransferase subunit 4	*Mus musculus*	Q99LX8	63.2	4.E-10	Protein turnover	2.04	0.016
Le_A01_P0_02B11	40S ribosomal protein S24	*Ictalurus punctatus*	Q90YQ0	135	6.E-32	Protein turnover	2.01	0.047
Le_A03_P0_24A09	60S ribosomal protein L22	*Ictalurus punctatus*	Q90YU6	125	9.E-29	Protein turnover	-2.10	0.014

3.3.4 Real time PCR

Five genes were arbitrarily selected for qPCR analysis for confirmation of the microarray results. All genes were upregulated in the 3°C treated group but not all could be identified by sequence similarity searching. The expression levels of all five genes after exposure to 0 and 3°C were estimated. Mean fold differences and associated standard errors between treated and control animals for five genes are plotted in figure 3.8. Fold changes are given in a logarithmic scale. These were calculated by dividing expression levels in the experimental group (3°C) by those in the control group (0°C). QPCR mean fold change estimates were calculated using the Pfaffl method (Pfaffl, 2001) and represent means for three biological replicates with associated standard errors. Microarray mean fold change estimates and associated standard errors were calculated using individual M-values from eight microarrays. Expression levels have been normalised using a control gene in the case of qPCR estimates and global normalization methods in the case of microarrays as previously discussed.

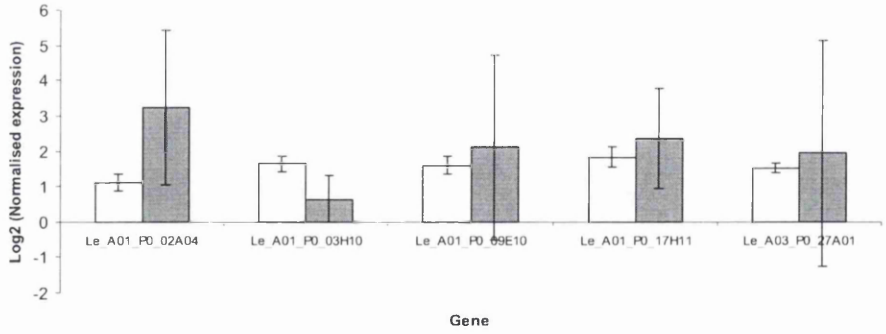


Figure 3.8 Mean fold changes and standard errors in gene expression for five genes each with three and eight biological replicates for qPCR and microarray estimates respectively. White and dark grey bars represent microarray and qPCR data respectively. All values are normalised and shown in a logarithmic scale.

QPCR results confirmed that all genes were apparently upregulated in the treated groups in relation to the control. With the exception of Le_A01 P0_03H10 at 3°C, mean fold differences between treated and control samples based on three biological replicates are estimated to be higher by qPCR than by microarrays. However, a

comparison between mean estimates using a t-test with unequal variances revealed that these differences are not significant (t = 2.53, 0.87, 2.50, 1.34, 1.21; P= 0.123, 0.471, 0.123, 0.312, 0.349 for Le_A01_P0_02A04, Le_A01_P0_03H10, Le_A01_P0_09E10, Le_A01_P0_17H11 and Le_A03_P0_27A01 respectively). There was a high level of individual variation in the qPCR estimates, as indicated by the large error bars.

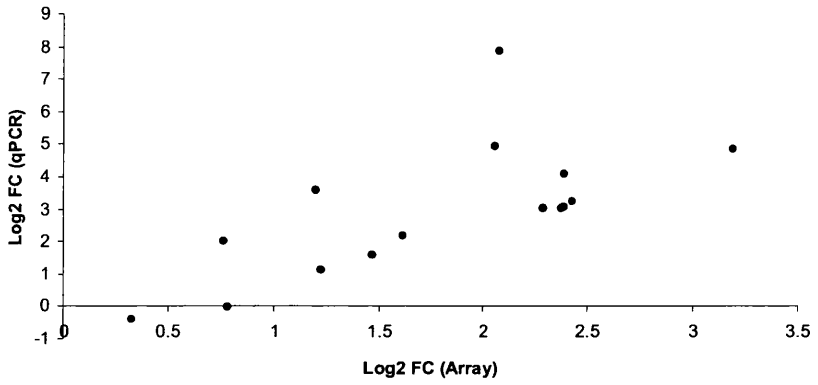


Figure 3.9 Correlation between $\log_2$ gene expression ratios measured by microarray and qPCR analysis for 3°C treated samples in relation to 0°C controls. Each point represents one of three biological replicates per gene. Pearson correlation r=0.69, P=0.005

The correlation between microarray and qPCR estimates of gene expression was also analysed. Figure 3.9 shows a significant positive correlation (Pearson's correlation coefficient r=0.69 and P<0.001) between normalised microarray and qPCR estimates of $\log_2$ gene expression ratios (treated/control) across fifteen pairs of values. Each point represents one of three biological replicates per gene, for all five genes analysed by qPCR.

3.3.5 Tissue specificity

Although the microarray was produced from libraries made from RNA extracted from different tissues, the only material hybridized was extracted from mantle. There was therefore the concern that i) only those probes prepared from mantle RNA

would cross-hybridize, resulting in a high proportion of missing values and ii) only features representing clones from mantle would show significant changes in expression revealing a mantle specific, rather than a generic organism response. An estimation of the level of tissue specificity was made by analysing the number of features flagged as passed in each library after applying quality control filters. Results are expressed as percentages of the total number of features representing each library. These were 45, 52 and 41% for Le_A01 (gill), Le_A02 (mantle) and Le_A03 (siphon) respectively. Given the proportion of passed features from each library, a high degree of tissue specificity was not apparent. Within the clones considered differentially expressed in response to temperature 28, 42 and 31% hybridized to targets printed from the gill, mantle and siphon libraries respectively. A higher proportion of features showing significant changes in expression between the treatments could suggest tissue specific responses for some transcripts.

3.3.6 Heterologous hybridisation

In order to determine the future potential use of the microarray with other bivalve species, two 500 ng aliquots from an RNA sample extracted from *M. edulis* mantle tissue were each labelled with either Cy3 or Cy5 dyes and hybridised onto the *L. elliptica* microarray. Samples were prepared as previously described with the stringency of the washes decreased by reducing the last wash to 2 min.

A comparable proportion of features was excluded during image analysis when hybridized with *M. edulis* (59.7%) or *L. elliptica* (50.2±3.72 %) RNA suggesting that a high proportion of probes hybridized with both target species. Density plots for signal distributions for both channels in all *L. elliptica* microarrays, and an average of both channels for the *M. edulis* microarray are shown in figure 3.10. The distribution of feature signal intensities in the *M. edulis* microarray approximates that of *L. elliptica* suggesting that hybridizing *M. edulis* RNA to the *L. elliptica* microarray does not result in a consistent decrease of signal intensities across spots.

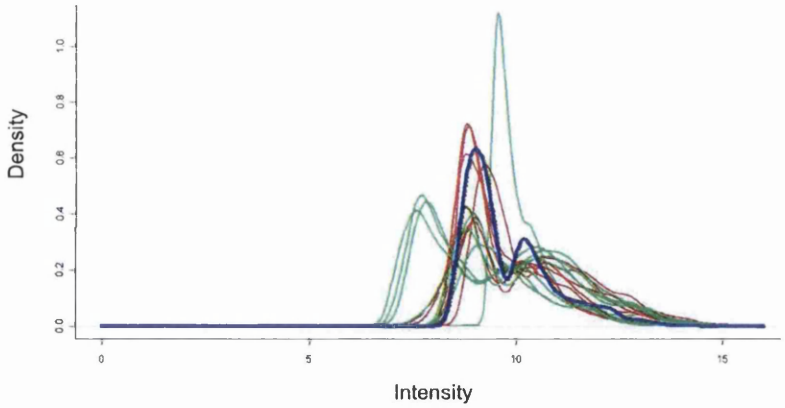


Figure 3.10 Density plots for $\log_2$ signal distributions for raw intensities. Each line represents one microarray and channel. Green and red lines correspond to the green (control) and red (3˚C treated) channels respectively in the *L. elliptica* microarrays. The blue line represents the intensity distribution of mean raw intensities for both channels of the *M. edulis* microarray.

The correlation between *L. elliptica* and *M. edulis* $\log_2$ raw signal intensities across features is shown in figure 3.11. Each point corresponds with one feature. *M. edulis* raw intensity represents signal intensities in the green channel from a single microarray. Signal intensities for *L. elliptica* are the average of raw intensities across all microarrays in the green channel. The comparison was made between untreated samples to avoid outliers due to highly differentially expressed genes in response to temperature. A strong positive correlation was observed between intensity signals across features (Pearson correlation r=0.86 P<0.001) suggesting that most gene transcripts have comparable levels of expression in both species.

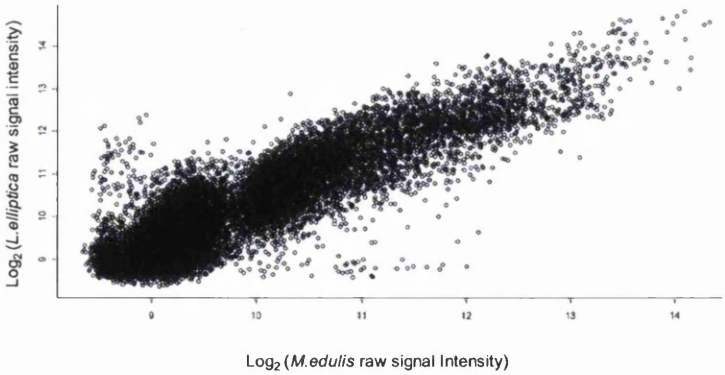


Figure 3.11 Correlation of $\log_2$ signal intensity across features between *M. edulis* and *L. elliptica* samples. Each point represents a single feature. *M. edulis* represents the

raw intensity of the green channel from a single microarray. *L. elliptica* represents the average raw signal across all microarrays in the green channel (control). Pearson correlation =0.86, $P<0.001$

3.4 DISCUSSION

The following sections include the analysis and discussion of a series of technical considerations related to RNA and microarray quality. This is followed by a discussion on the significance of the observed changes in gene expression following heat stress in *L. elliptica* and how these relate to current physiological theories.

3.4.1 Technical issues

3.4.1.1 Tissue specificity

The presence of three different tissues in the cDNA library was a concern in terms of the potential for a high level of tissue specificity. This could result in a large proportion of missing values, mainly from the limited binding of targets extracted from mantle to probes from the gill or siphon libraries. The proportion of features kept for analysis after quality control was similar for all three libraries, suggesting that tissue specificity was not of major concern. As expected, there was a slightly higher proportion of hybridised probes within the mantle library that in the other two (52% for mantle and 45% and 41% for gill and siphon respectively). The results are to be taken with care as a higher proportion of clones from the mantle library, successfully hybridised by targets prepared from mantle tissue could also reflect differences in the level of redundancy between the libraries. The lack of sequence information for most genes on the microarray does not allow one to draw further conclusions. The proportion of differentially expressed genes hybridized to probes representing each library also varied (28%, 42% and 31% for gill, mantle and siphon respectively). The higher proportion of features with significant changes in intensity in the mantle could suggest that some transcripts show a tissue specific response. An accurate measure of the level of tissue specificity can only be achieved by hybridizing RNA extracted from different tissues. Nonetheless, it appears that most of the transcripts showing significant changes in expression from mantle RNA are not specific to mantle but generally shared across the tissues involved in the study.

3.4.1.2 RNA quality

RNA quality is a critical determinant in the analysis of gene expression. Preparation of high quality RNA from many Antarctic molluscan species has proved challenging (G. Burns, G. Hillyard pers.comm.). RNA extraction from some Antarctic species currently under investigation in our laboratory has required a high level of optimization. Commercial tools for extraction (e.g. Qiagen RNeasy mini kit) and stabilization (e.g. Ambion RNA later) are not appropriate for some species, limiting the resources available. Modifications to existing methods are continuously being developed for novel species and preservation and storage methods are currently being investigated. The peculiarities in RNA preparation and storage for many Antarctic species could be largely related to the enhanced enzymatic activity at low temperatures. Cold adapted enzymes have been shown to present lower activation energies than warm adapted enzymes and, when measured at a common temperature, orthologs from cold adapted species catalyze conversions at higher rates (Somero, 2004). This potentially reduces the effectiveness of the use of cold temperatures to ensure preservation and minimise degradation.

Although intact RNA constitutes the best representation of the natural state of the transcriptome, gene expression analysis may still be possible on partially degraded RNA (Imbeaud et al., 2005). Auer et al. (2003) demonstrated that low levels of RNA degradation do not preclude microarray analysis if the comparison is done using samples of similar RNA integrity. The results presented here also suggest that despite the difficulties of working with *L. elliptica* RNA, meaningful results can be achieved with careful optimization of existing methodologies and stringent quality control.

Typical high quality RNA is characterised by a ribosomal RNA band ratio (28S/18S) ≥ 2 and a ratio of absorbances (A_{260}:A_{280}) ≥ 1.8 (Sambrook and Russell, 2001). Evaluation of the ribosomal RNA band ratio was not possible in the current study as invertebrates do not exhibit a 28S RNA band when purified total RNA is run on a gel (Barcia et al., 1997). Moreover, the reliability of these methods has been extensively questioned (Imbeaud et al., 2005; Copois et al., 2007). Visualization of RNA

samples on agarose gels revealed consistent RNA integrity in most samples. Treatment differences could not be observed based on these parameters, but were clear after qualitative examination of banding patterns of RNA samples in high resolution acrylamide gels. Considering that the same extraction protocol was applied to all samples, differences in the 9°C group must be the result of the treatment. The reduced banding in the 9°C group compared to the 0°C control group might relate to more degraded RNA, or a reduced level of transcription. Exposure to extreme environmental stressors in molluscs has been shown to induce DNA damage and result in cellular death via both apoptosis and necrosis (Terahara and Takahashi, 2008). In *L. elliptica* individuals exposed to acute heat stress, oxygen consumption and heart beat rate collapse at 9°C and individuals begin to die (Peck et al., 2002). It is therefore not surprising that mRNA stability is affected by this extreme treatment and that these changes in the transcriptional machinery, and a high level of RNA degradation may explain the differences in RNA banding patterns in the 9°C group. The cumulative effect of all these factors could explain the poor labelling and hybridization problems observed in this treatment group.

3.4.1.3 Confirmation of microarray results via qPCR

The qPCR estimated fold changes of the five tested genes were in the same direction but of different magnitude than the microarray estimates. For the majority of the genes, qPCR estimates were inflated compared to the microarray results. Additionally, although the qPCR and microarray results showed similar patterns of expression, differences in gene expression between the treatments were not significant for the tested genes based on qPCR results. Moreover, individual variability in gene expression was considerably higher for qPCR than for microarray estimates, but this is due to technical differences discussed below.

Differences in expression levels and variability estimates, together with the lack of significance in the qPCR data can be largely explained by the differences in sample sizes. Microarray estimates were based on measurements of gene expression based on eight individual samples for the treated groups and eight pooled samples for the control, whereas qPCR analysis was performed in three biological replicates with three technical replicates each. Technical replication on the microarray is provided

by duplicate probes for every gene. Differences in sample sizes between the methods will have an effect on the estimation of average gene expression, as well as on the variance. The lack of significance in the qPCR data is not surprising given that the noise of individual variation, particularly with small samples sizes, will result in increased variance and associated low power.

The divergence in the estimation of expression levels for individual genes between qPCR and microarray data is not exclusive to this study. Although qPCR is commonly used as a validation tool for microarray data, the results obtained from both methods are often in disagreement, the causes of which have deserved a great deal of attention (Holland, 2002; Kang et al., 2000; Skern et al., 2005; Morey et al., 2006; Karlen et al., 2007). Several factors have been shown to contribute to this discrepancy including failure to meet the assumptions necessary for successful qPCR results, differences in the dynamic range covered and the normalization methods used by both techniques and non specific binding and cross-hybridization of labelled targets to microarray probes.

As a general rule, microarrays provide less than three orders of magnitude of dynamic range between non specific hybridization noise and the highest intensity gene specific signal, whereas the dynamic range of transcript abundance is often far greater (Holland, 2002). Moreover, the samples used in the current study did not utilise the full dynamic range of the microarray (see figure 3.5). Good agreement between different methods for gene expression quantification has been demonstrated for highly expressed yeast metabolic genes (Kang et al., 2000), but fail for low abundance transcripts (Holland, 2002). The limited dynamic range as a result of low signal in the microarray data presented may have influenced the disagreement between microarray and qPCR results.

The type of normalization used can also have an effect on the agreement between microarray and qPCR data. Data normalization differs between the two methods. QPCR normalization uses the expression of a reference gene to individually calibrate gene expression against. Microarray normalization is usually performed in a global scale to remove non biological sources of variation within and between microarrays. Identifying suitable control genes is often challenging in qPCR experiments. These

are used to correct for any possible technical variations during the qPCR procedure, namely incorrect estimation of cDNA concentrations or pipetting errors. Their transcription should therefore not be affected within the experimental context being investigated. Although many genes are currently being used as standards, these may still show some degree of variability depending on the experimental treatment used. No gene has been shown to be completely invariant, and a slight variability translates to incorrect normalization of the gene of interest. The factors causing such variation are often unidentified, but two prevailing reasons have been highlighted: the cross-reactive response to pseudogenes and the presence of differentially expressed isoforms (Sturzenbaum and Kille, 2001). The printiploess and Gquantile normalization methods used, correct the expression ratios both for within microarray spatial variation and for intensity-based trends, and force all the reference values to have the same distribution, adjusting the values in the other channel accordingly. This results in adjustments which allow comparisons between microarrays but have an effect on the accuracy of the estimation of expression levels. Nonetheless, these are necessary for comparison between individuals, as well as between methods.

Despite the disagreement in the estimation of gene expression levels by both methods, it is clear that both methods produce similar patterns in gene expression for the five genes tested. A significant positive correlation between gene expression estimates by both methods has been shown (Pearson correlation $r=0.69$, $P=0.005$). Since both techniques are semi quantitative the results need to be interpreted in relative terms. Until a standard definition of validation of microarray results is established, agreement in the direction of the changes is of greater relevance than the assignation of numerical values.

3.4.2 Heterologous hybridization

The use of heterologous microarray experiments is gaining importance in non model organisms, for which few genomic resources are available. The benefits are clear, as the cost and time required to produce new microarray resources for every novel species is reduced. However, the effectiveness of this approach is affected primarily by the sequence divergence between the species for which the microarray was constructed, and the species from which a sample is hybridized (Buckley, 2007).

Limited sequence similarity between target and probe can be reflected in the microarray in different ways, related to the extent of sequence divergence (Kassahn, 2008). Firstly, high levels of sequence divergence will result in failure of some probes to hybridize, reducing the number of gene transcripts that can be measured. However, the number of missing features in the present study was comparable in both species, after lowering the stringency of the washes during the *M. edulis* hybridizations. This suggests that, although with some level of cross-reactivity, sequence divergence is not enough to prevent binding of all probes. It is likely that hybridization will be restricted to the most highly conserved genes. Secondly, for those transcripts presenting enough sequence similarity to cross-hybridize, intensities might be reduced due to poor binding. However, the distribution of feature signal intensities in the *M. edulis* microarray approximates that of *L. elliptica*, suggesting that hybridizing *M. edulis* RNA to the *L. elliptica* microarray does not result in a consistent decrease of signal intensities across all features. Finally, in the absence of the same gene transcripts in the target sample as those represented in the probes, cross-reactivity of unrelated transcripts can affect specificity. A strong positive correlation between *L. elliptica* and *M. edulis* signal intensities across features could suggest that most gene transcripts have comparable levels of expression in both species. If similar expression levels for the same gene in both species are assumed, then the observed correlation is unlikely to reflect a high degree of cross-reactivity of unrelated transcripts.

Although validation is required, the preliminary results presented here suggest that the *L. elliptica* microarray is likely to be useful to study responses to environmental stressors in *M. edulis* and perhaps other bivalve species. The use of the *L. elliptica* microarray with related species provides promising means by which the universality of the stress response can be further tested.

3.4.3 Effect of temperature increases in gene expression

Although a library of 8448 clones was used in this study, the potential redundancy of the library and limited sequence information means that the number of unique genes remains unknown. It is likely that other genes involved in the heat response are not represented in the microarray. In addition, only a small proportion of genes showing

significant changes in expression between the control and treated group were identified by sequence similarity searching. The lack of information for the rest of the genes highlights the limitations in the identification of a complete stress response in non model organisms. Notwithstanding the limitations, this study provides further confirmation for some previously described aspects of the stress response, and identifies novel genes potentially involved in the heat stress response of *L. elliptica.*

Genes showing significant changes in expression that were identified by sequence similarity searching were found to be associated with a wide range of functions. These have been assigned to a series of functional categories detailed below that include DNA synthesis and protein turnover, protein folding and chaperoning, antioxidant activity, intracellular signalling and trafficking, cytoskeletal activity and energy metabolism. The putative function of these genes is discussed below.

3.4.3.1 Protein turnover

Suppression of protein synthesis during heat stress has been shown in a number of model organisms (Causton et al., 2001; Sørensen et al., 2005) and has been suggested to provide energy savings to cells under stress, so the energy can be utilised to maintain critical cellular functions. Inhibition of protein synthesis occurs by reducing mRNA substrate availability, or by inhibition of the ribosomal translational machinery (Storey and Storey, 2004). However, the changes in expression of a large number of transcripts involved in transcription and translation observed in this study, together with genes involved in protein degradation, suggests an increase in protein turnover during acute exposure to elevated temperatures in *L. elliptica.* This observation matches those of other microarray studies in non model marine species (Teranishi and Stillman, 2007; Podrabsky and Somero, 2004; Buckley et al., 2006; Buckley and Somero, 2009; Kassahn et al., 2007).

Evidence for changes in transcription in response to temperature may be provided by the increased expression of a high mobility group protein gene (HMGB). HMG proteins are highly conserved, abundant non-histone nuclear proteins that bind to DNA or chromatin and affect DNA replication, repair, transcription and chromatin compaction (Bustin, 1999). They are divided into three families, one of which, the

HMGB (HMGB1, B2 and B3) is a family of very closely related genes containing a HMG-box domain that binds DNA affecting transcription (Thomas, 2001). Due to the short length of the *L. elliptica* EST sequence (240nts) and the high level of conservation in the HMGB protein family (70-80% similarity between B1, B2 and B3 from human) it was not possible to determine with confidence which member of the gene family had the closest match to the *L. elliptica* sequence.

Podrabsky and Somero (2004) observed a striking response in the expression of the HMGB1 transcript with cyclic temperatures in the killifish, where expression patterns and magnitude were highly negatively correlated with temperature. The observed patterns, together with the high thermal sensitivity of HMGB1 and its role in the global regulation of transcription led them to propose that HMGB1 protein may be a critical part of a compensatory transcriptional response to temperature and may indeed be a highly sensitive temperature sensor. In the current study, a HMGB transcript was significantly positively correlated with temperature (FC=2.89, P=0.001). It is not possible to determine whether such differences in behaviour observed in the two studies involve different members of the HMG protein family, or relate to differences in the length and intensity of the exposure. Nonetheless, this study provides further support for the involvement of HMG proteins in temperature mediated changes in transcription. The HMG proteins could be excellent candidates for further work to determine their behaviour in response to different levels of stress.

Further evidence for changes in transcription could be provided by the upregulation of a gene coding histone protein H2A. Histones are basic chromosomal proteins extremely conserved throughout evolution. They are essential structural proteins involved in chromatin packing, which actively participate in the regulation of gene transcription. Although histones are putative housekeeping proteins, they have been previously shown to respond to environmental stressors. Dondero et al. (2006) found the expression of histone genes in mussels to be significantly upregulated in response to copper pollution. Previous work on the yeast *S. cerevisiae* has suggested the participation of histone H2B in repair of UV-induced DNA damage (Martini et al., 2002) and in *S. Pombe* histone H2B has been shown to decrease significantly in response to heat stress (Chen et al., 2003). In addition to their *de novo* synthesis, post translational modifications of histones, such as phosphorylation of histone H3,

constitute an apparent response to various environmental stressors (Burkhart et al., 2007). Histone phosphorylation also occurs during DNA damage and apoptosis (Fernandez-Capetillo et al., 2004). The role of histones in preventing and repairing damage from environmental stress is poorly understood. Nonetheless, several studies have pointed out that the increase in either histone synthesis or modification, seen in response to abiotic stressors, may result in chromatin remodelling and potential changes in gene expression. Although the link between increased histone genes synthesis and heat stress is currently unknown, this may arise from the need to respond to a demand for increased transcription, and/or repair of damaged DNA.

A link between transcription and translation might be provided by the upregulation of the pre-mRNA-processing factor 19 transcript. Genes are transcribed under the control of transcription factors into pre-messenger RNA (pre-mRNA), which is then processed by splicing out introns into a mature mRNA. This is then transported to the cytoplasm, where it serves as a template for protein synthesis (Orphanides and Reinberg, 2002). Pre-mRNA processing involves capping of the 5′end to protect it from exonucleases; editing mainly through deamination; splicing to ensure the removal of introns; and 3′end processing involving cleavage and polyadenylation (Soller, 2006). The splicing reaction occurs in the spliceosome (Brody and Abelson, 1985), a large complex comprising the precursor mRNA, five U-class small nuclear ribonucleoprotein particles (snRNPs) and a number of associated protein factors required for the different steps of the splicing reaction. Pre-mRNA-processing factor 19 (PRP19) has been shown to be essential as a spliceosome component in yeast, where mutants are defective in spliceosome assembly; as it is required for the first cleavage-ligation of the splicing reaction (Cheng et al., 1993). In addition to this role in splicing, PRP19 has been shown to possess E3 activity in the presence of certain ubiquitin domains (Ubc2B or Ubc3) acting as E2 (ubiquitin conjugating enzymes). This ubiquitin ligase activity is due to the existence of a U-box domain in the amino terminus of the gene (Hatakeyama and Nakayama, 2003). Proteins with E3 activity associate with molecular chaperones that mediate the recognition of target proteins for ubiquitylation. Based on this activity, PRP19 has been suggested to be involved in the regulation of RNA splicing by mediating the ubiquitylation and degradation of spliceosomal components (Hatakeyama and Nakayama, 2003). In addition to the role of PRP19 in splicing, the complex formed by Prp19 and its allelic form Pso4,

has been shown to have a direct role in mediating the cellular response to DNA damage in mammals, and to reduce the levels of apoptosis after exposure of cells to DNA damage (Lu and Legerski, 2007). As for the histone genes previously discussed, it is not possible to determine whether the changes in expression of this gene reflect an attempt to accommodate to new environmental conditions, or to repair the damage caused by them.

Evidence of increased translation was provided by changes in expression of the dolichyl-diphosphooligosaccharide-protein glycosyltransferase and bifunctional aminoacyl-tRNA synthetase genes. The enzyme dolichyl-diphosphooligosaccharide-protein glycosyltransferase catalyzes the co-translational transfer of an oligosaccharide chain from a dolichol-linked oligosaccharide donor to certain asparagine residues present on the nascent polypeptide. This highly conserved pathway is part of the mechanism of N-glycosylation, the most ubiquitous protein co-translational modification in the endoplasmatic reticulum (ER) (Yan and Lennarz, 2005). N-linked glycans can play an important role in the folding and stability of many proteins, favouring the generation of the native structure. Certain glycans may also have an effect on the refolding of denatured proteins (Ceriotti et al., 1998), often produced in greater quantities in animals undergoing stressful conditions. Aminoacyl-tRNA synthetases are enzymes that catalyze the formation of aminoacylated tRNA (Aa-tRNA) in the presence of ATP. Aa-tRNA synthetases recognise individual amino acids and the tRNA specific to them and bind to their respective codons taking part in bond formation. The charged tRNAs are then available for polypeptide chain elongation in the ribosome (Cusack, 1997). Aa-tRNAs are therefore essential to ensure accurate translation of the genetic message and the increase in expression levels of aa-tRNA synthetase indicates active transcription and maturation of tRNA genes.

High translational activities demand a high number of ribosomal proteins for their ribosomes, which constitute the core of translation, necessary to decode the information contained in mRNA into amino acid chains. Ribosomes therefore constitute the link between the transcriptome and the proteome. They are made up of two subunits of unequal size. In eukaryotes, these are a small 40S subunit, with a decoding function; and a large 60S subunit with peptidyl transferase activity. The

small and large subunits contain 32 and 46 ribosomal proteins respectively in yeast, a number that varies slightly between species. It has become clear that the active sites in the ribosome consist of RNA, whose assembly and functions are assisted by ribosomal proteins (Lafontaine and Tollervey, 2001)

The assemblage of ribosomal proteins and RNA into ribosomal units is a very efficient process in that it involves equimolar quantities of each ribosomal protein and rRNA (Tsay et al., 1988). In this study, this equimolar requirement was not obvious at the transcript level, where highly variable mRNA levels of various ribosomal protein genes were observed. This observation agrees with previous studies (Karsi et al., 2002; Buckley et al., 2006; Buckley and Somero, 2009), which propose that since ribosomal proteins have a stabilising function in the ribosome, their repression or expression during heat stress may be an effort to protect ribosomal structure and function through replacement of damaged sub regions. This is possible, as the cellular response to different biotic and abiotic factors results in depression or activation of individual ribosomal protein genes, whose transcription is regulated independently (Nomura, 1999). It is also possible that some differences in expression of ribosomal protein genes at the transcript level relate to differences in mRNA half lives. The regulatory processes that ensure the specific stochiometric ratio of ribosomal proteins in the ribosome are complex and not fully understood but, given the differences in expression of individual genes, these must involve post transcriptional mechanisms. It is striking that all upregulated ribosomal proteins in this study are components of the small 40S subunit, whereas all downregulated proteins belong to the large 60S subunit. It is possible that the small ribosomal subunit is more prone to damage during heat stress in *L. elliptica*. However, previous studies showing mixed expression patterns within each subunit would suggest that both units require protein replacements. Under conditions of mRNA accumulation in yeast, translation is not repressed, and excess proteins undergo rapid degradation (Maicas et al., 1988) thus any potential overproduction of 40S ribosomal protein transcripts would result in degradation of excess proteins.

Evidence of ribosome biogenesis or repair was also provided by Rlp24, a highly conserved protein isolated in yeast that belongs to the family of eukaryotic Rpl24e proteins (Saveanu et al., 2001). Rlp24 is a pre-ribosomal protein associated with pre-

60S complexes along the 60S ribosomal subunit biogenesis pathway from its assembly in the nucleolus to its maturation in the cytoplasm (Saveanu et al., 2003). The increase in expression of the ribosome biogenesis protein Rlp24 suggests that despite the decrease in expression levels of the 60S ribosomal protein genes, 60S ribosomal subunit biogenesis still occurs.

Some evidence of the synthesis of essential amino acids has also been observed. Betaine--homocysteine S-methyltransferase (BHMT) is best known as a regulator of homocysteine metabolism. It catalyzes resynthesis of methionine using homocysteine and the methyl donor betaine, thus its upregulation could indicate that homocysteine is being utilised to synthesise the essential amino acid methionine, required for protein synthesis. In a study of the rat liver proteome, BHMT was shown to be upregulated following heat stress (Rajaseger et al., 2009), and although the authors suggested a link to protection from oxidative stress, the mechanisms for this remain unknown.

Elevated temperatures are known to cause damage and misfolding of proteins (Parsell and Lindquist, 1993). Some of these damaged proteins can be rescued via the activity of molecular chaperones, as discussed in the following section, whereas irreversibly damaged proteins are ubiquinated and removed by proteolytic degradation (Kültz, 2005). In the present study, evidence for protein damage is provided by the increase in synthesis of molecular chaperones, as well as indicators of increased protein degradation. Proteasome subunit beta type-5, a proteasome component involved in ATP dependent proteolytic activity, was found to be significantly upregulated. The induction of this protease suggests a need to remove denatured proteins, potentially damaged through increased temperature.

An upregulated gene also potentially involved in protein degradation was identified by sequence similarity searching as a BTB/POZ domain-containing protein KCTD7. The high sequence variability of the BTB makes the identification of BTB proteins in databases challenging (Perez-Torrado et al., 2006). Nonetheless, a domain search of the translated cDNA sequence in the NCBI CDD indicated that the BTB domain is present in the cDNA sequence ($5e^{-11}$). The BTB or POZ domain is a protein-protein interaction motif present in viruses and throughout eukaryotes. Although some

proteins are made up of just a BTB, it is more frequent for it to be combined with other domains; such as MATH, found in proteins involved in cytoplasmatic signal; Kelch, a structure that can interact with actin filaments; NPH3, which is plant specific; Ion transport domains, such as the voltage-gated channel for potassium ions; and Zinc finger motifs, which are involved in transcriptional regulation. In spite of their functional variety, they all use the BTB domain as a protein–protein interface, either for oligomerization or for interaction with other proteins (reviewed in Perez-Torrado et al. (2006). Despite their multiple functions, BTB-containing proteins have at least one widely shared function: the regulation of ubiquitin dependent protein degradation by recruiting proteins targeted for degradation and acting as adaptor molecules for cullin-based E3 ubiquitin ligases (Van Den Heuvel, 2004). Given the lack of a full length sequence to determine the full structure and presence of other domains, it was not possible to determine the specific function of the BTB protein gene upregulated in the present study. This gene could potentially be involved in one or more of the relevant functions in stress response, including transcriptional regulation, protein degradation, signalling and trafficking through K^+ ion channels or actin regulation.

The changes in expression of genes involved in protein turnover in thermally stressed individuals are not surprising. Enhanced protein synthesis may reflect the necessity to accommodate to new external conditions, whereas increased degradation suggests that, despite the rescue of misfolded proteins discussed below, heat induced irreversible protein damage also occurs.

3.4.3.2 Protein folding and chaperoning

The induction of heat shock proteins and their role as molecular chaperones has been shown to be central to the cellular stress response, not only to temperature, but to other environmental stressors (Feder and Hofmann, 1999). Heat shock proteins have been shown to be amongst the most strongly inducible chaperones at the gene and protein level during the stress response of marine species including fish (Dietz and Somero, 1992; Podrabsky and Somero, 2004; Das et al., 2005; Buckley et al., 2006) , molluscs (Smerdon et al., 1995; Tomanek and Somero, 2000; Place et al., 2008), crustaceans (Teranishi and Stillman, 2007) and cnidaria (Sharp et al., 1997).

Antarctic species, however, represent a special case in the study of thermal stress. Some Antarctic species constitutively express high levels of the inducible form, which is not significantly upregulated in response to elevated temperatures. For those species that do present a response, induction temperatures tend to be higher than predicted from observations in temperate species (reviewed in Clark and Peck, 2009). *L. elliptica* is able to induce transcription of heat shock protein genes following acute exposure to 8-10°C (Park et al., 2007; Clark et al., 2008). Based on these studies, HSP genes were not expected to be induced at 3°C. However, exposure to heat is expected to result in damaged and denatured proteins (Parsell and Lindquist, 1993) and so the induction of other genes acting as chaperones and assisting in protein folding is not surprising.

A prefoldin gene transcript was upregulated in this study. Prefoldin is a hexameric molecular chaperone complex encoded by at least six different genes, which is found in all archaea and eukaryotes. The eukaryotic prefoldin complex is built of two α and four β subunits (Siegert et al., 2000). Prefoldin binds specifically to cytosolic chaperonins, multisubunit assemblages that promote the ATP dependent folding of proteins both under normal and stress conditions by providing a cavity in which non-native polypeptides can be enclosed and assisted along their folding pathway. Chaperonins can prevent aggregation of unfolded polypeptides, bind and structurally rearrange polypeptides that fail to fold, and release unfolded proteins into enclosed folding compartments (Hartl, 1996). By transferring target proteins to chaperonins, prefoldin promotes folding of newly synthesised or denatured proteins in an environment in which there are many competing pathways for non-native proteins (Vainberg et al., 1998). Further confirmation of the changes in expression of chaperonins is provided at the protein level and discussed in chapter four.

3.4.3.3 Antioxidant activity

Eukaryotic cells depend on the availability of oxygen for their survival. It is used as the final electron acceptor in the mitochondrial respiratory chain during the normal metabolism of the aerobic cell. It ultimately allows oxidative phosphorylation and the generation of ATP. However, by successive transfer of electrons, the stable oxygen molecule generates reactive oxygen species (ROS) which have a strong

reactivity and can cause a series of oxidations and eventually interact with all cellular components. ROS are beneficial in that they are potential stress sensors (Toescu, 2004) and act as messengers in signalling pathways triggering the stress response, but under certain conditions, they can be highly deleterious (Kültz, 2005). During normal aerobic metabolism, organisms utilize a pool of antioxidant defences including enzymes (e.g. superoxide dismutase, catalase, and glutathione related enzymes) and other non-enzymatic molecules (e.g. β-carotene, glutathione, uric acid, vitamins A, E and C). An increase in ROS, triggered by environmental stimuli, leads to oxidative stress and to a corresponding increase in antioxidant defences. Under prolonged oxidative stress, the antioxidant systems may be exceeded leading to DNA damage, lipid peroxidation, protein degradation, metabolic malfunction and cellular damage (Winston and Digliulio 1991). Heat stress induces oxidative stress through the disruption of electron flow during mitochondrial electron transport (Davidson and Schiestl, 2001). The accumulation of ROS and induction of antioxidant defences during heat stress has been shown in previous studies in *L. elliptica* (Heise et al., 2003) and other marine species (Abele et al., 1998; Verlecar et al., 2007). Although a significant increase in ROS production in *L. elliptica* isolated mitochondria was not observed below 7°C *in vitro* (Heise et al., 2003), the results presented here suggested that exposure to 3°C was sufficient to produce an increase in ROS *in vivo*, as indicated by increased antioxidant activities.

Despite their low metabolic rate (Clarke, 2003) and reduced mitochondrial respiration and ROS output (Heise et al., 2003), polar bivalves have been shown to have elevated levels of antioxidants compared to temperate species (Camus et al., 2005). *L. elliptica* has been shown to have significantly higher levels of the antioxidants catalase and glutathione (Philipp et al., 2005a) and tocopherol and β-carotene (Estevez et al., 2002) than its temperate counterparts throughout its lifetime. Regardless of the initial high levels when compared to temperate species, the significant increase in the transcription of antioxidant genes observed in this study suggests that heat stressed *L. elliptica* may also undergo some degree of oxidative stress.

Peroxiredoxin-1 belongs to a family of ubiquitous antioxidant proteins with peroxidase activity, that reduce and detoxify H_2O_2, peroxynitrite and a range of

organic hydroperoxides. Their peroxidase activity overlaps with that of other antioxidant enzymes with higher catalytic efficiencies such as glutathione peroxidases and catalases (reviewed in Wood et al., 2003). The diversity in their function, from antioxidant activity to regulation of signal transduction, is reflected in evolutionary modifications in sequence and structure, built around a common peroxidatic active site. Many organisms produce more than one isoform of peroxiredoxins, with at least six peroxiredoxin isoforms found in humans, PrxI–PrxVI (Hofmann et al., 2002). Heat stress has been previously shown to induce the expression of two peroxiredoxin isoforms in *L. elliptica* (Park et al., 2008a). In the same study, H_2O_2 induced an increase in the expression of the same isoforms, and a change in pI values, suggesting a change to an acidic state, in some isoforms but not others. This may suggest that the different isoforms play different roles in the mechanisms to protect against oxidative stress in this species. In the current study, the EST identified as peroxiredoxin was not similar to the two annotated *L. elliptica* isoforms (ACE76885, ACE76884). However, given the short length of the sequences, it cannot be determined with confidence whether it represents a novel isoform.

A glutathione S-transferase gene was also significantly upregulated. The glutathione S-transferases (GSTs) are a large protein superfamily with several subfamilies identified in eukaryotes including soluble cytosolic, Kappa and microsomal GSTs, involved in glutathione metabolism (Frova, 2006). The complexity of the GST superfamily has not allowed a clear definition of the subfamilies or the multigene families they are encoded by. GSTs catalyze the conjugation of tripeptide glutathione with electrophilic compounds to detoxify them, thus they are essential for the protection against oxidative stress (Rushmore and Pickett, 1993). In addition, ROS have been shown to activate gene transcription through a regulatory element in the GST sequence, the antioxidant-responsive element (ARE) (Rushmore et al., 1991). This has been suggested to be part of a signal transduction pathway through which eukaryotic cells respond to oxidative stress. Activation of genes containing an ARE could lead to the induction of enzymes that protect the cells from compounds that undergo redox cycling and form reactive oxygen species. Changes in GST activity have also been demonstrated at the protein level in thermally stressed *L.elliptica*. Park et al. (2008b) measured the activity of antioxidant enzymes in

individuals exposed to 10°C and suggested that GSH and GSH related enzymes, including GST, play a major role in the protection against the toxicity of ROS during acute thermal stress. The lack of sequence information in their study did not allow identification of the GSTs involved in the measured response and their use of 1-chloro-2, 4-dinitbenzene and GSH as substrates meant that the activity of several members of the family could have been measured simultaneously. Sequence alignment with each of the three *L. elliptica* GST gene sequences in NCBI (accession numbers ACM44934, ACM44933, ABV44413) showed limited identity at the amino acid level, which is likely to be related to the length of the sequences, while reflecting the complexity of the GST superfamily. Nonetheless, some evidence of the involvement of GSTs in the heat stress response of *Lelliptica* has now been provided both at the mRNA and protein levels.

The results presented here suggest that exposure to 3°C was sufficient to produce an increase in ROS *in vivo* at the organismal level, as indicated by increased antioxidant activity.

3.4.3.4 Intracellular trafficking and signalling

Cells sense changes in the environment and convert an external signal into a series of signalling cascades that eventually leads to a response. Several trafficking and signalling related genes showed changes in expression in response to temperature in *L. elliptica* in this study, which is in accordance with previous studies of temperature stress in non model species (Buckley et al., 2006; Buckley and Somero, 2008; Desalvo et al., 2008)

A 14-3-3 gene homolog was upregulated in this study. 14-3-3 proteins are a large family of acidic proteins with a high degree of sequence identity. They coordinate multiple signalling pathways by binding specific proteins and regulating their activity (Takahashi, 2003). As a result, they are involved in a vast range of processes such as response to stress, redox regulation, cell cycle, transcription, cytoskeletal reorganization and further cellular trafficking (Kjarland et al., 2006). Another upregulated gene was serine/threonine-protein kinase (PAK 3) part of a ubiquitous family of serine/threonine p21 activated kinases. They have been shown to be

involved in the regulation of cellular processes such as gene transcription, cell morphology or cytoskeletal activities (Bagrodia and Cerione, 1999). Both 14-3-3 and PAK genes have been shown to be differentially expressed in heat stressed Antarctic fish (Buckley and Somero, 2009). Their involvement in a large variety of processes classically related to stress may suggest a fundamental role in stress signalling

A low-density lipoprotein receptor (LDLR) was also upregulated. LDLR belongs to an evolutionarily conserved group of cell-surface receptors. LDLR binds to proteases, protease inhibitors, signalling molecules, heat-shock proteins, vitamin carriers, toxins and antibiotics amongst other molecules. This results in endocytic uptake of the ligands, and ultimately affects cellular functions including migration, signal transduction, proteolysis, antigen presentation and synaptic activity. In addition, they present unusual properties for endocytic receptors, as they also present motifs for cytosolic adaptor or scaffold proteins, which in turn interact with other cellular factors (e.g. actin cytoskeleton, tyrosine kinases, transcription factors and enzymes) involved in cellular trafficking, signalling and other cellular pathways (reviewed in Nykjaer and Willnow, 2002). The multifunctional nature of the LDLR does not allow further characterization of its function but it appears to be involved in trafficking and signalling in response to heat stress.

Of particular interest in this study is the upregulation of a number of genes identified as associated with calcium (Ca^{2+}) transport and homeostasis. The epithelial chloride channel protein is a voltage operated Ca^{2+} channel, with a specific Ca^{2+} permeability, that allows the movement of Ca^{2+} between membranes. Calmodulin (CaM), centrin and the sarcoplasmic calcium-binding protein all contain structural Ca^{2+} binding motifs known as EF-hands. These belong to a group of Ca^{2+} binding proteins that act as intermediaries coupling the calcium signals to cellular and biochemical changes. Calmodulin has received special interest as in binding Ca^{2+} and in turn other proteins it serves as an adaptor of Ca^{2+} signals, and it is therefore involved in a wide range of functions.

Ca^{2+} is a universal second messenger, which controls a vast number of cellular reactions and adaptive responses, and is crucial for a range of biological functions.

The molecular systems responsible for Ca^{2+} signalling are limited to a few very well conserved and ubiquitous protein families of Ca^{2+} channels and transporters. These work and interact in different manners, creating multiple Ca^{2+} signalling pathways to suit the physiology of the cell. As a universal messenger in signalling, calcium is involved in muscle contraction, regulation of neural activity and cell proliferation amongst other processes (Bootman et al., 2001). In the ER, changes in Ca^{2+} concentrations constitute essential signalling triggers in diverse aspects of ER function. These include protein folding, lipid synthesis, protein secretion and post-translational modifications, ER stress responses, sensitivity to apoptosis, and modulation of transcriptional processes and developmental pathways (Michalak et al., 2002). A potentially significant process, in the context of heat stress, in which Ca^{2+} signalling and Ca^{2+} binding proteins are largely involved, is hypoxia sensing and reactivity to changes in oxygen concentrations in the cell (reviewed by Toescu, 2004). In certain cells, changes in partial oxygen pressure lead to changes in Ca^{2+} concentrations. A series of signalling pathways involving Ca^{2+} are initiated and as the changes are detected elsewhere, the signal is converted into functions that maintain oxygen homeostasis by altering protein and gene expression levels to increase oxygen delivery to hypoxic tissues and organs. Although the mechanisms for hypoxia sensing and those that link sensors to effects remain unclear, alterations in intracellular Ca^{2+} concentrations and activation of Ca^{2+} signalling pathways constitute an early response of many cell types to low levels of hypoxia (Toescu, 2004).

An important aspect of ionic regulation is the regulation of pH. The imidazole α-stat (alphastat) hypothesis (Reeves, 1972) postulates that the degree of ionisation (alpha) of the imidazole groups of intracellular proteins remains constant despite changes in temperature. Ectotherms would regulate the pH in their body fluids to maintain protonation of these groups and ensure the structural integrity of proteins in changing temperature conditions. In order for this to occur, a buffering system is required that allows changes in pH to occur linearly with temperature. Mobilization of $CaCO_3$ from the shell and the use of calcium as a buffer mechanism to compensate acidosis during emersion have been previously documented in bivalves (Byrne and McMahon, 1991; Byrne et al., 1991). A series of studies have linked calcium, pH and the neutralization of anaerobiosis by-products (Crenshaw, 1972; Akberali et al.,

1977). However, these studies relate the pH buffering properties of shell Ca^{2+} to the sequestration and buffering of anaerobic end products during aerial exposure, and provide no evidence of the involvement of calcium in alphastat regulation. In fact, alphastat regulation and the linear relationship between pH and temperature have been shown not to hold once critical temperatures are reached (Sommer et al., 1997), when the transition to anaerobic metabolism occurs. pH buffering due to temperature induced anaerobiosis is unlikely, as *L. elliptica* has been shown to start accumulating anaerobiosis end products at higher temperatures (Peck et al., 2002). Moreover, in order to minimise ATP expenditure, Ca^{2+} channel activity decreases progressively with length of anoxia, until channel closure occurs. This prevents the rising intracellular Ca^{2+} arising from shell dissolution to enter tissues where it can be highly deleterious (Storey and Storey, 2004)

The occurrence of alphastat pH regulation has been demonstrated in some cold adapted species, whereas others fail to show any evidence of this regulatory mechanism (reviewed in Pörtner et al, 1998). Nonetheless, this hypothesis cannot be excluded without further analysis involving measurement of intracellular pH changes in *L. elliptica* exposed to different temperatures. Hybridization of different tissues to the microarray could also determine whether the involvement of Ca^{2+} pathways in the stress response is limited to the mantle, the organ which actively secretes the shell matrix, which could provide some evidence that Ca^{2+} might be involved in alphastat regulation.

The exact process or processes in which calcium signalling is involved in *L. elliptica* exposed to increased temperatures can not be elucidated from this study. Nonetheless, an increase in expression of the genes involved in calcium homeostasis appears to be triggered by heat stress. It is possible that heat leads to oxidative stress, which triggers Ca^{2+} signalling pathways involved in a variety of functions implicated in maintaining homeostasis. These results suggest that Ca^{2+} signalling, and in fact Ca^{2+} itself, deserve attention as potential indicators of early and late signs of heat stress in *L. elliptica* and, given the universality of Ca^{2+} as a messenger, perhaps in other taxa.

Calcium homeostasis has been previously suggested to be affected by thermal stress and bleaching in the Caribbean coral *M .faveolata* through the changes in expression of genes related to calcium transport (Desalvo et al., 2008). CaM and an unidentified EF-hand protein were downregulated in bleached corals, which was interpreted as a disruption of Ca^{2+} homeostasis during heat stress and bleaching. However, in the present study both genes were upregulated which may suggest an active response to maintain Ca^{2+} homeostasis. Given the differences in the severity of the treatment between the two studies, it is possible that the expression of CaM, EF-hand protein genes and other genes involved in Ca^{2+} transport may be modulated by the different stages of heat stress. CaM and EF-hand genes may be potential candidates for future work on protein function in relation to heat stress.

The induction of signalling genes suggests that the stress response of *L. elliptica* involves the synthesis of cell signalling proteins that may result in biochemical changes needed to cope with elevated temperatures. The diversity of the signalling pathways induced by heat stress reflects the complexity of the response.

3.4.3.5 Cytoskeletal activity

Several cytoskeletal protein transcripts were upregulated in response to temperature in this study. These include the structural components actin and myosin; and the actin regulatory proteins tropomyosin and calponin. In the present study two different actin isoforms were identified by sequence similarity searching. Actin belongs to a multigene superfamily with at least eight actin and actin related protein subfamilies conserved from human to yeast (Goodson and Hawse, 2002). They perform roles in fundamental aspects of cell biology, of which cytoskeletal function has received the most attention. Given the diversity within the actin superfamily, it is not possible to determine the orthologues represented in the microarray, but it is clear that at least some isoforms are involved in the *L. elliptica* response to elevated temperatures.

The upregulation of cytoskeletal protein transcripts is in agreement with previous studies in marine species, which have shown cytoskeletal proteins to be regulated in response to stress (Ju et al., 2002, Kassahn et al., 2007; Podrabsky and Somero,

2004; Buckley et al., 2006; Sarmiento et al., 2000; Buckley and Somero, 2009; Teranishi and Stillman, 2007). Buckley and colleagues proposed two explanations for the induced expression of cytoskeletal proteins in relation to temperature. They argued that certain cytoskeletal components may be more susceptible to damage by temperature than other proteins, hence the need for newly synthesised proteins to replace those degraded. An alternative explanation suggests that heat stress results in extensive cytoskeletal organization, required to overcome osmotic stress, a potential secondary effect of heat stress, and regain osmotic balance. Although both explanations have been proposed in previous studies, the latter is favoured by the suggestion that the temperature dependent induction of cytoskeletal components may be tissue specific. In addition, the induction of genes involved in Ca^{2+} signalling may support this idea due to the direct link between Ca^{2+} homeostasis and cytoskeletal rearrangements (Orrenius et al., 1992).

An increase in the level of expression of these genes may also be required to respond to a demand in muscle contraction. Actin, tropomyosin and myosin are functionally linked, as tropomyosin controls muscle contraction through regulation of the thin filament interactions with myosin motors and actin-binding proteins (Skorzewski et al., 2009) a process in which Ca^{2+} plays a major role (Ebashi and Endo, 1968). *L. elliptica* is an infaunal species with a burying mechanism typical of other bivalves. It starts by extending its foot to penetrate into the sediment, followed by a foot dilation to form an anchor, and drawing down the shell by contraction of the pedal muscles (Peck et al., 2004). During preliminary experiments, individuals exposed to 3°C were observed extending their foot and actively attempting to bury. Indeed, despite being above its optimal temperature, the fastest burying rate in *L. elliptica* occurs at 3°C (Peck et al., 2007). Burrowing movements could partially explain the demand for contractile proteins. This is potentially linked to the identification of Ca^{2+} transport genes detailed before.

3.4.3.6 Energy metabolism

As expected, the level of expression of genes for ATP generation (ATP synthase) increased significantly in response to heat stress. This may be related to the need for energy associated with the changes required to cope with heat stress. Ribosome

biogenesis is a major consumer of cellular energy, but protein chaperoning and degradation also demand large amounts, as do DNA repair, muscle contraction and cytoskeletal activity.

3.4.3.7 Acute phase proteins

Inter-alpha-trypsin inhibitors (IαI) are a family of structurally related plasma protease inhibitors. They are protein-glycosaminoglycan-protein complexes that combine several types of heavy chains (HC), covalently linked with the common light chain bikunin, a proteoglycan responsible for the protease inhibitor activity. All polypeptides are encoded by distinct genes (Diarramehrpour et al., 1989). Under certain pathological conditions, bikunin free complexes also occur which comprise one or more heavy chains covalently bound to hyaluronic acid (HA), a glycosaminoglycan (Zhuo et al., 2004). In this form, IαI have been associated with oozyte maturation, development and acute inflammation (Salier et al., 1996). Acute inflammation is the tissue response to non lethal injury caused by physical, chemical, or biological agents. It results in a vast number of physiological, metabolic and endocrine changes known as the acute phase response, responsible for ensuring survival immediately after the insult. This involves the synthesis of acute phase proteins, including antiproteases, and allows survival following injury, through the containment or destruction of infectious agents, removal of damaged tissue, and restoration of the healthy state (Kushner, 1982).

An IαI heavy chain mRNA was upregulated in this study. Although this was identified as H4, given the short length of the sequence and the high sequence similarity between the HCs (Chan et al., 1995) it was not possible to determine with confidence which member of the family this is. Nonetheless, induction of HCs genes in response to heat stress has been previously shown in Antarctic species. Buckley and Somero (2009) found H2 to be upregulated in response to heat stress in the gill of the Antarctic fish *Trematomus bernacchii*, and a microarray study on heat stress in the Antarctic fish *Harpagifer antarcticus* revealed changes in the expression on the H3 gene (M. Thorne pers.com.). Evidence of an acute phase response in bivalves is given by studies in host defence and immune response, and IαI have been shown to be expressed at the mRNA level in the zhikong scallop *Chlamys farreri* (Wang et al.,

2009) and the clam *Ruditapes decussatus* (Gestal et al., 2007). Evidence for an acute phase response triggered by small temperature changes is limited and it is possible that the heavy chains of the IαI have yet unknown functions in the heat stress response.

3.4.4 Gene regulation and physiological theories

The results presented here could provide further support for the oxygen limitation hypothesis under acute exposure to elevated temperatures. Thermal tolerance has been suggested to be hierarchical with narrowing windows from molecular, to cellular, to systemic levels. Thermal limitation would thus be set at the highest level of complexity, the whole organism, through capacity limitations in ventilation and circulation (Pörtner, 2002). Given this paradigm of a decreasing thermal window at the higher levels of complexity, it is clear that some components contributing to the complex system will display early signs of failure before they become apparent at the organism level. Mitochondria may be particularly relevant in this respect, as they are the final element of the oxygen delivery system, and will reflect early signs of whole organism failure in circulatory and ventilatory systems. The results presented here support this, as the increase in antioxidant defences observed at 3°C, reflects signs of oxidative stress at this temperature. This would suggest that the upper pejus threshold of 4.3°C suggested for *L. elliptica*, based on a significant drop in oxygen supply is an overestimate. Further evidence is provided by the approximately 40% decrease in haemolymph oxygen tensions between 0 and 4°C in *L. elliptica*, which the authors also suggested was evidence for overestimation of Tp in that study (Pörtner et al., 2006). The observation that >75% of individuals failed to bury at 3°C in the conditions used, also provides support for this overestimation. Both reduced oxygen in the haemolymph and failure in essential biological functions indicate some decrement in aerobic scope before 3°C is reached (Peck et al., 2004). Combining this with the data presented here would indicate that the pejus threshold in *L. elliptica* has already been crossed at 3°C.

The concept of pejus thresholds is of considerable interest in that they mark the limits of the conditions for long term survival, and are thus more meaningful in an environmental context than critical temperatures. However, the physiological

mechanisms that have been used to define pejus thresholds provide ambiguous estimations of the temperatures that mark the transition between optimum and pejus, and thus the point at which aerobic capacity starts to decline. Combined analysis of oxygen supply to tissues coupled with measures of whole organism functional performance has been suggested to be most suitable to determine upper pejus thresholds (Pörtner et al., 2006). In *L. elliptica*, the induction of genes involved in the defence against oxidative stress reflect a worsening in oxygen supply and occur at temperatures around which the ability to perform essential biological functions is compromised. Although further analysis is needed at temperatures in the range 0-3°C, it can be proposed that the significantly increased expression of antioxidant genes, namely peroxiredoxin or GST, marks the upper pejus threshold in *L. elliptica* under acute thermal stress and could help define pejus thresholds during rapid warming in other species. Molecular analysis of the mechanisms determining pejus threshold at slower rates of change could help understanding the aspects involved in acclimation and thermal adaptation.

3.5 SUMMARY

The results presented here confirm the occurrence of a temperature triggered response in *L. elliptica* when exposed to elevated temperatures. This study suggests that molecular mechanisms characteristic of the stress response, such as changes in protein turnover through reversible and irreversible protein damage; increased chaperoning activity and oxidative stress are activated in *L. elliptica* when exposed to 3°C. The control of transcription appears critical in surviving exposure to elevated temperatures.

The induction of antioxidant genes at temperatures at which whole organism physiological responses start to differ from those at ambient temperatures, also indicates that molecular responses can be accurate predictors of the different stages of stress. The changes in expression in antioxidant genes reflect the declining capacity in aerobic systems with small elevations in temperature in *L. elliptica* and provide general support for the oxygen and capacity limitation hypothesis in response to acute heat challenges.

Markers for oxidative stress, rather than the universal heat shock proteins, may be better indicators of heat stress in *L. elliptica*. In addition, calcium and calcium signalling related genes and proteins may deserve attention in the future as markers for environmental stress. An understanding of how Ca^{2+} homeostasis and Ca^{2+} signalling pathways relate to stress may result in the discovery of new stress biomarkers.

The chip generated here will be used in future studies of the stress response in *L. elliptica* and related species. The universality of the stress response can be tested by subjecting the animals to different stressors and observing their effect on the genes described as differentially expressed in this study. Additionally, temporal and geographical variations in the stress response of *L. elliptica* can be studied, in order to assess whether different populations of Antarctic clams differ in their ability to respond to stress. This has potential implications for the discovery of universal biomarkers for stress, which may in turn offer clues into responses to climate change.

3.6 REFERENCES

Abele D, Burlando B, Viarengo A, Pörtner HO (1998) Exposure to elevated temperatures and hydrogen peroxide elicits oxidative stress and antioxidant response in the Antarctic intertidal limpet *Nacella concinna*. Comparative Biochemistry and Physiology B-Biochemistry and Molecular 120:425-435

Akberali HB, Marriott KRM, Trueman ER (1977) Calcium utilization during anaerobiosis induced by osmotic shock in a bivalve mollusk. Nature 266:852-853

Ashburner M, Ball CA, Blake JA, Botstein D, Butler H, Cherry JM, Davis AP, Dolinski K, Dwight SS, Eppig JT et al. (2000) Gene ontology: tool for the unification of biology. The Gene Ontology Consortium. Nature Genetics 25:25-9

Auer H, Lyianarachchi S, Newsom D, Klisovic MI, Marcucci U, Kornacker K (2003) Chipping away at the chip bias: RNA degradation in microarray analysis. Nature Genetics 35:292-293

Bagrodia S, Cerione RA (1999) PAK to the future. Trends in Cell Biology 9:350-355

Bairoch A, Apweiler R, Wu CH, Barker WC, Boeckmann B, Ferro S, Gasteiger E, Huang H, Lopez R, Magrane M, et al. (2007) The universal protein resource (UniProt). Nucleic Acids Research 35:D193-D197

Barcia R, Lopez-García JM, Ramos-Martínez JI (1997) The 28S fraction of rRNA - in molluscs displays electrophoretic behaviour different from that of mammal cells. Biochemistry and Molecular Biology International 42:1089-1092

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society series B-methodological 57:289-300

Bootman MD, Collins TJ, Peppiatt CM, Prothero LS, MacKenzie L, De Smet P, Travers M, Tovey SC, Seo JT, Berridge MJ, et al. (2001) Calcium signalling - an overview. Seminars in Cell and Developmental Biology 12:3-10

Brody E, Abelson J (1985) The spliceosome - yeast pre-messenger rna associates with a 40s complex in a splicing-dependent reaction. Science 228:963-967

Buckley BA (2007) Comparative environmental genomics in non-model species: using heterologous hybridization to DNA-based microarrays. Journal of Experimental Biology 210:1602-1606

Buckley BA, Somero GN (2009) cDNA microarray analysis reveals the capacity of the cold-adapted Antarctic fish *Trematomus bernacchii* to alter gene expression in response to heat stress. Polar Biology 32:403-415

Buckley BA, Gracey AY, Somero GN (2006) The cellular response to heat stress in the goby *Gillichthys mirabilis*: a cDNA microarray and protein-level analysis. Journal of Experimental Biology 209:2660-2677

Burkhart BA, Kennett SB, Archer TK (2007) Osmotic stress-dependent repression is mediated by histone H3 phosphorylation and chromatin structure. Journal of Biological Chemistry 282:4400-4407

Bustin M (1999) Regulation of DNA-dependent activities by the functional motifs of the high-mobility-group chromosomal proteins. Molecular and Cellular Biology 19:5237-5246

Byrne RA, McMahon BR (1991) Acid-base and ionic regulation, during and following emersion, in the fresh-water bivalve, *Anodonta-grandis-simpsoniana* (Bivalvia:Unionidae). The Biological Bulletin 181:289-297

Byrne RA, Shipman BN, Smatresk NJ, Dietz TH, McMahon RF (1991) Acid-base-balance during emergence in the fresh-water bivalve *Corbicula fluminea*. Physiological Zoology 64:748-766

Camus L, Gulliksen B, Depledge MH, Jones MB (2005) Polar bivalves are characterized by high antioxidant defences. Polar Research 24:111-118

Castilho PC, Buckley BA, Somero GN, Block BA (2009) Heterologous hybridization to a complementary DNA microarray reveals the effect of thermal acclimation in the endothermic bluefin tuna (*Thunnus orientalis*). Molecular Ecology 18:2092-2102

Causton HC, Ren B, Koh SS, Harbison CT, Kanin E, Jennings EG, Lee TI, True HL, Lander ES, Young RA (2001) Remodeling of yeast genome expression in response to environmental changes. Molecular Biology of the Cell 12:323-337

Ceriotti A, Duranti M, Bollini R (1998) Effects of N-glycosylation on the folding and structure of plant proteins. Journal of Experimental Botany 49:1091-1103

Chan P, Risler JL, Raguenez G, Salier JP (1995) The 3 heavy-chain precursors for the inter-alpha-inhibitor family in mouse - new members of the multicopper oxidase protein group with differential transcription in liver and brain. Biochemical Journal 306:505-512

Chen D, Toone WM, Mata J, Lyne R, Burns G, Kivinen K, Brazma A, Jones N, Bähler J (2003) Global transcriptional responses of fission yeast to environmental stress. Molecular Biology of the Cell 14:214-229

Cheng SC, Tarn WY, Tsao TY, Abelson J (1993) PRP19 - A novel spliceosomal component. Molecular and Cellular Biology 13:1876-1882

Clark MS, Peck LS (2009) HSP70 heat shock proteins and environmental stress in Antarctic marine organisms: A mini-review. Marine Genomics 2:11-18

Clark MS, Fraser KPP, Peck LS (2008) Antarctic marine molluscs do have an HSP70 heat shock response. Cell Stress and Chaperones 13:39-49

Clarke A (2003) Costs and consequences of evolutionary temperature adaptation. Trends in Ecology and Evolution 18:573-581

Copois V, Bibeau F, Bascoul-Mollevi C, Salvetat N, Chalbos P, Bareil C (2007) Impact of RNA degradation on gene expression profiles: Assessment of different methods to reliably determine RNA quality. Journal of biotechnology 127:549-559

Crenshaw M (1972) Inorganic composition of molluscan extrapallial fluid. The Biological Bulletin 143:506-512

Cusack S (1997) Aminoacyl-tRNA synthetases. Current Opinion in Structural Biology 7:881-889

Das P, Gupta A, Manna S (2005) Heat shock protein 70 expression in different tissues of *Cirrhinus mrigala* (Ham.) following heat stress. Aquaculture Research 36:525-529

Davidson J, Schiestl R (2001) Mitochondrial respiratory electron carriers are involved in oxidative stress during heat stress in *Saccharomyces cerevisiae.* Molecular and cellular biology 21:8483-8489

Desalvo MK, Voolstra C, Sunagawa S, Schwarz JA, Stillman JH, Coffroth MA, Szmant AM, Medina M (2008) Differential gene expression during thermal stress and bleaching in the Caribbean coral *Montastraea faveolata.* Molecular Ecology 17:3952-3971

Diarramehrpour M, Bourguignon J, Sesboue R, Mattei M, Passage E, Salier J (1989) Human-plasma inter-alpha-trypsin inhibitor is encoded by 4 genes on 3 chromosomes. European journal of biochemistry 179:147-154

Dietz TJ, Somero GN (1992) The threshold induction temperature of a 90-KDa heat-shock protein is subject to acclimatization in eurythermal goby fishes. Molecular Biology of the Cell 3:A111-A111

Dondero F, Dagnino A, Jonsson H, Capri F, Gastaldi L, Viarengo A (2006) Assessing the occurrence of a stress syndrome in mussels (*Mytilus edulis*) using a combined biomarker/gene expression approach. Aquatic Toxicology 78:S13-S24

Drummond AJ, Ashton B, Cheung M, Heled J, Kearse M, Moir R, Stones-Havas S, Thierer T, Wilson A (2007) Geneious v3.0. http://www.geneious.com/

Ebashi S, Endo M (1968) Calcium ion and muscle contraction. Progress in Biophysics and Molecular Biology 18:123-83

Estevez MS, Abele D, Puntarulo S (2002) Lipid radical generation in polar (*Laternula elliptica*) and temperate (*Mya arenaria*) bivalves. Comparative Biochemistry and Physiology B-Biochemistry and Molecular Biology 132:729-737

Feder ME, Hofmann GE (1999) Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annual Review of Physiology 61:243-282

Fernandez-Capetillo O, Lee A, Nussenzweig M, Nussenzweig A (2004) H2AX: the histone guardian of the genome. DNA Repair 3:959-967

Frova C (2006) Glutathione transferases in the genomics era: New insights and perspectives. Biomolecular Engineering 23:149-169

Gao HC, Wang Y, Liu XD, Yan TF, Wu LY, Alm E, Arkin A, Thompson DK, Zhou JZ (2004) Global transcriptome analysis of the heat shock response of *Shewanella oneidensis*. Journal of Bacteriology 186:7796-7803

Gestal C, Costa MM, Figueras A, Novoa B (2007) Analysis of differentially expressed genes in response to bacterial stimulation in hemocytes of the carpet-shell clam *Ruditapes decussatus*: Identification of new antimicrobial peptides. Gene 406:134-143

Goodson H, Hawse W (2002) Molecular evolution of the actin family. Journal of cell science 115:2619-2622

Hahn JS, Hu ZZ, Thiele DJ, Iyer VR (2004) Genome-wide analysis of the biology of stress responses through heat shock transcription factor. Molecular and Cellular Biology 24:5249-5256

Hartl FU (1996) Molecular chaperones in cellular protein folding. Nature 381:571-580

Hatakeyama S, Nakayama KI (2003) U-box proteins as a new family of ubiquitin ligases. Biochemical and Biophysical Research Communications 302:635-645

Heise K, Puntarulo S, Pörtner HO, Abele D (2003) Production of Reactive Oxygen Species by isolated Mitochondria of the Antarctic Bivalve *Laternula elliptica* (King and Broderip) under Heat Stress. Comparative Biochemistry and Physiology C-Toxicology and Pharmacology 134:79-90

Hofmann B, Hecht HJ, Flohe L (2002) Peroxiredoxins. Biological Chemistry 383:347-364

Hogstrand C, Balesaria S, Glover CN (2002) Application of genomics and proteomics for study of the integrated response to zinc exposure in a non-model fish species, the rainbow trout. Comparative Biochemistry and Physiology B-Biochemistry and Molecular Biology 133:523-535

Holland MJ (2002) Transcript abundance in yeast varies over six orders of magnitude. The Journal of Biological Chemistry 277:14363-6

Imbeaud S, Graudens E, Boulanger V, Barlet X, Zaborski P, Eveno E, Mueller O, Schroeder A, Auffray C (2005) Towards standardization of RNA quality assessment using user-independent classifiers of microcapillary electrophoresis traces. Nucleic Acids Research 33:e56

Ju Z, Dunham RA, Liu Z (2002) Differential gene expression in the brain of channel catfish (*Ictalurus punctatus*) in response to cold acclimation. Molecular Genetics and Genomics 268:87-95

Kang JJ, Watson RM, Fisher ME, Higuchi R, Gelfand DH, Holland MJ (2000) Transcript quantitation in total yeast cellular RNA using kinetic PCR. Nucleic Acids Research 28:e2

Karlen Y, McNair A, Perseguers S, Mazza C, Mermod N (2007) Statistical significance of quantitative PCR. BMC Bioinformatics. 8:131

Karsi A, Patterson A, Feng J, Liu Z (2002) Translational machinery of channel catfish: I. A transcriptomic approach to the analysis of 32 40S ribosomal protein genes and their expression. Gene 291:177-186

Kassahn KS (2008) Microarrays for comparative and ecological genomics: beyond single-species applications of array technologies. Journal of Fish Biology 72:2407-2434

Kassahn KS, Caley MJ, Ward AC, Connolly AR, Stone G, Crozier RH (2007) Heterologous microarray experiments used to identify the early gene response to heat stress in a coral reef fish. Molecular Ecology 16:1749-1763

Kim M, Ahn IY, Cheon J, Park H (2009) Molecular cloning and thermal stress-induced expression of a pi-class glutathione S-transferase (GST) in the Antarctic bivalve *Laternula elliptica*. Comparative Biochemistry and Physiology A-Molecular and Integrative 152:207-213

Kjarland E, Keen TJ, Kleppe R (2006) Does isoform diversity explain functional differences in the 14-3-3 protein family? Current Pharmaceutical Biotechnology 7:217-223

Koide T, Vencio RZN, Gomes SL (2006) Global gene expression analysis of the heat shock response in the phytopathogen *Xylella fastidiosa*. Journal of Bacteriology 188:5821-5830

Krogh, A (1916) The Respiratory Exchange of Animals and Man. Longmans, Green and Co, London p. 173.

Kültz D (2005) Molecular and evolutionary basis of the cellular stress response. Annual Review of Physiology 67:225-257

Kushner I (1982) The phenomenon of the acute phase response. Annals of the New York Academy of Sciences 389:39-48

Lafontaine DLJ, Tollervey D (2001) The function and synthesis of ribosomes. Nature Reviews Molecular Cell Biology 2:514-520

Lu X, Legerski RJ (2007) The Prp19/Pso4 core complex undergoes ubiquitylation and structural alterations in response to DNA damage. Biochemical and Biophysical Research Communications 354:968-974

Maicas E, Pluthero FG, Friesen JD (1988) The Accumulation of 3 Yeast ribosomal-Proteins under Conditions of Excess messenger-RNA is determined primarily by fast Protein Decay. Molecular and Cellular Biology 8:169-175

Marchler-Bauer A et al. (2009) CDD: specific functional annotation with the Conserved Domain Database. Nucleic Acids Research 37:D205-210

Martini EMD, Keeney S, Osley MA (2002) A role for histone H2B during repair of UV-induced DNA damage in *Saccharomyces cerevisiae*. Genetics 160:1375-1387

Michalak M, Parker JMR, Opas MA (2002) Ca^{2+} signalling and calcium binding chaperones of the endoplasmic reticulum. Cell Calcium 32:269-278

Morey JS, Ryan JC, Van Dolah FM (2006) Microarray validation: factors influencing correlation between oligonucleotide microarrays and real-time PCR. Biological Procedures Online:175-193

Morley SA, Peck LS, Miller AJ, Pörtner HO (2007) Hypoxia tolerance associated with activity reduction is a key adaptation for *Laternula elliptica* seasonal energetics. Oecologia 153:29-36

Nomura M (1999) Regulation of Ribosome Biosynthesis in Escherichia coli and *Saccharomyces cerevisiae*: Diversity and Common Principles. Journal of Bacteriology. 181:6857-6864

Nykjaer A, Willnow TE (2002) The low-density lipoprotein receptor gene family: a cellular Swiss army knife? Trends in Cell Biology 12:273-280

Orphanides G, Reinberg D (2002) A unified theory of gene expression. Cell 108:439-451

Orrenius S, Kass GEN, Dypbukt JM, Nicotera P (1992) Calcium-ions and oxidative cell injury. Annals of neurology 32:S33-S42

Osuna-Jimenez I, Williams TD, Prieto-Alamo MJ, Abril N, Chipman JK, Pueyo C (2009) Immune- and stress-related transcriptomic responses of *Solea senegalensis* stimulated with lipopolysaccharide and copper sulphate using heterologous cDNA microarrays. Fish and Shellfish Immunology 26:699-706

Park H, Ahn IY, Lee HE (2007) Expression of heat shock protein 70 in the thermally stressed Antarctic clam *Laternula elliptica*. Cell Stress and Chaperones 12:275-282

Park H, Ahn IY, Kim H, Cheon J, Kim M (2008a) Analysis of ESTs and expression of two peroxiredoxins in the thermally stressed Antarctic bivalve *Laternula elliptica*. Fish and Shellfish Immunology 25:550-559

Park H, Ahn IY, Park KI, Hyun S (2008b) Response of antioxidant defence systems to thermal stress in the Antarctic clam *Laternula elliptica*. Antarctic Science 20:521-526

Parsell DA, Lindquist S (1993) The function of heat-shock proteins in stress tolerance; degradation and reactivation of damaged proteins. Annual Review of Genetics 27:437-496

Peck LS, Pörtner HO, Hardewig I (2002) Metabolic demand, oxygen supply, and critical temperatures in the Antarctic bivalve *Laternula elliptica*. Physiological and Biochemical Zoology 75:123-133

Peck LS, Ansell A, Webb KE, Hepburn L, Burrows M (2004) Movements and burrowing activity in the Antarctic bivalve molluscs Laternula elliptica and *Yoldia eightsi*. Polar Biology 27:357-367

Peck LS, Webb KE, Bailey DM (2004) Extreme sensitivity of biological function to temperature in Antarctic marine species. Functional Ecology 18:625-630

Peck LS, Morley SA, Pörtner HO, Clark MS (2007) Thermal limits of burrowing capacity are linked to oxygen availability and size in the Antarctic clam *Laternula elliptica*. Oecologia 154:479-484

Perez-Torrado R, Yamada D, Defossez P (2006) Born to bind: the BTB protein-protein interaction domain. Bioessays Advances In Molecular Cellular And Developmental Biology 28:1194-1202

Pertea G, Huang X, Liang F, Antonescu V, Sultana R, Karamycheva S, Lee Y, White J, Cheung F, Parvizi B, Tsai J, Quackenbush J (2003) TIGR Gene Indices clustering tools (TGICL): a software system for fast clustering of large EST datasets. Bioinformatics 19:651-2

Petalidis L, Bhattacharyya S, Morris GA, Collins VP, Freeman TC, Lyons PA (2003) Global amplification of mRNA by template-switching PCR: linearity and application to microarray analysis. Nucleic Acids Research 31:e142

Pfaffl MW (2001) A new mathematical model for relative quantification in real-time RT-PCR. Nucleic Acids Research 29:e45

Pfaffl MW, Horgan GW, Dempfle L (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

Philipp E, Brey T, Pörtner HO, Abele D (2005a) Chronological and physiological ageing in a polar and a temperate mud clam. Mechanisms of Ageing and Development 126:598-609

Philipp E, Pörtner H, Abele D (2005b) Mitochondrial ageing of a polar and a temperate mud clam. Mechanisms of Ageing and Development 126:610-619

Place SP, O'Donnell MJ, Hofmann GE (2008) Gene expression in the intertidal mussel *Mytilus californianus*: physiological response to environmental factors on a biogeographic scale. Marine Ecology-Progress Series 356:1-14

Podrabsky JE, Somero GN (2004) Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. Journal of Experimental Biology 207:2237-2254

Pörtner H (2001) Climate change and temperature-dependent biogeography: oxygen limitation of thermal tolerance in animals. Naturwissenschaften 88:137-146

Pörtner H (2002) Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. Comparative Biochemistry and Physiology a-Molecular and Integrative Physiology 132:739-761

Pörtner H, Peck L, Hirse T (2006) Hyperoxia alleviates thermal stress in the Antarctic bivalve, *Laternula elliptica*: evidence for oxygen limited thermal tolerance. Polar Biology 29:688-693

R Development Core Team (2005) R: a language and environment for statistical computing. R Foundation for Statistical Computing. Vienna, Austria. http://www.R-project.org.

Rajaseger G, Lim C, Wui L, Saravanan P, Tang K, Gopalakrishnakone P (2009) A study on the differential protein profiles in liver cells of heat stress rats with and without turpentine treatment. Proteome Science 7:1

Reeves R (1972) Imidazole alphastat hypothesis for vertebrate acid-base regulation - tissue carbon-dioxide content and body-temperature in bullfrogs. Respiration Physiology 14:219

Ritchie ME, Silver J, Oshlack A, Holmes M, Diyagama D, Holloway A, Smyth GK (2007) A comparison of background correction methods for two-colour microarrays. Bioinformatics 23:2700-7

Rozen S, Skaletsky H (2000) Primer3 on the WWW for general users and for biologist programmers. Methods in Molecular Biology 132:365-386

Rushmore TH, Morton M, Pickett CB (1991) The antioxidant responsive element - activation by oxidative stress and identification of the DNA consensus sequence required for functional activity. Journal of Biological Chemistry 266:11632-11639

Rushmore TH, Pickett CB (1993) Glutathione S-transferases, structure, regulation, and therapeutic implications. Journal of Biological Chemistry 268:11475-11478

Salier J, Rouet P, Raguenez G, Daveau M (1996) The inter-alpha-inhibitor family: From structure to regulation. The Biochemical journal 315:1-9

Sambrook J, Russell DW (2001) Molecular Cloning: A laboratory manual, 3rd ed, Cold Spring Harbor Laboratory Press, New York.

Sarmiento J, Leal S, Quezada C, Kausel G, Figueroa J, Vera M, Krauskopf M (2000) Environmental acclimatization of the carp modulates the transcription of beta-actin. Journal of Cellular Biochemistry 80:223-228

Saveanu C, Bienvenu D, Namane A, Gleizes PE, Gas N, Jacquier A, Fromont-Racine M (2001) Nog2p, a putative GTPase associated with pre-60S subunits and required for late 60S maturation steps. EMBO Journal 20:6475-6484

Saveanu C, Namane A, Gleizes PE, Lebreton A, Rousselle JC, Noaillac-Depeyre J, Gas N, Jacquier A, Fromont-Racine M (2003) Sequential protein association with nascent 60S ribosomal particles. Molecular and Cellular Biology 23:4449-4460

Sharp VA, Brown BE, Miller D (1997) Heat shock protein (hsp 70) expression in the tropical reef coral *Goniopora djiboutiensis*. Journal of Thermal Biology 22:11-19

Siegert R, Leroux MR, Scheufler C, Hartl FU, Moarefi I (2000) Structure of the molecular chaperone prefoldin: Unique interaction of multiple coiled coil tentacles with unfolded proteins. CELL 103:621-632

Skern R, Frost P, Nilsen F (2005) Relative transcript quantification by Quantitative PCR: Roughly right or precisely wrong? BMC Molecular Biology 6:10

Skorzewski R, Sliwinska M, Borys D, Sobieszek A, Moraczewska J (2009) Effect of actin C-terminal modification on tropomyosin isoforms binding and thin filament regulation. Biochimica et Biophysica Acta-Proteins and Proteomics 1794:237-243

Smerdon GR, Chapple JP, Hawkins AJS (1995) The simultaneous immunological detection of 4 Stress-70 protein isoforms in *Mytilus edulis*. Marine Environmental Research 40:399-407

Smyth GK (2004) Linear models and empirical bayes methods for assessing differential expression in microarray experiments. Statistical Applications in Genetics and Molecular Biology 3:Article3

Smyth GK (2005) Limma: linear models for microarray data. In: Gentleman R, Carey V, Dudoit S, Irizarry R, Huber W (Eds.), Bioinformatics and Computational Biology Solutions Using R and Bioconductor. Springer, New York, pp. 397–420

Smyth GK, Speed T (2003) Normalization of cDNA microarray data. Methods 31:265-73

Smyth GK, Michaud J, Scott HS (2005) Use of within-array replicate spots for assessing differential expression in microarray experiments. Bioinformatics 21:2067-2075

Soller M (2006) Pre-messenger RNA processing and its regulation: a genomic perspective. Cellular and molecular life sciences 63:796-819

Somero GN (2004) Adaptation of enzymes to temperature: searching for basic "strategies". Comparative Biochemistry and Physiology B-Biochemistry and Molecular Biology 139:321-333

Sommer A, Klein B, Pörtner HO (1997) Temperature induced anaerobiosis in two populations of the polychaete worm *Arenicola marina* (L). Journal of Comparative Physiology B-Biochemical Systemic and Environmental Physiology 167:25-35

Sørensen JG, Nielsen MM, Kruhoffer M, Justesen J, Loeschcke V (2005) Full genome gene expression analysis of the heat stress response, in *Drosophila melanogaster*. Cell Stress and Chaperones 10:312-328

Storey KB, Storey JM (2004) Metabolic rate depression in animals: transcriptional and translational controls. Biological Reviews 79: 207-233

Sturzenbaum S, Kille P (2001) Control genes in quantitative molecular biological techniques: the variability of invariance. Comparative biochemistry and physiology. Part B Biochemistry and molecular biology 130: 281-28

Takahashi Y (2003) The 14-3-3 proteins: Gene, gene expression, and function. Neurochemical Research 28:1265-1273

Terahara K, Takahashi KG (2008) Mechanisms and immunological roles of apoptosis in molluscs. Current Pharmaceutical Design 14:131-137

Teranishi KS, Stillman JH (2007) A cDNA microarray analysis of the response to heat stress in hepatopancreas tissue of the porcelain crab *Petrolisthes cinctipes*. Comparative Biochemistry and Physiology D-Genomics and Proteomics 2:53-62

Thomas J (2001) HMG 1 and 2: architectural DNA-binding proteins. Biochemical Society Transactions 29:395-401

Toescu E (2004) Hypoxia sensing and pathways of cytosolic Ca^{2+} increases. Cell Calcium 36:187-199

Tomanek L, Somero G (2000) Time course and magnitude of synthesis of heat-shock proteins in congeneric marine snails (Genus *Tegula*) from different tidal heights. Physiological and Biochemical Zoology 73:249-256

Tsay YF, Thompson JR, Rotenberg MO, Larkin JC, Woolford JL (1988) Ribosomal-protein synthesis is not regulated at the translational level in *Saccharomyces cerevisiae* - balanced accumulation of ribosomal protein-l16 and protein-rp59 is mediated by turnover of excess protein. Genes and Development 2:664-676

Vainberg IE, Lewis SA, Rommelaere H, Ampe C, V J, ekerckhove, Klein HL, Cowan NJ (1998) Prefoldin, a Chaperone that Delivers Unfolded Proteins to Cytosolic Chaperonin. Cell 93:863-873

Van Den Heuvel S (2004) Protein degradation: CUL-3 and BTB - Partners in proteolysis. Current Biology 14:R59-R61

Verlecar XN, Jena KB, Chainy GBN (2007) Biochemical markers of oxidative stress in *Perna viridis* exposed to mercury and temperature. Chemico Biological Interactions 167:219-226

Wang LL, Song LS, Zhao JM, Qiu LM, Zhang H, Xu W, Li HL, Li CH, Wu LT, Guo XM (2009) Expressed sequence tags from the zhikong scallop (*Chlamys farreri*): Discovery and annotation of host-defense genes. Fish and Shellfish Immunology 26:744-750

Winston GW, Digiulio R (1991) Prooxidant and antioxidant mechanisms in aquatic organisms. Aquatic Toxicology 19:137-161

Wood ZA, Schroder E, Harris JR, Poole LB (2003) Structure, mechanism and regulation of peroxiredoxins. Trends in Biochemical Sciences 28:32-40

Yan AY, Lennarz WJ (2005) Unraveling the mechanism of protein N-glycosylation. Journal of Biological Chemistry 280:3121-3124

Zhuo L, Hascall V, Kimata K (2004) Inter-alpha-trypsin inhibitor, a covalent protein-glycosaminoglycan-protein complex. Journal of Biological Chemistry 279:38079-38082

4 Analysis of Changes in Protein Expression following Exposure to Thermal Stress

4.1 INTRODUCTION

4.1.1 Proteomics in non model marine species

Proteomics allows high-throughput analysis for the study of protein mixtures in complex systems. The two main proteomic tools are two dimensional electrophoresis (2DE) and mass spectrometry (MS), which have been extensively used in model species. Studies are particularly abundant in the fields of drug discovery, biomarker identification in disease, and protein-protein interaction studies in human disease processes (Hanash, 2003). However, studies using discovery driven proteomics to analyse global patterns in protein expression in non model species are less abundant. The effects of environmental stressors such as salinity and temperature variations on protein expression have been analyzed in different tissues of aquatic organisms (Tunsjo et al., 2007; Gardeström et al., 2007; Lee et al., 2006). 2DE has also been used in studies on marine and estuarine pollution, particularly using bivalves due to their capacity to concentrate and accumulate pollutants. Several studies have been carried out to study the effects of marine pollutants in protein expression profiles in bivalves (Rodriguez-Ortega et al., 2003; Manduzio et al., 2005; Apraiz et al., 2006; Jonsson et al., 2006). These studies provide information on the processes governing the mechanisms of the stress response, and often have as their ultimate aim the identification of proteins induced or repressed upon exposure to a stressor, which can become candidates for stress biomarkers. However, conclusions are often limited by the number of available protein identifications as a result of the poor characterization of non model species in the databases (Nunn and Timperman, 2007). The limited representation of *M. edulis* and *L. elliptica* on public databases has been discussed in previous chapters

The use of proteomics in the marine environment has not been limited to the study of stress comparing two experimental conditions with the aim of identifying biomarkers. Indeed, studies on protein expression profiles in non model marine species do not always rely on protein identifications. For example, within the genus *Mytilus*, protein expression patterns have been compared between *M. edulis* and *M. galloprovincialis* to study the molecular differentiation between species of the same

taxa. Protein expression profiles were compared between *M. edulis* from the Netherlands and *M. galloprovincialis* from the Iberian Peninsula to generate reference protein maps for foot tissue for both species (Lopez et al., 2002). Similarly, analysis of 2DE patterns was applied to elucidate differences between *M. edulis*, *M. galloprovincialis* and a hybrid of the two species. Underlying variation in protein expression was assessed and compared between species (Diz and Skibinski, 2007). Comparisons between the proteome of wild and cultured mussels have been made, and the results linked to the different ecological conditions of their habitats (Lopez, 2007). These examples illustrate that, despite the limited success of protein identifications in non model species when compared to model organisms, proteomics is a valuable tool to analyse differences between species and within species adapted to different environments.

4.1.2 Selection and the environment

Thermal tolerance windows are narrower in tropical than in temperate ectothermic species, a phenomenon described under the climate variability hypothesis (Stevens, 1989). In this respect, the Antarctic marine environment is equivalent to the tropics, as both represent environments where the magnitude of temperature variations is largely reduced compared to that observed in temperate habitats. In a constant, highly stable environment, stabilizing selection acts reducing variation favoring a single optimum phenotype (Lande and Shannon, 1996). If environmental conditions do not change, evolutionary theory predicts the loss of plasticity and evolution of a canalized phenotype (Pigliucci et al., 2006).

However, we are currently undergoing a period of global warming (Solomon, 2007), which is affecting the stability of even the most constant environments. To predict the effects of climate change on biodiversity, it is essential to determine whether species from different habitats have different susceptibility to warming temperatures. When individuals are exposed to changing conditions, they respond through physiological flexibility (Gienapp et al., 2008). Such flexibility is a component of phenotypic plasticity, a property of individual genotypes, to produce different phenotypes under different environmental conditions. Phenotypic plasticity can be

adaptive, when it places populations closer to a new phenotypic optimum, or non adaptive, when it ultimately causes the mean phenotype to be further from the optimum (Ghalambor et al., 2007). Physiological flexibility constitutes an important mechanism to cope with environmental changes, but there are limits to plastic responses thus these are unlikely to provide long term solutions for populations under continued directional environmental change (Gienapp et al., 2008). However, if there is enough genetic variation underlying the plasticity, this can be selected directionally, resulting in genetic assimilation of adaptive traits, a process by which environmentally induced plasticity becomes constitutively produced (Pigliucci et al., 2006). This can promote establishment of the species in the new environment enabling persistence.

In the present study, a measure of biological variation is provided for a eurythermal and a stenothermal ectotherm upon exposure to heat stress. A comparison is made between species and the potential relevance of the observed changes in variation in both species is discussed in relation to their respective environments.

4.1.3 Aims

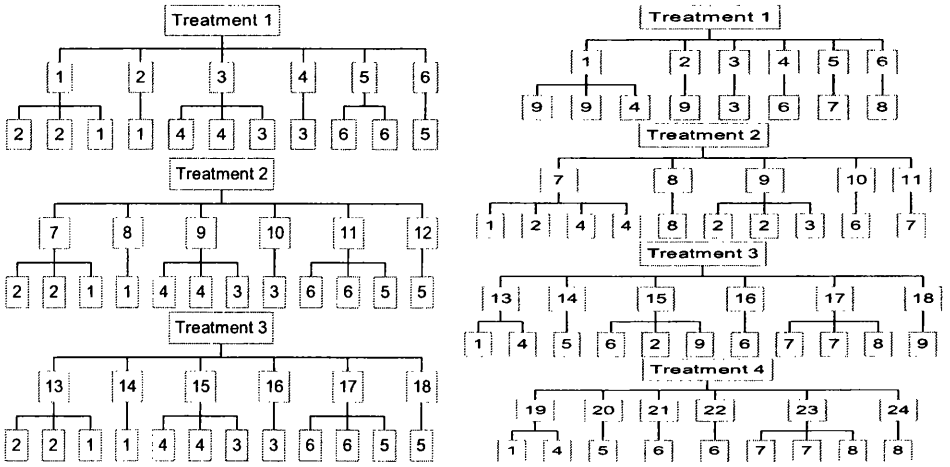
In this study, two dimensional gel electrophoresis was used to investigate differences in the stress proteome of two marine ectotherms, the eurythermal *Mytilus edulis* and the highly stenothermal *Laternula elliptica*. The aims of this analysis were to i) analyze changes in protein expression patterns in response to different intensities of thermal stress in both species, ii) determine whether potential differences in protein expression patterns between species may be related to their respective habitats, iii) identify proteins associated with the stress response, which can be proposed as stress biomarkers in both species. This study is the first analysis of global protein expression patterns in an Antarctic marine species, and provides an interesting insight into the differences between the effects of thermal stress in temperate and Antarctic species at the protein level.

4.2 MATERIALS AND METHODS

4.2.1 Experimental design

In order to minimise bias resulting from technical differences in the processes, randomisation was performed at several levels: random selection of individuals and treatments in each extraction batch, allocation of samples to strips and strips to gels, strip and gel placement in the first and second dimension respectively, and assignation of gels to trays during silver staining.

The characteristics of the equipment determined that a maximum of six gels could be run at the same time. Biological replicates are independent biological samples whereas technical replicates are repeated measurements on one biological sample. Six independent biological replicates were used per each of the three and four treatments used in *M. edulis* and *L. elliptica* respectively (with the exception of *L. elliptica* 3°C buried where only five individuals were available). For *M. edulis*, each run of six gels contained either two individuals from each treatment or one individual from each treatment with two technical replicates of each of the three individuals. There were six runs of six gels each. For *L. elliptica*, each run of six gels was designed to contain one individual from each of the four treatments, with two of these replicated on two separate gels. Modifications were made to the initial experimental design for *L. elliptica*, resulting from problems when running the gels and poor gel quality. This resulted in an unbalanced dataset. For *M. edulis*, a between day technical replicate for a single individual was excluded from analysis due to poor quality. Final experimental designs are shown in figure 4.1. This design allows quantification of technical and biological errors, improving precision in the analysis, and minimises bias introduced by heterogeneity of the apparatus and variability between days.



a) b)

Figure 4.1 Experimental design for a) *M. edulis* and b) *L. elliptica* 2DE. Within and between days replicates are included. Treatment is indicated in the first row of each chart, followed by animal numbers in each box on the second row and day when they were run on the third row. For *L. elliptica* A1 to A6 (labelled 1-6 in figure for brevity) are the six control animals, A2 to A11 moderate heat shock (reburied), A13 to A18 moderate heat shock (not buried) and A19 to A24 extreme heat shock. A12 is missing as only five animals were used in the treatment (only five reburied). For *M. edulis* A1 to A6, A7 to A12, and A13 to A18 correspond to control, moderate and extreme heat shocks respectively.

4.2.2 Preliminary tests of some experimental methods

A series of tests were carried out to optimise the protein extraction method in *L. elliptica*. For *M. edulis*, protein extractions had previously been optimised in our laboratory at Swansea University and the protocols detailed here had been described in Diz and Skibinski (2007). The possibility of using different *L. elliptica* tissues was also considered. At the time the experiments were carried out (2006-07), no recommendations were found in the literature regarding the preparation of protein extracts for 2DE from *L. elliptica* or, to my knowledge, other Antarctic species. All tests were carried out using lab-cast gels, which were prepared as described below.

4.2.2.1 Preparation of lab-cast gels

A 12.5% acrylamide solution was prepared for six 1.5 mm gels consisting of 250 ml bisacrylamide stock 30% (Severn Biotech), 150 ml 1.5 M Tris-HCl (pH 8.8), 187 ml milliQ H_2O (Millipore) and 6 ml 10% SDS combined in a Buchner flask. The solution was vacuum degassed for 5-10 min before the addition of 830 µl 10% TEMED (Sigma Aldrich) and 6 ml 10% APS (Sigma Aldrich). Immediately after, the previously assembled gel casters (Ettan Dalt six gel system, GE Healthcare) were filled by pouring the solution into the filling channel and allowing it to fill all casters until the final desired height was reached (approximately 1 cm below the top edge of the short plate). Water saturated butanol (5:1, Sigma Aldrich) was then pipetted onto each of the gels, which were allowed to polymerize at room temperature for 2 h. After polymerization, the butanol was removed and the surface of the gels cleaned with water to prevent gel damage. The gels were covered with paraffin tape and left at 4°C overnight. Running conditions were as described below for precast gels. Cathode (2x SDS) and anode (1x SDS) running buffers for lab-cast gels were diluted from a stock of 10x SDS electrophoresis buffer (250 mM Tris, 1.92 M glycine and 1% (w/v) SDS)

4.2.2.2 Preparation of protein extracts

Preliminary experiments were carried out to select the optimal composition of the lysis buffer. Two different buffers were made in which the proportion of CHAPS (32.4 mM (2%) and 64.8 mM (4%), Sigma-Aldrich), dithiothreitol (DTT, 64.8 mM (1%) and 194.4 mM (3%), GE Healthcare) and immobilized pH gradient (IPG) ampholytes (1 and 2% (v/v), GE Healthcare) was varied based on recommendations on (Khoudoli et al., 2004). One lysis buffer contained the highest and the other the lowest concentrations of all described reagents. A single sample was processed using the two lysis buffers and the extracts were analyzed on lab-cast 2D gels. The buffer that resulted in the best pattern (i.e. the highest number of spots with the lowest distortion) was used throughout the experiment.

4.2.2.3 Tissue selection

As explained in chapter two, the tissue used for each species was chosen according to previous recommendations. The foot was selected in *M. edulis* as it has been determined that it is the least likely to be subjected to contaminants or variation arising from changes in the seasonal reproductive cycle. For *L.elliptica*, protein extracts were prepared from several tissues to compare the quality in the 2DE patterns. Protein mixtures were prepared from 40 mg siphon, gill, digestive gland and mantle, and separated on lab-cast gels using 2DE following the protocol detailed below.

4.2.3 Preparation of protein mixtures

4.2.3.1 Protein extraction

Proteins were extracted from approximately 30 mg of tissue, crushed in liquid nitrogen with a pestle and mortar. The resulting homogenate was immediately suspended in 700 µl lysis buffer containing 7M urea, 2M thiourea (Sigma-Aldrich), 64.8 mM CHAPS, 194.4 mM DTT and 2% (v/v) IPG ampholytes. Proteins were solubilised with a sonicator (Branson digital sonicator 250) using 10 blasts of 15% amplitude and 5 sec each, with 10 sec pauses. All steps were carried out on ice to minimize protein damage and degradation. After centrifugation at 4°C for 30 min (15000*g*), the pellet was discarded and supernatant containing the protein was stored at -80°C in 150 µl aliquots until further use.

4.2.3.2 Protein quantification

Protein concentration was measured using the Protein 2-D Quant Kit (GE Healthcare) according to manufacturer instructions. Briefly, a standard curve was first prepared using a 2 mg ml^{-1} bovine serum albumin (BSA) standard solution by setting up a dilution series consisting of 0, 10, 20, 30, 40, and 50 µg protein, with the 0 µg (milliQ H_2O, Millipore) tube serving as the assay blank. Both the standard curve tubes and the sample tubes were set up in duplicate. For quantification, 500 µl

precipitant solution was added to 15 µl of each sample and to the standard curve tubes, followed by vortexing and incubating at room temperature for 3 min. An equal volume of co-precipitant solution was then added, followed by brief mixing by inversion. Samples were subsequently centrifuged at 10,000*g* for 5 min to precipitate the protein, after which the supernatant was discarded. Pellets were resuspended in 100 µl copper solution and 400 µl milliQ H_2O, to which 1ml working color solution (100 parts color solution A: 1 part color solution B) was added, followed by 20 min incubation at room temperature. The absorbance of each sample was then read using a spectrophotometer at 480 nm (A_{480}) using milliQ H_2O as a blank. Average readings were calculated for each tube and a standard curve generated from which the concentration of each protein sample was determined. Protein concentrations typically ranged between 1.5 and 2.5 µg μl^{-1}.

4.2.3.3 Protein clean-up

After quantification, 90 µg of total protein from each individual was cleaned to remove interfering substances using a 2-D Clean-Up Kit (GE Healthcare) according to manufacturer instructions. Briefly, 300 µl precipitant was added to the appropriate volume of the protein sample. The mixture was incubated on ice for 15 min prior to the addition of 300 µl co-precipitant. Samples were then vortexed and centrifuged at 4°C and 13000*g* for 5 min, after which a small pellet was visible. The supernatant was discarded and the pellet was washed with 40 µl co-precipitant for 5 min, during which period samples were kept on ice. After discarding the co-precipitant, the pellet was dispersed in 25 µl milliQ H_2O by vortexing and washed with 1ml wash buffer pre chilled at -20°C and 5 µl wash additive. Samples were incubated at -20°C for at least 2 h, vortexing every 30 min for 20 sec. At this stage, the pellet is dislodged but does not dissolve. Finally, the supernatant was discarded and the resulting pellet containing the protein sample was dissolved in 450 µl Destreak Rehydration solution (GE Healthcare) and 4.5 µl IPG ampholytes. After a final centrifugation at 4°C and 13000*g* for 5 min to precipitate any left over contaminants, samples were pipetted in the reswelling tray, over which the gel strips were placed facing downwards and covered with Immobiline DryStrip cover fluid (GE

Healthcare). Gel strips were allowed to rehydrate passively at room temperature overnight.

4.2.4 2-DE electrophoresis

After rehydration, strips were transferred to 24 cm Ettan IPGphor ceramic strip holders (GE Healthcare), to which two electrodes were fitted. Strips were covered with Immobiline DryStrip cover fluid. The first dimension (IEF) electrophoresis was carried out on non linear IPG Immobiline DryStrip gels (24 cm, pH 3-10NL, GE Healthcare) with a horizontal electrophoresis apparatus (Ettan IPGphor, GE Healthcare). Voltage was increased linearly in 1 h to 500 V, followed by a 1 h linear increase to 1000 V and finalized at 8000 V for 62500 Vh. Following IEF, strips were transferred to plastic tubes where they were equilibrated in two steps in 10 ml equilibration buffer (6M urea, 2% (w/v) SDS, 75 mM Tris-HCl (pH 8.8), 30% glycerol and 0.001% (w/v) bromophenol blue). In the first step, the strips were washed in equilibration buffer containing 100 mg DTT for 15 min with light shaking. In the second step, the first equilibration buffer was decanted and the wash procedure repeated with an extra 10 ml equilibration buffer with 250 mg iodoacetamide (IAA, Sigma-Aldrich) replacing the DTT. Following equilibration, the strips were washed in cathode running buffer (DALT buffer kit, GE healthcare) and placed along the top of the acrylamide gels previously assembled in the DALT gel cassettes (GE Healthcare) according to manufacturer instructions. The second dimension of gel electrophoresis was carried out on 12.5% DALT precast polyacrylamide gels (260x200x1 mm, GE Healthcare) using an Ettan Dalt six Electrophoresis System (GE Healthcare). Gels were run for 30 min at 2.5 W per gel, followed by 4 to 5 h at 17 W per gel up to a maximum of 100 W. The temperature within the chamber was maintained at 25°C using a MultiTemp III thermocirculator set to 10°C according to manufacturer instructions (GE healthcare). The DALT buffer kit provided with the precast gels was used throughout the experiment. Electrophoresis was stopped when the dye front reached the bottom of the gel. Immediately after electrophoresis, the gels were fixed in fixing solution (40% ethanol, 10% acetic acid) overnight.

4.2.5 Silver staining

Silver staining of the gels was used to visualize protein spots using a modification of the protocol developed by Heukeshoven and Dernick (1985). Briefly, after incubation in fixing solution overnight, the gels were incubated for 30 min in sensitization solution (30% ethanol, 500 mM sodium acetate trihydrate, 15.65 mM sodium thiosulfate, 0.5% glutaraldehyde), followed by three washes of 15 min each in milliQ H_2O, incubation for 40 min in staining solution (5.9 mM silver nitrate, 0.02% formalin) and incubation for 15 min in developing solution (230 mM sodium carbonate, 0.01% formalin) until the desired staining intensity was reached. Gels were then incubated twice for 10 min in stop solution (50 mM EDTA) and washed three times for 5 min in milliQ water before scanning. The sensitization and staining steps interspersed washes were carried out at 4°C. All reagents were provided by Sigma-Aldrich (SigmaUltra range). For short term storage, gels were kept at 4°C in 1% acetic acid.

4.2.6 Computer analysis of 2-DE patterns

Silver-stained gels were scanned to TIFF files using an Image Scanner (Amersham Pharmacia Biotech). Progenesis Samespots version 3.0 software (Nonlinear Dynamics Ltd) was used for alignment, spot detection and filtering, matching between gels and spot volume measurement. Manual intervention was essential to carry out the alignment. After automatic spot detection and matching, a visual spot by spot filtering was performed, and spots which did not meet the pre-established criteria were excluded from the analysis. The criteria used for spot selection were applied consistently throughout the gels:

- Spots in highly crowded areas, at the borders of the gels or areas showing streaking were excluded from analysis.
- Overstained spots were excluded from further analysis as accurate quantification would not be possible.
- Artifacts having no defined outline or which were difficult to discern from the background were also eliminated.

- In each treatment, only spots that were present on >50% individuals were selected. Where technical replicates were available, only those spots present in all replicates were included.

The number of detected spots was dramatically reduced by this spot filtering from over one thousand to 378 and 264 for *M. edulis* and *L. elliptica* respectively. After spot detection and filtering, the lowest on boundary method for background subtraction was applied, as recommended by Nonlinear Dynamics Ltd. The background level for a spot is equal to the lowest intensity value of the image pixels outside the outline of the spot. This background value was subtracted from the spot volume, the summed intensity values of every pixel inside the spot outline.

4.2.7 Normalization and transformation

Absolute spot volumes from 2DE gels are likely to be affected by variations in experimental conditions intrinsic to the use of the technique. Differences in amount of sample loading, gel staining process or batch variation must be accounted for (Chang et al., 2004), thus a normalization process is needed to standardize the distribution of intensities across gels. In the current study, a modification of a normalization method previously described for one channel microarray data was used (Bolstad et al., 2003). The gel with the lowest median spot volume was selected as a baseline. Individual spot volumes were normalized according to the following equations:

$$\beta_i = \frac{(\sum x_{ibase} / n)}{(\sum x_i / n)}$$

$$x_i' = \beta_i x_i$$

Where x_{ibase} are the spot volumes in the baseline gel, x_i represents spot volumes of any gel to be normalised, n is the total number of spots in the gel retained after filtering, and x_i' represents normalised spot volumes. Normalised spot volumes were $\log_2$ transformed for further analysis.

4.2.8 Statistical analysis

4.2.8.1 Normality and variance homogeneity

Statistical analysis was carried out using R version 2.8.1 (R Development Core Team, 2008) and MINITAB 15.0 for Windows. Normalized volumes were used throughout the analysis. A one way ANOVA with spot volume as dependent variable and treatment as factor or independent variable was carried out. The distribution of the residuals was examined visually, by checking the fit of the data to a straight line in a Q-Q plot, as well as statistically with the Anderson-Darling test. Variance homogeneity of the residuals was also checked visually using boxplots and statistically using Levene's test.

4.2.8.2 Technical error

The choice of statistical design was determined by the properties of the dataset. Technical replicates, repeated measurements of the same individual, allow calculation of the technical error and thus provide some estimation of the precision of the results. In the present study, different number of technical replicates have been used for each biological replicate, thus the experimental design is not balanced. There are several proposed methods to deal with this problem, which include eliminating the replicates, averaging them for each individual and carrying out statistical analysis on the means or using a statistical analysis method which can deal with unequal number of technical replicates for each biological replicate, such as Maximum Likelihood (ML) (Crawley, 2007). The possibility of averaging the technical replicates was explored by comparing the between and within day technical error of the replicates for each individual that had been run on several gels. For each animal, two factorial two-way ANOVAs were carried out of $\log_2$ normalized spot volume with *Spot* and *Gel* as random factors. *Spot* represented all selected spots within a gel and *Gel* represented the two technical replicates for a single individual. The interaction *Spot*Gel* thus measures the residual technical error. Separate ANOVAs were carried out to extract the within day and between day technical errors (i.e. the same animal run on different gels either twice on the same day or on

different days). The ratio of the mean square of each pair of measures of the technical error (within and between days) and associated probabilities were calculated. P values <0.05 indicated significant differences between the technical variation measured between and within days. If this occurs it would not be legitimate to average technical replicates.

4.2.8.3 Spot by spot analysis using mixed effects models

The use of Maximum Likelihood was the preferred choice as it allowed analysis of the unbalanced dataset, where all replicates could be included. The function lme in R fits linear mixed effects models with specified mixtures of fixed and random effects. Using lme, the dataset for each spot (i) could be modeled as:

```
modelspoti = lme(vol ~ treatment, random = list(~ +1|animal,
~+1|day),data = data[spot.number == i, ])
```

where `modelspoti` is the name given to that particular analysis, and *vol* is the response or dependent variable, or normalized spot volume for that particular spot. The variable *treatment* represents the fixed factor and has three and four levels for *M. edulis* and *L. elliptica* respectively. The variables *animal* and *day* are the random factors. *Animal* has 18 levels (3 treatments x 6 animals) for *M. edulis* and 23 ((3x6)+(1x5)) for *L. elliptica*. *Day* has six and nine levels for *M. edulis* and *L. elliptica* respectively. In simple terms, the model above can be read as "spot volume for spot *i* is modelled as a function of the treatment, given day and animal as random factors". The default estimation method is Restricted Maximum Likelihood (REML). ML tends to underestimate variance, thus the REML method is preferred by many analysts (Pinheiro and Bates, 2000). In addition, REML estimators are less sensitive to outliers than ML estimators (Crawley, 2007).

A spot by spot analysis was carried out to determine which spots differed significantly in their level of expression between at least two of the treatments. The null hypothesis that for each individual spot there are no treatment differences was tested by constructing two linear mixed models with *day* and *animal* as random

factors, thus allowing the components of variation to be estimated. For the first model, *treatment* was not included as a factor, and for the second model *treatment* was specified as a fixed effect and every other variable remained unchanged.

```
model1 = lme(vol ~ +1, random = list(~ +1|animal, ~+1|day), method =
"ML", data = data[spot.number == i, ])

model2 = lme(vol ~ treatment, random = list(~ +1|animal, ~+1|day),
method = "ML", data = data[spot.number == i, ])
```

A likelihood ratio test was then used to compare the fit of the two models represented by *model1* and mode*l2*. The *anova* function in R provides that capability. Because the default estimation method in lme is REML and likelihood comparisons between REML fits with different fixed effects structures are not meaningful, the two models are fitted using maximum likelihood (ML), before calling *anova* (Pinheiro and Bates, 2000). The results also produce the Akaike information criterion (AIC) for each model, the log likelihood ratio between the two models and associated probability. Since the model with the most parameters will always fit at least as well as the simplest model, the probability distribution of the test statistic is used to determine whether the more complex model fits the data significantly better. The likelihood ratio test would thus reject the null hypothesis of no treatment differences if $P<0.05$ and accept it when $P>0.05$. The AIC (Akaike, 1974) balances the number of parameters in the model, against the maximised likelihood for the model according to the following formulation:

$$AIC = 2k - 2 \ln (ML)$$

where k is the number of independently adjusted parameters in the model and ML is the maximum likelihood value. It provides a mathematical formulation of the principle of parsimony in the field of model construction and can thus be used to determine the simplest model that best explains the data. Given a dataset and several competing models to explain it, the model with the lowest AIC can be considered the best.

4.2.8.4 Variance components

Variance components were calculated on a spot by spot basis according to *modelv*:

```
modelv = lme(vol ~ treatment, random = list(~ +1|treatment, ~
+1|animal, ~+1|day),data = data[spot.number == i, ])
```

In this case, all factors are specified as random to obtain variance components for all variables. *Day* is nested within *animal* which is itself nested within *treatment.* Components were averaged across spots to obtain global values.

Comparisons of variance values across spots for *animal* were made between the temperature treated groups of the two species using a Mann-Whitney test. In order to account for potential artefacts introduced by the different sample sizes between the two species (i.e. dividing the total variance by a different number of spots in each case), the number of spots was equalized by deleting 114 randomly selected spots from the *M. edulis* raw data. Data were normalised and variance components calculated as previously described. In addition, in order to compare biological variance between species in the natural conditions of their respective environments, biological variance values were calculated for control animals. For this purpose, the number of spots was equalised as before and comparisons of *animal* variance values for the control group only were made between species via a Mann-Whitney test.

4.2.9 Mass Spectrometry

Gel preparation and mass spectrometric analysis was carried out at the mass spectrometry facility at the Cambridge Center for Proteomics at Cambridge University. Gels were prepared as previously described and stained with a modified protocol (Carpentier et al., 2005), compatible with mass spectrometry. Briefly, the gels were fixed overnight in a fixing solution (40% ethanol, 10% acetic acid), washed for 20 min in 30% ethanol, 20 min in 20% ethanol, and 20 min in Milli-Q water. After washing the gels were sensitized in 1.26 mM sodium thiosulfate for 1 min and incubated in silver stain (11.77 mM silver nitrate, 0.02% formamide) for 20

min. The gels were developed in developing solution (283 mM sodium carbonate, 31.62 µM sodium thiosulfate, 0.05% formamide) for approximately 5 min until the desired intensity was reached and washed in milliQ water.

Protein spots of interest were manually excised from the gel using a scalpel and transferred to a 96-well PCR plate. All subsequent sample preparation was performed in a MassPrep Station liquid handling robot (Micromass). Briefly, the gel fragments were destained, reduced with DTT, alkylated with iodoacetamide and subjected to enzymatic digestion with trypsin overnight at 37°C. After digestion, 10 µl of supernatant was pipetted into a sample vial and analyzed by LC-MS/MS.

All LC-MS/MS experiments were performed using an Eksigent NanoLC-1D Plus (Eksigent Technologies, Dublin) HPLC system and an LTQ Orbitrap mass spectrometer (ThermoFisher). Separation of peptides was performed by reverse-phase chromatography using at a flow rate of 300 nl min^{-1} and an LC-Packings (Dionex, Sunnyvale) PepMap 100 column (C18, 75 µM i.d. x 150 mm, 3 µM particle size). Peptides were loaded onto a precolumn (Dionex Acclaim PepMap 100 C18, 5 µM particle size, 100 Å, 300 µM i.d x 5 mm) from the autosampler with 0.1% formic acid (v/v) for 5 min at a flow rate of 10 µl min^{-1}. After this period, the ten port valve was switched to allow elution of peptides from the precolumn onto the analytical column. Solvent A consisted of water and 0.1% formic acid (v/v), and solvent B contained 50% acetonitrile (v/v) and 0.1% formic acid (v/v). The gradient employed was 5-50% B in 50 min. The LC eluant was sprayed into the mass spectrometer by means of a New Objective nanospray source. All m/z values of eluting ions were measured in the Orbitrap mass analyzer, set at a resolution of 7500. Peptide ions with charge states of 2^+ and 3^+ were then isolated and fragmented in the LTQ linear ion trap by collision-induced dissociation and MS/MS spectra were acquired.

Post-run, the data was processed using Bioworks Browser (version 3.3.1 SP1, ThermoFisher). Briefly, all MS/MS data were converted from raw output to dta (text) files using the Sequest Batch Search tool within Bioworks. The dta files for each spot were converted to a single Mascot generic file (mgf) file using a SSH script

in the SSH Secure Shell Client program (Version 3.2.9 Build 283, SSH Communications Corp.). These combined files were then submitted to the Mascot search algorithm (Matrix Science) and searched against the NCBI (all entries) database. A fixed modification of carbamidomethyl and a variable modification of methionine oxidation were used, to account for chemical modifications on the amino acids resulting from sample handling and processing. Cysteine reacts with iodoacetamide resulting in the covalent addition of a carboxamidomethyl group (+57.07 Da), thus this is specified as a fixed modification (Lapko et al., 2000). The extent of methionine oxidation is largely dependent on sample preparation, thus this is specified as a variable modification (Kussmann et al., 1997). Up to two internal cleavages sites were allowed, to account for incomplete digestion. Average atomic masses were used in the searches and a tolerance of 0.8 Da for fragment ions and 2.0 Da for precursor ions was allowed.

Mascot compares the observed peptide spectra to a database of theoretical translations from primary sequence databases and determines the most likely matches. It provides protein scores for each putatively identified protein, which are the sum of individual ion scores from each peptide contributing to the match, with a small correction to reduce the contribution of low scoring random matches. In this experiment, such correction was not applied to the majority of identifications or was negligible (<2) when applied. If a single peptide is detected more than once, all occurrences are included as peptide matches for as many times as the peptide has been detected. Each peptide has an ion score and associated probability that the observed match is a random event. A 5% confidence threshold for identity or extensive homology is provided. Ion scores below the homology threshold are non-significant. Duplicate peptides do not contribute to the protein score, whereas non-significant peptides do. In order to assign the correct protein match to a spot with a high degree of confidence, putatively identified proteins were examined and a series of steps were followed to increase confidence in the identification:

- Matches to trypsin and keratin were automatically excluded as these are present in all identifications and are likely to arise from contamination from

human cells when manipulating the gels (keratin), and from the digestion process (trypsin).

- For each protein, only peptides with individual ion scores higher than the homology threshold were selected ($P<0.05$). Duplicate peptides were also excluded. Protein matches with two or more unique and significant peptide matches were accepted.
- Approximate protein scores were calculated after eliminating non-significant peptides by adding individual ion scores from the retained peptides. No correction to account for low scoring random matches was applied.
- When different isoforms/homologies were identified, these were ranked according to their calculated protein score and the match with the highest score was accepted.
- For single peptide identifications that matched the same peptide in several taxa, the peptide sequence was searched in the NCBI conserved domains database, and only those peptides with a significant match (e value $<e^{-10}$) to a conserved domain were accepted.
- Observed molecular weights (MW) based on the spot position on the gel were used for confirmation with a 10 kDa tolerance.

4.3 RESULTS

4.3.1 Preliminary tests

4.3.1.1 Preparation of protein mixtures

As stated in the methodology, two lysis buffers were made using different concentrations of CHAPS, DTT and IPG ampholytes according to previous recommendations. Protein mixtures were separated on silver stained lab-cast 2D gels. Results are shown in figure 4.2. The sample prepared with the lysis buffer with the highest concentration of reagents (figure 4.2a) resulted in a clearer pattern with a higher number of spots and less distortion than that prepared with the buffer with the lower concentrations (figure 4.2b).

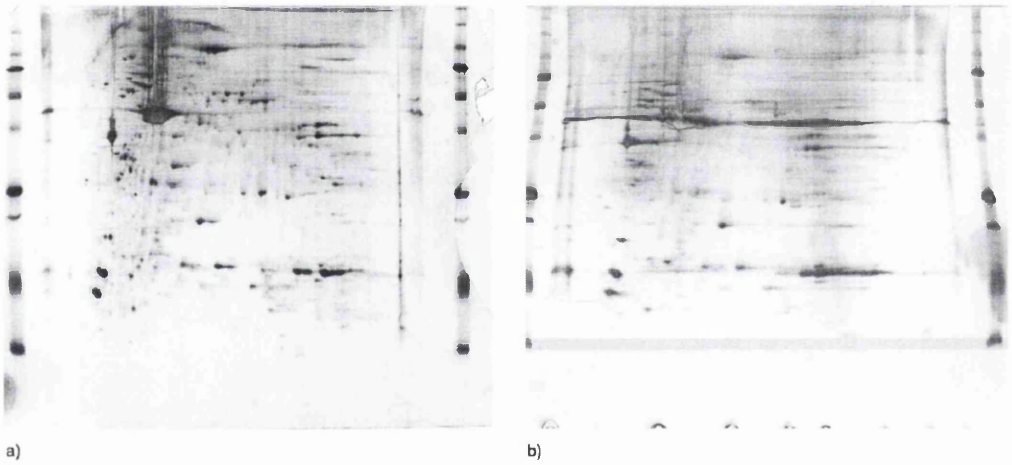


Figure 4.2 2D protein patterns on silver stained lab-cast gels obtained from *L. elliptica* mantle prepared with two different lysis buffers containing a) 7M urea, 2M thiourea, 64.8 mM CHAPS, 194.4 mM DTT and 2% (v/v) IPG ampholytes; and b) 7M urea, 2M thiourea, 32.4 mM CHAPS, 64.8 mM DTT and 1% (v/v) IPG ampholytes.

4.3.1.2 Tissue selection

2DE protein patterns obtained from *L. elliptica* siphon, gill, mantle and digestive gland are shown in figure 4.3. All tissues except digestive gland resulted in a large

number of visible spots and relatively low streaking. Mantle and siphon produced the most similar patterns. The mantle was selected for comparison of gene and protein expression changes in response to temperature in *L. elliptica*.

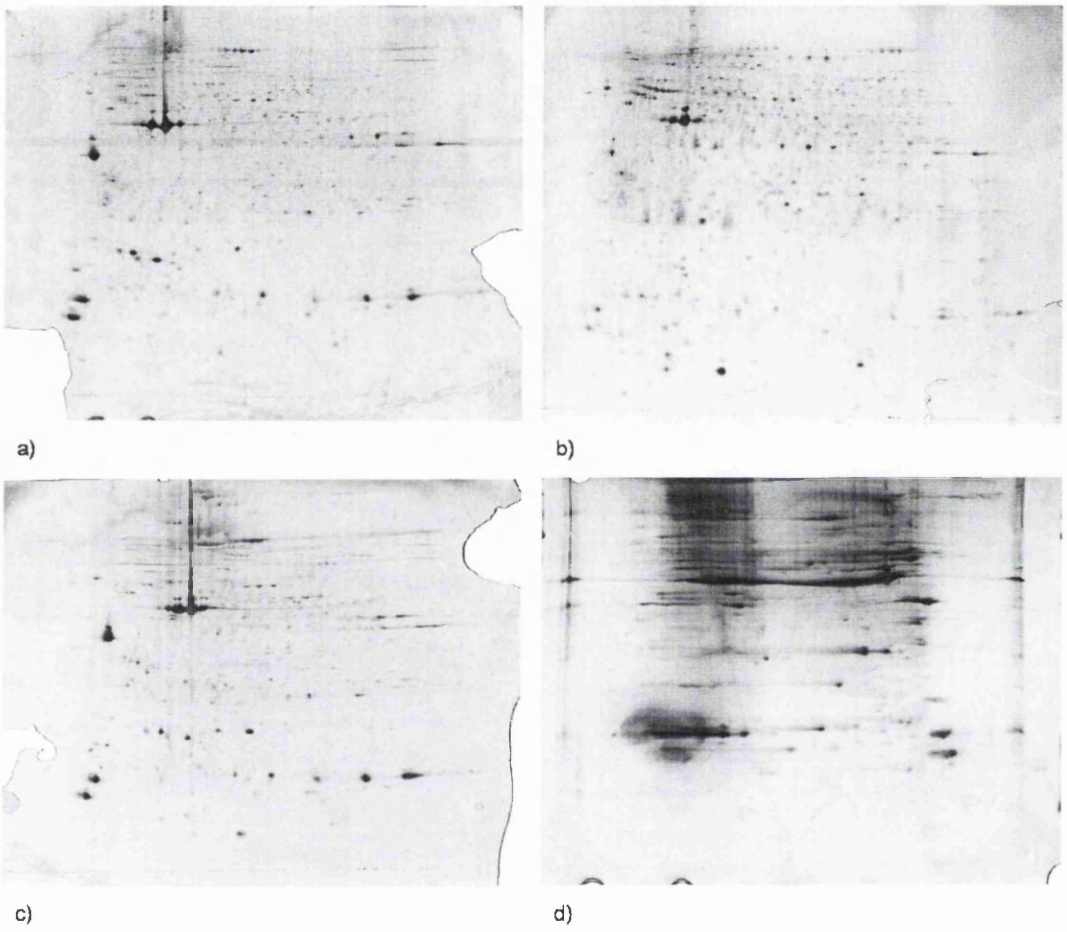


Figure 4.3 Two dimensional protein patterns on silver stained lab-cast gels obtained from different *L. elliptica* tissues including: a) siphon, b) gill, c) mantle and d) digestive gland.

4.3.2 Statistical analysis

4.3.2.1 Normality and variance homogeneity

The vast majority of spots gave acceptable fit to normality based on visual inspection using Q-Q plots and statistically with the Anderson-Darling test. Only 13.64% and 12.17% of spots did not give acceptable fit to normality for *L. elliptica* and *M. edulis*

respectively. Variance values for individual spots were homogeneous for the majority of spots. Only 3.41% and 3.44% of spots presented heterogeneous variances for *L. elliptica* and *M. edulis* respectively, as determined using Levene's test.

4.3.2.2 Technical error

The possibility of averaging the technical replicates for each individual was explored as indicated in the methodology section. Results are shown in table 4.1.

Table 4.1 Results from ANOVA analyses of $\log_2$ normalized spot volume with *Spot* and *Gel* as random factors. The *Spot*Gel* interaction for within day technical replicates (MS_{within}) and between day technical replicates ($MS_{between}$) is indicated for each animal and species. Degrees of freedom (df), F ratios for between divided by within *Spot*Gel* mean square values and associated probabilities (P) are also shown.

Mytilus edulis						*Laternula elliptica*					
Animal	**MS_{within}**	**$MS_{between}$**	**df**	**F ratio**	**P**	**Animal**	**MS_{within}**	**$MS_{between}$**	**df**	**F ratio**	**P**
A1	0.08	0.20	377	2.49	<0.001	A1	0.15	0.35	263	2.26	<0.001
		0.21	377	2.64	<0.001			0.40	263	2.59	<0.001
A3	0.11	0.20	377	1.76	<0.001	A7	0.53	0.36	263	0.68	0.999
		0.18	377	1.63	<0.001			0.42	263	0.79	0.972
A7	0.10	0.23	377	2.28	<0.001	A9	0.42	0.56	263	1.32	0.012
		0.23	377	2.30	<0.001			0.39	263	0.93	0.722
A9	0.13	0.17	377	1.30	0.006	A17	0.28	0.27	263	0.97	0.597
		0.15	377	1.19	0.047			0.39	263	1.38	0.005
A11	0.07	0.09	377	1.38	<0.001	A23	0.22	0.23	263	1.06	0.319
		0.09	377	1.38	<0.001			0.23	263	1.06	0.319
A13	0.10	0.16	377	1.52	<0.001						
		0.11	377	1.05	0.321						
A15	0.14	0.09	377	0.67	0.999						
		0.10	377	0.70	0.999						
A17	0.09	0.11	377	1.34	0.002						
		0.10	377	1.21	0.031						

For *M. edulis*, the technical error computed as *Spot*Gel* mean square for between day replicates for the same individual was significantly higher than that for within day replicates for all animals except A15. For *L. elliptica*, apart from A1 there was no clear evidence that the mean square for between day replicates was significantly greater than that within days. For A9 and A17 one comparison was significant and one was not. Interestingly, for A7 in *L. elliptica* and A15 in *M. edulis* the F ratio and associated probability indicate that technical variation measured within days is

greater than that measured between days. This suggests that although differences between runs are most relevant for most replicates, there are added differences between individual samples that can account for much of the technical variation. In both species, there was evidence that technical variation differed between replicates, hence a decision was made not to average technical replicates.

4.3.2.3 Variance components

Variance components for all factors were calculated on a spot by spot basis. Mean and medians were calculated across 378 and 264 spots in *M. edulis* and *L. elliptica* respectively to give global values. Mean and median variance components for treatment variance, biological variance and technical variance as a percentage of the total variance are given in table 4.2.

Table 4.2 Mean and median variance components across spots for treatment variance, biological variance and technical variance (both components, day and residual are indicated). Values are given for both species as a percentage of the total variance.

		Treatment (%)	Animal (%)	Day (%)	Residual (%)
L. elliptica	**Mean**	30.01	12.22	17.68	40.10
	Median	36.86	4.86	9.03	49.24
M. edulis	**Mean**	18.50	34.70	16.60	30.30
	Median	20.79	29.56	15.54	34.11

Treatment is the variation due to the differences between treatments, *animal* is the variation due to the differences between animals, which is the biological variation, and *day* is the variation due to differences between days, which is part of the technical variation of the experiment. The *residual* is the technical variation remaining after the *day* effect is accounted for and thus will be due to technical variation between replicate gels within days. In both cases, the largest variation is technical with 59.65% in *M. edulis* and 58.27% in *L. elliptica*. Biological variance is second largest in *M. edulis* (29.56%), whereas this represents the smallest component in *L. elliptica* (4.86%), where treatment variance is larger (36.86% compared to 20.79% in *M. edulis*).

Statistical comparisons were made of absolute values for biological variation (variance due to *animal*) between species across treatments using a Mann-Whitney test. In addition, comparisons of absolute values for *animal* variance for the control groups only were made between species. All comparisons were carried out using a cut down dataset for *M. edulis* to equalize the number of spots in the two species (n=264). Results are shown in table 4.3. Median *animal* variance for 264 spots was significantly higher in *M. edulis* than in *L. elliptica* suggesting significantly larger overall biological variation in *M. edulis*. However, no significant differences were found in median *animal* variance for the control group only between the two species.

Table 4.3 Pairwise comparisons of absolute values of biological variation (variance due to *animal*) between species. Comparisons were made across treatments (excluding the control group) and for the control group only. Median variances, Mann-Whitney test statistics (W) and associated probabilities (P) are shown for both comparisons.

	Median			
	L. elliptica	***M. edulis***	**W**	**P**
Across treatments	0.024	0.058	76267.5	<0.001
Control	0.048	0.039	113945.5	0.069

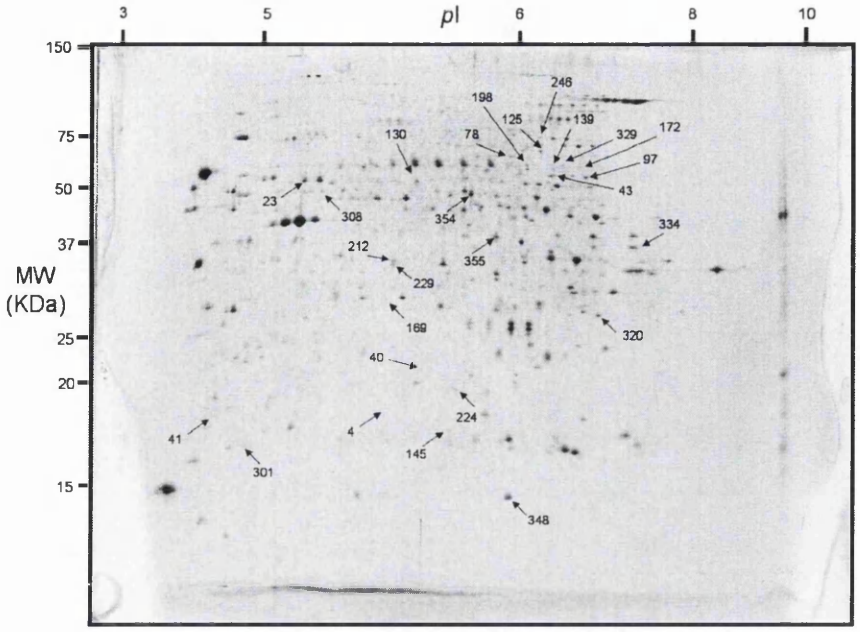
4.3.2.4 Spot by Spot analysis

Mean protein expression levels in treatment groups containing six individuals each (five in the case of *L. elliptica* 3°C buried) were compared for each species. 378 and 264 spots passed quality screening in *M. edulis* and *L. elliptica* respectively. Of these, 26 and 14 spots showed significant differences in expression between at least two of the treatments at a p-value $P<0.05$ in *a priori* tests. This represents 6.9% and 5.3% of the total number of spots in *M. edulis* and *L. elliptica* respectively. No further criteria based on the magnitude of expression differences were applied to reassess these spots and all were considered for further discussion. In *L. elliptica* no spots remained significant after FDR corrections. In *M. edulis* only five remained at a false discovery rate of 50% (spots numbers 139, 169, 224, 320, 329). Thus potentially 50% of these five might be true positive results. The chance of all five being false positives would be $0.5^5 = 0.03$ and thus the chance of at least one of these

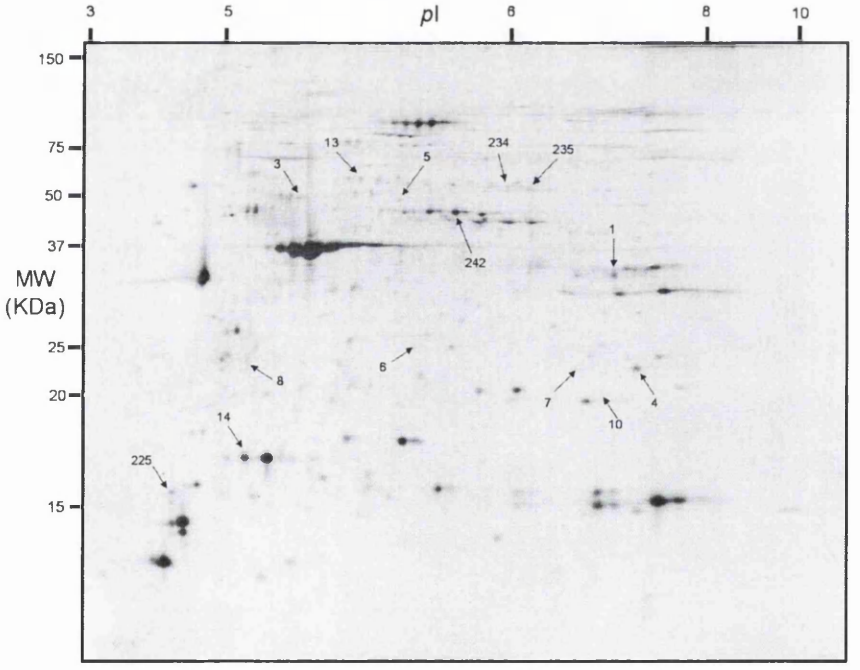
spots being a true positive is quite high at 1-0.03=0.97. Statistical analysis suggested there is a correlation between the treatments across spots in the direction of changes in expression in relation to the control (for both species $P<0.001$, one tailed Fisher's exact test). That is spots are upregulated (or downregulated) compared with the control in both treatments more often than this is expected by chance, assuming random variation in expression. This suggests that at least some spots responded to the treatment, making it unlikely that all the spots identified as showing significant changes in expression are false positives. Results from the lme tests for both species are shown in table 4.4. The AIC for both models, log likelihood ratios and associated p-values are indicated for each spot. The position of all significant spots is shown in figure 4.5. These are spread throughout the gel, although there appears to be a higher density of high molecular weight spots showing changes in expression in *M. edulis*.

Table 4.4 Results from the lme tests for model comparisons for *M. edulis* and *L. elliptica*. Spot numbers for spots with expression levels significantly different between treatments, Akaike information criterion (AIC) for models M1 (no fixed effect) and M2 (treatment as a fixed effect), log likelihood ratios (L. Ratio) and associated *a priori* p-values (P) are given.

M. edulis					*L. elliptica*				
Spot	**AIC M1**	**AIC M2**	**L. Ratio**	**P**	**Spot**	**AIC M1**	**AIC M2**	**L. Ratio**	**P**
4	67.06	62.99	8.07	0.018	1	68.56	66.27	8.28	0.041
23	68.92	65.35	7.58	0.023	3	85.14	79.50	11.64	0.009
40	26.73	23.93	6.80	0.033	4	118.06	116.02	8.04	0.045
41	66.93	64.10	6.83	0.033	5	71.65	67.44	10.21	0.017
43	49.40	46.15	7.24	0.027	6	86.79	83.49	9.29	0.026
78	35.16	32.99	6.17	0.046	7	91.37	87.70	9.67	0.022
97	29.73	27.42	6.32	0.043	8	91.04	88.61	8.43	0.038
125	38.79	35.57	7.22	0.027	10	81.29	78.46	8.82	0.032
130	40.65	36.24	8.41	0.015	13	74.00	71.68	8.32	0.040
139	49.90	42.69	11.20	0.004	14	63.74	60.32	9.42	0.024
145	31.67	29.11	6.57	0.038	225	115.29	112.91	8.38	0.039
169	48.20	40.03	12.17	0.002	234	83.74	81.03	8.71	0.034
172	20.44	17.18	7.26	0.027	235	43.73	40.95	8.78	0.032
198	6.52	1.24	9.29	0.010	242	46.87	42.22	10.65	0.014
212	28.48	25.15	7.32	0.026					
224	49.16	42.48	10.68	0.005					
229	34.97	32.60	6.37	0.041					
246	45.85	43.62	6.23	0.044					
301	77.09	74.02	7.06	0.029					
308	68.32	63.86	8.47	0.015					
320	17.12	7.65	13.48	0.001					
329	34.00	26.33	11.66	0.003					
334	64.29	59.90	8.38	0.015					
348	38.76	35.17	7.58	0.023					
354	14.40	12.22	6.19	0.045					
355	29.97	26.66	7.32	0.026					



a)

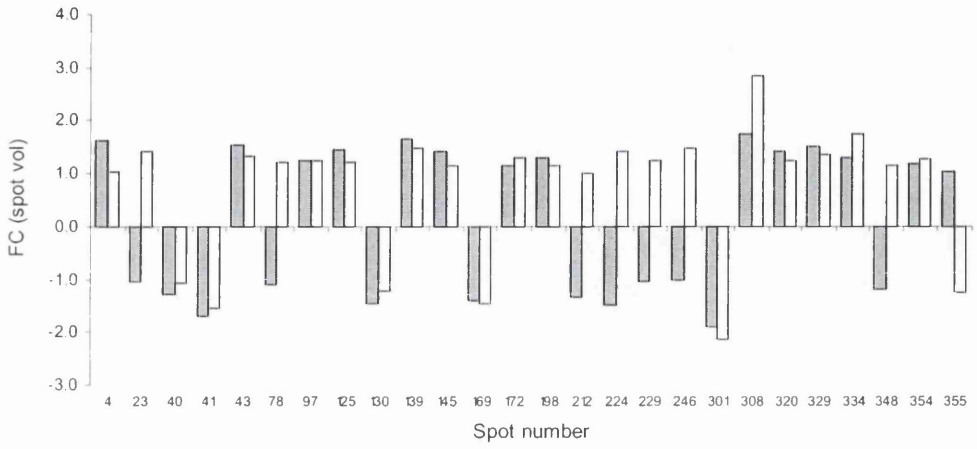


b)

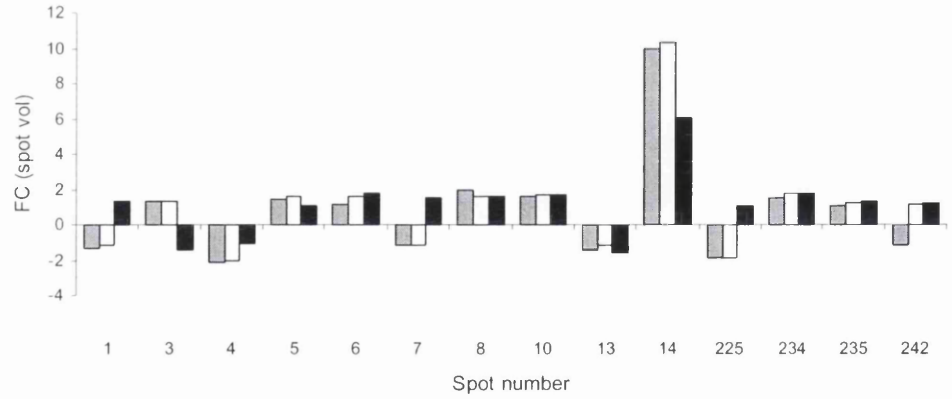
Figure 4.4 Two dimensional protein patterns on silver stained precast gels obtained from a) *M. edulis* foot and b) *L. elliptica* mantle. The isoelectric point (pI) and molecular weight (MW) in kilodaltons are indicated on the horizontal and vertical axis respectively. Arrows indicate 26 and 14 spots that show significant differences between at least two of the treatments in *M. edulis* and *L. elliptica* respectively.

4.3.2.5 Treatment differences

Fold changes in protein expression for the 26 and 14 spots showing significant differences between treatments in *a priori* tests for *M. edulis* and *L. elliptica* respectively are shown in figure 4.6. Fold changes were calculated by dividing mean expression levels in the experimental groups by those in the control group. Mean expression levels for each spot and treatment were ML estimates (i.e. from lme output).



a)



b)

Figure 4.5 Mean fold changes in protein expression for all proteins showing significant differences in expression between treatments for a) *M. edulis*. Grey and white bars represent fold changes in the 27°C and 37°C treatments respectively in relation to the control (12°C) and b) *L. elliptica* mantle. Grey, white and black bars represent fold changes in the 3°C buried, 3°C non buried and 9°C treatments respectively in relation to the control (0°C). All values are normalized.

Out of the 26 proteins that showed significant changes in expression between treatments in *M.edulis* foot tissue, 13 and five were up- and down-regulated respectively in both treatments in relation to the control. The remaining eight proteins showed a different pattern in the two treatments with seven proteins being downregulated in the moderate heat shock and upregulated again in the extreme heat shock, whereas one protein showed the inverse pattern. Of the 14 proteins showing significant changes in expression between treatments in *L. elliptica* mantle tissue, seven and two were up- and down-regulated respectively in all the treatments in relation to the control, whereas the remaining five showed variable patterns in expression. In all cases but one, protein expression in the 3°C treatments changed in the same direction in relation to the control.

Maximum likelihood statistics and fold changes for all treatments in relation to the control, as well as fold changes between treatments are shown in table 4.5. Only spots that showed significant differences between at least two of the treatments in *a priori* tests are included. ML statistics and associated p-values are shown for all treatment comparisons. Significant p-values ($P<0.05$) are shown in bold. In *M. edulis*, most significant differences are found between the heat shock treatments and the control, whereas the majority of spots did not show significant differences in expression between the two treatments. Six proteins out of 26 differed significantly between the treatments, while 15 and 14 differed significantly between the 27°C and 37°C treatments and the control respectively. For *L. elliptica*, as expected, most significant changes in expression occurred between the 9°C treatment and all others, with the least significant differences found between the two 3°C treatments.

Table 4.5 Mean fold changes (FC) in protein expression of each treatment in relation to the control, and to each other for a) *M. edulis* b) *L. elliptica*. Results are shown for significantly up- or down- regulated spots between at least two of the treatments in *a priori* tests. Positive and negative values indicate up- and down-regulation respectively. Maximum likelihood statistics (t-value) and associated probability (P) are also shown. Bold probabilities indicate significant differences between the treatments in the level of expression of a given spot (P<0.05)

a)

Spot	FC 27°C/12°C	t-value	P	FC 37°C/12°C	t-value	P	FC 37°C/27°C	t-value	P
4	1.63	2.59	**0.021**	1.03	0.18	0.861	-1.58	2.48	**0.026**
23	-1.05	0.34	0.742	1.41	2.19	**0.045**	1.49	2.59	**0.021**
40	-1.27	2.53	**0.023**	-1.07	0.69	0.500	1.19	1.89	0.079
41	-1.71	2.45	**0.027**	-1.56	2.02	**0.061**	1.10	0.43	0.676
43	1.55	2.68	**0.017**	1.32	1.70	0.109	-1.17	0.99	0.339
78	-1.11	0.87	0.397	1.20	1.54	0.145	1.34	2.48	**0.026**
97	1.25	2.21	**0.043**	1.24	2.16	**0.047**	1.00	0.05	0.962
125	1.44	2.71	**0.016**	1.22	1.48	0.160	-1.18	1.25	0.231
130	-1.46	2.96	**0.010**	-1.22	1.54	0.144	1.20	1.43	0.174
139	1.67	3.46	**0.004**	1.47	2.60	**0.020**	-1.14	0.86	0.402
145	1.42	2.53	**0.023**	1.15	1.00	0.335	0.81	1.55	0.143
169	-1.40	2.97	**0.010**	-1.46	3.37	**0.004**	-1.05	0.41	0.690
172	1.15	1.41	0.179	1.31	2.70	**0.016**	1.14	1.32	0.205
198	1.30	3.29	**0.005**	1.16	1.88	0.079	-1.12	1.43	0.174
212	-1.33	2.37	**0.032**	1.00	0.02	0.984	1.32	2.36	**0.032**
224	-1.48	3.16	**0.006**	1.42	2.79	**0.014**	1.05	0.38	0.708
229	-1.05	0.45	0.660	1.24	1.91	**0.076**	1.30	2.42	0.029
246	-1.02	0.10	0.923	1.49	2.10	0.053	1.52	2.21	0.043
301	-1.89	2.12	0.051	-2.13	2.52	**0.023**	-1.13	0.40	0.695
308	1.75	1.62	0.126	2.83	2.99	**0.009**	1.61	1.38	0.189
320	1.42	4.08	**0.001**	1.24	2.51	**0.024**	-1.15	1.60	0.131
329	1.49	3.45	**0.004**	1.36	2.66	**0.018**	-1.10	0.79	0.439
334	1.30	1.47	0.162	1.76	3.17	**0.006**	1.35	1.74	0.102
348	-1.18	1.53	0.148	1.14	1.17	0.260	1.35	2.77	**0.014**
354	1.18	1.69	0.113	1.28	2.50	**0.025**	1.08	0.82	0.427
355	1.05	0.42	0.683	-1.26	2.06	0.057	-1.32	2.50	**0.024**

b)

Spot	FC 3°CB/0°C	t-value	P	FC 3°CNB/0°C	t-value	P	FC 9°C/0°C	t-value	P
1	-1.35	1.34	0.195	-1.12	0.54	0.594	1.35	1.41	0.176
3	1.31	1.26	0.223	1.36	1.50	0.151	-1.39	1.54	0.141
4	-2.10	2.12	**0.047**	-2.03	2.09	0.050	-1.06	0.18	0.862
5	1.41	2.02	0.058	1.56	2.69	**0.015**	1.06	0.32	0.751
6	1.16	0.72	0.481	1.59	2.24	**0.037**	1.75	2.61	**0.017**
7	-1.19	0.77	0.452	-1.15	0.65	0.522	1.54	1.90	0.072
8	1.94	2.87	**0.010**	1.59	2.04	0.055	1.59	1.96	0.065
10	1.59	2.36	**0.029**	1.70	2.77	**0.012**	1.67	2.55	**0.020**
13	-1.46	2.05	0.054	-1.17	0.89	0.383	-1.61	2.56	**0.019**
14	9.94	0.03	0.976	10.31	0.21	0.836	6.02	2.40	**0.027**
225	-1.89	1.88	0.075	-1.84	1.86	0.079	1.10	0.28	0.786
234	1.52	1.91	0.071	1.78	2.71	**0.014**	1.77	2.58	**0.018**
235	1.03	0.28	0.780	1.20	1.56	0.135	1.37	2.56	**0.019**
242	-1.19	1.28	0.216	1.17	1.21	0.242	1.24	1.60	0.127

Spot	FC 3°CNB/3°CB	t-value	P	FC 9°C/3°C B	t-value	P	FC 9°C/3°CNB	t-value	P
1	1.20	0.85	0.408	1.82	2.73	**0.013**	1.52	1.99	0.061
3	1.04	0.20	0.846	-1.82	2.89	**0.009**	-1.90	3.20	**0.005**
4	1.04	0.11	0.913	1.98	2.00	0.060	1.91	1.96	0.065
5	1.11	0.67	0.513	-1.33	1.75	0.097	-1.48	2.43	**0.025**
6	1.36	1.60	0.125	1.51	2.02	0.058	1.11	0.51	0.618
7	1.03	0.14	0.891	1.83	2.80	**0.011**	1.78	2.73	**0.013**
8	-1.23	0.94	0.357	-1.23	0.90	0.377	1.00	0.00	1.000
10	1.07	0.38	0.709	1.05	0.26	0.801	-1.02	0.11	0.916
13	1.24	1.29	0.213	-1.11	0.57	0.576	-1.38	1.84	0.081
14	1.04	0.17	0.865	-1.65	2.36	**0.029**	-1.71	2.66	**0.016**
225	1.03	0.10	0.925	2.08	2.21	**0.040**	2.02	2.20	**0.041**
234	1.17	0.76	0.457	1.16	0.71	0.486	1.00	0.01	0.989
235	1.16	1.35	0.192	1.32	2.42	**0.026**	1.14	1.16	0.261
242	1.39	2.59	**0.018**	1.47	2.94	**0.008**	1.06	0.45	0.656

4.3.2.6 Identification of candidate spots by MS/MS

Protein identifications by mass spectrometry were attempted on the 40 spots showing significant differences between treatments prior to FDR corrections. Of these, ten were identified by MS/MS based on the chosen parameters, six of which were identified in *M. edulis* foot (38.5%), and four in *L. elliptica* mantle (28.6%). Results are shown in table 4.6. All spots were identified with two or more peptides, except M212. The latter was identified with a single peptide, which presented strong sequence similarity to a known conserved domain (e^{-17}). Eight of the ten spots gave

significant matches to different proteins, whereas the other two (M145 and L242) were putatively identified as enolase. These could represent the same protein in both species or different isoforms or allelic forms. All spots were found approximately within their theoretical MW (±10 kDa) except M145, which was present in the gel between 15-20 kDa, while expected at about 45 kDa.

In *M. edulis*, six spots were putatively identified as gamma actin, enolase-1, a member of the glycosyltransferase GTB type superfamily, fructose-biphosphate aldolase, S-adenosylhomocysteine hydrolase and malate dehydrogenase. These have been allocated categories based on their GO annotations and previous literature: structural constituent of the cytoskeleton, energy metabolism and protein turnover. In *L. elliptica*, four spots were putatively identified as alpha tubulin, aldehyde dehydrogenase, TCP-1 alpha subunit and enolase. The assigned categories are: structural constituent of the cytoskeleton, energy metabolism, redox regulation and protein folding.

Table 4.6 Protein spots identified by MS/MS in *M. edulis* foot (M) and *L. elliptica* mantle (L). Mascot protein score and number of peptides with results after filtering in brackets, % sequence cover, theoretical molecular weight (MW), sequences for retained peptides and allocated function are shown. For single peptide identifications e-values are shown in brackets.

Spot	Protein	Uniprot ID	Protein Score	Peptides	Sequence cover (%)	Theoretical MW (Da)	Peptide Sequence	Category
M23	Actin, gamma	P20359	129 (129)	2 (2)	7.5	41850	K.AGFAGDDAPR.A R.VAPEEHPVLLTEAPLNPK.A	Structural constituent of cytoskeleton
M145	Enolase 1	P00924	627 (563)	10 (6)	23.3	46830	R.GNPTVEVELTTEK.G K.AVDDFLISLDGTANK.S K.TAGIQIVADDLTVTNPK.R R.SGETEDTFIADLVVGLR.T R.SIVPSGASTGVHEALEMR.D R.IEEELGDNAVFAGENFHHGDK.L	Energy metabolism
M212	Glycosyltransferase superfamily	(e^{-17}) cl10013	55	1	5.3	30973	K.VIPAADLSEQISTAGTEASGTGNMK.F	Glycosylation
M334	Fructose-bisphosphate aldolase	Q6PTI8	198 (198)	4 (2)	17.4	21807	K.VTEQVLAFTYK.A K.GVVPLAGTDGESTTQGLDGLAER.C	Energy metabolism
M354	S-adenosylhomocysteine hydrolase	Q9SDP1	204 (204)	8 (3)	9.9	34216	R.ATDVMIAGK.V K.VAFVAGYGDVGK.G K.SKFDNLYGCR.H	Protein turnover
M355	Malate dehydrogenase	Q9NHX3	155 (155)	5 (2)	7	36590	R.IQNAGTEVVEAK.A R.DDLFNTNAGIVR.D	Energy metabolism
L3	Tubulin, alpha 1c	Q6P8G7	686 (549)	17 (7)	31.4	50532	K.DVNAAIATIK.T R.QLFHPEQLITGK.E R.TIQFVDWCPTGFK.V K.VGINYQPPTVVPGGDLAK.V R.AVCMLSNTTAIAEAWAR.L R.AFVHWYVGEGMEEGEFSEAR.E R.FDGALNVDLTEFQTNLVPYPR.I	Structural constituent of cytoskeleton
L5	Aldehyde dehydrogenase	A7XZK3	153 (115)	3 (2)	6	53544	K.ILGLIESGK.K R.EEIFGPVQQILK.Y	Redox regulation Energy metabolism
L235	T-complex protein 1, alpha subunit	Q4AE76	269 (226)	5 (3)	7.6	59962	R.LLEVEHPAAK.V R.ICDDELIIIK.G R.TQNVMAASSIANIVK.S	Chaperone
L242	Enolase	A8DU76	233 (191)	9(2)	10.8	40012	K.IQIGMDVAASEFCK.D R.AAVPSGASTGIYEALEMR.D	Energy metabolism

4.4 DISCUSSION

The following sections include the analysis of a series of technical considerations related to protein and gel quality. This is followed by a discussion on the observed biological variance in *L. elliptica* and *M. edulis*. Finally, the significance of the observed patterns in protein expression following heat stress in both species is discussed.

4.4.1 Technical issues

4.4.1.1 2DE patterns, protein and gel quality

Preliminary analysis of different tissues was carried out in *L. elliptica* to determine whether the choice of mantle tissue for analysis was appropriate and significantly better 2DE patterns were not achieved using other tissues. Visual inspection of 2DE patterns in different *L. elliptica* tissues (figure 4.3) suggests a similar overall proteome in gill, mantle and siphon, whereas high quality 2DE patterns were not achieved from digestive gland tissue, potentially as a result of the presence of exogenous contaminants or interfering substances (Lopez, 2005). The extent of the similarities between different 2DE patterns varies between species. A study of protein expression in osmoregulatory tissues in the euryhaline spiny dogfish shark using 2DE revealed that tissues with osmoregulatory function (rectal gland, gill, intestine and kidney) were more similar in their overall proteome than non-osmoregulatory tissues. However, the proportion of proteins found in a single type of tissue was very low (4%) when compared to that of proteins equally abundant in all tissues (36%) (Lee et al., 2006). In humans, however, protein patterns from different tissues are highly dissimilar (Cagney et al., 2005). Similarities in the expressed proteome of the different tissues do not necessarily guarantee that these will show similar responses to a given treatment. Since no other tissue resulted in significantly better 2DE patterns, the mantle was retained to allow direct comparisons with the gene expression data.

Visually, there were clear differences in the quality of the 2DE patterns between the two species, with *L. elliptica* gels presenting less proteins and more streaking than *M. edulis* (figure 4.5). This was reflected in the number of spots retained for final analysis after quality filtering (378 and 264 in *M. edulis* and *L. elliptica* respectively). The reason why overall gel quality is lower in *L.elliptica* is unclear. Streaking is a common problem with 2DE gels, which can occur as a result of individual properties of the proteins, the range of proteins in the sample or the 2DE technique used. The presence of spots with streaks in areas with well defined spots could indicate that this is due to the properties of specific proteins, either as a result of high protein concentration or because these are insoluble before reaching their isoelectric point (Lopez, 2005). Reducing protein concentrations tends to minimise streaking, with the disadvantage that low abundance proteins cannot be visualised, thus a compromise between these factors needs to be found. The streaking and reduced number of spots and reproducibility observed in *L. elliptica* could also be a result of low quality of the protein mixtures, as with the low RNA quality discussed in chapter three. It is possible that enhanced enzymatic activity at low temperatures reduces the effectiveness of the use of cold temperatures to ensure preservation and a high degree of protein degradation occurs.

Overall reproducibility was assessed by determining the proportion of the total variance explained by technical variation, and by comparing the between and within day technical error. In both cases, the largest variance component was technical (59.65 and 58.27% in *M. edulis* and *L. elliptica* respectively) confirming the high degree of technical variation intrinsic to the technique, which arises from the combined errors generated, starting with sample preparation to matching the 2D gels (c.f. Challapalli et al., 2004). The coefficient of determination for normalised spot volume for each animal for which technical replicates were run was calculated across spots between days and within days. Values ranged from 0.97 to 0.93 for between day technical error, and from 0.98 and 0.96 for within day technical error for *M. edulis*. These values are within the range reported in previous studies (Lopez et al., 2001; Diz and Skibinski, 2007). For *L. elliptica*, values were slightly lower ranging from 0.97 to 0.92 for between day technical error, and from 0.97 to 0.95 for within day technical error. This indicates that *M. edulis* gels were generally more

reproducible, and that technical variation between replicates of the same individual run in different days was higher than between replicates run on the same day for both species. The majority of the comparisons of between and within day technical variation were statistically significant (see table 4.1). The phenomenon of run to run variation is intrinsic to the technique (Ong and Pandey, 2001), thus it is convenient to run as many gels as possible in the same batch. This is limited by the equipment, which allows a maximum of six gels to be run at once, and by the characteristics of the sample preparation procedure. Protein mixtures undergo several steps before electrophoresis, and manipulating more than six to ten samples at once, even for an experienced researcher, is likely to increase protein damage due to the time elapsed between processing the first and last samples.

4.4.1.2 Selection of candidate proteins

The criteria established to select candidate proteins were reduced compared to those applied to select candidate genes. The only requirement for a spot to be selected as a candidate protein for further analysis via MS was that it was significantly upregulated between at least two of the treatments prior to FDR correction ($P<0.05$). In the study of gene expression in *L. elliptica* discussed in chapter three, a threshold in the level of expression was used to identify differentially expressed genes, i.e. a gene was said to be differentially expressed if the change in expression between the control and the heat shock treatments exceeded a two fold change. This approach is common in microarray studies and has some advantages. The size of systematic errors compared to changes in expression will be greater for low than for high abundance genes, thus significant changes in genes expressed at lower levels may be missed (Hack and Lopez, 2004). In addition, the use of a threshold expression change allows the simplification of vast amounts of data often generated and the significance of the changes in expression of the genes showing differences of larger magnitude can be discussed. Such a threshold, however, was not applied to the analysis of the stress proteome as the change in expression of the majority of significantly up- or down-regulated spots was below this threshold (e.g. mean fold changes in treated animals in relation to control were $+1.38\pm0.05$ and -1.37 ± 0.07 for *M. edulis* and $+2.24\pm0.45$ and -1.47 ± 0.10 for *L. elliptica*; mean±SE). Moreover, the

use of a two fold cutoff criterion has been described as a conservative approach to identifying genes differentially expressed under different experimental conditions, because physiologically important effects may arise from shifts in gene expression that are less than two fold (Podrabsky and Somero, 2004). This could potentially be more significant at the level of protein expression since the proteome is closer to the phenotype than the transcriptome (Biron et al., 2006), thus changes in protein expression of any magnitude may directly reflect changes observed in the phenotype.

The interpretation of the outcome of statistical analysis also resulted in reduced stringency in the selection of candidate spots. When multiple tests are undertaken, each at the same significance level, the probability of achieving at least one significant result increases with the number of independent tests (Verhoeven et al., 2005). Therefore, there is an increased probability of producing false positives, or rejecting a null hypothesis when it would be inappropriate to do so. A solution to this problem is to control and maintain at a desired level the proportion of significant results that are in fact type I errors or false positives through false discovery rate (FDR) (Benjamini and Hochberg, 1995). In the present study, comparisons between treatments were carried out spot by spot thus 378 and 264 independent tests were performed for *M. edulis* and *L. elliptica* respectively. FDR corrections were carried out at false discovery rates of 5, 20 and 50%, which revealed a high probability that the majority of spots showing *a priori* significant changes in expression between treatments could in fact be false positives. The list of differentially expressed spots was therefore taken as a ranking of proteins showing the most consistent changes in expression between treatments. All proteins with p-values $P<0.05$ prior FDR corrections were retained for mass spectrometric analysis. Given the uncertainty in statistical significance, the use of the identified proteins as stress biomarkers would require further validation. However, even with high confidence in the statistical significance of the treatment differences, validation using other methodologies (i.e. western blots or quantitative MS) would always be required to ensure independent and accurate results. As more data are collected, the probability of finding some spurious results is quite high, but the chance of all the results being spurious is extremely improbable (Moran, 2003). These potentially spurious results should not

be of great concern in exploratory research, as they would be eliminated in future experiments.

4.4.1.3 MS identifications

As previously discussed, the sequence information available for *M. edulis* and *L. elliptica* is currently limited to 1470 and 14 protein sequences respectively. Protein identification by MS relies on database searches which, when analysing poorly characterised species, are dependent on the degree of sequence similarity with better characterised species. The limited success of the identifications in this study suggests significant aminoacid sequence variation between molluscan proteins and their homologues in model species. Even if homologous peptide sequences differ by a single amino acid, the success of the identification of the corresponding protein will be reduced because peptides will not match (Kültz et al., 2007). The proteins identified in the current study are well described, highly conserved proteins. Nonetheless, the proportion of proteins identified from those showing significant changes in expression between treatments was higher than the proportion of *L. elliptica* genes identified in the microarray (14%). It is possible that this relates to the nature of 2DE, which is biased towards high abundance proteins (Wilkins et al., 1996). These are often housekeeping proteins, thus highly conserved across taxa. Indeed, many changes observed at the transcript level involved regulatory and signaling proteins, which cannot often be visualized in a gel due to their low abundance (Zivy and de Vienne, 2000)

4.4.2 Biological variation: species differences and the effects of temperature

In the present study, biological variation in protein expression was compared between species both in the control group and across treatments, excluding the control group. From the viewpoint of analysis of causes of variation, protein expression can be considered similar to any other phenotypic character, such as shell length (Diz and Skibinski, 2007). In this respect, the variation underlying this expression should reflect differences between the two species under stress and under the natural conditions of their respective environments. Biological variation across

treatments was significantly higher in *M. edulis*, a eurythermal species with a wide distribution, than in *L. elliptica*, a highly stenothermal species with a corresponding narrow distribution. Interestingly, there were no significant differences in biological variation between species when taking into account the control group only. This suggests that the increased biological variation observed in *M. edulis* across treatments could be stress induced.

Increased variation induced by stressful conditions has been extensively reported (reviewed in Badyaev, 2005). This variation can be hidden, when it is not seen under normal conditions but becomes apparent when a stressor threatens organismal homeostasis (Badyaev, 2005). The factors contributing to the increased variation in the response to stress in *M. edulis* are unclear. It is possible that populations of this species display larger biological variation in at least some stress responsive proteins, whereas *L. elliptica* populations do not have the individual variation underlying the expression of a similar set of proteins. In this case, variation would be hidden in the control individuals and become apparent as a result of a gene-environment interaction.

Differences in phenotypic and genetic variation in certain traits between species from different environments have been previously observed. For example, a recent study using *Drosophila* species demonstrated that narrowly distributed tropical species have low genetic variation for tolerance traits (i.e. desiccation and cold) compared with widely distributed species (Kellermann et al., 2009). The reasons for this are unclear, although a potential explanation is that traits subjected to strong directional or stabilizing selection display reduced genetic variability (Merilä, 2009). On this basis, it can be hypothesized that the observed differences in individual variation at the protein level between *L. elliptica* and *M. edulis* may be linked to the characteristics of their habitats. Stabilizing selection acting on stress related proteins under the extremely stable temperatures of the Antarctic environment may result in the decreased variation observed in *L. elliptica*. *M. edulis* experiences variable conditions in its environment, thus there might be an element of disruptive selection leading to increased genetic variation producing several genotypes under different

conditions. Alternatively, there could be a single genotype which responds to stress with physiological flexibility.

The differences in biological variation between the two species upon exposure to elevated temperatures might be relevant to their potential to adapt in a warming climate. Low levels of genetic variation have been shown to reduce the evolutionary potential of targeted ecological traits relevant during climate change (Hoffmann et al., 2003), whereas increased phenotypic and genotypic variance under stress is a source of novel adaptations under changed environments (Badyaev, 2005). In the present study, it might be expected that much of the observed phenotypic variation is environmentally induced rather than genetic. This distinction is particularly important and the relevance of distinguishing between genetic (evolutionary) and phenotypic (including a plastic component) responses to climate change has been previously discussed (Gienapp et al., 2008). To draw further conclusions, the measured variation should be partitioned into genetic and environmental components. Nonetheless, based on the points presented above, it can be hypothesized that *M. edulis* would have an evolutionary advantage over *L. elliptica* when adapting to a changing environment.

The potential inherent differences between species must be taken into account. To make comparisons possible, both species were exposed to temperatures representing similar thresholds, corresponding with temperatures around which they become unable to acclimate, and those close to their upper lethal limits. However, differences in the length and intensity of exposure and divergence between the genera are likely to be responsible for some of the differences observed between the two species, which may not be solely attributed to differential responses evolved as a result of the characteristics of their environment. The responses to heat stress observed in *M. edulis* and *L. elliptica* in this study must therefore be carefully evaluated before further conclusions can be made, regarding their susceptibility to climate change. In order to extrapolate the differences observed between species to their respective environments, comparative studies on the variation underlying protein and gene expression in other closely related Antarctic and temperate species is needed. Evaluation of the effects of slower rates of warming on biological

variance is also required. The results presented here provide an interesting line of research to compare the susceptibility to climate change of Antarctic and temperate species through quantitative genetic studies on variable heritable and non-heritable traits.

4.4.3 Effects of temperature increases in protein expression

Despite reducing the criteria for selection of candidate proteins, compared to those established when selecting candidate genes, the number of proteins showing changes in expression in response to heat stress in both species was low. This may be attributed to several factors. Firstly, the large variability, both technical and biological, results in lower power and fewer proteins being retained for analysis. This reduces the number of proteins showing statistically significant changes in expression between treatments. Secondly, the lack of significant changes in the expression of the majority of spots does not imply that changes in the proteome do not occur in response to heat stress. There are other properties that control protein regulation and function, namely post-translational modifications and association with other proteins, nucleic acids or micromolecules (inorganic ions, organic osmolytes, metabolic intermediates) (Kültz et al., 2007). These are not immediately detected on 2DE gels and require different sample preparation or staining techniques from those applied when identifying changes in expression. As an example, the phosphorylation of stress-activated protein kinases in the tissues of *M. galloprovincialis* is thought to initiate a signalling cascade essential in the induction of HSP genes and potentially other aspects of the stress response (Anestis et al., 2007). Thirdly, the response to stress for a large number of proteins may be delayed when compared to that at the level of gene expression. A time shift between mRNA expression and that of the corresponding protein has been observed in previous studies (Aaronson et al., 1995; Le Roch et al., 2004; Foth et al., 2008). This would prevent many of the changes observed at the mRNA level from being detected at the protein level, even if the corresponding proteins were present on the gels and selected for analysis. Nonetheless a number of candidate proteins showing differences in expression between treatments have been identified, and their putative function and relevance of the observed changes are discussed below.

4.4.3.1 Global patterns in changes in expression

In *M. edulis* foot, most significant changes in protein expression were found between the two temperature treatments and the control (15 and 14 proteins) whereas significant differences between the two treatments were only found in six proteins. In *M. edulis* 27°C is just below the temperature at which individuals become unable to acclimate (Chapple et al., 1998) with the upper lethal temperature estimated at 31 to 41°C (Davenport and Davenport, 2005; Henderson, 1929). It is possible that the temperature treatments applied in this study do not represent an increase in stress along a gradient and that a physiological threshold is exceeded before or as the lowest temperature is reached. This phenomenon has been previously suggested in a proteomic study on the thermal stress response in dogwhelks exposed to 26.5 and 30°C treatments, where similar protein expression patterns were observed between the two treatments (Gardeström et al., 2007). The existence of a physiological threshold before or at the lowest treatment temperature would point at 27°C being close to the critical temperature for the sampled *M. edulis* populations.

In *L. elliptica* mantle, more proteins showed significant changes in their level of expression between 9°C and all other temperature treatments (six, seven and five), than between any of the treatments and the control (three and four). The two 3°C treatments (buried and not buried) showed the lowest number of proteins changing significantly in expression between them (only one). At 9°C, the critical temperature has been exceeded and the transfer to anaerobic metabolism has occurred (Peck et al., 2002), potentially resulting in different protein expression profiles and a larger set of protein showing significant changes in expression.

4.4.3.2 Candidate proteins in *L. elliptica*

Aldehyde dehydrogenase and enolase changed in their level of expression in *L.elliptica* individuals exposed to elevated temperatures. Aldehyde dehydrogenase increased in both treatments, with the most significant upregulation observed in 3°C

treated clams. Enolase decreased initially, and then increased in expression with the intensity of the treatment. Both proteins have been described as part of the minimal stress proteome of cellular organisms, a set of proteins that participate in different aspects of the cellular stress response and are ubiquitously conserved in all three superkingdoms (Kültz, 2005). In the absence of stress, aldehyde dehydrogenase is an enzyme in the glycolytic pathway, thus essential for energy production. However, under acute stress, this enzyme acts as an oxidoreductase, indirectly involved in redox regulation, by detoxifying aldehydes, toxic intermediary metabolites during oxidative stress. Enolase-1 (αα- or non neural- enolase) is a glycolytic enzyme, found in all organisms, that catalyzes the dehydration of 2-phosphoglycerate to phosphoenolpyruvate in the glycolytic pathway, and the reverse reaction in gluconeogenesis (Tracy and Hedges, 2000). Induction of these and other enzymes involved in energy metabolism during stress may be necessary for generating reducing equivalents (NADH, NADPH) that are needed for cellular antioxidant systems, or to respond to the energetic requirements of protein degradation, protein chaperoning, and DNA repair (Kültz, 2005). Both are in accordance with the observed decrease in aerobic capacities and the induction of antioxidant defenses at 3°C.

Tubulin is a cytoskeletal protein and main constituent of microtubules, which consists of many isoforms (Williams et al., 1999; Forer, 2008). The upregulation of tubulin in clams exposed to 3°C is consistent with the induction of cytoskeletal protein transcripts reported in chapter three. Although there was no evidence that tubulin was induced at the transcript level, it is possible that either this gene was not represented on the microarray, or the change in expression was below the two fold threshold. The reason why tubulin is downregulated at 9°C is unclear but consistent with both hypotheses discussed in chapter three. If the demand for this protein is a response to damage of cytoskeletal components due to temperature, a decrease in its expression at 9°C could be related to an attempt to save energy to maintain other cellular processes. If the changes in expression of tubulin arise mainly as a response to increased muscle contraction in an attempt to bury at 3°C, the down-regulation of cytoskeletal proteins is also expected, as *L. elliptica* is unable to bury at 9°C (Peck et al., 2004). Indeed attempts to bury at this temperature were not observed in the

present study or during experiments carried out in our laboratories (S. Morley, pers. comm.)

T-complex polypeptide-1 (TCP-1) is a chaperonin found in archaebacteria and in the eukaryotic cytosol which belongs to the TRiC family. It is thought that members of this family are not stress inducible in eukaryotes and have a restricted range of substrates (Hartl, 1996). However, it holds striking similarity with the archaebacteria thermophilic factor 55 (TF55), a molecular chaperone induced by thermal stress, which increases thermotolerance (Trent et al., 1991). TCP-1 functions as a chaperone in eukaryotes directing folding of cytoskeletal proteins (reviewed in Liang and MacRae, 1997). Interestingly, it is key in the biogenesis of tubulin in the cytosol, where it binds both the α and β tubulin subunits releasing them in a native-like assembled form in a process than requires ATP (Yaffe et al., 1992). It is thus possible that the upregulation of TCP-1 responds to the demand for tubulin and, based on the gene expression data, other cytoskeletal proteins. The induction of TCP-1 is likely linked with the induction of a prefoldin transcript observed and reported in chapter three. Indeed, the biogenesis of the cytoskeletal proteins actin and tubulin involves the interaction of nascent chains of each of the two proteins with prefoldin and their subsequent transfer to the cytosolic chaperonin containing TCP-1 (Llorca et al., 1999; Martín-Benito et al., 2002). However, despite the obvious association with cytoskeletal proteins, given the similarity with TF55, the function of TCP-1 as a stress protein in *L. elliptica* cannot be rejected. In the absence of inducible HSPs at environmentally relevant temperatures, these chaperonins may have a role in protein folding of temperature damaged proteins. TCP-1 deserves further attention as a potential stress induced chaperone in *L. elliptica.*

4.4.3.3 Candidate proteins in *M.edulis*

Actin was downregulated in the 27°C treatment in relation to the control, although the change in expression was very small (-1.05 fold) and differences were not significant. In the 37°C treated mussels, however, it was significantly upregulated in relation to the control and to the 27°C treatment. Actin is a cytoskeletal protein

whose function has been discussed in chapter three. The upregulation of actin during acute heat stress is in accordance with previous studies, and suggests the need for newly synthesized proteins either to replace those degraded by high temperatures, or to perform the necessary cytoskeletal rearrangements to maintain osmotic balance following disruption by heat (Buckley et al., 2006; Buckley and Somero, 2009).

A number of proteins associated with energy metabolism changed in expression after exposure to elevated temperatures. Amongst these, those enzymes that catalyze reactions during glycolysis (enolase and aldolase) were upregulated in both treatments in relation to the control. The function of these enzymes and their demand during heat stress has been discussed in the previous section. Malate dehydrogenase, however, involved in the Krebs cycle, showed a negligible increase in expression (1.05 fold) at 27°C with a decrease at 37°C in relation to the control (-1.26 fold). Differences were only significant between the two treatments. Malate dehydrogenases are highly conserved multimeric enzymes that catalyze the reversible conversion of oxaloacetate and malate. Oxaloacetate plays a crucial role in many metabolic pathways including the tricarboxylic acid cycle, amino acid synthesis, gluconeogenesis and maintenance of oxidation/reduction balance amongst others (Goward and Nicholls, 1994). Its suppression has been previously observed in *M. edulis* specimens exposed to marine pollutants, although the reasons are unclear (Manduzio et al., 2005).

In general, cellular responses to stress are very energy demanding processes (Fiol et al., 2006), thus activation of energy metabolism is essential to initiate the necessary processes to maintain homeostasis. Mussels constitute a special case in their response to extreme environmental parameters. Marine intertidal zones regularly subject gill-breathing animals to anoxia stress and hyperosmotic conditions. To cope with regular exposure to harsh conditions, intertidal ectotherms present an adaptive response which involves the production and conservation of ATP via anaerobic pathways and metabolic rate depression (Brooks and Storey, 1997). The ability to utilize both aerobic and anaerobic sources for energy metabolism is termed facultative anaerobiosis. It must overcome the fundamental problems of conservation of fermentable substrate and self-pollution by anaerobic end products.

Glycogen depletion is minimized in *Mytilus* by the increased efficiency of fermentation pathways and, most notably, the depression of ATP turnover rates (Hochachaka, 1986). In this respect, there would be no *a priori* expectation for glycolytic enzymes to increase in expression, whereas some increase in enzymes involved in Krebs cycles could have been expected.

The results observed here ought to be interpreted with care and further conclusions cannot be drawn as increased expression of glycolytic enzymes is not the unique factor controlling glycolysis. Reversible protein phosphorylation to control the activation of glycolytic enzymes, changes in fructose 2,6-bisphosphate concentrations to regulate carbohydrate utilization, alterations in allosteric control of regulatory enzymes through changes in metabolite concentrations, and changes in the proportion of enzymes associated with subcellular structures are also used to regulate energy producing pathways (Brooks and Storey, 1997). In addition, treatment conditions did not mimic those under which the use of anaerobic metabolic pathways in mussels has been described. Facultative anaerobiosis in intertidal species constitutes an adaptive response to intermittent aerial exposure, and it is possible that responses to elevated water temperatures differ. Although exact reasons behind the observed patterns in changes in expression of these proteins are unclear, they are likely to reflect an alteration of aerobic metabolism during acute exposure to extreme temperatures. It is not possible to determine whether these changes constitute an adaptive response to heat stress characteristic of facultative anaerobes, or non adaptive changes resulting from the conditions and intensity of the treatment.

Other proteins whose expression changed significantly in response to heat stress in *M. edulis* were S-adenosylhomocysteine hydrolase and a member of the glycosyltransferase superfamily. S-adenosylhomocysteine hydrolase was upregulated in both treatments in relation to the control, although changes in expression were only significant at 37°C. S-adenosylmethionine (AdoMet) is a universal methyl donor in a variety of biological functions. S-adenosylhomocysteine (AdoHcy) is formed after the donation of a methyl group from AdoMet to a methyl acceptor and it is subsequently hydrolyzed to adenosine and L-homocysteine by AdoHcy hydrolase (Chiang et al., 1996). Methylation of

proteins, phospholipids and nucleic acids results in a vast number of functions. The biological function in which these enzymes are involved cannot be determined.

Glycosyltransferase was significantly downregulated in the 27°C treatment and returned to control levels in the 37°C treated animals. Glycosyltransferases catalyze the transfer of a sugar moiety from an activated donor sugar onto saccharide and nonsaccharide acceptors. They are classified in many families, reflecting the variety of molecules that can be used as acceptors and which are involved in a number of functions (Breton et al., 2006). A member of this superfamily, dolichyl-diphosphooligosaccharide-protein glycosyltransferase, involved in N-glycosylation, was induced by temperature at the transcript level. However, the protein putatively identified here as a glycosyltransferase was downregulated at the protein level, suggesting a different behavior for different members of the family. The specific function of this enzyme cannot be determined.

4.5 SUMMARY

The number of proteins showing statistically significant changes in expression after exposure to elevated temperatures was relatively small in both species. Moreover, the proportion of identified proteins was low. The limitation in the number of identified proteins is typical for proteomic studies in non model organisms, which face unique challenges minimised when working with model species. The identification of a greater number of proteins would have undoubtedly increased the success of this study in terms of the identification of candidate proteins that could potentially function as biomarkers of stress. Nonetheless, some of the changes in energy metabolism observed in both species confirm those observed at the mRNA level. In addition, the chaperonin TCP-1 has been identified as a potential substitute for HSPs in *L. elliptica*. Although its induction may be an indirect response to heat stress, their role in the stress response of Antarctic species deserves further attention.

Some differences were observed between species in the effect of temperature stress in biological variation. There is some indication that the increased biological variation in *M. edulis* in stress induced, which could indicate that stress responsive proteins present higher underlying variation in *M. edulis* than in *L. elliptica*. In order to draw further conclusions, comparative studies of the genetic and phenotypic variability underlying changes in traits associated with the response to elevated environmental temperatures are necessary. These will likely provide information on the intensity of selection acting in each environment, and help to quantify the susceptibility of each species to climate change.

4.6 REFERENCES

Aaronson RM, Graven KK, Tucci M, McDonald RJ, Farber HW (1995) Non-neuronal enolase is an endothelial hypoxic stress protein. Journal of Biological Chemistry 270:27752-27757

Akaike H (1974) New look at statistical model identification. IEEE Transactions on Automatic Control 19:716-723

Anestis A, Lazou A, Pörtner HO, Michaelidis B (2007) Behavioral, metabolic, and molecular stress responses of marine bivalve *Mytilus galloprovincialis* during long-term acclimation at increasing ambient temperature. American Journal of Physiology-Regulatory Integrative and Comparative Physiology 293:R911-R921

Apraiz I, Mi J, Cristobal S (2006) Identification of proteomic signatures of exposure to marine pollutants in mussels (*Mytilus edulis*). Molecular & Cellular Proteomics 5:1274-1285

Badyaev A (2005) Stress-induced variation in evolution: from behavioural plasticity to genetic assimilation. Proceedings of the Royal Society B 272:877-886

Benjamini Y, Hochberg Y (1995) Controlling the false discovery rate - a practical and powerful approach to multiple testing. Journal of the Royal Statistical Society series B-methodological 57:289-300

Biron DG, Brun C, Lefevre T, Lebarbenchon C, Loxdale HD, Chevenet F, Brizard JP, Thomas F (2006) The pitfalls of proteomics experiments without the correct use of bioinformatics tools. Proteomics 6:5577-5596

Bolstad BM, Irizarry RA, Astrand M, Speed TP (2003) A comparison of normalization methods for high density oligonucleotide array data based on variance and bias. Bioinformatics 19:185-193

Breton C, Snajdrová L, Jeanneau C, Koca J, Imberty A (2006) Structures and mechanisms of glycosyltransferases. Glycobiology 16:29R-37R

Brooks SPJ, Storey KB (1997) Glycolytic controls in estivation and anoxia: A comparison of metabolic arrest in land and marine molluscs. Comparative Biochemistry and Physiology a-Molecular & Integrative 118:1103-1114

Buckley BA, Somero GN (2009) cDNA microarray analysis reveals the capacity of the cold-adapted Antarctic fish *Trematomus bernacchii* to alter gene expression in response to heat stress. Polar Biology 32:403-415

Buckley BA, Gracey AY, Somero GN (2006) The cellular response to heat stress in the goby *Gillichthys mirabilis*: a cDNA microarray and protein-level analysis. Journal of Experimental Biology 209:2660-2677

Cagney G, Park S, Chung C, Tong B, O'Dushlaine C, Shields DC, Emili A (2005) Human tissue profiling with multidimensional protein identification technology. Journal of Proteome Research 4:1757-1767

Carpentier SC, Witters E, Laukens K, Deckers P, Swennen R, Panis B (2005) Preparation of protein extracts from recalcitrant plant tissues: An evaluation of different methods for two-dimensional gel electrophoresis analysis. Proteomics 5:2497-2507

Challapalli KK, Zabel C, Schuchhardt J, Kaindl AM, Klose J, Herzel H (2004) High reproducibility of large-gel two-dimensional electrophoresis. Electrophoresis 25:3040-3047

Chang JS, Van Remmen H, Ward WF, Regnier FE, Richardson A, Cornell J (2004) Processing of data generated by 2-dimensional gel electrophoresis for statistical analysis: Missing data, normalization, and statistics. Journal of Proteome Research 3:1210-1218

Chapple JP, Smerdon GR, Berry RJ, Hawkins AJS (1998) Seasonal changes in stress-70 protein levels reflect thermal tolerance in the marine bivalve *Mytilus edulis* L. Journal of Experimental Marine Biology and Ecology 229:53-68

Chiang P, Gordon R, Tal J, Zeng G, Doctor B, Pardhasaradhi K (1996) S-adenosylmethionine and methylation. FASEB 10:471-480

Crawley MJ (2007) The R Book. John Wiley, New York. 942pp.

Davenport J, Davenport J (2005) Effects of shore height, wave exposure and geographical distance on thermal niche width of intertidal fauna. Marine Ecology-Progress Series 292:41-50

Diz AP, Skibinski DOF (2007) Evolution of 2-DE protein patterns in a mussel hybrid zone. Proteomics 7:2111-2120

Fiol DF, Chan SY, Kultz D (2006) Identification and pathway analysis of immediate hyperosmotic stress responsive molecular mechanisms in tilapia (*Oreochromis mossambicus*) gill. Comparative Biochemistry and Physiology D-Genomics and Proteomics 1:344-356

Forer A (2008) The diversity of cytoplasmic microtubules. Journal of Cell Science 121:7-9

Foth BJ, Zhang N, Mok S, Preiser PR, Bozdech Z (2008) Quantitative protein expression profiling reveals extensive post-transcriptional regulation and post-translational modifications in schizont-stage malaria parasites. Genome Biology 9:R177

Gardeström J, Elfwing T, Lof M, Tedengren M, Davenport JL, Davenport J (2007) The effect of thermal stress on protein composition in dogwhelks (*Nucella lapillus*)

under non-normoxic and hyperoxic conditions. Comparative Biochemistry and Physiology a-Molecular and Integrative 148:869-875

Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Functional Ecology 21:394-407

Gienapp P, Teplitsky C, Alho JS, Mills JA, Merilä J (2008) Climate change and evolution: disentangling environmental and genetic responses. Molecular Ecology 17:167-178

Goward CR, Nicholls DJ (1994) Malate dehydrogenase: a model for structure, evolution, and catalysis. Protein Science 3:1883-1888

Hack CJ, Lopez JA (2004) An exploration of some factors affecting the correlation of mRNA and proteomic data. Knowledge Exploration in Life Science Informatics, Proceedings 3303:9-19

Hanash S (2003) Disease proteomics. Nature 422:226-232

Hartl FU (1996) Molecular chaperones in cellular protein folding. Nature 381:571-580

Heukeshoven J, Dernick R (1985) Simplified method for silver staining of proteins in polyacrylamide gels and the mechanism of silver staining. Electrophoresis 6:103-112

Hochachaka PW (1986) Defense strategies against hypoxia and hypothermia. Science 231:234-241

Hoffmann AA, Hallas RJ, Dean JA, Schiffer M (2003) Low potential for climatic stress adaptation in a rainforest Drosophila species. Science 301:100-102

Jonsson H, Schiedek D, Grosvik BE, Goksoyr A (2006) Protein responses in blue mussels (Mytilus edulis) exposed to organic pollutants: A combined CYP-antibody/proteomic approach. Aquatic Toxicology 78:S49-S56

Kellermann V, Van Heerwaarden B, Sgro CM, Hoffmann AA (2009) Fundamental evolutionary limits in ecological traits drive drosophila species distributions. Science 325:1244-1246

Khoudoli GA, Porter IM, Blow JJ, Swedlow JR (2004) Optimisation of the two-dimensional gel electrophoresis protocol using the Taguchi approach. Proteome Sci. 2:6

Kültz D (2005) Molecular and evolutionary basis of the cellular stress response. Annual Review of Physiology 67:225-257

Kültz D, Fiol D, Valkova N, Gomez-Jimenez S, Chan SY, Lee J (2007) Functional genomics and proteomics of the cellular osmotic stress response in `non-model' organisms. Journal of Experimental Biology 210:1593-1601

Kussmann M, Nordhoff E, RahbekNielsen H, Haebel S, RosselLarsen M, Jakobsen L, Gobom J, Mirgorodskaya E, KrollKristensen A, Palm L, Roepstorff P (1997) Matrix-assisted laser desorption/ionization mass spectrometry sample preparation techniques designed for various peptide and protein analytes. Journal of Mass Spectrometry 32:593-601

Lande R, Shannon S (1996) The role of genetic variation in adaptation and population persistence in a changing environment. Evolution 50:434-437

Lapko VN, Smith DL, Smith JB (2000) Identification of an artifact in the mass spectrometry of proteins derivatized with iodoacetamide. Journal of Mass Spectrometry 35:572-575

Le Roch KG, Johnson JR, Florens L, Zhou YY, Santrosyan A, Grainger M, Yan SF, Williamson KC, Holder AA, Carucci DJ, Yates JR, Winzeler EA (2004) Global analysis of transcript and protein levels across the *Plasmodium falciparum* life cycle. Genome Research 14:2308-2318

Lee J, Valkova N, White MP, Kultz D (2006) Proteomic identification of processes and pathways characteristic of osmoregulatory tissues in spiny dogfish shark (*Squalus acanthias*). Comparative Biochemistry and Physiology D-Genomics and Proteomics 1:328-343

Liang P, MacRae TH (1997) Molecular chaperones and the cytoskeleton. Journal of Cell Science 110:1431-1440

Llorca O, McCormack EA, Hynes G, Grantham J, Cordell J, Carrascosa JL, Willison KR, Fernandez JJ, Valpuesta JM (1999) Eukaryotic type II chaperonin CCT interacts with actin through specific subunits. Nature 402:693-696

Lopez JL (2005) Role of proteomics in taxonomy: the *Mytilus* complex as a model of study. Journal of Chromatography B-Analytical Technologies in the Biomedical 815:261-274

Lopez JL (2007) Two-dimensional electrophoresis in proteome expression analysis. Journal of Chromatography B-Analytical Technologies in the Biomedical 849:190-202

Lopez JL, Mosquera E, Fuentes J, Marina A, Vazquez J, Alvarez G (2001) Two-dimensional gel electrophoresis of *Mytilus galloprovincialis*: differences in protein expression between intertidal and cultured mussels. Marine Ecology-Progress Series 224:149-156

Lopez J, Marina A, Vazquez J, Alvarez G (2002) A proteomic approach to the study of the marine mussels *Mytilus edulis* and *M. galloprovincialis*. Marine Biology 141:217-223

Manduzio H, Cosette P, Gricourt L, Jouenne T, Lenz C, Andersen OK, Leboulenger F, Rocher B (2005) Proteome modifications of blue mussel (*Mytilus edulis* L.) gills as an effect of water pollution. Proteomics 5:4958-4963

Martín-Benito J, Boskovic J, Gómez-Puertas P, Carrascosa JL, Simons CT, Lewis SA, Bartolini F, Cowan NJ, Valpuesta JM (2002) Structure of eukaryotic prefoldin and of its complexes with unfolded actin and the cytosolic chaperonin CCT. EMBO 21:6377-6386

Merilä J (2009) Genetic Constraints on Adaptation? Science 325:1212-1213

Moran MD (2003) Arguments for rejecting the sequential Bonferroni in ecological studies. Oikos 100:403-405

Nunn B, Timperman A (2007) Marine proteomics. Marine Ecology Progress Series 332:281-289

Ong SE, Pandey A (2001) An evaluation of the use of two-dimensional gel electrophoresis in proteomics. Biomolecular Engineering 18:195-205

Peck LS, Pörtner HO, Hardewig I (2002) Metabolic demand, oxygen supply, and critical temperatures in the Antarctic bivalve *Laternula elliptica*. Physiologial and Biochemical Zoology 75:123-133

Peck LS, Webb KE, Bailey DM (2004) Extreme sensitivity of biological function to temperature in Antarctic marine species. Functional Ecology 18:625-630

Pigliucci M, Murren CJ, Schlichting CD (2006) Phenotypic plasticity and evolution by genetic assimilation. Journal of Experimental Biology 209:2362-2367

Pinheiro JC, Bates DM (2000) Mixed-effects models in S and S-PLUS. New York: Springer 528pp.

Podrabsky JE, Somero GN (2004) Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus*. Journal of Experimental Biology 207:2237-2254

R Development Core Team (2008) R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Rodriguez-Ortega MJ, Grosvik BE, Rodriguez-Ariza A, Goksoyr A, Lopez-Barea J (2003) Changes in protein expression profiles in bivalve molluscs (*Chamaelea gallina*) exposed to four model environmental pollutants. Proteomics 3:1535-1543

Solomon S, et al. (2007) Climate change 2007: the physical science basis. Contribution of working group I to the fourth assessment report of the Intergovernmental Panel on Climate Change. Cambridge University Press, Cambridge

Stevens GC (1989) The latitudinal gradient in geographical range: how so many species coexist in the tropics. American Naturalist 133:240-256

Tracy MR, Hedges SB (2000) Evolutionary history of the enolase gene family. Gene 259:129-138

Trent JD, Nimmesgern E, Wall JS, Hartl FU, Horwich AL (1991) A molecular chaperone from a thermophilic archaebacterium is related to the eukaryotic protein T-complex polypeptide-1. Nature 354:490-493

Tunsjo HS, Paulsen SM, Mikkelsen H, L'Abee-Lund TM, Skjerve E, Sorum H (2007) Adaptive response to environmental changes in the fish pathogen *Moritella viscosa*. Research in Microbiology 158:244-250

Verhoeven KJF, Simonsen KL, McIntyre LM (2005) Implementing false discovery rate control: increasing your power. Oikos 108:643-647

Wilkins MR, Pscuali C, Appel RD, Ou K, Golaz O, Sanchez JC, Yan JX, Godey AA, Hughes G, Humphery Smith I et al. (1996) From proteins to proteomes: Large scale protein identification by two-dimensional electrophoresis and amino acid analysis. Bio-Technology 14:61-65

Williams RC, Shah C, Sackett D (1999) Separation of Tubulin Isoforms by Isoelectric Focusing in Immobilized pH Gradient Gels. Analytical Biochemistry 275:265-267

Yaffe MB, Farr GW, Miklos D, Horwich AL, Sternlicht ML, Sternlicht H (1992) TCP-1 complex is a molecular chaperone in tubulin biogenesis. Nature 358:245-248

Zivy M, de Vienne D (2000) Proteomics: a link between genomics, genetics and physiology. Plant Molecular Biology 44:575-580

5 STRESS HORMONES IN MOLLUSCS

5.1 INTRODUCTION

5.1.1 Stress hormones

The stress response is triggered in an individual when changes in the external environment occur, beyond the window to which individuals are adapted. These changes are detected by sensory organs which trigger a response in the neuroendocrine system including the release of hormones. These in turn induce secondary responses aimed at maintaining homeostasis, which are part of the stress response. Stress hormones are therefore potentially useful tools for assessing the physiological status of bivalves in their natural habitat, under laboratory conditions and in aquaculture. A better understanding of the role of such hormones in invertebrates is thus relevant not only in the study of heat stress, but other stressors such as pollution. These could have potential as markers in ecotoxicology. Two major systems are stimulated in vertebrates in response to stress: the pituitary-adrenal axis, and the sympathetic nervous system. The former pathway involves neuropeptides and steroids as major messengers and the latter involves catecholamines (CA) (Chrousos and Gold, 1992).

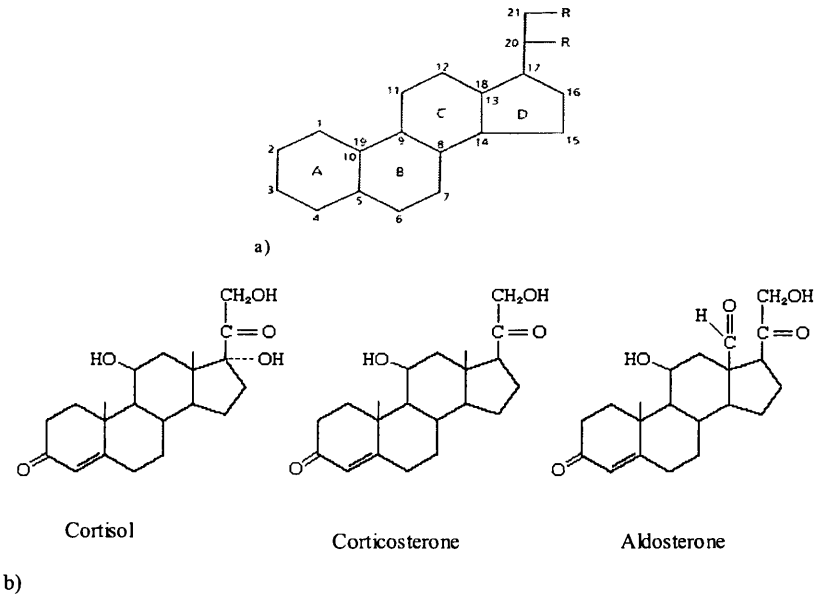


Figure 5.1 a) Configuration of the basic corticosteroid structure. R indicates functional groups in the side chain; b) Structure of some abundant vertebrate corticosteroid hormones.

The steroids synthesised by the adrenal cortex in mammals, are classified on the basis of their biological function as glucocorticoids (e.g. cortisol, corticosterone), mineralocorticoids (e.g. aldosterone) and sex steroids (e.g. estrogen, testosterone). The first two are also known as corticosteroids. Based on the number of carbons in the molecule, corticosteroids are classified as C_{21} steroids. The sex steroid progesterone, a precursor of corticosteroids and produced in the gonad, is also a C_{21} steroid, while androgens and estrogens are C_{19} and C_{18} steroids respectively. This study focuses on corticosteroids, whose basic structure, metabolism in mammals and relationship to other steroids is simplified in figures 5.1 and 5.2. The basic structure of a C_{21} steroid is perhydrocyclopenytanephenantrene, which consists of four rings (labelled A-D), three hexanes and one pentane, and a side chain (figure 5.1.a). Oxidations and reductions at specific sites give rise to a wide variety of hormones, as illustrated in figure 5.1.b.

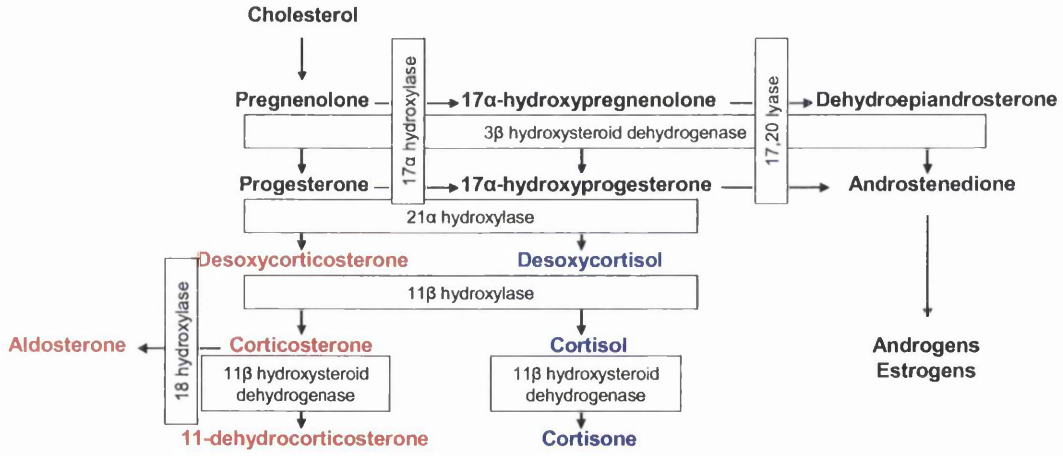


Figure 5.2 Simplified pathways of corticosteroid biosynthesis in mammals. Mineralocorticoids and glucocorticoids are indicated in red and blue respectively. Main enzymes are shown in boxes.

Cholesterol, a 27 carbon compound, is the basic precursor of steroid hormones. Metabolism of cholesterol leads to a reduction in the number of carbons to produce steroids. The first step converts cholesterol to pregnenelone and the related compound progesterone, reducing the number of carbons to 21. Progesterone is an important intermediate in the synthesis of glucocorticoids, mineralocorticoids and

sex steroids. The former two groups undergo a series of hydroxylations that give rise to several compounds without reducing the number of carbons (e.g. 17-hydroxylated steroids such as cortisol or cortisone). Androgenic hormones, however, are produced first by 17-hydroxylation followed by removal of two carbons reducing the number to 19. Estrogenic molecules (18 carbons) are produced from androgens by aromatisation of the steroid A-ring and loss of a methyl group attached to C10 (loss of C19) (Bird and Clark, 1979). A simplified diagram of the pathways for steroid biosynthesis in mammals is shown in figure 5.2. The exact structure of these compounds in invertebrates has not been fully characterised.

The mechanism that controls the synthesis of corticosteroids and initiates hormonal signalling is well understood in mammals. The stimuli that cause stress induce the release of corticotrophin releasing hormone (CRH) by the hypothalamus, which provokes corticotrophin (ACTH) release from the pituitary. This enters the bloodstream and reaches the adrenal gland, where it binds specific receptors causing the release of steroid hormones, which in turn are involved in regulating the synthesis and release of CA by the adrenal glands (Johnson et al., 1992). Since invertebrates lack the organs related to the stress response in vertebrates (hypothalamus, pituitary and adrenal glands), the mechanisms of their neuroendocrine stress response remain unknown. However, invertebrates are able to biosynthesise at least some molecules of the stress response comparable to those observed in mammals: CRH, ACTH and CA.

Evidence for the synthesis of CRH and ACTH-like molecules, and their involvement in the stress response, has been provided in molluscs, annelids and insects (reviewed in Ottaviani and Franceschi, 1996). The presence of CA is well documented in many invertebrate phyla including arthropods (Klemm, 1983; Shimizu et al., 1991), cnidaria (Pani and Anctil, 1994), annelids (Anctil et al., 1990) and molluscs (Sloley et al., 1990; Voronezhskaya et al., 1999; Pani and Croll, 2000; Cao et al., 2007). Little is known, however, about corticosteroid production in invertebrates. Immunoreactive corticosteroids have been recorded in several species of insects by radioimmunoassay (Bradbrook et al., 1990). However, the authors detected similar levels in insect food samples, thus the possibility that corticosteroids are sequestered from their food could not be excluded. Both cortisol and corticosterone have been

recorded in the insect *Calliphora vicina* by autoradiography (Bidmon and Stumpf, 1991) and cortisol has been detected in immunocytes from molluscs using immunocytochemistry (Ottaviani et al., 1998). None of these cortisol-like molecules have been characterised in invertebrates, and their endogenous synthesis has not been confirmed (Hooper et al., 2007). Therefore, the report of corticosteroid-like molecules in invertebrates needs further investigation.

5.1.2 Aim

The aim of this study was to determine whether molluscs synthesise corticosteroids, the classical stress hormones in vertebrates in response to heat stress. A combination of radioimmunoassays and thin layer chromatography was used for this purpose, and the potential biosynthesis of corticosteroids in several *Mytilus galloprovincialis* tissues was tested using radioactive precursors *in vitro*.

5.2 MATERIALS AND METHODS

5.2.1 Sample collection

L. elliptica and *M. edulis* specimens were collected, acclimated to laboratory conditions and heat shocked as described in chapter two. Haemolymph samples were in all cases <1 ml, thus preliminary experiments were performed using *Mytilus galloprovincialis* samples from Faro. Twenty *M. galloprovincialis* specimens were collected at different stages during a tidal cycle. Ten mussels were collected from the upper shore and ten from the lower shore at Ilha de Faro (36°59′26.16″N, 7°58′9.55″W) in June 2007. Mussels were found attached to the base of the pillars of a bridge over Ria Formosa, a tidal lagoon with easy access. In order to obtain mussels in putatively different stress conditions, collections were carried out at low tide ensuring that those from the upper shore had been emmersed for approximately 6 h (stressed). *M. galloprovincialis* specimens collected from the lower shore were permanently immersed (not stressed). Collections were repeated on two consecutive days. Haemolymph was extracted from each individual mussel as described in chapter two.

M. galloprovincialis specimens were used for the optimization of the sample preparation method, corticosteroid radioimmunoassays (RIA) and *in vitro* steroidogenesis. Corticosteroid radioimmunoassays were also carried out in *L. elliptica* and *M. edulis*. The general RIA procedure is described below, followed by a description of the experiments carried out.

5.2.2 Radioimmunoassay procedure

The radiolabelled hormones used in this study were [1,2,6,7,^{3}H]-cortisol (specific activity 2.74 TBq mmol^{-1}), [1,2(n)-^{3}H]-cortisone (specific activity 1.78 TBq mmol^{-1}) and [1,2,6,7,^{3}H]-corticosterone (specific activity 2.78 TBq mmol^{-1}) all purchased from Amersham, GE Healthcare Ltd. Rabbit anti-cortisone and sheep anti-corticosterone polyclonal antibodies were purchased from Chemicon International.

Anti-cortisol polyclonal antibody was purchased from Fitzgerald Industries International. The cross-reactivity of the antisera is presented in table 5.1.

Radioimmunoassays were carried out in duplicate tubes of equal final volume and results were expressed as mean values from both replicates. Preparations were carried out on ice. For each assay, a standard curve was produced using cold hormone standards at known concentrations (500, 250, 125, 100, 50, 25, 10, 5, 2.5, 2, 1, 0.5 pg $tube^{-1}$) and constant radiolabelled tracer and antibody concentrations. A blank tube containing only the tracer, a zero hormone (maximum binding) containing tracer and antibody, and total counts tube containing only tracer (to which activated charcoal was not added) were also run for each individual assay. Unknowns contained the sample, tracer and antibody.

Table 5.1 Percentage cross-reactions of selected steroids with the three antisera used in the RIAs in this study. Data obtained from the manufacturers.

	Antibody		
	Anti-cortisone	Anti-corticosterone	Anti-cortisol
Cortisol	3.4	0.12	—
Cortisone	—	<0.01	10
11-Deoxycortisol	0.6	—	54
11-Dehydrocorticosterone	0.6	0.67	—
Tetrahydrocortisone	0.2	—	—
Cortisoterone	<0.1	—	—
Deoxycorticosterone	<0.1	1.5	—
Pregesterone	<0.1	0.004	—
Aldosterone	<0.1	0.2	—
Allotetrahydrocortisol	<0.01	—	—
Tetrahydrocortisol	<0.01	—	—
Cortol	<0.01	—	—
Cortolone	<0.01	—	—
11-Hydroxyetiochoranolone	<0.01	—	—
18-Hydroxydeoxycorticosterone	—	0.02	—
17,21-Dihydroxy-5-pregnan-3,11,20-trione	—	—	16
11,17,21-Trihydroxy-5-pregnan-3,20-dione	—	—	5
11-Hydroxytestosterone	—	—	0.05
Testosterone	—	—	<0.001

For each sample, duplicate 100 µl aliquots of heated haemolymph (100 µl haemolymph in 400-900 µl RIA buffer, containing 0.05 M Na_2PO_4, pH 7.6 and 1% gelatine) were transferred to polypropylene RIA tubes (12x75 mm). Label and antibody were mixed in appropriate concentrations. The label was added so that each sample contained ca 1200 cpm of tritium. The antibody concentration was selected so that maximum binding was 40-50% of the total, and 100 µl of the mixture

was added to all tubes. After overnight incubation at 4°C, 250 µl dextran-coated charcoal suspension (0.5 g charcoal, 0.05 g Dextran T-70 and 100 ml RIA buffer) was added to each tube to separate bound from unbound tracer. Samples were left on ice for 12 min after addition of charcoal to the first tube, and centrifuged at 4°C for 12 min at 2000 rpm. The supernatant was then poured into plastic scintillation vials, to which 4 ml scintillation fluid (Ecolite, ICN) was added and radioactivity was measured in a Beckman scintillation counter (Beckman).

5.2.3 Preparation of steroid extracts from *M. galloprovincialis*

Optimization of the preparation of extracts was carried out prior to analysis. Haemolymph samples were tested for the presence of free and conjugated steroids, and for interfering substances by denaturation. The procedure is outlined below.

5.2.3.1 Test for presence of conjugates

Free and conjugated steroids (sulphates and glucuronides) were extracted separately from 5 ml haemolymph pooled from a mixed population of six mussels containing emmersed and immersed *M. galloprovincialis* collected from Ilha de Faro, in order to determine what proportion was present in each form. Extractions were performed using solid phase Sep-Pak C18 cartridges (Waters Associates, Inc.) according to manufacturer instructions. Briefly, cartridges were equilibrated with two washes of 2 ml methanol, followed by two washes of 2 ml distilled water to remove any methanol excess. The haemolymph was then run through the cartridge, which was subsequently eluted with three washes of 1 ml ethanol and collected into a glass tube. At this stage, both free and conjugated steroids, if present, were contained within the eluate. The ethanol was evaporated in a heat block at 40°C using a constant flow of nitrogen (N_2(g)) to accelerate evaporation. In order to separate free steroids, 200 µl distilled water and 4 ml diethyl ether was added to the C18 eluate dry residue, which was then vortexed for 10 min and centrifuged for 5 min at 700 rpm to separate aqueous and organic phases. The aqueous phase was frozen by immersion of the tube containing the sample in N_2(l) for 7 sec, and the supernatant containing free steroids was transferred to a clean glass tube. The diethyl ether was evaporated in a

heat block at 40°C using a constant flux of N_2(g) to accelerate evaporation. After evaporation, 1 ml RIA buffer was added to the sample, which was stored at –20°C. After removing the supernatant, the aqueous phase was evaporated in a heat block at 40°C using a constant flux of N_2(g). After evaporation, 1 ml TFA:ethyl acetate (1:100) was added to the residue and left overnight in a water bath at 40°C. The solvent was then evaporated and 500 µl acetate buffer (0.1 M CH_3CO_2Na, pH 4) added. Hydrolysed sulphates were extracted with diethyl ether as before and stored. Remaining traces of diethyl ether in the acetate buffer were allowed to evaporate, and 10 µl of the enzyme β-glucuronidase (Sigma-Aldrich) was added to the sample, which was incubated overnight at 37°C to release glucuronides. A further extraction with diethyl ether recovered the free steroid moieties from the glucuronides which were transferred to a separate tube, the diethyl ether was evaporated, and 1 ml RIA buffer was added to the sample, which was frozen until further use. Free and conjugated steroids were analysed for cortisol immunoreactivity via RIA.

5.2.3.2 Test for presence of binding proteins and interference

In order to test for the presence of binding proteins, RIA for immunoreactive cortisol was performed on heat denatured samples compared to intact samples. Denatured samples were prepared as previously described. In parallel, a similar procedure was used without heating. Furthermore, to test for potential interference from substances in the haemolymph, different volumes (20 µl, 50 µl and 100 µl) of haemolymph in RIA buffer to a total volume of 1 ml were used. If no interfering substances were present the dilution curve should be parallel to the standard curve.

5.2.4 Corticosteroid immunoreactivity of *M. galloprovincialis* haemolymph

5.2.4.1 Radioimmunoassays

Emmersed and immersed *M. galloprovincialis* samples (n=10) from Ilha de Faro were examined for cortisol immunoreactivity. 100 µl of each haemolymph sample was mixed with 900 µl RIA buffer and denatured in a heat block at 80°C for 1 h. Samples were assayed for immunoreactive cortisol as previously described.

5.2.4.2 Thin-layer chromatography (TLC) scans

In order to detect and confirm the identity of potential corticosteroid immunoreactivity, 10 ml haemolymph pooled from a mixed population of samples were extracted using solid phase Sep-Pak C18 cartridges as previously described. After evaporation, the extract was resuspended in 70 µl dichloromethane and applied onto a lane of a silica gel TLC plate (Silica gel 60 F_{254}, Merck) along 4 µl tritiated cortisone and cortisol loaded on a separate lane. The plate was developed for 50 min in chloroform:methanol (47.5:2.5). Radioactivity was scanned using a radiochromatograph Bioscan (Lablogic) and the position where the tritiated cortisol and cortisone peaks were located in the plate was noted. Fractions (0.75 cm) were individually scraped off the silica gel strips (haemolymph lane) and eluted with 80:20 dicholoromethane:ethanol (3x1 ml). The solvent was evaporated and 1 ml RIA buffer was added to each fraction and frozen at –20°C until assayed for immunoreactive cortisol by RIA as described above.

5.2.5 Corticosteroid immunoreactivity of *M. edulis* and *L. elliptica* haemolymph

Control and heat stressed *M. edulis* and *L. elliptica* samples were also examined for cortisol, cortisone and corticosterone immunoreactivity. 100 µl of each haemolymph sample was mixed with 400 µl RIA buffer and denatured in a heat block at 80°C for 1 h. Samples were assayed for immunoreactive corticosteroids as previously described.

5.2.6 *In vitro* steroid metabolism in *M. galloprovincialis*

In order to test for possible endogenous production of corticosteroids in *M. galloprovincialis*, different mussel tissues (foot, mantle edge, gill, mantle (gonad), digestive tissue and haemolymph) were incubated with a radioactive precursor. Small fragments (5-10 mg) of each tissue or 200 µl haemolymph were incubated in individual wells in a sterile 24 well plate (Cellstar, Greiner bio-one), each of them containing 1 ml incubation media (132.95 mM NaCl, 3.40 mM $CaCl_2.2H_2O$, 4 mM

KCl, 0.98 mM $MgCl_2.6H_2O$, 30.00 mM NaOH, 40mM HEPES, 1.55 mM glucose) and 5 μCi of radiolabelled 17α-Hydroxy[1,2,6,7-^{3}H]-progesterone (17P) (70 Ci $mmol^{-1}$; Amersham Biosciences, UK). Incubations were carried out for 6 h at 20°C in an oxygenated atmosphere with constant gentle agitation. At the end of the incubation period free steroids were extracted from each incubate separately using solid phase C18 cartridges as previously described.

Steroid metabolites were separated by normal phase TLC as before, with steroid standards (10 μl cortisol, cortisone, desoxycorticosterone and 17P) added to each extract prior to loading them onto a silica plate. Chromatograms for each tissue were stored and edited using Winscan v 1.51 (Lablogic). The radioactive peaks were collected and eluted with dichloromethane:ethanol (80:20) as previously described for microchemical analysis. The solvent was evaporated and redissolved directly in acetic and chromic acid as outlined below.

Chromic acid oxidations were used to test for the presence of a corticosteroid nucleus in the metabolite products of tissue incubations. Peaks from different tissues with similar mobility were pooled and subject to chromic acid oxidation. For the reaction, 200 μl acetic acid (Merck) and 120 μl chromic acid (Aldrich- Sigma) diluted in distilled water (20 mg ml^{-1}) was added to each sample, which was covered to avoid exposure to light and left for 1 to 2 h. 1 ml 1% sodium metabisulfite and 1 mg sodium bicarbonate was then added to each sample in order to stop the reaction, followed by extraction twice with 4 ml diethyl ether as described before. After evaporation of diethyl ether, the dried organic extracts were resuspended in 70 μl dichloromethane and separated on TLC with chloroform as the mobile phase. Tritiated androstenedione and androstenotrione were used as standards.

5.3 RESULTS

5.3.1 Presence of free and conjugated cortisol immunoreactivity in *M. galloprovincialis* haemolymph

The highest proportion of immunoreactive cortisol detected in *M. galloprovincialis* haemolymph was found in a free unconjugated form (70.79%) with 14.67 and 14.54% present as sulphates and glucuronides respectively.

5.3.2 Test for presence of steroid binding proteins

Results for cortisol immunoreactivity obtained by denaturing the samples for 1h at 80°C differed from those obtained by assaying haemolymph previous to denaturation. All of the haemolymph samples that were not denatured were below the detection threshold. Immunoreactive cortisol in the same denatured samples was low with 33% of the samples below detection and estimates of immunoreactive cortisol of 9.47±0.20 pg ml^{-1} (mean±SE).

5.3.3 Corticosteroid immunoreactivity in *M. galloprovincialis* haemolymph

Levels of immunoreactive cortisol were below the detection threshold in 45 and 60% of the specimens from the emmersed and immersed populations respectively. For those samples that had detectable levels, concentrations were 14.45±2.65 pg ml^{-1} and 10.53±1.61 pg ml^{-1} (mean±SE) in the emmersed and immersed mussels respectively.

A TLC scan for cortisol immunoreactivity is shown in figure 5.3. Cortisol immunoreactivity was similar and at background level in all fractions analysed. There was no immunoreactivity peak coincident with any of the corticosteroid standards.

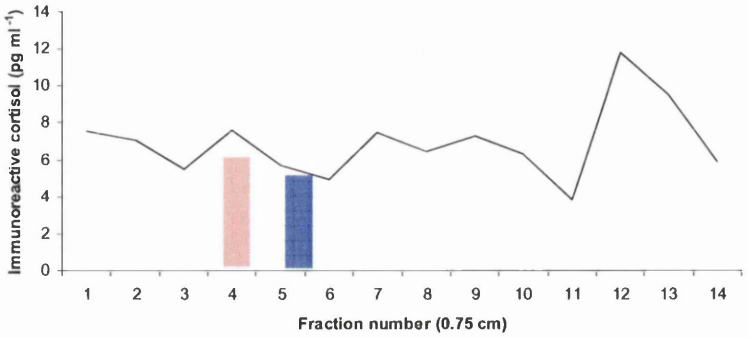


Figure 5.3 TLC scan of cortisol immunoreactivity in a pool of haemolymph from 20 *M. galloprovincialis* specimens. Red and blue shading shows the positions of tritiated cortisol and cortisone standards respectively.

5.3.4 Corticosteroid immunoreactivity of *M. edulis* and *L. elliptica* haemolymph

The possibility that corticosteroids could be produced by other bivalve species or under induced stress was studied by measuring immunoreactive cortisol, cortisone and corticosterone in *M. edulis* and *L. elliptica*. Over 90% of the measurements for immunoreactive corticosterone and cortisone were below detectable levels and thus these could not be analysed any further. Levels of immunoreactive cortisol were close to the detection limit of the assay. In *L. elliptica*, only one individual in the control treatment presented immunoreactive cortisol concentrations below the detection threshold. For those samples above the detection threshold, mean immunoreactive cortisol concentrations were 11.07±1.08, 13.04±2.97, and 13.02±1.95 pg ml^{-1} for control (0°C), 3°C and 9°C treated samples respectively (mean±SE). Immunoreactive cortisol concentrations in *M. edulis* haemolymph were below detection in one individual in the 27°C group and in 40% individuals in the 37°C group. Mean immunoreactive cortisol concentrations were 505.08±204.07, 11.65±4.90, and 6.52±0.44 pg ml^{-1} for control, 27°C and 37°C treated samples respectively (mean±SE). Individual cortisol measurements are shown in appendix II.

5.3.5 *In vitro* steroid metabolism in *M. galloprovincialis*

TLC chromatograms from *in vitro* incubations of different *M. galloprovincialis* tissues with 17P are shown in figure 5.4.

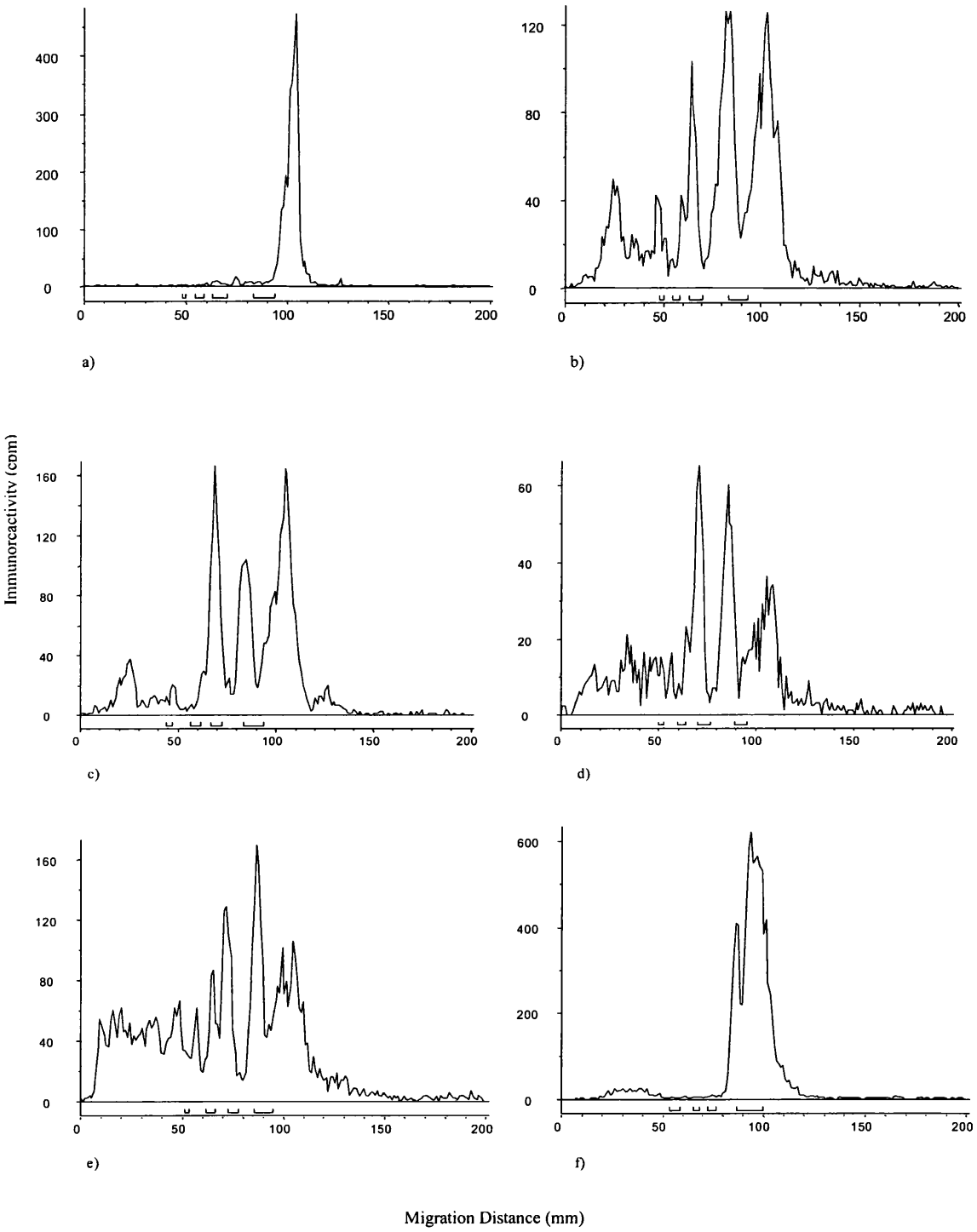


Figure 5.4 TLC scans of steroid immunoreactivities in extracts of *in vitro* incubates of *M. galloprovincialis* tissue corresponding with a) gill, b) foot, c) mantle edge, d) mantle (gonad), e) digestive tissue and f) haemolymph. Radiolabelled 17P was used as a precursor. The position of the standards cortisol, cortisone, corticosterone and 17P (from left to right) is indicated by four markers on the x axes.

In incubates from gill, the chromatogram showed a single peak, which was less polar than any of the standards or the precursor. Several peaks were observed in incubates from the gonad, but none coincided with the standards. In haemolymph, two peaks were observed, none of which coincided with the corticosteroid standards. In the case of the digestive tissue, peaks were observed in the regions of cortisone and corticosterone but these were not entirely coincident. In the foot incubates, peaks appeared to be within the cortisol and corticosterone regions, and in the mantle, one peak was coincident with corticosterone.

5.3.6 Analysis of metabolites in *M. galloprovincialis*

With 17P as precursor, the potential, more common products should be 17-hydroxylated C_{21} steroids, androgens or estrogens (figure 5.2). The typical corticosteroid nucleus when fully oxidized yields androgens with oxo groups at position 3, 11 and 17 (androstene-3,11,17-trione, see figure 5.5), while the progesterone nucleus (such as in 17P) yields androgens lacking oxygen at position 11 (Bush, 1961). Since these can be readily distinguished on TLC, oxidations with chromic acid were carried out on the TLC fractions above to search for evidence of a corticosteroid type steroid nucleus. The oxidized material was run on TLC alongside cold androstenetrione and androstenedione (Sigma-Aldrich).

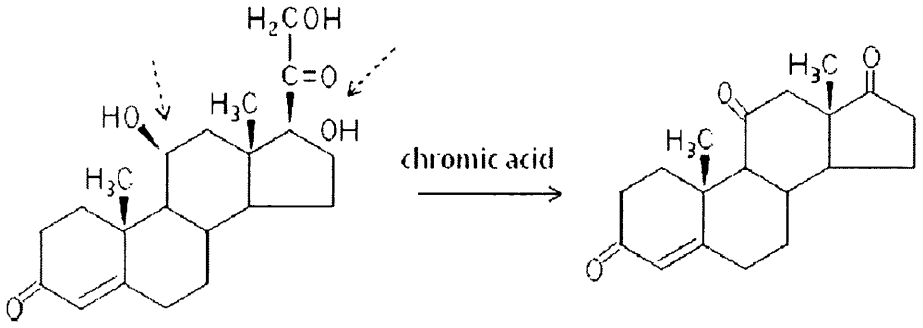


Figure 5.5 Conversion of cortisol to androstene-3,11,17-trione by chromic acid oxidation. The sites of oxidation are indicated by the broken arrows.

If products of chromic acid oxidation co-migrated with androstenetrione, the presence of a corticosteroid type steroid nucleus was possible, indicating the

presence of enzymes able to synthesize corticosteroids (e.g. 11β-hydroxylase. See figure 5.2). In contrast, the absence of a compound co-migrating with androstenetrione, or metabolites of close similarity indicates that the parent compound was not a corticosteroid. A compound co-migrating with androstenedione would indicate the presence of a similar nuclear structure to the precursor. TLC scans of oxidation products are shown in figure 5.6.

The fractions from gill (figure 5.6 a), when oxidized, yielded products less polar (i.e. migrated faster) than any of the standards. These compounds have a suggested androstenedione 5α/β-reduced configuration and there was no indication of the presence of a corticosteroid-type nucleus. The fractions from foot (figure 5.6 b) yielded more than one product, indicating a mixture. In both cases one product coincided with androstenedione and the other was less polar indicating they were not related to corticosteroids. The fractions from mantle (figure 5.6 c) gave three main co-migrating products. Interestingly, one of the products from fractions 77-88 and 100-113 matched andostenetrione suggesting that this tissue has the potential to produce corticosteroids. The fractions from haemolymph (figure 5.6 d) gave products matching androstenedione, excluding the possibility of a corticosteroid nucleus.

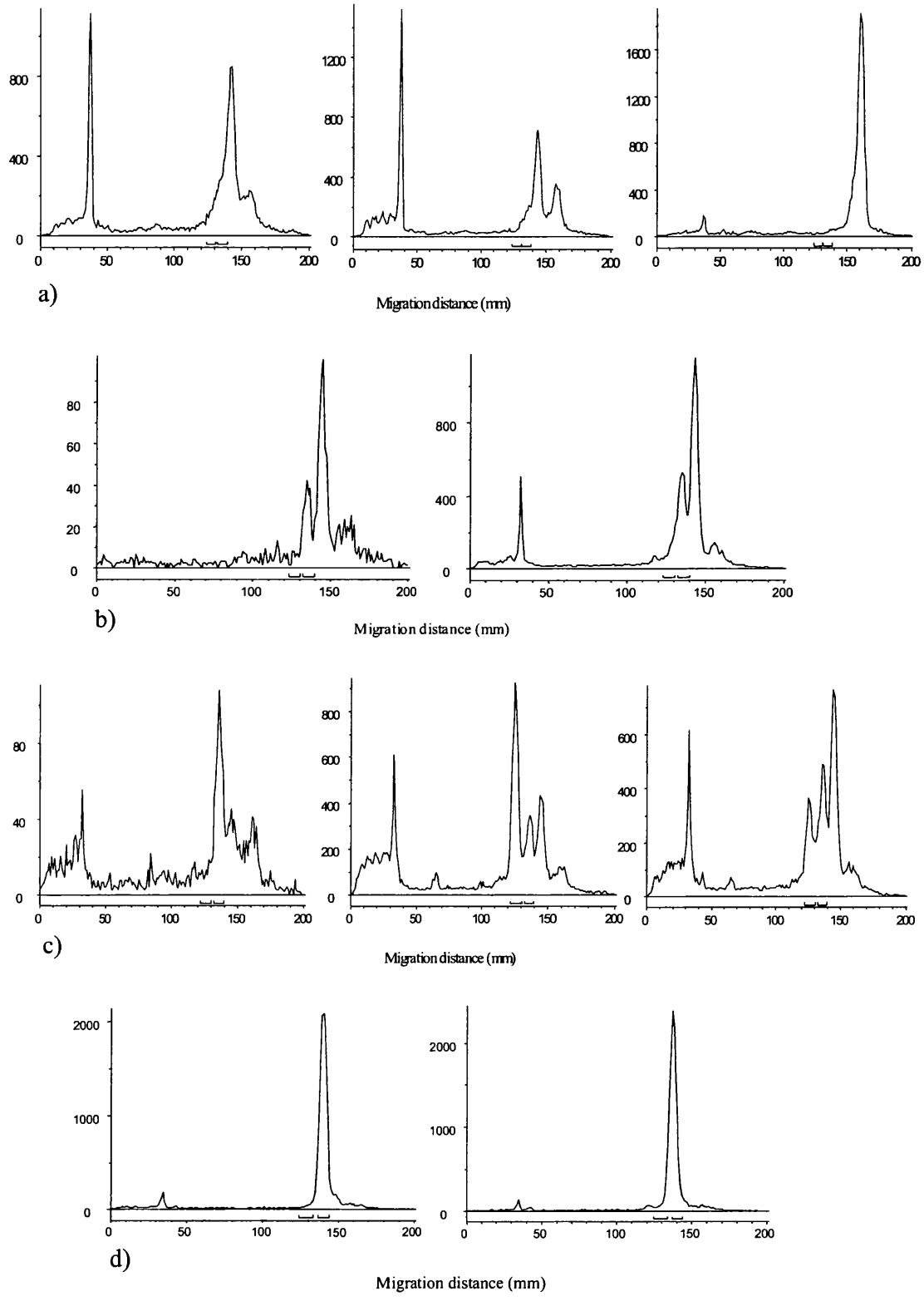


Figure 5.6 TLC scans of chromic acid products of fractions in figure 5.4 corresponding to a) gill, from left to right products of 62-69 mm fraction, 74-87 mm fraction and 100-115 mm fractions. Standards androstenedione and androstenetrione

ran on positions 132-139 mm and 124-131 mm respectively b) foot, from left to right products of 56-70 mm and 75-90 mm fractions. Standards androstenedione and androstenetrione ran on positions 133-140 mm and 123-130 mm respectively. c) mantle edge, from left to right products of fractions 65-74 mm, 77-88 mm and 100-113 mm. Standards were in the same position as in foot. d) haemolymph, from left to right products of fractions 79.5-89 mm and 89-107 mm. Standards androstenedione and androstenetrione ran on positions 137-145 mm and 126-134 mm respectively.

5.4 Discussion

In this study, a small amount of immunoreactivity was detected in bivalve haemolymph by a cortisol radioimmunoassay, although immunoreative fractions on TLC suggest that it is unlikely that the low level of immunoreactivity detected belongs to cortisol or any of the other corticosteroids tested: corticosterone and cortisone. However, the potential for corticosteroid biosynthesis cannot be ruled out since radioactive precursor studies with mantle originated a product which based on chromic acid oxidation and mobility is compatible with a corticosteroid.

5.4.1 Detection of corticosteroid immunoreactivity

In order to account for factors that could preclude the detection of steroids by immunoreactivity, haemolymph samples were analysed for the presence of conjugated immunoreactive corticosteroid and binding proteins. The proportion of putative conjugated corticosteroid was negligible compared to free steroids. Conjugation involves the addition of sulphate or glucuronide groups, which facilitates solubility and reduces the biological activity of the free steroid (Gomes et al., 2009). The results presented here suggest that immunoreactive putative steroids in *M. galloprovincialis* haemolymph are largely present in a free (unconjugated) form. The low level of conjugated immunoreactive corticosteroids detected, could indicate that they are rapidly excreted (Lowartz et al., 2003).

Several hormones have been documented to bind rapidly to vertebrate proteins (Powis, 1975). To explore the possibility that that corticosteroids were present bound to haemolymph proteins, haemolymph samples were denaturated to release steroids from endogenous binding proteins, freeing them to bind to their specific antibodies (Scott et al., 1982). Results obtained from assaying haemolymph samples directly, differed from those obtained when heating the samples. These results suggested that binding proteins might be present in bivalve haemolymph, which should be considered in future experiments.

5.4.2 Corticosteroids in bivalve molluscs

The endogenous synthesis of corticosteroid-like compounds by different *M. galloprovincialis* tissues was tested *in vitro*. TLC chromatograms from *in vitro* incubations did not provide conclusive evidence that corticosteroids were synthesised in any of the tissues analysed. 17P was not metabolized to cortisol, cortisone or corticosterone in the gill, gonad, haemolymph or digestive gland of *M. galloprovincialis* as suggested by the lack of co-migration with the standards. In the gill and haemolymph, the products were less polar than the precursor, making it unlikely that corticosteroids could be produced. Although there was no evidence that any of the analysed corticosteroids were produced in these tissues, the possibility of further metabolization of corticosteroids, if they were produced, could not be ruled out since smaller, more polar peaks were noticed. Interestingly, peaks coincident with the standards in foot and mantle incubate products, provided some indication that corticosteroid-like compounds could be produced in these tissues. The radioactively labelled compounds separated by TLC were, by definition, derivatives of the radioactively labelled precursor 17P. Therefore, although we may not know the structure of the metabolites detected in these tissues, these unidentified compounds are likely to be steroid-like hormones, or derivatives of steroids.

For confirmation, chromic acid oxidations were carried out on the TLC fractions from the incubation products. The results confirmed that haemolymph, gonad, foot and digestive gland of *M. galloprovincialis* cannot synthesize corticosteroids from the precursor 17P, which is largely metabolized to androstenedione or less polar compounds. The synthesis of corticosteroids through a different pathway cannot be ruled out. The presence of a compound co-migrating with androstenetrione in the mantle is interesting. However, this result was based only on a single micro-chemical reaction and co-migration on TLC and further work is required to confirm its structure. Unfortunately, time constraints precluded further work at the time. Although these results provide potentially interesting insights into the endocrinology of *M. galloprovincialis*, they do not confirm the presence of corticosteroids.

Although the results from the incubations indicated that there was no evidence of endogenous biosynthesis of the cortisol, cortisone or corticosterone in *M.*

galloprovincialis, the cortisol radioimmunoassays detected some level of immunoreactivity in the majority of the samples (14.45±2.65 and 10.53±1.61 pg ml^{-1} in emmersed and immersed *M. galloprovincialis*, 11.07±1.08, 13.04±2.97 and 13.02±1.95 pg ml^{-1} in control, 3°C and 9°C treated *L. elliptica*, and 505.08±204.07, 11.65±4.90 and 6.52±0.44 pg ml^{-1} in control, 27°C and 37°C treated *M. edulis* (mean±SE)). This suggests the presence of an alternative compound binding the antibody, whose concentration is highest in the *M. edulis* control group. It is possible that this compound is negatively correlated with stress, although further conclusions cannot be drawn. Moreover, concentrations are highly variable between individuals and results ought to be interpreted with care. In general, the concentration of this unidentified compound is low compared to the concentrations of other steroids reported in invertebrates. For example, reported levels of androgens and progesterone in tissues from most invertebrates are in the range 0.1-10 ng g^{-1} (reviewed in Janer and Porte, 2007). It is not possible to determine whether the cross-reacting compound was present at extremely low concentrations in the majority of the samples, or it presents a structure that the utilized extraction method could not separate efficiently in a consistent manner. There is no evidence that the detected compound presents a corticosteroid structure, or that it is synthesised endogenously.

It must be noted, that immunological techniques to detect steroids in invertebrates rely on antibodies raised against their known vertebrate variants, which may not be entirely appropriate (Lafont and Mathieu, 2007). It should also be kept in mind that steroids are always present in food items, thus it is also possible that the detected compound is indeed a corticosteroid-like hormone with dietary origin. To date, the presence of cortisol-like steroids in marine invertebrates has only been detected by Ottaviani et al. (1998) using an immunocytochemical approach, which is subjected to the limitations detailed above and does not provide evidence of the endogenous origin of the detected compounds. Combinations of *in vitro* steroidogenesis and mass spectrometric techniques are necessary to characterise invertebrate steroids, and ensure their endogenous nature.

Questions also remain concerning the presence of the enzymes and receptors necessary for the synthesis and signalling of corticosteroids in invertebrates. Estrogen and other vertebrate-type sex steroids have been extensively identified in

molluscs and their endogenous synthesis has been further confirmed by the presence of most of the enzymes required to produce them (reviewed in Janer and Porte, 2007). These include, amongst others, 3β-hydroxysteroid dehydrogenase and 17α-hydroxylase, also required in the synthesis of corticosteroids from the precursor pregnenolone (see figure 5.2). However, confirmation of the presence of other enzymes involved in costicosteroid metabolism such as 21α-hydroxylase or 11β-hydroxylase has not been provided in invertebrates. Likewise, to my knowledge, there is currently no evidence of the presence of corticosteroid receptors in invertebrates. Intracellular or cell surface receptors mediate the effects of steroids in vertebrates by acting as transcription factors modulating gene expression or leading to the production of more rapid non-genomic responses (Norman et al., 2004). Vertebrate genomes present nuclear steroid hormone receptors for estrogen (ERα and ERβ), androgens (AR), progestins (PR), glucocorticoids (GR) and mineralocorticoids (MR), which are absent from the main fully sequenced invertebrate models *D. melanogaster*, *C. elegans* and *C. intestinalis* (Thornton et al., 2003). Orthologs of the estrogen receptors, however, are present in other invertebrate phyla, highlighting some of the problems in drawing conclusions from model species to the arena of non model organisms. The presence of sex steroid receptors has been previously demonstrated by different methods in many bivalve species (reviewed in Croll and Wang, 2007). This provides further confirmation for the presence of sex steroids in bivalves. It is therefore unclear whether the lack of evidence of GR and MR in molluscs reflects limited efforts in the study of these hormones, or provides further evidence for their absence. New generation sequencing technologies and the rapid increase in the number of full sequenced genomes will provide unprecedented opportunities for the characterization, if present, of these receptors.

5.4.3 Stress hormones in marine invertebrates

Despite the apparent lack of corticosteroids, marine invertebrates are characterised by an otherwise very complex steroid metabolism, which varies between classes. Indeed, the more primitive the animal the greater the variety of steroids (Walton and Pennock, 1972). Steroid composition also varies significantly between classes, with lamellibranches containing complex mixtures of C_{26}, C_{27}, C_{28}, C_{29} and C_{30} sterols together with cholesterol (Kanazawa, 2001). In the mussel *M. edulis,* evidence for

the synthesis of cholesterol and cholesterol-like compounds from mevalonate has been provided by *in vitro* steroid metabolism (Teshima and Kanazawa, 1974). Given the complexity of the invertebrate sterol composition, it is likely that pathways for the synthesis of alternative stress hormones remain unknown. The use of vertebrate steroids as precursors, by ignorance of the metabolites utilized by invertebrates, potentially precludes the characterization of such pathways and their final products (Lafont and Mathieu, 2007). Nonetheless, examples of vertebrate-type hormones whose concentration varies in response to stress are available throughout the marine invertebrates, where catecholamine responses to abiotic stressors have been documented in previous studies. These include the increase in circulating noradrenaline and dopamine concentrations in response to stress caused by mechanical disturbances, and temperature and salinity variations in oysters (Lacoste et al., 2001) and increased hemolymphatic levels of noradrenaline and dopamine under mechanical stress in abalones (Malham et al., 2003).

Additionally, an increasing amount of evidence suggests that environmental pollutants can interfere with the reproductive endocrine system of many invertebrates (Janer and Porte, 2007). These endocrine disruptive chemicals are thought to interfere with steroid receptors, blocking the effects of endogenous hormones (DeRosa et al., 1998). The exact mechanisms of action, however, have not been elucidated due to the limited knowledge of the invertebrate endocrine system.

Based on current knowledge, it appears that a universal stress response does not occur at the endocrine level. There are important differences between taxa in the composition and mechanisms of action of stress related hormones. In molluscs, the presence of the "classic" stress hormones has not been conclusively demonstrated and it is possible that stressor specific responses occur. The behaviour of sex steroids could be modulated by chemical pollutants, whereas catecholamines could respond to other abiotic stressors such as temperature or salinity fluctuations. However, the information available on the endocrine responses to stress in molluscs is still limited and further conclusions cannot be made.

5.5 SUMMARY

The experiments presented here to demonstrate the endogenous nature of corticosteroids in mussels were inconclusive, but suggested that key enzymes for the production of corticosteroids from their immediate precursor, as described in vertebrates, are not present in these animals. Either they have an alternative biosynthetic pathway, or the detected corticosteroid-like compounds have a dietary origin. Although the possibility that corticosteroids are synthesised and metabolized further cannot be ruled out, the currently available data are not sufficient to confirm their endogenous production in bivalve molluscs.

5.6 REFERENCES

Anctil M, Dewaele JP, Miron MJ, Pani AK (1990) Monoamines in the nervous system of the tube worm *Chaetopterus variopedatus* (polychaeta): biochemical detection and serotonin immunoreactivity. Cell and Tissue Research 259:81-92

Bidmon HJ, Stumpf WE (1991) Uptake, distribution and binding of vertebrate and invertebrate steroid-hormones and time dependence of ponasterone A binding in *Calliphora vicina.* Comparisons among cholesterol, corticosterone, cortisol, dexamethasone, 5-alpha-dihydrotestosterone,1,25-dihydroxyvitamin D3, ecdysone, estradiol-17-beta, ponasterone-a, progesterone, and testosterone. Histochemistry 96:419-434

Bird CE and Clark AF (1973) The Adrenals. In: Ezrin C, Godden JO, Volpe R Wilson R (Eds.) Systematic Endocrinology. Harper and Row Publishers, London.

Bradbrook DA, Clement CY, Cook B, Dinan L (1990) The occurrence of vertebrate-type steroids in insects and a comparison with ecdysteroid levels. Comparative Biochemistry and Physiology B-Biochemistry & Molecular 95:365-374

Bush IE (1961) The Chromatography of Steroids. Elmsford, N.Y., Pergamon Press

Cao A, Ramos-Martinez JI, Barcia R (2007) In hemocytes from *Mytilus galloprovincialis* Lmk., treatment with corticotropin or growth factors conditions catecholamine release. International Immunopharmacology 7:1395-1402

Chrousos GP, Gold PW (1992) The concepts of stress and stress system disorders: overview of physical and behavioral homeostasis. Journal of the American Medical Association 267:1244-1252

Croll RP, Wang C (2007) Possible roles of sex steroids in the control of reproduction in bivalve molluscs. Aquaculture 272:76-86

DeRosa C, Richter P, Pohl H, Jones DE (1998) Environmental exposures that affect the endocrine system: Public health implications. Journal of Toxicology and Environmental Health Part B Critical Reviews 1:3-26

Gomes RL, Meredith W, Snape CE, Sephton MA (2009) Analysis of conjugated steroid androgens: deconjugation, derivatisation and associated issues. Journal of Pharmaceutical and Biomedical Analysis 49:1133-1140

Hooper C, Day R, Slocombe R, Handlinger J, Benkendorff K (2007) Stress and immune responses in abalone: Limitations in current knowledge and investigative methods based on other models. Fish & Shellfish Immunology 22:363-379

Janer G, Porte C (2007) Sex steroids and potential mechanisms of non-genomic endocrine disruption in invertebrates. Ecotoxicology 16:145-160

Johnson EO, Kamilaris TC, Chrousos GP, Gold PW (1992) Mechanisms of stress: a dynamic overview of hormonal and behavioral homeostasis. Neuroscience and Biobehavioral Reviews 16:115-130

Kanazawa A (2001) Sterols in marine invertebrates. Fisheries Science 67:997-1007

Klemm N (1983) Monoamine-containing neurons and their projections in the brain (supraesophageal ganglion) of cockroaches. An aldehyde fluorescence study. Cell and Tissue Research 229:379-402

Lacoste A, Malham SK, Cueff A, Poulet SA (2001) Stress-induced catecholamine changes in the hemolymph of the oyster *Crassostrea gigas*. General and Comparative Endocrinology 122:181-188

Lafont R, Mathieu M (2007) Steroids in aquatic invertebrates. Ecotoxicology 16:109-130

Lowartz S, Petkam R, Renaud R, Beamish FWH, Kime DE, Raeside J, Leatherland JF (2003) Blood steroid profile and in vitro steroidogenesis by ovarian follicles and testis fragments of adult sea lamprey, *Petromyzon marinus*. Comparative Biochemistry and Physiology - Part A: Molecular & Integrative Physiology 134:365-376

Malham SK, Lacoste A, Gelebart F, Cueff A, Poulet SA (2003) Evidence for a direct link between stress and immunity in the mollusc *Haliotis tuberculata*. Journal of Experimental Zoology Part A-Comparative Experimental Biology 295A:136-144

Norman AW, Mizwicki MT, Norman DPG (2004) Steroid-hormone rapid actions, membrane receptors and a conformational ensemble model. Nature Reviews Drug Discovery 3:27-41

Ottaviani E, Franceschi C (1996) The neuroimmunology of stress from invertebrate to man . Progress in Neurobiology 49:285-285

Ottaviani E, Franchini A, Franceschi C (1998) Presence of immunoreactive corticotropin-releasing hormone and cortisol molecules in invertebrate haemocytes and lower and higher vertebrate thymus. Histochemical Journal 30:61-67

Pani AK, Anctil M (1994) Evidence for biosynthesis and catabolism of monoamines in the sea pansy *Renilla koellikeri* (cnidaria). Neurochemistry International 25:465-474

Pani AK, Croll RP (2000) Catechol concentrations in the hemolymph of the scallop, *Placopecten magellanicus*. General and Comparative Endocrinology 118:48-56

Powis G (1975) Binding of catecholamines to human-serum proteins. Biochemical Pharmacology 24:707-712

Scott AP, Sheldrick E, Flint A (1982) Measurement of 17-alpha,20-beta-Dihydroxy-4-pregnen-3-one in plasma of trout (*Salmo gairdneri* Richardson): seasonal changes

and response to salmon pituitary extract. General and comparative endocrinology 46:444-451

Shimizu T, Mihara M, Takeda N (1991) High-performance liquid chromatography of biogenic amines in the corpus cardiacum of the american cockroach, *Periplaneta americana*. Journal of Chromatography 539:193-197

Sloley BD, Juorio AV, Durden DA (1990) High-performance liquid chromatographic analysis of monoamines and some of their gamma-glutamyl conjugates produced by the brain and other tissues of *Helix aspersa* (Gastropoda). Cellular and Molecular Neurobiology 10:175-192

Teshima SI, Kanazawa A (1974) Biosynthesis of sterols in abalone, *Haliotis gurneri*, and mussel, *Mytilus edulis*. Comparative Biochemistry and Physiology Part B: Biochemistry and Molecular Biology 47:555-561

Thornton JW, Need E, Crews D (2003) Resurrecting the ancestral steroid receptor: Ancient origin of estrogen signalling. Science 301:1714-1717

Voronezhskaya EE, Hiripi L, Elekes K, Croll RP (1999) Development of catecholaminergic neurons in the pond snail, *Lymnaea stagnalis*: I. Embryonic development of dopamine-containing neurons and dopamine-dependent behaviors. Journal of Comparative Neurology 404:285-296

Walton MJ, Pennock JF (1972) Some studies on biosynthesis of ubiquinone, isoprenoid alcohols, squalene and sterols by marine invertebrates. Biochemical Journal 127:471-479

6 Conclusions, Perspective and Future Directions

During the past decade, much has changed in our understanding of how biodiversity is responding in a changing world. We now have evidence that climatic changes are affecting species worldwide with consequences to the phenology and physiology of organisms, distributional ranges, species composition, and ecosystems structure and dynamics (Peñuelas and Filella, 2001; Walther et al., 2002; Parmesan and Yohe, 2003). It is clear that we need to be able to predict the effects of a warming climate in ecosystems worldwide and identify those species most susceptible to change, in order to concentrate conservation efforts. The aim of the present study was to analyse and compare the heat stress response of an Antarctic and a temperate marine ectotherm. The final goal was to increase our understanding of the relative susceptibility of Antarctic and temperate marine species in a changing climate. With this purpose, experiments were set up to identify the molecular mechanisms involved in the acute response to elevated temperatures in the eurythermal and highly stenothermal bivalves *M. edulis* and *L. elliptica*, with the understanding that acute exposure could be used as a proxy for climate change. Further aims included the generation of genomics resources for an Antarctic bivalve that could be used to study responses in closely related species; the identification of novel stress biomarkers in *L. elliptica*, and the study of how transcriptomics, proteomics, metabolomics and physiology can be linked to achieve a better understanding of species responses to change.

In the next sections, the major findings that have been presented throughout this document are evaluated in relation to the original aims. Conclusions from each chapter are not presented independently (refer to the summary section in each chapter) but their relevance in an environmental context is discussed. An assessment of the methodology is also presented and current and future work suggested.

6.1 CANDIDATE STRESS BIOMARKERS

The results in this study, as presented in chapter three, demonstrate that under the conditions examined, molecular responses can be used to identify when an organism is undergoing stress. The occurrence of a temperature triggered response in *L. elliptica* when exposed to elevated temperatures was observed at the transcript level. Molecular mechanisms characteristic of the stress response, such as changes in

protein turnover through reversible and irreversible protein damage and increased chaperoning activity and oxidative stress are activated in *L. elliptica* when exposed to 3°C.

The changes in expression in antioxidant genes reflect the declining capacity in aerobic systems with small elevations in temperature in *L. elliptica* and provide general support for the oxygen and capacity limitation hypothesis in response to acute heat challenges (c.f. Portner, 2002). Accordingly, markers for oxidative stress, rather than the universal heat shock proteins, have been proposed as candidates for heat stress biomarkers in *L. elliptica.* In this species, HSPs present higher threshold induction temperatures than the temperature at which the first signs of stress appear (Clark et al., 2008). The induction of antioxidant defences in *L. elliptica* exposed to 3°C suggests that these could be excellent candidates as heat stress markers in this species. The stress and tissue specificity of the different antioxidants needs to be studied to determine which might be most appropriate as heat stress markers in *L. elliptica.*

Calcium and calcium signalling related genes and proteins may deserve attention in the future as markers for environmental stress in *L. elliptica.* Highly conserved and ubiquitous proteins such as calmodulin may also be studied as markers across taxa. In addition, chaperonins in *L. elliptica* appear to be involved in the stress response. Although the mechanism or mechanisms dictating their induction are not clear in the present study, these deserve further attention as alternatives to HSPs in *L. elliptica* and potentially in other Antarctic marine species, in which HSPs are not induced at environmentally relevant temperatures, if at all induced by thermal stress (Clark and Peck, 2009)

Genes and proteins identified as potential biomarkers for stress in *L.elliptica* should be further tested. The expression of antioxidants (e.g. peroxiredoxin and glutathione-S-transferase), calcium signalling and transport proteins (e.g. calmodulin and other EF-hand domain proteins) and chaperonins (TCP-1) can be monitored both at the transcript and protein levels via qPCR and western blotting respectively. Their expression should be measured at different intensities of stress and lengths of exposure to determine whether these are modulated by stress. It is important to

determine whether the expression of these genes and corresponding proteins is induced to the same extent by stressors other than temperature, to determine whether they are stressor specific. Additionally, it is necessary to determine whether markers are present and induced by stress in other tissues, or are tissue specific.

The potential universality of the markers can then be tested by studying the conditions under which they are expressed across species. All the identified candidates are ubiquitous across taxa and highly conserved, thus relatively easy to clone using degenerative PCR techniques. Their expression can thus be analysed in a range of species, tissues and conditions to determine whether they are induced under the same conditions across species. Finally, they must be observed to overcome some of the limitations presented by HSPs when used as stress biomarkers discussed in chapter one such as induction at environmentally relevant temperatures, independence of thermal histories, and stress specificity.

As previously stated, some aspects of the cellular stress response, such as the expression of HSPs have been extensively studied (Feder and Hofmann, 1999), whereas other aspects of the stress response in non model organisms remain unidentified. An increasing number of studies in the stress response of non model organisms reveal responses to stress different from what would be expected from studies in model organisms. For example, the expression of genes involved in translation and protein synthesis is suppressed in response to acute heat stress in *S. cerevisiae* (Causton et al., 2001), a phenomenon previously described as part of the heat shock response and apparently required to save energy for other processes to maintain homeostasis, or for the production of specific proteins such as HSPs (Lindquist, 1986). However, transcriptomic studies on non model species such as *P. cinctipes* and *C. mirabilis*, show increases in the transcriptional regulation of genes involved in protein synthesis in response to thermal stress (Teranishi and Stillman, 2007; Buckley et al., 2006). In addition, although some mechanisms are consistently identified as involved in the stress response across studies, the behaviour of specific transcripts or proteins varies significantly. Examples of this are the previously discussed differences observed in the behaviour of calmodulin and HMGB between this and previous studies (Desalvo et al., 2008; Podrabsky and Somero, 2004). Given the scarcity of comparative analyses of the response at different levels of stress in non

model species, it is not possible to determine whether observed differences are linked to the organism or tissue under study, or depend on the intensity and time of exposure to the stressor. In the present study, some of the aspects of the cellular stress response, such as alterations in energy metabolism were observed in both *M. edulis* and *L. elliptica*. However, other molecular mechanisms that appear to be involved in their stress response, such as calcium signalling, have not been extensively described in other species.

Whereas the possibility of the existence of a universal stress response has been advocated (Kültz, 2005), the search for a single universal stress biomarker for each individual stressor continues. This may be precluded by the complexity of the response, and the immense diversity of organisms and stressors in the natural environment. As in other disciplines in which biomarkers are currently used extensively (i.e. disease diagnosis), I would argue that experimental biologists might need to resort to the identification of different markers for each species, stressor and environment. A better understanding of the biology and physiology of the species of interest in each case, and of the molecular mechanisms governing the response to stress is thus needed.

6.2 EVALUATION OF THE METHODOLOGY

6.2.1 The link between the proteome and transcriptome

The scientific literature is replete with studies that have tried to establish the link between the transcriptome and the proteome by determining the correlation between mRNA and protein expression levels (Anderson and Seilhamer, 1997; Gygi et al., 1999; Feder and Walser, 2005; Lu et al., 2007). When identifying genes involved in the stress response, based on changes in mRNA abundance, changes in protein abundance are not always observed (Hack and Lopez, 2004). Here, discovery driven transcriptomics and proteomics have been used to independently analyse changes in gene and protein expression using a global approach via microarrays and 2DE. This allows the study of hundreds of proteins expressed at any one time. However, as observed in the current study, limitations arise from the use of these techniques in

isolation, when attempting to establish a link between gene and protein expression. The changes observed at the transcript level in terms of the mechanisms involved could not be confirmed at the protein level. There are several reasons for this, some of which have been discussed in chapter four. These can be briefly summarised as i) the small number of proteins that were maintained for analysis due to the reduced reproducibility of the technique and high variability in individual responses ii) the lack of correlation between mRNA and protein expression as a result of post-translational modifications iii) a time shift between mRNA and protein expression in response to temperature iv) limited identification due to the poor characterization of the genome and proteome of non model compared to model species in public databases v) differences in the number of genes (>8000) and proteins (<400) studied and vi) the lack of previous information about any genes or proteins on the array or gels, which could have allowed a comparison of changes observed at the protein level, on the corresponding gene on the array and *vice versa*.

It is thus clear that if the interest lies in establishing the relationship between a specific gene and its protein product, further analysis of target proteins encoded by the identified genes is necessary. Techniques available for expression analysis of target proteins often rely on immunodetection, but commercially available antibodies generated for model species are often inappropriate for non model organisms (Kültz et al., 2007). This limits the number of proteins whose expression can be analysed under different treatments. An alternative would be to raise specific antibodies for the species of interest, which is labour intensive and time consuming. Therefore, if the main interest lies in analysing a biological process, independent transcriptome and proteome analysis is less biased and more informative than a targeted candidate protein approach.

In terms of the success of each technique in isolation, both techniques were similar in identifying differentially expressed genes and proteins when expressed as a proportion of the total number of features and spots retained after quality filtering (5.3 and 6.9% of proteins for *L. elliptica* and *M. edulis*, and 7.5% of transcripts in *L. elliptica*). The success of the protein identifications by database searches was higher when taken as a proportion of the total number of differentially expressed genes and proteins (28.6 and 38.5% identified proteins in *L. elliptica* and *M. edulis* respectively, as opposed to 14%

identified genes in *L. elliptica*). It must be remembered that 2DE is biased towards high abundant proteins, many of which are housekeeping proteins that tend to be highly conserved. Overall, however, microarray analysis was more successful in identifying differentially expressed genes as the number of clones represented on the array was only limited by time and financial constraints (microarrays can include several tens of thousands of genes) and was relatively large with over 8000 clones.

In contrast, 2DE can only resolve a maximum of around 2000 spots on a single gel, and as the number of gels increases, the number of spots present in all gels is reduced (Challapalli et al., 2004). This can result in a significant reduction in the number of spots to be retained for analysis. Different strategies can be used to increase the number of spots that can be visualised. Technical modifications include changing the pI gradient, changing extraction protocol, a variety of fractionation tools etc (for possible modifications see Lopez, 2007). However, these are expensive, time consuming, and involve running separate gels increasing the cost and labour and potentially decreasing reproducibility. Technical variability can also be minimised by reducing the number of gels using pooled samples. This possibility also has the benefit of increasing statistical power through a reduction in biological variance, although this is not as large as expected (Diz et al., 2009). In addition, pooling in proteomics is only recommended when access to the samples is easy and their cost low (Karp and Lilley, 2008), which is not the case for *L. elliptica*.

Although with reduced success when compared to equivalent studies in model species, sets of genes and proteins expressed in non model species can be identified using global analysis of gene and protein expression and sequence similarity searches. Despite the limited identifications, transcriptomics and proteomics have been used here in these poorly characterised non model species, revealing information about some of the mechanisms involved in the stress response and their underlying variability. An even more limiting situation is perhaps the reliance on information based on model species regarding biochemical pathways in non models. An example of this is the hypothalamus-pituitary-adrenal (HPA) axis in mammals and equivalent hypothalamus-pituitary-interrenal (HPI) axis in fish. This is an ancient physiological system that plays a central role in the vertebrate stress response, which function has been briefly summarized in chapter five. Studies of extant species show that many of

the genes of the HPA axis are present in urochordates and cephalochordates and it has been suggested that strong positive selection has maintained the structure and functions of this axis owing to the pivotal adaptive role that stress hormones play in individual survival (Denver, 2009). However, despite presenting some equivalent hormones to those synthesised by the HPA, this axis is not present in invertebrates, thus the pathways leading to the synthesis of stress hormones in invertebrates are not clear. This was reflected in chapter five, where well described stress hormones in mammals were targeted in bivalves. Their presence was not detected, confirming that previous knowledge that well conserved pathways and/or compounds are present in model species does not imply that the same or equivalent pathways exist across taxa. A more global approach via MS would be necessary to identify potential stress hormones in marine ectotherms.

Clearly, the microarray approach was the most successful in terms of the identification of molecular mechanisms potentially involved in the heat stress response, but the relevance of the observed changes in expression of candidate genes must be evaluated at the protein level.

6.2.2 Experimental temperatures and length of exposure

At the time the experiments were planned, there was some evidence from previous preliminary work that the 2-6 h exposure time used when evaluating the classic heat shock protein response in temperate species was not sufficient to initiate a HSP response in Antarctic ectotherms (L. Peck pers. comm.). Based on this observation and on knowledge of the slower metabolic rates observed in Antarctic species (Peck et al., 2002), the length of exposure was extended. However, later studies on HSP expression in *L. elliptica* showed that HSPs are induced after a short 2 h exposure, but the temperature required to induce the response is higher than predicted (Clark et al., 2008). However, the use of a 24 h exposure period did not preclude analysis of the pathways involved in the response and ensured that a delayed response was not missed. The temperatures to which the clams were exposed were based on physiological studies on this species as reviewed in chapter three. At 3°C *L.elliptica* suffers 50% failure in essential biological activities and survives only a few days at 9°C (Peck et al., 2004). There was therefore an interest in analysing the response to

an extreme and a moderate shock. In addition, another physiological measure was introduced by allowing the clams at 3°C to bury with the aim of comparing those that buried with those that were unable to burrow, to determine whether differences in the physiological status of these two groups could be observed at the molecular level.

Throughout the development of this project, knowledge on the molecular analysis of Antarctic species and on the efficiency of the techniques available increased significantly. If the experiment were to be repeated, the experimental set up would have been as follows:

- Exposure temperatures would have been 2°C, 3°C and 6°C. Exposure to temperatures lower than 3°C could have helped to determine the induction threshold for antioxidants and provide a better understanding of pejus thresholds in this species. An extreme response could have been studied at 6°C, the approximate critical temperature in *L. elliptica* (Peck et al., 2004).

- Physiological measurements: Problems with tissue storage resulted in degraded RNA, thus heat shocks had to be repeated to obtain new tissue samples for microarray analysis. Due to the limited number of specimens available at the time, differences between buried and non buried clams could not be analysed at the transcript level. Proteomic analysis did not allow the identification of a sufficient number of proteins to be able to compare expression between the treatments. However, global changes in protein expression revealed few differences. Based on this, the burrowing component would be excluded and substituted by a different treatment. Measurement of oxygen consumption rates or haemolymph oxygenation would have strengthened the link with the oxygen limitation theory.

6.2.3 Identification of mechanisms involved in the heat stress response at the transcript level

In the present study, only those genes identified as showing significant changes in expression in the treated clams when compared to the control group were sequenced. This approach is less time consuming and more cost effective than sequencing the full

library and is sufficient to identify genes exhibiting changes in expression as candidates for follow up studies on biomarker discovery. However, if the experiment were to be continued, efforts would be directed to sequence the full library. This could allow the analysis of gene networks, reflecting specific biological processes involved in the stress response, and identification of those than remain unchanged. This holistic approach could help in determining the significance in the observed changes in gene expression at higher levels of organization. The most common method of grouping genes relies on GO annotations, and thus on sequence identifications, which makes it less effective in poorly characterised non model species. Nonetheless, this approach can identify changes in gene expression that may be hidden when considering the expression of individual genes in isolation. In addition, between species comparisons can be made based on biological processes, rather than comparing orthologous genes (Gracey, 2007).

6.2.4 Acute stress as a proxy for climate change

The outcome of this study must be interpreted with care when making predictions on the effects of climate change on biodiversity. The reasons for this are presented below.

Organisms in their natural environment seldom undergo one stress at a time (Feder & Hofmann, 1999) but are generally exposed to a series of stresses, the cumulative effects of which are difficult, if at all possible, to quantify (Feder, 1999). Also, the rate at which species are likely to be exposed to rising temperatures is much slower than that used in short term experiments. This raises the concern as to whether conclusions drawn from such experiments are applicable in an environmental context. Rapid rates of change do not necessarily throw light on the ability of organisms to modulate physiological functions through processes such as acclimation. The rate of warming has a marked effect on temperature limits and it has been suggested that mechanisms limiting survival may differ at slower rates of change (Peck et al., 2009). Molecular analyses of the mechanisms determining the thresholds at slower rates of warming could thus aid in understanding the factors involved in acclimation and thermal adaptation that are more relevant to climate change. There is a need to understand organismal responses during periods of chronic stress, as little information

is currently available for persistent stressors. Evaluating the difference between acute, chronic or seasonal stress in the same species, would assist in the understanding of its overall stress response and help determine which mechanisms are likely to be compromised during climate change.

Despite the caveats of short term experiments when extrapolating their results to environmentally relevant processes, the analysis of molecular responses under acute exposure to elevated temperatures is essential. It provides an excellent tool for identifying mechanisms that allow animals to cope with stress, and those that fail under changing conditions.

6.3 *M. edulis* AND *L. elliptica*: PROSPECTS FOR SURVIVING CLIMATE CHANGE

When faced when new selection pressures, such as those exerted by climate change, organisms can respond in three main ways. Firstly, they can evade by migrating to different habitats. Secondly, they can cope with physiological flexibility either within their tolerance range or within their ability to acclimate without altering their genetic makeup. Thirdly, populations can adapt to the new conditions through genetic changes (Peck, 2005; Gienapp et al., 2008). The relative importance of these mechanisms will vary depending on the timescale of the change, species life history, rate and extent of environmental change, and availability of alternative habitats and dispersal ability. These strategies ought to be evaluated when analysing a species potential to survive climate change.

Migration of species along latitudinal and longitudinal gradients has been observed for many species, a phenomenon which is causing shifts in species distributions towards higher latitudes (Parmesan and Yohe, 2003). *M. edulis* populations, for example, have settled in the Arctic Archipelago of Svalbard after a thousand year absence as a result of an unusually high northward mass transport of warm Atlantic water resulting in elevated sea surface temperatures in the North Atlantic and along the west coast of Svalbard (Berge et al., 2005). Antarctic marine species, however, are limited in their ability to migrate. The circular outline of the Antarctic continent

and its isolation from other oceans by the circumpolar current make it very difficult for migrating species, even those with extended larval phases (Peck, 2005). In addition, the present study has shown that *L. elliptica* undergoes failure in essential biological functions and aerobic capacity at temperatures 2°C higher than current summer maxima. Although similar thresholds ought to be measured in larvae to determine their dispersal ability, such thermal sensitivity would leave few options for Antarctic organisms to colonize new suitable habitats in their lifespan. This is particularly relevant where migration requires individuals to cross thermal thresholds (Barnes and Peck, 2008).

This study has provided confirmation, at the molecular level, that rapid temperature increases of a similar magnitude to those predicted by current models over the next 50 years will result in oxygen and capacity limitations to the ability of *L. elliptica* to acclimate. Other essential mechanisms, such as calcium signalling and antioxidant defences have been shown to be affected by rapid warming. Essential biological functions such as burrowing are also impaired with small temperature elevations. Plasticity might allow an organism to overcome the first challenge when exposed to a changing environment: the ability to survive and persist. However, it is possible that the adaptive responses that allow survival are coupled with non adaptive changes which can affect growth, reproduction, or other essential biological functions, considerably limiting the species potential to survive. An example of this is HSP expression under severe stress. HSP induction is beneficial for it increases heat resistance, an adaptive response to thermal stress. However, it presents associated costs including the shut down of other cellular function, the extensive use of energy and the toxic effects of high HSP concentrations. These have detrimental effects in growth, cell division and reproduction (reviewed in Sørensen et al., 2003). It must be remembered that further work is needed at slower rates of change, but physiological flexibility without genetic assimilation of the adaptive traits might not to suffice for this species ability to survive climate change. Under a long-term directional change in the environment, evolutionary adaptation might therefore be essential.

It remains to be determined whether the predicted rate of warming may be too rapid for populations to adapt. Firstly, underlying genetic variation increases the capacity to respond and adapt to selective challenge (Bürger and Lynch, 1995). The results

presented here suggest a higher level of variation in the heat stress response of *M. edulis* compared to *L. elliptica*, which could imply that *L. elliptica* has reduced underlying genetic and/or phenotypic variability in its response to elevated temperatures when compared to *M. edulis*, thus the potential for this species to adapt to new conditions might be reduced. Secondly, evolutionary change must occur fast enough to keep up with the rapid rate of change predicted. There is ample evidence that high rates of evolution, observable within the lifetime of a researcher, do indeed occur in nature (reviewed in Reznick and Ghalambor, 2001). These, however, have been observed for species with short generation times with potential for rapid population growth such fruit flies or guppies. A change in global climate will be experienced as a severe deterioration in the environment by organisms with long generations, but as a gradual worsening by organisms with short generations (Holt, 1990). Antarctic benthic species have slow physiological rates and deferred sexual maturity, resulting in extended life histories and long generation times (Peck, 2002). Thus, in this respect, Antarctic species would also have a disadvantage when compared to temperate species.

Predicting the consequences of climate change on the evolutionary potential of species is extremely difficult. Identifying mechanisms that allow for plasticity is only the first step. The potential for such mechanisms to evolve must be evaluated independently, and as part of the complex system. After evaluating individual responses, these must be analysed at the population level, followed by the analysis of species interactions and other ecological responses, which require complex modelling tools. However, comparative studies as presented here are of value, as they are informative in terms of the relative susceptibility of species to climate change. Our efforts to protect biodiversity should be directed to those species that are least likely to adapt to a warming world.

6.4 DIRECTIONS FOR FUTURE WORK

6.4.1 Use of resources and ongoing work

The *L. elliptica* microarray generated here is currently being used to study the response to oxidative stress in *L. elliptica*. The overlap in the response to different stressors can be tested and mechanisms unique to each response can be identified. This has potential implications for the discovery of universal biomarkers for specific stressors in *L. elliptica*. Successful hybridizations of *M. edulis* samples on the microarray have shown that hybridizations of samples from closely related species are possible. This opens the possibility if testing the universality of the stress response by subjecting closely related species to a given stressor and analysing where responses overlap. Sequencing of a large number of genes in *L. elliptica* is currently being carried out in our laboratory, which will likely allow the identification of some of the differentially expressed genes that remain unknown. This will throw more light into the mechanisms involved in the response.

The proteome expressed by *M. edulis* under a series of stressors has been analysed and the relevance of the observed changes is currently being evaluated. This includes responses to other laboratory induced stressors, as well as naturally occurring environmental stress. Comparisons between the changes induced will be made.

6.4.2 Future directions

Living systems are complex, as are the responses to environmental stimuli, as reflected in this study. The value of identifying genes and proteins whose expression is modulated by stress lies in identifying mechanisms involved in the response. These can be used to formulate hypotheses about the physiological and cellular processes that increase organismal performance in response to changing environmental conditions. In the present study, transcriptomics and proteomics have been used to study the stress response of the Antarctic and temperate marine ectotherms *L. elliptica* and *M. edulis*. Genes whose expression is induced by heat stress have been identified in *L. elliptica* and proteins involved in the response to different levels of stress have

been identified in both species. These suggest mechanisms involved in the stress response, which can be used in turn to formulate hypotheses, not only about processes involved in the response, but about the underlying variation in traits that might be relevant to the prospect of surviving climate change in these species.

In order to expand current knowledge and make predictions about the effects of a warming climate, further work should be directed to the full characterization of the response to environmentally relevant temperatures under slower rates of change and over longer exposure times. It remains to be determined whether the observed changes in gene and protein expression constitute adaptive responses to change. As previously stated, adaptive responses should place the phenotype closer to the optimum range for physiological performance (Ghalambor et al., 2007). It thus follows that the plasticity that allows for acclimation to occur would represent an adaptive response to the changing conditions. Indeed, physiological changes involved in acclimation have often been in the direction predicted from an adaptive hypothesis, thus acclimation has been previously considered to be adaptive (Sørensen et al., 2003). Adaptive responses may be subjected to evolutionary selection over many generations enabling persistence during climate change (c.f. the Beneficial Acclimation Hypothesis; Leroi et al., 1994). Acclimation experiments to 2-3°C in *L. elliptica* and 27°C in *M. edulis* extending over several months and analysis of mechanisms involved in acclimation to elevated temperatures via microarray analysis could help determine which mechanisms involved in the stress response constitute adaptive responses. Validation of the observed changes at the level of gene expression would then be required at the protein level.

As the next step, once a complete description of the mechanistic machinery involved in the process of interest is achieved, measurements of components of fitness under field and controlled conditions, comparing the relative fitness of wild types and mutants that are deficient for the putative plasticity gene should be made (Pigliucci, 1996). Currently, gene manipulations are most easily addressed using model species (Cossins et al., 2006). The challenge of developing such tools for non model bivalve species lies ahead. However, with genomic resources increasing at the current rate, the scope for these types of studies will increase substantially over the next few years.

This could help determine the role of the identified genes and mechanisms involved in generating environmentally adaptive phenotypes.

Heterologous hybridizations can be used to provide further comparisons between closely related species across latitudinal gradients. Analysis of the plastic traits that allow species to respond to change, and their underlying variability, together with a good understanding of physiological responses to change could aid in understanding ecosystem resilience and predicting the impact of environmental stress on species worldwide. Efforts should concentrate on the study of species living close to their physiological limits such as Antarctic, tropical or high intertidal species, which have been suggested to be at higher risk (c.f. Parsons, 1991). Comparative studies with highly eurythermal, plastic species will reveal differences in their response relevant to persistence and adaptation in a warming climate.

6.5 Conclusion

It is currently accepted that the global ocean is warming and there is evidence that biodiversity is already being affected. It is also clear that some ecosystems and species will be at considerable higher risk than others. Our ability to predict how species will cope or adapt to a warming climate is limited by the lack of information of the mechanisms involved in the response to changing environmental conditions. Molecular analysis of responses to change can successfully identify such mechanisms, but this is only the first step in the evaluation of their susceptibility to change. Assessment of the genetic and phenotypic variance underlying the traits involved in the response to increased temperature, together with extensive knowledge of physiology and life histories could provide excellent tools to assess the susceptibility of individual species to elevated temperatures. Biomarkers can help determine whether an individual is undergoing stress, but these are usually identified using acute experiments and mechanisms involved in the response are likely to change under slower rates of change. Acclimation experiments over weeks, months or years; an evaluation of the ability of species to acclimate to elevated temperatures; and continued assessment of species responses in their natural environment would be needed.

6.6 REFERENCES

Anderson L, Seilhamer J (1997) A comparison of selected mRNA and protein abundances in human liver. Electrophoresis 18:533-537

Barnes DKA, Peck LS (2008) Vulnerability of Antarctic shelf biodiversity to predicted regional warming. Climate Research 37:149-163

Berge J, Johnsen G, Nilsen F, Gulliksen B, Slagstad D (2005) Ocean temperature oscillations enable reappearance of blue mussels *Mytilus edulis* in Svalbard after a 1000 year absence. Marine Ecology Progress Series 303:167-175

Buckley BA, Gracey AY, Somero GN (2006) The cellular response to heat stress in the goby *Gillichthys mirabilis*: a cDNA microarray and protein-level analysis. Journal of Experimental Biology 209:2660-2677

Bürger R, Lynch M (1995) Evolution and extinction in a changing environment: a quantitative-genetic analysis. Evolution 49:151-163

Causton HC, Ren B, Koh SS, Harbison CT, Kanin E, Jennings EG, Lee TI, True HL, Lander ES, Young RA (2001) Remodelling of yeast genome expression in response to environmental changes. Molecular Biology of the Cell 12:323-337

Challapalli KK, Zabel C, Schuchhardt J, Kaindl AM, Klose J, Herzel H (2004) High reproducibility of large-gel two-dimensional electrophoresis. Electrophoresis 25:3040-3047

Clark MS, Fraser KPP, Peck LS (2008) Antarctic marine molluscs do have an HSP70 heat shock response. Cell Stress and Chaperones 13:39-49

Clark MS, Peck LS (2009) HSP70 heat shock proteins and environmental stress in Antarctic marine organisms: A mini-review. Marine Genomics 2:11-18

Cossins A, Fraser J, Hughes M, Gracey A (2006) Post-genomic approaches to understanding the mechanisms of environmentally induced phenotypic plasticity. Journal of Experimental Biology 209:2328-2336

Denver R (2009) Structural and Functional Evolution of Vertebrate Neuroendocrine Stress Systems. Trends in Comparative Endocrinology and Neurobiology 1163:1-16

Desalvo MK, Voolstra CR, Sunagawa S, Schwarz JA, Stillman JH, Coffroth MA, Szmant AM, Medina M (2008) Differential gene expression during thermal stress and bleaching in the Caribbean coral *Montastraea faveolata*. Molecular Ecology 17:3952-3971

Diz AP, Truebano M, Skibinski DOF (2009) The consequences of sample pooling in proteomics: An empirical study. Electrophoresis 30:2967-2975

Feder ME, Hofmann GE (1999) Heat-shock proteins, molecular chaperones, and the stress response: Evolutionary and ecological physiology. Annual Review of Physiology 61:243-282

Feder ME (1999) Organismal, ecological, and evolutionary aspects of heat-shock proteins and the stress response: Established conclusions and unresolved issues. American Zoologist 39:857-864

Feder ME, Walser JC (2005) The biological limitations of transcriptomics in elucidating stress and stress responses. Journal of Evolutionary Biology 18:901-910

Ghalambor CK, McKay JK, Carroll SP, Reznick DN (2007) Adaptive versus non-adaptive phenotypic plasticity and the potential for contemporary adaptation in new environments. Functional Ecology 21:394-407

Gienapp P, Teplitsky C, Alho JS, Mills JA, Merilä J (2008) Climate change and evolution: disentangling environmental and genetic responses. Molecular Ecology 17:167-178

Gracey AY (2007) Interpreting physiological responses to environmental change through gene expression profiling. Journal of Experimental Biology 210:1584-1592

Gygi SP, Rochon Y, Franza BR, Aebersold R (1999) Correlation between protein and mRNA abundance in yeast. Molecular and Cellular Biology 19:1720-1730

Hack CJ, Lopez JA (2004) An exploration of some factors affecting the correlation of mRNA and proteomic data. Knowledge Exploration in Life Science Informatics, Proceedings 3303:9-19

Holt RD (1990) The microevolutionary consequences of climate change. Trends in Ecology and Evolution 5:311-315

Karp NA, Lilley KS (2008) Investigating sample pooling strategies for DIGE experiments to address biological variability. Proteomics 9:388-397

Kültz D (2005) Molecular and evolutionary basis of the cellular stress response. Annual Review of Physiology 67:225-257

Kültz D, Fiol D, Valkova N, Gomez-Jimenez S, Chan SY, Lee J (2007) Functional genomics and proteomics of the cellular osmotic stress response in `non-model' organisms. Journal of Experimental Biology 210:1593-1601

Leroi AM, Bennett AF and Lenski RE (1994) Temperature-Acclimation and Competitive Fitness - an Experimental Test of the Beneficial Acclimation Assumption. Proceedings of the National Academy of Science USA 91:1917-1921

Lindquist S (1986) The heat shock response. Annual Review of Biochemistry 55:1151-1191

Lopez JA (2007) Two-dimensional electrophoresis in proteome expression analysis. Journal of Chromatography B-Analytical Technologies in the Biomedical 849:190-202

Lu P, Vogel C, Wang R, Yao X, Marcotte E (2007) Absolute protein expression profiling estimates the relative contributions of transcriptional and translational regulation. Nature Biotechnology 25:117-124

Parmesan C, Yohe G (2003) A globally coherent fingerprint of climate change impacts across natural systems. Nature 421:37-42

Parsons PA (1991) Evolutionary rates - stress and species boundaries. Annual Review of Ecology and Systematics 22:1-18

Peck LS (2002) Ecophysiology of Antarctic marine ectotherms: limits to life. Polar Biology 25:31-40

Peck LS (2005) Prospects for surviving climate change in Antarctic aquatic species. Frontiers in Zoology 2:9

Peck LS, Pörtner HO, Hardewig I (2002) Metabolic demand, oxygen supply, and critical temperatures in the Antarctic bivalve *Laternula elliptica.* Physiologial and Biochemical Zoology 75:123-133

Peck LS, Webb KE, Bailey DM (2004) Extreme sensitivity of biological function to temperature in Antarctic marine species. Functional Ecology 18:625-630

Peck LS, Clark MS, Morley SA, Massey A, Rossetti H (2009) Animal temperature limits and ecological relevance: effects of size, activity and rates of change. Functional Ecology 23:248-256

Peñuelas J, Filella I (2001) Phenology - Responses to a warming world. Science 294:793

Pigliucci M (1996) How organisms respond to environmental changes: From phenotypes to molecules (and vice versa). Trends in Ecology and Evolution 11:168-173

Podrabsky JE, Somero GN (2004) Changes in gene expression associated with acclimation to constant temperatures and fluctuating daily temperatures in an annual killifish *Austrofundulus limnaeus.* Journal of Experimental Biology 207:2237-2254

Pörtner HO (2002) Climate variations and the physiological basis of temperature dependent biogeography: systemic to molecular hierarchy of thermal tolerance in animals. Comparative Biochemistry and Physiology A Molecular and Integrative Physiology 132:739-761

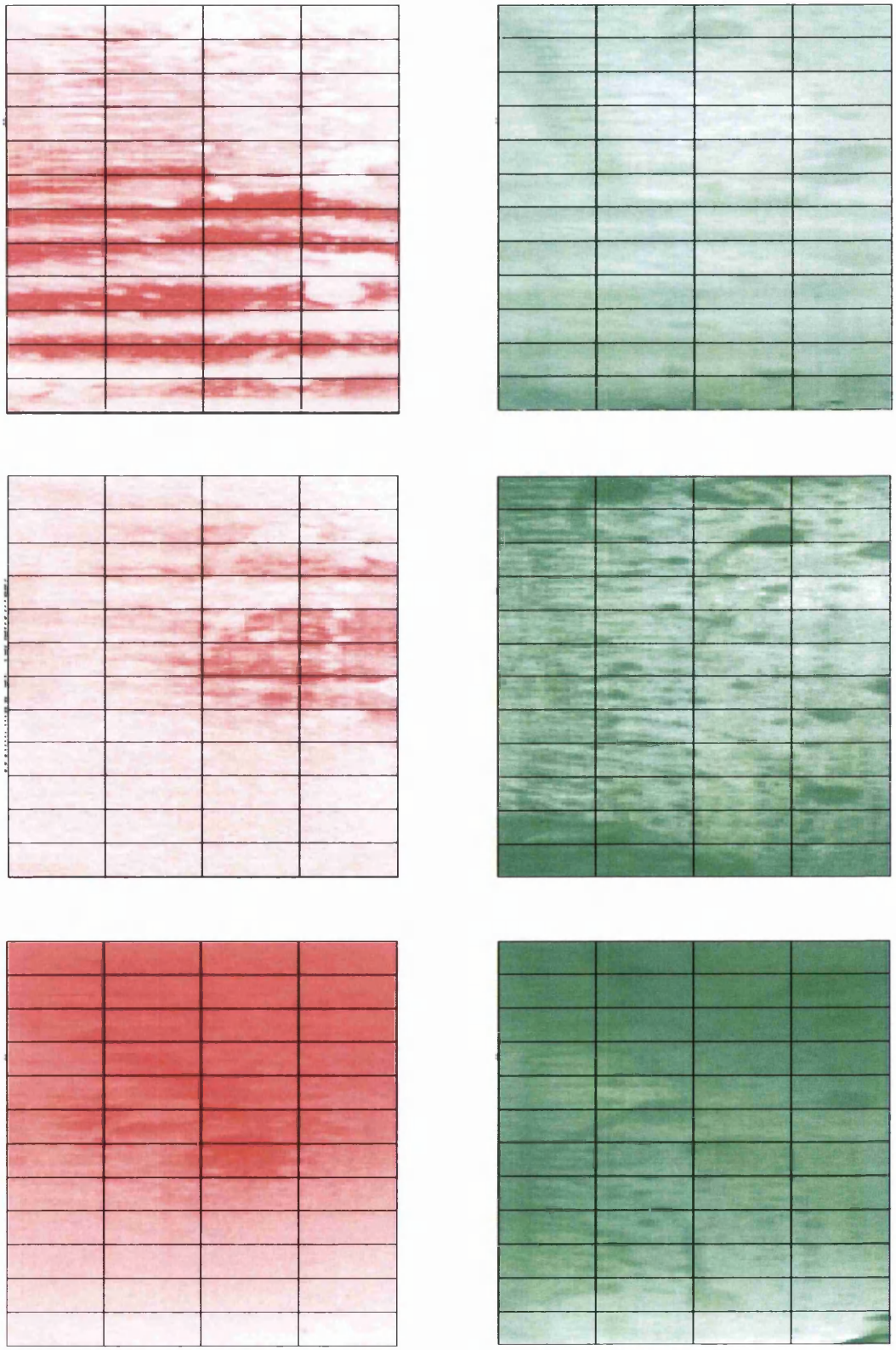
Reznick DN, Ghalambor CK (2001) The population ecology of contemporary adaptations: what empirical studies reveal about the conditions that promote adaptive evolution. Genetica 112:183-198

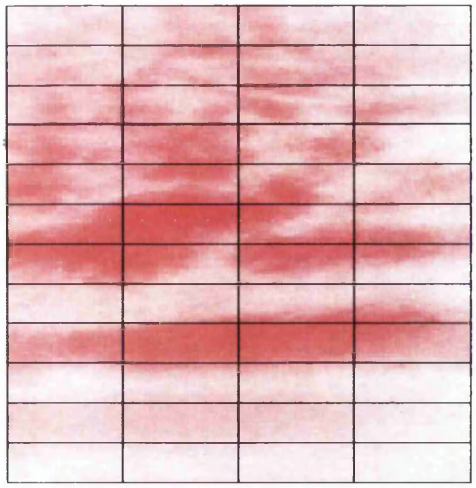
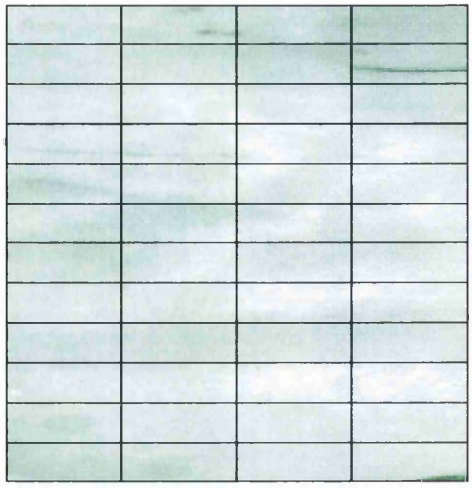
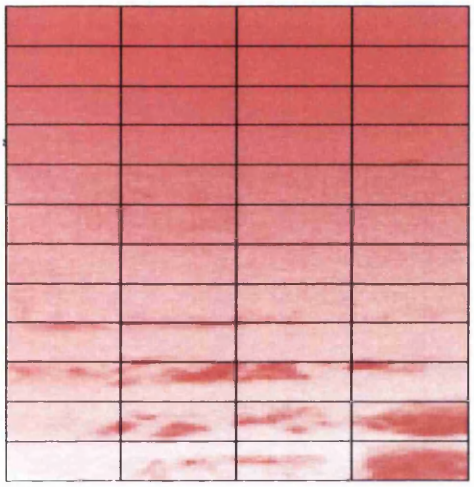
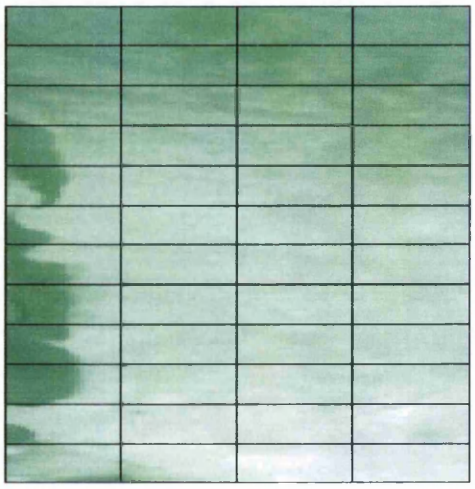
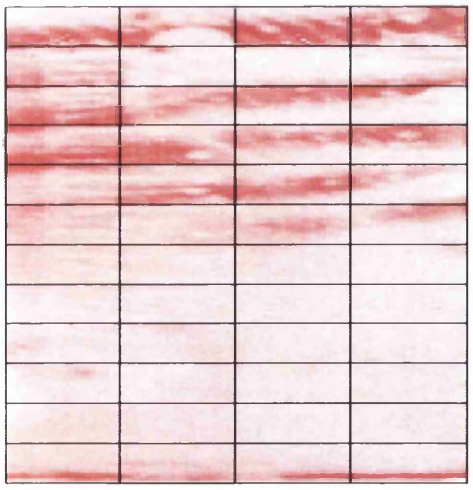
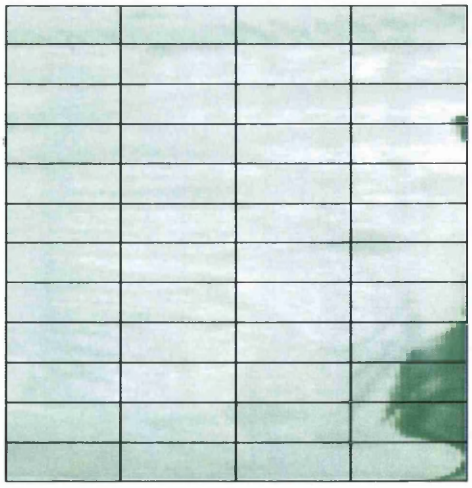
Sørensen JG, Kristensen TN, Loeschcke V (2003) The evolutionary and ecological role of heat shock proteins. Ecology Letters 6:1025-1037

Teranishi KS, Stillman JH (2007) A cDNA microarray analysis of the response to heat stress in hepatopancreas tissue of the porcelain crab *Petrolisthes cinctipes*. Comparative Biochemistry and Physiology D Genomics and Proteomics 2:53-62

Walther G, Post E, Convey P, Menzel A, Parmesan C, Beebee T, Fromentin J, Hoegh-Guldberg O, Bairlein F (2002) Ecological responses to recent climate change. Nature 416:389-395

APPENDIX I. MICROARRAY QUALITY SCREENING



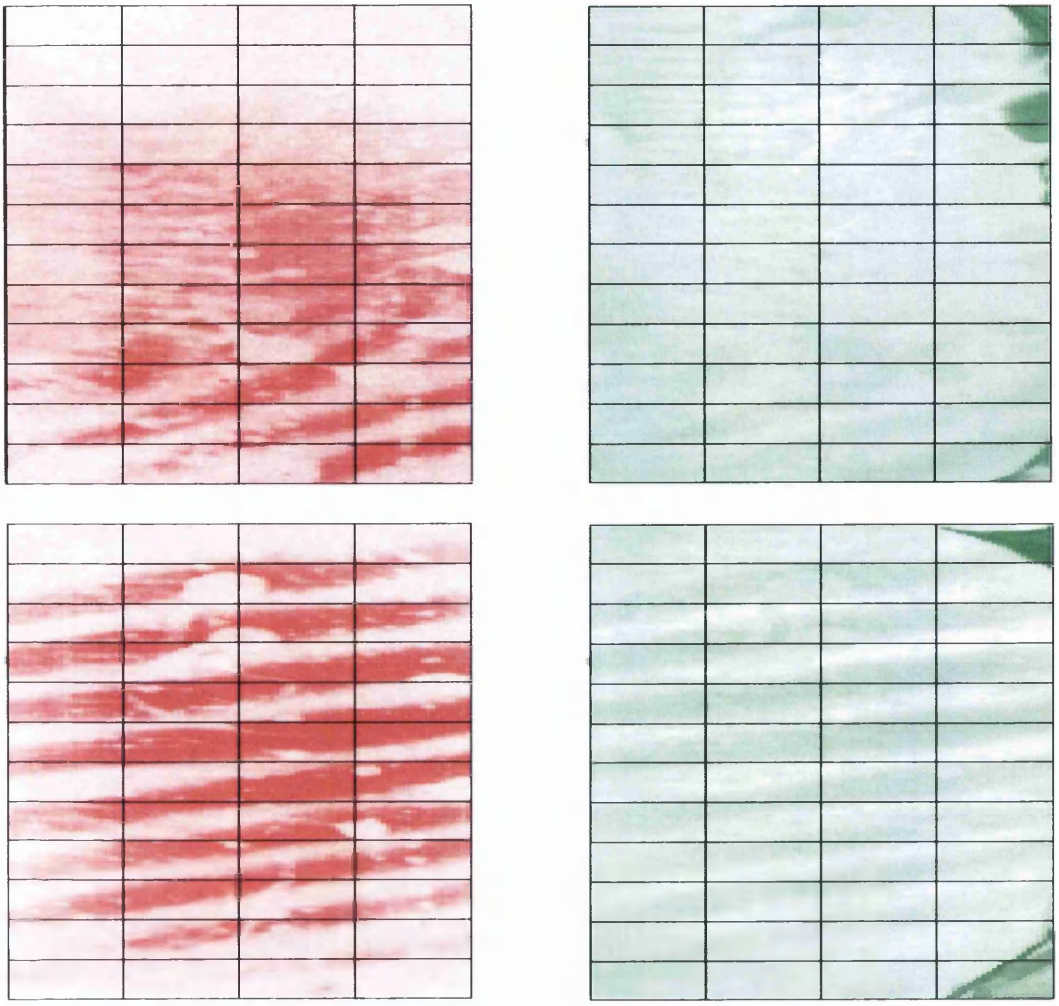
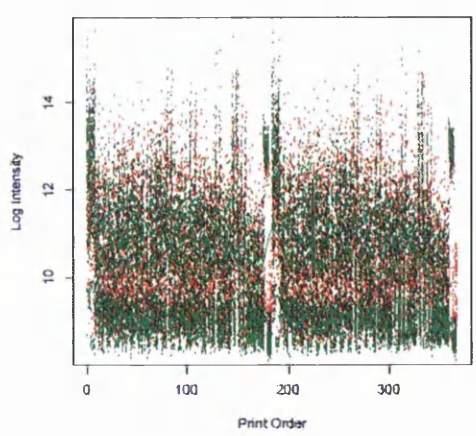
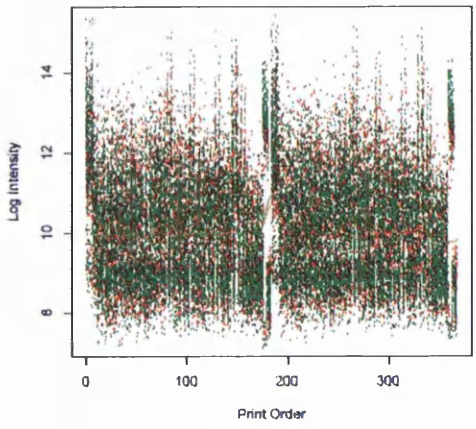


Figure A.1 Background signal images for microarrays 1-8 in both channels. Cy5 (635nm) and corresponding Cy3 (532nm) are shown in the first and second columns respectively.





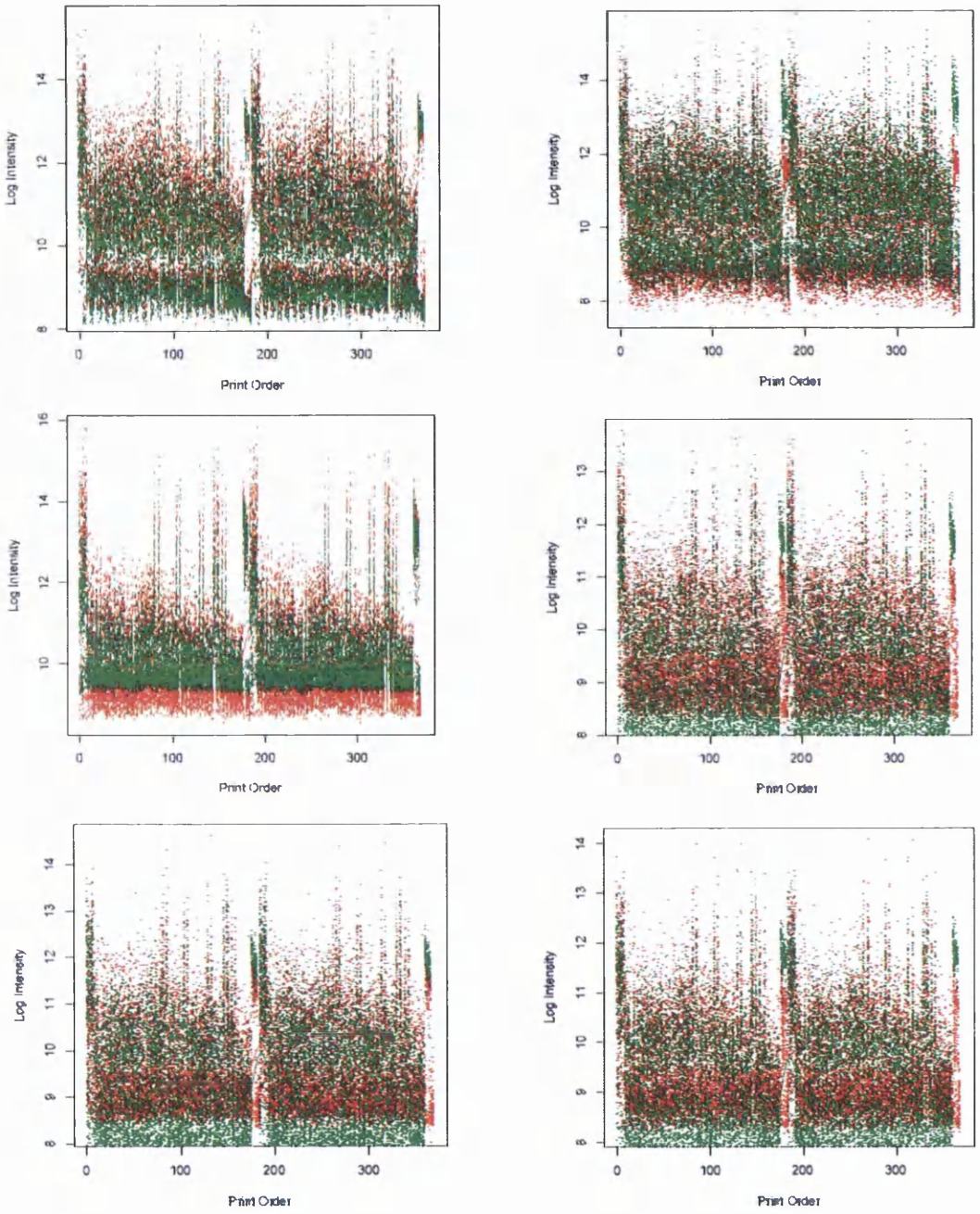
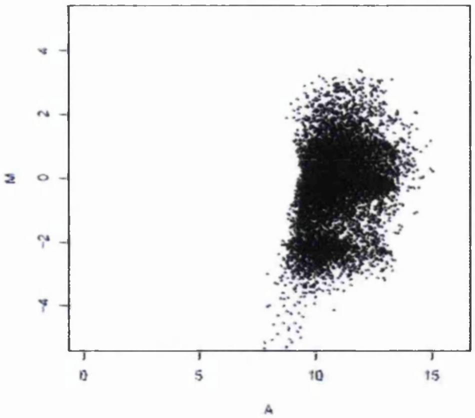
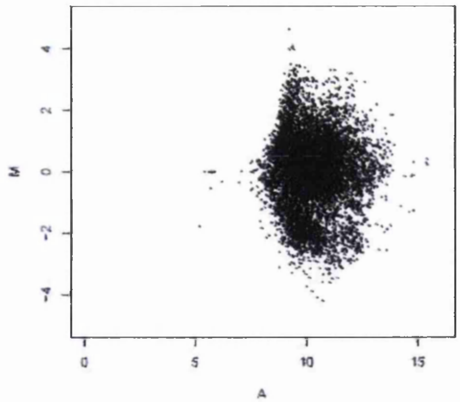
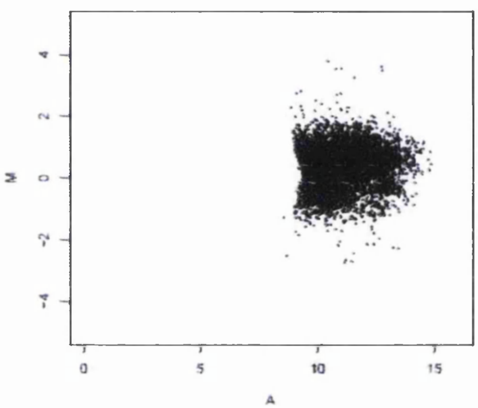
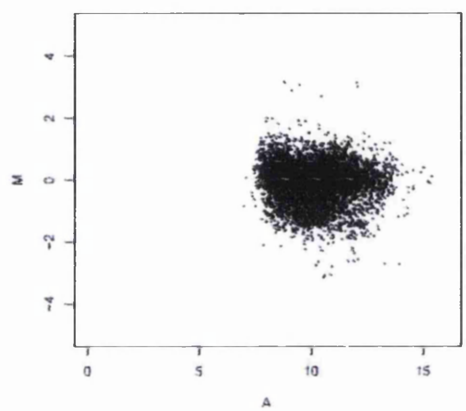


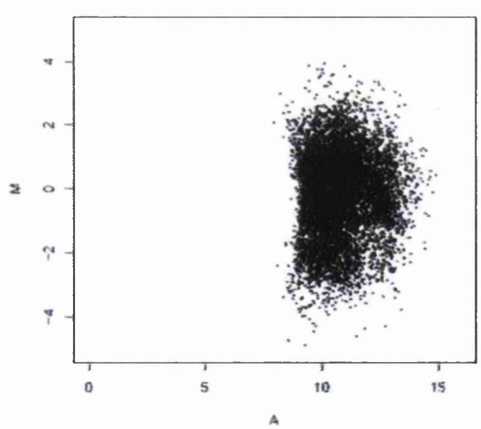
Figure A.2 Print tip order for microarrays 1-4 and 5-8 in the left and right columns respectively. Both channels are shown in the same plot. Cy3 (532nm) is shown in red and Cy5 (635nm) in green.

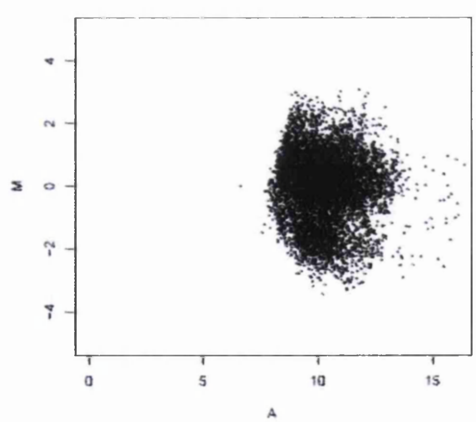

M
4
2
0
-2
-4
0
5
10
15
A

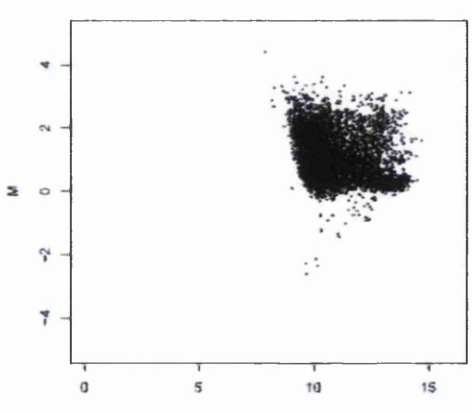

M
4
2
0
-2
-4
0
5
10
15
A

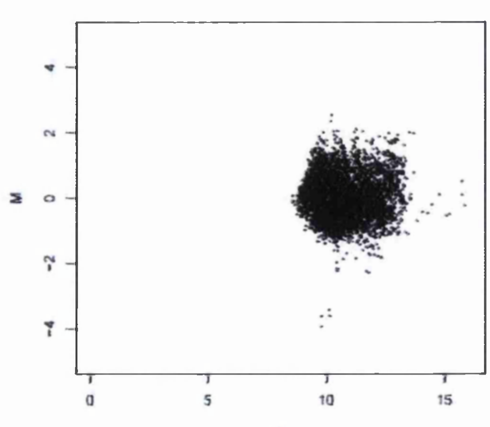

M
4
2
0
-2
-4
0
5
10
15
A


M
4
2
0
-2
-4
0
5
10
15
A


M
4
2
0
-2
-4
0
5
10
15
A


M
4
2
0
-2
-4
0
5
10
15
A


M
4
2
0
-2
-4
0
5
10
15
A


M
4
2
0
-2
-4
0
5
10
15
A

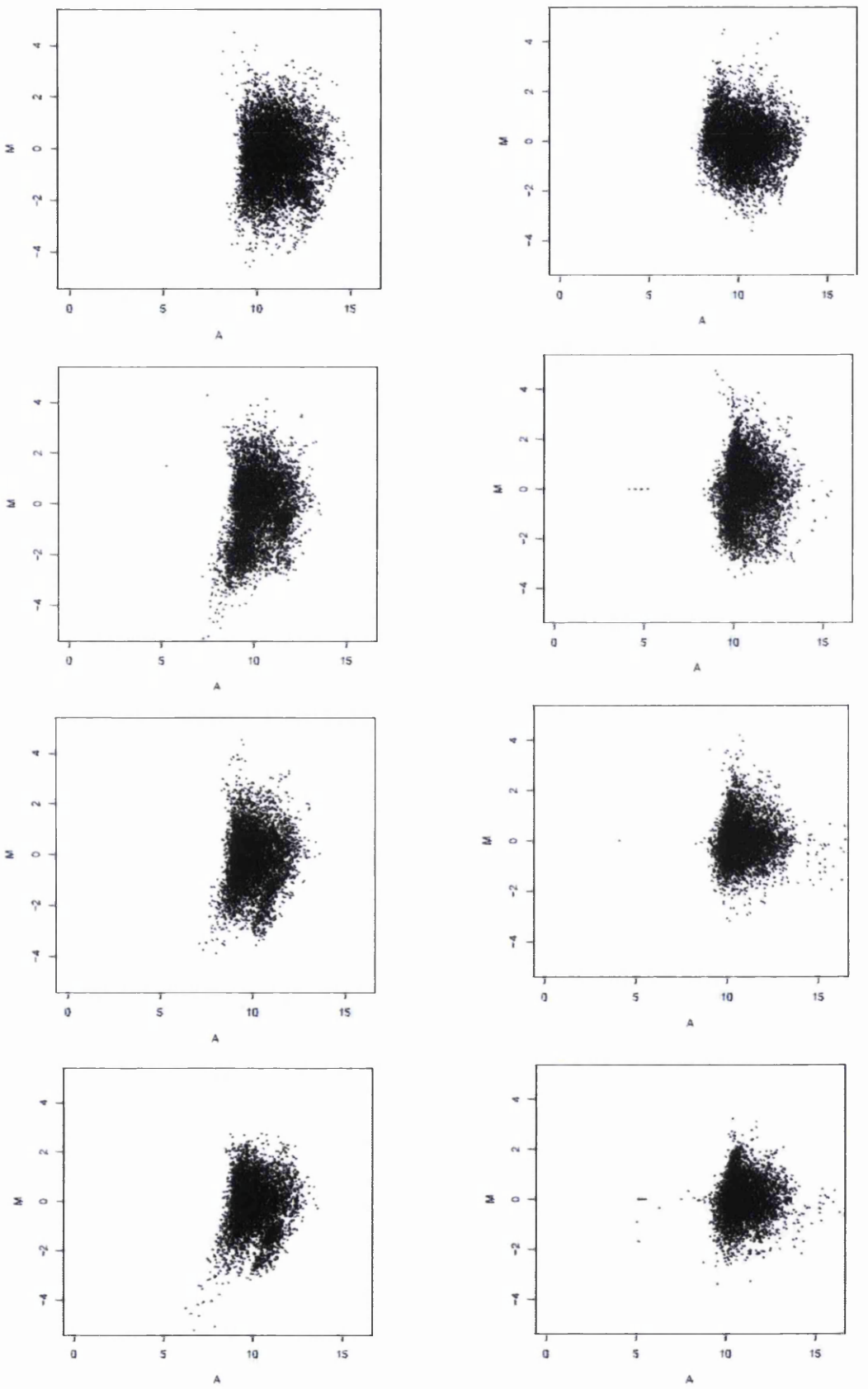


Figure A.3 MA plots showing the relationship of signal ratios to signal intensities. Plots are shown for arrays 1-8 in consecutive order. Raw data and corresponding normalised values are shown in the left and right columns respectively.

APPENDIX II. IMMUNOREACTIVE CORTISOL MEASUREMENTS

Figure II. Immunoreactive cortisol measurements for *L. elliptica* (control, 3°C and 9°C treated samples) and *M. edulis* (control, 27°C and 37°C treated samples). In all cases n=10, except for 3°C treated group for *L. elliptica* (n=20). Individual samples below detection level are not shown. Concentrations for each individual samples are shown in pg ml^{-1}

L. elliptica				*M. edulis*	
0°C	**3°C**	**9°C**	**12°C**	**27°C**	**37°C**
6.02	67.06	9.99	40	13.97	4.20
9.26	7.59	10.42	1941.32	9.22	6.29
14.39	5.68	19.62	1662.57	7.66	7.49
14.68	6.02	8.00	495.57	4.14	7.59
8.84	7.16	14.97	450.74	4.13	6.84
9.36	15.65	26.12	266.34	49.96	6.72
8.39	8.26	9.28	113.80	5.00	
12.68	15.37	16.25	45.32	4.64	
15.97	9.87	8.84	25.10	6.16	
	6.06	6.76	10.00		
	16.84				
	7.63				
	9.04				
	7.16				
	16.03				
	9.26				
	16.58				
	8.61				
	7.63				
	13.34				