VICTORIA UNIVERSITY OF WELLINGTON

Te Whare Wānanga o te Ūpoko o te Ika a Māui

AVAILABILITY OF THESIS

AUTHOR'S NAME Adele Lorene Hineaiu Whyte

TITLE OF THESIS Environmental Toxicology of Perna canaliculus

DEGREE PhD

SUBJECT Cell and Molecular Biology

YEAR 2006

I hereby consent to the above thesis being consulted, borrowed, copied or reproduced in accordance with the provisions of the Library Regulations from time to time made by the Academic Board.

(Signature) ..... AUTHOR

24/10/06 ..... DATE

Environmental Toxicology of *Perna canaliculus*

Adele Lorene Hineao Whyte

A thesis submitted to Victoria University of Wellington in fulfilment of the requirements for the degree of Doctor of Philosophy in Cell and Molecular Biology

Victoria
UNIVERSITY OF WELLINGTON
Te Whare Wānanga o te Ūpoko o te Ika a Māui

2006

Toitu te marae o Tane,

Toitu te marae o Tangaroa,

Toitu te Iwi.

Protect and strengthen the realms of the land and sea and they will protect and strengthen the people.

Whakatauki: Māori Proverb

ABSTRACT

New Zealand does not have a major problem with marine pollution but there is still a need to develop methods to monitor the environment and protect ecosystems. Although some previous studies in New Zealand have measured the concentrations of chemicals in tissues from marine organisms, few to date have developed biomarkers of contaminant exposure. In the current study attempts were made to develop biomarkers for heavy metal contamination in the endemic New Zealand greenshell mussel *Perna canaliculus*.

Metallothionein (MT) gene nucleotide sequences were isolated from *P. canaliculus* by cloning PCR products from genomic DNA. Nine MT exon 2 amino acid sequences were deduced, some of which were characterised by unusual features, including the presence of atypical tyrosine and histidine residues and lower than usual numbers of metal binding cysteine residues. MT sequences isolated in the current study were compared with those from other mollusc species worldwide.

A 2-D gel DIGE proteomic approach was used to detect proteins involved in response to low salinity or heavy metal contamination. In the salinity study, control mussels were killed at the start of the experiment and others were exposed to ambient (32 ppt) and reduced (14 ppt) salinity for 3 days. Approximately 115 proteins showed significant (*t*-test $p < 0.01$) differences in abundance between the three experimental groups. Two isoforms of tropomyosin and one isoform of actin were identified and these proteins have been implicated in previous studies in response to reduced salinity. The low number of proteins identified in this study and the heavy metal experiment highlights the difficulty in working with invertebrate species that are presently underrepresented in the DNA and protein sequence databases.

In the heavy metal experiment *P. canaliculus* were exposed to either 34.3 $\mu g\ l^{-1}$ Hg or 0.486 $mg\ l^{-1}$ Cd in the laboratory for 3 days. Control mussels were held in identical conditions without added metal. Over 100 proteins were detected which showed significant ($p < 0.01$) differences in abundance between control and metal treated groups but these proteins could not be identified using MALDI-TOF mass fingerprinting or tandem mass spectrometry. Tissue and time specific differences in metal uptake were observed.

Proteins which responded to heavy metals under laboratory conditions were compared to field samples from the Bay of Islands. Approximately 30 proteins were detected which appeared to be associated with the presence of heavy metals under both field and laboratory conditions. These results suggest that it may be possible to develop biomarkers for heavy metal contamination in *P. canaliculus*.

Based on the average concentrations of metals detected in the Bay of Islands, the amount of metal consumed through a typical diet containing shellfish would be below the provisional tolerable weekly intake (PTWI). However, because Māori, Pacific Islanders and Asians consume a greater quantity of seafood than the general New Zealand public a risk assessment for these groups was calculated. A survey of the frequency, amount and species consumed by these groups is suggested to enable an adequate risk assessment to be made.

ACKNOWLEDGEMENTS

First and foremost I would like to thank my whanau for their love, support, encouragement, advice and inspiration. I would especially like to thank my husband Jason Allen. Jason thanks for providing me with understanding and motivation. Thank you for picking me up and dropping me off late nights and weekends. I know you are not a taxi service, yet you did it without complaining (for the most part!). I couldn't have done this without you. Thank you to Mum, Dad, my 3 younger brothers: Dan, James and JK and extended whanau (uncles, aunties and cousins) for travelling to Wellington from Hawke's Bay to support me in the TPMF interview process and for encouraging me the whole way through. Thanks also to my 'in-laws': Ma, Pa, Lisa and Jon and the rest who I "acquired" during my PhD. See you all at the graduation celebrations!

I would like to thank my wonderful supervisors for their support, encouragement, advice and for editing and providing useful comments for this thesis. Thank you to Dr. Bill Jordan for the initial encouragement and advice in the development of my research topic. Bill you have been a mentor, supporter and a friend. Thank you to Dr. Jonathan Gardner for being involved in my TPMF application and reports and for providing useful feedback for this thesis. Thank you to Prof. Gary Raumati Hook for being my Māori mentor and an additional source of advice and support for this project. Thanks to Dr. Alan Clark for providing additional input into this project. Thank you to Dr. Gail Greening (ESR) who gave me the fantastic opportunity to become involved in her research programme with Ngāpuhi in the Bay of Islands. I have enjoyed working with such a wonderful group of scientists! Special thanks to Dr. Shirley Pledger who provided statistical support.

I would like to thank Ngahiwi Tomoana the CEO of Ngāti Kahungunu Iwi Incorporated (NKII) and Marei Apautu, Te Kaihautu at Te Taiwhenua o Heretaunga for their support and encouragement of this project. I would also like to thank Emma Gibbs, spokesperson for Ngāpuhi, and the generosity of the Ngāpuhi people for allowing me to work with mussels from their rohe. I would especially like to thank

Harry Mahanga for picking the mussels for me. These investigations have been initiated by the communities of Ngāti Kahungunu and Ngāpuhi, who necessarily have a stake in the outcomes. Publication of the results of these studies suitable for communication at a community level are planned, and will be separate to those planned for publication in scientific journals.

Thank you Liz Richardson, Dawn Reeves and Āwhina for your support. Liz I would not have gotten this far without you. Thank you.

Thank you to VUW for a PhD completion scholarship and FRST who provided funding: TPMF (VUWX0202) and ESR sub-contract (CO3X0301). I would also like to acknowledge NPOTM and Mai ki Poneke for support to attend a Māori doctoral writing retreat.

To all my 'non-science' friends. Thank you for your encouragement and support of me and this project. I think I can be a better friend now that this thesis is done!

Finally, I would like to thank the staff and students of SBS and the Science Faculty for their support and encouragement. Mary Murray and the admin staff for letting me know when that new important parcel has arrived and Lisa Bognuda for personal and professional support. Thanks also to Delwyn Carter, Patricia Stein, Celia Simpson and Shona de Sain. I would also like to specifically mention and thank my office mates, James Matthews, Rudi Schnitzler and Bruce Norris. Guys, your unique sense of humour while sometimes off-beat is always amusing – even if somewhat disturbing at times. It has been great sharing this experience together. I know that I'll never forget it. You have each endeared yourselves to me in your own special ways. Bruce with your offbeat positive nature, computer know-how and love of rugby, beer, whiskey and wine. It is nice to share interests. James for your sharing. Your enthusiasm for your research is always inspiring! Rudi for your caring. You always looked out for me and asked how I was doing, organised coffee, cakes and chocolate and tried to make our office homely. We really do have one of the best offices in SBS. Thank you to anyone I may have inadvertently left out.

Whaia Te Iti Kahurangi. To the stars through hard work.

CONTENTS

FIGURES

TABLES

CHAPTER 1: Introduction

1.1 Environmental toxicology

Ecotoxicology and environmental toxicology are two related fields of science which have evolved from classical toxicology i.e. the study of toxic substances. Ecotoxicology brings together the fields of ecology and toxicology and focuses primarily on the toxic effects of chemicals and radiation on levels of biological organisation from individual organisms to communities (Truhaut 1977). Environmental toxicology embraces both the disciplines of classical toxicology and ecotoxicology and includes risk management and regulation (Wright and Welbourn 2002; Figure 1.1). Environmental toxicology can include any toxic hazard in any environment but this thesis has focused on marine environmental toxicology.

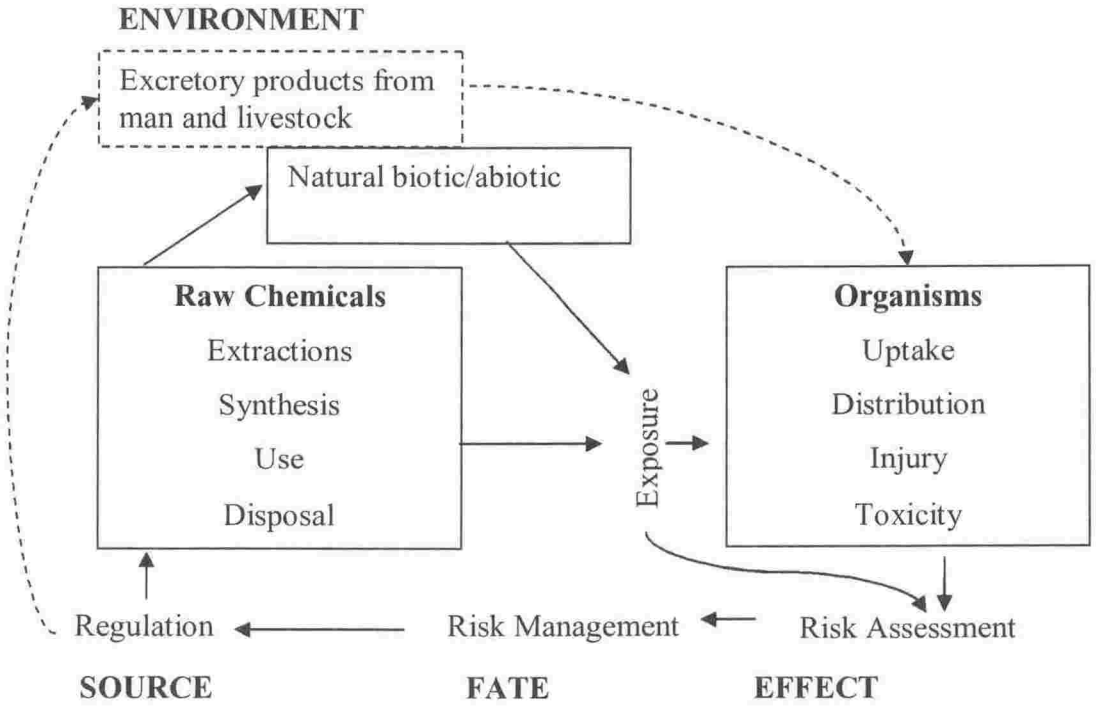


Figure 1.1: Factors included in the field of environmental toxicology
Diagram has been modified from Wright and Welbourn (2002)

1.2 Sources and types of marine contaminants

Contaminants enter waterways when rainwater flows over impervious urban surfaces (e.g. roads, roofs, car parks and construction sites), as a result of diffuse

agricultural run off (non-point sources) or from sewage and stormwater discharges (point source discharges). Potential contaminants include organic matter (e.g. leaf litter, sewage), heavy metals, organic chemicals (e.g. polycyclic aromatic hydrocarbons, organochlorines, dioxins), nutrients, pathogens and sediment, all of which can have negative impacts on marine ecosystems (Vincent and Thomas 1997; Bolton-Ritchie 2003). For example, low dissolved oxygen concentrations in many rivers in the United Kingdom, as a result of sewage discharges, led to a dramatic decline in fish fauna (Matthiessen and Law 2002). Polychlorinated biphenyls (PCBs) are persistent bioaccumulative pollutants which have been associated with developmental, neurological and immunological disorders in humans (Judd et al. 2004). Metals have been used in agriculture, industry and medicine and as such environmental contamination is relatively widespread (Caussy et al. 2003). Some metals such as Fe, Cu, Co, Mn, Zn and Cr are essential and can exert positive or negative biological effects dependent on concentration (Caussy et al. 2003). Other 'heavy metals' such as Hg, Pb, Cd and As are not known to be necessary in any biochemical processes and can cause severe adverse effects in organisms (e.g. Gochfeld 2003; Centeno et al. 2005; Phillips and Fowles 2005). This thesis used the following definitions: contaminant, an artificial increase over background concentrations, as opposed to pollution which implies harm to living resources or risks to human health (Preston 1989).

1.3 Monitoring marine contaminants

In the early phases of marine environmental monitoring, programmes principally sampled water or sediments from different sites and measured directly the amount of each relevant toxin without the inclusion of biological samples (Gregory et al. 1999; Rittschof and McClellan-Green 2005). However in most cases, except for the most extremely polluted sites, the amount of chemicals present in the water column was below the level of detection (Lam 2003), therefore providing limited information for monitoring studies. Shellfish are known to bioaccumulate (i.e. concentrate) trace contaminants from the environment to many thousands of times background concentrations (Sheehan and Power 1999). Bivalve molluscs in particular accumulate metals in their tissues approximately in proportion to the degree of environmental contamination (Goldberg 1983), providing a direct measure

of the aquatic pollution present in a region (Phillips 1986). Once established, many shellfish are long lived, easily sampled and their sedentary nature means that their geographical relationship to a pollution source can be easily ascertained. Shellfish therefore exhibit desirable characteristics to serve as sentinel organisms for environmental pollutants.

Because molluscs can indirectly monitor the state of the environment and act as an indicator of environmental health, they are often referred to as an example of a biomonitor or bioindicator species. Programmes using bioindicators to monitor marine pollution have been developed worldwide. The international Mussel Watch study is an example of one such monitoring program that gained substantial support (Goldberg 1975). Goldberg (1975) argued for a global 'Mussel Watch' to measure the concentration of certain contaminants in the tissues of different species of *Mytilus* (Bayne 1976). This research was initiated in the United States of America in 1960 and continues to the present day. Although most studies are concentrated in Europe e.g. Italy (Lionetto et al. 2001) and France (Roméo et al. 2003), and North America (Goldberg 1983; Roper et al. 1997; Beliaeff et al. 1998) biomonitoring programmes and preliminary studies have also been initiated in Australia (Richardson et al. 1994; Prest et al. 1995; King et al. 2006), South America (Baraj et al. 2003), South Africa (Gregory et al. 1999; Gregory et al. 2002) and countries in the Asia-Pacific region: India, Thailand, Indonesia, Hong Kong, Taiwan (Tanabe et al. 2000; Hung et al. 2001; Boonyatumanond et al. 2002) and Malaysia (Yap et al. 2002; Yap et al. 2003c; Yap et al. 2004a). In New Zealand there is limited data available from bioaccumulation studies of bioindicator species (reviewed in Section 1.6.3).

1.4 Biomarkers for marine contaminants

Despite providing valuable information, the 'Mussel Watch' programme is slowly being complemented and replaced by more sophisticated testing regimes. Chemical analysis, of water, sediment or of organisms, only provides a limited understanding of environmental well-being because the ecological consequences of pollution are biological, not chemical (Bayne 1976). Focus has therefore shifted to the development of biomarkers. *"Biomarkers are measurements of body fluids, cells, tissues or animals that biochemically, cellularly, physiologically, behaviourally or in energetic terms indicate the presence of contaminants in the environment*" (Chan et

al. 2002). Biomarkers should provide an early warning of deleterious effects on biological systems and also estimate biological effects due to contaminants (Lam 2003). To be useful a biomarker should respond in a predictable manner and only in response to a known stimulus (i.e. pollution).

There are a wide range of methodological techniques and biomarkers that can be used to indicate environmental contamination. Techniques include differential gene expression (Tanguy et al. 2005); RAPD (De Wolf et al. 2004); immunochemical (Boutet et al. 2002) and proteomics (Bradley et al. 2002). Biomarkers include genotoxin damage e.g. DNA adduct formation, gene mutation, DNA breaks (Depledge 1996); allozymes (Yap et al. 2004c), protein synthesis in general or of specific proteins e.g. glutathione S-transferase (Fitzpatrick 1995; Hoarau et al. 2002; Lyons et al. 2003; Yang et al. 2004), heat-shock proteins (Buckley et al. 2001; Snyder 2001) and metallothionein (Geffard 2002). Metallothioneins (MTs) ability to sequester heavy metals has led researchers to suggest that MTs are potentially useful as specific biomarkers for trace metal exposure (Mourgaud et al. 2002). Some attempts have been made in New Zealand to use the endemic greenshell mussel *Perna canaliculus* (Gmelin 1791) as a biomonitor (e.g. for monitoring trace elements, Kennedy 1986). Few, if any, studies to date have endeavoured to develop specific biological markers (biomarkers) within a local species, to indicate the presence of a particular pollutant (e.g. heavy metals).

1.5 Motivation for this study

1.5.1 Environmental contamination in New Zealand

The level of marine pollution around Aotearoa (New Zealand) is relatively low compared to that encountered around many other countries, particularly in the industrialised nations of the Northern Hemisphere (Ridgway and Glasby 1984). However, New Zealand's production and manufacturing industries, e.g. agriculture, horticulture, and pulp and paper mills, create a number of potentially damaging by-products such as heavy metal waste, pesticide and fertilizer run-off (Ridgway and Glasby 1984; Furness 2001; Quilbe et al. 2004). Additional man-made sources of contaminants include heavy metals (e.g. Pb, Zn) from road runoff and paint leaching into stormwater drains (Bolton-Ritchie et al. 1999; Botherway and Gardner 2002), raw or partially treated sewage (Rogers 1999), tailings and discharges from mines

(Livingston 1987; Harding 2005) and waste from galvanising and wood-treatment plants (Smith 1986). Heavy metals also enter the environment 'naturally' as a consequence of erosion from ore-bearing rock, wind blown dust, volcanic/geothermal activity, and forest fires (Gregory et al. 2002). These harmful contaminants make their way into waterways and have negative impacts on the biota of streams, rivers, lakes and oceans. For example, eels in the Waitangi River below the Ngāwha springs in the Bay of Islands were found to contain relatively high concentrations of mercury (Hg) due to contamination of riverine food webs (Chisnall 1997). Even moderate levels of metal pollutants and environmental stress in molluscs result in modified behaviour and life history traits (Calow 1991). There is therefore a need to develop effective methods for detecting and measuring the effects of pollution on aquatic biota.

1.5.2 Impact of environmental contaminants on human communities

The apparent effects of pollution on the marine environment cause anxiety within human communities. Māori people from the iwi of Ngāti Kahungunu in Hawke's Bay are concerned that contaminants may have led to the noticeable decrease in abundance or total absence of kaimoana (seafood) in historic shellfish gathering areas (pers comm. Ngahiwi Tomoana, CEO Ngāti Kahungunu Iwi Incorporated). In addition the Māori world-view recognises important relationships embedded in whakapapa (genealogies) between people and all living things (Williams 2001, pp 16). Maintaining a healthy marine ecosystem is therefore of major spiritual importance to Māori and is integral to their role as kaitiaki (guardians) of the environment. Possible detrimental side-effects to human health from eating shellfish containing unknown quantities of dangerous substances are also cause for concern.

In addition to the threats from heavy metals due to geothermal activity, seasonal influxes of visitors to the Bay of Islands occasionally overwhelm local sewage facilities causing fecal material to contaminate waterways. Runoff from septic tank systems at Okiato Point and the Kawakawa sewerage treatment plant are suspected of causing contamination of Waikare Inlet oyster farms (The Northern News 16 May 2002), which resulted in the closure of all 20 oyster farms in 2001 (The Northern Advocate 6 March 2006). Farmers are currently seeking $12m in compensation from the Far North District Council (The Northern Advocate 6 March

2006). In Auckland the discharge of sewage by commercial and recreational boats is suspected of causing the contamination of commercial oyster growing areas (Simmons et al. 2001). Similar incidents have been reported in the Bay of Islands (Jarman 1995). Sewage spills have led to many documented cases of viral gastroenteritis (food poisoning), which have been traced back to eating shellfish contaminated by sewage from treatment plant failures or boat discharges (Simmons et al 2001). Similar outbreaks have also occurred overseas (e.g. Morse et al. 1986). An effective and efficient way to monitor the levels of contamination within the marine environment could protect the environment and safeguard human health.

1.5.3 The development of this project: Methods to detect and identify heavy metal biomarkers within *P. canaliculus*

A scientific research team led by Dr. Gail Greening from Environmental Science Research Ltd (ESR) was contracted by the Foundation for Research, Science and Technology (FRST) to examine 'shellfish health' (Contract # C03X0301). Specifically, 1) by investigating viral contamination of shellfish in collaboration with Ngāpuhi (the results of which will not form any part of this thesis), and 2) by studying the effects of heavy metal contamination on greenshell mussels (this is a VUW sub-contract to ESR, the results of which will form part of this thesis). The broad aim of this study was to identify biomarkers for heavy metal contamination in *P. canaliculus* and to provide preliminary data on the uptake, distribution and effect of heavy metals on *P. canaliculus*.

The remainder of this chapter will introduce key concepts and issues, discuss their general relevance to this thesis, and highlight gaps in the current knowledge that this thesis will address. The thesis structure and the outcomes of this research as they relate to each of the remaining chapters are outlined at the end of this Introduction (Section 1.7).

1.6 Background to this project

1.6.1 Method of induction for biomarker development

1.6.1.1 Heavy Metals

The reason for focusing on heavy metal contamination has already being outlined (Section 1.5.1 and 1.5.2). Briefly, there are many potential sources of heavy metal contamination in New Zealand and this issue is of concern to iwi, especially in the Bay of Islands where Hg occurs naturally from local geothermal sources, potentially effecting human and environmental health. Biomarkers of elevated heavy metal concentration detected in this study focused on responses induced by either Hg or Cd. Hg was chosen because it is relevant to the field work conducted on greenshell mussel samples from the Bay of Islands. In previous studies Cd has been demonstrated to induce the largest response of metal binding proteins such as MT (e.g. Geret 2002). Properties of these chemicals, in brief, are given below.

Mercury (Hg)

All forms of Hg are toxic. In seafood, Hg is most commonly found in the most toxic organic form, methyl-mercury (MeHg). In the natural environment inorganic Hg released from industrial processes undergoes chemical transformations, the most significant of which is methlylation; resulting in the organic substances, monomethlymercury and dimethylmercury (Wright and Welbourn 2002). Freshwater and saltwater ecosystems are major sites of MeHg production (Wright and Welbourn 2002). MeHg can cause severe disruption of the developing central nervous system, resulting in retarded mental and physical development and is a serious cumulative toxin that bioaccumulates up the food chain. Predatory fish, for example tuna, often contain high levels of MeHg e.g. in the Pacific Islands including Samoa (Kraepiel et al. 2003). In Minamata (Japan) close to 100 deaths, over a period of several years, and many serious illnesses were caused by exposure to fish that had accumulated Hg released from a chemical plant (Wright and Welbourn 2002). The NZ Ministry of Health advises pregnant women against eating certain types of fish, because foetuses and infants are much more sensitive than adults to Hg poisoning and are therefore at higher risk (Vannoort 2000). Hg is of particular relevance to this study because of links between the Ngāwhā geothermal region and

the Waitangi River in the Bay of Islands (see Chapters 7). In the present studies, experimental induction of biomarkers using $HgCl_2$ has been undertaken and is compared with 'wild' samples taken from sites with varying degrees of contamination from the Bay of Islands.

Cadmium (Cd)

Cd has no known function in biochemical processes and is toxic, even in low concentrations. For example, inhalation causes respiratory tract and kidney problems which can be fatal. Ingestion of Cd causes immediate poisoning and damage to liver and kidneys, and osteomalacia (softening of the bones) may also result (Vannoort 2000). Compounds containing Cd are carcinogenic. Cd also bioaccumulates in organisms and ecosystems, and marine animals are especially sensitive to Cd toxicity (Material Safety Data Sheet, Sigma-Aldrich). One possible reason for Cd toxicity is that it interferes with the action of zinc-containing enzymes. Zinc (Zn) is very important in biological systems, and Cd, being similar chemically, acts as a competitive inhibitor of Zn dependent processes (Brzoska and Moniuszko-Jakoniuk 2001). Cd is also a very powerful inducer of MT, a potential biomarker of heavy metal pollution (Section 1.4). As Cd is often associated with geothermal activity, significant contamination at the Waitangi site may also occur. Cd biomarkers have also been investigated in the current study.

1.6.1.2 Low salinity

The overall aim of my PhD research was to assess the potential to develop biomarkers for heavy metal contamination in *P. canaliculus*. However, the study of proteomic expression profiles generated by heavy metal stress in *P. canaliculus* required the use of Cd and Hg, two potentially dangerous heavy metals. Therefore, before conducting the heavy metal experiment a pilot study was performed using low salinity as the stress inducer. The same experimental procedures, including DIGE were used for the salinity and heavy metal experiments (Chapter 4).

1.6.2 A proteomic approach

In metal bioaccumulation surveys (e.g. Mussel Watch) analysis of individual chemicals of concern are being replaced by analyses of groups of pollutants which

have a common impact upon marine organisms (Goldberg 2000). The same rationale can be applied to the development of biomarkers, i.e. suites of biomarkers which show reproducible responses to groups of pollutants should be sought. To date, most previous studies generally focused on only one or a very small range of specific markers. For example glutathione S-transferase, (e.g. Fitzpatrick 1995) or heat-shock proteins (for a review see Feder 1999), in response to a single challenge (e.g. temperature, toxic chemicals, salinity or heavy metals).

Proteomics is one technique that is currently being developed to identify a range of biomarkers which respond to specific inducers (Chapter 4). In the laboratory, environmental proteomic techniques have been applied to mussels exposed to copper, Aroclor and salinity stress (Shepard et al. 2000), rainbow trout treated with diazinon, nonylphenol, propetamphos and sewage treatment plant effluents (Bradley et al. 2002) and clams exposed to model pollutants (Rodríguez-Ortega et al. 2003b). More recently field experiments have been performed to identify protein patterns associated with *Mytilus edulis* exposed to heavy metals and polyaromatic hydrocarbons (Knigge et al. 2004) and to identify differences in carbonylation patterns between *M. edulis* from polluted and reference sites (McDonagh et al. 2005). This study reports changes in protein abundance patterns in response to either elevated heavy metal concentrations or decreased salinity. Gene transcription is measured indirectly through protein abundance. Although the term 'protein expression' is often used in the literature, this term is technically incorrect and will not be used in the current study unless in reference to work by previous groups.

1.6.3 Marine monitoring in New Zealand

Heavy metals, pesticides and organochlorides are found in marine sediments in inlets and harbours around New Zealand. Most previous marine monitoring studies in New Zealand have focused on marine sediment pollution in main urban areas such as Manukau and Waitemata Harbours in Auckland (Glasby et al. 1988; Williamson et al. 1996; Ahrens and Depree 2004) and Wellington Harbour (Stoffers et al. 1986; Bolton-Ritchie et al. 1999). Some work has also been conducted on sediment cores from Mapua (Waimea Inlet; Hendy and Peake 1996) and geothermal marine sediment enrichment in the Bay of Plenty (Propp et al. 1994) and an off-shore extension of the Taupo volcanic zone (Stoffers et al. 1999). However, as indicated

in Sections 1.3 and 1.4 most studies now focus on biological indicators and biomarkers.

Some attempts have been made to combine sediment, water and biological data to assess the biomonitoring capabilities of New Zealand aquatic organisms. Previous studies have measured the effect and accumulation of heavy metals in a range of organisms including barnacles (*Elminius modestus*) from Waitemata and Manukau Harbours (Zauke 1992); amphipods (*Transorchestia chiliensis*, *Talorchestia quoyana* and *Orchestia tenis*) from Dunedin (Rainbow et al. 1993) and Christchurch (Marsden et al. 2003); copepods (mesozooplankton; Mackie and Hunter 2005) from subtropical and subantarctic waters and gastropods (*Lepsiella scobina*) from throughout New Zealand (Smith and McVeagh 1991; Stewart et al. 1992; Smith 1996). Heavy metal concentrations have also been estimated from blood of South Island Pied oystercatchers (*Haematopus ostralegus finschi*) in Auckland (Thompson and Dowding 1999). High Hg concentrations have been observed in eels (*Anguilla dieffenbachii*) near the Ngāwhā geothermal field (Chisnall 1997) and rivers in Otago (Redmayne et al. 2000). MT mRNA levels have been investigated in the common New Zealand bully (Laurie 2004) and this work is discussed elsewhere (Chapter 2).

Studies on bivalves have included cockles from Whangateau and Manukau harbours (De Luca-Abbott 2001) and Dunedin (Peake et al. 2006); oysters (*Crassostrea gigas*) from Manukau Harbour (Pridmore et al. 1990; Roper et al. 1991) and freshwater mussels (*Hyridella menziesi*) from the Waikato river. Studies have also been conducted on Bluff oysters (Brooks and Rumsby 1967; Frew and Hunter 1996), wild and farmed oysters from throughout New Zealand (Winchester 1980; King and Miller 1989), greenshell mussels (*P. canaliculus*) from Wellington (Kennedy 1986) and *Ostrea chilensis* (formerly known as *Ostrea sinuata*), *Pecten novaezelandiae* and *Mytilus galloprovincialis* (formerly known as *M. edulis aoteanus*) from Nelson (Brooks and Rumsby 1965). A comprehensive survey measuring the concentrations of metals in 13 species of edible molluscs from 199 sites throughout New Zealand was conducted (Nielsen and Nathan 1975). Cd binding proteins have been studied in Bluff oysters (Nordberg et al. 1986). To date, I am aware of only one study that has conducted laboratory based metal uptake experiments using a New Zealand species which was conducted on amphipods (King et al. 2006).

1.6.4 *P. canaliculus* as a biomonitor

Although the blue mussel (*M. galloprovincialis*) is widespread around New Zealand and has been used in other studies world-wide, this study focuses on developing biomarkers for heavy metal pollution in the endemic greenshell mussel *P. canaliculus*. The greenshell mussel is very important in New Zealand economically, generating almost NZ$200 million worth of aquaculture exports each year (www.seafood.co.nz), culturally, being a species of significance to Māori and ecologically, being widespread throughout New Zealand. According to a recent overview and bibliography of research conducted on *P. canaliculus* (Jeffs et al. 1999) and my own literature review there has not been any development of biomarkers in *P. canaliculus* to date. Although previous studies have suggested that there is a need to develop suitable marine biomarkers in New Zealand, attempts to date have focused on other species (amphipods, barnacles, cockles and oysters, see Section 1.6.3).

P. canaliculus is endemic to New Zealand and consequently the metal accumulation rates obtained in this study cannot be directly compared to those obtained from other studies worldwide. It is not valid to compare absolute accumulated metal concentrations in biomonitors interspecifically, although cross correlations of relative bioavailabilities of heavy metals to different biomonitors at the same sites can be calculated (Rainbow 1995). Widespread cosmopolitan species such as the algae *Ulva lactuca,* mussels of the genera *Mytilus* and *Perna*, oysters of the genera *Ostrea* and *Crassostrea*, the barnacles *Balanus amphitrite* and *Tetraclita squamosa*, and the talitrid amphipod *Platorchestia platensis* can however be used for intraspecific comparisons of metal bioavailabilities between geographical areas (Rainbow 1995). A major caution in the use of such multinational biomonitors remains the need for reliable taxonomic identification (Rainbow 1995). For example, in New Zealand the specific status of blue mussels (Genera *Mytilus*) is still unclear (Gardner 2004). Until such taxonomic issues are resolved in New Zealand *P. canaliculus* is a good choice for monitoring because it is easily distinguished from *Mytilus* spp. *P. canaliculus* also has the advantage of being more widely distributed around New Zealand than the blue mussel (Morton and Miller 1968) which increases its attractiveness as a biomonitoring species. In short, *P. canaliculus* was considered

a better choice for biomarker development in New Zealand than the blue mussel *M. galloprovincialis*.

1.7 Thesis structure, Aims and Outcomes

1.7.1 Thesis structure

This chapter was written to give some background information on subjects relevant to this thesis topic in general and is intended to highlight gaps in current knowledge that this research will address. Each of the subsequent chapters (excluding CHAPTER 3: General Methods and CHAPTER 8: Conclusions/Future Directions) contain results from the current study and will provide data on 'stress' biomarkers in *P. canaliculus*.

1.7.2 Aim

The aim of this study was to detect and identify 'stress' biomarkers in *P. canaliculus*, with the ultimate goal of determining the suitability of *P. canaliculus* to act as a biomonitor or indicator of contaminants in the New Zealand coastal marine environment. To determine the potential of *P. canaliculus* to act as a biomonitor, relevant biomarkers within this organism must first be identified and characterised. Biomarkers chosen for analysis in this study were induced by exposure to extreme salinity changes and heavy metals.

1.7.3 Methodology

Molecular biological techniques have been used to provide new information on MT genes present in *P. canaliculus* (Chapter 2). MT is an important and highly relevant protein for studies on metal toxicity. Proteomic tools have been used to visualize, detect and identify proteins that show a response to stress induced by salinity changes or exposure to heavy metals (Chapters 4, 5 and 6). The methods used in this research are an important feature of this project as very few previous studies in this field have used proteomics to identify suites of protein biomarkers in molluscs (Olsson et al. 2004). Instead, most studies have focused only on certain proteins that are recognised as relevant in dealing with a known stress. A holistic approach, as used in this study, allows for the detection of novel proteins and pathways involved in detoxification. Additionally, this study uses two-dimensional

gels, and new DIGE fluorescent dyes (CyDyes™) and DeCyder™ software (Amersham). DIGE is very sensitive technique that can detect small changes in protein abundance (Chapter 3). Body burden data (i.e. that amount of metal accumulated in mussel tissue) has also been collected (Chapter 5, 6 and 7). Sub-samples were dissected during the metal experiment and sent to Environmental Laboratory Services (Lower Hutt, New Zealand) for chemical analysis, providing data on the rate of metal accumulation and providing complementary data to the biomarker analysis. The responses to heavy metals have been studied in both an experimental laboratory setting, and using mussels from field sites with varying degrees of metal contamination. Again, this is an important component of this research as similarities were sought between protein expression profiles produced in the lab trials, versus those generated from 'wild' mussel populations, to see if potential biomarkers could withstand a real life test. In addition, this study has assessed the potential health risk to human consumers of dietary heavy metal intake from *P. canaliculus*.

1.7.4 Outcomes

- Partial MT sequences from *P. canaliculus*, allowing comparisons with other mollusc species worldwide (Chapter 2).
- Detection of proteins involved in a 'salinity stress response' (Chapter 4)
- Detection of proteins involved in 'heavy metal response' (Chapter 5).
- Protein abundance profiles in response to heavy metals under laboratory conditions compared with wild populations from the Bay of Islands (Chapter 6).
- Recommendations on proteins that could possibly be used as biomarkers in *P. canaliculus* (Chapter 6)
- Estimation of human dietary intake of metal from *P. canaliculus* in the Bay of Islands (Chapter 7)
- Summary of results and conclusions (Chapter 8)
- Recommendations for areas of further study (Chapter 8).

CHAPTER 2: Isolation and characterisation of partial metallothionein gene sequences from *P. canaliculus*

Abstract

Metallothioneins (MTs) are potentially useful biomarkers for environmental heavy metal contamination, yet to date limited information is available on MT variability and response to contamination in New Zealand organisms. This research provides preliminary MT amino acid and nucleotide sequence data from the New Zealand endemic green mussel *P. canaliculus* isolated by cloning PCR products from genomic DNA. Nine MT exon 2 sequences were isolated, some of which are characterised by unusual features, including the presence of atypical tyrosine and histidine residues and lower than usual numbers of metal binding cysteine residues. A putative promoter sequence, active protein binding sites and a single metal responsive element (MRE) were identified in the 5' untranslated region of the *P. canaliculus MT* gene sequence, based on alignment with *P. viridis* sequence, providing information about transcriptional control. Sequencing results from uncloned samples show evidence of heterozgosity, suggesting *MT* allelic polymorphism. The predicted MT amino acid sequences from *P. canaliculus* have been aligned with MT sequences from other molluscs. Comparing *P. canaliculus* MT amino acid and genetic sequences with other species using FASTA and BLAST reveals greatest similarities with other invertebrates.

2.1 Introduction

2.1.1 MT: a biomarker for heavy metal pollution

Mussels are long lived, easily sampled, sedentary and known to accumulate metals in proportion to their environmental exposure (Goldberg 1983) making them ideally suited to act as sentinel species for pollution (e.g. Mussel Watch, Goldberg 1975). For over 30 years, tissues from mollusc species has been chemically analysed to detect and monitor pollutants within the marine environment. However, chemical analyses, of water, sediment or organisms, only provides a limited understanding of environmental health because the ecological consequences of pollution are

biological, not chemical (Bayne 1976). Therefore the current focus has shifted away from the analysis of metal content in mussel tissue toward developing biomarkers for contamination, such as MT.

MTs are low molecular weight (6-7 kDa), mainly cytoplasmic metal-binding proteins containing a large number of cysteine residues (30%) and typically containing low levels of aromatic residues and histidine. The cysteine residues conform to various conserved Cys-X_n-Cys motifs (where X is an amino acid other than cysteine; Kagi and Shaffer 1988). Although the exact physiological role(s) of MTs are still unclear it is widely accepted that MTs are involved in cellular essential metal (Cu, Zn) regulation and heavy metal (Cd, Hg) detoxification (Kojima and Kagi 1978). Numerous studies have reported increased MT expression in organisms after exposure to heavy metals (reviewed by Cosson 2000), suggesting a protective mechanism for MTs and indicating that MTs could act as a biomarker for heavy metal contamination (Mourgaud et al. 2002). For the purpose of this chapter references to *MT* genes will be italised whereas MT proteins or mention of MT in general will not.

2.1.2 *MT* genes

MT genes form relatively complex families consisting of multiple paralogues. Paralogues are genes which are created by duplications within the genome of a single species. Paralogous sequences complicate primer design because primers designed in exons will amplify numerous paralogues simultaneously resulting in ambiguous sequencing reactions (Section 2.2.2.1). *MT* genes are typically characterised by 3 coding exons (Tanguy et al. 2003; see Figure 2.1), although some additional non-coding exons have been observed in the 3' region of the sea urchin *Strongylocentrotus purpuratus* (Nemer et al. 1991) and the 5' region of the pacific oyster *C. gigas* (Tanguy et al. 2001). These 3 exons usually code for 2 structural metal binding domains: an α-domain that commonly contains 11 or 12 cysteine residues and binds 4 bivalent ions, and a β-domain that contains 9 cysteine residues and binds 3 bivalent ions. Domains are capable of binding metals independently and are separated by a short 'linker or hinge' region to form a 'dumb-bell' like structure (Vašák 2005). Some invertebrate MTs deviate from this typical structure, e.g. crustaceans possess two β-domains capable of binding six metal cations (Lerch et al. 1982) and earthworms contain a single α-domain capable of binding 4 Cd ions

(Gruber et al. 2000). In many invertebrates, the alpha domain is N-terminally encoded, whereas in vertebrates the reverse is the case, with the beta domain being N-terminally encoded (Nemer et al. 1985).

MT nomenclature splits the MT superfamily into three major classes (Nordberg and Kojima 1979; Fowler et al. 1987). Class I includes MTs with cysteine residue arrangements closely related to those identified in mammals, Class II comprises MTs without conserved cysteine locations, and Class III contains the remaining sequences, including plant phytochelatin. As the number of MT protein and gene sequences obtained continued to grow the nomenclature was updated to include sequence similarities and phylogenetic relationships (Binz and Kagi 1999). Currently MTs are separated into 15 families among the animal, plant, prokaryote and fungal kingdoms, further divided into subfamilies, subgroups and isolated isoforms and alleles. Mollusc MTs are Class I, Family 2 MT sequences and most mussel MTs fit into subfamily mo1 (21 conserved cysteines) and mo2 (23 conserved cysteines). For more details refer to: http://www.biochem.unizh.ch/mtpage/classif.html

Study of the structure and regulation of *MT* genes is important for better understanding of both their physiological roles and the utilisation of this metalloprotein as a biomarker for heavy metal exposure (Ceratto et al. 2002). Among molluscs 17 MT isoforms have been characterised in the blue mussel *M. edulis* (Mackay et al. 1993; Barsyte et al. 1999; Leignel et al. 2003, GenBank, unpublished) 2 in the tropical green mussel *P. viridis* (Khoo and Patel 1999), 1 in the freshwater mussel *Dreissena polymorpha* (Engelken and Hildebrandt 1999), 16 in the eastern oyster *C. virginica* (Tanguy et al. 2001; Tanguy et al. 2002, GenBank, unpublished) and 5 in the Pacific oyster *C. gigas* (Tanguy and Moraga 2001).

2.1.3 MT characterisation in New Zealand aquatic species

MT protein abundance and gene sequences have been determined in a few New Zealand fish, including the common New Zealand freshwater bully *Gobiomorphus contidianus* (Laurie 2004). Wide variations in individual hepatic MT mRNA levels occurred in field samples from different locations, leading to difficulties in correlating MT levels with environmental metal concentrations. Contaminant exposure in fish is uncertain due to their mobility, thus limiting their utility as biomonitors. Shellfish are considered more suitable bioindicators because

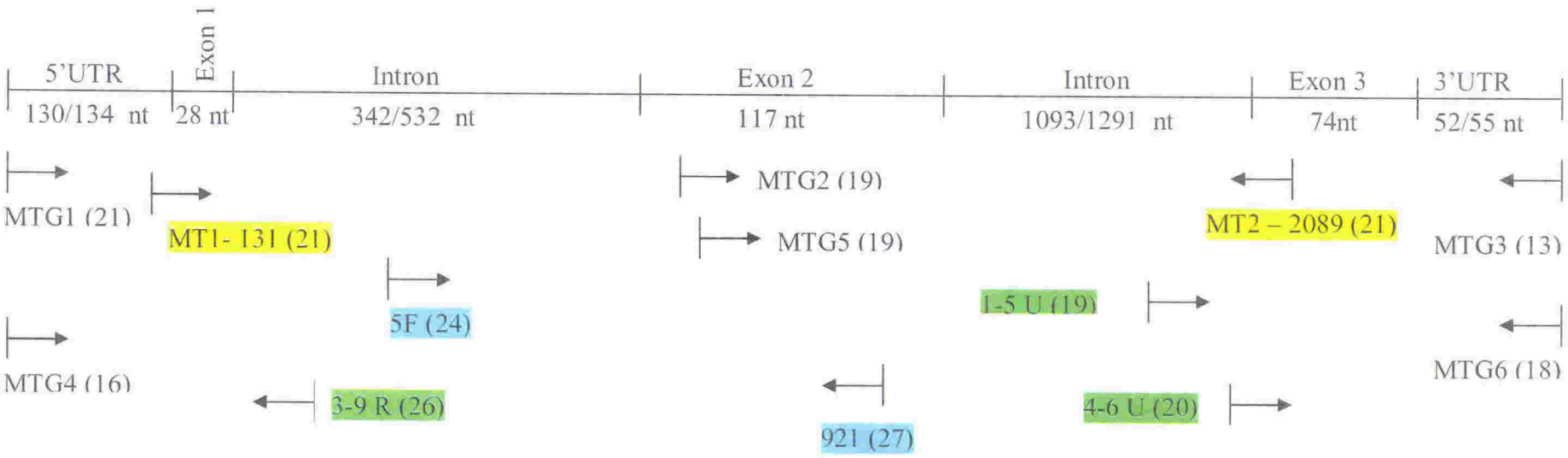
of their sedentary lifestyles and propensity to accumulate toxins in direct proportion to their exposure (Goldberg 1983). Despite this no biomarkers for heavy metals in New Zealand shellfish species have been developed to date. Attempts to characterise heavy metal biomarkers in the New Zealand endemic green mussel *P. canaliculus* are described elsewhere in this thesis (Chapters 4-6). This chapter focuses on characterising *MT* genes from *P. canaliculus*.

2.1.4 Scope of this research

This study reports preliminary MT amino acid and nucleotide sequence data from the New Zealand endemic green mussel *P. canaliculus*. The predicted MT amino acid sequences from *P. canaliculus* were aligned with MT sequences from other molluscs, providing a summary of what is currently known about MT amino acid sequences in these organisms. Comparisons of *P. canaliculus* MT sequences with other species were performed using FASTA and BLAST to examine evolutionary relationships between MT sequences.

Figure 2.1: Generic metallothionein gene

Approximate lengths of exons, introns and UTR's were deduced using two published *P. viridis* genomic MT DNA sequences (Khoo and Patel 1999).
Approximate locations of primers used in the current study are indicated. The number in parenthesis after the primer name indicates the length of the primer. Only those primers which generated sequences are shown below.



Key: Primers as designed by Khoo and Patel (unshaded); Primers designed based on *P. viridis* sequence; Specific *P. canaliculus* primers based on cloned sequences; Primers designed for genomic walk (Section 2.2.4) based on sequences from *P. canaliculus*

2.2 Methods

2.2.1 Extraction of DNA

P. canaliculus samples were collected from Seatoun Harbour, Wellington, New Zealand (n = 5). DNA was extracted from gill, digestive gland and mantle edge tissue using a DNeasy® Tissue Kit (Qiagen) following the directions of the manufacturer. DNA extracted from digestive gland tissue was found to be degraded and unsuitable for use, probably due to the action of digestive enzymes present within this tissue. In contrast, DNA isolated from gill and mantle samples was consistently found to be of high quality (i.e. high molecular weight and no evidence of degradation). For this reason, the protocols outlined in this chapter were developed using DNA isolated from gill and mantle samples. Once methods for the isolation of *MT* gene sequences from *P. canaliculus* were established, additional samples from South Island specimens were also sequenced to see if there were any genetic differences compared to North Island specimens. DNA from South Island samples had previously been isolated from gonad tissue using a Proteinase K extraction method. [For details refer to Smita Apte's PhD thesis (Apte 2001).]

2.2.2 PCR

Initially, primers developed by Khoo and Patel (1999) to amplify *MT* sequences from the tropical green mussel species *P. viridis* were trialled. The primers by Khoo and Patel (1999) were based on 5' flanking sequence data (MTG1 and MTG4) and cDNA data (MTG2, 3, 5 and 6; Figure 2.1). *P. viridis* DNA was isolated from mantle edge tissue of specimens collected from the Philippines and Vietnam that were obtained from Smita Apte (at Victoria University) and were used as a positive control.

2.2.2.1 Primers

The primers developed by Khoo and Patel (1999) failed to generate PCR products, in both *P. canaliculus* and the positive control, so a range of additional primers were designed based on *P. viridis* genomic MT sequences using the primer design software Oligo 6 (Figure 2.1 and Table 2.1). Primer design focused on exon MT sequences which were more likely to be conserved between *P. viridis* and *P. canaliculus* and also allow for the simultaneous amplification of multiple MT

paralogues from *P. canaliculus* (Figure 2.1). After *P. canaliculus* specific sequence had been generated by cloning, additional primers specific to *P. canaliculus* MT paralogues were designed within intronic regions (Figure 2.1). Primers which successfully generated PCR products are reported in Table 2.1, whereas the full list of primers trialled is reported in Appendix 1. Due to difficulties experienced during primer design, RT-PCR quantification of MT transcript levels was not attempted.

Primer Name	Primer Sequence	Length	Tm	Ref
MTG1	5' TGC CAC GGC GTG CAC ACG GCG 3'	21 nt	80 °C	a
MTG2	5' CTG TCT GTG TGG TAC TGG G 3'	19 nt	50 °C	a
MTG3	5' ATG GGA ACT AAA G 3'	13 nt	30 °C	a
MTG4	5' GCAAGC GGC ACA CGG C 3'	16 nt	65 °C	a
MTG5	5' ATC TGT GGT ACT GGG TGC A 3'	19 nt	51 °C	a
MTG6	5' ATT ATG GGA ACT AAA GAA 3'	18 nt	43 °C	a
131	5' AAA TAT GCC TAG CCC TTG TAA 3'	21 nt	54 °C	b
2089	5' ACA CTC TCC TGA AAA TAC ATC 3'	21 nt	49 °C	b
5F	5' AAA TAG CAC CAA AAA CAA AAT ACG 3'	24 nt	57 °C	c
921	5' TAC TTC CAC ACT TAC AGG CAT CAC CAC 3'	27 nt	62 °C	c
3-9 L	5' TTA TGT TAG CGG TTT CAC TTA CTT TC 3'	26 nt	70 °C	d
1-5 U	5' TCG CAT GAT AGT GAA GTT A 3'	19 nt	41 °C	d
4-6 U	5' GAT GTA TTT TCA GGA GAG TG 3'	20 nt	41 °C	d

a: Primers as designed by Khoo and Patel (1999) for *P. viridis*
b: Generic primers designed for the current project using Oligo 6, based on *P. viridis* sequences
c: Specific primers based on *P. canaliculus* sequences isolated by cloning PCR products from genomic DNA
d: Specific primers used for 'Genomic Walk' (Section 2.2.4).

Table 2.1: Primers

2.2.2.2 PCR conditions

PCR reactions were performed in 25 µl using the following reagents and concentrations; 10X Qiagen® PCR buffer, dNTPs (2.5 mM each, Pharmacia®), forward and reverse primers 10 µM (Biostrategy), Qiagen® *Taq* Polymerase (0.75U), Template DNA (~20 ng). The target MT sequences were amplified in an Applied Biosystems GeneAmp® PCR system 2700 cycler set to the following conditions: 94°C for 3 min, then 25 cycles of denaturation at 94 °C for 30 s, primer annealing at 55 °C for 30 s, extension at 72 °C for 30 s. Alternative cycling parameters were also trialled e.g. touchdown PCR; 60/50 °C, 55/45 °C etc, gradient PCR using a MJ

Research Gradient PTC-200 Peltier thermal cycler, increasing the number of cycles (up to 40X), increasing the concentration of *Taq* (up to 1.25 U), and use of additional primer pairs (Appendix 1). *P. viridis* DNA obtained from Smita Apte (Apte 2001) was used as a positive control.

2.2.2.3 Horizontal agarose electrophoresis

The purity and quantity of the MT PCR products were analysed using electrophoresis on 0.8% LE® agarose gels (FMC Bioproducts, Appendix 2). A 5 μl aliquot of the PCR product from each sample plus 0.5 μl of gel loading buffer (Blue Juice, Invitrogen) was loaded into each well. One or more lanes were loaded with 3 μl of 123 bp ladder (Invitrogen) size standard to determine the relative size of the DNA products. The gel was electrophoresed at 90 V for 90 min in Tris Borate Buffer (TBE: Appendix 2) containing 0.4 μg ml^{-1} ethidium bromide. Polaroid® and digital photographs of the gels were taken under UV transillumination with a Polaroid® MP4 Land camera using f5.6 aperture and 1 s exposure settings, or a digital Kodak Gel logic 100 imaging system.

If the results above were satisfactory, a 1% NuSieve® GTG® (Cambrex BioScience) low melting point agarose gel electrophoresed in Tris Acetate (TAE) buffer (Appendix 2) was used to purify the remaining (~20 μl) PCR products. Single PCR bands corresponding to a single PCR product were excised and purified using MinElute™ (Qiagen) spin columns, following manufacturer's instructions.

2.2.2.4 DNA cycle sequencing protocol

Reactions for sequencing were carried out using 8.0 μl Big Dye Terminator Mix™ (Applied Biosystems), 3.2 μl either forward or reverse primer (diluted to 1mM), purified DNA template (30-90 ng) and ddH_20 to make up to required volume (20 μl). Quarter reactions (using quarter of the recommended volumes shown above) were routinely used without compromising sequences obtained. The cycling parameters were 25 cycles of 94 °C for 30 s, 50 °C for 15 s and 60 °C for 4 min. Products were precipitated using an ethanol/sodium acetate protocol and were either sequenced in-house using the ABI Prism Sequencer 377 (Applied Biosystems), or samples were sent to be commercially sequenced using an ABI3730 Capillary Sequencer (Applied Biosystems) at the Allan Wilson Centre at Massey University in Palmerston North, New Zealand.

2.2.3 Separation of multiple PCR products

PCR reactions resulting in clear strong bands that were unable to be sequenced using direct sequencing of amplification products were cloned using a TA® TOPO® PCR cloning kit (Invitrogen) for sequencing. Gill (1Gi) and mantle (5Ma) DNA isolated from Seatoun samples was amplified and cloned initially. Gonad DNA from South Island samples, Riverton (RIV31) and Fiordland (FIO18), was also cloned for comparison. Primers 131 and 2089 were used for amplifying MT sequences isolated from all 4 mussels (Figure 2.1, Table 2.1) whereas primers MTG5 and 2089 were used for Seatoun mussels only (Figure 2.1, Appendix 4). Primer MTG5 is abbreviated to G5. The PCR reaction conditions used for primer pair 131/2089 followed the protocol outlined in section 2.2.2.2, however 40 cycles were performed and the annealing temperature was lowered to 45 °C. Directions for cloning followed those outlined by the manufacturer. Briefly, amplified PCR products (purified using MinElute™) were mixed with a TOPO® vector (supplied) and transformed into competent DH5α *E. coli*. Transformants were grown on ampicillin selective plates overnight at 37 °C. Single colonies were picked and grown overnight in L-broth with additional ampicillin added. Plasmids were isolated using a FastPlasmid™ Mini kit (Eppendorf) following the instructions of the manufacturer. Primers provided with the kit were used to amplify the inserts from the plasmids and sequencing proceeded as described in section 2.2.2.4.

2.2.4 Genomic Walk

Genome walking identifies unknown regions flanking a known DNA sequence by PCR and is useful in defining gene-regulatory regions. The Universal Vectorette™ System (Sigma) is a unidirectional PCR method that allows specific amplification of any uncharacterized sequence adjacent to a known region (Lilleberg and Patel 1998). Briefly, *P. canaliculus* DNA was digested with a restriction enzyme (*Eco*RI, *Hind*III, *Bam*HI, or *Cla*I) and ligated to the appropriate corresponding vectorette unit. PCR was conducted using one primer directed at the vectorette unit and a custom primer targeting the known *MT* DNA sequence (determined using clone sequence from section 2.2.3. above; Appendix 1). This allowed PCR amplification of DNA between the known sequence and the restriction site used to cut the target DNA. This fragment was then sequenced using the

vectorette primer provided with the kit, to provide 5' or 3' sequence surrounding the targeted MT gene. Primers were designed within the intron sequence of the same paralogue (Figure 2.1) but a nested primer set was not used. Because the primers were not sufficiently 'inset' a confident alignment with previously isolated exon 2 sequences was not possible. For this reason, the Exon 1 and Exon 3 sequences obtained from the genomic walk have been recorded separately in Figure 2.7.

2.2.5 Enrichment and isolation of MT DNA sequences

The following protocol was modified from that used by Hamilton et al. (1999) for the enrichment and isolation of microsatellite sequences. Although it was not expected that this technique would transfer to genomic MT sequences, it was an opportunity to develop additional technical skills. Briefly, *P. canaliculus* genomic DNA was digested with restriction enzymes to create fragments ranging in size from 200-1000 base pairs. A SNX linker was added to these fragments (SNX forward 5' – CTAAGGCCTTGCTAGCAGAACG-3'; SNX reverse 3' – AAAAGATTCCGGAA CGATCGTCTTCGp -5'). The poly-A tail polarizes the linker so that only one end can be used as a blunt end for ligation. These genomic–linker ligated DNA fragments were amplified using PCR, made single stranded by heating to 95 °C and then biotinylated MT oligos were added (Appendix 1). The biotinylated MT oligos should anneal to genomic fragments that have complementary sequence. Hybridised MT-oligo and genomic DNA were captured with magnetic beads and PCR amplified, thereby enriching for fragments of DNA containing MT sequences. This subset of DNA fragments were ligated into a plasmid (pBluescript) and transformed into *E. coli*. Colonies were picked and screened for MT gene inserts using PCR. Inserts containing MT sequence should amplify when the biotinylated MT primer is used in conjunction with the plasmid primer. However, no PCR products were produced, suggesting that none of the plasmids contained MT sequences. Despite this, some plasmids were selected at random for sequencing because previous work isolating microsatellites has shown that suitable sequences can sometimes be obtained regardless of a failure during the PCR screening test (Dr Ann Wood, pers. comm.). Plasmids were extracted using a FastPlasmid™ Mini kit (Eppendorf) following the instructions of the manufacturer. The inserts from the plasmids were PCR amplified using one plasmid primer (i.e. either T3 or T7) and the specific MT oligo, and sequencing proceeded as described in 2.2.2.4.

2.2.6 Sequence alignment using SeqMan™

MT nucleotide sequences were aligned using SeqMan™ (DNAStar Inc. Lasergene 99 package). SeqMan™ distinguished 3 different groups of sequences which have been labelled contig 1, 2 and 3. Consensus sequences from contigs 1 and 3 are reported in Appendix 3, and SNPs from individual clones are highlighted in red. Sequences in contig 2 were highly variable, and no consensus sequence was generated, hence each individual clones in contig 2 has been reported in Appendix 3.

2.2.7 FASTA and BLAST amino acid and nucleotide sequence comparisons

Both BLAST and FASTA search for local sequence similarity, but use different algorithms and statistical approaches. It is beyond the scope of this thesis to discuss these different approaches and statistics in detail, suffice to say that both BLAST and FASTA obtain roughly equivalent search results for query sequences with high homology to sequences already within the databases (Dufresne et al. 2002). However, FASTA may be better for less similar sequences because FASTA makes a full local alignment (Smith-Waterman algorithm).

2.2.7.1 FASTA amino acid sequence comparisons

Exon 2 amino acid query sequences from *P. canaliculus* were compared to protein sequence databases (SwissProt and NCBI) using the FASTA (version 3.4t25 Sept 2, 2005) algorithm (Pearson 1988). The optimised FASTA score, identities and similarities compared with representative examples from molluscs, other invertebrates and vertebrates are reported (Table 2.2).

2.2.7.2 BLAST amino acid, nucleotide and consensus sequence comparisons

Exon 2 amino acid query sequences (39 or 40 amino acids) from *P. canaliculus* were compared to protein sequence databases (all non-redundant GenBank CDS translations + PDB+ SwissProt + PIR +PRF) using BLASTp. Nucleotide searches, using exon 2 query sequences (114 or 117 nts) were compared to nucleotide sequence databases (all non-redundant GenBank + EMBL + DDBJ + PDB sequences) using BLASTn . Finally, consensus query sequences i.e. contig 1 (798nt) and contig 3 (739 nt; Appendix 5) were also compared to nucleotide

sequence databases using BLASTn. There was no consensus sequence for contig 2 so contig 2 sequences were not compared.

2.2.8 PCR without cloning step

After identifying *P. canaliculus MT* sequences using cloning, new primers were designed based on these sequences attempting to isolate MT sequences without a cloning step (Figure 2.1,and Table 2.1). This was achieved using primers 5F and 921, although the resulting sequence was short (~200 bp) and did not include the entire exon 2 sequence. Other primer pairs (Appendix 1) which would have produced larger sequences were also designed using Oligo and trialled. However, additional sequences were not produced using these primers, which could be the result of designing primers for different paralogues.

2.3 Results

2.3.1 PCR results

PCR using the primers and conditions described by (Khoo and Patel 1999) did not produce high quality products. PCR products were 1) not discernable i.e. no products, 2) comprised of multiple non-specific bands, or 3) what appeared to be a single band was actually multiple sequences that were unable to be interpreted using direct sequencing of amplification products. The results suggest that there are multiple copies of the *MT* gene and primers were binding to more than one of these genes creating ambiguous sequence, even when only a single band was cut from the gel for sequencing (Figure 2.2).

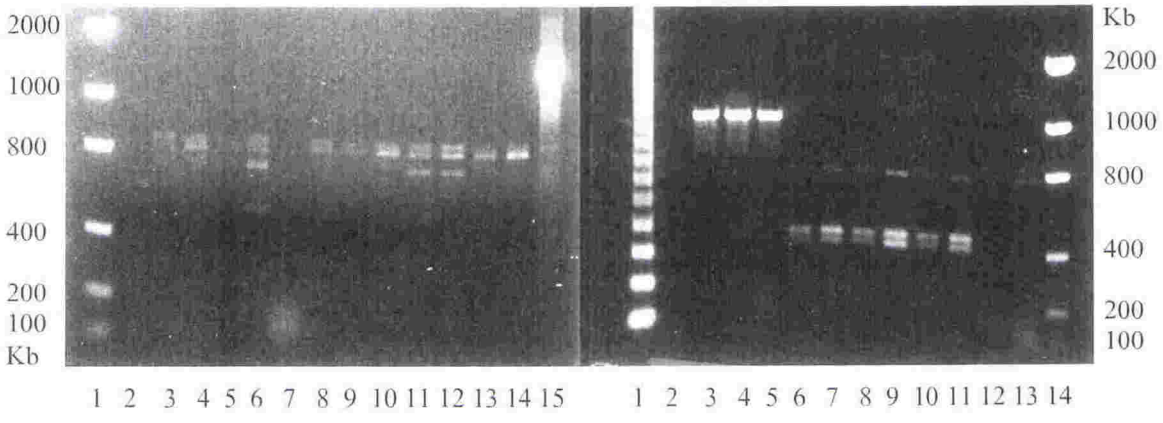


Figure 2.2: Examples of PCR reactions
A: Primers 131/2089; Lane 1 = Low Mass Ladder, Lane 2 = -ve control; Lane 3-14 = *P. canaliculus*; Lane 15 = +ve control (*P. viridis*)
B: Primers G5/2089; Lane 1 =123 bp Ladder, Lane 2 = -ve control; Lane 3-5 = +ve control (*P. viridis*); Lanes 6- 13 = *P. canaliculus*; Lane 14 = Low Mass Ladder

2.3.2 Cloning Results

2.3.2.1 Primer 131/2089

At least 9 different *MT* sequences were obtained (Figure 2.3) using primers 131/2089 and TA® TOPO cloning. The SeqMan™ sequence alignment facility within the DNAStar Inc. Lasergene 99 package separated these *MT* sequences into 3 distinct groups or contigs (Appendix 3).

2.3.2.2 Primer G5/2089

Genomic DNA samples purified from both gill (1Gi) and mantle (5Ma) samples where used as PCR templates in combination with the primer pair G5/2089. Amplified products were ligated into a T-tailed cloning vector and transformed into DH5α cells. Nucleotide sequences were obtained from eight transformants. Predicted protein sequences were derived from these nucleotide sequences and aligned with the predicted protein sequences derived form the PCR products amplified using the 131/2089 primer pair (Figure 2.4). Three MT amino acid sequences generated from G5/2089 clones (3, 4 and 8; Appendix 4) matched 131/2089 clones (5a, 6 and 3, respectively; Figure 2.4).

2.3.2.3 Putative exon 2 MT amino acid sequences

Nine putative exon 2 MT amino acid sequences have been isolated from *P. canaliculus* (Figure 2.3). Gaps have been introduced at the end of the sequences for optimum alignment. The full exon 2 nucleotide sequences and the corresponding amino acid sequences are shown in comparison with *P. viridis MT* sequence (Figure 2.4). Full-length nucleotide sequences including introns are shown in Appendix 3.

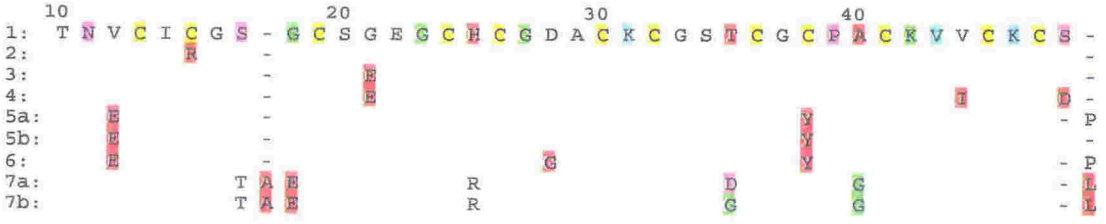


Figure 2.3: *P. canaliculus* MT sequences

Exon 2 amino acid sequence (1) from the mussel *P. canaliculus* is shown in full and only the amino acids in other isoforms which depart from this sequence are shown. Yellow = Conserved cysteine residue; Red = Unique in *P. canaliculus*; Pink = Observed in other invertebrates; Green = Observed in mussels/molluscs; Blue = Observed in vertebrates. Sequence 1-4 = Contig 1; Seq 5a/b &6 = Contig 2; Seq 7a/b = Contig 3; possibly indicating 3 different genes. Full nucleotide sequences are available in Appendix 3. NB: For amino acid residue labelling the last number lines up with the residue it is marking e.g. the 'S' under the '0' in '20' is residue 20. Numbering assumes exon 1 has 9 amino acids (Figure 2.7).

Figure 2.4: Mollusc MT Alignment *P. canaliculus* Exon Two cloned sequences and predicted amino acid sequences using primers 131/2089

Key:

Yellow = Conserved Cysteine residue
Red = Unique in *P. canaliculus*
Green = Observed in other mussels
Pink = Observed in other invertebrates
Blue = Observed in vertebrates

Equivalent sequences

Primers 131/2089	Primers G5/2089
3	8
5a	3
6	4
(See above)	(Appendix 4)

Samples used for cloning

The prefix indicates what sample the clone came from. i.e. clones 1-1 and 1-4 are from sample 1Gi

1 = 1Gi - Gill tissue

2 = 5Ma - Mantle tissue

3 = RIV31 - Gonad tissue

4 = FIO18 - Gonad tissue

MT = *P. viridis* MT sequence

NB: For full sequences (including introns), Appendix 3

3 Major sequence groupings using SeqMan™

Contig 1
Sequence 1 = 1-1, 1-6, 1-8, 1-10, 1-13, 2-3; 3-4, 3-6, 4-8
Sequence 2 = 1-4
Sequence 3 = 2-7, 2-13
Sequence 4 = 2-8

Contig 2
Sequence 5a= 2-12, 2-14; 3-4, 3-5, 3-9, 4-2, 4-3, 4-6
Sequence 5b = 3-7
Sequence 6 = 1-5, 1-7

Contig 3 NB: Y = C or T
Sequence 7a = 1-11, 2-1, 2-4, 2-9, 3-10, 4-7,
Sequence 7b = 4-9

2.3.3 Vectorette Results

Using *Hind*III restriction enzyme and the corresponding *Hind*III vectorette, the universal vectorette system generated 5' and 3' untranslated region (UTR) MT sequences from *P. canaliculus* (Figure 2.5). Partial exon 1 and 3 MT sequences were obtained from the genomic walk of the 5' and 3' UTR sequences respectively (Figure 2.7). Various transcription and enhancer site sequences identified in *P. viridis* were also found in analogous positions within the 5' untranslated region of *P. canaliculus* e.g. promoter sequence, active protein binding site (AP1) and metal responsive element (MRE: Figure 2.5). When subject to BLAST searches parts of these MT UTR sequences matched with 5' and 3' MT sequences from *M. edulis* (results not shown). Primers were designed in the 5' and 3' regions (Appendix 1), with the objective of generating full-length MT sequences. However, PCR amplification was unsuccessful, perhaps indicating that these 5' and 3' sequences were from different MT paralogues.

2.3.4 Biotin Enrichment Results

Despite trialling a range of different biotin labelled MT specific oligos, PCR conditions and hybridisation temperatures, biotin enrichment was unable to isolate MT sequences. These results are not discussed further.

2.3.5 FASTA Results

MT exon 2 amino acid sequences from *P. canaliculus* (Figure 2.3) were compared to other MTs from the NCBI database using FASTA (Table 2.2). The highest similarity scores were obtained with molluscs and other invertebrates. Representative FASTA alignments of *P. canaliculus* MT sequences from Table 2.2 with mussel, human and sea urchin are shown in Figure 2.6. The conserved invertebrate motif is underlined (Dallinger et al. 1993), the central conserved motif (Nemer et al. 1985) is highlighted red in both mussel and sea urchin sequences, and blue in humans. The polarity of the central conserved motif appears to be inverted in mussel and sea urchin sequences compared to human (Nemer et al. 1985).

Figure 2.5: *P. canaliculus* MT 5' and 3' untranslated regions (UTR)

5'UTR **Samples: 1** = 5Ma and RIV *Hind*III 5' Sequence. RIV SNP G to A at position 473. **2** = FJO18, evidence of Heteroplasmy (N)

```
 -430       -420       -410       -400       -390       -380       -370       -360       -350       -340       -330       -320       -310
1 CAACGTGGATCCGAATTCAAGCTTCATGGGGTTCTCCTGCAGCTTATTTGTAGGTCATTAAACCATCCCTTATTCTATTGATATGTCACATGTCGAAGCCGGTAATATGTTTG AATGACGCATC
2                                          GCTTATTTGTAGGTCATTAAACCATCCCTTATTCTATTGATATGTCACATGTCGAAGCCGGTAATATGTTTG AATGACGCATC

      -300       -290       -280       -270       -260       -250       -240       -230       -220       -210       -200       -190
1 AACATGAGGTGCCCTTATGG ATAAAAGAAATCGCGGGATATGTCATATAGATTGTCTATTTTTTCAAGCCTTTGAAAATAAACTTGATTTTTAAAGGAAACCAATGTAAAATACAAGGATAACT
2 AACATGAGGTGCCCTTATGG ATAAAAGAAATCGCGGGATATGTCATAT GATTGTCT TTTTTTCAAGCCTTTGAAAATAA CTTGATTTTTAAAGGAAA CAATGTAAAATACAAGGATAACT

 -180       -170       -160       -150       -140       -130       -120       -110       -100        -90        -80        -70        -60
1 AGAAAAAATGAGGGGACTCATTTGTAAATGAAATAAAGATTTTGAATAATTTATGACGCCTGACATAGAACATGGGGTCCCGACGTACCGCGTGCAGAGATTATGTTTTGCACACGGCATACTTA
2 AGAAAAAATGAGGGGACTCATTTGTAAATGAAATAAAGATTTTG ATAATT ATGACGCCTGACATAGAACATG GGTCCC ACGTACCG GTGCAGA ATTATGTTTTGCACACGGCATACTTA

      -50        -40        -30        -20        -10         ***        10         20         30         40         50         60
1 AAATCGATATATAAACCGGCATAACATTAACGGCAGCATATACGCAGTGAATTGAATAAAAGAGAAACATATATACCGATTTGAAATATGCCTACACCTTGCAACTGTATCCAAAGTAAGTGAAA
2 AAATCGATATATAAACCGGCATAACATT ACGGCAGCATATACGC GT AATTGAATAAAAGAGAAACATATATACCAATT GA ATATGCCTACAC

         530
1 CCGCTAACATAA  Key: AP1; AP1; MRE; Promoter; start of exon 1; end of exon 1; GT splicing site underlined
```

3'UTR **Samples: 1** = 5Ma, RIV31, FJO18, All *Hind*III Vectorettes, either 1-5U* or 4-6U primers * 1-5U gives slightly longer sequences

```
          10         20         30         40         50         60         70         80         90        100        110        120
1 AAAATTTCTCAGATTTGATCATTGCTTTATTTATTTTCAGGAGAGTGTGAATGTG ACA-GAGTTGTACCGGTCCAAA ACGTGTAATTGTGATTCCAATTGCTCCTGCAAATAACATAGAACTA

    130        140        150        160        170        180
1 TGCAATATATTCTAAGCTGCAGGAGAACCCCATGAAGCTTGAATTCGGATCCACGTTGA Key: start of exon 3; end of exon 3; AG splicing site underlined
```

56 del G FJO18 *Hind*III 60 ins G FJO *Hind*III 79 T to C 5Ma *Hind*III

NB: For nucleotide labelling the last number lines up with the residue it is marking e.g. the 'C' under the '0' in '10' is residue 10. Numbering assumes exon 1 has 9 amino acids

Figure 2.6: Representative FASTA alignment results

```
P. viridis: gi|18203491|sp|Q9U568|MT_PERVI Metallothionein (MT)                    (73 aa)
 initn: 275 init1: 275 opt: 275  Z-score: 324.7  bits: 63.1 E(): 5.4e-11
Smith-Waterman score: 275;  81.579% identity (94.737% similar) in 38 aa overlap (1-38:10-47)
>gi|182    1-  38:------------------------------------------------------------ :

                      10        20        30
QUERY           TNVCICGSGCSGEGCHCGDA          ACKVVCKCS - Exon 2
                :.:::::.:::::::.::::::.: ::: .:::::::
gi|182 MPSPCNCIETQVCICGTGCSGEGCRCGDACKCSSGCGCSGCKVVCKCQPGECACGKQCTGPDTCKCDSSCSCK
              10        20        30        40        50        60        70

-------------------------------------------------------------------------------
Human MT-4: gi|1346594|sp|P47944|MT4_HUMAN Metallothionein-4 (MT-4) (Metallothionein-I   (62 aa)
 initn: 151 init1:  78 opt: 139  Z-score: 178.0  bits: 35.8 E(): 0.0079
Smith-Waterman score: 139;  50.000% identity (61.111% similar) in 36 aa overlap (4-38:6-37)
>gi|134    4-  38:     ------------------------------------------------------- :

               10        20        30
QUERY     TNVCICGSGCSGEGCHCGDA          AC-KVVCKCS
          :.: ::     : ::: ::: .::.: .: :  : :
gi|134 MDPRECVCMSGGI---CMCGDNCKC-TTCNCKTCRKSCCPCCPPGCAKCARGCICKGGSDKCSCCP
              10          20        30        40        50        60

-------------------------------------------------------------------------------
Sea urchin: gi|2497877|sp|Q26497|MTA_SPHGR Metallothionein-A (MTA)                 (67 aa)
 initn: 222 init1: 120 opt: 137  Z-score: 175.5  bits: 35.4 E(): 0.011
Smith-Waterman score: 137;  47.059% identity (76.471% similar) in 34 aa overlap (3-36:36-67)
>gi|249    3-  36:    ----------------------------------------------------------:

                                          10        20        30
QUERY                                TNVCICGSGCSGEGCHCGDA          ACKVVCKCS
                                     .: ::. ::. .:.:.:.::::: :.:    . .:
gi|249 PGPDVKCVCCQDGKECPCKGGECCITGSCCKEGDGLC-CGK-CSNAACKCADGCKCGSGCSCTLGNCTC
              10        20        30        40        50        60
```

KEY:
invertebrate motif = underlined
central conserved motif = mussel and sea urchin (in red), human (reversed, in blue)

Table 2.2: Summary FASTA Results

FASTA analysis of *P. canaliculus* MT sequences with other MT sequences. Sequences are ranked by highest FASTA optimised score.

Seq#	Group	Organism	Isoform	Optimised FASTA Score	Identity	Similarity	# Residues in overlap	References
MOLLUSCS:								
ALL	Mussel	*P. viridis*		242-283	73.7-87.2%	86.8-94.9%	37-39	(Khoo and Patel 1999)
1-6,7b		*M. edulis*	MT20IA&B	233-274	71.1-84.6%	84.2-92.3%	38, 39	(Barsyte et al. 1999)
7a			MTIII	269	82.5%	90%	40	(Mackay et al. 1993)
1-6,7b		*M .edulis*	MT20II	233-274	71.1-84.6%	84.2-92.3%	38, 39	
7a			MT20IA&B	265	82.1%	89.7%	39	
ALL	Oyster	*C. virginica*	MT	175-199	46.5-53.8%	73.7-79.5%	36-39	(Unger et al. 1991) (Roesijadi 1989) (Roesijadi et al. 1991)
1-5b	Roman snail	*H. pomatia*	Cd-MT	144-162	47.1-51.6%	67.7-75%	31-34	(Dallinger et al. 1993)
7a,7b	Land snail	*A. arbustorum*	MT	155, 164	48.6-51.4%	77.1, 80%	35	(Berger 1995)
1-5b	Land snail	*A. arbustorum*	MT	141-159	39.5-43.8%	76.5-80%	32-38	
6-7b	Roman snail	*H. pomatia*	Cd-MT	144-155	47.1-50%	69.7-73.5%	33, 34	
OTHER INVERTEBRATES:								
1-4,7ab	Sea urchin	*S. granularis*	MTA&B	132-147	42.8-48.6%	68.4-77.1%	34,35,38	(Ciaramella et al. 1997)
5a,b,6	Silk moth	*Bombyx mori*	Chorion HCA	133-142	50.0-52.8%	61.1. 63.9%	36	(Iatrou et al. 1984)
1,3,4	Sea urchin	*S. neumayeri*	MT	133-134	45.5%	72.7,75.8%	33	(Ciaramella et al. 1997)
2	Blue crab	*C. sapidus*	Cu-MTII	131	50.0%	73.5%	34	(Brouwer et al. 1995)
5ab,6	Sea urchin	*S. granularis*	MTA or B	125-116	47.2-44.4%	69.4, 66.7%	36	
7a,b	Silk moth	*Bombyx mori*	Chorion HCA	140-149	48.6-51.4%	59.5,62.2%	37	

1	Silk moth	*Bombyx mori*	Chorion HCA	130-136	44.7-50%	57.9,61.1%	38,36	
2,5ab 7ab	Sea urchin	*S. neumayeri*	MT	118-144	37.5-47.1%	62.8-76.5%	40,34	
3	Blue crab	*C. sapidus*	Cu-MTII	131	48.6%	67.6%	37	
4,6	Fruit fly	*D. ananassae*	MT-1	129-111	47.2-48.6%	63.9, 62.9%	36,35	(Stephan et al. 1994)
VERTEBRATE:								
1,3-5b	Human	*H. sapiens*	MT-4	129-139	48.6-50%	59.5-63.6%	33,36,37	(Quaife et al. 1994)
2,7ab	Eelpout	*T. cereberus*	MT-A	130-139	51.6-53.1%	68.8,71.0%	31,32	(Kille and Olsson 1996)
6	Arctic char	*S. alpinus*	MTB	128	54.8%	71.0%	31	(Gerpe et al. 1998)
1,5ab	Dog	*C. familiaris*	MT-IV	121-131	48.6-50%	56.8-59.5%	36, 37	(Kobayashi et al. 1999)
2,3,7a	Pike	*E. lucius*	MT	130-133	53.1-54.8%	68.8-74.2%	31,32	(Kille et al. 1991)
4	Eelpout	*T. cereberus*	MT-A	130	53.6%	78.6 %	28	
6	Eelpout	*Z. viviparous*	MT	124	51.6%	71.0%	31	
7b	Human	*H. sapiens*	MT-3	128	50.0%	64.7%	34	
1,5ab, &7b	Mouse	*M. musculus*	MT-IV	121-131	45.9-50%	56.8-59.5%	36,37	(Quaife et al. 1994)
2	Human	*H. sapiens*	MT-4	128	54.8%	64.5%	31	
3	Eelpout	*T. cereberus*	MT-A	133	51.6%	74.2%	31	
4,7a	Arctic char	*S. alpinus*	MTB	126,131	53.1-53.6%	65.6,75.0%	28,32	
6	Pike	*E. lucius*	MT	123	51.5%	66.7%	33	

NB: All hits within a section (e.g. Molluscs) are arranged in decreasing order of significance. FASTA version 3.4t25 Sept 2, 2005.
Molluscs: Top 3 mussel hits are shown, then ignoring mussel hits, the next top 3 mollusc hits are shown.
Other sections: Top 3 hits, from 3 different species are shown (i.e. if 2 of the top 3 hits are from the same species, the second hit is not shown).
References: Appendix 6
Identity is the degree of correspondence between two sub-sequences (no gaps between the sequences). An identity of 25% or higher implies similarity of function, while 18-25% implies similarity of structure or function.
Similarity is the degree of resemblance between two sequences when they are compared. This is dependant on their identity.

2.3.6 BLAST amino acid, nucleotide and consensus sequence comparisons

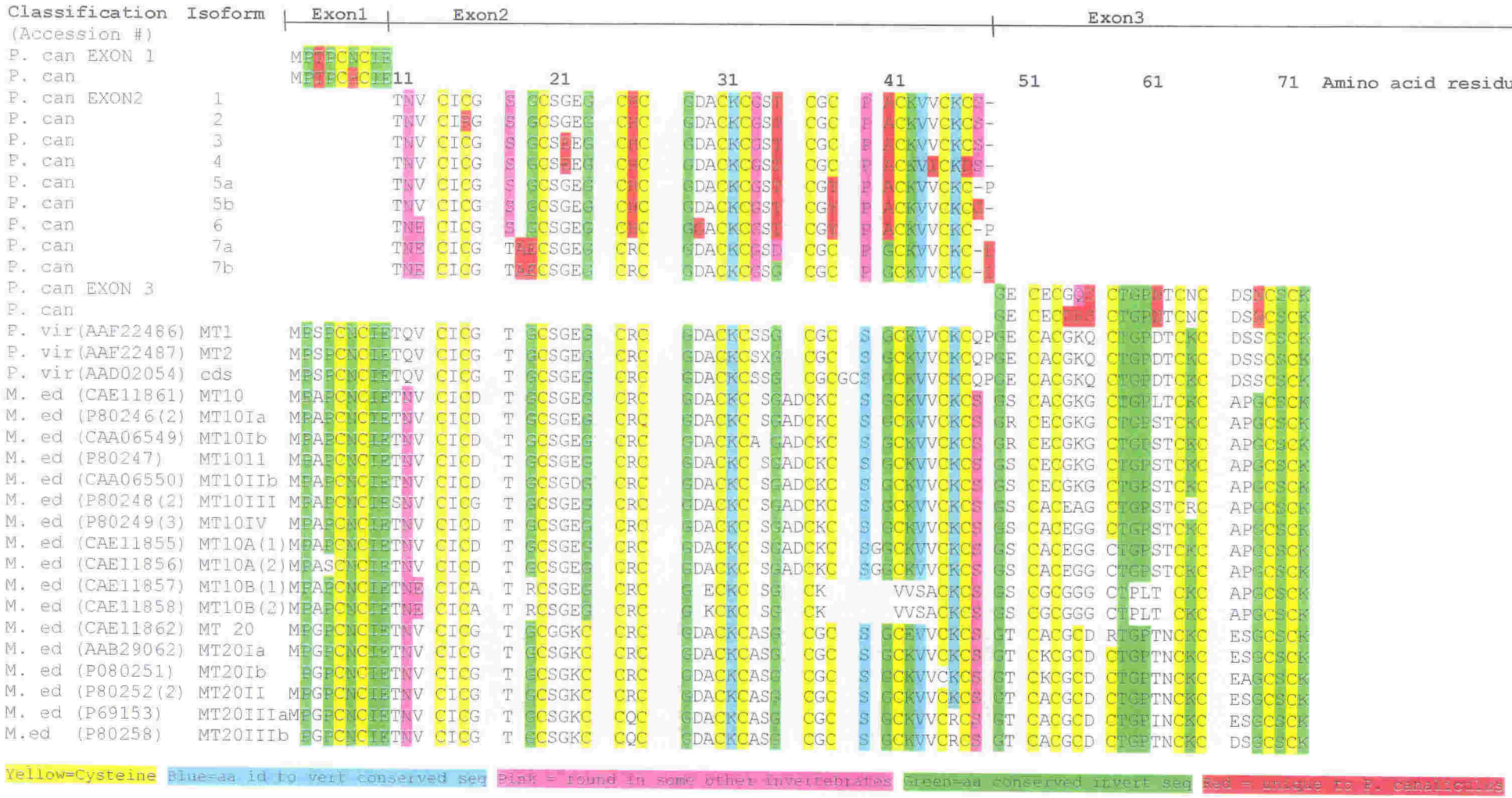
Exon 2 MT nucleotide, amino acid and consensus sequences from *P. canaliculus* were compared to those of other species using BLAST (Appendix 5). In general, amino acid comparisons with protein sequences gave the most significant results and all of the hits resulted from similarities with MTs from other species. The expect (E) values ranged from a low of 5e-06 (top match), which represent potentially interesting relationships, to a high of 8.9 (bottom match), which may be false matches. The E-value describes the likelihood that a sequence with a similar score will occur in the database by chance, so the smaller the e-value the more significant the alignment. The most significant matches were with *Bathymodiolus azoricus* (Hydrothermal vent mussel) MT and *M. edulis* (blue mussel) MT (multiple isoforms) both of which had the same expect values (i.e. they were equally likely to be the most significant match).

Comparison of *P. canaliculus MT* nucleotide sequences with those of other species using BLASTn gave poor results. In most cases few of the hits were for *MT*, and all of the E-values were relatively high in comparison to those obtained for protein amino acid comparisons (Appendix 5). Using consensus sequences 1 and 3, most of the hits were not for *MT*. Instead random matches were obtained with cloned sequences from mice and other model species which are present in high numbers within the databases. Consensus sequence 3 gave a similar number of matches as consensus sequence 1 (~100), but a greater number of MT matches were obtained with consensus sequence 3 (~20 compared to ~4; Appendix 5). There was no consensus sequence 2, as sequences within this contig were too diverse to generate a consensus sequence.

2.3.7 Mollusc MT alignment

Partial MT amino acid sequences isolated from *P. canaliculus* during this study were aligned for comparison with other mollusc MT sequences (Figure 2.7). MT sequences from *P. canaliculus* (n = 4) showed a large amount of variation (at least 9 MT sequences), although still behind the number isolated from *M. edulis* (17), the most extensively studied mollusc species to date.

Figure 2.7: Mollusc MT Alignment



Key:
P. can = *P. canaliculus;* New Zealand green mussel
P. vir = *P. viridis;* Tropical green mussel
M. ed = *M. edulis;* Blue mussel

NB: For amino acid residue labelling the last number lines up with the residue it is marking e.g. the 'N' under the '1' in '11' is residue 11. Numbering is based on *P. canaliculus* exon 2 amino acid isoform 1.

Key contd:

M. gallo = *M. galloprovincialis;* Blue mussel
D.poly = *D. polymorpha;* Freshwater zebra mussel
C.angu = *C. angulata;* Portuguse oyster
C.gigas = *C. gigas;* Pacific oyster
C.rhiz = *C. rhizoporae;* Mangrove oyster

C.vir = *C. virginica;* Eastern oyster
O.ed = *Ostrea edulis;* European flat oyster
A.gran = *Andara granosa;* Cockle
V.pull = *Venerupis pullastra;* Yellow carpetshell

R.duc = *Ruditapes decussates;* Grooved carpetshell
R.phil = *Ruditapes phillippinarum;* short necked clam
H.pom = *H. pomatia;* Roman snail (edible)

References: Appendix 7

2.2.8 PCR sequencing results without cloning

After identifying *P. canaliculus MT* sequences using cloning (above), new primers were designed to permit isolation of MT sequences without a cloning step (Appendix 1). This was achieved using primers 5F and 921, although the resulting sequence was short (only ~200 bp) and did not include the entire exon 2 region. Several of these sequences showed evidence of heterozygosity, suggesting that there are different MT alleles, and that mussels can be heterozygous or homozygous for those alleles (Figure 2.3). Most of the allelic variants predicted in this electropherogram have also been identified in the cloned sequences (Figure 2.7).

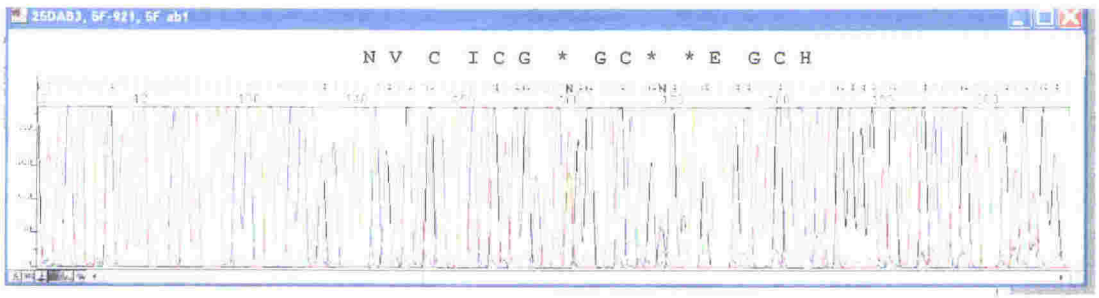


Figure 2.8: Electropherogram showing an example of a suspected heterozygous individual (DAB3).
Sequence was generated using primers 5F and 921. Protein residues stipulated by the nucleotides are shown above and heterozygosity is marked *. TCN could be TCA or TGT, in both cases these residues code for serine (S). AGN could be AGC or ACT, again both code for serine. In the next codon, GAA has been called, although GGA is also possible based on the electropherogram. GAA codes for glutamic acid (E), whereas GGA codes for glycine (G). All of these variants appear in Figure 2.7.

2.4 Discussion

2.4.1 *P. canaliculus* MT sequences

The present study reports MT partial gene nucleotide sequences and deduced amino acid sequences from the New Zealand greenshell mussel *P. canaliculus*. At least nine different MT exon 2 sequences and their corresponding intron sequences have been isolated from *P. canaliculus*. Each of these MT amino acid sequences contains 9 Cys-X-Cys motifs, as well as the conserved central segment, Cys-X-Cys-X-X-X-Cys-X-Cys-X, which was proposed by Nemer et al. (1985) as the common feature among all MTs. The characteristic high number of cysteine residues (~30%) and their alignment with MT sequences from other molluscs (Figure 2.7) increase the likelihood that the amino acid sequences deduced from the DNA sequences isolated

in the current study are indeed MTs and that they belong to the mollusc MT type 2 family. Consensus sequence alignment using SeqMan™ (DNAStar Inc. Lasergene 99 package) has split these 9 MT sequences into 3 different contigs (Contig 1 = Seqs 1-4; Contig 2 = Seqs 5a, 5b, 6 and Contig 3 = 7a and 7b, Appendix 3), which suggests that there are at least 3 MT genes, each with several different alleles.

2.4.2 Polymorphism of MT isoforms

Polymorphism of mussel MT isoforms between individuals is plausible. Previous studies used a large number of pooled mussel (*M. edulis*) samples to generate MT isoform sequences (Mackay et al. 1993). It was suggested that the large number of isoforms (5 dimeric and 4 monomeric) isolated in this previous study was possibly due to polymorphism between individuals. In the current study, single mussel samples were used to generate MT sequences, with the result that differences in the number and types of isoforms present between individuals can be recognised. It appears that some isoforms in the present study may be individual specific (Appendix 3). For example, sequence 1 was isolated from all 4 individuals perhaps suggesting that it is a ubiquitous MT variant which is required by all individuals. In contrast sequence 2 was only found in the gill tissue of 1 individual and sequence 3 was only found in the mantle of another. Whilst these results are suggestive of different isoforms of MT it is recognised that other explanations are also possible.

Isoforms isolated from different individuals may result as artefacts of the cloning process, due to preferential binding of the primers to some isoforms rather than others, or the relatively low fidelity of *Taq* polymerase may be generating artifactual variation. As some of these exon 2 MT sequences are based on the isolation of only one cloned sequence caution is advised. However, despite these uncertainties, it is likely that some of these sequence differences represent MT paralogues. Some of of these isoforms may be tissue-specific as has been demonstrated in humans e.g. brain specific MT-3 (Masters et al. 1994) and squamous epithelium specific MT-4 (Quaife et al. 1994) and snails (*H. pomatia*) e.g. mid-gut Cd-isoform and mantle Cu-isoform (Dallinger et al. 1997). However, tissue specific isoforms can not be proven in the current study because studies are based on genomic DNA and not cDNA.

Results to date indicate extensive *MT* gene polymorphism between species, as evidenced by the large number of different MT sequences (Figure 2.7). However to

date allelic polymorphism between individuals has not been demonstrated clearly in either mammals or mussels. Increasing evidence suggests that allozyme polymorphism in organisms (animals and plants) is linked to adaptation to specific environments (Yap et al. 2004c). For example, in *P. viridis*, sites with high metal pollution indices (MPI) favour certain alleles at the phosphoglucomutase (PGM; E.C. 2.7.5.1.) locus suggesting an adaptive response to heavy metals (Yap et al. 2004c). Similar results in clams (*Ruditapes decussatus* and *R. phillippinarum*) correlated specific PGM alleles with MT expression and the level of metal pollution (Moraga et al. 2002). It is possible that some individuals carry MT isoform and allelic variants that have superior capabilities for managing heavy metals. In the oyster *C. gigas* certain polymorphisms in two exons of the MT gene *CgMT1* have been associated with tolerance to metals using PCR-SSCP (Tanguy et al. 2002). There is also some evidence of homozygosity and heterozygosity at different MT loci in *C. gigas* MT genes using RFPL analysis (Tanguy et al. 2001). Genetic differences in the metal binding capabilities of alleles (and by inference the individuals that carry them) will lead to exciting new applications in the field of environmental protection. If allelic polymorphisms are identified and characterised in marine organisms, variants that provide greater protection against heavy metal toxicity or low tolerance variants that provide an 'early warning' system (i.e. sentinels) for contamination could be utilised for environmental monitoring.

In *P. canaliculus*, certain individuals showed evidence of having heterozygous allelic polymorphisms in their MT coding region. Primers 5F and 921 generated short *MT* nucleotide sequences without a cloning step. Heterozygous individuals were characterised by electropherograms having two different nucleotide peaks of similar height at the same nucleotide position (Figure 2.8). This is the first evidence of putative *MT* allelic diversity demonstrated in mussels using a sequencing approach. At this stage it is unknown whether a heterozygous MT state is advantageous to the mussel and how heterozygosity correlates with *MT* sequences results reported in Figure 2.3, further work is needed to confirm and extend these results, although there is a clear theoretical basis for polymorphism and heterozygote advantage in other species.

2.4.3 Unusual amino acid residues

2.4.3.1 Tyrosine

MT sequences from *P. canaliculus* possess a number of unusual amino acid residues when compared to MT sequences from other species. Three of nine amino acid sequences isolated from *P. canaliculus* contain the aromatic amino acid tyrosine (Y). To date, very few MT sequences have contained any of the aromatic amino acids phenylalanine (F), tryptophan (W) and tyrosine (Y), indeed low aromatic residue content is an identifying feature of MTs. However, there is one conserved tyrosine residue in nematodes (Imagawa et al. 1990; Slice et al. 1990; Freedman et al. 1993) and a tyrosine residue in exon 3 has also been discovered in the MTA gene of the European flat oyster *Ostrea edulis* (Tanguy et al. 2003; see also Table 2.7). Although phenylalanine has been observed in sea urchins (Riek et al. 1999), to the best of my knowledge, the current study is the first reported incidence of an aromatic amino acid in any mussel MT (Figure 2.7). At this stage it is unknown what the addition of a tyrosine residue at the expense of a cysteine residue will mean for the binding capabilities of the *P. canaliculus* MT isoforms. However, both cysteine and tyrosine are non-polar/hydrophobic residues with hydrogen binding capabilities, so metal binding capabilities may be conserved depending on the function of the cysteine residue that has been replaced (metal binding or bridging?). Although, MT isoforms containing tyrosine have only 20 cysteine residues, compared to the usual 21 or 23 in *Mytilus* (Figure 2.5) this would not necessarily mean a reduction in metal binding capabilities (discussed later).

2.4.3.2 Histidine

Histidine (H) is a polar/hydrophilic amino acid that shares many characteristics with aromatic amino acids and is rarely found in MT sequences, having an estimated prevalence of 1% in molluscan MT (Bebianno and Langston 1998). In another first for mussel MT, histidine was observed in 7 of 9 MT exon 2 sequences isolated from *P. canaliculus*. The only other incidence of histidine in molluscan MT was reported in the roman snail (*Helix pomatia*) CuMT amino acid sequence (Figure 2.7). In both mussel and snail MT the histidine occurs between a pair of cysteine residues although the significance of this feature is unknown. An

unusual distribution of histidine and cysteine residues is found in fungal MT (Dujon et al. 2004), but because the structure of this protein is very different to mollusc MT the histidine in common is probably coincidental and not indicative of any evolutionary relationships. Interestingly, a single exon 1 MT amino acid sequence of *P. canaliculus* (Figure 2.7) also has a histidine residue but again the significance of this substitution is unknown.

Mutagenesis studies of mammalian MT and its separated domains have shown that substitution of a metal-ligating cysteine by a non-coordinating residue has a detrimental effect on metal binding properties (Cismowskil and Huang 1991; Cismowskil et al. 1991), whereas substitution by a coordinating histidine preserves binding (Romero-Isart et al. 1999). However in both instances, in exon 1 and 2, of the MT isolated from *P. canaliculus* (Figure 2.7) the histidine residue occurs between two cysteine residues not replacing them. It is possible that histidine could be performing a structural or functional role e.g. by making contacts with the β-domain, or with other proteins (Riek et al. 1999). This proposition is not without precedence, for example in the α-domain of the sea urchin MTA phenylalanine is located in a fully solvent exposed location on the protein surface, despite evidence to show that this aromatic ring could be accommodated in a hydrophobic position, perhaps suggesting that phenylalanine could be there to perform a specific function (Riek et al. 1999). Given the proximity of histidine to flanking cysteine residues and the importance of cysteine in metal binding, an additional function for histidine, however intriguing, may be impossible given metal binding constraints. It would be interesting to model the 3D structures of *P. canaliculus* MTs and to predict their metal-binding capabilities, but it is beyond the scope of the present investigation.

2.4.3.3 Proline

The presence of a conserved proline residue at nucleotide position ~39 in all of the *P. canaliculus* MT sequences isolated in the current project is unexpected because proline can have a pronounced effect on 3-dimensional conformation. The side chain of proline contains a portion of the main chain and thus tends to influence the 3-dimentional structure. Proline occurs in a similar position in the MT amino acid sequences of Pacific and Flat oysters (Figure 2.7) suggesting that this substitution, although unusual, compared with the majority of species is still functional. In mammals, the occurrence of two proline residues in exposed peptide

loops of the β-domain in the mammalian MT-3 subclass specifically determines their inhibitory effect on the survival of cortical neurons (Sewell et al. 1995). Although interesting to speculate, there is no evidence at this stage to suggest that similar interaction or inhibition functions are occurring via proline in mussels.

2.4.3.4 Glycine

The large number of glycine residues typically identified in mussels and other bivalves is unusual for MTs in general and may represent a mollusc specific MT variant (Dohi 1986). In *M. edulis* there are 11 glycine residues (Figure 2.7) but this number drops to 9 or 10 in *P. canaliculus*. Glycine is replaced by threonine at residue position 34, although this may be partially compensated for by a glycine residue replacing serine at position 32. However, another glycine replacement by alanine at position 40 is not compensated by another glycine replacement elsewhere in the sequence. Although each *P. canaliculus* MT isoform varies in the number and position of glycine residues, overall there are still more glycine residues than are found in vertebrates (e.g. 5 or 6 in mammals), but less than in some other molluscs.

2.4.3.5 Methionine

The absence of methionine at the amino-terminal end of molluscan MT is a very common feature (Viarengo 2000). Although methionine (M) has been predicted to occur in *P. canaliculus* MT amino acid sequence based on isolated exon 2 nucleotide sequences, previous studies indicate that molluscs and other non-mammalian MTs are subject to amino-terminal modifications following initial synthesis. The first eight residues predicted from the nucleotide sequence in yeast MT for example, are not present in the final form of the protein and are not required for structural integrity or metal-binding properties (Wright et al. 1987). The amino-terminal region of MT protein appears labile when considered from a comparative perspective because the NH_2 –terminal of various non-vertebrates exhibit significant deviations from the mammalian forms (Roesijadi 1989). Methionine residues are reported in *P. canaliculus* sequences shown in Figure 2.7 deduced from the nucleotide sequences obtained from the current study. However, the presence of methionine would need to be confirmed by isolating and characterising full-length MT proteins from *P. canaliculus*. MT protein sequences from other invertebrates reported in Figure 2.7 have been obtained from GenBank, and are recorded in Figure

2.7 as they were reported in the database (i.e. some previous studies reported the presence of methionine in the MT sequence while others did not).

2.4.4 Isoform classes

Previous studies on mussel MT proteins from the blue mussel *M. edulis* have identified 10 kDa isoforms containing 21 conserved cysteine residues and 20 kDa isoforms containing 23 conserved cysteine residues (MacKay et al. 1993). In the current study *P. canaliculus* isoforms contain either 20 or 21 cysteine residues dependent on isoform. It is unknown whether isoforms from *P. canaliculus* form two separate weight classes. However cysteine content alone suggests that *P. canaliculus* does not possess sufficient numbers of cysteine residues to form 20 kDa isoforms. The additional cysteine residues present in the 20 kDa forms of *M. edulis* are postulated to be involved in the intermolecular linkages between the monomeric units of the MT20 isoforms either through a disulphide bridge or through S-Cd-S bridging. The number and size of MT isoforms in *P. canaliculus* could be determined using gel-permeation or anion exchange chromatography but such work is beyond the scope of this study.

P. canaliculus MT amino acid sequences 1, 3, 7a and 7b (Figure 2.7) have 21 conserved cysteine residues (including exon 1 and 3) suggesting that these MT sequences belong to the mollusc mo1 MT subfamily (Binz and Kagi 1999). *P. canaliculus* MT amino acid sequences 2, 4, 5a, 5b and 6 contain only 20 cysteine residues (including exon 1 and 3) in contrast to MT sequences from other mussels (genus *Mytilus*) which usually contain either 21 or 23 conserved cysteine residues (Figure 2.7). Typical features of the mo subfamily include at least 18 cysteine residues, high cysteine conservation with mo1 (mussel MT-1, 21 cysteines), mo2 (mussel MT-2, 23 cysteines) and mog (gastropod MT, 18 cysteines) MT sequences. These cysteine deficient *P. canaliculus* MT amino acid sequences with 20 cysteine residues should be included in the mollusc subfamily mo.

Mussel MT isoforms containing 21 or 23 cysteine residues bind 7 metal ions. *P. canaliculus* MT isoforms containing 20 cysteine residues may bind less metal ions as a decrease in the number of cysteine residues present within a MT isoform can be associated with a decrease in the number of metal ions that can be bound. For example, crustaceans have 18 cysteine residues that bind 6 metal ions (Roesijadi 1989). However, vertebrate and echinoderm MT sequences are characterised by 20

cysteine residues and still bind seven bivalent metal ion equivalents (Binz and Kagi 1999; Isani 2000).

2.4.5 Exon 3

The exon 3 sequences (Figure 2.7) correspond with the mollusc specific MT motif sequence C-X-C-X(3)-C-T-G-X(3)-C-X-C-X(3)-C-X-C-K (Tanguy et al. 2001), providing supporting evidence that the sequences isolated from *P. canaliculus* are genuine MT exon 3 sequences. When compared to each other the two exon 3 sequences isolated in the current study had 2 amino acid differences. This low level of carboxy-terminal MT sequence variability is among molluscs is comparable to what has been observed in other species of molluscs.

2.4.6 Intron/exon boundaries

MT intron and exon boundaries in *P. canaliculus* appear to follow the GT/AG splicing rule (Breathnach and Chambon 1981) as was observed by Khoo and Patel (1999) for *P. viridis*. The donor splice site (GT) is after the 5' end of the intron sequences and the acceptor site (AG) is before the 3' end of intron sequences (Figure 2.1). *P. canaliculus* does not appear to have the same splicing signals that were identified in *M. edulis* (Leignel and Laulier 2006). Refer to Figure 2.5 and Appendix 3 for examples related to *P. canaliculus* sequences.

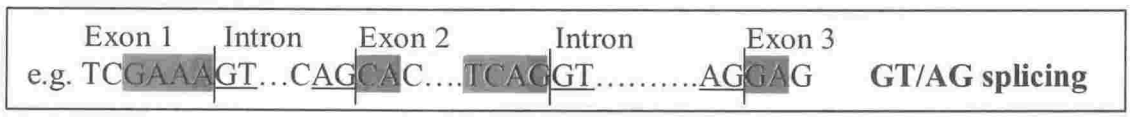


Figure 2.9: Intron and exon boundaries

The GT/AG splice sites are underlined. The first amino acid of exon 2 is ACA = Threonine (T), the first nucleotide of this amino acid occurs at the end of exon 1 (highlighted green). Exon 2 ends with TCA = Serine (S). The first amino acid of exon 3 is GGA. The first G is from the end of the exon 2 (highlighted).

2.4.7 FASTA Results

P. canaliculus MT sequences show strong homology (as demonstrated by high FASTA scores) with amino acid sequences from other molluscs, especially the blue mussel *M. edulis* and the tropical green mussel *P. viridis*. These results suggest,

unsurprisingly, that mollusc MT sequences are closely related. Excluding mollusc sequences, invertebrate sequences in general have the next highest FASTA rankings. This is in contrast to previous studies of MTs in the blue mussel (Mackay et al. 1993) and the roman snail (Dallinger et al. 1993) that reported highest FASTA similarity scores with vertebrates. However previous studies used full length MT sequences whereas the current study only uses exon 2 sequences. It may be possible to filter the full NCBI dataset to obtain only MT exon 2 sequences to run a 'fair' comparison against other species.

In mammalian MT the β-domain binds 3 metal ions and is N-terminally encoded by exons 1 and 2, and the α-domain binds 4 metal ions, is characterised by Cys-Cys pairs and coded by C-terminal exon 3 sequence. The sequential order of the α- and β- domains appears to be inverted in sea urchins compared to mammals (Nemer et al. 1985; Riek et al. 1999), with the α- domain being N-terminally encoded and containing Cys-Cys pairs. This 'domain switch' appears to be the result of an inversion of exons 2 and 3 in sea urchin and includes a reversal of the polarity of the central segment (Nemer et al. 1985; Figure 2.2).

In mussels and humans, exon 2 alignment occurs at the start of the amino acid sequence and in sea urchin the alignment occurs at the end of the sequence. These differences in alignment patterns may explain why when comparing full-length MT sequences, sea urchins do not score as highly as vertebrates, as only part of the sequence (i.e. exon 2) shows significant alignment between mussels and sea urchins.

Unlike mammal and sea urchin MT, mussel MT does not include any Cys-Cys pairs to guide assignment of α- and β- domains. However, the position of the 'linker' sequence (KVV), which often separates α and β domains, suggests that like sea urchins the α-domain of mussel MT is N-terminally encoded and contains 12 cysteine residues (1 more than sea urchin). Despite sharing α- and β- domain order and similar central domain sequences, previous papers suggest that mussel MTs do not share significant homology with sea urchin MTs. The current study indicates some similarity between *P. canaliculus* exon 2 sequences and the carboxy-terminal sequences of sea urchin MTs. This similarity may be less obvious in other studies when comparisons are made using the entire MT coding region instead of just the exon 2 region as used here. Full-length MT sequences from *P. canaliculus* should be obtained and analysed before conclusions on the significance of these alignments can be drawn. In particular metal binding experiments would be useful to determine the

binding capabilities of each MT mussel domain in comparison to mammals and sea urchin to ensure that α-and β- domains have been correctly assigned. The observed contradictions in MT binding domain structures and sequence homology could shed light on the evolution of MT in these organisms.

2.4.8 BLAST results

BLAST comparisons for MT amino acid sequences from *P. canaliculus* resulted in hits for MT in other species, with the top hits coming from mussels (e.g. the common blue mussel *M. edulis* and the hydrothermal vent mussel *B. azoricus*). However, nucleotide comparisons for MT resulted in numerous false hits for proteins other than MT. Because FASTA is the best comparison tool to use BLAST results will not be discussed further.

2.5 Conclusions

Nine partial MT amino acid sequences have been isolated from the endemic New Zealand greenshell mussel *P. canaliculus* for the first time. Uncharacteristically for MT, some of these isoforms have an aromatic amino acid (tyrosine) replacing a cysteine residue. At this stage the significance of this substitution on metal-binding is unknown. MT genes from *P. canaliculus* appear to follow the AG/GT splicing rule as was also identified in *P. viridis* (Khoo and Patel 1999). In the 5' untranslated region AP1, metal responsive element (MRE) and promoter sequences have been found in analogous positions to those found in *P. viridis* providing some information on MT transcriptional control. When MT exon 2 amino acid sequences from *P. canaliculus* were compared to other species using FASTA, MTs from other invertebrates scored highly. This is in contrast to a number of previous studies in which mollusc MT sequences were most similar to those from vertebrates (Dallinger et al. 1993; Mackay et al. 1993). However, previous results are based on full-length MT sequence comparison, whereas the current study only uses exon 2. It appears that *P. canaliculus* exon 2 MT sequences share greater homology with invertebrates than vertebrates. These results are interesting and indicate that the arrangement of metal-binding domains within different species may be of significance for determining evolutionary relationships for MT sequences.

CHAPTER 3: General Methods

3.1 Introduction to proteomics

Proteomics is the systematic, large-scale separation, identification and quantification of a large number of proteins and is based on the concept of the proteome. The proteome is the complete PROTEin complement expressed by the genOME of a given cell or organism under a defined set of conditions ("Proteome" first coined by Wasinger et al. 1995). Alterations in the environment experienced by an organism (or cell) can lead to qualitative and/or quantitative changes in proteome patterns. In the current study, a proteomic approach was used to identify proteins that respond to low salinity and/or heavy metal stress in the endemic NZ greenshell mussel.

Two-dimensional electrophoresis (2-DE) is a commonly used method for proteome analysis, because it is capable of simultaneously resolving thousands of proteins using relatively simple separation procedures. The first dimension step, isoelectric focusing (IEF), separates proteins according to their isoelectric points (pI), the second-dimension step, SDS polyacrylamide gel electrophoresis (SDS-PAGE), separates proteins according to their molecular weights (M_r relative molecular weight). Each spot on the resulting 2-D gel typically corresponds to a single major protein, or sometimes a mixture of proteins with similar pI and Mr values.

After proteins have been separated using 2-DE, image analysis software (e.g. ImageMaster) can be applied to the gel images to detect, quantify, match and analyse protein patterns. Protein quantification statistics can identify protein spots that show variation across gels that may be related to a particular treatment or experimental condition. However, problems arise due to inherent variation between gels. For example, protein patterns may not match if gels in a second set were separated using slightly different electrophoresis conditions, or if some samples enter the gel matrix more efficiently in a few gels and not others, thereby skewing protein abundance comparisons. To ensure meaningful results, differences in protein abundance should be due to biological variations and not gel inconsistencies.

It is also necessary to separate the induced biological changes within an experiment from the inherent biological variation present within organisms i.e. the genetic differences between two individuals that are present irrespective of the

applied experimental test conditions or separation techniques. To achieve this, multiple biological sample replicates must be incorporated within the experimental design. This requires the separation and analysis of a large number of samples and can be a slow process if each sample has to be separated on a different gel. The Ettan™ DIGE system utilised in this project largely overcomes gel-to-gel discrepancies by multiplexing samples and using an internal standard to calculate differences in protein abundance. The need to run gel replicates is also addressed, reducing the overall number of gels required per experiment.

3.2 Ettan™ DIGE

Ettan™ DIGE (Differential In Gel Electrophoresis) is a method for pre-labelling protein samples with fluorescent dyes (CyDyes™) prior to 2-D electrophoresis, allowing the multiplexing of differently labelled protein mixtures on the same 2-D gel. The system includes CyDye DIGE Fluor Cy2, Cy3 and Cy5 minimal labelling dyes, which are mass and charge-matched, spectrally resolvable, fluorescent dyes. CyDye DIGE Fluors have an NHS ester reactive group, and are designed to covalently attach to the epsilon amino group of lysine of proteins. The quantity of dye added to the sample is limiting in the reaction hence this method is referred to as 'minimal' labelling. This ensures that the dyes label approx 1-2% of the available lysine per protein molecule. The lysine amino acid in proteins carries a +1 charge at neutral or acidic pH. CyDye DIGE Fluors also carry an intrinsic +1 charge which, when coupled to the lysine, replaces the lysine charge with its own ensuring that the pI of the protein is not altered significantly. When coupled to the protein CyDye DIGE Fluors add approximately 500 Da to the protein mass in a uniform manner. A protein sample labelled with any of the CyDye DIGE Fluor dyes will migrate to the same position on a 2-D gel allowing multiplexing.

Multiplexing, the co-migration of more than one sample on a 2-D gel enables the inclusion of an internal standard. Ideally, the internal standard should consist of a pool taken from all of the samples within the experiment. This creates an image that is the average of all experimental samples, with all the proteins in the experiment represented. The presence of the internal standard on every gel provides an intrinsic link between samples. The DeCyder 2D™ automated image analysis software can use the internal standard to derive statistical data on protein changes within and between gels. Protein abundance within each sample is compared against the

standard on the same gel to generate a ratio of relative protein levels. Quantitative comparisons of samples between gels are made based on the relative change in abundance of proteins within each sample to its in-gel internal standard. Using the internal standard effectively eliminates gel-to-gel variation, allowing detection of even very small changes in protein abundance (Figure 3.1).

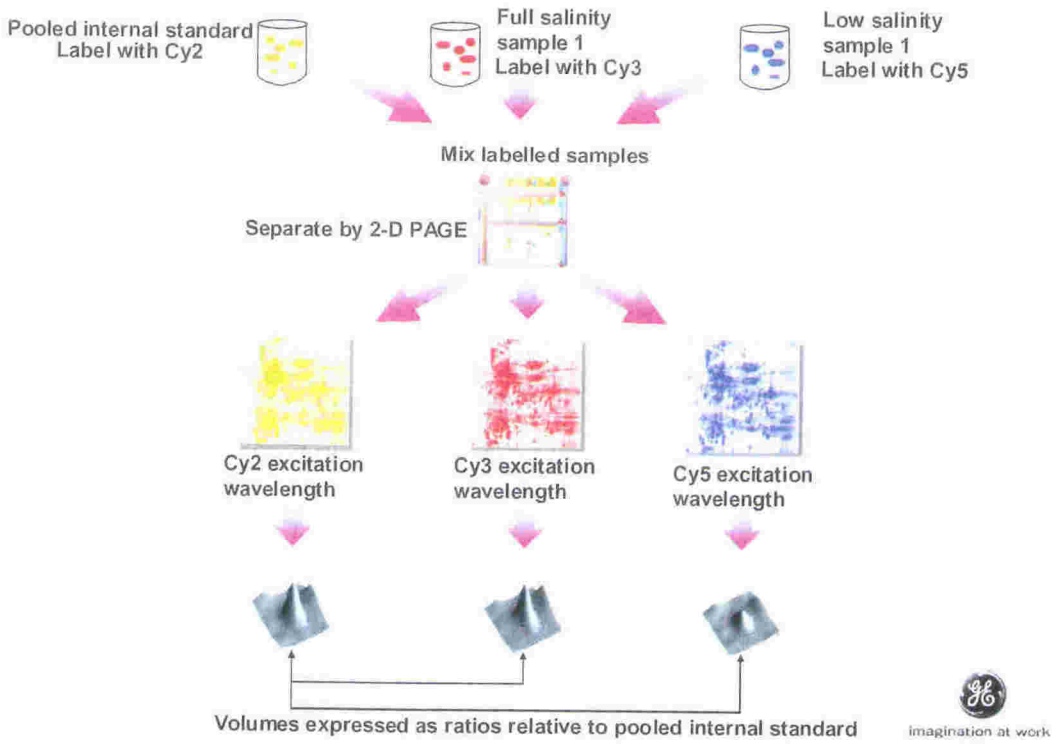


Figure 3.1: Stages of Ettan™ DIGE
Figure reproduced from GE Healthcare (Sweden) manual.

According to the GE Healthcare (Sweden) manual the Ettan™ DIGE system offers a number of advantages over traditional 2D systems including:

- Accurate quantitation and statistical analysis of protein abundance changes (by comparison with standard).
- High sensitivity and wide dynamic range (5 orders of magnitude).
- Minimisation of system (i.e. gel-to-gel) variation.
- Easier matching between gels, with increased confidence.
- Fewer gels required per experiment.
- Faster analysis due to a fully automated gel-processing workflow.

Samples in this research project have been analysed using Ettan™ DIGE, arguably the most technologically advanced and accurate quantification technique currently available, and the only system that uses a gel-based internal standard consisting of the experimental samples themselves.

3.3 Factors for a successful proteomic analysis

A successful proteomic analysis depends on a number of factors, including experimental design, sample selection and preparation, isoelectric focusing, 2-D electrophoresis, image detection and analysis, spot identification, and database searches. The purpose of this chapter is to provide general details of these methodologies as used in the proteomic experiments within this project. Additional details specific to each experiment are outlined in each of the subsequent chapters (Chapters 4-6).

3.4 Experimental design

Assessing Statistical significance

Univariate methods such as the Student's *t*-test calculate the probability (p) that groups to be compared are the same (i.e. there is no different in protein abundance between groups) and that any differences arise from sampling variation. A change in protein abundance is deemed significant if the calculated p value falls below a prescribed level, typically 0.01 (the 'nominal significance level'; Karp and Lilley 2005). Two types of errors are possible: Type I (α): A false-positive error occurs when a change in protein spot abundance is erroneously detected, and a Type II (β): A false-negative error occurs when the test fails to detect authentic changes in spot abundance. Power ($1 - \beta$), is the ability of a univariate test to detect change, which depends on the variance (noise), effect size (change in protein abundance), number of replicates and the nominal significance that the researcher sets. A power analysis by Karp and Lilley (2005) using DIGE suggests that 10 replicates provide sufficient sensitivity to detect changes in protein abundance of 1.5 fold (Figure 3.2). In the current study, the experimental design was based on the use of 10 biological replicates to ensure that the detected changes in protein abundances reported are genuine.

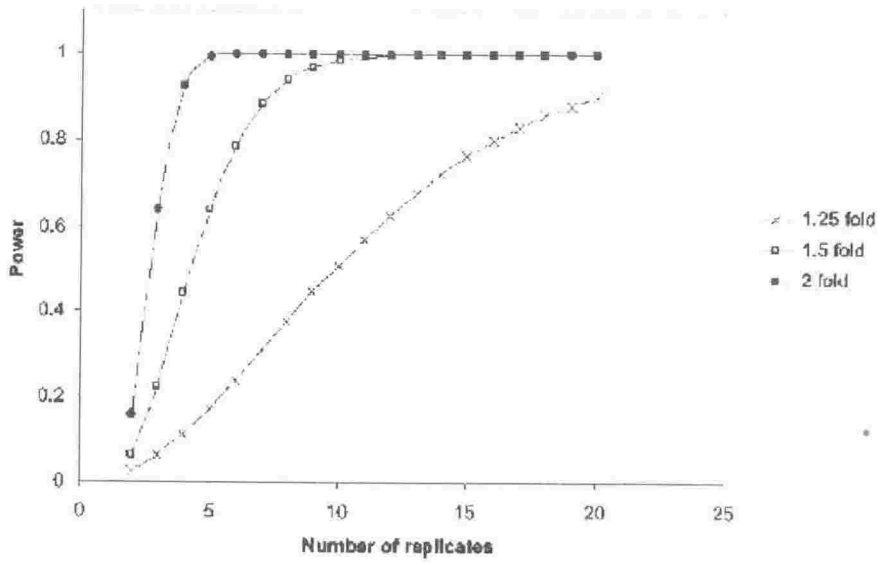


Figure 3.2: Power analysis
Figure reproduced from Karp and Lilley (2005)

Assessing Biological significance

Once a protein has been shown to exhibit statistically significant changes in abundance, the biological significance of the change should be assessed. Proteins should be identified and if possible manipulated to demonstrate a biologically relevant effect on pathways that are being investigated. However, this is beyond the scope of the current study.

3.5 Mussel soft tissue dissection for proteomic analysis

Gill tissue from *P. canaliculus* was used for the majority of the proteomic analysis as it is the most physiologically relevant tissue, being the site of initial contact with seawater. Gills, as well as being a site for metal uptake, are an important reservoir for metal storage (Roesijadi 1986). Foot tissue from *P. canaliculus* was also used in the salinity study to provide comparison for the protein patterns shown in gill. Foot tissue has been used in prior published proteomics studies to distinguish between different species of blue mussels because it produces clear spot maps (López et al. 2002a; López 2005). Protein spot identifications from foot tissue in the current study can be compared with those obtained previously. However, previous studies used different mussels (*Mytilus* spp.), different pI ranges (3-11) and different staining techniques (silver). Comparisons are only possible between abundant proteins that can be accurately recognised in both sets of gels.

Digestive tissue was also used because it is known to highly express LAP proteins, which were expected to be up regulated under the experimental conditions employed during the low salinity project. However, gut tissue also contains high levels of digestive enzymes which degrade protein producing poorly resolved and smeared spot maps (Figure 3.3). For this reason digestive tissue was not used for proteomic analysis. Mantle tissue was also dissected and found to produce clear spot maps (Figure 3.3). But because the principle role of mantle tissue is shell deposition (mantle edge) or gonad development, in comparison to gill tissue, mantle was not thought to be physiologically relevant to this experiment.

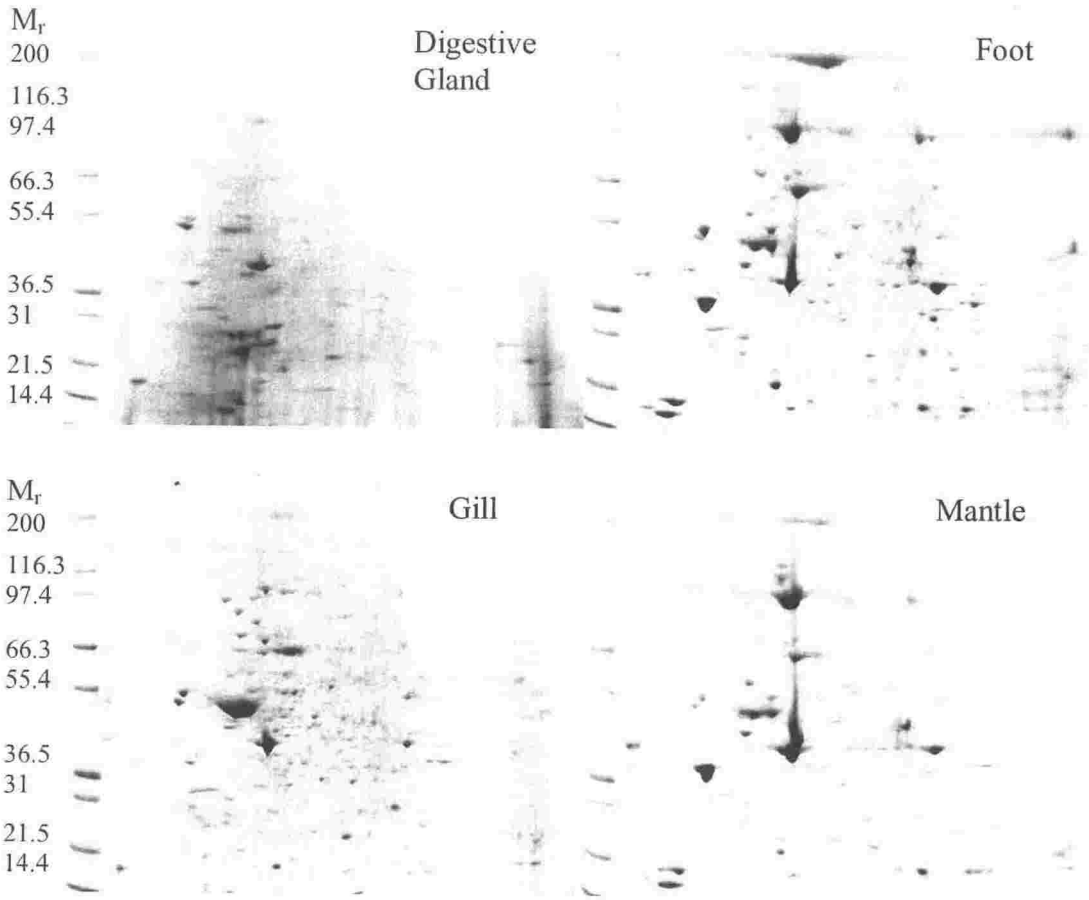


Figure 3.3: 2-D gel trials using different *P. canaliculus* tissues
pI 3-10 7 cm IPG strips were used. The numbers to the left of the gels indicate approximate relative molecular weight (kDa) estimated using a Mark12 (Invitrogen) ladder

3.6 Heavy metal analysis

In the heavy metal studies the levels of selected elements accumulated in mussel tissue and seawater were determined by commercial laboratories,

Environmental Laboratory Services (ELS) and Hill Laboratories, as specified in each of the relevant chapters (Chapter 5, 6 and 7). Methodologies used to determine element concentrations and their detection limits are described below.

3.6.1 Mussel tissue analysis (ELS)

All mussel soft tissue samples (whole animal or dissected tissue, depending on experiment) were stored in metal-free containers and frozen at -80 °C immediately after dissection. Samples were sent to ELS on dry-ice for analysis. Care was taken to ensure that samples arrived within an hour at ELS and therefore did not have a chance to thaw, which could affect metal detection. At ELS, samples were homogenised and total metal digestion was performed following an in-house method using nitric acid. All analyses were carried out using Inductively Coupled Plasma – Mass Spectrometry (ICP-MS). An ICP-MS is an instrument capable of determining the concentrations of around 40-70 elements simultaneously. The sample is introduced into the plasma, where it is vaporised, atomised, and ionised then passed through a magnetic quadrupole to the detector. The instrument is capable of ultra low detection limits of parts per trillion (ppt) for some elements. Detection limits for the elements tested were as follows:

Test	**Detection Limit**	**Test**	**Detection Limit**
Arsenic –Total	0.25 mg kg^{-1}	Chromium - Total	0.25 mg kg^{-1}
Cadmium – Total	0.005 mg kg^{-1}	Copper - Total	0.05 mg kg^{-1}
Lead – Total	0.01 mg kg^{-1}	Nickel - Total	0.05 mg kg^{-1}
Mercury – Total	0.01 mg kg^{-1}	Zinc - Total	0.1 mg kg^{-1}
Tin -Total	1 mg kg^{-1}		

Table 3.1: Detection limits for mussel tissue samples analysed by ICP-MS at ELS

3.6.2 Seawater sample analysis (Hill Laboratories)

Seawater samples required a specialised matrix for analysis which is currently unavailable at ELS. Seawater samples were analysed by Hill Laboratories instead. The seawater was collected in 50 mL metal-free plastic tubes containing a small amount of nitric acid for metal preservation. Seawater was sampled, thoroughly mixed and kept refrigerated at 4 °C until sent for analysis. Most samples were digested with nitric acid and analysed using ICP-MS. However, samples for mercury analysis were digested using permanganate/persulphate and analysed using

an automated cold vapour Hg analyser (FIMS). Detection limits for the elements tested were as follows:

Test	Detection Limit	Test	Detection Limit
Arsenic –Total	0.004 mg l^{-1}	Chromium - Total	0.001 mg l^{-1}
Cadmium – Total	0.0002 mg l^{-1}	Copper - Total	0.001 mg l^{-1}
Lead – Total	0.001 mg l^{-1}	Nickel - Total	0.006 mg l^{-1}
Mercury – Total	0.00008 mg l^{-1}	Zinc - Total	0.004 mg l^{-1}
Tin -Total	0.0016 mg l^{-1}		

Table 3.2: Detection limits for seawater analysed by ICP-MS at Hill's Laboratory

3.7 Protein Extraction

Proteins were extracted from mussel tissue (gill and foot as specified in each study) using DIGE lysis buffer (30 mM TrisCl, 7 M urea, 2 M thiourea, 4% CHAPS, Appendix 2). Urea solubilises and denatures proteins, unfolding most proteins to their fully random conformation, with all ionisable groups exposed to solution. The use of thiourea in addition to urea has been found to further improve solubilisation. A non-ionic or zwitterionic detergent (e.g. CHAPS) was included in the sample solution to enhance solubilisation of hydrophobic proteins and to prevent aggregation through hydrophobic interactions (for a review of sample preparation see Herbert 1999 and references therein). In some cases, buffers or bases (e.g. 40 mM Tris base) were added to the sample solution. This was done when basic conditions were required for solubilisation or to minimise proteolysis. Reducing agents are frequently included to reduce disulfide bonds and to maintain proteins in a reduced state. The most commonly used reductant is dithiothreitol (DTT) at concentrations ranging from 20 to 100 mM. However, the addition of DTT is known to reduce the labelling efficiency of CyDye because thiols will bind at increased concentrations, therefore DTT was not added until after the labelling reaction had been performed (Section 3.8).

Proteins were solubilised on ice for 2 min with a pestle in a microcentrifuge tube containing pre-wet sand. Additional lysis buffer was added to make the sample up to the required volume (as specified in each chapter). Samples were mixed by inversion, left at room temperature for at least 30 min and then centrifuged at 10,000 *g* for 5 min. The supernatant was transferred to a clean tube. Protein concentration was estimated using a 2D Quantification Kit following the manufacturer's

instructions (GE Healthcare, Sweden). Protein concentrations of 5-10 μg/μl were required for DIGE labelling.

3.8 DIGE Fluorescent labelling

An Ettan™ DIGE protein labelling kit was purchased from GE Healthcare and proteins extracted from mussel tissue (above) were labelled according to the manufacturer's directions. Briefly, high quality anhydrous N, N-dimethylformamide (99.8%, Sigma-Aldrich) was used to prepare the stock CyDye™ fluorescent solutions (1000 pmol/μl). Working fluorescent solutions were prepared from these stocks and mixed with the appropriate amount of protein. Working fluorescent dye to protein ratios used for labelling were equivalent to those recommended (400 pmol dye/50 μg protein - recommended c.f 80 pmol/10 μg - actual). Volumes were decreased to reduce costs given that 10 μg of labelled proteins per sample per gel is sufficient to generate suitable analytical gels (Dr P. Rawson, VUW pers. comm.). For each gel a total of 30 μg of labelled protein was loaded (i.e. 3 X 10 μg per sample, each labelled with a different dye). To ensure efficient labelling it was critical to check samples were at pH 8.5 before adding the dye. If required the pH was adjusted using 1.5 M Tris pH 8.8. For preparative gels 75 μg of a single protein sample was loaded per gel to provide enough protein for mass spectrometry analysis. Although 75 μg is at the low end of the range recommended for preparative samples (240 μg maximum), 75 μg was sufficient for analysis in most cases. It was found that overloading protein samples on the small 7 cm strips/gels as used in this study resulted in poor spot resolution, especially when using basic (pI 6-11) strips. The resultant spot maps produced by overloading were often unable to be matched with analytical gels (results not shown) and were therefore of little use.

A pooled sample comprising all the individual samples in the experiment (e.g. n = 30 for the Salinity experiment, Chapter 4) was labelled with Cy2 and was used as the internal standard. For each treatment (e.g. control, ambient or low salinity; Salinity experiment) half of the samples were labelled with Cy3 and the other half with Cy5 (see labelling Table associated with each individual chapter for clarification). Samples were mixed thoroughly by vortex and centrifuged briefly to bring all of the solution to the bottom of the tube. Samples were left on ice in the dark for 30 min before 1 μl of 10 mM lysine was used to stop the labelling reaction by binding excess dye. Samples were centrifuged briefly at room temperature and

left on ice in the dark for 10 min. At this point samples could be frozen at -80 °C until required, or prepared immediately for 2-DE (Section 3.9).

3.9 Sample preparation for 2-DE

3.9.1 Combine appropriately labelled samples

Samples labelled with different CyDyes™ were combined in a microcentrifuge tube as described in the Tables presented in Chapters 4-6. An equal volume of DIGE 2X sample buffer (7 M urea, 2 M thiourea, 2% 3-10 IPG buffer, 2% DTT, 4% CHAPS) was added, and the solution was left on ice for at least 10 min. Carrier ampholytes or IPG buffer (up to 2% v/v) enhance protein solubility by minimising protein aggregation due to charge-charge interactions. Ampholytes are primary amines and, as such, would compete with protein for CyDye; hence they were not added until after labelling. Samples could be frozen at -80°C or preparation for electrophoresis could continue (Section 3.9.2).

3.9.2 Isoelectric focusing

Isoelectric focusing (IEF) is an electrophoretic method that separates proteins according to their isoelectric points (pI). Modern systems use an immobilized pH gradient (IPG) that is created by covalently incorporating a gradient of acidic and basic buffering groups into a polyacrylamide gel that has a plastic backing to enhance gel stability and experimental reproducibility. For this study Immobiline™ dry strips (7 cm) with a pI range of 4-7 and 6-11 were purchased from GE Healthcare. These two strip ranges provided protein coverage over the major pI's with good protein resolution. Wide range (3-10) strips were trialled, but did not give good gels (Figure 3.3). Basic pI 6-11 strips were rehydrated overnight in 125 μl of 6-11 rehydration buffer (7 M urea, 2 M thiourea, 2% CHAPS, 10% isopropanol, 5% glycerol, 1% pI 6-11 IPG buffer, 2.5% DTT) using a GE Healthcare reswelling tray to promote even uptake of the solution. This 6-11 rehydration buffer is slightly modified from the standard buffer by the addition of glycerol and isopropanol to counteract water transport during isoelectric focusing (Hoving et al. 2002). Strips were overlaid with 3 mL of PlusOne dry strip cover fluid (GE Healthcare), an electrically isolating and inert fluid, which ensures good heat conductance during isoelectric focusing and protects the strips from drying out and the samples from oxidation. Labelled protein samples combined with 2X sample buffer (Section 3.9.1)

were made up to a total volume of 100 µl using 6-11 rehydration buffer and left at room temperature overnight. The maximum recommended protein loading volume (100 µl) for a 7 cm strip was used to facilitate protein entry into the strip and to prevent point precipitation of proteins during the focusing process. The following day the protein samples were centrifuged briefly at 10,000 g, cup-loaded onto the prepared strips and focused using the Ettan IPGphor Isoelectric Focusing system using the protocol outlined in section 3.9.

To isolate proteins with a pI range between 4-7, labelled samples combined with 2X sample buffer (Section 3.9.1) were made up to 125 µl with 4-7 rehydration buffer (8 M urea, 1% 4-7 IPG buffer, 0.2% DTT, 4% CHAPS). Immobiline™ dry strips (4-7, 7 cm) were rehydrated overnight using this 4-7 rehydration buffer plus sample solution in a GE Healthcare reswelling tray (i.e. rehydration loading). Strips were overlaid with 3 mL of PlusOne dry strip cover fluid. Samples were separated by electrophoresis on an Ettan IPGphor Isoelectric focusing system (Section 3.10).

3.10 First dimension (1-D) isoelectric focusing (IEF)

Strips were focused using an Ettan IPGphor Isoelectric Focusing system. The manifold and strips were prepared according to the manufacturer's instructions, but instead of preparing wicks with 150 µl MilliQ ddH_2O, the cathodic wicks for 6-11 strips were soaked with 150 µl of 6-11 rehydration solution plus 3% DTT (Hoving et al. 2002). Basic proteins are difficult to separate in the first dimension because depletion of DTT at the cathode can lead to reformation of intra- and inter-molecular disulphide bridges due to the oxidation of sulphydryl groups (Pennington et al. 2004). This can cause proteins within the sample to become less soluble, leading to horizontal streaking within the gels and thus poor resolution of the proteins (Pennington et al. 2004). The incorporation of DTT into the wick allowed DTT to flow into the strip, keeping the proteins in their reduced state, resulting in improved focusing. Both 4-7 and 6-11 IPG strips were used with the same focusing protocol (Table 3.3).

	Voltage	Minutes
Step	300	30
Gradient	1000	30
Gradient	5000	90
Step	5000	25
TOTAL		2 hr 55 mins

Table 3.3: Isoelectric focusing conditions for pI 4-7 and 6-11 IPG strips.

3.11 Second dimension (2-D) SDS-PAGE

SDS-PAGE (SDS-polyacrylamide gel electrophoresis) is an electrophoretic method for separating proteins according to their molecular weights. The technique is performed in polyacrylamide gels containing sodium dodecyl sulphate (SDS). SDS masks the charge of the proteins themselves and the anionic complexes formed have a roughly constant net negative charge per unit mass. The amount of SDS bound to a protein and the additional negative charge is directly proportional to the mass of the protein, thus electrophoresis of proteins through a sieving gel in the presence of SDS separates proteins on the basis of molecular mass.

Prior to electrophoresis in the second dimension strips were equilibrated for 10 min in 2.5 mL/strip equilibration solution [1.5 M Tris Cl pH 8.8 (final concentration 50 mM), 6 M urea, 30% glycerol (v/v), 2% SDS (w/v), 0.001% bromophenol blue (w/v): Appendix 2] plus 0.25% DTT, then for 10 min in 2.5 mL/strip equilibration solution plus 4.5% iodoacetamide. Equilibration buffer contains 50 mM Tris-HCl, pH 8.8 which maintains the IPG strip pH in a range appropriate for electrophoresis. Urea (6 M) together with glycerol (30%) reduces the effects of electroendosmosis by increasing the viscosity of the buffer, thus improving transfer of protein from the first to the second dimension. DTT preserves the fully reduced state of denatured, unalkylated proteins. SDS denatures proteins and forms negatively charged protein-SDS complexes. Iodoacetamide alkylates thiol groups on proteins, preventing their reoxidation during electrophoresis which can result in streaking and other artefacts. Iodoacetamide also alkylates residual DTT to prevent point streaking and other silver-staining artefacts and is used to minimise unwanted reactions of cysteine residues (i.e. when mass spectrometry is to be performed on the separated proteins).

While the IPG strips were equilibrating NuPAGE 4-12% Bis-Tris gels (Invitrogen) were rinsed with MilliQ ddH_2O and positioned in the NOVEX mini-cell tanks (X Cell *Sure*Lock™, Invitrogen). The tanks were filled with 1X MOPS buffer (Invitrogen) and 500 μl of antioxidant was put in the middle chamber to maintain proteins in their reduced state. Gels were electrophoresed at 200 V, 140 mA for 60 min.

3.12 DeCyder™ analyses

Immediately following electrophoresis, gels were rinsed with MilliQ ddH_2O and scanned using a Fuji film FLA-5100 scanner. Each gel was scanned at 3 different wavelengths, 473 nm for Cy2, 532 nm for Cy3 and 635 nm for Cy5. Gels were fixed overnight (50% ethanol, 2% phosphoric acid: Appendix 2), and then stained using colloidal Coomassie blue stain (Appendix 2). DIGE gel images were cropped and converted to the appropriate file format using ImageQuant tools (V2003) and loaded into DeCyder™ (Version 6.5 running on an Oracle 10g server) for analysis. Each image was assigned to the appropriate experimental group, e.g. control, high salinity and low salinity. Spots were detected and quantified using the DIA (Differential In-Gel Analysis) module on a set of images from the same gel. During this phase inaccurately matched spots were corrected and filtering was used to exclude background 'noise'. Matching multiple images from different gels was done using the BVA (Biological Variation Analysis) module. Although matching between gels was done automatically, manual checks were made to ensure that the matching was correct. In the BVA module, statistics could be used to identify spots that showed significant variation in protein abundance between different experimental groups. For this project, a Student *t*-test p value of < 0.01 was considered significant. Some protein spots that showed significant and large (i.e. greater than 2-fold) variations in protein abundance were picked from the gels for identification using Matrix Assisted Laser Desorption/Ionisation – Time of Flight (MALDI-TOF) mass fingerprinting or tandem mass spectrometry. Further details on these analysis modules and statistics are discussed below.

3.12.1 Differential In-gel Analysis (DIA) Module

Information in the following sections has been summarised from the DeCyder manuals. For more information please refer to the GE Healthcare DeCyder 2D Software, Version 6.5 User Manual and GE Healthcare DeCyder Extended Data Analysis Module Version 1.0 User Manual (Module for DeCyder 2D version 6.5).

Each gel was scanned at 3 different wavelengths resulting in 3 linked images, incorporating a pooled internal standard (Cy2) and two samples (Cy3 and Cy5). Protein spots were co-detected by the Differential in-gel analysis module (DIA) to ensure that spots were represented on all images within the set. DIA algorithms then

quantitate spot protein abundance ratios by dividing the volume of the secondary image spot by the volume of primary image spot. This ratio indicates the change in spot volume between the two images and is referred to as the volume ratio. The volume ratio describes the relative abundance of proteins in terms of fold increase or fold decrease. Volume ratios of 1.5 or -1.5 would indicate a 1.5 fold increase or decrease respectively. The volume ratio was used to quantify protein spot abundance against the internal standard to allow accurate inter-gel protein spot comparisons.

3.12.2 Biological Variation Analysis (BVA) Module

Biological variation analysis (BVA) processed gel images that had undergone spot detection in DIA. The BVA module matched all images to a single master image, identifying common protein spots across gels. Statistical analysis tools were used to highlight proteins that demonstrate significant changes in protein abundance. The standardised volume ratio for each internal standard from the different gels is set to the value of 1.0. The protein abundance ratio for each sample spot is then related to its corresponding standard spot in the same gel, thus making it possible to compare ratios between matched protein spots in different gels. These standardised results are converted to standardised log abundances which are the basis of the statistical tests (Section 3.12.3).

3.12.3 Overview of statistical tests and terms used in this study

Volume ratio: Normalized volume ratio between co-detected spots in the primary and secondary images of DIA.

Average ratio: The average ratio gives a measure of the differences in protein abundance between two groups. For independent samples, the average ratio is calculated as the degree of difference in the means of the standardised abundance between protein spots in different groups and is reported as fold-changes:

Average Ratio = m_a/m_b

Where $m_{a,b}$ is the mean of the protein abundance values in group a or b.

A two-fold increase or decrease is represented by +2 and -2 respectively (not 2 and 0.5). Note that the log standardised protein abundance is the only variable utilised by the statistical analysis tools within the DeCyder BVA module, whereas protein abundance values are not logged in the calculation shown above.

***Student's t-test*:** The Student's *t*-test is used to test the hypothesis that a variable differs between two groups or populations. The Student's *t*-test is performed as an equal variance two-tailed test, therefore, direction of change (i.e. increases or decreases) in the standardised abundance parameter is considered. This test requires a minimum of two data points in each of the groups and assumes small sample size and normally distributed data. The Student's *t*-test null hypothesis is that there is no change in the protein abundance between experimental groups (i.e. that the average ratio between two groups is 1).

$$t = \frac{\mu_a - \mu_b}{\sigma_{a-b}}$$

where $\mu_a - \mu_b$ is the difference in means between two groups

and σ_{a-b} is the variance

Therefore, the *t*-test *p* value represents the probability of obtaining the observed data if the two groups had the same protein abundance. For example, if the *p* value is 0.01 then the probability of obtaining the observed difference in protein abundance by stochastic variation alone is 1 in 100. Protein abundance differences are generally considered to be statistically significant when $p < 0.05$. However, because DIGE is a sensitive technique, in this study $p < 0.01$ was used to designate significant changes in protein abundance.

***One-Way ANOVA (Analysis Of Variance)*:** At its lowest level, ANOVA is essentially an extension of the logic of the Student's *t*-test to those situations where the concurrent comparison of the means of three or more samples is required. Thus, when comparing two means, ANOVA will give the same results as the *t*-test for independent samples (if comparing two different groups or observations). In the current study One-Way ANOVAs were calculated to evaluate differences in protein abundance among all assigned groups.

***False discovery rate (FDR) correction*:** When testing thousands of proteins for statistical significance with Student's *t*-test or ANOVA, many of the proteins may appear to show statistically significant differences in abundance, but several of these proteins may have achieved this significance by chance alone. Where appropriate

FDR adjustments (Benjamini and Hochberg 2000) have been applied to data as stipulated in each chapter (Chapters 4-6).

3.12.4 Extended Data Analysis (EDA)

Extended Data Analysis (EDA) is an add-on module for the DeCyder 2D Software. It is used for multivariate analysis of protein abundance data derived from the BVA module. As well as the Student's *t*-test and ANOVA analyses that can be performed in the BVA module, there are a number of additional analyses in EDA. The two features that were used in the current study are described briefly below.

Principal Component Analysis (PCA)

PCA produces an overview of the data and can be used to find outlying data. Two analyses were used in the current study as described below.

1) Protein versus Spot Maps calculations: detect protein outliers and provide a rough overview of the relationship between protein abundance and spot maps (Figure 3.4; panels A and B). In panel A, any proteins outside of the 95% significance value (the black ellipse) can be considered outliers compared to the rest of the data and should be checked for authenticity. These proteins are usually proteins with high average ratio, or proteins that were mismatched within BVA. The positions of the gels within the loading plot (Panel B) indicate the relative abundance of proteins (Panel A). So for example, proteins in quadrant a2 and a4 are probably up-regulated in spot maps within b2 and b4 and down regulated in spot maps b1 and b4. In the example shown, proteins with high abundance in spot maps from group 1 (green dots) showed low abundance in spot maps from Groups 2 and 3 (pink and red dots) and vice versa.

2) Spot Maps versus Protein calculations: check if there are any spot map outliers. Spot maps from the same experimental group should be grouped together. In panel C the spot maps from group 1 (green dots) clearly cluster together, the spot maps from Groups 2 and 3 also form a discrete cluster separate from Group 1. Groups 2 and 3 can not be separated from each other. The associated protein loading plot (panel D), indicates general relationships between protein abundance and spot maps in the same manner as described for Protein versus Spot Maps calculations.

The difference between Score plots and Loading plots

DeCyder datasets consist of variables and observations. PCA is used to find combinations of variables that "explain" the variance seen in the observations (Table 3.4). Because DeCyder EDA allows proteins and spot maps to be selected as either observations (scores) or variables (loadings), the resulting PCA output will vary slightly depending on the parameters chosen. The score plots show observations plotted against the first two PCs and these graphs show which observations were "similar" with respect to the chosen variables. The loading plots show the variables plotted against the first two PCs which show the relationship between variables and can indicate what the most important variables are. Score plots provide the most valuable information in terms of illustrating relationships between experimental spot maps and protein outliers. The loading plots provide additional information.

Observations	**Variables**		
	Gel 1	**Gel 2**	**Gel 3**
Spot 1	std vol		
Spot 2			
Spot 3			

Table 3.4: Example dataset, grey shaded area indicates protein spot standardised abundance values
NB: This could also be shown with spots as the variables and gels as observations

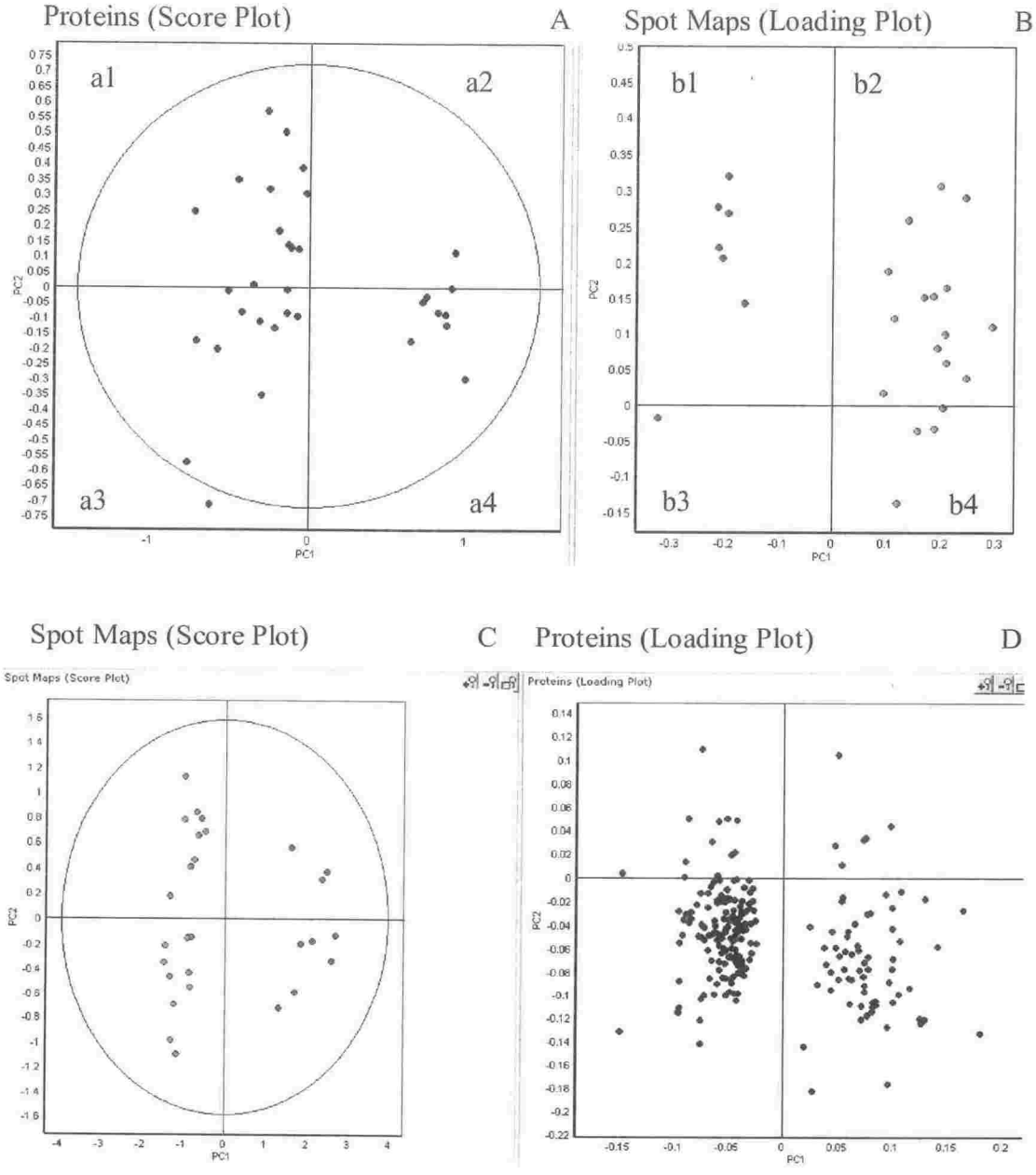


Figure 3.4: An example of a PCA

Key:

- Group 1
- Group 2
- Group 3

The black dots in panels A and D represent individual proteins and the coloured dots in panels B and C represent spot maps.

Pattern analysis

Pattern analysis groups together proteins and spot maps with similar protein profiles. Four different types of pattern algorithms were available, but in the current experiment Hierarchical clustering was utilized. Without having any known parameters, Hierarchical clustering can be used to find proteins that co-vary.

Default settings were used, Euclidean (Distance metrics) and Average Linkage (Linkage method). Data produced are displayed in the form of one or two dendrograms (depending on the calculations performed) together with a 'Heat Map'. The dendrogram orders the data so that similar data is displayed next to each other. The heat map is a coordinate system with proteins on the y-axis and spot maps/experimental groups on the x-axis (e.g. Figure 3.5). Each coordinate shows the relative abundance of a protein (log standard abundance) on a spot map on a colour scale. The colour scale varies from green (decreased protein abundance) to black (no change in protein abundance) to red (increased protein abundance). If no data exist for a coordinate (missing value) this coordinate is displayed in grey.

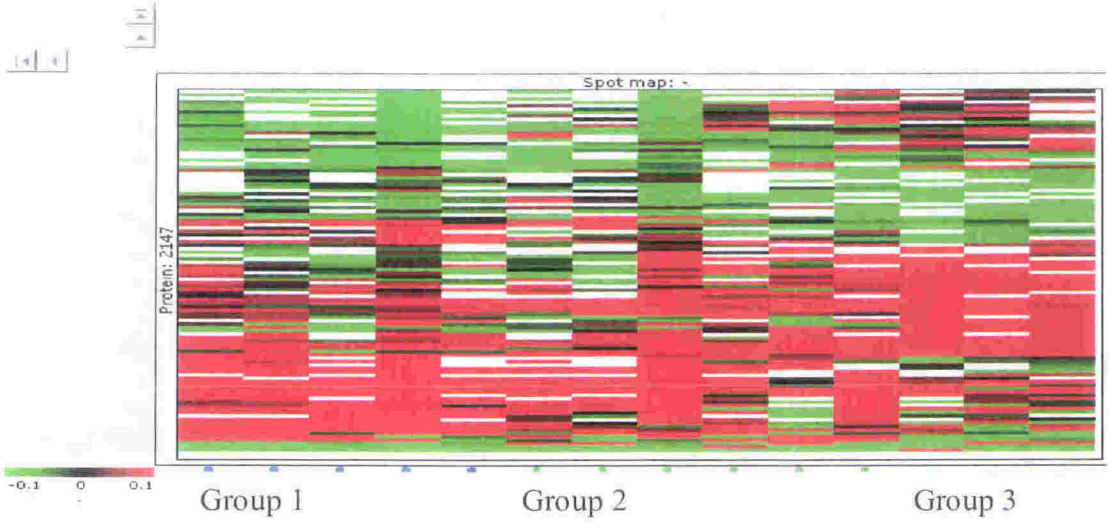


Figure 3.5: An example of a pattern analysis
Each coloured dot represents an individual spot map. In this example, individuals from the same groups cluster together as expected.

3.12.5 PERMANOVA

DeCyder could not perform all of the statistical tests required for analysis. In those instances permutational multivariate analysis of variance was conducted on the basis of Bray-Curtis dissimilarities using PERMANOVA v1.6 (Anderson 2001; McArdle and Anderson 2001) or an R programme developed specifically for these analyses (by Dr Shirley Pledger, School of Mathematics, Statistics and Computer Sciences, Victoria University of Wellington). Permutation randomly reorders the data by 'shuffling' factor labels. If there is no effect of the factor on the data then it is equally likely that any of the individual factors could have been associated with

any one of the observation units. If the null hypothesis were true (i.e. no effect of the factor), then the F-statistic actually obtained with the real ordering of the data will be similar to the values obtained under permutation. If, however, there is a significant effect of a factor on the data, then the value of F obtained with the real data ordering will appear large relative to the permutation (Manly 1997; Anderson 2001).

3.13 Sorting gels for analysis

Some gels contained spots that were poorly resolved and separated and were difficult to match with other gels. For example, many high molecular weight protein spots were removed from analysis because they tend to occur in a 'ridge' formation across the gel, rather than occurring as individual protein spots, meaning that accurate matching across gels could not be confirmed. In some gels, the software may divide a 'ridge' into 9 proteins, whereas in another gel only 6 'spots' were detected in the same ridge formation. The problem lies in determining whether the differences in the number of spots detected in each gel are 'real' representing a real difference in the abundance of these proteins, or an artefact of the DeCyder software. The 'spot' volumes that are compared between gels may not be equivalent simply because the ridge protein volumes have been divided among a different number of 'spots'. It may have been possible to merge spots within a ridge into a single spot, however, this would be very time consuming and it was easier to simply remove these spots from analysis. Also a lower volume of dye than required was added to some samples due to the additive effects of pipetting error (over 30 samples) which depleted the stock dye before all samples could be adequately labelled. A decrease in the amount of dye added to some samples would affect the normalisation of the spot data to the standard. Therefore, any gels where it was noted that fluorescent dye volumes were less than expected were also removed from analysis. The amount of dye added to samples was only a problem in the salinity experiments as experience led to better management in the later metal and Waitangi experiments. Removal of some gels as a precaution to increase confidence in the accuracy of the results led to a decrease in the number of samples retained for analysis. The number of samples retained for analysis is outlined in each of the relevant chapters (Chapters 4-6).

3.14 Identification of protein spots using MALDI-TOF

3.14.1 Destaining

Spots were prepared for MALDI-TOF using the following standard protocol. Spots extracted from 2D gels were washed 3X with 100 μl of MilliQ ddH_20 and destained using 1:1 ammonium bicarbonate (50 mM): acetonitrile (ACN). Spots were destained for as long as required, changing to fresh destaining solution every 45 min. When spots were fully destained spots were washed once with fresh destain solution. Approximately 100 μl of 100% ACN was added to the spot and incubated at room temperature for 3-5 min. At this stage, the spot was fully dehydrated and the gel piece was bright white. The ACN was evaporated in a SpeedVac for 5 min.

3.14.2 Trypsin digest

A fresh aliquot of trypsin (0.25 μg/μl trypsin in 50 mM ammonium bicarbonate) was prepared and 2 μl was added directly to each gel piece. It was vital that the trypsin was taken up by the gel piece as this significantly enhances protein digestion and peptide extraction. Gel pieces and trypsin were left for at least 10 min at room temperature and when the gel piece was fully re-swollen 10 μl of 50 mM ammonium bicarbonate was overlaid on the gel piece. The tube was capped securely and incubated at 30 °C overnight.

3.14.3 Extraction of peptides

The following day the digest solution was carefully removed into a fresh tube while being careful not to disturb the gel piece. Ten μl of a 1:1 mixture of 0.2% TFA: 100% ACN was added to the gel piece and the tube was shaken on a vortex for 30 min. The solution was removed and combined with the previous digest solution. Ten μl of 100% ACN was added to the gel piece and it was shaken for another 30 min. The ACN was combined with the previous digest solution and the gel piece was discarded. The peptide solution was dried in the SpeedVac at 35 °C until fully dry.

3.14.4 Purification

Peptides were resuspended in 10 μl of 0.2% TFA and thoroughly mixed for 10 min using a vortex mixer. Samples were cleaned up using ZipTips™ or PerfectPure C-18 tips following the advice of the manufacturers (Millipore or

Eppendorf respectively). Samples were eluted in 1.5 μl of elution buffer and combined with 1.5 μl of matrix (CHCA). The full sample was spotted onto a MALDI plate and left to dry in a laminar flow hood. When peptides were fully dry MALDI spectra could be acquired.

3.14.5 Acquiring spectra

Spectra were collected using an Applied Biosystems MALDI (Voyager DE PRO) and data acquisition was managed using the Voyager control panel. Laser intensity ranged between 1700 and 2000 and usually 4 spectra were obtained per spot.

3.14.6 Analysing spectra

Spectra were analysed using Data Explorer (Applied Biosystems 5.1) and two different search algorithms; ProFound or Xproteo. Before exporting peak data to these algorithms, the peptides were calibrated using either an internal standard (trypsin peaks) or known standard spectra (Calmix, PE Biosystems, CA) that were acquired at the same time as the unknown samples. Spectra were compared to empirical and theoretical databases e.g. NCBI and Swissprot.

3.15 Identification of proteins using tandem mass spectrometry

Destaining and digestion are the same for samples analyses by MALDI-TOF or tandem mass spectrometry (LTQ: section 3.13.1 and 3.13.2). After digestion peptides are extracted from gel pieces with approximately 100 μl 50 mM ammonium bicarbonate at 37 °C for 15 min. The supernatant is removed to a clean tube (1) and replaced with 50 μl of 5% formic acid. Gel pieces and formic acid are mixed by vortex at 37 °C for 15 min. The supernatant is again removed and added to (1) above. Approximately 100 μl of 100% acetonitrile was added to the gel piece and incubated at 37 °C for 15 min. The supernatant is pooled together with the rest (1) and lyophilised in a SpeedVac. The dried sample can be stored at -20 °C until analysis. Before HPLC the peptides are redissolved in 20-25 μl 0.1% formic acid in 2% acetonitrile. Samples are centrifuged and placed into an HPLC sample vial for analysis. The tandem mass spectrometry was carried out using an LC-Packings 3000 HPLC with a 15 cm C18 reverse phase column (PepMap, 3μm, 100 Å) and a Thermo LTQ linear ion trap mass spectrometer by Dr Lifeng Peng and the spectra were analysed using the associated software (Xcalibur, Bioworks v3.1: Thermo and Mascot v2.1.03: MatrixScience).

CHAPTER 4: A proteomic approach to study salinity induced protein expression signatures in the endemic New Zealand green mussel *P. canaliculus*

Abstract

In the current study the effect of low salinity on protein abundance was investigated in *P. canaliculus*. Control mussels were killed at the beginning of the study and experimental animals were either exposed to low salinity (14 ppt) or kept at ambient salinity (32 ppt) for 3 days. Proteins were isolated from mussel gill and foot tissue for proteomic analyses. Approximately 115 proteins showed significant (*t*-test $p < 0.01$) differences in abundance between the three experimental groups. Differences in the number of proteins showing statistically significant changes in abundance were found between tissues; foot (approximately 10%) and gill (approximately 1%) and between treatments. Changes in protein abundance among mussels exposed to low salinity compared to control and ambient groups were not as plentiful as expected, instead the largest number of differences in protein abundance was observed in comparisons between control and ambient groups. Variation in anoxic metabolism and ammonia and oxygen concentrations between treatments are discussed as possible explanations for these unexpected results. One isoform of actin and two isoforms of tropomyosin were identified. These proteins have been previously implicated in response to physiological stress caused by salinity.

4.1 Introduction

4.1.1 Environmental proteomics

Environmental proteomics assesses the ability of organisms to respond to various stress inducers (e.g. temperature, salinity, heavy metals), by analysing the protein complement of the organism's genome (i.e. the proteome). Protein expression profiles produced by exposed experimental organisms can be compared with profiles from control organisms that have not been subject to stressful conditions, to identify proteins involved in various 'stress responses'. Proteomics detects proteins that have had their abundance altered in response to experimental conditions without requiring any previous knowledge about the mode of action or function of the protein itself (López-Barea and Gomez-Ariza 2006). Once identified,

these proteins can be developed as new biological protein markers (biomarkers) to provide novel information about the biochemical pathways involved in a range of 'stress' response situations.

However, environmental proteomics has drawbacks. For example, protein identification is difficult if there are no related proteins or genes for comparison present in sequence databases. Initial studies focused on model organisms (e.g. yeast, mice and rats) exposed to model pollutants, including various heavy metals and organics, to facilitate protein identification (e.g. Hu et al. 2003; Noël-Georis et al. 2004). Many proteins were successfully identified in yeast (*Saccharomyces cerevisiae*) exposed to Cd^{2+}, including antioxidant enzymes, heat-shock proteins, proteolysis related compounds and general metabolic enzymes (Vido et al. 2001). To date, relatively few proteomic studies have been carried out in animals from ecosystems (López-Barea and Gomez-Ariza 2006).

4.1.2 Environmental biomarkers

Mussels are regarded as suitable sentinel organisms for monitoring the environment and are therefore prime candidates for environmental proteomics analysis. Blue mussels (*M. edulis*) exposed to pollutants and analysed by two-dimensional gel electrophoresis led to the identification of unique protein expression signatures (Shepard et al. 2000) which were correlated with the individual pollutants that were applied (i.e. PCB, salinity and Cu). Protein expression signatures are defined as the particular set of proteins observed in 2-D images that act as markers of physical, chemical or biological stress (Bradley et al. 2002). The difficulty in identifying proteins in invertebrates, which are poorly represented in the protein and DNA databases, was circumvented by the use of protein expression signatures (Shepard and Bradley 2000). The function and identity of many protein expression signature proteins cannot currently be determined due to the relative paucity of invertebrate protein and sequence information in the databases. However protein expression profiling creates a record of the proteins associated with a certain 'stress' state. Identification of the protein can occur at a later date when the databases become more complete.

In the laboratory, environmental proteomic techniques have been applied to mussels exposed to Cu, Aroclor and salinity stress (Shepard et al. 2000), rainbow trout treated with diazinon, nonylphenol, propetamphos and sewage treatment plant

effluents (Bradley et al. 2002) and clams exposed to model pollutants (Rodríguez-Ortega et al. 2003b). More recently field experiments have been performed to identify protein patterns associated with *M. edulis* exposed to heavy metals and polyaromatic hydrocarbons (Knigge et al. 2004) and to identify differences in carbonylation patterns between *M. edulis* from polluted and reference sites (McDonagh et al. 2005). An environmental proteomics approach using ecologically relevant species can contribute to the development of new ecosystem biomarkers that could help to monitor environmental health.

4.1.3 Salinity as a 'stress' inducer

Variations in salinity can be used to induce proteomic responses. For example, changes in protein abundance in *M. edulis* were induced by reducing salinity from 6.3 to 3 ppt (Shepard et al. 2000). Euryhaline species such as mussels may be exposed to variations in external salinity and oxygen availability throughout the day and must regulate cell volumes and amino acid concentrations to survive (Gosling 2003). Previous studies of *M. californianus* have demonstrated that mussels use a number of adaptations to cope under low salinity conditions and many of these responses are rapid (Silva and Wright 1992). Genetic adaptations to low salinity include the enzyme leucine aminopeptidase (LAP: E.C. 3.4.11.). In some populations mussels experience differential mortality depending on which LAP alleles they carry. *Lap-2* genotypes in *M. edulis* have been correlated with a heightened ability to cope with low salinity environments along a salinity gradient (e.g. Long Island Sound, Koehn et al. 1980). The Lap^{94} allele exists at highest frequency (approximately 0.55) in natural populations of *M. edulis* in high-salinity environments, and decreases in frequency to (approximately 0.12) with decreasing salinity over a distance of < 30 km (Koehn et al. 1976). Prior studies at VUW did not identify an association of adult LAP genotypes with a salinity gradient in *P. canaliculus* (Gardner and Kathiravetpillai 1997) and the 'universality' of a relationship between LAP genotype and salinity tolerance is far from established e.g. Lap^{94} occurs at similar frequencies in the Baltic (low salinity) and in the North Sea (high salinity) (Väinölä and Hvilsom 1991).

4.1.4 Current study

The overall aim of my PhD research was to assess the potential to develop biomarkers for heavy metal contamination in *P. canaliculus* (Chapters 5 and 6). However, the current chapter reports gill and foot protein expression signatures for *P. canaliculus* exposed to salinity stress. The study of proteomic expression profiles generated by heavy metal stress in *P. canaliculus* required the use of Cd and Hg (Chapter 5), two potentially dangerous heavy metals. Therefore, before conducting the heavy metal experiment this pilot study was performed using low salinity as the stress inducer. The same experimental procedures, including DIGE were used for the salinity and heavy metal experiments. Information generated during this experiment is relevant and useful to this project as a whole. For example, biomarkers in common between salinity and metal challenge experiments could indicate 'general' stress proteins. Specific biomarkers for salinity could potentially be developed to identify hyposaline sites subject to freshwater influence from rivers or heavy rain. It is also important to identify salinity stress markers because *P. canaliculus* is widespread throughout New Zealand, including lower salinity regions (Morton and Miller 1968). Previous studies have indicated that low salinity sites are associated with higher rates of metal accumulation (Chong and Wang 2001; Blackmore and Wang 2003). Additionally, this study could also lead to beneficial outcomes for the aquaculture industry. Freshwater runoff from rivers or heavy rain often creates a temporary layer of fresh water over aquaculture farms. *P. canaliculus* is less tolerant of low salinity than blue mussels (*M. galloprovincialis*) and is often overwhelmed and displaced on the top metre or so of the cultivation rope. Identifying *P. canaliculus* genotypes that have a high tolerance for low salinity may be of relevance to New Zealand's mussel aquaculture industry in the context of its selective breeding programme. However, this was not a major focus of this research. It was also of interest to determine whether LAP protein variants could be identified using 2D gel proteomics and whether there were any correlations with the ability of mussels to cope with low salinity.

4.2 Materials and Methods

4.2.1 Mussel sampling

Adult greenshell mussels (6-14 cm) were collected in December 2004 from Seatoun, Wellington Harbour, New Zealand. Epiphytes were removed and mussels were held overnight in circulating seawater at the Island Bay Marine laboratory. The following day 10 mussels were killed as control animals. The remaining mussels were split into 2 groups (n = 12 for each group) with similar biomass. Salinity is the total of all salts dissolved in water and is measured in parts per thousand (ppt). Half were maintained in 20 L unfiltered full salinity seawater (approximately 32 ppt; ambient salinity), the other half in low salinity seawater (unfiltered full salinity seawater diluted with distilled water; approximately 14 ppt). Salinity was estimated using a refractometer. Air was delivered to both tanks using a pump but the water was not changed (i.e. closed system). Air and water temperature were maintained at approximately 12 °C, similar to conditions at the site where the mussels were collected. After 3 days, 10 mussels from each group were killed. Foot, gill, mantle and digestive tissue were dissected from each mussel. Tissue was stored at -80 °C until required for proteomic analysis.

4.2.2 Protein Extraction

Proteins were extracted separately from 100 mg of gill and foot tissue using 100 µl of standard DIGE lysis buffer (30 mM TrisCl, 7 M urea, 2 M thiourea, 4% CHAPS, Appendix 2) and were processed as outlined in Chapter 3. Samples were made up to 1mL before measurements of protein content were made using a 2DQuant Kit (GE Healthcare). Protein concentrations of approximately 10 µg μl^{-1} were typically obtained.

4.2.3 DIGE Fluorescent labelling

Samples were labelled as outlined in Chapter 3. Briefly, an Ettan™ DIGE protein labelling kit was used. Samples were adjusted to pH 8.5 and 80 pmol of dye was added to 10 µg of protein. Samples were left on ice in the dark for 30 min, 1 µl of 10 mM lysine was used to stop the labelling reaction and samples were left on ice in the dark for a further 10 min.

4.2.4 Sample preparation for 2-DE

Individual mussel samples from each group were labelled separately, in contrast to most previous experiments where multiple animals from different experimental groups have been pooled to form a single sample for each treatment (e.g. Rodríguez-Ortega et al. 2003a; Knigge et al. 2004; McDonagh et al. 2005). Analysis of individual specific protein expression profiles may provide insight into the physiological differences and genetic polymorphisms that respond to this type of stress. Samples labelled with different CyDyes™ were combined in microcentrifuge tubes (Table 4.1; see Chapter 3 for further details). Immobiline™ dry strips of pI 4-7 were used to separate both gill and foot proteins but pI 6-11 strips were used for gill proteins only.

Gel #	Cy2	Cy3	Cy5	Gel #	Cy2	Cy3	Cy5
1	STD	C1	A1	9	STD	L4	C9
2	STD	C2	A2	10	STD	L5	C10
3	STD	C3	L1	11	STD	A6	L6
4	STD	C4	L2	12	STD	A7	L7
5	STD	C5	L3	13	STD	L8	A8
6	STD	A3	C6	14	STD	L9	A9
7	STD	A4	C7	15	STD	L10	A10
8	STD	A5	C8				

Table 4.1: Labelling strategy for the salinity DIGE experiment.
STD = Pooled internal standard (n = 30). C1-10 = Control; A1-10 = Ambient salinity (~32 ppt); L1-10 = Low salinity (~14 ppt).

4.2.5 2-DE

Approximately 10 µg of labelled protein (30 µg total) was loaded on to acidic strips (pI 4-7) using rehydration loading. Basic strips (pI 6-11) were cup-loaded. Strips were focused using an Ettan IPGphor Isoelectric system. Strips were focused using the following protocol: Step 300 V for 30 min, two Gradient steps at 1000 V and 5000 V for 30 min and 90 min respectively and a final Step of 5000 V for 25 min.

Before the second dimension, strips were equilibrated for 10 min in 2.5 mL/strip equilibration solution plus 0.25% DTT, then 10 min in 2.5mL/strip equilibration solution plus 4.5% iodoacetamide (Appendix 2). Strips were loaded

onto NuPAGE 4-12% Bis-Tris gels (Invitrogen) and were electrophoresed at 200 V, 140 mA for 1 hr.

4.2.6 Sorting gels for analysis

Only gels that produced good quality images with adequate protein separation were included in the data analysis. Of the original 45 images the number that were analysed for each tissue are indicated in Table 4.2.

	Low salinity (n = 10)	Full Salinity (n = 10)	Control (n =10)	Standard (n = 15)	Total # spot Maps (n = 45)
Gill 4-7	3	5	6	7	21
Gill 6-11	4	4	8	8	24
Foot 4-7	4	4	6	7	21

Table 4.2: Number of samples included for analysis for each tissue and pI.

Reducing the number of biological replicates used in the salinity experiment decreases the ability of DeCyder to detect significant changes. For example, power analysis shows that use of only three replicates in an experiment (biological or technical) will result in only a 65% chance of detecting a two-fold change, but by adding just one more replicate this can be increased to 95% (Karp and Lilley 2005). The effect of removing gels from the salinity analysis is clear when comparing the results of this experiment with the results of the heavy metal analysis (Chapter 5). In the heavy metal experiment many more gels were retained for analysis, leading to many more statistically significant results.

4.2.7 DeCyder 2D™ analysis

Initially, protein spot maps were analysed in the DeCyder BVA module. Protein spot expression for each dataset, (i.e. Gill 4-7, Gill 6-11 and Foot 4-7) was compared across different experimental groups. The types of interactions anticipated in this experiment are summarised in Table 4.3. Statistical comparisons between the three different experimental groups (control, ambient and low salinity) were performed as listed below (1-3) to detect proteins that varied among experimental groups.

1) Low salinity versus ambient salinity: Mussels were maintained for 3 days in either 14 ppt (Low salinity) or 32 ppt (Ambient salinity) seawater. Differences in protein abundance between these groups should be mainly due to salinity variation.

2) Low salinity versus control: Low salinity mussels were exposed to low salinity in the laboratory for 3 days. Control mussels were killed at Time 0. Therefore changes in protein abundance between these groups could be due to the effect of salinity and/or the effect of the experimental conditions.

3) Ambient salinity versus control: Control and ambient mussels were both exposed to 'normal' salinity (32 ppt). However, ambient mussels were maintained in the laboratory for 3 days, whereas control mussels were killed at Time 0. Because both groups were exposed to the same salinity, differences in protein abundance are most likely due to the effect of the experimental conditions on ambient samples.

Condition	Low salinity	Ambient	Control
Low salinity		Similarity = Experimental conditions in bucket ↑ammonia, lack of food, space etc.	Similarity = Constitutively expressed proteins that are not influenced by experimental conditions?
Ambient	Difference = Salinity Ambient vs. Low		Similarity = Salinity
Control	Differences = Salinity and experimental conditions.	Difference = Experimental conditions of bucket applied to ambient.	

Table 4.3: Potential differences and similarities in the experimental conditions experienced by each of the groups.

The statistical tools and filtering within BVA were utilised to identify proteins which showed significant differences in protein abundance between groups. Initially, proteins exhibiting changes in abundance associated with a *t*-test *p* value of < 0.05 were considered significant. However, because DIGE can detect small changes in protein abundance it is becoming standard practice to apply a *t*-test *p* value of < 0.01. Additionally, identification of proteins often focus on those proteins which exhibit relatively large differences in abundance i.e. average ratio ≥ 1.5 and $\leq$

-1.5. The results reported in this chapter therefore use these more conservative parameters ($p < 0.01$ and average ratio ≥ 1.5 and ≤ -1.5).

It is also important to note that the False Discovery Rate (FDR) correction (Benjamini and Hochberg 2000) has not been applied to the proteins reported within this chapter. When the FDR correction was applied, none of the protein variations were statistically significant using a conservative p-value ($p < 0.01$). However, 44 proteins from the Foot 4-7 dataset were significant using the FDR correction and a p-value of < 0.05, as reported in Table 4.7 Previous studies have noted apparent biological effects of salinity on HSP and GST protein abundances, as detected using antibodies (Lyons et al. 2003). It was decided to include all FDR uncorrected differences, for analysis of as many proteins as possible that may be involved in response to low salinity. The decision not to apply the FDR correction means that there is a possibility that some of the differences in protein abundance reported in this study may be due to chance. After the significant changes had been established for each dataset (i.e. Gill pI 4-7, Gill pI 6-11 and Foot pI 4-7), data for the varying proteins was imported into EDA and pattern analysis and principal component analyses were performed

4.2.8 Identification of proteins using MALDI-TOF mass fingerprinting

Protein spots were prepared for MALDI-TOF using a standard acetonitrile (ACN) and ammonium bicarbonate protocol, and were digested with trypsin at 30 °C as described in Section 3.14. MALDI spectra were collected and were analysed using Data Explorer (Applied Biosystems 5.1) and two different search algorithms; ProFound or Xproteo. Before exporting peak data to these algorithms, the peptides were calibrated using either an internal standard (trypsin peaks) or known standard spectra (Calmix, PE Biosystems, CA) that were acquired at the same time as the unknown samples.

4.3 Results

4.3.1 DeCyder results for gill pI 4-7

In gill pI 4-7 gels, 17 proteins exhibited significant ($p < 0.01$) differences in abundance between one or more experimental groups (Table 4.4). This was approximately 1% of the total number of protein spots separated in this dataset (approximately 1400 spots in total). The locations of proteins exhibiting significant differences in abundance are shown in Figure 4.1. The numbering on the spot map corresponds with the numbers shown in the 'Master No.' column in Table 4.4. In some instances DeCyder did not correctly detect spots on all gels or there were mismatches. In these cases spots were added and matching was corrected. Spots that required manual addition or other manipulations are indicated with an asterix (Table 4.4). A summary of the results presented in Table 4.4 is shown in Table 4.5. The larger number of proteins shown in the total column in Table 4.5 (n = 22) than Table 4.4 (n = 17) relates to the observation that the abundance of several proteins changed significantly in more than one experimental comparison (Figure 4.2).

Table 4.4: Gill pI 4-7 spots which showed significant ($p < 0.01$) differences in abundance
* Proteins spots were manually added or manipulated

Pos.	Master No.	Appearance	T-test	Av. Ratio	1-ANOVA
1*	1427	Low salinity vs. Control	0.0062	-2.08	0.012
2*	1728	Ambient vs. Control	0.0058	-1.60	0.013
3*	1932	Ambient vs. Control	0.00068	-1.89	0.0020
		Low salinity vs. Ambient	0.0094	2.31	
4	2264	Low salinity vs. Ambient	0.00099	2.16	0.0055
5*	2327	Low salinity vs. Control	0.0039	-2.13	0.0018
		Low salinity vs. Ambient	0.0084	-1.85	
6*	2332	Low salinity vs. Control	0.0017	2.36	0.00063
		Low salinity vs. Ambient	0.0029	2.58	
7	2354	Low salinity vs. Control	0.0040	-1.63	0.0041
		Low salinity vs. Ambient	0.0077	-1.55	
8	2373	Low salinity vs. Ambient	0.0022	1.83	0.0045
9*	2726	Low salinity vs. Control	0.0080	1.72	0.12
10*	2733	Low salinity vs. Control	0.0048	2.26	0.13
11*	2846	Low salinity vs. Ambient	0.00071	2.93	0.0070
12*	2993	Ambient vs. Control	0.0072	1.99	0.022
13	3069	Ambient vs. Control	0.0034	1.59	0.00077
		Low salinity vs. Ambient	0.00062	-2.19	
14*	3418	Low salinity vs. Control	0.0019	1.62	0.0068
15	3429	Low salinity vs. Control	0.0016	2.79	0.13
16	3478	Low salinity vs. Control	0.0092	-1.50	0.0040
17*	3572	Ambient vs. Control	0.0067	1.52	0.013

Group Comparison	# spots increased ↑	# spots decreased ↓	TOTAL (n = approx. 1400 spots)	
Average Ratio	**> 1.5**	**< -1.5**	**Total**	
Low salinity vs. Ambient	5	3	8 (0.6%)	62.5% ↑ 37.5% ↓
Low salinity vs. Control	5	4	9 (0.6%)	55% ↑ 45% ↓
Ambient vs. Control	2	3	5 (0.4%)	40% ↑ 60% ↓
TOTAL	**12**	**10**	**22**	

Table 4.5: Gill pI 4-7 summary of significant ($p < 0.01$) results

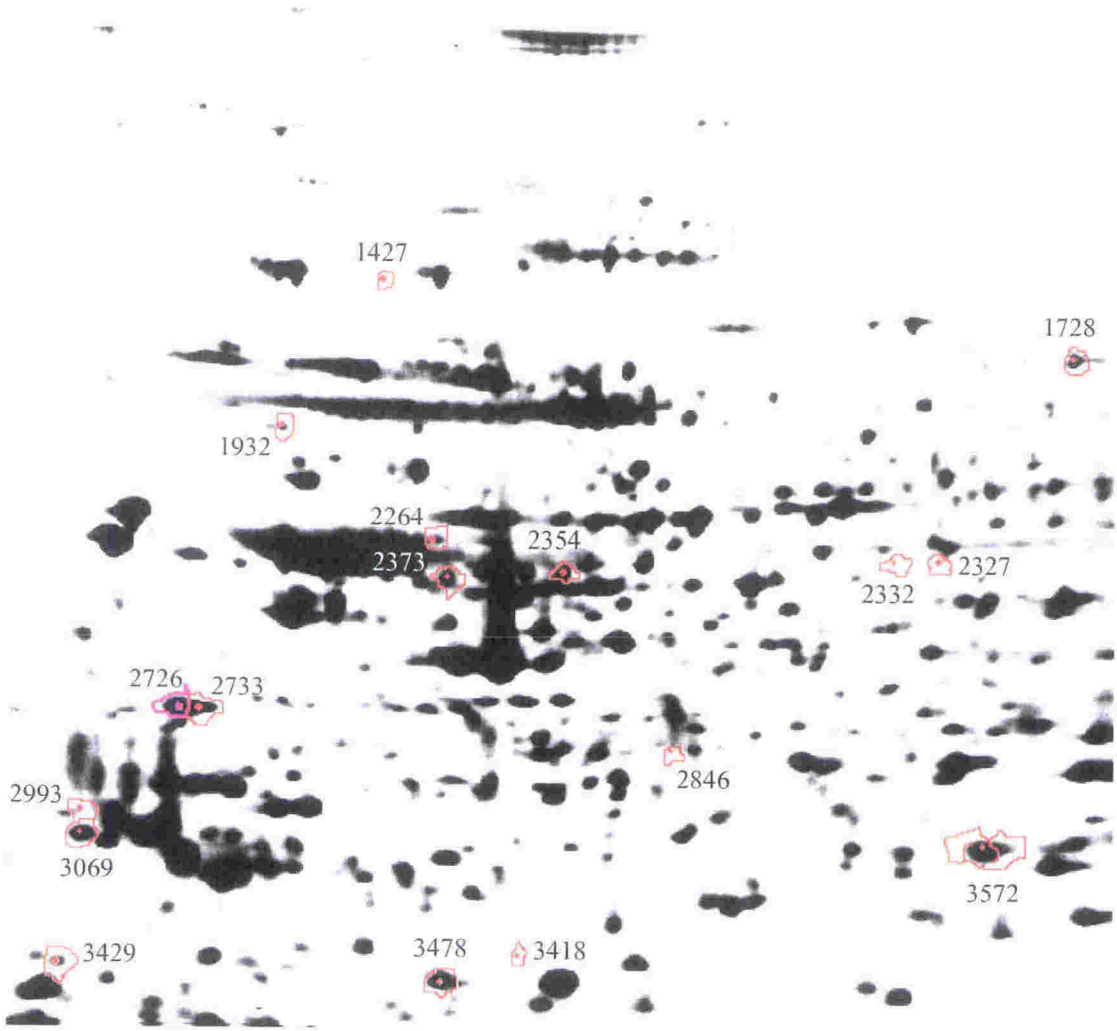


Figure 4.1: Gill pI 4-7 gel indicating spots which varied significantly among groups ($p < 0.01$)
The spot outlined in pink (2726) was identified as tropomyosin.

Some variation in protein abundance was significant in more than one comparison between different groups (Figure 4.2). For example, two proteins varied significantly in comparisons between low salinity versus ambient and ambient versus control groups (Figure 4.2). Low salinity and ambient groups were exposed to the 'same' experimental conditions in buckets for 3 days with salinity being the major

difference between the two groups (14 ppt and 32 ppt respectively). Therefore the main differences in protein expression in comparisons between these groups should be due to the effects of salinity. Comparing low salinity and ambient protein abundance revealed that there were 3 proteins which were likely to be important for dealing with low salinity (Figure 4.2). The remaining 5 proteins that were significant in low salinity and ambient comparisons were also significant in comparisons between other groups as illustrated by the intersecting circles. These proteins may give additional insight into the effects that the experimental conditions impose on mussel physiology.

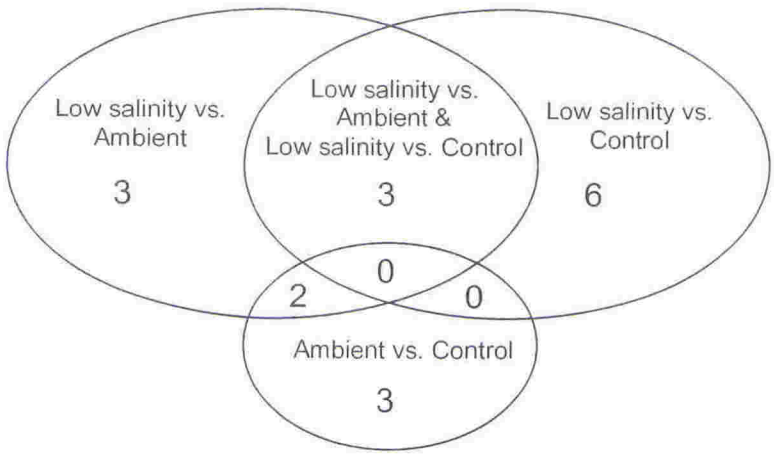


Figure 4.2: Venn diagram illustrating the distribution of significant protein spots and their overlap between treatments in gill tissue pI 4-7

Proteins which showed significant ($p < 0.01$) differences in abundance on gill pI 4-7 gels were used to generate a pattern analysis to visualise general relationships between experimental groups (Figure 4.3). Pattern analysis split the spot maps into 4 groups. One cluster predominantly contained ambient gels and two clusters exclusively contained either control or low salinity samples. One ambient gel, to the left of the yellow line, appears to be an outlier to the rest of the gels. Overall, the pattern analysis clustered gels into their expected groups.

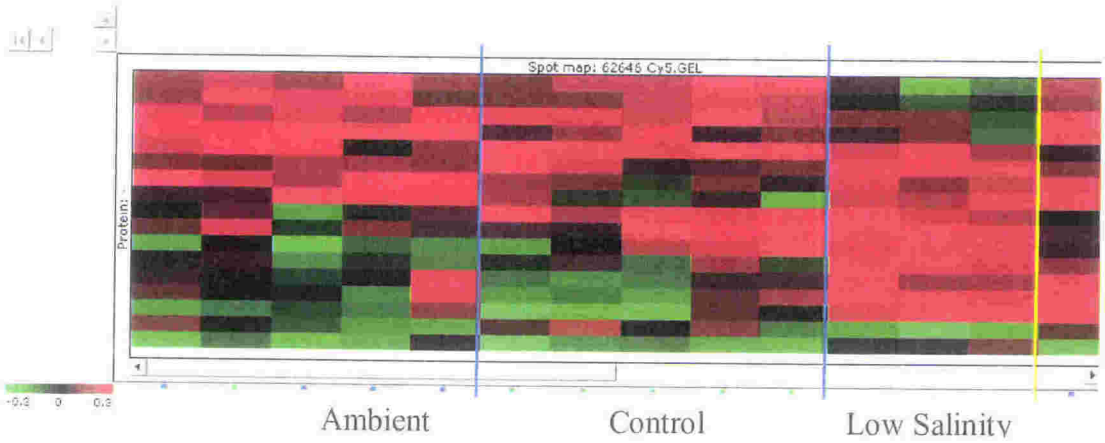


Figure 4.3: Gill pI 4-7 pattern analysis generated using proteins which showed significant ($p < 0.01$) differences in abundance between experimental groups

Key:

- Control
- Ambient
- Low salinity

Each dot represents an individual gel as indicated by the key. The yellow and blue lines have been added.

Proteins which showed significant ($p < 0.01$) differences in abundance among groups on gill pI 4-7 gels were exported to EDA and analysed using PCA (Figure 4.4). Control, ambient and low salinity spot maps cluster into 3 separate distinct groups with the exception of one ambient gel which is indicated by a red arrow (Figure 4.4, panel C). Control and ambient gels are both on the same half of the ellipse in Figure 4.4 (panel C) indicating that these proteins share similar patterns of protein abundance. This is also illustrated in the 'heat map' because the colour profiles for proteins from ambient and control spot maps are similar (Figure 4.3).

In panel A, PC1 describes 34.8% of the total protein variation in response to experimental conditions and PC2 an additional 22.4%, a total of 57.2%. Cumulative variance >90% is reached after 5 principal components (PCs). In panel C, PC1 describes 47.7% of the total spot map variation in response to experimental conditions, and PC2 describes an additional 23.7%, giving a cumulative total of 71.4%. Cumulative variance >90% is reached after 5 PCs. Panel B and D are 'loading plots' and are linked to the expression of proteins in the associated score plot as previously described (Chapter 3).

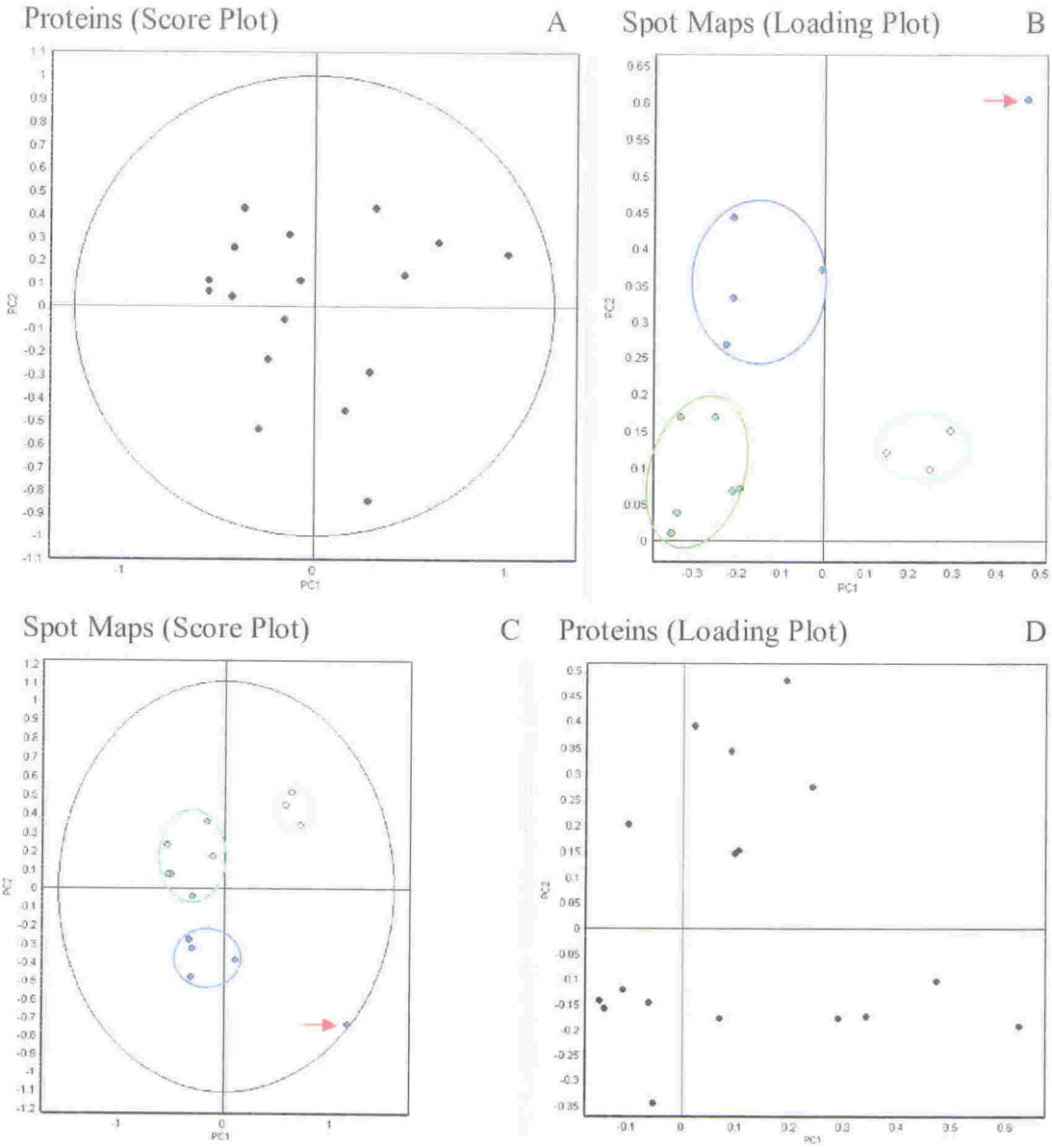


Figure 4.4: Gill pI 4-7 PCA using significant ($p < 0.01$) proteins

Key:

- Control
- Ambient
- Low salinity

The black dots in panels A and D represent individual proteins (n = 17), whereas the coloured dots in panels B and C represent spot maps (n = 14). Red arrows indicate outliers.

4.3.2 DeCyder results for gill pI 6-11

In gill pI 6-11 gels, only 2 proteins exhibited significant ($p < 0.01$) differences in protein abundance between one or more experimental groups (Table 4.6). This is approximately 0.25% of the total number of protein spots separated in this dataset (approximately 800 spots in total). The locations of proteins exhibiting significant differences in protein abundance are shown in Figure 4.5. Due to the low

number of significant proteins detected, pattern analysis and PCA were not performed on these proteins.

Pos.	Master No.		T-test	Av. Ratio	1-ANOVA
1*	2113	Low salinity vs. Control	0.0016	-2.17	0.0011
2	2658	Low salinity vs. Control	0.0022	2.65	0.0029
		Ambient vs. Control	0.031	1.86	

Table 4.6: Gill pI 6-11 table of significant ($p < 0.01$) proteins
* Proteins spot was manually added or manipulated

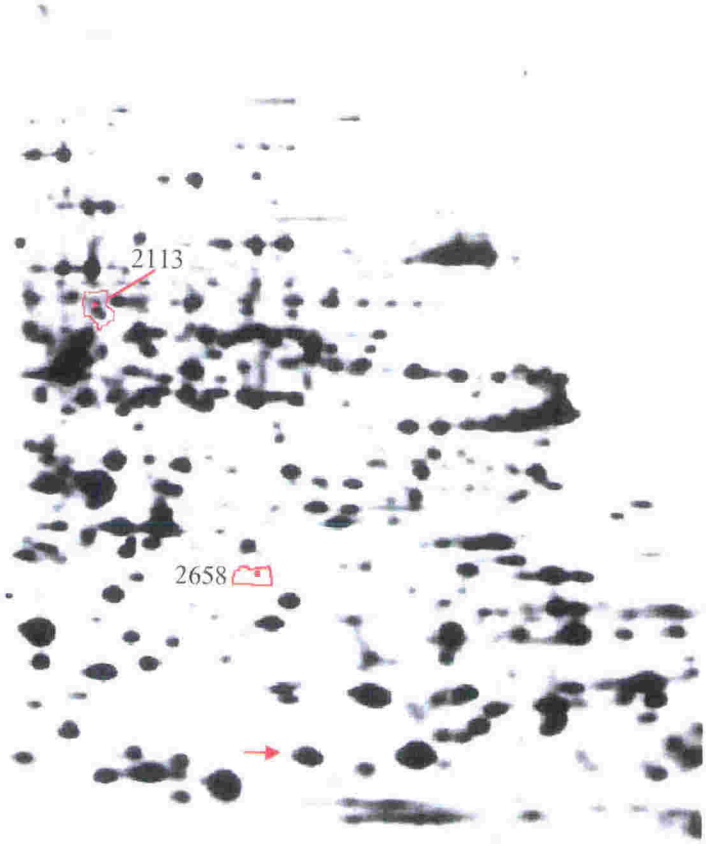


Figure 4.5: Gill 6-11 gel indicating spots which varied significantly ($p < 0.01$) among groups
The red arrow indicates a spot that was identified as actin, however this spot did not show significant changes in protein abundance between groups in the current experiment

4.3.3 DeCyder results for foot tissue, pI 4-7

In foot pI 4-7 gels, 96 proteins exhibited significant ($p < 0.01$) differences in protein abundance between one or more experimental groups (Table 4.7). This is approximately 10% of the total number of protein spots separated in this dataset (approximately 900 spots in total). The locations of proteins exhibiting significant

differences in protein abundance are shown in Figure 4.6, the numbering matches 'Master No' as shown in Table 4.7. A summary of the results are shown in Table 4.8. A much larger proportion of the total number of proteins show significant variation in expression in foot tissue (10.6%) compared to the gill datasets (1%). A larger number of proteins are shown in the total column in Table 4.8 (n = 99) than Table 4.7 (n = 96) because several proteins occur in more than one experimental comparison (Figure 4.7).

Table 4.7: Foot pI 4-7 table of significant ($p < 0.01$) proteins
* Proteins spots that were manually added or manipulated

Pos.	Master No.		T-test	Av. Ratio	1-ANOVA	FDR
1	265	Ambient vs. Control	1.2e-005	-1.53	0.018	0.014
2	518	Ambient vs. Control	0.0027	-2.18	0.0034	
3	626	Ambient vs. Control	0.0058	-2.85	0.0080	
4	639	Ambient vs. Control	0.0022	-3.32	0.0067	
5	641	Ambient vs. Control	0.0048	-1.75	0.058	
6*	644	Ambient vs. Control	0.0034	-1.83	0.011	
7	662	Ambient vs. Control	0.0078	-1.66	0.028	
8*	687	Ambient vs. Control	1.1e-005	-2.23	0.00067	0.014
9	692	Ambient vs. Control	0.0037	-2.74	0.021	
10*	724	Ambient vs. Control	0.0012	-1.99	0.0026	0.041
11	759	Ambient vs. Control	0.0030	-2.21	0.021	
12	798	Ambient vs. Control	0.0016	-1.82	0.0063	0.047
13	809	Ambient vs. Control	0.0081	-1.86	0.025	
14	814	Ambient vs. Control	0.0033	-2.39	0.0070	
15	820	Ambient vs. Control	0.00019	-3.04	0.0072	0.023
16	834	Low salinity vs. Ambient	0.0081	-1.75	0.042	
17*	859	Ambient vs. Control	0.00020	-3.41	0.00066	0.023
18	860	Ambient vs. Control	0.00024	-2.91	0.0022	0.023
19*	865	Ambient vs. Control	0.00030	-2.22	0.029	0.025
20*	875	Ambient vs. Control	0.0024	-2.04	0.029	
21	879	Ambient vs. Control	0.0017	-3.92	0.0050	0.048
22*	905	Ambient vs. Control	0.0012	-2.47	0.0057	0.041
23	930	Ambient vs. Control	0.00018	-2.76	0.011	0.023
24	960	Ambient vs. Control	0.0021	-2.17	0.032	
25	967	Ambient vs. Control	0.0013	-2.61	0.017	0.041
26	968	Ambient vs. Control	0.0021	-1.93	0.0069	
27	971	Ambient vs. Control	0.0033	-1.52	0.011	
28	976	Ambient vs. Control	0.0020	-2.75	0.027	
29	984	Ambient vs. Control	0.00074	-1.62	0.029	0.039
30	988	Ambient vs. Control	0.0091	-2.49	0.020	
31	995	Ambient vs. Control	0.00038	-2.08	0.0040	0.028
32	1003	Ambient vs. Control	0.0084	-1.63	0.10	
33*	1010	Ambient vs. Control	0.00024	-2.46	0.0095	0.023
34*	1015	Ambient vs. Control	0.00013	-3.24	0.0073	0.023
35	1020	Ambient vs. Control	0.0027	-2.41	0.11	
36	1022	Ambient vs. Control	0.0011	-1.61	0.022	0.041
37	1024	Ambient vs. Control	0.0029	-1.68	0.042	

38	1060	Ambient vs. Control	0.00087	-1.95	0.0017	0.041
39	1064	Ambient vs. Control	0.0061	-2.23	0.045	
40	1069	Ambient vs. Control	0.0040	-1.87	0.0081	
41*	1079	Ambient vs. Control	0.0037	-2.19	0.026	
42	1083	Ambient vs. Control	0.0039	-2.75	0.033	
43	1089	Ambient vs. Control	0.0028	-2.00	0.032	
44	1096	Ambient vs. Control	0.00062	-2.40	0.0032	0.03
45	1098	Ambient vs. Control	0.0012	-2.36	0.0099	0.041
46*	1102	Ambient vs. Control	0.00045	-2.15	0.028	0.030
47	1103	Ambient vs. Control	0.00034	-3.24	0.0094	0.026
48	1116	Ambient vs. Control	0.0025	-1.54	0.083	
49	1118	Ambient vs. Control	0.0027	-2.39	0.015	
50	1119	Ambient vs. Control	0.0098	-1.92	0.17	
51	1124	Ambient vs. Control	0.0035	-2.61	0.0038	
52	1125	Ambient vs. Control	0.0036	-2.89	0.014	
53	1129	Ambient vs. Control	0.0061	-1.63	0.016	
54	1131	Ambient vs. Control	0.0097	-2.26	0.022	
55	1135	Ambient vs. Control	0.0029	-3.34	0.013	
56	1139	Ambient vs. Control	0.0022	-3.09	0.0079	
57*	1164	Ambient vs. Control	0.00068	-3.12	0.0053	0.038
58	1168	Ambient vs. Control	0.0028	-2.22	0.0028	
		Low salinity vs. Ambient	0.0084	-1.74		
59	1169	Ambient vs. Control	0.0036	-2.20	0.026	
60	1177	Ambient vs. Control	0.00021	-1.86	0.00014	0.023
		Low salinity vs. Ambient	0.00077	-1.70		
61	1181	Ambient vs. Control	0.0042	-1.87	0.036	
62	1185	Ambient vs. Control	0.0015	-1.61	0.0076	0.044
63	1196	Ambient vs. Control	0.0095	-2.33	0.016	
64*	1206	Ambient vs. Control	0.0057	-1.76	0.0055	
65	1207	Ambient vs. Control	0.0010	-2.21	0.0072	0.041
66	1216	Ambient vs. Control	0.0016	-2.14	0.0042	0.047
67	1229	Ambient vs. Control	0.0011	-1.73	0.043	0.041
68	1259	Ambient vs. Control	0.0010	-2.74	0.025	0.041
69	1263	Ambient vs. Control	0.0082	-1.80	0.025	
70	1265	Ambient vs. Control	0.0011	-1.79	0.0039	0.041
71	1292	Ambient vs. Control	0.0047	-2.15	0.030	
72*	1300	Ambient vs. Control	0.0012	-3.52	0.0063	0.041
73	1301	Ambient vs. Control	0.0032	-2.22	0.018	
74	1321	Ambient vs. Control	0.0018	-1.81	0.0010	0.049
75	1346	Ambient vs. Control	0.00095	-3.39	0.0033	0.041
76	1357	Ambient vs. Control	0.0013	-3.43	0.012	0.041
77	1363	Ambient vs. Control	0.0022	-3.21	0.0087	
78	1427	Ambient vs. Control	0.00030	-1.89	0.00015	0.025
		Low salinity vs. Ambient	0.0038	-1.63		
79	1473	Ambient vs. Control	5.1e-005	-2.11	0.0022	0.023
80	1474	Ambient vs. Control	0.0079	-1.94	0.042	
81	1486	Ambient vs. Control	0.0071	-2.04	0.054	
82	1499	Ambient vs. Control	0.00089	-7.63	0.013	0.041
83	1537	Ambient vs. Control	0.00025	-16.74	0.0092	0.023
84	1570	Ambient vs. Control	0.0010	-2.93	0.0084	0.041
85	1591	Ambient vs. Control	0.0038	-7.34	0.0071	
86	1595	Ambient vs. Control	0.00049	-1.56	0.00022	0.030
87	1672	Ambient vs. Control	0.0021	-1.88	0.014	
88	1678	Ambient vs. Control	0.0034	-1.76	0.0081	
89	1780	Ambient vs. Control	0.00089	-2.05	0.0033	0.041
90	1829	Low salinity vs. Control	0.0069	-2.83	0.018	
91	1852	Low salinity vs. Control	0.0012	1.55	0.043	
92	2018	Ambient vs. Control	0.0093	-1.74	0.017	

93	2474	Ambient vs. Control	0.00024	-2.61	0.0013	0.023
94	2475	Ambient vs. Control	0.00013	-2.83	0.013	0.023
95	2477	Ambient vs. Control	0.00049	-3.20	0.0053	0.030
96*	2478	Ambient vs. Control	0.0078	-3.50	0.0074	

	# spots increased ↑	# spots decreased ↓	TOTAL (/ ~900 spots overall)	
Average ratio	**>1.5**	**<-1.5**	**Total**	
Low salinity vs. Ambient	4	0	4 (0.4%)	100% ↑
Low salinity vs. Control	1	1	2 (0.2%)	50% ↑ 50% ↓
Ambient vs. Control	0	93	93 (10.3%)	100% ↑
TOTAL	**5**	**94**	**99**	

Table 4.8: Foot pI 4-7 summary of significant ($p < 0.01$) results

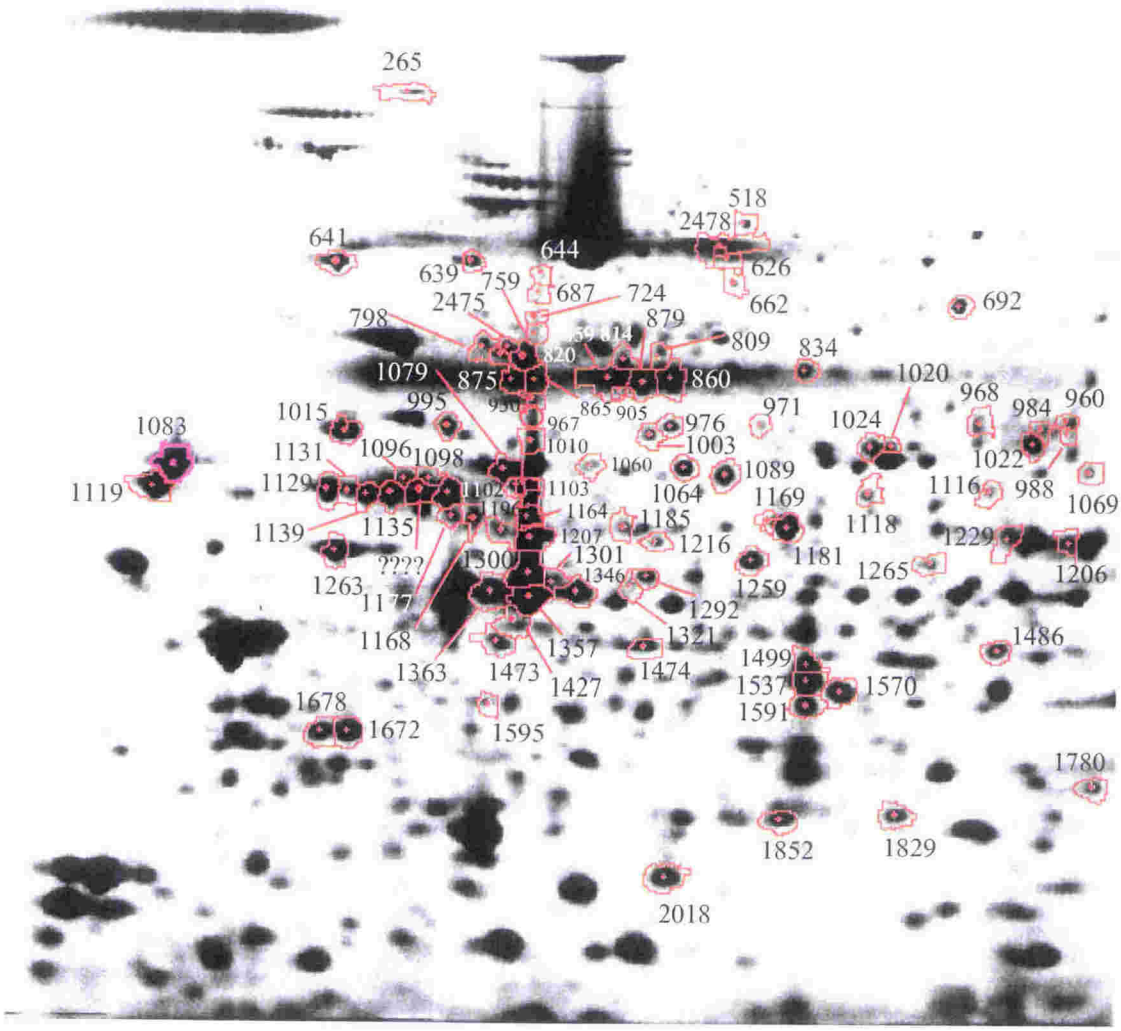


Figure 4.6: Foot pI 4-7 gel indicating spots which varied significantly ($p < 0.01$) among groups
The spot outlined in pink (1088) was identified as tropomyosin. Spot 1357 was identified as actin

Some changes in protein abundance were significant in more than one comparison between different groups (Figure 4.7). For example, 3 proteins show significant differences in abundance in comparisons between low salinity versus ambient and ambient versus control groups. The large number of proteins detected in the ambient versus control comparisons (93) was unexpected. Results indicate that experimental conditions may have affected ambient foot tissue to a greater extent than gill tissue.

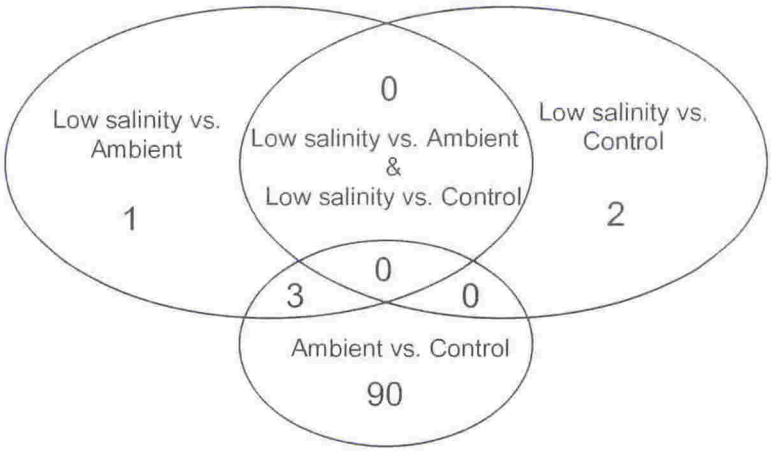


Figure 4.7: Venn diagram illustrating the distribution of significant ($p < 0.01$) protein spots and their overlap between treatments in foot tissue pI 4-7

Proteins which differed significantly ($p < 0.01$) between groups in foot tissue (pI 4-7) were used to generate a pattern analysis to visualise general relationships between different experimental groups (Figure 4.8). The results were unexpected and totally different from those results obtained for gill tissue. The most striking aspect was the clear difference in protein abundance between ambient and control groups which separated into two major clusters on either side of the yellow line. Protein expression was high in control samples and low in ambient samples as illustrated by red and green 'heat' patterns respectively (Figure 4.8). Control samples were split into two clusters on either site of the blue line. One cluster contains most of the control samples whereas the other contains a low salinity gel and a control gel. Low salinity mussels were split into each of these two main clusters (either side of the yellow line). Three low salinity samples had protein abundance profiles similar to control samples, whereas 1 sample showed a protein abundance profile similar to ambient samples.

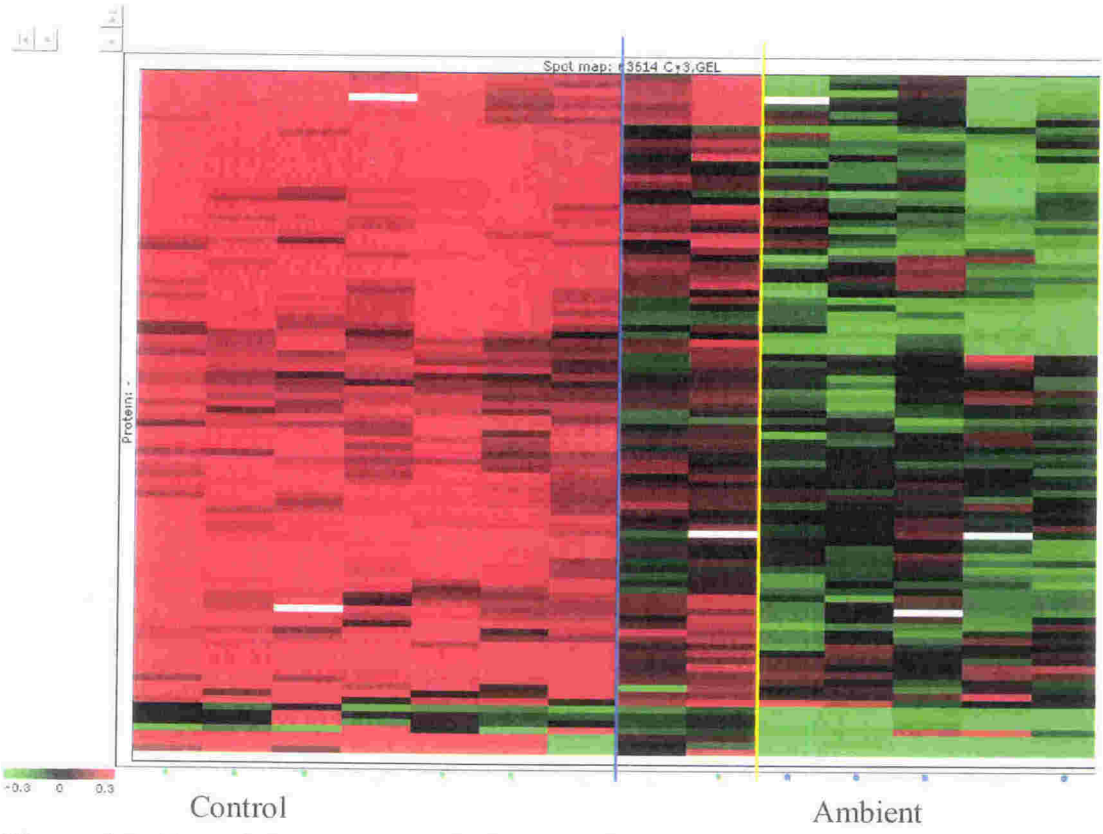


Figure 4.8: Foot pI 4-7 pattern analysis using significant ($p < 0.01$) proteins (n = 96)

Key:
- Control
- Ambient
- Low salinity

Each dot represents an individual gel as indicated by the key. The yellow and blue lines have been added.

The proteins which changed abundance significantly ($p < 0.01$) in foot tissue (pI 4-7) were analysed by PCA (Figure 4.9). Control and ambient spot maps clearly separate from each other (panel C), whereas low salinity spot maps overlap between control and ambient gels. One low salinity gel, indicated by the red arrow, is an outlier. PCA protein analysis in panel A revealed 7 protein outliers outside of the 95% confidence interval bounded by the black ellipse. Most of these protein spots showed strong differential expression e.g. protein numbers (average ratio value): 1321 (-1.81), 1829 (2.83), 1363 (-3.21), 1346 (-3.39), 1591 (-7.94), 1499 (-7.63), 1537 (-16.74). In the spot maps score plot (panel A) PC1 describes 36.9% of the variance, and PC2 describes a further 21.9 giving a cumulative total of 58.8%. Cumulative variance >90% is not reached after 5 PCs. In panel C, PC1 describes 76.1% of the total spot map variation in response to environmental conditions, and PC2 describes an additional 6.4%, giving a cumulative total of 82.5%. Cumulative

variance >90% is reached after 4 PCs, indicating that the PCA is a good representation of the raw data.

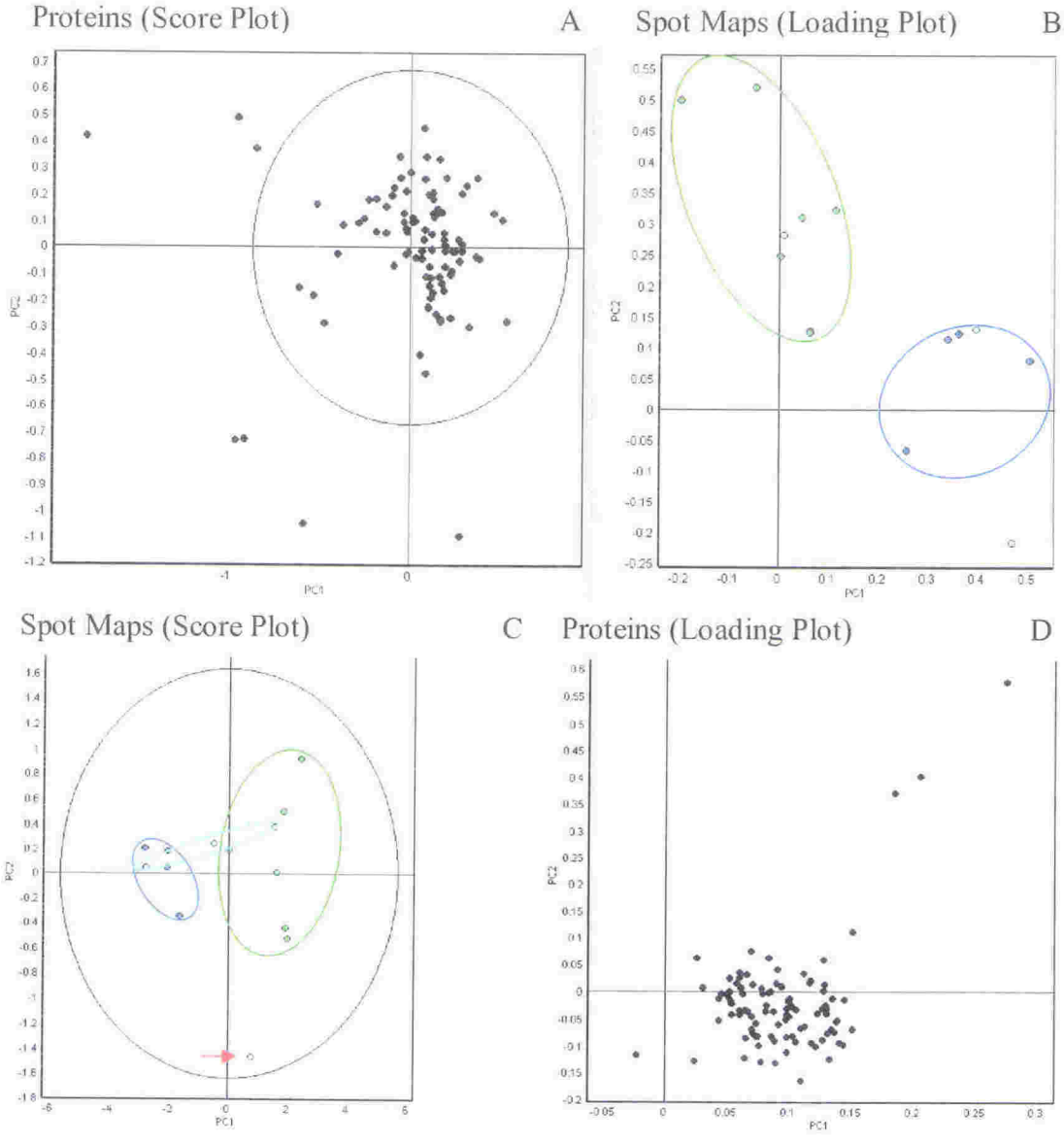


Figure 4.9: Foot pI 4-7 PCA using significant ($p < 0.01$) proteins

Key:

- Control
- Ambient
- Low salinity

The black dots in panels A and D represent individual proteins (n = 96), whereas the coloured dots in panels B and C represent spot maps (n = 14). The red arrow indicates an outlier.

4.3.4 Identified proteins

Despite obtaining good MALDI-TOF spectra (e.g. Figure 4.10) very few *P. canaliculus* proteins were positively identified. However, two different tropomyosin

and actin isoforms (four proteins in total) were able to be identified using MALDI mass fingerprinting and LTQ-MS/MS. The MALDI peptide mass fingerprint for one tropomyosin protein is shown in Figure 4.10 and the Profound results are shown in Figures 4.11 and 4.12. This match for tropomyosin was highly significant and the coverage and error rate were also very good. The identification of the other tropomyosin protein was obtained from a commercial protein sequencing service (Australian Proteome Analysis Facility; APAF, results not shown). The greatest rank and probability scores for both tropomyosin isoforms were from the tropical green mussel species *P. viridis*. In contrast, the two different actin isoforms identified had significant homology with actin isolated from either rabbit (*Oryctolagus cuniculus*) or duck (*Anas platyrhynchos*). However, the actin protein which had homology to duck actin was small (approximately 10 kDa) and basic (approximately pH 8) and did not show significant changes in abundance in the current study.

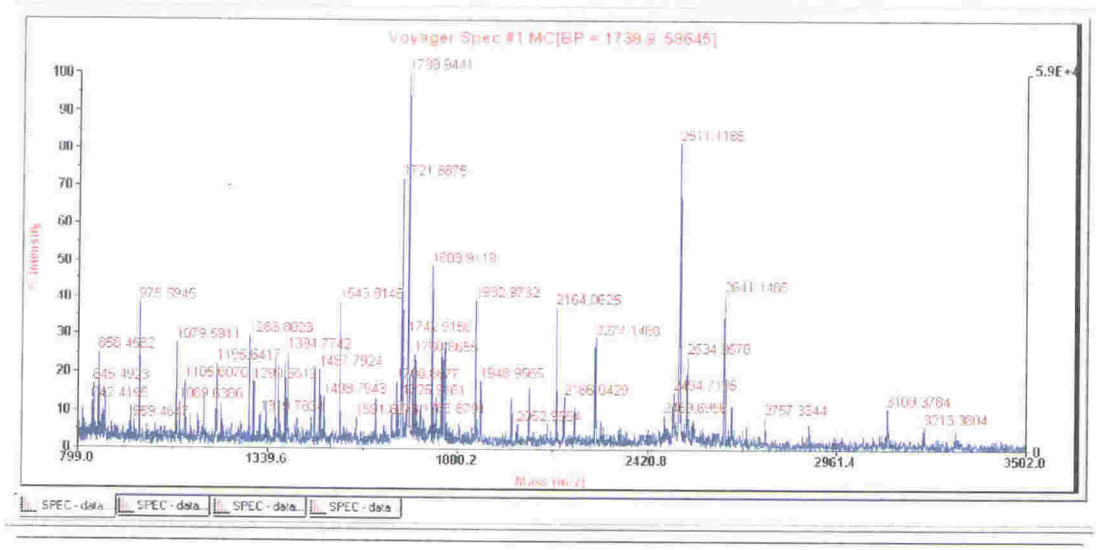


Figure 4.10: Protein mass fingerprint profile from *P. canaliculus*, identified as tropomyosin

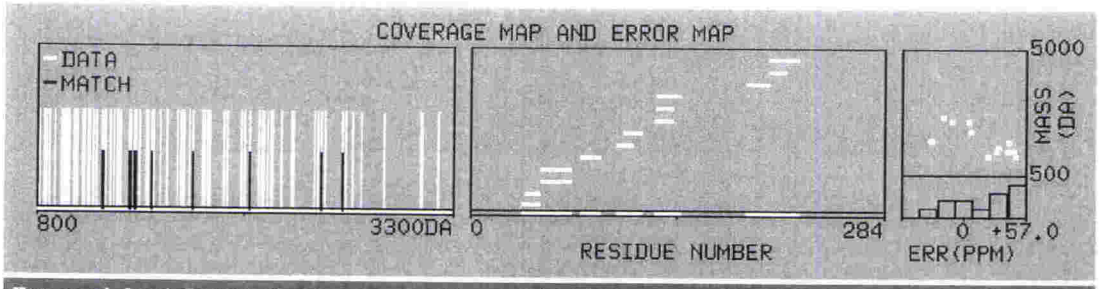
ProFound - Search Result Summary

Version 4.10.6
The Rockefeller University Edition

Protein Candidates for search 86DA3A98-08AC-7680A32F [549553 sequences searched]

Rank	Probability	Est'd Z	Protein Information and Sequence Analyse Tools (T)	%	pI	kDa	®
1	1.0e+000	2.13	T gi|42559692|sp|Q9GZ70|TPM_PERVI Tropomyosin	43	4.7	32.79	®
2	6.1e-008	0.12	T gi|3341980|gb|AAC27501.1| huntingtin-interacting protein HYPA/FBP11 [Homo sapiens]	27	5.8	45.58	®
3	5.3e-011	-	T gi|38087178|ref|XP_357774.1| similar to SOX6 [Mus musculus]	42	6.9	30.99	®
4	3.8e-011	-	T gi|13435135|ref|NP_072095.1| testis specific protein, Y-linked 2 [Homo sapiens]	26	5.3	35.40	®

Figure 4.11: Profound search result for *P. canaliculus*



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (ppm)	Residues Start	To	Missed Cut	Peptide sequence
1194.634	M	1194.576	48	39	48	0	IEDDYNSLVK
1195.611	M	1195.583	24	102	112	1	AEERFTTASGK
1358.697	M	1358.642	41	129	140	0	VLENLNSGNDER
1367.736	M	1367.693	32	106	118	1	FTTASGKLEEASK
1386.748	M	1386.684	46	206	217	0	TLQVQNDQASQR

Figure 4.12: Profound coverage and error map for tropomyosin isolated from *P. canaliculus*

4.4 Discussion

4.4.1 Proteins involved in response to low salinity in *P. canaliculus*

This experiment detected *P. canaliculus* proteins that responded to changes in salinity. Summarising the results from both pI ranges (4-7 and 6-11) and both tissues (gill and foot) approximately 115 proteins showed significant ($p < 0.01$) differences in protein abundance. However, only four of these proteins were identified, two isoforms of actin and two isoforms of tropomyosin.

Actin

Actin plays an important role in cell structure, cell motility and the generation of contractile force in both muscle and non-muscle cells (Rubenstein 1990). Multiple isoforms with diverse although very similar amino acid sequences have been isolated and are expressed in a conserved tissue specific fashion across many species (Rubenstein 1990). Actin is well represented in the NCBI databases, with 56687 nucleotide sequences, and 14884 protein sequences. To date, only 3 actin protein and DNA sequences from mussels (*M. edulis, M. galloprovincialis* and *D. polymorpha*) have been deposited in these databases. Interestingly, one of the actin isoforms isolated from *P. canaliculus* (Figure 4.1) that was identified using in-house tandem mass spectrometry matched with beta gamma non-muscle actin from rabbit (*Oryctolagus cuniculus*). The actin tandem mass spectra reported by APAF matched with beta-actin from duck (*Anas platyrhynchos*). Matches with unrelated species

demonstrate the diversity of actin sequences and the utility of having a large database to seek matches from.

Tropomyosin

Tropomyosin binds to actin and also helps to regulate the interaction of myosin with actin filaments (Patwary et al. 1990). Multiple tropomyosin isoforms are generally present in invertebrates and most of the work to date has focused on distinguishing muscle and non-muscle forms, rather than investigating hetero- and homo-dimer formation (Hooper and Thuma 2005). There are a relatively large number of tropomyosin DNA and protein sequences in the NCBI databases, 5298 and 1321 respectively. To date, tropomyosin has been identified in 3 mussel species (*M. edulis, M. galloprovincialis* and *P. viridis*). There is evidence for the presence of at least 2 tropomyosin isoforms in *P. canaliculus*, both of which give significant matches to tropomyosin from *P. viridis*. The location of tropomyosin (1083) on 2D gels from *P. canaliculus* foot tissue (Figure 4.6) is consistent with the spot location obtained from blue mussel (*M. edulis*) foot tissue (López et al. 2002a). However, the tropomyosin (2726) spot identified from *P. canaliculus* gill tissue (Figure 4.1) is in a different location (Figure 4.6), and appears to have a slightly different protein mass fingerprint when compared to the tropomyosin isolated from foot tissue (results not shown). Additionally, while the expression of tropomyosin in foot tissue is reduced in ambient *P. canaliculus* mussels compared to controls, in gill tissue the alternate isoform of tropomyosin is upregulated in low salinity mussels compared with controls. It is probable that these two tropomyosin proteins represent examples of tissue-specific isoforms, however further work is required to confirm this.

Tropomyosin and actin have also been identified in a number of previous studies and are thought to be associated with various 'stress' situations. For example, these proteins have been identified as indicators of oxidative stress in Venus striped clams (*Chamaelea gallina*) exposed to Aroclor 1254, tributyltin, As(III), Cu(II) and Cd (Rodríguez-Ortega et al. 2003a) and comparison between *M. edulis* and *M. galloprovincialis* foot proteins showed significant differences in abundance of actin and tropomyosin (López et al. 2002b). Among the proteins differently expressed by *M. edulis* exposed to crude oil five were identified, including actin and tropomyosin, both of which were downregulated (Manduzio et al. 2005).

4.4.2 Factors contributing to the low number of proteins identified in *P. canaliculus*

In the current study, out of approximately 1400 proteins isolated from *P. canaliculus* gill tissue (pI 4-7), 17 proteins were detected with significant ($p < 0.01$) differences in abundance. Of these proteins, only one (tropomyosin, 2726) was identified, a success rate of 5.9%. Out of approximately 900 proteins separated from *P. canaliculus* foot tissue, 96 proteins showed significant changes in protein abundance, of which only two (2.1%) could be identified. Some previous studies have reported a greater success rate for mussel protein identification than reported here. For comparison, a proteomic approach to study the differences between *M. edulis* and *M. galloprovincialis* separated approximately 1200 proteins spots of which 420 were chosen for comparison, significant differences (Mann-Whitney U-test, $p < 0.01$) in abundance were detected in 37 proteins (8.8%) and 15 or these proteins were able to be positively identified (41%) using a mixture of MALDI and MS/MS approaches (López et al. 2002b). The identified proteins were grouped into four broad functional classes: cytoskeletal and myofibrillar proteins, proteins associated with stress response, proteins associated with the storage or production of energy and proteins related to rearrangement in the synthesis of native structures (López et al. 2002b). In the current study, only cytoskeletal proteins were identified. Lopez et al. (2002b) note that identification of proteins from the *Mytilus* genus is a complicated process but their results were achieved using an identical approach to the current study (peptide mass fingerprinting and/or MS/MS) and therefore more protein identifications may have been expected in the current study than were achieved. Some of the factors that may have contributed to the lower than expected number of protein identifications in the current study are summarised below.

Summary of factors that may have contributed to the low number of identified proteins in the current study:

- Cross contamination (multiple proteins in a picked spot)
- Insufficient material for analysis
- Inconsistency in protein abundance between DIGE and Colloidal Coomassie (250-G) stained gels leading to errors in spot picking
- 'best' i.e. most abundant spots may not have been selected for identification
- Low number of proteins for comparison in invertebrate databases

Most of these factors are self-explanatory, however the most likely explanation for the low number of proteins identified in the current study is discussed in more detail (Section 4.4.2.1)

4.4.2.1 Low number of proteins for comparison in invertebrate databases

Often, good MALDI spectra from moderately abundant proteins spots were unable to be identified. Because protein sequences need to contain a large amount of similarity to produce significant matches (Johnson et al. 2005), it is difficult to identify invertebrate proteins using MALDI protein mass fingerprinting, the method utilised for the majority of this work. Although previous studies using similar techniques with blue mussels (*Mytilus* sp.) have managed to identify a relatively large number of proteins (Section 4.4.2) most of these experiments used protein mass fingerprinting supplemented with protein sequencing (MS/MS). Some of the samples in the current study were sent to APAF for commercial protein sequencing. Two samples were able to be identified using an Applied Biosystems 4700 in TOF/TOF at APAF but this method did not identify any proteins that had not already been identified using MALDI mass fingerprinting

Because *M. edulis* is found world-wide it is possible that there is more protein and DNA information available on this species. A survey of the NCBI databases recovers a total of 889 nucleotide and 780 protein sequences for "blue mussels". A search for "green mussels" returns 160 nucleotide sequences (many of which are microsatellites) and only approximately 30 proteins (database search April 2006). Due to the large amount of polymorphism in mussel species, it is possible that unless proteins are almost identical in *P. canaliculus* a successful identification will not be made. The identification of tropomyosin was made based on *P. viridis* sequence, a closely related tropical green mussel species. It is interesting to note that the proteins identified in the current study are both highly abundant and relatively prevalent in the mollusc databases, both factors which probably increase their chances of identification in this and previous studies.

4.4.2.2 Further directions

As databases become more complete it is hoped that further protein identifications will be made. The recent addition to the School of Biological Sciences of a tandem mass spectrometer (ThemoFinnigan LTQ) capable of protein

sequencing in MS/MS mode, lends itself to the possibility of using homology based searching to match closely related protein sequences with other species.

4.4.3 Protein expression signatures

Protein expression signatures for low salinity have been reported in the endemic New Zealand green mussel *P. canaliculus* for two different tissues. In gill tissue, at both pH ranges, approximately 1% of the proteins analysed showed differential expression between low salinity, ambient and control groups. Unexpectedly, in foot tissue a much larger number proportion (approximately 10%) of proteins displayed differential expression and most of these differences were obtained in ambient versus control comparisons. The experimental conditions encountered by the ambient group appear to have had a pronounced effect on ambient foot tissue while not affecting gill tissue to the same extent. Factors that could have affected conditions during the experiment and analysis are discussed in Section 4.4.4.

There is evidence to suggest that there are appreciable differences in protein abundance between experimental groups in the current study. However, protein expression profiles (PEPs) cannot differentiate between these groups. Although there are differences in relative protein abundances, spots that are universally present in certain groups and absent in others have not been identified. This effect may be an artefact of the DeCyder analysis process. For example if a protein was not able to be detected in both groups that were being compared, significance statistics could not be generated or if they are generated the *t*-test value may not be significant and these proteins would be excluded from further investigation. Alternatively, the lack of specific protein expression profiles could be an authentic outcome. McDonagh et al. (2005) found that protein expression profiles were not significantly different from polluted and control sites. Carbonylation was significantly higher in animals from polluted sites (McDonagh et al. 2005), but this was not investigated in the current study.

4.4.4 Factors influencing experimental outcomes

4.4.4.1 Experimental conditions

The large number of differences in protein abundance detected between control and ambient groups was unexpected. There are a range of factors which may have influenced experimental conditions leading to the disparities in protein abundance observed among groups. For example, temperature, salinity, water mixing and oxygen and ammonia concentrations. Temperature and water mixing were similar for both experimental groups and probably did not influence results. Salinity was reduced in the low salinity tank because it was the experimental variable under investigation. However, reduced salinity could also have affected oxygen and ammonia concentrations. Dissolved oxygen concentrations in ambient seawater are intrinsically lower than low salinity seawater (http://www.env.gov.bx.ca/wat/wq/BC guidelines/do/do-04.htm#tvk13) and ammonia excretion is lower at low salinities (Masilamoni et al. 2001). Previous studies have demonstrated that mussels can survive unfavourable conditions by switching to anaerobic metabolism, closing valves and depressing their basal metabolism (Shick and Widdows 1981; Sundaram and Shaffee 1989). The majority of the proteins detected as showing significant variation between ambient and control samples were down regulated, perhaps providing evidence that metabolic processes were reduced in ambient mussels. However, at this stage these ideas are speculative and further work is required.

4.4.4.2 Foot tissue

Foot tissue exhibited unusual patterns of protein abundance in the current experiment (i.e. a large number of differences between control and ambient samples). The morphology and function of foot tissue may have influenced the results (Figure 4.13). *M. edulis* foot tissue consists of as many as nine different kinds of glands, each with their own specific role in crawling and attachment (Lane and Nott 1975). During dissection the foot was cut into several pieces. It was possible that by chance all of the control samples were taken from a particular piece of the foot tissue (e.g. 'tip', Figure 4.13) whereas ambient analysis may have been performed on a different piece of foot tissue (e.g. 'base'). The differences in protein expression between control and ambient mussels could be due to the different glands present in the two different 'types' of foot tissue chosen for analysis. However, a range of foot tissue

types were used for analysis for both groups, suggesting that foot tissue 'type' was not an issue.

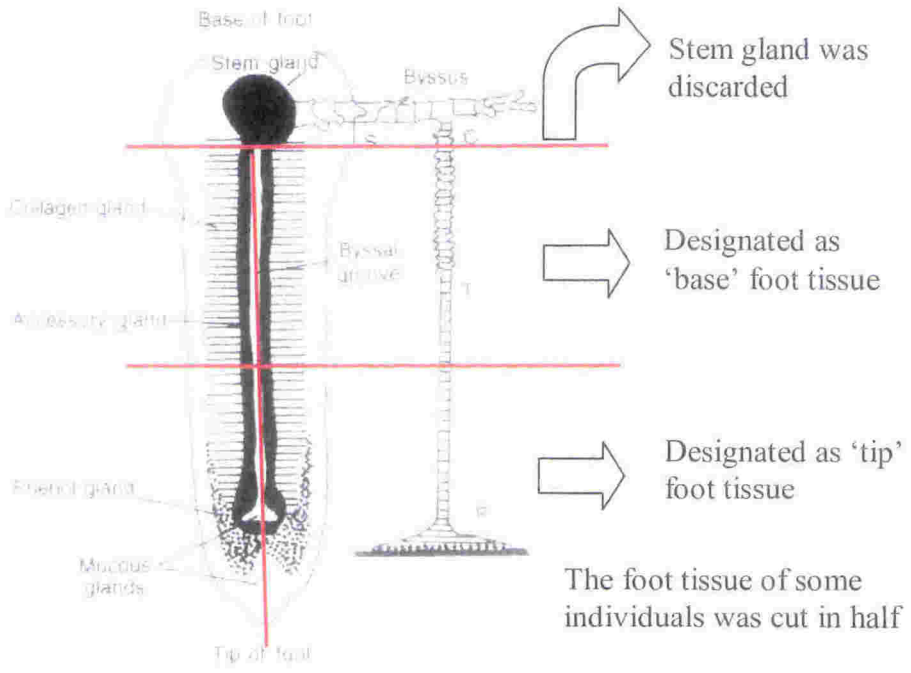


Figure 4.13: A schematic diagram of *M. edulis* foot
The names and approximate locations of the associated glands are labelled. Figure reproduced from Gosling (2003). Red lines demonstrate dissection lines and block arrows indicate specific foot tissue 'types', i.e. tip, base and stem gland (discarded).

4.5 Conclusions

Protein expression signatures for salinity stress in *P. canaliculus* were investigated by comparing protein expression between control samples, and samples kept under low or ambient salinity for 3 days. The protein expression signatures are an important outcome of this thesis and for this reason the 2-D gel location of gill and foot proteins that are up or down regulated as a result of salinity stress have been reported in detail. Approximately 115 proteins showed significant ($p < 0.01$) differences in protein abundance due to experimental conditions. However, in contrast to previous studies there does not appear to be any proteins exclusively associated with any of the groups compared in this study. Instead, all of the proteins reported are found in each of the groups and it is the relative abundance of these proteins that are influenced depending on experimental conditions. Foot tissue had a much greater reaction to experimental conditions than gill tissue, with approximately 10% of proteins showing significant differences in protein abundance compared to

only approximately 1% of the proteins isolated from gill tissue. In addition, differences in protein abundance between control and ambient groups accounted for most of this response. Three proteins which showed changes in abundance due to experimental conditions were identified, one isoform of actin and two isoforms of tropomyosin. These proteins have also been identified in previous environmental proteomic studies and may play a role in coping with salinity stress. Additional insight into the response of mussels to low salinity will be gained by the identification of more proteins.

CHAPTER 5: Cadmium and mercuric chloride bioaccumulation and proteomic induction in *P. canaliculus* under laboratory conditions

Abstract

P. canaliculus were exposed to either 34.3 μg l^{-1} Hg or 0.486 mg l^{-1} of Cd for 3 days. The bioaccumulation of Hg and Cd was measured in gill, digestive gland, foot and soma. Hg uptake was highest in gill tissue, and reached a maximum concentration of 5.21 mg kg^{-1} after 2 days. On day 3 the concentration of Hg in all tissues tested, including gill decreased, although concentrations were still elevated compared to pre-exposure. Cd uptake was maximal in digestive gland, and accumulation in this tissue was linear and was still increasing after 3 days (42.3 mg kg^{-1}). Protein response to both Hg and Cd was investigated in gill tissue using DIGE over a large pH range (pI 4-7 and 6-11). Over 100 proteins were detected which showed significant ($p < 0.01$) differences in abundance between control and metal treated groups but these proteins could not be identified using MALDI-TOF fingerprinting or tandem mass spectrometry. Overall, these data suggest that *P. canaliculus* can be used as a biomonitor for heavy metal contamination in New Zealand. This work provides preliminary information on proteins which vary in response to heavy metal contamination, and future studies should focus on identifying these proteins. Further work to determine whether these biomarkers are useful for identifying contamination by metals under field conditions is presented in Chapter 6.

5.1 Introduction

5.1.1 Biomonitoring in New Zealand

The use of marine mussels for biomonitoring heavy metal contamination has become widely established overseas (Chapter 1), but relatively little information is available on potential sentinel organisms in New Zealand. To date, the majority of the toxicity and conservation research carried out in New Zealand's aquatic environments have focused on freshwater systems including rivers and lakes (e.g. Hickey et al. 1995; Rowe and Chisnall 1997; Rowe et al. 1999), as discussed in the Australian and New Zealand guidelines for fresh and marine water quality (2000).

These studies have led to the development of the New Zealand National River Health Programme and the Macroinvertebrate Community Index (Stark et al. 2001), which is now widely used by regional councils to detect and monitor water quality (Stark et al. 2001). Marine waters in New Zealand have received less attention because most are thought to be reasonably clean, containing only low background concentrations of heavy metals from geothermal sources, compared to waters surrounding nations in the Northern Hemisphere which receive considerable heavy metal inputs from industry (Ridgway and Glasby 1984). There is evidence of pollution in some semi-enclosed coastal waters in New Zealand (e.g. sounds, estuaries and harbours), particularly those close to urban centres where point discharges of sewage, stormwater drains and industrial effluent occur (Ridgway and Glasby 1984). For these reasons most published marine studies have focused on sediment and water quality (Stoffers et al. 1986; Glasby et al. 1988; Williamson et al. 1996; Bolton-Ritchie et al. 1999; Grove 1999; Botherway and Gardner 2002; Ahrens and Depree 2004) and the accumulation of heavy metals by marine organisms (King and Miller 1989; Zauke 1992; Rainbow et al. 1993; Rogers 1999; Thompson and Dowding 1999) found near the main urban centres of Auckland, Wellington and Dunedin. Relatively few studies have surveyed metal content in shellfish throughout New Zealand (NZFSA 1981-2002; Nielsen and Nathan 1975; Kennedy 1986; Fenaughty et al. 1988) and very little information on the uptake of metals by New Zealand marine organisms under laboratory conditions has been published (King et al. 2006). Information on the metal accumulation patterns of a chosen biomonitor is required to establish the period over which the biomonitor reflects the ambient metal bioavailability (Rainbow 1995). Such data are most easily collected in laboratory experiments, but should be complemented by field data, particularly from transplant experiments (Richardson et al. 1994; Bodin et al. 2004; Do Amaral et al. 2004). This research provides preliminary information on the uptake of Cd and Hg into the tissues of *P. canaliculus*.

5.1.2 Development of heavy metal biomarkers for *P. canaliculus*

Cd and Hg accumulation patterns in *P. canaliculus* provide information that could be useful in the development of a marine biomonitoring programme in New Zealand. However, as recognised by international organisations and environmental agencies, risk assessment cannot be based solely on chemical analysis of

environmental samples because this approach does not provide any indication of deleterious effects of contaminants on the biota (Cajaravill et al. 2000). The measurement of the biological effects of pollutants (i.e. biomarkers, see Chapter 1) has become of major importance for the assessment of the quality of the environment. Conventional biomarkers are based on well known detoxicification pathways, e.g. glutathione S-transferase (Fitzpatrick 1995), MT (Geret 2002) and heat-shock proteins (Choresh 2001) and exclude many other proteins which also change abundance in response to pollutants, but whose mode of action is unknown. Proteomics measures the relative abundance of proteins and detects any proteins which are affected by pollutants or environmental conditions. Proteomic analysis does not rely on any previous knowledge about toxicity mechanisms and may help to identify new detoxification pathways (López-Barea and Gomez-Ariza 2006). The background and potential applications of environmental proteomics have been discussed elsewhere (Chapter 4). In this chapter, new proteomic signatures for Cd and Hg contamination in *P. canaliculus* are reported.

5.2 Materials and Methods

5.2.1 Experimental set-up

Adult greenshell mussels (8-15 cm) were collected from Seatoun, Wellington Harbour, New Zealand. Epiphytes were removed and mussels were held overnight in circulating seawater at the Island Bay Marine Laboratory. For heavy metal accumulation testing the mussels were split into 3 groups (Control, Cd treated and Hg treated) with similar biomass (n = 30 for each group), each group being maintained in 20 l of unfiltered aerated seawater at approximately 18 °C. Air was delivered to each tank using a pump, the water was not changed (i.e. closed system), and mussels were fed *Isochrysis galbana*. Experimental animals were exposed to either 34.3 μg l^{-1} Hg (0.13 μM) as $HgCl_2$ or 0.486 mg l^{-1} Cd (2.13 μM) as $CdCl_2$ for 72 h. Control animals were kept under the same conditions but without added metal. After the metal salts had been added to the seawater and mixed, but prior to the addition of mussels (Time 0), a water sample was taken from each of the three experimental tanks (i.e. control, Hg and Cd) for chemical analysis by Hill Laboratories (Hamilton). Initially tests for Hg, Cd, zinc (Zn), lead (Pb), copper (Cu), arsenic (As), chromium (Cr) and nickel (Ni) were conducted. Further water samples

were taken after 12, 24 and 72 h, but only Hg, Cd and Pb were analysed at these additional times. Pb was tested for comparative purposes because it was expected to be present at low levels and not to be influenced by experimentally altered levels of Cd and Hg i.e. it was expected to remain at unchanged concentrations during the experiment. The concentrations of the other metals were tested at the start of the experiment to establish baseline levels but they were not monitored during the course of the experiment because they were not strictly relevant to the current study.

Gill, foot, digestive gland and remaining soft tissue (soma) were dissected from 6 mussels and pooled into 3 replicates of 2 individuals for each tissue, to ensure there was enough material for analysis. Tissue samples were collected at time 0 and from each of the 3 groups (control, Cd and Hg) at 24, 48 and 72 h (n = 18 per group). Mussels were sent to ELS for analysis. At time 0 the mussels were tested for Hg, Cd, Zn, Pb, Cu, As, Cr and Ni to establish baseline levels for these heavy metals. At 24, 48 and 72 h, the mussels were tested for Hg, Cd (test metals) and Pb (comparison). After 72 h, 10 mussels from each group were killed for proteomic analysis. Foot, gill, and the digestive gland were dissected from each mussel and tissue was stored at -80°C until required.

The concentration of Cd used in the current experiment (0.486 mg l^{-1}) although relatively high, was less than the sub-lethal concentration of 1.25 mg l^{-1} determined for *P. viridis* (Yap et al. 2003c). Although the experimental dosage may be outside of ecologically relevant concentrations in New Zealand, it was required to ensure a strong proteomic response for the development of proteomic biomarkers. Chinese mitten crabs exposed to a similar concentration of Cd (500 μg Cd l^{-1} for 3 days) exhibited physiological, biochemical and ultrastructural damage of tissues (Slivestre et al. 2006). Substantial depletion of the lateral frontal cilia has also been noted in *P. perna* during early exposure to 50 μg l^{-1} Hg, a similar amount of Hg as used in the current study (Gregory et al. 2002). These studies indicate that the amount of metal used in the current study was likely to be sufficient to induce a proteomic response.

5.2.1.1 Bioconcentration factor (BCF) and rate of accumulation

The bioconcentration factor (BCF) is used to describe the accumulation of chemicals in aquatic organisms exposed to contaminated environments. A steady-

state BCF occurs when the organism is exposed for a sufficient length of time so that the BCF does not change substantially. The BCF in the current experiment was calculated in relation to metal concentration in seawater according to Taylor (1983):

$$BCF = \frac{Ce - Ci}{Cs}$$

where: Ce = metal concentration in the mussel tissue during exposure
Ci = initial metal concentration in the mussel tissue before exposure
Cs = experimental metal concentration in the test seawater

Rate of accumulation (RA) was calculated according to Yap et al. (2003c):

$$= \frac{\text{Metal level (exposed)} - \text{Metal level (control)}}{\text{Day(s) of metal exposure}}$$

5.2.2 PERMANOVA

Permutational multivariate analysis of variance (Chapter 3) was conducted on the basis of Bray-Curtis dissimilarities using PERMANOVA v1.6 (Anderson 2001; McArdle and Anderson 2001) to determine whether the differences in metal accumulation between tissues (gill, digestive gland, foot and soma) and time (24, 48 and 72 h) were significant. Data were not transformed or standardised. The number of permutations used was 999 and the integer used as a seed was 6. PRIMER v5 (Colwell 2005) was used to create non-Metric Multi Dimensional Scaling (nMDS) plots based on Bray-Curtis dissimilarity using non-standardised and non-transformed data.

5.2.3 Protein Extraction

Proteins were extracted from 100 mg of gill tissue in 100 µl of standard DIGE lysis buffer (30 mM Tris-Cl, 7 M urea, 2 M thiourea, 4% CHAPS, Appendix 2) and were processed as described in Chapter 3. Additional lysis buffer was added to make each sample up to 1 mL. Protein was purified before labelling using a 2-D Clean Up Kit (GE Healthcare).

5.2.4 DIGE labelling

Samples were labelled as outlined in Chapter 3. An Ettan™ DIGE protein labelling kit was used. Samples were adjusted to pH 8.5 and 80 pmol of dye was added to 10 µg of protein. Samples were left on ice in the dark for 30 min, 1 µl of 10 mM lysine was used to stop the labelling reaction and samples were left on ice in the

dark for a further 10 min. At this point samples could be frozen at -80 °C until required, or prepared for 2-DE (5.2.5).

5.2.5 Sample preparation for 2-DE

Samples labelled with different CyDyes™ were combined in a microcentrifuge tube (Table 5.1; see Chapter 3 for further details). Immobiline™ dry strips of pH range 4-7 and 6-11 were used to separate gill proteins.

Gel #	Cy2	Cy3	Cy5	Gel #	Cy2	Cy3	Cy5
1	STD	51	64	8	STD	75	56
2	STD	52	65	9	STD	76	57
3	STD	53	71	10	STD	77	58
4	STD	61	72	11	STD	78	69
5	STD	62	73	12	STD	66	80
6	STD	63	54	13	STD	67	79
7	STD	74	55	14	STD	59	70
				15	STD	60	68

Table 5.1: Labelling strategy for the heavy metal DIGE experiment.

STD = Pooled internal standard (n = 30). Samples 51-60 = Control; 61-70 = Mercury; 71-80 = Cadmium. Samples for gels 1-7 were labelled and separated using electrophoresis before gels 8-15.

5.2.6 2-DE

Approximately 10 µg of each labelled sample (30 µg total) was loaded on to pH 4-7 IPG strips, using rehydration loading. Basic strips (pH 6-11) were cup-loaded. Strips were focused using an Ettan IPGphor isoelectric focusing system using the following protocol: Step 300 V for 30 min, gradient steps at 1000 V and 5000 V for 30 min and 90 min respectively, and a final step of 5000 V for 25 min.

Before the second dimension, strips were equilibrated for 10 min in 2.5 mL/strip equilibration solution plus 0.25% DTT, then 10 min in 2.5 mL/strip equilibration solution plus 4.5% iodoacetamide (Appendix 2). Strips were loaded onto NuPAGE 4-12% Bis-Tris gels (Invitrogen) and proteins were separated by electrophoresis at 200 V, 140 mA for 1 h.

5.2.7 Selecting gels for DeCyder™ analysis

Gels were scanned using a Fujifilm FLA-5000 scanner, converted into the appropriate file format and imported into the DeCyder workspace. A few gels were excluded from analysis because of poor spot separation and matching. Data sets in the current study were reduced from 45 spot maps (i.e. 15 gels x 3 dyes) to the numbers shown in Table 5.2. The number of gels removed from this analysis was low compared to the salinity experiment (Chapter 4) thus the power of this experiment to identify statistically significant changes in protein abundance was greater than that of the salinity experiment.

	Control (n =10)	Cadmium (n =10)	Mercury (n =10)	Standard (n =15)	Total # spot Maps (n = 45)
Gill pI 4-7	9	9	10	14	42
Gill pI 6-11	8	9	9	13	39

Table 5.2: Number of samples included for analysis for each pI range

5.2.8 DeCyder 2-D™ analysis

5.2.8.1 Spot matching

Protein spots were checked to confirm accurate matching between gels. Spots that were not well matched or well resolved were removed from analysis (Chapter 3). In some cases spots were manually added to gels because they were not detected by the DeCyder software in all gels. Manually adding spots ensured that the same spot was detected in all gels, thereby giving a greater number of samples for statistical analysis. Spots that required manual addition or other manipulations are indicated with an asterix in Tables 5.4 and 5.6. The average ratio (Av. Ratio) values reported in these Tables relate to the differences in protein abundance between different experimental groups. The differences are reported as fold-change. For example, a value of 2 means a two-fold increase in abundance, whereas -2 indicates a two-fold decrease (see Chapter 3 for further details).

5.2.8.2 Students t-*test*

Initially, protein spot maps were analysed in the DeCyder BVA module. The Student's *t*-test was performed pairwise between each of the three different experimental groups (control, Cd, Hg), as listed below (1-3), to detect the proteins

that responded to experimental conditions. Protein spots which exhibited statistically significant (*t*-test $p < 0.01$) and relatively large changes in protein abundance (i.e. average ratios >1.5 and < -1.5) were designated as "Proteins of Interest" and were used for further analyses (5.8.2.3). The False Discovery Rate (FDR) correction (Benjamini and Hochberg 2000) was applied to all of the proteins reported in this chapter.

1) Control versus Cd: Both groups experienced similar experimental conditions but the mussels in the Cd tank were exposed to 0.486 mg l^{-1} Cd. Therefore, differences in protein abundance should be solely due to the effects of Cd. *t*-test calculations were performed between control and Cd groups and the dataset was filtered to retain only those proteins which were statistically significant (*t*-test $p < 0.01$) and which had a relatively large change in protein abundance (i.e. average ratios >1.5 and < -1.5).

2) Control versus Hg: Both groups experienced similar experimental conditions as above, but mussels in the Hg tank were exposed to 34.3 µg l^{-1} Hg. Therefore, differences in protein abundance should be solely due to the effects of Hg. *t*-test calculations were performed between control and Hg groups and the dataset was filtered as specified above to retain those proteins which showed statistically significant responses to Hg

3) Cd versus Hg: Both groups were exposed to similar experimental conditions, but the mussels in the Cd tank were exposed to 0.486 mg l^{-1} Cd whereas mussels in the Hg tank were exposed to 34.3 µg l^{-1} Hg. *t*-test calculations were performed between Cd and Hg groups and the dataset was filtered as specified above. Differences in protein abundance between these groups could indicate specific responses to either Cd or Hg at the concentrations used in the current experiment.

5.2.8.3 Extended Dataset Analysis (EDA)

Proteins which showed statistically significant changes in abundance were exported into EDA and pattern analysis and Principal Component Analysis (PCA) were performed. Pattern analysis uses hierarchical clustering to group together protein and spot maps with similar protein abundance profiles (see Chapter 3). The results of the pattern analysis give a visual overview of the relationship between experimental groups. PCA generates a more detailed overview of the data which can be used to visualise outliers and to conceptualise how these outliers relate to the rest of the data. Protein versus Spot Map calculations find protein outliers and perform a

comparison of the relationship between proteins and spot maps. Spot Map versus Protein calculations check if there are any spot map outliers. Spot maps in the same experimental group should be grouped together (see Chapter 3 for further details).

5.2.9 Identification of protein spots using MALDI-TOF mass fingerprinting

Protein spots were prepared for MALDI-TOF mass fingerprinting using a standard acetonitrile (ACN) and ammonium bicarbonate protocol, and were digested with trypsin at 30 °C overnight as described in Chapter 3.

5.2.10 MT PAGE kit

MT is a metal binding protein and as such it was expected that it would be induced by experimental conditions (i.e. exposure to heavy metals). The MT-PAGE kit (Ikzus, Italy) contains a fluorescent dye (unspecified by the manufacturer) which allows the visualisation of MT under UV light. Briefly, approximately 0.5 g of gill, mantle edge, digestive gland and foot tissue were individually homogenised on ice in buffer (supplied by Ikzus), using sand and a mortar and pestle. The homogenate was centrifuged briefly at room temperature, and the supernatant was transferred in to a clean microcentrifuge tube. The supernatant was centrifuged at 30,000 *g* at 4 °C for 20 min and then transferred to a clean tube. To concentrate the amount of MT isolated the protein was precipitated using ethanol at -20 °C for 1 h. The sample was then centrifuged at 16,000 *g* for 5 min at 4 °C. All of the supernatant was removed and the pellet was allowed to dry at room temperature for 5 min. The protein pellet was resuspended in buffer (supplied).

A MT standard from rabbit (supplied) and the samples were incubated at 65 °C for 15 min. While the samples were still at 65 °C the fluorescent agent was added and the tubes were transferred to room temperature. After 5 min, solution E (supplied SDS buffer, with added mercaptoethanol) was added to the samples and incubated for 5 min. Finally the samples were centrifuged at 10,000 *g* for 1 min at room temperature before being loaded on to a 1-D or 2-D gel. Protein spots labelled with dye were visualised using an UVP Inc. UV transilluminator (San Gabriel, Ca, USA) with a 5.6 aperture and a manual exposure setting. The Fuji scanner could not be used as it did not have the appropriate filters.

5.3 Results

5.3.1 Heavy metal in seawater

Hg stock solution was added to the experimental tank and after 5 min of mixing using air bubbles circulating through the tank, an initial Hg measurement of 0.0343 mg l^{-1} was obtained. Mussels (n = 30) were added to the tank and after 24 h the quantity of Hg in the seawater was measured at 0.00059 mg l^{-1}, representing an approximate >55-fold decrease compared to the initial Hg value. Greater than 98% of the initial Hg had been removed from the water by mussels and other processes (e.g. adhesion to sediments). After 48 and 72 h the amount of Hg detected in seawater was 0.00041 mg l^{-1} (approximately 0.7-fold decrease) and 0.00035 mg l^{-1} (approximately 0.9-fold decrease) respectively, relatively small decreases compared to the large initial drop in Hg concentration experienced during the first 24 h (Figure 5.1).

In contrast, the Cd level initially detected in the seawater was 0.486 mg l^{-1}. After adding mussels (n = 30) to the tank and waiting for 24 h, the amount of Cd detected in the seawater decreased to 0.301 mg l^{-1} an approximately 1.5-fold decrease. After 48 and 72 h the amount of Cd detected was 0.247 mg l^{-1} (approximately 0.8-fold decrease) and 0.208 mg l^{-1} (approximately 0.8-fold decrease) respectively (Figure 5.2).

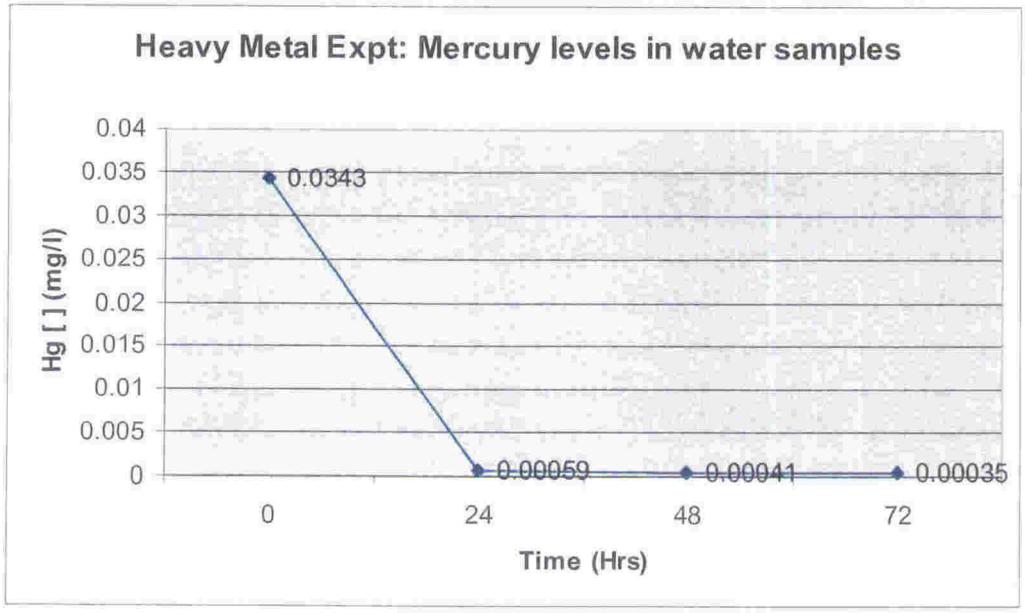


Figure 5.1: Hg measurements in seawater

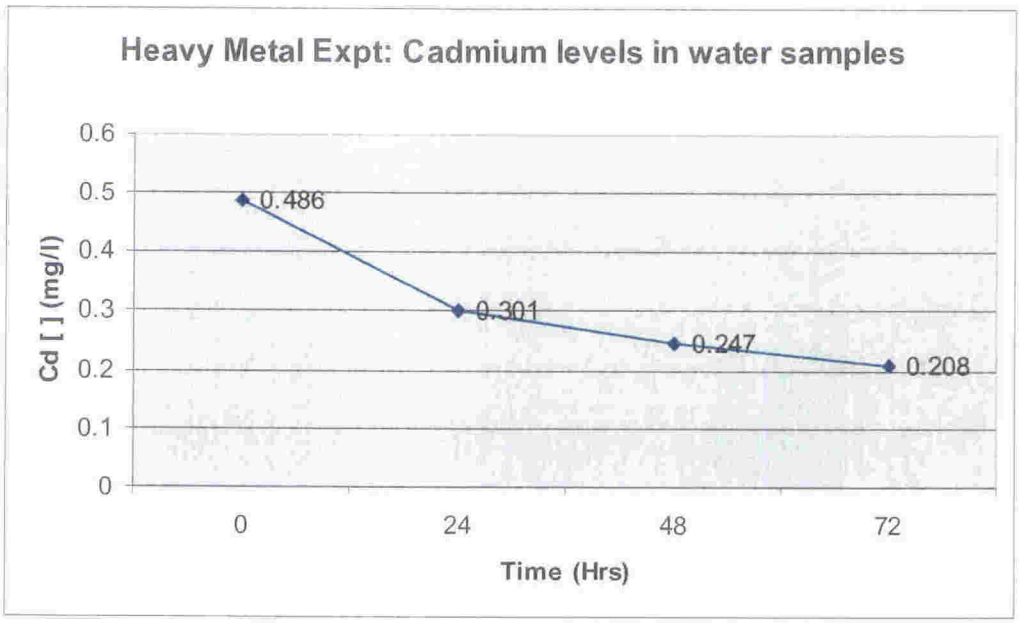


Figure 5.2: Cd measurements in seawater

5.3.2 Heavy metal accumulation in mussel tissue

Hg concentrations in gill, digestive gland, foot and the soma were measured at the start of the experiment and after 24, 48 and 72 h (Figure 5.3). Permutation testing revealed a significant influence of Time (p = 0.0480) and Tissue (p = 0.0010) on the amount of metal accumulated and also a significant interaction between Time and Tissue (p = 0.0020). In all four tissues, the highest concentration of Hg was detected after 48 h (e.g. gill = 5.22 mg kg^{-1}), with a decrease in Hg after 72 h (e.g. gill = 2.70 mg kg^{-1}). Gill tissue had the highest concentration of Hg detected at all time points, with digestive gland showing the next highest Hg levels. Foot and the remaining soma both had relatively low concentrations of accumulated Hg compared to gill and digestive gland (Figure 5.3). An nMDS plot illustrates the relationships between Hg accumulation and tissue type. Gill and digestive gland accumulated the highest concentration of Hg as represented by the size of the green circles (Figure 5.4). The low stress value (0.01) indicates that the nMDS plot is a good representation of the data (see Chapter 3).

In digestive gland tissue the concentration of Cd measured at each time point increased and was still increasing after 72 h (13.6 to 42.3 mg kg^{-1} from 24 to 72 h). Cd accumulation in the remaining tissues was relatively low and the concentration decreased slightly after 72 h (Figure 5.5). The highest concentration of Cd was

detected in digestive, tissue not in gill as for Hg. Permutation testing found a significant influence of Time (p = 0.0030) and Tissue (p = 0.0010) on the level of metal accumulation and also a significant interaction between Time and Tissue (p = 0.0270). An nMDS plot illustrates the relationships between Cd accumulation and tissue type. Digestive gland accumulated the highest concentration of Cd as represented by the size of the green circles (Figure 5.5). The low stress value (0.01) indicates that the nMDS plot is a good representation of the data (see Chapter 3).

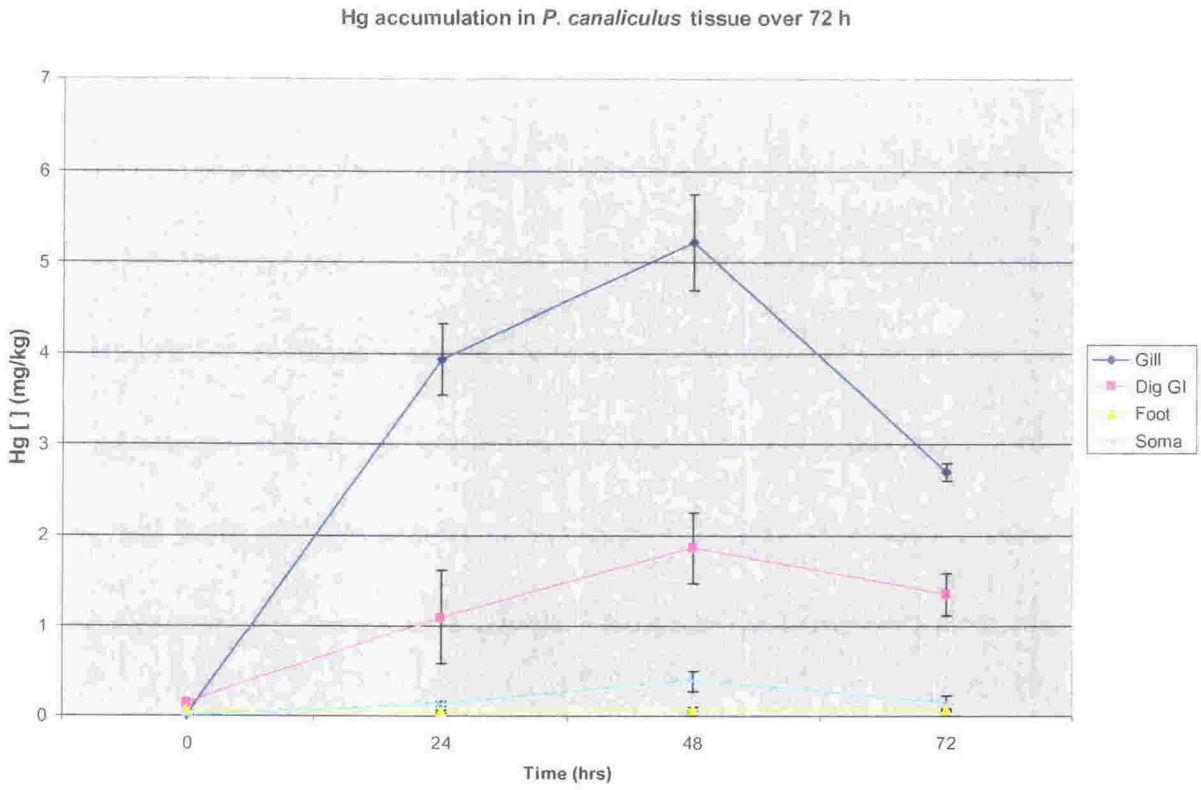


Figure 5.3: Hg accumulation in *P. canaliculus* tissue

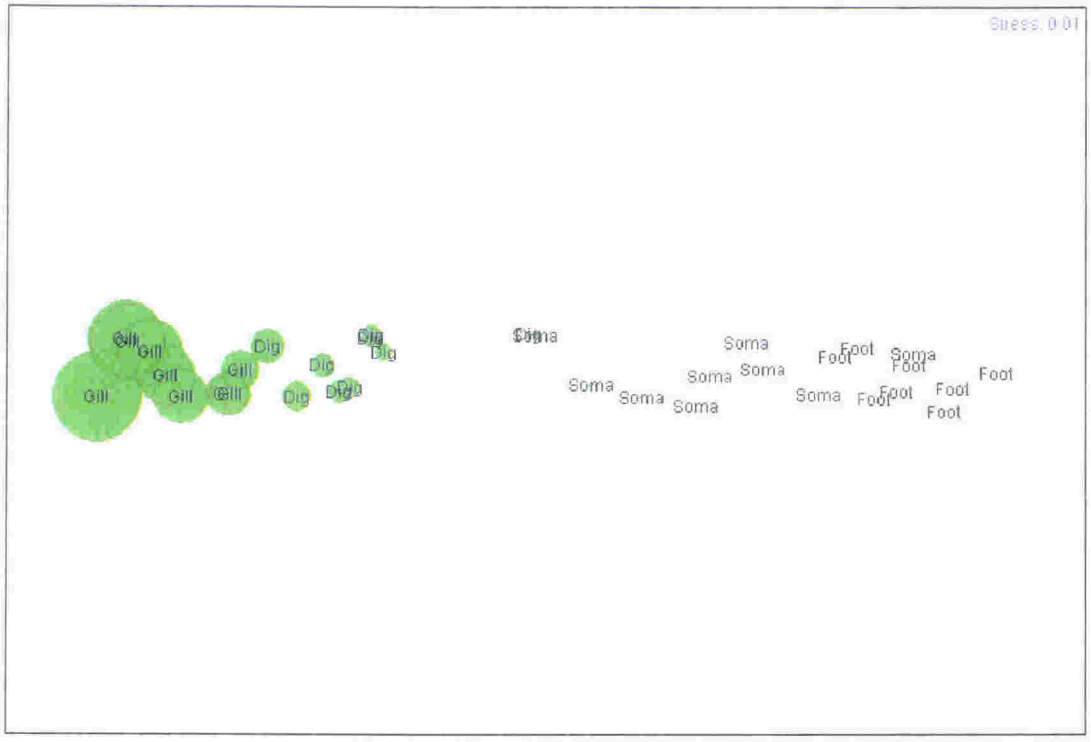


Figure 5.4: A nMDS plot to show the relationships between Hg accumulation and tissue type. Dig = Digestive gland.

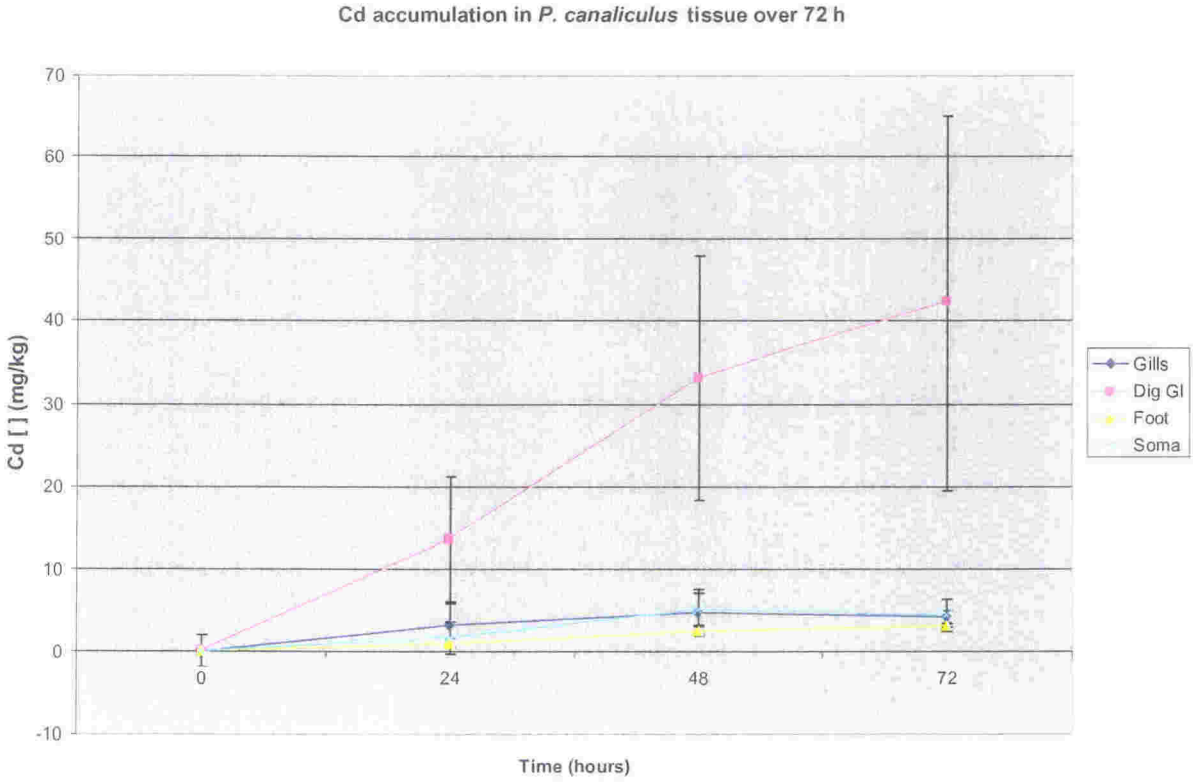


Figure 5.5: Cd accumulation in *P. canaliculus* tissue

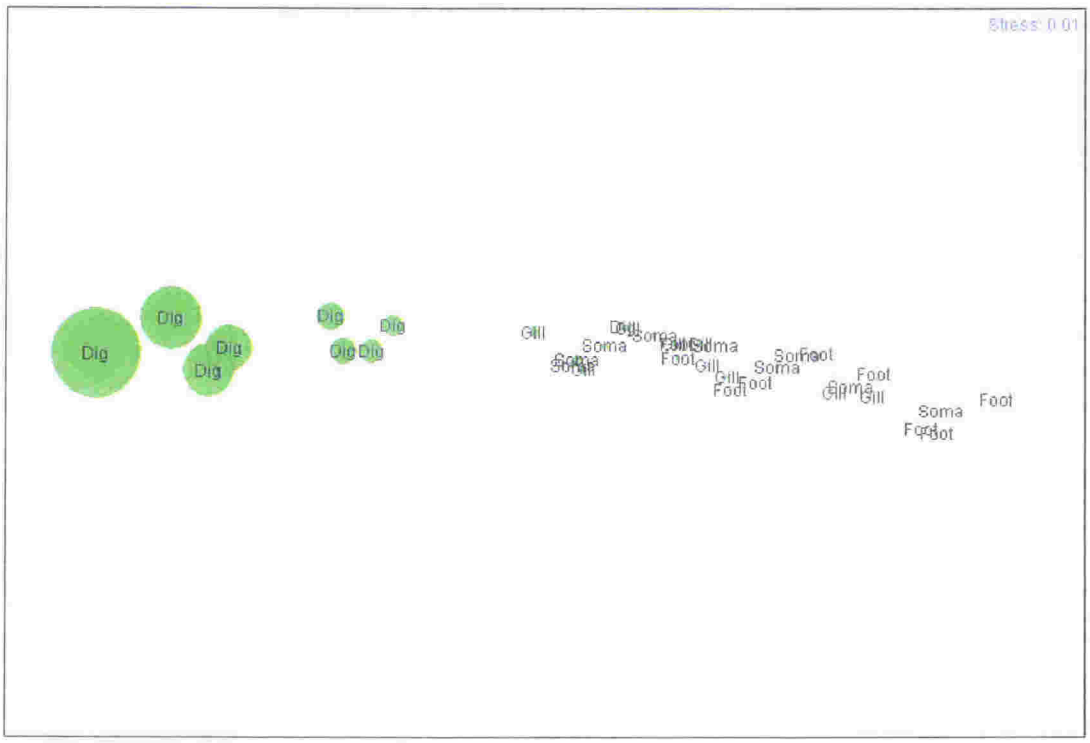


Figure 5.6: A nMDS plot to show the relationships between Cd accumulation and tissue type. Dig = Digestive gland.

5.3.2.1 Bioconcentration Factor (BCF) and Rate of accumulation (RA)

Hg reached a maximum BCF of 151.6 in gill tissue at 48 h (Table 5.3), which correlated with the maximum Hg concentration that was measured in gill tissue at the same time point. Other tissues followed the same general pattern, with a maximum BCF at 48 h and a reduced BCF at 72 h. In contrast the RA for gill tissue and most other tissues excluding the soma were at a maximum at 24 h. In whole tissue and the soma the BCF and the RA are both at a maximum at 48 h (Table 5.3). In contrast Cd reached a maximum BCF of 42.33 in digestive tissue at 72 h. Foot tissue also had a maximum BCF at 72 h whereas gill and soma reached a maximum BCF at 48 h. In digestive gland, foot and soma the maximum RA was at 48 h, whereas the maximum RA for gill tissue was at 24 h. Overall, the BCF for Cd is less that the BCF obtained for Hg. There are differences in BCF and RA of Cd and Hg among tissues.

	Mercury (Cs = 0.0343 mg l^{-1})				Cadmium (Cs = 0.486 mg l^{-1})			
Average Tissue [ ] (mg/kg)	**Time 0 (Ci)**	**24**	**48**	**72**	**Time 0 (Ci)**	**24**	**48**	**72**
Gill	0.01	3.93	5.21	2.70	0.02	3.21	4.83	4.25
± SD	± 0	± 0.39	± 0.53	± 0.10	± 0.003	±2.67	± 2.76	± 0.78
BCF		114.29	151.60	78.43		6.56	9.90	8.66
RA		3.92	2.60	0.90		3.19	2.41	1.41
Dig Gl	0.03	1.10	1.86	1.35	0.14	13.60	33.13	42.33
± SD	± 0.01	± 0.52	± 0.39	± 0.23	± 0.08	± 7.52	± 14.79	± 22.80
BCF		31.20	53.35	38.48		27.70	67.88	86.81
RA		1.07	0.92	0.44		13.46	16.50	14.06
Foot	0.01	0.05	0.07	0.07	0.06	1.02	2.52	3.20
± SD	± 0	± 0.01	± 0.02	± 0.01	± 0.05	± 0.18	± 1.30	± 0.66
BCF		1.12	1.75	1.75		1.98	5.06	6.46
RA		0.04	0.03	0.02		0.96	1.23	1.05
Soma	0.01	0.13	0.38	0.15	0.05	1.67	5.19	4.46
± SD	± 0.0006	± 0.03	± 0.11	± 0.08	± 0.03	± 0.59	±1.67	± 1.99
BCF		3.50	10.79	4.08		3.33	10.58	9.07
RA		0.12	0.19	0.05		1.62	2.57	1.47
Whole tissue *	< 0.01	0.49	0.97	0.45	0.1 (approx)	2.70	7.5	7.42
BCF		14.3	28.3	13.1		5.6	15.4	15.3
RA		0.49	0.49	0.15		2.7	3.75	2.47

Table 5.3: Metal accumulation in the soft tissues of *P. canaliculus* during 3 days of exposure.

Values reported are the average of 3 separate measurements based on 2 pooled samples per measurement (n = 6 total). *The whole tissue measurements were calculated based on the average contribution of each tissue to the whole soft tissues of the animal and are therefore approximate values.

5.3.3 DeCyder results

5.3.3.1 Gill tissue pI 4-7

In gill tissue pI 4-7, 78 proteins exhibited significant ($p < 0.01$) differences in protein abundance between one or more experimental groups (Table 5.4). This was approximately 6% of the total number of protein spots separated in this dataset (approximately 1400 spots in total). The locations of proteins exhibiting significant differences in abundance are shown in Figure 5.7. The numbering on the spot map corresponds with the numbers shown in the 'Master No.' column in Table 5.4. A summary of the results presented in Table 5.4 is shown in Table 5.5. The larger number of proteins shown in Table 5.5 (n = 123) than Table 5.4 (n = 78) was because the abundance of several proteins changed significantly in more than one experimental comparison (Figure 4.2). There were no statistically significant variations in the abundance of proteins in comparisons between Cd and Hg treated samples when the FDR adjustment was applied to the data, suggesting that these chemicals may have similar modes of action. However, there were significant ($p < 0.01$) differences in the abundance of proteins in comparisons between control and Cd treated, and control and Hg treated samples after FDR corrections had been applied. All of the data reported in this chapter have had FDR corrections applied.

Table 5.4: Gill pI 4-7 proteins which varied significantly ($p < 0.01$)
* Proteins spots that were manually added or manipulated

Pos.	Master No.	Group comparison	T-test	Av. Ratio	1-ANOVA
1	1171	Control vs. mercury	0.0083	-1.65	0.0022
		Control vs. cadmium	0.0081	-1.79	
2	1180	Control vs. mercury	0.00045	2.13	3.0e-005
		Control vs. cadmium	0.00025	1.76	
3*	1181	Control vs. mercury	0.00078	1.52	0.00068
4	1186	Control vs. mercury	0.00036	1.72	1.8e-005
		Control vs. cadmium	0.0019	1.57	
5	1224	Control vs. mercury	0.0018	-1.58	0.00031
		Control vs. cadmium	0.0039	-1.61	
6*	1229	Control vs. mercury	3.0e-006	-2.18	1.4e-006
		Control vs. cadmium	0.00029	-2.32	
7*	1343	Control vs. cadmium	0.0037	-1.88	0.0035
8*	1350	Control vs. mercury	0.0043	-2.03	0.0052
9*	1402	Control vs. mercury	0.0019	-2.01	0.00032
		Control vs. cadmium	0.0079	-1.80	
10	1406	Control vs. mercury	0.0090	-1.89	0.0022
		Control vs. cadmium	0.0015	-2.19	
11*	1414	Control vs. mercury	0.0031	-1.67	0.0085
12	1439	Control vs. mercury	7.4e-006	-2.22	8.9e-007

		Control vs. cadmium	0.00034	-2.01	
13*	1529	Control vs. mercury	0.00063	-2.17	0.0012
14*	1535	Control vs. mercury	0.0050	-1.71	0.0062
15	1661	Control vs. cadmium	0.0096	-1.83	0.0062
16*	1663	Control vs. cadmium	0.00095	1.63	0.00022
17*	1753	Control vs. mercury	0.0060	-1.84	0.0085
18	1772	Control vs. mercury	0.0068	1.69	0.011
19	1841	Control vs. mercury	5.5e-008	-3.43	4.2e-012
		Control vs. cadmium	1.2e-007	-3.25	
20	1844	Control vs. mercury	0.00034	1.78	5.7e-005
		Control vs. cadmium	0.0028	1.65	
21*	1861	Control vs. mercury	0.00078	1.70	0.00012
		Control vs. cadmium	0.0027	1.57	
22*	1864	Control vs. mercury	0.0010	1.51	0.00031
23*	1866	Control vs. mercury	0.0012	1.87	0.00022
		Control vs. cadmium	0.00095	1.86	
24	1871	Control vs. mercury	0.0086	1.83	0.0048
25	1938	Control vs. mercury	0.00059	-1.61	0.00057
26	1954	Control vs. mercury	0.00059	1.65	0.00012
		Control vs. cadmium	0.0066	1.54	
27*	1959	Control vs. mercury	0.00065	-1.95	0.00032
		Control vs. cadmium	0.0027	-2.16	
28	1973	Control vs. cadmium	0.00027	2.13	0.00073
29	1978	Control vs. cadmium	0.0053	-1.51	0.0064
30	2008	Control vs. mercury	0.0012	1.51	0.0012
		Control vs. cadmium	0.0066	1.51	
31	2012	Control vs. mercury	0.0010	1.85	0.00068
32	2021	Control vs. mercury	0.00096	-1.86	0.0022
33	2026	Control vs. mercury	0.0028	2.56	0.0015
34	2047	Control vs. mercury	4.3e-006	-3.61	1.3e-008
		Control vs. cadmium	3.2e-006	-3.99	
35*	2049	Control vs. mercury	0.0016	-1.61	0.00013
		Control vs. cadmium	0.00089	-1.80	
36*	2059	Control vs. mercury	0.0028	1.91	0.0012
		Control vs. cadmium	0.0080	1.68	
37	2068	Control vs. mercury	0.00049	-1.99	8.0e-006
		Control vs. cadmium	7.1e-005	-2.21	
38	2082	Control vs. mercury	4.5e-006	-2.28	1.6e-007
		Control vs. cadmium	2.2e-005	-2.80	
39	2084	Control vs. mercury	0.00039	-1.69	0.00050
		Control vs. cadmium	0.0056	-1.73	
40*	2085	Control vs. mercury	0.0012	-2.12	4.0e-005
		Control vs. cadmium	2.7e-005	-2.63	
41	2089	Control vs. mercury	0.0016	-1.72	0.00038
		Control vs. cadmium	0.0039	-1.93	
42	2095	Control vs. mercury	0.0012	-1.58	0.00029
		Control vs. cadmium	0.00049	-1.70	
43	2101	Control vs. mercury	0.00030	-2.34	2.1e-005
		Control vs. cadmium	9.0e-005	-2.73	
44	2140	Control vs. mercury	0.0039	1.61	0.00070
		Control vs. cadmium	0.0053	1.65	
45	2156	Control vs. mercury	0.00054	-1.88	0.00021
		Control vs. cadmium	0.0072	-1.78	
46	2177	Control vs. mercury	0.0098	1.80	0.0034
47	2186	Control vs. mercury	0.0074	1.66	0.00033
		Control vs. cadmium	0.00076	1.91	
48*	2203	Control vs. cadmium	0.0035	-1.76	0.0033
49	2205	Control vs. mercury	2.0e-005	-2.68	2.8e-007

		Control vs. cadmium	5.2e-005	-2.33	
50*	2215	Control vs. mercury	0.0044	1.56	0.0033
		Control vs. cadmium	0.0077	1.62	
51	2249	Control vs. mercury	0.0045	-1.73	0.0011
		Control vs. cadmium	0.0020	-1.78	
52	2254	Control vs. mercury	0.0012	-1.75	0.00015
		Control vs. cadmium	0.00037	-1.95	
53*	2298	Control vs. mercury	3.0e-006	-2.21	2.0e-007
		Control vs. cadmium	2.2e-005	-2.47	
54*	2301	Control vs. mercury	6.1e-007	-2.01	9.3e-008
		Control vs. cadmium	2.2e-005	-2.13	
55	2313	Control vs. mercury	0.0029	2.15	0.0035
56*	2330	Control vs. mercury	0.0096	1.95	0.0052
57*	2355	Control vs. mercury	0.0023	1.52	0.0036
58*	2372	Control vs. mercury	0.00027	1.58	7.0e-005
		Control vs. cadmium	0.0029	1.56	
59*	2396	Control vs. mercury	0.0015	-1.94	3.5e-005
		Control vs. cadmium	0.00049	-2.27	
60	2474	Control vs. mercury	0.0018	1.55	8.3e-005
		Control vs. cadmium	0.00073	1.58	
61	2488	Control vs. mercury	0.0015	1.52	0.00073
62*	2494	Control vs. mercury	0.0055	2.03	0.0039
		Control vs. cadmium	0.0092	1.86	
63	2541	Control vs. mercury	0.0098	1.55	0.0094
64*	2582	Control vs. mercury	0.00011	1.63	8.3e-005
		Control vs. cadmium	0.0016	1.63	
65	2595	Control vs. mercury	5.5e-008	2.84	4.2e-012
		Control vs. cadmium	1.2e-007	2.91	
66*	2596	Control vs. mercury	4.2e-005	1.96	2.8e-007
		Control vs. cadmium	2.2e-005	2.00	
67	2611	Control vs. cadmium	0.0024	-1.55	0.00077
68	2617	Control vs. cadmium	0.0013	1.50	0.0019
69*	2668	Control vs. mercury	0.00045	-2.00	8.8e-005
		Control vs. cadmium	0.00095	-2.34	
70*	2681	Control vs. cadmium	0.0053	-1.53	0.0039
71*	2733	Control vs. mercury	0.0011	2.17	0.00073
72*	2735	Control vs. mercury	0.00065	1.80	0.00050
73*	2737	Control vs. cadmium	0.00073	1.85	0.00026
74*	2834	Control vs. mercury	2.7e-005	-2.53	9.3e-008
		Control vs. cadmium	2.2e-005	-3.00	
75	2836	Control vs. mercury	0.00017	1.52	1.8e-005
		Control vs. cadmium	0.00049	1.54	
76*	2838	Control vs. mercury	2.0e-005	2.06	0.00010
		Control vs. cadmium	0.0074	1.81	
77*	2840	Control vs. mercury	0.00019	1.52	0.00012
78*	2854	Control vs. mercury	0.00059	1.81	0.0021

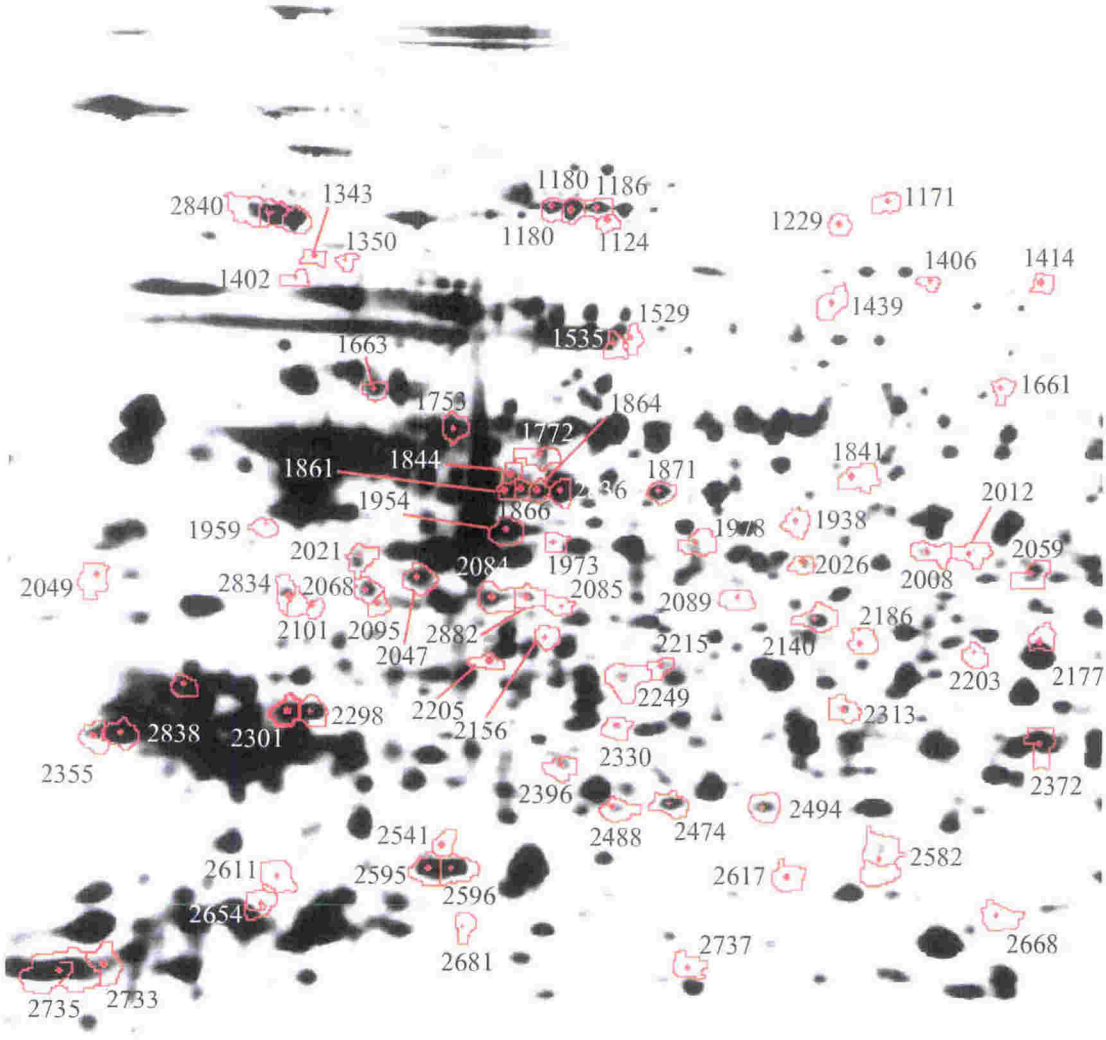


Figure 5.7: Gill pI 4-7 gel indicating spots which varied significantly between groups ($p < 0.01$) Spot 2595 was also significant in the salinity study Chapter 4 (spot 3478).

Group Comparison	# spots increased ↑			# spots decreased ↓			TOTAL
	1.5+	2.0+	Total	-1.5+	-2.0+	Total	(approx. 1400 spots)
Control vs. Cd	21	3	24	15	16	31	55
Control vs. Hg	28	7	35	18	15	33	68
TOTAL	49	10	59	33	31	64	123

Table 5.5: Gill pI 4-7 summary of significant ($p < 0.01$) results

Exposure to Hg resulted in a greater number of significantly up or down regulated proteins (68) than exposure to Cd (55) (Table 5.4). Hg had a greater effect on protein abundance than Cd because 23 proteins were significantly affected by Hg exposure and not Cd, compared to only 10 proteins exclusively affected by Cd (Figure 5.8). However, the variation in abundance of most proteins were

significantly affected by exposure to both Cd and Hg (45), suggesting a common mode of action for these metals. One protein spot (2595) also had significant changes in abundance in response to salinity (Chapter 4; spot 3478). However, this spot has not been identified.

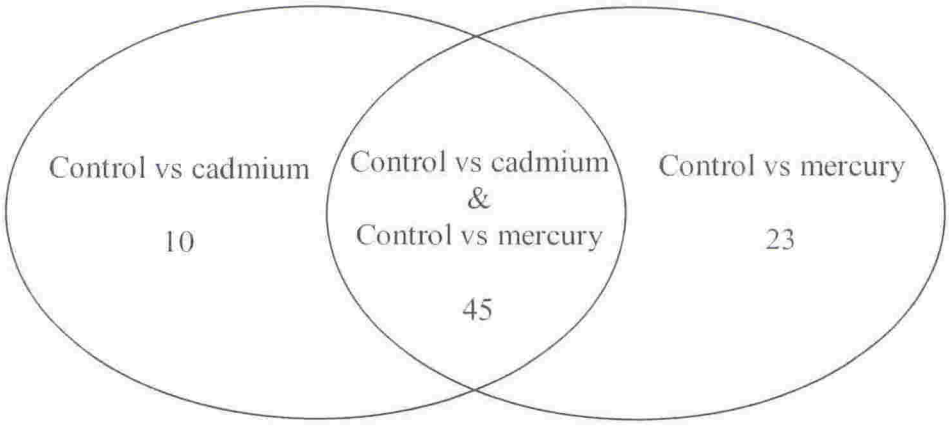


Figure 5.8: Venn diagram illustrating the distribution of significant ($p < 0.01$) protein spots and their overlap between treatments in gill tissue pI 4-7

Pattern analysis splits the spot maps into 2 major groups; control and metal exposed (Cd and Hg), shown on either side of the yellow line (Figure 5.9). Within the control samples, a single control spot map appears to be an outlier (to the immediate left of the yellow line). Within the Cd and Hg treated spot maps there were three distinct sub-groupings, which are indicated by blue lines. The major sub-group contained samples treated with Cd and samples treated with Hg (i.e. a mixture of both experimental classes), whereas the other sub-groups either exclusively contained individuals exposed to Hg or Cd.

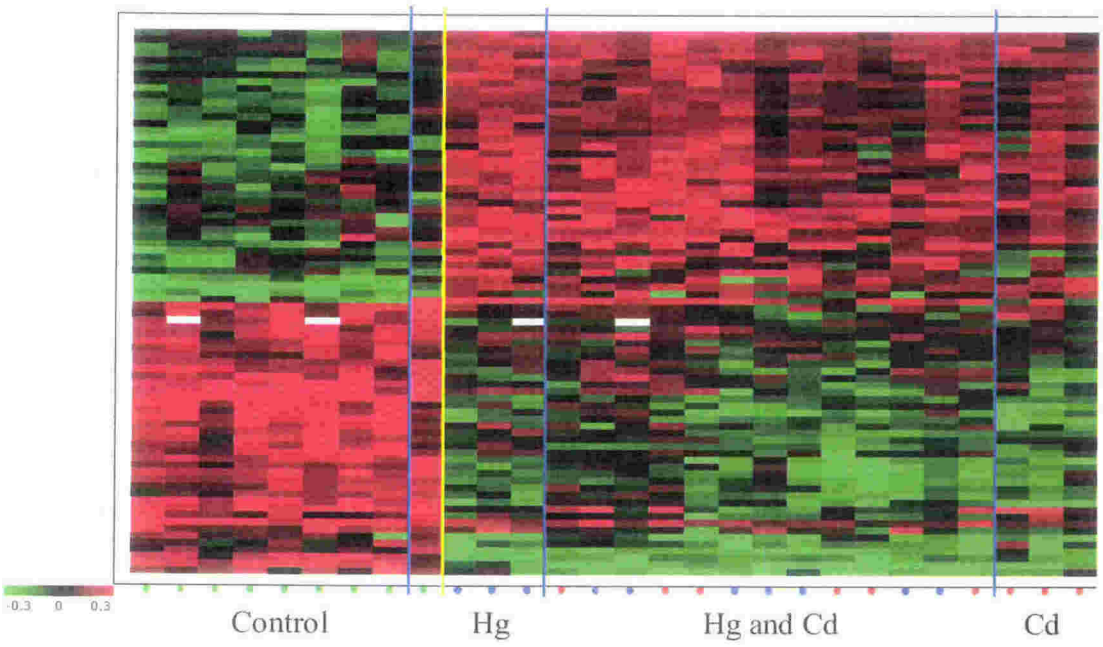


Figure 5.9: Gill pI 4-7 pattern analysis using significant ($p < 0.01$) proteins (n = 78)

Key:
- Control
- Hg
- Cd

Each dot represents an individual gel as indicated by the key.
The yellow and blue lines have been added.

PCA analysis separated the spot maps into two major groups in a similar manner to the pattern analysis (Figure 5.10; panels B and C). The control samples formed a single cluster (green ellipse) that was well separated from metal treated samples (red ellipse). However, Hg and Cd spot maps could not be clearly separated from each other. Although Hg samples formed a tighter cluster than Cd samples, they had overlapping distributions (Figure 5.10, panel C). In panel A, PC1 describes 57.8% of the total protein variation and PC2 an additional 11.6%, a total of 69.4%. In panel C, PC1 describes 60.5% of the total spot map variation, and PC2 describes an additional 6.8%, a total of 67.3%. The proteins outside the 95% significance circle in panel A (black ellipse) could be considered as outliers because they fall outside of the 'normal' range. These outliers fit into 2 categories: 1) highly significant spots with large average ratio values; e.g. protein numbers (average ratio value): 1439 (-2.01), 2494 (1.86) and 2) small spot volume; 1224 (-1.61), 1350 (-1.91). All of these spots have been checked and have been retained for analysis

because matching within BVA is correct, supporting the idea that the differences in the abundance of these spots are real.

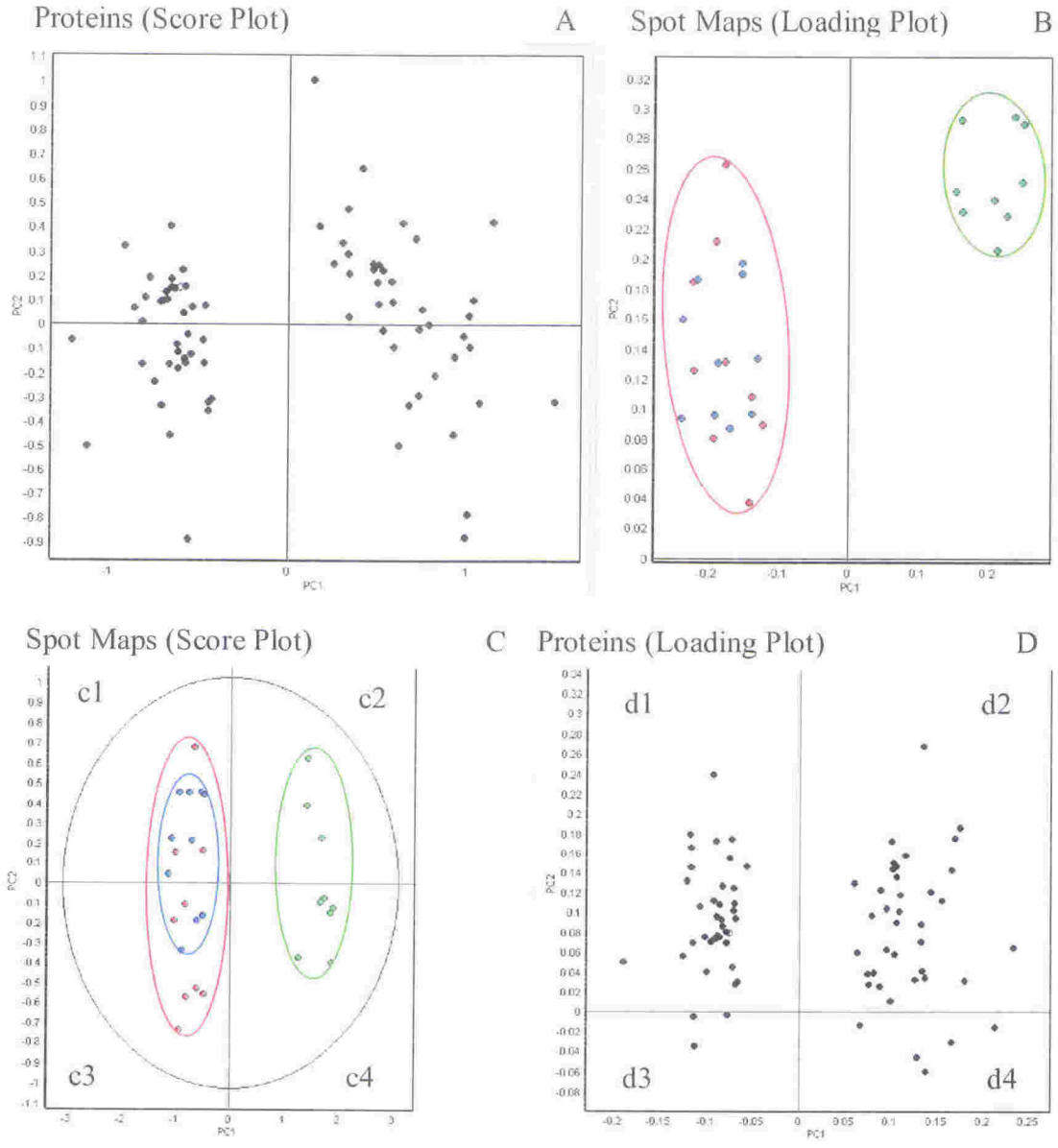


Figure 5.10: Gill pI 4-7 PCA using statistically significant ($p < 0.01$) proteins

Key:

- Control
- Hg
- Cd

The black dots in panels A and D represent individual proteins (n = 78), whereas the coloured dots in panels B and C represent spot maps (n = 28).

5.3.3.2 Gill tissue pI 6-11

In gill tissue pI 6-11, 33 proteins exhibited significant ($p < 0.01$) differences in protein abundance between one or more experimental groups (Table 5.6). This was approximately 3% of the total number of protein spots separated in this dataset (approximately 1200 spots in total). The locations of proteins exhibiting significant differences in abundance are shown in Figure 5.11. The numbering on the spot map corresponds with the numbers shown in the 'Master No.' column in Table 5.6. The larger number of proteins shown in the total column in Table 5.7 (n = 49) than Table 5.6 (n = 33) relates to the observation that the abundance of several proteins changed significantly in more than one experimental comparison (Figure 5.12). There were no significant variations in the abundance of proteins in comparisons between Cd and Hg when the FDR adjustment was applied to the data. Protein spot 1223 is low molecular weight and has a pI which may correspond to MT.

Table 5.6: Gill pI 6-11 proteins which varied significantly ($p < 0.01$)
* Proteins spots that were manually added or manipulated

Pos.	Master No.	Group Comparison	T-test	Av. Ratio	1-ANOVA
1*	87	Control vs. Cd	0.0051	-1.75	0.0084
		Control vs. Hg	0.0062	-1.79	
2*	88	Control vs. Cd	0.0025	-2.08	0.00082
		Control vs. Hg	0.00064	-2.24	
3	269	Control vs. Cd	0.0019	2.23	0.0034
		Control vs. Hg	0.0062	1.60	
4	319	Control vs. Cd	0.0034	-1.86	0.0077
		Control vs. Hg	0.0019	-1.88	
5	335	Control vs. Cd	0.0017	-1.96	0.012
6	339	Control vs. Cd	0.0025	1.68	0.0016
		Control vs. Hg	0.0019	1.59	
7	361	Control vs. Hg	0.00063	-1.64	0.00069
8	362	Control vs. Cd	0.00034	-1.74	4.8e-005
		Control vs. Hg	0.00028	-1.87	
9*	385	Control vs. Cd	0.0056	-1.58	0.015
		Control vs. Hg	0.0062	-1.58	
10*	423	Control vs. Hg	0.00064	-1.66	0.015
11	485	Control vs. Cd	0.0020	1.59	0.0047
12	489	Control vs. Hg	0.0020	-1.56	0.0029
13*	516	Control vs. Cd	0.0080	1.69	0.015
14*	535	Control vs. Cd	0.00047	-2.66	0.00013
		Control vs. Hg	0.00020	-2.45	
15	626	Control vs. Cd	0.0017	-2.54	0.012
16*	641	Control vs. Hg	0.0097	-2.61	0.014
17	643	Control vs. Cd	0.0045	1.50	0.0047
18	651	Control vs. Cd	0.00090	1.61	0.00030
		Control vs. Hg	0.00064	1.75	

19	660	Control vs. Cd	0.0016	-1.64	0.017
20	675	Control vs. Cd	0.0079	1.56	0.015
21*	698	Control vs. Cd	0.00018	-2.49	4.6e-005
		Control vs. Hg	0.00015	-1.93	
22*	702	Control vs. Cd	7.3e-005	-2.24	1.2e-006
		Control vs. Hg	3.1e-005	-1.92	
23	704	Control vs. Cd	0.0028	-1.64	0.0029
		Control vs. Hg	0.0014	-1.65	
24	714	Control vs. Cd	0.00018	-1.67	8.1e-006
		Control vs. Hg	0.00020	-1.52	
25*	874	Control vs. Cd	0.0074	-1.51	0.012
		Control vs. Hg	0.0039	-1.53	
26*	903	Control vs. Hg	0.0052	-1.56	0.029
27	932	Control vs. Cd	0.0073	-1.57	0.025
28*	937	Control vs. Cd	0.0019	1.80	0.00094
		Control vs. Hg	0.0017	2.25	
29	952	Control vs. Hg	0.0012	1.54	0.0026
30	979	Control vs. Hg	0.0031	-1.51	0.0041
31	982	Control vs. Cd	7.3e-005	-1.73	1.2e-006
		Control vs. Hg	3.1e-005	-1.96	
32	1223	Control vs. Cd	0.0080	-2.22	0.098
33	1241	Control vs. Hg	0.00064	-1.60	0.0018

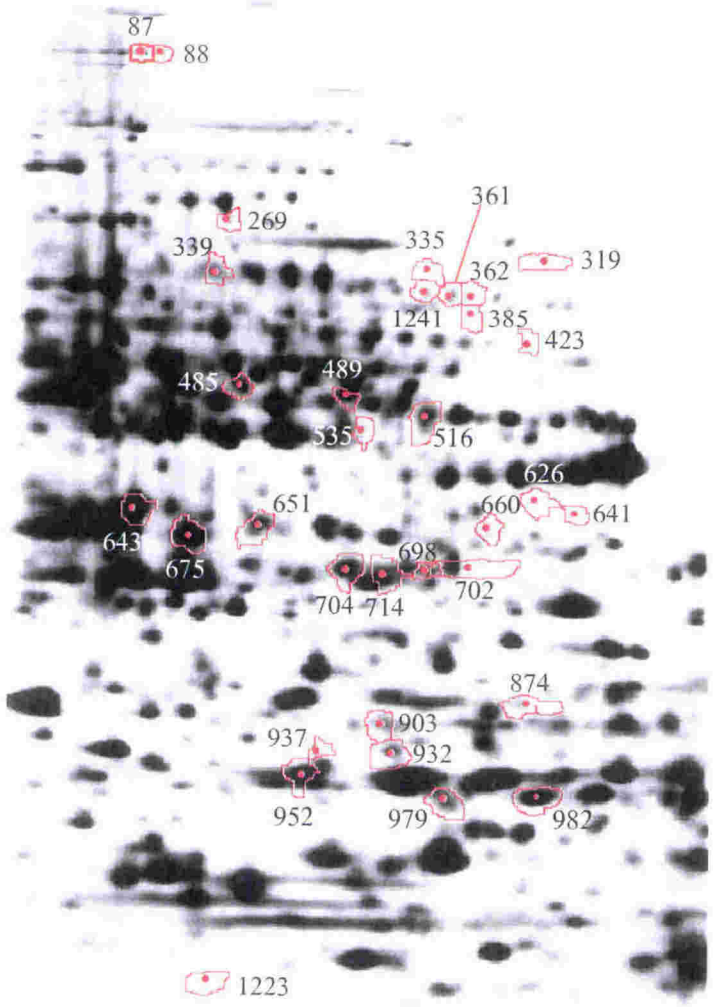


Figure 5.11: Gill pI 6-11 gel showing significant ($p < 0.01$) protein spots

Group Comparison	# spots increased ↑			# spots decreased ↓			TOTAL
	1.5+	2.0+	Total	-1.5+	-2.0+	Total	(approx. 1400 spots)
Control vs. Cd	8	1	9	10	6	16	25
Control vs. Hg	3	2	5	16	3	19	24
TOTAL	11	3	14	26	9	25	49

Table 5.7: Gill pI 6-11 summary of significant (< 0.01) results

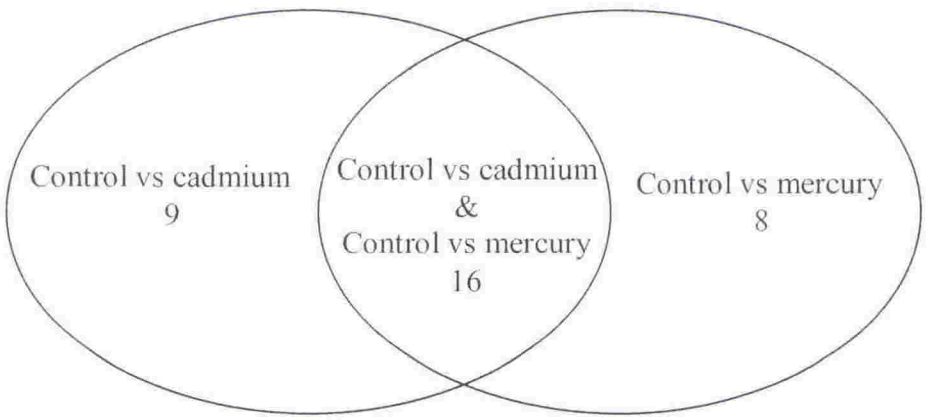


Figure 5.12: Venn diagram illustrating the distribution of significant ($p < 0.01$) protein spots and their overlap between treatments in gill tissue pI 6-11

Pattern analysis split the spot maps into 2 major groups; control and metal exposed (Cd and Hg), shown on either side of the yellow line (Figure 5.13). Within the Cd and Hg treated spot maps there were two distinct sub-groupings, indicated by the blue line. The major sub-group contains samples treated with Cd and samples treated with Hg (i.e. a mixture of both experimental classes), to the right of this, is a group of two Cd exposed gels. Within the control samples, one control spot map appears to be an outlier (to the right of the blue line).

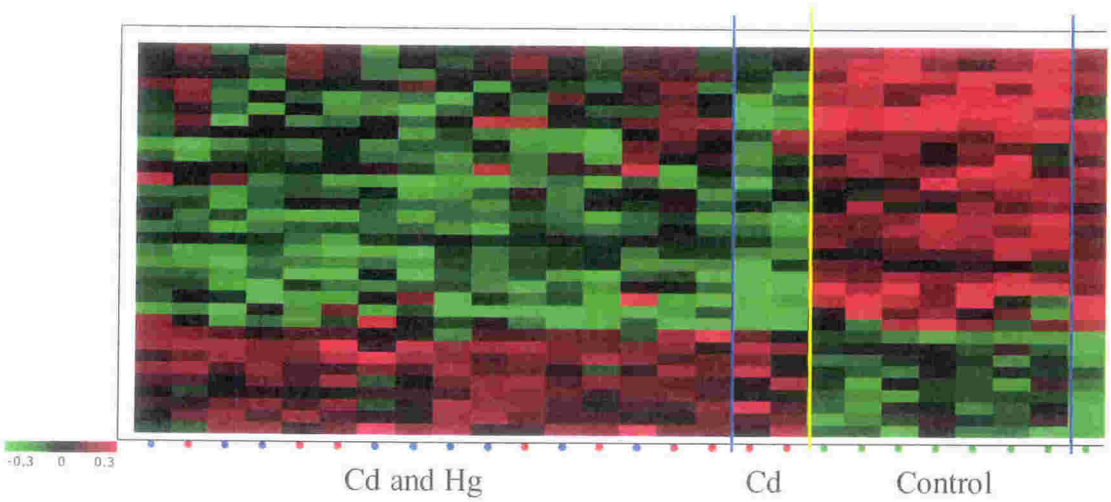


Figure 5.13: Gill pI 6-11 pattern analysis using significant ($p < 0.01$) proteins (n = 33)

Key:

- Control
- Hg
- Cd

Each dot represents an individual gel as indicated by the key. The yellow and blue lines have been added.

PCA analysis separated the spot maps into two major groups in a similar manner to the pattern analysis (Figure 5.14; panels B and C). Control samples (green ellipse) are clearly separated from metal treated samples (red ellipse). In panel A, PC1 describes 48.9% of the total protein variation PC2 an additional 12%, a total of 60.9%. In panel C, PC1 describes 57.4% of the total spot map variation and PC2 an additional 9%, a total of 66.4%. In panel C, one Cd spot map appears to be an outlier (indicated by the red arrow). The protein outside of the 95% significance circle in panel A (black ellipse) is spot 1223 (average ratio -2.22).

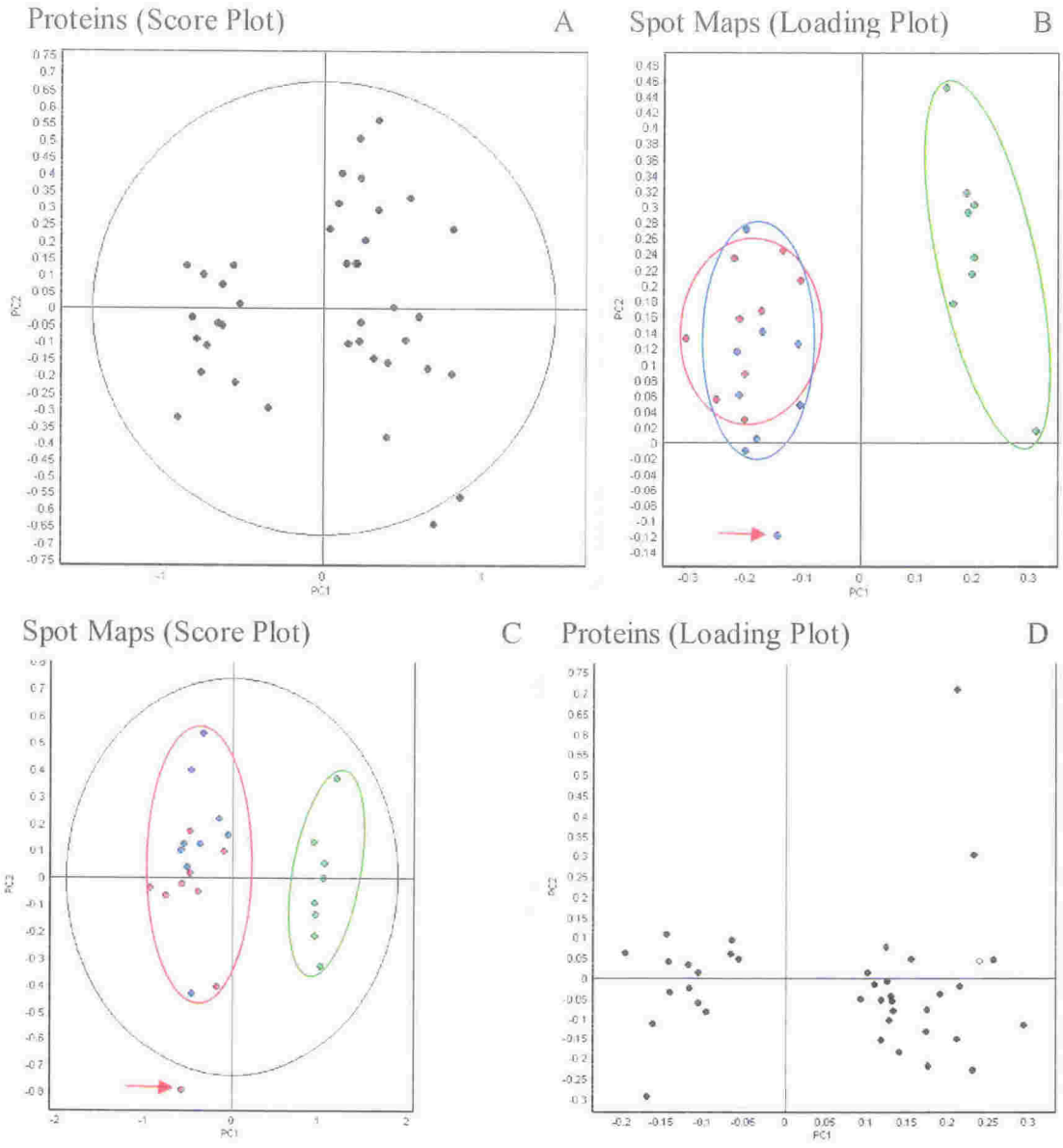


Figure 5.14: Gill pI 6-11 PCA using significant ($p < 0.01$) proteins

Key:

- Control
- Hg
- Cd

The black dots in panels A and D represent individual proteins (n = 33), whereas the coloured dots in panels B and C represent spot maps (n = 26). The red arrow indicates a spot map outlier.

5.3.4 Separation and identification of MT

MT was visualised on 1-D gels using an Ikzus MT fluorescent labelling kit (Figure 5.15). Rabbit MT was used as a positive control and these low molecular weight bands (lanes 9 and 10) were cut out for identification using MALDI mass fingerprinting. These bands were identified as MT (Figure 5.16 and 5.17), although the top match was for human rather than rabbit MT. The low molecular weight

mussel protein bands in lanes 1-8 were also cut out and prepared for MALDI mass fingerprinting. In contrast to the positive control, where a single clear band could be visualised (B: lanes 9 and 10), multiple low molecular weight bands appeared in the mussel samples (B: lanes 1-8). There were no noticeable differences in band intensities between samples exposed to metals and those that were not. Low molecular weight bands were cut out (as indicated in B) for MALDI mass fingerprinting, but none of these bands could be matched with database sequences.

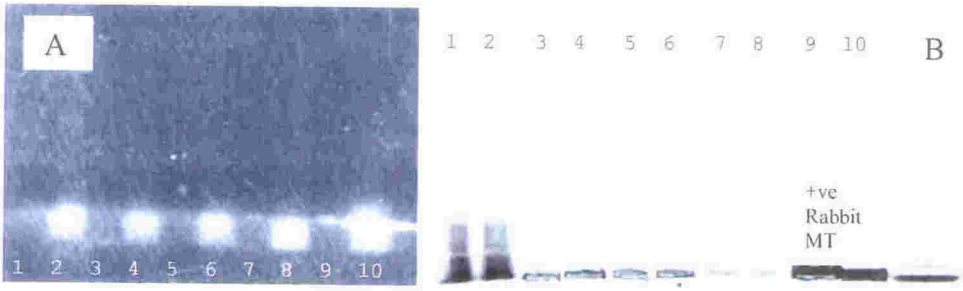


Figure 5.15: MT visualised under UV (A) and after staining with Coomassie G-250 (B)

Lanes 1 and 2 = control digestive gland tissue; 3 and 4 = Gill + Cd; 5 and 6 = Gill + Hg; 7 and 8 = control gill; 9 and 10 = positive control, Rabbit MT. Lane 11 = Ladder. Note that only samples in the even numbered lanes were labelled. The odd numbered lanes contain the same sample without the fluorescent label added.

ProFound - Search Result Summary

Version 4.10.5
The Rockefeller University Edition

Protein Candidates for search B70BC0F6-04C0-787CA51B [1970469 sequences searched]

Rank	Probability	Est'd Z	Protein Information and Sequence Analyse Tools (T)	%	pI	kDa	®
+1	1.0e+000	1.69	T gi|10835230|ref|NP_005941.1| metallothionein 1G [Homo sapiens]	62	10.0	7.19	®
+2	1.7e-006	-	T gi|127400|sp|P18055|MT2A_RABIT Metallothionein-IIA (MT-2A)	34	9.9	7.21	®
+3	1.2e-006	-	T gi|18202363|sp|P79376|MT1C_PIG Metallothionein-IC (MT-IC)	34	10.6	7.08	®
+4	2.4e-007	-	T gi|127385|sp|P11957|MT1A_RABIT Metallothionein-IA (MT-1A)	46	10.0	7.23	®

Figure 5.16: ProFound search summary showing a significant match for MT

ProFound - Search Result Details

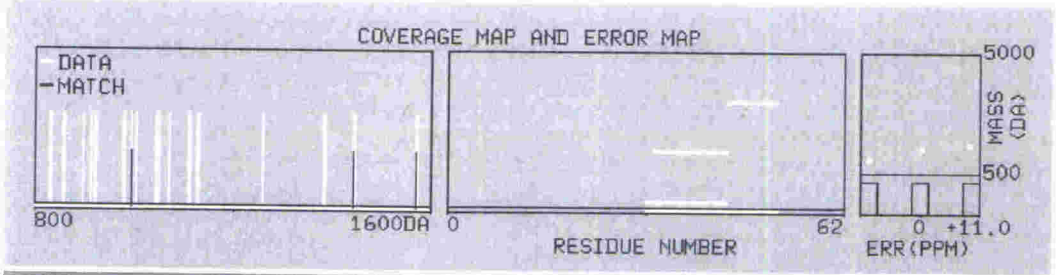
Version 4.10.5
The Rockefeller University Edition

Details for rank 2 candidate in search B70BC0F8-04C0-787CA51B

gi|127400|sp|P18055|MT2A_RABIT Metallothionein-IIA (MT-2A)

gi|1363074|pir||S54336 metallothionein-2a - rabbit

Sample ID : [Pass:0]
Measured peptides : 40
Matched peptides : 3
Min. sequence coverage: 34%



Measured Mass(M)	Avg/ Mono	Computed Mass	Error (ppm)	Residues Start	Residues To	Missed Cut	Peptide sequence
995.389	M	995.398	-9	45	52	0	CAQCCICK
1442.486	M	1442.486	1	33	44	0	SCCSCCPPGCAK
1570.595	M	1570.581	9	32	44	1	KSCCSCCPPGCAK

Figure 5.17: Sequence coverage of rabbit MT

5.4 Discussion

5.4.1 Tissue specific metal accumulation patterns

P. canaliculus tissue accumulated both Cd and Hg to high concentrations under laboratory conditions consistent with worldwide studies on metal accumulation in other marine organisms, in both laboratory (Allah et al. 2003; Wang and Rainbow 2005) and natural settings (Hung et al. 2001; Nicholson and Szefer 2003). There were pronounced time and tissue dependent differences in the uptake of Cd and Hg into *P. canaliculus* gill, digestive gland, foot and the soma. Gill and digestive tissue for example, concentrated heavy metals to a greater extent than the other tissues in accord with most previous studies (e.g. Janssen and Scholz 1979; Everaarts 1990). Gills were directly exposed to metals through a relatively large surface area and are known to function in metal uptake and storage (Bebianno and Serafim 1998), whereas digestive tissue is considered to be an important target in long-term exposure because it is a possible site for MT biosynthesis (Serra et al. 1999). In the current study digestive gland accumulated the highest levels of Cd, but in field studies of *M. galloprovincialis* the highest Cd concentrations were obtained from gill tissue (Serra et al. 1999). In other laboratory studies the highest concentration of Cd accumulated in the digestive tissue of *M. edulis* (Janssen and Scholz 1979; Everaarts

1990) and *P. viridis* (Yap et al. 2003c) in agreement with the current study, suggesting that the genus *Perna* utilises a similar Cd uptake route as *Mytilus*.

In *P. canaliculus* Hg accumulation was highest in gill tissue in agreement with studies conducted on other molluscs including *M. edulis* (Roesijadi 1982) and the hydrothermal vent mussel *B. azoricus* (Kádár et al. 2004). The concentrations of Cd and Hg in control mussels, accumulated naturally prior to the start of the experiment, were highest in the digestive gland for both metals (Table 5.3). Overall, these results suggest that gill is important in the initial uptake of Hg, but during long-term naturally occurring low levels of exposure, Hg and Cd are stored and accumulated in the digestive gland.

5.4.2 Time dependent metal accumulation

Hg concentrations in all tissues peaked after 48 h. By the end of the experiment (72 h), the Hg concentrations in all tissues had decreased and were below the concentration that was accumulated in the first 24 hr (although still above pre-experiment Hg concentrations). This is in contrast to previous studies in which mussels continued to accumulate Hg in soft tissues until the depuration phase when mussels were transferred to clean water (e.g. Kádár et al. 2004). Previous studies added Hg to the experimental tank throughout the experiment to maintain an approximately constant Hg concentration. In the current study mussels were exposed to a one-off spike of Hg at the beginning of the experiment. Cd was also only added to the tank at the beginning of the experiment, but in contrast to Hg, Cd concentrations continued to increase in digestive tissue and were still increasing at the end of the experiment (72 h). This time dependent difference in accumulation pattern between the two metals could be due to the large difference in the initial concentrations of Cd and Hg dissolved in seawater (Cd was 16x more concentrated). Cd levels in seawater decreased at a steady rate (Figure 5.2) but 43% of the Cd remained after 72 hr. The initial drop in Hg concentration in seawater was rapid, 98% within 24 hr, so there was very little Hg left to accumulate after this time and the mussel could focus on eliminating this metal.

The time-dependent differences in metal accumulation could be related to the elimination mechanisms of Cd and Hg from *P. canaliculus*. In general Cd has a relatively long half-life in bivalves, estimated at 96 to 190 d (Borchardt 1983) and 300 d (Bebianno and Langston 1993) in *M. edulis*, 1254 d in *M. galloprovincialis*

(Taylor 1983) and 70 d in the clam *Macoma balthica* (Langston and Zhou 1987). For comparison, the biological half-life of inorganic Hg in gills of the hydrothermal mussel *B. azoricus* was less than 21 d (Kádár et al. 2004) and estimates of 9 to 17 d were obtained from the oyster *C. virginica* (Cunningham and Tripp 1975). In the current study it only took 24 hr for the concentration of Hg in whole tissues to decrease to less than half of the maximum concentration obtained after 48 hr. A depuration period was not included in the current study but results suggest that the half-lives for Cd and Hg in *P. canaliculus* follow trends which have been observed in other studies i.e. a longer half-live for Cd than Hg.

Bioaccumulation rates and bioconcentration factors (BCF) calculated in the current study for Cd and Hg (Table 5.3) provide preliminary metal accumulation information on a local species. The whole tissue BCF for Cd was approximately 15, which is within the lower range for molluscs exposed to a similar concentration of Cd as illustrated in Taylor (1983). The maximum bioconcentration factor for Hg was 28.3 times but a 'steady state' BCF was not reached. Comparison with field information is discussed in Chapter 6.

5.4.3 Other metals

Additional insight into the physiological functioning of metal homeostasis in *P. canaliculus* could have been gained by including essential metals in this study. However the major focus of this research was toxic metals, with the aim of exploring the biomonitoring suitability of *P. canaliculus*. Essential metals (e.g. Cu, Zn) are regulated to some extent in molluscs (Chong and Wang 2001). Although Zn can be toxic at high levels (Abel 1976), previous studies indicate that mussels have the capability to regulate Zn. For example in *P. viridis*, when clean seawater is restored, after laboratory or transient local exposure, Zn quickly returns to baseline concentrations (Yap et al. 2003c). Cu is also thought to be regulated to some extent but is subject to severe fluctuations due to environmental conditions, making it unsuitable for biomonitoring (Phillips 1976).

5.4.4 Biomonitoring implications

The observed time and tissue dependent differences in Cd and Hg accumulation in *P. canaliculus* may have implications for the development of biomonitoring programmes in New Zealand. Based on the information gathered

during laboratory experiments digestive gland would be the most contaminated tissue during the initial uptake phase of a Cd spill, whereas gill would initially accumulate the highest levels of Hg. It is difficult to make a recommendation on which of these tissues should be used for monitoring programmes because although using both tissues would provide valuable information on metal accumulation patterns, dissecting both tissues for analysis is time consuming and expensive. Additional factors which are known to influence metal uptake include reproductive cycle, temperature, food availability and metabolic regulation (Isani 2000), the effects of which are not equal across tissues. For example digestive gland mass (and associated metal concentration) is influenced by the growth of gonad tissue whereas gill mass remains fairly constant throughout the reproductive cycle (Dragun 2004).

The tissue specific responses encountered using individual tissues for metal accumulation analysis could be avoided by analysing whole mussel tissues. But because gill and digestive gland are a relatively small proportion of the entire organism, whole tissues may not be as sensitive to low levels of heavy metals. The sensitivity issue is unlikely to be problem because metal concentrations that are below the level of detection in whole mussel tissues are probably insufficient to have an effect on the organism or human consumers. Whole body concentrations as low as 0.005 mg kg^{-1} can be detected for Cd and 0.05 mg kg^{-1} for Hg, whereas relevant health limits imposed by the Food Safety Authority New Zealand (FSANZ) are 2.0 mg kg^{-1} and 0.5 mg kg^{-1} respectively (see Chapter 7 for further details). Alternatively, previous studies have suggested that byssus could be an effective monitoring tissue (Nicholson and Szefer 2003). Not only does byssus accumulate high levels of metals, it is easy to handle, can be pulled out of the foot before or after the death of the mussel and can be handled like a human hair with no need for freezing (Yap et al. 2003b). In *P. viridis* gills and byssus were the best tissues for the monitoring Pb (Yap et al. 2004b). Unfortunately, there are too few studies on byssus to date (Nicholson and Szefer 2003; Yap et al. 2004b) to definitively establish the relationship between metal concentrations in the byssus and the water column, therefore its usefulness for monitoring, have yet to be conclusively established.

It is probable that once mussels are attached via byssal threads to solid substrate in the field that additional byssal threads are not continuously produced, therefore different individual byssus threads from different individual mussels may contain diverse temporal information about contaminant exposure. Intuitively, after

a one-off contamination, newly produced byssus near the foot gland would contain the largest amounts of metal, which may be quite different to the levels of metal detected in byssus produced under continual low background levels of metals. In laboratory based experiments the removal of existing byssus before metal exposure may overexpose mussels to metals while the mussels attempt to reattach (Rajagopal et al. 2005). Byssus was collected in the current experiment, but has not been analysed yet, so it is unknown whether byssus could be a good biomonitoring tool in *P. canaliculus*.

5.4.5 Proteomics analysis

Over 100 proteins were detected which showed significant ($p < 0.01$) differences in abundance between control and metal treated groups but these proteins could not be identified using MALDI-TOF fingerprinting or tandem mass spectrometry. Although identification of proteins from invertebrates is difficult, it is not impossible and many previous studies have successfully identified proteins from mussels and other invertebrates (e.g. Olsson et al. 2004; McDonagh et al. 2005). The factors which may have contributed to the poor protein identification success rate in this study have been discussed previously (Chapter 4, Section 4.4.2).

PCA and pattern analysis suggest that Hg and Cd have similar effects on protein abundances because the majority of proteins which showed significant changes in protein abundance were affected by both metals. Although metal treated groups were clearly separated from control samples, there was no clear separation between Cd and Hg treated samples. Hg appeared to have a slightly greater effect on protein abundance in gill tissue than Cd, because 31 proteins were significantly affected by Hg and not Cd, compared to only 19 proteins that were exclusively affected by Cd. Previous studies have proposed that Hg and Cd are processed in similar ways (Janssen and Scholz 1979; Nolan 1984; Roesijadi 1986), so similar effects on protein abundances were not unexpected. However, this is the first time a proteomic approach has been used to illustrate the similar effects of these metals.

Although a potential MT protein (~10 kDa and pI 6-8) was detected (Figure 5.11; protein 1223) it could not be positively identified as MT. Among molluscs 17 MT isoforms have been characterised in the blue mussel *M. edulis* (Mackay et al. 1993; Barsyte et al. 1999; Leignel et al., Genbank 2003, unpublished) 2 in the tropical green mussel *P. viridis* (Khoo and Patel 1999), 1 in the freshwater mussel *D.*

polymorpha (Engelken and Hildebrandt 1999), 16 in the eastern oyster *C. virginica* (Tanguy et al. 2001; Tanguy et al., GenBank 2002, unpublished) and 5 in the Pacific oyster *C. gigas* (Tanguy and Moraga 2001) which should provide adequate sequence similarities for matching peptide fingerprints. There may be a technical reason for the failure to detect MT in the current study (e.g. insufficient protein), or the protein detected in Figure 5.11 may not be MT. Previous proteomic studies have not identified MT from 2D gels, although previous studies have not used basic (pI 6-11) gels (e.g. Rodríguez-Ortega et al. 2003a; Slivestre et al. 2006).

The company that supplied the MT labelling kit (Ikzus) had not previously attempted to visualise MT on 2-D gels. However, in the current study rabbit MT was separated on 2-D gels and identified successfully, suggesting that MT could be isolated and identified using this technique. Identification of mussel MT from 2-D gels was difficult because the buffers supplied with the kit gave variable results. Also the 2-D patterns obtained using the buffer contained in the kit were different to those separated using standard DIGE buffers, so it was difficult to correlate the location of fluorescently labelled MT with the possible location on gels prepared using DIGE buffer and stained using Coomassie G-250. Despite these difficulties, MT was identified from a gill preparation labelled with fluorescent dye. Unfortunately, these data were lost during an unscheduled disk failure. Despite following the same preparation procedure the result could not be replicated, highlighting the variability of this system. Attempts were made to identify spots isolated from both 1-D and 2-D separation of rabbit MT and mussels using tandem mass spectrometry. However, no identifications were able to be made. There is limited evidence to suggest that traditional low molecular weight MT proteins were detected using proteomics in the current study. A single low molecular weight protein in the correct pH range (6-11) was detected as showing significant differences in protein abundance (Figure 5.11; protein 1223) but this protein was not able to be identified

An enzyme-linked immunosorbent assay (ELISA) test used to quantify MT protein expression in *O. edulis* exposed to metal under laboratory conditions showed no significant induction of MTs in gills or digestive gland, despite an increase in metal concentration observed in those tissues (Tanguy et al. 2003). In *M. edulis*, 50% of the cellular Cd was associated with particulate material, and the soluble metal was initially associated with a component of approx 60 kDa and later with MT

(Carpene and George 1981). MT may not be the main mechanism for coping with heavy metals in all organisms. Intracellular binding of Cd might be a 2 step process with initial binding to a high molecular weight protein and, after a lag period, bind[illegible]

CHAPTER 6: Protein abundance patterns in *P. canaliculus* exposed to naturally occurring contaminants in the Bay of Islands, Northland, New Zealand

Abstract

P. canaliculus protein abundance patterns produced by exposure to heavy metal and other contaminants in the field (Bay of Islands) were compared with those determined in the laboratory (Chapter 5) using 2D-DIGE. Differences in protein abundance were tested by permutation analysis. Approximately 30 protein changes were detected which appeared to be associated with the presence of heavy metal under both field and laboratory conditions. However, the control site from the Bay of Islands was unexpectedly found to contain relatively high concentrations of Cd and As, so that the association of these proteins with contaminants could not be unequivocally established. Although further work is required to validate these proteins, this study suggests that biomarkers for heavy metal pollution in the field could be developed for *P. canaliculus*.

6.1 Introduction

6.1.1 Environmental Proteomics

The concept of environmental proteomics has already been outlined (Chapters 4 and 5). Briefly, environmental proteomics assesses an organism's ability to respond to various stress inducers (e.g. temperature, salinity, heavy metals) by analysing the protein complement of the organism's genome (i.e. the proteome). Protein expression profiles produced by exposed experimental organisms can be compared with profiles from control organisms that have not been subject to stressful conditions, to identify proteins involved in various 'stress responses'. In Chapter 5, this technique was used to detect proteins in *P. canaliculus* which responded to Cd and Hg under laboratory conditions. Possible biomarkers for heavy metal pollution were detected using this methodology but it is important to test the suitability of these proteins under natural conditions (Petrovic et al. 2001; Mourgaud et al. 2002). In the current study, the proteins which responded to heavy metal under laboratory conditions were examined under field conditions, using *P. canaliculus* exposed to

naturally occurring low levels of metal contamination from 4 sites within the Bay of Islands.

6.2 Materials and Methods

6.2.1 Mussel and water sampling

Adult greenshell mussels (6-15 cm) were collected in December 2005 from 4 locations within the Bay of Islands, New Zealand (Figures 6.1 and 6.2). Urapukapuka-Rawhiti (abbreviated to Urapukapuka from this point forward) is an offshore island approximately 15 km away from the other sites, which are in close proximity to the Veronica Channel. Mussels from Urapukapuka were selected as controls because it was assumed that mussels from this site have not been exposed to elevated concentrations of heavy metals from industry or boating activities. Opua Marina and Opua Wharf were expected to be subject to pollutants from boating activities, whereas Waitangi River was thought to be influenced by geothermal inputs from Ngāwhā. Proteomic analysis was conducted on 6 individual mussel gill samples collected from each of the 4 locations (n = 24).

Six mussel samples from each site were sent to ELS for total tissue chemical analysis of Hg, Cd, As, Pb and Tin (Sn). Whole mussels were analysed i.e. mussels were not dissected into individual tissues as described previously (Chapter 5). Hg, Cd and As were analysed because these elements are associated with geothermal activity. Pb was expected to act as a control metal (as in Chapter 5), although Pb paint could be associated with boating activity. Tributyltin (TBT) is an antifouling coating which has been banned in many countries since the early 1990's but could still be associated with boating activity. Total tin (including TBT) concentrations were also tested.

A single water sample was collected from each location and was analysed by Hill Laboratories for the presence of the same heavy metals as listed above. Water and mussel samples were collected concurrently (Dec 2005) from Urapukapuka and Opua Marina. The remaining water samples were collected in mid-January 2005 from Waitangi Estuary and Te Haumi Estuary, locations which are in very close proximity to the mussel sampling locations (Waitangi Bridge and Opua Wharf) (Figures 6.1 and 6.2).

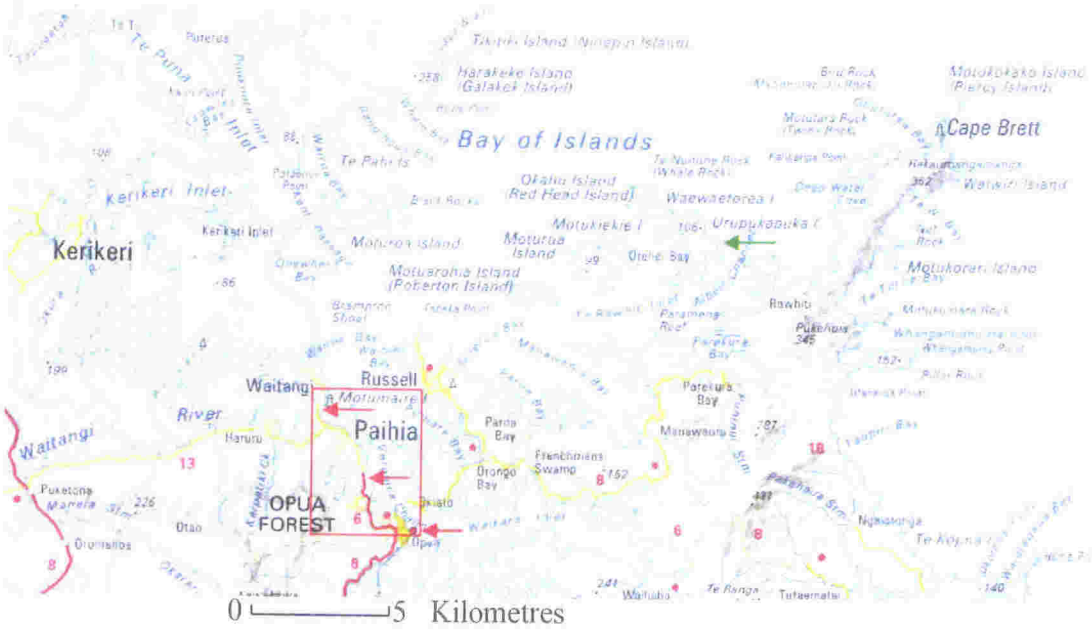


Figure 6.1: Map of the Bay of Islands
Sampling locations are indicated by arrows. The red rectangle is enlarged in Figure 6.2. The approximate sampling location for Urapukapuka-Rawhiti (the clean site) is indicated by the green arrow. Map reproduced from the Reed New Zealand Atlas (Dowling et al. 2004).

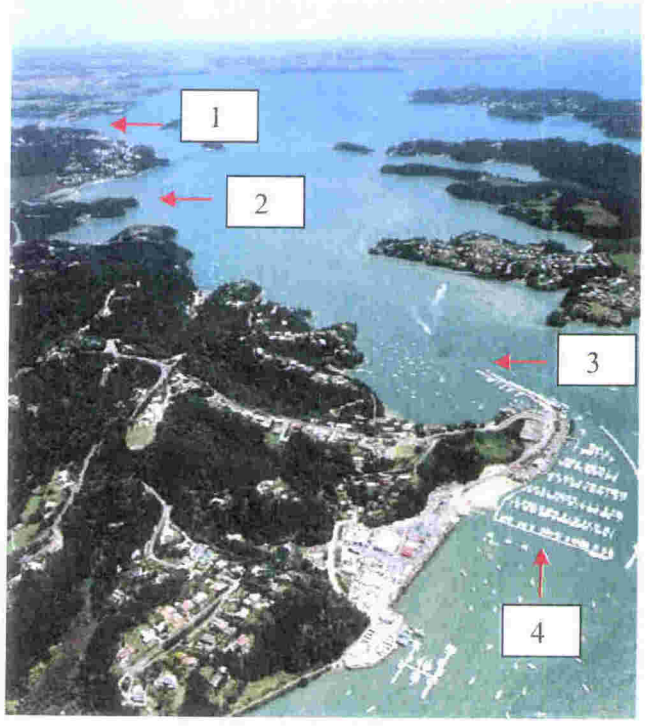


Figure 6.2: Enlarged view of sampling locations
(1)Waitangi Bridge/Estuary, (2) Te Haumi Estuary (Water sample only), (3) Opua Wharf (4) Opua Marina

6.2.2 PERMANOVA and Principal Component Analysis

Permutational multivariate analysis of variance was conducted on Bray-Curtis dissimilarity values using PERMANOVA v1.6 (Anderson 2001; McArdle and Anderson 2001) to determine whether the metal concentrations were significantly different between locations (Chapter 3). Data was not transformed or standardised. The number of permutations used was 999 and the integer used as a seed was 6. To get an overview of the Bay of Islands data, a principal component analysis was performed and the results have been displayed as a Biplot. PRIMER v5 (Colwell 2005) was used to create non-Metric Multi Dimensional Scaling (nMDS) plots based on Bray-Curtis dissimilarity values using non-standardised and non-transformed data (i.e. raw data). The MDS plots arrange the samples obtained from each location in 2D space to reproduce the observed dissimilarity distances i.e. they are a visual representation of the similarity/dissimilarity among sites. The amount of metal contained in each sample is also indicated by the size of the sphere used in the MDS plot.

6.2.3 Comparisons between Wellington and Bay of Islands data

In the current study whole mussels were analysed for the presence of heavy metals, but in the laboratory experiment (Chapter 5) gill, digestive gland, foot and remaining soft tissue were dissected from Wellington mussels and analysed individually. To enable the comparison of metal content from Wellington and Bay of Islands mussels, the metal concentrations in samples from Wellington were recalculated for the whole animals based upon the relative proportion of each tissue to the mussel. Whole body concentrations of Cd, Hg and Pb for Wellington control mussels were determined at each time point (Time 0, 24, 48, and 72) and then averaged to give a single value. As, Cu and Zn concentrations are an average of a single time point (Time 0) because additional measurements were not made during the course of the experiment.

6.2.4 Bioconcentration Factor (BCF)

BCF is used to describe the accumulation of chemicals in aquatic organisms exposed to contaminated environments. A steady-state BCF occurs when the organism is exposed for a sufficient length of time so that the BCF does not change

substantially. The bioaccumulation factor was calculated to allow comparison of the BCFs from field samples from the Bay of Islands with BCFs obtained from laboratory experiments. The BCF was calculated according to Taylor (1983):

$$BCF = \frac{Ce - Ci}{Cs}$$

where: Ce = metal concentration in the mussel tissue during exposure
Ci = initial metal concentration in the mussel tissue before exposure
Cs = experimental metal concentration in the test seawater

6.2.5 Protein Extraction

Gill proteins were extracted from approximately 80-140 mg of tissue using 100 µl of standard DIGE lysis buffer (30 mM TrisCl, 7 M urea, 2 M thiourea, 4% CHAPS). After homogenisation a variable amount of additional lysis buffer was added to ensure that all of the samples were approximately the same protein concentration (10 µg per µl). Protein was purified before labelling using a 2-D clean up kit (GE Healthcare).

6.2.6 Pooling samples for DIGE labelling

Protein spot abundances in mussels from the Bay of Islands were compared with those obtained from Wellington to identify possible biomarkers of heavy metal pollution. The number of samples within this experiment (n = 54) was simplified by combining samples from within each group (Table 6.1) to create seven pooled samples. Pooled samples were formed by mixing together equal amounts of protein from each individual within a group i.e. Urapukapuka, Opua Marina, Opua Wharf and Waitangi Bridge from Bay of Islands and control, Cd and Hg from Wellington. Pooling samples not only significantly cuts down cost and preparation time, but it also simplifies analysis by reducing the number of gels that need to be matched. There is however a loss of power compared with analysis of multiple individuals. Gill proteins in these pooled samples were labelled according to the manufacturer's directions as described previously (Chapter 3). A pooled internal standard comprising all the samples in the experiment (n = 54 individuals) was labelled with Cy2. Each pooled group was labelled with both Cy3 and Cy5 (Table 6.1).

6.2.7 Sample preparation for 2-DE

Samples labelled with different CyDyes™ were combined in a microcentrifuge tube (Table 6.1).

Gel #	Cy2	Cy3	Cy5
1	STD	C1	WB
2	STD	Cd	OW
3	STD	OM	Hg
4	STD	URA	Cd
5	STD	OW	C1
6	STD	WB	OM
7	STD	Hg	URA

Table 6.1: Labelling strategy for the Bay of Islands and Heavy Metal (Wellington) DIGE experiment.
STD = pooled internal standard (n = 54 individuals); Bay of Islands: URA = pooled Urapukapuka samples (n = 6 individuals); OM = pooled Opua Marina samples (n = 6 individuals); WB = pooled Waitangi Bridge samples (n = 6 individuals); OW = pooled Opua Wharf samples (n = 6 individuals); Wellington: C1 = pooled control samples (n = 10 individuals); Cd = pooled Cd samples (n = 10 individuals); Hg = pooled Hg samples (n = 10 individuals).

6.2.8 2-DE

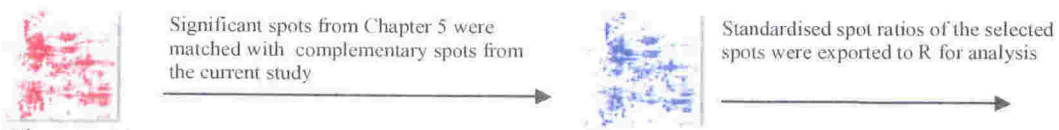
Immobiline™ dry strips of pI range 4-7 and 6-11 were used to separate gill proteins using an Ettan IPGphor Isoelectric focusing system and NuPAGE 4-12% Bis-Tris gels (Invitrogen) as described previously (Chapter 3). Gels were scanned using a Fuji film FLA-5000 scanner, converted into the appropriate file format and imported into the DeCyder workspace.

6.2.9 Protein Statistics

DeCyder analyses make a number of assumptions about data, such as an approximately normal distribution of protein abundance and independence of samples. DeCyder can handle slight deviations from 'normal' distributions because it uses log standardised protein abundances for statistical analysis (Chapter 3). DeCyder cannot currently analyse dependent data. The assumption of independence was violated in the current study because each pooled sample was used twice and labelled with Cy3 and Cy5 (Section 6.2.6 and Table 6.1). Discrepancies in the protein abundance values obtained from these two data points reflect differences in the labelling efficiency between the two dyes but not biological differences because the same sample is used. This reverse labelling procedure was standard practise, but the underlying statistical assumptions and the impact on experimental design in DIGE are now being questioned (Karp and Lilley 2005). For example, Karp et al. (2005) suggest that instead of having a single large pool, samples should be split in half e.g. if n =10, label 5 samples with Cy3 and label the other 5 samples with Cy5.

This slight modification ensures that DeCyder software can be used to analyse the data because the two pooled samples are independent. The current experiments were conducted prior to these publications. Therefore an alternative analysis procedure was developed in consultation with Dr Shirley Pledger (Victoria University of Wellington, School of Mathematics, Statistics and Computer Science).

Analyses for the current chapter were conducted using the statistical package R (R Development Core Team 2006). Gels from this study were matched with laboratory gels so that the analysis concentrated on those proteins that had been previously shown to respond to Cd or Hg (Chapter 5). Only proteins which demonstrated a statistically significant change in abundance with a small p value (< 0.01) and a large average ratio (> 1.5 and < -1.5) were reported in Chapter 5. However, the proteins picked for initial permutation analysis in the current study are based on a larger p value (< 0.05) and any average ratio (Figure 6.3). The larger set was used to allow for the detection of as many potential field biomarkers as possible. On pI 4-7 gels, 247 proteins showed statistically significant ($p < 0.05$) protein changes in the laboratory. Of those proteins 200 were able to be matched in the current study. On pI 6-11 gels, of 120 proteins which demonstrated significant ($p < 0.05$) differences in protein abundance in the laboratory 77 were able to be matched in the current study (Figure 6.3 for summary). Not all proteins were able to be matched across experiments, probably due to slightly different electrophoresis conditions.



Chapter 5
$p < 0.05$, any average ratio
n = 247, pI 4-7 and n = 120, pI 6-11

Chapter 6 (current study)
Matched spots
n = 200, pI 4-7 and n = 77, pI 6-11

Figure 6.3: Explanation of the data exported from DeCyder for R analysis

The standardised spot ratios of the selected proteins were exported from DeCyder for analysis using R. Initially a one-way ANOVA comparing the 7 groups was conducted, applying Bonferroni (Quinn and Keough 2002) and FDR correction on p values (Benjamini and Hochberg 2000). Permutation tests were then conducted to allow for the lack of independence among samples (Manly 1997). Preliminary

permutation tests used 1000 replications and a Bonferroni adjustment to eliminate the least significant spots. A more detailed analysis of fewer spots and 200,000 replicates was conducted. In addition to overall testing using all of the data, pairwise comparisons within the Bay of Islands and between Bay of Islands and Wellington data were performed. Pairwise testing within the Wellington (metal experiment) data used a Monte Carlo test because there were only 6 data points per spot. A Monte Carlo test calculates all possible permutations and gives them equal probability, which is less sensitive to random fluctuations and more accurate than the permutation testing (Manly 1997). One gel was removed from the pI 6-11 analysis because spots were not well separated and resolved. ANOVA testing was performed despite these missing values but pairwise comparisons between groups were limited by the missing values and for this reason the multiple comparison tests were not performed on the pI 6-11 dataset. The final comparison between the control group from Wellington (n = 2) and Bay of Islands and Wellington metal treated groups (n = 12) yielded the best results.

6.3 Results

6.3.1 Heavy metal concentrations in seawater samples from the Bay of Islands

Overall the concentrations of metals detected in seawater were relatively low (Table 6.2), although some Hg and Pb concentrations exceeded "trigger values" reported in the Australian and New Zealand fresh and marine water quality guidelines (Appendix 9, Table 9.1; Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand. 2000). Various trigger values are specified which offer alternate levels of protection for ecosystems e.g. metal concentrations < 90% trigger value will protect 90% of the species within the ecosystem. The 95% trigger level for Hg was exceeded at Waitangi Estuary and Te Hauini Estuary indicating a slightly-moderately disturbed system. The 99% trigger level for Pb was exceeded at Waitangi Estuary, whereas the 90% trigger level for Pb was exceeded at Te Haumi Estuary.

	Urapukapuka	Opua Marina	Waitangi Estuary	Te Haumi Estuary
Cd	<0.0002	<0.0002	< 0.0002	< 0.0002
Hg	0.00008	< 0.00008	0.00043 **	0.00054 **
As	0.006	<0.004	0.016	0.014
Pb	0.002	<0.001	0.004*	0.009***
Sn	<0.002	<0.002	< 0.002	< 0.002

Table 6.2: Metal concentrations (mg l^{-1}) in seawater from sites within the Bay of Islands
All are single measurements. Values greater than the recommended trigger values are indicated: * 99% ** 95% *** 90%

6.3.2 Bioconcentration Factors (BCF)

The BCFs for Cd accumulation in the Bay of Islands (Table 6.3) were much larger than the BCFs estimated in laboratory samples (Table 6.4), > 300 compared to approximately 15 respectively. In contrast the Hg BCFs were more comparable at the most Hg contaminated sites (Waitangi Estuary and Opua Wharf) because the maximum BCF calculated in the laboratory (28.3; Table 6.4) was similar to the BCF calculated in the field (30.9 and 96.9; Table 6.3). In contrast, at Urapukapuka and Opua Marina where Hg concentrations were below the level of detection, the BCF was appoximately 375.

Metal	Location	Average (Ce)	Background (Ci)	Seawater (Cs)	BCF
Cd	Urapukapuka	0.754	0.01	0.0002*	3720
	Opua Marina	0.114	0.01	0.0002*	520
	Waitangi Bridge	0.072	0.01	0.0002*	310
	Opua Wharf	0.091	0.01	0.0002*	405
Hg	Urapukapuka	0.05*	0.02	0.00008	375
	Opua Marina	0.05*	0.02	0.00008*	375
	Waitangi Bridge	0.062	0.02	0.00043	96.9
	Opua Wharf	0.037	0.02	0.00054	30.9
As	Urapukapuka	2.97	?	0.006	493
	Opua Marina	2.25	?	0.004*	560
	Waitangi Bridge	1.71	?	0.016	106.3
	Opua Wharf	1.56	?	0.014	110.7
Pb	Urapukapuka	0.03	?	0.002	1
	Opua Marina	0.072	?	0.001*	62
	Waitangi Bridge	0.065	?	0.004	13.8
	Opua Wharf	0.097	?	0.009	9.7

Table 6.3: Bioconcentration Factors for metal accumulation in the Bay of Islands
* The concentration used is the detection limit. The background concentration of Cd in mussels before exposure (Ci), used to calculate the BCF, was as reported by Hungspreugs et al. (1989). The Hg (Ci) was as reported by Hoggins (1972). When the Ci was unknown (?) a level of 0.01 was used, which is likely to overestimate the BCF

	Mercury (Cs = 0.0343 mg l^{-1})				Cadmium (Cs = 0.486 mg l^{-1})			
Time	**0**	**24**	**48**	**72**	**0**	**24**	**48**	**72**
BCF		14.3	28.3	13.1		5.6	15.4	15.3

Table 6.4: Metal accumulation in the soft tissues of *P. canaliculus* from Wellington

*The whole tissue measurements were calculated based on the average contribution of each tissue to the whole animal (Chapter 5)

6.3.3 Heavy metal concentrations in *P. canaliculus* samples from the Bay of Islands

Cd was below the level of detection in seawater (< 0.002 mg l^{-1}) at all 4 sites but was detectable in mussel tissue. The highest concentration of Cd in mussel samples (average value 0.754 mg kg^{-1}) was obtained from Urapukapuka, the 'clean site' (Table 6.5). Cd values at the remaining 3 sites were relatively low (0.072 – 0.114 mg kg^{-1}). Hg was detected in seawater from 3 of 4 locations but was below the level of detection in mussel tissue from Urapukapuka and Opua Marina (< 0.05 mg kg^{-1}). Waitangi Bridge and Opua Wharf recorded average Hg concentrations of 0.062 mg kg^{-1} and 0.037 mg kg^{-1} respectively (Table 6.5). As concentrations were highest in samples from Urapukapuka (2.97 mg kg^{-1}) and decreased at the remaining 3 locations, Opua Marina (2.25 mg kg^{-1}), Waitangi Bridge (1.71 mg kg^{-1}) and Opua Wharf (1.56 mg kg^{-1}; Table 6.3). Low concentrations of Pb were detected in mussels from Urapukapuka (0.03 mg kg^{-1}) and increasing concentrations were measured at Opua Marina (0.072 mg kg^{-1}), Waitangi Bridge (0.065 mg kg^{-1}) and Opua Wharf (0.097 mg kg^{-1}; Table 6.5) consistent with the trend reported for seawater. Sn was below the level of detection (< 1 mg kg^{-1}) in mussel samples from all sites.

Concentrations of iron (Fe), zinc (Zn), copper (Cu), aluminium (Al) and titanium (Ti) in mussel tissues showed trends which may be of relevance to this study. All of these metals occurred at low concentrations in samples from Urapukapuka, suggesting that despite relatively high Cd and As concentrations, this site may be relatively uncontaminated (Table 6.5). Waitangi Bridge had the highest concentrations of Fe, and Ti, whereas Opua Wharf had the highest concentrations of Zn and Cu. Opua Marina has the highest concentration of Al.

	Urapukapuka		Opua Marina		Waitangi Bridge		Opua Wharf	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Cd	0.754	0.332	0.114	0.025	0.072	0.016	0.091	0.019
Hg	< 0.05	-	< 0.05	-	0.062	0.012	0.037	0.008
As	2.965	0.427	2.248	0.666	1.707	0.173	1.555	0.143
Pb	0.03	0.015	0.072	0.01	0.065	0.042	0.097	0.024
Fe	22.8	9.6	86.5	8.6	105.2	33.3	78.7	22.8
Zn	9.8	2.1	14.6	5.5	10.6	1.2	15.0	3.2
Al	10.1	7.3	65.1	10.5	60.9	19.1	56.0	17.2
Cu	0.867	0.097	1.6	0.51	1.197	0.149	5.917	4.753
Ti	1.627	0.158	2.698	0.411	3.033	0.224	2.853	0.368
Sn	< 1	-	< 1	-	< 1	-	< 1	-

Table 6.5: Mean metal concentrations (mg kg^{-1}) in *P. canaliculus* from the Bay of Islands
n = 6 individuals per location

6.3.4 PERMANOVA analysis

PERMANOVA revealed significant (p = 0.0010) variation in metal concentrations in mussel tissue between locations. Multiple pairwise comparison of the sites determined that Urapukapuka was significantly ($p < 0.05$) different to the other 3 sites (shaded grey values, Table 6.6). Urapukapuka is also clearly separated from the other sites based on the dissimilarity data (Table 6.6) and Biplot (Figure 6.4). Urapukapuka and Waitangi Bridge have low within group dissimilarities (shaded light blue diagonal, Table 6.6) which is illustrated by compact clustering of samples in the Biplot (Figure 6.4). In contrast, Opua Marina and Opua Wharf have relatively high within-group variability. The arrows on the biplot indicate that Opua Marina and Opua Wharf are associated with most of the heavy metals, although Urapukapuka is more closely associated with Cd and As, and Waitangi Bridge is associated with Hg (Figure 6.4).

	Waitangi Bridge	Opua Wharf	Urapukapuka	Opua Marina
Waitangi Bridge	13.75	0.19	0.002*	0.470
Opua Wharf	22.37	28.08	0.002*	0.089
Urapukapuka	680.06	578.25	0.82	0.020*
Opua Marina	99.54	30.12	276.01	169.86

Table 6.6: Average dissimilarities within and between groups
Results in the grey shaded boxes are permanova p-values which indicate that Urapukapuka is significantly different to the other sites * Significant results ($p < 0.05$). Values in the blue shaded boxes are within site dissimilarity values. Note: The pair-wise tests have not been corrected for multiple comparisons

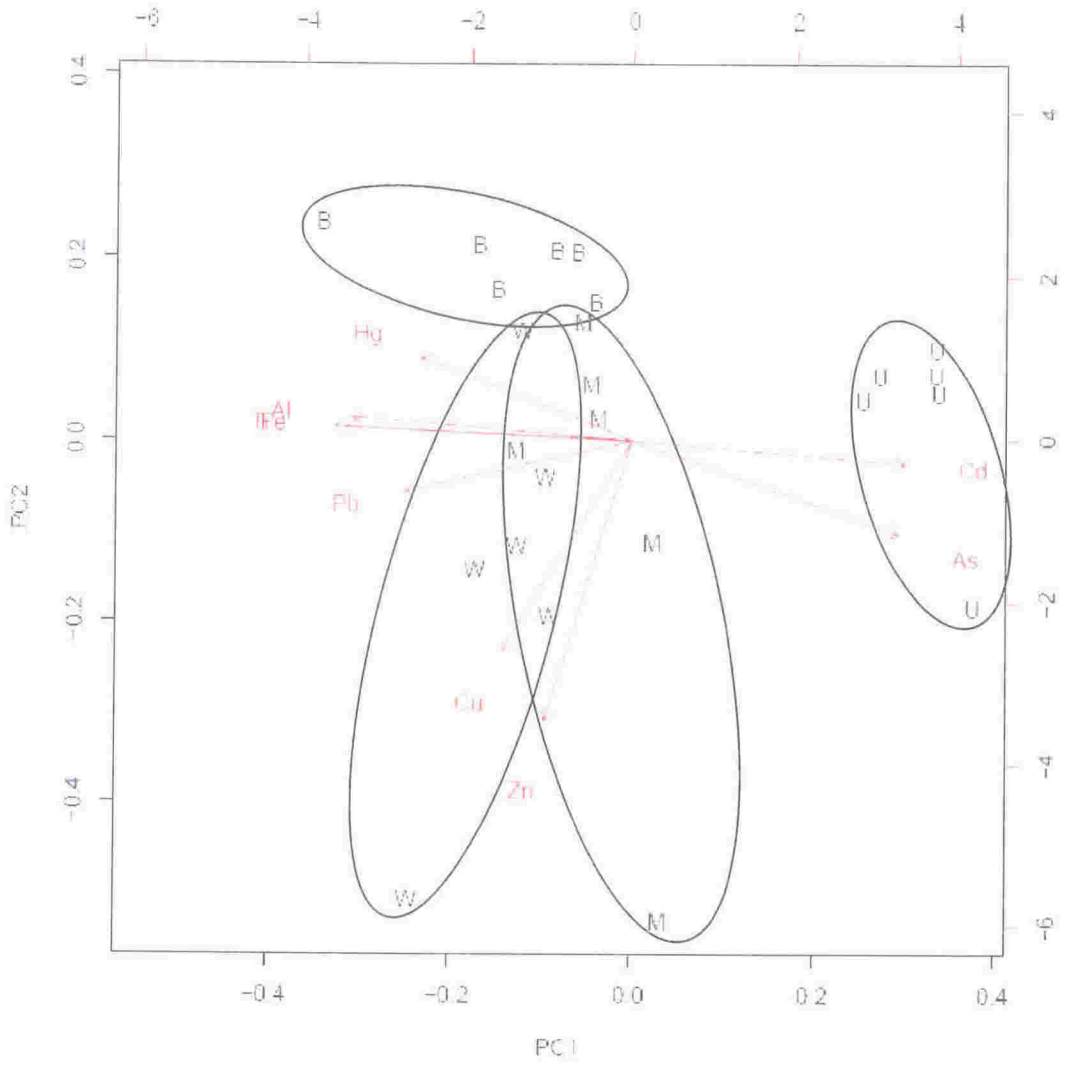


Figure 6.4: 2D Biplot of the of the Bay of Islands locations based on their first two Principal Components (PCs)

B = Waitangi Bridge, M = Opua Marina, W = Opua Wharf and U = Urapukapuka. The arrows show the direction of the highest metal values e.g. Cd and As are associated with Urapukapuka.

In all nMDS plots Urapukapuka samples clustered away from the other Bay of Island locations. All plots had low stress values of 0.02 indicating that the nMDS patterns shown are a good representation of the raw data. Average dissimilarity in the amount of metal detected among locations shows 3 different trends:

1) Urapukapuka has the highest concentration of Cd and the remaining sites have little or non-detectable amounts (Figure 6.5).

2) All sites have similar occurrence, although slightly different concentrations of As, Ti and Zn (Figure 6.6).

3) Opua Marina, Opua Wharf and Waitangi Bridge have high concentrations of Pb, Hg, Fe, Al and Cu but either low levels or non-detectable amounts were present at Urapukapuka (Figure 6.7). nMDS plots for Ti, Zn, Hg, Fe, Al and Cu are given in Appendix 10.

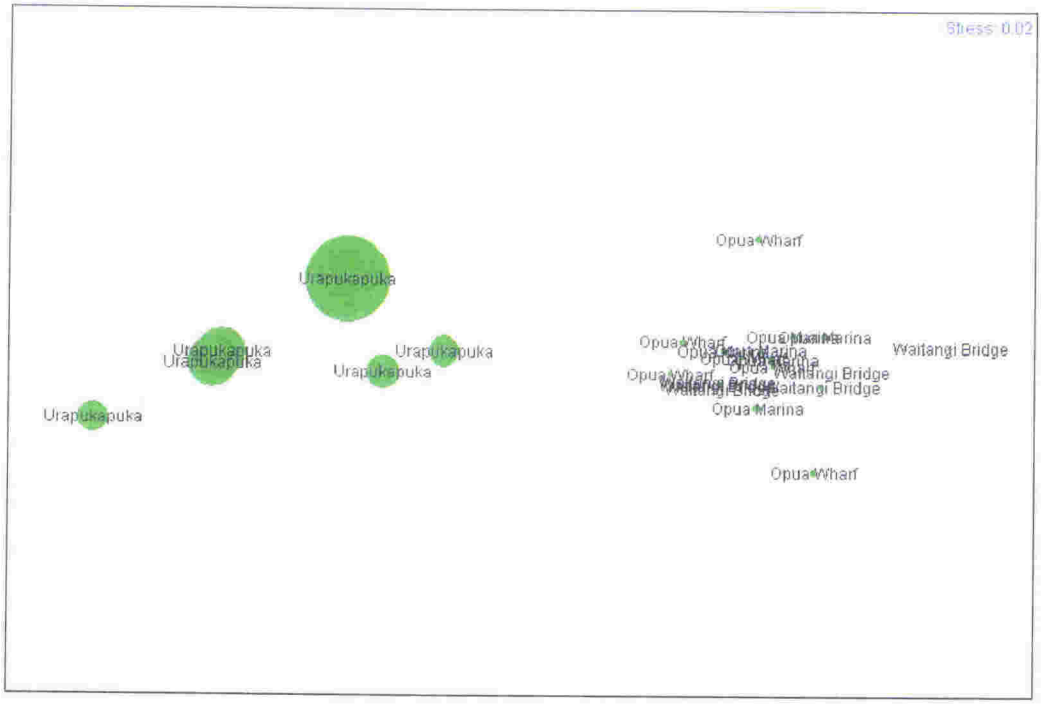


Figure 6.5: nMDS plot illustrating the relative dissimilarity between Bay of Island locations
The diameters of the green circles indicate the relative amount of Cd in individual mussels at the various sites.

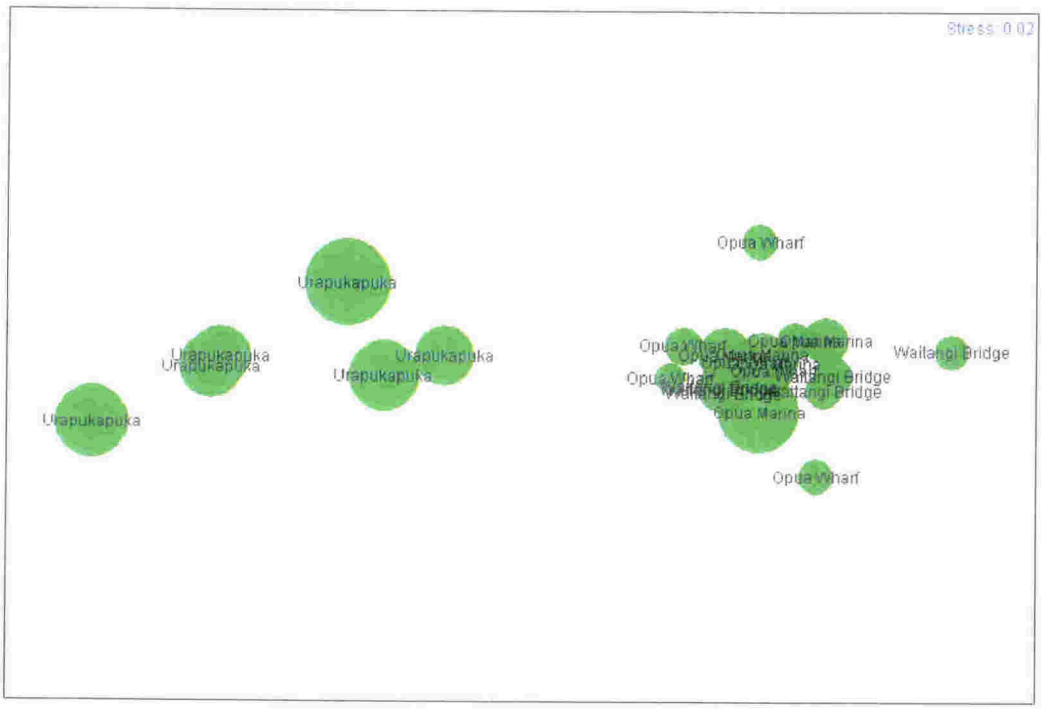


Figure 6.6: nMDS plot illustrating the relative dissimilarity between Bay of Island locations
The diameters of the green circles indicate the relative amount of As in individual mussels at the various sites.

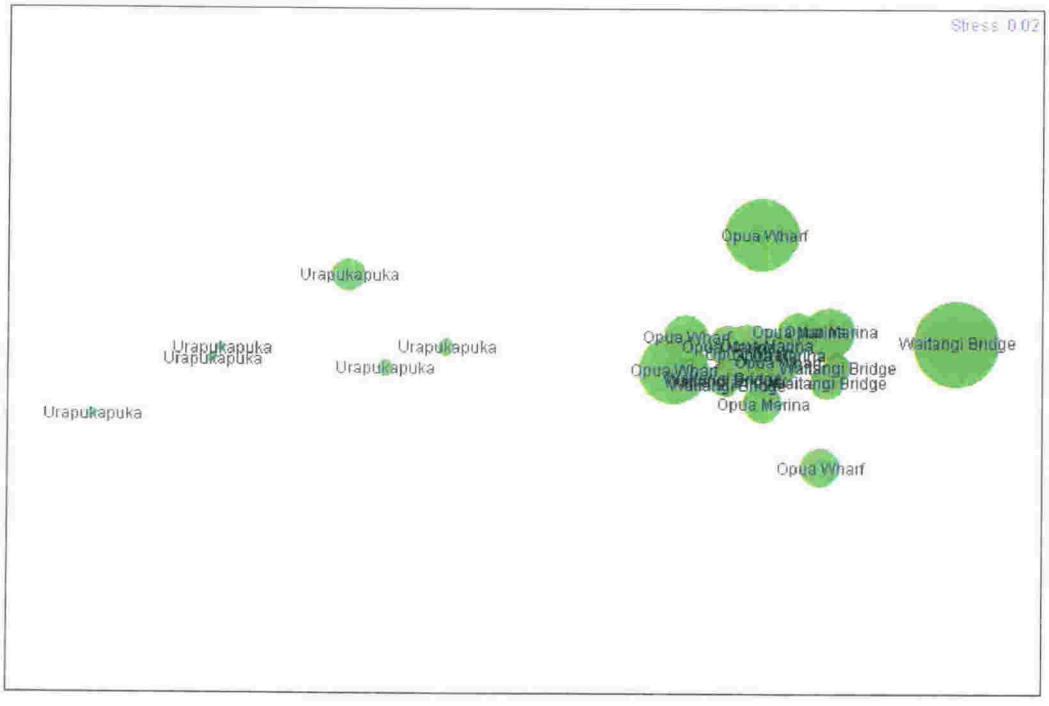


Figure 6.7: nMDS plot illustrating the relative dissimilarity between Bay of Island locations
The diameters of the green circles indicate the relative amount of Pb in individual mussels at the various sites.

6.3.5 Statistical analysis of protein abundance

Protein spots which showed statistically significant ($p < 0.05$) changes in abundance in Wellington (Chapter 5) samples were matched with the corresponding spots in the current Bay of Islands experiment. On pI 4-7 gels 200 spots (of 247) detected in Wellington samples were able to be matched with spots from the current Bay of Islands study and 77 spots (of 120) from pI 6-11 gels. Unless stated otherwise, the results reported below focus predominantly on pI 4-7 gels. One gel was removed from the pI 6-11 dataset in the current study because of poor spot separation and resolution and so the number of statistical tests that could be performed was reduced. In most cases after FDR adjustments, changes in protein abundance among groups were no longer significant. However, the results of these analyses are reported in full (Appendix 11) to indicate spots which may require further analysis.

One-way ANOVA plus sequential Bonferroni or FDR adjustment was used to test for proteins which exhibited significant ($p < 0.05$) changes in abundance among all seven groups. Five spots (spots; 2601, 1395, 1770, 1599, 2123) were detected initially but after permutation tests and FDR adjustments only one spot still demonstrated a significant difference in abundance among sites at the 5% level (spot 1395). Unadjusted p values are reported in Appendix 11 to indicate additional spot which may warrant further investigation (Table 11.1). Five spots which showed significant variation in abundance (spots; 574, 683, 929, 1014, 1203) from pI 6-11 gels are also reported in Appendix 11 (Table 11.2). After permutation testing and FDR adjustment the differences in abundance for these spots were no longer significant.

Permutation analyses were conducted on protein abundances from the 4 Bay of Islands sites (overall), all pairwise comparisons among the 4 Bay of Islands sites and the planned contrast between the least contaminated site (Urapukapuka) versus all other Bay of Islands sites (Appendix 11, Table 11.3). Permutation tests identified 9 spots showing significant ($p < 0.05$) changes in abundance for one or more of these comparisons but when the FDR adjustment was applied only one spot (1010) remained significant in comparisons between Opua Marina and Opua Wharf.

The normalised spot ratios for Wellington (n = 3) and Bay of Islands (n = 4) sites were averaged and compared (i.e. Wellington vs. Bay of Islands). Preliminary permutation tests with only 1000 replications selected 10 spots that appeared to show

significant differences in abundance for this contrast. Results for these 10 spots, with 200,000 replications, are shown in Appendix 11 (Table 11.4). However, after FDR adjustment p values were no longer significant.

The abundances of proteins in the control group from Wellington (n = 2) compared with the Bay of Islands and Wellington metal treated groups (n = 12) yielded the best results. Differences in protein abundance for 21 spots resulted in very low p values after FDR adjustment (p = 0.000), and another 10 showed significance at 8% (Table 6.7). Only three of these protein spots showed significant and large ($p < 0.01$; Av. Ratio ± 1.5) differences in abundance and were reported in Chapter 5, spot # in current study (spot # in Chapter 5): 2072 (2313), 2397 (2488) and 2486 (2595). The remaining spots were significant at the 5% level in Chapter 5. The locations of these protein spots are shown in Figure 6.8.

Spot	raw.p	FDR adj. p	Spot	raw p	FDR adj. p
863	0.000	0.000	2428	0.000	0.000
1395	0.000	0.000	2486	0.000	0.000
1492	0.000	0.000	2507	0.000	0.000
1591	0.000	0.000	2541	0.000	0.000
1599	0.000	0.000	2566	0.000	0.000
1734	0.000	0.000	1127	0.01098901	0.08453085
1737	0.000	0.000	1256	0.01098901	0.08453085
1770	0.000	0.000	1647	0.01098901	0.08453085
1865	0.000	0.000	1682	0.01098901	0.08453085
2014	0.000	0.000	1691	0.01098901	0.08453085
2072	0.000	0.000	1787	0.01098901	0.08453085
2123	0.000	0.000	2120	0.01098901	0.08453085
2148	0.000	0.000	2195	0.01098901	0.08453085
2151	0.000	0.000	2322	0.01098901	0.08453085
2376	0.000	0.000	2395	0.01098901	0.08453085
2397	0.000	0.000			

Table 6.7: PERMANOVA control (n = 2) versus all other groups (n = 12)

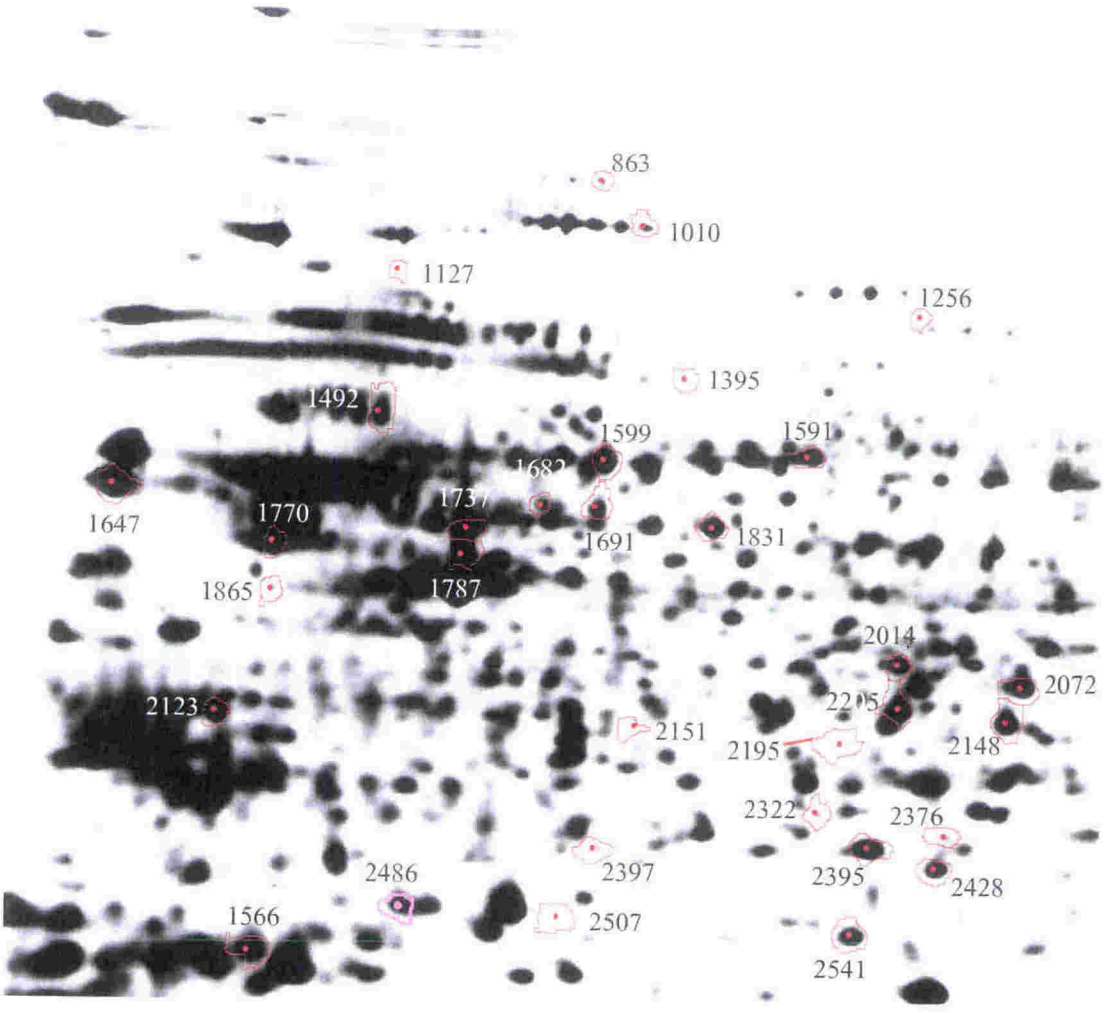


Figure 6.8: Gill pI 4-7 showing the spots with significant differences in abundance between control and all other groups. Spot 2486 also varied significantly in the salinity and heavy metal experiments.

Standardised log abundance plots illustrate the variation in protein abundance between sites (Figure 6.9). The example given is typical of the plots obtained from the spots shown in Table 6.8. In general standardised log abundance in control (n = 2) samples were lower than the other sites (n = 12) and Bay of Island samples tended to have higher overall standardised log abundance values than samples from Wellington (control, Cd and Hg). Opua Marina often had the highest standardised log abundance.

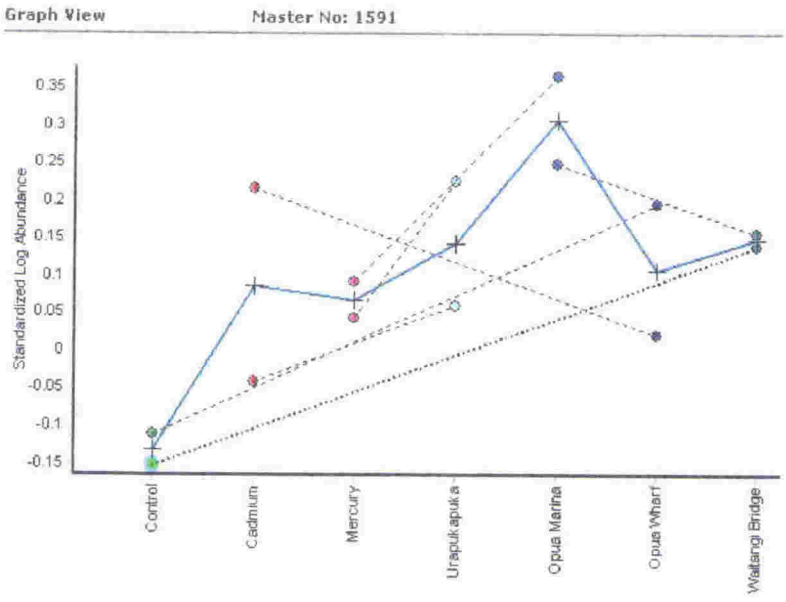


Figure 6.9: Graph view illustrating the variation in standardised log abundance between the 7 groups
Each spot represents a pooled sample labelled with either Cy3 or Cy5. The standardised log abundance of this protein was relatively low in the control group and higher in the other groups.

6.3.6 PCA

PCA gives a visual overview of the data, clearly separating the spot maps into four distinct groups: control, metal treated (Hg and Cd), Bay of Islands and outliers (Figure 6.10). In panel A, PC1 describes 33.2% of the total protein variation in response to environmental conditions and PC2 an additional 27.9%, a total of 61.1%. In panel C, PC1 describes 41% of the total spot map variation, and PC2 describes 23.7%, a total of 64.7%. There was a strong correlation between protein abundance and treatment (panels A and B respectively), because control (samples not exposed to metal under field or laboratory conditions) and other groups clearly separate. Control spot maps were on the opposite half of the quadrant (i.e. separated by the black line down the middle of panel B) from Wellington and Bay of Islands groups. For spot maps (panel C), the control and other groups are also clearly separated. It is also possible to separate Wellington (Control, Cd and Hg) and the Bay of Islands spot maps. Relative protein abundance between treatments is linked to the location of the spot maps and proteins within the quadrants of panels C and D respectively (as described previously, Chapter 3). PCA for pI 6-11 gels did not separate spots maps into groups as expected, which may be the result of removing one gel from analysis (Appendix 12).

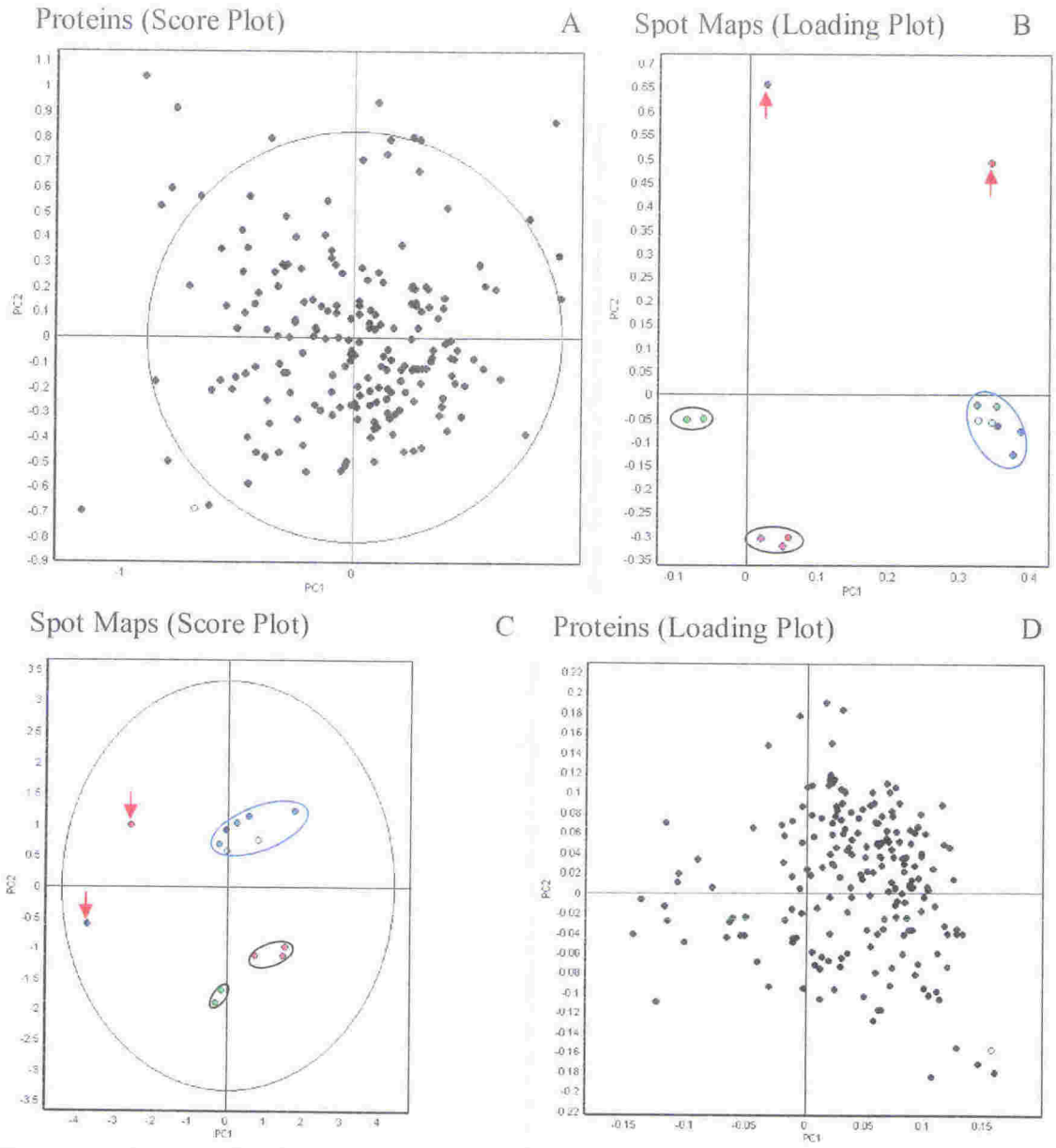


Figure 6.10: PCA for Waitangi samples (gill, pI 4-7)

Key:

- Opua Marina
- Opua Wharf
- Urapukapuka
- Waitangi Bridge
- Cd
- Control
- Hg

The black dots in panels A and D represent individual proteins (n = 200), whereas the coloured dots in panels B and C represent spot maps (n = 14). The red arrows indicate outliers.

6.3.7 Pattern analysis

Pattern analysis split the spot maps into 4 major groups; Control, Wellington, Bay of Islands, and outliers (Figure 6.11). Within the Wellington spot maps there were two distinct sub-groupings. One of these groups contained control samples whereas the other group contained individuals exposed to either Cd or Hg. The two spot maps to the right of the blue line at the extreme right of the figure (Cd and Opua

Wharf) appeared to be outliers. These two samples were separated on the same gel, and despite outward appearances of normality, it appears that this gel did not behave in the same manner as the other gels. In 6-11 gels, pattern analysis did not split the gels into groups as expected, which may be the result of removing one gel from analysis (Appendix 12).

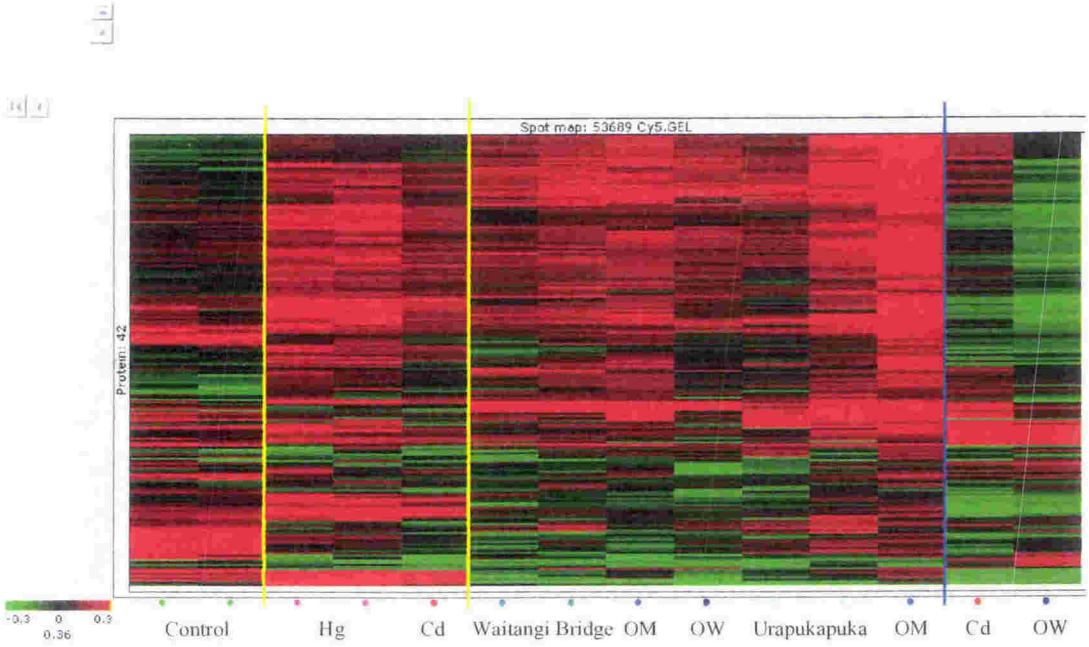


Figure 6.11: Pattern analysis for Waitangi samples (gill, pI 4-7)

Key:

- Opua Marina
- Opua Wharf
- Urapukapuka
- Waitangi Bridge
- Cd
- Control
- Hg

The yellow and blue lines have been added.
OM = Opua Marina; OW = Opua Wharf. Green = low protein abundance, red = high protein abundance.

6.4 Discussion

6.4.1 Metal concentrations in seawater from the Bay of Islands

The Hg concentrations detected in seawater from Te Haumi and Waitangi Estuaries exceed the 95% "Trigger Level" according to the Australian and New Zealand Guidelines for Fresh and Marine Water Quality, indicating that the Hg levels at these locations could cause distress for certain susceptible species (Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand. 2000). These results suggest that Hg concentrations in seawater from the Bay of Islands may be

sufficient to cause physiological responses in *P. canaliculus* which could be detected in the proteomic analysis (Section 6.4.3).

The "Trigger Level" for Pb was exceeded at Waitangi Estuary (within 95%), Urapukapuka and Te Haumi (>90% and <80%), indicating that Pb is a problem pollutant at some sites in the Bay of Islands region. Relatively high Pb concentrations were detected in seawater samples from Urapukapuka, but very low values were obtained from shellfish suggesting that the presence of Pb in seawater does not necessarily result in elevated Pb concentrations in mussels. However, previous studies suggest that *P. viridis* is a good biomonitoring organism for Pb based on field and laboratory experiments (Yap et al. 2004b). It is possible that a temporal spike of Pb was detected in seawater and the Pb concentrations determined during this study do not represent 'normal' levels. There are no established 'trigger levels' for As in seawater. Cd and Sn were not detected in seawater samples from the Bay of Islands so these metals were not a problem and are not discussed further.

6.4.1.1 Bioconcentration Factors

P. canaliculus BCFs calculated in field samples (current study) do not appear to correlate with BCFs determined in the laboratory (Chapter 5). The BCF for Cd was over 300 for field samples compared to a maximum of 15.4 under laboratory conditions. A Cd BCF of 15 is within the lower range for molluscs exposed to a similar concentration of Cd as illustrated in Taylor (1983). The Hg BCFs were more comparable, 28.3 in the laboratory compared to 30.9 and 96.9 in field samples. In previous laboratory experiments Hg BCF in hydrothermal vent mussels (*B. azoricus*) reached 10^4 (Kádár et al. 2004) and 10^5 in *M. galloprovincialis* (Casas et al. 2004). The BCF in the current study appear to be relatively low, although in previous studies Hg was accumulated in the laboratory for longer time periods (21 d) and using lower initial Hg concentrations (20 $\mu l\ l^{-1}$).

There are many possible reasons that could account for differences in field and laboratory BCFs, the most obvious being the experimental conditions. Cd was 2420 times more concentrated in the laboratory experiment (Chapter 5) than under natural conditions in the Bay of Islands, whereas Hg was a maximum of approximately 80 times more concentrated in the laboratory experiment. Furthermore, after 24 hours the Hg concentration in the laboratory metal experiment was only slightly more concentrated (0.00059) than Hg concentration in seawater

from Te Haumi (0.00054) and after 48 hours the amount of Hg left in the seawater during the metal experiment, was less than the amount of Hg in seawater from the 2 sites in the Bay of Islands (0.00041). It is possible that if Hg levels in seawater had been maintained at a constant low concentration, similar to natural conditions, that the BCFs between laboratory and field conditions could have been comparable.

6.4.2 Metal concentrations in *P. canaliculus*

The Hg and Pb concentrations reported in seawater and mussels in this study establish the idea that mussels from the Bay of Islands were naturally exposed to concentrations of heavy metals that could cause physiological responses suitable for proteomic analysis because the concentrations of these metals were above established "trigger levels". However, the concentration of Cd found at most sites was relatively low, for comparison, whole mussels tested from a contaminated site in a previous study had Cd concentrations of 1.1 mg kg^{-1} (Knigge et al. 2004). PCA and Biplot results indicated that metal concentrations at Urapukapuka were significantly different to the other Bay of Island sites. However, Urapukpuka was not the 'clean' site envisioned because it had relatively high concentrations of Cd and As although concentrations of Pb, Fe, Zn, Al, Cu and Ti were relatively low. Comparison of metal concentrations in *P. canaliculus* with other locations and other bivalves are discussed elsewhere (Chapter 7). The focus of this chapter is the development of field biomarkers for heavy metal exposure (Section 6.4.3).

6.4.3 Environmental Proteomics

Previous proteomic studies compared protein abundance profiles from polluted and reference sites (e.g. McDonagh et al. 2005; Mi et al. 2005) but in the current study there was no uncontaminated reference site from the Bay of Islands. Instead, protein abundances in the Bay of Island samples were compared with control (uncontaminated) samples from Wellington, and samples which were experimentally exposed to heavy metals in the laboratory (Chapter 5). Permutation testing of protein abundances in control mussels (from Wellington, n = 2) compared to all other groups (i.e. groups from the Bay of Islands and metal treated groups, n = 12) detected 21 proteins with extremely low (0.000) FDR corrected *p* values. In addition there were a further 10 proteins with relatively low FDR corrected *p* values (0. 085, Table 6.7). These proteins could be useful for determining whether *P. canaliculus* have been

exposed to heavy metals under field conditions. However, caution is advised as the link between these proteins and pollution although postulated, has not been unequivocally proven. Samples from a location in the Bay of Islands which had not been exposed to heavy metals would have helped to clarify this issue.

The proteins which showed significant differences in abundance as determined by permutation testing had low abundance in control samples and much higher standardised log abundance in samples from the remaining locations (e.g. Figure 6.9). In general, abundances for these proteins in mussels from the Bay of Islands were elevated compared to Cd and Hg treated samples (Figure 6.9), even though Cd and Hg concentrations were much higher during the metal experiment than in the field. It is unknown whether the high abundance of these proteins in sites from the Bay of Islands are the result of continued long-term exposure to heavy metals, or a genetic adaptation of mussels from these sites to cope with higher ambient metal concentrations in seawater in this region, or a coincidence that has nothing to do with exposure to metals. Mussels from the Bay of Islands may have naturally higher abundances of these proteins, even without stimulation from heavy metals within their environment. The results for the pattern analysis and PCA clearly differentiates between these Wellington and Bay of Islands groups, indicating clear 'genetic?' differences in the abundance of these proteins between these two populations (Figures 6.10 and 6.11). Mussels from a 'clean' site from the Bay of Islands may have been able to clarify this situation. The response of mussels from the Bay of Islands could have been scaled in relation to the Bay of Islands control to give a true indication of whether these proteins are influenced by heavy metals or not.

The statistical analysis of these proteins was made difficult by the failure of an ANOVA assumption (independent data). In addition, pooling samples into groups reduced the power of the experiment to detect significant changes in protein abundance. Having more than 2 data points per group would have ensured greater statistical power. Karp et al. (2005) suggest that instead of having a single large pool, samples should be split in half e.g. if n =10, label 5 samples with Cy3 and label the other 5 samples with Cy5. This slight modification ensures that DeCyder software can be used to analyse the data because the two pooled samples are independent. Despite these limitations, the proteins detected in the current study are likely to be involved in dealing with heavy metal stress under both laboratory and

field conditions. Further work is required to validate these findings before these proteins could be used as biomarkers of heavy metal pollution in *P. canaliculus*.

6.5 Conclusions

Approximately 30 proteins were detected which could potentially be used to determine whether greenshell mussels have been exposed to heavy metals under field conditions (Tables 6.7). However, validation of these proteins is still required, as a comparison with a clean field site was not possible during the current study. Additionally, this experiment has highlighted some of the statistical difficulties that can be encountered using DIGE and the DeCyder analysis package.

6.6 Acknowledgements

I would like to gratefully acknowledge the help of Dr. Shirley Pledger (Victoria University, School of Mathematics and Statistics) for writing and running the R programme for the protein permutations. I would also like to thank Rudi Schnitzler for his instructions on the use of the programme PRIMER for the nMDS plots.

CHAPTER 7: Dietary exposure to heavy metals in the Bay of Islands, Northland, New Zealand

Abstract

Cd, Hg, As, Pb and Sn concentrations were determined in *P. canaliculus* from Urapukapuka-Rawhiti, Opua Marina, Waitangi Bridge and Opua Wharf (Bay of Islands, New Zealand). Metal concentrations were compared to measurements taken by previous studies conducted throughout New Zealand. All individual samples in the current study had metal concentrations well below the Food Standards Australia and New Zealand (FSANZ) maximum limits and comparable to concentrations obtained in previous studies. Based on the average concentrations of metals detected in the current study, the amount of metal consumed through a typical diet containing shellfish would be below the provisional tolerable weekly intake (PTWI). However, Māori, Pacific Islanders and Asians consume a greater quantity of seafood than the general New Zealand public. The risks of heavy metal exposure from dietary intake for this sector of society are discussed. A survey of the frequency, amount and species consumed by this group is suggested to enable an adequate risk assessment to be made.

7.1 Introduction

7.1.1 Dietary exposure to heavy metals in the Bay of Islands

Iwi from the Bay of Islands (Northland, New Zealand) are concerned that heavy metal inputs into rivers from naturally occurring geothermal sources in this region, are having negative impacts on human health and the coastal biota (pers. comm. Emma Gibbs, Ngāpuhi). Volcanic emissions and geothermal activity in New Zealand pose some risk to human health through the introduction of metal contaminants, including Hg and As, into the environment (Cook and Weinstein 2005). In 2002, particular concerns were raised during the construction of a controversial prison on wāhi tapu (Māori cemetery/reserved) land at a geothermal site in Ngāwhā (Northland, New Zealand). Geological instability at the construction site, compounded by heavy rainfall and rising water levels, led to increased amounts of Hg in the Waiaruhe River (Hg reached 5.4 ppb, compared to a 'satisfactory' level of 2.0 ppb, Far North District Council). The Waiaruhe River flows into the Waitangi

River from which Waitangi and Paihia residents get their drinking water, so concerns were raised regarding the quality of this water for human consumption. In addition, Hg and other heavy metals are known to accumulate in shellfish (e.g. Irato et al. 2003), which provide a substantial proportion of the diet for many people within the Northland region (Hay 1996).

To establish the risk of heavy metal contamination for human consumers the concentration of heavy metals contained within kaimoana (seafood) must be measured. Although the concentrations of heavy metals in a variety of food sources were determined in the 2003/2004 New Zealand Total Diet Survey (NZTDS; Vannoort and Thomson 2005) this survey did not include shellfish from all regions of New Zealand. Regions subject to volcanism are likely to contain elevated concentrations of heavy metals. Metal concentrations reported in the NZTDS reflect averages from shellfish sampled from Auckland, Napier, Christchurch and Dunedin, locations which may not adequately reflect heavy metal concentrations found in shellfish sourced from the Bay of Islands. Additionally, the heavy metal risk assessment in the Total Diet survey is based on the amount of shellfish consumed in a 50th percentile diet which does not adequately reflect Māori and Pacific Island diets which include more shellfish. In calculating risk associated with shellfish consumption, a meal size of 100 g shellfish meat (or less) is generally used, but observation and discussion with Māori suggest that this is an underestimation of shellfish meal size for rural Māori in Northland, where kaimoana often forms a substantial part of the diet (pers. comm. 2006, Brenda Hay, AquaBio Consultants). In the present study the concentrations of heavy metals in *P. canaliculus* from the Bay of Islands were compared with samples from around New Zealand and the Food Standards Australia and New Zealand limits. The influence of shellfish in dietary metal accumulation in Māori and other high risk groups is discussed.

7.2 Materials and Methods

7.2.1 Mussel and water sampling

Adult greenshell mussels (6-15 cm) were collected in December 2005 from 4 locations within the Bay of Islands, New Zealand (Figures 7.1 and 7.2). Urapukapuka-Rawhiti (abbreviated to Urapukapuka from this point forward) is an offshore island approximately 15 km away from the other sites, which are in close

proximity to the Veronica Channel. Mussels from Urapukapuka were selected as controls because it was assumed that mussels from this site have not been exposed to elevated concentrations of heavy metals from industry and boating activities. It was expected that the remaining 3 sites would be exposed to varying degrees of heavy metal pollution from boating, geothermal and industrial activities.

Six mussel samples from each site were sent to ELS for total tissue chemical analysis of Hg, Cd, As, Pb and Tin (Sn). Whole mussels were analysed, i.e. mussels were not dissected into individual tissues as described previously (Chapter 5). Whole mussel metal concentrations are more appropriate for dietary risk assessments because mussels are consumed whole. Hg, Cd and As were analysed because these elements are associated with geothermal activity. Pb was expected to act as a control metal (as in Chapter 5), although Pb paint could be associated with boating activity. Tributyltin (TBT) is an antifouling coating which was subject to a partial ban in 1989 and a full ban in 1993 (Smith 1996). Because TBT lingers in sediments (Coelho et al. 2002), TBT could still be associated with boating activity therefore total Sn concentrations were also tested.

A single water sample was collected from each location and was analysed by Hill Laboratories for the presence of the same heavy metals as listed above. Water and mussel samples were collected concurrently (Dec 2005) from Urapukapuka and Opua Marina. The remaining water samples were collected in mid-January 2005 from Waitangi Estuary and Te Haumi Estuary, locations which are in very close proximity to the mussel sampling locations (Waitangi Bridge and Opua Wharf; Figures 6.1 and 6.2).

7.2.2 Dietary risk assessment

The maximum metal concentrations detected in mussels from the Bay of Islands were used to calculate the maximum weekly metal intake based on an 'average' (78 g per day) diet (Appendix 13 for an example calculation). This example provides a 'worst case scenario' because average metal concentrations are much less than maximum values. Based on average metal concentrations and an 'average' diet the Provisional Tolerable Weekly Intakes (PTWI) were not exceeded (results not shown). The amount of greenshell mussels required to exceed the PTWI for each metal was calculated based on the minimum and maximum average metal concentrations obtained from this study (Appendix 13 for an example calculation).

These calculations were performed to assess whether the amount of mussels required to exceed the PTWI were realistic for 'extreme' consumers.

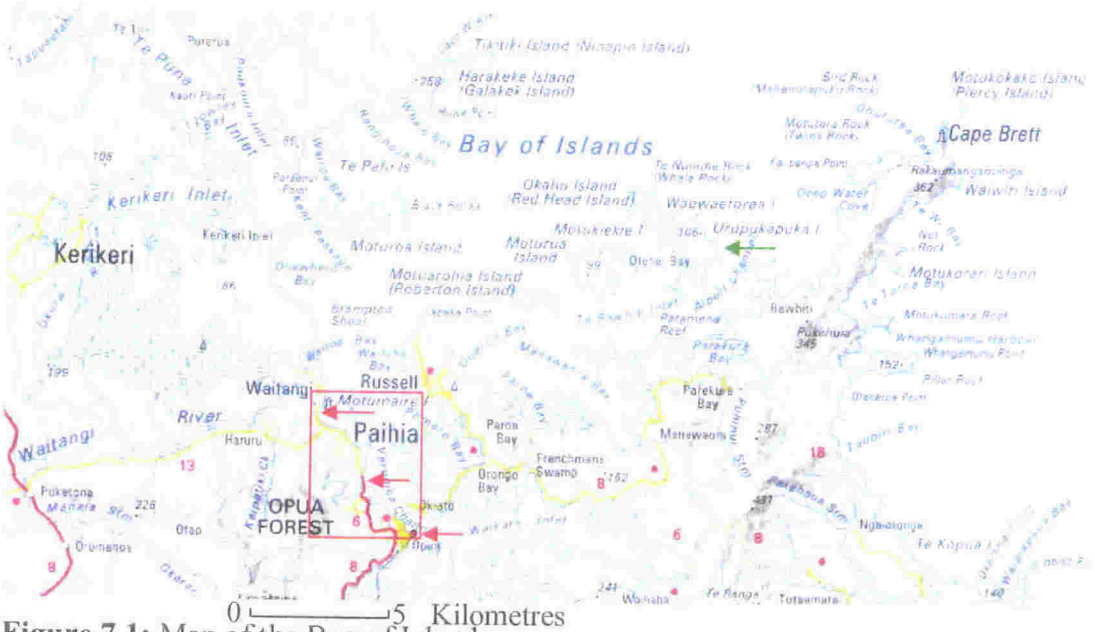


Figure 7.1: Map of the Bay of Islands

Sampling locations are indicated by arrows. The red rectangle is enlarged in Figure 6.2. The approximate sampling location for Urapukapuka-Rawhiti (the clean site) is indicated by the green arrow. Map reproduced from the Reed New Zealand Atlas (Dowling et al. 2004).

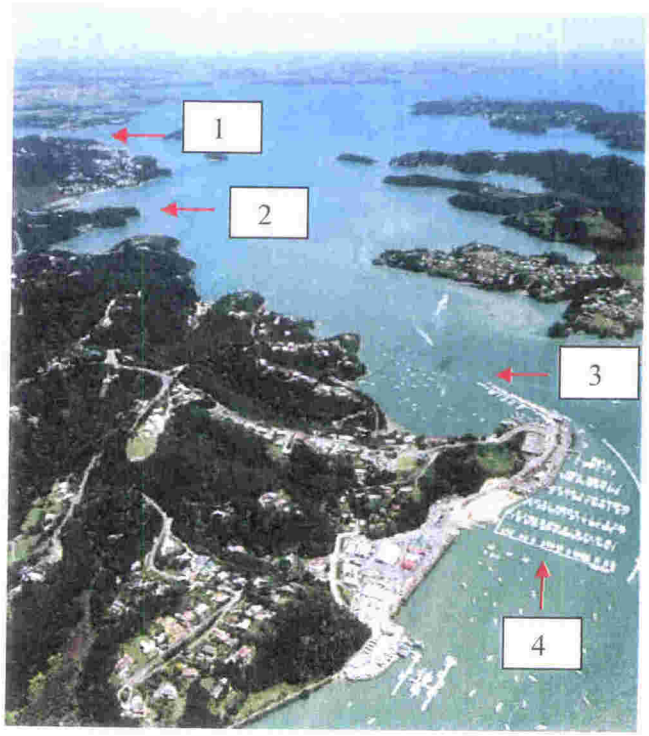


Figure 7.2: Enlarged view of sampling locations

(1)Waitangi Bridge/Estuary, (2) Te Haumi Estuary (Water sample only), (3) Opua Wharf (4) Opua Marina

7.3 Results

7.3.1 Heavy metal concentrations in seawater samples from the Bay of Islands

Concentrations of Cd and Sn were both below the level of detection at all sites tested (Table 7.1). Hg concentrations were relatively low at Urapukapuka (0.00008 mg l^{-1}) and Opua Marina (< 0.00008 mg l^{-1}) but were elevated at Waitangi Estuary (0.00043 mg l^{-1}) and Te Haumi Estuary (0.00054 mg l^{-1}; Table 7.1). As concentrations were relatively high at Waitangi Estuary (0.016 mg l^{-1}) and Te Haumi Estuary (0.014 mg l^{-1}) and were below the level of detection at Urapukapuka (0.006 mg l^{-1}) and Opua Marina (< 0.004 mg l^{-1}). Pb concentrations were higher at Te Haumi Estuary (0.009 mg l^{-1}) than Opua Marina (< 0.001 mg l^{-1}), Urapukapuka (0.002 mg l^{-1}) and Waitangi Estuary (0.004 mg l^{-1}; Table 7.1).

	Urapukapuka	Opua Marina	Waitangi Estuary	Te Haumi Estuary
Cd	<0.0002	<0.0002	< 0.0002	< 0.0002
Hg	0.00008	< 0.00008	0.00043 **	0.00054 **
As	0.006	<0.004	0.016	0.014
Pb	0.002	<0.001	0.004*	0.009***
Sn	<0.002	<0.002	< 0.002	< 0.002

Table 7.1: Metal concentrations (mg l^{-1}) in seawater from the Bay of Islands

7.3.2 Heavy metal concentrations in *P. canaliculus* from the Bay of Islands

The highest concentration of Cd in mussel samples (average value 0.754 mg kg^{-1}) was obtained from Urapukapuka, the 'clean site' (Table 7.2). Cd concentrations at the remaining 3 sites were relatively low (0.072 – 0.114 mg kg^{-1}). Hg was below the level of detection in samples from Urapukapuka and Opua Marina (< 0.05 mg kg^{-1}) whereas Waitangi Bridge and Opua Wharf recorded average Hg concentrations of 0.062 mg kg^{-1} and 0.037 mg kg^{-1} respectively (Table 7.2). As concentrations were highest in samples from Urapukapuka (2.97 mg kg^{-1}) and decreased at the remaining 3 locations, Opua Marina (2.25 mg kg^{-1}), Waitangi Bridge (1.71 mg kg^{-1}) and Opua Wharf (1.56; Table 7.2). Pb concentrations in mussel tissue exhibited the opposite trend to As with low concentrations at Urapukapuka (0.03 mg kg^{-1}) and increasing values at Waitangi Bridge (0.065 mg kg^{-1}), Opua Marina (0.072 mg kg^{-1}) and Opua Wharf (0.097 mg kg^{-1}; Table 7.2). Sn was below the level of detection (< 0.1 mg kg^{-1}; Table 7.2) at all sites tested.

	Urapukapuka		Opua Marina		Waitangi Bridge		Opua Wharf	
	Mean	SD	Mean	SD	Mean	SD	Mean	SD
Cd	0.754	0.332	0.114	0.025	0.072	0.016	0.091	0.019
Hg	< 0.05	-	< 0.05	-	0.062	0.012	0.037	0.008
As	2.965	0.427	2.248	0.666	1.707	0.173	1.555	0.143
Pb	0.03	0.015	0.072	0.01	0.065	0.042	0.097	0.024
Sn	< 0.1	-	< 0.1	-	< 0.1	-	< 0.1	-

Table 7.2: Mean metal concentrations (mg kg^{-1}) in *P. canaliculus* (n = 6 per location) from the Bay of Islands

7.3.3 Possible dietary intake of heavy metals from *P. canaliculus* based on measurements taken from the Bay of Islands

The concentrations of all metals tested in all individual *P. canaliculus* samples from the Bay of Islands were below the Food Standards Australia and New Zealand (FSANZ) maximum levels (Table 7.3). Based on average metal concentrations obtained from the Bay of Islands the amount of metal consumed through a typical diet containing shellfish were below the PTWI. Using maximum individual concentrations As, Pb and Hg were still all below the PTWI (Table 7.3). However, the PTWI obtained for Cd using the maximum concentration was 10.7 µg kg^{-1} body weight per week (bw $week^{-1}$), which exceeded the PTWI (Table 7.3). Although As values appear to exceed the PTWI, the FSANZ standards are for inorganic As not total As. Inorganic As only forms approximately 20% of total As, so the PTWI for total As (350 kg^{-1} bw $week^{-1}$) was not exceeded (Table 7.3). There are no standards for Sn and this metal was below the level of detection (< 0.1 mg kg^{-1}) and so is not discussed further.

In general large amounts (approximately 5 kg plus) of *P. canaliculus* from the Bay of Islands would need to be consumed each week in order to exceed the PTWI (Table 7.4). However, based on the maximum average Cd concentration obtained from Urapukapuka, the amount of mussels required to exceed the PTWI limit for Cd were plausible for high volume consumers (0.650 kg; Table 7.4). The results suggest that Cd could pose a risk to consumers who eat a large amount of greenshell mussels.

Metal	FSANZ Standards Maximum levels	FSANZ Provisional Tolerable Weekly Intake (μg kg^{-1} bw $week^{-1}$)	Maximum individual [ ] from Bay of Islands	Weekly intake based on maximum value (μg kg^{-1} bw $week^{-1}$)
As	1.0 mg kg^{-1}	15#	3.69 mg kg^{-1}	28#
Cd	2.0 mg kg^{-1}	7	1.37 mg kg^{-1}	10.7
Pb	2.0 mg kg^{-1}	25	0.15 mg kg^{-1}	1.2
Hg	0.5 mg kg^{-1}‡	5	0.08 mg kg^{-1}	0.6

Table 7.3: Summary of Australian and New Zealand Food standards and weekly intake based on maximum individual metal concentrations obtained from the Bay of Islands

Maximum levels of specified metal in molluscs exclude dredge/bluff oysters and queen scallops.
Note: There is no stated standard for dredge/bluff oysters and queen scallops.
The provisional tolerable weekly intake (PTWI) was based on a 70 kg adult and an average serving size of 78 g (i.e. 546 g per week). The grey shaded box indicates that the PTWI was exceeded

#The PTWI is for inorganic As, not total As. A maximum allowable daily body load for total As of 50 μg kg^{-1} (350 μg kg^{-1} bw $week^{-1}$) was set by World Health Organisation (WHO) in 1967.
‡ The limit applies to the average of five samples and individual samples may be higher and still comply.

Metal	Average concentration (mg kg^{-1})		Amount of mussel flesh per week required to exceed FSANZ limit (kg)	
	min	max	Based on max	Based on min
Cd	0.1	0.754	0.650	4.9
Hg	0.05*	0.062	5.6	7.0
As (total)	1.56	2.97	8.25	15.7
Pb	0.03	0.097	18.0	58.3

Table 7.4: Amount of *P. canaliculus* that would need to be consumed by a 70 kg adult (per week) to exceed FSANZ limits

min = minimum average concentration; max = maximum average concentration
Grey shaded box indicates an amount of mussel flesh which could realistically be exceeded in the diet of at risk groups.
* no analyte was found in the sample at the level of detection shown
The amount of mussel flesh required to exceed FSANZ limit was calculated using minimum and maximum average concentrations. The amounts of mussel flesh reported are green weight

7.3.4 Heavy metal concentrations in shellfish from New Zealand

The concentrations of metals found in *P. canaliculus* in the current study were comparable to, or slightly less than, those found in previous studies (Table 7.5). The elevated Cd concentration from Urapukapuka was the exception, but the value was below the FSANZ limits and was comparable with historical samples from the Marlborough Sounds. Selected heavy metal concentrations from other shellfish in New Zealand are provided for comparison (Table 7.5). Of note are the high concentrations of Cd in oysters and scallops. Dredge/bluff oysters and queen scallops are not included in the FSANZ standards (Table 7.3).

7.3.5 Heavy metal concentrations in shellfish worldwide

The concentrations of metals found in *P. canaliculus* in the current study were comparable to, or slightly less than, those found in worldwide studies (Table 7.6). However, the elevated Cd concentration from Urapukapuka was again the exception. Cd concentrations in *P. canaliculus* from Urapukapuka were higher than Cd concentrations found in *P. viridis* and *P. perna* from China, Malaysia and Brazil (Table 7.6). In general, concentrations of metals reported in worldwide studies are also below FSANZ limits but Hg concentrations in China, and Pb concentrations in some locations in Asia were above FSANZ maximum values. In particular, Hg concentrations in China were extremely high (Table 7.6).

Shellfish	Year	Location	n	Cd	Hg	Pb	As	Ref
Green mussel	2005	BoI	18	0.09	0.05	0.078	1.82	a
	2005	BoI	6	0.754	< 0.05	0.03	2.97	a
	2005	Seatoun		0.1	< 0.01	0.2	3	a
	2005	VAR	8	0.211	0.02	0.076	1.97	b
	1975	VAR		0.3	0.09	1.8		c
		WCNI	12		0.03			d
	1991	MS	17	0.8				e
	1992	MS	21	0.72				e
	1993	MS	38	0.9				e
	1994	MS	34	0.95				e
	1995	MS	35	0.68				e
	1996	MS	38	0.5				e
	1999	MS	19	0.22				e
	2002	MS	25	0.295				e
Blue mussel	1975	VAR		0.63	0.23	0.67		c
		WCNI	12		0.02			d
Pacific oyster		ECNI			0.02			d
Bluff oyster		SL		4.92	0.01	0.01		d
Rock oyster	1975	VAR		1.3	0.18	0.90		c
Dredge oyster	1975	VAR		3.9	0.19	0.69		c
Oysters	2005	VAR	8	2.92	0.01	0.022	1.92	b
Pipi	1975	VAR		0.13		0.40		c
Cockle	1975	VAR		0.19		1.8		c
Tuatua	1975	VAR		0.29		0.65		c
Toheroa	1975	VAR		0.11		0.76		c
Scallop			10		0.01			d
(muscle and roe)	1995	ECSI	10	4.0				d
(gut)	1975	VAR		0.18	0.05	1.14		c
(abductor)	1975	VAR		0.51				c
Paua		VAR		0.02		0.07		d

Table 7.5: Mean heavy metal concentrations in New Zealand shellfish (mg kg^{-1} wet wt) Grey shading indicates relatively high metal concentrations.

BoI, Bay of Islands; WCNI, North Island; SL, Southland; MS, Marlborough Sounds; ECSI; East Coast South Island; VAR, various locations

(a) Current study
(b) New Zealand Total Diet Survey. Averages from samples collected from Auckland, Napier, Christchurch and Dunedin.
(c) Nielsen and Nathan (1975)
(d) Fenaughty et al. (1988) summary of data collected by MAF mid-1970s to late 1980s. No dates specified with data. Detection limits were Cd 0.01 mg kg^{-1}, Hg 0.01 mg kg^{-1}, Pb 0.2 mg kg^{-1}
(e) Records from regional councils collected by New Zealand Food Safety Authority (NZFSA). Cd concentrations in samples from the North Island decreased between 1991 and 2003 from 0.705 to 0.15 mg kg^{-1}. However, these have not been reported on the table as averages were based on a small number of samples (1-3 samples/yr).

Shellfish	Location	Cd	Hg	Pb	As	Ref
Green mussel						
P. viridis	China	0.152 - 0.612	3.47 - 32.98	0.062 - 0.111	1.60 - 3.03	a
	Malaysia	0.120 - 0.220		0.430 - 1.49		b
	Malaysia	0.030 - 0.154		0.156 - 1.08		c
	various			0.123 - 3.08		d
P. perna	Brazil	0.065 - 0.123	0.002	0.119		e
	Ghana		0.04 - 0.05			f
Blue mussel						
M. edulis	China	bld - 0.294	0.42 - 13.93	bld - 0.3	bld - 2.62	a
	Bohai Sea	0.34 - 1.96				g
	Scotland	0.037 - 0.074				h
	Ireland	0.086 – 0.8				h
	UK	0.5 – 0.85				h
	Canada	0.111-0.123				h
	Portugal	0.086 – 0.111				h
M. galloprovincialis	various	< 0.37				h
	Italy	0.74 - 1.73				i
Mussel watch	France	0.123	0.02	0.300		j
Mussel watch	USA	0.370	0.01	0.222		j
Others						
C. tulipa	Ghana		0.02 - 0.03			f
C. talienwhanensis	Bohai Sea	0.51 - 3.29				g
Anadara senelis	Ghana		0.02 - 0.04			f
R. phillippinarum	Bohai Sea	0.14 - 0.63				g

Table 7.6: Maximum and minimum mean heavy metal concentrations in shellfish from worldwide locations (mg kg^{-1} wet wt)
bld = below the level of detection

(a) Fung et al. (2004). Concentrations reported are maximum and minimum averages from 4 locations, values were converted from dw to wet wt.
(b) Yap et al. (2004a). Maximum and minimum averages from 9 locations within Peninsular Malaysia.
(c) Yap et al. (2003a) Maximum and minimum averages from 20 locations within Peninsular Malaysia, values converted from dw to wet wt.
(d) Yap et al. and references therein (2004b) Maximum and minimum concentrations from various locations (India, Hong Kong, Malaysia and Thailand), values converted from dw to wet wt.
(e) Baraj et al. (2003). Mean maximum and minimum concentrations from 15 locations, values converted from dw to wet wt.
(f) Otchere et al. (2003). Concentrations reported are maximum and minimum average values from 3 locations.
(g) Liang et al. (2004) Maximum and minimum average concentrations from 7 sites within the Chinese Bohai Sea
(h) Bebianno and Machado and references within (1997). Cd concentrations in *Mytilus* spp. from different geographical areas are summarised. *M. galloprovincialis* samples were from Portugal, the north coast of Greece and the Adriatic, values converted from dw to wet wt.
(i) Bargagli et al. (1985) Italy *M galloprovincialis*
(j) Beliaeff et al. (1998) Summary of chemical concentrations in mussels and oysters from "Mussel Watch" monitoring programs in the United States and France. Values reported are the maximum average concentration over a period from 1986 to 1993.

7.4 Discussion

7.4.1 Metal concentrations in seawater

In general, maximum metal concentrations detected in seawater from the Bay of Islands were relatively low compared to overseas studies. In the current study both Cd and Sn were below the level of detection (< 0.0002 mg l^{-1}). Cd concentrations in seawater samples from Ria Formosa Lagoon (Portugal) ranged from approximately < 0.0002 to 0.001 mg l^{-1}, with the highest Cd concentrations occurring in the inner parts of the lagoon (Bebianno 1995). Total dissolved Cd concentrations at the most contaminated sites were greater than the recommended EEC Standards (Bebianno 1995). Pb concentrations at the same sites ranged from < 0.002 to 0.033 mg l^{-1} (Bebianno 1995), which are also relatively high in comparison to concentrations reported in the current study (< 0.002 – 0.009 mg l^{-1}). Heavy metal concentrations in seawater from the Gulf of Patras (Greece) ranged from < 0.002 to 0.0122 mg l^{-1} (Pb), < 0.0002 to 0.0005 mg l^{-1} (Cd) and from < 0.00008 to 0.014 mg l^{-1} (Hg) (Kalpaxis et al. 2004). Hg, Cd and Pb concentrations were elevated in Greece compared to the present study (Hg concentrations in the current study ranged from < 0.00008 to 0.00054 mg l^{-1}). The Bay of Islands does not have any major industries, so the low metal concentrations in seawater compared to overseas studies was not unexpected.

7.4.2 Metal concentrations in shellfish

It is not valid to compare absolute accumulated metal concentrations in biomonitors inter-specifically (Rainbow 1995), so metal accumulation comparisons focus on *P. canaliculus*. However, heavy metal concentrations reported in other bivalves from selected international studies are discussed briefly so that the relative risks to consumers in New Zealand are placed in a worldwide context. Emphasis was placed on providing comparative information from mussel species, although accumulation data from other bivalves are also reported in some instances to highlight differences in metal accumulation between species sampled from the same site. Where previous studies reported metal concentrations in mg kg^{-1} dry weight the concentrations were converted to wet weight equivalents using a conservative conversion factor of 8.1 (conversion factors vary from approximately 5 - 8.1; Dragun 2004; Yap et al. 2004a).

7.4.2.1 As and Hg concentrations in shellfish

In general As and Hg concentrations detected in *P. canaliculus* samples from the Bay of Islands were slightly elevated compared to the values reported in the NZTDS (Vannoort and Thomson 2005). Possible reasons for the higher concentrations of As and Hg in the Bay of Islands include the proximity of these sites to the Ngāwhā geothermal zone and/or mining activities within the vicinity of the Waitangi River catchments (Craw et al. 2005). Previous studies of eels have recorded elevated MeHg concentrations in flesh from eels caught in the central geothermal plateau (Chisnall 1997). In the current study the average Hg concentration ranged from < 0.05 to 0.062 mg kg^{-1}. The average Hg concentration in *P. canaliculus* was 0.09 mg kg^{-1} (Nielsen and Nathan 1975), while the highest Hg concentrations detected were at the Ngauranga Stream mouth in Wellington (0.14 mg kg^{-1}) and Nikau Bay wharf in Marlborough Sounds (0.19 mg kg^{-1}). Hg was not detected in samples from the Bay of Islands (Nielsen and Nathan 1975). Hg concentrations in mussels from the Marlborough Sounds were 0.03 mg kg^{-1} in the 1980's (NZFSA 1981; Fenaughty et al. 1988) and 0.3 mg kg^{-1} dry weight (equivalent to approximately 0.04 mg kg^{-1} wet weight) was recorded in mussels from Miramar Wharf in Wellington (Kennedy 1986). However, the measurement from Kennedy (1986) is based on digestive gland tissue and so is therefore not directly comparable. Hg was below the level of detection (< 0.01 mg kg^{-1}) in whole mussel samples from Seatoun Harbour in Wellington (current study, Chapter 6).

Most Hg concentrations reported in overseas studies were similar to those reported in *P. canaliculus* (Beliaeff et al. 1998; Baraj et al. 2003; Otchere et al. 2003) However, Hg concentrations in *P. viridis* (3.47 - 32.98 mg kg^{-1}) and *M. edulis* (0.42 - 13.93 mg kg^{-1}) from China were extremely high and pose a high level of risk to consumers (Fung et al. 2004).

As concentrations in *P. canaliculus* from the Bay of Islands ranged from 1.55 to 2.97 mg kg^{-1}. As was not measured in the study by Nielsen and Nathan (1975). The As concentration in *P. canaliculus* from Miramar Wharf in Wellington (Kennedy 1986) was 10.5 mg kg^{-1} dry weight (equivalent to approximately 1.30 mg kg^{-1} wet weight), which is less than contemporary (2005) estimates in mussels from Seatoun (Wellington) of 3 mg kg^{-1}. As noted previously, the concentrations reported by Kennedy (1986) are based on measurements taken from the digestive gland and are therefore likely to be overestimated compared to measurements from all mussel

tissues. Nielsen and Nathan (1975) illustrated the variability in metal concentrations that can occur over relatively short distances (e.g. oysters in the Brampton Shoals in the Bay of Islands have nearly twice the Cd levels of oysters from Wairoa Bay, less than 1 km away). Overall, Hg and As concentrations detected in the Bay of Islands, although slightly elevated compared to the NZTDS, are relatively low compared to measurements from other locations around New Zealand.

As concentrations in *P. viridis* and *M. edulis* from along the east coast of China were comparable to concentrations detected in *P. canaliculus*. Average As concentrations ranged from 1.60 to 3.03 mg kg^{-1} (*P. viridis*) and from below the limit of detection to 2.62 mg kg^{-1} (*M. edulis*; Fung et al. 2004). As concentrations were below the FSANZ limit (2.0 mg kg^{-1}, inorganic As).

7.4.2.2 Cd concentrations in shellfish

Cd concentrations in individual mussels from the Bay of Islands ranged from 0.09 to 1.37 mg kg^{-1}. Which are below the current FSANZ limits, although the Food and Agriculture organization of the United Nations are in the process of revising maximum Cd levels for molluscs down to 1.0 mg kg^{-1} (www.fao.org; September 2006). The average Cd concentration detected from the Bay of Islands was 0.1 mg kg^{-1} for the majority of sites, with the exception of the elevated average concentration at Urapukapuka (0.754 mg kg^{-1}). The high Cd concentrations at Urapukapuka were unexpected because this site is remote from obvious anthropogenic sources of pollution. One possible explanation is the presence of a thermal sea vent. Hydrothermal vents are common in the Bay of Plenty (Propp et al. 1994) but have yet to be reported in the Bay of Islands. The source of Cd at this remote site in the Bay of Islands may require further investigation. The average Cd concentration in *P. canaliculus* as surveyed by Nielsen and Nathan (1975) was 0.3 mg kg^{-1} (range 0.10 to 1.0 mg kg^{-1}). The average Cd concentration in mussel digestive tissue from Miramar (n = 18) was 0.5 mg kg^{-1} wet weight, equivalent to approximately 0.062 mg kg^{-1} (Kennedy 1986), compared to approximately 0.1 mg kg^{-1} in Seatoun (current study, Chapter 6). Unpublished data from the New Zealand Food Safety Authority (NZFSA) collected since 1991 shows a trend for decreasing Cd concentrations in *P. canaliculus* from various unspecified North Island locations and the Marlborough Sounds, from 0.8 mg kg^{-1} in the early 1990's to an average of approximately 0.2 mg kg^{-1} in 2003. In general, Cd concentrations at most sites in the

Bay of Islands are relatively low compared to sites around the rest of New Zealand. Excluding samples from Urapukapuka, Cd concentrations in *P. canaliculus* were comparable to or less than Cd concentrations measured in other mussel species worldwide. Slightly elevated Cd concentrations were detected in China (Fung et al. 2004), Bohai Sea (Liang et al. 2004), Scotland, Ireland and UK (Bebianno and Machado 1997) and Italy (Bargagli et al. 1985). However, all values were below the maximum limit set by FSANZ, with the exception of oysters from the Bohai Sea (Liang et al. 2004).

7.4.2.3 Pb concentrations in shellfish

The average Pb concentration in *P. canaliculus* from Wellington was 0.2 mg kg^{-1} (current study, Chapter 6). This is 2X higher than the highest value obtained from the Bay of Islands (0.1 mg kg^{-1}) but still well below the FSANZ guidelines (2.0 mg kg^{-1}). High Pb concentrations near cities have been recorded previously. For example, the average Pb concentration reported in *P. canaliculus* in Nielsen and Nathan (1975) was 1.8 mg kg^{-1} (range 0.1-7.8 mg kg^{-1}) with the highest Pb concentrations around the vicinity of major cities, including Wellington e.g. 7.8 mg kg^{-1} at the Ngauranga Stream mouth and 6.4 mg kg^{-1} at the Inter Island Ferry Wharf. Excluding the relatively high Pb concentrations in the Wellington region (all >1.5 mg kg^{-1}), the Pb concentrations at the other locations were all <1.5 with an average of 0.8 mg kg^{-1}. The Pb concentration at Russell in the Bay of Islands was 1.3 mg kg^{-1} (Nielsen and Nathan 1975), in contrast to contemporary measurements from the Bay of Islands which were all < 0.1 mg kg^{-1}. It may be useful to test shellfish from more locations within the Bay of Islands, including Russell to compare metal concentrations with historical measurements.

Pb concentrations in *P. canaliculus* were less than concentrations detected overseas. In *P. viridis* samples from various locations within Asia, some samples exceeded the FSANZ limit (Yap et al. 2004b), which could cause human health concerns.

7.4.2.4 Sn concentrations in shellfish

Sn was not detectable (< 0.1 mg kg^{-1}) in any mussel samples from the Bay of Islands. Because tributyltin was a common ingredient in anti-fouling paint it was expected that trace amounts would still remain. Although Sn was below the level of

detection in seawater and organisms, it is possible that Sn is still present in sediment (Coelho et al. 2002). There are no historical records for Sn concentrations in *P. canaliculus* for comparison.

7.4.3 Dietary Exposure to heavy metals

A sub-section of the New Zealand community, in particular comprising Māori, Pacific Island and Asian people, consume more than 'average' amounts of seafood (Ministry of Health 1999) and therefore experience potentially higher exposures to heavy metals and other seafood related pollutants. In some regions of New Zealand, particularly in low-socioeconomic areas of Northland and the Bay of Islands, kaimoana (seafood) is a major part of the diet. A survey in Northland found that 11% of households reported collecting seafood from the wild more than once a week, 31% collected seafood at least weekly, and 52% reported collecting seafood at least fortnightly (Hay 1996). Calculations based on the metal concentrations present in mussels collected from the Bay of Islands suggest that a large number of mussels would need to be consumed to exceed the PTWI for all heavy metals assayed here (Table 7.4). For example, based on a typical Cd concentration of 0.1 mg kg^{-1} a 70 kg adult from the Bay of Islands would need to consume approximately 4.9 kg of greenshell mussel flesh per week to attain a weekly Cd intake equivalent to the PTWI (Table 7.4). It is unknown whether this amount of mussels is realistic, even for large volume consumers. However, if mussels were collected from the most Cd contaminated site in the Bay of Islands (Urapukapuka), the amount of mussels required to exceed the PTWI decreases to 0.650 kg which is plausibly achievable. At this stage it is impossible to determine whether consumption of shellfish in some sectors of society is at a level that could cause disease. Further information on the type, amount and frequency of shellfish consumed by these groups is required to adequately assess the risk.

7.4.3.1 Influence of location on metal concentrations in P. canaliculus

The sites chosen for greenshell mussel testing in the current study are associated with most of the kaimoana gathered in this region. Therefore, the concentrations of metals reported here will provide a good indication of the amount of heavy metal exposure that human populations at these locations are subject to via the consumption of greenshell mussels. The locations were expected to show

variation in pollution, in relation to the fluctuating influence of and distance to anthropogenic pollution sources. For example, the Opua Marina and Wharf sites were expected to be influenced by boating activities whereas Waitangi Bridge is influenced by geothermal and industrial inputs from the Waitangi River. The high Cd and As concentrations measured in mussels from Urapukapuka are of concern. Locals assume that because Urapukapuka is approximately 15 km off-shore that mussels are not exposed to anthropogenic sources of pollution and are likely to be relatively uncontaminated. However, the results of this study have demonstrated that mussels from Urapukapuka contain elevated levels of Cd. Although Cd concentrations in mussels from Urapukapuka were below FSANZ limits, the concentration is sufficient to exceed the PTWI dependant on the amount of mussels collected and consumed from this site.

7.4.3.2 Influence of inclusion of other shellfish in diet

Shellfish accumulate heavy metals at different rates. For example, scallops and oysters, in particular Bluff oysters, contain much higher concentrations of Cd (1.3 – 5.3 mg kg^{-1}) than greenshell mussels (0.09 -0.95 mg kg^{-1})(Brooks and Rumsby 1965; Brooks and Rumsby 1967; Nielsen and Nathan 1975; Frew and Hunter 1996). Some previous studies suggest that Cd from Bluff oysters is not absorbed as efficiently as other dietary sources of Cd (Sharma et al. 1983; Nordberg et al. 1986; McKenzie-Parnell et al. 1988). The Food Standards Authority has deemed that Bluff oysters and Queen Scallops present a special case, and the higher levels of Cd in these species do not present an unacceptable risk to New Zealand consumers (pers. comm. 2006, John Reeve, Principal Adviser, Toxicology, NZ Food Safety Authority). This is why Bluff oysters and Queen Scallops are not subject to the FSANZ limits. Although Cd in whole blood and urine did not increase after consumption of oysters (Sharma et al. 1983), it is possible that Cd bound to MT may be quickly transported to the kidney (Nordberg et al. 1986) where damage to kidney function could occur. Although further studies were suggested, to my knowledge these have not been conducted.

Oysters were also highlighted in the New Zealand Total Diet Survey (NZTDS) as a potential source of Cd accumulation in humans. The inclusion of only 2 oysters per fortnight in simulated diets increased estimated Cd intake from 1.3 to 1.8 $\mu g\ kg^{-1}$ bw wk^{-1} (Vannoort and Thomson 2005). Cd dietary exposure calculated

in the NZTDS for an adult male (including oysters) are below those of the Republic of Korea, similar to those of the Czech Republic, and above those of Australia, the USA, the UK, France and Basque Country (Vannoort and Thomson 2005). If oysters are excluded from the diet, the 2003/2004 NZTDS exposures to Cd are then below or similar to all countries except France (Vannoort and Thomson 2005). These results suggest that oysters pose a greater risk to Cd accumulation than mussels, although the exact risk will again depend on the type, amount and frequency of oyster consumption.

7.4.3.3 Overseas studies

Compared to overseas studies, metal concentrations reported in *P. canaliculus* are comparable to, or less than, metal concentrations reported in other mussels. In Singapore, a study of 20 different seafood types found that the mean daily intake of contaminants from seafood for the general population, did not exceed the oral reference dose set by the U.S. Food and Drug Administration (Bayden et al. 2005). Intake of As, Cd Hg, and Pb in Spain from the consumption of 14 edible species was estimated and the PTWI for most of the population was within acceptable limits, although MeHg concentrations for boys were exceeded (Falcó et al. 2006). This study highlights the particular risk of heavy metal accumulation in children who because of their smaller body weight require less seafood to exceed regulatory limits. Consumption of TBT, Cu, Cd, Zn and As contaminated seafood were found to pose a potential health risk to fisherman in Taiwan (Chien et al. 2002), highlighting the increased risk of heavy metal accumulation in sectors of the population that consume more seafood than 'average'.

The Cd concentrations detected in mussels from Urapukapuka were relatively high compared to most international studies, indicating an increased risk of Cd exposure in mussels collected from some New Zealand sites. This increased Cd risk is reflected in the comparison of simulated New Zealand diets with overseas diets (Section 7.4.3.2). Although this diet comparison used oysters, the potential for increased Cd exposure in New Zealand depending on the amount of seafood consumed is demonstrated. Concentrations of Pb, Cd and Hg from some international sites exceeded FSANZ limits, and could pose a risk to human health depending on the amount of shellfish consumed. In New Zealand the risk of dietary exposure to heavy metals is predominantly from Cd and not Hg, Pb or As. Based on

the concentrations of metals detected in the current study, seafood consumers in New Zealand have a similar or lower risk of exceeding the PTWI for Pb, As and Hg than seafood consumers from other countries. However, the relative risk of Cd exposure from seafood consumption in New Zealand is higher than many overseas sites if the Cd concentration from Urapukapuka is used for comparison or comparable to other countries if the average from other sites is used.

7.4.4 Implications for human health in New Zealand

Currently throughout New Zealand, targeted public health messages are aimed at women of child bearing age and caregivers of toddlers and infants, making recommendations on the frequency and type of fish that should be consumed, thereby ensuring Hg exposure is kept to a minimum for these vulnerable groups (developing foetuses and children; Vannoort and Thomson 2005). A similar message may need to be considered for sectors of the population which are recognised as eating large amounts of shellfish to reduce exposure to Cd e.g. recommending that children do not consume Bluff oysters and limiting consumption of other shellfish. However, adequate information on the eating habits of 'at risk' groups should be collected and assessed before such actions are taken (Judd et al. 2004). It is acknowledged that established factors such as socio-economic status, poor diet and high smoking rates (approximately 40-49% of adult Māori; Statistics New Zealand 1998) contribute to the over-representation of Māori in poor health statistics. However, the effect of high shellfish consumption on the accumulation of heavy metals and possible associated effects on Māori health may require some attention.

In addition other culturally important foods such as watercress (Robinson et al. 2003) and eels (Chisnall 1997) have been found to contain high levels of As and Hg respectively. A link between traditional food sources high in Cd and diabetes has been postulated in Australian Aboriginals (Satarug et al. 2000). High Cd concentrations have also been associated with prostate cancer (Gray et al. 2005) and high Hg and low Zn concentrations have been linked to autism (Yorbik et al. 2004; Fido and Al-Saad 2005). Although previous studies have found a link between shellfish consumption and protection from certain diseases, e.g. myocardial infarction (Yuan et al. 2001) the relative risks and benefits of seafood consumption are hard to assess (Foran et al. 2006).

7.5 Conclusions

P. canaliculus from Urapukapuka-Rawhiti, Opua Marina, Waitangi Bridge and Opua Wharf in the Bay of Islands, all contained metal concentrations well below FSANZ limits. Based on the concentrations of metals detected in the current study, the amount of metal consumed through a typical diet containing shellfish would be below the PTWI. However, some sectors of the community, including Māori, Pacific Islanders and Asians, consume more seafood than the general public. Calculations suggest that as little as 0.650 kg of greenshell mussels are required from the most contaminated site (Urapukapuka) to exceed the PTWI. In addition some species of shellfish, e.g. Bluff oysters and scallops, contain higher concentrations of Cd than mussels and may present an even greater risk to dietary metal accumulation. It is possible that the PTWI could be exceeded in some populations groups, dependent on how much, how often, and which types of seafood are consumed. A comprehensive survey of the eating habits of Māori, Pacific Islanders and Asians, is suggested to determine if seafood is consumed in quantities that could cause health concerns.

CHAPTER 8: Summary/ Conclusions/ Future Directions

8.1 Overall summary of intent

The aim of this study was to investigate the suitability of *P. canaliculus* to act as a biomonitor for heavy metal pollution and to detect and identify protein biomarkers in *P. canaliculus* which indicate exposure to heavy metals. *MT* gene nucleotide sequences were isolated from *P. canaliculus* because of their known role in heavy metal binding and detoxification. A proteomic approach was developed to detect *P. canaliculus* proteins involved in response to salinity stress. The proteomic methodology was then applied to study the abundance of *P. canaliculus* proteins after exposure to Cd and Hg in the laboratory. The performance of *P. canaliculus* biomarkers selected from laboratory based heavy metal experiments were investigated under field conditions in mussels from the Bay of Islands. Finally, the risk of metal intake from *P. canaliculus* to human consumers from the Bay of Islands was discussed.

8.2 *MT* gene and amino acid sequences in *P. canaliculus*

8.2.1 Conclusions/Summary of findings

- At least 9 different MT amino acid sequences were determined from gene sequences isolated from *P. canaliculus*. Unusually for MT, these sequences contain some atypical residues (H and Y).
- AP1, metal responsive element (MRE) and promoter sequences were found in the 5' untranslated region, suggesting a complex transcription mechanism.
- There was some evidence of heterozygosity, which may have implications for tissue specific or individual specific response to heavy metals.
- Using the FASTA algorithm *P. canaliculus* MTs were most similar to those found in other molluscs indicating that *P. canaliculus* MT belongs to the Class I MTs.

8.2.2 Future Directions

MT sequences from *P. canaliculus* have been identified and reported for the first time. It would be interesting from an evolutionary perspective to determine whether there are 10 and 20 kDa MT molecular-weight classes in *P. canaliculus* as has been determined for MT proteins from the blue mussel *M. edulis*. The presence or absence of dimers could be established using gel-permeation or anion exchange chromatography (Mackay et al. 1999). Alternatively, the mass of purified MT could be analysed using mass spectrometry, and the amino acid sequence to differentiate between 10 and 20 kDa MT forms could be determined by tandem mass spectrometry.

Because partial *MT* sequences were isolated in the current study, none of the *MT* genes or their products were fully characterised. *MT* cDNA or mRNA produced by degenerate primers could have been used to isolate full length MT sequences as has been successfully applied in previous studies (e.g. Khoo and Patel 1999). This would allow the full length MT sequence from *P. canaliculus* to be characterised and could facilitate more accurate FASTA analysis to determine the evolutionary relationship of *P. canaliculus* MT to MTs from other species. Full length MT sequences could also be used to determine relationships between paralogous and orthologous MT sequences using phylogenetic trees. In addition primers designed from full length MT sequences could be used to develop RT-PCR approaches to quantify MT transcripts in response to Hg or Cd exposure. Or a cDNA library for *P. canaliculus* could be created using oligo (dT) or random hexamers as primers under conditions or interest. In the current study PCR products were amplified using *Taq* DNA polymerase, these results could be compared to those obtained using high fidelity proof-reading polymerase.

Although MT gene and protein polymorphism between individuals was suggested in the current study polymorphism was not proven. In the oyster *C. gigas*, single strand conformation polymorphism in exon 2 and 3 of the MT1 gene has been associated with tolerance to metals in both experimental and control populations (Tanguy et al. 2002). A similar PCR approach could be modified and applied to *P. canaliculus* to test for polymorphism in this species, and examine possible links between certain genotypes and increased tolerance of heavy metals.

8.3 Protein biomarkers for salinity stress

8.3.1 Conclusions/Summary of findings

- Approximately 115 gill and foot proteins showed significant ($p < 0.01$) differences in abundance due to salinity and experimental conditions, and could possibly act as biomarkers for salinity stress.
- Unexpectedly, experimental conditions affected foot tissue to a greater degree than gill tissue, inducing a response in approximately 10% of the proteins isolated from foot tissue compared to only approximately 1% from gill tissue. Foot tissue therefore accounted for most (96/115) of the reported differences.
- Differences in the protein abundance between control and ambient groups were responsible for most of the protein responses in foot tissue, which was contrary to predictions postulating that the greatest differences in protein abundance would be between control and low salinity groups.
- Variation in ammonia concentration and oxygen consumption by different tissues and between experimental groups are discussed as possible explanations to account for the differences in protein abundance, but no firm conclusions could be drawn.
- Two isoforms of actin and two isoforms of tropomyosin were identified, in this and previous proteomic studies. Further protein identifications could assist elucidation of possible roles for these proteins that differ among tissue with stress.

8.3.2 Future Directions

Very few proteins were identified in the current study. Although, many proteins produced good spectra for MALDI mass fingerprinting, the spectra were unable to be matched to protein and DNA databases. Initial trials using tandem mass-spectrometry were no more successful than MALDI. N-terminal sequencing (Edman degradation) was not used in the current study, although this method may have led to more successful protein identifications. However, this process is slow and expensive. In addition the N-terminus of some proteins may be biologically or chemically blocked. The ability of tandem mass spectrometry to generate protein sequences using relatively small amounts of protein suggests that it is a better option

for protein identification than either MALDI or Edman degradation. As shellfish and other invertebrate databases become more complete it is likely that further identifications will be possible. The spot maps obtained in this study might also be subsequently compared to spot maps and protein identifications from new research. However, this will only be possible for distinctive spot patterns (e.g. tropomyosin) or if future studies use similar isoelectric strips, gels and electrophoresis conditions. The MALDI mass fingerprints for candidate biomarker proteins could also be made available on a server for comparisons with profiles produced by other researchers. The creation of mussel cell lines could be an invaluable tool for studying biochemical and molecular responses to environmental challenges such as heavy metal and other toxicants. The challenges in protein identification in *P. canaliculus* may invite an opportunity to gain support for a national genome sequencing project.

Many gels were removed from the salinity experiment due to poor spot resolution, leading to reduced power in the experiment to detect significant changes in protein abundances. Protein purification, for example with a 2D clean-up kit, should produce better gels and results. Currently FDR adjustments were not applied to the data which means that some spots could have been detected as significant changes by chance. It would be relatively simple to repeat the DIGE experiment using tissue or protein samples from the current study. This would allow the authenticity of the foot protein results to be re-examined.

The unexpected results obtained during the salinity experiment, namely the large number of protein changes which occurred in foot tissue and the greatest number of differences in protein abundance being found in comparisons between control and ambient salinity groups, may warrant further investigation. The experiment could be repeated and physiological components including oxygen and ammonium consumption could be monitored to ensure that these factors do not unduly influence the study.

8.4 Protein biomarkers for heavy metal contamination

8.4.1 Conclusions/Summary of findings

- Approximately 117 gill proteins showed significant ($p < 0.01$) differences in abundance after exposure to 0.486 mg l^{-1} Cd or 34.3 μg l^{-1} Hg for 3 days in the laboratory.

- Hg uptake was highest in gill tissue, and reached an average maximum value of 5.21 mg kg^{-1} after 2 days, Hg concentrations decreased in all tissues thereafter.
- Cd uptake was maximal in digestive gland, and accumulation in this tissue was linear and was still increasing after 3 days (42.3 mg kg^{-1} maximum value obtained during the experiment).
- Metal accumulation patterns in *P. canaliculus* were dependent on tissue type and metal species. However, both Cd and Hg appear to have similar effects on protein change in gill tissue.
- Some of the varying proteins could possibly be used as biomarkers for heavy metal stress in *P. canaliculus* even prior to identification.

8.4.2 Future Directions

MT proteins from *P. canaliculus* were not isolated in the current study. There are a variety of ways that MT could be isolated or quantified e.g. affinity chromatography through metal-chelating columns, differential pulse polarography, radioimmunoassay, spectrophotometry, ELISA at the level of protein expression or as a function of the metals bound to MT (Cajaravill et al. 2000).

Identification of MT proteins from *P. canaliculus* tissue was trialled using a fluorescent labelling kit but the results were ambiguous and inconsistent. Boutet et al. (2002) synthesised a recombinant MT1 antigen and used it to develop rabbit anti-CgMT1 IgG, which reacted with MT1 isolated from *C. gigas*. It is possible that a similar Western Blot technique could be developed to react with MT from *P. canaliculus*.

Variations in protein separation patterns between individuals within the same treatment were obtained indicating significant genetic polymorphism in the *P. canaliculus* population (Figure 8.1). Allelic polymorphism was detected using DeCyder, although it did not reach significance in the *t*-test. In the example shown (Figure 8.2), one individual mussel has two proteins (A1 and A2) that were not highly abundant in any other individuals. The *t*-test value was not significant for either of these spots and for this reason these proteins would not have been investigated further if they had not been detected manually. There may be many such polymorphic proteins with rare alleles that have not been investigated because

the *t*-test value was not significant. This is of concern because MT may have multiple protein isoforms and alleles which may not have been detected as significant using the DeCyder software. Further work is required.

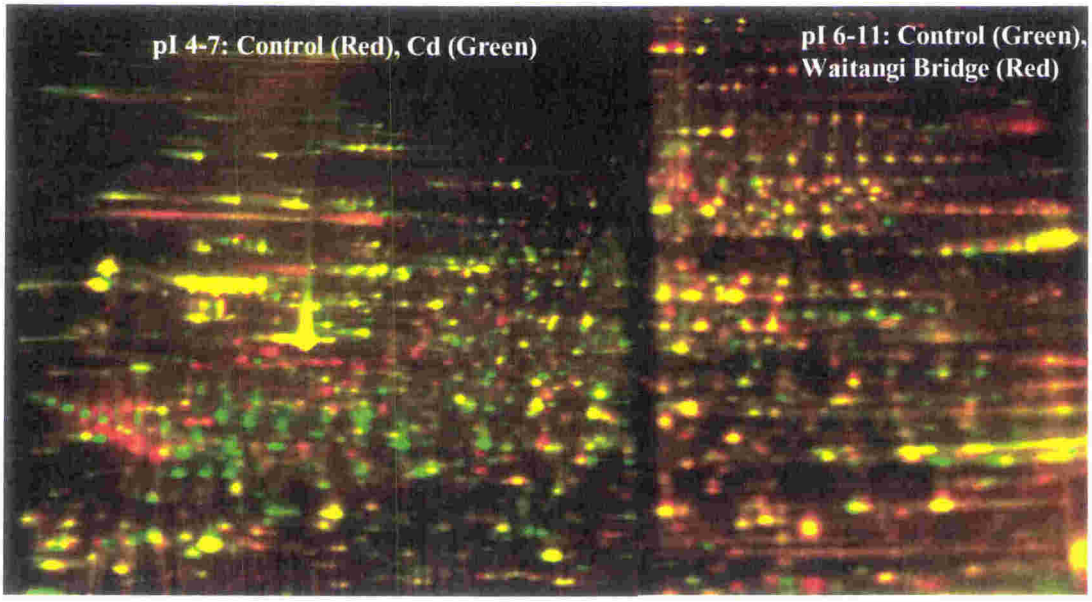


Figure 8.1: Gel-based evidence of genetic polymorphism in *P. canaliculus*
Protein spots which are yellow indicate similar protein abundance. Proteins which are different colours (i.e. red or green) indicate treatment induced differences in protein abundance (e.g. between control and treated) or may be unique proteins specific to the individual (i.e. genetic polymorphism).

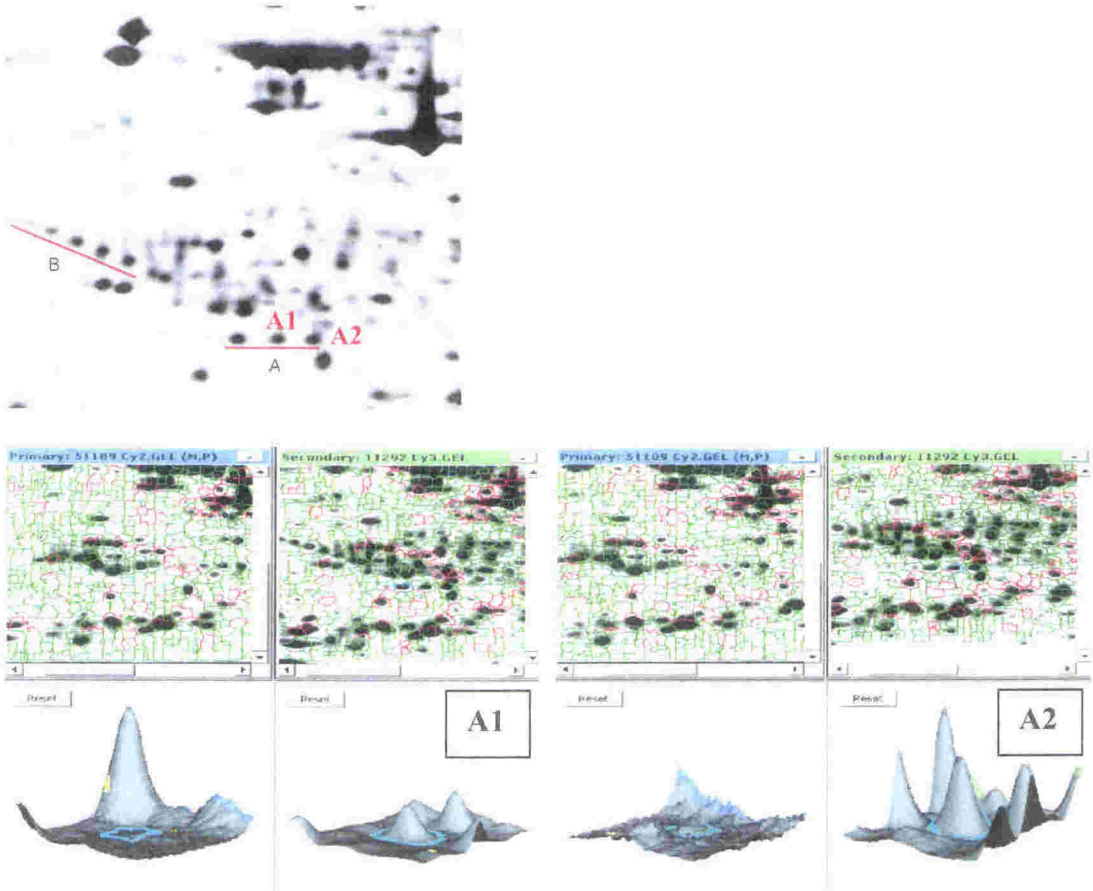


Figure 8.2: Further evidence of genetic polymorphism in *P. canaliculus*

8.5 Validation of protein biomarkers for heavy metal contamination from field sites within the Bay of Islands

8.5.1 Conclusions/Summary of findings

- Approximately 45 gill proteins showed significant ($p < 0.05$) differences in abundance among control samples, laboratory exposed heavy metal samples (Hg and Cd) and samples collected from 4 field sites within the Bay of Islands.
- A control site from the Bay of Islands without metals was absent from the current study so the association of these proteins with pollution could not be unequivocally established.
- Further work is required to validate these proteins as markers of heavy metal pollution, although this study suggests that biomarkers for heavy metal pollution in the field could be developed in *P. canaliculus*.

8.5.2 Future Directions

The detection of potential protein biomarkers from field sites in the current study was hampered by the lack of a control site from the Bay of Islands because unexpectedly the planned control site (Urapukapuka) had high levels of Cd and As. Most of the proteins which were detected as significant were at a low abundance in the control group (from the metal experiment) and were higher in the remaining six groups (Bay of Island sites and the laboratory heavy metal exposed groups). In general these proteins also occurred at a higher abundance in the Bay of Islands groups than Wellington groups (metal experiment), but the significance of this observation is unknown. Samples from sites around New Zealand which are not subject to geothermal influences or anthropogenic sources of pollution could shed some light on this issue, by providing examples for comparison, thereby providing further data to determine whether the selected proteins are appropriate biomarkers for heavy metal pollution.

The use of pooled samples in the current study reduced the power of the experiment and necessitated the use of R to deal with independence within the data. However, despite these problems, the proteins detected in the current study could offer a tool to help determine whether an area is subject to pollution. Further validation and testing from other regions throughout New Zealand is required.

8.6 Dietary exposure to heavy metals from mussels in the Bay of Islands

8.6.1 Conclusions/Summary of findings

- The concentrations of Cd, Hg, As, Pb and Sn in *P. canaliculus* samples taken from Urapukapuka-Rawhiti, Opua Marina, Waitangi Bridge and Opua Wharf (in the Bay of Islands) were well below the Food Standards Australia and New Zealand (FSANZ) limits and therefore pose a low risk to an 'average' diet including shellfish.
- Māori, Pacific Islanders and Asians consume a larger quantity of seafood than the general public which may put them at a greater risk. A survey of the frequency, amount and species of shellfish consumed is required to enable an adequate risk assessment for these groups to be made.

8.6.2 Future Directions

As discussed in the summary/conclusions, the impact of heavy metals through dietary intake of shellfish on Māori, Pacific Islanders and Asians can not currently be adequately assessed because not enough is known about the eating habits of these groups. Although anecdotal evidence suggests exposure could potentially reach levels for concern, there is no firm evidence. Cd levels in scallops in particular could be high enough to exceed the PTWI and the assumption that the absorption of dietary Cd from scallops by humans is poor, as is the case for Bluff oysters, has not to the best of my knowledge been verified for scallops.

Iwi in the Bay of Islands have expressed an interest in a continued partnership with ESR. This unlocks a number of potentially beneficial collaborations between scientists and the Māori community in the Bay of Islands. Currently in New Zealand many projects are developed without iwi input during the early planning stages. Iwi therefore do not get an opportunity to develop a project that would be of interest, relevance and benefit to them. Instead Māori generally get 'token' involvement once the scientists have already planned what they want to do. Attitudes are beginning to change which could lead to advantageous outcomes for everyone. Because Māori have a different worldview and scientific framework to 'western' science, I believe synergies exist which will give science in New Zealand a competitive edge.

The rich oral traditions of Māori provide records of historical ecological management practices which could be used to inspire and inform new studies, but many Māori are anxious about revealing their sacred knowledge to scientists most of whom are not Māori. Having a relationship built on trust between Māori and scientists can facilitate the sharing of ideas. This concept has been clearly illustrated to me during my involvement with ESR and their project with Ngāpuhi. During the end users meeting at ESR, Emma Gibbs from the Ngāpuhi Marae committee contributed a number of interesting comments and observations which could warrant further investigation, including taboo on harvesting certain species of fish and shellfish during particular seasons and a sudden respiratory illness in pregnant women bought on by consuming shellfish.

8.7 Concluding Remarks

Preliminary work has determined that *P. canaliculus* is suited to act as a biomonitor for marine pollution in New Zealand. A number of potential biomarkers were detected under laboratory conditions in *P. canaliculus* and initial trials in the Bay of Islands indicate that these biomarkers may be suitable in the field. However, further work is needed to validate these biomarkers.

REFERENCES

Abel P. (1976) Effect of some pollutants on the filtration rate of *Mytilus*. Marine Pollution Bulletin 7: 228-231

Ahrens M., Depree C. (2004) Inhomogeneous distribution of polycyclic aromatic hydrocarbons in different size and density fractions of contaminated sediment from Auckland Harbour, New Zealand: an opportunity for mitigation. Marine Pollution Bulletin 48: 341-350

Allah A., Wanas M., Thompson S. (2003) Dissolved heavy metals, lead, cadmium and mercury, accumulate in the body of the schistosome vector, *Biomphalaria glabrata* (Gastropoda: Pulmonata). Journal of Molluscan Studies 69: 35-41

Anderson M. (2001) A new method for non-parametric multivariate analysis of variance. Austral Ecology 26: 32-46

Apte S. (2001) Population genetic variation in the New Zealand greenshell mussel *Perna canaliculus*. PhD thesis. School of Biological Sciences, Wellington

Australian and New Zealand Environment and Conservation Council and Agriculture and Resource Management Council of Australia and New Zealand. (2000) Australian and New Zealand Guidelines for Fresh and Marine Water Quality, Canberra, ACT

Baraj B., Niencheski L., Corradi C. (2003) Trace metal content trend of the mussel *Perna perna* (Linnaeus, 1758) from the Atlantic coast of southern Brazil. Water, Air and Soil Pollution 145: 205-214

Bargagli R., Baldi R., Leonzio C. (1985) Trace metal assessment in sediment, molluscs and leaves in the Bay of Follonica (Italy). Marine Environmental Research 16: 281-300

Barsyte D., White K., Lovejoy D. (1999) Cloning and characterization of metallothionein cDNAs in the mussel *Mytilus edulis* L. digestive gland. Comparative Biochemistry and Physiology C 122: 287-296

Bayden S., Koroleva E., Lee H., Obbard J. (2005) Persistent organic pollutants and heavy metals in typical seafoods consumed in Singapore. Journal of Toxicology Environmental Health A 68: 151-166

Bayne B. (1976) Watch on mussels. Marine Pollution Bulletin 7: 217-219

Bebianno M., Langston W. (1993) Turnover rate of metallothionein and cadmium in *Mytilus edulis*. Biometals 6: 239-244

Bebianno M., Langston W. (1998) Cadmium and metallothionein turnover in different tissues of the gastropod *Littorina littorea*. Talanta 46: 301

Bebianno M., Serafim M. (1998) Comparison of metallothionein induction in response to cadmium in the gills of the bivalve molluscs *Mytilus galloprovincialis* and *Ruditapes decussatus*. The Science of the Total Environment 214: 123-131

Bebianno M.J. (1995) Effects of pollutants in the Ria Formosa Lagoon, Portugal. Science of the Total Environment 171: 107

Bebianno M.J., Machado L.M. (1997) Concentrations of metals and metallothioneins in *Mytilus galloprovincialis* along the south coast of Portugal. Marine Pollution Bulletin 34: 666

Beliaeff B., O'Connor T., Claisse D. (1998) Comparison of chemical concentrations in mussels and oysters from the United States and France. Environmental Monitoring and Assessment 49: 87-95

Benjamini Y., Hochberg Y. (2000) The adaptive control of the false discovery rate in multiple comparison problems. Journal of Educational and Behavioural Statistics 25: 60-83

Berger B., Hunziker,P.E., Hauer,C.R., Birchler,N. and Dallinger,R. (1995) Mass spectrometry and amino acid sequencing of two cadmium-binding metallothionein isoforms from the terrestrial gastropod *Arianta arbustorum*. Biochemical Journal 311 (Pt 3): 951-957

Binz P., Kagi J. (1999) Molecular evolution of class I metallothioneins: evidence from non-coding DNA regions Metallothionein IV. Birkhauser, Basel, pp 7-13

Blackmore G., Wang W. (2003) Inter-population differences in Cd, Cr, Se, and Zn accumulation by the green mussel *Perna viridis* acclimated at different salinities. Aquatic Toxicology 62: 205-218

Bodin N., Burgeot T., Stanisiere J.Y., Bocquene G., Menard D., Minier C., Boutet I., Amat A., Cherel Y., Budzinski H. (2004) Seasonal variations of a battery of biomarkers and physiological indices for the mussel *Mytilus galloprovincialis* transplanted into the northwest Mediterranean Sea. Comparative Biochemistry and Physiology Part C: Toxicology & Pharmacology 138: 411-427

Bolton-Ritchie L. (2003) The effect of stormwater discharge on the nearshore benthic environment of inner Wellington Harbour. PhD thesis. School of Biological Sciences, Victoria University of Wellington Wellington

Bolton-Ritchie L., Chague-Goff C., Gardner J., Goff J. (1999) Wellington's storm drain sediments - spreading it around. NZWWA 1999 Conference papers

Boonyatumanond R., Jaksakul A., Puncharoen P., Tabucanon M. (2002) Monitoring of organochlorine pesticides residues in green mussels (*Perna viridis*) from the coastal area of Thailand. Environmental Pollution 119: 245-252

Borchardt T. (1983) Influence of food quantity on the kinetics of cadmium uptake and loss via food and seawater in *Mytilus edulis*. Marine Biology 76: 67-76

Botherway K., Gardner J.P.A. (2002) Effect of storm drain discharge on the soft shore ecology of Porirua Inlet, New Zealand. New Zealand Journal of Marine and Freshwater Research 36: 241-255

Boutet I., Tanguy A., Auffret M., Riso R., Moraga D. (2002) Immunochemical quantification of MTs in marine mollusks: characterization of a metal exposure bioindicator. Environmental Toxicology and Chemistry 21: 1009-1014

Bradley B., Shrader E., Kimmel D., Meiller J. (2002) Protein expression signatures: an application of proteomics. Marine Environmental Research 54: 373-377

Breathnach R., Chambon P. (1981) Organization and expression of eukaryotic split genes coding for proteins. Annual Review of Biochemistry 50: 349-383

Brooks R., Rumsby M. (1965) The biogeochemistry of trace element uptake by some New Zealand bivalves. Limnology and oceanography 10: 521-527

Brooks R., Rumsby M. (1967) Studies on the uptake of cadmium by the oyster, *Ostrea sinuata* (Lamarck). Australian Journal of Marine and Freshwater Research 18: 53-61

Brouwer M., Enghild J., Hoexum-Brouwer T., Thogersen I., Truncali A. (1995) Primary structure and tissue-specific expression of blue crab (*Callinectes sapidus*) metallothionein isoforms. Biochemical Journal 311: 617-622

Brzoska M.M., Moniuszko-Jakoniuk J. (2001) Interactions between cadmium and zinc in the organism. Food and Chemical Toxicology 39: 967-980

Buckley B.A., Owen M.-E., Hofmann G.E. (2001) Adjusting the thermostat: the threshold induction temperature for the heat-shock response in intertidal mussels (genus *Mytilus*) changes as a function of thermal history. Journal of Experimental Biology 204: 3571-3579

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

Calow P. (1991) Physiological costs of combating chemical toxicants: ecological implications. Comparative Biochemistry and Physiology 100C: 3-6

Carpene E., George S. (1981) Absorption of cadmium by gills of *Mytilus edulis* (L.). Molecular Physiology 1: 23-34

Casas S., Cossa D., Gonzalez J.-L., Bacher C., Andral B. (2004) In situ kinetics of mercury bioaccumulation in *Mytilus galloprovincialis* estimated by transplantation experiments. 7th International Conference on mercury as a global pollutant, Ljubljana, Slovenia

Caussy D., Gochfeld M., Gurzau E., Neagu C., Ruedel H. (2003) Lessons from case studies of metals: investigating exposure, bioavailability, and risk. Ecotoxicology and Environmental Safety 56: 45-51

Centeno J., Gray M., Mullick F., Tchounwou P., Tseng C.-H. (2005) Arsenic in drinking water and health issues. In: *Metal contaminants in New Zealand: Sources, Treatments, and Effects on Ecology and Human Health*. Resolutionz Press, Christchurch, New Zealand, pp 415-439

Ceratto N., Dondero F., van de Loo J.-W., Burlando B., Viarengo A. (2002) Cloning and sequencing of a novel metallothionein gene in *Mytilus galloprovincialis* Lam. Comparative Biochemistry and Physiology C 131: 217-222

Chan M., Othman R., Zubir D., Salmijah S. (2002) Induction of a putative metallothionein gene in the blood cockle, *Anadara granosa*, exposed to cadmium. Comparative Biochemistriy and Physiology C 131: 123-132

Chien L.-C., Hung T.-C., Choang K.-Y., Yeh C.-Y., Meng P.-J., Shieh M.-J., Han B.-C. (2002) Daily intake of TBT, Cu, Zn, Cd and As for fishermen in Taiwan. The Science of the Total Environment 285: 177-185

Chisnall B., Rowe, DK. (1997) Heavy metals in the muscle of two eel species near the Ngawha geothermal field. NIWA client report TSA70201. National Institute of Water and Atmospheric Research, Hamilton.

Chong K., Wang W. (2001) Comparative studies on the biokinetics of Cd, Cr, and Zn in the green mussel *Perna viridis* and the Manila clam *Ruditapes philippinarum*. Environmental Pollution 115: 107-121

Choresh O., Ron, E. and Loya, Y. (2001) The 60-kDa Heat Shock Protein (HSP60) of the sea anemone *Anemonia viridis*: A potential early warning system for environmental changes. Marine Biotechnology 3: 501-508

Ciaramella M., Filosa S., Parisi E. (1997) PCR amplification and cloning of metallothionein complementary DNAs in temperate and Antarctic sea urchin characterized by a large difference in egg metallothionein content. Cellular and Molecular Life Science 53: 472-477

Cismowskil M., Huang P. (1991) Effect of cysteine replacements at positions 13 and 50 on metallothionein structure. Biochemistry 30: 6626-6632

Cismowskil M., Narula S., IM A., Chernaik M., Huang P. (1991) Mutation of invariant cysteines of mammalian metallothionein alters metal binding

capacity, cadmium resistance and ^{113}Cd NMR spectrum. The Journal of Biological Chemistry 266: 24390-24397

Coelho M.R., Bebianno M.J., Langston W.J. (2002) Routes of TBT uptake in the clam *Ruditapes decussatus*. I. Water and sediments as vectors of TBT uptake. Marine Environmental Research 54: 179-192

Colwell R. (2005) EstimateS: Statistical estimation of species richness and shared species from samples. Version 7.5. User's Guide and application published at:http://viceroy.ecb.uconn.edu/estimateS. September 2006

Cook A., Weinstein P. (2005) Volcanic emissions and health risks of metal contaminants. In: *Heavy Metal Contaminants in New Zealand: Sources, Treatments, and Effects on Ecology and Human Health*. Resolutionz Press, Christchurch, New Zealand

Coombs T., George S. (1978) Mechanisms of immobilization and detoxification of metals in marine organisms. In: *Physiology and behaviour of Marine Organisms*. Pergamon Press, Oxford and New York., pp p 179-189

Cosson R. (2000) Bivalve metallothionein as a biomarker of aquatic ecosystem pollution by trace metals: limits and perspectives. Cellular and Molecular Biology 46: 295-309

Craw D., Brown K., Webster-Brown J. (2005) Metals derived from gold mining and geothermal sources. In: *Metal contaminants in New Zealand: Sources, Treatments, and Effects on Ecology and Human Health*. Resolutionz Press, Christchurch, New Zealand, pp 231-245

Cunningham P., Tripp M. (1975) Factors affecting the accumulation and removal of mercury from tissues of the American oyster *Crassostrea virginica*. Marine Biology 31: 311-319

Dallinger R., Berger B., Hunziker P., Birchler N., Hauer C., Kagi J. (1993) Purification and primary structure of snail metallothionein. Similarity of the N-terminal sequence with histones H4 and H2A. European Journal of Biochemistry 216: 739-746

Dallinger R., Berger B., Hunziker P., Kagi J. (1997) Metallothionein in snail Cd and Cu metabolism. Nature 388: 237-238

De Luca-Abbott S. (2001) Biomarkers of sublethal stress in the soft-sediment bivalve *Austrovenus stutchburyi* exposed in-situ to contaminated sediment in an urban New Zealand harbour. Marine Pollution Bulletin 42: 817-825

De Wolf H., Blust R., Backeljau T. (2004) The use of RAPD in ecotoxicology. Mutation Research 566: 249-262

Depledge M. (1996) Genetic ecotoxicology: an overview. Journal of Experimental Biology and Ecology 200: 57-66

Do Amaral M.C.R., Rebelo M.D.-F., Torres J.P.M., Pfeiffer W.C. (2004) Bioaccumulation and depuration of Zn and Cd in mangrove oysters (*Crassostrea rhizophorae*, Guilding, 1828) transplanted to and from a contaminated tropical coastal lagoon. Marine Environmental Research In Press

Dohi Y., Kosaka, K., Ohba, K., Yoneyama, Y. (1986) Cadmium-binding proteins of three marine molluscs and characterization of two cadmium-binding glycoproteins from the hepatopancreas of a whelk, *Buccinum tenuissimum*. Environmental Health Perspectives 65: 49-55

Dowling P., Kootstra G., Chan E. (2004) Reed New Zealand Atlas in association with Terralink International. Reed Books, Reed Publishing, New Zealand Limited

Dragun Z., Erk, M., Raspor, B., Ivankovic, D., Pavicic, J. (2004) Metal and metallothionein level in the heat-treated cytosol of gills of transplanted mussels *Mytilus galloprovincialis* Lmk. Environmental International 30: 1019-1025

Dufresne G., Takács L., Heus H., Codani J.-J., Duval M. (2002) Patent searches for genetic sequences: How to retrieve relevant records from patented sequence databases. Nature Biotechnology 20: 1269-1271

Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S., Lafontaine I., de Montigny J., Marck C., Neuveglise C., Talla, E, Goffard, N., Frangeul, L., Aigle, M., Anthouard, V., Babour, A., Barbe, V., Barnay, S., Blanchin, S., Beckerich, J.-M., Beyne, E., Bleykasten, C., Boisrame, A., Boyer, J., Cattolico, L., Confanioleri, F., de Daruvar, A., Despons, L., Fabre, E., Fairhead, C., Ferry-Dumazet, H., Groppi, A., Hantraye, F., Hennequin, C., Jauniaux, N., Joyet, P., Kachouri, R., Kerrest, A., Koszul, R., Lemaire, M., Lesur, I., Ma, L., Muller, H., Nicaud, J.-M., Nikolski, M., Oztas, S., Ozier-Kalogeropoulos, O., Pellenz, S., Potier, S., Richard, G.-F., Straub, M.-L., Suleau, A., Swennen, D., Tekaia, F., Wesolowski-Louvel, M., Westhof, E., Wirth, B., Zeniou-Meyer, M., Zivanovic, Y., Bolotin-Fukuhara, M., Thierry, A., Bouchier, C., Caudron, B., Scarpelli, C., Gaillardin, C., Weissenbach, J., Wincker, P., Souciet, J.-L.; (2004) Genome evolution in yeasts. Nature 430: 35-44

Engelken J., Hildebrandt A. (1999) cDNA cloning and cadmium-induced expression of metallothionein mRNA in the zebra mussel *Dreissena polymorpha*. Biochemistry and Cellular Biology 77: 237-241

Everaarts J. (1990) Uptake and release of cadmium in various organs of the common mussel, *Mytilus edulis* (L.). Bulletin of Environmental Contamination and Toxicology 45: 560-567

Falcó G., Llobet J., Bocio A., Domingo J. (2006) Daily intake of arsenic, cadmium, mercury and lead by consumption of edible marine species. Journal of Agricultural and Food Chemistry 54: 6106-6112

Feder M.E., Hofmann, Gretchen E. (1999) Heat-shock proteins, molecular chaperones and the stress response: Evolutionary and Ecological Physiology. Annu. Rev. Physiol. 61: 243-282

Fenaughty C., Tracey D., Lock J. (1988) Heavy metal and organochlorine concentrations in New Zealand aquatic fish, crustaceans, and molluscs. New Zealand Fisheries Data Report No. 34, MAF Fish

Fido A., Al-Saad S. (2005) Toxic trace elements in the hair of children with autism. SAGE Publications and The National Autistic Society 9: 290-298

Fitzpatrick P.J., Sheehan, D., Livingstone, D.R. (1995) Studies on isoenzymes of glutathione S-transferase in the digestive gland of *Mytilus galloprovincialis* with exposure to pollution. Marine Environmental Research 39: 241-244

Foran J., Carpenter D., Good D., Hamilton M., Hites R., Knuth B., Schwager S. (2006) Risks and benefits of seafood consumption. American Journal of Preventative Medicine 30: 438-439

Fowler B., Hildbrand C., Kojima Y., Webb M. (1987) Nomenclature of metallothionein. Experientia 52: 19-22

Freedman J., Slice L., Dixon D., Fire A., Rubin C. (1993) The novel metallothionein genes of *Caenorhabditis elegans*. Structural organization and inducible, cell-specific expression. Journal of Biological Chemistry 268: 2554-2564

Frew R., Hunter K. (1996) Cadmium in oysters and sediments from Foveaux Strait, New Zealand Proceedings of the Trace Element Group of New Zealand. Waikato University

Fung C., Lam J., Zheng G., Connell D., Monirith I., Tanabe S., Richardson B., Lam P. (2004) Mussel-based monitoring of trace metal and organic contaminants along the east coast of China using *Perna viridis* and *Mytilus edulis*. Environmental Pollution 127: 203-216

Furness H. (2001) Mercury - In Soil, Food and Fertiliser. Fertiliser Research Report, Oct 2001

Gardner J. (2004) A historical perspective of the genus *Mytilus* (Bivalvia: Mollusca) in New Zealand: multivariate morphometric analyses of fossil, midden and contemporary blue mussels. Biological Journal of the Linnean Society 82: 329-344

Gardner J., Kathiravetpillai G. (1997) Biochemical genetic variation at a leucine aminopeptidase (LAP) locus in blue (*Mytilus galloprovincialis*) and greenshell (*Perna canaliculus*) mussel populations along a salinity gradient. Marine Biology 128: 619-625

Geffard A., Geffard, O., His, E., Amiare, J-C. (2002) Relationships between metal bioaccumulation and metallothionein levels in larvae of *Mytilus galloprovincialis* exposed to contaminated estuarine sediment elutriate. Marine Ecology Progress Series 233: 131-142

Geret F., Cosson, R. (2002) Induction of specific isoforms of Metallothionein in mussel tissues after exposure to cadmium or mercury. Archives of Environmental Contamination and Toxicology 42: 36-42

Gerpe M., Kling P., Olsson P. (1998) Metallothionein gene expression in arctic char (*Salvelinus alpinus*) following metal and PCB exposure. Marine Environmental Research 46: 551-554

Glasby G., Stoffers P., Walter P., Davis K., Renner R. (1988) Heavy-metal pollution in Manukau and Waitemata Harbours, New Zealand. New Zealand Journal of Marine and Freshwater Research 22: 595-611

Gochfeld M. (2003) Cases of mercury exposure, bioavailability, and absorption. Ecotoxicology and Environmental Safety 56: 174-179

Goldberg E., Bertine, K. (2000) Beyond the Mussel Watch - new directions for monitoring marine pollution. The Science of the Total Environment 247: 165-174

Goldberg E.D. (1975) The mussel watch - a first step in global marine monitoring. Marine Pollution Bulletin 6: 111

Goldberg E.D., Koide, M., Hodge, V., Flegal, A.R., Martin, J. (1983) United States mussel watch --1977-1978 results on trace metals and radionuclides. Estuarine Coastal and Shelf Science 16: 69-93

Gosling E. (2003) Bivalve Molluscs: Biology, Ecology and Culture. Fishing News Books, Blackwell Publishing, Oxford, UK

Gray M., Harris A., Centeno J. (2005) The role of cadmium, zinc, and selenium in prostate disease. In: *Metal contaminants in New Zealand: Sources, Treatments, and Effects on Ecology and Human Health*. Resolutionz Press, Christchurch, New Zealand, pp 393-414

Gregory M., George R., Marshall D., Anandraj A., Mcclurg T. (1999) The effects of mercury exposure on the surface morphology of gill filaments in *Perna perna* (Mollusca: Bivalvia). Marine Pollution Bulletin 39: 116-121

Gregory M., Marshall D., George R., Anandraj A., McClurg T. (2002) Correlations between metal uptake in the soft tissue of *Perna perna* and gill filament pathology after exposure to mercury. Marine Pollution Bulletin 45: 114-125

Grove S.L., Probert, P. K. (1999) Sediment macrobenthos of upper Otago Harbour, New Zealand. New Zealand Journal of Marine and Freshwater Research 33: 469-480

Gruber C., Sturzenbaum S., Gehrig P., Sack R., Hunziker P., Berger B., Dallinger R. (2000) Isolation and characterization of a self-sufficient one-domain protein: (Cd)-Metallothionein from *Eisenia foetida*. European Journal of Biochemistry 267: 573-582

Hamilton M., Pincus E., DiFore A., Flischer R. (1999) A universal linker and ligation procedures for construction of genomic DNA libraries enriched for microsatellites. Biotechniques 27: 500-502, 504-507

Harding J. (2005) Impacts of metals and mining on communties. In: *Metal contaminants in New Zeland: Sources, Treatments, and Effects on Ecology and Human Health*. Resolutionz Press, Christchurch, New Zealand

Hay B. (1996) Marine resources of Te Hiku o te Ika. In: *Sustainable development in Tai Tokerau Case Study Two: Te Hiku o te Ika.* Report published by James Henare Maori Research Centre, University of Auckland. pp. 24-55

Hendy E., Peake B. (1996) Organochlorine pesticides in a dated sediment core from Mapua, Waimea inlet, New Zealand. Marine Pollution Bulletin 32: 751-754

Herbert B. (1999) Advances in protein solubilisation for two-dimensional electrophoresis. Electrophoresis 20: 660-663

Hickey C., Roper D., Buckland S. (1995) Metal concentrations of resident and transplanted freshwater mussels *Hyridella menziesi* (Unionacea: Hyriidae) and sediments in the Waikato River, New Zealand. The Science of the Total Environment 175: 163-177

Hoarau P., Garello G., Gnassia-Barelli M., Roméo M., Girard J.-P. (2002) Purification and partial characterization of seven glutathione S -transferase isoforms from the clam *Ruditapes decussatus*. European Journal of Biochemistry 269: 4359-4366

Hoggins F.E., Brooks, R.R. (1972) Natural dispersion of mercury from Puhipuhi, Northland, New Zealand. New Zealand Journal of Marine and Freshwater Research 7: 125-132

Hooper S., Thuma J. (2005) Invertebrate Muscles: Muscle specific genes and proteins. Physiological Reviews 85: 1001-1060

Hoving S., Gerrits B., Voshol H., Muller D., Roberts R., van Oostrum J. (2002) Preparative two-dimensional gel electrophoresis at alkaline pH using narrow range immobilized pH gradients. Proteomics 2: 127-134

Hu Y., Wang G., Chen G., Fu X., Yao S. (2003) Proteome analysis of *Saccharomyces cerevisiae* under metal stress by two-dimensional differential gel electrophoresis. Electrophoresis 24: 1458-1470

Hung T., Meng P., Han B., Chuang A., Huang C. (2001) Trace metals in different species of mollusca, water and sediments from Taiwan coastal area. Chemosphere 44: 833-841

Hungspreugs M., Utoomprurkporn W., Dharmvanij S., Som-pongchaiyakul P. (1989) The present status of the aquatic environment of Thailand. Marine Pollution Bulletin 20: 327-332

Iatrou K., Tsitilou S., Kafatos F. (1984) DNA sequence transfer between two high-cysteine chorion gene families in the silkmoth *Bombyx mori*. Proceedings of the National Academy of Science 81: 4452-4456

Imagawa M., Onozawa T., Okumura K., Osada S., Nishihara T., Kondo M. (1990) Characterization of metallothionein cDNAs induced by cadmium in the nematode *Caenorhabditis elegans*. Biochemical Journal 268: 237-240

Irato P., Santovito G., Cassini A., Piccinni E., Albergoni V. (2003) Metal accumulation and binding protein induction in *Mytilus galloprovincialis*, *Scapharca inaequivalis*, and *Tapes philippinarum* from the lagoon of Venice. Archives of Environmental Contamination and Toxicology 44: 476-484

Isani G., Andreani, G., Kindt, M. and Carpene, E. (2000) Metallothioneins (MTs) in Marine Molluscs. Cellular Molecular Biology 46: 311-330

Janssen H., Scholz N. (1979) Uptake and cellular distribution of cadmium in *Mytilus edulis*. Marine Biology 55: 133-141

Jarman J., Brown, P (1995) Outbreaks of gastroenteritis associated with consumption of oysters in Northland. New Zealand Public Health Report

Jeffs A., Holland R., Hooker S., Hayden B. (1999) Overview and bibliography of research on the greenshell mussel, *Perna canaliculus*, from New Zealand waters. Journal of Shellfish Research 18: 347-360

Johnson R., Davis M., Taylor J., Patterson S. (2005) Informatics for protein identification by mass spectrometry. Methods 35: 223-236

Judd N., Griffith W., Faustman E. (2004) Consideration of cultural and lifestyle factors in defining susceptible populations for environmental disease. Toxicology 198: 121-133

Kádár E., Costa V., Santos R., Lopes H. (2004) Behavioural response to the bioavailability of inorganic mercury in the hydrothermal mussel *Bathymodiolus azoricus*. The Journal of Experimental Biology 208: 505-513

Kagi J., Shaffer A. (1988) Biochemistry of metallothioneins. Biochemistry 27: 8509-8515

Kalpaxis D., Theos C., Xaplanteri M., Dinos G., Catsiki A., Leotsinidis M. (2004) Biomonitoring of Gulf of Patras, N. Peloponnesus, Greece. Application of a biomarker suite including evaluation of translation efficiency in *Mytilus galloprovincialis* cells. Environmental Research 94: 211-220

Karp N., Lilley K. (2005) Maximising sensitivity for detecting changes in protein expression: Experimental design using minimal CyDyes. Proteomics 5: 3105-3115

Karp N., Spencer M., Lindsay H., O'Dell K., Lilley K. (2005) Impact of replicate types on proteomic expression analysis. Journal of Proteome Research 4: 1867-1871

Kennedy P. (1986) The use of molluscs for monitoring trace elements in the marine environment in New Zealand 1. The contribution of ingested sediment to the trace element concentrations in New Zealand molluscs. New Zealand Journal of Marine and Freshwater Research 20: 627-640

Khoo H., Patel K. (1999) Metallothionein cDNA, promoter, and genomic sequences of the tropical green mussel, *Perna viridis*. Journal of Experimental Zoology 284: 445-453

Kille P., Olsson P. (1996) The use of metallothionein genes for determining the phylogenetic and evolutionary relationship between extant teleosts GenBank entry

Kille P., Stephens P., Kay J. (1991) Elucidation of cDNA sequences for metallothioneins from rainbow trout, stone loach and pike liver using the polymerase chain reaction. Biochimica et Biophysica Acta 1089: 407-410

King C., Gale S., Hyne R., Stauber J., Simpson S., Hickey C. (2006) Sensitivities of Australian and New Zealand amphipods to copper and zinc in waters and metal-spiked sediments. Chemosphere 63: 1466-1476

King N., Miller M. (1989) Tributyl tin levels for sea water, sediment, and selected marine species in coastal Northland and Auckland, New Zealand. New Zealand Journal of Marine and Freshwater Research 23: 287-294

Knigge T., Monsinjon T., Andersen O.-K. (2004) Surface-enhanced laser desorption/ionization-time of flight-mass spectrometry approach to biomarker discovery in blue mussels (*Mytilus edulis*) exposed to polyaromatic hydrocarbons and heavy metals under field conditions. Proteomics 4

Kobayashi K., Morita T., Shimada A. (1999) *Canis familiaris* (dog) MT-IV SWISSPROT entry

Koehn R., Milkman R., Mitton J. (1976) Population genetics of marine pelecypods. IV. Selection, migration, and genetic differentiation in the blue mussel *Mytilus edulis*. Evolution 30: 2-32

Koehn R., Newell R., Immermann F. (1980) Maintenance of an aminopeptidase allele frequency cline by natural selection. Proceedings of the National Academy of Science 77: 5385-5389

Kojima Y., Kagi J. (1978) Metallothionein. Trends in Biochemical Sciences 3: 90-93

Kraepiel A., Keller K., Chin H., Malcolm E., Morel F. (2003) Sources and variations of mercury in tuna. Environmetal Science and Technology 37: 5551-5558

Lam P.K.S.a.G., J. S. (2003) The use of biomarkers in environmental monitoring programmes. Marine Pollution Bulletin 46: 182-186

Lane D., Nott J. (1975) A study of the morphology, fine structure and histochemistry of the foot of the pediveliger of *Mytilus edulis*. Journal of the Marine Biological Association UK 55: 477-495

Langston W., Zhou M. (1987) Cadmium accumulation, distribution and elimination in the bivalve *Macoma balthica*: Neither metallothionein nor metallothionein-like proteins are involved. Marine Environmental Research 21: 225-237

Laurie A. (2004) Quantitation of metallothionein mRNA from the New Zealand common bully (*Gobiomorphus contidianus*) and its implications for biomonitoring. New Zealand Journal of Marine and Freshwater Research 38: 869-877

Leignel V., Laulier M. (2006) Isolation and characterization of Mytilus edulis metallothionein genes. Comparative Biochemistry and Physiology Part C 142: 12-18

Lerch K., Ammer D., Olafson R. (1982) Crab metallothionein. Primary structures of metallothioneins 1 and 2. Journal of Biological Chemistry 257: 2420-2426

Liang L., He B., Jiang G., Chen D., Yao Z. (2004) Evaluation of mollusks as biomonitors to investigate heavy metal contaminations along the Chinese Bohai Sea. Science of the Total Environment 324: 105-113

Lilleberg S., Patel S. (1998) Isolation of DNA flanking retroviral integration sites using Vectorette II. Life Science Quarterly 3

Lionetto M., Giordano M., Caricato R., Pascariello M., Marinosci L., Schettino T. (2001) Biomonitoring of heavy metal contamination along the Salento coase (Italy) by metallothionein evaluation in *Mytilus galloprovincialis* and *Mullus*

barbatus. Aquatic Conservation:Marine and Freshwater Ecosystems 11: 305-310

Livingston M. (1987) Preliminary studies on the effects of past mining on the aquatic environment, Coromandel Peninsula. Water and Soil Miscellaneous Publication No. 104, Water and Soil Directorate, Ministry of Works and Development, Wellington

López-Barea J., Gomez-Ariza J. (2006) Environmental proteomics and metallomics. Proteomics 6: 51-62

López J. (2005) Role of proteomics in taxonomy: the *Mytilus* complex as a model of study. Journal of Chromatography B 815: 261-274

López J., Marina A., Alvarez G., Vazquez J. (2002a) Application of proteomics for fast identification of species-specific peptides from marine species. Proteomics 2: 1658-1665

López J., Marina A., Vazquez J., Alvarez G. (2002b) A proteomic approach to the study of the marine mussels *Mytilus edulis* and *M. galloprovincialis*. Marine Biology 141: 217-233

Lyons C., Dowling V., Tedengren M., Gardestrom J., Hartl M.G.J., O'Brien N., van Pelt F.N., O'Halloran J., Sheehan D. (2003) Variability of heat shock proteins and glutathione S-transferase in gill and digestive gland of blue mussel, *Mytilus edulis*. Marine Environmental Research 56: 585-597

Mackay E.A., Overnell J., Dunbar B., Davidson I., Hunziker P., Kagi J., Fothergill J. (1993) Complete amino acid sequences of five dimeric and four monomeric forms of metallothionein from the edible mussel *Mytilus edulis*. European Journal of Biochemistry 218: 183-194

Mackie D., Hunter K. (2005) The partitioning of zinc, manganese, cadmium and selenium within copepods in oceanic waters east of New Zealand. Marine and Freshwater Research 56: 95-100

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

Manly B. (1997) Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall, London

Marsden I., Rainbow P., Smith B. (2003) Trace metal concentrations in two New Zealand talitrid amphipods: effects of gender and reproductive state and implications for biomonitoring. Journal of Experimental Marine Biology and Ecology 290: 93-113

Masilamoni J., Azariah J., Nandakumar K., Jesudoss K., Satpathy K., Nair K. (2001) Excretory products of green mussel *Perna viridis* L. and their implications on power plant operation. Turkish Journal of Zoology 25: 117-125

Masters B., Quaife C., Erickson J., Kelly E., Froelick G., Zambrowicz B., Brinster R., Palmiter R. (1994) Metallothionein III is expressed in neurons that sequester zinc in synaptic vesicles. Journal of Neuroscience 14: 5844-5857

Matthiessen P., Law R. (2002) Contaminants and their effects on estuarine and coastal organisms in the United Kingdom in the late twentieth century. Environmental Pollution 120: 739-757

McArdle B., Anderson M. (2001) Fitting multivariate models to community data: a comment on distance-based redundancy analysis. Ecology 82: 290-297

McDonagh B., Tyther R., Sheehan D. (2005) Carbonylation and glutathionylation of proteins in the blue mussel *Mytilus edulis* detected by proteomic analysis of

Western blotting: Actin as a target for oxidative stress. Aquatic Toxicology 77: 315-326

McKenzie-Parnell J., Kjellstrom T., Sharma R., Robinson M. (1988) Unusually high intake and fecal output of cadmium, and fecal output of other trace elements in New Zealand adults consuming dredge oysters. Environmental Research 46: 1-14

Mi J., Orbea A., Syme N., Ahmed M., Cajaravill M., Cristóbal S. (2005) Peroxisomal proteomics, a new tool for risk assessment of peroxisome proliferating pollutants in the marine environment. Proteomics 5: 3954-3965

Ministry of Health (1999) NZ Food: NZ People Key results of the 1997 National Nutrition Survey, Wellington

Moraga D., Mdelgi-Lasram E., Romdhane M.S., Abed A.E., Boutet I., Tanguy A., Auffret M. (2002) Genetic responses to metal contamination in two clams: *Ruditapes decussatus* and *Ruditapes philippinarum*. Marine Environmental Research 54: 521-525

Morse D., Guzewich J., Hanrahan J., Stricof R., Shayegani M., Deibel R., Grabau J., Nowak N., Herrmann J., Cukor H. (1986) Widespread outbreaks of clam- and oyster- associated gastroenteritis. Role of Norwalk virus. The New England Journal of Medicine 314: 678-681

Morton J., Miller M. (1968) The New Zealand sea shore. Collins, London-Auckland

Mourgaud Y., Martinez E., Geffard A., Andral B., Stanisiere J.-Y., Amiard J.-C. (2002) Metallothionein concentration in the mussel *Mytilus galloprovincialis* as a biomarker of response to metal contamination: validation in the field. Biomarkers 7: 479-490

Nemer M., Thornton R., Stuebing E., Harlow P. (1991) Structure, spatial, and temporal expression of two sea urchin metallothionein genes, SpMTB1 and SpMTA. Journal of Biological Chemistry 266: 6586-6593

Nemer M., Wilkinson D., Travaglini E., Sternberg E., Butt R. (1985) Sea urchin metallothionein sequence: Key to an evolutionary diversity. Proceedings of the National Academy of Science 82: 4992-4994

Nicholson S., Szefer P. (2003) Accumulation of metals in the soft tissues, byssus and shell of the mytilid mussel *Perna viridis* (Bivalvia: Mytilidae) from polluted and uncontaminated locations in Hong Kong coastal waters. Marine Pollution Bulletin 46: 1040-1043

Nielsen S., Nathan A. (1975) Heavy metal levels in New Zealand molluscs. New Zealand Journal of Marine and Freshwater Research 9: 467-481

Noël-Georis I., Valleaeys T., Chauvaux R., Falmagne P., Mergeay M., Wattiez R. (2004) Global analysis of the *Ralstonia metallidurans* proteome: Prelude for the large-scale study of heavy metal response. Proteomics 4: 151-179

Nolan C., Duke, E. (1984) Heavy metal uptake and intracellular binding in isolated gel preparations of *Mytilus galloprovincialis* L. The Science of the Total Environment 40: 83-92

Nordberg M., Kojima Y. (1979) Metallothionein. Experientia 34: 48-55

Nordberg M., Nuottaniemi I., Cherian M., Nordberg G., Kjellstrom T., Garvey J. (1986) Characterization studies on the cadmium-binding proteins from two species of New Zealand oysters. Environmental Health Perspectives 65: 57-62

NZFSA (1981) Heavy metal concentrations in green mussels *Perna canaliculus*. Council reports

Olsson B., Bradley B., Gilek M., Reimer O., Shepard J., Tedengren M. (2004) Physiological and proteomic responses in *Mytilus edulis* exposed to PCBs and PAHs extracted from Baltic sea sediments. Hydrobiologia 514: 15-27

Otchere F., Joiris C., Holsbeek L. (2003) Mercury in the bivalves *Anadara (Senilia) senilis, Perna perna* and *Crassostrea tulipa* from Ghana. Science of the Total Environment 304: 369-375

Patwary M., Reith M., Kenchington E. (1990) Cloning and characterization of tropomyosin cDNAs from the sea scallop *Placopecten magellanicus* (Gmelin, 1791). Journal of Shellfish Research 18: 67-70

Peake B., Marsden I., Bryan A. (2006) Spatial and temporal variations in trace metal concentrations in the cockle, *Austrovenus stutchburyi* from Otago, New Zealand. Environmental Monitoring and Assessment 115: 119-144

Pearson W.R., Lipman, D.J. (1988) Improved tools for biological sequence comparison. Proceedings of the National Academy of Science 85: 2444-2448

Pennington K., McGregor E., Beasley C., Everall I., Cotter D., Dunn M. (2004) Optimization of the first dimension for separation by two-dimensional gel electrophoresis of basic proteins from human brain tissue. Proteomics 4: 27-30

Petrovic S., Ozetic B., Krajnovic-Ozretic M., Bobinac D. (2001) Lysosomal membrane stability and metallothioneins in digestive gland of mussels (*Mytilus galloprovincialis* Lam.) as biomarkers in a field study. Marine Pollution Bulletin 42: 1373-1378

Phillips D., Fowles J. (2005) The surveillance of heavy metals and human health outcomes In: *Metal contaminants in New Zealand: Sources, Treatments, and Effects on Ecology and Human Health*. Resolutionz Press, Christchurch, New Zealand, pp 441-463

Phillips D.J.H. (1976) The common mussel *Mytilus edulis* as an indicator of pollution by zinc, cadmium, lead and copper. I. Effects of environmental variables on uptake of metals. Marine Biology 38: 59-69

Phillips D.J.H., Segar, D.A. (1986) Use of bioindicators in monitoring conservative contaminants: programme design imperatives. Marine Pollution Bulletin 17: 10-17

Prest H., Richardson B., Jacobson L., Vedder J., Martin M. (1995) Monitoring organochlorines with semi-permeable membrane devices (SPMDs) and mussels (*Mytilus edulis*) in Corio Bay, Victoria, Australia. Marine Pollution Bulletin 30: 543-554

Preston M. (1989) Marine pollution. Chemical Oceanography. Academic Press, New York, pp 53-197

Pridmore R., Roper D., Hewitt J. (1990) Variation in composition and condition of the Pacific oyster, *Crassostrea gigas*, along a pollution gradient in Manukau Harbour, New Zealand. Marine Environmental Research 30: 163-177

Propp M., Odintsov V., Propp L. (1994) Marine sediments in a region of recent volcanic activity in the Bay of Plenty, New Zealand. New Zealand Journal of Marine and Freshwater Research 28: 219-226

Quaife C., Findley S., Erickson J., Froelick G., Kelly E.J., Zambrowicz B., Palmiter R. (1994) Induction of a new metallothionein isoform (MT-IV) occurs during differentiation of stratified squamous epithelia. Biochemistry 33: 7250-7259

Quilbe R., Pieri I., Wicherek S., Dugas N., Tasteyre A., Thomas Y., Oudinet J.-P. (2004) Combinatory Chemical and Biological Approaches to Investigate

Metal Elements in Agricultural Runoff Water. Journal of Environmental Quality 33: 149-153

Quinn G., Keough M. (2002) Experimental design and data analysis for biologists. Cambridge University Press, Cambridge, UK

R Development Core Team (2006) R: A Language and Environment for Statistical Computing, Vienna, Austria

Rainbow P. (1995) Biomonitoring of heavy metal availability in the marine environment. Marine Pollution Bulletin 31: 183-192

Rainbow P., Emson R., Smith B., Moore P., Mlandenov P. (1993) Talitrid amphipods as biomonitors of trace metals near Dunedin, New Zealand. New Zealand Journal of Marine and Freshwater Research 27: 201-207

Rajagopal S., van der Velde G., van der Gaag M., Jenner H. (2005) Byssal detachment underestimates tolerance of mussels to toxic compounds. Marine Pollution Bulletin 50: 20-29

Redmayne A., Kim J., Closs G., Hunter K.A. (2000) Methyl mercury bioaccumulation in long-finned eels, *Anguilla dieffenbachii*, from three rivers in Otago, New Zealand. The Science of the Total Environment 262: 37-47

Richardson B., Garnham J., Fabris J. (1994) Trace metal concentrations in mussels (*Mytilus edulis planulatus* L.) transplanted into southern Australian waters. Marine Pollution Bulletin 28: 392-396

Ridgway N., Glasby G. (1984) Sources of marine pollution around New Zealand. NZOI Oceanographic Summary 23

Riek R., Precheur B., Wang Y., Mackay E., Wider G., Guntert P., Liu A., Kagi J., Wuthrich K. (1999) NMR structure of the sea urchin (*Strongylocentrotus purpuratus*) metallothionein MTA. Journal of Molecular Biology 291: 417-428

Rittschof D., McClellan-Green P. (2005) Molluscs as multidisciplinary models in environment toxicology. Marine Pollution Bulletin 50: 369-373

Robinson B., Duwig C., Bolan N., Kannathasan M., Saravanan (2003) Uptake of arsenic by New Zealand watercress (*Lepidium sativum*). The Science of the Total Environment 301: 67-73

Rodríguez-Ortega M., Grøsvik B., Rodríguez-Ariza A., Goksøyr A., López-Barea J. (2003a) Changes in protein expression profiles in bivalve molluscs (*Chamaelea gallina*) exposed to four model environmental pollutants. Proteomics 3: 1535-1543

Rodríguez-Ortega M., Rodríguez-Ariza A., Amezcua O., López-Barea J. (2003b) Mutagenic activation of arylamines by subcellular fractions of *Chamaelea gallina* clams exposed to environmental pollutants. Environmental and Molecular Mutagenesis 41: 55-63

Roesijadi G. (1982) Uptake and incorporation of mercury into mercury-binding proteins of gills of *Mytilus edulis* as a function of time. Marine Biology 66: 151-157

Roesijadi G. (1986) Mercury-binding proteins from the marine mussel, *Mytilus edulis*. Environmental Health Perspectives 65: 45-48

Roesijadi G., Kielland, S., Kerks, P. (1989) Purification and properties of novel molluscan metallothioneins. Archives of Biochemistry and Biophysics 273: 403-413

Roesijadi G., Vestling M., Murphy C., Klerks P., Fenselau C. (1991) Structure and time-dependent behaviour of acetylated and non-acetylated forms of a molluscan metallothionein. Biochimica et Biophysica Acta 1074: 230-236

Rogers K.M. (1999) Effects of sewage contamination on macro-algae and shellfish at Moa Point, New Zealand using stable carbon and nitrogen isotopes. New Zealand Journal of Marine and Freshwater Research 33: 181-188

Roméo M., Mourgaud Y., Geffard A., Gnassia-Barelli M., Amiard J., Budzinski H. (2003) Multimarker approach in transplanted mussels for evaluating water quality in Charentes, France, coast areas exposed to different anthropogenic conditions. Environmental Toxicology 18: 295-305

Romero-Isart N., Cols N., Temansen M., Gelpi J., Gonzalez-Duarte R., Atrian S., Capdevila M., Gonzalez-Duarte P. (1999) Replacement of terminal cysteine with histidine in the metallothionein alpha- and beta- domains maintains its binding capacity. European Journal of Biochemistry 259: 519-527

Roper D., Pridmore R., Cummings V., Hewitt J. (1991) Pollution related differences in the condition cycles of pacific oysters *Crassostrea gigas* from Manukau Harbour, New Zealand. Marine Environmental Research 31: 197-214

Roper J., Cherry D., Simmers J., Tatem H. (1997) Bioaccumulation of PAHs in the zebra mussel at Times Beach, Buffalo, New York. Environmental Monitoring and Assessment 46: 267-277

Rowe D.K., Chisnall B.L. (1997) Distribution and conservation status of the dwarf inanga *Galaxias gracilis* (Teleostei: Galaxiidae) an endemic fish of Northland dune lakes. Journal of the Royal Society of New Zealand 27: 223-233

Rowe D.K., Chisnall B.L., Dean T.L., Richardson J. (1999) Effects of land use on

Shick J., Widdows J. (1981) Direct and indirect calorimetric measurement of metabolic rate in bivalve molluscs during aerial exposure. American Zoology 21: 983

Silva A., Wright S. (1992) Integumental taurine transport in *Mytilus* gill: Short-term adaptation to reduced salinity. Journal of Experimental Biology 162: 265-279

Slice L., Freedman J., Rubin C. (1990) Purification, characterization, and cDNA cloning of a novel metallothionein-like, cadmium-binding protein from *Caenorhabditis elegans*. Journal of Biological Chemistry 265: 256-263

Slivestre F., Dierick J.-F., Dumont V., Dieu M., Raes M., Devos P. (2006) Differential protein expression profiles in anterior gills of *Eriocheir sinensis* during acclimation to cadmium. Aquatic Toxicology 76: 46-58

Smith D. (1986) Heavy metals in the New Zealand aquatic environment: a review. National Water and Conservation Authority, Ministry of Works and Development

Smith P. (1996) Selective decline in imposex levels in the dogwhelk *Lepsiella scobina* following a ban on the use of TBT antifoulants in New Zealand. Marine Pollution Bulletin 32: 362-365

Smith P., McVeagh M. (1991) Widespread organotin pollution in New Zealand coastal waters as indicated by imposex in dog whelks. Marine Pollution Bulletin 22: 409-413

Snyder M., Girvetz, E. and Mulder, E. (2001) Induction of marine mollusc stress proteins by chemical or physical stress. Archives of Environmental Contamination and Toxicology 41: 22-29

Stark J., Boothroyd I., Harding J., Maxted J., Scarsbrook M. (2001) Protocols for sampling macroinvertebrates in wadeable streams. New Zealand Macroinvertebrate Working Group Report No. 1. Prepared for the Ministry for the Environment, Wellington, New Zealand. 57 p.

Statistics New Zealand (1998) New Zealand Official Yearbook 1998, GP Publications, Wellington

Stephan W., Rodriguez V., Zhou B., Parsch J. (1994) Molecular evolution of the metallothionein gene Mtn in the melanogaster species group: results from *Drosophila ananassae*. Genetics 138: 135-143

Stewart C., de Mora S., Jones M., Miller M. (1992) Imposex in New Zealand neogastropods. Marine Pollution Bulletin 24: 204-209

Stoffers P., Glasby G., Wilson C., Davis K., Walter P. (1986) Heavy metal pollution in Wellington Harbour. New Zealand Journal of Marine and Freshwater Research 20: 495-512

Stoffers P., Hannington M., Wright I., Herzig P., de Ronde C. (1999) Elemental mercury at submarine hydrothermal vents in the Bay of Penty, Taupo volcanic zone, New Zealand. The Geological Society of America 27: 931-934

Sundaram K., Shaffee M. (1989) Salinity tolerance of some bivalves of Ennore estuary. Journal of Marine Biology Association India 31: 299-302

Tanabe S., Prudente M., Kan-atireklap S., Subramanian A. (2000) Mussel watch: marine pollution monitoring of butyltins and organochlorines in coastal waters of Thailand, Philippines and India. Ocean and Coastal Management 43: 819-839

Tanguy A., Boutet I., Bonhomme F., Boudry P., Moraga D. (2002) Polymorphism of metallothionein genes in the Pacific oyster *Crassostrea gigas* as a biomarker of response to metal exposure. Biomarkers 7: 439-450

Tanguy A., Boutet I., Laroche J., Moraga D. (2005) Molecular identification and expression study of differentially regulated genes in the Pacific oyster *Crassostrea gigas* in response to pesticide exposure. FEBS Journal

Tanguy A., Boutet I., Riso R., Boudry P., Auffret M., Moraga D. (2003) Metallothionein genes in the European flat oyster *Ostrea edulis*: a potential ecological tool for environmental monitoring. Marine Ecology Progress Series 257: 87-97

Tanguy A., Mura C., Moraga D. (2001) Cloning of a metallothionein gene and characterization of two other cDNA sequences in the Pacific oyster *Crassostrea gigas* (CgMT1). Aquatic Toxicology 55: 35-47

Taylor D. (1983) The significance of the accumulation of cadmium by aquatic organisms. Ecotoxicology and Environmental Safety 7: 33-42

Thompson D., Dowding J. (1999) Site-specific heavy metal concentrations in blood of south island pied oystercatchers *Haematopus ostralegus finschi* from the Auckland region, New Zealand. Marine Pollution Bulletin 38: 202-206

Truhaut R. (1977) Ecotoxicology: objectives, principles and perspectives. Ecotoxicology and Environmental Safety 1: 151-173

Unger M., Chen T., Murphy C., Vestling M., Fenselau C., Roesijadi G. (1991) Primary structure of molluscan metallothioneins deduced from PCR-amplified cDNA and mass spectrometry of purified proteins. Biochimica et Biophysica Acta 1074: 371-377

Väinölä R., Hvilsom M. (1991) Genetic divergence and a hybrid zone between Baltic and North Sea *Mytilus* populations (Mytilidae: Mollusca). Biological Journal of the Linnean Society 43: 127-148

Vannoort R., Cressey, P, Silvers, C, Silvers, F.R. (2000) 1997/98 New Zealand Total Diet Survey Elements report - Selected Contaminants & Nutrients. ESR, Wellington

Vannoort R., Thomson B. (2005) 2003/04 New Zealand Total Diet Survey. Conducted by the Institute of Environmental Science and Research Limited a s part of a New Zealand Food and Safety Authority contract for scientific services. Published by the New Zealand Food Safety Authority Wellington, New Zealand

Vašák M. (2005) Advances in metallothionein structure and functions. Journal of Trace Elements in Medicine and Biology 19: 13-17

Viarengo A., Burlando, B., Ceratto, N., Panfoli, I. (2000) Antioxidant role of Metallothioneins: A comparative overview. Cellular and Molecular Biology 46: 407-417

Vido K., Spector D., Lagniel G., Lopez S., Toledano M., Labarre J. (2001) Proteome analysis of the cadmium response in *Saccharomyces cerevisiae*. Journal of Biological Chemistry 276: 8469-8474

Vincent V., Thomas M. (1997) Urban runoff pollution: An overview. Journal of Environmental Education and Information 16: 185-196

Wang W., Rainbow P. (2005) Influence of metal exposure history on trace metal uptake and accumulation by marine invertebrates. Ecotoxicology and Environmental Safety 61: 145-159

Wasinger V., Cordwell S., Cerpa-Poljak A., Yan J., Gooley A., Wilkins M., Duncan M., Harris R., Williams K., Humphery-Smith I. (1995) Progress with gene-product mapping of the Mollicutes: *Mycoplasma genitalium*. Electrophoresis 16: 1090-1094

Williams D. (2001) Matauranga Maori. Matauranga Maori and Taonga: The nature and extent of treaty rights held by iwi and hapu in indigenous flora and fauna, cultural heritage objects, valued traditional knowledge. Waitangi Tribunal Publication, Wellington, NZ

Williamson R., Van Dam L., Bell R., Green M., Kim J. (1996) Heavy metal and suspended sediment fluxes from a contaminated, intertidal inlet (Manukau Harbour, New Zealand). Marine Pollution Bulletin 32: 812-822

Winchester R.V., Keating, D.L. (1980) Trace metal and organochlorine pesticide residues in New Zealand farmed oysters: A preliminary survey. New Zealand Journal of Science 23: 161-169

Wright D., Welbourn P. (2002) Environmental toxicology. Cambridge University Press, Cambridge

Yang H., Zeng Q., Li E., Zhu S., Zhou X. (2004) Molecular cloning, expression and characterization of glutathione S-transferase from *Mytilus edulis*. Compative Biochemistry and Physiology B Biochemistry and Molecular Biology 139: 175-182

Yap C., Ismail A., Tan S. (2003a) Background concentrations of Cd, Cu, Pb and Zn in the green-lipped mussel *Perna viridis* (Linnaeus) from Peninsular Malaysia. Marine Pollution Bulletin 46: 1044-1048

Yap C., Ismail A., Tan S. (2003b) Can the byssus of green-lipped mussel *Perna viridis* (Linnaeus) from the west coast of Peninsular Malaysia be a biomonitoring organ for Cd, Pb and Zn? Field and laboratory studies. Environment International 29: 521-528

Yap C., Ismail A., Tan S. (2004a) Heavy metal (Cd, Cu, Pb and Zn) concentrations in the green-lipped mussel *Perna viridis* (Linnaeus) collected from some wild and aquacultural sites in the west coast of Peninsular Malaysia. Food Chemistry 84: 569-575

Yap C., Ismail A., Tan S., Ismail A. (2004b) Assessment of different soft tissues of the green-lipped mussel *Perna viridis* (Linnaeus) as biomonitoring agents of Pb: Field and laboratory studies. Water, Air and Soil Pollution 153: 253-268

Yap C., Ismail A., Tan S., Omar H. (2002) Correlations between speciation of Cd, Cu, Pb and Zn in sediment and their concentrations in total soft tissue of green-lipped mussel *Perna viridis* from the west coast of Peninsular Malaysia. Environment International 28: 117-126

Yap C., Ismail A., Tan S., Omar H. (2003c) Accumulation, depuration and distribution of cadmium and zinc in the green-lipped mussel *Perna viridis* (Linnaeus) under laboratory conditions. Hydrobiologia 498: 151-160

Yap C., Tan S., Ismail A., Omar H. (2004c) Allozyme polymorphisms and heavy metal levels in the green-lipped mussel *Perna viridis* (Linnaeus) collected from contaminated and uncontaminated sites in Malaysia. Environment International 30: 39-46

Yorbik O., Akay C., Sayal A., Cansever A., Söhmen T., Cavdar A. (2004) Zinc status in autistic children. The Journal of Trace Elements in Experimental Medicine 17: 101-107

Yuan J.-M., Ross R., Gao Y.-T., Yu M. (2001) Fish and shellfish consumption in relation to death from myocardial infarction among men in Shanghai, China. American Journal of Epidemiology 154: 809-815

Zauke G.P. (1992) Cadmium, lead, copper, and zinc in *Elminius modestus* Darwin (Crustacea, Cirripedia) from Waitemata and Manukau Harbours, Auckland,

New Zealand. New Zealand Journal of Marine and Freshwater Research 26: 405-415

GLOSSARY

Average Ratio

The average ratio gives a measure of the differences in protein abundances between two groups. For independent samples, the average ratio is calculated as the degree of difference in the means of the standarised abundance between protein spots in different groups and is report as fold-changes.

Biological Variation Analysis (BVA) Module

BVA processes gel images that have already undergone spot detection in DIA and uses statistical tools to detect proteins that demonstrate significant changes in abundance between treatments.

DeCyder™ 2D Software

Image analysis software developed for use with DIGE flurorescently labelled protein samples.

Differential In-gel Analysis (DIA) Module

Software module within the DeCyder suite, which co-detects protein spots within all gel images within an experimental set. Module also quantitates protein spot abundance ratios.

Differential In Gel Electrophoresis (DIGE)

A method for pre-labelling protein samples with flurorescent CyDyes™ prior to 2-D electrophoresis, allowing the multiplexing of differently labelled protein mixtures on the same 2-D gel.

Extended Data Analysis (EDA)

EDA is an add-on module for the DeCyder™ 2D software. It is used for multivariate analysis of protein abundance data derived from the BVA module. As well as the Student's *t*-test and ANOVA analyses that can be performed in the BVA module, there are a number of additional analyses that can be performed in EDA, including Principle Component Analysis (PCA) and Pattern Analysis.

False discovery rate (FDR) correction

When testing thousands of proteins for statistical significance with Student's *t*-test or ANOVA, many of the proteins may appear to show statistically significant differences in abundance, but several of these proteins may have achieved this significance by chance alone. FDR adjusts *t*-test values to decrease the probability of detecting changes in protein abudances by chance.

First dimension (1-D) isoelectric focusing (IEF)

1-D IEF is an electrophoretic method that separates proteins according to their isoelectric points (pI).

Loading Plot

A Loading Plot is a visual output of a PCA which shows the variables plotted against the first two Principal Components. A Loading Plot illustrates the relationship between variables and can indicate what the most important variables are.

One-Way Analysis of Variance (ANOVA)

ANOVA is an extension of the logic of the Student's *t*-test to those situations where the concurrent comparison of the means of three or more samples is required.

Pattern Analysis

Pattern analysis groups together proteins and spot maps with similar protein profiles using Hierarchical clustering.

PERMANOVA

Permuational multivariate analysis of variance (PERMANOVA) on the basis of Bray-Curtis dissimilarities was used to determine which protein spots showed significant changes in abundance, whenever some of the assumptions for DeCyder analysis were violated (i.e. independence of data, normal distribution, equal variance). Permutation randomly reorders the data by 'shuffling' factor labels. If there is no effect of the factor on the data then it is equally likely that any of the individual factors could have been associated with any one of the observation units.

Principal Component Analysis (PCA).

PCA produces an overview of the dataset by reducing the dimensionality of the data. PCA can be used to detect patterns of protein abundance which may not be obvious in the raw data and can be used to find outlying data (refer to Figure 3.4 and accompanying text for an example). Two analyses are used in the current study. Protein versus Spot Maps calculations and Spot maps versus Protein calculations.

Principal Component 1 (PC1)

The first principal component (PC1) of a PCA describes the majority of the variability in the data.

Principal Component 2 (PC2)

The second principal component (PC2) and each succeeding principal component in a PCA accounts for as much of the remaining variability as possible.

Protein versus Spot Maps calculations

A calculation which uses PCA to find protein oultiers.

Score Plot

A score plot is a visual output of a PCA and shows observations plotted against the first two Principal Components. These figures illustrate which observations are "similar" with respect to the chosen variables.

Second dimension (2-D) SDS polyacrylamide gel electrophoresis (PAGE)

2-D SDS-PAGE is an electrophoretic method for separating proteins according to their molecular weights. The technique is performed in polyacrylamide gels containing sodium dodecyl sulphate (SDS).

Spot Maps versus Protein calculations

A calculation which uses PCA to find spot map outliers.

Student's *t*-test

The Student's *t*-test is used to test the hypothesis that a variable differs between two groups or populations. It is performed as an equal variance two-tailed test, therefore, direction of change (i.e. increases or decreases) in the standarised abundance parameter is considered.

Volume Ratio

Normalised volume ratio between co-detected spots in the primary and secondary images of DIA

APPENDICES

Appendix 1: Primers

Primers MTG1-6 as specified in Khoo and Patel (1999)

MTG1 (Forward, F)
5' TGC CAC GGC GTG CAC ACG GCG 3'
Tm = 80 °C Primer length = 21 nt

MTG2 (F)
5' CTG TCT GTG TGG TAC TGG G 3'
Tm = 50 °C Primer length =19 nt

MTG3 (Reverse, R)
5' ATG GGA ACT AAA G 3'
Tm = 29.9 °C Primer length = 13 nt (very short oligo)

* Primer pairs MTG1/MTG3 or MTG1/MTG3.

MTG4 (F)
5' GCAAGC GGC ACA CGG C 3'
Tm = 64.8 °C Primer length = 16 nt

MTG5 (F)
5'ATC TGT GGT ACT GGG TGC A 3'
Tm = 51.1 °C Primer length = 19 nt

MTG6 (R)
5' ATT ATG GGA ACT AAA GAA 3'
Tm = 42.7 °C Primer length = 18 nt

* Primer pairs MTG4/MTG6 or MTG5/MTG6.

Primers designed using Oligo 6

Primers were designed based on *P. viridis* sequence, as the primers specified in Khoo and Patel (1999) did not amplify to give 'clean' sequences.

131 (Forward, F)
5' AAA TAT GCC TAG CCC TTG TAA 3' ***
Tm = 54.4°C Primer length =21 nt

***The Vectorette results (Chapter 2, Figure 2.5) show that this primer does not bind to the DNA as tightly as it could. Compared with above sequence based on *P. viridis* there are a couple of changes in the *P. canaliculus* sequence (underlined) i.e. CA and C respectively.

942
5' AGA TGA CGG ATA AAC AAT ACG 3'
Tm = 52.9 °C Primer length = 21 nt

1795
5' GGC CAA TGA ACC TAC AGA ATA 3'
Tm = 55.3 °C Primer length = 21 nt

588a
5' AAT AGA ACC ATA ACA AAC CTT 3'
Tm= 48.6 °C Primer length = 21 nt

588b
5' AAT AGA ACC ATA ACA AAC C 3'
Tm = 45 °C Primer length = 19 nt

1126
5' AGA TGC GGA TAA ACA ATA CG 3'
Tm=53.9 °C Primer length = 20 nt

1861
5' GAC GAG GAA TAC AAT AAA AG 3'
Tm = 47.8 °C Primer length 20 nt

2089 (R)
5' ACA CTC TCC TGA AAA TAC ATC 3'
Tm = 49.3 °C Primer length =21 nt

Few reverse primers above, so some more were ordered.

1757 (R)
5' TGG AGT CAC ATT TAC AGG 3'
Tm = 47.8 °C Primer length = 18 nt

599 (R)
5' GAC ATT TGC ACA CGA CTT TA 3'
Tm = 53.2 °C Primer length = 20 nt

Primers designed using Oligo 6 based on *P. canaliculus* obtained from cloning experiments.

As these primers have been designed based on *P. canaliculus* sequence it was expected that some of these primer should amplify 'single' unambiguous sequences. This was the case for primers 5F and 921.

5F
5' AAA TAG CAC CAA AAA CAA AAT ACG 3'
Tm = 57.4 °C Primer length = 24 nt

5R
5' CCC AAA ACA ACA AAA CAT ACA TAA 3'
Tm =56.7 °C Primer length =24 nt

4F
5' CTA ACT CAG ACA CAA GCA TTC ATT 3'
Tm= 56.3 °C Primer length =24 nt

4R
5' CTG GAC CAC TAT CGT AAA AAG CAT 3'
Tm = 59.4°C Primer length = 24 nt

2F
5' ATT TGA TTC CAT TTG TTT TAT TTT TG 3'
Tm = 56.4 °C Primer length = 26 nt

Primers developed for use with Universal Genome Walker Kit.

1-4 (Forward, F)
5' GAA ATG ATA ATA GAA ACT GCG ACT GC 3'
Tm = 71.3 °C, 72 °C, 54.8 °C Primer length = 26 nt

1-4 (Reverse, R)
5' CAC ACG TGG GTA AAC AA 3'
Tm = 41.5 °C, 61.1 °C, 50 °C Primer length = 17 nt (Did not order)

1-4 (R)
5' CAC ACG TGG GTA AAC AAA AAG ATT 3'
Tm = 68.8 °C, 66 °C, 54.1 °C Primer length = 24 nt

1-5 (R)
5' CAC CTG GTT CAC AAA AAG ACT ATG 3'
Tm = 52.4 °C Primer length = 24 nt

1-5 (U)
5' TCG CAT GAT AGT GAA GTT A 3'
Tm = 40.8 °C, 61 °C, 52 °C Primer length = 19 nt

2-4 (F)
5' GCT TTT TAC GAT AGT GGT C 3'
Tm = 63.2 °C, 54 °C, 41.4 °C Primer length =19 nt

2-4 (R)
5' TAA ACA CCC ACT TGA CGT ACA C 3'
Tm = 69.5 °C, 62 °C, 49.4 °C Primer length = 22 nt

2-14 (F)
5' TCG CGT GAT AGT GAA GTT A 3'
Tm = 63.2 °C, 54 °C, 43.8 °C Primer length = 19 nt

2-14 (R)
5' CAC TCG TTT GCT GGA ATG GTT A 3'
Tm = 69.5 °C, 64 °C, 54.3 °C Primer length = 22 nt

3-9 (R)
5' TTA TGT TAG CGG TTT CAC TTA CTT TC 3'
Tm = 69.7 °C, 70 °C Primer length = 26 nt

4-6 (U)
5' GAT GTA TTT TCA GGA GAG TG 3'
Tm = 40.7 °C, 64.2 °C, 56 °C Primer length = 20 nt

4-8 (F)
5' CTG CGA CTG CTT GAT GTA 3'
Tm = 44.5 °C, 64.5 °C, 54 °C Primer length = 18 nt

4-9 (F)
5' CAC CGT ATC GAG AAA GAA TG 3'
Tm = 66.2°C, 58°C Primer length = 20 nt

Vectorette Primers for use with the kit

47
5' TCG AAG CCG GTA ATA TGT TTG TAA TG 3'
Tm = 62.8 °C Primer length = 26 nt

346
5'CGG CAT AAC ATT AAC GGC AGC ATA 3'
Tm = 65.4 °C Primer length = 24 nt
New Primer should work in the conserved region before exon 1.

921
5' TAC TTC CAC ACT TAC AGG CAT CAC CAC 3'
Tm = 62.2°C Primer length = 26 nt
Should work with primer above to give a nice coverage of MT gene.

Primers for Biotin Oligo work

MTEx1Deg
5' CCT TGY AAY TGY TAT YGA AG 3'
Tm=50.8 °C (approx, according to Oligo 6 after adjusting for degeneracy).
Primer = 20nt

MTInt1Spec
5' AAA TAG CAC CAA AAA CAA AAT ACG 3'
Tm 65.8 °C (using Oligo 6). Primer length = 24 nt

MT131Spec
5'AAA TAT GCC TAG CCC TTG TAA 3'
Tm = 54.4 °C Primer length = 21nt

Appendix 2: Solutions and Gel Protocols

(Sambrook et al. 1989)

Molecular Biology Solutions

0.5 M EDTA (pH 8.0)
Dissolve 93.05 g disodium ethylenediaminetetraacetate.$2H_2O$ (EDTA, GibcoBRL®) in 400 ml ddH_2O. Adjust pH to 8.0 using NaOH pellets (approximately 10 g). Make up volume to 500 ml with ddH_2O and autoclave before use.

5X Tris Borate (TBE) Buffer
Dissolve 108 g of Tris Base and 55 g of Boric Acid. Add 40 ml of 0.5 EDTA pH 8.0. Make up to 2 L with ddH_2O.

1X TBE Buffer
1:5 ddH_2O dilution of stock solution made up as needed.

50X Tris Acetate (TAE) Buffer
Dissolve 484 g Tris base in ddH_2O, add 114.2 ml Glacial acetic acid and 200 ml of 0.5 EDTA (pH 8.0). Make up volume to 2 L with ddH_2O.

1X TAE Buffer
1:50 ddH_2O dilution of stock solution made up as needed.

Sequencing Loading Buffer
Aliquot 200 µl formamide and add 40 µl blue dextran. Vortex.

Agarose Gel Protocols

1% Agarose Gel
Dissolve 0.6 g SeaKem LE® agarose (FMC BioProducts) in 60 ml 1X TBE buffer.

Microwave for 1 min, mix by swirling. Microwave for another 15 s until all the agarose has dissolved. Add 2 µl ethidium bromide (10 mg/ml). Allow solution to cool to approximately 50 °C. Pour into taped minigel plate/comb apparatus. Leave to set for 30 min at room temperature. Gel contains 0.6 µg ml^{-1} ethidium bromide, used for DNA visualisation.

Proteomics Solutions

4-7 IPG Buffer
8 M urea, 1% 4-7 IPG buffer, 0.2% dithiothreitol (DTT*) and 4% 3-[(3-Cholamidophropyl)-dimethylammonio]-1-propanesulfonate (CHAPS) made up to required volume using MilliQ ddH_2O. Aliquot into microcentrifuge tubes (1mL) and store at -20°C until required. *Add DTT immediately before use.

6-11 IPG Buffer
7 M urea, 2 M thiourea, 2% CHAPS, 10% isopropanol, 5% glycerol and 1% 6-11 IPG buffer, *2.5% DTT made up to required volume using MilliQ ddH_2O. Aliquot into microcentrifuge tubes (1 mL) and store at -20 °C until required. *Add DTT

immediately before use. 3% DTT IPG buffer solution was added to the cathodic strips for cup loading.

Lysis Buffer
7 M urea, 2 M thiourea 30 mM TrisCl and 4% CHAPS made up to required volume using MilliQ ddH_2O. Aliquot into microcentrifuge tubes (1 mL) and store at -20 °C until required. Check protein samples are at pH 8.5 before labelling with fluorescent dyes. If required adjust pH using 1.5 M Tris.

2X Sample Buffer
7 M urea, 2 M thiourea, 2% pharmalyte, 2% DTT, 4% CHAPS made up to required volume using MilliQ ddH_2O. Aliquot into microcentrifuge tubes (1mL) and store at -20 °C until required.

SDS Equilibration buffer
1.5 M Tris Cl pH 8.8 (final concentration 50 mM), 6M urea, 30% glycerol (v/v), 2% SDS* (w/v), 0.001% bromophenol blue (w/v), made up to required volume using MilliQ ddH_2O. * Take precautions handling this chemical.

Fixing solution
50% Ethanol, 2% Phosphoric acid made up to required volume with MilliQ ddH_2O.

Coomassie stain
17% ammonium sulphate (w/v), 34% methanol (v/v), 2% phosphoric acid, 0.075 g/gel coomassie (G-250) make up to required volume with MilliQ ddH_2O

Appendix 3: *Perna canaliculus* Consensus Sequences (Clones), Primers 131/2089

Includes exon sequences (Figure 2.4) and associated intron sequences

CONTIG ONE

```
          10         20         30         40         50         60         70         80         90         100        110        120
1 CTGTATCGAAAGTAAGTACATTTTTAAACTATTTAAATCAGATATATTAAAAACATCTCTGAATATTACATGTTCATGTGTGAATTCGTTGGCAATTACGGACAATAAGATTAATTAAAATAAAA

    130        140        150        160        170        180        190        200        210        220        230        240        250
1 GTTAAATAAGAATTCCCGTTAAATGTGATCTTTAGTATAGAATATCCTAATCTTTTTGTTTACCCACGTGTGTTTACATTTCTTTAAAGGTCAAGAGGTTGTTAGTCTATAAATAGCACCAAAAA

          260        270        280        290        300        310        320        330        340        350        360        370
1 CAAAATACGTTGATAAAAAGAAAAATATGTTTGATTAGATCTTCACGTCAATCATACTTACAAAAGTTTTTATTTGACT-AAAATAAGATATTTGATTTTAACTAATTCAATTTTGTACTTGTTC

    380        390        400        410        420        430        440        450        460        470        480        490        500
1 TATGTTTATCTTAATCCATTCCAGCAAACGTGTGCATCTGTGGCTCAGGATGCAGCGCAGAAGGTTGCCACTGTGGTGATGCCTGTAAGTGTGGAAGTACTTGTGGATGTCCTGCGTGCAAGGTC

          510        520        530        540        550        560        570        580        590        600        610        620
1 GTTTGCAAGTGTCCAGGTGAGCTTTACTCTCAGTTATGTTTATCTTATGTATGTTTTGTTGTTTTGGGTTTATTTGTTATTTTAGAGAAAAAGGATTTCGATGGCTACCGGCCAGTATTGTTGAT

    630        640        650        660        670        680        690        700        710        720        730        740        750
1 ATCTTTTTGTTGATATTCAGGATAATGGCTCAATGTTGTGAAGAACCTTATGTTTTTTTCCTGGTATRCCGGTAGTTACAKTTAATTTCAACATATGCCTGATCACATGACTATAAATATTGTAT

          760        770        780        790
1 TTTTGATTTTATCACMTTTTGAAATGATAATAGAAACTSCGACTGCTT
```

Refer to Figure 2.4

Clone 1-1, 1-6, 1-8, 1-10, 1-13, 2-3; 3-4, 3-6, 4-8 = Sequence 1

Clone 1-4 = Sequence 2 Clone 2-7, 2-13 = Sequence 3 Clone 2-8 = Sequence 4

Key: = Start of exon 2

SNPs marked in red

40 A to G Clone 1-6
71 T to C Clone 1-8 (F)
188 C to G Clone 1-6
225 T to C Clone 1-10 (R)
239 T to A Clone 3-4 (R)
260 T to G Clone 2-8
263 A to G Clone 2-3
274 A to T Clone 2-3
308 T to C Clone 1-8 F

317 T del Clone 1-8 F
330 A(ins) Clone 2-8 (F)
414 T to C Clone 1-4
422 A to T 2-7, 2-8, 2-13
431 C to T 2-7, 2-8, 2-13
433 G to A 2-7, 2-8, 2-13
501 G to A 2-8
510 T to G 2-8
519 G to A 3-6

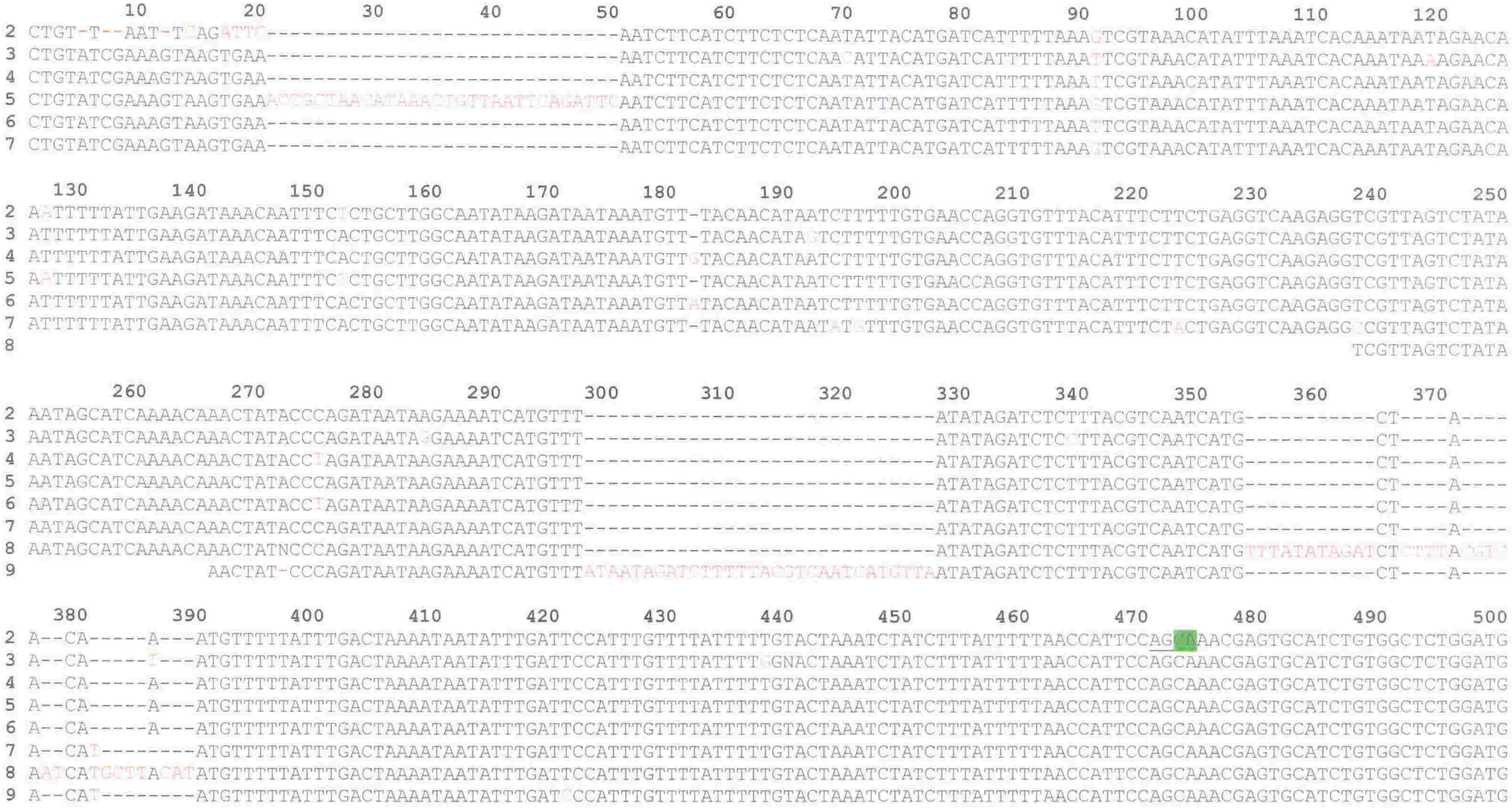
614 A to G Clone 3-6
693 R = G or A
A = 2-3, 2-7, 2-8, 2-13
G = 1-1, 1-4, 1-6, 1-13, 3-6, 4-8
706 K = T or G
G = 2-3, 2-7, 2-8, 2-13
T = 1-1, 1-4, 1-6, 1-13, 3-6, 4-8
732 A to G Clone 4-8
752 T to C 4-8

766 M = A or C
C = 2-3, 2-7, 2-8, 2-13,3-6
A = 1-1,1-4,1-6, 1-13,4-8
789 G to T Clone 3-6

NOTE:
GT at the beginning of an intron
AG at the end of an intron

GT/AG splicing rule (Breathnach 1981)

CONTIG TWO



Refer to Figure 2.4

#Current	Clone #	Figure 2.4
2	Clone 3-9	Sequence 5a
3	Clone 1-7	Sequence 6
4	Clone 4-6	Sequence 5a
5	Clone 3-5	Sequence 5a
6	Clone 4-2	Sequence 5a
7	Clone 3-7	Sequence 5b
8	Clone 2-14	Sequence 5a
9	Clone 3-4	Sequence 5a

CONTIG THREE

```
          10        20        30        40        50        60        70        80        90        100       110       120
10TTGTATCGAAAGTAAGTGAAACCTTATTACCATGTTTAGTGTTTACATTCCGATCAATATGATTTGTTGTAAACAWATTTCTGTAATGTCACATTCGATTATCATGGTATTGATTTTAAAAACGT

     130       140       150       160       170       180       190       200       210       220       230       240       250
10AATTACCACGAATAATGTCTTGATTTTGAATAATTTCGGGTAATTATAAAAATACATRAGAAAAAGATCTTATTCATGTATATTTGTGTACGTCAMGTGGGTGTTTATATTTGGTTGAGATTGAC

          260       270       280       290       300       310       320       330       340       350       360       370
10CCAGATGTAGGTCAAGTAGTCCTTATTCTATAAATAGAATCAAAACAAACAGTCGATAACAAAGTCCTATCCGCGTT-CTGTCACTTTTCATTTAACTCAGACACAAGCATTCATTTG--TTTTT

     380       390       400       410       420       430       440       450       460       470       480       490       500
10TTTATAAAATATAATTGAAGTGAACCATTTTCTTATCTTTCTATCAGCAAACGTTTGTATCTGTGGCACTGCAGAATGCAGCGGTGAAGGTTGTCGCTGTGGTGATGCTTGTAAATGTGGCAGCG

          510       520       530       540       550       560       570       580       590       600       610       620
10ATTGTGGTTGTCCTGGGTGTAAGGTCGTCTGCAAGTGTTTAGGTGAGAATTTTCTTCTATWAGTCTCTTTATTGTATGTGTTGTTTTTTGTTTTGGATTTTTCACCATATCGAGAAAGAATGATT

     630       640       650       660       670       680       690       700       710       720       730
10TCAATGGCTATTTTGATAAATTTTAATGCTTTTTACGATAGTGGTCCAGTACACTACATCACTYCAACTACACTCTTATGTTGATAATTGCTTGATGTATTTTCAGGAGAGTGT
```

SNPs marked in red

4 T to C 4-7
5 A to G 3-10
71 A to C 3-10
76 W = A or T
A = 3-10, 4-7, 4-9
T = 2-1, 2-4
144 C to T 3-10, 4-7
149 T to C 2-1
183 R = A or G
A = 2-1, 2-4, 4-9
G = 1-11, 3-10, 4-7
212 T to C 1-11
221 M = C or A
C = 1-11, 3-10, 4-7, 4-9
A = 2-1, 2-4, 2-9
243 A to G 2-1
252 C to A 4-9

299 A (del) 2-9
306 A to G 1-11
328 T (ins) 2-9
343 S = G or C
G = 1-11, 3-10. 4-7, 4-9
C = 2-1, 2-4, 2-9
369 T (ins) 3-10
369, 370 TT (ins)4-7
478 Y = C or T
C = 2-1, 2-4, 2-9, 4-9
T = 1-11, 3-10, 4-7
501 A to G 4-9
561 W = A or T
A = 1-11, 3-10, 4-7, 4-9
T = 2-1, 2-4, 2-9
572 C to T 1-11
579 T to G 4-9

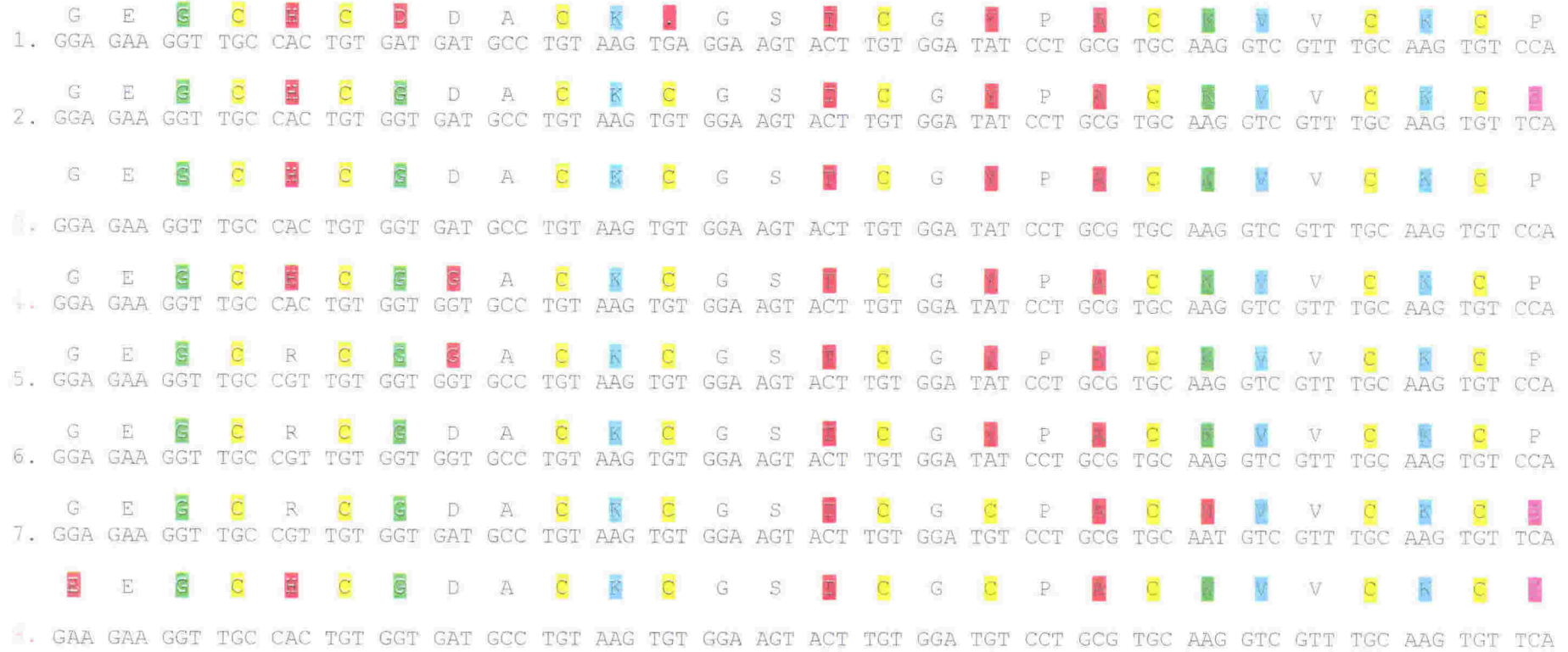
587 R = A or G
A = 2-1, 2-4, 2-9
G = 1-11, 3-10, 4-7, 4-9
597 A to G 4-7
607 A to G 4-9
646 T to C 4-9
665 A to G 3-10
667 T to A 3-10, 4-7
Y = C or T
C = 1-11, 2-4, 2-9
T = 2-1, 2-4, 2-9;
seq terminated here
for 2-1, 2-4, 2-9

BOLD = Primer 2089

Refer to Figure 2.4

Clone 1-11,2-1,2-4,2-9,3-10,4-7 = Seq 7a

Appendix 4: *P. canaliculus* partial exon two cloned sequences and predicted amino acid sequences using primers G5/2089



Key:

Yellow = Conserved Cysteine residue
Red = Unique in *P. canaliculus*
Pink = Observed in other invertebrates
Green = Observed in mussels
Blue = Observed in invertebrates

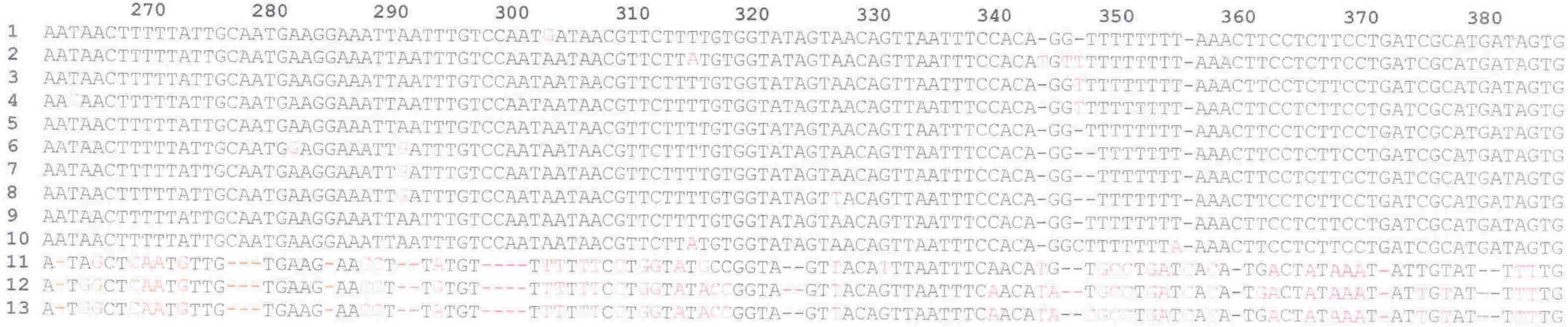
KEY: Clones

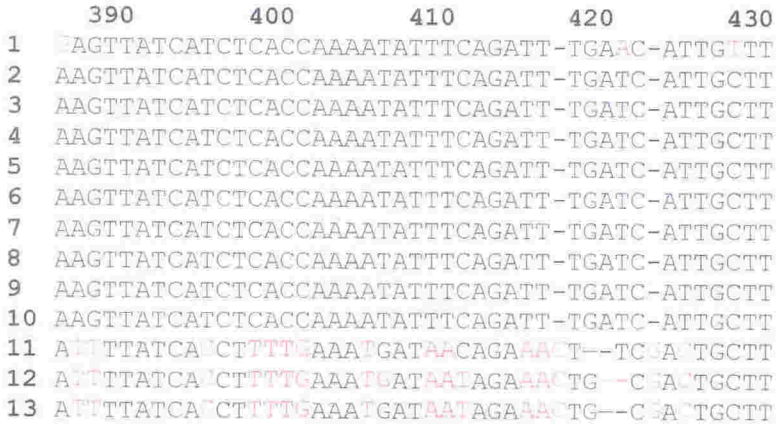
1. 2-15
2. 2-12
3. 2-1, 2-6, 2-13
4. 1-3, 1-4, 1-5, 1-7, 1-10, 1-13
5. 1-12
6. 2-10
7. 1-1
8. 2-9, 2-14

Equivalent sequence in Table 2.4

5 = #5a
4 = #6
3 = #3

Perna canaliculus sequences (Clones), Primers G5/2089





Sequence #	Clone #	Appendix 4 Seq #
1	2-15	1
2	2-12	2
3	2-1, 2-6	3
4	2-13	3
5	1-4	4
6	1-3	4
7	1-10	4
8	1-5, 1-7, 1-13	4
9	1-12	5
10	2-10	6
11	1-1	7
12	2-9	8
13	2-14	8

Appendix 5: BLAST Search Summaries

Summary Protein BLAST (BLASTp)

The top and bottom hit for each amino acid sequence are listed.

Sequence 1

43 matches. ALL MT. Top match is *B. azoricus* (Hydrothermal vent mussel) and *Mytilus* spp.

Other top scores = blue and green mussels, *Meretrix lusoria* (Hard clam), *B. thermophilus* (Hydrothermal vent mussel), M ed. 10kDa forms, *Tegillarca granosa* (Blood clam), *Megathura crenulate* (Giant keyhole limpet).

```
MT [B. azoricus] Length=69

Score = 45.8 bits (107),  Expect = 5e-04
Identities = 17/21 (80%), Positives = 17/21 (80%), Gaps = 0/21 (0%)

Query  19  DACKCGSTCGCPACKVVCKCS  39
           DACKC S CGC  CKVVCKCS
Sbjct  26  DACKCASGCGCSGCKVVCKCS  46

MT [C. rhizophorae] Length=54

Score = 34.7 bits (78),  Expect = 1.0
Identities = 11/19 (57%), Positives = 12/19 (63%), Gaps = 0/19 (0%)

Query  21  CKCGSTCGCPACKVVCKCS  39
           CKCG  C C  CKV C C+
Sbjct  27  CKCGDDCKCAGCKVKCNCT  45
```

Sequence 2

43 hits, as above. ALL MT.

```
MT [B. azoricus] Length=69

Score = 48.1 bits (113),  Expect = 9e-05
Identities = 29/38 (76%), Positives = 31/38 (81%), Gaps = 0/38 (0%)
```

MT [*Megathura crenulata*] Length=70

Score = 32.3 bits (72), Expect = 5.2
Identities = 12/25 (48%), Positives = 13/25 (52%), Gaps = 5/25 (20%)

Sequence 3

43 Hits all MT again.

MT [*B. azoricus*]Length=69

Score = 48.5 bits (114), Expect = 7e-05
Identities = 29/38 (76%), Positives = 31/38 (81%), Gaps = 0/38 (0%)

MT [*Megathura crenulata*] Length=70

Score = 32.3 bits (72), Expect = 5.2
Identities = 12/25 (48%), Positives = 13/25 (52%), Gaps = 5/25 (20%)

Sequence 4

28 Hits, all MT.

MT [*B. azoricus*]Length = 69

Score = 41.2 bits (95), Expect = 0.011
Identities = 26/37 (70%), Positives = 29/37 (78%), Gaps = 0/37 (0%)

MT, isoform MT-10b [*B. azoricus*] Length=73

Score = 33.9 bits (76), Expect = 1.8
Identities = 25/38 (65%), Positives = 28/38 (73%), Gaps = 1/38 (2%)

Sequence 5a

29 Hits, all MT

MT [*B. azoricus*] Length=69

Score = 40.0 bits (92), Expect = 0.025
Identities = 15/20 (75%), Positives = 15/20 (75%), Gaps = 0/20 (0%)

MT, isoform MT-10b [*B. azoricus*] Length=73

Score = 32.3 bits (72), Expect = 5.2
Identities = 13/21 (61%), Positives = 14/21 (66%), Gaps = 1/21 (4%)

Sequence 5b

31 Hits, all MT

MT [*B. azoricus*] Length=69

Score = 41.6 bits (96), Expect = 0.009
Identities = 16/21 (76%), Positives = 16/21 (76%), Gaps = 0/21 (0%)

MT [*C. virginica*] Length=76

Score = 31.6 bits (70), Expect = 8.9
Identities = 11/19 (57%), Positives = 11/19 (57%), Gaps = 0/19 (0%)

Sequence 6

11 Hits all MT.

MT [*Meretrix lusoria*] Length=76 (Hard clam)

Score = 34.3 bits (77), Expect = 1.4
Identities = 12/17 (70%), Positives = 12/17 (70%), Gaps = 0/17 (0%)

MT 20-IIIa - blue mussel Length=71

Score = 31.6 bits (70), Expect = 8.9
Identities = 11/17 (64%), Positives = 12/17 (70%), Gaps = 0/17 (0%)

Sequence 7a

95 Hits all MT. Much higher expect scores. Fewer gaps in seqs?

MT 10-III (MT-10-III)
MT 10 III [*M. edulis*]
MT 10-III [*M. galloprovincialis*] Length=73

Score = 52.4 bits (124), Expect = 5e-06
Identities = 23/27 (85%), Positives = 24/27 (88%), Gaps = 2/27 (7%)

MT-I Length=63

Score = 31.6 bits (70), Expect = 8.8
Identities = 9/19 (47%), Positives = 12/19 (63%), Gaps = 0/19 (0%)

Sequence 7b

29 Hits all MT

MT 10-III (MT-10-III)
MT 10 III [*M. edulis*]
MT 10-III [*M. galloprovincialis*]Length=73

Score = 43.1 bits (100), Expect = 0.003
Identities = 18/20 (90%), Positives = 18/20 (90%), Gaps = 1/20 (5%)

```
MT 20-IV [M. galloprovincialis]Length=72

Score = 33.5 bits (75),  Expect = 2.3
Identities = 14/20 (70%), Positives = 16/20 (80%), Gaps = 1/20 (5%)

Query  1   NVCICGTAECSGEGCRCGDA  20
           NVCICGT  CS + C+CGDA
Sbjct  11  NVCICGTG-CSEKCCQCGDA  29
```

Summary of BLASTn information

Using exon 2 information only.

Sequence 1

Total of 7 hits, but only 2 are for MT (the top two)

(*M. edulis* MT20-IV and *B.thermophilus* (Hydrothermal vent mussel) partial MT-10Bt(2))

Scores, E-value etc are relatively small.

```
Score = 48.1 bits (24),  Expect = 0.005
Identities = 60/72 (83%), Gaps = 0/72 (0%)
```

```
Score = 44.1 bits (22),  Expect = 0.081
Identities = 52/62 (83%), Gaps = 0/62 (0%)
```

Other hits are from small and are cloned non-MT fragments from mouse, rat and *Nicotiana tabacum*,

Sequence 2

Total of 7 hits, only 1 for MT this time, from M. edulis MT-20IV, the top 1

```
Score = 40.1 bits (20),  Expect = 1.3
Identities = 59/72 (81%), Gaps = 0/72 (0%)
```

Rest, same as above.

Sequence 3

16 hits, all small, and NONE MT. Human, mouse etc clones.

Sequence 4

17 hits, none MT, as above

Sequence 5a

20 hits, but only 1 MT, it is the 3rd highest ranked hit. *B. thermophilus*. All hits are very small matches.

```
Score = 40.1 bits (20),  Expect = 1.3
Identities = 50/60 (83%), Gaps = 0/60 (0%)
```

Sequence 5b

20 hits, one MT, 4th highest racked hit *B. thermophilus*. As above.

```
Score = 40.1 bits (20),  Expect = 1.3
Identities = 50/60 (83%), Gaps = 0/60 (0%)
```

Sequence 6

```
26 hits, none MT.
```

Sequence 7a

37 hits, top 20 hits are all for MT. P. viridis = top

```
Score = 60.3 bits (30),  Expect = 1e-06
Identities = 69/83 (83%), Gaps = 0/83 (0%)

MT10III (lowest, out of the MT hits, still better than most of the others)
Score = 46.3 bits (23),  Expect = 0.018
Identities = 27/29 (93%), Gaps = 0/29 (0%)
```

Sequence 7b

As above

```
For P. viridis MT2, then 1

Score = 68.3 bits (34),  Expect = 4e-09
Identities = 70/83 (84%), Gaps = 0/83 (0%)


M. edulis

Score = 46.3 bits (23),  Expect = 0.018
Identities = 27/29 (93%), Gaps = 0/29 (0%)
```

Summary Consensus Seq BLAST

The top and bottom hit for each nucleotide sequence are listed.

Consensus 1

101 hits. Most hits are NOT MT. (~4 out of 101 are MT)

Top Hit:

M. galloprovincialis MT 20-IV (MT-20-IV) mRNA,
complete cds Length = 449

```
Score = 52.2 bits (26),  Expect = 0.003
Identities = 62/74 (83%), Gaps = 0/74 (0%)
Strand=Plus/Plus

Query  399  CAAACGTGTGCATCTGTGGCTCAGGATGCAGCGGAGAAGGTTGCCACTGTGG  450
            ||||||||||| ||||||||  ||||||||||||   || ||||||| |||||
Sbjct  72   CAAACGTGTGTATCTGTGGTACAGGATGCAGCGAGAAATGTTGCCAATGTGG  131

Query  451  TGATGCCTGTAAGTGTGGAAGT  472
             ||||||||  || |||| ||||
Sbjct  122  AGATGCCTGCAAATGTGAAAGT  145
```

Other MT hits scattered, but near the top.

```
Bottom Hit:
Homo sapiens chromosome 18, clone RP11-346H17, complete sequence
Length=165591

Score = 42.2 bits (21),  Expect = 2.5
Identities = 24/25 (96%), Gaps = 0/25 (0%)
Strand=Plus/Minus

Query  116    TAAAATAAAAGTTAAATAAGAATTC  140
               ||||||||||||||| ||||||||
Sbjct  18982  TAAAATAAAAGTTAAAAAAGAATTC  18958
```

Consensus 3

104 Hits, most hits aren't MT. Approx 20 are MT. Top hit = Perna MT

```
Perna viridis MT 2 gene, complete cds Length=2231

Score = 60.4 bits (30),  Expect = 8e-06
Identities = 69/83 (83%), Gaps = 0/83 (0%)
Strand=Plus/Plus

Query  449  TGCAGCGGTGAAGGTTGTCGCTGTGGYGATGCTTGTAAATGTGGCAGCGATTGTGGTTGT  508
            ||||||||  ||||||||||| ||||| || || || |||||| |||| | |||||| |||
Sbjct  717  TGCAGCGGAGAAGGTTGTCGTTGTGGTGACGCGTGCAAATGTAGCAGNGGTTGTGGATGT  776

Query  509  CCTGGGTGTAAGGTCGTCTGCAA  531
             | ||||||||  ||||| |||||
Sbjct  777  TCAGGGTGTAAAGTCGTGTGCAA  799
```

```
Score = 46.4 bits (23),  Expect = 0.13
Identities = 23/23 (100%), Gaps = 0/23 (0%)
Strand=Plus/Plus

Query  714   TTGATGTATTTTCAGGAGAGTGT  736
             |||||||||||||||||||||||
Sbjct  2087  TTGATGTATTTTCAGGAGAGTGT  2109
```

Mus musculus BAC clone RP23-437C1 from chromosome 19, complete sequence
Length=180129

```
Score = 40.4 bits (20),  Expect = 8.0
Identities = 20/20 (100%), Gaps = 0/20 (0%)
Strand=Plus/Plus

Query  282   AAATAGAATCAAAACAAACA  301
             ||||||||||||||||||||
Sbjct  6130  AAATAGAATCAAAACAAACA  6149
```

Appendix 6: Summary of top ranking FASTA results

Seq	Group	Organism	Isoform	Optimised FASTA Score	Identity	Similarity	# Residues	References
MOLLUSCS:								
1	Mussel	*P. viridis*		275	81.6%	94.7%	38	Khoo and Patel 1999
2				258	78,9%	92.1%	38	
3				264	78.9%	92.1%	38	
4				250	75.7%	91,9%	37	
5a,b				251	76.3%	89.5%	38	
6				242	73.7%	86.8%	38	
7a				274	84.6%	92.3%	39	
7b				283	87.2%	94.9%	39	
1		*M. edulis*	MT-20IA&B	271	79.5%	92.3%	39	Mackay 1993
2				254	76.9%	89.7%	39	
3				260	76.9%	89.7%	39	
4				242	71.8%	87.2%	39	
5a				242	73.7%	86.8%	38	
5b				247	74.4%	87.2%	39	
6				233	71.1%	84.2%	38	
7a			MT-10III	269	82.5%	90.0%	40	Mackay 1993; Barstyte 1999
7b			MT-20IA&B	274	84.6%	92.3%	39	
1		*M. edulis*	MT-20II	271	79.5%	92.3%	39	Mackay 1993;Barstyte 1999
2				254	76.9%	89.7%	39	
3				260	76.9%	89.7%	39	
4				242	71.8%	87.2%	39	
5a				242	73.7%	86.8%	38	

5b		*M. edulis*	MT-20II	247	74.4%	87.2%	39	
6				233	71.1%	84.2%	38	
7a			MT20IA&B	265	82.1%	89.7%	39	* More mussel entries, *M edulis* and *M. galloprovincialis.*
7b			MT-20II	274	84.6%	92.3%	39	
1	Oyster	*C. virginica*	MT	193	48.7%	79.5%	39	Unger et al. 1991, Roesijadi et al. 1989, 1991.
2				176	46.2%	76.9%	39	
3				192	48.7%	76.9%	39	
4				177	50.0%	75.0%	36	
5a				174	47.4%	73.7%	38	
5b				176	46.2%	74.4%	39	
6				173	47.4%	73.7%	38	
7a				199	53.8%	76.9%	39	
7b				190	51.3%	74.4%	39	
1	Roman snail	*H. pomatia*	Cd-MT	162	50.0%	73.5%	34	Dallinger et al. 1993
2				151	50.0%	75.0%	32	
3				161	50.0%	70.6%	34	
4				146	51.6%	67.7%	31	
5a				144	48.5%	69.7%	33	
5b				146	47.1%	70.6%	34	
6				144	48.5%	69.7%	33	
7a	Land snail	*A. arbustorum*		155	48.6%	77.1%	35	Berger et al. 1995
7b				164	51.4%	80.0%	35	
1	Land snail	*A. arbustorum*		159	42.9%	80.0%	35	
2				147	39.5%	76.3%	38	
3				159	42.9%	80.0%	35	
4				144	43.8%	78.1%	32	
5a				141	41.2%	76.5%	34	
5b				143	40.0%	77.1%	35	
6	Land snail	*A. arbustorum*		141	41.2%	76.5%	34	

7a	Roman snail	*H. pomatia*	Cd-MT	146	47.1%	70.6%	34	
7b				155	50.0%	73.5%	34	
OTHER INVERTEBRATES:								
1	Sea urchin	*S. granularis*	MTA&B	137	47.1%	76.5%	34	Ciaramella et al. 1997
2				132	42.1%	68.4%	38	
3				137	47.1%	76.5%	34	
4				136	47.1%	73.5%	34	
5a	Silk moth	*Bombyx mori*	Chorion HCA12	133	50.0%	61.1%	36	Iatrou et al. 1984
5b				133	50.0%	61.1%	36	
6				142	52.8%	63.9%	36	
7a	Sea urchin	*S. granularis*	MTB	138	45.7%	74.3%	35	
7b			MTA	147	48.6%	77.1%	35	
1	Sea urchin	*S. neumayeri*	MT	134	45.5%	75.8%	33	Ciaramella et al. 1997
2	Blue crab	*C. sapidus*	Cu-MTII	131	50.0%	73.5%	34	Brouwer et al. 1995
3	Sea urchin	*S. neumayeri*	MT	134	45.5%	75.8%	33	
4				133	45.5%	72.7%	33	
5a	Sea urchin	*S. granularis*	MTA & B	125	47.2%	69.4%	36	
5b				125	47.2%	69.4%	36	
6				116	44.4%	66.7%	36	
7a	Silk moth	*Bombyx mori*	Chorion HCA12	140	48.6%	59.5%	37	
7b				149	51.4%	62.2%	37	
1	Silk moth	*Bombyx mori*	Chorion HCA12	136	50.0%	61.1%	36	
2	Sea urchin	*S. neumayeri*	MT	126	37.5%	65.0%	40	
3	Blue crab	*C. sapidus*	Cu-MTII	131	48.6%	67.6%	37	
4	Fruit fly	*D. ananassae*	MT-1	129	47.2%	63.9%	36	Stephan et al. 1994
5ab	Sea urchin	*S. neumayeri*	MT	118	39.5%	62.8%	43	
6	Fruit fly	*D. ananassae*	MT-1	111	48.6 %.	62.9%	35	
7a	Sea urchin	*S. neumayeri*	MT	135	44.1%	73.5%	34	
7b				144	47.1%	76.5%	34	

VERTEBRATES:

1	Human	*H. sapiens*	MT-4	139	50.0%	61.1%	36	Quaife et al. 1994
2	Eelpout	*T. cereberus*	MT-A	130	51.6%	71.0%	31	Kille and Olsson 1996
3	Human	*H. sapiens*	MT-4	139	50.0%	61.1%	36	
4				133	48.5%	63.6%	33	
5a				135	48.6%	62.2%	37	
5b				129	48.6%	59.5%	37	
6	Arctic char	*S. alpinus*	MTB	128	54.8%	71.0%	31	
7a	Eelpout	*T. cereberus*	MT-A	139	53.1%	68.8%	32	
7b				137	53.1%	68.8%	32	
1	Dog	*C. familiaris*	MT-IV	131	50.0%	58.3%	36	Kobayashi et al. 1999
2	Pike	*E. lucius*	MT	130	54.8%	71.0%	31	
3				133	54.8%	74.2%	31	
4	Eelpout	*T. cereberus*	MT-A	130	53.6%	78.6 %	28	
5a	Dog	*C. familiaris*	MT-IV	127	48.6%	59.5%	37	
5b			MT-IV	121	48.6%	56.8%	37	
6	Eelpout	*Z. viviparous*	MT	124	51.6%	71.0%	31	
7a	Pike	*E. lucius*	MT	132	53.1%	68.8%	32	
7b	Human	*H. sapiens*	MT-3	128	50.0%	64.7%	34**	
1	Mouse	*M. musculus*	MT-IV	131	50.0%	58.3%	36	Quaife et al. 1994
2	Human	*H. sapiens*	MT-4	128	54.8%	64.5%	31	
3	Eelpout	*T. cereberus*	MT-A	133	51.6%	74.2%	31	
4	Arctic char	*S. alpinus*	MTB	126	53.6%	75.0%	28	
5a	Mouse		MT-IV	127	48.6%	59.5%	37	
5b				121	48.6%	56.8%	37	
6	Pike	*E. lucius*	MT	123	51.5%	66.7%	33	
7a	Arctic char	*S. alpinus*	MTB	131	53.1%	65.6%	32	
7b	Mouse		MT-3	126	45.9%	59.5%	37	

Organisation of Table 4.

All hits within a section (e.g. Molluscs) are arranged in decreasing order of significance.
Molluscs: Top 3 mussel hits are shown, then ignoring mussel hits, the next top 3 mollusc hits are shown.
Other sections: Top 3 hits, from 3 different species are shown (i.e. if 2 of the top 3 hits are from the same species, the second hit is not shown).

NB: In contrast with other sequences, ** 7b has a lot of high ranking plant hits. However, they still didn't have sufficient rank to feature on the table.

References:

Barsyte D, White K, Lovejoy D. 1999. Cloning and characterization of metallothionein cDNAs in the mussel *Mytilus edulis* L. digestive gland. Comp Biochemistry Physiology C Pharmacology Toxicology Endocrinology 122(2):287-96.

Berger B, Hunziker,P.E., Hauer,C.R., Birchler,N. and Dallinger,R. 1995. Mass spectrometry and amino acid sequencing of two cadmium-binding metallothionein isoforms from the terrestrial gastropod *Arianta arbustorum*. Biochemisry Journal 311 (Pt 3):951-957.

Brouwer M, Enghild J, Hoexum-Brouwer T, Thogersen I, Truncali A. 1995. Primary structure and tissue-specific expression of blue crab (*Callinectes sapidus*) metallothionein isoforms. Biochemistry Journal 311(Pt 2):617-622.

Ciaramella M, Filosa S, Parisi E. 1997. PCR amplification and cloning of metallothionein complementary DNAs in temperate and Antarctic sea urchin characterized by a large difference in egg metallothionein content. Cellular and Molecular Life Science 53(5):472-477.

Dallinger R, Berger B, Hunziker P, Birchler N, Hauer C, Kagi J. 1993. Purification and primary structure of snail metallothionein. Similarity of the N-terminal sequence with histones H4 and H2A. European Journal of Biochemistry 216:739-746.

Gerpe M, Kling P, Olsson P. 1998. Metallothionein gene expression in arctic char (*Salvelinus alpinus*) following metal and PCB exposure. Marine Environmetal Research 46(1-5):551-554.

Iatrou K, Tsitilou S, Kafatos F. 1984. DNA sequence transfer between two high-cysteine chorion gene families in the silkmoth *Bombyx mori*. Proceedings of the National Academy of Science 81(14):4452-4456.

Khoo H, Patel K. 1999. Metallothionein cDNA, promoter, and genomic sequences of the tropical green mussel, *Perna viridis*. Journal of Experimental Zoology 284(4):445-53.

Kille P, Olsson P. The use of metallothionein genes for determining the phylogenetic and evolutionary relationship between extant teleosts. Unpublished.

Kille P, Stephens P, Kay J. 1991. Elucidation of cDNA sequences for metallothioneins from rainbow trout, stone loach and pike liver using the polymerase chain reaction. Biochimica et Biophysica Acta 1089(3):407-410.

Kobayashi K, Morita T, Shimada A. 1999. *Canis familiaris* (dog) MT-IV. Unpublished. SWISSPROT Entry.

Mackay EA, Overnell, J., Dunbar, B., Davidson, I., Hunziker, PE., Kagi, JHR., Fothergill, JE. 1993. Complete amino acid sequences of five dimeric and four monomeric forms of metallothionein from the edible mussel *Mytilus edulis*. European Journal of Biochemistry 218:183-194.

Mehra R, Garey J, Butt T, Gray W, Winge D. 1989. Candida glabrata metallothioneins. Cloning and sequence of the genes and characterization of proteins. Journal of Biology and Chemistry 264(33):19747-19753.

Quaife C, Findley S, Erickson J, Froelick G, Kelly EJ, Zambrowicz B, Palmiter R. 1994. Induction of a new metallothionein isoform (MT-IV) occurs during differentiation of stratified squamous epithelia. Biochemistry 33(23):7250-7259.

Roesijadi G, Kielland, S., Kerks, P. 1989. Purification and properties of novel molluscan metallothioneins. Archives of Biochemistry and Biophysics 273(2):403-413.

Roesijadi G, Vestling M, Murphy C, Klerks P, Fenselau C. 1991. Structure and time-dependent behavior of acetylated and non-acetylated forms of a molluscan metallothionein. Biochimica et Biophysica Acta 1074(2):230-6.

Stephan W, Rodriguez V, Zhou B, Parsch J. 1994. Molecular evolution of the metallothionein gene Mtn in the melanogaster species group: results from *Drosophila ananassae*. Genetics 138(1):135-143.

Unger M, Chen T, Murphy C, Vestling M, Fenselau C, Roesijadi G. 1991. Primary structure of molluscan metallothioneins deduced from PCR-amplified cDNA and mass spectrometry of purified proteins. Biochimica et Biophysica Acta 1074(3):371-7.

Appendix 7: References Mollusc MT Alignment

NB: Some sequences have both an amino acid (protein) accession number (as shown in Table 4) and a sequence accession number. Both are shown below if applicable in the references below. Some sequences, have more than one reference. For example, 10-IV appears in 3 separate studies: McKay et al.; Lemoine et al.; and Barstyte et al. Only 1 reference appears in the table above due to space constraints, full references are shown below.

P. canaliculus; current study

P. viridis; Accession numbers: MT cds = AAD02054 (AF036904), MT1 = AAF22486 (AF092971), MT2 = AAF22486 (AF092971); Khoo H, Patel K. 1999. Metallothionein cDNA, promoter, and genomic sequences of the tropical green mussel, Perna viridis. Journal of Experimental Zoology 284(4):445-53.

M. edulis; Accession numbers: MT10 = CAE11861 (AJ577130), MT20 = CAE11862 (AJ577131); Leignel,V. and Laulier,M. Isolation and characterization of Mytilus edulis metallothionein genes. Comparative Biochemistry and Physiology C 142 (1), 12-18 (2006)

M. edulis; Accession numbers: MT10A(2) = CAE11856 (AJ577125), MT10A(1)= CAE11855 (AJ577124). Leignel,V., Hardivillier,Y. and Laulier,M. Small metallothionein MT-10 genes in coastal and hydrothermal mussels. Marine Biotechnology 7 (3), 236-244 (2005)

M. edulis; Accession numbers: MT10B(2) = CAE11858 (AJ577127), MT10B(1) = CAE11857 (AJ577126). Leignel,V., Hardivillier,Y. and Laulier,M. MT-10 metallothionein genes without intron on coastal and hyhdrothermal mussels. Unpublished

M. edulis; Accession numbers: MT10IIb = CAA06550 (AJ005453), MT10Ib = CAA06549 (AJ005452, O62554), MT10Ia = P80246 (CAA06548, AJ005451), MT20II = P80252 (CAA06553, AJ005456), MT10IV = P80249 (CAA06552) (AJ005455 ALSO LEMOINE), MTIII = P80248 (CAA06551, AJ005454); Barsyte,D., White,K.N. and Lovejoy,D.A. 1999. Cloning and characterization of metallothionein cDNAs in the mussel Mytilus edulis L. digestive gland. Comparative Biochemistry and Physiology C 122, 287-296

M. edulis; Accession number MTIV = P80249. Lemoine,S., Bigot,Y., Sellos,D., Cosson,R.P. and Laulier,M. Metallothionein Isoforms in Mytilus edulis (Mollusca, Bivalvia): Complementary DNA Characterization and Quantification of Expression in Different Organs after Exposure to Cadmium, Zinc, and Copper. Marine Biotechnology 2 (2), 195-203 (2000)

M. edulis; MTIV = P80249 (CAA07546, AJ007506) SAME AS Barsyte. Lemoine,S., Bigot,Y., Sellos,D. and Laulier,M. Metallothionein in Mytilus edulis (Mollusca, bivalvia):cDNA characterization and quantification of the expression. Unpublished

M. edulis; Accession numbers: MT20Ia = AAB29062 (P080251, S39420), MT10Ia = P80246 (AAB29061), MTIV = P80249 (AAB29060), MT20II = P80252, MTIII = P80248, MT20IIIa = P69153 (S39422), MT10II = P80247, MT20IIIb = 80258 (S47577), MT20Ia&b = PO80251 (S47576); MacKay, E.A., Overnell, J., Dunbar, B., Davidson, I., Hunziker, P.E., Kagi, J.H.R., Fothergill, J.E. 1993. Complete amino acid sequences of five dimeric and four monomeric forms of metallothionein from the edible mussel Mytilus edulis. European Journal Biochemistry 218, 183-194.

M. galloprovincialis; Accession numbers: MT1 = AAG28538 (AF199020), MT20IIIa = P69154. Ceratto,N., Dondero,F., van de Loo,J.W., Burlando,B. and Viarengo,A. Cloning and sequencing of a novel metallothionein gene in Mytilus galloprovincialis Lam Comparative Biochemistry and Physiology C 131 (3), 217-222 (2002)

M. galloprovincialis; Accession numbers: MT10III = AAT72936 (AY56648), MT20IV = AAT72935 (AY566247). Dondero,F., Piacentini,L., Banni,M., Rebelo,M., Burlando,B. and Viarengo,A. Quantitative PCR analysis of two molluscan metallothionein genes unveils differential expression and regulation. Gene 345 (2), 259-270 (2005)

D. polymorpha; Accession number: U67347; Engelken,J. and Hildebrandt,A. 1999. cDNA cloning and cadmium-induced expression of metallothionein mRNAin the zebra mussel Dreissena polymorpha. Biochemistry and Cell Biology 77 (3), 237-241

C. gigas; Accession numbers: MT2 = CAC48045 (AJ297818), MT = CAB64869 (AJ242657); Tanguy,A. and Moraga,D. 2001. Cloning and characterization of a gene coding for a novel metallothionein in the Pacific oyster Crassostrea gigas (CgMT2): a case of adaptive response to metal-induced stress? Gene 273 (1), 123-130 (2001)

C. angulata; Accession number: MT = AAK155181 (AF349907). Funes,M.V., Lopez-Barea,J. and Dorado,G. Complete metallothionein cds (mRNA) from Crassostrea angulata. Unpublished

C. gigas; Accession number: MT3 = CAC82788 (AJ295157). Tanguy,A., Boutet,I. and Moraga,D. Cloning and characterization of a third metallothionein gene in the Pacific oyster Crassostrea gigas: a singular case. Unpublished

C. gigas; Accession number: MTmRNAp = CAB85588 (AJ243263). Tanguy,A. and Moraga,D. Molecular cloning and characterization of a new metallothionein gene in Crassostrea gigas. Unpublished

C. virginica; Accession numbers: mRNA = P23038 (CAA42522, X59862) Unger,M.E., Chen,T.T., Murphy,C.M., Vestling,M.M., Fenselau,C. and Roesijadi,G. Primary structure of molluscan metallothioneins deduced from PCR-amplified cDNA and mass spectrometry of purified proteins. Biochimica et Biophysica Acta 1074 (3), 371-377 (1991)

C. virginica; Accession numbers: MTA = AAM 90257 (AF506977), MTB = AAM90258 (AF506978 same as MTIIIC AAQ23919); Tanguy,A., Yu,Z., Guo,X. and Ford,S. 2002. Cloning and expression of two metallothionein genes in the Eastern oyster Crassostrea virginica, GenBank 2002, Unpublished.

C. virginica; Accession number: mRNA = P23038.Roesijadi,G., Kielland,S. and Klerks,P. Purification and properties of novel molluscan metallothioneins. Archives of Biochemistry and Biophysics 273 (2), 403-413 (1989)

C. virginica; Accession number: mRNA = P23038. Roesijadi,G., Vestling,M.M., Murphy,C.M., Klerks,P.L. and Fenselau,C.C. Structure and time-dependent behavior of acetylated and non-acetylated forms of a molluscan metallothionein. Biochimica et Biophysica Acta 1074 (2), 230-236 (1991)

C. virginica; Accession number: MTIIIA = AAQ23917 (AY331708), MTIIIB = AAQ23918 (AY331709), MTIIIC = AAQ23919 (AY331710 same as MTB AAM90258, MTIA = AAQ23907 (AAQ23906, AAQ23906, AAQ23904, AY331698, AY331697, AY221697, AY331695), MTIB = AAQ23908 (AY331699), MTIIA = AAQ23909 (AY331700), MTIIB = AAQ23910 (AY331701), MTIIC = AAQ23911 (AY331702), MTIIDp = AAQ23912 (AY331703, only partial on table as full sequence is 149aa), MTIIEp = AAQ23913 (AY331704, only partial on table as full sequence is 149aa), MTIIFp = AAQ23914 (AY331705, only partial on table as full sequence is 145aa), MTIIGp = AAQ23915 (AY331706, only partial on table as full sequence is 204aa), MTIIHp = AAQ23916 (AY331707), only partial on table as full sequence is 200aa),),. Jenny,M.J., Ringwood,A.H., Schey,K., Warr,G.W. and Chapman,R.W. Confirmation of Diversity and Differential Gene Expression of Two Distinct Metallothionein Gene Families from Crassostrea virginica. Unpublished 2003.

C. rhizphorae; Accession number: MT = AAK50565 (AY029512). Moraes,M.O., Rebelo,M.F., Junior,H.S., Sampaio,E.P. and Pfeiffer,W.C. Direct detection of metallothionein mRNA by RT-PCR in mangrove oysters (Crassostrea rhizophorae): in situ evaluation for a biomarker of metal contamination. Unpublished
O. edulis; Accession numbers: AJ306365, AJ306366; Tanguy,A., Boutet,I., Riso,R., Auffret,M. and Moraga,D. 2002. Cloning and characterization of two metallothionein genes in the european flat oyster Ostrea edulis, Unpublished

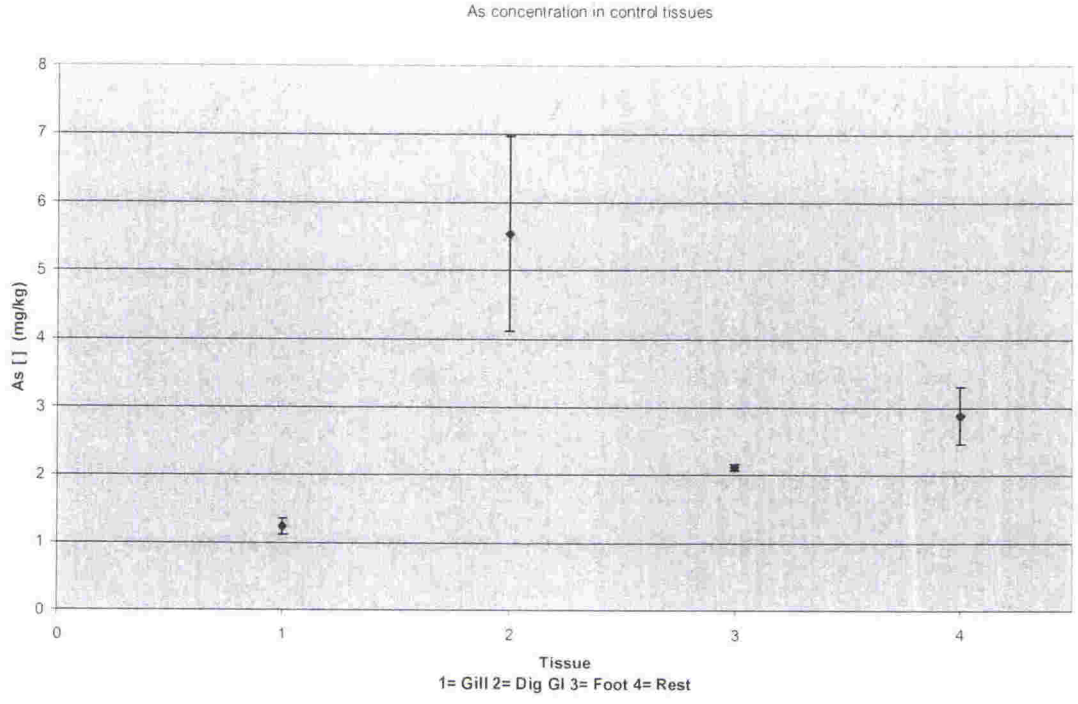
R. ducussates; Ruditapes phillipinarum; Venerupis pullastra; Accession numbers: CAB96402 (AJ249687), CAB96403 (AJ249686), CAB96419 (AJ249688, respectively): Moraga,D. and Tanguy,A. Characterisation of metallothioneins sequenies in three clams Ruditapes decussatus, Ruditapes philippinarum and Venerupis pullastra. Unpublished.

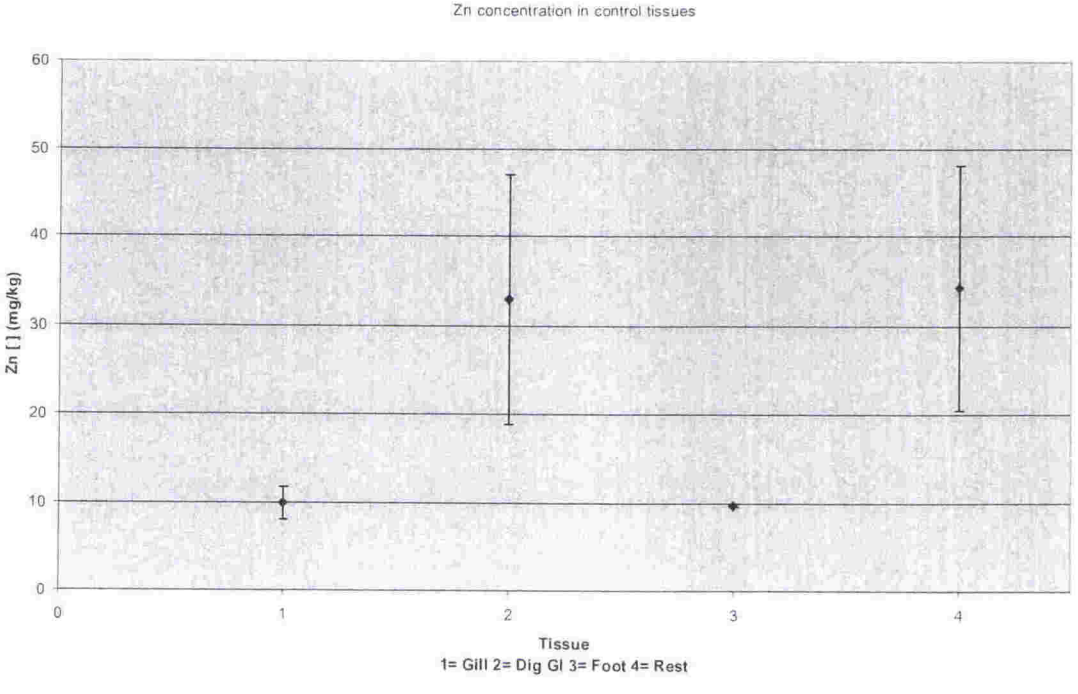
A. granosa; Accession number: mRNA = AAK39563. Mooi Kwai,C., Surif,S., Othman,R. and Din,Z. Induction of metallothionein gene in the blood cockles, Anadara granosa, exposed to cadmium. Unpublished

H. pomatia; Accession numbers: Cd-MT = P33187, MT-Cu = P55947. Dallinger,R., Berger,B., Hunziker,P.E., Birchler,N., Hauer,C.R. and Kagi,J.H. Purification and primary structure of snail metallothionein. Similarity of the N-terminal sequence with histones H4 and H2A. European Journal of Biochemistry 216 (3), 739-746 (1993)

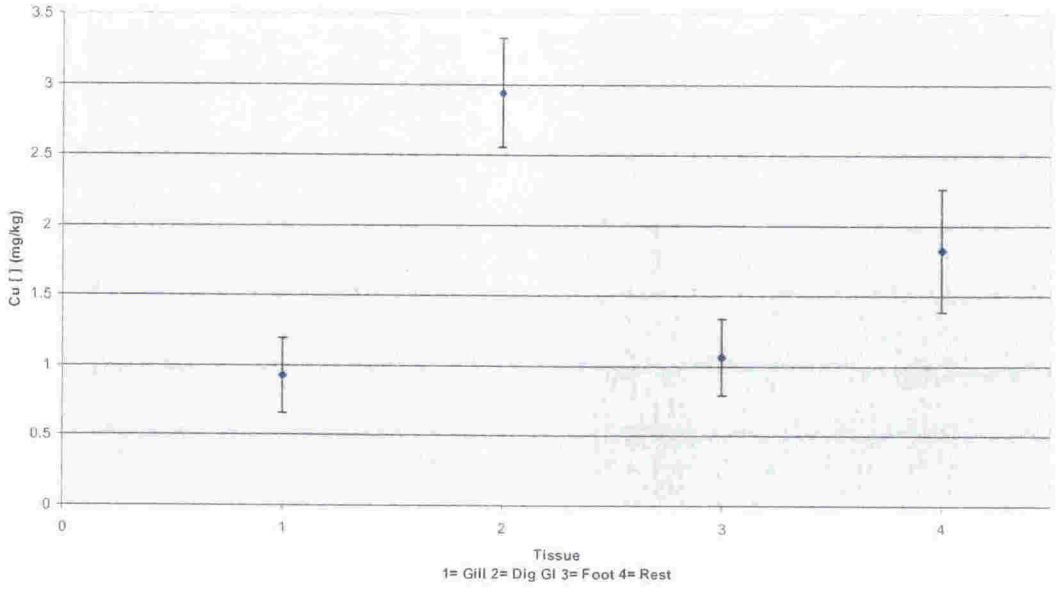
Appendix 8: Metal concentrations in control mussels at Time 0

Additional results showing the concentrations of metals detected in control samples at Time 0 in the heavy metal experiment (Chapter 5).

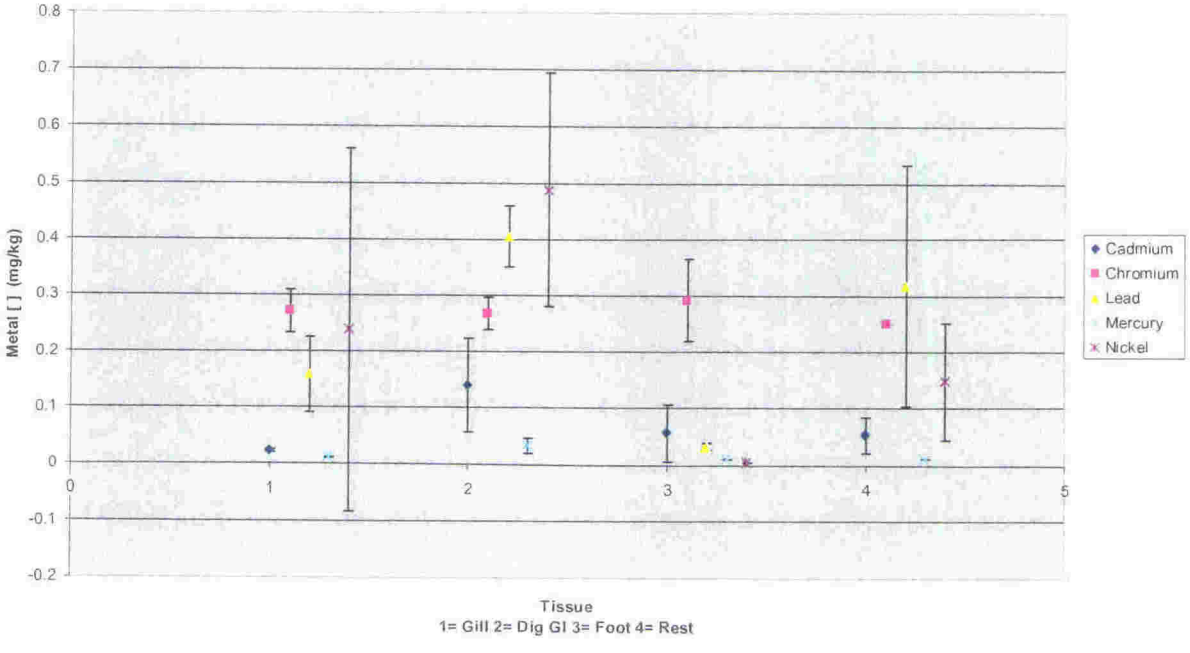




Cu concentration in control tissues



Cd, Cr, Pb, Hg and Ni concentrations in control tissue



Appendix 9: "Trigger" values for marine water protection

	Trigger values for marine water (mg l^{-1})			
	Level of protection (% species)			
Metals	**99%**	**95%**	**90%**	**80%**
Arsenic	n.d	n.d	n.d	n.d
Cadmium	0.0007 [B]	0.0055 [B,C]	0.014 [B,C]	0.036 [A,B]
Copper	0.0003	0.0013	0.003 [C]	0.008 [A]
Lead	0.0022	0.0044	0.0066 [C]	0.012 [C]
Mercury (inorganic)	0.0001	0.0004 [C]	0.0007 [C]	0.0014 [C]
Zinc	0.007	0.015 [C]	0.023 [C]	0.043 [C]

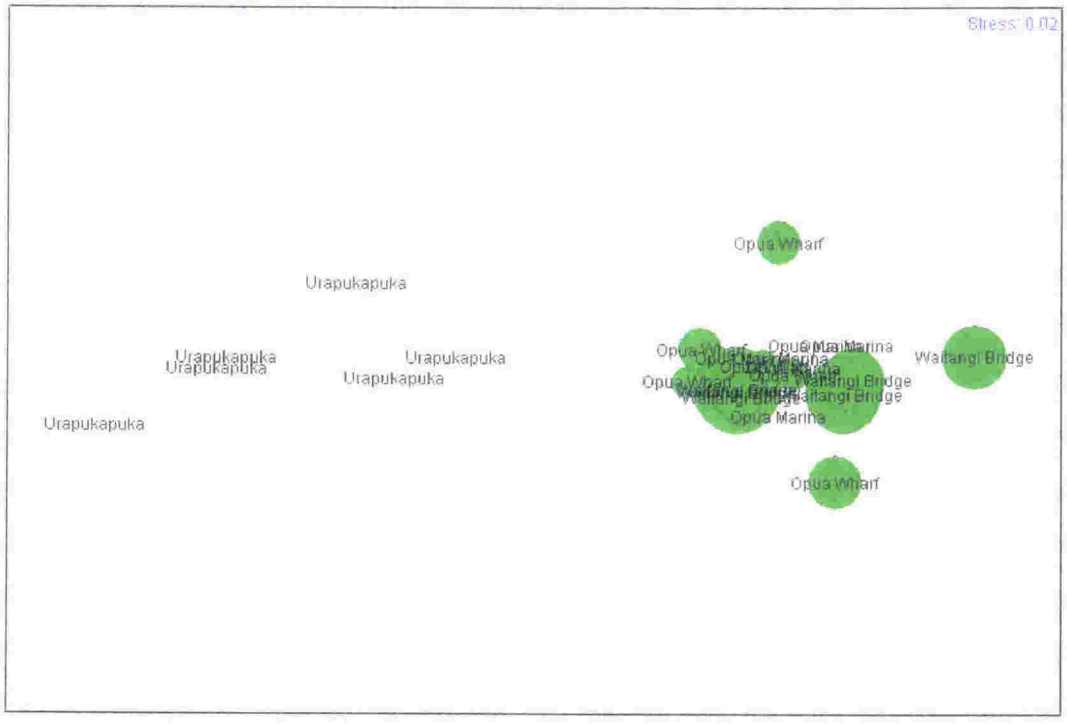
Table 9.1: Trigger values for selected metal toxicants at alternative levels of protection.

Values in grey shading are the trigger values applying to typical *slightly-moderately disturbed* systems. Table has been selectively reproduced from the Australian and New Zealand Guidelines for Fresh and Marine Water Quality (2000).

A = Figure may not protect key test species from acute toxicity (and chronic). 'A' indicates that trigger value > acute toxicity figure: note that trigger value should be <1/3 of acute figure.
B = Chemicals for which possible bioaccumulation and secondary poisoning effects should be considered.
C = Figure may not protect key test species from chronic toxicities.
For further details refer to the Australian and New Zealand Environment and Conservation Council. The Australian and New Zealand Guidelines for Fresh and Marine Water Quality.

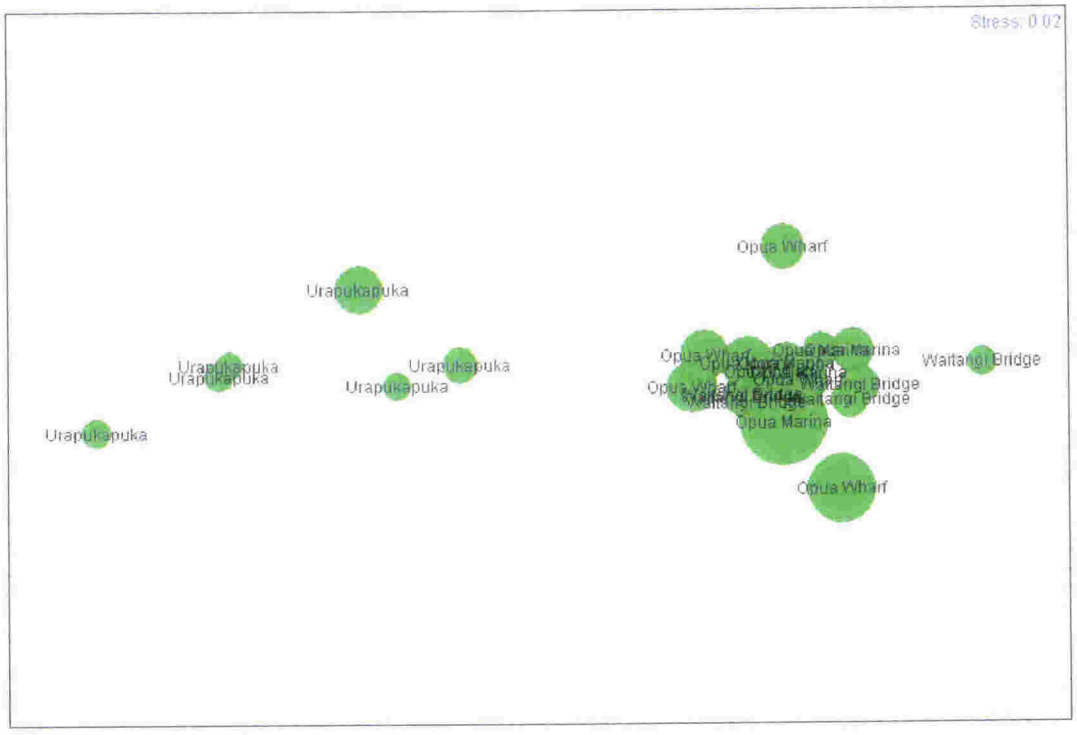
Appendix 10: Bay of Islands nMDS plots

Non-standardised and non-transformed data, multidimensional scaling based on Bray-Curtis dissimilarity.

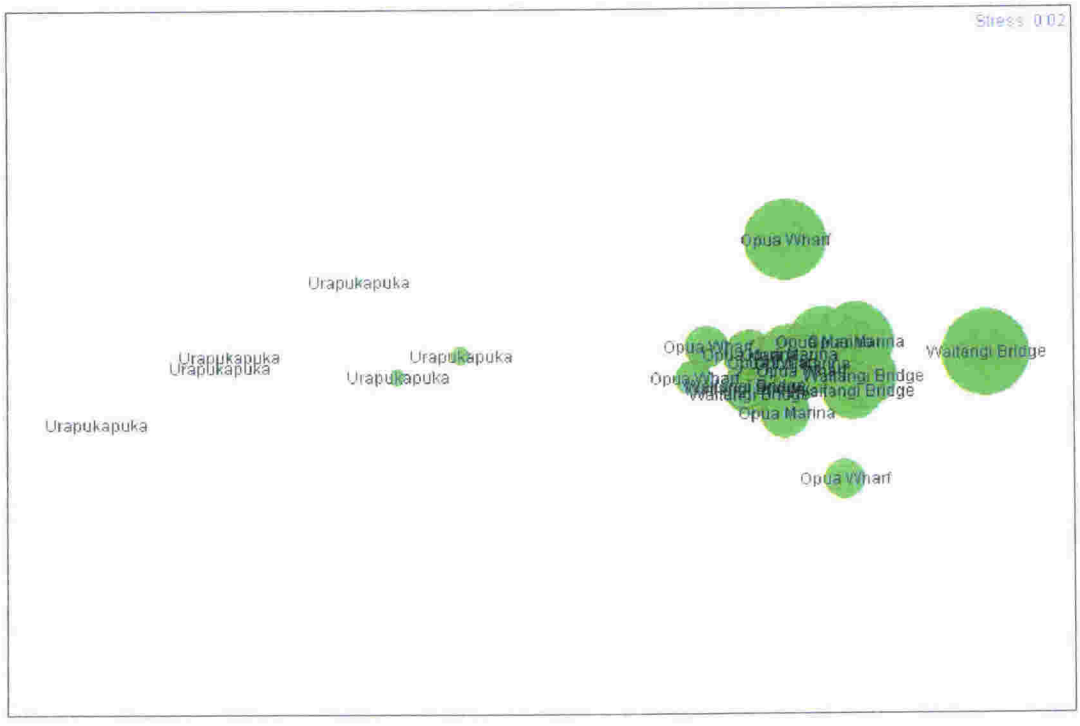


nMDs plot showing dissimilarities among Bay of Island locations
Green circles indicate relative concentrations of mercury

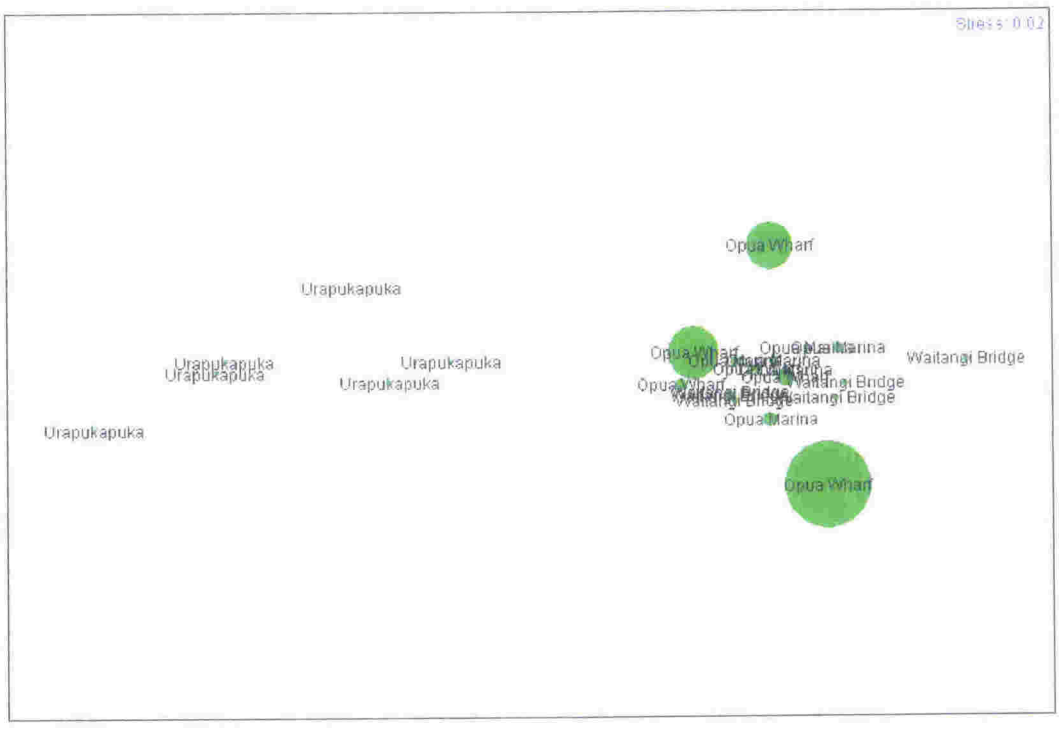
nMDs plot showing dissimilarities among Bay of Island locations
Green circles indicate relative concentrations of iron



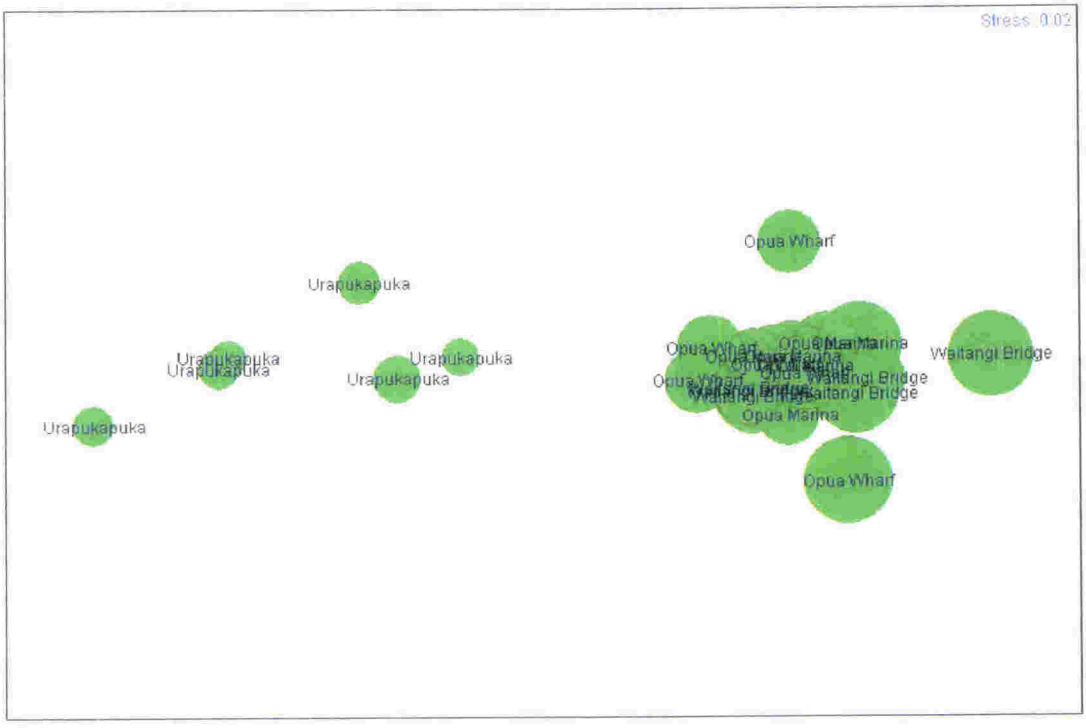
nMDs plot showing dissimilarities among Bay of Island locations
Green circles indicate relative concentrations of zinc



nMDs plot showing dissimilarities among Bay of Island locations
Green circles indicate relative concentrations of aluminium



nMDs plot showing dissimilarities among Bay of Island locations
Green circles indicate relative concentrations of copper



nMDs plot showing dissimilarities among Bay of Island locations
Green circles indicate relative concentrations of titanium

Appendix 11: Statistical Results, Bay of Islands Study

In most cases, after permutation testing and FDR (Benjamini and Hochberg 2000) adjustment, spots which had appeared initially significant were no longer statistically significant. The short listed spots have been reported in full to indicate additional spots which may warrant further investigation in future analyses (Figures 11.1 and 11.2).

Overall ANOVA and permutation results using data from all 7 sites (pI 4-7)

Initially, a basic ANOVA analysis was conducted. Assumptions for ANOVA are normality of data, equal variance and independence. Because of the non-independence of data, permutation (randomisation) tests were also performed as these tests do not rely on any of the ANOVA assumptions. Preliminary analysis was performed using 1000 replications to eliminate the least significant spots. Then 200000 replications were carried out on the 10 most significant spots in order to give enough decimal places for multiple testing adjustments by Bonferroni (Quinn and Keough 2002) and FDR (Benjamini and Hochberg 2000). Only one spot (1395) remained significant at the 5% level (*). The permutation tests give more conservative results than the basic ANOVA in cases where there is considerable dependence in the data (Manly 1997). There appears to be some failure of the ANOVA assumptions, because otherwise both of these tests should give similar results.

Spot	Spot # Chapter 5	1-ANOVA raw *p*-value	Bonferroni adjusted *p*-value	FDR adjusted *p*-value	Permutation raw *p*-value	Bonferroni adjusted *p*-value	FDR adjusted *p*-value
1256		0.0062719708	1.0000000000	0.1545664112	0.01565	1	0.41675
1342		0.0024767587	0.4829679521	0.0825586243	0.00807	1	0.376
1361		0.0159295792	1.0000000000	0.2744776594			
*1395		0.0000004585	0.0000912455	0.0000458520	0.00015	0.03	0.03
1599		0.0001016248	0.0200200925	0.0050812418	0.019755	1	0.538
1770		0.0000593794	0.0117571162	0.0039586250	0.01667	1	0.4575714
1906	2047	0.0232756111	1.0000000000	0.2982443523			
1924	2068	0.0066202163	1.0000000000	0.1655054084	0.016015	1	0.4435
1928		0.0164686596	1.0000000000	0.2896287134	0.00887	1	0.3811111
2123		0.0002177923	0.0426872852	0.0087116909	0.01715	1	0.5216667
2294	2396	0.0069554885	1.0000000000	0.1791991649	0.0094	1	0.3951
2397	2488	0.0109493018	1.0000000000	0.2189860368			
2601	2681	0.0000001897	0.0000379492	0.0000379492	0.00097	0.19303	0.097

Table 11.1: Overall ANOVA and permutation using data from all 7 sites (pI 4-7), large dataset

Overall ANOVA and permutation results using data from all 7 sites (Gel 6-11)

A basic ANOVA analysis and permutation (randomisation) test was conducted on proteins from 6-11 gels despite some missing values (1 gel had to be removed). Preliminary analysis used 1000 replications to eliminate the least significant spots, and then 200000 replications were carried out on the most significant spots in order to give enough decimal places for multiple testing adjustments by Bonferroni and FDR. No spots remained significant (Figure 11.2). The comparisons between groups were limited by the missing values in this dataset and were therefore not performed.

Spot	1-ANOVA raw *p*-value	Permutation Raw *p*-value	Bonferroni adjusted *p*-value	FDR adjusted *p*-value
574	0.003399	0.126338	1	1
683	0.004859	0.045883	1	1
929	0.005464	0.076338	1	1
1014	0.036765	0.061104	1	1
1203	0.071423	0.096545	1	1

Table 11.2: Overall ANOVA and permutation using data from all 7 sites (pI 6-11),

Pairwise comparisons and overall permutations for Bay of Islands data

Tests were conducted using only data from the Bay of Islands to determine whether there were any consistent differences between these sites. The following permutation tests were performed on the large dataset: 1) Among all 4 Bay of Islands sites (overall). 2) All pairwise comparisons among the 4 Bay of Islands sites. 3) Planned contrast between clean (Urapukapuka) versus the 3 other sites. Preliminary screening with only 1000 replications found 9 different spots with possible significance for at least 1 of the tests. A full analysis of these, with 200000 replications, gave the results reported in Table 11.3. Both raw and FDR adjusted values are shown. No spots were still significant.

Planned contrast of Wellington (n = 3) versus Waitangi (n = 4) sites

The average values for Wellington sites were compared with the averages for Waitangi sites. Preliminary screening was conducted using 1000 replications to select the 10 most significant spots, and repeated using 200000 replications. After FDR adjustment, no spots remained significant.

Spot	Spot # in Chapter 5	Permutation raw *p*-value	Permutation adjusted *p*-value	1-ANOVA values from DeCyder
1256		0.0006660007	0.1332001	0.0037
1342		0.0113220113	1.0000000	0.046
1395		0.4971694972	1.0000000	0.00095
1599		0.0093240093	1.0000000	0.028
1770		0.0223110223	1.0000000	0.048
1924	2068	0.0436230436	1.0000000	0.0059
1953	2095	0.0176490176	1.0000000	0.0070
2072		0.0029970030	0.5964036	0.063
2123		0.0459540460	1.0000000	0.12
2601	2681	0.1318681319	1.0000000	0.0015

Table 11.4: Planned contrast of Wellington (n = 3) versus Waitangi (n = 4) sites, large dataset

spot	Spot # Chapt. 5	overall	overall raw	OM vs. URA	OM vs. URA raw	OW vs. URA	OW vs. URA raw	URA vs. WB	URA vs. WB raw
1010		0.58	0.002900	1	0.220695	1.00000	0.140885	1.00	0.795785
1354		1.00	0.006845	1	0.091265	0.54128	0.002720	1.00	0.076165
1361		1.00	0.006105	1	0.078060	0.14800	0.000740	1.00	0.026515
1395		1.00	0.007765	1	0.072730	0.85338	0.004310	1.00	0.073790
1417		1.00	0.524020	1	0.932080	1.00000	0.405690	1.00	0.456630
2094	2215	1.00	0.007020	1	0.006545	1.00000	0.072835	1.00	0.089975
2294	2396	1.00	0.026430	1	0.315535	1.00000	0.014260	0.88	0.004400
2440	2541	1.00	0.037885	1	0.657585	1.00000	0.018835	1.00	0.681695
2601	2681	1.00	0.055725	1	0.020450	1.00000	0.022200	1.00	0.034730
spot	**Spot # Chapt. 5**	**OM vs. OW**	**OM vs. OW raw**	**OM vs. WB**	**OM vs. WB raw**	**OW vs. WB**	**OW vs. WB raw**	**URA vs. the rest**	**URA vs. the rest raw**
1010		0.02100*	0.000105	1.000	0.349505	1.000	0.065900	1	0.857640
1354		0.19701	0.000990	1.000	0.072555	1.000	0.010320	1	0.028780
1361		0.46844	0.002390	1.000	0.088885	1.000	0.007325	1	0.007665
1395		0.19998	0.001010	1.000	0.077725	0.527	0.002635	1	0.054890
1417		1.00000	0.377150	1.000	0.486600	1.000	0.134555	1	0.977235
2094	2215	0.25019	0.001270	0.616	0.003080	1.000	0.078230	1	0.095105
2294	2396	1.00000	0.018445	1.000	0.009035	1.000	0.142580	1	0.013430
2440	2541	1.00000	0.033290	1.000	0.077275	1.000	0.009210	1	0.070880
2601	2681	1.00000	0.880300	1.000	0.087595	1.000	0.089980	1	0.016370

Table 11.3: Pairwise comparisons and overall permutations for Waitangi data, large dataset.
Spots which are significant (p <0.05) after FDR adjustment are indicated an asterix (*). OM = Opua Marina; OW = Opua Wharf; URA = Urapukapuka; WB = Waitangi Bridge

Monte Carlo pairwise comparisons for Wellington data

Comparison between the 3 Wellington groups (control, cadmium and mercury) was performed using a Monte Carlo test to calculate all possible permutations (Manly 1997). Due to the low number of data points (6) there are too few different groups (only 90 different ones), so very low *p*-values are not possible. After FDR adjustment all of the *p*-values were 1 and were therefore no longer significant.

Spot	Spot # Chapter 5	1-ANOVA DeCyder	Control vs. Cadmium (raw)	Control vs. Mercury (raw)	Cadmium vs. Mercury (raw)
1906	2047	0.0060	0.73333333	0.97777778	0.82222222
1950	2082	0.0076	0.13333333	0.02222222	0.68888889
1691			0.28888889	0.75555556	0.02222222
1944			0.02222222	0.71111111	0.35555556
2177	2301	0.017	0.08888889	0.06666667	0.84444444
2428			0.13333333	0.02222222	0.68888889
1954	2084	0.00022	0.06666667	0.08888889	1.00000000
2566			0.02222222	0.13333333	0.75555556
2389	2474	0.14	0.02222222	0.51111111	0.17777778
1953			0.02222222	0.60000000	0.37777778
2397	2488	0.021	0.02222222	0.60000000	0.13333333
2294	2396	0.0052	0.13333333	0.02222222	0.68888889
2601	2681	0.0015	0.04444444	0.11111111	0.77777778
1803			0.02222222	0.42222222	0.22222222
2100			0.02222222	0.53333333	0.37777778
2632			0.13333333	0.02222222	0.68888889
1001	1180	0.10	0.13333333	0.02222222	0.68888889
2120			0.68888889	0.02222222	0.13333333
2352			0.13333333	0.02222222	0.71111111
1361			0.31111111	0.02222222	0.35555556
949			0.02222222	0.68888889	0.13333333
1226	1406	0.34	0.13333333	0.02222222	0.68888889

Table 11.5: Monte Carlo pairwise comparisons for Wellington data
Raw values without FDR corrections.

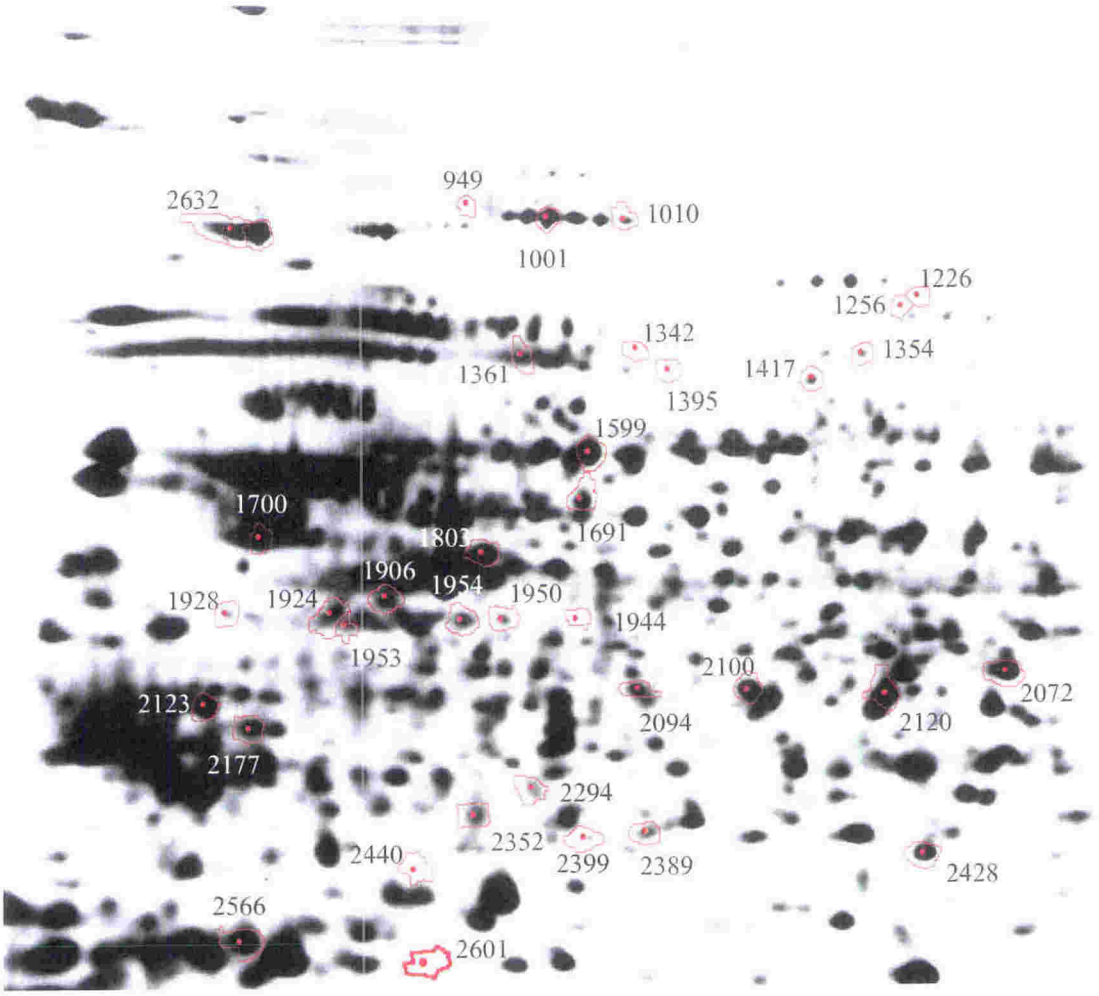


Figure 11.1: Gill pI 4-7 gel showing the spots with potentially significant differences in abundance

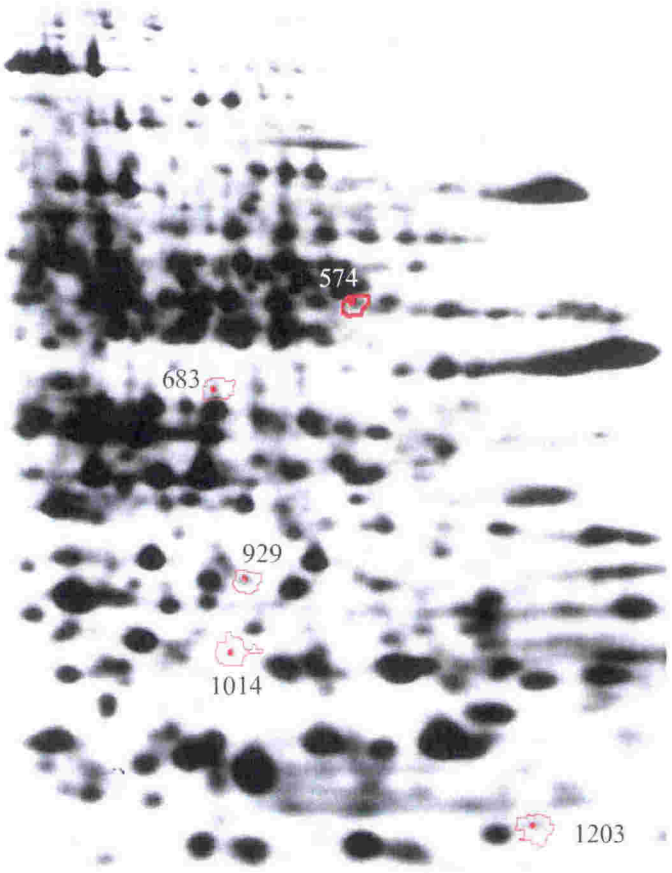


Figure 11.2: Gill pI 6-11 gel showing the spots with potentially significant differences in abundance

Appendix 12: PCA and Pattern Analysis, Bay of Islands Study

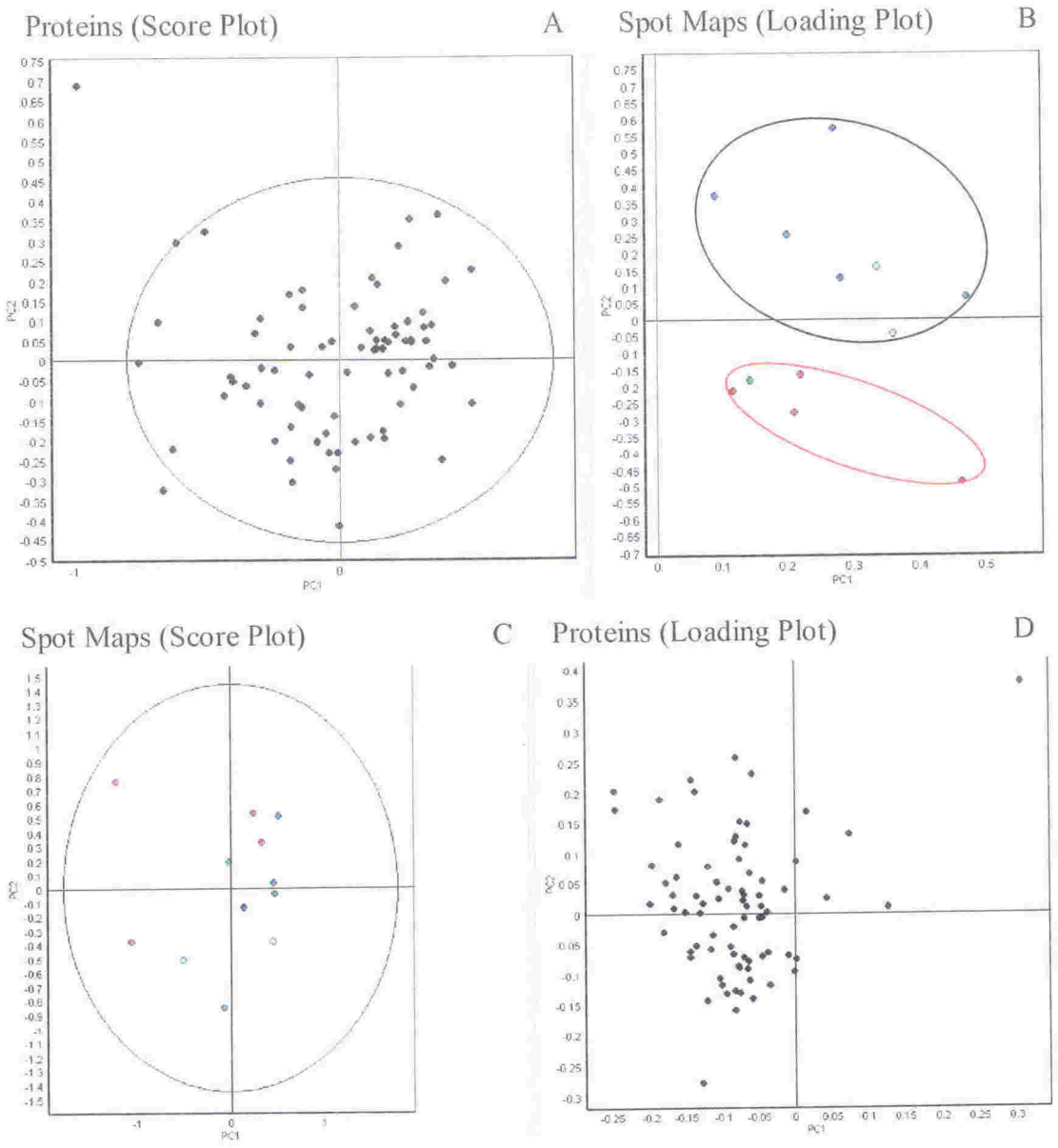


Figure 12.1: PCA for Waitangi samples (gill, pI 6-11)

Key:

- Opua Marina
- Opua Wharf
- Urapukapuka
- Waitangi Bridge
- Cd
- Control
- Hg

The black dots in panels A and D represent individual proteins (n = 77), whereas the coloured dots in panels B and C represent spot maps (n = 14).

Pattern analysis: pI 6-11

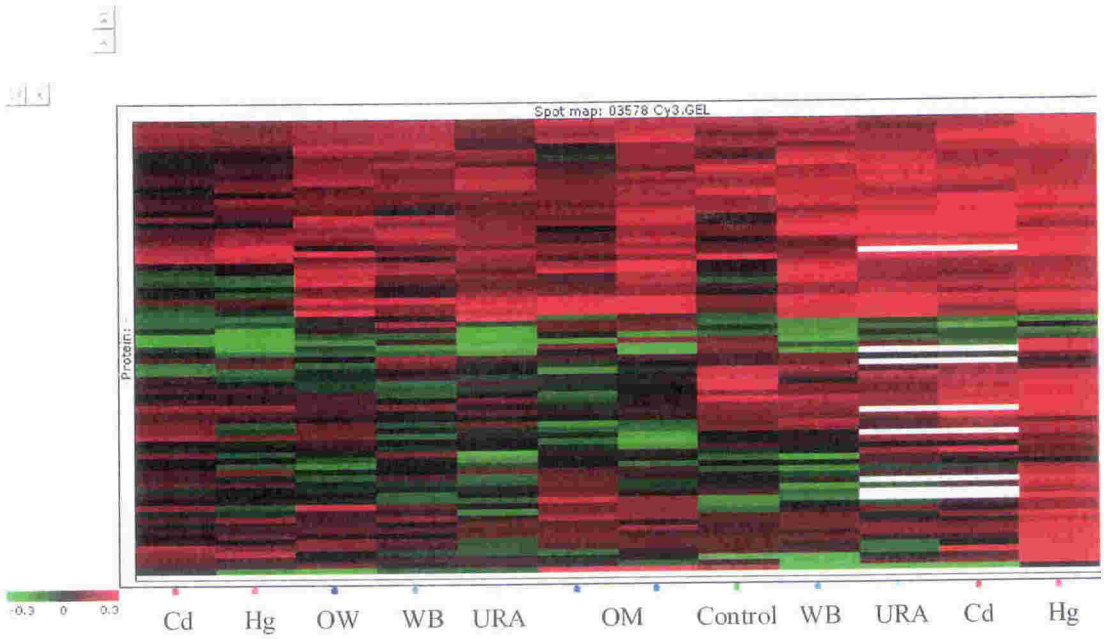


Figure 12.2: Pattern analysis for Waitangi samples (gill, pI 6-11)

Key:

- Opua Marina
- Opua Wharf
- Urapukapuka
- Waitangi Bridge
- Cd
- Control
- Hg

Appendix 13: Calculations estimating dietary metal intake

Example 1: Calculating the PTWI based on the maximum Cd concentration, using an "average" diet.

Maximum Cd concentration = 1.37 mg kg^{-1}

Average serving size quoted as 78 g, so assume the same for mussels so;

1.37 mg of Cd in 1kg
= 0.107 mg in 78 g (i.e. 1kg /78g = 12.8; 1.37/12.8 =0.107)

Therefore 0.107 mg Cd per serving of mussels:
0.107 X 7 (7 servings per week) = 0.749 mg

Divided by the weight of an average 70 kg adult:
0.749 mg / 70kg = 0.0107 mg/kg bw/wk

Convert milligrams to micrograms = **10.7 μg/kg bw/wk**

This calculation is over the limit of 7 μg/kg bw/wk

Example 2: Calculating the amount of mussels required to exceed the PTWI

Average Cd concentration at Urapukapuka = 0.754 mg kg^{-1}
which is equivalent to 754 μg/kg?

PTWI = 7 μg/kg bw/wk

Based on a 70 kg adult:
70 kg x 7 μg (PTWI) = 490 μg to exceed the PTWI

490 μg / 754 μg/kg (the concentration at Urapukapuka)
= 0.650 kg

So if a 70 kg adult consumes 0.650 kg of mussels from Urapukapuka the PTWI for Cd will be exceeded.